

UNIVERSIDADE DE ÉVORA

Doutoramento em Economia

TÁBUAS DE MORTALIDADE CONTEMPORÂNEAS E PROSPECTIVAS:  
MODELOS ESTOCÁSTICOS, APLICAÇÕES ACTUARIAIS  
E COBERTURA DO RISCO DE LONGEVIDADE

*Jorge Miguel Ventura Bravo*



ORIENTADOR Prof. Doutor Carlos Manuel Pereira da Silva  
(INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO - UTL)

CO-ORIENTADOR Prof. Doutor Carlos Alberto dos Santos Braumann  
(UNIVERSIDADE DE ÉVORA)

ÉVORA | 2007

**UNIVERSIDADE DE ÉVORA**

*Doutoramento em Economia*

**TÁBUAS DE MORTALIDADE CONTEMPORÂNEAS E PROSPECTIVAS:**

**MODELOS ESTOCÁSTICOS, APLICAÇÕES ACTUARIAIS**

**E COBERTURA DO RISCO DE LONGEVIDADE**

*Jorge Miguel Ventura Bravo*



168 650

Orientador: Prof. Doutor Carlos Manuel Pereira da Silva  
(INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO - UTL)

Co-orientador: Prof. Doutor Carlos Alberto dos Santos Braumann  
(UNIVERSIDADE DE ÉVORA)

Esta tese não inclui as críticas e sugestões feitas pelo júri

ÉVORA | 2007

**CONTEMPORANEOUS AND PROSPECTIVE LIFETABLES: STOCHASTIC  
MODELS, ACTUARIAL APPLICATIONS AND LONGEVITY RISK HEDGING**

Jorge Miguel Ventura Bravo

**Abstract**

The spectacular improvements in life expectancy documented over the last decades constitute an invaluable achievement of modern societies but pose new challenges in multiple areas, threatening in particular the financial sustainability of classic pay-as-you-go social security systems. Traditionally, actuaries have been calculating premiums and reserves using a deterministic mortality intensity, which is a function of the age only (i.e., considering moment lifetables), and a constant interest rate. This approach is simple and pragmatic but fails to account for the trends observed in mortality intensity. In this thesis, we aim to provide an integrated and systematic approach to the problem of measuring and managing mortality and longevity risks. The thesis offers a detailed review of classic and recent parametric and non-parametric methods used for graduating the mortality curve and for building moment lifetables, including the results of applying parametric methods to the actual mortality experience of life insured and pension beneficiaries in Portugal. We extensively review traditional methods used to project mortality and investigate new solutions for building dynamic or prospective lifetables. These tools were empirically tested and then used to build prospective lifetables for both the Portuguese and pensioners populations and to address the importance of adverse selection effects in the Portuguese market. We develop a dynamic and stochastic approach to longevity risk by admitting that the mortality intensity can be modelled, either by affine-jump stochastic processes, or by an affine function of latent factors whose dynamics is given by ordinary differential equations with jumps. This allows us to capture two important features of the mortality intensity: time dependency and uncertainty of the future development and provides us with closed-form solutions for the survival probability. The analytical tractability of affine processes is explored in actuarial applications. We evaluate a number of both internal and external strategies designed to hedge and transfer longevity and mortality risks, including capital markets-based solutions and the development of mortality-derivatives.

**TÁBUAS DE MORTALIDADE CONTEMPORÂNEAS E PROSPECTIVAS: MODELOS ESTOCÁSTICOS, APLICAÇÕES ACTUARIAIS E COBERTURA DO RISCO DE LONGEVIDADE**

Jorge Miguel Ventura Bravo

**Resumo**

Os espectaculares ganhos de esperança média de vida registados nas últimas décadas constituem uma inquestionável conquista das sociedades modernas mas colocam novos desafios em múltiplas áreas, ameaçando em particular a sustentabilidade dos tradicionais sistemas públicos e privados de segurança e protecção social. Nas companhias de seguros do ramo vida, a abordagem tradicional ao cálculo dos prémios e das reservas matemáticas baseia-se na utilização de uma intensidade de mortalidade determinística, função apenas da idade do indivíduo (tábua contemporânea), e de uma taxa de juro técnica constante. Esta abordagem simples e pragmática revela-se, no entanto, desajustada nos casos em que a mortalidade evolui no tempo. Esta dissertação procura constituir-se como uma abordagem integrada e sistemática sobre o problema da medição e gestão dos riscos de mortalidade e de longevidade. O documento oferece uma revisão detalhada sobre os principais métodos paramétricos e não-paramétricos de graduação de tábuas de mortalidade contemporâneas, incluindo um estudo sobre o seu desempenho na população de pessoas seguras e de beneficiários de fundos de pensões em Portugal. São analisadas técnicas clássicas de projecção da mortalidade e investigadas novas soluções para a construção de tábuas prospectivas. Os resultados são aplicados na elaboração das primeiras tábuas prospectivas para as populações portuguesa e de pensionistas e na avaliação do efeito selecção adversa. É desenvolvida uma abordagem simultaneamente dinâmica e estocástica do risco de longevidade admitindo que a intensidade de mortalidade pode ser modelada, quer por processos estocásticos do tipo afim, quer por uma função do tipo afim de um conjunto de factores latentes com dinâmica ditada por equações diferenciais estocásticas com saltos. Esta solução permite-nos captar duas características importantes da mortalidade: dependência temporal e incerteza sobre a trajectória futura, e possibilita a derivação de soluções fechadas para a probabilidade de sobrevivência, de fácil aplicação no contexto actuarial. São ainda analisadas um conjunto de estratégias internas e externas de cobertura do risco de longevidade, envolvendo a transferência do risco para os mercados de capitais e a criação de um mercado de derivados de mortalidade.

# Índice Geral

|   |           |
|---|-----------|
| Resumo  | i         |
| Abstract  | ii        |
| Índice Geral  | iii       |
| Índice de Figuras   | xii       |
| Índice de Quadros   | xix       |
| Agradecimentos  | xxiii     |
| <br>  |           |
| <b>I Tábuas de Mortalidade Contemporâneas</b>                               | <b>1</b>  |
| <b>1 Introdução</b>   | <b>2</b>  |
| 1.1 Relevância e oportunidade do tema . . . . .                             | 2         |
| 1.1.1 Tendências demográficas e consequências económicas . . . . .          | 2         |
| 1.1.2 Tábuas de mortalidade contemporâneas vs tábuas prospectivas . .       | 3         |
| 1.1.3 Risco de longevidade e risco de mortalidade . . . . .                 | 5         |
| 1.1.4 Modelos determinísticos vs modelos estocásticos . . . . .             | 6         |
| 1.2 Objectivos e principais contributos da dissertação . . . . .            | 7         |
| 1.3 Organização e estrutura do documento . . . . .                          | 9         |
| <br>  |           |
| <b>2 Modelos Probabilísticos para a Descrição da Duração da Vida Humana</b> | <b>11</b> |
| 2.1 Introdução . . . . .  | 11        |
| 2.2 Análise de sobrevivência . . . . .                                      | 12        |
| 2.2.1 Duração de vida restante . . . . .                                    | 12        |
| 2.2.2 Quocientes de mortalidade . . . . .                                   | 14        |

|          |  |           |
|----------|--|-----------|
| 2.2.3    | Taxa de mortalidade instantânea . . . . .                                  | 16        |
| 2.2.4    | Hipóteses para idades fraccionadas . . . . .                               | 18        |
| 2.2.5    | Pressupostos do modelo probabilístico . . . . .                            | 22        |
| 2.3      | Indicadores característicos . . . . .                                      | 23        |
| 2.3.1    | Esperança de vida . . . . .  | 23        |
| 2.3.2    | Variância da vida residual . . . . .                                       | 25        |
| 2.3.3    | Vida mediana e ponto de Lexis . . . . .                                    | 25        |
| 2.4      | Tábuas de mortalidade . . . . .  | 26        |
| 2.4.1    | Introdução e enquadramento histórico . . . . .                             | 26        |
| 2.4.2    | Definição e conceitos básicos . . . . .                                    | 28        |
| 2.4.3    | A tábua enquanto modelo probabilístico de sobrevivência . . . . .          | 30        |
| 2.4.4    | Classificação das tábuas de mortalidade . . . . .                          | 32        |
| 2.5      | Leis de mortalidade . . . . .  | 34        |
| 2.5.1    | Modelos de ajustamento para a função de sobrevivência . . . . .            | 36        |
| 2.5.2    | Modelos que parametrizam a força de mortalidade . . . . .                  | 40        |
| 2.5.3    | Modelos de ajustamento para quocientes de funções de mortalidade . . . . . | 49        |
| 2.5.4    | Modelos de ajustamento para a função de densidade . . . . .                | 51        |
| 2.5.5    | Combinação de leis de mortalidade . . . . .                                | 52        |
| 2.6      | Heterogeneidade e modelos de <i>frailty</i> . . . . .                      | 52        |
| <b>3</b> | <b>Métodos Paramétricos de Graduação</b> . . . . .                         | <b>53</b> |
| 3.1      | Introdução . . . . .   | 53        |
| 3.2      | Caracterização do problema . . . . .                                       | 54        |
| 3.2.1    | Estimação dos quocientes de mortalidade . . . . .                          | 57        |
| 3.2.2    | Estimação da taxa de mortalidade instantânea . . . . .                     | 58        |
| 3.2.3    | Distribuição da exposição ao risco . . . . .                               | 59        |
| 3.3      | Métodos paramétricos . . . . .   | 60        |
| 3.3.1    | Método de máxima verosimilhança . . . . .                                  | 61        |
| 3.3.2    | Método do mínimo Chi-Quadrado . . . . .                                    | 63        |
| 3.3.3    | Método dos mínimos quadrados . . . . .                                     | 65        |
| 3.3.4    | Escolha do método de estimação . . . . .                                   | 65        |
| 3.4      | Metodologia do CMIB . . . . .  | 67        |
| 3.5      | Modelos Lineares Generalizados . . . . .                                   | 70        |
| 3.5.1    | Introdução aos Modelos Lineares Generalizados . . . . .                    | 71        |
| 3.5.2    | Família exponencial natural e <i>link</i> canónico . . . . .               | 72        |

|          |   |            |
|----------|---|------------|
| 3.5.3    | Modelos Lineares Generalizados e graduação . . . . .                          | 75         |
| 3.5.4    | Abordagem dual . . . . .  | 79         |
| 3.5.5    | Modelos Não-Lineares Generalizados e graduação . . . . .                      | 81         |
| 3.6      | Modelo de Heligman-Pollard . . . . .  | 83         |
| 3.7      | Indicadores de qualidade do ajustamento . . . . .                             | 87         |
| <b>4</b> | <b>Métodos Não-Paramétricos de Graduação</b>                                  | <b>92</b>  |
| 4.1      | Introdução . . . . .  | 92         |
| 4.2      | Bin smoothers e médias móveis . . . . .                                       | 93         |
| 4.2.1    | Bin smoothers . . . . .   | 93         |
| 4.2.2    | Método das médias móveis . . . . .  | 94         |
| 4.3      | Técnica de Splines . . . . .  | 95         |
| 4.3.1    | Splines cúbicos . . . . .   | 95         |
| 4.3.2    | B-Splines . . . . .   | 99         |
| 4.3.3    | P-Splines . . . . .   | 101        |
| 4.4      | Método de Whittaker-Henderson . . . . .                                       | 103        |
| 4.5      | Kernel Smoothers . . . . .  | 105        |
| 4.5.1    | Definição e conceitos base . . . . .  | 105        |
| 4.5.2    | Graduação mediante <i>kernel smoothers</i> . . . . .                          | 106        |
| 4.5.3    | Estimador <i>kernel</i> com bandwith variável . . . . .                       | 110        |
| 4.6      | Graduação Polinomial Local . . . . .  | 112        |
| 4.7      | Modelos Aditivos Generalizados . . . . .                                      | 115        |
| 4.8      | Seleccção do parâmetro de alisamento e número de graus de liberdade . . . . . | 116        |
| 4.8.1    | Escolha do parâmetro de alisamento . . . . .                                  | 116        |
| 4.8.2    | Número de graus de liberdade . . . . .  | 118        |
| <b>5</b> | <b>Graduação Paramétrica da Mortalidade da População Segura em Portugal</b>   | <b>119</b> |
| 5.1      | Introdução . . . . .  | 119        |
| 5.2      | Descrição da amostra . . . . .  | 121        |
| 5.2.1    | Beneficiários de fundos de pensões . . . . .                                  | 121        |
| 5.2.2    | Seguros em caso de vida . . . . .   | 125        |
| 5.2.3    | Seguros em caso de morte . . . . .  | 128        |
| 5.3      | Modelos lineares e não-lineares generalizados . . . . .                       | 131        |
| 5.3.1    | Beneficiários de fundos de pensões . . . . .                                  | 131        |
| 5.3.2    | Seguros em caso de vida . . . . .   | 145        |

|  |  |            |
|--|--|------------|
| 5.3.3  | Seguros em caso de morte . . . . .   | 150        |
| 5.4  | GLM's: Abordagem Dual . . . . .  | 156        |
| 5.5  | Modelo de Heligman-Pollard . . . . .   | 159        |
| 5.6  | Análise comparativa dos diferentes métodos . . . . .   | 165        |
| 5.7  | Conclusão . . . . .  | 166        |
| <br><b>II Tábuas de Mortalidade Prospectivas</b> |  | <b>168</b> |
| <br><b>6 Modelos Dinâmicos</b>                   |  | <b>169</b> |
| 6.1  | Introdução . . . . .   | 169        |
| 6.2  | Tendências na mortalidade humana . . . . .   | 171        |
| 6.3  | Modelação da mortalidade num contexto dinâmico . . . . .                                       | 172        |
| 6.3.1  | Diagrama de Lexis e cálculo das estimativas brutas dos<br>indicadores de mortalidade . . . . . | 172        |
| 6.3.2  | Caracterização do problema . . . . .   | 178        |
| 6.3.3  | Tábuas projectadas unidimensionais . . . . .   | 181        |
| 6.3.4  | Classificação dos métodos de projecção . . . . .   | 183        |
| 6.3.5  | Risco de longevidade . . . . .   | 185        |
| 6.4  | Projecção mediante métodos paramétricos . . . . .  | 187        |
| 6.5  | Projecção por extrapolação dos perfis de mortalidade . . . . .                                 | 192        |
| 6.5.1  | Modelo exponencial (ou geométrico) . . . . .   | 194        |
| 6.5.2  | Metodologia do CMIB . . . . .  | 196        |
| 6.5.3  | Modelos de projecção sem tábua base . . . . .  | 199        |
| 6.6  | Método Lee-Carter . . . . .  | 201        |
| 6.6.1  | Definição do modelo . . . . .  | 201        |
| 6.6.2  | Estimação dos parâmetros por mínimos quadrados: análise de<br>componentes principais . . . . . | 203        |
| 6.6.3  | Estimação dos parâmetros por mínimos quadrados: método<br>iterativo . . . . .                  | 206        |
| 6.6.4  | Estimação dos parâmetros por mínimos quadrados ponderados . . . . .                            | 208        |
| 6.6.5  | Projecções de mortalidade . . . . .  | 209        |
| 6.6.6  | Apreciação do método e desenvolvimentos recentes . . . . .                                     | 210        |
| 6.7  | Modelos Lineares Generalizados com respeito à idade e ao tempo . . . . .                       | 214        |
| 6.8  | Modelação dos factores de redução da mortalidade . . . . .                                     | 217        |
| 6.8.1  | Caracterização do problema . . . . .   | 217        |
| 6.8.2  | Factores de redução e modelos GLM . . . . .  | 218        |

|          |  |            |
|----------|--|------------|
| 6.8.3    | Método Lee-Carter modificado . . . . .   | 222        |
| 6.8.4    | Método Lee-Carter com duplo previsor log-bilinear . . . . .  | 224        |
| 6.8.5    | Factores de redução com inclusão de efeitos geracionais . . . . .  | 225        |
| 6.9      | Modelo de Poisson . . . . .  | 228        |
| 6.9.1    | Definição . . . . .  | 228        |
| 6.9.2    | Fundamentação da utilização da distribuição de Poisson . . . . .   | 230        |
| 6.9.3    | Estimação dos parâmetros por máxima verosimilhança . . . . .   | 231        |
| 6.9.4    | Extensão do modelo com incorporação de P-Splines e estimação dos parâmetros por máxima verosimilhança penalizada . . . . . | 235        |
| 6.9.5    | Extensão de Renshaw-Haberman com duplo previsor log-bilinear . . . . .   | 238        |
| 6.9.6    | Extensão do modelo com inclusão de “Tábua Limite” . . . . .  | 240        |
| 6.9.7    | Construção de intervalos de confiança . . . . .  | 245        |
| 6.10     | Modelo Binomial bilinear . . . . .   | 249        |
| 6.11     | Modelos de projecção com “Tábua Limite” . . . . .  | 251        |
| 6.12     | Modelos de projecção com “Tábua Modelo” . . . . .  | 252        |
| 6.13     | Projecção via transformação de funções da tábua de mortalidade . . . . .   | 254        |
| 6.13.1   | Modelo relacional . . . . .  | 254        |
| 6.13.2   | Função de resistência . . . . .  | 256        |
| 6.14     | Métodos não paramétricos . . . . .   | 257        |
| 6.14.1   | P-Splines . . . . .  | 257        |
| 6.14.2   | Métodos baseados em <i>smoothers</i> do tipo Kernel . . . . .  | 259        |
| 6.15     | Métodos de projecção nas idades avançadas . . . . .  | 260        |
| 6.15.1   | Modelo de Heligman-Pollard . . . . .   | 261        |
| 6.15.2   | Modelo Logit . . . . .   | 261        |
| 6.15.3   | Método de Coale-Kisker . . . . .   | 262        |
| 6.15.4   | Método de Lindbergson . . . . .  | 264        |
| 6.15.5   | Método de Denuit e Goderniaux . . . . .  | 264        |
| 6.15.6   | Modelo logístico . . . . .   | 265        |
| 6.15.7   | Modelo de Kannistö . . . . .   | 266        |
| 6.15.8   | Método das gerações extintas . . . . .   | 266        |
| 6.15.9   | Método do rácio de sobrevivência . . . . .   | 267        |
| <b>7</b> | <b>Tábuas de Mortalidade Prospectivas para a População Portuguesa</b>  | <b>269</b> |
| 7.1      | Introdução . . . . .   | 269        |
| 7.1.1    | Bases técnicas . . . . .   | 270        |
| 7.1.2    | Necessidade de tábuas prospectivas . . . . .   | 271        |

|       |   |     |
|-------|---|-----|
| 7.2   | Evolução da mortalidade em Portugal . . . . .   | 272 |
| 7.2.1 | Rectangularização da função de sobrevivência . . . . .  | 273 |
| 7.2.2 | Evolução da esperança de vida . . . . .   | 278 |
| 7.2.3 | Bossa de acidentes . . . . .  | 280 |
| 7.2.4 | Diferencial de mortalidade entre homens e mulheres . . . . .  | 282 |
| 7.3   | Modelo de Poisson . . . . .   | 285 |
| 7.3.1 | Estatísticas de mortalidade . . . . .   | 285 |
| 7.3.2 | Estimativa bruta da força de mortalidade . . . . .  | 286 |
| 7.3.3 | Estimação dos parâmetros por máxima verosimilhança . . . . .  | 287 |
| 7.3.4 | Modelo com incorporação de P-Splines e estimação dos<br>parâmetros por máxima verosimilhança penalizada . . . . . | 295 |
| 7.3.5 | Extensão do modelo com inclusão de tábua limite . . . . .   | 299 |
| 7.3.6 | Extensão do modelo com duplo previsor log-bilinear . . . . .  | 303 |
| 7.4   | Método Lee-Carter . . . . .   | 309 |
| 7.5   | Extrapolação da tendência temporal . . . . .  | 313 |
| 7.5.1 | Estacionaridade da série . . . . .  | 314 |
| 7.5.2 | Identificação do modelo . . . . .   | 318 |
| 7.5.3 | Estimação dos parâmetros do modelo ARIMA . . . . .  | 320 |
| 7.5.4 | Análise dos resíduos . . . . .  | 320 |
| 7.5.5 | Previsão . . . . .  | 321 |
| 7.6   | Projecção da mortalidade nas idades avançadas . . . . .   | 323 |
| 7.6.1 | Fixação da idade biológica máxima . . . . .   | 323 |
| 7.6.2 | Método de fecho das tábuas . . . . .  | 324 |
| 7.7   | Evolução da mortalidade . . . . .   | 327 |
| 7.7.1 | Por ano cronológico . . . . .   | 327 |
| 7.7.2 | Por geração ou ano de nascimento . . . . .  | 329 |
| 7.7.3 | Estimativas da esperança de vida . . . . .  | 331 |
| 7.7.4 | Evolução do preço de uma renda vitalícia . . . . .  | 338 |
| 7.8   | Intervalos de confiança . . . . .   | 341 |
| 7.8.1 | Método de Simulação de Monte Carlo . . . . .  | 342 |
| 7.8.2 | Parametric bootstrap . . . . .  | 346 |
| 7.9   | Conclusão . . . . .   | 349 |
| 7.A   | Anexos . . . . .  | 351 |
| 7.A.1 | Estimativas dos parâmetros do modelo de Poisson . . . . .   | 351 |
| 7.A.2 | Evolução da esperança de vida à nascença e aos 65 anos . . . . .  | 353 |

|          |   |            |
|----------|---|------------|
| <b>8</b> | <b>Avaliação do Efeito Selecção Adversa no Mercado de Fundos de Pensões em Portugal</b> | <b>354</b> |
| 8.1      | Introdução . . . . .  | 354        |
| 8.2      | Relação entre a mortalidade das populações . . . . .                                    | 355        |
| 8.2.1    | Modelo relacional de tipo Brass . . . . .   | 358        |
| 8.2.2    | Modelo log-linear com intercepto variável com a idade . . . . .                         | 358        |
| 8.2.3    | Modelo log-linear com intercepto linearizado . . . . .                                  | 362        |
| 8.2.4    | Modelo log-linear com intercepto constante . . . . .                                    | 363        |
| 8.2.5    | Comparação entre os modelos . . . . .   | 366        |
| 8.3      | Tábuas de mortalidade prospectivas para a população de pensionistas . . . . .           | 367        |
| 8.4      | Estimativa do efeito selecção adversa . . . . .   | 370        |
| 8.4.1    | Na esperança de vida residual . . . . .   | 370        |
| 8.4.2    | No preço das rendas vitalícias . . . . .  | 372        |
| 8.5      | Tábuas prospectivas vs tábuas regulamentares . . . . .                                  | 372        |
| 8.6      | Intervalos de confiança . . . . .   | 377        |
| 8.7      | Conclusão . . . . .   | 380        |

**III Mortalidade Estocástica: Aplicações Actuarias e Cobertura do Risco de Longevidade** **382**

|          |  |            |
|----------|--|------------|
| <b>9</b> | <b>Mortalidade Estocástica</b>   | <b>383</b> |
| 9.1      | Introdução . . . . .   | 383        |
| 9.2      | Ferramentas matemáticas . . . . .  | 385        |
| 9.2.1    | Processos pontuais e intensidades aleatórias . . . . .                     | 386        |
| 9.2.2    | Processos do tipo afim . . . . .   | 392        |
| 9.3      | Modelação da mortalidade num contexto estocástico . . . . .                | 396        |
| 9.3.1    | Estrutura de mortalidade do tipo afim . . . . .                            | 397        |
| 9.3.2    | Revisão da literatura sobre mortalidade estocástica . . . . .              | 403        |
| 9.3.3    | Requisitos importantes nos modelos . . . . .                               | 410        |
| 9.4      | Modelos do tipo afim para a intensidade de mortalidade . . . . .           | 411        |
| 9.4.1    | Modelo de Vasicek . . . . .  | 412        |
| 9.4.2    | Modelo de Vasicek com incerteza ditada exclusivamente por saltos . . . . . | 416        |
| 9.4.3    | Modelo de Cox-Ingersoll-Ross . . . . .                                     | 419        |
| 9.4.4    | Modelo de Merton . . . . .   | 423        |
| 9.4.5    | Modelo de Merton com saltos . . . . .                                      | 424        |
| 9.4.6    | Modelo de Dothan . . . . .   | 426        |

|           |  |            |
|-----------|--|------------|
| 9.4.7     | Modelo de Feller . . . . .   | 428        |
| 9.4.8     | Modelo de Feller com saltos . . . . .  | 430        |
| 9.4.9     | Equação de Ornstein-Uhlenbeck com tendência positiva . . . . .                                     | 432        |
| 9.4.10    | Equação de Ornstein-Uhlenbeck com tendência positiva e saltos . . . . .                            | 434        |
| 9.5       | Modelos do tipo afim com factores latentes . . . . .   | 436        |
| 9.5.1     | Formulação teórica . . . . .   | 437        |
| 9.5.2     | Caso I: Modelo de Makeham . . . . .  | 440        |
| 9.5.3     | Caso II: Modelo de Siler . . . . .   | 446        |
| 9.5.4     | Caso III: Modelo de Thiele não Gaussiano . . . . .   | 449        |
| 9.A       | Anexos . . . . .   | 452        |
| <b>10</b> | <b>Ensaaios de Calibragem dos Modelos Estocásticos do Tipo Afim na População Portuguesa</b>        | <b>457</b> |
| 10.1      | Introdução . . . . .   | 457        |
| 10.2      | Dados e metodologia de calibragem . . . . .  | 458        |
| 10.2.1    | Caracterização da amostra . . . . .  | 458        |
| 10.2.2    | Metodologia de calibragem . . . . .  | 459        |
| 10.3      | Resultados . . . . .   | 460        |
| 10.4      | Conclusão . . . . .  | 475        |
| <b>11</b> | <b>Aplicações Actuariais e Cobertura do Risco de Longevidade</b>                                   | <b>477</b> |
| 11.1      | Introdução . . . . .   | 477        |
| 11.2      | Contratos de seguro e activos com risco de crédito . . . . .                                       | 479        |
| 11.3      | Estrutura do modelo de avaliação . . . . .   | 480        |
| 11.3.1    | Mercados financeiros . . . . .   | 480        |
| 11.3.2    | Mortalidade . . . . .  | 481        |
| 11.3.3    | Modelo combinado . . . . .   | 482        |
| 11.4      | Avaliação de contratos de seguro do ramo vida . . . . .  | 483        |
| 11.4.1    | Activos base . . . . .   | 484        |
| 11.4.2    | Contratos de seguro do ramo vida . . . . .   | 485        |
| 11.4.3    | <i>Risk-neutral</i> valuation e determinação do preço de mercado do risco de mortalidade . . . . . | 492        |
| 11.5      | Cobertura do risco de longevidade . . . . .  | 507        |
| 11.5.1    | Desenho dos produtos e estratégias de <i>pricing</i> . . . . .                                     | 508        |
| 11.5.2    | Conceito de solvência estocástica . . . . .  | 509        |
| 11.5.3    | Estratégia de diversificação e <i>natural hedging</i> . . . . .                                    | 512        |

## Índice Geral

---

|           |  |            |
|-----------|--|------------|
| 11.5.4    | Contratos de resseguro . . . . .   | 514        |
| 11.5.5    | Titularização do risco: <i>Longevity/Survivor Bonds</i> . . . . .                        | 520        |
| 11.5.6    | Derivados de mortalidade . . . . .   | 527        |
| 11.5.7    | Mercado de rendas vitalícias e cobertura individual do risco de<br>longevidade . . . . . | 531        |
| <b>12</b> | <b>Conclusão</b>   | <b>539</b> |
|           | <b>Bibliografia</b>  | <b>541</b> |

# Índice de Figuras

|   |     |
|---|-----|
| 5.2.1 Estrutura etária da população masculina de pensionistas . . . . .             | 122 |
| 5.2.2 Estrutura etária da população feminina de pensionistas . . . . .              | 122 |
| 5.2.3 Estimativas brutas dos quocientes $q_x$ , pensionistas . . . . .              | 124 |
| 5.2.4 Estimativas brutas dos quocientes $q_x$ , população portuguesa 2001/2002 . .  | 124 |
| 5.2.5 Seguros em caso de vida: estrutura etária da população masculina . . . .      | 126 |
| 5.2.6 Seguros em caso de vida: estrutura etária da população feminina . . . .       | 126 |
| 5.2.7 Estimativas brutas dos quocientes $q_x$ , seguros em caso de vida . . . . .   | 127 |
| 5.2.8 Seguros em caso de morte: estrutura etária da população masculina . . .       | 129 |
| 5.2.9 Seguros em caso de morte: estrutura etária da população feminina . . . .      | 129 |
| 5.2.10 Estimativas brutas dos quocientes $q_x$ , seguros em caso de morte . . . . . | 130 |
| 5.3.1 Modelo $LGM(2, 6)$ , pensionistas do sexo masculino . . . . .                 | 136 |
| 5.3.2 Resíduos de Pearson do modelo $LGM(2, 6)$ , pensionistas do sexo masculino    | 136 |
| 5.3.3 Modelo $LGM(0, 7)$ , pensionistas do sexo feminino . . . . .                  | 137 |
| 5.3.4 Resíduos do modelo $LGM(0, 7)$ , pensionistas do sexo feminino . . . . .      | 137 |
| 5.3.5 Modelo $GM(2, 6)$ , pensionistas do sexo masculino . . . . .                  | 142 |
| 5.3.6 Resíduos do modelo $GM(2, 6)$ , pensionistas do sexo masculino . . . . .      | 142 |
| 5.3.7 Modelo $GM(0, 7)$ , pensionistas do sexo feminino . . . . .                   | 143 |
| 5.3.8 Resíduos do modelo $GM(0, 7)$ , pensionistas do sexo feminino . . . . .       | 143 |
| 5.3.9 Modelo $LGM(3, 3)$ , seguros em caso de vida, homens . . . . .                | 148 |
| 5.3.10 Resíduos do modelo $LGM(3, 3)$ , seguros em caso de vida, homens . . . . .   | 148 |
| 5.3.11 Modelo $LGM(3, 3)$ , seguros em caso de vida, mulheres . . . . .             | 149 |
| 5.3.12 Resíduos do modelo $LGM(3, 3)$ , seguros em caso de vida, mulheres . . .     | 149 |
| 5.3.13 Modelo $LGM(2, 3)$ , seguros em caso de morte, homens . . . . .              | 154 |
| 5.3.14 Resíduos do modelo $LGM(2, 3)$ , seguros em caso de morte, homens . . .      | 154 |
| 5.3.15 Modelo $LGM(3, 2)$ , seguros em caso de morte, mulheres . . . . .            | 155 |
| 5.3.16 Resíduos do modelo $LGM(3, 2)$ , seguros em caso de morte, mulheres . .      | 155 |

|        |   |     |
|--------|---|-----|
| 5.4.1  | Comparação dos modelos $GM(0, s)$ Gamma, sexo masculino . . . . .   | 158 |
| 5.4.2  | Comparação dos modelos $GM(0, s)$ Gamma, sexo feminino . . . . .  | 159 |
| 5.5.1  | Graduação pela 2ª Lei de Heligman-Pollard, sexo masculino . . . . .   | 164 |
| 5.5.2  | Graduação pela 2ª Lei de Heligman-Pollard, sexo feminino . . . . .  | 164 |
| 6.3.1  | Diagrama de Lexis . . . . .   | 173 |
| 6.3.2  | Número de óbitos à idade $x$ entre os indivíduos da geração $t - x$ . . . . .   | 174 |
| 6.3.3  | Número de óbitos registados no ano $t$ de indivíduos nascidos em $t - x$ . . . . .  | 175 |
| 6.4.1  | Método vertical . . . . .   | 188 |
| 6.4.2  | Método diagonal . . . . .   | 191 |
| 6.5.1  | Método horizontal . . . . .   | 193 |
| 6.12.1 | Projeção com Tábuas Modelo . . . . .  | 254 |
| 7.2.1  | Densidade $x \mapsto d_x/l_0$ (homens à esquerda e mulheres à direita) . . . . .  | 273 |
| 7.2.2  | Função de sobrevivência $x \mapsto l_x/l_0$ , Homens . . . . .  | 274 |
| 7.2.3  | Função de sobrevivência $x \mapsto l_x/l_0$ , Mulheres . . . . .  | 274 |
| 7.2.4  | Esperança de vida $x \mapsto \hat{e}_x(t)$ (homens à esquerda e mulheres à direita) . . . . .   | 280 |
| 7.2.5  | Logaritmo do quociente de mortalidade $x \mapsto \ln(q_x)$ , Homens . . . . .   | 281 |
| 7.2.6  | Logaritmo do quociente de mortalidade $x \mapsto \ln(q_x)$ , Mulheres . . . . .   | 281 |
| 7.2.7  | Evolução da esperança de vida à nascença $\hat{e}_0$ e aos 65 anos $\hat{e}_{65}$ por sexo . . . . .  | 283 |
| 7.2.8  | Evolução do diferencial entre as esperanças de vida $(e_x^M - e_x^H)$ em função da idade $x$ . . . . .  | 284 |
| 7.2.9  | Ratio entre os quocientes de mortalidade $q_x^H/q_x^M$ em função da idade $x$ . . . . .   | 285 |
| 7.3.1  | Estimativas brutas de $\mu_{x,t}$ no período 1970-2004, Homens . . . . .  | 286 |
| 7.3.2  | Estimativas brutas de $\mu_{x,t}$ no período 1970-2004, Mulheres . . . . .  | 287 |
| 7.3.3  | Estimativas dos parâmetros $\alpha_x, \beta_x$ e $\kappa_t$ do modelo de Poisson aplicado a todas as idades (homens à esquerda e mulheres à direita) . . . . .  | 289 |
| 7.3.4  | Estimativas dos parâmetros $\alpha_x, \beta_x$ e $\kappa_t$ do modelo de Poisson aplicado apenas às idades $x \geq 50$ (homens à esquerda e mulheres à direita) . . . . .   | 290 |
| 7.3.5  | Diferença entre o número de óbitos observado em cada ano de calendário e o seu valor estimado pelo modelo aplicado a todas as idades (em cima) e apenas às idades $x \geq 50$ (em baixo) e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita) . . . . . | 293 |
| 7.3.6  | Resíduos de deviance padronizados do modelo de Poisson (homens à esquerda e mulheres à direita) . . . . .   | 294 |

|   |     |
|---|-----|
| 7.3.7 Estimativas do parâmetro $\beta_x$ pelo método de máxima verosimilhança penalizada para valores seleccionados de $\lambda_\beta$ (homens) . . . . .   | 296 |
| 7.3.8 Estimativas dos parâmetros $\alpha_x$ e $\kappa_t$ pelo método de máxima verosimilhança penalizada para valores seleccionados de $\lambda_\beta$ (homens) . . . . .                         | 297 |
| 7.3.9 Estimativas do parâmetro $\beta_x$ pelo método de máxima verosimilhança penalizada para valores seleccionados de $\lambda_\beta$ (mulheres) . . . . .                                       | 297 |
| 7.3.10 Estimativas dos parâmetros $\alpha_x$ e $\kappa_t$ pelo método de máxima verosimilhança penalizada para valores seleccionados de $\lambda_\beta$ (mulheres) . . . . .                      | 298 |
| 7.3.11 Tábua limite de Duchêne-Wunsch, função de sobrevivência $x \mapsto l_x/l_0$ . . . . .  | 300 |
| 7.3.12 Comparação entre as taxas $\mu_{x,t}$ da tábua limite de Duchêne-Wunsch e as estimativas brutas em 1970, 1985 e 2004 . . . . .   | 301 |
| 7.3.13 Estimativas dos parâmetros $\kappa_t$ do modelo de Poisson com tábua limite de Duchêne-Wunsch (homens à esquerda e mulheres à direita) . . . . .   | 301 |
| 7.3.14 Estimativas dos parâmetros $\alpha_x$ e $\beta_x$ do modelo de Poisson com tábua limite de Duchêne-Wunsch (homens à esquerda e mulheres à direita) . . . . .                               | 302 |
| 7.3.15 Trajectória projectada pelo modelo Poisson com tábua limite para a taxa $\hat{\mu}_{80,t}$ (homens à esquerda e mulheres à direita) . . . . .  | 303 |
| 7.3.16 Estimativas de $\alpha_x$ no modelo de Poisson com duplo previsor log-bilinear (Homens) . . . . .  | 304 |
| 7.3.17 Estimativas de $\beta_{x,1}$ , $\beta_{x,2}$ , $\kappa_{t,1}$ e $\kappa_{t,2}$ do modelo de Poisson com duplo previsor log-bilinear (Homens) . . . . .                                     | 304 |
| 7.3.18 Estimativas de $\alpha_x$ no modelo de Poisson com duplo previsor log-bilinear (Mulheres) . . . . .  | 305 |
| 7.3.19 Estimativas de $\beta_{x,1}$ , $\beta_{x,2}$ , $\kappa_{t,1}$ e $\kappa_{t,2}$ do modelo de Poisson com duplo previsor log-bilinear (Mulheres) . . . . .                                   | 305 |
| 7.3.20 Percentagem da variância explicada pelo modelo de Poisson com um e dois termos log-bilineares (homens em cima e mulheres em baixo) . . . . .   | 307 |
| 7.3.21 Resíduos do modelo de Poisson com um e dois termos log-bilineares (homens à esquerda e mulheres à direita) . . . . .   | 308 |
| 7.4.1 Estimativas dos parâmetros $\alpha_x$ , $\beta_x$ e $\kappa_t$ obtidas pelos métodos Lee-Carter e de Poisson aplicados a todas as idades (homens à esquerda e mulheres à direita) . . . . . | 310 |
| 7.4.2 Reestimação do parâmetro $\kappa_t$ de modo a satisfazer a restrição (6.6.11) . . . . .   | 311 |
| 7.4.3 Percentagem da variância explicada pelos modelos de Lee-Carter e de Poisson (homens à esquerda e mulheres à direita) . . . . .  | 311 |

|   |     |
|---|-----|
| 7.4.4 Resíduos de deviance estandardizados do método Lee-Carter (homens à esquerda e mulheres à direita) . . . . .  | 312 |
| 7.5.1 Séries iniciais $\hat{\kappa}_t$ e primeiras diferenças $(1 - B)\hat{\kappa}_t$ (homens à esquerda e mulheres à direita) . . . . .  | 315 |
| 7.5.2 Coeficientes de autocorrelação e de autocorrelação parcial para as séries diferenciadas $\Delta_1\kappa_t$ (homens à esquerda e mulheres à direita) . . . . .   | 318 |
| 7.5.3 QQ-plots normal dos resíduos dos modelos ARIMA (homens à esquerda e homens à direita) . . . . .   | 321 |
| 7.5.4 Projecção do parâmetro $\kappa_t$ para o horizonte temporal 2060 e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita) . . . . .   | 322 |
| 7.6.1 Quocientes $q_{x,t}$ com tábuas fechadas (Homens) . . . . .   | 326 |
| 7.6.2 Quocientes $q_{x,t}$ com tábuas fechadas (Mulheres) . . . . .   | 327 |
| 7.7.1 Evolução das $\mu_{x,t}$ ajustadas pelo modelo de Poisson aplicado a todas as idades (homens à esquerda e mulheres à direita) . . . . .   | 328 |
| 7.7.2 Evolução das $\mu_{x,t}$ ajustadas pelo modelo de Poisson aplicado ao intervalo de idades $x \in [50, 84]$ (homens à esquerda e mulheres à direita) . . . . .   | 328 |
| 7.7.3 Evolução dos $q_{x,t}$ entre 1970 e 2124 para algumas idades representativas (homens à esquerda e mulheres à direita) . . . . .   | 329 |
| 7.7.4 Evolução das $\mu_{x,t}$ ajustadas pelo modelo de Poisson para os indivíduos nascidos nos anos de 1970, 1980, 1990 e 2004 (homens à esquerda e mulheres à direita) . . . . .                                    | 330 |
| 7.7.5 Evolução das $\mu_{x,t}$ para as gerações de indivíduos que atingem os 50 anos em 1970, 1980, 1990 e 2004 (homens à esquerda e mulheres à direita) . . . . .  | 331 |
| 7.7.6 Comparação entre as taxas $\mu_{x,t}$ obtidas nas ópticas transversal e geracional em anos seleccionados (homens à esquerda e mulheres à direita) . . . . .   | 332 |
| 7.7.7 Evolução da esperança de vida $\hat{e}_x(t)$ segundo as ópticas transversal e longitudinal no período $t \in [1970, 2004]$ e idades $x = 0$ e 65 anos (homens à esquerda e mulheres à direita) . . . . .        | 336 |
| 7.7.8 Evolução projectada da esperança de vida à nascença e aos 65 anos calculada numa óptica transversal . . . . .   | 337 |
| 7.7.9 Evolução do preço de uma renda vitalícia $a_{65}(t)$ calculado segundo as ópticas transversal e longitudinal no período $t \in [1970, 2004]$ , com $i = 3\%$ (homens à esquerda e mulheres à direita) . . . . . | 341 |
| 7.8.1 Histogramas dos valores empíricos de $e_0(2004)$ e $e_{65}(2004)$ , MSMC . . . . .  | 345 |
| 7.8.2 Histogramas dos valores empíricos de $a_0(2004)$ e $a_{65}(2004)$ , MSMC . . . . .  | 345 |
| 7.8.3 Histogramas dos valores de $e_0(2004)$ e $e_{65}(2004)$ , Parametric Bootstrap . . . . .  | 347 |

|       |   |     |
|-------|---|-----|
| 7.8.4 | Histogramas dos valores de $a_0$ (2004) e $a_{65}$ (2004), Parametric Bootstrap   | 348 |
| 7.8.5 | Quocientes de mortalidade $q_{0+k}$ (2004 + $k$ ) simulados pelo método parametric bootstrap (homens à esquerda e mulheres à direita)   | 348 |
| 8.2.1 | Taxas brutas $\mu_{x,2004}$ ( $x \geq 50$ ) para a população portuguesa e para os beneficiários de fundos de pensões (homens à esquerda e mulheres à direita)   | 356 |
| 8.2.2 | Taxas brutas $\mu_{x,t}$ ( $x = 55, 65, 75, 85$ ) para a população portuguesa e para os beneficiários de fundos de pensões (homens à esquerda e mulheres à direita)   | 357 |
| 8.2.3 | Relação entre as taxas $\mu_{x,t}$ das populações portuguesa e de beneficiários de fundos de pensões em idades seleccionadas (homens em cima e mulheres em baixo)   | 359 |
| 8.2.4 | Estimativas do parâmetro $\eta_x$ do modelo (8.2.3) e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita)  | 361 |
| 8.2.5 | Resultado da regressão linear de $\hat{\eta}_x$ sobre a idade $x$ (homens à esquerda e mulheres à direita)  | 362 |
| 8.2.6 | Estimação do modelo log-linear com intercepto constante pelo método dos mínimos quadrados ponderados (homens à esquerda e mulheres à direita)   | 364 |
| 8.2.7 | Estimativa do modelo relacional (8.2.7) ligando as taxas de mortalidade da população e dos BFP (homens à esquerda e mulheres à direita)   | 365 |
| 8.3.1 | Evolução temporal dos quocientes $q_{x,t}^{BFP}$ ajustados pelo modelo Poisson-Relacional (homens à esquerda e mulheres à direita)  | 367 |
| 8.3.2 | Comparação entre as taxas $\mu_{x,t}$ brutas e estimadas para as populações portuguesa e de BFP em anos seleccionados (homens à esquerda e mulheres à direita)  | 368 |
| 8.3.3 | Quocientes de mortalidade $q_{x,2004}^{POP}$ e $q_{x,2004}^{BFP}$ relativos à geração que atinge a idade $x = 50$ em 2004 (homens à esquerda e mulheres à direita)  | 369 |
| 8.5.1 | Estimativas de $\hat{e}_{65}(t)$ (à esquerda) e de $a_{65}(t)$ (à direita) para as populações portuguesa e de BFP, segundo a óptica longitudinal  | 374 |
| 8.5.2 | Estimativas de $\hat{e}_{65}(t)$ (à esquerda) e de $a_{65}(t)$ (à direita) para as populações portuguesa e de BFP, segundo a óptica transversal   | 374 |
| 8.5.3 | Prémio puro $a_{65}^{POP}(t)$ e $a_{65}^{BFP}(t)$ calculado com base nas tábuas prospectivas obtidas para as populações portuguesa e de BFP, respectivamente, e com base nas tábuas regulamentares TV 73/77 e TV 88/90 (homens à esquerda e mulheres à direita) | 375 |

|         |   |     |
|---------|---|-----|
| 8.6.1   | Histogramas dos valores de $\hat{e}_{65}^{BFP}$ (2004) e $a_{65}^{BFP}$ (2004) gerados pelo método de Parametric Bootstrap (homens à esquerda e mulheres à direita) . . . . .   | 379 |
| 8.6.2   | Taxas de mortalidade $\mu_{50+k}^{BFP}$ (2004 + k) simuladas pelo método parametric bootstrap (homens à esquerda e mulheres à direita) . . . . .  | 380 |
| 10.3.1  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Vasicek (homens à esquerda e mulheres à direita) . . . . .  | 463 |
| 10.3.2  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Vasicek com saltos (homens à esquerda e mulheres à direita) . . . . .                                     | 464 |
| 10.3.3  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Cox-Ingersoll-Ross (homens à esquerda e mulheres à direita) . . . . .                                     | 465 |
| 10.3.4  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Merton (homens à esquerda e mulheres à direita) . . . . .   | 466 |
| 10.3.5  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Merton com saltos (homens à esquerda e mulheres à direita) . . . . .                                      | 467 |
| 10.3.6  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Dothan (homens à esquerda e mulheres à direita) . . . . .   | 468 |
| 10.3.7  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Feller (homens à esquerda e mulheres à direita) . . . . .   | 469 |
| 10.3.8  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pelo modelo de Feller com saltos (homens à esquerda e mulheres à direita) . . . . .                                      | 470 |
| 10.3.9  | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pela equação de Ornstein-Uhlenbeck com <i>drift</i> positivo (homens à esquerda e mulheres à direita) . . . . .          | 471 |
| 10.3.10 | Calibragem da probabilidade de sobrevivência ${}_{T-t}p_{65}(t)$ em função da idade $x + T - t$ para $t = 1970$ e $t = 2004$ pela equação de Ornstein-Uhlenbeck com <i>drift</i> positivo e saltos (homens à esquerda e mulheres à direita) . . . . . | 472 |

|         |   |     |
|---------|---|-----|
| 10.3.11 | Desvios entre a função de sobrevivência $T_{-t}p_{65}(1970)$ estimada pelos modelos e a implícita nas tábuas prospectivas, Homens . . . . .   | 473 |
| 10.3.12 | Desvios entre a função de sobrevivência $T_{-t}p_{65}(1970)$ estimada pelos modelos e a implícita nas tábuas prospectivas, Mulheres . . . . . | 474 |
| 10.3.13 | Desvios entre a função de sobrevivência $T_{-t}p_{65}(2004)$ estimada pelos modelos e a implícita nas tábuas prospectivas, Homens . . . . .   | 474 |
| 10.3.14 | Desvios entre a função de sobrevivência $T_{-t}p_{65}(2004)$ estimada pelos modelos e a implícita nas tábuas prospectivas, Mulheres . . . . . | 475 |

# Índice de Quadros

|   |     |
|---|-----|
| 3.5.1 Distribuições de probabilidade englobadas na família exponencial natural  | 73  |
| 3.5.2 Média e variância associada às distribuições de probabilidade englobadas na família exponencial natural . . . . .                   | 73  |
| 3.5.3 Funções de <i>link</i> canónico associadas às principais leis de probabilidade pertencentes à família exponencial natural . . . . . | 74  |
| 3.5.4 Funções de <i>link</i> canónico usadas em problemas de graduação . . . . .  | 76  |
| 5.2.1 Beneficiários de fundos de pensões ( $E_x; d_x$ ) . . . . .   | 121 |
| 5.2.2 Seguros em caso de vida ( $E_x; d_x$ ) . . . . .  | 125 |
| 5.2.3 Seguros em caso de morte ( $E_x; d_x$ ) . . . . .   | 128 |
| 5.3.1 LV e deviance para pensionistas do sexo masculino, $LGM(r, s)$ . . . . .  | 132 |
| 5.3.2 LV e deviance para pensionistas do sexo feminino, $LGM(r, s)$ . . . . .   | 132 |
| 5.3.3 Indicadores de qualidade do ajustamento, $LGM(r, s)$ , pensionistas do sexo masculino . . . . .                                     | 134 |
| 5.3.4 Indicadores de qualidade do ajustamento, $LGM(r, s)$ , pensionistas do sexo feminino . . . . .                                      | 135 |
| 5.3.5 Significância estatística dos parâmetros, $LGM(r, s)$ , pensionistas . . . . .  | 138 |
| 5.3.6 Testes de qualidade do ajustamento, $LGM(r, s)$ , pensionistas . . . . .  | 138 |
| 5.3.7 LV e deviance para pensionistas do sexo masculino, $GM(r, s)$ . . . . .   | 139 |
| 5.3.8 LV e deviance para pensionistas do sexo feminino, $GM(r, s)$ . . . . .  | 139 |
| 5.3.9 Indicadores de qualidade do ajustamento, $GM(r, s)$ , pensionistas do sexo masculino . . . . .                                      | 140 |
| 5.3.10 Indicadores de qualidade do ajustamento, $GM(r, s)$ , pensionistas do sexo feminino . . . . .                                      | 141 |
| 5.3.11 Significância estatística dos parâmetros, $GM(r, s)$ , pensionistas . . . . .  | 144 |
| 5.3.12 Testes de qualidade do ajustamento, $GM(r, s)$ , pensionistas . . . . .  | 144 |
| 5.3.13 LV e deviance, seguros em caso de vida, sexo masculino, $LGM(r, s)$ . . . . .  | 145 |

|        |  |     |
|--------|--|-----|
| 5.3.14 | LV e deviance, seguros em caso de vida, sexo feminino, $LGM(r, s)$ . . . .   | 145 |
| 5.3.15 | Indicadores de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de vida, sexo masculino . . . . .   | 146 |
| 5.3.16 | Indicadores de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de vida, sexo feminino . . . . .  | 147 |
| 5.3.17 | Significância estatística dos parâmetros, $LGM(r, s)$ , seguros em caso de vida . . . . .  | 150 |
| 5.3.18 | Testes de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de vida  | 150 |
| 5.3.19 | LV e deviance, seguros em caso de morte, sexo masculino, $LGM(r, s)$ . .   | 151 |
| 5.3.20 | LV e Deviance, seguros em caso de morte, sexo feminino, $LGM(r, s)$ . . .  | 151 |
| 5.3.21 | Indicadores de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de morte, sexo masculino . . . . .  | 152 |
| 5.3.22 | Indicadores de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de morte, sexo feminino . . . . .   | 153 |
| 5.3.23 | Significância estatística dos parâmetros, $LGM(r, s)$ , seguros em caso de morte . . . . .   | 156 |
| 5.3.24 | Testes de qualidade do ajustamento, $LGM(r, s)$ , seguros em caso de morte   | 156 |
| 5.4.1  | Abordagens Dual e de Poisson, $GM(0, s)$ , sexo masculino . . . . .  | 157 |
| 5.4.2  | Abordagens Dual e de Poisson, $GM(0, s)$ , sexo feminino . . . . .   | 157 |
| 5.5.1  | Funções de perda no modelo de Heligman-Pollard . . . . .   | 160 |
| 5.5.2  | Comparação entre as Leis de Heligman-Pollard, sexo masculino . . . . .   | 160 |
| 5.5.3  | Comparação entre as Leis de Heligman-Pollard, sexo feminino . . . . .  | 161 |
| 5.5.4  | Comparação entre as funções de perda, sexo masculino . . . . .   | 162 |
| 5.5.5  | Comparação entre as funções de perda, sexo feminino . . . . .  | 162 |
| 5.5.6  | Modelo de Heligman-Pollard, estimativa dos parâmetros do modelo, sexo masculino . . . . .  | 163 |
| 5.5.7  | Modelo de Heligman-Pollard, estimativa dos parâmetros do modelo, sexo feminino . . . . .   | 163 |
| 5.6.1  | Análise comparativa dos resultados dos testes de qualidade do ajustamento  | 165 |
| 6.3.1  | Quocientes de mortalidade anuais por idade e ano cronológico . . . . .   | 180 |
| 7.2.1  | Esperança de vida à nascença, variância e coeficiente de variação das tábuas de mortalidade da população portuguesa (I.N.E.) no período de 1930-31 a 2002-03 . . . . . | 275 |

|  |     |
|--|-----|
| 7.2.2 Entropia das tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003 . . . . .  | 277 |
| 7.2.3 Vida mediana subjacente às tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003 . . . . .  | 278 |
| 7.2.4 Intervalo inter-quartis <i>IQR</i> subjacente às tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003 . . . . .  | 278 |
| 7.2.5 Evolução da esperança de vida da população portuguesa entre 1930-1931 e 2002-2003 . . . . .  | 279 |
| 7.3.1 Percentagem da variância explicada pelo modelo Poisson em idades seleccionadas . . . . .   | 292 |
| 7.3.2 Estimativa dos parâmetros de penalização óptimos . . . . .   | 298 |
| 7.5.1 Testes de Phillips-Perron para a existência de raízes unitárias . . . . .  | 316 |
| 7.5.2 Testes ADF para a existência de raízes unitárias . . . . .   | 317 |
| 7.5.3 Testes de Portmanteau e Bartlett à hipótese de ruído branco . . . . .  | 317 |
| 7.5.4 Identificação da ordem do modelo ARIMA . . . . .   | 319 |
| 7.5.5 Ordem dos modelos ARIMA seleccionados . . . . .  | 319 |
| 7.5.6 Estimativas dos parâmetros do modelo ARMA . . . . .  | 320 |
| 7.5.7 Teste de Ljung e Box para os resíduos do modelo ARIMA . . . . .  | 321 |
| 7.7.1 Evolução da esperança de vida à nascença $e_0(t)$ . . . . .  | 333 |
| 7.7.2 Evolução da esperança de vida aos 65 anos $e_{65}(t)$ . . . . .  | 335 |
| 7.7.3 Evolução previsível da esperança de vida $e_0(t)$ e $e_{65}(t)$ segundo uma óptica transversal . . . . .   | 337 |
| 7.7.4 Evolução do preço de uma renda vitalícia à nascença $a_0(t)$ e aos 65 anos $a_{65}(t)$ , calculado segundo as ópticas transversal e longitudinal . . . . .   | 339 |
| 7.7.5 Taxa interna de rentabilidade . . . . .  | 340 |
| 7.8.1 Valores médios e percentis empíricos de $e_x(2004)$ e $a_x(2004)$ , $x = 0, 65$ , gerados pelo método de Simulação de Monte Carlo, onde $Q_\alpha$ denota o percentil empírico de ordem $\alpha$ . . . . . | 344 |
| 7.8.2 Valores médios e percentis empíricos de $e_x(2004)$ e $a_x(2004)$ , $x = 0, 65$ , simulados pelo método parametric Bootstrap, onde $Q_\alpha$ denota o percentil de ordem $\alpha$ . . . . .               | 347 |
| 7.A.1 Estimativas dos parâmetros $\alpha_x$ e $\beta_x$ do modelo de Poisson . . . . .   | 351 |
| 7.A.2 Estimativas do parâmetro $\kappa_t$ do modelo de Poisson . . . . .   | 352 |
| 7.A.3 Evolução da esperança de vida $e_x(t)$ no período 1970-2004 . . . . .  | 353 |

|   |     |
|---|-----|
| 8.2.1 Estimativas do parâmetro $\theta$ do modelo log-linear com intercepto variável com a idade . . . . .  | 360 |
| 8.2.2 Regressão linear do parâmetro $\eta_x$ sobre a idade $x$ . . . . .  | 361 |
| 8.2.3 Estimativas dos parâmetros do modelo log-linear com intercepto linearizado  | 363 |
| 8.2.4 Estimativas dos parâmetros do modelo log-linear com intercepto constante pelo método dos mínimos quadrados ponderados . . . . .   | 363 |
| 8.2.5 Estimativas dos parâmetros do modelo log-linear com intercepto constante no contexto dos modelos lineares generalizados . . . . .   | 365 |
| 8.2.6 Estimativas de $e_{65}(2004)$ e $a_{65}(2004)$ segundo as ópticas transversal e longitudinal calculadas com base nas tábuas geradas pelos modelos relacionais   | 366 |
| 8.4.1 Comparação entre as esperanças de vida aos 65 anos para as populações portuguesa e de BFP, segundo as visões longitudinal e transversal . . . . .   | 371 |
| 8.5.1 Comparação entre os prémios de rendas vitalícias calculados aos 65 anos para as populações portuguesa e de BFP, segundo as visões longitudinal e transversal . . . . .  | 373 |
| 8.5.2 Taxa de juro técnica necessária para igualar o prémio puro de uma renda vitalícia calculada em 2004 segundo as tábuas prospectivas ao seu valor deduzido a partir das tábuas TV 73/77 e TV 88/90 . . . . .      | 376 |
| 8.6.1 Valores médios e quartis empíricos de $e_x(2004)$ e $a_x(2004)$ para a população de BFP, simulados pelo método de parametric bootstrap, onde $Q_\alpha$ denota o percentil empírico de ordem $\alpha$ . . . . . | 378 |
| 10.3.1 Erro de calibragem dos diferentes modelos estocásticos do tipo afm . . . . .   | 461 |
| 10.3.2 Estimativas dos parâmetros do modelo de Vasicek . . . . .  | 463 |
| 10.3.3 Estimativas dos parâmetros do modelo de Vasicek com incerteza ditada exclusivamente por saltos . . . . .   | 464 |
| 10.3.4 Estimativas dos parâmetros do modelo de Cox-Ingersoll-Ross . . . . .   | 465 |
| 10.3.5 Estimativas dos parâmetros do modelo de Merton . . . . .   | 466 |
| 10.3.6 Estimativas dos parâmetros do modelo de Merton com saltos . . . . .  | 467 |
| 10.3.7 Estimativas dos parâmetros do modelo de Dothan . . . . .   | 468 |
| 10.3.8 Estimativas dos parâmetros do modelo de Feller . . . . .   | 469 |
| 10.3.9 Estimativas dos parâmetros do modelo de Feller com saltos . . . . .  | 470 |
| 10.3.10 Estimativas dos parâmetros da equação de Ornstein-Uhlenbeck com <i>drift</i> positivo . . . . .   | 471 |
| 10.3.11 Estimativas dos parâmetros da equação de Ornstein-Uhlenbeck com <i>drift</i> positivo e saltos . . . . .  | 472 |

# Agradecimentos

Aqueles que tiverem a paciência de ler na íntegra este trabalho rapidamente se interrogarão sobre a possibilidade de ele ter sido obra de uma única pessoa. A resposta a essa questão é simples: obviamente que não é. Das mais variadas formas, muitas pessoas contribuíram para que este barco chegasse a bom porto. Tentarei, por isso, não me esquecer de ninguém nesta hora de agradecer a enorme ajuda que me prestaram.

Em primeiro lugar, agradeço ao orientador desta dissertação, Prof. Dr. Carlos Manuel Pereira da Silva, Professor Catedrático do Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa, e ao co-orientador, Prof. Dr. Carlos Alberto dos Santos Braumann, Professor Catedrático da Universidade de Évora, o inestimável incentivo e apoio concedidos durante esta longa jornada. Respeitando, desde o início, as minhas opções e pontos de vista, contribuíram decisivamente com os seus sábios ensinamentos, críticas e sugestões para que a trajectória seguida não se afastasse do objectivo traçado. Todos os erros, imprecisões e/ou emissões que no resultado final subsistem são, naturalmente, da minha inteira responsabilidade.

À Fundação Eugénio de Almeida em Évora, agradeço o interesse manifestado pelo projecto de investigação, materializado na atribuição de uma Bolsa de Doutoramento.

A obtenção das estatísticas demográficas sobre a população portuguesa, sobre as populações cobertas por seguros de vida e sobre a população de beneficiários de fundos de pensões não teria sido possível sem a colaboração do Instituto Nacional de Estatística e do Instituto de Seguros de Portugal (ISP).

Aos meus pais, José António Perico Bravo e Luísa Maria David Ventura Bravo, e aos meus irmãos, Luís Bravo e Ruben Bravo, agradeço eternamente todo o apoio e incentivo que me deram para abraçar mais este difícil projecto, proporcionando-me todas as condições ao seu alcance para que a sua concretização tivesse sido possível. Foram eles, juntamente com Dra. Ana Isabel Gafo, que mais sofreram com a minha dedicação a esta causa. A todos eles dedico este trabalho. Espero que o futuro me permita compensá-los de alguma forma pela ausência prolongada que este projecto representou.

## 0. Agradecimentos

---

A extensa recolha bibliográfica não teria sido possível sem a ajuda da Dra. Otilia Silva e da Dra. Marta Guilherme, colaboradores na Biblioteca do ISP, de todos os colaboradores da Biblioteca Geral da Universidade de Évora e do Jorge, colaborador da Biblioteca Francisco Pereira de Moura no ISEG.

No Departamento de Economia da Universidade de Évora, agradeço a todos os colegas o seu incansável apoio e encorajamento, em particular aos Professores Doutores Joaquim Ramalho, Esmeralda Ramalho, Paulo Neto, João Vaz Rodrigues, Silvério Rocha e Cunha e Nuno Salter Cid. Nesta mesma escola, não posso deixar de agradecer o entusiasmo dos meus alunos de Licenciatura e de Mestrado, que com as suas questões e interrogações me incentivam todos os dias a aprofundar o conhecimento sobre estas matérias. Os meus amigos merecem uma palavra de apreço pela compreensão que manifestaram quando faltava, repetidamente, às *convocatórias*. Prometo desferrar-me.

A nota final é naturalmente para mim. Em 2002, aquando da redacção da dissertação de Mestrado, conclui esta mesma secção de agradecimentos desejando um dia conseguir realizar projectos que exigissem menos esforço e sacrifício pessoal. Volvidos quase cinco anos, e perante o documento que agora apresento, é por demais evidente que não consegui cumprir esse desiderato. Por isso, não me resta outra coisa que não seja renovar esse desejo reproduzindo aqui a mensagem com que então conclui.

Espero um dia conseguir fazer trabalhos que exijam...menos trabalho.

Évora, Setembro de 2007

Aos meus Pais,  
*José e Luisa*

**Parte I**

**Tábuas de Mortalidade  
Contemporâneas**

# Capítulo 1

## Introdução

*“O começo é a parte mais importante do trabalho.”* (Platão, 427-347 BC)

*“We don’t do it for the glory. We don’t do it for the recognition. We do it because it needs to be done. Because if we don’t, no one else will. And we do it even if no one knows what we’ve done. Even if no one knows we exist. Even if no one remembers we ever existed.”* (Supergirl (Kara Zor-El)  
- Christmas With The Superheroes #2, Dec 1989)

### 1.1 Relevância e oportunidade do tema

#### 1.1.1 Tendências demográficas e consequências económicas

A evolução da longevidade humana nos países desenvolvidos nas últimas décadas é marcada por um aumento notável da esperança média de vida em todas as idades, resultado de uma diminuição assinalável nas taxas de mortalidade em todo o arco da vida humana, em especial nas idades mais jovens, e por uma redução na variabilidade da idade de óbito, que conduziu a uma progressiva rectangularização da função de sobrevivência.

Estes ganhos na longevidade humana constituem, naturalmente, uma alteração positiva para os indivíduos e uma significativa conquista civilizacional das sociedades modernas. Tratando-se de um objectivo individual e colectivo inquestionável, o aumento da longevidade humana coloca, no entanto, novos desafios em múltiplas áreas de intervenção, quer no domínio público, quer na esfera da actividade privada.

Ao nível dos sistemas públicos de protecção e segurança social, as tendências demográficas registadas nas últimas décadas, caracterizadas por um envelhecimento acentuado, pela diminuição da taxa de fecundidade, pela diminuição da relação entre traba-

## 1.1. Relevância e oportunidade do tema

---

lhadores activos e dependentes e pelo significativo aumento da esperança média de vida em todas as idades, ameaçam a sustentabilidade dos tradicionais sistemas financiados com base no princípio da repartição contemporânea (“*pay-as-you-go*”). A progressiva moderação dos benefícios concedidos pelos sistemas de repartição, a conversão dos planos de pensões de benefício definido em planos de contribuição definida, baseados em contas individuais em regime de capitalização (real ou virtual), tenderão a reduzir a parcela do rendimento na reforma que é “garantida”. Neste contexto, a criação de um mercado privado de rendas vitalícias competitivo e eficiente emerge como um substituto natural da função de seguro social desempenhada pelas pensões públicas. No sector privado, o aumento da longevidade humana tem consequências óbvias no *pricing* e no valor das responsabilidades financeiras das companhias de seguros e de resseguro e dos fundos de pensões com produtos que envolvem benefícios de longo prazo em caso de vida (e.g., rendas vitalícias e pensões de velhice e/ou invalidez de benefício definido), constituindo um motivo de preocupação para todos os *stakeholders* envolvidos.

A um outro nível, é de prever que estas tendências demográficas afectem os níveis de crescimento económico, modificando a dimensão e a estrutura da oferta de trabalho e do emprego, a acumulação de capital e a produtividade total dos factores, a dimensão e a composição da despesa pública, o comportamento do consumo e da poupança e a evolução dos mercados financeiros. O prolongamento da vida humana aumenta igualmente a pressão sobre os sistemas de saúde e de apoio social e refunda o papel da família na sociedade.

### 1.1.2 Tábuas de mortalidade contemporâneas vs tábuas prospectivas

A determinação do prémio puro de um contrato de seguros do ramo vida assenta no princípio da equivalência, segundo o qual o valor actual esperado das prestações devidas pela seguradora deve ser igual ao valor actual esperado dos prémios puros pagos pela pessoa segura. Ao prémio puro assim obtido, é habitual as companhias adicionarem um conjunto de cargas (margem de segurança, comissão de gestão, comissão de emissão, etc.) de modo a obter um prémio comercial.

Nos últimos anos, a problemática dos métodos e pressupostos usados na avaliação actuarial dos activos e das responsabilidades ganhou uma importância acrescida, não apenas para as empresas, mas também para as entidades reguladoras do sector a nível nacional e internacional. Tradicionalmente, os actuários baseiam o cálculo dos prémios de seguro em tábuas contemporâneas ou de momento, construídas a partir de quocientes de mortalidade estimados com base nas estatísticas demográficas observadas num dado

## 1.1. Relevância e oportunidade do tema

---

momento. Esta abordagem simples e pragmática revela-se claramente desajustada nos casos em que a mortalidade evolui no tempo. Com efeito, os prémios assim calculados não reflectem o risco real de nenhuma das pessoas seguras da carteira (salvo na hipótese, pouco provável, de que a mortalidade permanece estática no tempo), pelo que em contextos de aumento da longevidade a realização do *pricing* nestes termos conduz a uma subestimação sistemática dos prémios e das responsabilidades com operações em caso de vida.

No passado, as companhias lidavam com o problema escolhendo taxas de juro e tábuas de mortalidade prudentes (do ponto de vista da empresa), o que fazia com que à medida que o retorno dos investimentos e o comportamento da mortalidade eram conhecidos se gerassem excedentes, posteriormente distribuídos (parcial ou totalmente) às pessoas seguras sob a forma de bónus. A natureza fixa e de longo prazo de muitos dos compromissos assumidos nos contratos significou que aquilo que no início se consideravam estimativas prudentes quanto à evolução dos factores de risco (financeiros e demográficos) se revele, hoje, perfeitamente desajustado da realidade e susceptível de gerar perdas técnicas consideráveis. No actual contexto, marcado pelo incremento da volatilidade nos mercados financeiros e pelo aumento da longevidade humana, a necessidade de garantir o *matching* entre os activos e as responsabilidades passa, em boa medida, pela utilização de bases técnicas que espelhem, de forma fiel, os factores de risco a que se encontram expostas as carteiras. Em particular, as companhias devem recorrer a tábuas de mortalidade dinâmicas ou prospectivas, que antecipam a evolução provável da lei de sobrevivência dos indivíduos de uma população. Neste campo, o debate centra-se em torno da selecção do método de projecção adequado de entre as múltiplas abordagens de natureza epidemiológica, explicativa ou extrapolativa adoptadas na prática.

Em Portugal, não existem tábuas de mortalidade (contemporâneas ou prospectivas) regulamentares, nem para a população portuguesa, nem para as subpopulações de pessoas cobertas por contratos de seguro ou de beneficiários de fundos de pensões. Em resultado, as companhias de seguros que operam o ramo vida são “forçadas” a recorrerem, nos seus pressupostos actuariais, a tábuas de mortalidade adoptadas noutros países (e.g., França, Suíça, Inglaterra, Espanha,...). Pese embora se trate de uma prática permitida pelas autoridades de supervisão, a utilização de uma lei de sobrevivência relativa à população de um outro país, tendencialmente desajustada das condições demográficas observadas no país e na carteira em que é aplicada, envolve naturalmente riscos de base significativos, em particular o risco de sobrestimação da taxa de mortalidade da população.

As novas disposições contidas nas normas contabilísticas IFRS e no projecto Solvên-

## 1.1. Relevância e oportunidade do tema

---

cia II reforçam a necessidade de se proceder a uma avaliação mais criteriosa das responsabilidades das empresas. Em particular, elas reforçam a importância dos modelos internos de avaliação do risco, abrindo a porta à possibilidade de as companhias de seguros adoptarem tábuas de mortalidade específicas, devidamente calibradas ao comportamento da lei de sobrevivência observado nas populações que constituem as suas carteiras, designadamente nos casos em que o actuário responsável considera que as características das populações envolvidas não se encontram devidamente espelhadas nas tábuas regulamentares. Neste contexto, é inadiável a construção de tábuas regulamentares em Portugal, quer para a população em geral, quer para as subpopulações de pessoas seguras e dos fundos de pensões.

A um outro nível, a presença do fenómeno de selecção adversa no mercado de seguros manifesta-se através de taxas de mortalidade superiores às da população em geral nos indivíduos que adquirem seguros de vida e de taxas inferiores nas carteiras cobertas por rendas vitalícias. Neste sentido, a desvalorização da importância deste risco de base pode gerar perdas técnicas importantes para as companhias, e requer uma adaptação das tábuas de mortalidade relativas à população em geral ao comportamento específico observado nas carteiras de seguros.

### 1.1.3 Risco de longevidade e risco de mortalidade

Os termos “risco de longevidade” e “risco de mortalidade” são muitas vezes usados de forma indistinta apesar de representarem, na prática, duas faces da mesma moeda. Com efeito, enquanto que a expressão risco de longevidade denota colectivamente o risco de desvios sistemáticos da mortalidade face aos valores esperados (projectados), em particular que um indivíduo ou uma população sobrevivam sistematicamente para além do esperado, o termo risco de mortalidade é usado, regra geral, para descrever o risco de que uma população viva, em termos agregados, menos do que o esperado, i.e., que a mortalidade seja superior ao previsto.

A incerteza em torno da longevidade humana pode decompor-se em diferentes fontes de risco: flutuações aleatórias nas taxas de mortalidade (risco de volatilidade), desvios sistemáticos (*trend risk*), mortalidade catastrófica (*jump risk*) e risco de base. Se é verdade que algumas destas fontes de incerteza podem ser eliminadas aumentando a dimensão das carteiras (homogéneas), apelando à lei dos grandes números (*pooling risk*), no caso do risco de longevidade a sua natureza sistemática faz com que não possa ser eliminado diversificando a carteira. Pelo contrário, o seu impacto financeiro aumenta à medida que a carteira cresce, na medida em que os desvios afectam todas as pessoas

## 1.1. Relevância e oportunidade do tema

---

seguras da mesma forma e no mesmo sentido. Em consequência, as companhias devem procurar novos mecanismos de cobertura do risco de longevidade, que envolvam soluções internas ou externas, usando técnicas clássicas no mercado de seguros ou clamando por soluções inovadoras, que passem pela titularização dos riscos ou pela criação de um novo mercado de derivados de mortalidade.

Os riscos de longevidade e de mortalidade têm uma natureza distinta da dos riscos financeiros. Com efeito, enquanto que a volatilidade dos preços dos activos financeiros tende a ser elevada, as alterações na longevidade ocorrem de forma lenta no tempo. Se a volatilidade nos mercados financeiros representa um problema imediato, as tendências na longevidade traduzem um desafio que, sendo de longo prazo, é persistente e permanente. Esta diferença, combinada com o facto de se tratar de um risco pouco conhecido nos mercados financeiros, medido e analisado de forma distinta da dos riscos financeiros, e com um impacto nas responsabilidades dos fundos de pensões ocultado durante muito tempo, fez com que a atenção dada ao risco de longevidade tenha sido menor no passado. Esperamos que ao catapultar a discussão sobre a mortalidade para um contexto dinâmico e estocástico esta dissertação contribua para atrair a atenção sobre a incerteza que subjaz à evolução da longevidade.

### 1.1.4 Modelos determinísticos vs modelos estocásticos

Uma das formas de lidar com o problema da longevidade passa pela utilização de tábuas de mortalidade prospectivas. A sua construção assenta num conjunto variado de metodologias que permite, ora em tempo discreto, ora em tempo contínuo, a implementação de uma abordagem que, sendo dinâmica é, na sua essência, determinística.

Nesta dissertação, adoptamos uma perspectiva mais ampla e consideramos que a intensidade de mortalidade pode ser modelada através de processos estocásticos. Esta abordagem, simultaneamente dinâmica e estocástica, permite-nos captar de forma mais realista duas características importantes da mortalidade: dependência temporal e incerteza sobre a trajectória futura. Esta opção comporta ainda outras vantagens importantes, desde logo o facto de permitir a derivação de prémios de seguro e reservas mais realistas face à natureza dos factores de risco (demográficos e financeiros) subjacentes. Por outro lado, ela corporiza as novas orientações emanadas pelo International Accounting Standards Board, nas quais se reconhece a necessidade de quantificar adequadamente todas as fontes de risco, incluindo os riscos não financeiros, sejam estes idiossincráticos ou não diversificáveis, e onde se advoga a contabilização das responsabilidades das companhias de seguros pelo seu valor de mercado.

## 1.2. Objectivos e principais contributos da dissertação

---

A adopção, no contexto actuarial, de uma estrutura conceptual mais rica para a intensidade de mortalidade, incorporando argumentos consolidados na avaliação de activos financeiros, em especial do princípio da ausência de oportunidades de arbitragem, permite ainda equacionar, de forma mais rigorosa, um conjunto de estratégias de cobertura do risco de longevidade, que envolvem designadamente a sua transferência para outras contrapartes. Entre estas, destacam-se a possibilidade de transferir o risco de longevidade para a pessoa segura (e.g., através de contratos com prémios e/ou benefícios indexados à evolução da intensidade de mortalidade), para o mercado (e.g., mediante a emissão de obrigações com cupões e/ou principal indexado à evolução da longevidade), as tradicionais opções de resseguro e securitização, e a adopção de estratégias de *hedging* assentes em derivados financeiros (opções, futuros, swaps, etc.) contingentes à evolução da intensidade de mortalidade.

## 1.2 Objectivos e principais contributos da dissertação

Esta dissertação procura constituir-se como uma abordagem integrada e sistemática sobre o problema da medição e gestão dos riscos de mortalidade e de longevidade em diferentes contextos, sejam estes estáticos ou dinâmicos, determinísticos ou estocásticos. Os métodos e técnicas discutidos neste documento devem, por um lado, contribuir para um melhor conhecimento sobre a natureza da exposição a este tipo de riscos de carteiras de seguros e dos fundos de pensões e, por outro, constituir-se como ferramentas de estruturação e de avaliação de estratégias de cobertura do risco de longevidade.

A dissertação inclui um conjunto amplo e diversificado de contributos originais. O documento oferece, em primeiro lugar, uma revisão crítica detalhada sobre os principais métodos paramétricos e não-paramétricos de graduação de tábuas de mortalidade contemporâneas e sobre os principais métodos de projecção da mortalidade e de construção de tábuas de mortalidade dinâmicas ou prospectivas. Nesta revisão são fornecidas pistas importantes para a implementação prática dos modelos, designadamente os algoritmos necessários à estimação dos parâmetros.

Em segundo, realizamos nesta dissertação o primeiro exercício sistemático de graduação paramétrica da mortalidade das pessoas seguras e dos beneficiários dos fundos de pensões efectuado em Portugal, tomando como referência os dados disponibilizados pelo Instituto de Seguros de Portugal para o período 2000-2004.

Em terceiro, são desenvolvidas duas variantes do modelo Poisson Lee-Carter para a projecção da mortalidade. Na primeira, é ensaiada uma nova técnica que incorpora as preocupações quanto ao alisamento das séries no procedimento de estimação dos

## 1.2. Objectivos e principais contributos da dissertação

---

parâmetros do modelo de Poisson. A solução, inspirada na técnica de P-Splines desenvolvida por Eilers e Marx (1996), envolve a estimativa dos parâmetros mediante um procedimento que envolve a maximização da função de log-verosimilhança penalizada, com a penalização a incidir sobre a diferença finita entre os coeficientes estimados do modelo. Na segunda variante do modelo Poisson Lee-Carter admitimos, explicitamente, que existem limites naturais para a longevidade humana, dos quais decorrem níveis de mortalidade mínimos abaixo dos quais se assume ser impossível descer no futuro. Em consequência, introduzimos no modelo uma restrição segundo a qual as projecções de mortalidade são balizadas por uma determinada tábua limite, que representa os limites impostos pela biologia à longevidade da espécie humana.

Em quarto, esta dissertação desenvolve pela primeira vez as bases que sustentam a adopção de uma abordagem longitudinal no estudo da mortalidade da população portuguesa. Em particular, comparamos o desempenho de um conjunto de métodos de projecção da mortalidade e elaboramos as primeiras tábuas de mortalidade prospectivas para a população portuguesa. Estas tornam possível, quer o estabelecimento de bases técnicas adequadas a uma correcta quantificação e controlo do risco de longevidade por parte das companhias de seguros, quer a avaliação da sustentabilidade financeira dos sistemas de segurança social.

Em quinto, investigamos pela primeira vez a importância do efeito selecção adversa no contexto do problema de projecção da mortalidade, considerando a carteira de beneficiários de fundos de pensões em Portugal. A adopção de uma metodologia assente num modelo relacional, incorporado num modelo linear generalizado com distribuição de Poisson, permitiu a derivação pela primeira vez de tábuas de mortalidade prospectivas para a população de beneficiários de fundos de pensões.

Em sexto, desenvolvemos uma abordagem simultaneamente dinâmica e estocástica do risco de longevidade admitindo que a intensidade de mortalidade pode ser modelada através de processos estocásticos. É estabelecido um paralelo entre o conceito de intensidade de incumprimento derivado na literatura sobre risco de crédito e o conceito de intensidade de mortalidade. Esta analogia permite-nos equacionar a adaptação de um conjunto de modelos do tipo afim propostos na literatura sobre taxas de juro na modelação da intensidade de mortalidade e a obtenção de soluções fechadas para a probabilidade de sobrevivência. Para testar a capacidade destes modelos para descrever adequadamente a trajectória da intensidade de mortalidade, conduzimos igualmente um primeiro ensaio de calibragem dos modelos considerando a realidade representada pelas tábuas de mortalidade prospectivas derivadas para a população portuguesa.

Em sétimo, exploramos uma nova abordagem teórica ao problema da modelação

da mortalidade num contexto estocástico onde consideramos, de forma simultânea, a evolução da intensidade em todas as idades que integram o arco da vida humana. Esta formulação assenta no princípio simples de que a intensidade de mortalidade pode ser representada por uma função do tipo afim de um conjunto de factores latentes, com dinâmica ditada por equações diferenciais estocásticas com saltos. São analisados três casos particulares, onde se catapulta para um contexto estocástico algumas das mais conhecidas leis de mortalidade determinísticas.

Em oitavo, analisamos a aplicação de modelos de mortalidade estocástica na avaliação de contratos de seguro com cash flows contingentes à sobrevivência ou morte do indivíduo e de produtos estruturados, dedicando especial atenção ao problema da determinação do preço de mercado do risco de longevidade. Neste ponto, desenvolvemos um novo método de estimação do preço do risco, recorrendo a um princípio clássico de determinação do prémio em contratos de seguro (princípio do desvio-padrão). Esta solução envolve a modelação da intensidade de mortalidade por processos estocásticos do tipo afim e a derivação de probabilidades de sobrevivência ajustadas pela inclusão de um prémio pelo risco de longevidade.

Por último, esta dissertação inclui uma discussão detalhada sobre as estratégias e soluções que particulares e empresas podem adoptar para gerir a sua exposição ao risco de longevidade. As estratégias analisadas incluem não apenas soluções de cobertura interna mas também mecanismos de transferência do risco para outras contrapartes (beneficiários das apólices, resseguradores e mercado de capitais). Neste último caso, equacionamos a aplicação de um conjunto de tratados de resseguro e o desenvolvimento de um mercado de activos financeiros com cash flows contingentes à evolução da longevidade/mortalidade, incluindo a oferta de derivados de mortalidade.

### **1.3 Organização e estrutura do documento**

Este documento está dividido em três partes e doze capítulos no total. Na Parte I introduzimos os principais conceitos e a formalização matemática adoptada ao longo de toda a dissertação (Capítulo 2) e efectuamos uma revisão crítica detalhada sobre os principais métodos paramétricos (Capítulo 3) e não-paramétricos (Capítulo 4) de graduação de tábuas de mortalidade contemporâneas. Esta parte inclui ainda, no Capítulo 5, a metodologia e os resultados dos testes de graduação paramétrica da mortalidade das pessoas seguras e dos beneficiários dos fundos de pensões em Portugal.

A Parte II compreende uma extensa análise dos diferentes modelos de projecção da mortalidade e de construção de tábuas prospectivas, incluindo a projecção da morta-

### 1.3. Organização e estrutura do documento

---

lidade nas idades avançadas (Capítulo 6). No Capítulo 7 comparamos o desempenho dos diferentes métodos de projecção da mortalidade e elaboramos as primeiras tábuas prospectivas para a população portuguesa. No Capítulo 8 investigamos a importância do efeito selecção adversa no contexto do problema de projecção da mortalidade e desenvolvemos tábuas prospectivas para a população de beneficiários de fundos de pensões.

Na Parte III consideramos a aplicação de processos estocásticos na modelação da intensidade de mortalidade e discutimos um conjunto de aplicações actuariais e de cobertura do risco de longevidade. No Capítulo 9 descrevemos o enquadramento teórico adoptado para modelar a mortalidade num contexto estocástico e equacionamos a adopção de processos do tipo afim com saltos na derivação de soluções analíticas para a probabilidade de sobrevivência. No Capítulo 10 testamos o desempenho empírico destes modelos na realidade representada pelas tábuas de mortalidade prospectivas derivadas para a população portuguesa. No Capítulo 11 ilustramos a aplicação de modelos de mortalidade estocástica no contexto actuarial, nomeadamente na avaliação de apólices tradicionais do ramo vida e de produtos estruturados. Neste capítulo, são ainda analisadas um conjunto de estratégias internas e externas de cobertura do risco de longevidade. Por fim, o Capítulo 12 resume as principais conclusões da dissertação.

## Capítulo 2

# Modelos Probabilísticos para a Descrição da Duração da Vida Humana

*“Most of us retain, consciously or unconsciously, a feeling that, underlying all the roughnesses in our data referable to errors of observation and an ever-changing environment, there may be an inherent mathematical system of law and order, which if it could be discovered would give such insight into the meaning of unadjusted figures...” (Perks, 1932)*

### 2.1 Introdução

No estado actual do conhecimento científico, o momento em que a vida humana cessa a sua existência física é certamente uma informação desconhecida. Apesar disso, é possível avaliar o risco de morte numa dada população e intervalo de tempo. Neste sentido, é natural o recurso ao formalismo matemático do cálculo de probabilidades para descrever a duração de vida dos indivíduos. O tempo que medeia entre a idade do indivíduo e o momento da sua morte é modelado como uma variável aleatória, à qual estão associadas, como é habitual, um conjunto de funções como sejam a função de densidade, a função de distribuição, a função de sobrevivência, entre outras.

Igualmente comum no contexto actuarial e demográfico é o recurso ao modelo probabilístico em tempo discreto, materializado nas denominadas tábuas de mortalidade ou de sobrevivência, um modelo tabular que sintetiza um conjunto de funções básicas que permitem analisar, numa determinada população, o fenómeno da longevidade e efectuar

juízos probabilísticos sobre a evolução da mortalidade. Num e noutro modelo é possível calcular diferentes indicadores de longevidade como sejam a esperança de vida, a vida mediana, a entropia, entre outros.

A caracterização da função de sobrevivência e/ou do padrão de mortalidade registados numa população pode, em alternativa, passar pela utilização de funções paramétricas que exprimam o comportamento do fenómeno em termos da idade e de um conjunto limitado de parâmetros, estimados com base nas estatísticas demográficas disponíveis.

Neste capítulo, introduzimos os principais conceitos e a formalização matemática adoptada ao longo de toda a dissertação. Estes elementos serão fundamentais na exposição dos diferentes modelos analisados e na avaliação dos resultados dos estudos empíricos realizados. Na Secção 2.2 são definidos os principais conceitos usados para caracterizar a duração de vida remanescente de um indivíduo e discutidas as principais hipóteses usadas para repartir os óbitos nas idades fraccionadas. Na Secção 2.3 são apresentados os principais indicadores adoptados na literatura demográfica e actuarial para resumir as características da distribuição de probabilidade da vida restante. Na Secção 2.4, é introduzido e discutido o conceito de tábua de mortalidade enquanto modelo tabular de análise de sobrevivência. Por fim, na Secção 2.5 é apresentada uma síntese crítica relativamente exaustiva das principais leis de mortalidade desenvolvidas desde o século XVIII até ao presente para parametrizar o comportamento da longevidade humana. Por fim, na Secção 2.6 discutimos brevemente o problema da heterogeneidade na população na modelação da mortalidade.

## 2.2 Análise de sobrevivência

### 2.2.1 Duração de vida restante

Designemos por  $T$  a variável aleatória positiva que representa a duração de vida de um indivíduo pertencente a uma dada população de referência. Definamos de seguida uma sequência de variáveis aleatórias  $\{T_x : x = 0, 1, 2, \dots, \omega\}$ , onde  $T_x$  simboliza a vida residual de um indivíduo que atinge a idade  $x$ , i.e.,

$$\mathbb{P}(T_x > t) = \mathbb{P}[T > x + t | T > x] \quad (2.2.1)$$

e onde  $\omega$  tem o significado de “idade extrema”.<sup>1</sup>

---

<sup>1</sup>Por idade extrema (ou máxima) entende-se a idade em que não é admitida a hipótese de sobrevivência do indivíduo. No limite, se atribuirmos ao evento  $\{T \geq \omega\}$  uma probabilidade nula ou muito reduzida, podemos escrever  $\omega = \infty$ .

## 2.2. Análise de sobrevivência

---

Daqui resulta que um sobrevivente à idade  $x$  falecerá à idade  $x + T_x$ . A probabilidade de um indivíduo viver  $t$  anos aumenta com a idade por este alcançada, i.e., para duas idades  $x_2 > x_1$  podemos escrever

$$\mathbb{P}(x_2 + T_{x_2} > t) \geq \mathbb{P}(x_1 + T_{x_1} > t), \quad \forall t > x_1. \quad (2.2.2)$$

Analisemos o caso particular de um recém-nascido (i.e., com idade  $x = 0$ ). A função de distribuição de probabilidade (f.d.) da sua vida residual  $T_0$ ,  $F_0(\cdot) : \mathbb{R}_+ \mapsto [0, 1]$ , vem dada por

$$F_0(x) = \mathbb{P}(T_0 \leq x), \quad x \geq 0. \quad (2.2.3)$$

A função  $F_0(x)$  representa a probabilidade de um indivíduo recém-nascido perecer antes de atingir os  $x$  anos. Na ciência actuarial (e na demografia), é habitual o recurso a uma função complementar de  $F_0(\cdot)$  para especificar a distribuição de  $T_0$ . Essa função, designada por *função de sobrevivência* (FS),  $S(\cdot) : \mathbb{R}_+ \mapsto [0, 1]$ , exprime a probabilidade de um indivíduo recém-nascido sobreviver pelo menos  $x$  anos. É definida da seguinte forma

$$S(x) = \mathbb{P}(T_0 > x) = 1 - \mathbb{P}(T_0 \leq x) = 1 - F_0(x). \quad (2.2.4)$$

A função  $x \mapsto S(\cdot)$  é decrescente (i.e.,  $S(x_1) \geq S(x_2)$ ,  $x_1 < x_2$ ) e satisfaz as habituais condições  $S(0) = 1$  e  $\lim_{x \rightarrow +\infty} S(x) = 0$ . Tal é assim porquanto se assume que  $F_0(0) = 0$  e que  $\lim_{x \rightarrow +\infty} F_0(x) = 1$ . Vamos admitir que a distribuição de probabilidade de  $T_0$  é contínua e dotada de função de densidade de probabilidade (f.d.p.)  $f_0(\cdot)$  (i.e.  $F_0(t) = \int_0^t f_0(u) du$  para  $t \geq 0$ ). Neste caso, podemos escrever

$$f_0(x) = \frac{dF_0(x)}{dx} = -\frac{dS(x)}{dx}. \quad (2.2.5)$$

O gráfico da função de densidade de probabilidade  $x \mapsto f_0(\cdot)$  é denominado de curva de óbitos (*curve of deaths*). Usando  $F_0(x)$  ou  $S(x)$  é possível efectuar juízos probabilísticos sobre a idade de falecimento do indivíduo. Por exemplo, a probabilidade de um recém-nascido perecer entre as idades  $x$  e  $z$  ( $x < z$ ) é dada por

$$\mathbb{P}\{x < T_0 \leq z\} = F_0(z) - F_0(x) = S(x) - S(z). \quad (2.2.6)$$

Consideremos agora, em termos mais gerais, um indivíduo de idade  $x$  e vida residual aleatória  $\{T_x : x \in \Lambda\}$ , onde  $\Lambda$  denota o conjunto de valores possíveis para a variável idade. Uma família particular é aquela em que  $\{T_x : x \in (0, \omega)\}$ . Por definição,  $T_x = (T - x) | T > x$ , pelo que a função de probabilidade acumulada de  $T_x$ ,  $F_x(t)$ , pode ser

definida por

$$\begin{aligned}
 F_x(t) &= \mathbb{P}(T_x \leq t) = \mathbb{P}\{T \leq x+t | T > x\} \\
 &= \frac{\mathbb{P}(x < T \leq x+t)}{\mathbb{P}(T > x)} = \frac{F_0(x+t) - F_0(x)}{1 - F_0(x)} \\
 &= \frac{S(x) - S(x+t)}{S(x)} = 1 - \frac{S(x+t)}{S(x)}, \quad t \geq 0,
 \end{aligned} \tag{2.2.7}$$

onde  $S(\cdot)$  representa a função de sobrevivência definida em (2.2.4). Constata-se assim que, a partir do conhecimento da função de distribuição de  $T$ , é possível determinar a função de distribuição de  $T_x$  para qualquer idade  $x > 0$ . Verifica-se igualmente que, no caso de  $T_x$  admitir função de densidade  $f_x(\cdot)$  (i.e.  $F_x(t) = \int_0^t f_x(u)du$  para  $t \geq 0$ ), esta é definida por<sup>2</sup>

$$f_x(t) = \frac{f_0(x+t)}{S(x)}. \tag{2.2.8}$$

Combinando (2.2.5) e (2.2.7), podemos escrever

$$f_x(t) = \frac{dF_x(t)}{dt} = -\frac{1}{S(x)} \frac{d}{dt} S(x+t) \tag{2.2.9}$$

para  $t \geq 0$ .

### 2.2.2 Quocientes de mortalidade

A comunidade internacional de actuários recorre a uma notação própria para caracterizar as variáveis de interesse acima apresentadas. Por exemplo, as probabilidades de sobrevivência  ${}_t p_x$  ( $t \geq 0$ ) e de falecimento  ${}_t q_x$  ( $t \geq 0$ ) são definidas, respectivamente, por<sup>3</sup>

$${}_t p_x = \mathbb{P}(T_x > t) = \mathbb{P}[T > x+t | T > x] = \frac{\mathbb{P}(T > x+t)}{\mathbb{P}(T > x)} \tag{2.2.10}$$

e por

$${}_t q_x = \mathbb{P}(T_x \leq t) = \mathbb{P}[T \leq x+t | T > x] = 1 - {}_t p_x. \tag{2.2.11}$$

O termo  ${}_t p_x$  define a probabilidade de um indivíduo vivo à idade  $x$  sobreviver até à idade  $x+t$ , enquanto que  ${}_t q_x$  determina a probabilidade de um indivíduo que atinge a

---

<sup>2</sup>Esta conclusão é facilmente verificável se notarmos que

$$F_x(t) = \frac{F_0(x+t) - F_0(x)}{1 - F_0(x)} = \int_x^{x+t} \frac{f_0(u)}{1 - F_0(x)} du = \int_0^t \frac{f_0(x+u)}{S(x)} du.$$

<sup>3</sup>Para uma introdução mais detalhada à análise de sobrevivência no contexto actuarial veja-se, por exemplo, Gerber (1995), Bowers *et al.* (1997), Pitacco (2000) e Promislow (2006).

## 2.2. Análise de sobrevivência

---

idade  $x$  falecer antes de alcançar a idade  $x + t$ . O termo  ${}_tq_x$  é igualmente conhecido por quociente de mortalidade. Uma definição alternativa de  ${}_tp_x$  e  ${}_tq_x$ , baseada na função de sobrevivência, é dada por<sup>4</sup>

$${}_tp_x = \frac{S(x+t)}{S(x)} \quad (2.2.12)$$

$${}_tq_x = 1 - \frac{S(x+t)}{S(x)}. \quad (2.2.13)$$

Os actuários usam frequentemente uma notação abreviada considerando o valor de  $t = 1$  (ano), omitindo nesse caso o índice  $t$ , i.e., pondo  ${}_1p_x \equiv p_x$  e  ${}_1q_x \equiv q_x$ . Note-se igualmente que as relações  ${}_0q_x = 0$  e  ${}_0p_x = 1$  são respeitadas.<sup>5</sup> Conhecida a sequência  $\{q_x : x = 1, 2, \dots\}$  (ou  $\{p_x\}$ ), é possível reconstruir a função de sobrevivência para idades  $x$  inteiras. Com efeito, se reformularmos a equação (2.2.13) obtemos, recapitulando a condição inicial  $S(0) = 1$ ,

$$S(x+1) = S(x)(1 - q_x), \quad x = 0, 1, 2, \dots \quad (2.2.14)$$

A probabilidade de um indivíduo com idade  $x$  sobreviver  $t$  anos e perecer nos seguintes  $s$  anos, conhecida na literatura por probabilidade diferida (ou *forward*), e designada por  ${}_{t|s}q_x$ , é definida da seguinte forma

$$\begin{aligned} {}_{t|s}q_x &= \mathbb{P}(t < T_x \leq t+s) = \mathbb{P}[x+t < T \leq x+t+s | T > x] \\ &= \frac{S(x+t) - S(x+t+s)}{S(x)} = {}_{t+s}q_x - {}_tq_x = {}_t p_x - {}_{t+s} p_x \\ &= ({}_t p_x) \times ({}_s q_{x+t}). \end{aligned} \quad (2.2.15)$$

No cálculo actuarial, existem ainda outras variáveis de interesse para a análise da longevidade do indivíduo. Uma delas, conhecida na literatura por *curtate-future-lifetime*,

<sup>4</sup>Refira-se ainda que as probabilidade de morte e de sobrevivência acima definidas podem ser expressas em termos da função de densidade de probabilidade  $f_x(t)$ , i.e., mediante

$${}_tp_x = \int_t^{+\infty} f_x(u) du \quad \text{e} \quad {}_tq_x = \int_0^t f_x(u) du.$$

<sup>5</sup>Para dois momentos  $t_1, t_2 > 0$  quaisquer, podemos igualmente escrever a relação

$${}_{t_1+t_2}p_x = {}_{t_1}p_x \cdot {}_{t_2}p_{x+t_1},$$

que exprime o facto de que, para sobreviver  $t_1 + t_2$  anos suplementares, é preciso, em primeiro lugar, sobreviver até à idade  $x + t_1$ , e depois mais  $t_2$  anos.

## 2.2. Análise de sobrevivência

---

$K_x$ , respeita ao número de anos inteiros futuros vividos por um indivíduo com idade  $x$  antes de falecer. Como  $K_x$  corresponde ao maior valor inteiro em  $T_x$ , a sua distribuição de probabilidade pode ser escrita da seguinte forma<sup>6</sup>

$$\begin{aligned}\mathbb{P}[K_x = k] &= \mathbb{P}(k \leq T_x < k + 1) \\ &= {}_k p_x - {}_{k+1} p_x \\ &= {}_k p_x q_{x+k} = {}_k | q_x \quad k = 0, 1, 2, \dots\end{aligned}\tag{2.2.16}$$

### 2.2.3 Taxa de mortalidade instantânea

#### Definição

A taxa de mortalidade instantânea à idade  $x + t$  (também conhecida por força de mortalidade), representada por  $\mu_{x+t}$ , é definida pelo seguinte limite

$$\mu_{x+t} = \lim_{\Delta t \rightarrow 0^+} \frac{\mathbb{P}[t < T_x \leq t + \Delta t | T_x > t]}{\Delta t}.\tag{2.2.17}$$

De forma intuitiva, podemos interpretar  $\mu_{x+t}$  como a probabilidade de um indivíduo que sobrevive até à idade  $x + t$  falecer antes de alcançar a idade  $x + t + \Delta t$ , i.e., no intervalo  $[x + t, x + t + \Delta t)$ .

#### Relação com as probabilidades de sobrevivência e óbito

Se recordarmos que

$$\mathbb{P}[t < T_x \leq t + \Delta t | T_x > t] = \frac{\mathbb{P}[t < T_x \leq t + \Delta t]}{\mathbb{P}[T_x > t]} = \frac{{}_{t+\Delta t} q_x - {}_t q_x}{{}_t p_x}$$

é possível estabelecer a seguinte ligação entre  $\mu_{x+t}$  e as probabilidades de sobrevivência  ${}_t p_x$  e de óbito  ${}_t q_x$

$$\mu_{x+t} = \lim_{\Delta t \rightarrow 0^+} \frac{{}_{t+\Delta t} q_x - {}_t q_x}{{}_t p_x \Delta t} = \frac{1}{{}_t p_x} \frac{d}{{}dt} {}_t q_x.\tag{2.2.18}$$

Com base numa expansão em série de Taylor de primeira ordem e na equação (2.2.18), podemos escrever as seguintes relações

$$\Delta t q_x \approx {}_0 q_x + [{}_t p_x \cdot \mu_{x+t}]_{t=0} \Delta t + o(\Delta t) = \mu_x \Delta t + o(\Delta t)\tag{2.2.19}$$

---

<sup>6</sup>A troca de desigualdades em (2.2.16) é possível na medida em que a hipótese de que  $T_x$  é uma variável aleatória contínua garante que  $\mathbb{P}[T_x = k] = \mathbb{P}[T_x = k + 1] = 0$ . Refira-se ainda que a equação (2.2.16) é um caso particular de (2.2.15) em que  $s = 1$  e  $k$  assume valores inteiros não negativos.

## 2.2. Análise de sobrevivência

---

e

$$\Delta t p_x = {}_0 p_x - [{}_t p_x \cdot \mu_{x+t}]_{t=0} \Delta t + o(\Delta t) = 1 - \mu_x \Delta t + o(\Delta t),$$

onde  $o(\Delta t)$  é uma quantidade tal que  $\lim_{\Delta \rightarrow 0} [o(\Delta t) / \Delta] = 0$ . De forma intuitiva, podemos considerar  $o(\Delta t)$  negligenciável, pelo que as aproximações

$$\begin{aligned} \Delta t q_x &\simeq \mu_x \Delta t \\ \Delta t p_x &\simeq 1 - \mu_x \Delta t \end{aligned}$$

são válidas em intervalos  $\Delta t$  suficientemente pequenos. Em suma, as taxas instantâneas de mortalidade governam (numa aproximação de primeira ordem) as probabilidades  ${}_t p_x$  e  ${}_t q_x$  em intervalos de tempo reduzidos.

Se admitirmos que  $T_x$  é uma variável aleatória contínua, é possível deduzir a sua função de densidade de probabilidade a partir de (2.2.18) segundo

$$\frac{d}{dt} {}_t q_x = {}_t p_x \cdot \mu_{x+t}. \quad (2.2.20)$$

A relação (2.2.20) pode ser interpretada como uma equação diferencial. Recordando a condição inicial  ${}_0 q_x = 0$ , da resolução de (2.2.20) deduz-se a seguinte representação para  ${}_t q_x$ <sup>7</sup>

$${}_t q_x = \int_0^t {}_\tau p_x \mu_{x+\tau} d\tau.$$

Raciocínio análogo pode ser usado para derivar as probabilidades de sobrevivência em função das taxas de mortalidade instantânea. De facto, partindo de (2.2.18) podemos escrever

$$\frac{d}{dt} \ln({}_t p_x) = -\mu_{x+t}.$$

Se considerarmos novamente esta relação como uma equação diferencial da função  $t \mapsto {}_t p_x$ , com  $x$  fixo, e recapitularmos a condição inicial,  ${}_0 p_x = 1$ , obtemos a seguinte

---

<sup>7</sup>Também com base na equação (2.2.9) e na relação

$$\mu_x = -\frac{dS(x)/dx}{S(x)} = -\frac{d}{dx} \ln S(x),$$

podemos ainda escrever a função de densidade de probabilidade de  $T_x$  em termos da força de mortalidade

$$\begin{aligned} f_x(t) &= -\frac{1}{S(x)} \frac{d}{dt} S(x+t) = -\frac{S(x+t)}{S(x)} \left[ \frac{1}{S(x+t)} \frac{d}{dt} S(x+t) \right] \\ &= {}_t p_x \cdot \mu_{x+t}, \quad t \geq 0, \text{ com } \int_0^\infty {}_\tau p_x \mu_{x+\tau} d\tau = 1. \end{aligned}$$

solução

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+\tau} d\tau\right). \quad (2.2.21)$$

### Relação com o conceito de taxa de mortalidade

A taxa de mortalidade instantânea pode ser deduzida a partir do conceito clássico de taxa de mortalidade. A taxa de mortalidade no intervalo  $[x, x+t)$ , designada por  ${}_t m_x$ , é definida por

$${}_t m_x = \frac{{}_t d_x}{{}_t E_x}, \quad (2.2.22)$$

onde  ${}_t d_x$  e  ${}_t E_x$  denotam, respectivamente, o número de óbitos registado e a respectiva exposição ao risco. A taxa de mortalidade distingue-se do quociente de mortalidade pelo facto do número de óbitos se reportar a uma exposição ao risco média, e não ao efectivo inicial. A taxa  ${}_t m_x$  exprime o número de óbitos por pessoa e por ano, sendo por vezes denominada de taxa central de mortalidade.

Dito isto, partindo de (2.2.22) podemos escrever

$$\Delta t m_x = \frac{\Delta t d_x}{\Delta t E_x} = \frac{l_x - l_{x+\Delta t}}{\int_0^{\Delta t} l_{x+\tau} d\tau}$$

onde  $l_x$  simboliza o número de sobreviventes de uma geração inicial  $l_0$ .

Em termos assintóticos, demonstra-se finalmente que<sup>8</sup>

$$\lim_{\Delta t \rightarrow 0^+} \Delta t m_x = -\frac{\frac{d}{dx} l_x}{l_x} = \mu_x. \quad (2.2.23)$$

### 2.2.4 Hipóteses para idades fraccionadas

A prática actuarial requer, frequentemente, o cálculo de probabilidades de sobrevivência e/ou de morte para idades e/ou intervalos de tempo não inteiros, por exemplo  ${}_u p_x$ ,  ${}_{1-u} p_{x+u}$ , com  $x$  inteiro e  $0 \leq u \leq 1$ . Na maioria dos casos, os dados de que dispõem os actuários são fruto de um agrupamento por período temporal. Esta limitação impõe,

---

<sup>8</sup>Na derivação deste resultado, note que

$$\lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \int_{\tau=0}^{\Delta t} l_{x+\tau} d\tau = l_x$$

e que

$$\frac{d}{dx} l_x = \lim_{\Delta x \rightarrow 0} \frac{l_{x+\Delta x} - l_x}{\Delta x} = l_x \lim_{\Delta x \rightarrow 0} \frac{\Delta x p_x - 1}{\Delta x} = l_x \lim_{\Delta x \rightarrow 0} \frac{1 - \mu_x \Delta x - 1}{\Delta x} = -l_x \mu_x.$$

## 2.2. Análise de sobrevivência

---

regularmente, o recurso a modelos probabilísticos em tempo discreto, em particular a tábuas de mortalidade tabuladas apenas para idades inteiras.

Na ausência de informação sobre a repartição dos óbitos e outras variáveis de interesse no intervalo de tempo considerado, o actuário recorre a premissas de trabalho de modo a prolongar a aplicação do modelo probabilístico a idades fraccionadas. Uma solução alternativa, que analisaremos mais adiante neste capítulo, envolve a adopção de um modelo paramétrico ou lei de mortalidade. Resumimos, nesta secção, as principais hipóteses avançadas na literatura para resolver este problema.<sup>9</sup>

### Hipótese de repartição uniforme dos óbitos durante o ano

Numa das primeiras tentativas para descrever a evolução da mortalidade mediante funções paramétricas, Abraham de Moivre propôs, em meados do século XVIII, que a função de sobrevivência ( $l_x$ ) de uma tábua de mortalidade fosse representada por uma linha recta. De acordo com esta hipótese, o número de sobreviventes de uma determinada geração (*cohort*) inicial diminui de forma linear e uniforme todos os anos, até à sua extinção total. Por outras palavras, segundo esta hipótese todos os anos registar-se-ia o mesmo número de óbitos, um pressuposto que é rejeitado pela generalidade dos estudos empíricos.

No entanto, uma variante desta hipótese, segundo a qual o número de sobreviventes entre duas idades inteiras consecutivas diminui linear e uniformemente durante o ano, é hoje largamente adoptada pelos actuários. Segundo esta formulação, admite-se que, em média, os óbitos são registados a meio do ano. Em termos formais, a hipótese estabelece que a função  ${}_uq_x = {}_uq_x$  é linear no intervalo  $u \in [0, 1)$ , i.e.,

$${}_uq_x = a + bu, \quad u \in [0, 1). \quad (2.2.24)$$

Se avaliarmos a função em  $u = 1$  ( ${}_1q_x = a + b = q_x$ ) e  $u = 0$  ( ${}_0q_x = a = 0$ ) e resolvermos o sistema de equações obtém-se  $(a, b) = (0, q_x)$ . Substituindo em (2.2.24), deduz-se finalmente que

$${}_uq_x = u \cdot q_x, \quad u \in [0, 1). \quad (2.2.25)$$

Com base na hipótese (2.2.25), prova-se-se que (Batten, 1978)

$$\mu_{x+u} = \frac{q_x}{1 - u \cdot q_x}, \quad u \in [0, 1)$$

---

<sup>9</sup>Para uma análise detalhada deste tema recomendamos a leitura de, por exemplo, Batten (1978), Elandt-Johnson e Johnson (1980), Hürlimann (1990), Gerber (1995), Bowers *et al.* (1997) e Pitacco (2000).

## 2.2. Análise de sobrevivência

---

e que

$$\begin{aligned}\mu_x &= q_x, \\ \mu_{x+1} &= \frac{q_x}{1 - q_x} > q_x, \\ l_{x+u} &= l_x - u \cdot d_x, \quad u \in [0, 1), \text{ com } d_x = {}_1d_x \\ {}_{1-s}q_{x+u} &= \frac{(1-s)q_x}{1 - u \cdot q_x} = (1-s)\mu_{x+u}, \quad (0 \leq u \leq s \leq 1) \\ {}_up_x \cdot \mu_{x+u} &= q_x \quad u \in [0, 1).\end{aligned}$$

Entre os resultados acima derivados, destacamos o facto de  $\mu_{x+u}$  se constituir como uma função (hipérbole) crescente de  $u$ , um comportamento plausível em termos empíricos se exceptuarmos o extremo inferior da curva de mortalidade. Realce ainda para o facto da probabilidade de óbito em períodos inferiores a um ano ser função linear da duração.

Por fim, e para  $u \in [0, 1)$ , é possível deduzir as seguintes relações equivalentes

$${}_um_x = \frac{2uq_x}{u(2 - uq_x)} \quad (2.2.26)$$

$${}_uq_x = \frac{2u{}_um_x}{2 + u{}_um_x}. \quad (2.2.27)$$

### Hipótese de Balducci

A hipótese sugerida pelo actuário italiano Gaetano Balducci (dita de *interpolação harmónica*) é a de que a função  $g(u) = {}_{1-u}q_{x+u}$  é linear no intervalo  $u \in [0, 1)$ , i.e.,

$${}_{1-u}q_{x+u} = a + bu, \quad u \in [0, 1) \quad (2.2.28)$$

Avaliando a função em  $u = 0$  ( ${}_1q_x = a = q_x$ ) e  $u = 1$  ( ${}_0q_{x+1} = a + b = 0$ ) e resolvendo o sistema de equações, obtém-se  $(a, b) = (q_x, -q_x)$ . Substituindo em (2.2.28), deduz-se

$${}_{1-u}q_{x+u} = (1 - u)q_x, \quad u \in [0, 1) \quad (2.2.29)$$

Da aplicação do pressuposto de Balducci, decorre que

$$\begin{aligned}\mu_{x+u} &= \frac{q_x}{1 - (1 - u)q_x}, \quad u \in [0, 1) \\ \mu_{x+1} &= q_x, \\ l_{x+u} &= \frac{l_x \cdot l_{x+1}}{l_{x+1} + u \cdot d_x}, \quad u \in [0, 1)\end{aligned}$$

e que

$$\begin{aligned} {}_{1-s}q_{x+u} &= \frac{(1-s)q_x}{1-(s-u)q_x}, & (0 \leq u \leq s \leq 1) \\ {}_uq_x &= \frac{u \cdot q_x}{1-(1-u)q_x}, & u \in [0, 1) \\ {}_{1-u}p_{x+u} \cdot \mu_{x+u} &= q_x & u \in [0, 1). \end{aligned}$$

Como se observa, a hipótese de Balducci impõe um comportamento para a função  $\mu_{x+u}$  oposto ao anteriormente derivado a partir da hipótese de distribuição uniforme dos óbitos. Com efeito, segundo Balducci a representação de  $\mu_{x+u}$  configura uma hipérbole decrescente no intervalo  $u \in [0, 1)$ , ou seja, produz taxas de mortalidade instantâneas decrescentes entre duas idades consecutivas, um resultado que apenas se observa nas idades infantis e que é, por esta razão, evidentemente criticável.

### Hipótese de força de mortalidade constante

Uma outra hipótese, alternativa às anteriores, consiste em supor que a taxa de mortalidade instantânea é constante em cada intervalo de tempo unitário, i.e., para qualquer idade inteira se observa

$$\mu_{x+u} = \mu_x, \quad u \in [0, 1). \quad (2.2.30)$$

Esta hipótese equivale a admitir, para  $u \in [0, 1)$ , que

$$\begin{aligned} {}_uq_x &= 1 - \exp(-\mu_x \cdot u), \\ {}_{1-u}q_{x+u} &= 1 - \exp[-\mu_x(1-u)], \\ \ln_u p_x &: \text{ é linear em } u, \\ l_{x+u} &= l_x \exp(-\mu_x \cdot u), \\ {}_um_x &= -\ln(1-q_x) = \mu_x. \end{aligned}$$

Assinale-se, em particular, a igualdade entre os conceitos de taxa de mortalidade e de taxa de mortalidade instantânea. Tal permite estimar  $\mu_x$  com base no ratio entre o número de óbitos observado e a correspondente exposição ao risco.

### Comparação das hipóteses

A opção por uma das hipóteses (2.2.25), (2.2.29) ou (2.2.30) não é neutra do ponto de vista actuarial, como se evidencia em seguida. Designemos por  $T_x^A$ ,  $T_x^B$  e  $T_x^C$  a duração de vida restante de um indivíduo com idade  $x$  calculada com base nas hipóteses (2.2.25),

## 2.2. Análise de sobrevivência

---

(2.2.29) e (2.2.30), respectivamente. Seja  $K_x$  uma variável que mede o número de anos inteiros que um indivíduo sobrevive antes de alcançar a idade  $x$ , i.e.,

$$K_x = \lfloor T_x \rfloor = \max \{n \text{ inteiro} \mid T_x \geq n\}.$$

Definamos igualmente por  $U_x$  a fração do ano que esse indivíduo viverá para além dos  $K_x$  anos, i.e.,  $U_x = T_x - K_x$ . Neste caso, para qualquer idade inteira  $x$ , as expressões das funções de sobrevivência associadas às três variáveis aleatórias são dadas por

$$\begin{aligned} \mathbb{P} [T_x^A > t] &= {}_k p_x (1 - u q_{x+k}) \\ \mathbb{P} [T_x^B > t] &= {}_k p_x \left( \frac{1 - q_{x+k}}{1 - (1-u) q_{x+k}} \right) \\ \mathbb{P} [T_x^C > t] &= {}_k p_x (1 - q_{x+k})^u. \end{aligned}$$

Demonstra-se que, para qualquer idade inteira  $x$  e  $t \geq 0$ , a seguinte relação existe

$$\mathbb{P} [T_x^A > t] \geq \mathbb{P} [T_x^C > t] \geq \mathbb{P} [T_x^B > t].$$

Como se observa, a hipótese de repartição uniforme dos óbitos é aquela que proporciona a maior duração de vida restante, enquanto que a hipótese de Balducci produz o menor valor. Tal significa, por exemplo, que nos contratos de seguros em caso de vida a utilização da hipótese (2.2.25) produzirá um prémio único mais elevado. Pelo contrário, nos seguros de vida em caso de morte a hipótese de Balducci será aquela que gerará um prémio puro mais elevado, uma vez que é neste caso que se estima que as pessoas seguras sobreviverão, em média, menos tempo.

### 2.2.5 Pressupostos do modelo probabilístico

O modelo probabilístico acima formulado incorpora três princípios fundamentais que importa definir de forma clara. São eles o princípio da homogeneidade, o princípio da independência e o princípio da estacionariedade.

O princípio da homogeneidade pressupõe que todos os indivíduos que integram o grupo sobre o qual se analisa o fenómeno da sobrevivência são equivalentes no que se refere à mortalidade, ou seja, têm a mesma função de distribuição de probabilidade para a variável idade de morte  $\xi$ . Tal equivale a afirmar que o grupo é homogéneo e que as probabilidades de morte e de sobrevivência dependem unicamente da idade.

Dizem-se equivalentes, na medida em que têm a mesma função de distribuição de probabilidade para a variável  $\xi$ . Refira-se, contudo, que na prática é comum a diferen-

## 2.3. Indicadores característicos

---

ciação das probabilidades de acordo com outros factores, que não apenas a idade (e.g. sexo, profissão, tipo de seguro, estado de saúde ou a pertença a determinados grupos de risco).<sup>10</sup>

O princípio da independência estabelece que os indivíduos que integram o grupo em análise constituem variáveis estocasticamente independentes. Este princípio estipula que a probabilidade de um determinado indivíduo sobreviver até uma determinada idade não depende da sobrevivência de qualquer outro elemento no grupo. O alcance desta hipótese traduz-se na inexistência de qualquer tipo de interacção ou “contágio” entre os elementos que integram uma mesma população.

Segundo o princípio da estacionariedade, a probabilidade de um indivíduo não sobreviver a uma idade concreta é independente do seu ano de cálculo. Esta hipótese é muito forte na medida em que, como veremos nos Capítulos 6 e 7, as probabilidades de sobrevivência oscilam, não apenas com o tempo biológico, mas também com o tempo cronológico (tempo físico ou de calendário), ou seja, apresentam um comportamento dinâmico no tempo. Refira-se no entanto que, regra geral, se admite que as probabilidades estimadas com base numa experiência recente são válidas para um determinado momento posterior. Ao longo deste capítulo, e salvo referências explícitas, admitiremos que estamos perante uma população homogénea, para a qual a probabilidade de sobrevivência é função apenas da idade, e um período de tempo no qual as probabilidades podem ser consideradas constantes.

## 2.3 Indicadores característicos

A literatura demográfica e actuarial propõe diversos valores característicos para sintetizar a distribuição de probabilidade de  $T_x$ . Nesta secção descrevemos, de forma breve, alguns dos principais indicadores usados para este efeito.

### 2.3.1 Esperança de vida

O valor esperado de  $T_x$ , denotado por  $\dot{e}_x$ , representa a esperança de vida (completa) de um indivíduo com idade  $x$ <sup>11</sup>

$$\dot{e}_x = \mathbb{E}[T_x] = \int_0^\infty \mathbb{P}[T_x > t] dt = \int_0^\infty {}_t p_x \mu_{x+t} dt = \int_0^\infty {}_t p_x dt. \quad (2.3.1)$$

---

<sup>10</sup>Para uma breve introdução a este tema veja-se a Secção 2.6.

<sup>11</sup>Para compreender este resultado, recorde-se que

$$\frac{d}{dt} ({}_t p_x) = \frac{d}{dt} \frac{S(x+t)}{S(x)} = \frac{S(x+t)}{S(x)} \frac{\frac{d}{dt} S(x+t)}{S(x+t)} = -{}_t p_x \mu_{x+t}.$$

### 2.3. Indicadores característicos

Em particular, quando  $x = 0$ ,  $\dot{e}_0$  representa a esperança de vida à nascença.<sup>12</sup>

Conforme salientámos na Secção 2.2.4, o cálculo de  $\dot{e}_x$  exige a formulação de hipóteses quanto à repartição dos óbitos entre idades inteiras e períodos de observação. Se designarmos por  $\dot{e}_x^A$  a esperança de vida  $\dot{e}_x$  calculada a partir da hipótese (2.2.25) de repartição uniforme dos óbitos durante o ano, prova-se que a sua expressão é dada por

$$\dot{e}_x^A = \frac{1}{2} + \sum_{k=1}^{\infty} k p_x. \quad (2.3.2)$$

Se considerarmos a hipótese de Balducci (2.2.29), a esperança de vida correspondente  $\dot{e}_x^B$  é calculada mediante

$$\dot{e}_x^B = \frac{p_x \ln p_x}{p_x - 1} + \sum_{k=1}^{\infty} \left( \prod_{j=0}^{k-1} p_{x+j} \right) \frac{p_{x+k} \ln p_{x+k}}{p_{x+k} - 1}. \quad (2.3.3)$$

Por fim, se usarmos a hipótese de força de mortalidade constante (2.2.30), a esperança de vida correspondente  $\dot{e}_x^C$  é calculada pela expressão

$$\dot{e}_x^C = \frac{p_x - 1}{\ln p_x} + \sum_{k=1}^{\infty} \left( \prod_{j=0}^{k-1} p_{x+j} \right) \frac{p_{x+k} - 1}{\ln p_{x+k}}. \quad (2.3.4)$$

Integrando por partes, obtemos

$$\dot{e}_x = - \int_0^{\infty} t \frac{d}{dt} ({}_t p_x) dt = -t {}_t p_x \Big|_0^{\infty} + \int_0^{\infty} {}_t p_x dt.$$

Uma função de sobrevivência realista respeita a condição  $\lim_{t \rightarrow \infty} t S(x+t) = 0$  ou, de forma equivalente,  $\lim_{t \rightarrow \infty} t {}_t p_x = 0$ . Neste sentido,  $-t {}_t p_x \Big|_0^{\infty} = 0$ , obtendo-se assim a equação (2.3.1). Para mais detalhes veja-se, por exemplo, Gerber (1995), Bowers *et al.* (1997) e Pitacco (2000).

<sup>12</sup>No caso em que se consideram apenas idades (discretas) inteiras, é comum a utilização do conceito de *esperança de vida truncada* (*Curtate Expectation of Life*),  $e_x$ :

$$e_x = \mathbb{E}[K_x] = \sum_{k=0}^{\infty} k \mathbb{P}[K_x = k] = \sum_{k=0}^{\infty} k {}_k p_x - \sum_{k=0}^{\infty} k {}_{k+1} p_x = \sum_{k=1}^{\infty} {}_k p_x,$$

onde, na última equação, fizemos uso da condição  $\lim_{t \rightarrow \infty} k (-{}_k p_x) \Big|_0^{\infty} = 0$ . Por definição,  $K_x$  satisfaz as desigualdades  $K_x \leq T_x \leq K_x + 1$ , pelo que aplicando o operador valor esperado obtemos  $\mathbb{E}(K_x) \leq \mathbb{E}(T_x) \leq \mathbb{E}(K_x) + 1$ , donde é possível escrever, para todas as idades  $x$

$$e_x \leq \dot{e}_x \leq e_x + 1.$$

Denotemos novamente por  $U$  a fracção do ano da morte durante o qual ( $x$ ) vive. Nesse caso,  $T_x = K_x + U$ .  $U$  é uma variável aleatória com distribuição contínua entre 0 e 1. Aproximando o seu valor esperado por  $1/2$ , podemos escrever

$$\dot{e}_x \simeq e_x + \frac{1}{2}.$$

### 2.3. Indicadores característicos

---

Demonstra-se igualmente que os valores da esperança de vida completa calculados mediante (2.3.2), (2.3.3) e (2.3.4) respeitam a seguinte relação

$$\dot{e}_x^A \geq \dot{e}_x^C \geq \dot{e}_x^B.$$

#### 2.3.2 Variância da vida residual

A variância de  $T_x$  fornece informação sobre a dispersão da duração da vida residual do indivíduo. Definida por  $\text{Var}[T_x]$ , o seu valor é calculado mediante

$$\begin{aligned} \text{Var}[T_x] &= \mathbb{E}[(T_x)^2] - (\dot{e}_x)^2 = \int_0^\infty t^2 f_x(t) dt - (\dot{e}_x)^2 \\ &= -\frac{1}{S(x)} \int_0^\infty t^2 S'(x+t) dt - (\dot{e}_x)^2 \\ &= -\frac{1}{S(x)} \left[ t^2 S(x+t) \Big|_0^\infty - 2 \int_0^\infty t S(x+t) dt \right] - (\dot{e}_x)^2 \\ &= 2 \int_0^\infty t {}_t p_x dt - (\dot{e}_x)^2, \end{aligned} \tag{2.3.5}$$

onde, na última equação, se admite que  $\lim_{t \rightarrow \infty} t^2 S(x+t) = 0$ .<sup>13</sup>

#### 2.3.3 Vida mediana e ponto de Lexis

A vida mediana, ou vida provável, corresponde à idade que um indivíduo com  $x$  anos alcançará com 50% de probabilidade. Dito de outro modo, um indivíduo mediano tem igual probabilidade de alcançar essa idade e de não a atingir. Designada por  $vm_x$ , a vida provável à idade  $x$  é definida por

$$\mathbb{P}[T_x > vm_x] = \frac{S(x + vm_x)}{S(x)} = {}_{vm_x} p_x = \frac{1}{2}. \tag{2.3.6}$$

No caso particular em que  $x = 0$ ,  $vm_0$  é obtido resolvendo  $S(vm_0) = 1/2$ . A vida mediana à idade  $x$  pode ainda ser interpretada como a idade que 50% dos indivíduos da população com essa idade alcançarão.

A moda da distribuição de  $T_x$ , conhecida na literatura actuarial por *ponto de Lexis*, é determinada encontrando o valor  $t$  que maximiza a grandeza  ${}_t p_x \mu_{x+t}$ . Designada por  $m_x^{Lexis}$ , é calculada em cada idade  $x$  resolvendo o problema

$$m_x^{Lexis} = \max_t \{ {}_t p_x \mu_{x+t} \}. \tag{2.3.7}$$

---

<sup>13</sup>Veja-se Pitacco (2000).



## 2.4 Tábuas de mortalidade

### 2.4.1 Introdução e enquadramento histórico

A tábua de vida ou de sobrevivência (*life table*) ou, como é usualmente conhecida entre os actuários portugueses, tábua de mortalidade, é um modelo tabular de análise demográfica que sintetiza um conjunto de funções básicas que permitem analisar, numa determinada população, o fenómeno da longevidade e efectuar juízos probabilísticos sobre a evolução da mortalidade. A tábua de mortalidade constitui uma ferramenta estatística usada frequentemente por actuários, demógrafos, médicos e outros investigadores no domínio da saúde pública.

Entre outras aplicações, as tábuas de mortalidade são fundamentais para a veracidade das bases técnicas em que assentam os cálculos relativos às operações de seguros do ramo vida, devendo por isso constituir uma das preocupações primordiais do actuário. Destaque ainda para a importância que as tábuas assumem na elaboração de projecções demográficas para a população, um tema que retomaremos no capítulo 6.

O interesse pela análise da longevidade humana remonta ao início da civilização moderna.<sup>14</sup> Os primeiros escritos sobre análise de sobrevivência e sobre aquilo que mais tarde viria a ser chamado de tábua de mortalidade surgem na época dos Césares, na antiga Roma, com a chamada Tábua Ulpiana, datada do ano 220, assim designada em tributo ao seu criador, o romano Domitius Ulpianus, prefeito da cidade de Roma. Apesar de rudimentares, e de revelarem uma imagem provavelmente distorcida da mortalidade na época, estas primeiras tábuas serviram ainda assim de base ao cálculo de rendas associadas a legados, tendo sido usadas até finais do século III no norte da Itália.

A elaboração de tábuas de mortalidade com algum rigor só se tornou possível a partir do momento em que se iniciou o processo de recolha de informação estatística sobre a mortalidade das populações. Deste modo, não é de estranhar que as primeiras tábuas com algum rigor matemático surjam associadas às denominadas “Tontinas”<sup>15</sup> e aos chamados “*bills of mortality*”, uma espécie de certificados de óbito, coligidos pelas paróquias de Londres a partir do século XVI, onde se registava o número de óbitos, as

---

<sup>14</sup>Para uma análise histórica detalhada sobre a evolução da ciência actuarial recomenda-se a leitura de Haberman (1996) e Haberman e Sibbett (1995).

<sup>15</sup>O termo “Tontina” designa um sistema idealizado pelo banqueiro napolitano Lorenzo Tonti, que consistia num fundo comum constituído pelas contribuições dos subscritores que, mediante a entrega de um capital, adquiriam o direito a uma renda anual vitalícia. Na morte de um subscritor, o seu quinhão era adquirido pelos restantes até que, falecido o último beneficiário, o capital revertia para o Estado, gestor da Tontina. As Tontinas aparecem pela primeira vez em Itália em 1650, na Holanda em 1671, em França em 1684, em Inglaterra em 1692 e na Prússia em 1698.

## 2.4. Tábuas de mortalidade

---

respectivas causas e, mais tarde, a idade do falecido. É com base nestes registos que John Graunt (1620-1674), um merceeiro londrino, considerado por muitos o precursor da demografia, publica em 1662 o primeiro tratamento estatístico sobre dados demográficos e constrói a primeira tábua de mortalidade. Na sua construção, foram usados 230 000 registos ocorridos entre 1629-1638 e 1647-1658, procurando-se desta forma inferir a distribuição de probabilidade da mortalidade em grandes populações. Apesar de apresentar algumas fragilidades metodológicas, o seu trabalho pioneiro despoletou múltiplas investigações em vários domínios da demografia e conduziu, em última análise, à criação de organismos estatísticos governamentais.

Christian e Lodewijk Huygens formulam uma interpretação probabilística da tábua de mortalidade elaborada por Graunt, desenvolvendo os conceitos de idade média de morte, de esperança de vida e de mediana da vida residual. Em 1671, o primeiro-ministro holandês Jan de Witt publica um relatório onde apresenta um método de cálculo do valor de anuidades, onde combina a taxa de juro e a idade do subscritor com estimativas sobre a probabilidade de sobrevivência extraídas de uma tábua de mortalidade hipotética.

Mais tarde, em 1693, Edmund Halley, um astrónomo real britânico, publica uma das primeiras tábuas de mortalidade conhecidas, construída a partir do número de óbitos, classificados por causa de morte, observados na cidade polaca de Breslau, entre 1687 e 1691. A metodologia adoptada assentava no pressuposto de que a população da cidade era estacionária, i.e., com taxa de crescimento nulo. No entanto, é a Milne que, em 1815, é imputada a elaboração da primeira tábua considerada cientificamente correcta, apoiada em população e óbitos classificados por idade de duas paróquias da cidade britânica de Carlisle, para o período 1779-1787.

Nos cem anos que se seguiram aos trabalhos de Halley foram divulgadas várias tábuas, com destaque para a tábua londrina de John Smart em 1738, para a primeira tábua francesa de William Kersseboom e Antoine Deparcieux em 1746, para as tábuas de Estocolmo, Northampton e da Suécia divulgadas por Richard Price (1783), e para a primeira tábua elaborada para os Estados Unidos da América por Edward Wigglesworth em 1789, com base em dados para os Estados do Massachusetts e New Hampshire. Desde então, a publicação de tábuas de mortalidade tornou-se prática comum um pouco por todo o mundo, em especial nos momentos em que ocorrem censos de população.

Em Portugal, o Instituto Nacional de Estatística (INE) publica tábuas de mortalidade desde 1930. As primeiras tábuas utilizavam a população censitária, dada a maior fiabilidade da estrutura etária da população observada nestes momentos. Existem tábuas completas para o período de 1930-31, 1939-42, 1949-52, 1959-62, 1969-72 e 1979-82, períodos em que a população utilizada foi a do Censo de 1930, 1940, 1950, 1960, 1970

## 2.4. Tábuas de mortalidade

---

e 1980, respectivamente, e os óbitos correspondiam, na maioria dos casos, à média dos valores registados nos quatro anos centrados em cada Censo.<sup>16</sup>

Carrilho (1980) calcula tábuas abreviadas de mortalidade para Portugal (território continental, Açores e Madeira), desagregadas por sexo e por distrito, para o período 1941-1975. As tábuas abreviadas de mortalidade utilizavam as tabelas de Reed e Merrell, tendo ainda sido ensaiados outros métodos de transformação da taxa de mortalidade em quocientes (e.g. métodos de Canedo, Chiang e da O.N.U.). Actualmente, as tábuas de mortalidade, sobretudo as abreviadas, calculam-se anualmente, sendo a esperança de vida um dos indicadores chave na análise da evolução da situação demográfica (Carrilho e Patrício, 2004).

### 2.4.2 Definição e conceitos básicos

Imaginemos um conjunto (*cohort*) de  $l_\alpha$  indivíduos recém-nascidos no mesmo ano. O valor de  $l_\alpha$  é chamado de *radix* (ou raiz) da tábua e toma, regra geral, valores convenientes (e.g. 100.000 ou 1.000.000) de forma a facilitar a interpretação da informação contida na tábua.<sup>17</sup> Admitamos que se trata de um *cohort* fechado a novos ingressos (i.e., não são admitidos novos nascimentos e não são observados fenómenos de emigração ou imigração), pelo que apenas pelo falecimento é possível deixar de fazer parte desta 'geração'. Com o passar do tempo, a dimensão deste *cohort* diminui gradualmente em consequência da morte dos seus membros. A informação recolhida permite-nos então construir uma tábua de mortalidade geracional, uma forma simples de representar a evolução da mortalidade de um *cohort* ideal.

Como se depreende, a construção de tábuas geracionais só é possível após a extinção integral da geração. Por esta razão, a maioria das tábuas publicadas funda-se, por um lado, na utilização de dados sobre a população contemporânea e, por outro, na definição de um *cohort* fictício, que apresenta uma evolução da mortalidade estimada a partir dos valores observados na população.

Designemos por  $T_\alpha^{(j)}$  a duração (aleatória) residual da vida do  $j^{\text{ésimo}}$  indivíduo,  $j = 1, \dots, l_\alpha$ . Admitamos que a probabilidade de sobrevivência de cada um dos  $l_\alpha$  membros deste grupo é descrita por  $S(x)$ . O número de sobreviventes (de um conjunto inicial de

---

<sup>16</sup>Marques (1970) publicadas tábuas abreviadas de mortalidade para os períodos 1950-51, 1960-61, 1962, 1963, 1964, 1965, 1966 e 1967, para o total do país e por sexo. Também foram publicadas pela primeira vez tábuas de mortalidade abreviadas por estado civil e sexo. Segundo o autor, as tábuas foram elaboradas utilizando as tabelas de Reed e Merrell e as relações funcionais aproximadas entre taxas e quocientes de mortalidade (no sentido probabilístico) de mortalidade.

<sup>17</sup>No caso particular em que a tábua tem início nas idades referentes aos recém-nascidos, o radix da tábua é representado por  $l_0$ .

## 2.4. Tábuas de mortalidade

---

$l_\alpha$ ) à idade  $x$  ( $x > \alpha$ ),  $\mathcal{L}_x$ , é dado por

$$\mathcal{L}_x = \sum_{j=1}^{l_\alpha} \mathbf{1}_x^{(j)}, \quad (2.4.1)$$

onde  $\mathbf{1}_x^{(j)}$  é um indicador de sobrevivência (função indicatriz) da vida  $j$ , i.e.,

$$\mathbf{1}_x^{(j)} = \begin{cases} 1 & \text{se } T_\alpha^{(j)} > (x - \alpha) \\ 0 & \text{se } T_\alpha^{(j)} \leq (x - \alpha) \end{cases}.$$

Representemos por  ${}_tD_x$  o número aleatório de falecimentos entre as idades  $x$  e  $x + t$ . Nesse caso,

$${}_tD_x = \mathcal{L}_x - \mathcal{L}_{x+t}. \quad (2.4.2)$$

Com base nesta informação, é possível calcular os seguintes valores esperados

$$\mathbb{E}(\mathcal{L}_x) = \mathbb{E}\left(\sum_{j=1}^{l_\alpha} \mathbf{1}_x^{(j)}\right) = \sum_{j=1}^{l_\alpha} \mathbb{E}\left(\mathbf{1}_x^{(j)}\right) = l_\alpha \frac{S(x)}{S(\alpha)} = l_\alpha \cdot {}_{x-\alpha}p_\alpha \quad (2.4.3)$$

$$\mathbb{E}({}_tD_x) = \mathbb{E}(\mathcal{L}_x) - \mathbb{E}(\mathcal{L}_{x+t}) = l_\alpha ({}_{x-\alpha}p_\alpha - {}_{x+t-\alpha}p_\alpha) = l_\alpha \cdot {}_{x-\alpha|t}q_\alpha. \quad (2.4.4)$$

Admitindo que as variáveis aleatórias  $T_\alpha^{(j)}$  ( $j = 1, \dots, l_\alpha$ ) são mutuamente independentes, é fácil de ver que  $\mathcal{L}_x$  segue uma distribuição Binomial,<sup>18</sup> i.e.,

$$\mathbb{P}[\mathcal{L}_x = k] = \binom{l_\alpha}{k} ({}_{x-\alpha}p_\alpha)^k (1 - {}_{x-\alpha}p_\alpha)^{l_\alpha - k}. \quad (2.4.5)$$

Representemos por  $l_x$  o número esperado de sobreviventes à idade  $x$  de um conjunto inicial de  $l_\alpha$ , i.e.,  $\mathbb{E}(\mathcal{L}_x) = l_x$ . Nesse caso, a equação (2.4.3) dá lugar a

$$l_x = l_\alpha \frac{S(x)}{S(\alpha)} \quad (2.4.6)$$

e, em particular, para  $\alpha = 0$ ,  $l_x = l_0 S(x)$ . Denotemos por  $P_x$  a proporção de sobreviventes na idade exacta  $x$ , i.e.,

$$P_x = \frac{l_x}{l_0} = S(x). \quad (2.4.7)$$

---

<sup>18</sup>Veja-se, por exemplo, Bowers *et al.* (1997).

## 2.4. Tábuas de mortalidade

De modo análogo, podemos escrever

$${}_t d_x = \mathbb{E}({}_t D_x) = l_x - l_{x+t} \quad (2.4.8)$$

e, em particular, para  $t = 1$ ,  $d_x = l_x - l_{x+1}$ . A tabulação das funções  $l_x$  para idades  $x$  inteiras,  $\{l_x\}_{x=\alpha}^{x=\omega}$ , com  $\omega$  definido de forma a que  $l_\omega \simeq 0$ , constitui a base das tábuas de mortalidade (ou de sobrevivência). Registe-se igualmente que a sequência  $\{d_x\}_{x=\alpha}^{x=\omega-1}$  respeita a relação  $\sum_{x=\alpha}^{x=\omega-1} d_x = l_\alpha$ .

### 2.4.3 A tábua enquanto modelo probabilístico de sobrevivência

Na secção anterior descrevemos a construção de uma tábua  $\{l_x\}$  usando um procedimento puramente dedutivo, elaborado a partir de um modelo probabilístico baseado na função de sobrevivência  $S(x)$ . Na prática, a construção de  $\{l_x\}$  assenta na recolha e tratamento de informação estatística. Contudo, o conhecimento de  $\{l_x\}$  comporta em si mesmo um modelo probabilístico construído em tempo discreto, em particular para idades e duração de vida inteiras.

Admitamos que a função de sobrevivência  $S(x)$  é descrita pela sequência  $\{l_x\}_{x=\alpha}^{x=\omega}$  da tábua de mortalidade e consideremos, por simplificação, que  $\alpha = 0$ . Com base na equação (2.4.6), é possível expressar as probabilidades condicionadas de sobrevivência e de morte no intervalo  $(x, x + t]$  para um indivíduo com idade exacta  $x$  segundo

$${}_t p_x = \prod_{j=0}^{t-1} p_{x+j} = \frac{l_{x+t}}{l_x} \quad (2.4.9)$$

$${}_t q_x = 1 - {}_t p_x = \frac{{}_t d_x}{l_x} = \frac{l_x - l_{x+t}}{l_x}. \quad (2.4.10)$$

Com base em (2.4.10), podemos redefinir o conceito de força ou intensidade de mortalidade  $\mu_x$  segundo<sup>19</sup>

$$\mu_x = -\frac{1}{l_x} \frac{dl_x}{dx} = -\frac{d \ln l_x}{dx}. \quad (2.4.11)$$

<sup>19</sup>Na derivação deste resultado recorde-se que

$$\begin{aligned} \mu_x &= \lim_{\Delta x \rightarrow 0^+} \left( \frac{l_x - l_{x+\Delta x}}{l_x \Delta x} \right) = -\lim_{\Delta x \rightarrow 0^+} \left( \frac{l_{x+\Delta x} - l_x}{l_x \Delta x} \right) \\ &= -\frac{1}{l_x} \lim_{\Delta x \rightarrow 0^+} \left( \frac{l_{x+\Delta x} - l_x}{\Delta x} \right) = -\frac{1}{l_x} \frac{dl_x}{dx} = -\frac{d \ln l_x}{dx}. \end{aligned}$$

## 2.4. Tábuas de mortalidade

---

A função

$$-\frac{dl_x}{dx} = \mu_x l_x \quad (2.4.12)$$

representa, neste contexto, a função de densidade ou, como referimos anteriormente, a curve de mortes. Integrando (2.4.12), obtém-se

$$l_x = \int_x^\infty \mu_y l_y dy = \int_0^\infty \mu_{x+t} l_{x+t} dt \quad (2.4.13)$$

De forma análoga, se integrarmos (2.4.11) obtemos

$$l_x = l_0 \exp\left(-\int_0^x \mu_y dy\right) \quad (2.4.14)$$

e, se integrarmos (2.4.11) entre  $x$  e  $x+n$ , obtemos

$$l_x - l_{x+n} = \int_x^{x+n} l_y \mu_y dy. \quad (2.4.15)$$

As tábuas de mortalidade comportam ainda outras funções básicas. O símbolo  ${}_nL_x$  é usado para denotar o número total (esperado) de anos vividos entre as idades  $x$  e  $(x+n)$  pelos  $l_x$  sobreviventes de um grupo inicial de  $l_0$ . Formalmente<sup>20</sup>

$${}_nL_x = \int_0^n l_{x+t} dt. \quad (2.4.16)$$

Com base na função  ${}_nL_x$ , podemos igualmente redefinir o conceito de taxa de mortalidade<sup>21</sup> no intervalo  $(x, x+n)$ ,  ${}_n m_x$ . O seu valor representa, neste contexto, a taxa de mortalidade média experimentada pelos  $l_x$  sobreviventes de um grupo inicial de  $l_0$

$${}_n m_x = \frac{\int_0^n l_{x+t} \mu_{x+t} dt}{\int_0^n l_{x+t} dt} = \frac{l_x - l_{x+n}}{{}_nL_x}. \quad (2.4.17)$$

Denotemos por  $\omega$  o ano de vida tal que  $l_{\omega-1} > l_\omega = 0$ . O símbolo  $T_x$  denota, na literatura actuarial, o número total de anos vividos para além da idade  $x$  pelos

---

<sup>20</sup> Este resultado é derivado notando que

$$\begin{aligned} {}_nL_x &= \int_0^n t l_{x+t} \mu_{x+t} dt + n l_{x+n} = -\int_0^n t \frac{\partial l_{x+t}}{\partial t} dt + n l_{x+n} \\ &= -t l_{x+t} \Big|_0^n + \int_0^n l_{x+t} dt + n l_{x+n} = \int_0^n l_{x+t} dt. \end{aligned}$$

<sup>21</sup> Também conhecido por taxa de mortalidade central.

## 2.4. Tábuas de mortalidade

---

sobreviventes de um grupo inicialmente constituído por  $l_0$  membros<sup>22</sup>

$$T_x = \int_0^{\infty} {}_t l_{x+t} \mu_{x+t} dt = - \int_0^{\infty} t \frac{d l_{x+t}}{dt} dt = \int_0^{\infty} l_{x+t} dt \quad (2.4.18)$$

ou, em alternativa

$$T_x = \sum_{t=0}^{\infty} L_{x+t}. \quad (2.4.19)$$

onde  $L_x = {}_1 L_x$ .

Registe-se que  $T_x = \lim_{n \rightarrow \infty} ({}_n L_x)$ . O número médio de anos de vida futura dos  $l_x$  sobreviventes de um grupo inicial de  $l_0$  (i.e. a sua *esperança de vida completa*,  $\dot{e}_x$ ) é dado, para cada idade  $x$ , por<sup>23</sup>

$$\dot{e}_x = \frac{T_x}{l_x} = \frac{\int_0^{\infty} l_{x+t} dt}{l_x} = \int_0^{\infty} {}_t p_x dt = \frac{1}{l_x} \sum_{t=0}^{\infty} L_{x+t}. \quad (2.4.20)$$

### 2.4.4 Classificação das tábuas de mortalidade

As tábuas de mortalidade diferenciam-se pelas características da população analisada (sexo, tipo de seguro, grupo de risco, etc.) e por um conjunto de aspectos metodológicos que presidem à sua construção. Entre estes factores, destacamos a referência a gerações reais ou fictícias, a amplitude do intervalo de idades no qual assenta a informação, a possibilidade de considerar a idade em que o indivíduo passa a integrar a população, a indexação (ou não) da tábua a um tempo cronológico e a consideração de outras formas de exclusão da população. A possibilidade de nos referirmos a gerações reais ou fictícias permite-nos efectuar a distinção entre tábuas contemporâneas (transversais ou de momento) e tábuas geracionais (ou longitudinais):

- **Tábuas contemporâneas (transversais ou de momento):** assentam na análise *cross-section* de uma geração fictícia (composta, normalmente, por 100 000 indivíduos), que é sujeita às condições de mortalidade observadas para cada idade (ou agrupamentos de idades) num determinado ano (ou num período de tempo).

---

<sup>22</sup>O significado do símbolo  $T_x$  não deve aqui ser confundido com o de duração aleatória da vida humana usado desde o início deste capítulo.

<sup>23</sup>Um conceito próximo na literatura actuarial é aquele que se refere ao número médio de anos de vididos entre as idades  $x$  e  $(x+n)$  pelos  $l_x$  sobreviventes de um grupo inicial de  $l_0$ . Designado por *esperança de vida completa temporária* de  $(x)$ ,  $\dot{e}_{x:n}$ , é definido por

$$\dot{e}_{x:n} = \frac{{}_n L_x}{l_x} = \frac{\int_0^n l_{x+t} dt}{l_x} = \frac{T_x - T_{x+n}}{l_x} = \int_0^n {}_t p_x dt.$$

## 2.4. Tábuas de mortalidade

---

Estas tábuas baseiam-se na utilização de estatísticas referentes a um período de tempo curto quer para os óbitos, quer para a população exposta ao risco. Se a informação se reporta a um único ano de calendário, denominam-se tábuas *anuais*. Para contornar as oscilações anuais no número de óbitos, recorre-se frequentemente a uma média de dois ou mais anos, sendo que nesse caso as tábuas são chamadas de *plurianuais*.

- **Tábuas geracionais (ou longitudinais):** assentam no cálculo das taxas de mortalidade averbadas por uma mesma geração, ou seja, por indivíduos nascidos no mesmo ano. A construção deste tipo de tábuas pressupõe o acompanhamento dos indivíduos desde o seu nascimento até à morte, naquilo que se denomina em demografia por análise longitudinal.

Se considerarmos a amplitude do intervalo de idades em que assenta a informação, podemos distinguir entre tábuas completas e tábuas abreviadas:

- **Tábuas completas:** contém dados para cada idade singular (idade ano a ano) desde o nascimento até ao limite superior (idade máxima) adoptado.
- **Tábuas abreviadas:** reportam dados para grupos de idades (quinquenais ou decenais).

A incorporação na análise da idade em que o indivíduo passa a integrar a população (i.e., é seleccionado) permite-nos construir tábuas bidimensionais, em que os índices são a *idade de selecção* (e.g., a idade em que se realiza a apólice de seguro) e a *duração* desde esse momento. Neste caso, podemos distinguir entre tábuas seleccionadas e tábuas finais:

- **Tábuas seleccionadas** (*select life tables*): são tábuas em que as probabilidades são tabuladas considerando a idade  $x$  do indivíduo e outras informações (e.g. tempo decorrido desde a realização de uma apólice de seguro).
- **Tábuas finais** (*ultimate life tables*): corresponde à última coluna de uma tábua seleccionada, ou seja, aquela em que se admite que a duração (ou período de selecção) deixa de ter efeito sobre a mortalidade.

A possibilidade de considerar o tempo cronológico na elaboração e apresentação das tábuas permite-nos distinguir entre tábuas estáticas e tábuas dinâmicas ou prospectivas:

- **Tábuas estáticas:** são tábuas unidimensionais onde todas as funções se reportam apenas à idade biológica  $x$ .

- **Tábuas dinâmicas (ou prospectivas)**: são tábuas bidimensionais em que todas as funções estão indexadas, em linha, pela idade biológica e, em coluna, pelo ano de calendário (tempo cronológico).

Por fim, a consideração de outras formas de exclusão (decremento) da população que não apenas a morte (e.g., retirada de plano de pensões) ou de outros factores (ou estados) que afectam materialmente o valor de uma apólice ou obrigação do plano de pensões (e.g., estado de saúde, estado civil), permite-nos falar em **tábuas multi-decremento** e **tábuas multi-estado**. Por exemplo, uma tábua multi-estado pode indicar a probabilidade de um indivíduo transitar da condição de saudável para a condição de doente.

## 2.5 Leis de mortalidade

Uma solução alternativa para caracterizar a função de sobrevivência e o padrão de mortalidade registados numa população passa pelo recurso às denominadas “leis de mortalidade”, expressões analíticas que descrevem as características de mortalidade em termos da idade e de um conjunto limitado de parâmetros.

O interesse pelo desenvolvimento de funções paramétricas no contexto da ciência actuarial e na demografia é justificado por razões de ordem filosófica e por motivos de natureza prática (Tenembein e Vanderhoof, 1980; Keyfitz, 1982; Pitacco, 2000).

Por razões de ordem eminentemente “filosófica”, na medida em que alguns autores sustentam, com recurso a argumentos de natureza biológica, que o fenómeno da sobrevivência humana é governado por uma “lei”, universal e imutável, semelhante àquelas que é possível identificar, por exemplo, na física ou na química.

Por motivos de ordem prática, uma vez que a utilização de leis de mortalidade:

1. Facilita a apresentação e análise da função de sobrevivência de uma população, expressa neste caso em termos de um número restrito de parâmetros e não através de centenas de valores, como no caso de uma tábua de mortalidade;
2. Proporciona um método de graduação (*smoothing*) do comportamento errático dos dados, assim como a remoção e eliminação de irregularidades (e.g., erros amostrais) na informação estatística;
3. Facilita a elaboração das próprias tábuas de mortalidade;
4. Possibilita a realização de inferência estatística nos casos em que a informação é incompleta;

5. Permite comparar de forma simples experiências de mortalidade no tempo;
6. Constitui um método simples de realizar projecções demográficas;
7. Permite efectuar juízos probabilísticos envolvendo mais do que uma vida.

Mais recentemente, é possível encontrar na literatura a referência a novas áreas de aplicação dos modelos paramétricos. Entre elas, destaco a abordagem epidemiológica da mortalidade (Van den Berg Jeths *et al.*, 2001), a modelação da heterogeneidade, estocasticidade e homeostasis (Yashin, 2001) e o desenvolvimento de modelos demográficos com inclusão de efeitos geracionais (Renshaw e Haberman, 2005).

A investigação sobre leis de mortalidade remonta à primeira metade do século XVIII e não mais terminou. Nesta secção, apresentamos uma síntese crítica dos principais modelos desenvolvidos até ao presente. Para uma análise mais detalhada deste tema recomenda-se a leitura de, por exemplo, Benjamin e Pollard (1980), Elandt-Johnson e Johnson (1980), Keyfitz (1982), Gavrilov e Gavrilova (1991), Gerber (1995), Haberman e Sibbett (1995), London (1997), Bowers *et al.* (1997), Pitacco (2000, 2004), Tabeau (2001) e López (2001).

Os modelos analíticos propostos podem ser agrupados segundo inúmeros critérios. Em primeiro lugar, é possível identificar modelos *explicativos*, i.e., modelos que sugerem um mecanismo teórico para explicar a mortalidade, e modelos meramente *descritivos*, ou seja, modelos que se limitam a ajustar os dados num determinado momento do tempo e para um dado intervalo de idades.

Em segundo, os modelos podem ser distinguidos de acordo com o número de componentes que incorporam. Os modelos assentes numa única componente não efectuem qualquer distinção quanto às “causas” de morte dominantes nas diferentes etapas em que se decompõe a vida de um indivíduo, algo que os modelos multi-função permitem. Em terceiro, podemos restringir entre leis que cobrem a totalidade do espectro da vida humana e leis que se reportam apenas a idades (ou intervalos de idades) seleccionadas.

Por fim, as leis de mortalidade diferenciam-se segundo a função escolhida para modelar a mortalidade humana. Assim, é possível encontrar modelos de ajustamento para a função de sobrevivência  $S(x)$ , modelos que parametrizam a força de mortalidade  $\mu_x$ , modelos de ajustamento para quocientes de funções de mortalidade (e.g.  $q_x/p_x$ ) e modelos que ajustam a função de densidade de probabilidade  $f_x(\cdot)$ . Adoptaremos este último critério na apresentação dos modelos, por considerarmos que é aquele que introduz maior clareza na exposição ao mesmo tempo que facilita a identificação das semelhanças e diferenças entre as abordagens.

### 2.5.1 Modelos de ajustamento para a função de sobrevivência

#### Lei de De Moivre

A primeira tentativa de formular uma lei de mortalidade é atribuída a Abraham De Moivre (1725). O autor admite a existência de uma relação proporcional entre o número de sobreviventes numa dada idade  $x$  e o número de anos que a separam da idade máxima admissível  $\omega$  ( $\omega = 86$  na versão original). Em termos mais precisos, admite-se que a vida residual  $T_x$  é uniformemente distribuída entre as idades 0 e  $\omega - x$ , o que pressupõe uma função de sobrevivência linear do tipo

$$S(x) = \begin{cases} 1 - \frac{x}{\omega} & x \in [0, \omega] \\ 0 & x > \omega. \end{cases} \quad (2.5.1)$$

A expressão (2.5.1) estabelece que o número de sobreviventes de uma geração decresce linearmente com a idade, uma hipótese irrealista face aos valores observados, postula que a força de mortalidade é crescente para todas as idades e, não menos significativo, determina uma função de densidade constante. Com efeito, com base em (2.2.23) e em  $f_x(\cdot) = \mu_x S(x) = -\frac{d}{dx} S(x)$ , deduz-se que

$$\mu_x = -\frac{\frac{d}{dx} S(x)}{S(x)} = \frac{1}{\omega - x}, \quad x \in [0, \omega] \quad (2.5.2)$$

$$f_x(t) = \frac{1}{\omega - x} \text{ para } x < t < \omega. \quad (2.5.3)$$

Em termos de tábua de mortalidade, o modelo de De Moivre estipula um número de óbitos constante entre duas idades actuariais consecutivas, i.e.,

$$d_x = l_x - l_{x+1} = l_0 [S(x) - S(x+1)] = l_0 \left( \frac{1}{\omega} \right), \quad (2.5.4)$$

um resultado dificilmente aceitável em termos empíricos, e a seguinte expressão para as probabilidades de sobrevivência (para  $x + t \leq \omega$ )<sup>24</sup>

$${}_t p_x = 1 - \frac{t}{\omega - x} \text{ para } x < t < \omega - x. \quad (2.5.5)$$

---

<sup>24</sup>Demonstra-se ainda que os indicadores  $\hat{e}_x$  e  $vm_x$  derivados do modelo são iguais e dados por

$$\hat{e}_x = vm_x = \frac{(\omega - x)}{2}.$$

### Lei de De Graaf

De Graaf (1729)<sup>25</sup> desenvolveu uma generalização do modelo de De Moivre em que admite que a vitalidade de um indivíduo diminui a uma taxa crescente à medida que a idade  $x$  avança. Esta teoria materializou-se na seguinte formulação para  $S(x)$

$$S(x) = \begin{cases} 1 - \left(\frac{x}{\omega}\right)^n & x \in [0, \omega] \\ 0 & x > \omega, \end{cases} \quad (2.5.6)$$

com  $n \geq 1$  e, na versão original,  $\omega = 92$ . Note-se que, no caso particular em que  $n = 1$ , obtemos a lei de De Moivre. O parâmetro  $n$  representa a velocidade de diminuição da vitalidade do indivíduo, ou seja, determina o grau de concavidade da função de sobrevivência e condiciona a força de mortalidade que apresenta, como referimos anteriormente, um crescimento exponencial, sobretudo nas idades mais avançadas. Pode ver-se que  $\mu_x$  é dado, neste caso, por

$$\mu_x = \frac{nx^{n-1}}{\omega^n - x^n} \quad x \in [0, \omega]. \quad (2.5.7)$$

### Lei de Lambert

Lambert (1776) desenvolveu uma fórmula para aproximar uma tábua de mortalidade construída a partir dos óbitos registados nos *London Bills of Mortality* no período 1728-1757. O seu modelo é decalcado da lei de De Graaf, com  $n = 2$ , corrigido com duas funções exponenciais

$$S(x) = \begin{cases} \left(1 - \frac{x}{\omega}\right)^2 - a(e^{-b_1x} - e^{-b_2x}) & x \in [0, \omega] \\ 0 & x > \omega, \end{cases}$$

com  $\omega = 96$  e  $a, b_1, b_2 \geq 0$ . O autor não fundamentou a sua lei, mas verifica-se que a sua concepção básica é a de que a capacidade do ser humano para “resistir” à morte vai diminuindo com a idade de forma quadrática (e não linear, como assumia De Moivre), uma explicação que seria retomada cinquenta anos mais tarde por Gompertz.

### Lei de Babbage

Em alternativa aos modelos anteriores, Babbage (1823) sugere a seguinte função polinomial de segundo grau

---

<sup>25</sup>Citado por Pitacco (2000) e López (2001).

$$S(x) = \begin{cases} 1 - \left(\frac{x}{\omega}\right)^2 + ax \left(1 - \frac{x}{\omega}\right) & x \in [0, \omega] \\ 0 & x > \omega, \end{cases} \quad (2.5.8)$$

com  $a > 0$ .<sup>26</sup> Prova-se que a expressão de  $\mu_x$  é dada, neste caso, por

$$\mu_x = -\frac{2x(1+aw) - a\omega^2}{x^2(1+aw) - \omega^2(1+ax)} \quad x \in [0, \omega]. \quad (2.5.9)$$

### Lei de Sang

Sang (1868)<sup>27</sup> propõe uma função de sobrevivência do tipo exponencial, com valores nulos para a idade extrema  $\omega$

$$S(x) = \begin{cases} \frac{a^x - a^\omega}{1 - a^\omega} & x \in [0, \omega] \\ 0 & x > \omega, \end{cases} \quad (2.5.10)$$

com  $a > 0$ . O parâmetro  $a$  condiciona a forma da função de sobrevivência e, em particular, determina a sua convexidade ( $0 < a < 1$ ) ou concavidade ( $a > 1$ ). Demonstra-se que  $\mu_x$  é dado, neste caso, por

$$\mu_x = \frac{\ln(a)}{a^{\omega-x} - 1} \quad a > 0 \text{ e } x \in [0, \omega]. \quad (2.5.11)$$

### Leis de Dormoy

Em alternativa à Lei de Sang, Dormoy (1878) alvitra uma expressão mais simples para a função de sobrevivência, mantendo contudo a sua natureza exponencial

$$S(x) = \exp(-\eta x) \quad x \geq 0, \quad (2.5.12)$$

com  $\eta > 0$ . Note-se que a equação (2.5.12) pode ser considerada um caso particular de (2.5.10) com  $a < 1$  e  $\omega \rightarrow +\infty$ . A expressão de  $\mu_x$  é dada, neste caso, por  $\mu_x = \eta$ ,  $\eta > 0$  e  $x \geq 0$ . Assinale-se, em especial, o facto de o modelo estipular uma força de mortalidade constante ao longo de todo o espectro de vida. A expressão da probabilidade  ${}_tq_x$  é independente da idade atingida pelo indivíduo

$${}_tq_x = 1 - \frac{S(x+t)}{S(x)} = 1 - e^{(-\eta t)}. \quad (2.5.13)$$

---

<sup>26</sup>Littrow (1852) desenvolve uma extensão da lei de Babbage em que considera um polinómio de grau superior a 2.

<sup>27</sup>Citado por Pitacco (2000) e López (2001).

## 2.5. Leis de mortalidade

---

um resultado que, no campo actuarial, é aceitável apenas como aproximação “local” ao fenómeno da mortalidade, nomeadamente no cálculo de probabilidades de sobrevivência ou de morte para idades fraccionadas e intervalos de tempo de curta duração.

Em resposta às críticas dirigidas à constância de  $\mu_x$ , o autor propõe uma segunda lei, conhecida por 2ª Lei de Dormoy, assente em exponenciais de polinómio para  $S(x)$

$$S(x) = \exp[-(\eta x + \gamma x^2)] \quad x \geq 0. \quad (2.5.14)$$

com  $\eta \geq 0$  e  $\gamma > 0$ , donde resulta que  $\mu_x$  e  ${}_tq_x$  são dados por

$$\mu_x = \eta + 2\gamma x \quad x \geq 0 \quad (2.5.15)$$

$${}_tq_x = 1 - e^{[-t(\eta+2\gamma x+t)]}. \quad (2.5.16)$$

Note que, neste caso, a força de mortalidade é linear e crescente com a idade enquanto que  ${}_tq_x$  depende agora da idade  $x$  atingida pelo indivíduo.

### Lei de Laurent

Laurent (1892)<sup>28</sup> propõe uma extensão das Leis de Dormoy ao sugerir um modelo exponencial generalizado para  $S(x)$

$$S(x) = \sum_{j=1}^k a_j \exp(-b_j x) \quad x \geq 0,$$

com  $a_j \geq 0$ ,  $b_j \geq 0$  ( $j = 1 \dots k$ ). Vê-se que  $\mu_x$  é dado, neste caso, por

$$\mu_x = \frac{1}{S(x)} \left[ \sum_{j=1}^k a_j b_j \exp(-b_j x) \right] \quad x \geq 0. \quad (2.5.17)$$

### Lei de Moser

Moser (1939) sugere uma função com cinco parâmetros do tipo

$$S(x) = 1 - ax^{\frac{1}{4}} + bx^{\frac{9}{4}} - cx^{\frac{17}{4}} - dx^{\frac{25}{4}} + ex^{\frac{33}{4}}, \quad (2.5.18)$$

onde  $(a, b, c, d, e)$  são parâmetros a estimar.<sup>29</sup>

---

<sup>28</sup>Ibidem.

<sup>29</sup>Uma solução distinta para a modelação da mortalidade é proposta por Steffenson (1930) e por Harper (1936) (citados por Tabeau (2001)), que adoptam uma transformação logarítmica para a função

### Lei de Petrioli

A Lei de Petrioli (1979) baseia-se na seguinte representação para a função  $S(x)$

$$S(x) = \left[ x^a (\omega - x)^{-b} \exp (cx^2 + dx + k) + 1 \right]^{-1}, \quad (2.5.19)$$

onde  $\omega$  representa a idade máxima admissível e  $a, b, c, d$  e  $k$  são parâmetros a estimar. A lei de Petrioli é criticada por Anson (1988), que considera desnecessária e contraproducente a hipótese de uma dada idade máxima admissível estática. Em alternativa, propõe a seguinte parametrização

$$S(x) = \exp \left( -b_0x + \frac{b_1x^2}{2} + \frac{b_2x^3}{3} + \frac{b_3x^4}{4} + \frac{b_4x^5}{5} \right). \quad (2.5.20)$$

Mais recentemente, Hannerz (1999, 2001) propõe, entre outros modelos, uma reformulação da lei de Anson, com o objectivo de ajustar a evolução recente da mortalidade da população feminina. O modelo baseia-se na seguinte formulação analítica

$$S(x) = \left[ 1 + e^{G(x)} \right]^{-1} \quad (2.5.21)$$

onde

$$G(x) = b_0 - \frac{b_1}{x} + \frac{b_2}{2}x^2 + \frac{b_3}{c}e^{cx}.$$

### 2.5.2 Modelos que parametrizam a força de mortalidade

As leis de mortalidade incluídas nesta categoria apresentam um notável interesse, tanto histórico como actual. Na maioria dos casos, são leis que incorporam hipóteses biologicamente plausíveis, mas carecem de algumas correcções de modo a possibilitar a sua utilização em todo o arco da vida humana.

#### Lei de Gompertz

O modelo desenvolvido por Gompertz (1825), um dos mais influentes e utilizados em áreas tão distintas como a ciência actuarial ou a biologia, é considerado o primeiro modelo explicativo da mortalidade. Recorde-se que os modelos explicativos estão alicerçados em teorias epidemiológicas ou biomédicas para descrever os mecanismos que influenciam o processo de envelhecimento do indivíduo. Estes modelos incorporam ainda abordagens de sobrevivência.

## 2.5. Leis de mortalidade

---

que encaram o processo de envelhecimento do organismo como uma acumulação de defeitos ou falhas, que reduzem a capacidade deste para evitar a morte.

Gompertz assume que existem apenas duas causas de morte:

“It is possible that death may be the consequence of two generally co-existing causes; the one, chance, without previous disposition to death or deterioration; the other, a deterioration, or an increased inability to withstand destruction.” (Gompertz, 1825, pp.7)

O autor admite ainda que a capacidade do organismo para “resistir à morte” diminui exponencialmente com a idade. Em consequência, o autor admite que  $\mu_x$  aumenta (também de forma exponencial) com a idade.

“If the average exhaustion’s of a man’s power to avoid death were such that at the end of equal infinitely small intervals of time, he lost equal portions of his remaining power to oppose destruction which he had at the commencement of those intervals, then at age  $x$  his power to avoid death, or the intensity of his mortality might be denoted by  $a \cdot q^x \dots$ ” (Gompertz, 1825, pp.8)

Em termos formais, se admitirmos que  $\frac{1}{\mu_x}$  simboliza aquilo que Gompertz designa por capacidade para “resistir à morte”, a lei pode ser escrita como  $\frac{d}{dx} \left( \frac{1}{\mu_x} \right) = -\beta \left( \frac{1}{\mu_x} \right)$ . Resolvendo esta equação, obtém-se a seguinte parametrização para  $\mu_x$

$$\mu_x = \alpha \exp(\beta x) \quad x \geq 0, \quad (2.5.22)$$

com  $\alpha > 0$  e  $\beta > 0$ .<sup>30</sup> Como se observa,  $\mu_x$  tende, no limite, para o infinito, pelo que o modelo não estipula uma idade máxima para a vida humana. A função de sobrevivência decorrente do modelo é facilmente derivável recorrendo a (2.2.21)

$$S(x) = \exp \left( - \int_0^x \mu_s ds \right) = \exp \left[ \frac{\alpha}{\beta} \left( 1 - e^{\beta x} \right) \right]. \quad (2.5.23)$$

O modelo de Gompertz dominou a modelação da mortalidade humana durante mais

---

<sup>30</sup> A lei de Gompertz é frequentemente adoptada na ciência actuarial com base na sua formulação original

$$\mu_x = Bc^x$$

com  $B > 0$  e  $c > 1$ . Sublinhe-se, no entanto, que as duas representações são equivalentes considerando  $B = \alpha$  e  $\beta = \ln(c)$ .

## 2.5. Leis de mortalidade

---

de um século sendo, ainda hoje, extensamente aplicado noutras espécies animais.<sup>31</sup> Como veremos em seguida, uma boa parte das leis de mortalidade que se lhe sucederam incluem o modelo de Gompertz como uma das suas componentes. Não obstante a sua simplicidade, o modelo apresenta um desempenho empírico assinalável, sobretudo nas idades adultas intermédias.

Um dos principais atractivos do modelo de Gompertz reside na sua flexibilidade e abrangência. Willekens (2001) procede a uma análise detalhada das características do modelo e demonstra: (i) que este é facilmente generalizável de modo a incluir outras causas de morte; (ii) que este é um caso particular da distribuição logística generalizada; (iii) que este é uma versão truncada ( $x \in [0, +\infty)$ ) de uma distribuição de valores extremos do Tipo I; e (iv) que uma transformação (logarítmica) da variável idade permite relacionar o modelo de Gompertz com a distribuição de Weibull.

Pollard e Valkovics (1992) notam que a força de mortalidade no ponto corresponde à moda da função de densidade  $l_x \mu_x$  (curva de mortes) coincide com o parâmetro  $\beta$  da lei de Gompertz, sugerindo a seguinte reformulação de (2.5.22)

$$\mu_x = \beta \exp [(x - m) \beta] \quad x \geq 0,$$

onde  $m$  denota a moda da curva de mortes.<sup>32</sup>

### Leis de Makeham

Reconhecendo as dificuldades da lei de Gompertz em captar o comportamento da mortalidade nas idades extremas do arco da vida humana, William Makeham (1860, 1867), um actuário britânico, apresenta em 1860 uma primeira generalização da lei de Gompertz.<sup>33</sup> O modelo, conhecido na literatura por lei de Makeham, considera que existem outras influências sobre a mortalidade para além do envelhecimento, nomeadamente causas accidentais, cuja incidência é, em certa medida, independente da idade do indivíduo.

---

<sup>31</sup>Para uma análise detalhada do modelo de Gompertz recomenda-se a leitura de Pollard e Valkovics (1992), Carriere (1992, 1994), Olshansky e Carnes (1997), Willemse e Koppelaar (2000), Willekens (2001) e referências citadas.

<sup>32</sup>Veja-se Carriere (1992) e Willemse e Koppelaar (2000) para duas reformulações alternativas da lei de Gompertz.

<sup>33</sup>Gompertz (1960) desenvolve, ele próprio, uma segunda formulação destinada a representar a mortalidade em todo o arco da vida humana. O modelo estava assente na seguinte parametrização da função de sobrevivência

$$S(x) = -Bc^x + GH^x - xDF^x - JK^{m^x(x-n)}.$$

Trata-se, como se observa, de uma fórmula assente em dez parâmetros, bastante mais complexa e difícil de estimar.

“The formula for  $F_x$  (“ $\mu_x$ ”), according to Mr. Gompertz is  $Bq^x$ . For this I propose to substitute  $A + Bq^x$ , where  $A$  is the sum of certain partial forces which we assume to be, in the aggregate, of equal amount at all ages.”  
(Makeham, 1867, pp. 335)

A solução encontrada por Makeham para incorporar estes factores na explicação da mortalidade passou pela inclusão de um termo constante  $\theta$  em (2.5.22), i.e.,

$$\mu_x = \theta + \alpha \exp(\beta x) \quad x \geq 0, \quad (2.5.24)$$

com  $\alpha, \beta > 0$  e  $\theta \geq 0$ .<sup>34</sup> No caso particular em que  $\theta = 0$  obtemos, naturalmente, a lei de Gompertz. No caso em que  $\beta = 0$ , a equação (2.5.24) traduz uma força de mortalidade constante, uma hipótese irrealista como referimos anteriormente a propósito da primeira lei de Dormoy. A função de sobrevivência correspondente à lei de Makeham é definida por

$$S(x) = \exp \left[ \frac{\alpha}{\beta} (1 - e^{\beta x}) - \theta x \right]. \quad (2.5.25)$$

À semelhança da lei de Gompertz, a lei de Makeham revelou-se desajustada para descrever a mortalidade em todo o espectro da vida humana, sendo aplicável apenas às idades adultas. Ainda assim, o modelo de Makeham é pioneiro na medida em que sugere, pela primeira vez, a estruturação da força de mortalidade por causas de morte.<sup>35</sup>

Após os trabalhos de Makeham, surgiram múltiplas tentativas de adaptar a lei de Gompertz. O próprio Makeham (1889) generaliza o seu modelo adicionando um termo linear à equação (2.5.24), dando assim lugar à, menos conhecida e divulgada, *segunda Lei de Makeham*

$$\mu_x = \theta + \phi x + \alpha \exp(\beta x) \quad x \geq 0, \quad (2.5.26)$$

com  $\alpha, \beta > 0$  e  $\theta \geq 0$ . A função de sobrevivência correspondente é definida por

$$S(x) = \exp \left\{ - \left[ \theta x + \frac{1}{2} \phi x^2 - \frac{\alpha}{\beta} (1 - e^{\beta x}) \right] \right\}. \quad (2.5.27)$$

---

<sup>34</sup>Na formulação original, a lei de Makeham baseava-se na seguinte equação

$$\mu_x = A + Bc^x,$$

com  $A, B > 0$  e  $c > 1$ .

<sup>35</sup>A análise da mortalidade por causas de morte constitui hoje um campo de investigação importante na demografia, representando igualmente uma área de interesse crescente no domínio actuarial, nomeadamente em relação a formas de seguro onde é desejável a discriminação do capital seguro não apenas segundo a duração da vida humana, mas também de acordo com a causa de morte.

## 2.5. Leis de mortalidade

---

Lidstone (1935), King e Reid (1936) e Olifiers (1937) formulam uma lei duplamente geométrica com base na seguinte função de cinco parâmetros

$$\mu_x = A + Bc^x + Mn^x, \quad (2.5.28)$$

numa generalização do modelo proposto uns anos antes por Henderson (1915) para o caso em que  $A \neq 0$ .

Reformulações adicionais da fórmula de Gompertz-Makeham podem ainda ser encontradas em Van der Maen (1943),<sup>36</sup>

$$\mu_x = \theta + \phi x + \eta x^2 + \frac{I}{N-x} \quad (2.5.29)$$

$$\mu_x = \theta + \beta c^x + \frac{c}{N-x} \quad (2.5.30)$$

e em El Shaarawi *et al.* (1974), numa versão quadrática da equação de Gompertz

$$\mu_x = \theta + \alpha \exp(\beta x + \xi x^2) \quad x \geq 0, \quad (2.5.31)$$

com  $\alpha, \beta > 0$  e  $\theta \geq 0$ . Em alternativa, Le Bras (1976) e Keyfitz (1982) propõem a utilização de exponenciais de polinómio para a força de mortalidade

$$\mu_x = \theta + \alpha \exp\left(\sum_{k=0}^n \beta_k x^k\right) \quad x \geq 0, \quad (2.5.32)$$

com  $\alpha > 0$  e  $\theta \geq 0$ .

Gavrilov e Gavrilova (1991) desenvolvem um modelo, que apelidam de *lei de mortalidade binomial generalizada*, formulado através da seguinte equação

$$\mu_x = \theta + (\alpha + \beta x)^n \quad x, n \geq 0. \quad (2.5.33)$$

### Lei de Quiquet

Quiquet (1893) propõe um modelo assente na seguinte representação paramétrica

$$\mu_x = \sum_{i=1}^n P_i(x) \exp(\beta_i x), \quad (2.5.34)$$

---

<sup>36</sup>Citado por Tabeau (2001).

## 2.5. Leis de mortalidade

---

onde  $P_i(x)$  representa um polinómio de grau  $k$  ( $k = 0, \dots, K$ ) e  $\beta_i$  é constante para cada  $i$ . Registe-se, em particular, que com polinómios de grau 0 ou 1 e valores apropriados para as constantes  $\beta_i$  é possível deduzir a lei de Gompertz (com  $n = 1$ ), a lei de Makeham (com  $n = 2$ ) e a lei da Lazarus e segunda lei de Makeham (com  $n = 3$ ).

### Lei de Lazarus

Lazarus (1862) expande a lei de Makeham com a adição de um termo exponencial, do tipo “Gompertz negativo” a respeito da modelação da mortalidade nas idades infantis e juvenis

$$\mu_x = \alpha_1 \exp(-\beta_1 x) + \alpha_3 + \alpha_2 \exp(\beta_2 x) \quad x \geq 0, \quad (2.5.35)$$

com  $\alpha_i, \beta_i > 0$  ( $i = 1, 2$ ) e  $\alpha_3 \geq 0$ . A função de sobrevivência correspondente é definida, neste caso, por

$$S(x) = \exp \left\{ -\alpha_3 x - \frac{\alpha_1}{\beta_1} (1 - e^{-\beta_1 x}) + \frac{\alpha_2}{\beta_2} (1 - e^{\beta_2 x}) \right\}. \quad (2.5.36)$$

Um século mais tarde, Siler (1979, 1983) desenvolve um modelo de causas de morte competitivas para a mortalidade animal que é virtualmente idêntico ao modelo de Lazarus.

### Lei de Oppermann

Oppermann (1870) adopta uma representação alternativa para a mortalidade infantil e juvenil, nomeadamente

$$\mu_x = \theta_1 x^{-1/2} + \theta_2 + \theta_3 x^{1/2} \quad x \in [0, 20], \quad (2.5.37)$$

donde resulta a seguinte expressão para a função de sobrevivência

$$S(x) = \exp \left[ - \left( 2\theta_1 x^{1/2} + \theta_2 x + \frac{2}{3} \theta_3 x^{3/2} \right) \right]$$

Em alternativa, Wittstein (1883) propõe a seguinte expressão para a função  $q_x$

$$q_x = \frac{1}{m} a^{-(mx)^n} + a^{-(\omega-x)^n}, \quad (2.5.38)$$

onde  $\omega$  denota a idade máxima admissível e  $a$  e  $m$  são parâmetros a estimar. A lei de Wittstein é desenhada de modo a reflectir a mortalidade decrescente na infância, seguida de um ligeiro aumento na adolescência e de um aumento crescente nas idades avançadas.

### Lei de Thiele

Thiele (1872) sugere uma parametrização da força de mortalidade que cobre todas as idades que compõem o arco da vida humana. A sua abordagem assenta no pressuposto de que as causas de morte podem ser classificadas em três categorias distintas, que exercem a sua influência de forma particular sobre a mortalidade infantil, juvenil e adulta, respectivamente. Em termos formais, a lei é definida por

$$\mu_x = \alpha_1 \exp(-\beta_1 x) + \alpha_2 \exp\left(-\frac{1}{2}\beta_2(x - \xi)^2\right) + \alpha_3 \exp(\beta_3 x) \quad x \geq 0, \quad (2.5.39)$$

com  $\alpha_i, \beta_i > 0$  ( $i = 1, 2$ ) e  $\alpha_3, \beta_3, \xi \geq 0$ .

O primeiro termo no lado esquerdo de (2.5.39) representa a mortalidade decrescente nas idades mais jovens. O segundo termo, que traduz genericamente a função de densidade de probabilidade para uma distribuição normal, é usado para modelar a mortalidade juvenil, muito afectada pela mortalidade acidental (o denominado *accident hump*). O terceiro termo espelha a mortalidade nas idades adultas, e replica a tradicional lei de Gompertz. Registe-se ainda que, no caso particular em que  $\beta_1 = \alpha_2 = 0$ , obtemos a lei de Makeham.

Mode e Busby (1982) adoptam a formulação original de Thiele mas substituem o segundo termo por uma função do tipo parabólico

$$\mu_x = \begin{cases} \alpha_1 \beta_1 \exp(-\beta_1 x) & x \in [0, 10] \\ \alpha_2 - \beta_2 (x - \xi)^2 & x \in (10, 30] \\ \alpha_3 + \gamma_3 \beta_3 \exp(-\beta_3 x) & x \geq 30. \end{cases} \quad (2.5.40)$$

Em alternativa, Rogers e Planck (1984) propõem a adição de um termo independente da idade e a substituição do segundo termo por uma função duplamente exponencial

$$\mu_x = \alpha_0 + m_1(x) + m_2(x) + m_3(x) \quad x \geq 0, \quad (2.5.41)$$

onde

$$\begin{aligned} m_1(x) &= \alpha_1 \exp(-\beta_1 x) \\ m_2(x) &= \alpha_2 \exp\{-\beta_2(x - \xi) - \exp[-\lambda_2(x - \xi)]\} \\ m_3(x) &= \alpha_3 \exp(\beta_3 x). \end{aligned}$$

Alguns anos mais tarde, Rogers e Little (1994) agregam todas as tentativas ante-

## 2.5. Leis de mortalidade

---

rios e desenvolvem um modelo multiexponencial, suficientemente geral para capturar as múltiplas configurações possíveis da mortalidade ao longo de todo o arco da vida humana. O modelo assenta na seguinte especificação paramétrica

$$\mu_x = \alpha_0 + m_1(x) + m_2(x) + m_3(x) + m_4(x) \quad x \geq 0, \quad (2.5.42)$$

onde

$$\begin{aligned} m_1(x) &= \alpha_1 \exp(-\beta_1 x) \\ m_2(x) &= \alpha_2 \exp\{-\beta_2(x - \xi) - \exp[-\lambda_2(x - \xi)]\} \\ m_3(x) &= \alpha_3 \exp\{-\beta_3(x - \eta) - \exp[-\lambda_3(x - \eta)]\} \\ m_4(x) &= \alpha_4 \exp(\beta_4 x). \end{aligned}$$

As tábuas de mortalidade Inglesas (ELT 11 e 12), relativas ao período 1950-52 e 1960-62, foram graduadas usando a seguinte função de sete parâmetros

$$m_x = \theta_1 + \theta_3 e^{-\beta(x-x_2)^2} + \frac{\theta_2}{1 + e^{-\alpha(x-x_1)}} \quad (2.5.43)$$

aplicada num intervalo de idades limitado (no caso da ELT 12, para  $x \geq 27$ ).

### Lei Logística

Em meados dos anos 30 do século XX, os demógrafos interessaram-se pelo ajustamento de tábuas de mortalidade com recurso a uma função logística. Os diferentes modelos propostos na literatura podem ser encaixados na seguinte formulação geral

$$\mu_x = \theta_1 + \frac{\theta_2 \exp(\theta_3 x)}{1 + \sigma^2 \frac{\theta_2}{\theta_3} (\exp(\theta_3 x) - 1)}, \quad (2.5.44)$$

com quatro parâmetros  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  e  $\sigma^2$ . Note-se, em particular, que fixando  $\sigma^2 = 0$ , a lei logística se reduz à lei de Makeham. O efeito do denominador em (2.5.44) é claramente o de atenuar o crescimento exponencial decorrente do termo Gompertziano no numerador, sobretudo nas idades mais avançadas, uma característica bem patente na mortalidade actual.

Perks (1932), um actuário britânico, é o primeiro a recorrer a uma lei do tipo (2.5.44). O autor reconhece que a  $\mu_x$  não aumenta indefinidamente com a idade (como pressupõe a lei de Gompertz) e detecta um padrão de desaceleração da mortalidade a partir dos 84 anos de idade.

## 2.5. Leis de mortalidade

---

Mais tarde, Beard (1959, 1971) desenvolve um modelo em que admite que o risco de morte segue uma trajetória exponencial crescente (segundo a lei de Gompertz), reconhecendo contudo que a desaceleração da mortalidade nas idades adultas é explicável pelas diferenças nos níveis iniciais de risco dos indivíduos (heterogeneidade). O modelo produz uma função logística para a mortalidade que se aproxima da função exponencial nas idades mais jovens, desacelerando progressivamente até atingir um patamar nas idades extremas. Formalmente, fixando  $\theta_1 = 0$  o autor reduz (2.5.44) a três parâmetros usando<sup>37</sup>

$$\mu_x = \frac{\phi_1 \exp(\phi_2 x)}{1 + \phi_3 \exp(\phi_2 x)}, \quad (2.5.45)$$

onde  $\phi_1, \phi_2, \phi_3 \geq 0$ .

Mais recentemente, Kannistö (1992) simplifica a expressão (2.5.44) mantendo apenas dois parâmetros

$$\mu_x = \frac{\phi_1 \exp(\phi_2 x)}{1 + \phi_1 (\exp(\phi_2 x) - 1)}, \quad (2.5.46)$$

onde  $\phi_1, \phi_2 \geq 0$ . A lei (2.5.46) é um caso particular de (2.5.44) fixando  $\theta_1 = 0$  e  $\sigma^2 = \theta_3$ . Conforme veremos no capítulo 6, o modelo proporciona, nalguns casos, estimativas adequadas para o comportamento da mortalidade nas idades elevadas.

### Lei de Weibull

A força de mortalidade pode ser representada pelo modelo de Weibull (1939)

$$\mu_x = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1}, \quad (2.5.47)$$

onde  $\alpha > 0$  é um parâmetro de forma e  $\beta > 0$  um parâmetro de escala.

A lei de Weibull tem a sua origem em modelos aplicados na análise da resistência dos materiais e da duração de vida dos equipamentos. À semelhança dos engenhos que deixam de funcionar quando um dos seus componentes principais avaria, assume-se que o organismo humano se extingue quando uma das suas funções vitais deixa de funcionar.

---

<sup>37</sup>Martinelle (1987) sugere uma generalização dos modelo de Perks e Beard mediante a adição de um termo do tipo Gompertziano de modo a obter uma representação mais adequada para as idades adultas. O modelo é definido por

$$\mu_x = \frac{\theta_1 + \theta_2 e^{\theta_3 x}}{1 + \theta_4 e^{\theta_3 x}} + \theta_5 e^{\theta_3 x}.$$

A probabilidade e a função de sobrevivência associadas a (2.5.47) são dadas por

$${}_t p_x = \exp\left(\left(\frac{x}{\beta}\right)^\alpha - \left(\frac{x+t}{\beta}\right)^\alpha\right) \quad (2.5.48)$$

$$S(x) = \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right). \quad (2.5.49)$$

Prova-se ainda facilmente que a vida mediana  $vm_x$  é calculada mediante

$$vm_x = \beta \left( \left( \frac{x}{\beta} \right)^\alpha + \ln 2 \right)^{\frac{1}{\alpha}} - x. \quad (2.5.50)$$

### Lei de Brillinger

Brillinger (1961) desenvolve um modelo explicativo para a evolução da mortalidade onde sustenta que o óbito é o resultado da quebra num dos subsistemas do corpo humano indispensáveis à vida. Em particular, o autor admite que o corpo humano é composto por um conjunto de componentes cuja durabilidade é independente e identicamente distribuída, e que em consequência a probabilidade  $q_x$  é dada pela soma das probabilidades de quebra de cada um dos subsistemas que compõem o corpo humano. O modelo estipula a seguinte expressão geral para a força da mortalidade

$$\mu_x = \sum_i H_i (x - B_i)^{c_i - 1} + \sum_j \frac{A_j}{(b_j - x)^{c_j + 1}} + \sum_k E_k d_k^x, \quad (2.5.51)$$

onde os parâmetros são função de um conjunto de componentes e podem respeitar apenas a alguns intervalos de idades. Prova-se que o modelo contém, como casos particulares, a lei de Gompertz e a primeira e segunda leis de Makeham.

### 2.5.3 Modelos de ajustamento para quocientes de funções de mortalidade

Algumas das formulações analíticas exprimem o andamento da mortalidade em termos do quociente  $q_x/p_x$  (*"mortality odds" ratio*). Note-se que definindo  $\phi(x) = q_x/p_x$ , podemos escrever

$$q_x = \frac{\phi(x)}{1 + \phi(x)},$$

donde resulta que, para uma dada função  $\phi(x) \geq 0$ , se garante que  $0 \leq q_x < 1$ . Esta propriedade proporciona uma ampla liberdade na escolha de  $\phi(x)$  e explica, em boa medida, o interesse por modelos com esta particularidade.

### Modelo do CMIB

O CMIB<sup>38</sup> (1974) propõe a seguinte fórmula com quatro parâmetros para graduar a mortalidade das pessoas seguras no Reino Unido no período 1967-70

$$\frac{q_x}{p_x} = \theta - \phi x + \alpha \beta^x, \quad (2.5.52)$$

com  $\theta, \phi, \alpha, \beta > 0$ . O modelo tem a particularidade de consentir uma diminuição na mortalidade para o início do intervalo das idades adultas e pode ser confrontado com a segunda lei de Makeham expressa em (2.5.26).

Wilkie (1976) e CMIB (1976) propõem uma família de modelos para estudar a mortalidade dos beneficiários de rendas vitalícias e pensões, assente na seguinte representação

$$q_x = \frac{f(x)}{1 + f(x)},$$

onde

$$f(x) = \exp \left( \sum_{i=1}^s \alpha_i x^{i-1} \right). \quad (2.5.53)$$

Os estudos empíricos realizados para o período 1967-70 permitiram concluir que o modelo revela um comportamento especialmente adequado nas idades mais avançadas. Forfar *et al.* (1988) propõem uma família de modelos do tipo Gompertz-Makeham incorporando tantos parâmetros quantos os necessários para ajustar a mortalidade, em que a equação (2.5.53) é um caso particular. Pela sua relevância, este modelo será analisado em detalhe no capítulo seguinte.

### Lei de Heligman-Pollard

Numa das mais recentes tentativas de representar a mortalidade ao longo de todo o arco da vida humana usando uma única expressão analítica, Heligman e Pollard (1980) propõem a seguinte fórmula

$$\frac{q_x}{p_x} = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + GH^x, \quad (2.5.54)$$

onde  $A, B, \dots, H$  são parâmetros positivos a estimar. A ideia subjacente ao modelo, que como se observa apresenta uma estrutura semelhante à lei de Thiele, é a de que as causas de morte podem ser subdivididas em três classes, nomeadamente aquelas que afectam

---

<sup>38</sup>Continuous Mortality Investigation Bureau.

## 2.5. Leis de mortalidade

---

a mortalidade infantil, a juvenil e a das idades mais avançadas. O modelo proporciona bons resultados em termos de ajustamento e será analisado em detalhe no Capítulo 3.

### 2.5.4 Modelos de ajustamento para a função de densidade

Uma solução alternativa para caracterizar a evolução da mortalidade consiste em desenvolver expressões analíticas para a função de densidade de probabilidade. De entre os inúmeros modelos propostos, destacamos o modelo de Lexis e o modelo exponencial.<sup>39</sup>

#### Modelo de Lexis

O denominado modelo de Lexis admite que a idade de morte do indivíduo segue uma distribuição do tipo normal (gaussiano), i.e.,

$$f_0(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x - x^{\max})^2\right\} \quad x^{\min} \leq x \leq x^{\max}, \quad (2.5.55)$$

onde  $x^{\max}$  representa o ponto máximo da curva de mortes (ponto de Lexis) e  $x^{\min}$  denota a idade mínima de aplicação da equação.

#### Modelo Exponencial

A função de densidade da distribuição exponencial é dada por

$$f_0(x) = \theta_1 \exp(-\theta_1 x), \quad (2.5.56)$$

onde  $\theta_1, x > 0$ . A função e a probabilidade de sobrevivência são dadas, respectivamente, por

$$S(x) = \exp(-\theta_1 x) \quad (2.5.57)$$

$${}_t p_x = \exp(-\theta_1 t). \quad (2.5.58)$$

O modelo gera taxas instantâneas de mortalidade independentes das idade, i.e.

$$\mu_x = \theta_1. \quad (2.5.59)$$

Refira-se, por fim, que a lei exponencial é um caso particular da lei de Makeham (2.5.24) com  $\alpha = 0$ .

---

<sup>39</sup>Para uma descrição mais exaustiva dos modelos propostos neste contexto recomenda-se a leitura de Elandt-Johnson e Johnson (1980).

### 2.5.5 Combinação de leis de mortalidade

Carriere (1992,1994) e Rempala e Szatzschneider (2004) propõem a graduação da mortalidade ao longo de todo o arco da vida humana com recurso a combinações de funções paramétricas. Os autores argumentam que desta forma é possível conjugar as diferentes formulações analíticas que melhor ajustam cada componente da mortalidade.

O modelo proposto é formalizado mediante

$${}_x p_0 = \sum_{k=1}^n w_k {}_x p_0^{(k)}, \quad \sum_{k=1}^n w_k = 1 \quad (2.5.60)$$

onde os ponderadores  $w_k$  ( $k = 1, \dots, n$ ) são interpretados como probabilidades de óbito devido a diferentes causas, as quais estão geralmente ligadas a um período específico da vida (infância, adolescência, idade adulta,...). A cada um destes ponderadores corresponde uma função de sobrevivência (para um recém-nascido)  ${}_x p_0^{(k)}$  com uma forma particular. Mais concretamente, Carriere (1992,1994) propõe associar às idades jovens uma lei de Weibull do tipo (2.5.47), enquanto que para as idades adultas sugere uma lei de Gompertz (2.5.22). O autor encontra alguma dificuldade em modelar a chamada “bossa de acidentes”.

## 2.6 Heterogeneidade e modelos de *frailty*

Regra geral, as tábuas de mortalidade não diferenciam os indivíduos a não ser em função do sexo. Existem, no entanto, outros factores que afectam a sua sobrevivência. Estudos empíricos realizados (e.g., Mesrine (2000), Morris *et al.* (1994)) demonstram que factores como o desemprego, o nível de estudos, a categoria sócio-profissional ou o estado civil influenciam o nível de mortalidade. Esta forte heterogeneidade potencia o fenómeno da selecção adversa e deve ser tida em conta pelas companhias de seguros.

Uma forma de levar em linha de conta os factores observáveis que influenciam a mortalidade passa, por exemplo, por considerar um modelo de regressão de Poisson, inspirado no modelo de riscos proporcionais de Cox (1972), captando a heterogeneidade residual através de um modelo de *frailty*.

Segundo este modelo, proposto por Vaupel *et al.* (1979), a força de mortalidade é dada por  $\mu_x = \theta \mu_x^{\text{ref}}$ , onde  $\mu_x^{\text{ref}}$  denota uma força de mortalidade de referência e  $\theta$  é uma variável aleatória positiva que dita o nível de risco relativo (não observável) do indivíduo em particular. Como se observa, a inclusão da heterogeneidade no modelo faz-se mediante a simples aplicação de um factor multiplicativo  $\theta$ .

## Capítulo 3

# Métodos Paramétricos de Graduação

*“Science is the great antidote to the poison of enthusiasm and superstition.”*  
(Adam Smith, 1723-1790)

### 3.1 Introdução

A literatura actuarial e demográfica consagra uma atenção especial à construção de tábuas de mortalidade, ou seja, à estimação dos quocientes de mortalidade  $q_x$  no intervalo de idades considerado. Partindo de uma estimativa bruta dos quocientes, calculada a partir da relação entre o número de óbitos registado e a respectiva exposição ao risco, actuários e demógrafos procuram muitas vezes suavizar estas quantidades de modo a melhor reflectir as características da mortalidade da população analisada, que admitem ser relativamente regular.

Este procedimento de alisamento das estimativas brutas é conhecido na literatura actuarial por graduação. Em termos gerais, a graduação compreende o conjunto de princípios e métodos mediante os quais um conjunto de probabilidades brutas é ajustado de modo a fornecer uma base adequada, quer para a realização de inferência estatística, quer para a realização de cálculos aplicados (e.g., para a construção de tábuas de mortalidade).

Na literatura actuarial (veja-se, e.g., Benjamin e Pollard (1980)), os métodos de graduação são habitualmente classificados em métodos de ajustamento gráfico, métodos de natureza paramétrica e métodos ditos não-paramétricos. Neste capítulo, centramos a nossa atenção exclusivamente nos métodos paramétricos.

## 3.2. Caracterização do problema

---

Os métodos paramétricos de graduação baseiam-se no pressuposto de que as estatísticas de mortalidade podem ser expressas sob a forma de uma função analítica da idade do indivíduo e de um conjunto restrito de parâmetros. Estes parâmetros são apurados com base nas estimativas brutas, calculadas a partir dos dados disponíveis, usando procedimentos estatísticos formais. Em suma, a abordagem paramétrica resume toda a informação disponível sobre o fenómeno num conjunto limitado de parâmetros.

Este capítulo está organizado da seguinte forma. Na Secção 3.2, efectuamos uma primeira caracterização do problema de graduação, delimitando claramente o modelo probabilístico adequado à estimação dos indicadores de mortalidade (quocientes ou taxas instantâneas) e à realização de inferência estatística, e precisando os conceitos de população alvo e de exposição ao risco. Na Secção 3.3, caracterizamos formalmente a abordagem paramétrica e discutimos a aplicação de diferentes critérios de optimização na obtenção de estimativas para os parâmetros do modelo. Em particular, descrevemos a aplicação dos métodos de máxima verosimilhança, do mínimo chi-quadrado e dos mínimos quadrados em problemas de graduação. Na Secção 3.4, examinamos a metodologia desenvolvida pelo *Continuous Mortality Investigation Bureau* do Instituto e da Faculdade de Actuários de Londres para graduar indicadores de mortalidade. Esta metodologia amplia os tradicionais modelos de Gompertz e Makeham, e a sua flexibilidade torna-a facilmente compatível com a utilização de modelos lineares generalizados na construção, quer de tábuas contemporâneas, quer de tábuas prospectivas. Na Secção 3.5, fornecemos uma breve introdução aos modelos lineares generalizados e discutimos a sua aplicação em problemas de graduação. Em particular, destacamos o facto de estes modelos facultarem um tratamento unificado de um conjunto de distribuições de probabilidade (incluídas na chamada família exponencial natural) habitualmente usadas para modelar a mortalidade, investigamos a sua aplicação nos casos em que o predictor é linear e nos casos em que este é não-linear, e estudamos a denominada abordagem dual. Na Secção 3.6, analisamos em detalhe o modelo paramétrico assente nas leis de mortalidade de Heligman-Pollard. Por último, na Secção 3.7 consideramos a utilização de um conjunto de indicadores estatísticos para aferir a qualidade de ajustamento proporcionada pelos métodos paramétricos.

## 3.2 Caracterização do problema

A graduação de dados de mortalidade por intermédio de métodos paramétricos inicia-se com a definição da população alvo e com a delimitação do período temporal do estudo. Regra geral, assume-se que a população sob observação é homogênea, excepto no que se

### 3.2. Caracterização do problema

refere à idade dos indivíduos, ao seu sexo, à duração das apólices ou, nalguns casos, à presença de determinados factores de risco específicos.

Posto isto, para estimar  $q_x$  devemos observar os indivíduos desde a idade  $x$  até à idade  $x + 1$ , procedendo ao registo do número de óbitos (e outras características relevantes), subdividindo-os, regra geral, por idade mais próxima do aniversário. Se todas as observações forem completas, entendendo-se por tal que foram seguidas desde o início do estudo (idade  $x$ ) até ao fim do período de análise (idade  $x + 1$ ), a análise simplifica-se bastante. Contudo, nos estudos em que o objecto é uma população de pessoas seguras, existe a possibilidade real de se registarem entradas ou saídas do grupo sob observação, ou até mesmo a perda de informação sobre um indivíduo, antes da sua morte.

Designemos por  $x + t_i$  a idade em que o  $i^{\text{ésimo}}$  indivíduo é incluído no estudo e por  $x + s_i$  a idade em que este é retirado do mesmo, com  $0 \leq t_i \leq s_i \leq 1$ . Neste sentido, o tempo de permanência de cada elemento na amostra é dado por  $(s_i - t_i)$ , enquanto a soma  $E_x = (s_1 - t_1) + (s_2 - t_2) + \dots + (s_n - t_n)$  define a *exposição* total dos  $n$  indivíduos.

Denotemos por  $D_x$  a variável aleatória que representa o número de óbitos e por  $d_x$  o seu número efectivamente observado. Seja  $I$  o conjunto de subíndices correspondente aos indivíduos que acabam por falecer. Para estimar  $q_x$  a partir dos dados observados, dispomos de diferentes alternativas.

1. A primeira consiste em igualar os óbitos esperados aos seus valores observados. O valor esperado de  $D_x$  é dado por

$$\sum_{i=1}^n {}_{1-t_i}q_{x+t_i} - \sum_{i \notin I} {}_{1-s_i}q_{x+s_i}, \quad (3.2.1)$$

onde  ${}_{1-t_i}q_{x+t_i}$  ( ${}_{1-s_i}q_{x+s_i}$ ) denota a probabilidade de um indivíduo com idade  $x + t_i$  ( $x + s_i$ ) não alcançar a idade  $x + 1$ .<sup>1</sup>

Se recorrermos à hipótese de Balducci (2.2.28) para idades fraccionadas,  ${}_{1-u}q_{x+u} = (1 - u)q_x$ , e considerarmos que  $E_x$  pode expressar-se da seguinte forma

$$\begin{aligned} E_x &= [(1 - t_1) - (1 - s_1)] + [(1 - t_2) - (1 - s_2)] + \dots + [(1 - t_n) - (1 - s_n)] \\ &= \sum_{i=1}^n (1 - t_i) - \sum_{i=1}^n (1 - s_i), \end{aligned}$$

<sup>1</sup>Note-se que somando todas as probabilidades  ${}_{1-t_i}q_{x+t_i}$  obteríamos a duração de vida total na eventualidade de observarmos todos os óbitos. Como os dados são incompletos, a este valor é necessário subtrair o tempo de exposição contabilizado a mais relativo aos indivíduos para os quais não é possível observar o óbito, ou seja, a soma das probabilidades de que estes não alcançarão a idade  $x + 1$ , depois de terem atingido  $x + s_i$  ( ${}_{1-s_i}q_{x+s_i}$ ).

### 3.2. Caracterização do problema

---

a expressão (3.2.1) dá lugar a

$$\begin{aligned}
 & \sum_{i=1}^n (1 - t_i) q_x - \sum_{i \notin I} (1 - s_i) q_x = \\
 &= q_x \left( \sum_{i=1}^n (1 - t_i) - \sum_{i \notin I} (1 - s_i) \right) \\
 &= q_x \left( \sum_{i=1}^n (1 - t_i) + \sum_{i \in I} (1 - s_i) - \sum_{i \in I} (1 - s_i) - \sum_{i \notin I} (1 - s_i) \right) \\
 &= q_x \left( E_x + \sum_{i \in I} (1 - s_i) \right)
 \end{aligned}$$

pelo que

$$D_x = q_x \left( E_x + \sum_{i \in I} (1 - s_i) \right).$$

Igualando os óbitos esperados aos valores observados, obtemos o estimador clássico de  $q_x$

$$\hat{q}_x = \frac{d_x}{E_x + \sum_{i \in I} (1 - s_i)}. \quad (3.2.2)$$

O denominador da expressão (3.2.2) define o número de indivíduos *inicialmente expostos ao risco*. Se admitirmos, por hipótese, que os óbitos ocorrem aproximadamente a meio do intervalo  $(x, x + 1]$ , i.e., à idade  $x + \frac{1}{2}$ , o número de expostos inicialmente é estimado por

$$E_x + \frac{1}{2}d_x. \quad (3.2.3)$$

O símbolo  $E_x$  designa o número de indivíduos *expostos centralmente ao risco*. Conforme verificaremos em seguida, as definições alternativas do conceito de exposição ao risco correspondem, naturalmente, estimativas distintas para os indicadores de mortalidade  $q_x$  e  $\mu_x$ . Regra geral, assume-se que o estimador (3.2.2) é adequado apenas nos casos em que o número de observações é significativo. No pressuposto de que as hipóteses usadas para o derivar são adequadas, demonstra-se igualmente que o estimador é não enviesado (London, 1997).

2. Um método alternativo consiste, como vimos no capítulo anterior, em admitir que a força de mortalidade é constante em cada intervalo unitário. Designemos simplesmente por  $\mu_x$  o valor constante de  $\mu_{x+u}$  ( $u \in [0, 1]$ ). Neste caso, o número

### 3.2. Caracterização do problema

---

esperado de óbitos é dado por  $\mathbb{E}[D_x] = E_x \mu_x$ . Igualando de novo os valores de  $d_x$  e  $\mathbb{E}[D_x]$ , obtém-se um estimador para a força de mortalidade

$$\hat{\mu}_x = \frac{d_x}{E_x}, \quad (3.2.4)$$

com base no qual e na hipótese (2.2.30), é possível derivar uma estimativa para o quociente  $q_x$

$$\hat{q}_x = 1 - \exp(-\hat{\mu}_x) = 1 - \exp\left(-\frac{d_x}{E_x}\right). \quad (3.2.5)$$

Na realidade, tanto  $D_x$  como  $E_x$  são variáveis aleatórias. Apesar disso, e por razões de conveniência, a quantidade  $E_x$  é muitas vezes tratada como não aleatória, sendo o seu valor estimado com base nas estatísticas disponíveis.

#### 3.2.1 Estimação dos quocientes de mortalidade

Considere-se um conjunto  $E_x$  de indivíduos expostos inicialmente ao risco. Admitamos que a probabilidade de morte no intervalo de tempo unitário (ano) é definida por  $q_x$ , e que o óbito ou sobrevivência de cada um dos indivíduos desta população é independente do óbito ou sobrevivência dos restantes.

Neste caso, é apropriado considerar que  $D_x$ , a variável aleatória que representa o número de óbitos que ocorrem durante o ano, segue uma distribuição Binomial com parâmetros  $E_x$  e  $q_x$ . Nesse caso, a probabilidade de se registarem  $d_x$  óbitos é dada por

$$\mathbb{P}(D_x = d_x) = \left( \frac{E_x!}{d_x!(E_x - d_x)!} \right) q_x^{d_x} (1 - q_x)^{E_x - d_x} \quad (3.2.6)$$

com valor esperado  $\mathbb{E}(D_x) = E_x q_x$  e variância  $\text{Var}(D_x) = E_x q_x (1 - q_x)$ .

Apesar de o valor de  $q_x$  ser desconhecido, sabemos que, de um total de  $E_x$  indivíduos expostos ao risco,  $d_x$  falecem, donde se infere que  $E_x - d_x$  sobrevivem. O primeiro factor em (3.2.6) não depende de  $q_x$ , pelo que podemos simplificar a expressão e descrever a função de verosimilhança de  $d_x$  em termos de  $q_x$ ,  $\mathcal{L}(q_x)$ , onde

$$\mathcal{L}(q_x) = q_x^{d_x} (1 - q_x)^{E_x - d_x}$$

ou, mediante a correspondente transformação logarítmica  $\mathcal{L}^*(q_x) = \ln \mathcal{L}(q_x)$ ,

$$\mathcal{L}^*(q_x) = d_x \ln(q_x) + (E_x - d_x) \ln(1 - q_x). \quad (3.2.7)$$

### 3.2. Caracterização do problema

Da maximização de (3.2.7), obtém-se o seguinte estimador de máxima verosimilhança de  $q_x$

$$\hat{q}_x = \frac{d_x}{E_x}, \quad (3.2.8)$$

cujas propriedades de não-enviesamento e consistência são bem conhecidas. De facto, demonstra-se que o valor esperado e a variância de  $\hat{q}_x$  são dados por  $\mathbb{E}[\hat{q}_x] = q_x$  e por  $\text{Var}[\hat{q}_x] = q_x(1 - q_x)/E_x$ , respectivamente.

Nos casos em que o número de óbitos observado é suficientemente grande relativamente ao número de indivíduos expostos ao risco podemos, recorrendo à aproximação da distribuição Binomial pela distribuição Normal, derivar uma solução apropriada para os intervalos de confiança das estimativas dos parâmetros. Com efeito, nestas condições o intervalo de confiança para  $q_x$  a  $(1 - 2\alpha) \cdot 100\%$  é aproximado por

$$\left[ \hat{q}_x - z_\alpha \left( \frac{\hat{q}_x(1 - \hat{q}_x)}{E_x} \right)^{1/2}, \hat{q}_x + z_\alpha \left( \frac{\hat{q}_x(1 - \hat{q}_x)}{E_x} \right)^{1/2} \right],$$

onde  $z_\alpha$  denota o quantil de ordem  $(1 - \alpha)$  da distribuição Normal padronizada.

Nos casos em que  $d_x$  é muito pequeno relativamente a  $E_x$ , a aproximação não é adequada, sendo por isso necessário recorrer às propriedades da distribuição Binomial para obter um intervalo de confiança exacto,  $q_x^{\text{inf}} \leq q_x \leq q_x^{\text{sup}}$ , i.e.,<sup>2</sup>

$$q_x^{\text{inf}} \text{ tal que } \sum_{k=d_x}^{E_x} \frac{E_x!}{d!(E_x - k)!} (q_x^{\text{inf}})^k (1 - q_x^{\text{inf}})^{E_x - k} = \alpha, \quad d_x > 0$$

$$q_x^{\text{sup}} \text{ tal que } \sum_{k=0}^{d_x} \frac{E_x!}{k!(E_x - k)!} (q_x^{\text{sup}})^k (1 - q_x^{\text{sup}})^{E_x - k} = \alpha, \quad d_x < E_x.$$

Note que  $q_x^{\text{inf}} = 0$  quando  $d_x = 0$ , e que  $q_x^{\text{sup}} = 1$  no caso em que  $d_x = E_x$ . As soluções para estas equações podem ser obtidas recorrendo a métodos iterativos de aproximação sucessiva.

#### 3.2.2 Estimação da taxa de mortalidade instantânea

Considere-se um grupo de  $E_x$  indivíduos observados entre as idades  $x$  e  $x+1$ . Admitamos que os eventos (óbito ou sobrevivência) dos indivíduos são independentes entre si, e que a força de mortalidade é constante no período de observação. A quantidade  $E_x$  representa, neste caso, o número de expostos centralmente ao risco. Nestas condições, a variável

<sup>2</sup>Para mais detalhes veja, por exemplo, Forfar *et al.* (1988).

### 3.2. Caracterização do problema

---

aleatória  $D_x$  segue uma distribuição de Poisson com média e variância  $\mu_x E_x$  (Brillinger, 1986). A probabilidade de ocorrerem  $d_x$  óbitos é dada por

$$\mathbb{P}(D_x = d_x) = \exp(-\mu_x E_x) \frac{(\mu_x E_x)^{d_x}}{d_x!} \quad (3.2.9)$$

Se ignorarmos os termos que não incluem  $\mu_x$  em (3.2.9), e aplicarmos novamente a transformação logarítmica obtemos a seguinte função de log-verosimilhança

$$\mathcal{L}^*(\mu_x) = -\mu_x E_x + d_x \ln(\mu_x)$$

Da maximização desta função, obtém-se o seguinte estimador de máxima verosimilhança de  $\mu_x$

$$\hat{\mu}_x = \frac{d_x}{E_x} \quad (3.2.10)$$

cujas propriedades de não-enviesamento e consistência estão, de novo, asseguradas.

Com efeito, prova-se que o valor esperado e a variância de  $\hat{\mu}_x$  são dados, respectivamente, por  $\mathbb{E}[\hat{\mu}_x] = \mu_x$  e por  $\text{Var}[\hat{\mu}_x] = \mu_x/E_x$ . Se o número de óbitos for razoavelmente grande, podemos assumir que  $D_x$  segue uma distribuição Normal com média e variância  $\mu_x E_x$  e derivar o seguinte intervalo de confiança para  $\mu_x$  a  $(1 - 2\alpha) \cdot 100\%$

$$\left[ \hat{\mu}_x - z_\alpha \left( \frac{\hat{\mu}_x}{E_x} \right)^{1/2}, \quad \hat{\mu}_x + z_\alpha \left( \frac{\hat{\mu}_x}{E_x} \right)^{1/2} \right]$$

De forma análoga ao referido a propósito da estimação de  $q_x$ , quando a aproximação não é adequada torna-se necessário recorrer às propriedades da distribuição de Poisson para derivar um intervalo de confiança exacto,  $\mu_x^{\text{inf}} \leq \mu_x \leq \mu_x^{\text{sup}}$ , onde

$$\begin{aligned} \mu_x^{\text{inf}} \quad \text{tal que} \quad & \sum_{k=d_x}^{\infty} \exp(-\mu_x^{\text{inf}} E_x) \frac{(\mu_x^{\text{inf}} E_x)^k}{k!} = \alpha, \quad d_x > 0 \\ \mu_x^{\text{sup}} \quad \text{tal que} \quad & \sum_{k=0}^{d_x} \exp(-\mu_x^{\text{sup}} E_x) \frac{(\mu_x^{\text{sup}} E_x)^k}{k!} = \alpha, \quad d_x < E_x \end{aligned}$$

Note que  $\mu_x^{\text{inf}} = 0$  quando  $d_x = 0$ , e que  $\mu_x^{\text{sup}} = 1$  no caso em que  $d_x = E_x$ .

#### 3.2.3 Distribuição da exposição ao risco

Sem prejuízo dos casos anteriores, é possível tratar  $D_x$  como uma quantidade não aleatória fixando o seu valor a partir do número de óbitos observados  $d_x$ , assumindo,

### 3.3. Estimação dos parâmetros

---

em contrapartida, que  $E_x$  segue uma distribuição Gamma com parâmetros  $\alpha = D_x$  e  $\beta = \mu_x$ . Esta interpretação alternativa é sugerida por Gerber (1995) e utilizada por Renshaw e Haberman (1996) e Renshaw *et al.* (1997) no contexto dos modelos lineares generalizados, que analisaremos mais adiante.

Admitindo que os  $n$  indivíduos sob observação são independentes, a função de verosimilhança das observações é dada por

$$\prod_{i \in I} \mu_{x+s_i} p_{x+t_i} \cdot \prod_{i \notin I} \mu_{x-t_i} p_{x+t_i}.$$

Se retomarmos a hipótese de que a força de mortalidade é constante em cada intervalo de tempo unitário, donde resulta  ${}_s q_x = 1 - \exp(-\mu_x s)$ , é possível simplificar esta equação para dar lugar a

$$(\mu_x)^{D_x} \exp(-\mu_x E_x), \quad (3.2.11)$$

expressão esta que é maximizada no ponto  $\hat{\mu}_x = D_x/E_x$ .

### 3.3 Estimação dos parâmetros

A representação da mortalidade por intermédio de leis de mortalidade atraiu, como vimos no capítulo 2, a atenção de actuários, demógrafos e académicos de outras áreas ao longo dos últimos três séculos.

Os métodos paramétricos baseiam-se no pressuposto de que os quocientes de mortalidade  $q_x$  (ou outras estatísticas de mortalidade apropriadas) podem ser expressos sob a forma de uma função  $f(x; \theta)$  da idade  $x$  e de um conjunto (limitado)  $\theta = (\theta_1, \theta_2, \dots, \theta_k)'$  de parâmetros desconhecidos. Estes parâmetros são derivados a partir das estimativas brutas  $\hat{Q} = \{\hat{q}_x, x = x_{\min}, \dots, x_{\max}\}$ , calculadas com base nas estatísticas de mortalidade e de população exposta ao risco disponíveis. Se representarmos por  $\hat{\theta}$  o estimador de  $\theta$ , os métodos paramétricos de graduação consistem em substituir a série  $\hat{Q}$  por

$$\hat{Q}^{gra} = \{\hat{q}_x^{gra} = f(x; \hat{\theta}), x = x_{\min}, \dots, x_{\max}\}.$$

Daqui resulta que toda a informação disponível é resumida num conjunto restrito de parâmetros. Conforme salientámos no capítulo anterior, são inúmeros os métodos paramétricos propostos na literatura.<sup>3</sup> Podemos, por exemplo, recorrer a uma das muitas

---

<sup>3</sup>Deve registar-se que as técnicas de graduação descritas neste ponto para alisamento dos dados de mortalidade não são exclusivas do campo actuarial, sendo largamente utilizadas em áreas tão distintas como a economia, as finanças, a biologia ou a física.

### 3.3. Métodos paramétricos

---

leis de mortalidade ali descritas. Na medida em que o modelo paramétrico seja bem escolhido, i.e., na medida em que este descreva adequadamente a mortalidade da população considerada, a abordagem paramétrica apresenta, como realçámos no capítulo anterior, um conjunto de atractivos face a metodologias alternativas. A selecção de um modelo desajustado tem, pelo contrário, consequências bastantes nefastas e pode conduzir a resultados desastrosos.

O procedimento de graduação mediante métodos paramétricos é balizado por dois objectivos nucleares. Em primeiro lugar, a necessidade de conseguir uma boa qualidade de ajustamento (*goodness of fit*) entre a série  $\hat{Q}^{gra}$  e as estimativas brutas  $\hat{Q}$ . Em segundo, a procura de uma certa regularidade ou alisamento (*smoothness*) para a série  $\hat{Q}^{gra}$  ao longo de todo o espectro de idades que compõem o arco da vida humana.<sup>4</sup>

Estes dois objectivos colidem, de certa forma, entre si. Com efeito, a obtenção de uma boa qualidade de ajustamento pode requerer um número significativo de parâmetros, um requisito que pode comprometer a regularidade das estimativas. Neste sentido, e como salienta Keyfitz (1982), na implementação desta abordagem é necessário encontrar um equilíbrio entre o número de parâmetros do modelo e a bondade do ajustamento.<sup>5</sup>

Para estimar o valor dos parâmetros da função  $f(x; \theta)$ , são considerados distintos critérios de optimização. Entre os principais, descrevemos nesta secção o método de máxima verosimilhança, o método do mínimo  $\chi^2$  e o método dos mínimos quadrados.

#### 3.3.1 Método de máxima verosimilhança

Este procedimento baseia-se na determinação do vector de parâmetros  $\theta$  que maximiza a função de verosimilhança de  $d_x$  no intervalo de idades  $x \in [x_{\min}, x_{\max}]$ , e será naturalmente distinto no caso de estarmos perante a graduação de  $q_x$  ou de  $\mu_x$ .

---

<sup>4</sup>Esta regularidade é uma característica muito valorizada no campo actuarial, entre outras razões pelo facto de se traduzir, regra geral, em prémios de seguros que apresentam uma progressão “razoável” com a idade do beneficiário da apólice.

<sup>5</sup>Muitos estudos de graduação demográfica enfatizam demasiado a qualidade do ajustamento sem considerar convenientemente a estabilidade estatística dos parâmetros (i.e., sem ter em conta os desvios-padrão e as correlações entre os parâmetros, a correcta especificação dos erros, ou de uma função de perda, etc.). Associado a este enfoque, surge muitas vezes o problema da sobreparametrização do modelo, posto em evidência, por exemplo, nos casos em que se observam desvios-padrão demasiado grandes para fornecer significância estatística aos parâmetros, correlações elevadas entre os parâmetros e dificuldades em fazer convergir as rotinas iterativas usadas para estimar os modelos não lineares. Acresce que a sobreparametrização tem implicações práticas sobre a técnica de graduação, em particular em termos de comparação das séries temporais dos parâmetros no tempo, e na elaboração de projecções demográficas. Neste último caso, a verificação de flutuações erráticas irregulares nos parâmetros dificulta a elaboração de previsões. Existe, com efeito, uma forte relação entre a sobreparametrização e a instabilidade nos parâmetros ao longo do tempo, pelo que do ponto de vista estatístico são muitas vezes preferidos os modelos analíticos parsimoniosos (com um reduzido número de parâmetros), ainda que isso implique o sacrifício de alguma qualidade de ajustamento.

### 3.3. Métodos paramétricos

---

#### Graduação de $q_x$

No caso da graduação de  $q_x$ , a distribuição da variável aleatória número de óbitos  $D_x$  pode ser descrita por  $D_x \sim \text{Bin}(E_x, q_x(\theta))$ , onde  $E_x$  denota os indivíduos expostos inicialmente ao risco e  $\text{Bin}(N, q)$  define a distribuição Binomial com parâmetros  $N$  e  $q$ . Seja

$$q_x = q_x(\theta) = f(x; \theta),$$

onde  $f(x; \theta)$  representa uma determinada função paramétrica.

Admitamos que os valores de  $d_x$  e  $E_x$  são conhecidos no intervalo  $x \in [x_{\min}, x_{\max}]$ , e que  $d_x/E_x$  denota o quociente de mortalidade bruto. A função de verosimilhança (desprezando os termos constantes) é definida por

$$\mathcal{L}(\theta) \propto \prod_{x=x_{\min}}^{x_{\max}} [q_x(\theta)]^{d_x} [1 - q_x(\theta)]^{E_x - d_x}. \quad (3.3.1)$$

Maximizar  $\mathcal{L}(\theta)$  equivale a maximizar o seu logaritmo  $\mathcal{L}^*(\theta) = \ln \mathcal{L}(\theta)$ , pelo que as estimativas do modelo são derivadas resolvendo o seguinte problema de otimização

$$\hat{\theta} = \arg \max_{\theta} \left\{ \mathcal{L}^*(\theta) \propto \sum_{x=x_{\min}}^{x_{\max}} \{d_x \ln [q_x(\theta)] + (E_x - d_x) \ln [1 - q_x(\theta)]\} \right\}. \quad (3.3.2)$$

#### Graduação de $\mu_x$

No caso da graduação de  $\mu_x$ , a distribuição da variável aleatória  $D_x$  pode ser descrita por  $D_x \sim \text{Poisson}(\mu_x E_x)$ , onde  $E_x$  representa agora os expostos centralmente ao risco e  $\text{Poisson}(\lambda)$  define a distribuição de Poisson com parâmetro  $\lambda$ . Seja

$$\mu_x = \mu_x(\theta) = f(x; \theta),$$

onde  $f(x; \theta)$  denota, uma vez mais, um determinado modelo paramétrico. A função de verosimilhança correspondente é expressa por

$$\mathcal{L}(\theta) = \prod_{x=x_{\min}}^{x_{\max}} \exp(-E_x \mu_x(\theta)) \frac{(E_x \mu_x(\theta))^{d_x}}{d_x!}, \quad (3.3.3)$$

enquanto que o seu logaritmo é dado por

$$\ln \mathcal{L}(\theta) = \sum_{x=x_{\min}}^{x_{\max}} [-E_x \mu_x(\theta) + d_x \ln (E_x \mu_x(\theta)) - \ln (d_x!)].$$

### 3.3. Estimação dos parâmetros

---

Como  $d_x$  e  $E_x$  não dependem de  $\theta$ , as estimativas do modelo são derivadas resolvendo o seguinte problema de optimização

$$\hat{\theta} = \arg \max_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} [-E_x \mu_x(\theta) + d_x \ln(E_x \mu_x(\theta))] \right\}. \quad (3.3.4)$$

#### 3.3.2 Método do mínimo Chi-Quadrado

A tradicional estatística  $\chi^2$  mede, neste contexto, a discrepância entre os óbitos observados e o seu valor esperado, no pressuposto de que o modelo probabilístico seleccionado para estimar as medidas  $q_x$  e  $\mu_x$  é adequado. Em termos práticos, o método assenta na estimação do vector de parâmetros  $\theta$  que minimiza o desvio global entre os valores observados e os valores produzidos pelo modelo  $f(x; \theta)$ .

#### Gradação de $q_x$

Neste caso, o número de óbitos na idade  $x$  segue uma distribuição Binomial com média  $E_x q_x(\theta)$  e variância  $E_x q_x(\theta)(1 - q_x(\theta))$ . O *desvio relativo* a cada idade  $x$ ,  $z_x$ , é definido por

$$z_x = \frac{d_x - E_x q_x(\theta)}{[E_x q_x(\theta)(1 - q_x(\theta))]^{1/2}}, \quad (3.3.5)$$

onde  $q_x(\theta) = f(x; \theta)$  e  $z_x$  segue aproximadamente uma distribuição Normal padronizada, i.e.,  $z_x \sim \mathcal{Nor}(0, 1)$ .

Representemos por  $\chi^2(\theta)$  a estatística

$$\chi^2(\theta) = \sum_{x=x_{\min}}^{x_{\max}} z_x^2 = \sum_{x=x_{\min}}^{x_{\max}} \frac{[d_x - E_x q_x(\theta)]^2}{E_x q_x(\theta)(1 - q_x(\theta))}. \quad (3.3.6)$$

O seu valor segue (aproximadamente) uma distribuição de Chi-quadrado com  $[(x_{\max} - x_{\min} + 1) - p - n]$  graus de liberdade, onde  $(x_{\max} - x_{\min} + 1)$  denota o número de idades objecto de gradação,  $p$  simboliza o número de parâmetros da função  $f(x; \theta)$  e  $n$  traduz o número de restrições impostas aos parâmetros, caso existam. As estimativas dos parâmetros produzidas por este método correspondem aos valores  $\theta$  que minimizam  $\chi^2(\theta)$ , i.e., a  $\hat{\theta} = \arg \min_{\theta} \{\chi^2(\theta)\}$ .<sup>6</sup>

---

<sup>6</sup> Alguns estudos sobre o comportamento da mortalidade da população segura baseiam-se em dados sobre o número de apólices e não em dados sobre o número de indivíduos expostos ao risco. Nestes casos, torna-se necessário introduzir uma correcção no cálculo da estatística  $\chi^2(\theta)$  de modo a considerar o chamado efeito de sobredispersão, ou seja, a possibilidade de um indivíduo deter duas ou mais apólices.

### 3.3. Estimação dos parâmetros

#### Graduação de $\mu_x$

Neste caso, o número de óbitos na idade  $x$  segue uma distribuição de Poisson com média e variância iguais a  $E_x \mu_x(\theta)$ . O desvio relativo a cada idade  $x$ ,  $z_x$ , é definido por

$$z_x = \frac{d_x - E_x \mu_x(\theta)}{[E_x \mu_x(\theta)]^{1/2}}, \quad (3.3.7)$$

onde  $\mu_x(\theta) = f(x; \theta)$  e  $z_x$  é aproximado por  $\mathcal{N}(0, 1)$ . A estatística  $\chi^2(\theta)$  correspondente é definida pela equação

$$\chi^2(\theta) = \sum_{x=x_{\min}}^{x_{\max}} z_x^2 = \sum_{x=x_{\min}}^{x_{\max}} \frac{[d_x - E_x \mu_x(\theta)]^2}{E_x \mu_x(\theta)}. \quad (3.3.8)$$

Na ausência de duplicidades, as estimativas dos parâmetros geradas pelo método são novamente os valores  $\theta$  que resolvem o problema de optimização  $\hat{\theta} = \arg \min_{\theta} \{\chi^2(\theta)\}$ .<sup>7</sup>

Seja  $r_x$  o “rácio de variância” na idade  $x$ , definido por

$$r_x = \frac{\sum_i i^2 f_x^i}{\sum_i i f_x^i},$$

onde  $f_x^i$  ( $i = 1, 2, \dots$ ) denota a proporção de indivíduos na idade  $x$  que detêm um total de  $i$  apólices de seguro, com  $\sum_i f_x^i = 1$ . Nos casos em que se registam duplicidades (i.e., para  $r_x > 1$ ), o valor esperado da distribuição mantém-se inalterado mas a variância é incrementada pelo factor  $r_x$ , i.e.,  $\text{Var}(D_x) = r_x E_x q_x(\theta) (1 - q_x(\theta))$ . Em consequência, a definição apropriada da estatística  $\chi^2(\theta)$  a usar neste caso é dada por

$$\chi^2(\theta) = \sum_{x=x_{\min}}^{x_{\max}} z_x^2 = \sum_{x=x_{\min}}^{x_{\max}} \frac{[d_x - E_x q_x(\theta)]^2}{r_x E_x q_x(\theta) (1 - q_x(\theta))}.$$

<sup>7</sup>Quando existem duplicidades na amostra, a variável  $d_x$  representa o número de apólices activadas na idade  $x$  durante o período de observação, enquanto  $E_x$  denota os expostos centralmente ao risco calculados a partir do número de apólices. Neste caso, é possível escrever,  $d_x = \sum_i i \eta_x^i$ , onde  $\eta_x^i$  denota o número de indivíduos que falecem na idade  $x$  e detêm um total de  $i$  apólices. De forma similar, podemos escrever  $E_x = \sum_i i \xi_x^i$ , onde  $\xi_x^i$  representa os expostos centralmente ao risco calculados com base no número de vidas, associados aos casos em que o segurado detém um total de  $i$  apólices. À semelhança do referido anteriormente para a graduação de  $q_x$ , demonstra-se (Forfar *et al.*, 1988) que a presença de duplicidades não altera o valor esperado da distribuição mas modifica a respectiva variância, i.e.,  $\text{Var}(D_x) = r_x E_x \mu_x(\theta)$ , onde

$$r_x = \frac{\sum_i i^2 \xi_x^i}{\sum_i i \xi_x^i}$$

Na presença de duplicidades, os parâmetros  $\theta$  são determinados pela minimização da estatística

$$\chi^2(\theta) = \sum_{x=x_{\min}}^{x_{\max}} z_x^2 = \sum_{x=x_{\min}}^{x_{\max}} \frac{[d_x - E_x \mu_x(\theta)]^2}{r_x E_x \mu_x(\theta)}.$$

#### 3.3.3 Método dos mínimos quadrados

A aplicação do tradicional método dos mínimos quadrados em problemas de graduação envolve a minimização da soma dos desvios (quadráticos) entre os valores brutos observados  $\hat{Q}$  e os valores ajustados pelo modelo paramétrico  $f(x; \theta)$ . Em termos formais, a estimativa dos parâmetros é o resultado do problema de otimização

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} (\hat{q}_x - f(x; \theta))^2 \right\}. \quad (3.3.9)$$

A estimativa gerada pelo método dos mínimos quadrados (não ponderados) só pode considerar-se adequada nos casos em que os  $\hat{q}_x$  apresentam idêntica variância ao longo de todo o intervalo de idades em análise, uma característica que não é confirmada pelos estudos empíricos. De facto, na maioria dos casos a variância oscila com a idade, pelo que é desejável que a curva graduada atribua maior importância (relativa) às idades com menor variabilidade.

A solução encontrada passa pela utilização do método dos mínimos quadrados ponderados, i.e., pela resolução do seguinte problema de otimização

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} w_x (\hat{q}_x - f(x, \theta))^2 \right\}, \quad (3.3.10)$$

onde os ponderadores  $w_x$  são, regra geral, expressos por um valor inversamente proporcional à variância do estimador.<sup>8</sup>

#### 3.3.4 Escolha do método de estimação

A opção por um destes três métodos de estimação depende, essencialmente, de critérios de natureza pessoal ou computacional porquanto, como veremos em seguida, os métodos proporcionam resultados muito semelhantes. O método de máxima verosimilhança apresenta, no entanto, uma vantagem face aos seus concorrentes, pelo facto de proporcionar uma estimativa da matriz de variâncias-covariâncias dos parâmetros do modelo.

Quando o número de indivíduos expostos ao risco é grande em todas as idades, a aproximação Normal à distribuição Binomial de  $q_x$  pode ser usada. Nesse caso, a função

---

<sup>8</sup>Por exemplo, no caso da graduação de  $q_x$  a sua variância é dada por  $q_x(1 - q_x)/E_x$ . Na prática, como  $q_x$  assume valores relativamente pequenos, a variância pode ser aproximada simplesmente por  $\frac{q_x}{E_x}$ .

### 3.3. Métodos paramétricos

---

de verosimilhança para  $d_x$  é dada por

$$\mathcal{L}(\theta) = \prod_{x=x_{\min}}^{x_{\max}} \frac{1}{(2\pi)^{1/2} (E_x q_x (1 - q_x))^{1/2}} \exp \left\{ -\frac{1}{2} \left( \frac{(d_x - E_x q_x)^2}{E_x q_x (1 - q_x)} \right) \right\}, \quad (3.3.11)$$

onde  $q_x = q_x(\theta)$ . A correspondente transformação logarítmica  $\mathcal{L}^*(\theta) = \ln \mathcal{L}(\theta)$  é dada por

$$\mathcal{L}^*(\theta) = -\frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} \left\{ \ln(2\pi \cdot E_x q_x (1 - q_x)) + \frac{(d_x - E_x q_x)^2}{E_x q_x (1 - q_x)} \right\}.$$

Se omitirmos os termos que não dependem de  $\theta$ , obtemos

$$\mathcal{L}^*(\theta) = -\frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} \left\{ \ln[q_x (1 - q_x)] + \frac{(d_x - E_x q_x)^2}{E_x q_x (1 - q_x)} \right\}.$$

Forfar *et al.* (1988) sustentam que o primeiro termo entre parêntesis é menos sensível a mudanças nos parâmetros  $\theta$  do que o segundo, pelo menos na vizinhança da solução óptima. Deste modo, uma boa aproximação pode ser conseguida desprezando o termo  $\ln[q_x (1 - q_x)]$  e maximizando

$$\mathcal{L}^{**}(\theta) = -\frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} \left\{ \frac{(d_x - E_x q_x)^2}{E_x q_x (1 - q_x)} \right\} = -\frac{1}{2} \chi^2(\theta). \quad (3.3.12)$$

Demonstra-se, assim, que é possível reduzir a maximização da log-verosimilhança a uma (aproximadamente equivalente) minimização do Chi-quadrado.

Por outro lado, se admitirmos que a variância do número de óbitos pode ser aproximada por  $E_x q_x$ , a função de verosimilhança a maximizar resume-se a

$$\mathcal{L}(\theta) = \prod_{x=x_{\min}}^{x_{\max}} \frac{1}{(2\pi \cdot E_x q_x)^{1/2}} \exp \left\{ -\frac{1}{2} \left( \frac{(d_x - E_x q_x)^2}{E_x q_x} \right) \right\}, \quad (3.3.13)$$

onde  $q_x = q_x(\theta)$ .

Aplicando a transformação logarítmica e invocando novamente os argumentos atrás enunciados, a função a maximizar será

$$\mathcal{L}^*(\theta) = -\frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} \left\{ \frac{(d_x - E_x q_x)^2}{E_x q_x} \right\},$$

### 3.4. Metodologia do CMIB

o que, em alternativa, equivale aproximadamente a minimizar

$$\begin{aligned}
 \mathcal{L}^{**}(\theta) &= \sum_{x=x_{\min}}^{x_{\max}} \left\{ \frac{(d_x - E_x q_x)^2}{E_x q_x} \right\} \\
 &= \sum_{x=x_{\min}}^{x_{\max}} \left\{ \left[ \frac{d_x - E_x q_x}{E_x} \right]^2 \cdot \frac{E_x^2}{E_x q_x} \right\} \\
 &= \sum_{x=x_{\min}}^{x_{\max}} \left\{ \left( \frac{d_x}{E_x} - q_x \right)^2 \cdot \frac{E_x}{q_x} \right\} \\
 &= \sum_{x=x_{\min}}^{x_{\max}} \left( \frac{E_x}{q_x} \right) \left[ \frac{d_x}{E_x} - q_x(\theta) \right]^2, \tag{3.3.14}
 \end{aligned}$$

com  $q_x(\theta) = f(x; \theta)$ . Em suma, a análise permite concluir que a maximização da verosimilhança pode ser reduzida a uma (aproximadamente equivalente) minimização dos desvios quadráticos ponderados.

### 3.4 Metodologia do CMIB

Nesta secção, analisamos a metodologia usada pelo Continuous Mortality Investigation Bureau (CMIB)<sup>9</sup> para graduar indicadores de mortalidade. O interesse especial dedicado a esta abordagem, descrita em detalhe em Forfal *et al.* (1988), é justificado por três ordens de razões. Em primeiro lugar, ela constitui uma generalização dos tradicionais modelos de Gompertz e de Makeham, ainda hoje bastante populares nalgumas aplicações. Em segundo lugar, ela pode ser facilmente incorporada na graduação mediante modelos lineares generalizados, conforme teremos oportunidade de verificar na Secção 3.5. Por fim, a metodologia do CMIB afirmou-se como um ponto de partida chave para o desenvolvimento de modelos dinâmicos, um tema que analisaremos em detalhe no capítulo 6.

Considere-se um conjunto de idades consecutivas, observáveis em intervalos de tempo inteiros, e as respectivas séries  $\{E_x\}$  e  $\{d_x\}$ , relativas à idade  $x$ . Admite-se igualmente que não existem duplicidades, ou seja, a investigação recolhe dados sobre pessoas seguras e apólices de forma a que a cada indivíduo corresponde uma única apólice, de modo a

<sup>9</sup>O Continuous Mortality Investigation Bureau (CMIB), do Instituto e da Faculdade de Actuários de Londres, tem a sua origem em 1924, quando se inicia o processo de recolha sistemática de estatísticas demográficas (mortalidade e morbilidade) sobre as pessoas cobertas por contratos de seguro de longo prazo. O CMIB tem a seu cargo a elaboração de tábuas de mortalidade standard para uso na indústria seguradora, cobrindo os principais tipos de seguros do ramo vida: beneficiários de rendas vitalícias, pensionistas e seguros de saúde.

### 3.4. Metodologia do CMIB

que os óbitos sejam independentes. Como vimos nas secções precedentes, a presença de duplicidades não afecta o valor esperado das distribuições Binomial e de Poisson normalmente utilizadas na graduação, mas influencia a sua variância, pelo que nestes casos é necessário introduzir um parâmetro de sobredispersão.

Para efectuar a graduação, o CMIB recorre a uma família de funções paramétricas, denominadas de *Gompertz-Makeham de tipo*  $(r, s)$ ,  $GM(r, s)$ , definidas por

$$GM_{\theta}^{r,s}(x) = \sum_{i=0}^{r-1} \alpha_i x^i + \exp\left(\sum_{j=0}^{s-1} \beta_j x^j\right), \quad (3.4.1)$$

onde  $r$  e  $s$  assumem valores inteiros não negativos (a possibilidade de ambos assumirem o valor zero é excluída) e  $\theta = (\alpha_0, \dots, \alpha_{r-1}; \beta_0, \dots, \beta_{s-1})$  é um vector de parâmetros. Os modelos do tipo  $GM(r, s)$  incluem, como casos particulares, algumas das principais leis de mortalidade apresentadas no capítulo anterior. Por exemplo, é fácil de verificar que

- $GM(0, 2) \rightarrow$  Lei de Gompertz
- $GM(1, 2) \rightarrow$  Lei de Makeham
- $GM(2, 2) \rightarrow$  2ª Lei de Makeham

Note-se ainda que a equação (3.4.1) deve ser interpretada com a convenção de que, no caso particular em que  $r = 0$ , ela inclui apenas o termo exponencial, e no caso em que  $s = 0$  a função se resume ao termo polinomial. Nalgumas aplicações, resulta conveniente definir a seguinte função *Logit-Gompertz-Makeham de tipo*  $(r, s)$  (ou simplesmente  $LGM(r, s)$ )

$$LGM_{\theta}^{r,s}(x) = \frac{GM_{\theta}^{r,s}(x)}{1 + GM_{\theta}^{r,s}(x)}. \quad (3.4.2)$$

Dado um vector de parâmetros  $\theta$ , estimados a partir das estatísticas observadas, a metodologia preconizada pelo CMIB sustenta que uma fórmula do tipo

$$q_x \text{ ou } \mu_x = GM_{\theta}^{r,s}(x) \text{ ou } LGM_{\theta}^{r,s}(x)$$

proporciona um ajustamento adequado.<sup>10</sup>

De acordo com as circunstâncias, o intervalo de valores admissíveis para  $\mu_x$  ou  $q_x$  condiciona a opção por uma equação do tipo Gompertz-Makeham ( $GM$ ) ou Logit-

<sup>10</sup>Modelos do tipo  $GM(r, s)$  têm sido usados pelo CMIB no Reino Unido desde 1990 para graduar a força de mortalidade e construir tábuas de mortalidade contemporâneas. Veja-se, por exemplo, CMIB (1990).

### 3.4. Metodologia do CMIB

---

Gompertz-Makeham (*LGM*). Com efeito, o intervalo admissível para  $\mu_x$  é  $[0, \infty]$ , enquanto que  $q_x$  assume valores no intervalo  $[0, 1]$ .

Estas restrições tornam a equação *LGM* potencialmente mais adequada para graduar  $q_x$ , na medida em que esta garante valores no intervalo  $[0, 1]$ , ao passo que, em determinadas condições<sup>11</sup>, a equação *GM* admite valores no intervalo  $[0, \infty)$ , o que a torna propícia à graduação de  $\mu_x$ .<sup>12</sup>

A metodologia do CMIB recomenda ainda a utilização de uma base ortogonal para os polinómios da equação *GM* ( $r, s$ ).<sup>13</sup> Em particular, são destacadas as potencialidades oferecidas pelos polinómios de Chebycheff  $\{C_i(x)\}$ , definidos pelas equações

$$\begin{aligned}C_0(x) &= 1 \\C_1(x) &= x\end{aligned}$$

e pela seguinte equação de recorrência

$$C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x) \quad (n \geq 1), \quad (3.4.3)$$

que são ortogonais no intervalo  $[-1, +1]$  quando é adoptada a função ponderadora  $w(x) = (1 - x^2)^{-1/2}$ . Ainda no mesmo intervalo, quando usamos a função ponderadora  $w(x) = 1$  obtemos os conhecidos polinómios de Legendre  $\{L_i(x)\}$ , definidos pelos valores iniciais

$$\begin{aligned}L_0(x) &= 1 \\L_1(x) &= x\end{aligned}$$

e pela equação de recorrência

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - nL_{n-1}(x) \quad (n \geq 1). \quad (3.4.4)$$

---

<sup>11</sup>Nomeadamente nos casos em que o termo exponencial (de ordem  $s$ ) da equação *GM* compensa eventuais elementos negativos na componente polinomial (de ordem  $r$ ).

<sup>12</sup>Salienta-se, contudo, que estas regras não são absolutas e servem apenas de referência.

<sup>13</sup>A utilização de uma base ortogonal permite isolar o impacto sobre o ajustamento que decorre da adição de coeficientes aos polinómios da equação *GM* ( $r, s$ ) ou *LGM* ( $r, s$ ). Com efeito, se adoptarmos a base natural  $\{1, x, x^2, \dots, x^{n-1}\}$ , os parâmetros do polinómio de grau  $n$ ,  $P(x) = \sum_{i=0}^n \theta_i x^i$ , que melhor ajustam os dados segundo o critério dos mínimos quadrados, serão diferentes daqueles que se obtém considerando um polinómio de grau  $n+1$ . Se, em alternativa, usarmos uma base ortogonal,  $P(x) = \sum_{i=0}^n \theta_i C_i(x)$ , o polinómio de grau  $n+1$  que melhor ajusta os dados é obtido a partir do polinómio de grau  $n$  pela simples adição de um múltiplo ( $\theta_{n+1}$ , por exemplo) da  $(n+1)^{\text{ésima}}$  função base.

### 3.5. Modelos Lineares Generalizados

---

Assim, por exemplo

$$\begin{aligned} 2L_2(x) &= 3xL_1(x) - L_0(x) = 3x^2 - 1 \\ 3L_3(x) &= 5xL_2(x) - 2L_1(x) = 5xL_2(x) - 2x \\ 4L_4(x) &= 7xL_3(x) - 3L_2(x). \end{aligned}$$

A utilização de polinómios ortogonais requer o rescalonamento das idades no intervalo  $[-1, 1]$ . Para tal o CMIB sugere a seguinte transformação

$$x \rightarrow x' = \frac{x - u}{v} \quad \text{com } u = \frac{x^{\max} + x^{\min}}{2} \quad \text{e } v = \frac{x^{\max} - x^{\min}}{2}$$

donde decorre a seguinte relação  $x \in [x^{\min}, x^{\max}] \rightarrow x' \in [-1, +1]$ .

Se recapitularmos todos estes aspectos, a equação a estimar será

$$\{p\}_{u,v} GM_{\theta}^{r,s}(x) = \sum_{i=0}^{r-1} \alpha_i p_i \left( \frac{x - u}{v} \right) + \exp \left[ \sum_{j=0}^{s-1} \beta_j p_j \left( \frac{x - u}{v} \right) \right], \quad (3.4.5)$$

onde  $\{p_{i,j}(x) : i, j = 0, 1, 2, \dots\}$  define a base de polinómios adoptada e os demais parâmetros têm o significado habitual ou, em alternativa,

$$\{p\}_{u,v} LGM_{\theta}^{r,s}(x) = \frac{\{p\}_{u,v} GM_{\theta}^{r,s}(x)}{1 + \{p\}_{u,v} GM_{\theta}^{r,s}(x)}. \quad (3.4.6)$$

Para estimar os parâmetros do modelo, são considerados como critérios de optimização os métodos de máxima verosimilhança e o método do mínimo Chi-quadrado sendo certo que, como referimos anteriormente, ambos os critérios produzem resultados muito semelhantes.

### 3.5 Modelos Lineares Generalizados

Uma boa parte dos modelos usados pelos actuários para graduar dados de mortalidade pode ser especificada no contexto dos denominados modelos lineares generalizados (GLMs), uma característica unificadora explorada por autores como Renshaw (1991), Haberman e Renshaw (1996) e Verrall (1996), entre outros.

O interesse pelos GLMs é facilmente perceptível se recordarmos que formulações do tipo  $GM(r, s)$  não podem ser estimadas pelos modelos lineares clássicos, dada a natureza eminentemente não linear das funções paramétricas adoptadas. Naquilo que resta desta

secção, procedemos a uma breve introdução aos GLMs, para em seguida discutirmos a sua aplicação na construção de tábuas de mortalidade.

#### 3.5.1 Introdução aos Modelos Lineares Generalizados

Os GLMs constituem uma extensão dos modelos lineares para distribuições da variável de resposta não normais e transformações não lineares.<sup>14</sup> Em termos genéricos, considere-se um vector de observações (*response variables*)  $\mathbf{y} = (y_i : i = 1, 2, \dots, n)$ , tratadas como realizações de uma variável aleatória  $\mathbf{Y} = (Y_i)$ , independentes e idênticamente distribuídas com média  $m_i = \mathbb{E}(\mathbf{Y})$ .

O modelo de regressão linear clássico fornece uma estimativa do valor esperado da variável de resposta em termos de uma função linear de um conjunto limitado de variáveis explicativas (*predictor variables*)  $x_1, x_2, \dots, x_p$ . Formalmente

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i \quad \text{com } \varepsilon_i \sim \text{Nor} \left( 0, \frac{\sigma^2}{w_i} \right), \quad (3.5.1)$$

onde  $w_i$  é um ponderador associado à observação  $i$ . De maneira equivalente, temos

$$Y_i \sim \text{Nor} \left( \beta_0 + \sum_{j=1}^p \beta_j x_{ij}, \frac{\sigma^2}{w_i} \right).$$

Admite-se assim que as observações  $Y_i$ , condicionalmente aos valores das variáveis explicativas, seguem uma distribuição Normal com média  $\mathbb{E}(\mathbf{Y}|x) = m_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}$  e variância  $\sigma^2/w_i$ , e que os termos aleatórios são independentes.

Em alternativa, os GLMs providenciam um método para estimar uma função da média da variável de resposta, expressa em termos de uma combinação linear das variáveis explicativas. Em termos formais,

$$g(\mathbb{E}(\mathbf{Y}|x)) = g(m_i) = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} = \eta(x), \quad (3.5.2)$$

onde  $g$  denota uma função de ligação (denominada *link function*) e a variância,  $\text{Var}(Y) = \phi V(m_i)/\omega$ , compreende uma função de variância  $V(m_i)$ , um parâmetro de escala ( $\phi > 0$ ) e ponderadores prévios  $\omega$ . A função linear das variáveis explicativas,  $\eta(x)$ , é chamada de previsor linear (*linear predictor*).

---

<sup>14</sup>Para uma análise detalhada dos modelos lineares generalizados veja-se, por exemplo, McCullagh e Nelder (1989) e Lindsey (1997).

Em suma, o modelo linear generalizado caracteriza-se por três ingredientes básicos:

- Uma variável de resposta,  $Y$ , observada independentemente para valores fixos das variáveis explicativas
- Um conjunto de variáveis explicativas incorporadas num predictor linear  $\eta(x) = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}$
- Uma função de ligação invertível, que relaciona a variável de resposta com o predictor linear,  $g(m_i) = \eta(x)$ ,  $m = g^{-1}(\eta(x))$

#### 3.5.2 Família exponencial natural e *link* canónico

Admitamos que as observações  $y_i$  são independentes mas não idênticamente distribuídas. A função densidade de probabilidade  $f(\cdot | \theta_i, \phi)$  de  $y_i$  provém da família exponencial natural com parâmetro  $\theta_i$  e é representada, de forma genérica, por

$$f(y_i, \theta_i, \phi) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{\phi} w_i + c\left(y_i, \frac{\phi}{w_i}\right)\right) \quad y_i \in \mathcal{S}, \quad (3.5.3)$$

onde  $\mathcal{S}$  denota o suporte (subconjunto dos números inteiros  $\mathbb{N}$  ou dos reais  $\mathbb{R}$ ),  $w_i$  simboliza um ponderador para cada observação,  $\phi$  representa um parâmetro (constante) de escala, que expressa a magnitude e importância do fenómeno de sobredispersão na amostra e  $\theta_i$  denota o chamado parâmetro natural. As funções  $b(\cdot)$  e  $c(\cdot)$  determinam a lei probabilística (Poisson, Normal, Binomial, Gamma,...) particular em análise.

Demonstra-se que o valor esperado e a variancia de  $Y_i$  são dados, respectivamente, por (McCullagh e Nelder, 1989)

$$\begin{aligned} m_i &= \mathbb{E}(Y_i) = \frac{d}{d\theta_i} b(\theta_i) \\ \text{Var}(Y_i) &= \frac{\phi}{w_i} \frac{d^2}{d\theta_i^2} b(\theta_i) = \frac{\phi}{w_i} V(m_i). \end{aligned}$$

Note-se, em particular, que a função de variância depende de duas parcelas: uma primeira,  $V(m)$ , correspondente à função de variância, que depende apenas do parâmetro  $\theta$  (e, portanto, da média  $m$ ); uma segunda, que depende unicamente do parâmetro de sobredispersão  $\phi$  e dos ponderadores  $w_i$ .

Os GLMs proporcionam um tratamento estatístico unificado de um conjunto de distribuições de probabilidade habitualmente usadas para modelar a mortalidade.

### 3.5. Modelos Lineares Generalizados

Como se observa no Quadro 3.5.1, estas distribuições possuem uma função de densidade que pode ser encaixada na formulação geral (3.5.3).

| Lei de Probabilidade                   | Suporte $\mathcal{S}$ | Densidade de Probabilidade                                    | $\theta = \theta(m)$ | $b(\theta)$          | $\phi$          |
|--|-----------------------|---|----------------------|----------------------|-----------------|
| Bernoulli<br><i>Bin</i> (1, $m$ )      | $\{0, 1\}$            | $m^y (1 - m)^{1-y}$   | $\ln \frac{m}{1-m}$  | $\ln(1 + e^\theta)$  | 1               |
| Poisson<br><i>Poisson</i> ( $m$ )      | $\{0, 1, 2, \dots\}$  | $\frac{m^y e^{-m}}{y!}$                                       | $\ln m$              | $e^\theta$           | 1               |
| Normal<br><i>Nor</i> ( $m, \sigma^2$ ) | $\mathbb{R}$          | $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y-m)^2}$ | $m$                  | $\frac{\theta^2}{2}$ | $\sigma^2$      |
| Gamma<br><i>Gam</i> ( $m_1, m_2$ )     | $\mathbb{R}^+$        | $\frac{1}{\Gamma(m_1)} m_2^{m_1} y^{m_1-1} e^{-m_2 y}$        | $-\frac{m_2}{m_1}$   | $-\ln(-\theta)$      | $\frac{1}{m_1}$ |

Quadro 3.5.1: Distribuições de probabilidade englobadas na família exponencial natural

As funções de variância correspondentes às distribuições mencionadas no Quadro 3.5.1 são sintetizadas no Quadro 3.5.2:

| Lei de Probabilidade | Média<br>$m = b'(\theta)$                 | Função de variância<br>$V(m) = b''(\theta)$ | Variância<br>$\phi V(m_i)/w_i$ |
|----------------------|---|---|--------------------------------|
| Binomial             | $m = \frac{\exp \theta}{1 + \exp \theta}$ | $m(1 - m)$                                  | $\frac{m(1-m)}{w}$             |
| Poisson              | $m = \exp \theta$                         | $m$   | $\frac{m}{w}$                  |
| Normal               | $m = \theta$                              | 1   | $\frac{\sigma^2}{w}$           |
| Gamma                | $m = -1/\theta$                           | $m^2$                                       | $\frac{m^2}{m_1 w}$            |

Quadro 3.5.2: Média e variância associada às distribuições de probabilidade englobadas na família exponencial natural

Cada uma das leis de probabilidade acima enunciadas e pertencentes à família exponencial natural possui uma *link function* específica. Esta função, dita de *link* canônico, é definida por  $\theta = \eta$ , onde  $\theta$  é o parâmetro canônico. A ligação canônica é tal que  $g(m_i) = \theta_i$ . As funções canônicas são apresentadas no Quadro 3.5.3. Como se observa, a lei de Poisson privilegia o *link* logarítmico, enquanto a lei Binomial adota o *link logit*.

### 3.5. Modelos Lineares Generalizados

| Lei de Probabilidade | Função de <i>link</i> canónico            |
|----------------------|---|
| Poisson              | $\eta = \ln m$                            |
| Binomial             | $\eta = \ln \left( \frac{m}{1-m} \right)$ |
| Normal               | $\eta = m$                                |
| Gamma                | $\eta = 1/m$                              |

Quadro 3.5.3: Funções de *link* canónico associadas às principais leis de probabilidade pertencentes à família exponencial natural

Nos GLM's, as estimativas dos parâmetros do previsor linear obtêm-se com base num procedimento iterativo, usando o método dos mínimos quadrados ponderados.<sup>15</sup> Em termos formais, os estimadores de máxima verosimilhança dos parâmetros  $\beta_0, \beta_1, \dots, \beta_p$  maximizam a seguinte função de log-verosimilhança

$$\begin{aligned}
 \ln \mathcal{V}(\boldsymbol{\theta} | \mathbf{y}, \phi) &= \sum_{i=i_{\min}}^{i=i_{\max}} \ln f(y_i | \theta_i, \phi) \\
 &= \sum_{i=i_{\min}}^{i=i_{\max}} \frac{y_i \theta_i - b(\theta_i)}{\phi/w_i} w_i + \sum_{i=i_{\min}}^{i=i_{\max}} c(y_i, \phi) \\
 &= \frac{1}{\phi} \sum_{i=i_{\min}}^{i=i_{\max}} w_i (y_i \theta_i - b(\theta_i)) + \text{constante},
 \end{aligned}$$

ou seja, são soluções das equações

$$\frac{\partial \ln \mathcal{V}(\boldsymbol{\theta} | \mathbf{y}, \phi)}{\partial \beta_j} = 0, \quad j = 0, 1, \dots, p.$$

Para avaliar a qualidade do ajustamento, i.e., a capacidade do modelo para representar os dados observados, McCullagh e Nelder (1989) defendem a utilização de uma medida de bondade do ajustamento entre os dados observados e os valores ajustados gerados pelo modelo denominada de (*scaled*) *deviance*,  $\mathcal{D}(\mathbf{y}; \mathbf{m})$ , cujo valor se obtém a partir da expressão

$$\mathcal{D}(\mathbf{y}; \hat{\mathbf{m}}) = 2\mathcal{L}(\mathbf{y}; \mathbf{y}) - 2\mathcal{L}(\hat{\mathbf{m}}; \mathbf{y}), \quad (3.5.4)$$

onde  $\mathcal{L}(\cdot) = \ln \mathcal{V}(\boldsymbol{\theta} | \mathbf{y}, \phi)$ ,  $\mathbf{m} = (m_1, \dots, m_n)$ ,  $\mathcal{L}(\mathbf{y}; \mathbf{y})$  representa a log-verosimilhança quando os valores observados coincidem com os valores estimados (modelo saturado, em que existem tantos parâmetros quantas as observações), e  $\mathcal{L}(\hat{\mathbf{m}}; \mathbf{y})$  denota a log-

<sup>15</sup>Para mais detalhes sobre o método de estimação veja-se, e.g., McCullagh e Nelder (1989).

### 3.5. Modelos Lineares Generalizados

---

verossimilhança do modelo ajustado  $\hat{m}$ . Como  $\mathcal{L}(\mathbf{y}; \mathbf{y})$  não depende do modelo utilizado e a *deviance* mede o desvio entre os valores observados e os estimados, maximizar a verossimilhança  $\mathcal{L}(\hat{m}; \mathbf{y})$  equivale a minimizar  $\mathcal{D}(\mathbf{y}; \hat{m})$  com respeito a  $m$ , sujeito às restrições impostas pelo modelo.

O valor de  $\mathcal{D}(\mathbf{y}; \hat{m})$  é aproximado assintoticamente por uma distribuição Chi-quadrado com  $(n - p)$  graus de liberdade,  $\chi_{n-p}^2$ , onde  $p$  denota o número de parâmetros do modelo ajustado. A deviance desempenha um papel importante, não tanto como medida absoluta da qualidade do ajustamento, mas sobretudo pela sua magnitude relativa, permitindo efectuar comparações entre os vários modelos analisados. Por exemplo, se quisermos avaliar o impacte resultante da adição de uma variável explicativa ao modelo, no caso em que  $H_0$  é o modelo a validar e  $H_1$  a sua correspondente extensão e  $\hat{m}_0$  e  $\hat{m}_1$  os respectivos valores ajustados, a diferença

$$\begin{aligned} \mathcal{D}(\mathbf{y}; \hat{m}_0) - \mathcal{D}(\mathbf{y}; \hat{m}_1) &= [2\mathcal{L}(\mathbf{y}; \mathbf{y}) - 2\mathcal{L}(\hat{m}_0; \mathbf{y})] - [2\mathcal{L}(\mathbf{y}; \mathbf{y}) - 2\mathcal{L}(\hat{m}_1; \mathbf{y})] \\ &= 2\mathcal{L}(\hat{m}_1; \mathbf{y}) - 2\mathcal{L}(\hat{m}_0; \mathbf{y}) \end{aligned} \quad (3.5.5)$$

corresponde à estatística rácio de verossimilhança para contrastar  $H_0$  vs  $H_1$ , distribuída aproximadamente por  $\chi_1^2$ .

O diagnóstico do modelo pode ainda ser feito com base na análise de outras estatísticas (e.g., estatística de Pearson, Pseudo- $R^2$ , ...) e na avaliação do comportamento dos resíduos do modelo.<sup>16</sup>

#### 3.5.3 Modelos Lineares Generalizados e graduação

A aplicação dos GLMs em problemas de graduação é analisada na literatura actuarial por Renshaw (1991), Renshaw e Hatzopoulos (1996), Haberman e Renshaw (1996), Verrall (1996), Delwarde *et al.* (2004), entre outros. Nesta secção, ilustramos a aplicação desta metodologia analisando, em particular, o problema da graduação de  $q_x$  com base em

---

<sup>16</sup>Registe-se ainda que, como salientámos anteriormente, o problema da modelação requer um equilíbrio entre a qualidade do ajustamento e a sua complexidade, associada ao aumento do número de parâmetros a estimar. Uma forma de considerar este potencial *trade-off* consiste em recorrer à estatística  $C_p$  de Mallows (Venables e Ripley, 2002), definida da seguinte forma

$$C_p = \frac{RSS}{\hat{\sigma}^2} + 2p - n,$$

onde  $RSS$  representa a soma dos quadrados residuais,  $\hat{\sigma}^2$  denota a variância do erro e  $p$  simboliza novamente o número de parâmetros do modelo. O modelo será tanto melhor quanto menor for o valor de  $C_p$ , ou seja, a medida de qualidade do ajustamento do modelo ( $RSS$ ) é penalizada pela complexidade associada ao número de parâmetros. Indicadores como o critério BIC (*Bayesian Information Criterion*) e AIC (*Akaike Information Criterion*) podem igualmente ser usados para este efeito.

### 3.5. Modelos Lineares Generalizados

funções do tipo  $LGM(r, s)$  sem componente polinomial e considerando a transformação logit, e o problema da graduação de  $\mu_x$  adotando funções do tipo  $GM(r, s)$ ,  $r = 0$  e transformação logarítmica. São ainda abordados alguns casos mais gerais.

#### Graduação de $q_x$

Consideremos em primeiro lugar o problema da graduação de  $q_x$ . Tal como na Secção 3.4, assumimos que a variável aleatória  $D_x$  segue uma distribuição Binomial com óbitos independentes entre si, i.e.,  $D_x \sim \text{Bin}(E_x, q_x)$ .

A expressão geral da função de verosimilhança é dada por

$$\begin{aligned} \mathcal{L}(q) &= \prod_{x=x_{\min}}^{x_{\max}} \binom{E_x}{d_x} q_x^{d_x} (1 - q_x)^{E_x - d_x} \\ &\propto \prod_{x=x_{\min}}^{x_{\max}} q_x^{d_x} (1 - q_x)^{E_x - d_x} \\ &= \exp \left\{ \sum_{x=x_{\min}}^{x_{\max}} [d_x \log(q_x) + (E_x - d_x) \log(1 - q_x)] \right\} \\ &= \exp \left\{ \sum_{x=x_{\min}}^{x_{\max}} \left[ d_x \log \left( \frac{q_x}{1 - q_x} \right) + E_x \log(1 - q_x) \right] \right\}. \end{aligned}$$

No que respeita às *link functions*, o Quadro 3.5.4 resume as possibilidades sugeridas por Renshaw (1991):<sup>17</sup>

| Função               | <i>link</i> canónico                               | Função inversa                                |
|----------------------|--|---|
| Logit                | $\eta_x = \log \left( \frac{q_x}{1 - q_x} \right)$ | $q_x = \frac{\exp(\eta_x)}{1 + \exp(\eta_x)}$ |
| Log-Log              | $\eta_x = -\log(-\log(q_x))$                       | $q_x = \exp(-\exp(-\eta_x))$                  |
| Log-Log Complementar | $\eta_x = \log(-\log(1 - q_x))$                    | $q_x = 1 - \exp(-\exp(\eta_x))$               |
| Probit               | $\eta_x = \Phi^{-1}(q_x)$                          | $q_x = \Phi(\eta_x)$                          |

Quadro 3.5.4: Funções de *link* canónico usadas em problemas de graduação

onde  $\Phi$  designa a função de distribuição cumulativa de uma variável com distribuição normal padronizada.

A título demonstrativo, analisemos em profundidade a aplicação da transformação *logit* considerando um predictor linear do tipo Gompertz-Makeham, i.e., uma função do

<sup>17</sup>Registe-se, a este propósito, que na teoria dos GLMs é mais usual a definição do *link* em termos de  $\mathbb{E}[D_x] = E_x q_x$ , uma reparametrização trivial que não altera a essência do modelo.

### 3.5. Modelos Lineares Generalizados

---

tipo  $LGM(r, s)$ , no caso particular em que  $r = 0$ . Em termos formais,

$$q_x = LGM(0, s) = \frac{GM(0, s)}{1 + GM(0, s)}.$$

Neste caso, o modelo a ajustar é definido por

$$\log\left(\frac{q_x}{1 - q_x}\right) = \eta_x = GM(0, s) = \sum_{j=0}^{s-1} \beta_j x^j, \quad (3.5.6)$$

donde se retira que as probabilidades graduadas são calculadas mediante

$$\hat{q}_x = \frac{\exp(\hat{\eta}_x)}{1 + \exp(\hat{\eta}_x)}. \quad (3.5.7)$$

Para avaliar a qualidade do ajustamento, a medida de bondade adequada (*scaled deviance*) é definida por

$$\mathcal{D}(\mathbf{y}; \hat{\mathbf{m}}) = 2 \sum_{x=x_{\min}}^{x_{\max}} \left[ d_x \log\left(\frac{d_x}{\hat{m}_x}\right) + (E_x - d_x) \log\left(\frac{E_x - d_x}{E_x - \hat{m}_x}\right) \right], \quad (3.5.8)$$

onde  $\hat{m}_x = g^{-1}(\hat{\eta}_x)$  representa os valores ajustados pelo modelo.<sup>18</sup>

#### Gradação de $\mu_x$

Neste caso, a gradação é realizada no pressuposto de que o número de óbitos observado,  $d_x$ , é modelado em termos de realizações independentes de variáveis aleatórias Poisson

---

<sup>18</sup>Partindo da clássica lei de Gompertz para a força de mortalidade (i.e.,  $\mu_x = Bc^x$ ), Renshaw (1991) demonstra ainda que é possível transformar um modelo não linear num modelo linearizável, por exemplo em termos de  $q_x$ , recorrendo ao *link* log-log complementar. Em termos mais específicos

$$\eta_x = \log(-\log(1 - q_x)) = \beta_0 + \beta_1 x,$$

onde  $c = \exp(\beta_1)$  e  $B = \log(c) \exp(\beta_0) / (c - 1)$ . O autor prova, deste modo, que é possível estimar o modelo de Gompertz recorrendo a um modelo linear generalizado, envolvendo a distribuição Binomial, *link* log-log complementar e uma função linear do tipo  $GM(0, 2)$ . Para melhorar a qualidade do ajustamento, o autor sugere ainda a extensão do modelo mediante a adição de termos ao predictor linear, ou seja, adoptando um modelo do tipo

$$\hat{\eta}_x = \sum_{j=0}^{s-1} \hat{\beta}_j x^j,$$

com valores graduados determinados por

$$\hat{q}_x = 1 - \exp(-\exp(\hat{\eta}_x)).$$

### 3.5. Modelos Lineares Generalizados

---

$D_x$ , condicionadas por  $E_x$ , o número de expostos centralmente ao risco, i.e.,

$$D_x \sim \text{Poisson}(E_x \mu_x). \quad (3.5.9)$$

Os elementos necessários para implementar o modelo GLM apropriado são, neste caso, dados por

$$\text{variável de resposta } \{d_x\}, \text{ com } m_x = E_x \mu_x, V(m) = m_x, \phi = 1, w_x = 1 \quad (3.5.10)$$

ou, de forma equivalente,

$$\text{variável de resposta } \left\{ \frac{d_x}{E_x} \right\}, \text{ com } m_x = \mu_x, V(m) = m_x, \phi = 1, w_x = E_x. \quad (3.5.11)$$

Nos casos, já mencionados, em que a análise da mortalidade da população segura se baseia em dados sobre o número de apólices e existe a possibilidade de um indivíduo deter duas ou mais apólices, o parâmetro de sobredispersão  $\phi$  toma um valor superior à unidade.

A expressão da função de verosimilhança correspondente é dada por

$$\begin{aligned} \mathcal{L}(\mu) &= \prod_{x=x_{\min}}^{x_{\max}} \frac{(E_x \mu_x)^{d_x} \exp(-E_x \mu_x)}{d_x!} \\ &\propto \exp \left[ \sum_{x=x_{\min}}^{x_{\max}} (d_x \log(E_x) + d_x \log(\mu_x) - E_x \mu_x) \right]. \end{aligned}$$

Admitamos agora que  $\mu_x$  é graduada usando um previsor linear do tipo  $GM(0, s)$  e *link* logarítmico, com variável de resposta  $d_x$ , i.e.,

$$\begin{aligned} \log(m_x) &= \eta_x = \log(E_x \mu_x) = \log(E_x) + \log(\mu_x) \\ &= \log(E_x) + \sum_{j=0}^{s-1} \beta_j x^j. \end{aligned} \quad (3.5.12)$$

Como se observa, neste modelo o previsor linear é aumentado de um regressor com coeficiente conhecido (designado por *offset*, com valor igual a  $(+1)$ ,  $\log(E_x)$ ). Este procedimento permite a inclusão do termo respeitante à exposição  $d_x \log(E_x)$  na função de verosimilhança.

Neste caso, os valores graduados são estimados mediante

$$\hat{\mu}_x = \exp(\hat{\eta}_x) = \exp\left(\sum_{j=0}^{s-1} \hat{\beta}_j x^j\right). \quad (3.5.13)$$

A medida de bondade do ajustamento (*unscaled deviance*) adequada é, neste caso, definida por

$$\mathcal{D}(\mathbf{y}; \hat{\mathbf{m}}) = 2 \sum_{x=x_{\min}}^{x_{\max}} \left[ d_x \log\left(\frac{d_x}{\hat{m}_x}\right) - (d_x - \hat{m}_x) \right], \quad (3.5.14)$$

onde  $\hat{m}_x = g^{-1}(\hat{\eta}_x)$  representa, uma vez mais, os valores ajustados pelo modelo.

Nos casos em que o fenómeno de sobredispersão está presente, Renshaw *et al.* (1996) propõem a aferição da qualidade do ajustamento pela denominada *scaled deviance*, definida por

$$\mathcal{S}(\mathbf{y}; \hat{\mathbf{m}}) = \frac{\mathcal{D}(\mathbf{y}; \hat{\mathbf{m}})}{\phi}, \quad (3.5.15)$$

com parâmetro de escala  $\phi$  estimado por

$$\hat{\phi} = \frac{\mathcal{D}(\mathbf{y}; \hat{\mathbf{m}})}{\nu} \quad (3.5.16)$$

ou, em alternativa, por

$$\hat{\phi} = \frac{1}{\nu} \sum_{x=x_{\min}}^{x_{\max}} \left[ \frac{(d_x - \hat{m}_x)^2}{\hat{m}_x} \right], \quad (3.5.17)$$

onde  $\nu$  representa os graus de liberdade da estimação.

#### 3.5.4 Abordagem dual

Renshaw e Haberman (1996) e Renshaw *et al.* (1997) desenvolvem uma metodologia alternativa (*dual approach*), em que assumem que o número de indivíduos expostos centralmente ao risco pode ser expresso em termos de realizações independentes de variáveis aleatórias Gamma, condicionadas pelo correspondente número de óbitos, ou seja,

$$E_x \sim \text{Gamma}(d_x, \mu_x). \quad (3.5.18)$$

Os elementos necessários para implementar o modelo GLM apropriado são, neste caso, dados por

$$\text{variável de resposta } \{E_x\}, \text{ com } m_x = d_x \frac{1}{\mu_x}, V(m) = m_x^2, \phi = 1, w_x = d_x \quad (3.5.19)$$

ou, de forma equivalente,

$$\text{variável de resposta } \left\{ \frac{E_x}{d_x} \right\}, \text{ com } m_x = \frac{1}{\mu_x}, V(m) = m_x^2, \phi = 1, w_x = d_x. \quad (3.5.20)$$

Admitamos novamente que  $\mu_x$  é graduada usando um previsor linear do tipo *GM* (0,  $s$ ) e *link* logarítmico, com variável de resposta  $E_x$ , i.e.,

$$\begin{aligned} \log(m_x) &= \eta_x = \log\left(\frac{d_x}{\mu_x}\right) = \log(d_x) - \log(\mu_x) \\ &= \log(d_x) + \sum_{j=0}^{s-1} \beta_j x^j. \end{aligned} \quad (3.5.21)$$

Como se observa, o objectivo da graduação é agora a denominada *força de vitalidade*,  $1/\mu_x$ , e não a força de mortalidade. No demais, a estrutura do modelo é análoga à apresentada em (3.5.12), com a diferença de que neste caso  $\log(d_x)$  é declarado como *offset*. Em resultado, os coeficientes estimados para os parâmetros das equações (3.5.12) e (3.5.21) são idênticos em magnitude, mas opostos em termos de sinal.

Em suma, os autores demonstram que a graduação de  $\mu_x$  mediante funções paramétricas, admitindo que os óbitos observados são modelados em termos de realizações independentes de variáveis aleatórias do tipo Poisson equivale, na prática, à abordagem dual, na qual os indivíduos expostos centralmente ao risco são modelados como realizações independentes de variáveis aleatórias gamma, condicionadas pelo número observado de óbitos.

A equivalência entre as duas abordagens resulta evidente quando constatamos que a função de verosimilhança  $\mathcal{V}(\mu_x)$  associada a (3.5.18)

$$\mathcal{V}(\mu_x) = \exp(-E_x \mu_x) \frac{(E_x \mu_x)^{d_x}}{E_x \Gamma(d_x)}$$

difere da que se obtém para a distribuição de Poisson (3.5.9)

$$\mathcal{V}(\mu_x) = \exp(-E_x \mu_x) \frac{(E_x \mu_x)^{d_x}}{d_x!}$$

apenas por um factor multiplicativo.<sup>19</sup>

---

<sup>19</sup>As vantagens proporcionadas pela abordagem dual, face à graduação tradicional, na graduação de tábuas seleccionadas são discutidas em detalhe por Renshaw *et al.* (1997).

### 3.5.5 Modelos Não-Lineares Generalizados e graduação

Analisemos de novo a graduação de  $q_x$ , considerando agora o caso mais geral em que o previsor é não-linear e do tipo  $LGM(r, s)$

$$q_x = LGM(r, s) = \frac{GM(r, s)}{1 + GM(r, s)},$$

donde facilmente se deriva

$$\frac{q_x}{1 - q_x} = GM(r, s) \quad \text{e} \quad \frac{1}{1 - q_x} = 1 + GM(r, s).$$

Logaritmizando a função de verosimilhança, obtemos

$$\begin{aligned} \log \mathcal{L}(q) &= \sum_{x=x_{\min}}^{x_{\max}} \left[ \log \left( \frac{E_x}{d_x} \right) + d_x \log(q_x) + E_x \log(1 - q_x) - d_x \log(1 - q_x) \right] \\ &= \sum_{x=x_{\min}}^{x_{\max}} \left[ \log \left( \frac{E_x}{d_x} \right) + d_x \log \left( \frac{q_x}{1 - q_x} \right) + E_x \log(1 - q_x) \right] \\ &= \sum_{x=x_{\min}}^{x_{\max}} \left[ \log \left( \frac{E_x}{d_x} \right) + d_x \log \left( \frac{q_x}{1 - q_x} \right) - E_x \log \left( \frac{1}{1 - q_x} \right) \right] \\ &= \sum_{x=x_{\min}}^{x_{\max}} \left[ \log \left( \frac{E_x}{d_x} \right) + d_x \log(GM(r, s)) - E_x \log(1 + GM(r, s)) \right]. \end{aligned}$$

Se desprezarmos os termos constantes, os coeficientes  $\theta = \{\alpha_r; \beta_s\}$  das funções  $GM(r, s)$  são estimados por máxima verosimilhança ou, o que é equivalente, resolvendo o seguinte problema de minimização

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} [E_x \log(1 + GM(r, s)) - d_x \log(GM(r, s))] \right\}. \quad (3.5.22)$$

Em suma, a utilização do previsor não linear com o *link* identidade (*odds*)

$$\hat{\eta}_x = GM(r, s) = \frac{q_x}{1 - q_x} \quad (3.5.23)$$

permite obter valores graduados mediante a função inversa

$$\hat{q}_x = \frac{\hat{\eta}_x}{1 + \hat{\eta}_x}. \quad (3.5.24)$$

### 3.5. Modelos Lineares Generalizados

---

Retomemos o caso da graduação da força de mortalidade  $\mu_x$ . Se considerarmos um previsor não linear do tipo  $GM(r, s)$ , distribuição de Poisson e *link* identidade, i.e.,

$$\eta_x = \mu_x = GM(r, s) \quad (3.5.25)$$

e substituímos na função de verosimilhança, obtemos

$$\mathcal{L}(\mu) = \exp \left[ \sum_{x=x_{\min}}^{x_{\max}} (d_x \log(E_x) + d_x \log(GM(r, s)) - E_x GM(r, s) - \log(d_x!)) \right].$$

Ignorando os termos constantes (que não dependem de  $\mu_x$ ), os coeficientes  $\theta$  das funções  $GM(r, s)$  obtém-se por máxima verosimilhança, i.e., resolvendo o seguinte problema de minimização

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} [E_x GM(r, s) - d_x \log(GM(r, s))] \right\}. \quad (3.5.26)$$

Renshaw (1991) desenvolve uma abordagem alternativa à graduação de modelos não lineares, envolvendo a utilização de uma aproximação linear aos termos não lineares do modelo paramétrico em análise. Consideremos, a título de exemplo, a clássica lei de Makeham  $\eta_x = A + Bc^x$ . Introduzindo uma reparametrização trivial desta equação, podemos escrever

$$\eta_x = \alpha + \beta \exp(\phi x), \quad (3.5.27)$$

onde  $\alpha = A$ ,  $\beta = B$  e  $\phi = \ln(c)$ .

Representemos por

$$g(x; \phi) = \exp(\phi x)$$

o termo não linear num qualquer previsor linear. Se efectuarmos uma aproximação de Taylor (de 1ª ordem) em torno dos valores iniciais  $\phi_0$ ,

$$g(x; \phi) \simeq g(x; \phi_0) + (\phi - \phi_0) \left( \frac{\partial g}{\partial \phi} \right)_{\phi=\phi_0}$$

o termo não linear em (3.5.27),  $\beta \exp(\phi x)$ , pode ser substituído por

$$\beta u + \gamma \xi, \quad (3.5.28)$$

onde

$$u = g(x; \phi_0), \quad \xi = \left( \frac{\partial g}{\partial \phi} \right)_{\phi=\phi_0} \quad \text{e} \quad \gamma = \beta(\phi - \phi_0).$$

O procedimento de graduação sugerido envolve então as seguintes etapas:

1. Conhecido o valor inicial  $\phi_0$ , são calculados os valores das variáveis explicativas  $u = \exp(\phi_0 x)$  e  $\xi = x \exp(\phi_0 x)$ ;
2. Estimam-se os parâmetros  $\beta$  e  $\gamma$  mediante um qualquer procedimento de ajustamento;
3. O valor de  $\phi_0$  é actualizado usando

$$\phi_1 = \phi_0 + \frac{\hat{\gamma}}{\hat{\beta}}.$$

4. O procedimento é repetido até que seja alcançada a convergência. Renshaw (1991) adverte que a convergência do processo iterativo não é garantida caso os valores iniciais de  $\phi_0$  sejam muito distintos do seu “verdadeiro” valor.<sup>20</sup>

Na medida em que o objectivo é graduar  $\mu_x$ , é natural o recurso à distribuição de Poisson e ao *link* logarítmico.

### 3.6 Modelo de Heligman-Pollard

Inspirado nos trabalhos de Thiele (1872), o modelo de Heligman e Pollard (1980) representa uma formulação paramétrica alternativa às discutidas nas secções anteriores, sendo amplamente utilizado em estudos de mortalidade. Aplicada originalmente a dados de mortalidade da população australiana no período pós-guerra, esta lei serviu de base a inúmeras investigações nos Estados Unidos (Mode e Busby, 1982), Inglaterra (Forfar e Smith, 1987; Renshaw, 1992; Benjamin e Soliman, 1993), Suécia (Hartmann, 1987; Kostaki, 1992) e Espanha (Felipe e Guillén, 1999), entre outros, em especial a partir do momento em que as Nações Unidas promoveram o ajustamento da mortalidade através da denominada primeira lei de Heligman-Pollard.

Na sua formulação mais genérica, o modelo adopta a seguinte representação

$$\frac{q_x}{1 - q_x} = \sum_{i=1}^n A_i \exp \left[ -B_i (f_i(x) - C_i)^{D_i} \right] \quad 0 \leq x < \infty, \quad (3.6.1)$$

---

<sup>20</sup>Nos estudos empíricos realizados o autor conclui que um valor inicial de  $\phi_0 = 0,0005$  assegura uma convergência rápida do procedimento iterativo.

### 3.6. Modelo de Heligman-Pollard

---

onde  $A_i$ ,  $B_i$ ,  $C_i$  e  $D_i$  ( $i = 1, 2, \dots, n$ ) são parâmetros a estimar e  $f_i(x)$  é uma função de  $x$  (normalmente  $\ln x$  ou simplesmente  $x$ ). Os autores argumentam que  $n = 3$  é suficiente para obter bons níveis de ajustamento e evita o problema da sobreparametrização.

Na realidade, as três formulações efectivamente usadas pelos autores foram as seguintes:

- 1ª Lei (versão A) de Heligman-Pollard

$$\frac{q_x}{1 - q_x} = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + GH^x \quad (3.6.2)$$

- 1ª Lei (versão B) de Heligman-Pollard

$$q_x = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + \frac{GH^x}{1 + GH^x} \quad (3.6.3)$$

- 2ª Lei de Heligman-Pollard

$$q_x = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + \frac{GH^x}{1 + KGH^x} \quad (3.6.4)$$

- 3ª Lei de Heligman-Pollard

$$q_x = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + \frac{GH^{x^K}}{1 + GH^{x^K}} \quad (3.6.5)$$

O modelo de Heligman-Pollard apresenta algumas vantagens face à maioria das leis de mortalidade apresentadas no capítulo anterior. Com efeito, a análise da equação base do modelo, enunciada em (3.6.2), permite verificar que:

- A função é contínua
- Aplicável a todo o intervalo de idades  $0 \leq x < \infty$
- A probabilidade graduada  $q_x$  está compreendida no intervalo  $[0, 1]$
- Os parâmetros têm uma interpretação demográfica e biológica
- É suficientemente flexível de modo a ajustar uma grande variedade de experiências de mortalidade
- O número de parâmetros é relativamente diminuto, se considerarmos que a função é aplicável a todo o intervalo de idades que compõem o arco da vida humana.

### 3.6. Modelo de Heligman-Pollard

---

A equação base (3.6.2) inclui três termos distintos, cada um deles espelhando uma componente distinta da mortalidade.

O **primeiro termo**, uma função exponencial rapidamente decrescente, reflecte a diminuição da mortalidade durante os primeiros anos de vida. Esta componente é modelada por três parâmetros, com a seguinte interpretação:

- O parâmetro  $A$ , que assume valores no intervalo  $[0, 1]$  e é aproximadamente igual a  $q_1$ , está directamente relacionado com o nível da mortalidade infantil e juvenil.
- O parâmetro  $B$  é indicativo do nível de mortalidade infantil  $q_0$  que, de acordo com o modelo, assume valores no intervalo  $[q_1, 0.5]$ . Na prática,  $B$  assume valores muito reduzidos, pelo que o seu efeito é negligenciável para  $x \geq 1$ . No caso particular em que  $B = 0$ ,  $q_0 = 0.5$ , independentemente dos valores dos parâmetros  $A$  e  $C$ .
- O parâmetro  $C$ , que assume igualmente valores no intervalo  $[0, 1]$ , está associado à taxa de decréscimo na mortalidade. Quanto maior for o seu valor, mais rápido é o decréscimo na mortalidade.

O **segundo termo**, uma função semelhante à lognormal, representa a mortalidade nas idades adultas intermédias (entre os 10 e os 40 anos). Está referenciado na literatura demográfica e actuarial como descrevendo a mortalidade acidental para ambos os sexos (*accident hump*) e a mortalidade maternal para a população feminina. Trata-se, na prática, de uma componente de mortalidade adicional, sobreposta à curva natural de mortalidade, expressa pelos restantes dois termos da equação (3.6.2). Os seus parâmetros têm a seguinte interpretação:

- O parâmetro  $D$ ,  $D \in [0, 1]$ , estabelece a severidade da parábola correspondente à mortalidade acidental. Quanto maior for o seu valor maior será o impacto dos acidentes sobre a mortalidade total.
- O parâmetro  $E$ ,  $E \in [0, \infty)$ , condiciona a dispersão em torno do pico de mortalidade, pelo que a um valor mais elevado de  $E$  corresponde um pico de mortalidade concentrado num intervalo de idades mais estreito.
- O parâmetro  $F$  indica a localização do pico de acidentes.

O **terceiro termo** traduz a tradicional lei de Gompertz, que reflecte o crescimento exponencial da mortalidade nas idades mais avançadas. Integra dois parâmetros com a seguinte interpretação:

### 3.6. Modelo de Heligman-Pollard

---

- O parâmetro  $G$  indica o nível de base para a mortalidade sénior e toma valores no intervalo  $[0, \infty)$ .
- O parâmetro  $H$  espelha a taxa de crescimento na mortalidade sénior.

Hartmann (1987) demonstra que todos os parâmetros da primeira lei de Heligman-Pollard são positivos e que a função apenas está definida para  $x > 0$ , pelo que é necessária uma ligeira modificação para a especificar à idade 0. O autor sugere, em particular, o seguinte valor

$$\lim_{x \rightarrow 0} \left( \frac{q_x}{1 - q_x} \right) = A^{BC} + G = \frac{q_0}{p_0}.$$

A segunda e terceira leis de Heligman-Pollard, que diferem da primeira apenas no termo concernente à mortalidade sénior, apresentam variações destinadas a captar de forma mais adequada o comportamento nas idades mais avançadas. Na segunda lei, o parâmetro  $K$  é o único a poder assumir valores negativos. Ainda assim, Heligman e Pollard (1980) sustentam que a interpretação dos parâmetros permanece válida.

Os autores salientam ainda que os valores de  $G$  e  $H$  na terceira lei são bastante afectados pela direcção e extensão da curvatura da mortalidade nas idades avançadas, e que não é claro que preservem algum significado demográfico. Contudo,  $x^K$  pode ser entendido como uma transformação da escala de idades. Refira-se ainda que, no caso particular em que  $K = 1$ , as segunda e terceira leis se reduzem a (3.6.3), com  $K$  a determinar o grau de curvatura da função.

As leis de Heligman-Pollard podem ser usadas, quer para graduar tábuas de mortalidade contemporâneas quer, como veremos no Capítulo 6, para construir tábuas dinâmicas. As estimativas dos parâmetros do modelo obtêm-se através do método dos mínimos quadrados ponderados não lineares, i.e., minimizando

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{x=x_{\min}}^{x_{\max}} w_x (q_x - f(x; \theta))^2 \right\},$$

onde  $w_x$  denota os ponderadores,  $f(x; \theta)$  a função a ajustar e as demais variáveis assumem o significado habitual.

Numa análise sobre a evolução temporal dos parâmetros do modelo, Congdon (1993) argumenta que este está sobreparametrizado. Em consequência, fixa alguns dos parâmetros da segunda lei (considera, por exemplo, que  $A$  e  $B$  são iguais a  $q_0$  e  $q_1$ , respectivamente), que adopta em estudos empíricos de graduação. Hannerz (1999, 2001) propõe uma alternativa à lei de HP destinada a ajustar a evolução recente da mortalidade da população feminina sueca.

Outra solução, sugerida por autores como Carriere (1992, 1994) e Rempala e Szatzschneider (2004) consiste, como vimos no Capítulo 2, em graduar usando combinações de funções paramétricas. Mais recentemente, Dellaportas *et al.* (2001) adoptam a análise Bayesiana e usam técnicas de simulação de *Markov Chain Monte Carlo* (MCMC) para simular o comportamento dos parâmetros da lei de Heligman-Pollard.

## 3.7 Indicadores de qualidade do ajustamento

Uma vez ajustado o modelo, i.e., estimados os parâmetros da função  $f(x; \theta)$  através de um dos métodos acima enunciados, é fundamental avaliar a qualidade do ajustamento, ou seja, a sua capacidade de descrever adequadamente a realidade observada. Essa aferição pode ser feita com o auxílio de indicadores, que medem a qualidade do ajustamento e possibilitam a comparação dos resultados proporcionados pelas diferentes metodologias. Nesta secção, seguimos de perto Benjamin e Pollard (1980) e Forfar *et al.* (1988) e descrevemos, de forma sucinta, os principais indicadores sugeridos na literatura e usados em estudos de graduação.

### Desvios absolutos e desvios relativos

O procedimento de graduação permite obter os valores ajustados para os diferentes indicadores de mortalidade ( $q_x$  ou  $\mu_x$ ), com base nos quais é possível calcular facilmente o número de óbitos esperados,  $d_x^{\text{exp}}$ , bastando para tal a multiplicação pelos expostos (inicialmente ou centralmente) ao risco. Designemos por  $Dev_x$  o *desvio absoluto* para cada idade  $x$ , i.e.,

$$Dev_x = d_x - d_x^{\text{exp}} \quad (3.7.1)$$

e por  $z_x$  os correspondentes *desvios relativos* (*Pearson residuals*), com

$$z_x = \frac{Dev_x}{\sqrt{\text{Var}(d_x)}}, \quad (3.7.2)$$

onde  $\text{Var}(d_x)$  denota a variância do número de óbitos, calculada segundo as hipóteses assumidas para a distribuição de probabilidade da variável aleatória. Posto isto, a qualidade do ajustamento é aferida averiguando se os desvios estão distribuídos de forma aleatória (quando considerados sequencialmente), e se a sua distribuição está de acordo com as hipóteses subjacentes ao modelo (Binomial, Poisson, etc.) adoptado.<sup>21</sup>

---

<sup>21</sup> A maior parte dos testes estatísticos assume que o número de óbitos segue uma distribuição aproximadamente Normal, uma hipótese razoável apenas nos casos em que  $d_x^{\text{exp}}$  é suficientemente grande

### Teste dos sinais

O teste dos sinais (*signs test*) considera apenas o sinal (positivo ou negativo) dos desvios. Se admitirmos que em cada idade a distribuição do número de óbitos tem média igual à mediana (como é o caso da distribuição Normal, mas já não acontece com a distribuição de Poisson), a probabilidade do sinal de cada desvio ser positivo ou negativo é igual a  $1/2$ , e o número de sinais positivos (ou negativos) segue uma distribuição  $\text{Bin} \sim (n, \frac{1}{2})$ . Seja  $n_1$  ( $n_2$ ) o número de sinais positivos (negativos) e  $n = n_1 + n_2$  o número de idades consideradas. Neste caso, temos

$$\mathbb{P}(n_1 = k) = \frac{n!}{k!(n-k)!} \cdot \left(\frac{1}{2}\right)^n \quad (3.7.3)$$

com valor esperado e variância iguais a  $n/2$  e  $n/4$ , respectivamente. Quando  $n$  é grande, podemos assumir que  $n_1$  segue uma distribuição aproximadamente Normal e calcular a estatística

$$W^{ST} = \frac{(\bar{p} - \pi_0)}{\sigma_p}, \text{ com } \sigma_p = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} \text{ e } W^{ST} \sim \mathcal{N}or(0, 1) \quad (3.7.4)$$

onde  $\pi_0 = 0.5$  e  $\bar{p}$  denota a proporção de valores abaixo (ou acima) de  $\pi_0$ .<sup>22</sup>

### Teste das sequências ou das mudanças de sinal (*Runs Test*)

Admitamos que o número de óbitos em cada idade tem uma distribuição Normal, que os desvios calculados em cada idade são independentes, que os sinais dos desvios são distribuídos aleatoriamente e que as sequências de desvios sucessivos com o mesmo sinal não são, nem muito grandes, nem muito curtas. Denotemos por  $R$  o número de sequências (*runs*) de um ou mais desvios consecutivos com o mesmo sinal. Se  $n_i \geq 20$  ( $i = 1, 2$ ), a distribuição do número de sequências é aproximadamente Normal com média

$$\mathbb{E}(R) = \frac{2n_1n_2}{n} + 1$$

---

(e.g.,  $d_x^{\text{exp}} \geq 5$ ). Quando tal não acontece (e.g., nas idades extremas), recomenda-se que previamente à aplicação dos testes seja feito o agrupamento de idades até que este valor seja atingido.

<sup>22</sup>Em alternativa, podemos recorrer à distribuição Binomial e calcular a probabilidade,  $p(+)$ , de que o número de sinais positivos não exceda o valor observado. Se o valor de  $p(+)$  for muito reduzido (e.g.,  $p(+)$  < 0.025), o número de desvios positivos é inesperadamente baixo, enquanto que se for muito elevado (e.g.,  $p(+)$  > 0.975), o número é anormalmente alto. Em qualquer dos casos, a conclusão a retirar é a de que as probabilidades graduadas estão muito afastadas (para um dos lados) dos valores observados, não representando por isso de forma adequada a experiência sob investigação. Em alternativa ao teste dos sinais, o teste de Wilcoxon (*Wilcoxon signed-rank test*) considera, quer o sinal, quer a magnitude dos desvios.

### 3.7. Indicadores de qualidade do ajustamento

e variância

$$\text{Var}(R) = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}.$$

Nestas condições, a estatística de teste utilizada é<sup>23</sup>

$$W^{RT} = \frac{R - \mathbb{E}(R)}{\sqrt{\text{Var}(R)}}, \quad W^{RT} \sim \mathcal{N}_{or}(0, 1). \quad (3.7.5)$$

#### Teste de Kolmogorov-Smirnov

O teste de Kolmogorov-Smirnov (KS) considera a distribuição do desvio máximo absoluto entre duas distribuições acumuladas. Pode ser usado na graduação para comparar as distribuições dos óbitos esperadas e observados entre as idades  $x_{\min}$  e  $x_{\max}$ . Seja  $D^{act} = \sum_{k=x_{\min}}^{x_{\max}} d_k$  o número observado de óbitos e  $D^{exp} = \sum_{k=x_{\min}}^{x_{\max}} d_k^{exp}$  o correspondente valor esperado. As respectivas distribuições acumuladas são definidas por

$$F(x) = \frac{1}{D^{act}} \left( \sum_{k=x_{\min}}^x d_k \right) \quad \text{e por} \quad G(x) = \frac{1}{D^{exp}} \left( \sum_{k=x_{\min}}^x d_k^{exp} \right).$$

A diferença máxima (absoluta) entre as duas distribuições é igual a

$$Dev^{\max} = \max_x |F(x) - G(x)|.$$

A estatística do teste de Kolmogorov-Smirnov,  $W^{KS}$ , é definida por

$$W^{KS} = Dev^{\max} \left( \frac{D^{act} D^{exp}}{D^{act} + D^{exp}} \right)^{\frac{1}{2}} \quad (3.7.6)$$

e tem distribuição conhecida, de modo que a probabilidade,  $p(W^{KS})$ , pode ser calculada e tabulada. Assim, uma vez determinado o nível de significância do teste, é possível comparar  $W^{KS}$  com o respectivo valor crítico e concluir sobre a rejeição ou não da hipótese nula de que as distribuições são iguais.<sup>24</sup>

<sup>23</sup> Em alternativa, podemos sempre calcular a probabilidade exacta,  $p(\text{runs})$ , de que o valor  $R$  seja menor ou igual ao valor observado de sequências. Um valor demasiado pequeno para  $p(\text{runs})$  significa que o número de mudanças de sinal é muito reduzido, o que é típico de uma graduação demasiado linear face aos dados observados, ou seja, sinaliza um ajustamento desajustado do modelo.

<sup>24</sup> A aplicação do teste de KS neste contexto deve ser feita com cautela, na medida em que a hipótese de que as séries são independentes não se verifica na prática. Com efeito, recorde-se que os óbitos esperados são calculadas com base em indicadores estimados a partir dos valores observados para as variáveis. Ainda assim, Forfar *et al.* (1988) argumentam que o teste é útil para despistar modelos desadequados com, por exemplo, valores muito elevados para  $Dev^{\max}$  (ou, de forma equivalente, valores muito baixos

### Teste do Chi-Quadrado

O teste do Chi-quadrado é um dos testes não-paramétricos mais usados para avaliar a qualidade do ajustamento. A estatística do teste

$$\chi^2 = \sum_{x_{\min}}^{x_{\max}} (z_x)^2 = \sum_{x_{\min}}^{x_{\max}} \frac{(d_x - d_x^{\text{exp}})^2}{\text{Var}(d_x)} \quad (3.7.7)$$

tem distribuição de  $\chi_{n-k}^2$  com  $(n - k)$  graus de liberdade, onde  $n$  representa o número de idades (ou grupos de idades) graduadas e  $k$  denota o número de parâmetros da função paramétrica usada. A probabilidade de obter um valor  $\chi^2$  superior a um valor observado,  $p(\chi^2)$ , pode ser calculada de forma exacta recorrendo a algoritmos matemáticos. Em alternativa, se o número de graus de liberdade for grande (e.g.,  $n - k \geq 30$ ), a probabilidade pode ser estimada com recurso à estatística

$$t(\chi^2) = \sqrt{2\chi^2} - \sqrt{2(n - k) - 1}, \quad t(\chi^2) \sim \text{Nor}(0, 1) \quad (3.7.8)$$

Uma vez fixado o nível de significância  $\alpha$ , a decisão sobre a rejeição da hipótese nula é feita com base na comparação entre o valor da estatística de teste e o valor crítico  $\chi_{n-k}^2(\alpha)$ .<sup>25</sup>

### Testes de autocorrelação

Este teste baseia-se na análise do comportamento das correlações entre os desvios relativos em cada idade. Cada desvio relativo,  $z_x$ , tem uma distribuição aproximadamente Normal com média nula e variância unitária. Os valores de  $z_x$  formam uma sequência, permitindo o cálculo dos sucessivos coeficientes de correlação  $\rho_l$  ( $l = 1, \dots, m$ ) entre os valores de  $z_x$  e  $z_{x-l}$ , i.e.,

$$\hat{\rho}_l = \frac{\sum_{k=l+1}^T (z_k - \bar{z})(z_{k-l} - \bar{z})}{\sum_{k=1}^T (z_k - \bar{z})^2}, \quad 0 \leq l \leq T - 1. \quad (3.7.9)$$

---

para  $p(W^{KS})$  a sinalizarem graduações insatisfatórias.

<sup>25</sup>Sublinhe-se, neste ponto, que o teste  $\chi^2$  é tanto um teste de independência das observações como um teste à qualidade do ajustamento, pelo que na prática muitas graduações consideradas satisfatórias apresentam valores de  $\chi^2$  elevados. Uma inspecção detalhada aos valores individuais de  $z_x$  pode, neste caso, ajudar a detectar padrões de irregularidade que indiquem problemas com os dados e/ou com o modelo.

### 3.7. Indicadores de qualidade do ajustamento

---

Se  $\{z_x\}$  é uma sequência de variáveis aleatórias e identicamente distribuídas com  $\mathbb{E}(z_x^2) < \infty$ , então  $\rho_l$  é (assintoticamente) normalmente distribuído com média nula e variância igual a  $1/n$ . Neste caso, podemos calcular a estatística

$$t_l = \frac{\hat{\rho}_l}{n^{-1/2}}, \quad t_l \sim \mathcal{N}or(0, 1) \quad (3.7.10)$$

e comparar o seu valor com o da distribuição Normal unitária. Valores demasiados elevados para  $t_l$  indiciam que os desvios sucessivos  $z_x$  estão relacionados de forma próxima, pelo que a graduação deve ser considerada insatisfatória.

Para testar a hipótese nula de que os coeficientes de autocorrelação para diferentes níveis de defasamento são conjuntamente nulos (i.e.,  $H_0 : \rho_1 = \dots = \rho_l = 0$ ), contra a hipótese alternativa  $H_1 : \rho_l \neq 0$  para algum  $l \in \{1, \dots, m\}$ , podemos recorrer à estatística de Portmanteau (Box e Pierce, 1970), definida por

$$Q^*(m) = n \sum_{l=1}^m \hat{\rho}_l^2, \quad Q^*(m) \sim \chi_m^2 \quad (3.7.11)$$

ou, em alternativa, adoptar a estatística de Ljung e Box (1978), definida por

$$Q(m) = n(n+2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{n-l} \quad (3.7.12)$$

que, como se observa, modifica a estatística (3.7.11) de maneira a melhorar o poder do teste em amostras finitas.

## Capítulo 4

# Métodos Não-Paramétricos de Graduação

*“Things should be made as simple as possible, but not any simpler.”* (Albert Einstein, 1879-1955)

### 4.1 Introdução

Os métodos paramétricos descritos no capítulo anterior constituem uma solução particularmente eficiente quando se dispõe de informação sobre o comportamento esperado das variáveis subjacentes. Em contrapartida, estes métodos introduzem uma fonte de erro adicional no problema, na medida em que existe sempre o risco de se seleccionar uma função paramétrica inadequada.

Em alternativa aos métodos paramétricos, inúmeros métodos de graduação ditos não-paramétricos foram sendo propostos na literatura actuarial para suavizar as estimativas brutas dos quocientes de mortalidade. O seu objectivo passa, novamente, pela substituição da série bruta  $\hat{Q} = \{\hat{q}_x, x = x_{\min}, \dots, x_{\max}\}$  por uma série graduada  $\hat{Q}^{gra} = \{\hat{q}_x^{gra}, x = x_{\min}, \dots, x_{\max}\}$ , mais regular ao longo de todo o arco da vida humana.

A principal diferença entre as abordagens reside no facto do comportamento da mortalidade não se resumir, neste caso, a um conjunto reduzido de parâmetros, continuando a ser descrito pela própria série de  $\mathcal{N} = (x_{\max} - x_{\min} + 1)$  quocientes anuais. Não existe, por isso, qualquer diminuição na dimensão do problema ou perda de informação. No passado, a complexidade numérica da maioria dos métodos não paramétricos constituía um sério entrave à sua utilização prática na ciência actuarial. O desenvolvimento da capacidade computacional e de aplicações informáticas cada vez mais sofisticadas eliminaram

definitivamente esta barreira.

Neste capítulo, analisamos os principais métodos não paramétricos usados em problemas de graduação. O capítulo está organizado da seguinte forma. Na Secção 4.2, exemplificamos a aplicação do clássico método de médias móveis no alisamento de uma série de indicadores de mortalidade. Na Secção 4.3, examinamos a popular técnica de splines, discutindo as variantes que decorrem da adopção de diferentes bases (polinomial, B-Splines,...). Analisamos igualmente as diferentes soluções propostas para mitigar o tradicional conflito entre os objectivos qualidade de ajustamento e alisamento da série. Na Secção 4.4 reavaliámos o método de Whittaker-Henderson, um método que deriva do modelo de regressão linear clássico mas que se distingue deste pelo facto de introduzir na função objectivo um termo que penaliza as diferenças entre valores graduados adjacentes. Uma solução semelhante é adoptada pelo método P-Splines, com a diferença de que neste caso a penalização incide sobre as diferenças finitas entre os coeficientes estimados das B-Splines adjacentes.

Na Secção 4.5, investigamos a aplicação de *smoothers* do tipo *kernel* em problemas de graduação, analisando as potencialidades oferecidas por esta variante, mais flexível, do método das médias móveis ponderadas. Na Secção 4.6, consideramos o método de regressão polinomial local ponderada, um método que envolve a aplicação de técnicas de regressão linear local numa vizinhança definida em torno de cada ponto da série. Na Secção 4.7, exploramos a flexibilidade oferecida pelos modelos aditivos generalizados enquanto formulação geral que abrange a maioria dos métodos não paramétricos de graduação analisados no capítulo. Por último, na Secção 4.8 discutimos alguns dos principais métodos propostos na literatura para seleccionar os parâmetros de alisamento dos modelos, um aspecto importante na medida em que estes regulam a importância dos objectivos qualidade do ajustamento e regularidade da série nos problemas de graduação.

## 4.2 Bin smoothers e médias móveis

### 4.2.1 Bin smoothers

Os *smoothers* do tipo “bin” baseiam-se na partição do intervalo de valores para a variável previsora  $x$  num conjunto mais ou menos numeroso de regiões adjacentes não sobrepostas (*bins*), após o qual se procede ao alisamento da variável de resposta  $y$  mediante cálculo da sua média em cada região. Em termos formais, seja  $R_k$  a região

$$R_k = \{c_k \leq x_i \leq c_{k+1}\} \quad k = 0, 1, \dots, n - 1,$$

onde  $c_0 < \dots < c_n$  define os pontos de corte, escolhidos, regra geral, de modo a que exista um número de pontos aproximadamente igual em cada região. O smoother  $s = S(y|x)$  é então definido pela sequência de valores

$$s(x_0) = \bar{y}_k = \underset{x_i \in R_k}{\text{média}}(y_i) \quad x_0 \in R_k. \quad (4.2.1)$$

A utilização deste método gera, na maior parte dos casos, um nível de alisamento da série desajustado uma vez que, por definição, a sequência de valores exibe descontinuidades em cada ponto de corte. Acresce que a qualidade da graduação é afectada, quer pela dimensão das regiões, quer pelo ponto de corte inicial.

Uma solução simples para melhorar o ajustamento da série consiste em aumentar o número de regiões. Tal implica, no entanto, uma diminuição do alisamento da série e coloca-nos perante o tradicional conflito entre qualidade de ajustamento e alisamento da série, ou entre enviesamento e variância. Uma solução alternativa para melhorar o alisamento da série passa pela utilização de regiões sobrepostas, como é o caso do método das médias móveis que a seguir introduzimos.

#### 4.2.2 Método das médias móveis

O método das médias móveis (MWA - *Moving Weighted Average*) foi um dos primeiros a ser adoptado na graduação de dados de mortalidade, constituindo uma forma prática e flexível de suavizar uma sequência de estimativas brutas. A sua aplicação assenta na definição, para cada valor de  $x_i$ , de uma vizinhança simétrica de dimensão  $k$ ,  $N_k^S(x_i)$ ,

$$N_k^S(x_i) = \{x_j : j \in \mathcal{I}_k^s(x_i)\}, \quad \mathcal{I}_k^s(x_i) = \{\max(1, i-k), \dots, i, \dots, \min(n, i+k)\},$$

após o que se procede à substituição da observação  $y_i$  pela média (simples ou ponderada) das observações registadas na sua vizinhança, i.e.,

$$s(x_i) = \bar{y}_k = \underset{j \in \mathcal{I}_k^s(x_i)}{\text{média}}(y_j). \quad (4.2.2)$$

O exemplo mais famoso da aplicação do método das médias móveis na graduação de dados de mortalidade é devido a Spencer (1904), com a sua fórmula de 21 pontos

$$\begin{aligned} \hat{q}_x = & \frac{1}{350} [60\hat{q}_x + 57(\hat{q}_{x+1} + \hat{q}_{x-1}) + 47(\hat{q}_{x+2} + \hat{q}_{x-2}) + 33(\hat{q}_{x+3} + \hat{q}_{x-3}) \\ & + 18(\hat{q}_{x+4} + \hat{q}_{x-4}) + 6(\hat{q}_{x+5} + \hat{q}_{x-5}) - 2(\hat{q}_{x+6} + \hat{q}_{x-6}) - 5(\hat{q}_{x+7} + \hat{q}_{x-7}) \\ & - 5(\hat{q}_{x+8} + \hat{q}_{x-8}) - 3(\hat{q}_{x+9} + \hat{q}_{x-9}) - (\hat{q}_{x+10} + \hat{q}_{x-10})], \end{aligned} \quad (4.2.3)$$

### 4.3. Técnica de Splines

---

onde  $\hat{q}_x$  denota o quociente bruto e  $\hat{q}_x$  o respectivo valor graduado.

Não obstante a sua simplicidade, o método apresenta uma limitação importante, que decorre do facto de nos pontos iniciais e finais da série, que distem menos de  $k$  unidades dos extremos, a vizinhança conter um número de observações inferior aos restantes introduzindo, desta forma, algum enviesamento na estimativa.<sup>1</sup>

## 4.3 Técnica de Splines

A utilização da técnica de splines fornece uma solução alternativa para a graduação da série bruta. Um spline é uma curva definida matematicamente por dois ou mais pontos de controlo, designados por nós (*knots*). A graduação mediante splines baseia-se na concatenação de funções polinomiais de determinado grau, definidas por troços, e sujeitas a determinadas condições de modo a garantir a continuidade da função ao longo de todo o intervalo.<sup>2</sup>

Os splines podem ser divididos em duas grandes categorias: (i) splines de interpolação, que passam por todos os pontos de controlo, e (ii) splines de aproximação, que passam perto de todos os pontos de controlo. Os splines podem ainda ser distinguidos segundo as funções de base escolhidas para representar a função interpoladora. De entre as alternativas disponíveis, são analisadas nesta secção, pelo seu interesse em problemas de graduação, as seguintes: Splines cúbicos, B-Splines e P-Splines.

### 4.3.1 Splines cúbicos

De entre as múltiplas configurações possíveis, uma das escolhas mais populares compreende a utilização de polinómios cúbicos definidos por troços, com primeira e segunda derivadas contínuas nos nós da função. Uma restrição adicional, destinada a garantir a linearidade da função para além dos nós de fronteira, é aquela que exige que a segunda

---

<sup>1</sup>Para lidar com o problema da aplicação do método das médias móveis aos extremos da série, autores como Greville (1981a,b) e Hoem e Linnemann (1988), entre outros, sugerem duas soluções tipo. A primeira, envolve o ajustamento de um polinómio aos pontos extremos da sequência de estimativas iniciais, usando em seguida os valores ajustados como valores graduados. A segunda, implica o alargamento da série mediante a adição de pontos aos extremos iniciais (e.g., por ajustamento de uma função e extrapolação), para em seguida utilizar as médias móveis em todo o intervalo desejado. As limitações do método e a emergência de soluções mais sofisticadas, tornada possível pelo assinalável desenvolvimento dos meios informáticos, conduziu a uma progressiva diminuição da utilização do método das médias móveis em problemas de graduação.

<sup>2</sup>Para uma descrição mais pormenorizada da técnica de splines e da sua aplicação no contexto da graduação veja-se, por exemplo, Greville (1969), Lancaster e Salkaukas (1986), Benjamin e Pollard (1980) e De Boor (2001).

### 4.3. Técnica de Splines

---

derivada da função nos extremos do intervalo seja nula. Os splines assim definidos são denominados de splines cúbicos naturais.

Em termos formais, considere-se um intervalo de idades arbitrário  $\mathcal{I} = [a, b]$ , particionado em  $n$  sub-intervalos  $\mathcal{I}_i = [x_i, x_{i+1}]$  ( $i = 0, \dots, n-1$ ) por  $n+1$  nós  $a = x_0 < x_1 < \dots < x_n = b$ . Nos casos em que os nós são distribuídos de forma equidistante no intervalo  $[a, b]$ , o spline diz-se *uniforme*, caso contrário é designado de *não uniforme*. Uma função  $s$  é denominada spline de grau  $k$  com nós nos pontos  $\{x_i\}_{i=0}^n$  se é definida por um polinómio de grau  $k$  em cada intervalo  $[x_i, x_{i+1}]$  e é  $(k-1)$  vezes diferenciável em todo o intervalo  $[a, b]$ .

Dito isto, podemos definir a interpolação de uma função  $g(x)$ , explicitada em  $[a, b]$ , por intermédio de splines cúbicos naturais com nós em  $x_0, \dots, x_n$ , como uma função  $s(x)$  que verifica as seguintes condições:

1.

$$s(x) \in C^2[a, b].$$

2. Em cada intervalo  $[x_i, x_{i+1}]$ , a função é definida por um polinómio cúbico do tipo:

$$s(x) = s_i(x) = \sum_{l=0}^3 a_l^{(i)} (x - x_i)^l, \quad i = 0, \dots, n-1.$$

3. Em cada um dos nós  $x_i$  ( $i = 1, \dots, n$ )

$$s(x_i) = g(x_i).$$

4. Condições de continuidade nos nós interiores

$$\begin{cases} s_i(x_{i+1}) = s_{i+1}(x_{i+1}) \\ s'_i(x_{i+1}) = s'_{i+1}(x_{i+1}) \\ s''_i(x_{i+1}) = s''_{i+1}(x_{i+1}). \end{cases}$$

5. São respeitadas as seguintes condições limite

$$s''(a) = s''(b) = 0.$$

Dada a série  $(x_i, y_i)$ , os coeficientes  $a_l^{(i)}$  ( $l = 0, \dots, 3$ ) do polinómio são estimados

minimizando uma função objectivo do tipo

$$\hat{a}_l^{(i)} = \arg \min_{a_l^{(i)}} \left\{ \sum_{i=1}^n w_i (y_i - s(x_i))^2 \right\}, \quad (4.3.1)$$

onde  $w_i$  representa um conjunto de ponderadores pré-definido.

Em face do habitual conflito entre qualidade do ajustamento e alisamento da série, um dos aspectos mais sensíveis na implementação deste método prende-se com a escolha do número de nós e respectiva localização no intervalo. A opção mais simples consiste em escolher apenas o número de nós interiores, posicionando-os de seguida de modo a obter uma distribuição uniforme ao longo do intervalo dos dados. Um esquema alternativo consiste em posicionar os nós nos quartis da variável independente ou em pontos de inversão da concavidade da função. Mais recentemente, Friedman e Silverman (1989) desenvolvem um mecanismo adaptativo em que o número e a localização dos nós é otimizado com recurso a um algoritmo.<sup>3</sup>

Uma forma de obter um compromisso entre os dois objectivos passa pela inclusão de um termo adicional em (4.3.1) de modo a penalizar a sobreparametrização do modelo ou, o que é o mesmo, as oscilações excessivas na curvatura da série. Os splines cúbicos possuem, a este propósito, uma propriedade conhecida e interessante: são, de entre todas as funções  $s(x)$  com duas derivadas contínuas, aqueles que minimizam a seguinte soma dos quadrados penalizados

$$\hat{a}_l^{(i)} = \arg \min_{a_l^{(i)}} \left\{ \sum_{i=1}^n w_i (y_i - s(x_i))^2 + \lambda \int_a^b (s''(u))^2 du \right\}, \quad (4.3.2)$$

onde  $\lambda$  ( $\lambda \geq 0$ ) é um parâmetro de *smoothing* fixo. Como se observa, o primeiro termo da função objectivo (4.3.2) mede o grau de ajustamento face aos valores observados, enquanto que o segundo introduz uma penalização directamente proporcional às oscilações na curvatura da função interpoladora. No limite, à medida que  $\lambda \rightarrow \infty$  o termo correspondente à penalização domina, forçando  $s''(x_i) = 0$  em todo o domínio, tendendo a solução para a convencional aproximação pelos mínimos quadrados. No extremo oposto, quando  $\lambda \rightarrow 0$ , o termo penalizador torna-se irrelevante e a solução tende para uma função interpoladora duplamente diferenciável.

A solução  $\hat{s}_\lambda(\cdot)$  do problema de minimização (4.3.2) é um spline cúbico em que os nós são os pontos  $x_0, x_1, \dots, x_n$ , ou seja,  $\hat{s}_\lambda(\cdot)$  traduz um polinómio de terceiro grau em

<sup>3</sup>Para uma aplicação da técnica de splines em problemas de graduação veja-se, por exemplo, Forfar *et al.* (1988), McCutcheon (1987) e referências neles incluídas.

### 4.3. Técnica de Splines

---

cada intervalo  $[x_i, x_{i+1}]$ , com primeira e segunda derivadas contínuas em cada um dos  $x_i$ . O problema (4.3.2) equivale, em termos matriciais, a

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \{(\mathbf{y} - \mathbf{s})' \mathbf{W}(\mathbf{y} - \mathbf{s}) + \lambda \mathbf{s}' \mathbf{P} \mathbf{s}\}, \quad (4.3.3)$$

onde

$$\mathbf{P} = \mathbf{\Omega}' \mathbf{Q}^{-1} \mathbf{\Omega}$$

e onde  $\mathbf{\Omega}$  é uma matriz tri-diagonal de dimensão  $(n-2) \times n$  definida por

$$\mathbf{\Omega} = \begin{pmatrix} \frac{1}{\xi_1} & -\left(\frac{1}{\xi_1} + \frac{1}{\xi_2}\right) & \frac{1}{\xi_2} & \cdots & 0 \\ 0 & \frac{1}{\xi_2} & -\left(\frac{1}{\xi_2} + \frac{1}{\xi_3}\right) & \ddots & 0 \\ 0 & 0 & \frac{1}{\xi_3} & \ddots & \frac{1}{\xi_{n-1}} \\ \vdots & \vdots & \vdots & \ddots & -\left(\frac{1}{\xi_{n-2}} + \frac{1}{\xi_{n-1}}\right) \\ 0 & 0 & 0 & \cdots & \frac{1}{\xi_{n-2}} \end{pmatrix}$$

$\xi_i = x_{a+i} - x_{a+i-1}$ , e  $\mathbf{Q}$  é uma matriz tri-diagonal simétrica de dimensões  $(n-2) \times (n-2)$  definida por

$$\mathbf{Q} = \begin{pmatrix} 2(\xi_1 + \xi_2) & \xi_2 & \cdots & 0 \\ \xi_2 & 2(\xi_2 + \xi_3) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \xi_{n-2} \\ 0 & \cdots & \xi_{n-2} & 2(\xi_{n-2} + \xi_{n-1}) \end{pmatrix}$$

e  $\mathbf{W}$  é uma matriz diagonal contendo os ponderadores.

A solução  $\hat{s}_\lambda(\cdot)$  deriva-se anulando o gradiente da função objectivo (4.3.3), i.e.,

$$-2\mathbf{W}(\mathbf{y} - \mathbf{s}) + 2\lambda\mathbf{P}\mathbf{s} = 0,$$

donde se obtém finalmente

$$\hat{s}_\lambda = (\mathbf{W} + \lambda\mathbf{P})^{-1} \mathbf{W}\mathbf{y}. \quad (4.3.4)$$

O parâmetro  $\lambda$  regula, como vimos, a importância relativa dos objectivos na gradação e assume, como veremos mais adiante, um significado semelhante ao conceito de *kernel bandwidth*. Pode ser determinado de forma subjectiva ou, como explicitaremos em seguida, resultar de um procedimento automático como seja o de validação cruzada. Registe-se, por fim, que o método preconizado pela função objectivo (4.3.2) se encontra

amplamente difundido na literatura sobre graduação com a designação de fórmula de Whittaker-Henderson, que analisaremos em detalhe na Secção 4.4.

### 4.3.2 B-Splines

O ajustamento mediante splines cúbicos depara-se, muitas vezes, com um problema de quase colinearidade da matriz de regressores. Para contornar este problema, o recurso às denominadas *Basis Splines* (B-Splines) fornece uma solução estável e computacionalmente mais eficiente para aproximar uma determinada função.

Considere-se um intervalo de idades  $[x_{\min}, x_{\max}]$ , particionado em  $n-1$  sub-intervalos por  $n$  nós. Cada sub-intervalo é coberto por  $m+1$  B-Splines de grau  $m$ . É garantida a união dos segmentos polinomiais nos nós interiores, onde as derivadas de ordem  $m-1$  são contínuas.

Dada uma sequência de nós  $\{x_i\}_{i=0}^n$  e  $p = 0, \dots, n-4$ , uma função B-Spline cúbica ( $m = 3$ ) é definida por (Powell, 1981)<sup>4</sup>

$$B_p(x) = \sum_{j=p}^{p+4} \left[ \left( \prod_{i=p, i \neq j}^{p+4} \frac{1}{(x_i - x_j)} \right) \right] [\max(x - x_p, 0)]^3. \quad (4.3.5)$$

As funções  $B_p(x)$  dependem unicamente da sequência de nós e apenas assumem valores não nulos no intervalo  $[x_p, x_{p+4}]$ , i.e.,

$$B_p(x) = \begin{cases} 0 & : x \in (-\infty, x_p) \\ B_p(x) & : x \in [x_p, x_{p+4}] \\ 0 & : x \in (x_{p+4}, +\infty). \end{cases}$$

O número total de nós necessário para derivar as B-Splines é igual a  $n + 2m$ .<sup>5</sup> O número total de B-Splines na regressão será de  $s = n - 1 + m$ . Dada a sequência de nós, uma função B-Spline cúbica é uma representação paramétrica resultante da combinação linear das funções base  $B_p(x)$ , i.e.,

$$s(x) = \sum_{p=1}^s a_p B_p(x), \quad (4.3.6)$$

---

<sup>4</sup>Veja-se De Boor (2001) para uma definição alternativa das funções base B-Spline, assente na aplicação de uma fórmula recursiva que depende apenas da sequência de nós.

<sup>5</sup>Tal é assim porquanto são precisas  $n+3$  funções para formar uma base. Como pela equação (4.3.5) só conseguimos derivar  $(n-3)$  B-Splines cúbicas usando os nós  $\{x_i\}_{i=0}^n$ , são necessárias  $2m$  funções adicionais. Uma solução *ad hoc* envolve a imposição de  $2m$  nós adicionais  $\{x_{-3}, x_{-2}, x_{-1}, x_{n+1}, x_{n+2}, x_{n+3}\}$ , adjacentes aos extremos do intervalo original.

### 4.3. Técnica de Splines

---

onde  $a_p$  são coeficientes a estimar.

Considere a sequência de  $N$  pontos  $(x_i, y_i)$  e o conjunto de funções base  $B_p(x)$ . Em problemas de graduação,  $x_i$  e  $y_i$  correspondem, por exemplo, à idade actuarial da pessoa segura e a um indicador da mortalidade ( $q_x$  ou  $\mu_x$ ), respectivamente. As estimativas dos parâmetros  $a_p$  obtêm-se pela resolução do seguinte problema de optimização

$$\hat{a}_p = \arg \min_{a_p} \left\{ \sum_{i=1}^N \left( y_i - \sum_{p=1}^s a_p B_p(x) \right)^2 \right\}. \quad (4.3.7)$$

Em alternativa, podemos optar pela formulação matricial do problema (4.3.7)

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \{ (\mathbf{y} - \mathbf{B}\mathbf{a})' (\mathbf{y} - \mathbf{B}\mathbf{a}) \}, \quad (4.3.8)$$

onde

$$\mathbf{B} = \begin{pmatrix} B_1(x_1) & B_2(x_1) & \cdots & B_p(x_1) \\ B_1(x_2) & B_2(x_2) & \cdots & B_p(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(x_N) & B_2(x_N) & \cdots & B_p(x_N) \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix}'$$

$$\mathbf{y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_p \end{pmatrix}'.$$

A solução do problema de optimização é então dada por

$$\hat{\mathbf{a}} = (\mathbf{B}'\mathbf{B})^{-1} \mathbf{B}'\mathbf{y}. \quad (4.3.9)$$

Por fim, a sequência de valores graduados é representada por  $\hat{\mathbf{y}} = \mathbf{B}\hat{\mathbf{a}} = \mathbf{H}\mathbf{y}$ , onde  $\mathbf{H}$  designa a denominada “*hat-matrix*”

$$\mathbf{H} = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \mathbf{B}'. \quad (4.3.10)$$

A utilização de B-Splines fornece uma ferramenta poderosa no contexto dos métodos não paramétricos de graduação mas enfrenta, à semelhança dos splines cúbicos, o dilema da escolha do número óptimo de nós e sua respectiva localização no intervalo da variável independente.

Uma primeira solução consiste em introduzir na função objectivo uma penalização  $\lambda$  (constante) que incida, tal como anteriormente, sobre a segunda derivada da curva

ajustada

$$\hat{a}_p = \arg \min_{a_p} \left\{ \sum_{i=1}^N \left( y_i - \sum_{p=1}^s a_p B_p(x) \right)^2 + \lambda \int_{x_{\min}}^{x_{\max}} \left( \sum_{p=1}^s a_p B_p''(x) \right)^2 dx \right\}. \quad (4.3.11)$$

### 4.3.3 P-Splines

Em alternativa à função objectivo definida em (4.3.11), Eilers e Marx (1996) propõem a utilização de um número relativamente grande de nós e a imposição de uma penalização (de ordem  $d$ ) incidindo sobre a diferença finita entre os coeficientes estimados das B-Splines adjacentes.

O método, “baptizado” pelos autores de P-Splines (*Penalized Splines*), baseia-se na estimativa dos parâmetros pela resolução do seguinte problema de optimização

$$\hat{a}_p = \arg \min_{a_p} \left\{ \sum_{i=1}^N \left( y_i - \sum_{p=1}^s a_p B_p(x) \right)^2 + \lambda \sum_{p=d+1}^s \left( \Delta^d a_p \right)^2 \right\}, \quad (4.3.12)$$

onde  $\Delta^d$  denota a  $d^{\text{ésima}}$  diferença de  $a_p$  e  $\lambda$  representa novamente um termo de penalização, que regula o grau de alisamento da função. Recorde-se que  $\Delta a_p = a_p - a_{p-1}$ ,  $\Delta^2 a_p = \Delta(\Delta a_p) = a_p - a_{p-1} - (a_{p-1} - a_{p-2}) = a_p - 2a_{p-1} + a_{p-2}$ , e assim sucessivamente para valores de ordem superior para  $d$ .

Os autores demonstram que a penalização das diferenças finitas constitui uma boa aproximação ao integral da segunda derivada em (4.3.11), com a vantagem de que é de mais fácil incorporação no procedimento de optimização.<sup>6</sup> No caso particular em que  $\lambda = 0$ , obtemos o tradicional modelo de regressão linear (4.3.9) com uma base formada por B-Splines.

Em termos matriciais, a função objectivo a minimizar é agora dada por

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \{ (\mathbf{y} - \mathbf{B}\mathbf{a})' (\mathbf{y} - \mathbf{B}\mathbf{a}) + \lambda \mathbf{a}' \mathbf{D}'_d \mathbf{D}_d \mathbf{a} \}, \quad (4.3.13)$$

onde  $\mathbf{D}_d$  representa o operador de diferenças  $\Delta^d$  em notação matricial.

<sup>6</sup> Conforme veremos na secção seguinte, a penalização das diferenças está igualmente na raiz do clássico modelo de Whittaker (1923) e Henderson (1924).

### 4.3. Técnica de Splines

---

Por exemplo, no caso em que  $d = 2$ ,  $D_d$  é uma matrix de segundas diferenças do tipo

$$D_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix}$$

A solução do problema de optimização é dada por

$$\hat{\mathbf{a}} = (\mathbf{B}'\mathbf{B} + \lambda\mathbf{D}'_d\mathbf{D}_d)^{-1} \mathbf{B}'\mathbf{y} \quad (4.3.14)$$

A sequência de valores graduados é representada por  $\hat{\mathbf{y}} = \mathbf{B}\hat{\mathbf{a}} = \mathbf{H}\mathbf{y}$ , onde  $\mathbf{H}$  é neste caso definida por

$$\mathbf{H} = \mathbf{B} (\mathbf{B}'\mathbf{B} + \lambda\mathbf{D}'_d\mathbf{D}_d)^{-1} \mathbf{B}'. \quad (4.3.15)$$

Eilers e Marx (1996) sugerem a incorporação das P-Splines no contexto dos modelos lineares generalizados (GLMs), desenvolvendo um novo método de gradação não paramétrica. A metodologia assenta nos seguintes princípios: (i) utilização das B-Splines como base para a regressão, e (ii) modificação da função de log-verosimilhança mediante a inclusão de uma penalização baseada na diferença entre os coeficientes.

Em termos formais, seja  $E_x$  o número de expostos ao risco com idade  $x$ ,  $d_x$  o correspondente número de óbitos registados e admitamos que  $D_x$  segue uma distribuição de Poisson com média  $m_x = E_x\mu_x$ . Considere-se o seguinte predictor linear  $\eta_i = \sum_{p=1}^s a_p B_{ip}(x)$  e o respectivo *link* logarítmico  $\eta_i = \ln(\mu_i)$ . A introdução da penalização no modelo conduz-nos à seguinte função objectivo

$$\mathcal{L}_P = \mathcal{L}(\mathbf{y}; \mathbf{a}) - \frac{1}{2} \lambda \sum_{p=d+1}^s (\Delta^d a_p)^2 \quad (4.3.16)$$

ou à sua equivalente em termos matriciais

$$\mathcal{L}_P = \mathcal{L}(\mathbf{y}; \mathbf{a}) - \frac{1}{2} \lambda \mathbf{a}' \mathbf{D}'_d \mathbf{D}_d \mathbf{a}, \quad (4.3.17)$$

onde  $\mathcal{L}(\mathbf{y}; \mathbf{a})$  representa a função de log-verosimilhança. A estimação do vector de parâmetros  $\mathbf{a}$  é conseguida recorrendo a um procedimento iterativo em que

$$\mathbf{a}_{t+1} = (\mathbf{B}'\mathbf{W}_t\mathbf{B} + \lambda\mathbf{D}'_d\mathbf{D}_d)^{-1} \mathbf{B}'\mathbf{W}_t\mathbf{z}_t, \quad (4.3.18)$$

onde  $\mathbf{z} = \boldsymbol{\eta} + \mathbf{W}^{-1}(\mathbf{d} - \mathbf{y})$  representa a variável de controlo.  $\mathbf{W} = \text{diag}(\mathbf{y})$  denota a matrix diagonal contendo os ponderadores, cujo elemento genérico é dado por

$$w_{ii} = \frac{1}{v_i} \left( \frac{\partial y_i}{\partial \eta_i} \right)^2,$$

onde  $v_i$  denota a variância de  $d_i$ , dado  $y_i$ .

Os principais vantagens do método das P-Splines são a sua flexibilidade, a facilidade de computação e a sua ligação com os métodos de regressão polinomial e “smoothing splines”. Acresce que, conforme teremos oportunidade de verificar no capítulo 6, o método pode, entre outras aplicações, ser adoptado na graduação de vectores pluridimensionais de dados e na elaboração de projecções de mortalidade. O próprio método de Whittaker-Henderson, que analisaremos em profundidade na secção seguinte, pode ser considerado um caso particular das P-Splines.

## 4.4 Método de Whittaker-Henderson

Whittaker (1923) e Henderson (1924) desenvolveram um dos primeiros métodos não paramétricos usados em problemas de graduação.<sup>7</sup> O método assenta, tal como no caso das P-Splines, na minimização de uma função objectivo que conjuga os objectivos de qualidade do ajustamento e alisamento da série.

Designemos por  $x$  a idade do indivíduo, com  $x_{\min} \leq x \leq x_{\max}$ , por  $d_x$  o número observado de óbitos e por  $E_x$  a estimativa da exposição (inicial ou central) ao risco, com estimativas brutas de  $q_x$  ou  $\mu_x$  dadas por  $d_x/E_x$ . Designemos por  $\theta_x$  as “verdadeiras” taxas de mortalidade e por  $u_x$  as estimativas brutas iniciais (não graduadas). O método de Whittaker-Henderson consiste em determinar os valores graduados,  $y_x$ , que minimizam a função objectivo

$$\mathcal{Q}(y) = \sum_{x=x_{\min}}^{x_{\max}} w_x (y_x - u_x)^2 + \lambda \sum_{x=x_{\min}+d+1}^{x_{\max}} \left( \Delta^d y_x \right)^2, \quad (4.4.1)$$

onde  $w_x$  representa os ponderadores,  $\lambda$  ( $\lambda \geq 0$ ) denota uma vez mais um parâmetro de smoothing e  $\Delta^d y_x$  simboliza o operador de diferenças finitas aplicado  $d$  vezes à quantidade  $y_x$ . Nos casos em que se deseja maior controlo sobre o ajustamento, o método admite a possibilidade do parâmetro de penalização poder variar com a idade, de modo a

---

<sup>7</sup>Para uma aplicação recente do método de Whittaker-Henderson na graduação de dados de mortalidade veja-se, por exemplo, IAA (2001) e Panjer e Tan (1995).

#### 4.4. Método de Whittaker-Henderson

---

capturar comportamentos específicos ao longo da curva de mortalidade. Os ponderadores  $\{w_x\}$  são escolhidos de modo a dar mais importância aos valores mais significativos de  $u_x$ , sendo normalmente aproximados por  $w_x = 1/\text{Var}(u_x)$ .

O primeiro termo em (4.4.1) valoriza o nível de ajustamento entre as estimativas iniciais e os valores graduados, ponderando os desvios quadráticos mediante a atribuição de pesos distintos a cada desvio. O segundo termo representa, uma vez mais, uma medida da suavidade da sequência das estimativas graduadas.

As semelhanças entre este método e as variantes da técnica de splines acima descritas (em especial das P-Splines) são evidentes, com a diferença de que neste caso as estimativas de  $y_x$  são obtidas directamente ao passo que no método P-Splines elas derivam, indirectamente, da estimação dos coeficientes  $a_j$  dos polinómios da função interpoladora.

A função objectivo (4.4.1) pode ser escrita em termos matriciais da seguinte forma

$$\begin{aligned} Q(y) &= (\mathbf{y} - \mathbf{u})' \mathbf{W} (\mathbf{y} - \mathbf{u}) + \lambda (\mathbf{D}_d \mathbf{y})' (\mathbf{D}_d \mathbf{y}) \\ &= (\mathbf{y} - \mathbf{u})' \mathbf{W} (\mathbf{y} - \mathbf{u}) + \lambda \mathbf{y}' (\mathbf{D}'_d \mathbf{D}_d) \mathbf{y} \end{aligned} \quad (4.4.2)$$

onde

$$\begin{aligned} \mathbf{y} &= \begin{pmatrix} y_{x_{\min}} & y_{x_{\min}+1} & \cdots & y_{x_{\max}} \end{pmatrix}' \\ \mathbf{u} &= \begin{pmatrix} u_{x_{\min}} & u_{x_{\min}+1} & \cdots & u_{x_{\max}} \end{pmatrix}' \\ \mathbf{W} &= \begin{pmatrix} w_{x_{\min}} & 0 & \cdots & 0 \\ 0 & w_{x_{\min}+1} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & w_{x_{\max}} \end{pmatrix}, \quad \mathbf{W} = \mathbf{W}' \end{aligned}$$

e onde  $\mathbf{D}_d$  representa o operador de diferenças  $\Delta^d$  em notação matricial. Das condições de primeira ordem para a minimização de  $Q(y)$ , resulta

$$\frac{dQ(y)}{dy} = 0 \Leftrightarrow 2\mathbf{W}(\mathbf{y} - \mathbf{u}) + 2\lambda (\mathbf{D}'_d \mathbf{D}_d) \mathbf{y} = 0$$

ou, de forma equivalente,

$$[\mathbf{W} + \lambda (\mathbf{D}'_d \mathbf{D}_d)] \mathbf{y} = \mathbf{W} \mathbf{u}.$$

A estimativa dos valores graduados é dada por

$$\mathbf{y} = [\mathbf{W} + \lambda (\mathbf{D}'_d \mathbf{D}_d)]^{-1} \mathbf{W} \mathbf{u}. \quad (4.4.3)$$

Demonstra-se que a matriz  $[\mathbf{W} + \lambda(\mathbf{D}'_d\mathbf{D}_d)]$  é definida positiva, pelo que a estimativa de  $y$  pode ser obtida, por exemplo, recorrendo à factorização de Cholesky.

Chan *et al.* (1986) demonstram que o método de Whittaker-Henderson pode ser formulado em termos de um problema de programação linear ou de um problema de programação quadrática. Taylor (1992) e Verrall (1993a) desenvolvem uma derivação bayesiana formal da fórmula de Whittaker-Henderson, onde argumentam que a função objectivo do modelo pode ser interpretada como uma função de perda semelhante a uma verosimilhança penalizada. Verrall (1993a) prova ainda que a graduação pode ser formulada em termos de um modelo linear generalizado dinâmico, em que a estimação é feita recursivamente mediante o filtro de Kalman.

Lowrie (1982) introduz na fórmula de Whittaker-Henderson a possibilidade do ajustamento ser feito em relação a uma tábua standard. Seja  $\hat{\theta}_x$  o valor retirado de uma tábua base. A função objectivo preconizada pelo modelo é definida por

$$\mathcal{Q}_1(y) = (1 - \alpha) \sum_{x=x_{\min}}^{x_{\max}} w_x(y_x - u_x)^2 + \alpha \sum_{x=x_{\min}}^{x_{\max}} w_x(y_x - \hat{\theta}_x)^2 + \lambda \sum_{x=x_{\min}+d+1}^{x_{\max}} (\Delta^d y_x)^2, \quad (4.4.4)$$

onde  $\alpha$  ( $0 \leq \alpha \leq 1$ ) determina a importância dada ao ajustamento face às estimativas iniciais e face aos valores da tábua standard, enquanto as demais variáveis preservam o seu significado anterior.

Schuette (1978) desenvolve uma variante do modelo (4.4.1) em que sugere a minimização dos desvios absolutos em detrimento da minimização dos desvios quadráticos. Em termos formais, a função objectivo correspondente é definida por

$$\mathcal{Q}_2(y) = \sum_{x=x_{\min}}^{x_{\max}} w_x |y_x - u_x| + \lambda \sum_{x=x_{\min}+d+1}^{x_{\max}} |\Delta^d y_x|. \quad (4.4.5)$$

Segundo o autor, quando os erros de graduação não seguem uma distribuição normal, a utilização de desvios quadráticos produz mais estimativas anómalas do que no caso em que se recorre aos desvios e diferenças em termos absolutos.

## 4.5 Kernel Smoothers

### 4.5.1 Definição e conceitos base

Os *smoothers* do tipo *kernel* constituem uma variante do método das médias móveis em que, para cada ponto focal  $x_0$  no espaço da variável explicativa, a estimativa ajustada

$\hat{y}_i$  é obtida com recurso a uma média ponderada dos pontos na sua vizinhança,

$$\hat{y}_i = \sum_{j=1}^n S_{0j} y_j, \quad (4.5.1)$$

com a diferença de que os pesos são determinados por uma função *kernel*

$$S_{0j} = c_0 K\left(\frac{|x_0 - x_j|}{b}\right), \quad \sum_{j=1}^n S_{0j} = 1. \quad (4.5.2)$$

onde  $c_0$  é uma constante escolhida de modo a garantir que a soma dos pesos iguala a unidade, e o parâmetro  $b$  (habitualmente chamado de *bandwith*) determina a largura da vizinhança em torno do ponto focal. A suavidade do ajustamento é directamente proporcional ao valor do parâmetro  $b$ , que desempenha aqui um papel semelhante ao referido anteriormente para o parâmetro  $\lambda$  na técnica de splines ou no método de Whittaker-Henderson.

A função *kernel*  $K(u)$ , decrescente em  $|u|$ , é usada para calcular os pesos e cumpre, regra geral, as seguintes condições<sup>8</sup>

- a)  $K(u) \geq 0, \forall u$
- b)  $\int_{-\infty}^{\infty} K(u) du = 1$
- c)  $K(-u) = K(u), \forall u$  (simetria).

A cada escolha da função  $K(u)$  corresponde um tipo diferente de *kernel*. Entre os mais populares incluem-se o Normal, o Triângulo, o Parzen, o Epanechnikov, entre outros. Se é verdade que existem funções *kernel* com maior justificação teórica que outras, a verdade é a que a literatura sobre esta matéria é unânime em concluir que a escolha da função *kernel* é menos importante que a selecção do parâmetro de *bandwith*. Independentemente do *kernel* escolhido, a contribuição de cada observação  $y_j$  para a estimativa  $\hat{y}_i$  é tanto maior quanto mais próximo estiver  $x_j$  do ponto focal, sendo que a região de interesse é determinada pelo parâmetro  $b$ .

#### 4.5.2 Graduação mediante *kernel* smoothers

A introdução dos *kernel smoothers* na literatura actuarial é devida a Copas e Haberman (1983) e Ramlau-Hansen (1983). Numa série de trabalhos, Gavin *et al.* (1993,1994,1995)

---

<sup>8</sup>As condições (a) e (b) são habitualmente respeitadas pelas funções de densidade de probabilidade. Refira-se ainda que a condição (a) é uma condição frequente mas não obrigatória, existindo casos (e.g., Epanechnikov Kernel) em que são admitidos valores negativos.

analisam as potencialidades proporcionadas pela aplicação de *kernel smoothers* como método de graduação não paramétrica.

Em Gavin *et al.* (1993), os autores estabelecem a conexão entre os *kernel smoothers* e a graduação mediante médias móveis ponderadas. Em Gavin *et al.* (1994), os autores descrevem a aplicação do mecanismo de validação cruzada (*cross-validation*) na selecção do parâmetro de bandwith no contexto da graduação não paramétrica. Em Gavin *et al.* (1995), os autores desenvolvem um modelo que contempla de forma explícita a variação do número de expostos ao risco com a idade, introduzem a possibilidade de transformar os dados como forma de reduzir o enviesamento introduzido pela estimação *kernel* nos extremos da tábua de mortalidade, e discutem as potencialidades de uma função *kernel* desenhada de modo a melhorar as estimativas obtidas nos extremos do intervalo de idades.

Considere-se um conjunto específico de quocientes de mortalidade (brutos)  $\hat{q}_i = (d_x/E_x)$ , estimados num intervalo de idades a partir das séries de óbitos  $\{d_i\}_{i=1}^n$  e  $\{E_i\}_{i=1}^n$  expostos ao risco. Gavin *et al.* (1994) descrevem e comparam o desempenho de dois estimadores clássicos, atribuídos a Copas e Haberman (1983) e a Nadaraya (1964) e Watson (1964).

O estimador de Copas-Haberman (CH) para  $q_x$ ,  $q_x^{CH}$ , é definido por

$$\hat{q}_x^{CH} = \frac{\sum_{i=1}^n d_i K_b(x - x_i)}{\sum_{i=1}^n E_i K_b(x - x_i)}, \quad (4.5.3)$$

onde  $K_b(x) = b^{-1}K(x/b)$  é uma função *kernel*. Como se observa, o estimador usa a mesma função *kernel* e o mesmo bandwith no numerador e no denominador. No caso em que  $K(x) \geq 0$ , garante-se que o valor estimado de  $q_x$  apenas assume valores no intervalo  $[0, 1]$ .

Se admitirmos que o número de óbitos na idade  $x$  segue uma distribuição binomial com média  $E_x q_x$  e variância  $E_x q_x (1 - q_x)$ , o estimador do Copas-Haberman é obtido pela minimização da seguinte função de log-verosimilhança local

$$\sum_{i=1}^n K_b(x - x_i) \{d_i \log(\hat{q}_i) + (E_i - d_i) \log(1 - \hat{q}_i)\}. \quad (4.5.4)$$

Se ignorarmos a binomialidade inerente a  $q_x^{CH}$ , os  $d_i$  óbitos registados num total de  $E_i$  expostos ao risco convertem-se numa única observação  $d_i/E_i$  em cada idade

## 4.5. Kernel Smoothers

$x_i$ . Por outras palavras, as sequências  $\{d_i; E_i\}_{i=1}^n$  são condensadas em  $n$  observações  $\{d_i/E_i\}_{i=1}^n$  igualmente espaçadas, e o estimador (4.5.3) dá lugar ao denominado estimador de Nadaraya-Watson (NW)

$$\hat{q}_x^{NW} = \frac{\sum_{i=1}^n \left(\frac{d_i}{E_i}\right) K_b(x - x_i)}{\sum_{i=1}^n K_b(x - x_i)} = \frac{\sum_{i=1}^n \hat{q}_i K_b(x - x_i)}{\sum_{i=1}^n K_b(x - x_i)}. \quad (4.5.5)$$

Este estimador minimiza os desvios quadráticos médios locais

$$\frac{1}{n} \sum_{i=1}^n K_b(x - x_i) (\hat{q}_i - \hat{q}_i)^2. \quad (4.5.6)$$

Uma transformação simples de (4.5.5) permite entender o estimador *kernel* de Nadaraya-Watson como uma forma de graduação contínua pelo método de médias móveis ponderadas<sup>9</sup>

$$\hat{q}_x = \sum_{i=1}^n S_{xi} \hat{q}_i \quad (4.5.7)$$

onde

$$S_{xi} = \frac{K_b(x - x_i)}{\sum_{i=1}^n K_b(x - x_i)} \quad \text{com} \quad \sum_{i=1}^n S_{xi} = 1.$$

A utilização de *kernel smoothers* em problemas de graduação apresenta uma limitação importante, relacionada com a introdução de um enviesamento crescente da função nos extremos do intervalo de idades, um problema analisado em detalhe por Greville (1981a,b) e Hoem e Linnemann (1988) no âmbito do método MWA.

Bloomfield e Haberman (1987) e Gavin *et al.* (1993) observam que o estimador de Nadaraya-Watson tende a subestimar a mortalidade nas idades mais avançadas, onde a verdadeira curva é côncava, e a sobreestimar a mortalidade nas idades mais jovens, onde a verdadeira curva apresenta um padrão convexo. Conclusão semelhante é obtida por

<sup>9</sup>A conexão entre uma função kernel e o método de médias móveis ponderadas (em particular com a conhecida fórmula de Spencer) é explicada em detalhe por Gavin *et al.* (1993). A graduação mediante kernel smoothers apresenta algumas vantagens face ao método MWA. Em primeiro lugar, o método é aplicável mesmo nos casos em que os valores a graduar não estão igualmente espaçados. Em segundo, o método é mais flexível uma vez que permite a interpolação das taxas de mortalidade entre idades para as quais não se conhecem estimativas brutas. Por último, neste método o parâmetro de bandwith pode ser ajustado de forma contínua, ao passo que na graduação mediante MWA a vizinhança de referência apenas pode ser ajustada de forma discreta.

Copas e Haberman (1983) utilizando o seu respectivo estimador.<sup>10</sup>

Uma forma de atenuar este problema consiste simplesmente em restringir a vizinhança de interesse, diminuindo o parâmetro de bandwidth. Todavia, esta opção depara-se com a escassez de dados nos extremos do intervalo de idades, que tende a produzir flutuações erráticas em  $\hat{q}_x$ , um fenómeno que se reproduz naturalmente nos valores graduados e torna a curva pouco fiável neste segmento.

Uma solução alternativa consiste em truncar a curva de modo a excluir as idades com poucos óbitos registados. Esta técnica elimina as oscilações erráticas nos extremos da curva, mas permite que as observações registadas nestes segmentos contribuam para a graduação das taxas de mortalidade nas idades centrais. Uma outra opção passa por delimitar o intervalo de idades a graduar antes de aplicar o procedimento. Uma abordagem alternativa envolve o agregação dos dados nas idades em que o número de óbitos é diminuto. Os valores acumulados do número de expostos ao risco e de óbitos são atribuídos ao ponto médio do intervalo, excluindo as demais idades do processo de graduação.

Gavin *et al.* (1995) e Verrall (1996) propõem uma metodologia distinta, que passa pela transformação (logarítmica, logit, Gompertz,...) das estimativas brutas antes da graduação.<sup>11</sup> Esta transformação visa atenuar a concavidade/convexidade das estimativas brutas nos extremos do intervalo, reduzindo o enviesamento originado pela graduação. Esta solução levanta, no entanto, um problema nas idades em que não se registam óbitos (e.g., no caso da transformação logarítmica), um entrave que pode, todavia, ser contornado através do agrupamento de dados.

Copas e Haberman (1983) testam uma solução diferente, recorrendo a um procedimento de estimação repartido por duas fases. O método assenta na seguinte equação

$$\hat{q}_x = q_x^* + \frac{\sum_{i=1}^n (Z_i - q_{xi}) K_b(x - x_i)}{\sum_{i=1}^n K_b(x - x_i)}, \quad (4.5.8)$$

<sup>10</sup>Para uma comparação entre o enviesamento dos estimadores kernel de Nadaraya-Watson e de Copas-Haberman veja-se Gavin *et al.* (1994).

<sup>11</sup>Gavin *et al.* (1995) analisam um conjunto de transformações e advertem que a opção por uma delas é uma decisão de certo modo subjectiva, que depende da performance conseguida pelo modelo na amostra. Os autores utilizam finalmente a transformação logit

$$\hat{q}_i^t = \log \left( \frac{\hat{q}_i}{1 - \hat{q}_i} \right) \quad i = 1, \dots, n,$$

cuja inversa:

$$\hat{q}_i = \frac{\eta_i}{1 + \eta_i}, \quad \eta_i = \exp \left\{ \sum_{i=1}^n S_{xi} \hat{q}_i^t \right\}$$

garante que  $0 \leq \hat{q}_i \leq 1$ .

onde  $q_x^*$  denota uma estimativa prévia dos quocientes de mortalidade em cada idade  $x$  e  $Z_i$  é uma função indicatriz que assume o valor 1 em caso de óbito à idade  $x$  e o valor 0 caso contrário. O procedimento compreende as seguintes etapas: (i) cálculo das estimativas preliminares  $q_x^*$  usando, por exemplo, resultados de estudos anteriores ou ajustando uma função paramétrica, (ii) subtração de  $q_x^*$  às estimativas brutas iniciais, (iii) ajustamento da diferença ( $\hat{q}_x - q_x^*$ ) mediante *kernel smoothers*, e (iv) adição do valor obtido em (iii) a  $q_x^*$  de modo a obter a estimativa final  $\hat{q}_x$ .

Mais recentemente, Gavin *et al.* (1995) propõem um método baseado numa combinação linear de duas funções *kernel* distintas com valores diferentes para o parâmetro de bandwidth. Esta combinação visa atenuar o enviesamento da estimação pela eliminação do termo associado à inclinação da curva de mortalidade. O método permite ainda a derivação de funções *kernel* transformadas para os extremos direito e esquerdo do intervalo de idades, que atenuam o enviesamento da graduação nestas regiões.

### 4.5.3 Estimador *kernel* com bandwidth variável

Gavin *et al.* (1995) desenvolvem um estimador *kernel* variável (ou adaptativo), ao admitir que o parâmetro de bandwidth não é constante e oscila ao longo do intervalo de idades graduadas de acordo com o número de expostos ao risco  $E_x$  (dimensão da amostra).

Em consequência, nas regiões onde o número de expostos ao risco é elevado, i.e., onde a fiabilidade das estimativas brutas é maior, um valor mais baixo para  $b$  produz valores graduados que reflectem de forma mais aproximada os valores brutos. Pelo contrário, nas regiões onde o número de expostos ao risco é reduzido, como é por exemplo o caso das idades mais elevadas, um valor mais elevado para o parâmetro de bandwidth proporciona a obtenção de uma curva graduada mais suave, eliminando uma boa parte do comportamento errático das estimativas brutas nesta região.

Na prática, esta solução determina que nas idades mais avançadas, os valores graduados são calculados mediante médias ponderadas locais numa vizinhança que inclui um maior número de observações, o que reduz naturalmente a variância dos valores graduados mas introduz potencialmente maior enviesamento.

Os autores sugerem que a incorporação de  $E_x$  no modelo adaptativo pode ser feita de múltiplas formas. Uma solução passa por considerar um bandwidth diferente  $b_i$  para cada idade  $x_i$  em que a curva deve ser estimada, caso em que o modelo a estimar é definido por

$$\hat{q}_i = \sum_{j=1}^n S_{ij} \hat{q}_j \quad \text{onde} \quad S_{ij} = \frac{K_{b_i}(x_i - x_j)}{\sum_{j=1}^n K_{b_i}(x_i - x_j)}, \quad i = 1, \dots, n. \quad (4.5.9)$$

#### 4.5. Kernel Smoothers

---

Em alternativa, podemos considerar um bandwidth diferente  $b_j$  para cada idade observada  $x_j$  ( $j = 1, \dots, n$ ). Para cada idade observada, recorre-se ao bandwidth correspondente para medir a distância entre essa idade e a idade a ser graduada. O resultado é uma novo modelo, definido por

$$\hat{q}_i = \sum_{j=1}^n S_{ij} \hat{q}_j \quad \text{onde} \quad S_{ij} = \frac{K_{b_j}(x_i - x_j)}{\sum_{j=1}^n K_{b_j}(x_i - x_j)}, \quad i = 1, \dots, n. \quad (4.5.10)$$

O parâmetro de bandwidth local  $b_i$  relativo a cada idade é dado pelo produto do parâmetro de bandwidth global  $b$  por um factor de proporcionalidade local  $l_i^s$ , i.e.,  $b_i = bl_i^s$  ( $i = 1, \dots, n$ ). Uma forma simples de determinar este factor de proporcionalidade, considerando as diferenças na exposição ao risco entre as idades jovens e as idades avançadas, é dada por

$$l_i^s \propto (\hat{f}_i)^{-s}, \quad \hat{f}_i = \frac{E_i}{\sum_{j=1}^n E_j}, \quad i = 1, \dots, n, \quad (4.5.11)$$

onde  $s \in [0, 1]$  representa um parâmetro de sensibilidade. No caso extremo em que  $s = 0$ , os dois modelos acima referidos são reduzidos ao tradicional ajustamento com bandwidth fixo. No extremo oposto, a selecção de  $s = 1$  pode conduzir a oscilações significativas em  $b_i$  ao longo do intervalo de idades.<sup>12</sup>

Uma solução alternativa para determinar o factor de proporcionalidade local envolve a utilização do quociente entre os expostos ao risco, i.e.,

$$l_{ij}^s = \left( \frac{E_j}{E_i} \right)^s, \quad i, j = 1, \dots, n. \quad (4.5.12)$$

Neste caso, o peso a atribuir à  $j$ ésima probabilidade bruta no cálculo do valor graduado correspondente à idade  $i$  é ditado pela exposição relativa, i.e.,  $b_{ij} = bl_{ij}^s$  ( $i, j = 1, \dots, n$ ). Neste caso, o modelo é definido por

$$\hat{q}_i = \sum_{j=1}^n S_{ij} \hat{q}_j \quad \text{onde} \quad S_{ij} = \frac{K_{b_{ij}}(x_i - x_j)}{\sum_{j=1}^n K_{b_{ij}}(x_i - x_j)} \quad i = 1, \dots, n. \quad (4.5.13)$$

---

<sup>12</sup>A relação entre o factor de proporcionalidade, o número de expostos ao risco e o parâmetro de bandwidth é clara. Por exemplo, se o número de expostos ao risco numa determinada idade  $x_i$  for reduzido face ao total, o valor do factor  $l_i^s$  será elevado, o que aumenta o valor do parâmetro  $b_i$ . Em consequência, o procedimento de graduação diminui o peso atribuído à probabilidade bruta nessa idade, o que aumenta o alisamento da série na região em análise. Raciocínio análogo, com a devida inversão de argumentos, pode ser aduzido no caso em que o número de expostos ao risco é elevado.

## 4.6. Graduação Polinomial Local

---

Jones (1993) propõe um modelo alternativo, definido por

$$\hat{q}_i = \frac{\sum_{j=1}^n w_j \hat{q}_j K_b(x_i - x_j)}{\sum_{j=1}^n w_j K_b(x_i - x_j)}, \quad (4.5.14)$$

onde  $w_j$  são ponderadores dependentes do valor da exposição em cada idade.

A finalizar, refira-se que a implementação destes modelos exige a fixação prévia do valor dos parâmetros de sensibilidade,  $s$ , e de bandwith global,  $b$ . Embora seja possível recorrer a procedimentos automáticos de optimização, Gavin *et al.* (1995) recomendam, por razões de estabilidade das estimativas, a fixação subjectiva do parâmetro  $s$  numa primeira etapa, e o recurso a um procedimento de validação cruzada para determinar o valor de  $b$ .

## 4.6 Graduação Polinomial Local

Entre os inúmeros métodos não paramétricos com interesse em problemas de graduação, o método de regressão local ponderada (também conhecido por *Loess - Locally-weighted regression smoother*) merece especial destaque. O método, desenvolvido por Cleveland (1979) e analisado em detalhe por Fan e Gijbels (1996), baseia-se na aplicação da regressão linear local a uma vizinhança definida em torno de cada ponto da série.

Considere-se uma variável de resposta  $\mathbf{y}' = (y_{\min}, \dots, y_{\max})$  com distribuição Normal que depende de um previsor único  $\mathbf{x}' = (x_{\min}, \dots, x_{\max})$ . Para um dado ponto  $x_i$ , designemos por  $s(x_i)$  o estimador resultante da aplicação deste método, calculado da seguinte forma:

1. Para cada observação  $x_i$ , é identificado o conjunto de  $k$  pontos mais próximos de  $x_i$ , i.e., a sua vizinhança, designada por  $\mathcal{V}_x$ . Por exemplo, o valor de  $k$  pode ser definido escolhendo a percentagem de pontos a incluir em  $\mathcal{V}_x$  (o chamado *span*). Nesse caso, se admitirmos que  $\mathcal{V}_x$  deve conter uma proporção  $v$  das  $\mathcal{N}_x = x_{\max} - x_{\min} + 1$  observações, i.e.,

$$\mathcal{V}_x = \left\{ x_j \mid j = \max \left( x - \frac{v\mathcal{N}_x - 1}{2}, x_{\min} \right), \dots, \min \left( x + \frac{v\mathcal{N}_x - 1}{2}, x_{\max} \right) \right\}$$

para  $v\mathcal{N}_x$  ímpar. Neste caso, consideramos a observação  $x_i$  e  $(v\mathcal{N}_x - 1)/2$  dados em torno de um lado e outro desse valor.

#### 4.6. Graduação Polinomial Local

---

2. A cada ponto  $x_j$  na vizinhança  $\mathcal{V}_x$  de  $x_i$  é associado um ponderador  $w_j(i)$ , definido da seguinte forma:

$$w_j(i) = P\left(\frac{|x_i - x_j|}{\Delta(x_i)}\right),$$

onde

$$\Delta(x_i) = \max_{x_j \in \mathcal{V}_x} |x_i - x_j|$$

$$P(u) = \begin{cases} (1 - u^3)^3, & 0 \leq u \leq 1 \\ 0, & \text{outros casos.} \end{cases}$$

O termo  $\Delta(x_i)$  denota a distância máxima entre  $x_i$  e qualquer outro ponto na sua vizinhança e  $P(u)$  é uma função ponderadora (neste caso é usada a função tri-cúbica).

3. Os estimadores  $\hat{\theta}$  dos parâmetros do modelo obtém-se efectuando uma regressão de  $\mathbf{y}$  sobre  $\mathbf{x}$  na vizinhança  $\mathcal{V}_x$ , usando o método dos mínimos quadrados ponderados, i.e.,

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{j \in \mathcal{V}_x} w_j(i) (y_j - f(x_j, \theta))^2 \right\}, \quad (4.6.1)$$

onde  $f(x_i, \theta)$  representa um polinómio de ordem  $p$  que determina a natureza (linear, quadrática, cúbica,...) da regressão local

$$f(x_i, \theta) = a + \sum_{j=0}^p \beta_j (x_i - x_0)^j + \varepsilon_i, \quad x_i \in \mathcal{V}_x \quad (4.6.2)$$

O valor graduado em cada ponto focal  $x_i$  é dado simplesmente por  $\hat{f}(x_i, \theta) | x_i = a$ . Em cada ponto focal, os pesos alteram-se assim como as estimativas dos parâmetros da função. Em termos matriciais, seja

$$\mathbf{X}_{\mathcal{V}_x} = \begin{pmatrix} 1 & (x_1 - x_i) & \dots & (x_1 - x_i)^p \\ \vdots & & & \vdots \\ 1 & (x_{\mathcal{N}} - x_i) & \dots & (x_{\mathcal{N}} - x_i)^p \end{pmatrix}, \quad \mathbf{y}_{\mathcal{V}_x} = \begin{pmatrix} y_1 \\ \vdots \\ y_{\mathcal{N}} \end{pmatrix}, \quad \boldsymbol{\beta}_{\mathcal{V}_x} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix}$$

e  $\mathbf{W}_{\mathcal{V}_x}$  a matriz diagonal dos ponderadores.

#### 4.6. Graduação Polinomial Local

---

O problema de mínimos quadrados ponderados (4.6.1) é definido por

$$\arg \min_{\beta} (\mathbf{y}_{\mathcal{V}_x} - \mathbf{X}_{\mathcal{V}_x} \beta_{\mathcal{V}_x})' \mathbf{W}_{\mathcal{V}_x} (\mathbf{y}_{\mathcal{V}_x} - \mathbf{X}_{\mathcal{V}_x} \beta_{\mathcal{V}_x}), \quad (4.6.3)$$

donde se deriva o seguinte vector de estimadores

$$\hat{\beta}_{\mathcal{V}_x} = \left( \mathbf{X}'_{\mathcal{V}_x} \mathbf{W}_{\mathcal{V}_x} \mathbf{X}_{\mathcal{V}_x} \right)^{-1} \mathbf{X}'_{\mathcal{V}_x} \mathbf{W}_{\mathcal{V}_x} \mathbf{y}_{\mathcal{V}_x}, \quad (4.6.4)$$

admitindo que  $\left( \mathbf{X}'_{\mathcal{V}_x} \mathbf{W}_{\mathcal{V}_x} \mathbf{X}_{\mathcal{V}_x} \right)$  é invertível.

O grau de alisamento da série é determinado pelo parâmetro de span  $\nu$  que representa, como vimos, uma proporção do número total de observações compreendida na vizinhança de cada ponto, i.e.,  $\nu = \#\mathcal{V}_x / \mathcal{N}_x$ . No caso em que  $\nu = 2$ , cada intervalo  $\mathcal{V}_x$  contém a totalidade dos dados, independentemente de  $x_i$ . Neste caso, a graduação é conduzida por intermédio do tradicional método dos mínimos quadrados ponderados. No caso extremo em que  $\nu = 1/\mathcal{N}_x$ , cada vizinhança é reduzida a um único ponto  $x_i$  e a graduação é feita por simples interpolação, i.e.,  $\hat{f}(x_i, \theta) = y_i$ . Daqui resulta que o valor de  $\nu$  deve ser escolhido entre  $1/\mathcal{N}_x$  e 2 de modo a encontrar um compromisso entre o envasamento e a variância do estimador.

Por definição, o método determina que a importância relativa de cada ponto na estimação é dada pela distância face ao ponto  $x_i$  correspondente. Se exceptuarmos o caso em que os previsores estão igualmente espaçados, o método *Loess* é preferível aos métodos baseados em vizinhanças simétricas (como é o caso das médias móveis). Tal é assim porquanto numa vizinhança com um número fixo de pontos, a distância destes face ao valor objectivo é, em termos relativos, menor. Refira-se ainda que o método *Loess* pode ser expandido de modo a incorporar múltiplos regressores.

Este método possui algumas similitudes com os modelos lineares generalizados dinâmicos sugeridos por Verrall (1993b) e com a graduação pelo método de Whittaker-Henderson. Estes métodos podem, segundo Verrall (1994), ser considerados modelos baseados numa recta cujo gradiente varia suavemente com a idade.

A graduação mediante modelos lineares generalizados dinâmicos proposta por Verrall (1993b) não é mais do que uma extensão da teoria dos GLM's a modelos onde se admite a possibilidade dos coeficientes não serem constantes ao longo de todo o intervalo de idades. Esta solução gera uma regressão que é localmente linear, mas não o é em termos globais. Por esta razão, o método pode ser considerado uma combinação de métodos paramétricos, uma vez que considera uma função analítica dependente de parâmetros, e

de métodos não paramétricos, na medida em que utiliza um procedimento de alisamento, nomeadamente o filtro de Kalman.

## 4.7 Modelos Aditivos Generalizados

A maioria dos métodos não paramétricos acima descritos pode ser reformulada no contexto dos denominados Modelos Aditivos Generalizados (*Generalized Additive Models* (GAM)). Os modelos GAM expandem a teoria subjacente aos tradicionais modelos de regressão linear múltipla ao admitir que algumas (ou a totalidade) das funções lineares que associam os regressores à variável independente podem ser substituídas por funções arbitrárias suaves (paramétricas ou não) da variável explicativa.<sup>13</sup>

Dada a sequência de observações,  $(x_i, y_i)$ ,  $i = i_{\min}, \dots, i_{\max}$ , a influência de  $x_i$  sobre a variável de resposta  $y_i$  é modelada com o auxílio de um modelo aditivo

$$y_i = \alpha + \sum_{j=1}^p f_j(x_j) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}or(0, \sigma^2), \quad (4.7.1)$$

onde  $f_j(\cdot)$  são funções de alisamento arbitrárias (splines cúbicos, B-Splines, regressão local, polinómios, etc.), uma para cada variável explicativa, e  $\alpha$  é uma constante.<sup>14</sup> Assume-se ainda que os erros  $\varepsilon_i$  são independentes dos regressores  $x_j$ , e que no caso gaussiano  $\varepsilon_i \sim \mathcal{N}or(0, \sigma^2)$ . Implícita em (4.7.1) está ainda a hipótese  $\mathbb{E}\{f_j(x_j)\} = 0$ , caso contrário o modelo admitiria constantes livres nas funções  $f_j(\cdot)$ . Das condições de identificação do modelo resulta que  $\mathbb{E}(Y) = \alpha$ .

No caso particular em que as funções de alisamento são lineares, i.e., em que  $\hat{f}(x_i)$  pode ser expresso em termos de uma combinação linear dos valores  $y_{i_{\min}}, \dots, y_{i_{\max}}$ , temos

$$\hat{f}(x_i) = \sum_{j=i_{\min}}^{i_{\max}} h_{ij} y_j, \quad (4.7.2)$$

onde os ponderadores  $h_{ij} = (x_i, x_j)$  dependem das observações com base nas quais a

<sup>13</sup>Para uma descrição pormenorizada dos modelos GAM veja-se, por exemplo, Hastie e Tibshirani (1990).

<sup>14</sup>As funções  $f_j(\cdot)$  podem incluir formulações paramétricas, o que significa que incorporam, por exemplo, o tradicional modelo linear generalizado (GLM) analisado no capítulo 3. Os modelos GAM utilizam, na prática, a mesma estrutura que os GLM, com a diferença de que os previsores lineares podem incluir funções não paramétricas.

## 4.8. Selecção do parâmetro de alisamento e número de graus de liberdade

---

variável de resposta  $f(x_i)$  deve ser estimada. Definindo os vectores

$$\begin{aligned}\hat{\mathbf{f}} &= \left( \hat{f}(x_{i_{\min}}), \dots, \hat{f}(x_{i_{\max}}) \right)' \\ \mathbf{y} &= \left( y_{i_{\min}}, \dots, y_{i_{\max}} \right)'\end{aligned}$$

e a matriz de ponderadores  $h_{ij}$ ,  $\mathbf{H}$ , podemos reescrever a relação (4.7.2) como

$$\hat{\mathbf{f}} = \mathbf{H} \mathbf{y}. \quad (4.7.3)$$

O vector dos resíduos é dado por

$$\hat{\boldsymbol{\varepsilon}} = (\mathbf{I} - \mathbf{H}) \mathbf{y}. \quad (4.7.4)$$

As funções estimadas nos modelos GAM têm uma interpretação análoga aos coeficientes estimados no modelo linear clássico. Para além disso, os modelos GAM retêm uma importante propriedade interpretativa do modelo clássico de regressão linear. Com efeito, a sua estrutura aditiva permite, na ausência de interacções, analisar o efeito isolado de cada variável explicativa sobre a variável dependente, ou seja, reduz o problema a uma série de regressões lineares parciais.

No contexto actuarial, os modelos GAM permitem, à semelhança do que enunciámos para o caso dos GLMs, a utilização da distribuição exacta da variável de interesse ( $q_x$  ou  $\mu_x$ ), e não o recurso a aproximações, como no caso do método de Whittaker-Henderson (Verrall, 1996). Dito de outro modo, para graduar  $q_x$  continua a ser possível usar a distribuição Binomial e o *link* logit, enquanto que para graduar  $\mu_x$  recorreremos à distribuição de Poisson e ao *link* logarítmico, mudando apenas a forma adoptada pelo predictor linear.

## 4.8 Selecção do parâmetro de alisamento e número de graus de liberdade

### 4.8.1 Escolha do parâmetro de alisamento

A técnica de splines e os métodos de Whittaker-Henderson, de *kernel* e de regressão polinomial local dependem de um parâmetro que controla a regularidade das estimativas obtidas. A selecção do parâmetro de alisamento (também chamado de parâmetro de *bandwidth*, *smoothing* ou de *span*) é crucial, porquanto ele estipula a importância relativa dos objectivos qualidade do ajustamento e alisamento da série em problemas de gra-

#### 4.8. Selecção do parâmetro de alisamento e número de graus de liberdade

duação.<sup>15</sup> Pode ser feita de maneira subjectiva ou, como descrevemos nesta secção, por intermédio de um procedimento automático.

Gavin *et al.* (1994) e Verrall (1996) analisam a aplicação do critério de validação cruzada (*cross validation*) na escolha do parâmetro de alisamento em problemas de graduação. Dada a sequência de  $(x_i, y_i)$  de idades  $x_i$  ( $i = 1, \dots, \mathcal{N}_x$ ) e indicadores de mortalidade  $y_i$ , o critério de validação cruzada selecciona o parâmetro de bandwidth  $b$  minimizando a seguinte soma de quadrados

$$\hat{b} = \arg \min_{b_j} \left\{ CV(b_j) = \frac{1}{\mathcal{N}_x} \sum_{i=x_{\min}}^{x_{\max}} \left( y_i - \hat{f}_{b_j}^{-i}(x_i) \right)^2 \right\}, \quad j = 1, \dots, k, \quad (4.8.1)$$

onde, para um dado valor do parâmetro de bandwidth,  $\hat{f}_b^{-i}(x_i)$  representa a estimativa de  $f(x_i)$  calculada usando o subconjunto  $\{(x_j, y_j), j \neq i\}$ , de dimensão  $\mathcal{N}_x - 1$ . O critério de validação cruzada privilegia o poder explicativo do modelo em análise. De facto,  $\hat{f}_b^{-i}(x_i)$  corresponde à estimativa de  $y_i$  gerada por todas as observações excluindo a  $i$ ésima, donde o desvio  $y_i - \hat{f}_b^{-i}(x_i)$  é um indicador da qualidade de ajustamento do modelo.

A aplicação do critério de validação cruzada simplifica-se enormemente no caso de métodos de alisamento lineares (4.7.2), onde uma aproximação simples de  $\hat{f}_b^{-i}(x_i)$  é dada por

$$\hat{f}_b^{-i}(x_i) = \sum_{i=x_{\min}, j \neq i}^{x_{\max}} \frac{h_{ij}}{1 - h_{ii}} y_j,$$

onde  $\sum_{j \neq i} \frac{h_{ij}}{1 - h_{ii}} = 1$ . Com base nesta equação deduz-se, após alguma manipulação algébrica, que

$$\hat{f}_b^{-i}(x_i) = \frac{h_{ij}}{1 - h_{ii}} \hat{f}_b(x_i) - \frac{h_{ii}}{1 - h_{ii}} y_j,$$

donde

$$CV(b) = \frac{1}{\mathcal{N}_x} \sum_{i=x_{\min}}^{x_{\max}} \left( \frac{y_i - \hat{f}_b(x_i)}{1 - h_{ii}} \right)^2. \quad (4.8.2)$$

Da equação (4.8.2) resulta que é possível estimar  $\hat{f}_b^{-i}(x_i)$  sem necessidade de eliminar a observação relativa a  $x_i$ .

Em alternativa ao critério (4.8.1), é possível recorrer ao critério de validação cruzada

<sup>15</sup>Para uma análise detalhada deste tema veja-se, por exemplo, Hastie e Tibshirani (1990).

generalizado, definido por (Hastie e Tibshirani, 1990)

$$GCV(b) = \frac{1}{\mathcal{N}_x \left(1 - \frac{1}{\mathcal{N}_x} \sum_{i=x_{\min}}^{x_{\max}} h_{ii}\right)^2} \sum_{i=x_{\min}}^{x_{\max}} \left(y_i - \hat{f}_b(x_i)\right)^2. \quad (4.8.3)$$

### 4.8.2 Número de graus de liberdade

A informação sobre o número de graus de liberdade dos diferentes métodos de alisamento é fundamental para comparar a sua complexidade e capacidade de alisamento. Para uma dada matriz de alisamento  $\mathbf{H}_b$ , os graus de liberdade  $df(b)$  são definidos por

$$df(b) = tr(\mathbf{H}_b),$$

onde  $tr(\cdot)$  denota o traço da matriz.

Hastie e Tibshirani (1990) apresentam definições alternativas para os graus de liberdade, desenvolvidas por analogia com os modelos de regressão linear. Seja  $y_i = \hat{f}(x_i) + \varepsilon_i$ ,  $\varepsilon_i \sim \mathcal{Nor}(0, \sigma^2)$ . A partir da definição de valor esperado da soma dos quadrados residuais, os autores sugerem a seguinte definição para os graus de liberdade do erro

$$df^*(b) = \mathcal{N}_x - tr(2\mathbf{H}_b - \mathbf{H}_b\mathbf{H}_b'),$$

onde a analogia com o modelo de regressão linear clássico é clara, para o qual, recordamos,  $df = (\mathcal{N}_x - p)$ , com  $p$  a denotar o número de parâmetros do modelo.

## Capítulo 5

# Graduação Paramétrica da Mortalidade da População Segura em Portugal

*“Não basta conquistar a sabedoria, é preciso usá-la.” (Cícero)*

### 5.1 Introdução

As tábuas de mortalidade constituem um elemento essencial na fixação das tarifas dos contratos de seguro do ramo vida e na avaliação das responsabilidades das companhias de seguros e dos fundos de pensões. Nos últimos anos, a problemática dos métodos e pressupostos usados na avaliação actuarial dos activos e das responsabilidades ganhou uma importância acrescida, não apenas para as empresas, mas também para as entidades reguladoras do sector a nível nacional e internacional. Com efeito, num contexto marcado pelo incremento da volatilidade nos mercados financeiros e pelo aumento da longevidade humana, a necessidade de garantir o *matching* entre os activos e as responsabilidades financeiras das empresas passa, em boa medida, pela utilização de bases técnicas que espelhem, de forma adequada, os factores de risco (demográficos, financeiros, operacionais) a que se encontram expostas as carteiras.

Não obstante as melhorias verificadas nos sistemas de informação das empresas, a recolha sistemática pelo Instituto de Seguros de Portugal (ISP) de dados estatísticos sobre a mortalidade das pessoas seguras e dos beneficiários dos fundos de pensões teve início apenas em 1995. Como tal, não existe um referencial histórico que sirva de base a uma análise comparativa sobre o comportamento da mortalidade neste tipo de populações.

## 5.1. Introdução

---

A dimensão curta da série dificulta igualmente a identificação de uma tendência clara para a dinâmica da mortalidade, uma limitação importante em estudos prospectivos.

Se exceptuarmos os ensaios exploratórios publicados nos Relatórios do Sector Segurador e Fundos de Pensões de 1999 e 2002 (ISP, 2000, 2003) e as referências neles citadas, nunca foi realizado um estudo empírico sobre a mortalidade destas populações. Em Portugal, não existem tábuas de mortalidade (contemporâneas ou prospectivas) regulamentares, nem para a população portuguesa, nem para as subpopulações de pessoas seguras. Em resultado, as companhias de seguros que operam o ramo vida são forçadas a recorrerem, nos seus pressupostos actuariais, a tábuas de mortalidade adoptadas noutros países (França, Suíça, Inglaterra, Espanha,...). A utilização de uma lei de sobrevivência relativa a outro país, tendencialmente desajustada das condições demográficas do país e da carteira em que é aplicada, envolve, naturalmente, riscos significativos, nomeadamente o risco de sobrestimação da taxa de mortalidade da população.

As novas disposições contidas nas normas contabilísticas IFRS e no projecto Solvência II reforçam a necessidade de se proceder a uma avaliação mais criteriosa das responsabilidades das empresas. Em particular, elas abrem a porta à possibilidade de serem adoptadas tábuas de mortalidade específicas, ajustadas à população representada nas carteiras das companhias de seguros, designadamente nos casos em que o actuário responsável considera que as características das populações envolvidas não se encontram devidamente espelhadas nas tábuas regulamentares. Estes desenvolvimentos tornam inadiável a construção de tábuas regulamentares em Portugal, quer para a população em geral, quer para as subpopulações de pessoas seguras e dos fundos de pensões.

Para tentar colmatar esta lacuna, realizamos neste capítulo o primeiro exercício sistemático de graduação paramétrica da mortalidade das pessoas seguras e dos beneficiários dos fundos de pensões efectuado em Portugal, tomando como referência os dados relativos ao período 2000-2004. O capítulo está organizado da seguinte forma. Na Secção 5.2, descrevemos as principais características das populações abrangidas pelo estudo. Na Secção 5.3, apresentamos os resultados obtidos pela aplicação de modelos lineares e não lineares generalizados na graduação paramétrica de quocientes e taxas de mortalidade. Na Secção 5.4, exemplificamos o emprego da abordagem dual descrita no Capítulo 3 em problemas de graduação, considerando a população coberta por seguros em caso de vida. Na Secção 5.5, avaliamos o desempenho das diferentes leis de Heligman-Pollard (HP) e aferimos a importância da escolha da função objectivo na eficácia do modelo. Na Secção 5.6, comparamos o desempenho dos diferentes métodos paramétricos testados usando um conjunto de indicadores de qualidade do ajustamento. Por último, a Secção 5.7 resume as principais conclusões do estudo.

## 5.2 Descrição da amostra

Nesta secção descrevemos, de forma sumária, as principais características das populações que servem de base ao estudo. Em particular, analisamos a estrutura etária da população exposta ao risco e algumas particularidades observadas no comportamento da mortalidade no período investigado.

### 5.2.1 Beneficiários de fundos de pensões

O presente estudo sobre a mortalidade dos Beneficiários de Fundos de Pensões (BFP) tem por base os dados coligidos pelo Instituto de Seguros de Portugal (ISP) para o universo dos Fundos de Pensões a operar no mercado português. A categoria BFP compreende os pensionistas com direito ao recebimento de uma pensão a título de pré-reforma, reforma antecipada, reforma por velhice, reforma por invalidez ou em caso de sobrevivência (e.g., viuvez e orfandade). A informação estatística usada na análise compreende os dados sobre o número de pessoas expostas (centralmente) ao risco ( $E$ ) e sobre o número de óbitos registados ( $d$ ), devidamente discriminados por sexo e idade actuarial, no período compreendido entre os anos de 2000 e 2004.

O Quadro 5.2.1 resume algumas características das populações. No período em estudo, a população anualmente exposta ao risco de morte aumentou 8% de 95,4 mil para cerca de 103 mil indivíduos, sobretudo devido ao crescimento dos pensionistas do sexo feminino que aumentaram 16%, atingindo um total de 38770 vidas em 2004. Em comparação, no mesmo período a população do sexo masculino apenas cresceu 4%, somando 64172 vidas em 2004. Após um período de crescimento acentuado do número de pensionistas, observado na segunda metade da década de noventa (o total de vidas expostas ao risco ascendia a 63,4 mil em 1995) e no início do novo milénio, nos últimos anos o número de pensionistas manteve-se relativamente constante.

| Ano   | Homens  |       | Mulheres |       | Total   |        |
|-------|---------|-------|----------|-------|---------|--------|
|       | $E$     | $d$   | $E$      | $d$   | $E$     | $d$    |
| 2000  | 61 918  | 1 753 | 33 486   | 876   | 95 404  | 2 629  |
| 2001  | 60 027  | 1 624 | 30 614   | 731   | 90 641  | 2 355  |
| 2002  | 66 278  | 2 107 | 36 024   | 1 830 | 102 302 | 3 937  |
| 2003  | 65 420  | 1 914 | 35 969   | 1 059 | 101 389 | 2 973  |
| 2004  | 64 172  | 1 961 | 38 770   | 1 075 | 102 941 | 3 036  |
| Total | 317 812 | 9 356 | 174 848  | 5 562 | 492 659 | 14 918 |

Quadro 5.2.1: Beneficiários de fundos de pensões ( $E_x; d_x$ )

As Figuras 5.2.1 e 5.2.2 revelam algumas particularidades sobre a distribuição etária

## 5.2. Descrição da amostra

---

da população masculina e feminina exposta ao risco.

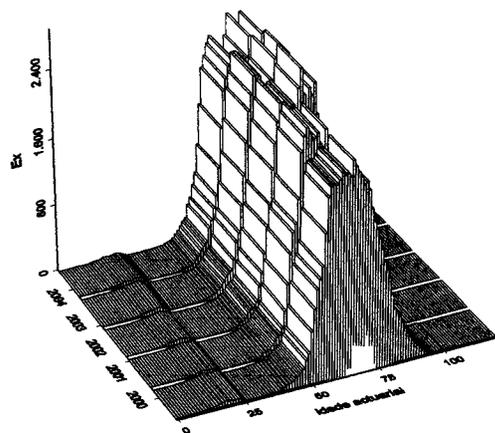


Figura 5.2.1: Estrutura etária da população masculina de pensionistas

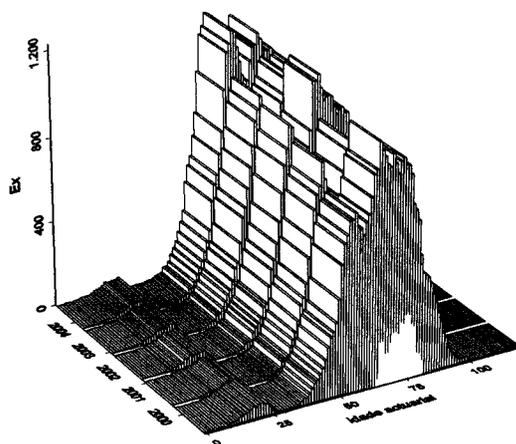


Figura 5.2.2: Estrutura etária da população feminina de pensionistas

## 5.2. Descrição da amostra

---

Em primeiro lugar, e atendendo a natureza do benefício em causa, observa-se que a população está concentrada em duas faixas etárias distintas. Uma primeira, de menor importância em termos de vidas expostas ao risco, que abrange as idades entre os 15 e os 25 anos e que representa, essencialmente, os órfãos de beneficiários e participantes nos fundos.<sup>1</sup> Uma segunda, de maior significado e interesse, correspondente aos indivíduos com mais de 40 anos, com um pico natural em torno dos 65 anos, a idade “normal” de reforma em Portugal. Em ambos os sexos, são observadas vidas expostas ao risco entre os 0 e os 104 anos. No entanto, cerca de 95% da exposição está concentrada numa faixa etária que vai dos 50 aos 85 anos.

A Figura 5.2.3 dá-nos uma primeira impressão sobre o comportamento da mortalidade no período em análise. As estimativas brutas dos quocientes  $q_x$  foram calculadas com base na hipótese de repartição uniforme dos óbitos durante o ano, mediante a fórmula

$$\hat{q}_x = \frac{\sum_{t=2000}^{2004} d_{x,t}}{\sum_{t=2000}^{2004} (E_{x,t} + \frac{1}{2}d_{x,t})}. \quad (5.2.1)$$

A Figura 5.2.4 representa, a título de comparação, os quocientes de mortalidade da população portuguesa extraídos da tábua de mortalidade referente ao período 2001/02, elaborada pelo Instituto Nacional de Estatística (I.N.E.).

A primeira característica a salientar refere-se à mortalidade anormalmente elevada nas idades mais jovens, em especial no intervalo entre os 15 e os 25 anos. Este comportamento, devidamente assinalado pelo ISP (2000), é explicado em parte pela forma como os dados são compilados em Portugal, que equiparam à morte enquanto pensionista, todos os eventos que conduzem um indivíduo a perder a sua qualidade de beneficiário do fundo de pensões. Significa isto que a mortalidade registada tanto pode corresponder à morte física, propriamente dita, como a outros factos que impliquem a perda do direito à pensão, entre os quais se incluem, por exemplo, a extinção da pensão de orfandade por se atingir a idade máxima elegível, ou a perda do direito à pensão de viuvez em resultado da contracção de novo casamento pelo cônjuge sobrevivente. Deste modo, os picos de mortalidade observados em ambos os sexos em torno dos 25 anos de idade não podem ser dissociados do facto de que os planos de pensões que prevêem benefícios de orfandade e/ou viuvez estipularem, regra geral, o vigésimo quinto aniversário como a idade limite para cessação do direito à pensão de orfandade.<sup>2</sup> Em suma, uma boa fa-

---

<sup>1</sup> Assinala-se, a este propósito, que uma boa parte dos planos de pensões existentes em Portugal estipula benefícios de orfandade em caso de falecimento dos beneficiários e/ou participantes nos fundos.

<sup>2</sup> Registe-se que esta idade pode, nalguns casos, ser prolongável, em particular se o órfão for inválido ou se estiver em fase de conclusão dos estudos com bom aproveitamento.

## 5.2. Descrição da amostra

---

tia da “mortalidade” registada nestas idades não corresponde, de facto, à mortalidade física, mas simplesmente à perda do direito à pensão e, como tal, deve ser analisada com cautela.

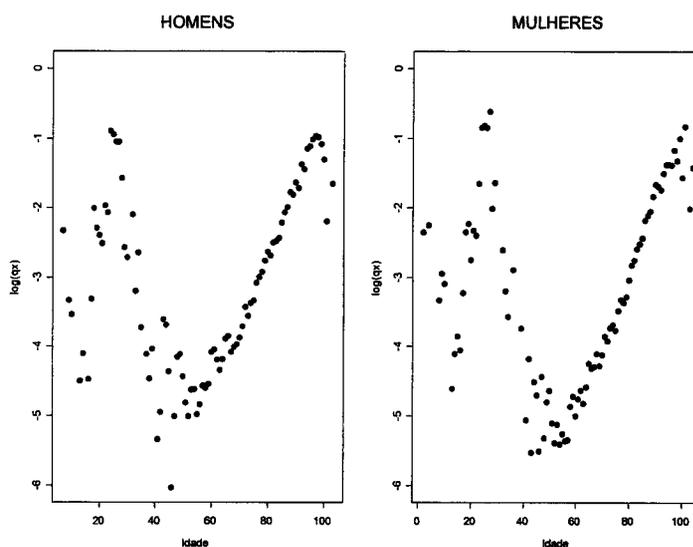


Figura 5.2.3: Estimativas brutas dos quocientes  $q_x$ , pensionistas

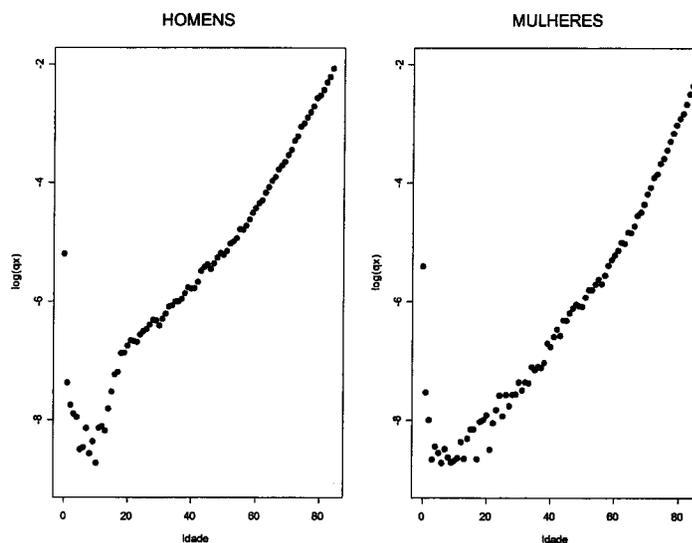


Figura 5.2.4: Estimativas brutas dos quocientes  $q_x$ , população portuguesa 2001/2002

## 5.2. Descrição da amostra

---

A observação da Figura 5.2.3 permite ainda constatar que nas idades em que o número de expostos ao risco é (estatisticamente) pouco significativo (até aos 50 anos), os quocientes  $\hat{q}_x$  apresentam um padrão gráfico bastante errático e volátil. Em compensação, nas faixas etárias em que se concentra a maior parte da exposição ao risco é possível registar, em ambos os sexos, uma tendência crescente nos quocientes, muito próxima daquela que é observada na população portuguesa. Nos indivíduos com mais de 50 anos, destacamos ainda o facto de os quocientes relativos aos pensionistas do sexo feminino serem menores do que os observados nos pensionistas do sexo masculino, um resultado esperado atendendo ao comportamento observado na Figura 5.2.4 para a população portuguesa, e ao diferencial na esperança de vida entre homens e mulheres registado na generalidade dos países ocidentais.

### 5.2.2 Seguros em caso de vida

Em Portugal, a população exposta ao risco nos Seguros de vida em Caso de Vida (SCV) é constituída, essencialmente, por seguros de capitais (capitais diferidos, planos de poupança-reforma (PPR) e planos de poupança-reforma-educação (PPR/E), seguros do tipo “*Universal Life*”, vida inteira e mistos), operações de capitalização e seguros de rendas.

Como se observa no Quadro 5.2.2, a população coberta pelos seguros em caso de vida cresceu significativamente (82.8%) no período em análise, passando das 1198,7 mil vidas em 2000 para os cerca de 2,2 milhões de indivíduos em 2004. Esse crescimento foi, uma vez mais, desigual entre sexos, com a população feminina a quase duplicar no quinquénio em estudo (cresceu 98%), enquanto que na população do sexo masculino o aumento foi de apenas 75%. Neste período, a repartição da população por sexos registou flutuações importantes. Ainda assim, é patente o maior peso das vidas do sexo masculino, que representavam em 2004 62% do total. Este último número contrasta com detectado na população portuguesa, onde a população feminina representa cerca de 51.5% do total.

| Ano   | Homens    |        | Mulheres  |       | Total     |        |
|-------|-----------|--------|-----------|-------|-----------|--------|
|       | $E$       | $d$    | $E$       | $d$   | $E$       | $d$    |
| 2000  | 782 830   | 2 652  | 415 883   | 1 401 | 1 198 713 | 4 053  |
| 2001  | 1 011 050 | 3 280  | 501 776   | 974   | 1 512 826 | 4 254  |
| 2002  | 1 112 976 | 3 544  | 575 253   | 1 110 | 1 688 229 | 4 654  |
| 2003  | 959 341   | 2 928  | 545 358   | 1 034 | 1 504 699 | 3 962  |
| 2004  | 1 368 112 | 5 338  | 822 904   | 1 812 | 2 191 016 | 7 150  |
| Total | 5 234 308 | 17 742 | 2 861 174 | 6 331 | 8 095 483 | 24 073 |

Quadro 5.2.2: Seguros em caso de vida ( $E_x; d_x$ )

## 5.2. Descrição da amostra

---

Nas Figuras 5.2.5 e 5.2.6 é possível verificar que, apesar do forte crescimento da carteira registado neste período, a distribuição etária das populações (masculina e feminina) apresenta um perfil relativamente estável.

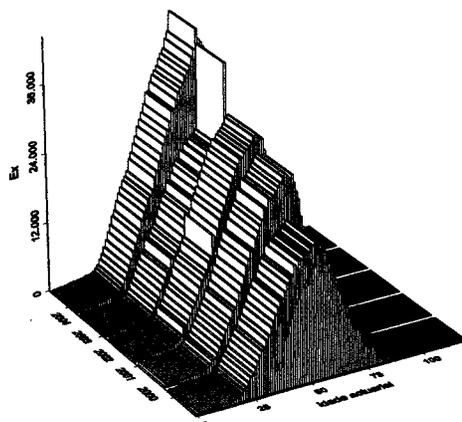


Figura 5.2.5: Seguros em caso de vida: estrutura etária da população masculina

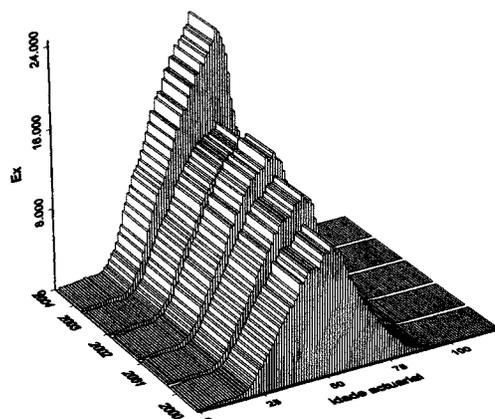


Figura 5.2.6: Seguros em caso de vida: estrutura etária da população feminina

## 5.2. Descrição da amostra

Com efeito, a maioria das vidas concentra-se numa faixa etária entre os 45 e os 65 anos de idade, ou seja, relativamente próximo da idade normal de reforma.<sup>3</sup> Em ambos os sexos, foram registadas vidas expostas ao risco entre os 0 e os 108 anos, sendo que cerca de 95% da exposição se concentrava numa faixa etária entre os 25 e os 75 anos.

No que se refere ao comportamento da mortalidade, a Figura 5.2.7 mostra que a população segura em caso de vida apresenta perfis semelhantes aos registados na população portuguesa em geral, embora se constate que os quocientes  $q_x$  são claramente inferiores. Destaque ainda para as já mencionadas diferenças de mortalidade entre sexos, assim como para os habituais picos de mortalidade no intervalo 15-30 anos (sobretudo na população do sexo masculino), atribuídos à denominada “mortalidade accidental”.

Nas idades em que o número de expostos ao risco é (estatisticamente) pouco significativo, a irregularidade nos quocientes é naturalmente mais acentuada.

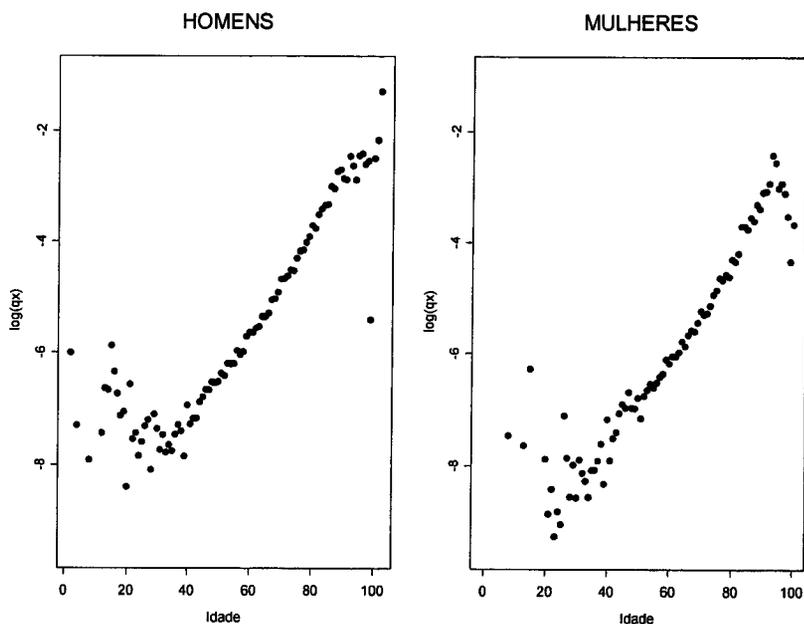


Figura 5.2.7: Estimativas brutas dos quocientes  $q_x$ , seguros em caso de vida

<sup>3</sup>Este perfil, revelador de um mercado próximo da maturidade, é em boa medida explicado pelo facto de no período 2000-2004 se ter assistido a uma estabilização das características da procura (por principais tipos de produtos), com a quase totalidade da produção anual concentrada em produtos de elevada componente financeira e idade de subscrição relativamente próxima da idade de reforma (capitais diferidos com contraseguro, PPR/E).

### 5.2.3 Seguros em caso de morte

A população exposta ao risco nos Seguros de vida em Caso de Morte (SCM) é constituída, maioritariamente, por seguros temporários, um tipo de seguro usualmente exigido como garantia em operações de crédito, sobretudo em operações de crédito à habitação, operações de crédito pessoal ou associadas a garantias de cartões de crédito. Para a maioria da população segura, o fundamento que conduz à celebração do contrato de seguro é, assim, completamente independente da preocupação para com um falecimento prematuro face à média, um comportamento que pode introduzir alguma selecção adversa na composição das carteiras das companhias de seguros.

A evolução da carteira de vidas cobertas por seguros em caso de morte no período 2000-2004 é posta em evidência no Quadro 5.2.3. Como se observa, a dimensão da carteira cresceu significativamente (cerca de 1.39 milhões de vidas ou 43.3%), atingindo em 2004 um universo de 4.6 milhões de indivíduos de ambos os sexos. Este crescimento é consistente com a evolução dos contratos e dos montantes de crédito à habitação observada em Portugal neste período. Sublinha-se, uma vez mais, o maior contributo das vidas do sexo feminino para o crescimento da carteira global, com o número de expostos ao risco deste sexo a quase duplicar no período em apreço. Apesar desta evolução, em 2004 a população feminina continuava a representar apenas 35% do total da carteira.

| Ano   | Homens     |        | Mulheres  |       | Total      |        |
|-------|------------|--------|-----------|-------|------------|--------|
|       | $E$        | $d$    | $E$       | $d$   | $E$        | $d$    |
| 2000  | 2 376 533  | 4 082  | 839 862   | 1 338 | 3 216 395  | 5 420  |
| 2001  | 2 469 984  | 3 211  | 1 003 407 | 767   | 3 473 391  | 3 978  |
| 2002  | 2 597 102  | 3 263  | 1 130 075 | 715   | 3 727 177  | 3 978  |
| 2003  | 2 266 837  | 3 699  | 1 186 668 | 795   | 3 453 506  | 4 494  |
| 2004  | 2 991 054  | 5 207  | 1 616 957 | 1 330 | 4 608 011  | 6 537  |
| Total | 12 701 511 | 19 462 | 5 776 969 | 4 945 | 18 478 480 | 24 407 |

Quadro 5.2.3: Seguros em caso de morte ( $E_x; d_x$ )

Como se observa pela análise das Figuras 5.2.8 e 5.2.9, os diferentes ritmos de crescimento nas carteiras do sexo masculino e feminino tiveram correspondência na evolução da sua distribuição etária. Com efeito, a carteira de vidas do sexo masculino apresenta perfis típicos de populações mais maduras, evidenciando um comportamento relativamente estável ao longo de todo o quinquénio em estudo, com uma forma senoidal ligeiramente enviesada à esquerda, com média, moda e mediana situadas numa faixa etária em torno dos 40 anos. Esta evolução contrasta com o observado nas vidas do sexo feminino, onde

## 5.2. Descrição da amostra

---

o crescimento da carteira introduziu mudanças estruturais substanciais.<sup>4</sup>

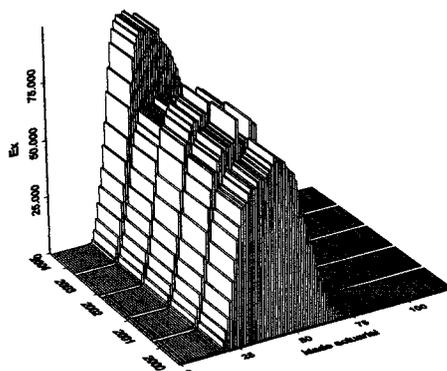


Figura 5.2.8: Seguros em caso de morte: estrutura etária da população masculina

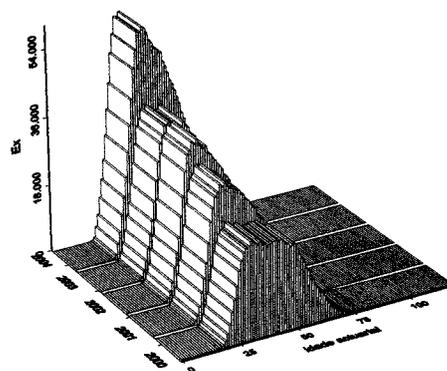


Figura 5.2.9: Seguros em caso de morte: estrutura etária da população feminina

<sup>4</sup>Como se observa na Figura 5.2.9, no transcurso deste período a estrutura etária tendeu para uma forma senoidal bastante enviesada à esquerda, com moda situada em torno dos 30 anos, recuada em aproximadamente 10 anos face à média e à mediana da distribuição. Perante esta evolução, é inquestionável que se assistiu a uma forte entrada de mulheres mais jovens para a carteira global, um facto que está invariavelmente associado à evolução da política de crédito à habitação das instituições financeiras nos últimos anos.

## 5.2. Descrição da amostra

No período em estudo, são observadas vidas expostas ao risco entre os 0 e os 110 anos no caso do sexo masculino, e entre os 0 e os 105 anos no caso do sexo feminino. Contudo, mais de 99% da exposição em ambos os sexos está concentrada numa faixa etária que vai dos 20 aos 75 anos de idade.

Na Figura 5.2.10 é possível verificar que os perfis gráficos dos quocientes de mortalidade se aproximam dos da população em geral. No entanto, observa-se que as tendências de mortalidade da população segura em caso de morte são anormalmente mais brandas do que as registadas na população em geral (e até na própria população segura em caso de vida), e para todo o intervalo de idades considerado. Ainda mais surpreendente é o facto dos quocientes relativos à população masculina serem, num número significativo de idades, inferiores aos da população feminina, um padrão inédito e invulgar face ao exposto na Secção 5.2.1.

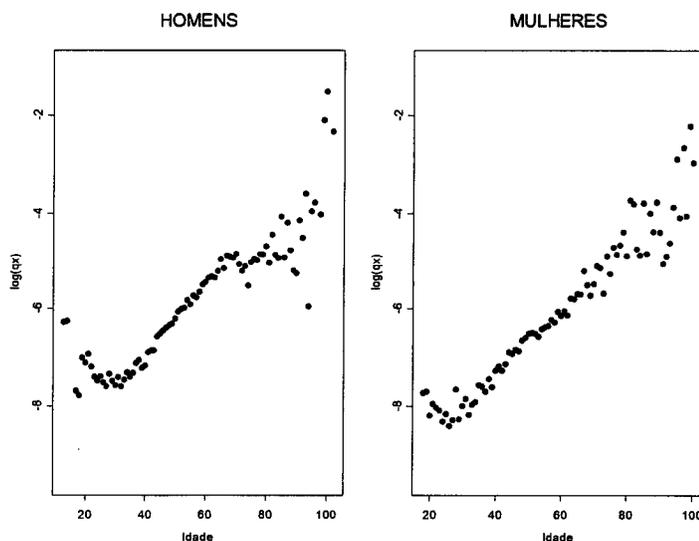


Figura 5.2.10: Estimativas brutas dos quocientes  $q_x$ , seguros em caso de morte

Este resultado inesperado, devidamente assinalado pelo ISP (2003), é explicado pela existência de erros na recolha de informação estatística sobre a mortalidade efectiva neste tipo de população.<sup>5</sup> A presença de erro na amostra introduz um enviesamento não

<sup>5</sup>Estes erros estão relacionados, desde logo, com atrasos na comunicação entre os bancos e as companhias de seguros do ramo vida aquando do falecimento de clientes que contraíram empréstimos para a aquisição de habitação, um facto que pode levar à perda de informação sobre os óbitos. Por outro lado, algumas coberturas do risco de morte associadas a operações de crédito pessoal ou a cartões de

negligenciável na análise e condiciona, inevitavelmente, os resultados e as conclusões a retirar quanto ao comportamento da mortalidade.<sup>6</sup>

## 5.3 Modelos lineares e não-lineares generalizados

Nesta secção, reportamos os principais resultados obtidos pela aplicação dos modelos lineares e não lineares generalizados descritos no Capítulo 3 na graduação da experiência relativa às populações de beneficiários de fundos de pensões e de pessoas seguras em caso de vida e em caso de morte, no período 2000-2004. O procedimento de graduação de  $q_x$  e  $\mu_x$  baseia-se na utilização de modelos com predictor linear do tipo  $LGM(r, s)$  ou  $GM(r, s)$ , com *link* logit ou logarítmico, respectivamente.

As estimativas dos parâmetros dos modelos (obtidas com base no método de máxima-verosimilhança) e os resultados dos testes de qualidade do ajustamento foram obtidos com recurso a uma rotina informática especialmente escrita para o efeito, implementada no *package* S-PLUS 2000 (Release 3). A exposição dos resultados compreende, nas três subpopulações em estudo, a descrição dos ajustamentos efectuados aos dados de base, os testes usados para seleccionar o modelo adequado e a análise da sua significância estatística.

### 5.3.1 Beneficiários de fundos de pensões

Os dados usados na graduação da mortalidade dos beneficiários de fundos de pensões correspondem apenas as idades compreendidas entre os 50 e os 100 anos, no período que decorre de 2000 a 2004, com centro no ano de 2002. Esta limitação do intervalo de análise é justificada por duas ordens de razões. Em primeiro lugar, pelo facto do número de vidas expostas ao risco ser reduzido nas restantes idades e apresentar as idiosincrasias referidas na Secção 5.2, razão pela qual a sua inclusão afectaria a significância estatística dos resultados. Em segundo lugar, recordamos que esta faixa etária é aquela que apresenta maior interesse relativo para as Sociedades Gestoras de Fundos de Pensões (SGFP), quer por concentrar a maioria da exposição ao risco, quer por ditar as responsabilidades financeiras de maior valor. Nas idades em que o número de óbitos registados no período foi nulo ou escasso (e.g., nas idades mais avançadas), procedeu-se ao agrupamento das

---

crédito não são, por vezes, accionadas, seja por corresponderem a situações de exclusão contratualmente estipuladas, levando a que a companhia de seguros não seja informada da ocorrência do falecimento, seja por desconhecimento ou até esquecimento dos beneficiários.

<sup>6</sup> A presença de erro na amostra pode igualmente ajudar a explicar o comportamento «estranho» dos quocientes na população do sexo masculino a partir dos 70 anos de idade, com uma quebra na natural tendência de crescimento dos quocientes com o aumento da idade.

### 5.3. Modelos lineares e não-lineares generalizados

idades contíguas até que a soma dos óbitos esperados ( $d_x^{\text{exp}}$ ) fosse maior ou igual a 5 (i.e.,  $d_x^{\text{exp}} \geq 5$ ).

#### Graduação de $q_x$

**Escolha da ordem do modelo** A primeira etapa no procedimento de graduação envolveu a determinação da ordem ( $r, s$ ) da fórmula de Gompertz-Makeham (3.4.1) que melhor ajusta os dados. Em ambos os sexos, foram testadas 30 combinações diferentes considerando  $s \in [2, 7]$  e  $r \in [0, 4]$ . Os valores da função de Log-Verossimilhança (LV) e da (*unscaled*) deviance são apresentados nos Quadros 5.3.1 e 5.3.2.

| Log-verossimilhança |          |          |          |          |          |          |
|---------------------|----------|----------|----------|----------|----------|----------|
| $r$                 | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                   | 37091.44 | 36996.57 | 36994.80 | 36992.12 | 36992.10 | 36991.06 |
| 1                   | 36994.54 | 36994.10 | 36993.64 | 36991.97 | 36989.69 | 36971.31 |
| 2                   | 36993.02 | 36991.19 | 36989.86 | 36985.38 | 36977.66 | 36970.45 |
| 3                   | 36991.73 | 36991.19 | 36989.34 | 36973.52 | 36970.59 | 36970.57 |
| 4                   | 36988.81 | 36986.34 | 36981.16 | 36969.42 | 36969.42 | 36969.51 |
| Deviance            |          |          |          |          |          |          |
| $r$                 | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                   | 342.14   | 152.39   | 148.84   | 143.49   | 143.45   | 141.38   |
| 1                   | 148.34   | 147.44   | 146.52   | 143.18   | 138.62   | 101.87   |
| 2                   | 145.21   | 141.63   | 138.97   | 130.02   | 114.56   | 100.16   |
| 3                   | 142.71   | 141.63   | 137.93   | 106.30   | 100.44   | 100.40   |
| 4                   | 136.87   | 131.92   | 121.57   | 98.09    | 98.08    | 98.26    |

Quadro 5.3.1: LV e deviance para pensionistas do sexo masculino,  $LGM(r, s)$

| Log-verossimilhança |          |          |          |          |          |          |
|---------------------|----------|----------|----------|----------|----------|----------|
| $r$                 | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                   | 19686.30 | 19641.51 | 19630.38 | 19625.97 | 19620.98 | 19616.32 |
| 1                   | 19633.47 | 19632.63 | 19629.44 | 19621.08 | 19616.36 | 19615.97 |
| 2                   | 19633.46 | 19628.21 | 19616.39 | 19615.67 | 19615.67 | 19615.57 |
| 3                   | 19633.21 | 19619.33 | 19615.78 | 19615.66 | 19615.65 | 19615.56 |
| 4                   | 19621.72 | 19616.39 | 19615.77 | 19615.64 | 19615.37 | 19615.36 |
| Deviance            |          |          |          |          |          |          |
| $r$                 | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                   | 182.70   | 93.13    | 70.87    | 62.03    | 52.07    | 42.75    |
| 1                   | 77.07    | 75.36    | 68.99    | 52.26    | 42.81    | 42.04    |
| 2                   | 77.01    | 66.52    | 42.87    | 41.44    | 41.44    | 41.23    |
| 3                   | 76.52    | 48.76    | 41.66    | 41.43    | 41.41    | 41.23    |
| 4                   | 53.54    | 42.88    | 41.65    | 41.37    | 40.84    | 40.83    |

Quadro 5.3.2: LV e deviance para pensionistas do sexo feminino,  $LGM(r, s)$

Como seria de antever, regista-se uma diminuição paulatina nos valores da Log-verosimilhança (LV) e da deviance à medida que nos movemos ao longo de uma linha ou ao longo de uma coluna da tabela, ou seja, à medida que aumenta o número de parâmetros do modelo. Para determinar a ordem óptima da fórmula é necessário avaliar se os decréscimos na LV e na deviance são estatisticamente significativos, ou seja, é imprescindível verificar se a melhoria na qualidade da graduação conseguida com a adição de parâmetros ao modelo é compensada pelo inevitável aumento da sua complexidade. A este propósito, recordamos que a diferença entre os valores de duas colunas (ou linhas) adjacentes da deviance pode ser aproximada por uma distribuição de  $\chi^2$  com um grau de liberdade. A análise dos quadros permite identificar os modelos  $LGM(1, 7)$ ,  $LGM(2, 6)$  e  $LGM(3, 5)$  como potenciais candidatos no caso da população de pensionistas do sexo masculino, e os modelos  $LGM(0, 7)$ ,  $LGM(1, 6)$ ,  $LGM(2, 4)$ ,  $LGM(3, 4)$  e  $LGM(4, 3)$  na população do sexo feminino.

Para seleccionar o modelo óptimo, foram considerados outros indicadores de qualidade do ajustamento, reportados nos Quadros 5.3.3 e 5.3.4.<sup>7</sup> Nestes, as primeiras três colunas indicam, respectivamente, a ordem  $(r, s)$  da fórmula, o valor (*unscaled*) da estatística  $\chi^2$  e o correspondente *p-value*. No caso dos pensionistas do sexo masculino, o *p-value* do teste  $\chi^2$  foi ajustado por um factor de escala de modo a contemplar a previsível presença de sobredispersão na amostra. As três colunas seguintes reportam, por esta ordem, os *p-values* do teste dos sinais (*Signs Test*), do teste das sequências ou das mudanças de sinal (*Runs Test*) e do teste de Kolmogorov-Smirnov para a igualdade entre a distribuição observada e a distribuição estimada. As duas colunas seguintes incluem informação sobre os *p-values* dos testes de autocorrelação efectuados, nomeadamente dos testes de Portmanteau e de Ljung-Box.

A coluna intitulada “Par. Sign.” sinaliza o resultado da análise à significância estatística dos parâmetros dos modelos. Deste modo, na eventualidade de todos os parâmetros estimados se revelarem, num teste bilateral, significativamente diferentes de zero, a sigla “S” é apresentada. Caso contrário, é indicada a sigla “N”. A obtenção de estimativas não significativas é um primeiro sinal de possível sobre-parametrização do modelo.

A coluna intitulada “Conf.” evidencia se o modelo testado apresenta (S) ou não (N) uma boa configuração, entendendo-se por tal a capacidade deste para gerar estimativas plausíveis fora do intervalo de graduação (nomeadamente para  $x > 100$ ), em particular quocientes de mortalidade crescentes com a idade.<sup>8</sup> Por fim, a coluna intitulada

---

<sup>7</sup>Para mais detalhes sobre a definição dos indicadores e a forma de cálculo, veja-se o Capítulo 3.

<sup>8</sup>No capítulo seguinte, são apresentadas soluções alternativas para contornar o problema da irregularidade da mortalidade nas idades avançadas, que passam, em particular, pela adopção de métodos de

### 5.3. Modelos lineares e não-lineares generalizados

“Int. Conf.” apresenta informação sobre a qualidade do ajustamento proporcionada pelo modelo nas idades intermédias, onde se concentra a maioria da exposição ao risco. Para qualificar esta característica, considerámos que um modelo apresenta um desempenho adequado (S) se os intervalos de confiança estimados nas idades intermédias são relativamente estreitos e incluem a maioria das estimativas brutas dos quocientes  $q_x$ .

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 367.7    | 0.000       | 0.984           | 0.000                    | 1                | 0.000               | 0.000                | S             | S     | S             |
| (0, 3)   | 155.0    | 0.142       | 0.556           | 0.317                    | 1                | 0.380               | 0.138                | S             | S     | S             |
| (0, 4)   | 151.8    | 0.143       | 0.336           | 0.700                    | 1                | 0.171               | 0.034                | S             | S     | S             |
| (0, 5)   | 146.2    | 0.165       | 0.556           | 0.668                    | 1                | 0.169               | 0.031                | S             | S     | S             |
| (0, 6)   | 146.2    | 0.140       | 0.556           | 0.668                    | 1                | 0.175               | 0.032                | N             | S     | S             |
| (0, 7)   | 143.5    | 0.138       | 0.336           | 0.700                    | 1                | 0.151               | 0.024                | N             | S     | S             |
| (1, 2)   | 151.3    | 0.172       | 0.336           | 0.700                    | 1                | 0.134               | 0.022                | S             | S     | S             |
| (1, 3)   | 150.2    | 0.156       | 0.336           | 0.700                    | 1                | 0.167               | 0.031                | N             | S     | S             |
| (1, 4)   | 149.3    | 0.140       | 0.336           | 0.700                    | 1                | 0.145               | 0.025                | N             | S     | S             |
| (1, 5)   | 145.8    | 0.144       | 0.556           | 0.668                    | 1                | 0.159               | 0.028                | N             | S     | S             |
| (1, 6)   | 140.7    | 0.160       | 0.444           | 0.676                    | 1                | 0.241               | 0.053                | S             | S     | S             |
| (1, 7)   | 103.0    | 0.641       | 0.444           | 0.659                    | 1                | 0.334               | 0.124                | N             | N     | S             |
| (2, 2)   | 147.9    | 0.176       | 0.336           | 0.700                    | 1                | 0.212               | 0.046                | S             | S     | S             |
| (2, 3)   | 144.3    | 0.181       | 0.556           | 0.668                    | 1                | 0.188               | 0.036                | S             | S     | S             |
| (2, 4)   | 141.9    | 0.177       | 0.556           | 0.886                    | 1                | 0.286               | 0.071                | S             | S     | S             |
| (2, 5)   | 131.7    | 0.252       | 0.556           | 0.886                    | 1                | 0.319               | 0.093                | N             | S     | S             |
| (2, 6)   | 116.3    | 0.426       | 0.240           | 0.967                    | 1                | 0.300               | 0.116                | S             | S     | S             |
| (2, 7)   | 101.2    | 0.628       | 0.556           | 0.198                    | 1                | 0.146               | 0.040                | N             | N     | S             |
| (3, 2)   | 145.6    | 0.170       | 0.444           | 0.676                    | 1                | 0.203               | 0.041                | N             | S     | S             |
| (3, 3)   | 144.3    | 0.155       | 0.556           | 0.668                    | 1                | 0.189               | 0.036                | N             | S     | S             |
| (3, 4)   | 140.5    | 0.162       | 0.556           | 0.886                    | 1                | 0.264               | 0.062                | N             | S     | S             |
| (3, 5)   | 107.6    | 0.566       | 0.443           | 0.481                    | 1                | 0.166               | 0.050                | N             | S     | S             |
| (3, 6)   | 101.6    | 0.621       | 0.336           | 0.922                    | 1                | 0.122               | 0.032                | S             | N     | S             |
| (3, 7)   | 101.6    | 0.578       | 0.234           | 0.590                    | 1                | 0.129               | 0.034                | N             | N     | S             |
| (4, 2)   | 139.0    | 0.205       | 0.556           | 0.668                    | 1                | 0.231               | 0.052                | S             | S     | S             |
| (4, 3)   | 133.6    | 0.230       | 0.444           | 0.676                    | 1                | 0.288               | 0.077                | S             | S     | S             |
| (4, 4)   | 123.3    | 0.322       | 0.240           | 0.535                    | 1                | 0.411               | 0.174                | N             | S     | S             |
| (4, 5)   | 98.6     | 0.671       | 0.444           | 0.877                    | 1                | 0.191               | 0.058                | N             | N     | S             |
| (4, 6)   | 98.6     | 0.628       | 0.444           | 0.877                    | 1                | 0.188               | 0.057                | N             | N     | S             |
| (4, 7)   | 99.1     | 0.576       | 0.664           | 0.877                    | 1                | 0.090               | 0.020                | N             | N     | S             |

Quadro 5.3.3: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , pensionistas do sexo masculino

Como se observa, em ambos os sexos são inúmeros os modelos que satisfazem todos os critérios de bondade do ajustamento. No entanto, para os pensionistas do sexo masculino verifica-se que, de entre os modelos anteriormente seleccionados como candidatos a óptimo, o único que respeita todos os indicadores de qualidade é o modelo  $LGM(2, 6)$ . Os modelos  $LGM(1, 7)$  e  $LGM(3, 5)$  são descartados pelo facto de gerarem parâmetros fecho de tábuas.

### 5.3. Modelos lineares e não-lineares generalizados

não significativos e/ou uma configuração pouco adequada em termos de graduação, não obstante proporcionarem, por exemplo, valores inferiores para a estatística deviance. Conclui-se, assim, pela análise da bateria de indicadores acima indicada, que o modelo que melhor ajusta o comportamento da mortalidade dos pensionistas do sexo masculino é o modelo  $LGM(2, 6)$ . No caso dos pensionistas do sexo feminino, um raciocínio análogo identifica o modelo  $LGM(0, 7)$  como sendo o único que, de entre os candidatos a óptimo, respeita todos os critérios de qualidade do ajustamento. Trata-se de um modelo que, no contexto dos GLMs, assume uma estrutura linear, correspondente a um ajustamento mediante um polinómio de grau 6.

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Signf. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|----------------|-------|---------------|
| (0, 2)   | 218.3    | 0.000       | 0.919           | 0.035                    | 1                | 0.000               | 0.000                | S              | S     | N             |
| (0, 3)   | 95.5     | 0.000       | 0.500           | 0.259                    | 1                | 0.026               | 0.002                | S              | S     | S             |
| (0, 4)   | 70.8     | 0.014       | 0.390           | 0.553                    | 1                | 0.022               | 0.001                | S              | S     | S             |
| (0, 5)   | 63.8     | 0.042       | 0.610           | 0.569                    | 1                | 0.471               | 0.167                | S              | S     | S             |
| (0, 6)   | 52.4     | 0.210       | 0.500           | 0.569                    | 1                | 0.249               | 0.047                | S              | N     | S             |
| (0, 7)   | 42.6     | 0.533       | 0.712           | 0.084                    | 1                | 0.037               | 0.002                | S              | S     | S             |
| (1, 2)   | 77.2     | 0.005       | 0.712           | 0.267                    | 1                | 0.029               | 0.001                | S              | S     | S             |
| (1, 3)   | 75.3     | 0.006       | 0.390           | 0.980                    | 1                | 0.014               | 0.001                | N              | S     | S             |
| (1, 4)   | 69.0     | 0.016       | 0.610           | 0.569                    | 1                | 0.064               | 0.006                | N              | S     | S             |
| (1, 5)   | 53.5     | 0.181       | 0.610           | 0.256                    | 1                | 0.281               | 0.061                | S              | N     | S             |
| (1, 6)   | 42.7     | 0.530       | 0.390           | 0.084                    | 1                | 0.034               | 0.002                | S              | N     | S             |
| (1, 7)   | 41.8     | 0.522       | 0.500           | 0.089                    | 1                | 0.026               | 0.001                | N              | S     | S             |
| (2, 2)   | 77.1     | 0.004       | 0.712           | 0.267                    | 1                | 0.023               | 0.001                | N              | S     | S             |
| (2, 3)   | 69.2     | 0.015       | 0.712           | 0.553                    | 1                | 0.523               | 0.209                | S              | S     | S             |
| (2, 4)   | 42.8     | 0.564       | 0.500           | 0.089                    | 1                | 0.039               | 0.003                | S              | N     | S             |
| (2, 5)   | 41.3     | 0.590       | 0.610           | 0.089                    | 1                | 0.024               | 0.001                | N              | S     | S             |
| (2, 6)   | 41.2     | 0.547       | 0.610           | 0.089                    | 1                | 0.024               | 0.001                | N              | S     | S             |
| (2, 7)   | 41.1     | 0.511       | 0.390           | 0.084                    | 1                | 0.018               | 0.001                | N              | N     | S             |
| (3, 2)   | 77.0     | 0.003       | 0.500           | 0.573                    | 1                | 0.039               | 0.003                | N              | S     | S             |
| (3, 3)   | 49.3     | 0.306       | 0.712           | 0.553                    | 1                | 0.169               | 0.026                | S              | S     | S             |
| (3, 4)   | 41.5     | 0.580       | 0.288           | 0.074                    | 1                | 0.022               | 0.001                | N              | S     | S             |
| (3, 5)   | 41.2     | 0.548       | 0.610           | 0.089                    | 1                | 0.024               | 0.002                | N              | S     | S             |
| (3, 6)   | 41.2     | 0.505       | 0.610           | 0.089                    | 1                | 0.029               | 0.002                | N              | S     | S             |
| (3, 7)   | 41.1     | 0.467       | 0.610           | 0.089                    | 1                | 0.019               | 0.001                | N              | S     | S             |
| (4, 2)   | 52.8     | 0.198       | 0.201           | 0.890                    | 1                | 0.428               | 0.140                | S              | S     | S             |
| (4, 3)   | 42.8     | 0.523       | 0.610           | 0.089                    | 1                | 0.030               | 0.002                | N              | S     | S             |
| (4, 4)   | 41.5     | 0.537       | 0.288           | 0.074                    | 1                | 0.021               | 0.001                | N              | S     | S             |
| (4, 5)   | 41.2     | 0.507       | 0.610           | 0.089                    | 1                | 0.022               | 0.001                | N              | S     | S             |
| (4, 6)   | 40.8     | 0.480       | 0.610           | 0.047                    | 1                | 0.028               | 0.002                | N              | N     | S             |
| (4, 7)   | 40.8     | 0.425       | 0.610           | 0.089                    | 1                | 0.019               | 0.001                | N              | S     | S             |

Quadro 5.3.4: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , pensionistas do sexo feminino

As Figuras 5.3.1 e 5.3.3 dão-nos uma percepção clara quanto à qualidade do ajustamento proporcionada pelos modelos óptimos.

### 5.3. Modelos lineares e não-lineares generalizados

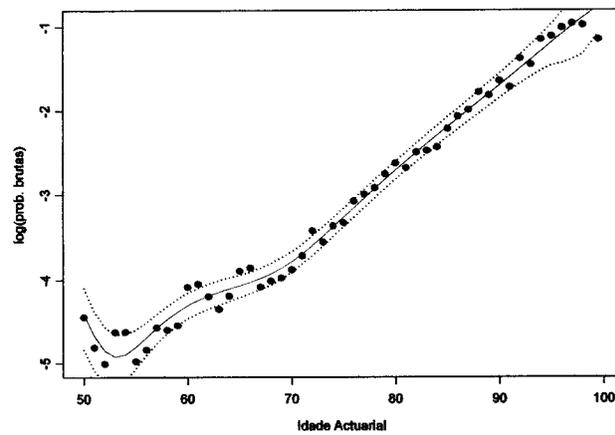


Figura 5.3.1: Modelo  $LGM(2,6)$ , pensionistas do sexo masculino

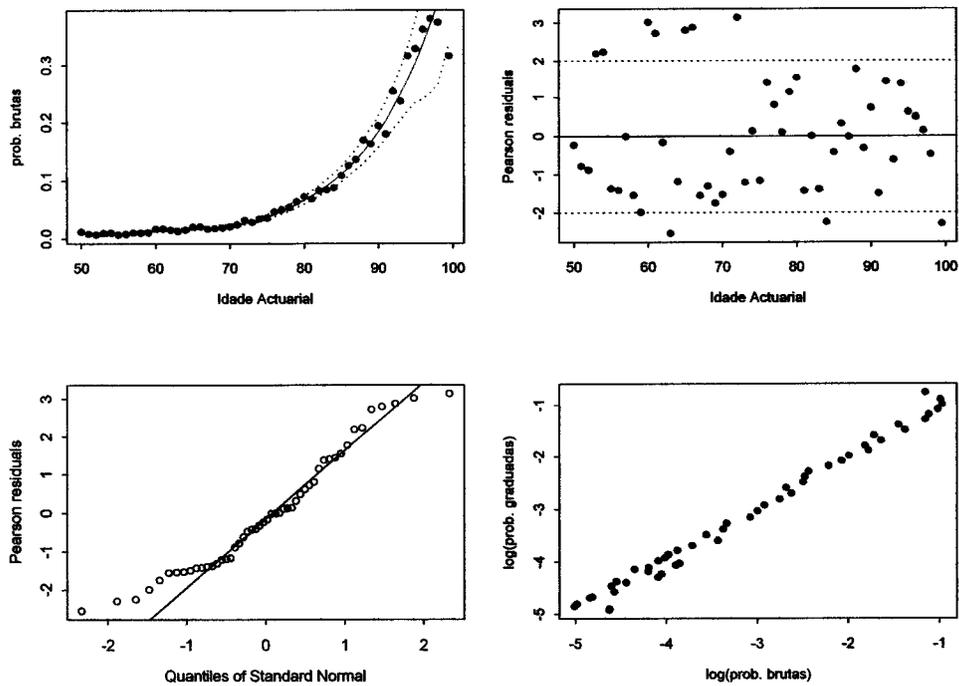


Figura 5.3.2: Resíduos de Pearson do modelo  $LGM(2,6)$ , pensionistas do sexo masculino

### 5.3. Modelos lineares e não-lineares generalizados

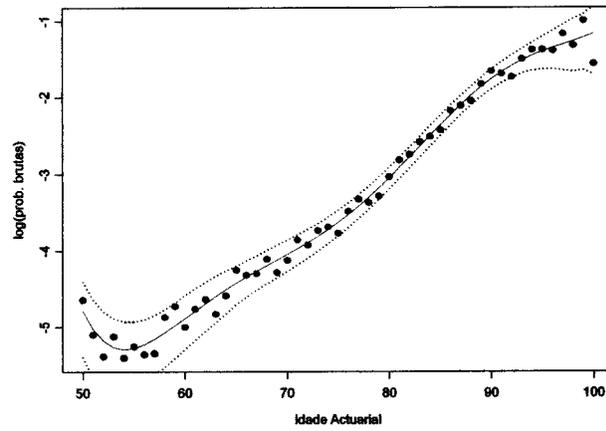


Figura 5.3.3: Modelo  $LGM(0,7)$ , pensionistas do sexo feminino

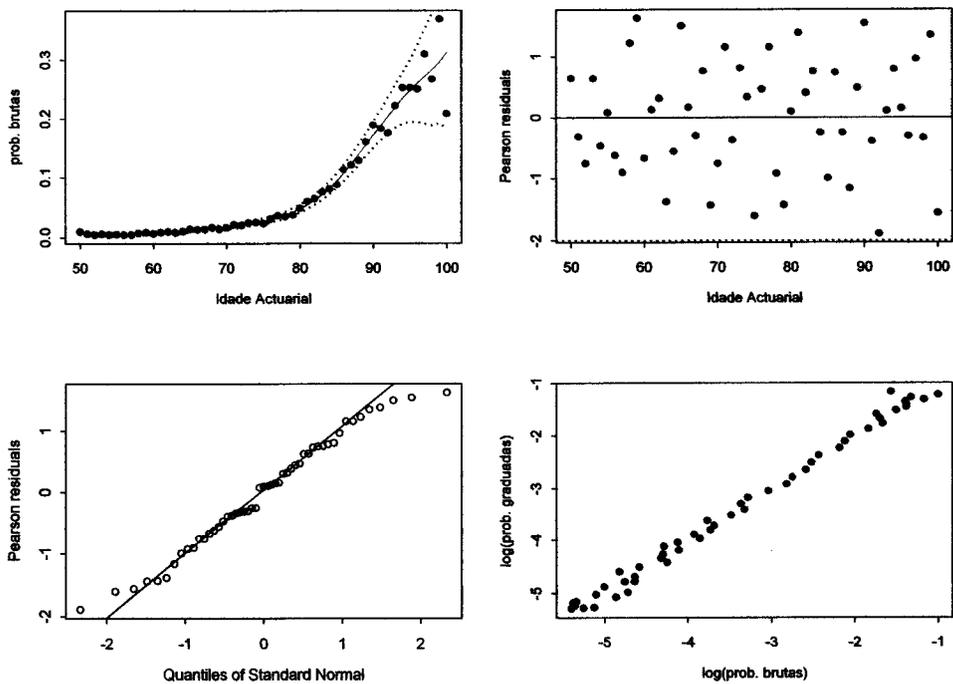


Figura 5.3.4: Resíduos do modelo  $LGM(0,7)$ , pensionistas do sexo feminino

### 5.3. Modelos lineares e não-lineares generalizados

Nestas são representados, no mesmo referencial, os quocientes brutos e graduados e os respectivos intervalos de confiança a 95%. Nas Figuras 5.3.2 e 5.3.4 juntamos mais alguns elementos de diagnóstico, representando o perfil dos resíduos de Pearson do modelo por idade. Comparamos ainda os quantis da distribuição dos resíduos com os da distribuição Normal padronizada. No caso dos pensionistas do sexo masculino, assinalamos alguns desvios relativos elevados, situados fora do intervalo de confiança a 95%, bem como a presença de algumas observações nas caudas da distribuição Normal. A explicação mais plausível para este padrão gráfico encontra-se na presença de duplicidades na amostra.

**Significância estatística do modelo** Para completar a análise, o Quadro 5.3.5 apresenta algum detalhe sobre os modelos óptimos escolhidos na secção anterior, nomeadamente o valor dos parâmetros estimados (Coef.), o desvio-padrão das estimativas (se), o *t*-ratio e correspondente *p*-value.

|            | Homens: $LGM(2, 6)$ |       |                 |                 | Mulheres: $LGM(0, 7)$ |       |                 |                 |
|------------|---------------------|-------|-----------------|-----------------|-----------------------|-------|-----------------|-----------------|
|            | Coef.               | se    | <i>t</i> -ratio | <i>p</i> -value | Coef.                 | se    | <i>t</i> -ratio | <i>p</i> -value |
| $\alpha_0$ | -1.652604           | 0.258 | -6.40           | <0.0001         |                       |       |                 |                 |
| $\alpha_1$ | -1.26309            | 0.180 | -7.01           | <0.0001         |                       |       |                 |                 |
| $\beta_0$  | 0.525325            | 0.153 | 3.44            | <0.0001         | -3.594738             | 0.028 | -127.76         | <0.0001         |
| $\beta_1$  | 0.812968            | 0.029 | 28.10           | <0.0001         | 2.529659              | 0.115 | 22.09           | <0.0001         |
| $\beta_2$  | -0.216991           | 0.018 | -11.72          | <0.0001         | 2.24867               | 0.411 | 5.47            | <0.0001         |
| $\beta_3$  | 0.135882            | 0.014 | 9.86            | <0.0001         | 1.242792              | 0.525 | 2.37            | 0.00448         |
| $\beta_4$  | -0.097409           | 0.022 | -4.50           | <0.0001         | -4.481736             | 1.272 | -3.52           | <0.0001         |
| $\beta_5$  | 0.173482            | 0.044 | 3.90            | <0.0001         | -1.781275             | 0.510 | -3.50           | <0.0001         |
| $\beta_6$  |                     |       |                 |                 | 3.049402              | 0.993 | 3.07            | 0.0001          |

Quadro 5.3.5: Significância estatística dos parâmetros,  $LGM(r, s)$ , pensionistas

|   | Óbitos ( <i>A, E</i> ) |                 | Deviance              |                 |                 | Teste $\chi^2$  |           |                 |
|---|------------------------|-----------------|-----------------------|-----------------|-----------------|-----------------|-----------|-----------------|
|   | ( <i>A - E</i> )       | <i>A/E</i> (%)  | Null                  | Unscaled        | <i>p</i> -value | Unsc. $\chi^2$  | g.l.      | <i>p</i> -value |
| H | 0                      | 100.0           | 7268.17               | 114.56          | 0.453           | 116.26          | 42        | 0.426           |
| M | 0                      | 100.0           | 6078.44               | 42.75           | 0.525           | 42.56           | 44        | 0.533           |
|   | Kolmogorov-Smirnov     |                 | Signs Test            |                 |                 |                 |           |                 |
|   | Desv Max               | <i>p</i> -value | Desv.(+)              | Desv.(-)        | <i>p</i> -value | $z_x > 2$       | $z_x > 3$ |                 |
| H | 0.0063365              | 1               | 22                    | 28              | 0.240           | 10              | 2         |                 |
| M | 0.0040279              | 1               | 27                    | 24              | 0.712           | 0               | 0         |                 |
|   | Runs Test              |                 | Testes Autocorrelação |                 |                 |                 |           |                 |
|   | N.º runs               | <i>p</i> -value | Portm.                | <i>p</i> -value | Ljung-Box       | <i>p</i> -value |           |                 |
| H | 25                     | 0.9676          | 22.769                | 0.300           | 27.73           | 0.116           |           |                 |
| M | 32                     | 0.0839          | 32.671                | 0.037           | 42.17           | 0.003           |           |                 |

Quadro 5.3.6: Testes de qualidade do ajustamento,  $LGM(r, s)$ , pensionistas

### 5.3. Modelos lineares e não-lineares generalizados

Como se observa, os parâmetros estimados em ambos os sexos são significativamente diferentes de zero. Por fim, o Quadro 5.3.6 inclui informação detalhada sobre os resultados dos testes de qualidade do ajustamento efectuados aos modelos óptimos. Neste, a coluna ( $A - E$ ) denota a diferença entre os óbitos observados ( $A$ ) e estimados ( $E$ ).

#### Graduação de $\mu_x$

**Escolha da ordem do modelo** Para determinar a ordem do predictor linear  $GM(r, s)$  usado na graduação de  $\mu_x$ , foram testadas 30 variantes considerando  $s \in [2, 7]$  e  $r \in [0, 4]$ . Os valores da Log-verosimilhança e da deviance são reportados nos Quadros 5.3.7 e 5.3.8.

| Log-verosimilhança |          |          |          |          |          |          |
|--------------------|----------|----------|----------|----------|----------|----------|
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 37080.05 | 37005.51 | 36999.86 | 36996.17 | 36996.15 | 36995.39 |
| 1                  | 36999.55 | 36999.41 | 36998.17 | 36996.00 | 36994.18 | 36976.04 |
| 2                  | 36999.39 | 36995.11 | 36994.34 | 36988.71 | 36982.84 | 36974.78 |
| 3                  | 36996.59 | 36995.11 | 36993.91 | 36977.45 | 36974.71 | 36974.71 |
| 4                  | 36994.06 | 36989.50 | 36985.36 | 36973.68 | 36973.67 | 36973.87 |
| Deviance           |          |          |          |          |          |          |
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 310.76   | 161.70   | 150.39   | 143.00   | 142.97   | 141.45   |
| 1                  | 149.78   | 149.49   | 147.00   | 142.68   | 139.02   | 102.75   |
| 2                  | 149.44   | 140.89   | 139.35   | 126.87   | 115.70   | 100.22   |
| 3                  | 143.85   | 140.89   | 138.48   | 105.57   | 100.09   | 100.08   |
| 4                  | 129.06   | 127.50   | 121.37   | 98.02    | 98.00    | 97.52    |

Quadro 5.3.7: LV e deviance para pensionistas do sexo masculino,  $GM(r, s)$

| Log-verosimilhança |          |          |          |          |          |          |
|--------------------|----------|----------|----------|----------|----------|----------|
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 19681.72 | 19652.38 | 19634.12 | 19629.05 | 19624.93 | 19619.69 |
| 1                  | 19641.61 | 19637.50 | 19632.88 | 19624.83 | 19619.78 | 19619.29 |
| 2                  | 19640.57 | 19631.57 | 19619.94 | 19618.96 | 19618.96 | 19618.84 |
| 3                  | 19639.48 | 19622.73 | 19619.12 | 19618.95 | 19618.94 | 19618.84 |
| 4                  | 19629.89 | 19619.58 | 19619.11 | 19618.91 | 19618.65 | 19618.83 |
| Deviance           |          |          |          |          |          |          |
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 166.73   | 108.06   | 71.54    | 61.40    | 53.17    | 42.69    |
| 1                  | 86.73    | 78.29    | 69.06    | 52.95    | 42.86    | 41.88    |
| 2                  | 84.44    | 66.44    | 43.19    | 41.23    | 41.22    | 40.98    |
| 3                  | 82.27    | 48.76    | 41.55    | 41.20    | 41.18    | 40.61    |
| 4                  | 48.23    | 42.47    | 41.47    | 41.12    | 40.59    | 38.79    |

Quadro 5.3.8: LV e deviance para pensionistas do sexo feminino,  $GM(r, s)$

### 5.3. Modelos lineares e não-lineares generalizados

Conforme se antevia, os indicadores de qualidade do ajustamento são, em todos os modelos testados, muito semelhantes aos valores obtidos na graduação dos quocientes  $q_x$  mediante um previsor linear do tipo  $LGM(r, s)$ . Assim sendo, não surpreende que os modelos  $GM(1, 7)$ ,  $GM(2, 6)$  e  $GM(3, 5)$ , no caso dos pensionistas do sexo masculino, e os modelos  $GM(0, 7)$ ,  $GM(1, 6)$ ,  $GM(2, 4)$ ,  $GM(3, 4)$  e  $GM(4, 3)$ , no caso dos pensionistas do sexo feminino, se constituam, uma vez mais, como sérios candidatos a óptimo. Para ajudar a identificar o modelo óptimo, os Quadros 5.3.9 e 5.3.10 resumem os resultados dos testes de qualidade do ajustamento para as 30 variantes consideradas. A análise dos indicadores permite concluir, sem estranheza, que o modelo  $GM(2, 6)$  ( $GM(0, 7)$ ) proporciona o melhor desempenho para os pensionistas do sexo masculino (feminino).

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 334.8    | 0.000       | 0.966           | 0.000                    | 1                | 0.000               | 0.000                | S             | S     | S             |
| (0, 3)   | 162.5    | 0.094       | 0.444           | 0.676                    | 1                | 0.527               | 0.259                | S             | S     | S             |
| (0, 4)   | 153.2    | 0.133       | 0.444           | 0.322                    | 1                | 0.165               | 0.033                | S             | S     | S             |
| (0, 5)   | 145.8    | 0.169       | 0.556           | 0.668                    | 1                | 0.170               | 0.031                | S             | S     | S             |
| (0, 6)   | 145.8    | 0.144       | 0.556           | 0.668                    | 1                | 0.174               | 0.032                | N             | S     | S             |
| (0, 7)   | 143.6    | 0.137       | 0.556           | 0.668                    | 1                | 0.147               | 0.023                | N             | S     | S             |
| (1, 2)   | 152.0    | 0.166       | 0.336           | 0.337                    | 1                | 0.198               | 0.043                | S             | S     | S             |
| (1, 3)   | 152.0    | 0.142       | 0.444           | 0.322                    | 1                | 0.165               | 0.032                | N             | S     | S             |
| (1, 4)   | 149.7    | 0.137       | 0.336           | 0.670                    | 1                | 0.141               | 0.024                | N             | S     | S             |
| (1, 5)   | 145.3    | 0.147       | 0.556           | 0.668                    | 1                | 0.160               | 0.028                | N             | S     | S             |
| (1, 6)   | 141.1    | 0.158       | 0.444           | 0.676                    | 1                | 0.224               | 0.047                | S             | S     | S             |
| (1, 7)   | 103.9    | 0.626       | 0.444           | 0.895                    | 1                | 0.344               | 0.129                | N             | N     | S             |
| (2, 2)   | 151.4    | 0.147       | 0.336           | 0.337                    | 1                | 0.243               | 0.062                | N             | S     | S             |
| (2, 3)   | 143.6    | 0.188       | 0.556           | 0.668                    | 1                | 0.193               | 0.038                | S             | S     | S             |
| (2, 4)   | 142.0    | 0.175       | 0.556           | 0.143                    | 1                | 0.256               | 0.059                | N             | S     | S             |
| (2, 5)   | 129.8    | 0.274       | 0.444           | 0.676                    | 1                | 0.331               | 0.100                | S             | S     | S             |
| (2, 6)   | 117.5    | 0.407       | 0.240           | 0.967                    | 1                | 0.335               | 0.133                | S             | S     | S             |
| (2, 7)   | 101.3    | 0.626       | 0.556           | 0.668                    | 1                | 0.148               | 0.041                | N             | N     | S             |
| (3, 2)   | 146.6    | 0.162       | 0.444           | 0.676                    | 1                | 0.208               | 0.043                | N             | S     | S             |
| (3, 3)   | 143.6    | 0.162       | 0.556           | 0.668                    | 1                | 0.192               | 0.037                | N             | S     | S             |
| (3, 4)   | 140.8    | 0.160       | 0.556           | 0.886                    | 1                | 0.238               | 0.052                | N             | S     | S             |
| (3, 5)   | 106.8    | 0.579       | 0.444           | 0.481                    | 1                | 0.161               | 0.047                | N             | N     | S             |
| (3, 6)   | 101.2    | 0.628       | 0.240           | 0.590                    | 1                | 0.121               | 0.032                | S             | N     | S             |
| (3, 7)   | 101.2    | 0.584       | 0.240           | 0.590                    | 1                | 0.124               | 0.033                | N             | N     | S             |
| (4, 2)   | 140.8    | 0.187       | 0.444           | 0.676                    | 1                | 0.247               | 0.058                | S             | S     | S             |
| (4, 3)   | 131.4    | 0.255       | 0.556           | 0.886                    | 1                | 0.304               | 0.086                | S             | S     | S             |
| (4, 4)   | 123.0    | 0.326       | 0.240           | 0.535                    | 1                | 0.407               | 0.170                | N             | S     | S             |
| (4, 5)   | 98.5     | 0.672       | 0.444           | 0.877                    | 1                | 0.194               | 0.060                | S             | N     | S             |
| (4, 6)   | 98.5     | 0.630       | 0.444           | 0.877                    | 1                | 0.188               | 0.057                | N             | N     | S             |
| (4, 7)   | 99.2     | 0.574       | 0.664           | 0.877                    | 1                | 0.076               | 0.016                | N             | N     | S             |

Quadro 5.3.9: Indicadores de qualidade do ajustamento,  $GM(r, s)$ , pensionistas do sexo masculino

### 5.3. Modelos lineares e não-lineares generalizados

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 196.8    | 0.000       | 0.920           | 0.035                    | 1                | 0.000               | 0.000                | S             | S     | N             |
| (0, 3)   | 108.8    | 0.000       | 0.500           | 0.004                    | 1                | 0.002               | 0.000                | S             | S     | S             |
| (0, 4)   | 71.0     | 0.014       | 0.390           | 0.553                    | 1                | 0.009               | 0.000                | S             | S     | S             |
| (0, 5)   | 63.1     | 0.048       | 0.610           | 0.570                    | 1                | 0.455               | 0.155                | S             | S     | S             |
| (0, 6)   | 53.7     | 0.175       | 0.500           | 0.569                    | 1                | 0.264               | 0.052                | N             | N     | S             |
| (0, 7)   | 42.4     | 0.540       | 0.712           | 0.084                    | 1                | 0.037               | 0.002                | S             | S     | S             |
| (1, 2)   | 84.9     | 0.001       | 0.610           | 0.259                    | 1                | 0.007               | 0.000                | S             | S     | S             |
| (1, 3)   | 77.2     | 0.004       | 0.288           | 0.616                    | 1                | 0.003               | 0.000                | S             | S     | S             |
| (1, 4)   | 68.6     | 0.017       | 0.610           | 0.569                    | 1                | 0.044               | 0.003                | N             | S     | S             |
| (1, 5)   | 54.4     | 0.160       | 0.712           | 0.553                    | 1                | 0.302               | 0.068                | N             | N     | S             |
| (1, 6)   | 42.8     | 0.524       | 0.390           | 0.084                    | 1                | 0.037               | 0.002                | S             | N     | S             |
| (1, 7)   | 41.6     | 0.530       | 0.500           | 0.089                    | 1                | 0.025               | 0.001                | N             | S     | S             |
| (2, 2)   | 82.5     | 0.001       | 0.610           | 0.090                    | 1                | 0.000               | 0.000                | N             | S     | S             |
| (2, 3)   | 69.0     | 0.016       | 0.712           | 0.553                    | 1                | 0.519               | 0.208                | S             | S     | S             |
| (2, 4)   | 43.3     | 0.544       | 0.500           | 0.089                    | 1                | 0.048               | 0.003                | S             | N     | S             |
| (2, 5)   | 41.0     | 0.600       | 0.610           | 0.089                    | 1                | 0.023               | 0.001                | N             | S     | S             |
| (2, 6)   | 41.0     | 0.558       | 0.610           | 0.089                    | 1                | 0.022               | 0.001                | N             | S     | S             |
| (2, 7)   | 40.9     | 0.521       | 0.390           | 0.084                    | 1                | 0.017               | 0.001                | N             | N     | S             |
| (3, 2)   | 81.5     | 0.001       | 0.500           | 0.573                    | 1                | 0.005               | 0.000                | N             | S     | S             |
| (3, 3)   | 49.3     | 0.304       | 0.610           | 0.569                    | 1                | 0.173               | 0.027                | S             | S     | S             |
| (3, 4)   | 41.4     | 0.583       | 0.288           | 0.074                    | 1                | 0.022               | 0.001                | N             | N     | S             |
| (3, 5)   | 41.0     | 0.560       | 0.610           | 0.089                    | 1                | 0.024               | 0.001                | N             | S     | S             |
| (3, 6)   | 40.9     | 0.516       | 0.500           | 0.156                    | 1                | 0.027               | 0.002                | N             | S     | S             |
| (3, 7)   | 40.9     | 0.493       | 0.610           | 0.089                    | 1                | 0.017               | 0.001                | N             | N     | S             |
| (4, 2)   | 60.8     | 0.058       | 0.288           | 0.616                    | 1                | 0.260               | 0.060                | S             | S     | S             |
| (4, 3)   | 42.4     | 0.542       | 0.500           | 0.089                    | 1                | 0.026               | 0.001                | N             | S     | S             |
| (4, 4)   | 41.4     | 0.541       | 0.288           | 0.074                    | 1                | 0.020               | 0.001                | N             | N     | S             |
| (4, 5)   | 40.9     | 0.519       | 0.610           | 0.089                    | 1                | 0.021               | 0.001                | N             | S     | S             |
| (4, 6)   | 40.6     | 0.490       | 0.610           | 0.047                    | 1                | 0.026               | 0.002                | N             | N     | S             |
| (4, 7)   | 40.8     | 0.434       | 0.500           | 0.089                    | 1                | 0.018               | 0.001                | N             | S     | S             |

Quadro 5.3.10: Indicadores de qualidade do ajustamento,  $GM(r, s)$ , pensionistas do sexo feminino

As Figuras 5.3.5, 5.3.6, 5.3.7 e 5.3.8 evidenciam a qualidade do ajustamento proporcionada pelos modelos óptimos identificados para os pensionistas de ambos os sexos.

Em resumo, a análise conduzida nesta secção permite-nos concluir, em primeiro lugar, que um modelo do tipo  $GM(2, 6)$  (ou do tipo  $LGM(2, 6)$ ) proporciona os melhores resultados em termos de graduação para os pensionistas do sexo masculino. Em segundo, um modelo linear do tipo  $GM(0, 7)$  (ou do tipo  $LGM(0, 7)$ ) revela-se adequado à graduação da mortalidade registada nos pensionistas do sexo feminino.

Por último, os resultados obtidos com predictor linear do tipo *Logit-Gompertz-Makeham de tipo*  $(r, s)$  são, na população de beneficiários de fundos de pensões, indistinguíveis dos conseguidos com o recurso a predictores lineares do tipo *Gompertz-Makeham de tipo*  $(r, s)$ .

### 5.3. Modelos lineares e não-lineares generalizados

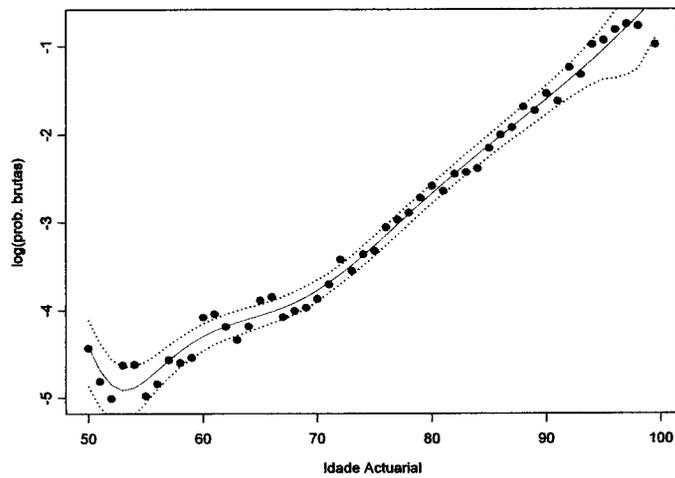


Figura 5.3.5: Modelo  $GM(2,6)$ , pensionistas do sexo masculino

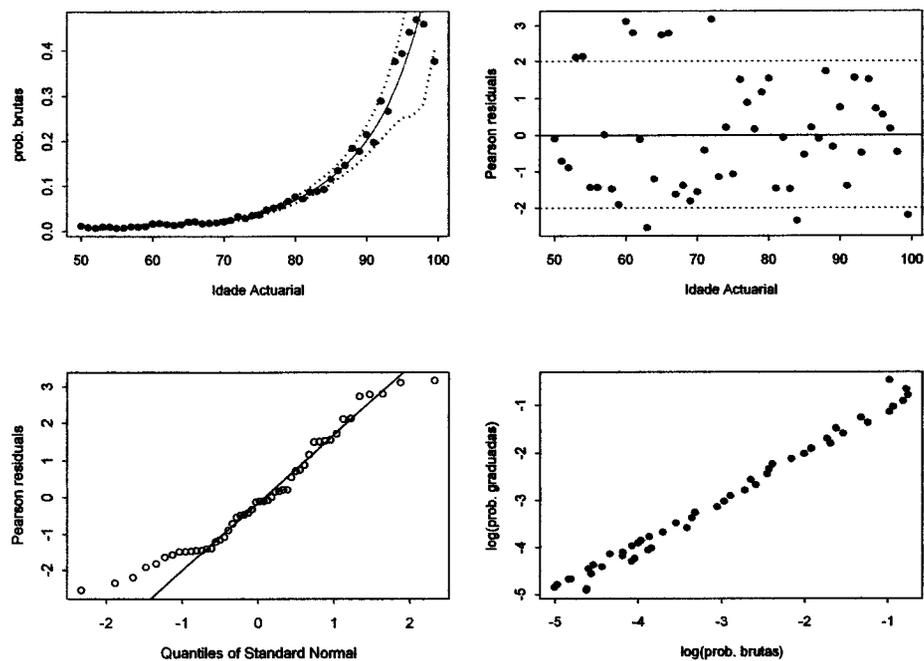


Figura 5.3.6: Resíduos do modelo  $GM(2,6)$ , pensionistas do sexo masculino

### 5.3. Modelos lineares e não-lineares generalizados

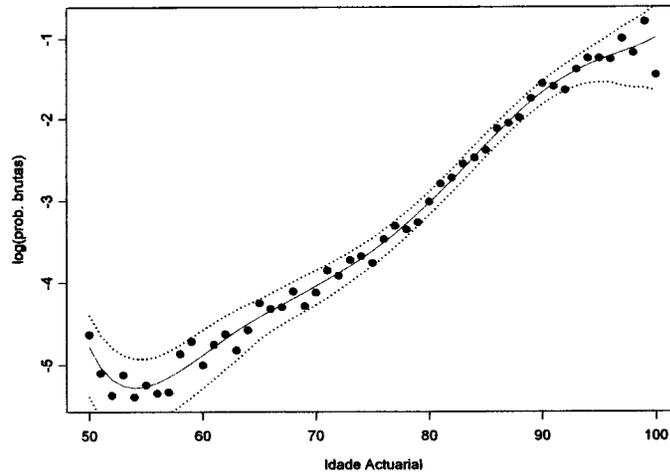


Figura 5.3.7: Modelo  $GM(0,7)$ , pensionistas do sexo feminino

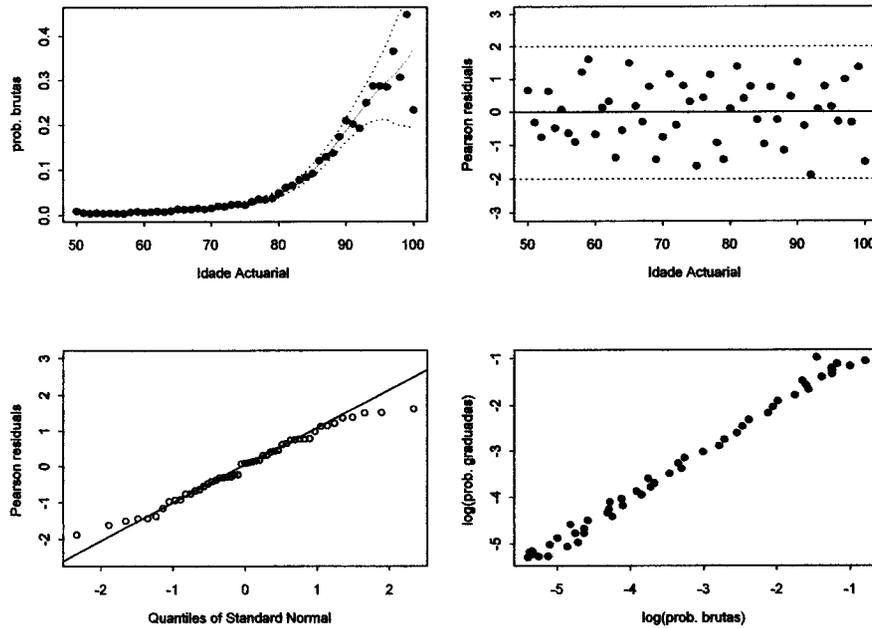


Figura 5.3.8: Resíduos do modelo  $GM(0,7)$ , pensionistas do sexo feminino

### 5.3. Modelos lineares e não-lineares generalizados

Esta última conclusão é confirmada pelos resultados obtidos nos testes efectuados considerando a população de pessoas seguras em caso de vida e em caso de morte. Em consequência, e por uma questão de redundância, daqui em diante optamos por reportar apenas os resultados referentes aos modelos do tipo *Logit-Gompertz-Makeham*, i.e., modelos com predictor linear do tipo  $LGM(r, s)$  e *link* logit. Esta escolha é justificada essencialmente por motivos de conveniência prática na elaboração das tábuas de mortalidade. Por último, verifica-se que os quocientes  $q_x$  respeitantes aos indivíduos do sexo feminino são inferiores aos observados nos elementos do sexo masculino, à semelhança do referido anteriormente para a população em geral.

**Significância estatística do modelo** Para concluir a análise, os Quadros 5.3.11 e 5.3.12 resumem a informação sobre a significância estatística dos modelos óptimos e sobre os resultados dos testes de qualidade do ajustamento.

|            | Homens: $GM(2, 6)$ |       |         |         | Mulheres: $GM(0, 7)$ |       |         |         |
|------------|--------------------|-------|---------|---------|----------------------|-------|---------|---------|
|            | Coef.              | se    | t-ratio | p-value | Coef.                | se    | t-ratio | p-value |
| $\alpha_0$ | -4.496179          | 0.726 | -6.19   | <0.0001 |                      |       |         |         |
| $\alpha_1$ | -2.955564          | 0.432 | -6.84   | <0.0001 |                      |       |         |         |
| $\beta_0$  | 1.5115289          | 0.160 | 9.44    | <0.0001 | -3.606946            | 0.027 | -131.21 | <0.0001 |
| $\beta_1$  | 0.6754037          | 0.026 | 25.82   | <0.0001 | 2.496212             | 0.112 | 22.26   | <0.0001 |
| $\beta_2$  | -0.1872704         | 0.015 | -12.69  | <0.0001 | 2.133815             | 0.393 | 5.42    | <0.0001 |
| $\beta_3$  | 0.0934068          | 0.007 | 13.17   | <0.0001 | 1.087732             | 0.507 | 2.14    | 0.0135  |
| $\beta_4$  | -0.0597252         | 0.011 | -5.53   | <0.0001 | -4.458174            | 1.204 | -3.70   | <0.0001 |
| $\beta_5$  | 0.0557507          | 0.016 | 3.47    | <0.0001 | -1.689576            | 0.486 | -3.48   | <0.0001 |
| $\beta_6$  |                    |       |         |         | 3.048325             | 0.934 | 3.27    | <0.0001 |

Quadro 5.3.11: Significância estatística dos parâmetros,  $GM(r, s)$ , pensionistas

|   | Óbitos ( $A, E$ )  |           | Deviance              |          |           | Teste $\chi^2$ |           |         |
|---|--------------------|-----------|-----------------------|----------|-----------|----------------|-----------|---------|
|   | $(A - E)$          | $A/E(\%)$ | Null                  | Unscaled | p-value   | Unsc. $\chi^2$ | g.l.      | p-value |
| H | 0                  | 100.0     | 7260.22               | 115.70   | 0.434     | 117.47         | 42        | 0.407   |
| M | 0                  | 100.0     | 6072.04               | 42.69    | 0.528     | 42.40          | 44        | 0.540   |
|   | Kolmogorov-Smirnov |           | Signs Test            |          |           |                |           |         |
|   | Desv Max           | p-value   | Desv.(+)              | Desv.(-) | p-value   | $z_x > 2$      | $z_x > 3$ |         |
| H | 0.00725            | 1         | 22                    | 28       | 0.234     | 10             | 2         |         |
| M | 0.00428            | 1         | 27                    | 24       | 0.712     | 0              | 0         |         |
|   | Runs Test          |           | Testes Autocorrelação |          |           |                |           |         |
|   | N.º runs           | p-value   | Portm.                | p-value  | Ljung-Box | p-value        |           |         |
| H | 25                 | 0.968     | 22.11                 | 0.335    | 27.096    | 0.134          |           |         |
| M | 32                 | 0.084     | 32.66                 | 0.037    | 42.187    | 0.003          |           |         |

Quadro 5.3.12: Testes de qualidade do ajustamento,  $GM(r, s)$ , pensionistas

### 5.3. Modelos lineares e não-lineares generalizados

#### 5.3.2 Seguros em caso de vida

Por razões que se prendem com a reduzida exposição ao risco e a significância estatística dos resultados, os dados usados na graduação restringiram-se às idades  $x$  no intervalo  $x \in [0, 100]$ . Nas idades em que o número de óbitos registados no período foi nulo ou insignificante (idades infantis e mais avançadas), procedeu-se ao agrupamento dos dados em idades contíguas até que a condição  $d_x^{\text{exp}} \geq 5$  fosse respeitada.

**Escolha da ordem do modelo** Foram testadas, em ambos os sexos, diferentes combinações considerando  $s \in [2, 7]$  e  $r \in [0, 4]$ . Os valores da função de Log-verosimilhança e da (*unscaled*) deviance são apresentados nos Quadros 5.3.13 e 5.3.14.

| Log-verosimilhança |          |          |          |          |          |          |
|--------------------|----------|----------|----------|----------|----------|----------|
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 106611.8 | 106375.6 | 106333.0 | 106330.1 | 106304.9 | 106301.4 |
| 1                  | 106349.3 | 106332.0 | 106324.3 | 106319.2 | 106304.5 | 106299.7 |
| 2                  | 106348.0 | 106332.0 | 106311.3 | 106302.1 | 106299.9 | 106299.6 |
| 3                  | 106324.8 | 106304.5 | 106302.4 | 106300.7 | 106299.7 | 106297.6 |
| 4                  | 106311.3 | 106304.5 | 106302.4 | 106301.2 | 106297.0 | 106295.0 |

| Deviance |         |         |         |         |         |         |
|----------|---------|---------|---------|---------|---------|---------|
| $r$      | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | $s = 7$ |
| 0        | 781.53  | 309.11  | 223.96  | 218.17  | 167.79  | 160.66  |
| 1        | 256.59  | 221.82  | 206.50  | 196.32  | 166.93  | 157.37  |
| 2        | 253.95  | 221.80  | 180.55  | 162.01  | 157.62  | 157.09  |
| 3        | 207.58  | 166.94  | 162.67  | 159.25  | 157.22  | 153.10  |
| 4        | 180.54  | 166.85  | 162.61  | 160.36  | 151.81  | 147.55  |

Quadro 5.3.13: LV e deviance, seguros em caso de vida, sexo masculino,  $LGM(r, s)$

| Log-verosimilhança |         |         |         |         |         |         |
|--------------------|---------|---------|---------|---------|---------|---------|
| $r$                | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | $s = 7$ |
| 0                  | 40342.3 | 40259.6 | 40258.2 | 40256.8 | 40243.1 | 40242.1 |
| 1                  | 40277.6 | 40258.9 | 40258.1 | 40245.5 | 40237.9 | 40237.9 |
| 2                  | 40259.9 | 40255.4 | 40239.2 | 40239.1 | 40235.0 | 40233.1 |
| 3                  | 40252.6 | 40241.4 | 40239.2 | 40233.9 | 40232.9 | 40233.1 |
| 4                  | 40242.5 | 40241.4 | 40237.0 | 40233.2 | 40232.2 | 40232.0 |

| Deviance |         |         |         |         |         |         |
|----------|---------|---------|---------|---------|---------|---------|
| $r$      | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | $s = 7$ |
| 0        | 331.03  | 165.62  | 162.86  | 159.95  | 132.56  | 130.70  |
| 1        | 201.54  | 164.31  | 162.67  | 137.40  | 122.18  | 122.15  |
| 2        | 166.24  | 157.17  | 124.75  | 124.62  | 116.35  | 112.71  |
| 3        | 151.56  | 129.29  | 124.75  | 114.24  | 110.20  | 112.71  |
| 4        | 131.33  | 129.26  | 120.50  | 112.86  | 110.79  | 110.46  |

Quadro 5.3.14: LV e deviance, seguros em caso de vida, sexo feminino,  $LGM(r, s)$

### 5.3. Modelos lineares e não-lineares generalizados

A análise dos quadros permite identificar os modelos  $LGM(0,7)$  e  $LGM(3,3)$  como candidatos a óptimo no caso das vidas do sexo masculino, enquanto que nas vidas do sexo feminino a selecção recai sobre os modelos  $LGM(1,6)$ ,  $LGM(2,4)$ ,  $LGM(3,3)$  e  $LGM(3,5)$ . Para identificar o modelo mais adequado recorreremos, uma vez mais, à extensa bateria de testes estatísticos mencionada na secção anterior. Os Quadros 5.3.15 e 5.3.16 resumem os resultados obtidos.

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 1777.2   | 0.000       | 0.999           | 0.000                    | 1                | 0.000               | 0.000                | S             | S     | N             |
| (0, 3)   | 381.7    | 0.000       | 0.586           | 0.276                    | 1                | 0.000               | 0.000                | S             | S     | S             |
| (0, 4)   | 227.1    | 0.010       | 0.414           | 0.016                    | 1                | 0.001               | 0.000                | S             | S     | S             |
| (0, 5)   | 216.6    | 0.019       | 0.332           | 0.133                    | 1                | 0.004               | 0.000                | S             | S     | S             |
| (0, 6)   | 167.2    | 0.340       | 0.414           | 0.376                    | 1                | 0.165               | 0.052                | S             | S     | S             |
| (0, 7)   | 160.1    | 0.414       | 0.414           | 0.667                    | 1                | 0.155               | 0.049                | S             | N     | S             |
| (1, 2)   | 285.2    | 0.000       | 0.500           | 0.016                    | 1                | 0.000               | 0.000                | S             | S     | S             |
| (1, 3)   | 218.1    | 0.021       | 0.414           | 0.388                    | 1                | 0.019               | 0.004                | S             | S     | S             |
| (1, 4)   | 205.5    | 0.044       | 0.332           | 0.200                    | 1                | 0.045               | 0.010                | S             | S     | S             |
| (1, 5)   | 220.2    | 0.017       | 0.668           | 0.836                    | 1                | 0.459               | 0.275                | S             | N     | S             |
| (1, 6)   | 166.6    | 0.320       | 0.139           | 0.296                    | 1                | 0.149               | 0.046                | N             | S     | S             |
| (1, 7)   | 157.9    | 0.416       | 0.414           | 0.505                    | 1                | 0.117               | 0.037                | N             | N     | S             |
| (2, 2)   | 273.5    | 0.000       | 0.414           | 0.050                    | 1                | 0.000               | 0.000                | N             | S     | S             |
| (2, 3)   | 218.2    | 0.017       | 0.500           | 0.383                    | 1                | 0.018               | 0.004                | N             | S     | S             |
| (2, 4)   | 179.5    | 0.192       | 0.064           | 0.879                    | 1                | 0.147               | 0.046                | S             | S     | S             |
| (2, 5)   | 163.0    | 0.371       | 0.258           | 0.949                    | 1                | 0.109               | 0.031                | S             | S     | S             |
| (2, 6)   | 158.2    | 0.412       | 0.414           | 0.654                    | 1                | 0.112               | 0.032                | N             | N     | S             |
| (2, 7)   | 157.5    | 0.390       | 0.414           | 0.505                    | 1                | 0.115               | 0.038                | N             | N     | S             |
| (3, 2)   | 206.9    | 0.040       | 0.414           | 0.280                    | 1                | 0.117               | 0.039                | S             | S     | S             |
| (3, 3)   | 168.2    | 0.326       | 0.332           | 0.974                    | 1                | 0.114               | 0.032                | S             | S     | S             |
| (3, 4)   | 163.4    | 0.366       | 0.414           | 0.654                    | 1                | 0.106               | 0.029                | S             | S     | S             |
| (3, 5)   | 160.2    | 0.382       | 0.332           | 0.974                    | 1                | 0.125               | 0.038                | N             | S     | S             |
| (3, 6)   | 157.8    | 0.387       | 0.332           | 0.638                    | 1                | 0.113               | 0.033                | N             | N     | S             |
| (3, 7)   | 154.3    | 0.410       | 0.500           | 0.512                    | 1                | 0.104               | 0.031                | N             | N     | S             |
| (4, 2)   | 179.4    | 0.193       | 0.139           | 0.465                    | 1                | 0.164               | 0.055                | N             | S     | S             |
| (4, 3)   | 168.0    | 0.301       | 0.414           | 0.654                    | 1                | 0.108               | 0.030                | N             | S     | S             |
| (4, 4)   | 163.2    | 0.338       | 0.414           | 0.654                    | 1                | 0.109               | 0.030                | N             | S     | S             |
| (4, 5)   | 161.0    | 0.340       | 0.332           | 0.638                    | 1                | 0.117               | 0.034                | N             | S     | S             |
| (4, 6)   | 152.0    | 0.446       | 0.500           | 0.826                    | 1                | 0.064               | 0.016                | N             | N     | S             |
| (4, 7)   | 147.5    | 0.486       | 0.586           | 0.512                    | 1                | 0.037               | 0.008                | N             | N     | S             |

Quadro 5.3.15: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de vida, sexo masculino

No caso das vidas do sexo masculino, o modelo  $LGM(0,7)$  obtém bons resultados em todos os testes estatísticos mas exhibe um comportamento inadequado fora do intervalo de estimação (gera valores decrescentes para  $q_x$ ), uma característica que o torna pouco adequado para a elaboração de uma tábua de mortalidade realista. Em contrapartida,

### 5.3. Modelos lineares e não-lineares generalizados

o modelo  $LGM(3, 3)$  não só passa com sucesso todos testes, como extrapola um perfil de evolução dos quocientes  $q_x$  nas idades avançadas considerado mais realista. Idênticos argumentos podem ser invocados na população do sexo feminino, com a diferença de que entre os dois modelos que ultrapassam com sucesso todos testes estatísticos, apenas a variante  $LGM(3, 3)$  apresenta um perfil gráfico adequado. A grau de ajustamento entre os quocientes brutos e os quocientes graduados pelos modelos ótimos é posto em evidência nas Figuras 5.3.9 e 5.3.11.

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 415.0    | 0.000       | 0.938           | 0.000                    | 0.999            | 0.000               | 0.000                | S             | S     | S             |
| (0, 3)   | 169.5    | 0.096       | 0.453           | 0.288                    | 1                | 0.358               | 0.174                | S             | S     | S             |
| (0, 4)   | 169.0    | 0.085       | 0.205           | 0.343                    | 1                | 0.471               | 0.260                | S             | S     | S             |
| (0, 5)   | 172.0    | 0.059       | 0.144           | 0.595                    | 1                | 0.548               | 0.324                | S             | S     | S             |
| (0, 6)   | 136.3    | 0.404       | 0.205           | 0.343                    | 1                | 0.538               | 0.313                | N             | S     | S             |
| (0, 7)   | 133.1    | 0.423       | 0.097           | 0.272                    | 1                | 0.534               | 0.313                | N             | S     | S             |
| (1, 2)   | 205.9    | 0.005       | 0.278           | 0.000                    | 1                | 0.000               | 0.000                | S             | S     | S             |
| (1, 3)   | 168.9    | 0.086       | 0.362           | 0.569                    | 1                | 0.385               | 0.194                | N             | S     | S             |
| (1, 4)   | 168.3    | 0.076       | 0.205           | 0.343                    | 1                | 0.465               | 0.255                | N             | S     | S             |
| (1, 5)   | 144.3    | 0.282       | 0.144           | 0.310                    | 1                | 0.545               | 0.312                | S             | S     | S             |
| (1, 6)   | 122.9    | 0.603       | 0.278           | 0.169                    | 1                | 0.546               | 0.335                | S             | N     | S             |
| (1, 7)   | 122.9    | 0.567       | 0.278           | 0.169                    | 1                | 0.549               | 0.338                | N             | S     | S             |
| (2, 2)   | 178.0    | 0.045       | 0.278           | 0.952                    | 1                | 0.513               | 0.298                | S             | S     | S             |
| (2, 3)   | 159.4    | 0.138       | 0.205           | 0.640                    | 1                | 0.541               | 0.324                | N             | S     | S             |
| (2, 4)   | 125.4    | 0.592       | 0.144           | 0.310                    | 1                | 0.546               | 0.334                | S             | N     | S             |
| (2, 5)   | 125.2    | 0.561       | 0.144           | 0.310                    | 1                | 0.548               | 0.335                | N             | S     | S             |
| (2, 6)   | 116.8    | 0.674       | 0.278           | 0.169                    | 1                | 0.333               | 0.156                | N             | S     | S             |
| (2, 7)   | 113.1    | 0.704       | 0.278           | 0.591                    | 1                | 0.265               | 0.120                | N             | N     | S             |
| (3, 2)   | 155.7    | 0.172       | 0.097           | 0.112                    | 1                | 0.561               | 0.338                | S             | S     | S             |
| (3, 3)   | 128.4    | 0.539       | 0.362           | 0.181                    | 1                | 0.558               | 0.347                | S             | S     | S             |
| (3, 4)   | 125.4    | 0.559       | 0.144           | 0.310                    | 1                | 0.544               | 0.332                | N             | S     | S             |
| (3, 5)   | 114.1    | 0.719       | 0.362           | 0.701                    | 1                | 0.272               | 0.122                | S             | N     | S             |
| (3, 6)   | 112.4    | 0.715       | 0.278           | 0.676                    | 1                | 0.274               | 0.121                | N             | N     | S             |
| (3, 7)   | 113.1    | 0.671       | 0.278           | 0.591                    | 1                | 0.266               | 0.120                | N             | N     | S             |
| (4, 2)   | 132.0    | 0.477       | 0.278           | 0.676                    | 1                | 0.561               | 0.345                | N             | S     | S             |
| (4, 3)   | 128.3    | 0.507       | 0.278           | 0.169                    | 1                | 0.556               | 0.346                | N             | S     | S             |
| (4, 4)   | 121.5    | 0.593       | 0.205           | 0.343                    | 1                | 0.377               | 0.183                | S             | S     | S             |
| (4, 5)   | 113.4    | 0.699       | 0.362           | 0.926                    | 1                | 0.252               | 0.112                | S             | N     | S             |
| (4, 6)   | 111.6    | 0.697       | 0.362           | 0.701                    | 1                | 0.276               | 0.117                | N             | N     | S             |
| (4, 7)   | 111.3    | 0.669       | 0.547           | 0.192                    | 1                | 0.265               | 0.111                | N             | N     | S             |

Quadro 5.3.16: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de vida, sexo feminino

Nas Figuras 5.3.10 e 5.3.12 completamos o diagnóstico dos modelos com a análise do comportamento dos resíduos. Como se observa, em ambos os sexos a maioria dos resíduos é de valor reduzido e exhibe um padrão condizente com a distribuição Normal.

### 5.3. Modelos lineares e não-lineares generalizados

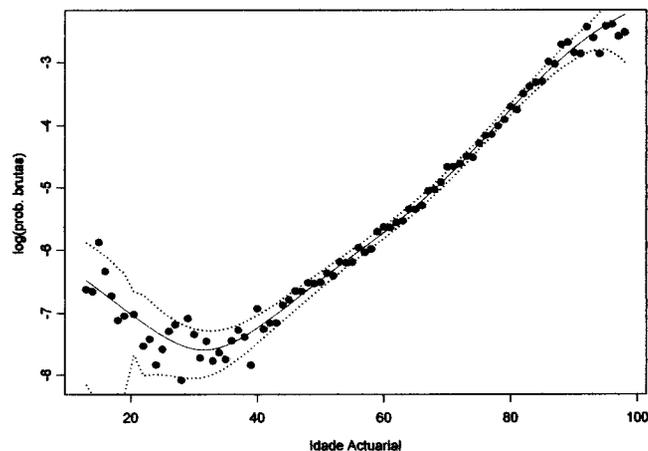


Figura 5.3.9: Modelo  $LGM(3, 3)$ , seguros em caso de vida, homens

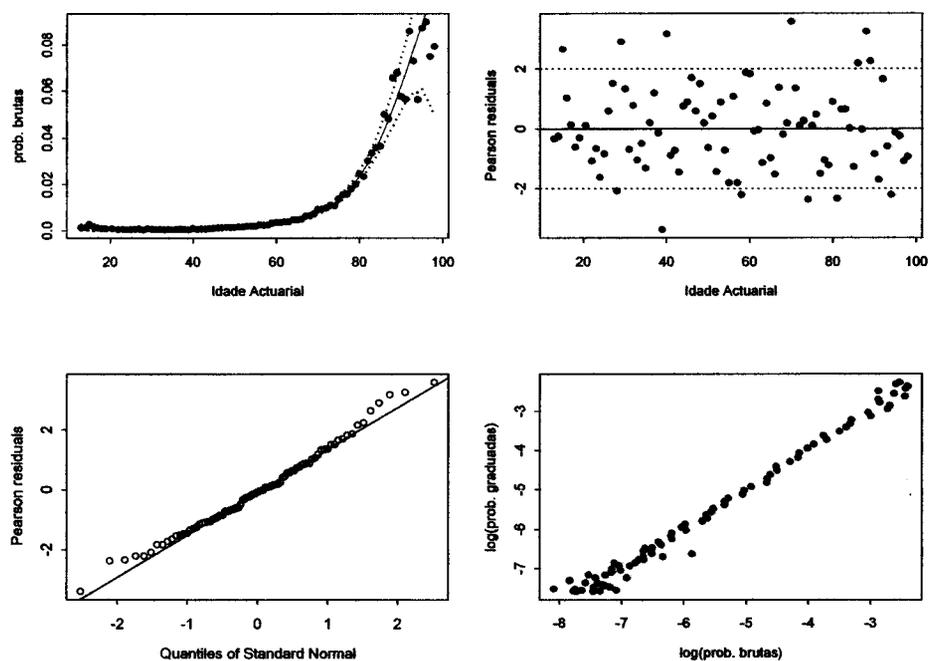


Figura 5.3.10: Resíduos do modelo  $LGM(3, 3)$ , seguros em caso de vida, homens

### 5.3. Modelos lineares e não-lineares generalizados

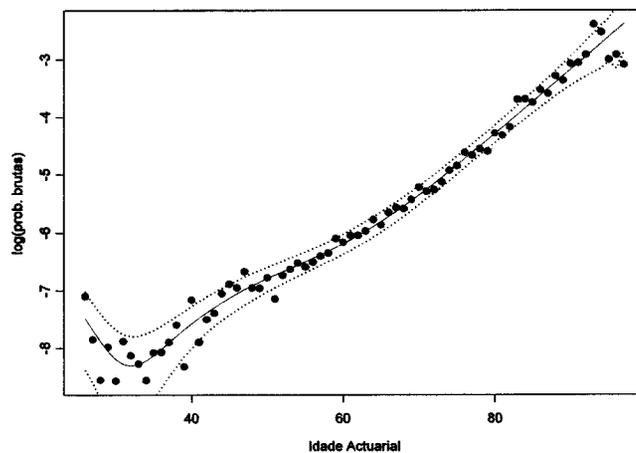


Figura 5.3.11: Modelo  $LGM(3, 3)$ , seguros em caso de vida, mulheres

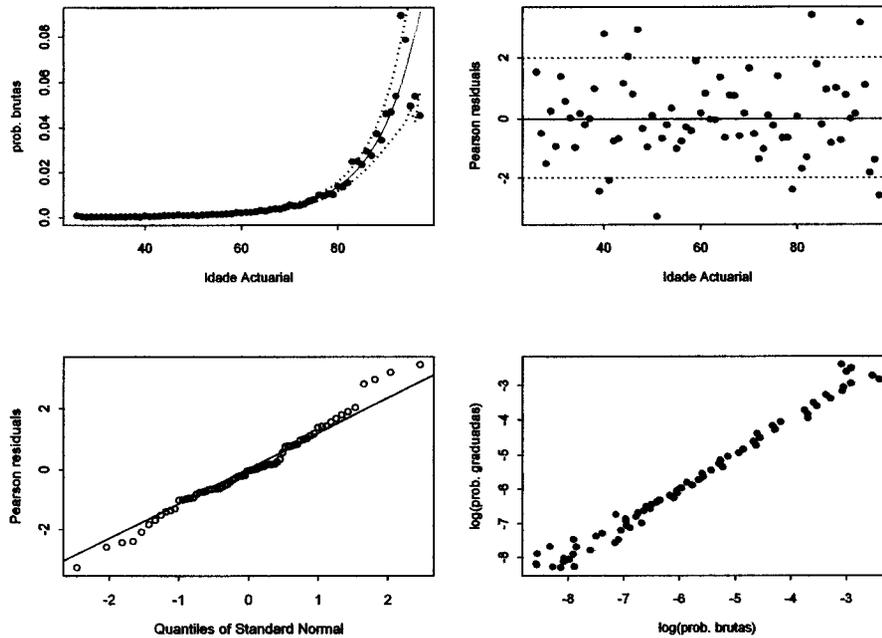


Figura 5.3.12: Resíduos do modelo  $LGM(3, 3)$ , seguros em caso de vida, mulheres

### 5.3. Modelos lineares e não-lineares generalizados

**Significância estatística do modelo** Os Quadros 5.3.17 e 5.3.18 concluem a análise condensando informação sobre as estimativas dos parâmetros dos modelos ótimos e sobre os resultados dos testes de qualidade do ajustamento.

|            | Homens: $LGM(3, 3)$ |        |            |            | Mulheres: $LGM(3, 3)$ |       |            |            |
|------------|---------------------|--------|------------|------------|-----------------------|-------|------------|------------|
|            | Coef.               | se     | $t$ -ratio | $p$ -value | Coef.                 | se    | $t$ -ratio | $p$ -value |
| $\alpha_0$ | 0.00223264          | 0.0001 | 36.16      | <0.0001    | -0.01105485           | 0.005 | -2.23      | 0.0084     |
| $\alpha_1$ | 0.00613485          | 0.0003 | 19.44      | <0.0001    | -0.00863382           | 0.003 | -2.65      | 0.0008     |
| $\alpha_2$ | 0.00544452          | 0.0004 | 14.59      | <0.0001    | -0.01192961           | 0.004 | -3.13      | <0.0001    |
| $\beta_0$  | -9.0250236          | 0.3088 | -29.23     | <0.0001    | -4.3163919            | 0.370 | -11.66     | <0.0001    |
| $\beta_1$  | 10.9826454          | 0.9012 | 12.19      | <0.0001    | 1.0901205             | 0.165 | 6.59       | <0.0001    |
| $\beta_2$  | -4.216191           | 0.6787 | -6.21      | <0.0001    | 1.2012087             | 0.168 | 7.13       | <0.0001    |

Quadro 5.3.17: Significância estatística dos parâmetros,  $LGM(r, s)$ , seguros em caso de vida

|   | Óbitos ( $A, E$ )  |            | Deviance              |            |            | Teste $\chi^2$ |           |            |
|---|--------------------|------------|-----------------------|------------|------------|----------------|-----------|------------|
|   | $(A - E)$          | $A/E(\%)$  | Null                  | Unscaled   | $p$ -value | Unsc. $\chi^2$ | g.l.      | $p$ -value |
| H | 0                  | 100.0      | 24424.62              | 166.94     | 0.344      | 168.23         | 79        | 0.326      |
| M | 0                  | 100.0      | 9006.29               | 129.29     | 0.524      | 128.43         | 66        | 0.539      |
|   | Kolmogorov-Smirnov |            | Signs Test            |            |            |                |           |            |
|   | Desv Max           | $p$ -value | Desv.(+)              | Desv.(-)   | $p$ -value | $z_x > 2$      | $z_x > 3$ |            |
| H | 0.00478            | 1          | 40                    | 45         | 0.332      | 13             | 4         |            |
| M | 0.01033            | 1          | 34                    | 38         | 0.362      | 10             | 3         |            |
|   | Runs Test          |            | Testes Autocorrelação |            |            |                |           |            |
|   | N.º runs           | $p$ -value | Portm.                | $p$ -value | Ljung-Box  | $p$ -value     |           |            |
| H | 43                 | 0.974      | 27.80                 | 0.114      | 33.19      | 0.032          |           |            |
| M | 42                 | 0.182      | 18.44                 | 0.558      | 21.87      | 0.347          |           |            |

Quadro 5.3.18: Testes de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de vida

#### 5.3.3 Seguros em caso de morte

Os dados usados na graduação da mortalidade da população segura em caso de morte restringem-se às idades compreendidas entre os 0 e os 100 anos. Esta opção é fundamentada, uma vez mais, pelo reduzido número de vidas expostas ao risco e pela necessidade de garantir a significância estatística dos resultados. Foi igualmente adoptado o procedimento de agrupamento dos dados relativos a idades contíguas (com a condição  $d_x^{\text{exp}} \geq 5$ ) nos casos em que o número de óbitos registados no período 2000-2004 foi nulo ou insuficiente.

### 5.3. Modelos lineares e não-lineares generalizados

**Escolha da ordem do modelo** Foram testadas, em ambos os sexos, um total de 30 combinações diferentes, fazendo variar os parâmetros  $s$  e  $r$  nos intervalos  $s \in [2, 7]$  e  $r \in [0, 4]$ . Os valores da função de Log-verosimilhança e da (*unscaled*) deviance são reportados nos Quadros 5.3.19 e 5.3.20.

| Log-verosimilhança |          |          |          |          |          |          |
|--------------------|----------|----------|----------|----------|----------|----------|
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 139699.9 | 139686.7 | 139528.3 | 139435.2 | 139434.0 | 139430.6 |
| 1                  | 139697.5 | 139494.6 | 139460.5 | 139435.1 | 139432.9 | 139430.0 |
| 2                  | 139460.2 | 139452.9 | 139449.4 | 139434.0 | 139430.4 | 139430.4 |
| 3                  | 139460.3 | 139433.3 | 139430.2 | 139428.6 | 139425.6 | 139413.1 |
| 4                  | 139460.4 | 139430.7 | 139429.2 | 139421.7 | 139413.5 | 139411.0 |
| Deviance           |          |          |          |          |          |          |
| $r$                | $s = 2$  | $s = 3$  | $s = 4$  | $s = 5$  | $s = 6$  | $s = 7$  |
| 0                  | 697.47   | 670.96   | 354.19   | 168.07   | 165.50   | 158.73   |
| 1                  | 692.66   | 286.70   | 218.65   | 167.70   | 163.29   | 157.68   |
| 2                  | 218.07   | 203.43   | 196.34   | 165.64   | 158.31   | 158.29   |
| 3                  | 218.16   | 164.14   | 157.95   | 154.70   | 148.73   | 123.81   |
| 4                  | 218.30   | 158.94   | 156.07   | 141.05   | 124.50   | 119.60   |

Quadro 5.3.19: LV e deviance, seguros em caso de morte, sexo masculino,  $LGM(r, s)$

| Log-verosimilhança |         |         |         |         |         |         |
|--------------------|---------|---------|---------|---------|---------|---------|
| $r$                | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | $s = 7$ |
| 0                  | 38325.1 | 38324.5 | 38324.4 | 38320.7 | 38315.1 | 38311.5 |
| 1                  | 38324.3 | 38324.2 | 38318.3 | 38312.7 | 38314.3 | 38305.6 |
| 2                  | 38322.9 | 38314.7 | 38314.6 | 38313.9 | 38312.8 | 38305.4 |
| 3                  | 38315.6 | 38313.8 | 38309.7 | 38309.5 | 38305.6 | 38305.3 |
| 4                  | 38313.6 | 38312.7 | 38308.8 | 38308.6 | 38305.4 | 38305.2 |
| Deviance           |         |         |         |         |         |         |
| $r$                | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | $s = 7$ |
| 0                  | 110.69  | 109.35  | 109.31  | 101.77  | 90.66   | 83.34   |
| 1                  | 108.93  | 108.90  | 96.97   | 85.78   | 89.01   | 71.69   |
| 2                  | 106.17  | 89.88   | 89.58   | 88.14   | 85.95   | 71.28   |
| 3                  | 91.52   | 88.08   | 79.85   | 79.34   | 71.67   | 70.97   |
| 4                  | 87.59   | 85.80   | 78.05   | 77.64   | 71.17   | 70.76   |

Quadro 5.3.20: LV e Deviance, seguros em caso de morte, sexo feminino,  $LGM(r, s)$

A análise dos quadros permitiu identificar os modelos  $LGM(3, 4)$ ,  $LGM(3, 7)$  e  $LGM(4, 5)$  como os mais sérios candidatos a óptimo no caso das vidas do sexo masculino. No caso das vidas do sexo feminino, a selecção recaiu sobre os modelos  $LGM(3, 2)$  e  $LGM(3, 4)$ . Para escolher o modelo óptimo, considerámos uma vez mais os resulta-

### 5.3. Modelos lineares e não-lineares generalizados

dos obtidos nos testes de qualidade do ajustamento, apresentados nos Quadros 5.3.21 e 5.3.22.

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 723.3    | 0.000       | 0.452           | 0.000                    | 1                | 0.000               | 0.000                | S             | N     | N             |
| (0, 3)   | 743.3    | 0.000       | 0.137           | 0.000                    | 1                | 0.000               | 0.000                | S             | N     | N             |
| (0, 4)   | 416.6    | 0.000       | 0.358           | 0.021                    | 1                | 0.002               | 0.000                | S             | N     | S             |
| (0, 5)   | 168.0    | 0.335       | 0.452           | 0.908                    | 1                | 0.050               | 0.008                | S             | S     | S             |
| (0, 6)   | 165.9    | 0.329       | 0.452           | 0.708                    | 1                | 0.037               | 0.005                | N             | S     | S             |
| (0, 7)   | 159.0    | 0.385       | 0.272           | 0.761                    | 1                | 0.021               | 0.002                | S             | S     | S             |
| (1, 2)   | 700.2    | 0.000       | 0.548           | 0.000                    | 1                | 0.000               | 0.000                | S             | N     | N             |
| (1, 3)   | 317.9    | 0.000       | 0.642           | 0.181                    | 1                | 0.012               | 0.002                | S             | N     | S             |
| (1, 4)   | 220.6    | 0.020       | 0.452           | 0.546                    | 1                | 0.095               | 0.025                | S             | N     | S             |
| (1, 5)   | 167.8    | 0.306       | 0.548           | 0.903                    | 1                | 0.040               | 0.006                | N             | S     | S             |
| (1, 6)   | 163.7    | 0.324       | 0.358           | 0.692                    | 1                | 0.056               | 0.009                | N             | S     | S             |
| (1, 7)   | 158.0    | 0.364       | 0.272           | 0.761                    | 1                | 0.030               | 0.003                | N             | S     | S             |
| (2, 2)   | 213.2    | 0.039       | 0.198           | 0.135                    | 1                | 0.032               | 0.007                | S             | S     | S             |
| (2, 3)   | 206.1    | 0.051       | 0.452           | 0.719                    | 1                | 0.141               | 0.040                | S             | S     | S             |
| (2, 4)   | 197.5    | 0.072       | 0.358           | 0.187                    | 1                | 0.052               | 0.010                | N             | S     | S             |
| (2, 5)   | 165.5    | 0.302       | 0.358           | 0.692                    | 1                | 0.017               | 0.002                | N             | S     | S             |
| (2, 6)   | 158.7    | 0.355       | 0.452           | 0.908                    | 1                | 0.046               | 0.006                | N             | S     | S             |
| (2, 7)   | 158.6    | 0.322       | 0.548           | 0.903                    | 1                | 0.044               | 0.006                | N             | S     | S             |
| (3, 2)   | 214.3    | 0.030       | 0.137           | 0.150                    | 1                | 0.023               | 0.004                | N             | S     | S             |
| (3, 3)   | 161.2    | 0.389       | 0.137           | 0.347                    | 1                | 0.716               | 0.488                | S             | N     | S             |
| (3, 4)   | 153.0    | 0.468       | 0.198           | 0.318                    | 1                | 0.660               | 0.406                | S             | N     | S             |
| (3, 5)   | 150.1    | 0.474       | 0.198           | 0.318                    | 1                | 0.678               | 0.427                | N             | S     | S             |
| (3, 6)   | 147.2    | 0.480       | 0.452           | 0.719                    | 1                | 0.255               | 0.082                | N             | S     | S             |
| (3, 7)   | 123.3    | 0.785       | 0.452           | 0.388                    | 1                | 0.320               | 0.135                | S             | N     | S             |
| (4, 2)   | 213.5    | 0.026       | 0.198           | 0.135                    | 1                | 0.032               | 0.007                | N             | S     | S             |
| (4, 3)   | 159.9    | 0.373       | 0.272           | 0.582                    | 1                | 0.620               | 0.384                | S             | N     | S             |
| (4, 4)   | 157.5    | 0.371       | 0.272           | 0.582                    | 1                | 0.578               | 0.340                | S             | N     | S             |
| (4, 5)   | 139.5    | 0.595       | 0.272           | 0.582                    | 1                | 0.526               | 0.281                | N             | S     | S             |
| (4, 6)   | 124.4    | 0.771       | 0.452           | 0.708                    | 1                | 0.288               | 0.114                | N             | N     | S             |
| (4, 7)   | 119.9    | 0.798       | 0.452           | 0.708                    | 1                | 0.143               | 0.040                | N             | N     | S             |

Quadro 5.3.21: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de morte, sexo masculino

No caso das vidas do sexo masculino, o modelo  $LGM(3, 7)$  obtém bons resultados em todos os testes estatísticos mas apresenta um desempenho inadequado fora do intervalo de estimação. Por seu turno, o modelo  $LGM(4, 5)$  apresenta parâmetros não significativos e um perfil gráfico desajustado. O modelo  $LGM(3, 4)$  ultrapassa todos testes estatísticos com sucesso, mas gera uma configuração errática nas idades adultas, motivada é claro pelos já aludidos problemas de base da amostra. Em consequência

### 5.3. Modelos lineares e não-lineares generalizados

destes problemas, optámos pela selecção de um modelo com pior desempenho em termos de ajustamento (medido pela estatística deviance), mas que gerasse séries graduadas plausíveis em todas as idades. A escolha recaiu sobre o modelo  $LGM(2, 3)$ .

| $(r, s)$ | $\chi^2$ | $p(\chi^2)$ | Signs<br>$p(+)$ | Runs<br>$p(\text{runs})$ | KS<br>$p(\cdot)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ | Par.<br>Sign. | Conf. | Int.<br>Conf. |
|----------|----------|-------------|-----------------|--------------------------|------------------|---------------------|----------------------|---------------|-------|---------------|
| (0, 2)   | 116.1    | 0.122       | 0.549           | 0.035                    | 1                | 0.033               | 0.004                | S             | S     | N             |
| (0, 3)   | 112.6    | 0.142       | 0.269           | 0.119                    | 1                | 0.014               | 0.001                | N             | S     | S             |
| (0, 4)   | 112.6    | 0.123       | 0.356           | 0.037                    | 1                | 0.016               | 0.001                | N             | S     | S             |
| (0, 5)   | 105.4    | 0.196       | 0.731           | 0.731                    | 1                | 0.138               | 0.030                | N             | S     | S             |
| (0, 6)   | 94.5     | 0.371       | 0.731           | 0.400                    | 1                | 0.976               | 0.926                | S             | S     | S             |
| (0, 7)   | 87.7     | 0.494       | 0.644           | 0.895                    | 1                | 0.978               | 0.931                | S             | S     | S             |
| (1, 2)   | 112.0    | 0.150       | 0.549           | 0.107                    | 1                | 0.014               | 0.001                | N             | S     | S             |
| (1, 3)   | 112.0    | 0.130       | 0.549           | 0.107                    | 1                | 0.014               | 0.001                | N             | S     | S             |
| (1, 4)   | 100.9    | 0.271       | 0.644           | 0.715                    | 1                | 0.576               | 0.304                | S             | S     | S             |
| (1, 5)   | 89.5     | 0.488       | 0.549           | 0.172                    | 1                | 0.972               | 0.925                | S             | S     | S             |
| (1, 6)   | 93.2     | 0.365       | 0.731           | 0.400                    | 1                | 0.979               | 0.934                | N             | S     | S             |
| (1, 7)   | 74.4     | 0.776       | 0.134           | 0.948                    | 1                | 0.829               | 0.672                | N             | S     | S             |
| (2, 2)   | 110.3    | 0.150       | 0.644           | 0.108                    | 1                | 0.061               | 0.009                | N             | S     | S             |
| (2, 3)   | 94.2     | 0.412       | 0.731           | 0.731                    | 1                | 0.966               | 0.897                | N             | S     | S             |
| (2, 4)   | 93.8     | 0.386       | 0.731           | 0.731                    | 1                | 0.969               | 0.905                | N             | S     | S             |
| (2, 5)   | 92.1     | 0.389       | 0.731           | 0.400                    | 1                | 0.981               | 0.940                | N             | S     | S             |
| (2, 6)   | 90.0     | 0.403       | 0.731           | 0.400                    | 1                | 0.984               | 0.951                | N             | N     | S             |
| (2, 7)   | 74.0     | 0.755       | 0.134           | 0.948                    | 1                | 0.815               | 0.653                | N             | N     | S             |
| (3, 2)   | 95.3     | 0.388       | 0.731           | 0.731                    | 1                | 0.894               | 0.734                | S             | S     | S             |
| (3, 3)   | 91.9     | 0.431       | 0.644           | 0.389                    | 1                | 0.966               | 0.901                | S             | S     | S             |
| (3, 4)   | 83.0     | 0.612       | 0.549           | 0.385                    | 1                | 0.967               | 0.907                | S             | N     | S             |
| (3, 5)   | 82.6     | 0.585       | 0.356           | 0.925                    | 1                | 0.955               | 0.875                | N             | N     | S             |
| (3, 6)   | 74.4     | 0.745       | 0.134           | 0.948                    | 1                | 0.821               | 0.661                | N             | N     | S             |
| (3, 7)   | 73.3     | 0.740       | 0.195           | 0.617                    | 1                | 0.817               | 0.657                | N             | N     | S             |
| (4, 2)   | 91.8     | 0.433       | 0.805           | 0.419                    | 1                | 0.981               | 0.939                | N             | S     | S             |
| (4, 3)   | 90.2     | 0.434       | 0.805           | 0.759                    | 1                | 0.975               | 0.925                | N             | S     | S             |
| (4, 4)   | 81.7     | 0.608       | 0.644           | 0.715                    | 1                | 0.957               | 0.886                | N             | N     | S             |
| (4, 5)   | 81.4     | 0.579       | 0.356           | 0.686                    | 1                | 0.946               | 0.858                | N             | N     | S             |
| (4, 6)   | 73.2     | 0.741       | 0.195           | 0.617                    | 1                | 0.848               | 0.701                | N             | N     | S             |
| (4, 7)   | 73.0     | 0.713       | 0.270           | 0.657                    | 1                | 0.839               | 0.690                | N             | N     | S             |

Quadro 5.3.22: Indicadores de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de morte, sexo feminino

Nas vidas do sexo feminino, o modelo  $LGM(3, 2)$  apresenta o desempenho mais adequado entre os candidatos a óptimo.

A qualidade de ajustamento conseguida pelos modelos óptimos pode ser visualizada nas Figuras 5.3.13 e 5.3.15. Nas Figuras 5.3.14 e 5.3.16 completamos o diagnóstico dos modelos com informação sobre o comportamento dos resíduos.

### 5.3. Modelos lineares e não-lineares generalizados

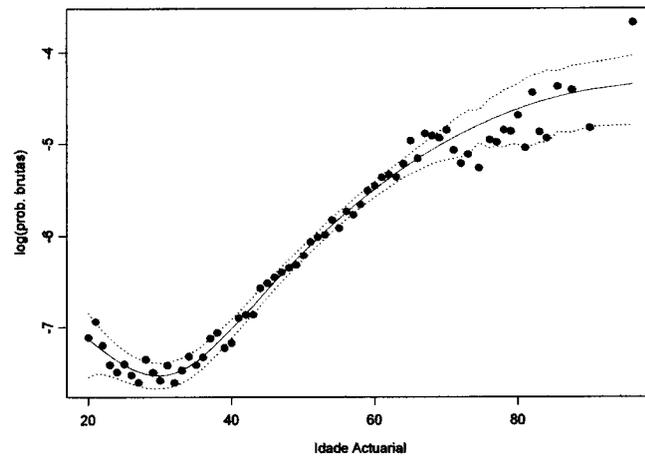


Figura 5.3.13: Modelo  $LGM(2,3)$ , seguros em caso de morte, homens

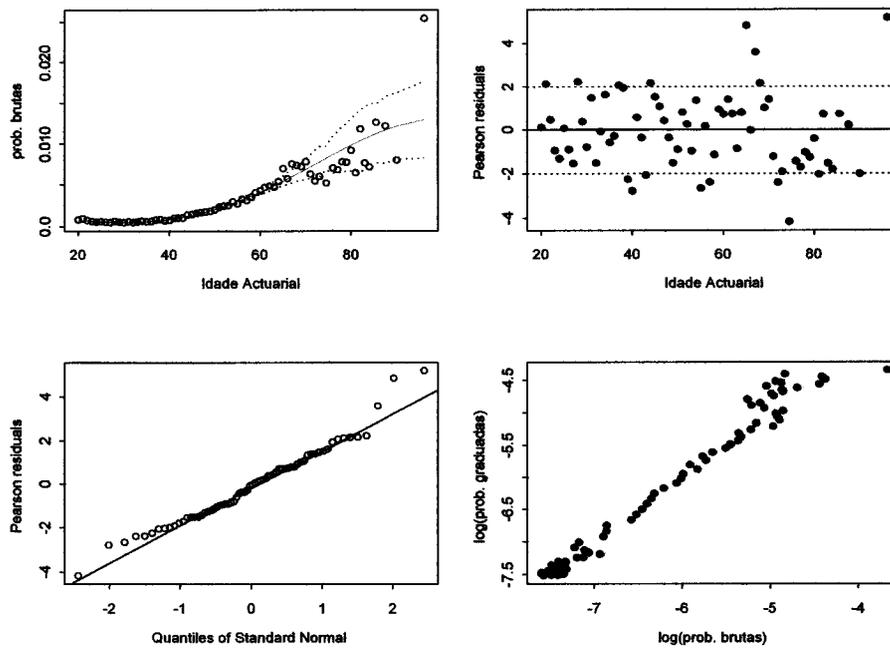


Figura 5.3.14: Resíduos do modelo  $LGM(2,3)$ , seguros em caso de morte, homens

### 5.3. Modelos lineares e não-lineares generalizados

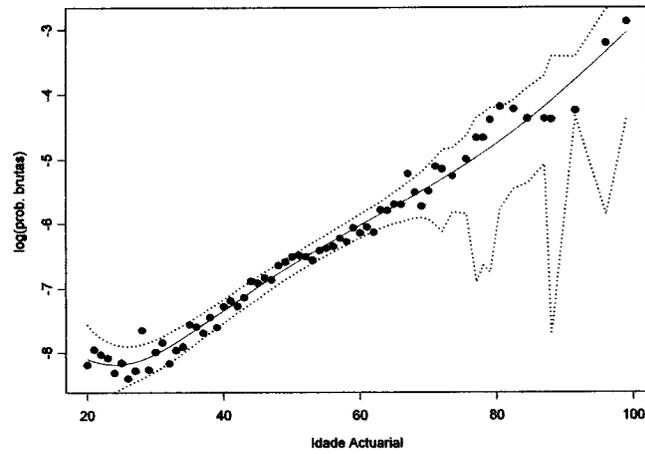


Figura 5.3.15: Modelo  $LGM(3,2)$ , seguros em caso de morte, mulheres

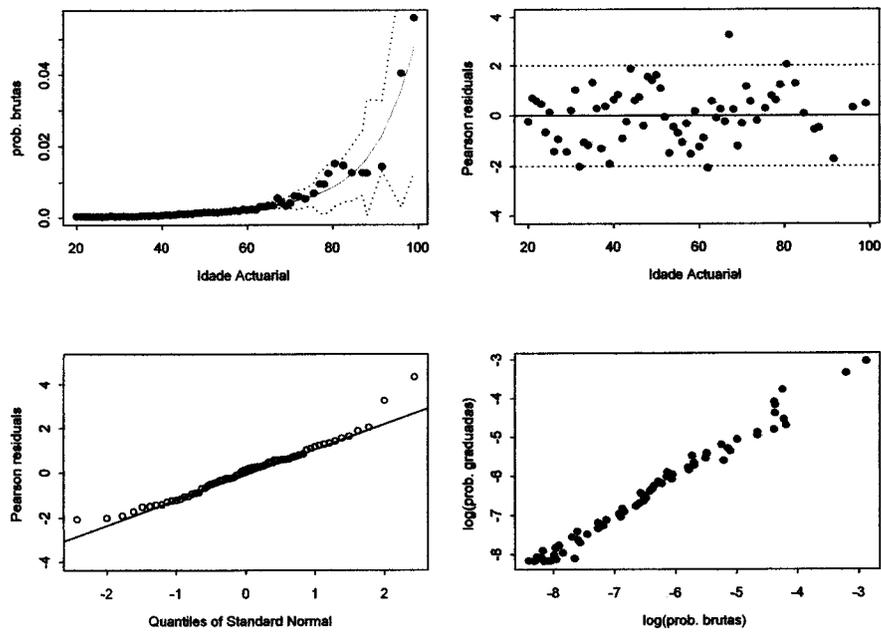


Figura 5.3.16: Resíduos do modelo  $LGM(3,2)$ , seguros em caso de morte, mulheres

## 5.4. GLM's: Abordagem Dual

**Significância estatística do modelo** Os Quadros 5.3.23 e 5.3.24 concluem a análise apresentando informação detalhada sobre as estimativas dos parâmetros dos modelos ótimos e sobre os resultados dos testes de qualidade do ajustamento.

|            | Homens: $LGM(2, 3)$ |        |         |         | Mulheres: $LGM(3, 2)$ |        |         |         |
|------------|---------------------|--------|---------|---------|-----------------------|--------|---------|---------|
|            | Coef.               | se     | t-ratio | p-value | Coef.                 | se     | t-ratio | p-value |
| $\alpha_0$ | -0.0039084          | 0.0015 | -2.63   | 0.0009  | 0.00248805            | 0.0003 | 9.86    | <0.0001 |
| $\alpha_1$ | 0.0040061           | 0.0011 | -3.57   | <0.0001 | 0.00379351            | 0.0004 | 8.81    | <0.0001 |
| $\alpha_2$ |                     |        |         |         | 0.00189214            | 0.0003 | 6.97    | <0.0001 |
| $\beta_0$  | -5.107605           | 0.2526 | -20.22  | <0.0001 | -8.6570002            | 0.6246 | -13.86  | <0.0001 |
| $\beta_1$  | 1.6985577           | 0.2029 | 8.37    | <0.0001 | 5.4868489             | 0.7770 | 7.06    | <0.0001 |
| $\beta_2$  | -0.4532155          | 0.1195 | -3.79   | <0.0001 |                       |        |         |         |

Quadro 5.3.23: Significância estatística dos parâmetros,  $LGM(r, s)$ , seguros em caso de morte

|   | Óbitos ( $A, E$ )  |           | Deviance              |          |           | Teste $\chi^2$ |           |         |
|---|--------------------|-----------|-----------------------|----------|-----------|----------------|-----------|---------|
|   | $(A - E)$          | $A/E(\%)$ | Null                  | Unscaled | p-value   | Unsc. $\chi^2$ | g.l.      | p-value |
| H | 0                  | 100.0     | 11984.83              | 203.4265 | 0.003     | 206.10         | 63        | 0.0024  |
| M | 0                  | 100.0     | 3092.429              | 91.52    | 0.476     | 95.26          | 61        | 0.388   |
|   | Kolmogorov-Smirnov |           | Signs Test            |          |           |                |           |         |
|   | Desv Max           | p-value   | Desv.(+)              | Desv.(-) | p-value   | $z_x > 2$      | $z_x > 3$ |         |
| H | 0.00716            | 1         | 33                    | 35       | 0.452     | 17             | 4         |         |
| M | 0.0152             | 1         | 35                    | 31       | 0.731     | 5              | 2         |         |
|   | Runs Test          |           | Testes Autocorrelação |          |           |                |           |         |
|   | N.º runs           | p-value   | Portm.                | p-value  | Ljung-Box | p-value        |           |         |
| H | 33                 | 0.719     | 26.81                 | 0.141    | 32.36     | 0.0396         |           |         |
| M | 32                 | 0.731     | 12.59                 | 0.894    | 15.71     | 0.734          |           |         |

Quadro 5.3.24: Testes de qualidade do ajustamento,  $LGM(r, s)$ , seguros em caso de morte

## 5.4 GLM's: Abordagem Dual

Nesta secção, ilustramos a aplicação da abordagem dual descrita no Capítulo 3 na graduação das estimativas brutas de  $\mu_x$  relativas à população coberta por seguros de vida em caso de vida. O ajustamento proporcionado pelo modelo Gamma, onde se considera como variável de resposta os expostos centralmente ao risco e se estima a força de vitalidade, a inversa da força de mortalidade, é então comparado com o gerado pelo tradicional modelo Poisson assente em previsores lineares do tipo Gompertz-Makeham.

Em ambos os sexos, foram testados diferentes modelos fazendo variar  $s$  no intervalo  $2 \leq s \leq 10$ . As estimativas dos parâmetros e os resultados dos testes de qualidade de

#### 5.4. GLM's: Abordagem Dual

ajustamento foram obtidos com recurso a uma rotina informática escrita e implementada no *package* S-PLUS 2000 (Release 3).

Os valores da (*unscaled*) deviance, da estatística  $\chi^2$  e do ratio entre a exposição ao risco (mortalidade) real e a exposição (mortalidade) estimada - ratio  $A/E(\%)$ , para a abordagem dual (Poisson), são apresentados nos Quadros 5.4.1 e 5.4.2.

| <i>GM</i><br>(0, <i>s</i> ) | Gamma    |          |           | Poisson  |          |           |
|-----------------------------|----------|----------|-----------|----------|----------|-----------|
|                             | Deviance | $\chi^2$ | $A/E(\%)$ | Deviance | $\chi^2$ | $A/E(\%)$ |
| Null                        | 24423.94 | -        | -         | 24423.94 | -        | -         |
| (0, 2)                      | 766.98   | 480.65   | 77.93     | 766.98   | 1747.27  | 100.00    |
| (0, 3)                      | 319.89   | 322.61   | 99.42     | 319.89   | 397.17   | 100.00    |
| (0, 4)                      | 223.99   | 241.18   | 100.69    | 223.99   | 226.71   | 100.00    |
| (0, 5)                      | 219.44   | 239.49   | 100.49    | 219.44   | 217.86   | 100.00    |
| (0, 6)                      | 167.49   | 178.20   | 100.03    | 167.49   | 166.91   | 100.00    |
| (0, 7)                      | 160.62   | 169.01   | 100.11    | 160.62   | 160.07   | 100.00    |
| (0, 8)                      | 159.69   | 166.63   | 100.08    | 159.69   | 159.72   | 100.00    |
| (0, 9)                      | 156.64   | 163.15   | 99.99     | 156.64   | 156.96   | 100.00    |
| (0, 10)                     | 156.61   | 163.08   | 99.99     | 156.61   | 156.96   | 100.00    |

Quadro 5.4.1: Abordagens Dual e de Poisson,  $GM(0, s)$ , sexo masculino

| <i>GM</i><br>(0, <i>s</i> ) | Gamma    |          |           | Poisson  |          |           |
|-----------------------------|----------|----------|-----------|----------|----------|-----------|
|                             | Deviance | $\chi^2$ | $A/E(\%)$ | Deviance | $\chi^2$ | $A/E(\%)$ |
| Null                        | 9006.12  | -        | -         | 9006.12  | -        | -         |
| (0, 2)                      | 326.25   | 260.72   | 87.00     | 326.25   | 410.17   | 100.00    |
| (0, 3)                      | 165.81   | 179.68   | 100.25    | 165.81   | 169.69   | 100.00    |
| (0, 4)                      | 163.84   | 178.36   | 100.04    | 163.84   | 169.45   | 100.00    |
| (0, 5)                      | 160.31   | 168.32   | 99.58     | 160.31   | 172.60   | 100.00    |
| (0, 6)                      | 132.42   | 134.22   | 100.45    | 132.42   | 136.40   | 100.00    |
| (0, 7)                      | 130.56   | 133.97   | 100.44    | 130.56   | 133.19   | 100.00    |
| (0, 8)                      | 110.38   | 115.11   | 99.99     | 110.38   | 111.15   | 100.00    |
| (0, 9)                      | 110.02   | 114.85   | 99.88     | 110.02   | 110.85   | 100.00    |
| (0, 10)                     | 109.91   | 114.88   | 99.85     | 109.91   | 110.71   | 100.00    |

Quadro 5.4.2: Abordagens Dual e de Poisson,  $GM(0, s)$ , sexo feminino

Como se observa, os valores da estatística deviance são idênticos em termos de magnitude e de sinal. Este resultado, antecipado em razão da discussão realizada no Capítulo 3, torna a abordagem dual virtualmente equivalente à 'convencional' graduação de  $\mu_x$  com recurso a GLM's, considerando que os óbitos são modelados em termos de realizações independentes de variáveis aleatórias do tipo Poisson. As ligeiras diferenças nos valores da estatística  $\chi^2$  e do ratio  $A/E(\%)$  são naturalmente explicadas pelas diferenças

## 5.4. GLM's: Abordagem Dual

na definição dos desvios absolutos (e relativos) em ambas as metodologias.

Recordamos, a este propósito, que as estatísticas de qualidade do ajustamento na abordagem dual são definidas por:

$$Dev_x^G = E_x - E_{x,\text{exp}}, \quad Var_x = E_{x,\text{exp}}^2/d_x, \quad z_x = \frac{Dev_x^G}{\sqrt{Var_x}}, \quad A/E = \frac{E_x}{E_{x,\text{exp}}} \quad (5.4.1)$$

onde  $E_{x,\text{exp}}$  denota a exposição ao risco estimada.

A relação entre os desvios absolutos de ambos os modelos é dada por:

$$Dev_x = \frac{-Dev_x^G}{\mu_x} \quad (5.4.2)$$

A análise dos quadros permite ainda concluir que o modelo correspondente a exponenciais de polinómio de grau 6, no caso dos homens, e de grau 7, no caso das mulheres, é suficiente para ajustar correctamente os dados.

Nas figuras 5.4.1 e 5.4.2 é feita uma comparação gráfica do desempenho do modelo Gamma em diferentes dimensões ( $s = 2, \dots, 10$ ).

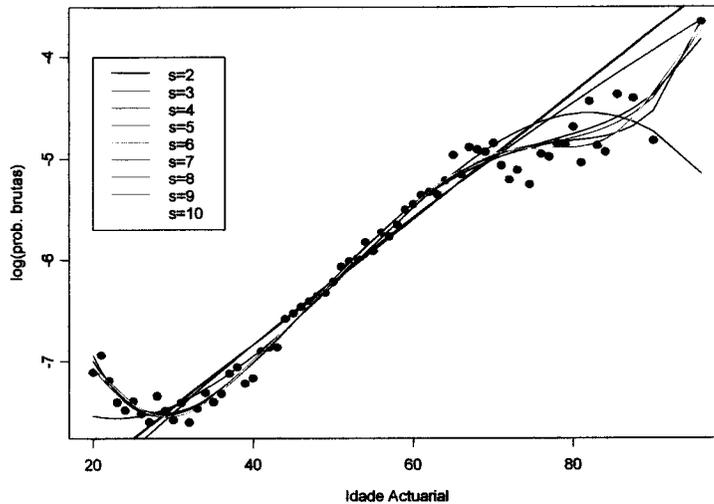


Figura 5.4.1: Comparação dos modelos  $GM(0, s)$  Gamma, sexo masculino

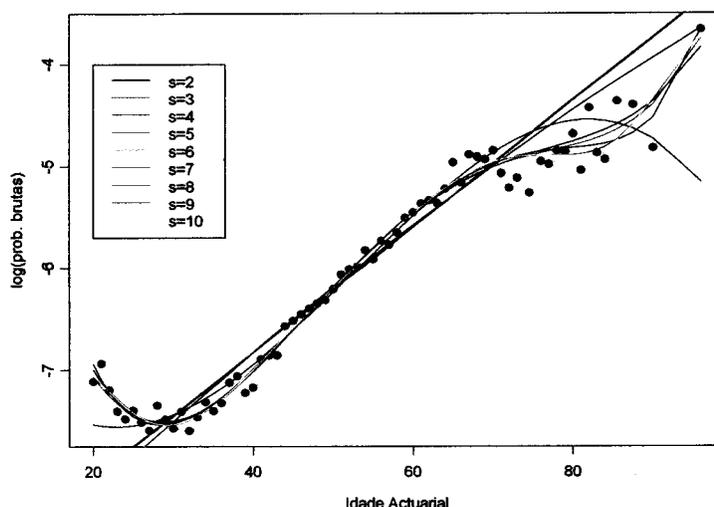


Figura 5.4.2: Comparação dos modelos  $GM(0, s)$  Gamma, sexo feminino

## 5.5 Modelo de Heligman-Pollard

Nesta secção, testamos a eficácia das Leis de Heligman-Pollard (HP) descritas em detalhe no Capítulo 3, na graduação da experiência relativa à população de indivíduos abrangidos por seguros de vida em caso de vida. Perante a presença de heterocedasticidade nos dados, os parâmetros do modelo foram estimados através do método dos mínimos quadrados ponderados não lineares, minimizando  $\sum_{x=x_{\min}}^{x_{\max}} w_x (\hat{q}_x - f(x; \theta))^2$ , onde  $w_x$  denota os ponderadores,  $f(x; \theta)$  representa a função a ajustar,  $\hat{q}_x$  simboliza as probabilidades brutas e as demais variáveis assumem o seu significado habitual.<sup>9</sup>

Para seleccionar os ponderadores, investigámos o desempenho de doze funções de perda distintas  $L_i$  ( $i = 1, \dots, 12$ ), definidas no Quadro 5.5.1. A função de perda  $L_1$  representa o modelo *naïve* em que é admitida a hipótese de igualdade entre as variâncias. As funções de perda  $L_2, L_3, L_4, L_5$  e  $L_{12}$  têm em comum o facto de usarem como ponderadores expressões que derivam da função inversa da variância estimada de  $q_x$  ( $\text{Var}(\hat{q}_x) = \frac{\hat{q}_x(1-\hat{q}_x)}{E_x}$ ). A função de perda  $L_6$  foi usada por Tenenbein e Vanderhoof (1980). Os critérios  $L_7$  e  $L_{11}$  baseiam-se na medida de divergência de Kullback's, re-

<sup>9</sup>As estimativas dos parâmetros e os resultados dos testes de qualidade de ajustamento foram obtidos com o auxílio de uma rotina informática escrita e implementada no *package* S-PLUS 2000 (Release 3).

## 5.5. Modelo de Heligman-Pollard

comendada por Brockett (1991) devido à sua interpretação no contexto da teoria da informação. As restantes funções de perda são usadas por Carriere (1992, 1994).

| # | Função de perda   | #  | Função de perda   |
|---|---|----|---|
| 1 | $L_1 = \sum_{x=x_{\min}}^{x_{\max}} (\hat{q}_x - f(x; \theta))^2$   | 7  | $L_7 = \sum_{x=x_{\min}}^{x_{\max}} \log\left(\frac{\hat{q}_x}{f(x; \theta)}\right) (\hat{q}_x - f(x; \theta))$ |
| 2 | $L_2 = \sum_{x=x_{\min}}^{x_{\max}} \frac{1}{\hat{q}_x} (\hat{q}_x - f(x; \theta))^2$                                   | 8  | $L_8 = \sum_{x=x_{\min}}^{x_{\max}} \left(1 - \frac{d_x^{\text{exp}}}{d_x}\right)^2$                            |
| 3 | $L_3 = \sum_{x=x_{\min}}^{x_{\max}} \frac{E_x}{\hat{q}_x} (\hat{q}_x - f(x; \theta))^2$                                 | 9  | $L_9 = \sum_{x=x_{\min}}^{x_{\max}} \left[\log\left(\frac{d_x^{\text{exp}}}{d_x}\right)\right]^2$               |
| 4 | $L_4 = \sum_{x=x_{\min}}^{x_{\max}} \frac{E_x}{\hat{q}_x^2} (\hat{q}_x - f(x; \theta))^2$                               | 10 | $L_{10} = \sum_{x=x_{\min}}^{x_{\max}} \frac{1}{d_x} (d_x - d_x^{\text{exp}})^2$                                |
| 5 | $L_5 = \sum_{x=x_{\min}}^{x_{\max}} \frac{1}{\hat{q}_x^2} (\hat{q}_x - f(x; \theta))^2$                                 | 11 | $L_{11} = \sum_{x=x_{\min}}^{x_{\max}} (d_x - d_x^{\text{exp}}) \log\left(\frac{d_x}{d_x^{\text{exp}}}\right)$  |
| 6 | $L_6 = \sum_{x=x_{\min}}^{x_{\max}} \left\{ \log\left[\frac{\log(1-f(x; \theta))}{\log(1-\hat{q}_x)}\right] \right\}^2$ | 12 | $L_{12} = \sum_{x=x_{\min}}^{x_{\max}} E_x (\hat{q}_x - f(x; \theta))^2$  |

Quadro 5.5.1: Funções de perda no modelo de Heligman-Pollard

Num primeiro momento, testámos o desempenho das diferentes leis de HP com intuito de seleccionar aquela que proporciona melhores resultados em termos de graduação. Os Quadros 5.5.2 e 5.5.3 resumem a informação sobre os indicadores de qualidade do ajustamento obtidos para cada uma das funções de perda e para cada uma das leis de mortalidade analisadas, num total de 36 combinações distintas.

| Critério | 1ª Lei (versão B) |                |                        | 2ª Lei   |                |                        | 3ª Lei   |                |                        |
|----------|-------------------|----------------|------------------------|----------|----------------|------------------------|----------|----------------|------------------------|
|          | $\chi^2$          | Ratio<br>(A/E) | Signs<br><i>p(pos)</i> | $\chi^2$ | Ratio<br>(A/E) | Signs<br><i>p(pos)</i> | $\chi^2$ | Ratio<br>(A/E) | Signs<br><i>p(pos)</i> |
| $L_1$    | 415.3             | 93.1           | 0.001                  | 185.1    | 100.5          | 0.414                  | 228.8    | 97.1           | 0.064                  |
| $L_2$    | 198.5             | 100.3          | 0.668                  | 189.0    | 100.9          | 0.807                  | 197.0    | 100.2          | 0.807                  |
| $L_3$    | 179.9             | 101.0          | 0.500                  | 173.8    | 100.9          | 0.666                  | 180.3    | 101.0          | 0.586                  |
| $L_4$    | 201.6             | 101.7          | 0.936                  | 186.4    | 102.7          | 0.861                  | 226.7    | 102.3          | 0.807                  |
| $L_5$    | 210.4             | 102.4          | 0.936                  | 211.8    | 101.7          | 0.959                  | 211.1    | 101.7          | 0.959                  |
| $L_6$    | 232.3             | 100.3          | 0.139                  | 177.6    | 100.2          | 0.332                  | 178.2    | 99.6           | 0.096                  |
| $L_7$    | 207.6             | 99.4           | 0.041                  | 173.3    | 99.5           | 0.139                  | 209.9    | 102.3          | 0.904                  |
| $L_8$    | 210.4             | 102.4          | 0.936                  | 211.8    | 101.7          | 0.959                  | 210.5    | 102.0          | 0.936                  |
| $L_9$    | 180.8             | 100.6          | 0.414                  | 177.5    | 100.2          | 0.414                  | 183.7    | 100.2          | 0.414                  |
| $L_{10}$ | 179.9             | 101.0          | 0.500                  | 179.4    | 100.9          | 0.586                  | 180.4    | 101.0          | 0.586                  |
| $L_{11}$ | 170.9             | 100.3          | 0.256                  | 170.4    | 100.3          | 0.332                  | 170.9    | 100.3          | 0.258                  |
| $L_{12}$ | 223.4             | 99.9           | 0.064                  | 173.3    | 99.2           | 0.258                  | 171.6    | 99.1           | 0.193                  |

Quadro 5.5.2: Comparação entre as Leis de Heligman-Pollard, sexo masculino

## 5.5. Modelo de Heligman-Pollard

| Critério | 1ª Lei (versão B) |                |                   | 2ª Lei   |                |                   | 3ª Lei   |                |                   |
|----------|-------------------|----------------|-------------------|----------|----------------|-------------------|----------|----------------|-------------------|
|          | $\chi^2$          | Ratio<br>(A/E) | Signs<br>$p(pos)$ | $\chi^2$ | Ratio<br>(A/E) | Signs<br>$p(pos)$ | $\chi^2$ | Ratio<br>(A/E) | Signs<br>$p(pos)$ |
| $L_1$    | 584.9             | 79.8           | 0.000             | 411.1    | 103.9          | 0.278             | 655.0    | 85.2           | 0.000             |
| $L_2$    | 158.0             | 100.7          | 0.278             | 184.8    | 101.8          | 0.795             | 185.3    | 99.2           | 0.856             |
| $L_3$    | 140.9             | 102.1          | 0.795             | 139.1    | 101.1          | 0.547             | 172.3    | 102.5          | 0.547             |
| $L_4$    | 188.2             | 103.5          | 0.903             | 169.5    | 104.2          | 0.938             | 248.1    | 104.8          | 0.978             |
| $L_5$    | 213.0             | 103.0          | 0.938             | 167.8    | 103.6          | 0.962             | 182.5    | 103.6          | 0.994             |
| $L_6$    | 138.9             | 100.6          | 0.453             | 138.5    | 100.2          | 0.453             | 152.6    | 100.2          | 0.722             |
| $L_7$    | 129.1             | 99.1           | 0.144             | 127.1    | 99.3           | 0.144             | 129.1    | 99.0           | 0.144             |
| $L_8$    | 213.0             | 103.0          | 0.938             | 168.1    | 103.6          | 0.962             | 185.1    | 103.6          | 0.994             |
| $L_9$    | 138.6             | 100.7          | 0.453             | 136.1    | 100.4          | 0.453             | 138.0    | 100.6          | 0.362             |
| $L_{10}$ | 140.9             | 102.1          | 0.795             | 139.2    | 102.0          | 0.638             | 153.1    | 102.1          | 0.903             |
| $L_{11}$ | 211.9             | 100.8          | 0.722             | 128.6    | 100.5          | 0.144             | 129.6    | 100.5          | 0.362             |
| $L_{12}$ | 132.3             | 99.8           | 0.453             | 164.9    | 99.5           | 0.278             | 153.0    | 99.2           | 0.278             |

Quadro 5.5.3: Comparação entre as Leis de Heligman-Pollard, sexo feminino

A observação atenta dos quadros permite extrair as seguintes conclusões: (i) para a maior parte das funções de perda, a 2ª Lei de HP proporciona o melhor desempenho entre as funções paramétricas analisadas; (ii) nos casos em que tal não se verifica, a diferença face à 'melhor' lei é pouco significativa; (iii) a função de perda "óptima" em cada sexo é, aparentemente, independente da lei de mortalidade seleccionada. Em face destes resultados, concluímos que, na amostra em estudo, a 2ª Lei de HP é aquela que melhor performance apresenta.

A etapa seguinte do estudo consistiu em seleccionar, nas duas subpopulações analisadas, a função de perda que reporta o melhor desempenho. Recorrendo à 2ª Lei de HP, testámos as 12 funções de perda acima definidas e recolhemos informação sobre os respectivos indicadores de qualidade do ajustamento. Os Quadros 5.5.4 e 5.5.5 sumarizam os resultados deste procedimento.

A análise dos quadros permite concluir, desde logo, que a performance do modelo de HP não pode ser dissociada da escolha da função de perda a minimizar no contexto da aplicação do método dos mínimos quadrados ponderados não lineares. Pelo contrário, são observadas diferenças significativas no desempenho do modelo em razão da aplicação de distintas funções objectivo. Por outro lado, verifica-se que o desempenho das funções objectivo depende da população em estudo. Com efeito, enquanto que nas vidas do sexo masculino a função de perda  $L_{11}$  apresenta os melhores resultados, seguida de perto pelos critérios  $L_3$  e  $L_{12}$ , nas vidas do sexo feminino a função de perda  $L_7$  supera as demais, seguida no entanto de muito perto pela função  $L_{11}$ .

## 5.5. Modelo de Heligman-Pollard

| Crítério | $\chi^2$ | $p(\chi^2)$ | Ratio<br>100(A/E) | Signs<br>$p(pos)$ | Runs<br>$p(runs)$ | KS<br>$p(KS)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ |
|----------|----------|-------------|-------------------|-------------------|-------------------|---------------|---------------------|----------------------|
| $L_1$    | 185.1    | 0.095       | 100.5             | 0.414             | 0.991             | 1             | 0.117               | 0.033                |
| $L_2$    | 189.0    | 0.074       | 100.9             | 0.807             | 0.615             | 1             | 0.260               | 0.103                |
| $L_3$    | 173.8    | 0.184       | 100.9             | 0.666             | 0.991             | 1             | 0.124               | 0.035                |
| $L_4$    | 186.4    | 0.088       | 102.7             | 0.861             | 0.737             | 1             | 0.160               | 0.050                |
| $L_5$    | 211.8    | 0.013       | 101.7             | 0.959             | 0.061             | 1             | 0.217               | 0.080                |
| $L_6$    | 177.6    | 0.150       | 100.2             | 0.332             | 0.974             | 1             | 0.170               | 0.055                |
| $L_7$    | 173.3    | 0.189       | 99.5              | 0.139             | 0.874             | 1             | 0.141               | 0.043                |
| $L_8$    | 211.8    | 0.013       | 101.7             | 0.959             | 0.061             | 1             | 0.217               | 0.080                |
| $L_9$    | 177.5    | 0.150       | 100.2             | 0.414             | 0.654             | 1             | 0.167               | 0.054                |
| $L_{10}$ | 179.4    | 0.135       | 100.9             | 0.586             | 0.382             | 1             | 0.134               | 0.039                |
| $L_{11}$ | 170.4    | 0.220       | 100.3             | 0.332             | 0.638             | 1             | 0.128               | 0.037                |
| $L_{12}$ | 173.3    | 0.189       | 99.2              | 0.258             | 0.707             | 1             | 0.132               | 0.038                |

Quadro 5.5.4: Comparação entre as funções de perda, sexo masculino

| Crítério | $\chi^2$ | $p(\chi^2)$ | Ratio<br>100(A/E) | Signs<br>$p(pos)$ | Runs<br>$p(runs)$ | KS<br>$p(KS)$ | Portm<br>$p(\cdot)$ | Lj-Box<br>$p(\cdot)$ |
|----------|----------|-------------|-------------------|-------------------|-------------------|---------------|---------------------|----------------------|
| $L_1$    | 411.1    | 0.000       | 103.90            | 0.278             | 0.000             | 1             | 0.000               | 0.000                |
| $L_2$    | 184.8    | 0.010       | 101.76            | 0.795             | 0.591             | 1             | 0.119               | 0.035                |
| $L_3$    | 139.1    | 0.267       | 101.10            | 0.547             | 0.406             | 1             | 0.634               | 0.422                |
| $L_4$    | 169.5    | 0.035       | 104.20            | 0.938             | 0.903             | 1             | 0.788               | 0.610                |
| $L_5$    | 167.8    | 0.040       | 103.62            | 0.962             | 0.779             | 1             | 0.738               | 0.543                |
| $L_6$    | 138.5    | 0.275       | 100.22            | 0.453             | 0.453             | 1             | 0.570               | 0.340                |
| $L_7$    | 127.1    | 0.457       | 99.25             | 0.144             | 0.962             | 1             | 0.562               | 0.349                |
| $L_8$    | 168.1    | 0.039       | 103.62            | 0.962             | 0.779             | 1             | 0.737               | 0.542                |
| $L_9$    | 136.1    | 0.309       | 100.40            | 0.453             | 0.717             | 1             | 0.580               | 0.358                |
| $L_{10}$ | 139.2    | 0.265       | 102.00            | 0.638             | 0.717             | 1             | 0.637               | 0.425                |
| $L_{11}$ | 128.6    | 0.431       | 100.50            | 0.144             | 0.310             | 1             | 0.558               | 0.344                |
| $L_{12}$ | 164.9    | 0.051       | 99.54             | 0.278             | 0.135             | 1             | 0.043               | 0.009                |

Quadro 5.5.5: Comparação entre as funções de perda, sexo feminino

Elegido o modelo analítico adequado, seleccionada a função de perda óptima e estimados os coeficientes do modelo, o passo seguinte consistiu em verificar a significância dos parâmetros. Neste campo, a singularidade da matriz de correlações resultante do procedimento de estimação do modelo inviabilizava, à partida, este objectivo. Para contornar esta dificuldade, recorreremos ao método dos mínimos quadrados não lineares generalizados (*Generalised Non-Linear Least Squares*, GNLS). Em ambos os sexos, a opção foi a de fixar o valor dos parâmetros  $B$  e  $F$  estimando, de seguida, os restantes, usando como estrutura de variâncias a proporcionada pelos valores ajustados pelo próprio modelo. Esta técnica, que consiste em fixar o valor de alguns parâmetros e ajustar os

## 5.5. Modelo de Heligman-Pollard

restantes é sugerida, entre outros, por Congdon (1993). Os valores pré-definidos de  $B$  e  $F$  foram fixados tendo por base as suas estimativas no modelo sem restrições. As estimativas dos parâmetros foram obtidas com o concurso de uma rotina informática escrita e implementada no *package* S-PLUS 2000 (Release 3).

Os Quadros 5.5.6 e 5.5.7 resumem os principais elementos necessários à análise da significância estatística dos parâmetros. Se exceptuarmos o parâmetro  $G$ , no caso das mulheres, e o parâmetro  $K$ , sobretudo no caso dos homens, a generalidade dos parâmetros é significativamente diferente de zero.

| Sexo Masculino |            |          |            |            |
|----------------|------------|----------|------------|------------|
| Parâmetro      | Coef.      | se       | $t$ -ratio | $p$ -value |
| $A$            | 0.08767139 | 0.031389 | 2.79       | 0.0066     |
| $B$            | 0.5        | —        | —          | —          |
| $C$            | 0.3616615  | 0.05077  | 7.12       | <0.0001    |
| $D$            | 0.03698067 | 0.01007  | 3.67       | 0.0004     |
| $E$            | 40.86645   | 8.19676  | 4.99       | <0.0001    |
| $F$            | 98         | —        | —          | —          |
| $G$            | 0.00002235 | 0.000007 | 3.04       | 0.0032     |
| $H$            | 1.087364   | 0.006427 | 169.18     | <0.0001    |
| $K$            | 4.757039   | 6.130646 | 0.78       | 0.4401     |

Quadro 5.5.6: Modelo de Heligman-Pollard, estimativa dos parâmetros do modelo, sexo masculino

| Sexo Feminino |           |           |            |            |
|---------------|-----------|-----------|------------|------------|
| Parâmetro     | Coef.     | se        | $t$ -ratio | $p$ -value |
| $A$           | 0.7964925 | 0.1855756 | 4.29       | 0.0001     |
| $B$           | 0.0561181 | —         | —          | —          |
| $C$           | 1.0683812 | 0.3099428 | 3.45       | 0.0001     |
| $D$           | 0.0009236 | 0.0001509 | 6.12       | <0.0001    |
| $E$           | 5.9904770 | 1.1518110 | 5.20       | <0.0001    |
| $F$           | 57.07553  | —         | —          | —          |
| $G$           | 0.0000005 | 0.0000004 | 1.16       | 0.2496     |
| $H$           | 1.1385477 | 0.0131924 | 86.30      | <0.0001    |
| $K$           | 5.8599126 | 3.3189544 | 1.77       | 0.0822     |

Quadro 5.5.7: Modelo de Heligman-Pollard, estimativa dos parâmetros do modelo, sexo feminino

Por fim, as Figuras 5.5.1 e 5.5.2 ilustram a qualidade do ajustamento e os resíduos dos modelos. Como se observa, a qualidade do ajustamento proporcionada pelo modelo de HP é, em ambos os sexos, considerada adequada, captando as principais características do comportamento da mortalidade.

## 5.5. Modelo de Heligman-Pollard

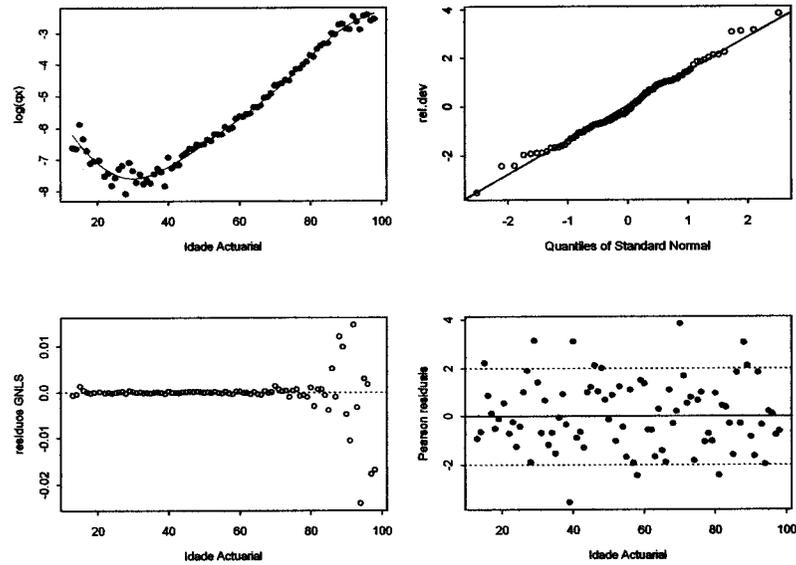


Figura 5.5.1: Graduação pela 2ª Lei de Heligman-Pollard, sexo masculino

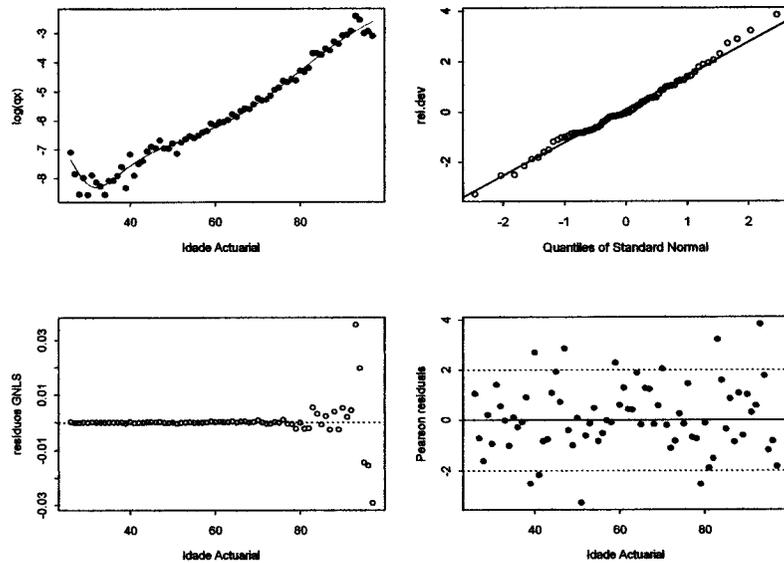


Figura 5.5.2: Graduação pela 2ª Lei de Heligman-Pollard, sexo feminino

## 5.6 Análise comparativa dos diferentes métodos

Para completar o estudo, comparamos nesta secção os resultados obtidos pelos diferentes métodos paramétricos testados no capítulo na graduação da amostra relativa à população de pessoas seguras em caso de vida. A avaliação dos méritos relativos de cada método tem por base a informação transmitida pelos resultados alcançados nos diferentes testes de qualidade do ajustamento, sumariados no Quadro 5.6.1.

| Ordem          | $(r, s)$   | $LGM(r, s)$ |         | $GM(0, s)$ Gamma |         | Heligman-Pollard |         |
|----------------|------------|-------------|---------|------------------|---------|------------------|---------|
|                |            | H           | M       | H                | M       | H                | M       |
|                |            | (3, 3)      | (3, 3)  | (0, 7)           | (0, 6)  | -                | -       |
| Teste $\chi^2$ | $\chi^2$   | 168.2       | 128.4   | 169.0            | 132.4   | 170.4            | 127.1   |
|                | g.l.       | 79          | 66      | 78               | 66      | 76               | 63      |
|                | $p$ -value | 0.326       | 0.539   | 0.288            | 0.439   | 0.220            | 0.457   |
| $z_x$          | $> 2$      | 13          | 10      | 11               | 11      | 11               | 9       |
|                | $> 3$      | 4           | 3       | 2                | 3       | 5                | 3       |
| Signs Test     | pos.(neg.) | 40 (45)     | 34 (38) | 44 (41)          | 40 (32) | 40 (45)          | 31 (41) |
|                | $p$ -value | 0.332       | 0.362   | 0.668            | 0.856   | 0.332            | 0.144   |
| Runs Test      | N.º runs   | 43          | 42      | 46               | 40      | 45               | 36      |
|                | $p$ -value | 0.974       | 0.182   | 0.505            | 0.343   | 0.638            | 0.962   |
| KS             | max. dev   | 0.0048      | 0.0103  | 0.0115           | 0.0237  | 0.0055           | 0.0114  |
|                | $p$ -value | 1.0         | 1.0     | 1.0              | 1.0     | 1.0              | 1.0     |

Quadro 5.6.1: Análise comparativa dos resultados dos testes de qualidade do ajustamento

A observação atenta do quadro permite-nos concluir, em primeiro lugar, que na generalidade dos indicadores de análise o desempenho dos vários métodos é muito semelhante. Com efeito, constata-se que na amostra em estudo as diferentes abordagens superaram com sucesso os distintos testes de qualidade de ajustamento, pelo que podem ser consideradas ferramentas válidas para efeitos de graduação.

Ainda assim, se atendermos unicamente às diferenças existentes em termos de graus de liberdade, podemos concluir que o modelo de Heligman-Pollard apresenta os piores resultados em termos relativos, sobretudo quando comparado com os modelos lineares e não lineares generalizados com previsor linear do tipo  $LGM(r, s)$  (ou, equivalentemente, do tipo  $GM(r, s)$ ). Estes últimos revelam-se, entre os métodos paramétricos testados, os mais flexíveis e adequados, captando adequadamente as principais incidências da mortalidade ao longo do arco da vida humana.

### 5.7 Conclusão

A implementação de um modelo de supervisão baseado na avaliação dos riscos em que incorrem as companhias de seguros e os fundos de pensões requer uma correcta identificação e mensuração dos diversos factores de risco a que estão expostos, quer os activos financeiros que compõem o património das empresas, quer as responsabilidades decorrentes dos contratos de seguros ou dos planos de pensões.

A actividade nos seguros de vida assenta num equilíbrio entre as prestações do segurador e do segurado. Com efeito, os prémios pagos pelo conjunto dos segurados devem permitir ao segurador pagar a totalidade dos sinistros à medida que estes ocorrem. Em particular, as companhias devem calcular os prémios com base em estimativas realistas quanto à probabilidade de sobrevivência das pessoas seguras incluídas nas suas carteiras. Para tal, é imprescindível a existência de tábuas de mortalidade (contemporâneas ou prospectivas) que estimem, de forma adequada, a longevidade da população.

A um outro nível, é importante fomentar e generalizar as práticas de gestão conjunta de activos financeiros e de responsabilidades. De facto, num contexto marcado pela volatilidade dos mercados financeiros e pela incerteza em torno da evolução das variáveis demográficas, é fundamental ajustar continuamente as carteiras das empresas de modo a melhorar a eficácia das estratégias de cobertura do risco salvaguardando, deste modo, os direitos em formação e/ou adquiridos pelos participantes e beneficiários dos planos de pensões.

Neste capítulo, realizámos o primeiro estudo empírico de graduação paramétrica da mortalidade das pessoas seguras e dos beneficiários dos fundos de pensões efectuado em Portugal. O estudo tem como ponto de partida as estatísticas disponibilizadas pelo Instituto de Seguros de Portugal relativas ao período 2000-2004.

Os resultados obtidos permitem concluir, em primeiro lugar, que nas três subpopulações em estudo os quocientes de mortalidade dos indivíduos do sexo feminino são inferiores aos observados nos elementos do sexo masculino, à semelhança daquilo que se observa na população portuguesa em geral. No entanto, assinala-se que a diferença no comportamento da mortalidade entre os sexos tende a diminuir à medida que a idade aumenta.

Em segundo lugar, os estudos confirmam a versatilidade e a eficácia dos modelos lineares e não-lineares em problemas de graduação. Com efeito, os resultados obtidos demonstram que a qualidade do ajustamento depende essencialmente de uma escolha adequada da ordem do previsor linear do Gompertz-Makeham, combinada com a análise dos resultados obtidos num conjunto de testes de ajustamento. A flexibilidade dos

modelos torna-os compatíveis com diferentes experiências de mortalidade.

Em terceiro lugar, a análise realizada nas três subpopulações provou que os resultados obtidos mediante modelos com previsor linear do tipo  $LGM(r, s)$  e *link* logit são indistinguíveis dos alcançados com o recurso a um previsor linear do tipo  $GM(r, s)$  e *link* logarítmico. Por esta razão, a opção por uma das duas alternativas deve obedecer essencialmente a razões de ordem prática

Em quarto lugar, os resultados empíricos obtidos pela abordagem dual confirmam a hipótese teórica de que a graduação de  $\mu_x$  mediante funções paramétricas considerando que a exposição é modelada em termos de realizações independentes de variáveis aleatórias Gamma, condicionadas por  $d_x$ , é equivalente à graduação de  $\mu_x$  considerando que  $d_x$  é modelado em termos de realizações independentes de variáveis aleatórias do tipo Poisson.

O teste à capacidade das leis de Heligman-Pollard (HP) para graduar experiências de mortalidade permitiu concluir, em primeiro lugar, que a segunda lei de HP proporciona o melhor desempenho entre as funções paramétricas analisadas e, em segundo, que a eficácia do modelo depende, em certa medida, da função de perda adoptada no contexto da aplicação do método dos mínimos quadrados ponderados não lineares. Constatou-se ainda que a performance das funções objectivo depende da população em estudo.

A comparação dos resultados proporcionados pelos vários modelos permitiu concluir que, nas amostras consideradas, o desempenho é muito semelhante, constatando-se porém uma ligeira superioridade dos modelos lineares e não-lineares generalizados.

A existência de fontes de erro na amostra, associadas a problemas na recolha de informação estatística, não pode ser negligenciada na avaliação dos resultados e deve motivar a revisão dos procedimentos por parte da entidade reguladora.

**Parte II**

**Tábuas de Mortalidade**  
**Prospectivas**

## Capítulo 6

# Modelos Dinâmicos

*“The limit to the possible duration of life is a subject not likely ever to be determined, even should it exist.”* (Gompertz, 1825)

### 6.1 Introdução

As profundas transformações sócio-económicas ocorridas no último século nos países desenvolvidos reflectiram-se numa melhoria assinalável das condições de vida e de bem-estar das populações e num aumento significativo da esperança de vida à nascença e nas idades adultas. Estes ganhos na longevidade humana são naturalmente entendidos como uma alteração positiva para os indivíduos e uma significativa conquista civilizacional das sociedades modernas.

O declínio da mortalidade é marcado por duas tendências principais: uma redução na mortalidade provocada por doenças infecciosas, especialmente nas idades mais jovens, e um decréscimo da mortalidade nas idades avançadas, reflexo dos avanços no combate às doenças cardiovasculares e ao cancro. Tratando-se de um objectivo individual e colectivo inquestionável, o aumento da longevidade humana coloca, no entanto, novos desafios em múltiplos domínios, em particular na definição e sustentabilidade dos sistemas públicos e privados de protecção e segurança social, tradicionalmente financiados com base no princípio da repartição contemporânea (*“pay-as-you-go”*).

Num contexto em que a mortalidade se assume definitivamente como um fenómeno dinâmico, a necessidade de identificar a sua trajectória e de projectar com exactidão o seu valor futuro afiguram-se cruciais, constituindo condição essencial a um correcto aprovisionamento e gestão das responsabilidades financeiras futuras com os sistemas de segurança e protecção social. Esta tarefa envolve, em boa medida, a utilização de

modelos de projecção adequados que conduzam à construção de tábuas de mortalidade prospectivas (ou geracionais) para uso, por exemplo, no campo actuarial na quantificação e controlo do risco de longevidade.

Na Parte I, examinámos a aplicação de métodos paramétricos e não-paramétricos na gradação de dados de mortalidade, centrando-nos na influência que a idade do indivíduo tem sobre esta. Neste capítulo, reconhecemos o carácter dinâmico da mortalidade e incorporamos o tempo cronológico na análise, discutindo as características e a aplicação de um conjunto extenso e diversificado de métodos de projecção, desenvolvidos nos campos actuarial e demográfico.

O capítulo está organizado da seguinte forma. Nas Secções 6.2 e 6.3 são analisadas as principais tendências registadas na mortalidade humana e identificados os aspectos cruciais a ter em conta na modelação dinâmica do fenómeno. Nas Secções 6.4 e 6.5 discutimos a aplicação dos métodos de projecção vertical, horizontal e diagonal, considerando métodos paramétricos com ou sem recurso a tábua base.

Na Secção 6.6 analisamos em detalhe as características e a estimação do popular método Lee-Carter. Nas Secções 6.7 e 6.8 é investigada a aplicação dos modelos lineares generalizados na modelação da mortalidade num contexto em que a idade e o tempo cronológico são simultaneamente considerados na estimação, tendo em vista a projecção de factores de redução.

Na Secção 6.9 é examinada em detalhe uma extensão do método Lee-Carter designada por modelo de Poisson, as suas variantes, métodos de estimação dos parâmetros e de construção de intervalos de confiança. Em particular, é ensaiada uma nova técnica, inspirada nas P-Splines, que incorpora as preocupações quanto ao alisamento das séries no procedimento de estimação dos parâmetros. Nesta Secção, é ainda desenvolvida uma extensão do modelo onde se admite que as projecções de mortalidade são balizadas por uma determinada tábua limite.

Na Secção 6.10 estudamos uma variante do método Lee-Carter que recorre à distribuição Binomial. Nas Secções 6.11 e 6.12 descrevemos os métodos de projecção assentes em tábuas limite e tábuas modelo. Nas Secções 6.13 e 6.14 analisamos os métodos de projecção que envolvem a transformação das variáveis de interesse e os métodos não-paramétricos, respectivamente. Por fim, na Secção 6.15 discutimos a aplicação de um conjunto de métodos na projecção da mortalidade nas idades avançadas.

## 6.2 Tendências na mortalidade humana

A análise das tendências na mortalidade da população atraiu desde sempre a atenção de investigadores e organismos oficiais a nível internacional, interessados em compreender as múltiplas implicações do fenómeno no desenvolvimento económico, social e humano das sociedades.<sup>1</sup> O comportamento da mortalidade nos países desenvolvidos durante o século XX é marcado, em termos genéricos, por uma assinalável redução nas taxas de mortalidade em todas as idades, em especial nas idades mais jovens, e por um aumento notável da esperança de vida, quer à nascença, quer nas idades adultas.<sup>2</sup> Extraordinária foi ainda a redução na mortalidade infantil, que assume hoje em dia valores insignificantes face aos apresentados no início do século passado.

Se é verdade que este declínio foi contínuo ao longo de todo o século XX, também é verdade que a distribuição etária desse decréscimo não foi simétrica. Com efeito, se na primeira metade do século a redução se concentrou principalmente nas idades jovens, devido à diminuição na mortalidade infantil e juvenil, na segunda metade os ganhos na esperança de vida são explicados, sobretudo, pela diminuição da mortalidade nas idades adultas e avançadas. Estas modificações têm, desde logo, uma explicação pragmática. A redução alcançada na mortalidade infantil e juvenil para níveis tão baixos conduziu-nos a um patamar em que, mesmo que se registem ganhos extraordinários neste segmento, o seu impacto sobre a esperança de vida será reduzido, sobretudo quando comparado com o que pode advir da redução da mortalidade nas idades adultas.

Por outro lado, estas alterações acompanharam, em boa medida, as principais tendências registadas durante este período no campo epidemiológico. Na primeira metade do século XX, as doenças infecciosas (tuberculose, pneumonia, gripe, tifoide, difteria, etc.) constituíam a principal causa de morte. O desenvolvimento de vacinas e a introdução de antibióticos após a II Guerra-Mundial possibilitaram, em conjunto com a melhoria dos níveis de nutrição e com a expansão dos sistemas de higiene e saúde públicos, a quase erradicação desta causa de morte. Assim, não surpreende que na segunda metade do século os ganhos na mortalidade tenham estado associados a factores distintos, como sejam as conquistas no combate às doenças cardiovasculares e ao cancro. De facto, a

---

<sup>1</sup>Veja-se, por exemplo, MacDonald (1997), MacDonald *et al.* (1998), Cutler e Meara (2001), Willets *et al.* (2004) e referências neles citadas.

<sup>2</sup>A título de exemplo, a esperança de vida à nascença em Inglaterra e Gales era, em 1901, de 46 anos para os homens e de 50 para as mulheres, sendo que um século depois, em 2001, estes valores haviam aumentado para 76 e 81 anos, respectivamente. Em Portugal, a tábua de mortalidade relativa ao período 1939-42 (I.N.E.) reportava uma esperança de vida à nascença de 48.58 anos para os homens e de 52.82 anos para as mulheres. Em 2002-03, ou seja, 60 anos mais tarde, a evolução positiva da mortalidade fez com que este indicador alcançasse os 73.93 anos para os homens e os 80.23 anos para as mulheres (Carrilho e Patrício, 2004).

inovação tecnológica no combate às doenças agudas do coração (e.g., cirurgia de bypass, angioplastia, medicação para reduzir a pressão arterial e o colesterol, etc.), os avanços no tratamento do cancro e algumas alterações do ponto de vista comportamental (e.g., tabagismo, maior preocupação com a nutrição e com a prática de desporto) ajudam a compreender a redução da mortalidade neste período. O facto de este tipo de patologias incidir, principalmente, sobre as idades adultas e avançadas explica, em boa parte, as modificações no perfil etário dos ganhos de longevidade neste período.

Uma forma alternativa de compreender as tendências na mortalidade humana consiste em analisar o perfil gráfico de funções que representam a mortalidade em termos da idade alcançada pelo indivíduo, como sejam a função de sobrevivência ou a “curva de mortes” (MacDonald *et al.*, 1998; Horiuchi e Wilmoth, 1998). A análise destes perfis no tempo coloca em evidência a crescente concentração dos óbitos em torno da moda (nas idades adultas) da curva de mortes, ou seja, a função de sobrevivência tende para uma forma rectangular, num fenómeno que é por isso apelidado de «rectangularização». Por outro lado, observa-se que a moda da curva de mortes (que, devido à rectangularização, tende para a idade máxima admissível  $\omega$ ) se move em direcção às idades mais avançadas, numa trajectória que é apelidada de «expansão» da função de sobrevivência (Cheung e Robine, 2007). Por fim, recentemente têm-se registado níveis mais elevados e uma maior dispersão das mortes acidentais nas idades jovens (a chamada bossa de acidentes ou *young mortality hump*). No Capítulo 7 teremos oportunidade de verificar que estas tendências são igualmente observadas na população portuguesa.

## 6.3 Modelação da mortalidade num contexto dinâmico

### 6.3.1 Diagrama de Lexis e cálculo das estimativas brutas dos indicadores de mortalidade

O tempo intervém de três formas distintas num estudo sobre a mortalidade: na idade dos indivíduos que integram a população, na geração da qual estes emanam e no instante em que é realizada a observação. Existem, assim, três coordenadas que permitem caracterizar um evento demográfico (e.g., um óbito): (i) a data em que este se produz, (ii) a idade do indivíduo abrangido pelo evento e (iii) a sua data de nascimento. Pese embora estas três informações sejam obviamente redundantes, uma vez que bastam apenas duas para caracterizar o evento, elas são importantes para caracterizar o fenómeno em estudo.

O diagrama de Lexis, introduzido em 1875 por Wilhelm Lexis (1837-1914), pode ser definido como um sistema de eixos rectangulares que permite precisar a forma como se

### 6.3. Modelação da mortalidade num contexto dinâmico

combinam as três coordenadas temporais na representação gráfica dos acontecimentos demográficos.<sup>3</sup> A sua construção baseia-se num conjunto de paralelas a dois eixos perpendiculares. No eixo das abcissas, inscrevem-se o tempo ou as datas de observação das variáveis, habitualmente um ano coincidente com o ano civil, e no eixo das ordenadas a idade dos indivíduos ou a duração do acontecimento, em anos completos. Cada acontecimento demográfico corresponde a um ponto cujas coordenadas são o tempo em que o mesmo ocorreu e a idade do indivíduo sujeito ao referido evento. Os pontos referentes aos acontecimentos demográficos são contabilizados em triângulos que representam a dupla classificação, i.e., por idade e por data de nascimento. As diagonais limitam as linhas de vida dos indivíduos nascidos no mesmo ano, ou seja, que fazem parte da mesma geração.<sup>4</sup> Na Figura 6.3.1 representamos a linha de vida de um indivíduo no diagrama de Lexis.

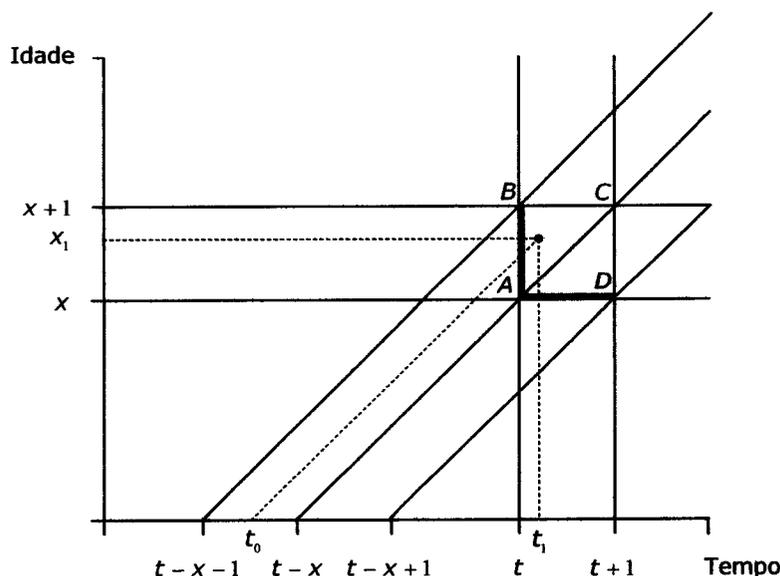


Figura 6.3.1: Diagrama de Lexis

Ele nasce no momento  $t_0$  e morre no momento  $t_1$  à idade  $x_1$ , entre as idades exactas (datas de aniversário)  $x$  e  $x+1$ . O segmento de recta vertical a azul representa o número de indivíduos que no momento  $t$  tem  $x$  anos de idade, i.e., o número de sobreviventes  $P_{x,t}$  de idade  $x$  no instante  $t$ . O segmento de recta horizontal a vermelho representa o

<sup>3</sup>Para uma introdução ao diagrama de Lexis veja-se, por exemplo, Vandeschrisck (2001).

<sup>4</sup>Doravante, designaremos uma geração pelo ano de nascimento dos indivíduos que a integram. Por exemplo, a geração  $g$  é formada pelos indivíduos nascidos no ano de calendário  $g$ .

### 6.3. Modelação da mortalidade num contexto dinâmico

---

número de indivíduos que no ano  $t$  cumpriram o  $x^{\text{ésimo}}$  aniversário, i.e., o número de sobreviventes à idade  $x$  provenientes da geração  $t - x$ .

O diagrama de Lexis permite identificar facilmente o efectivo de uma geração: (i) nas intersecções com os eixos verticais obtemos o efectivo no dia 1 de Janeiro de anos consecutivos, data na qual todos os indivíduos têm a mesma idade em anos decorridos, i.e., em número de anos inteiros vividos até ao último aniversário; (ii) nas intersecções com os eixos horizontais, onde se obtém o efectivo em datas de aniversários sucessivas. O diagrama permite ainda resgatar o número de óbitos localizados nas células delimitadas pelo sistema, seja no quadrado circunscrito pelas coordenadas  $ABCD$ , seja nos triângulos  $ABC$  e  $ACD$  demarcados pelas diagonais.

No decurso do ano  $t$ , são registados  $d_{x,t}$  óbitos entre os indivíduos com  $x$  anos inteiros vividos, localizados num quadrado delimitado pelas coordenadas  $ABCD$ . Esses  $d_{x,t}$  óbitos correspondem a indivíduos provenientes de duas gerações distintas,  $t - x$  e  $t - x - 1$ , e correspondem às linhas de vida que terminaram, respectivamente, nos triângulos inferior e superior.

Na Figura 6.3.2 vemos que é igualmente possível isolar o número de óbitos  $d_{x,t,g}$  à idade  $x$  de indivíduos pertencentes a uma geração, neste caso à geração  $g$  nascida no ano  $g = t - x$ .

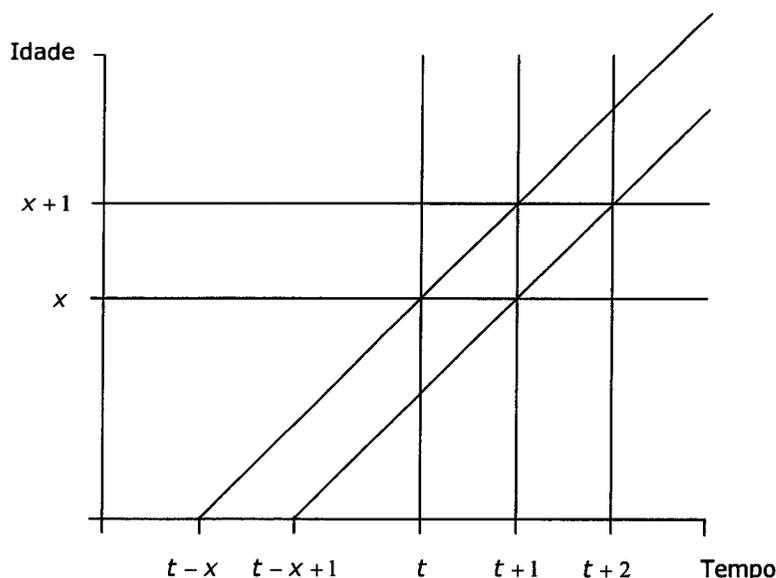


Figura 6.3.2: Número de óbitos à idade  $x$  entre os indivíduos da geração  $t - x$

Esses óbitos estão localizados num paralelograma na horizontal, situado na inter-

### 6.3. Modelação da mortalidade num contexto dinâmico

secção entre o segmento sombreado na diagonal (correspondente à geração  $g$ ) e o segmento sombreado na horizontal (relativo à idade  $x$ ). Como se observa, os óbitos são registados no decurso de dois anos de calendário consecutivos,  $t$  e  $t + 1$ .

Nalguns estudos, estamos interessados em isolar o número de óbitos registados num único ano de indivíduos oriundos de uma única geração. Como se observa na Figura 6.3.3, esses óbitos estão localizados num paralelograma na vertical, situado na intersecção entre o segmento sombreado na diagonal (correspondente à geração  $g$  nascida no ano  $t - x$ ) e o segmento sombreado na vertical (corresponde ao ano de calendário  $t$ ). Assinale-se que esses óbitos ocorrem no decurso de um único ano  $t$  mas abrangem duas idades consecutivas  $x$  e  $x - 1$ .

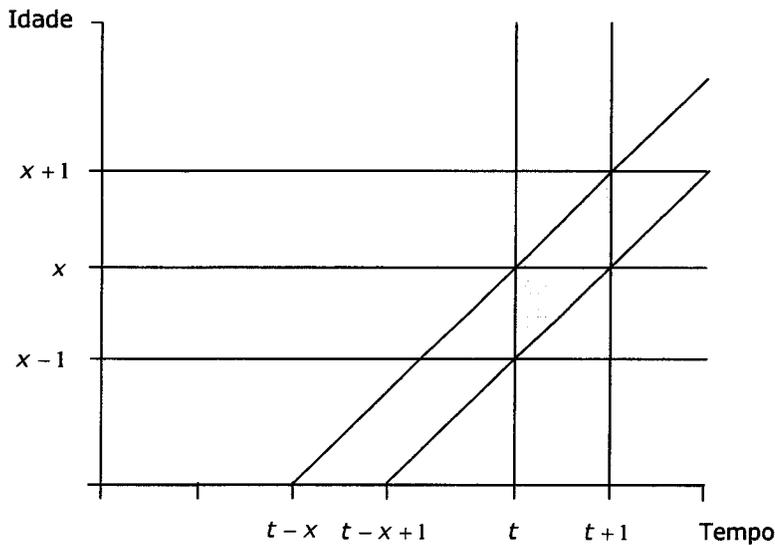


Figura 6.3.3: Número de óbitos registados no ano  $t$  de indivíduos nascidos em  $t - x$

As Figuras 6.3.1, 6.3.2 e 6.3.3 traduzem, como vimos, formas distintas de decompor a informação sobre o evento demográfico. Nos casos em que os óbitos têm dupla classificação (por idade e ano de nascimento), elas dão origem a definições alternativas para os conceitos de exposição ao risco, quocientes e taxas de mortalidade. Consideremos, em primeiro lugar, a estimativa do quociente de mortalidade relativo ao quadrado  $ABCD$  da Figura 6.3.1. Este quadrado pode ser decomposto em dois triângulos: um triângulo inferior  $ACD$ , que agrega os óbitos  $d_{x,t,g}^L$  registados no ano  $t$  entre os indivíduos oriundos da geração nascida em  $g = t - x$ ; um triângulo superior  $ABC$ , que agrupa os óbitos  $d_{x,t,g-1}^U$  registados no ano  $t$  entre os indivíduos oriundos da geração nascida em

### 6.3. Modelação da mortalidade num contexto dinâmico

$g - 1 = t - x - 1$ . Denotemos por  $d_{x,t}$  o número total de óbitos registados em cada um desses quadrados  $(x, t)$ , com

$$d_{x,t} = d_{x,t,g}^L + d_{x,t,g-1}^U. \quad (6.3.1)$$

Se admitirmos que os óbitos se distribuem de forma uniforme nos triângulos em que se decompõe o quadrado de Lexis, demonstra-se que a estimativa da taxa de mortalidade no intervalo  $[x, x + 1)$ , definida em (2.2.22), é dada por<sup>5</sup>

$$\hat{m}_{x,t} = \frac{d_{x,t}}{E_{x,t}} = \frac{d_{x,t,g}^L + d_{x,t,g-1}^U}{\frac{1}{2} [P_{x,t} + P_{x,t+1}] + \frac{1}{6} [d_{x,t,g}^L - d_{x,t,g-1}^U]}. \quad (6.3.2)$$

As estimativas dos quocientes de mortalidade  $q_{x,t}$  correspondentes são calculadas, regra geral, convertendo (6.3.2) em probabilidades de óbito, sob a hipótese de repartição uniforme dos óbitos durante o ano, i.e.,

$$\hat{q}_{x,t,\bullet} = \frac{\hat{m}_{x,t}}{1 + (1 - a_{x,t}) \cdot \hat{m}_{x,t}}, \text{ com } a_{x,t} = \frac{1}{2}, \quad (6.3.3)$$

onde  $a_{x,t}$  denota o número médio de anos vividos no intervalo  $[x, x + 1)$  pelos indivíduos que falecem à idade  $x$  no momento  $t$ . No caso do paralelograma representado na Figura 6.3.2, a probabilidade de óbito no intervalo  $[x, x + 1)$ ,  $q_{x,\bullet,g}$ , é estimada de forma directa mediante

$$\hat{q}_{x,\bullet,g} = \frac{d_{x,t,g}^L + d_{x,t+1,g}^U}{P_{x,t+1} + d_{x,t,g}^L}. \quad (6.3.4)$$

Prova-se que a estimativa de  $m_x$  é dada, neste caso, de forma exacta por<sup>6</sup>

$$\hat{m}_{x,t} = \frac{d_{x,t}}{E_{x,t}} = \frac{d_{x,t,g}^L + d_{x,t+1,g}^U}{P_{x,t+1} + \frac{1}{3} [d_{x,t,g}^L - d_{x,t+1,g}^U]}. \quad (6.3.5)$$

Uma alternativa envolve o recurso ao método dos quocientes parciais de Vallin (1973). Esta técnica consiste em estimar os quocientes de mortalidade parciais para os triângulos inferior e superior na Figura 6.3.1. Na ausência de migração,<sup>7</sup> o quociente relativo ao

<sup>5</sup> Para mais detalhes sobre o cálculo da exposição ao risco  $E_{x,t}$  veja-se Wilmoth *et al.* (2005).

<sup>6</sup> A única excepção refere-se à idade 0, para a qual é usual efectuarem-se alguns ajustamentos no cálculo da exposição ao risco  $E_{x,t}$ , uma vez que não se observa a hipótese de distribuição uniforme dos óbitos, nem no interior dos quadrados de Lexis, nem no interior dos triângulos.

<sup>7</sup> As migrações podem facilmente ser incorporadas na estimativa admitindo a hipótese de que elas se repartem uniformemente em intervalos de um ano (Vallin, 1973).

### 6.3. Modelação da mortalidade num contexto dinâmico

---

triângulo inferior  $ACD$  é estimado por

$$\hat{q}_{x,t,g} = \frac{d_{x,t,g}^L}{P_{x,t+1} + d_{x,t,g}^L}, \quad (6.3.6)$$

enquanto que no triângulo superior  $ABC$  este é estimado por

$$\hat{q}_{x,t,g-1} = \frac{d_{x,t,g-1}^L}{P_{x,t}}. \quad (6.3.7)$$

A estimativa do quociente relativo ao paralelograma representado na Figura 6.3.2 é então dada por

$$\hat{q}_{x,\bullet,g} = 1 - (1 - \hat{q}_{x,t,g})(1 - \hat{q}_{x,t+1,g}). \quad (6.3.8)$$

O método de Vallin pode igualmente ser usado para calcular os quocientes de mortalidade  $\hat{q}_{\bullet,t,g}$  relativos ao paralelograma representado na Figura 6.3.3 e ao quadrado na Figura 6.3.1. Os seus valores são estimados por

$$\hat{q}_{\bullet,t,g} = 1 - (1 - \hat{q}_{x,t,g})(1 - \hat{q}_{x-1,t,g}) \quad (6.3.9)$$

e por

$$\hat{q}_{x,t,\bullet} = 1 - (1 - \hat{q}_{x,t,g})(1 - \hat{q}_{x,t,g-1}), \quad (6.3.10)$$

respectivamente.

A escolha da combinação adequada para calcular os quocientes ou taxas de mortalidade depende dos dados disponíveis e do objectivo a perseguir com o estudo. As configurações representadas nas Figuras 6.3.1 e 6.3.3 são, à partida, mais adequadas a estudos periódicos que isolem as condições registadas num único ano civil. Esta vantagem não é, no entanto, decisiva, se as condições de momento forem susceptíveis de oscilações importantes entre anos consecutivos, que exerçam efeitos significativos sobre a mortalidade. Em contrapartida, a configuração na Figura 6.3.2 considera dois anos consecutivos, uma característica que não constitui qualquer obstáculo quando o objectivo do estudo é o de proceder a uma análise sobre a evolução temporal da mortalidade.

No cálculo dos quocientes  $\hat{q}_{x,t,\bullet}$  correspondentes à Figura 6.3.1, são considerados óbitos de um ano civil relativos a duas gerações distintas, registados num intervalo de idades entre duas datas de aniversário exactas. Esta opção garante a homogeneidade por ano civil e por idade, mas não por geração. Por outro lado, ela não permite, como vimos, o cálculo directo das probabilidades  $\hat{q}_{x,t,\bullet}$ , sendo necessário recorrer à aproximação dada por 6.3.3, nem a utilização directa em estudos prospectivos.

No cálculo dos quocientes  $\hat{q}_{x,\bullet,g}$  correspondentes à Figura 6.3.2, são considerados óbitos de dois anos civis relativos a uma única geração, registados num intervalo de idades entre duas datas de aniversário exactas. As principais vantagens desta solução são a garantia de homogeneidade por idade e por geração, assim como o facto de ela permitir o cálculo directo das probabilidades  $\hat{q}_{x,\bullet,g}$  e o seu uso em estudos prospectivos. No caso de optarmos por considerar um intervalo de tempo mais alargado (e.g., 3 anos de observação), ela permite igualmente suavizar eventuais oscilações erráticas no número de óbitos. Em contrapartida, o método não assegura a homogeneidade por ano civil.

No cálculo dos quocientes  $\hat{q}_{\bullet,t,g}$  correspondentes à Figura 6.3.3, são considerados óbitos de um ano civil relativos a uma única geração, registados num intervalo de idades entre três datas de aniversário exactas, ou seja, entre dois anos vividos completos. Esta opção garante a homogeneidade por ano civil e por geração, mas apresenta uma limitação importante na construção de tábuas de mortalidade. Com efeito, ela não garante a homogeneidade por idade, ou seja, não fornece uma estimativa da probabilidade de óbito entre duas datas de aniversário consecutivas.

#### 6.3.2 Caracterização do problema

Num contexto dinâmico, a ideia base subjacente aos vários modelos de projecção é a de que a mortalidade deve ser expressa mediante uma função do tempo (futuro) cronológico  $t$ . Se nos basearmos num único indicador, o modelo dinâmico é definido por uma função  $\Psi(t)$ .<sup>8</sup> Neste campo, existe um interesse particular pela análise da evolução no tempo de indicadores sintéticos (“*markers*”), dada a sua capacidade de sintetizar características importantes no comportamento passado da mortalidade e sua respectiva distribuição de probabilidade. Os diferentes indicadores podem, segundo Pitacco (2004), ser classificados em “medidas de localização”, fornecendo informação sobre o fenómeno de expansão, e “medidas de variabilidade”, condensando elementos sobre o fenómeno de rectangularização da função de sobrevivência.

Entre as principais medidas de localização, merecem destaque a esperança de vida (à nascença,  $\hat{e}_0$ , ou à idade  $x$ ,  $\hat{e}_x$ ) e o ponto de Lexis (a moda da “*curve of deaths*”). Entre as medidas de variabilidade realce para:<sup>9</sup>

1. a variância da vida residual,  $\text{Var}(T_x)$ , ou o respectivo desvio-padrão,  $\sigma_x = \sqrt{\text{Var}(T_x)}$ ;

---

<sup>8</sup>Por exemplo, a esperança de vida à nascença para um recém-nascido é representada, num contexto dinâmico, pela função  $\hat{e}_0(t)$  do ano cronológico  $t$  (neste caso do ano de nascimento). De forma análoga, o quociente de mortalidade numa dada população pode ser expressa pela função  $q_x(t)$ , onde  $t$  denota o ano de calendário em que a população é analisada.

<sup>9</sup>Para uma lista exaustiva de medidas de variabilidade da idade de morte e de avaliação do fenómeno de rectangularização recomenda-se a leitura de Wilmoth e Horiuchi (1999).

### 6.3. Modelação da mortalidade num contexto dinâmico

---

2. o intervalo inter-quartis,  $IQR = x'' - x'$ , onde  $x'$  e  $x''$  são as idades para as quais  $l_{x'} = 0.75$  e  $l_{x''} = 0.25$  (com  $l_0 = 1$ ). Assinale-se que o intervalo  $IQR$  é tanto maior quanto mais dispersa é a distribuição da vida residual da população;
3. a denominada “entropia” da função de sobrevivência,  $H$ , definida por (Keyfitz, 1985; Wilmoth e Horiuchi, 1999)

$$H = -\frac{\int_0^{\infty} S(x) \ln S(x) dx}{\int_0^{\infty} S(x) dx} \quad (6.3.11)$$

consistindo numa média ponderada de  $\ln S(x)$ , usando como pesos a própria função de sobrevivência. Quanto mais concentrados forem os óbitos menor será o valor de  $H$ . Em particular,  $H = 0$  quando a função de sobrevivência apresenta uma configuração gráfica perfeitamente rectangular;

4. o coeficiente  $G$ , definido por

$$G = \frac{1}{e_0} \int_0^{\infty} [S(x)]^2 dx$$

que mede, neste caso, a desigualdade na distribuição da idade de óbito. O seu valor está contido no intervalo  $[0, 1]$ , com um valor mais baixo de  $G$  a corresponder a menor variabilidade na idade de óbito.

Informação complementar sobre a distribuição de probabilidade pode ainda ser obtida a partir de indicadores como a mortalidade infantil, ou através da chamada “*endurance*”, definida como a idade  $\xi$  para qual se observa que  $l_{\xi} = 0.90$ .

Num contexto dinâmico, todos os indicadores acima apresentados devem ser indexados por ano de calendário (e.g.,  $e_x(t)$ ,  $IQR(t)$ ,  $H(t)$ , etc.). A complexidade do problema aumenta significativamente no contexto actuarial, onde na maior parte dos casos são necessários indicadores de localização e de variabilidade específicos para cada idade. Dito de outro modo, num contexto dinâmico a mortalidade é uma função  $\Gamma(x, t)$ , não apenas da idade  $x$  mas também do tempo cronológico  $t$ .<sup>10</sup>

O modelo de projecção pode ser definido pela condição  $\Gamma(x, t) \mid t > t_0$ , onde  $t_0$  simboliza o ano actual ou o ano para o qual estão disponíveis as últimas tábuas contemporâneas. Admitamos que  $x$  e  $t$  são variáveis discretas e que  $\Gamma(x, t)$  representa os quocientes de mortalidade  $q_x(t)$ . Neste caso, a função  $\Gamma(x, t)$  pode ser representada por

---

<sup>10</sup>Em termos mais precisos, a função  $\Gamma(x, t)$  pode representar indicadores como a taxa de mortalidade, a força de mortalidade, a função de sobrevivência, transformações da função de sobrevivência, entre outros.

### 6.3. Modelação da mortalidade num contexto dinâmico

uma matriz cujas colunas correspondem aos anos cronológicos e cujas linhas definem as idades actuariais, i.e.,

| Idade        | Ano cronológico |                       |                   |                       |     |
|--------------|-----------------|-----------------------|-------------------|-----------------------|-----|
|              | ...             | $t - 1$               | $t$               | $t + 1$               | ... |
| 0            | ...             | $q_0(t - 1)$          | $q_0(t)$          | $q_0(t + 1)$          | ... |
| 1            | ...             | $q_1(t - 1)$          | $q_1(t)$          | $q_1(t + 1)$          | ... |
| ...          | ...             | ...                   | ...               | ...                   | ... |
| $x$          | ...             | $q_x(t - 1)$          | $q_x(t)$          | $q_x(t + 1)$          | ... |
| $x + 1$      | ...             | $q_{x+1}(t - 1)$      | $q_{x+1}(t)$      | $q_{x+1}(t + 1)$      | ... |
| ...          | ...             | ...                   | ...               | ...                   | ... |
| $\omega - 1$ | ...             | $q_{\omega-1}(t - 1)$ | $q_{\omega-1}(t)$ | $q_{\omega-1}(t + 1)$ | ... |

Quadro 6.3.1: Quocientes de mortalidade anuais por idade e ano cronológico

Esta representação dos quocientes  $q_x$  num contexto dinâmico permite três leituras distintas:

1. uma leitura “vertical” (i.e., por colunas),

$$q_0(t), \dots, q_x(t), \dots, q_{\omega-1}(t),$$

correspondendo a uma sequência de tábuas contemporâneas, uma por cada ano cronológico  $t$ ;

2. uma leitura “horizontal” (i.e., por linhas),

$$\dots, q_x(t - 1), q_x(t), q_x(t + 1), \dots,$$

fornecendo informação sobre os perfis de mortalidade em cada uma das idades  $x$ ;

3. uma leitura “diagonal”,

$$q_0(t), q_1(t + 1), \dots, q_x(t + x), \dots,$$

correspondendo a uma sequência de tábuas de mortalidade geracionais, uma por cada ano de nascimento (geração)  $t$ .

Para um dado ano base  $t_0$  e um determinado horizonte temporal de previsão  $t_{\max}$ , designa-se por tábua projectada a sub-matriz

$$\{q_x(t)\}; \quad x = 0, 1, \dots, \omega - 1; \quad t = t_0, t_0 + 1, \dots, t_{\max}.$$

Para a construção de uma tábua projectada são necessários três elementos base: um suporte estatístico, compreendendo uma amostra temporal com dados sobre a população

e sobre a mortalidade real (ou um conjunto de tábuas de mortalidade), um modelo de projecção (i.e., uma estrutura formal a adoptar no procedimento de extrapolação) e, eventualmente, uma tábua base, sobre a qual incidirá o modelo de projecção.

#### 6.3.3 Tábuas projectadas unidimensionais

Como se depreende da definição acima, uma tábua de mortalidade projectada é uma matriz bidimensional  $\{q_x(t)\}_{x \in [0, \omega-1]; t \in [t_0, t_{\max}]}$ . O uso apropriado de uma tábua projectada pressupõe a adopção de uma perspectiva geracional, ou seja, implica uma leitura diagonal da matriz para cada ano de nascimento. Por exemplo, a probabilidade, calculada no momento  $t$ , de um indivíduo com  $x$  anos de idade (i.e., nascido no ano  $\tau = t - x$ ) sobreviver  $k$  anos é dada pela fórmula

$${}_k p_x^{(\tau)} = \prod_{j=0}^{k-1} [1 - q_{x+j}(t+j)]. \quad (6.3.12)$$

O uso de múltiplas tábuas geracionais não suscita qualquer complicação do ponto de vista conceptual, mas comporta algumas desvantagens do ponto de vista prático. Em particular, em cada ano cronológico é necessário empregar tantas tábuas geracionais quantos os anos de nascimento das pessoas seguras incluídas na carteira. Por outro lado, a ser aceite pelas entidades reguladoras, este procedimento implicaria que todos os anos seriam usadas tábuas de mortalidade distintas no cálculo dos prémios dos contratos relativos aos novos membros da carteira, ou seja, as bases técnicas demográficas usadas pelas companhias mudariam todos os anos.

Apesar de os actuais sistemas de informação terem capacidade mais do que suficiente para resolver os problemas acima levantados, na prática actuarial continua a ser frequente o recurso a uma única tábua unidimensional, extraída, de certa forma, da tábua projectada. Entre as várias alternativas que se nos oferecem para a sua construção, destacamos as duas seguintes:

1. Escolhe-se um determinado ano de nascimento  $\bar{\tau}$  e considera-se exclusivamente a tábua correspondente a essa geração,  $\{q_x^{[\bar{\tau}]}\}$ , e portanto as probabilidades

$$q_{x_{\min}}(\bar{\tau} + x_{\min}), q_{x_{\min}+1}(\bar{\tau} + x_{\min} + 1), \dots, q_x(\bar{\tau} + x), \dots,$$

onde  $x_{\min}$  representa a idade mínima de interesse e  $\bar{\tau}$  pode corresponder, por exemplo, ao ano médio de nascimento dos indivíduos de uma dada carteira, calculado num intervalo de tempo alargado.

### 6.3. Modelação da mortalidade num contexto dinâmico

---

2. Em alternativa, pode seleccionar-se um determinado ano de calendário futuro  $s$  :  $s > t$  e considerar-se unicamente a tábua de mortalidade contemporânea correspondente, i.e.,

$$q_{x_{\min}}(s), q_{x_{\min}+1}(s), \dots, q_x(s), \dots,$$

onde  $s$  pode representar, por exemplo, o horizonte temporal de médio-longo prazo em que se admite que a população coberta pela carteira estabilizará.

De acordo com a modalidade 1, a probabilidade de um indivíduo sobreviver até à idade  $x + k$  (qualquer que seja o ano de nascimento  $\tau = t - x$ ) é dada por

$${}_k p_x^{(\bar{\tau})} = \prod_{j=0}^{k-1} [1 - q_{x+j}(\bar{\tau} + x + j)], \quad (6.3.13)$$

enquanto que na modalidade 2, esta mesma probabilidade viria dada por

$${}_k p_x^{(s)} = \prod_{j=0}^{k-1} [1 - q_{x+j}(s)]. \quad (6.3.14)$$

Como se depreende, ambos os métodos configuram soluções aproximadas que introduzem distorções relativamente à correcta aplicação de uma tábua projectada. Ainda assim, é possível introduzir ajustamentos nos modelos de modo a minimizar os desvios e a não renunciar à simplicidade implícita no uso de uma única tábua unidimensional. O ajustamento mais conhecido, adoptado nalguns países (e.g., Alemanha, Áustria, Itália, França,...) consiste em reintroduzir, na modalidade 1, alguma dependência das probabilidades face ao ano de nascimento  $\tau$ . Para tal, recorre-se a um mecanismo de correcção da idade desenvolvido por Rueff (1955), conhecido na ciência actuarial por “*age-shift mechanism*”<sup>11</sup> ou correcção de Rueff.

Para uma dada tábua geracional unidimensional  $\{q_x^{(\bar{\tau})}\}$ , em cada ano de nascimento  $\tau$  ( $\tau \lesseqgtr \bar{\tau}$ ) o procedimento preconiza a leitura dos quocientes de mortalidade relativos a cada idade  $x$ ,  $q_x^{(\tau)}$ , na linha que resulta da aplicação de um factor de correcção da idade, i.e.

$$q_x^{(\tau)} = q_{x+h(\tau,x)}^{(\bar{\tau})},$$

onde a função  $h(\tau, x)$  fornece o ajustamento na idade (*age-shift*) considerado necessário para adequar a probabilidade  $q_x^{(\bar{\tau})}$  estimada pela tábua unidimensional  $\{q_x^{(\bar{\tau})}\}$  à projec-

---

<sup>11</sup> *Altersverschiebung* na definição original. Para mais detalhes sobre este procedimento veja-se, por exemplo, Pitacco (2000).

### 6.3. Modelação da mortalidade num contexto dinâmico

tada para a geração nascida no ano  $\bar{\tau}$ . Por exemplo, no caso em que é projectada uma diminuição da mortalidade a função  $h(\tau, x)$  é definida por

$$h(\tau, x) = \begin{cases} > 0 & \text{se } \tau < \bar{\tau} \\ 0 & \text{se } \tau = \bar{\tau} \\ < 0 & \text{se } \tau > \bar{\tau}. \end{cases}$$

Na prática, é usada normalmente uma função de correcção  $h(\tau)$  que não depende da idade mas apenas do ano de nascimento.<sup>12</sup> Segundo este mecanismo de *age-shift*, a probabilidade de um indivíduo nascido no ano  $\tau$  sobreviver até à idade  $x + k$  é calculada por

$${}_k p_x^{[\bar{\tau}; h(\tau)]} = \prod_{j=0}^{k-1} [1 - q_{x+h(\tau)+j}(\bar{\tau} + x + j)]. \quad (6.3.15)$$

#### 6.3.4 Classificação dos métodos de projecção

A literatura sobre métodos de projecção da mortalidade é extensa e diversificada.<sup>13</sup> Não existe, contudo, unanimidade quanto à melhor forma de os classificar no contexto da demografia e da ciência actuarial. Nesta secção, efectuamos uma primeira arrumação dos métodos distinguindo-os entre métodos biomédicos e epidemiológicos, métodos explicativos e métodos extrapolativos. Mais adiante, procedemos a uma análise detalhada dos principais métodos usados na prática.

Os métodos epidemiológicos concentram-se nos factores que determinam o óbito e procuram modelar as taxas de mortalidade numa perspectiva biomédica.<sup>14</sup> Tratam-se de

<sup>12</sup>Por exemplo, o ajustamento preconizado pela tábua italiana RG48 para indivíduos do sexo masculino ( $\tau = 1948$ ) é o seguinte:

| $\tau$      | $h(\tau)$ |
|-------------|-----------|
| 1908 – 1932 | 2         |
| 1933 – 1941 | 1         |
| 1942 – 1951 | 0         |
| 1952 – 1965 | -1        |
| 1966–       | -2        |

Como se observa, a função procede a um envelhecimento das probabilidades para as gerações nascidas antes de  $\bar{\tau} = 1948$ , e a um rejuvenescimento para as gerações nascidas após essa data.

<sup>13</sup>Para uma revisão da literatura sobre esta matéria veja-se, por exemplo, Pollard (1987), Benjamin e Soliman (1993), Tuljapurkar e Boe (1998), Tabeau (2001), GAD (2001), Wong-Fupuy e Haberman (2004) e Booth (2006).

<sup>14</sup>Por exemplo, a hipótese de que o óbito é o resultado da acumulação de “defeitos” conduz-nos a uma descrição matemática da mortalidade.

### 6.3. Modelação da mortalidade num contexto dinâmico

---

métodos que não são, na prática, usados pela generalidade dos organismos oficiais para realizar projecções, mas que se revelam úteis quando é necessário balizar os resultados proporcionados pelos métodos extrapolativos. A sua principal limitação prende-se com o facto da sua eficácia estar condicionada ao conhecimento integral e cabal das múltiplas interações que se estabelecem entre as diferentes causas de morte, algo que os constantes avanços nas ciências médicas poderá tornar possível no futuro.<sup>15</sup>

Os métodos explicativos empregam um mecanismo de projecção com base em relações de causa-efeito recorrendo, por exemplo, a técnicas econométricas baseadas em variáveis de natureza económica ou a factores ambientais. Estes métodos são igualmente pouco usados na elaboração de projecções oficiais, porquanto o seu poder explicativo não está suficientemente consolidado. Uma limitação adicional destes métodos diz respeito ao facto de ser tão ou mais difícil prever as variáveis explicativas do que a própria variável dependente (mortalidade).<sup>16</sup>

Os métodos extrapolativos assentam na projecção para o futuro das tendências históricas registadas na mortalidade e são, na prática, os mais usados. São inúmeros os métodos usados para extrapolar as tendências, a maioria carecendo, ou de informação sobre as principais forças biomédicas que determinam a mortalidade, ou de uma avaliação subjectiva, por exemplo, em relação ao período temporal em que as tendências devem ser captadas. Conforme teremos oportunidade de verificar ao longo do capítulo, as principais diferenças entre os métodos situam-se na forma de sumariar os perfis de mortalidade.

Os métodos paramétricos envolvem o ajustamento de uma determinada lei de mortalidade durante um número consecutivo de anos, e a subsequente extrapolação para o futuro das tendências detectadas no valor estimado dos parâmetros. Os métodos são aplicados, sobretudo, no contexto do chamado “método vertical” de projecção, que em seguida aprofundaremos, mas a sua flexibilidade torna-os igualmente compatíveis com abordagens diagonais visando a construção de tábuas geracionais. A adopção de uma mesma função para descrever a mortalidade no passado e para prever o futuro constitui

---

<sup>15</sup> Com efeito, as causas de morte não são, em muitos casos, independentes entre si, o que dificulta a obtenção de projecções fiáveis. Esta metodologia apresenta ainda outros problemas como sejam a dificuldade em identificar a causa última do óbito, sobretudo nas pessoas idosas, as mudanças nos agrupamentos das causas de morte ou a própria má classificação dos óbitos.

<sup>16</sup> Assinale-se ainda que o facto da metodologia efectuar projecções a partir da análise do impacto produzido pelos vários factores de risco (e.g., rendimento, condição socioeconómica) na mortalidade pode exigir uma desagregação dos dados por subgrupos. Esta decomposição pode revelar-se, na prática, impossível, para além de que é provável que a composição dos subgrupos sofra modificações ao longo do tempo, modificações essas difíceis de vaticinar. Ainda assim, o estudo sobre os efeitos induzidos no passado pelas principais variáveis explicativas pode ajudar a compreender as tendências subjacentes no comportamento da mortalidade.

uma das principais limitações desta abordagem.<sup>17</sup>

Os métodos de projecção de perfis de mortalidade enquadram-se no denominado “método horizontal” e envolvem a extrapolação, desagregada, das tendências observadas nas taxas de mortalidade (ou outros indicadores) relativas a cada idade. Os vários métodos incluídos nesta categoria diferenciam-se, essencialmente, pelo modelo matemático usado para ajustar os dados (exponencial, logarítmico, polinomial, etc.), pela sua natureza eminentemente determinística (i.e., os valores projectados baseiam-se num conjunto pré-determinado de parâmetros) ou estocástica (i.e., os valores projectados contém alguma componente aleatória), pelas hipóteses assumidas relativamente à distribuição estatística do número de óbitos (Normal, Poisson, Binomial, etc.), pela introdução de transformações nas funções base da tábua (logit, função de resistência) ou, por fim, pela necessidade ou não de uma tábua base.

Noutro domínio, os métodos de projecção com referência a um determinado objetivo (*target*) envolvem a interpolação entre as taxas de mortalidade actuais e um conjunto de metas pré-definidas, que se assume serão observadas no futuro. Estes métodos subdividem-se, regra geral, em duas categorias: métodos que envolvem a interpolação em relação a uma tábua limite ou “óptima”, para a qual convergem os aumentos na longevidade verificados ao longo do tempo, e métodos que encerram a interpolação em relação a uma tábua “modelo”.

Uma outra categoria refere-se aos métodos de projecção por causa de morte, que pode envolver a utilização de técnicas incluídas em qualquer umas das três abordagens anteriores. A extrapolação das tendências nas taxas de mortalidade por causa de morte tem vindo a ser defendida com cada vez maior insistência, porquanto ela possibilita um melhor conhecimento sobre a dinâmica da mortalidade (Gutterman e Vanderhoof, 1998). A utilização desta abordagem não se encontra, no entanto, isenta de problemas.

#### 6.3.5 Risco de longevidade

As tendências observadas na longevidade humana devem ser acomodadas ao nível do *pricing* e do cálculo das responsabilidades financeiras, em particular nos produtos que incluem benefícios de longo prazo em caso de vida (e.g., rendas vitalícias). Com efeito, o cálculo do valor actual esperado dos cash flows futuros dos produtos requer uma projecção adequada das variáveis demográficas, tendo em vista a minimização dos riscos de subestimação das responsabilidades com a carteira. Como vimos na secção anterior,

---

<sup>17</sup>Uma alternativa passa pelo uso de métodos não-paramétricos, que preservam a natureza extrapolativa mas são mais flexíveis na medida em que não dependem de hipóteses sobre a lei analítica que governa a mortalidade.

### 6.3. Modelação da mortalidade num contexto dinâmico

---

são diversos os modelos propostos para esta finalidade e usados na prática actuarial corrente.

A utilização de modelos dinâmicos não elimina, todavia, a possibilidade de ocorrerem desvios face aos valores projectados. Esses desvios podem ficar a dever-se a quatro fontes de risco distintas:

- uma primeira, associada à própria natureza estocástica do fenómeno (número de óbitos), que faz com que em cada ano de calendário  $t$  se registem flutuações aleatórias nas taxas de mortalidade observadas face aos respectivos valores esperados (i.e., projectados). O risco de flutuações aleatórias é um risco bem conhecido na indústria seguradora, em ambos os ramos (vida e não-vida), sendo usualmente designado por “risco de processo”. Um dos resultados fundamentais na teoria do risco estipula que a severidade deste tipo de risco diminui à medida que a dimensão da carteira (homogénea) aumenta, ou seja, corresponde a um denominado “*pooling risk*”.
- uma segunda, relacionada com o facto de a tendência futura da mortalidade ser, ela própria, aleatória, o que significa que independentemente do modelo adoptado poderão sempre ocorrer desvios sistemáticos face à mortalidade projectada. Os desvios são explicados, neste caso, pela incerteza em torno dos valores dos parâmetros do modelo de projecção (o chamado “risco de parâmetro”) e pela incerteza relativamente ao próprio modelo seleccionado para captar as tendências na mortalidade (o denominado “risco de modelo”). O risco de desvios sistemáticos não pode ser eliminado aumentando a dimensão da carteira. Pelo contrário, o seu impacto financeiro aumenta à medida que a carteira cresce na medida em que os desvios afectam todas pessoas seguras (ou beneficiários de rendas vitalícias e de pensões) numa carteira da mesma forma.
- uma terceira, correspondente à denominada mortalidade catastrófica, ou seja, aos desvios associados a “estados da natureza” extremos (e.g., conflitos militares, sismos, inundações, seca).
- uma quarta, relativa ao chamado risco de base, i.e., ao risco de que o comportamento da mortalidade na população que serviu de base à estimação dos modelos seja substancialmente diferente da observada na carteira da seguradora.

A avaliação da incerteza em relação às tendências futuras constitui um output de alguns dos modelos estatísticos usados para analisar o comportamento passado e para

projectar a mortalidade. Num contexto mais geral, quando ignoramos a análise de risco proveniente do modelo estatístico e nos concentramos nos desvios sistemáticos face à estimativa (pontual) da mortalidade futura, em especial nas idades adultas e mais avançadas, é usual referirmo-nos a este tipo de risco como traduzindo um “risco de longevidade” (Marocco e Pitacco, 1998; Olivieri, 2001).

A rectangularização da função de sobrevivência, com o conseqüente aumento da concentração dos óbitos em torno da moda da curva de óbitos, tende a diminuir o risco de flutuações aleatórias nas taxas de mortalidade, independentemente da localização da moda da distribuição. Esta trajectória no comportamento da mortalidade poderia levar-nos a concluir que, a prazo, a natureza aleatória do fenómeno se esvaneceria e os riscos envolvidos se tornariam negligenciáveis. No entanto, devido à incerteza relativamente à localização da própria moda da distribuição, uma nova fonte de risco, que apelidámos de risco de longevidade, atribuível aos desvios sistemáticos da mortalidade face aos valores projectados, emerge e ganha importância. A incorporação do risco de longevidade no cálculo actuarial revela-se, deste modo, de vital importância, visando uma correcta avaliação das responsabilidades financeiras futuras das empresas.

### 6.4 Projecção mediante métodos paramétricos

Os métodos paramétricos de projecção envolvem o ajustamento de uma determinada lei de mortalidade durante um número consecutivo de anos, e a subsequente extrapolação do valor estimado dos parâmetros para o futuro. Desta forma, conseguem-se obter, de forma indirecta, estimativas para a medida de mortalidade seleccionada. Em termos formais, seja  $\chi = \{x_{\min}, \dots, x_{\max}\}$  o intervalo de idades considerado e  $\zeta = \{t_1, t_2, \dots, t_n\}$  o período temporal das observações. A base de dados necessária à elaboração de projecções é definida pela matriz

$$\{\mu_x(t)\}_{x \in \chi; t \in \zeta} = \{\mu_x(t_1), \mu_x(t_2), \dots, \mu_x(t_n)\}_{x \in \chi}.$$

Considere-se uma dada lei representando, por exemplo, a força de mortalidade

$$\mu_x = \phi(x; \underbrace{\alpha, \beta, \dots}_{\text{parâmetros}}).$$

Num contexto dinâmico, a introdução do tempo cronológico  $t$  na projecção de  $\mu_x$  é

#### 6.4. Projecção mediante métodos paramétricos

feita de forma indirecta, via parâmetros da lei de mortalidade, ou seja,

$$\mu_x(t) = \phi(x; \alpha(t), \beta(t), \dots).$$

A obtenção de projecções é então baseada no denominado “*método vertical*”, representado na Figura 6.4.1. Como se observa, este método requer, numa primeira etapa, a estimação dos parâmetros do modelo em cada um dos anos  $t_i$  ( $i = 1, \dots, n$ ) da amostra, i.e.,

$$\mu_x(t_i) = \phi(x; \alpha_i, \beta_i, \dots),$$

obtendo-se em resultado séries temporais de parâmetros  $\{\alpha_i; \beta_i; \dots\}_{i=1, \dots, n}$ , indexados por ano cronológico.

Numa segunda fase, é aplicada a metodologia de Box-Jenkins às séries temporais no sentido de identificar o modelo ARIMA( $p, d, q$ ) adequado. Estimados os coeficientes do modelo dinâmico, é possível projectar o valor futuro (no ano  $n + t$ ) dos parâmetros ( $\alpha_{n+t}, \beta_{n+t}, \dots$ ), após o que, substituindo na lei de mortalidade, estamos em condições de projectar os indicadores de mortalidade.

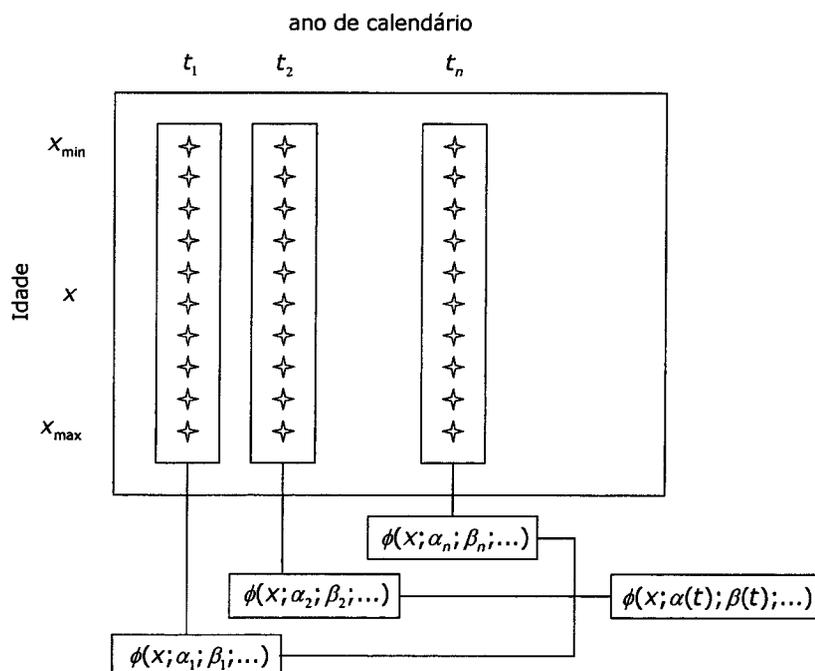


Figura 6.4.1: Método vertical

#### 6.4. Projecção mediante métodos paramétricos

Entre os modelos pioneiros neste domínio, destaca-se o proposto por Blaschke (1923), fundado na popular lei de Makeham e definido por

$$\mu_x(t) = \gamma(t) + \alpha(t)\beta(t)^x, \quad (6.4.1)$$

onde os três parâmetros são dependentes do ano de calendário.

Uma proposta alternativa é sugerida pelo CMIB (1976), com base na seguinte lei para o “odds ratio”

$$\frac{q_x(t)}{p_x(t)} = \exp[P_x(t)], \quad (6.4.2)$$

onde  $P_x(t)$  denota um polinómio de grau  $k$ . Consideremos agora um factor de extrapolação exponencial, i.e.,

$$\frac{q_x(t)}{p_x(t)} = \frac{q_x(t_0)}{p_x(t_0)} r^t, \quad (6.4.3)$$

onde  $t_0$  designa o ano base,  $t > t_0$  e  $r < 1$ . Se admitirmos um modelo linear para o ano base, i.e.,

$$P_x(t_0) = \ln\left(\frac{q_x(t_0)}{p_x(t_0)}\right) = \alpha + \beta x,$$

e substituírmos em (6.4.3), obtemos

$$\begin{aligned} \ln\left(\frac{q_x(t)}{p_x(t)}\right) &= \alpha + \beta x + t \ln r \\ &= \alpha + \beta(x - wt) = P_{x-wt}(t_0), \end{aligned} \quad (6.4.4)$$

onde  $w = \frac{-\ln r}{\beta}$ . O andamento das probabilidades  $q_x$  e  $p_x$  permite inferir que  $\beta > 0$ . Por hipótese  $r < 1$ , donde resulta que  $w > 0$ . Daqui se conclui que a aplicação de um coeficiente de redução constante ao rácio  $q_x/p_x$  equivale, na prática, a um rejuvenescimento  $w$  na idade (“age-shift”) para cada um dos anos projectados  $t$ .<sup>18</sup> Se convertermos este resultado em termos de probabilidades, obtemos

$$q_x(t) \cong q_{x-wt}(t_0).$$

O popular modelo de Heligman-Pollard (HP), descrito em detalhe no Capítulo 3, constitui outro exemplo profícuo da aplicação de métodos paramétricos na projecção da mortalidade (Heligman e Pollard, 1980; Forfar e Smith, 1987; McNown e Rogers, 1989; Benjamin e Soliman, 1993; Felipe *et al.*, 2002).<sup>19</sup>

<sup>18</sup>Um modelo deste tipo foi aplicado no Reino Unido (CMIB, 1978) considerando  $w \cong 1/20$ .

<sup>19</sup>Para uma aplicação do modelo logístico no contexto dos GLM's veja-se Butt e Haberman (2004).

#### 6.4. Projecção mediante métodos paramétricos

A adaptação do modelo envolve o ajustamento, em cada um dos anos que compõem a amostra, de uma das leis de HP, por exemplo a segunda

$$q_x(t, \theta_t) = A_t^{(x+B_t)C_t} + D_t \exp \left[ -E_t (\ln x - \ln F_t)^2 \right] + \frac{G_t H_t^x}{1 + K_t G_t H_t^x}, \quad (6.4.5)$$

onde  $\theta_t = (A_t, \dots, H_t)$  representa o vector de parâmetros a estimar com recurso ao método dos mínimos quadrados ponderados não lineares.<sup>20</sup>

Um outro exemplo interessante é proposto por Menthonnex e Wanner (1998), baseado no integral da força de mortalidade

$$R_x(t) = -\ln \{l_x(t)\} = \int_0^x \mu_\xi(t) d\xi, \quad (6.4.6)$$

com

$$R_x(t) = b_0(t) + b_1(t)x + b_2(t)\{b_3(t)\}^x, \quad (6.4.7)$$

onde  $l_x(t)$  denota a função de sobrevivência. Esta formulação, próxima da de Makeham, inclui quatro parâmetros dependentes do tempo de calendário, com a seguinte interpretação:  $b_0(\cdot)$  representa a taxa de mortalidade infantil,  $b_1(\cdot)$  traduz o coeficiente de proporcionalidade que acumula o risco ao qual estão igualmente expostos jovens e idosos<sup>21</sup>,  $b_2(t)$  tem um efeito de intensidade do risco de mortalidade ligado ao envelhecimento e  $b_3(t)$  exerce um efeito de aceleração do risco com o avanço da idade, uma vez que está directamente ligado à curvatura da função  $R_x(t)$ .

Um dos principais atractivos dos métodos paramétricos prende-se com a flexibilidade que estes oferecem na análise da evolução do fenómeno e na produção de intervalos de confiança que reflectam as principais tendências na mortalidade das populações. Por exemplo, se admitirmos que a função de densidade de probabilidade da vida residual de um indivíduo à nascença é representada pela lei de Weibull (2.5.47), a correspondente força de mortalidade é definida, num contexto dinâmico, por

$$\mu_x(t) = \frac{\alpha(t)}{\beta(t)} \left( \frac{x}{\beta(t)} \right)^{\alpha(t)-1}. \quad (6.4.8)$$

O facto de os valores característicos (média, moda, variância) da distribuição de pro-

<sup>20</sup> Conforme demonstrámos no Capítulo 5, a qualidade do ajustamento é sensível à escolha da função de perda. Entre os outros aspectos a ter em conta na estimação dos parâmetros, merecem atenção especial a escolha dos valores iniciais e a selecção do intervalo de idades. A previsão no modelo pressupõe, tal como anteriormente, a extrapolação das tendências detectadas nas séries temporais relativas a cada um dos parâmetros da lei de mortalidade.

<sup>21</sup> Este parâmetro é particularmente influenciado pelo número de óbitos registado nas idades jovens, principalmente os atribuíveis aos acidentes ou suicídios.

## 6.4. Projecção mediante métodos paramétricos

babilidade poderem, neste caso, ser expressos em termos analíticos permite, mediante a formulação de cenários para os parâmetros  $\alpha(t)$  e  $\beta(t)$ , projectar funções de mortalidade que espelhem, por exemplo, os fenómenos de expansão e de rectangularização mencionados na Secção 6.2.

A flexibilidade dos métodos paramétricos de projecção torna-os facilmente compatíveis com uma abordagem “diagonal”, especialmente adequada quando se pretende obter tábuas de mortalidade geracionais. O procedimento de projecção, ilustrado na Figura 6.4.2, é muito semelhante ao acima descrito para o método vertical, com a única diferença de que os parâmetros da lei analítica são agora indexados pelo ano de nascimento da geração.

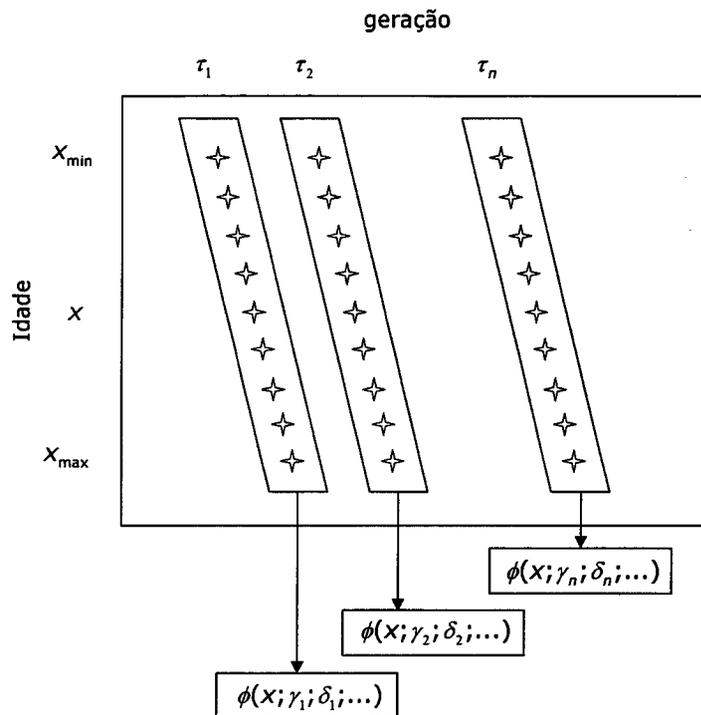


Figura 6.4.2: Método diagonal

Um dos primeiros exemplos desta abordagem é devido a Davidson e Reid (1927), que propõem um modelo baseado na lei de Makeham

$$\mu_x(\tau) = \gamma(\tau) + \alpha(\tau)\beta(\tau)^x, \quad (6.4.9)$$

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

onde  $\tau = t - x$  denota o ano de nascimento.<sup>22</sup>

Ainda neste contexto, Pollard (1949) sugere uma abordagem interessante onde admite que a força de mortalidade depende da idade alcançada  $x$  e do ano de nascimento  $\tau$ , num comportamento ditado pela equação

$$\mu_x(x + \tau) = Q(x) R(\tau), \quad (6.4.10)$$

onde  $Q(x)$  e  $R(\tau)$  são funções que dependem exclusivamente da idade e do ano de nascimento, respectivamente.

A principal vantagem dos métodos paramétricos refere-se, porventura, à redução na dimensão do problema de projecção.<sup>23</sup> Com efeito, o problema cinge-se, neste contexto, à extrapolação das tendências temporais de um conjunto restrito de parâmetros e não de um intervalo alargado de idades. Em contrapartida, o facto do método assentar num procedimento de extrapolação univariada, quando se sabe que os parâmetros estimados são normalmente fortemente dependentes, constitui uma limitação séria desta abordagem, que pode conduzir a projecções de mortalidade inverosímeis. As próprias estimativas anuais dos parâmetros do modelo são normalmente bastante instáveis, o que aumenta a volatilidade das projecções.<sup>24</sup>

Por outro lado, a adopção de uma lei particular introduz uma fonte de risco adicional na análise, na medida em que esta pode não ser a que melhor representa o fenómeno em todos os anos considerados. A escolha da função paramétrica condiciona ainda a análise às idades em que esta é aplicável. Por fim, o facto do método não considerar informação colateral (e.g., causas de morte, epidemias) na previsão constitui uma restrição relevante.

## 6.5 Projecção por extrapolação dos perfis de mortalidade

O método de projecção por extrapolação dos perfis de mortalidade assenta no pressuposto simples de que as tendências verificadas no passado podem, com recurso a um determinado modelo, prolongar-se no futuro. A extrapolação dos perfis  $q_x$  representa, na

---

<sup>22</sup>Na implementação prática do modelo os autores admitem que  $\beta(\tau) = \beta \forall \tau$ , enquanto que os parâmetros  $\gamma(\tau)$  e  $\alpha(\tau)$  são estimados através do método diagonal. Cramer e Wold (1935) discutem a aplicação desta lei na graduação e projecção de taxas de mortalidade da população Sueca comparando, em particular, o desempenho dos métodos vertical e diagonal.

<sup>23</sup>As vantagens dos métodos paramétricos são analisadas em detalhe no Capítulo 3.

<sup>24</sup>Uma das respostas possíveis a este problema passa pelo recurso a uma abordagem de natureza multivariada, uma solução que sendo teoricamente possível, é de difícil e complexa implementação. Outra alternativa envolve, como veremos mais adiante no capítulo, a consideração simultânea da idade actuarial e do tempo cronológico na estimação dos parâmetros do modelo.

## 6.5. Projecção por extrapolação dos perfis de mortalidade

realidade, um caso particular do chamado “*método horizontal*” de projecção, ilustrado na Figura 6.5.1.<sup>25</sup>

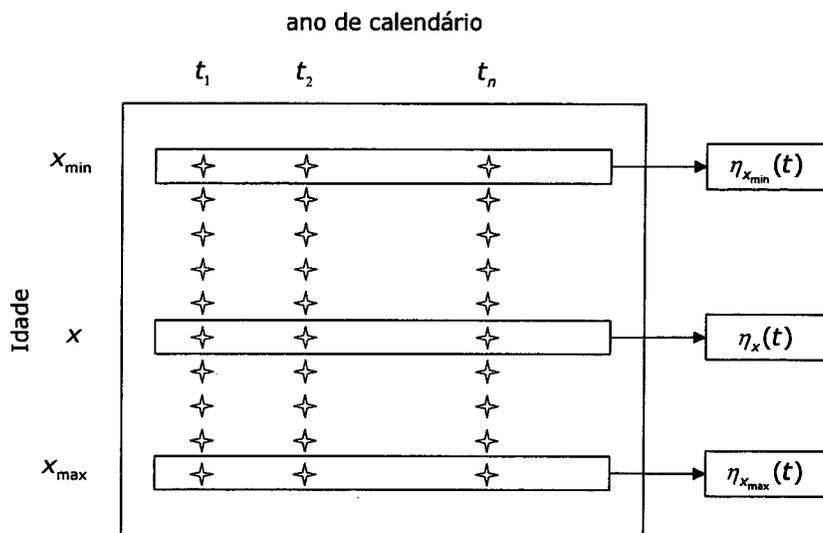


Figura 6.5.1: Método horizontal

Os diferentes modelos incluídos nesta categoria distinguem-se, desde logo, pelo facto de recorrerem ou não a uma determinada tábua base. Em termos gerais, o modelo de projecção fundado numa tábua base (ou de momento) pode ser definido por

$$q_{x,t} = q_{x,t_0} \times RF(x,t), \quad (6.5.1)$$

onde  $RF(x,t)$  representa a relação entre os quocientes de mortalidade para a idade  $x$  no ano base  $t_0$  e num determinado ano futuro  $t$  ( $t > t_0$ ). Para evitar a confusão entre a notação usada no contexto dos modelos dinâmicos e a adoptada no Capítulo 9 a propósito dos modelos de mortalidade estocástica adoptaremos, daqui em diante, e salvo menção em contrário, o índice inferior  $t$  para expressar a dependência temporal das variáveis.

O termo  $RF(x,t)$  é designado por factor de redução (se  $RF(x,t) < 1$ , caso contrário é chamado de “factor de variação”) da mortalidade. Uma das estruturas mais simples é aquela em que o factor de redução para cada idade depende apenas da distância temporal

<sup>25</sup>O método horizontal de projecção pode ser aplicado a outros indicadores de mortalidade como sejam o ratio  $q_x/p_x$ , a taxa de mortalidade  $m_x$ , a esperança de vida  $e_x$ , etc.

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

entre o ano base e o ano projectado, i.e.,

$$RF(x, t_0, t) = RF(x, t - t_0).$$

No caso particular em que o factor de redução é independente da idade, temos

$$RF(x, t_0, t) = RF(t - t_0).$$

### 6.5.1 Modelo exponencial (ou geométrico)

O modelo exponencial ou geométrico foi desenvolvido por Nolfi (1959) e adoptado, por exemplo, na construção das tábuas de mortalidade suíças (GRM/F 1995), espanholas (PERM/F 2000) e austríacas (AVÖ 2005R). O método admite, em cada idade  $x$ , que a mortalidade decresce de forma exponencial segundo a relação

$$q_{x,t} = q_{x,t_0} e^{-\lambda_x(t-t_0)}, \quad (6.5.2)$$

onde  $q_{x,t_0}$  denota a mortalidade base,  $e^{-\lambda_x(t-t_0)}$  o factor de redução e  $\lambda_x > 0$  representa a intensidade de decréscimo anual da mortalidade, um parâmetro estimado a partir das tendências registadas no passado (e.g., através do método dos mínimos quadrados). As hipóteses do modelo pressupõem, na prática, que a diferença temporal entre os logaritmos dos quocientes é linear (i.e.,  $\ln q_{x,t} - \ln q_{x,t_0} = -\lambda_x(t - t_0)$ ).<sup>26</sup>

O modelo exponencial descrito pela equação (6.5.2) estabelece que, no caso de  $\lambda_x > 0$ , os quocientes  $q_{x,t}$  tendem, em todas as idades, para zero (i.e.,  $q_{x,\infty} = 0$ ), um comportamento assintótico inverosímil. Neste sentido, embora se admita que as projecções realizadas com base no modelo apenas podem ser consideradas aceitáveis num horizonte temporal de projecção limitado, parece-nos mais realista a presunção de que a mortalidade assintótica deve assumir um valor positivo. Uma generalização simples da equação (6.5.2) permite-nos atingir este objectivo

$$q_{x,t} = q_{x,t_0} \left[ \alpha_x + (1 - \alpha_x) e^{-\lambda_x(t-t_0)} \right], \quad (6.5.3)$$

onde,  $\forall x$ ,  $0 < \alpha_x \leq 1$  e  $\lambda_x > 0$ . No caso particular em que  $\alpha_x = 0$ , obtém-se o modelo base (6.5.2). Para  $\alpha_x > 0$ , o modelo projecta uma mortalidade limite positiva,

---

<sup>26</sup>A fórmula proposta em 1929 pelo actuário alemão Sachs (citado por Pitacco (2004)) representa um caso particular de (6.5.2)

$$q_{x,t} = q_{x,t_0} \left( A \frac{(t-t_0)}{(x+B)} \right),$$

onde  $A$  e  $B$  são parâmetros constantes e  $A^{(t-t_0)/(x+B)}$  desempenha o papel de factor de redução.

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

correspondente a uma percentagem  $\alpha_x$  do valor de partida, i.e.,

$$q_{x,\infty} = \lim_{t \rightarrow \infty} q_{x,t} = \alpha_x \cdot q_{x,t_0}.$$

O parâmetro  $\lambda_x$  determina, neste caso, a velocidade a que o quociente  $q_{x,t}$  decresce para o seu valor limite.

O modelo (6.5.3) é susceptível de uma interessante versão alternativa, que combina informação objectiva com opiniões qualitativas avalizadas sobre o andamento futuro da mortalidade. Seja  $RF(x, t)$  o factor de redução à idade  $x$  para um horizonte de projecção de  $t$  anos

$$RF(x, t) = \frac{q_{x,t_0+t}}{q_{x,t_0}} = \alpha_x + (1 - \alpha_x) e^{-\lambda_x(t-t_0)}. \quad (6.5.4)$$

A redução total (assintótica) na mortalidade, contada a partir do ano base, é dada por  $(q_{x,t_0} - q_{x,\infty})$ , enquanto que a mesma diminuição nos primeiros  $m$  anos é expressa por  $(q_{x,t_0} - q_{x,t_0+m})$ . Seja  $\rho_x(m)$  a fracção da diminuição total da mortalidade registada nos primeiros  $m$  anos, i.e.,

$$\rho_x(m) = \frac{q_{x,t_0} - q_{x,t_0+m}}{q_{x,t_0} - q_{x,\infty}}. \quad (6.5.5)$$

Se dividirmos o numerador e o denominador da equação (6.5.5) por  $q_{x,t_0}$ , obtemos

$$\rho_x(m) = \frac{1 - RF(x, m)}{1 - RF(x, \infty)} = \frac{(1 - \alpha_x)(1 - e^{-\lambda_x m})}{1 - \alpha_x} = 1 - e^{-\lambda_x m}. \quad (6.5.6)$$

Se admitirmos que  $\lambda_x = \lambda, \forall x$ , temos que  $\rho_x(m) = \rho(m)$ , donde resulta

$$e^{-\lambda} = [1 - \rho(m)]^{\frac{1}{m}}. \quad (6.5.7)$$

Em suma, a escolha de  $(m, \rho(m))$  determina univocamente o valor do factor de redução. Por fim, se substituirmos esta última equação em (6.5.4) podemos escrever

$$RF(x, t) = \alpha_x + (1 - \alpha_x) [1 - \rho(m)]^{\frac{(t-t_0)}{m}}. \quad (6.5.8)$$

Por exemplo, se supusermos que 50% do declínio na mortalidade ocorre nos primeiros 20 anos de projecção, i.e.,  $(m, \rho(m)) = (20, 0.5)$ , obtemos um factor de redução anual igual a  $e^{-\lambda} = 0.5^{\frac{1}{20}} = 0.96594$ , donde se conclui que  $\lambda = 0.034657$ . Esta versão do modelo com mortalidade assintótica é adoptada, como veremos em seguida, no Reino Unido (CMIB, 1990, 1999).

### 6.5.2 Metodologia do CMIB

No Reino Unido, o Continuous Mortality Investigation Bureau (CMIB) efectua periodicamente projecções para os factores de redução da mortalidade das populações de pensionistas e de beneficiários de rendas vitalícias. O procedimento desenrola-se em duas fases. Numa primeira etapa, as estimativas brutas relativas a um determinado período de investigação são graduadas com base no procedimento descrito na Secção 3.4 do Capítulo 3, usando funções *Gompertz-Makeham de tipo*  $(r, s)$ . Em resultado, são construídas tábuas contemporâneas. Numa segunda etapa, são construídas tábuas de mortalidade projectadas pela aplicação de factores de redução  $RF(x, t)$ , estimados essencialmente com base no comportamento passado das séries, mas com a possibilidade da introdução de juízos subjectivos sobre as trajectórias futuras esperadas para a mortalidade.

Seja  $t$  o tempo cronológico, com origem em  $t = 0$ , situado no centro do período base, e  $q_{x,t}$  a probabilidade de óbito entre as idades  $x$  e  $x + 1$  no momento  $t \geq 0$ . O factor de redução é definido pela equação

$$q_{x,t} = q_{x,0} \times RF(x, t), \quad t \geq 0, \quad (6.5.9)$$

onde  $q_{x,0}$  denota o valor no ano base.

A primeira fórmula usada no Reino Unido (CMIB, 1978) para a projecção da mortalidade dos pensionistas e dos beneficiários de rendas vitalícias ( $x \geq 60$ ) adoptava uma estrutura exponencial com mortalidade limite nula. O modelo era definido por

$$q_{x,t} = q_{x,0} (r^t), \quad \text{com } RF(x, t) = r^t = (0.9)^{\frac{t}{20}}, \quad (6.5.10)$$

e tábuas base relativas à experiência observada no período 1967-1970 (tábuas designadas por  $PA(90)$  e  $a(90)$ ).

Como se constata, o modelo considera reduções uniformes na mortalidade futura, independentes da idade atingida pelo indivíduo.<sup>27</sup> Estudos subsequentes realizados no Reino Unido puseram em evidência uma correlação elevada entre os factores de redução e a idade. Em consequência, o CMIB propôs um novo modelo de projecção (CMIB, 1990), com tábuas base relativas à experiência observada no quinquénio 1978-1982, com centro em 1980. O modelo assentava numa definição para o factor de redução derivada da equação (6.5.8)

$$RF(x, t) = \alpha(x) + [1 - \alpha(x)] (1 - f_m)^{\frac{t}{m}}, \quad (6.5.11)$$

---

<sup>27</sup> As aplicações efectuadas com base no modelo permitiram concluir que o efeito acumulado das reduções na mortalidade ao fim de 20 anos seria equivalente ao rejuvenescimento de um ano na idade na tábua base, i.e.,  $q_{x,t} = q_{x-t/20,0}$ .

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

com  $m = 20$ ,  $f_m = 0.60$  e

$$\alpha(x) = \begin{cases} 0.50 & \text{se } x < 60 \\ \frac{x-10}{100} & \text{se } 60 \leq x \leq 110 \\ 1 & \text{se } x > 110. \end{cases}$$

Como se observa, o modelo estipula que 60% do decréscimo na mortalidade ocorre nos primeiros 20 anos. Por outro lado, estabelece que a mortalidade limite, determinada pelo parâmetro  $\alpha(x)$ , cresce linearmente com a idade entre os 60 e os 110 anos, com valores fixos de 0.5 nas idades inferiores a 60 anos, e de 1 nas idades superiores a 110 anos. Em consequência, nas idades compreendidas entre os 60 e os 110 anos o factor de redução depende simultaneamente da idade e do ano cronológico.

O CMIB reestruturou, recentemente, o seu modelo de projecção (CMIB, 1999), passando a usar tábuas base relativas à experiência observada no quadriénio 1991-1994, com centro em 1992. O novo modelo reformula igualmente o factor de redução, que é agora definido por

$$RF(x, t) = \alpha(x) + [1 - \alpha(x)] [1 - f_m(x)]^{\frac{t}{m}}, \quad (6.5.12)$$

onde

$$\alpha(x) = \begin{cases} c & \text{se } x < 60 \\ 1 + (1 - c) \frac{(x-110)}{50} & \text{se } 60 \leq x \leq 110 \\ 1 & \text{se } x > 110 \end{cases}$$

e

$$f_m(x) = \begin{cases} h & \text{se } x < 60 \\ \frac{(110-x)h + (x-60)k}{50} & \text{se } 60 \leq x \leq 110 \\ k & \text{se } x > 110. \end{cases}$$

com  $m = 20$ ,  $c = 0.13$ ,  $h = 0.55$  e  $k = 0.29$ . Os parâmetros  $c$ ,  $h$  e  $k$  são estimados de modo a reflectir as tendências passadas no comportamento da mortalidade e incorporam opiniões subjectivas quanto à sua evolução futura, tendo sido ajustados de modo a tomar o ano de 1992 como ano base.

Como se observa, apesar de continuar a assumir que  $m = 20$ , o modelo apresenta diferenças significativas nos parâmetros. Com efeito, em contraste com os factores de redução definidos em (6.5.11), o modelo determina agora que a velocidade de convergência dos quocientes de mortalidade para o seu valor limite (ditada pela função  $f_m(x)$ ) depende da idade do indivíduo. De facto, o parâmetro  $f_m(x)$  varia linearmente entre 0.55 e 0.29 para  $x \in [60, 110]$ . Abaixo dos 60 anos e acima dos 110 as funções assumem valores

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

fixos.<sup>28</sup> O modelo de projecção adoptado pelo CMIB foi criticado por Willets (2004) pelo facto de não incorporar os efeitos geracionais na fórmula. O autor recomenda ainda a estimação dos factores de redução a partir do comportamento observado na totalidade da população do país.

Nos Estados Unidos da América (E.U.A.), a Society of Actuaries Group Annuity Valuation Task Force (1995) propôs, pela primeira vez, um método de projecção aquando da publicação das tábuas de mortalidade GAR94 relativas aos beneficiários de rendas vitalícias. O modelo, que assenta num mecanismo de progressão geométrica para os factores de redução relativos a cada idade, é definido pela equação

$$q_{x,t} = q_{x,1994} \cdot (1 - AA_x)^{t-1994}, \quad (6.5.13)$$

onde  $AA_x$  representa, de acordo com a notação original, o factor anual de redução na mortalidade para a idade  $x$ , e as demais variáveis assumem o seu significado habitual. As tábuas base, graduadas segundo o método de Whittaker-Henderson descrito no Capítulo 4, são relativas ao ano base de 1994.

Os valores de  $AA_x$  são tabulados e respeitam os seguintes limites

$$\begin{cases} 0 < AA_x \leq 0.02, & 1 \leq x \leq 100 \\ 0 & x > 100. \end{cases}$$

Uma reformulação simples da equação (6.5.13) segundo

$$q_{x,t} = q_{x,1994} \cdot e^{(t-1994) \log(1-AA_x)} \quad (6.5.14)$$

permite-nos facilmente concluir que (6.5.13) assume, à semelhança do modelo exponencial, uma relação log-linear entre os quocientes de mortalidade e o tempo cronológico.

Os modelos de projecção com recurso a factores de redução constituem, como vimos, uma solução prática bem aceite pelos actuários, razão pela qual mereceram nos últimos anos uma atenção especial por parte dos investigadores. Na Secção 6.8 são discutidos os desenvolvimentos mais recentes na modelação dos factores de redução.

---

<sup>28</sup>Demonstra-se facilmente que, quer no que se refere aos valores limite, quer no que respeita às percentagens de redução verificadas nos primeiros 20 anos, a nova fórmula incorpora aumentos na longevidade a uma velocidade superior à implícita na anterior fórmula. Nas idades compreendidas entre os 60 e os 110 anos, o factor de redução continua a depender simultaneamente da idade e do ano cronológico. Saliente-se, por fim, que o CMIB recomendou a aplicação dos mesmos factores de redução a todas as experiências analisadas e aos indivíduos de ambos os sexos.

### 6.5.3 Modelos de projecção sem tábua base

Os modelos descritos até ao momento fundam-se, grosso modo, na aplicação do modelo exponencial a uma tábua base mediante o recurso a diferentes definições para os factores de redução. Os modelos analisados nesta secção baseiam-se, pelo contrário, em procedimentos de projecção dos perfis de mortalidade que fazem uso de formulações que não têm em conta, de forma explícita, o valor dos indicadores de mortalidade no ano base. Dentro desta categoria, podemos ainda incluir o rudimentar método de extrapolação gráfica, que envolve a mera extrapolação do comportamento dos perfis a partir do ajuste gráfico das suas tendências passadas.<sup>29</sup>

#### Modelo logarítmico

O modelo logarítmico, analisado em detalhe por Benjamin e Soliman (1993) e Felipe e Guillén (1999), é definido por uma equação que relaciona directamente os perfis de mortalidade e o tempo cronológico  $t$  através de um conjunto restrito de parâmetros específicos para cada idade

$$q_{x,t} = \alpha_x (\beta_x)^t, \quad (6.5.15)$$

onde  $\alpha_x$  denota o nível base (inicial) de  $q_x$  e  $\beta_x$  representa a sua taxa anual de variação. Aplicando a transformação logarítmica, obtemos

$$\ln q_{x,t} = \ln(\alpha_x) + t \ln(\beta_x) \quad (6.5.16)$$

Como se observa, o modelo estipula que a redução anual em  $\ln q_{x,t}$  é constante, uma hipótese admissível em intervalos de tempo curtos, mas inverosímil no longo prazo. Os parâmetros  $\alpha_x$  e  $\beta_x$  dependem da idade e são estimados usando critérios de optimização convencionais. A obtenção de valores projectados para  $q_{x,t}$  pressupõe a substituição dos parâmetros estimados e a consideração de um valor para  $t$  posterior ao período de estimação.<sup>30</sup>

Os principais inconvenientes do modelo estão relacionados, em primeiro lugar, com facto de este menosprezar a possibilidade de ocorrerem alterações estruturais (e.g., epidemias, avanços médicos) no período em análise. Em segundo lugar, e dadas as características do fenómeno em estudo, a (provável) presença de autocorrelação no erro

---

<sup>29</sup>Para mais detalhes sobre este método veja-se, por exemplo, Benjamin e Pollard (1980).

<sup>30</sup>Nalguns casos, a aplicação do modelo é precedida de um procedimento de gradação das séries relativas a cada idade, aplicando um dos métodos não paramétricos descritos no Capítulo 4. A título de exemplo, o método logarítmico com séries alisadas mediante a técnica de splines foi adoptado na construção das tábuas alemãs DAV1994R.

## 6.5. Projecção por extrapolação dos perfis de mortalidade

---

aumenta a variância dos estimadores perturbando, deste modo, a fidelidade das projecções realizadas. Em terceiro lugar, é perfeitamente plausível que o termo aleatório do modelo para diferentes idades seja afectado por um mesmo factor, pelo menos em intervalos de idades contíguos.<sup>31</sup> Por fim, o procedimento de alisamento da série retira informação ao modelo, e a sua aplicabilidade é restringida aos casos em que as séries têm uma dimensão temporal alargada.

### Outros modelos

Entre as múltiplas formulações sugeridas neste contexto, destaque para o modelo de índole exponencial proposto em Cramér e Wold (1935), definido pela equação

$$q_{x,t} = a_x + b_x c_x^t, \quad (6.5.17)$$

onde  $a_x$ ,  $b_x$  e  $c_x$  são parâmetros a estimar com base nos perfis observados. Outros exemplos propostos no contexto actuarial incluem o modelo hiperbólico

$$q_{x,t} = a_x + \frac{b_x}{t} \quad (6.5.18)$$

e o modelo polinomial

$$q_{x,t} = \sum_{k=0}^p a_{x,k} t^k \quad (6.5.19)$$

onde, para  $p = 1$ , se obtém um modelo de extrapolação linear

$$q_{x,t} = a_{x,0} + a_{x,1}t,$$

com  $a_{x,1} < 0$ , de modo a expressar declínios na mortalidade.<sup>32</sup>

Outro modelo, bastante utilizado, recorre a uma função baseada em exponenciais de polinómio

$$q_{x,t} = \frac{e^{G_x(t)}}{1 + e^{G_x(t)}} \quad \text{com} \quad G_x(t) = \sum_{k=0}^s c_{x,k} t^k. \quad (6.5.20)$$

Se admitirmos que  $G_x(t)$  é linear em  $t$  (i.e.,  $s = 1$ ) e aplicarmos a transformação

---

<sup>31</sup>Este facto coloca-nos perante um conjunto de regressões aparentemente não relacionadas, para as quais a estimação mais eficiente é aquela que é conseguida com recurso a um modelo de equações simultâneas e não com recurso ao método dos mínimos quadrados ordinários.

<sup>32</sup>Note-se, no entanto, que a aplicabilidade deste modelo é limitada a intervalos de previsão reduzidos na medida em que, para horizontes longínquos, se obtém  $q_{x,t} < 0$ .

*logit*, podemos expressar o modelo (6.5.20) da seguinte forma

$$\ln \frac{q_{x,t}}{p_{x,t}} = G_x(t) = c_{x,0} + c_{x,1}t,$$

donde se deriva o seguinte caso particular de (6.5.20)

$$q_x(t) = \frac{e^{c_{x,0} + c_{x,1}t}}{1 + e^{c_{x,0} + c_{x,1}t}}. \quad (6.5.21)$$

Este modelo de projecção é adoptado, por exemplo, em França.

## 6.6 Método Lee-Carter

Lee e Carter (1992) desenvolvem um novo método de extrapolação das tendências e perfis etários da mortalidade no longo prazo, assente na combinação de um modelo demográfico para a mortalidade com métodos normalmente aplicados às séries temporais. O método, conhecido na literatura actuarial e demográfica por método Lee-Carter (LC), foi aplicado originalmente à população dos EUA e converteu-se, desde então, no método de referência usado por organismos oficiais e investigadores de todo o mundo na elaboração de projecções de mortalidade para a população.<sup>33</sup> Em Portugal, as últimas projecções de população residente publicadas pelo INE (2003) para o horizonte de 2050 utilizam o método LC na previsão da componente mortalidade.<sup>34</sup>

### 6.6.1 Definição do modelo

O método LC assenta na decomposição da mortalidade em duas parcelas: um modelo demográfico, específico para cada idade, e um indicador de tendência temporal, sobre o qual são aplicados métodos de séries temporais no sentido de projectar valores futuros para os indicadores de mortalidade.

Formalmente, o método define-se pela seguinte representação log-bilinear para a taxa de mortalidade  $m_{x,t}$

$$\ln m_{x,t} = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t}, \quad (6.6.1)$$

onde:

<sup>33</sup>Veja-se, por exemplo, Wilmoth (1993, 1995a), Bell (1997), Figoli (1998), Lee (2000) e Tuljapurkar *et al.* (2000).

<sup>34</sup>Veja-se igualmente Coelho (2001) para uma aplicação do método LC na modelação e previsão da mortalidade da população portuguesa no período 1942-1999, com sexos agregados e grupos etários quinquenais.

## 6.6. Método Lee-Carter

$\alpha_x$  : define, para cada idade, o nível médio das taxas de mortalidade no tempo

$\kappa_t$  : descreve as tendências temporais do nível de mortalidade

$\beta_x$  : simboliza o padrão de oscilação etária nas taxas de mortalidade, ou seja, dá-nos uma ideia da sensibilidade de  $m_{x,t}$  às variações em  $\kappa_t$ , i.e.,  $\frac{d \ln m_{x,t}}{dt} = \beta_x \frac{d\kappa_t}{dt}$

$\epsilon_{x,t}$  : termo aleatório do tipo ruído branco  $\epsilon_{x,t} \sim \mathcal{N}or(0, \sigma_\epsilon^2)$ , incluído de modo a refletir influências históricas particulares em cada idade não capturadas pelo modelo

Como se observa, o modelo é sobreparametrizado no sentido de que a sua estrutura base é invariante às seguintes transformações

$$\begin{aligned} \{\alpha_x, \beta_x, \kappa_t\} &\mapsto \{\alpha_x, \beta_x/c, c\kappa_t\} \quad \forall c \in \mathbb{R}, c \neq 0 \\ \{\alpha_x, \beta_x, \kappa_t\} &\mapsto \{\alpha_x - c\beta_x, \beta_x, \kappa_t + c\} \quad \forall c \in \mathbb{R} \end{aligned}$$

para uma dada constante  $c$ .

O sistema definido por (6.6.1) é indeterminado se não forem impostas condições adicionais. Neste sentido, os autores adicionam duas restrições ao modelo de modo a garantir a sua identificação e a obtenção de uma solução única para as estimativas dos parâmetros

$$\sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1 \quad (6.6.2)$$

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0. \quad (6.6.3)$$

Em resultado destas restrições, o parâmetro  $\alpha_x$  é calculado simplesmente pela média de  $\ln m_{x,t}$  no período de observação.<sup>35</sup>

<sup>35</sup>Para compreender esta conclusão, note que:

$$\begin{aligned} \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \ln(m_{x,t}) &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} (\alpha_x + \beta_x \kappa_t) = \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \alpha_x + \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \beta_x \kappa_t \\ &= \sum_{x=x_{\min}}^{x_{\max}} \alpha_x \sum_{t=t_{\min}}^{t_{\max}} 1 + \left( \sum_{t=t_{\min}}^{t_{\max}} \kappa_t \right) \left( \sum_{x=x_{\min}}^{x_{\max}} \beta_x \right). \end{aligned}$$

Se admitirmos que  $\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0$  e  $\sum_{x=x_{\min}}^{x_{\max}} \beta_x \neq 0$  obtemos, após alguma simplificação,

$$\sum_{x=x_{\min}}^{x_{\max}} \alpha_x = \frac{1}{T} \left[ \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \ln(m_{x,t}) \right],$$

### 6.6.2 Estimação dos parâmetros por mínimos quadrados: análise de componentes principais

O modelo definido em (6.6.1) não pode ser estimado pelos tradicionais métodos de regressão linear uma vez que nenhum dos regressores é observável. A estimação dos parâmetros é efectuada com recurso ao método dos mínimos quadrados ordinários, ou seja, resolvendo o seguinte problema de optimização

$$(\hat{\alpha}_x, \hat{\beta}_x, \hat{\kappa}_t) = \arg \min_{\alpha_x, \beta_x, \kappa_t} \left\{ \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} (\ln m_{x,t} - \alpha_x - \beta_x \kappa_t)^2 \right\} \quad (6.6.4)$$

Na versão original, Lee e Carter (1992) propõem a utilização de uma decomposição em valores singulares (*Singular Value Decomposition*, SVD) para obter uma solução para (6.6.4). Segundo este procedimento, a estimação dos parâmetros do modelo subdivide-se em três etapas distintas. Em resultado das restrições (6.6.2) e (6.6.3), a solução obtida é, no entanto, única.

#### Etapa 1: Estimação de $\alpha_x$

Os parâmetros  $\alpha_x$  são estimados a partir das condições de primeira ordem para a minimização da função objectivo (6.6.4), i.e.,

$$\sum_{t=t_{\min}}^{t_{\max}} \ln m_{x,t} = (t_{\max} - t_{\min} + 1) \alpha_x,$$

donde se deriva, considerando  $T = (t_{\max} - t_{\min} + 1)$ ,

$$\hat{\alpha}_x = \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \ln (m_{x,t}).$$

onde  $T = t_{\max} - t_{\min} + 1$ . Prova-se facilmente que esta restrição é cumprida quando:

$$\alpha_x = \frac{1}{T} \left[ \sum_{t=t_{\min}}^{t_{\max}} \ln (m_{x,t}) \right].$$

Lee e Carter (1992) investigam ainda a normalização alternativa pressuposta pela restrição  $\sum_{x=x_{\min}}^{x_{\max}} \beta_x = 0$ , mas rejeitam-na por razões empíricas. Segundo os autores, neste caso os coeficientes  $\{\beta_x\}$  poderiam assumir valores negativos, indiciando que a mortalidade poderia subir para umas idades e descer noutras, um comportamento que embora possível não é confirmado pelos estudos empíricos realizados em séries longas. Wilmoth (1993) argumenta ainda que a restrição  $\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0$  facilita a comparação entre populações.

**Etapa 2: Estimação de  $\beta_x$  e  $\kappa_t$**

Seja  $\mathcal{N} = (x_{\max} - x_{\min} + 1)$ . Designemos por  $\mathbf{Z}$  a matriz de dimensão  $(\mathcal{N} \times \mathcal{T})$ , com elemento representativo  $z_{x,t}$  definido pelo valor centrado de  $\ln \hat{m}_{x,t}$  com relação à sua média temporal  $\hat{\alpha}_x$ , i.e.,

$$z_{x,t} = \ln \hat{m}_{x,t} - \hat{\alpha}_x.$$

O objectivo é então o de aproximar  $\mathbf{Z}$  pelo produto de duas matrizes

$$\mathbf{Z} \simeq \hat{\boldsymbol{\beta}} \hat{\boldsymbol{\kappa}}', \quad (6.6.5)$$

onde  $\hat{\boldsymbol{\beta}} = (\hat{\beta}_{x_{\min}}, \dots, \hat{\beta}_{x_{\max}})'$  e  $\hat{\boldsymbol{\kappa}} = (\hat{\kappa}_{t_{\min}}, \dots, \hat{\kappa}_{t_{\max}})'$  são determinados segundo o princípio dos mínimos quadrados, ou seja, pela minimização de

$$\sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} (z_{x,t} - \beta_x \kappa_t)^2,$$

usando uma decomposição em valores singulares(SVD).

Designemos por  $\mathbf{u}_i$  o  $i^{\text{ésimo}}$  vector próprio normalizado da matriz  $\mathbf{Z}'\mathbf{Z}$  de dimensões  $(\mathcal{T} \times \mathcal{T})$  correspondente ao valor próprio  $\lambda_i$ . Seja  $\mathbf{v}_i$  o  $i^{\text{ésimo}}$  vector próprio da matriz  $\mathbf{Z}\mathbf{Z}'$  de dimensões  $(\mathcal{N} \times \mathcal{N})$  correspondente ao valor próprio  $\lambda_i$ . Prova-se que a SVD da matriz  $\mathbf{Z}$  é dada por

$$\mathbf{Z} = \sum_{i \geq 1} \sqrt{\lambda_i} \mathbf{v}_i \mathbf{u}_i'. \quad (6.6.6)$$

Nos casos em que o primeiro valor próprio  $\lambda_1$  domina os restantes, obtém-se a seguinte aproximação

$$\mathbf{Z} \simeq \sqrt{\lambda_1} \mathbf{v}_1 \mathbf{u}_1'. \quad (6.6.7)$$

A qualidade da aproximação (6.6.7) pode ser aferida pela chamada “taxa de inércia”  $\tau_1$  (ou percentagem de variância explicada), calculada mediante

$$\tau_1 = \frac{\lambda_1}{\sum_{i \geq 1} \lambda_i}. \quad (6.6.8)$$

Combinando as equações (6.6.5) e (6.6.7), obtemos as seguintes estimativas de  $\hat{\boldsymbol{\beta}}$  e  $\hat{\boldsymbol{\kappa}}$

$$\hat{\boldsymbol{\beta}} = \frac{\mathbf{v}_1}{\sum_j v_{1j}} \quad (6.6.9)$$

$$\hat{\boldsymbol{\kappa}} = \sqrt{\lambda_1} \left( \sum_j v_{1j} \right) \mathbf{u}_1, \quad (6.6.10)$$

admitindo que  $\sum_j v_{1j} \neq 0$ . As estimativas  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  respeitam as restrições (6.6.2) e (6.6.3) do modelo. A decomposição em valores singulares só é aplicável a matrizes de dados rectangulares completas. Significa isto que não pode ser usada nos casos em que a idade máxima disponível não é constante no período de observação.

### Etapa 3: Reajustamento dos $\hat{\kappa}_t$

Para finalizar, é necessário proceder a um reajustamento dos  $\hat{\kappa}_t$  de modo a garantir que o total de óbitos observados em cada ano  $\sum_x d_{x,t}$  iguala o seu número estimado.<sup>36</sup> Em termos formais,

$$\sum_{x=x_{\min}}^{x_{\max}} d_{x,t} = \sum_{x=x_{\min}}^{x_{\max}} E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t) \quad \forall t. \quad (6.6.11)$$

Os novos estimadores  $\check{\kappa}_t$  obtêm-se resolvendo a equação (6.6.11), que pode ser reformulada segundo

$$\sum_{x=x_{\min}}^{x_{\max}} \left[ E_{x,t} \exp(\hat{\alpha}_x) \exp(\hat{\beta}_x \check{\kappa}_t) - d_{x,t} \right] = 0.$$

O método iterativo de Newton-Raphson constitui uma escolha adequada para resolver este problema. Recorde-se, a este propósito, que o método se baseia no seguinte esquema iterativo

$$\kappa_t^{(\nu+1)} = \kappa_t^{(\nu)} - \frac{F(\kappa_t^{(\nu)})}{\left. \frac{\partial F}{\partial \kappa_t} \right|_{\kappa_t = \kappa_t^{(\nu)}}}, \quad (6.6.12)$$

onde  $F(\kappa_t)$  representa a função a anular. Na iniciação do procedimento iterativo, consideramos os valores brutos de  $\kappa_t$  estimados na etapa anterior, i.e.,  $\kappa_t^{(0)} = \hat{\kappa}_t$ . A restrição (6.6.3) é violada pelas novas estimativas  $\check{\kappa}_t$  resultantes do algoritmo iterativo. Em consequência, torna-se necessário impor a seguinte transformação

$$\begin{aligned} \hat{\kappa}_t &\leftarrow \check{\kappa}_t - \bar{\kappa}_t \\ \hat{\alpha}_x &\leftarrow \hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t, \end{aligned}$$

onde  $\bar{\kappa}_t$  denota a média dos  $\check{\kappa}_t$  estimados. A qualidade do ajustamento é aferida por indicadores como a percentagem de variância explicada pelo primeiro componente principal da SVD e pela análise dos resíduos do modelo.

<sup>36</sup>Para além de melhorar a qualidade do ajustamento, esta correcção apresenta uma segunda vantagem. Com efeito, nos casos em que não dispomos de dados demográficos discriminados por idade em determinados anos, mas apenas do número total de óbitos  $\sum_x d_{x,t}$ , o método permite a dedução de um valor para o índice temporal  $\kappa_t$  recorrendo aos valores de  $\alpha_x$  e  $\beta_x$  estimados com base nos restantes anos de calendário.

### 6.6.3 Estimação dos parâmetros por mínimos quadrados: método iterativo

Uma solução alternativa para estimar os parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ , que não está condicionada pela disponibilidade de matrizes de dados rectangulares completas, passa pela utilização de um algoritmo iterativo do tipo Newton-Raphson (6.6.12). Este algoritmo é definido, no presente problema, pelas seguintes equações

$$\begin{aligned}
 \hat{\alpha}_x^{(v+1)} &= \hat{\alpha}_x^{(v)} + \frac{\sum_{t=t_{\min}}^{t_{\max}} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(v)} - \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)} \right]}{\mathcal{T}} \\
 \hat{\kappa}_t^{(v+1)} &= \hat{\kappa}_t^{(v)} + \frac{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(v)} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(v+1)} - \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)} \right]}{\sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_x^{(v)})^2} \\
 \hat{\beta}_x^{(v+1)} &= \hat{\beta}_x^{(v)} + \frac{\sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(v+1)} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(v+1)} - \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v+1)} \right]}{\sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_t^{(v+1)})^2}.
 \end{aligned} \tag{6.6.13}$$

A implementação prática do algoritmo requer uma iniciação e a escolha de um critério que defina a solução óptima. Em relação à iniciação, a atribuição de um valor nulo a  $\hat{\beta}_x^{(0)}$  implica a divisão por zero na estimação de  $\hat{\kappa}_t^{(1)}$ . Se exceptuarmos esta limitação, a escolha dos parâmetros iniciais é livre desde que garanta a convergência do algoritmo.

No que diz respeito ao critério de óptimo, são múltiplas as escolhas possíveis. Uma solução simples envolve, por exemplo, a fixação de um valor mínimo para a diminuição relativa na soma dos quadrados do erro. Formalmente, o algoritmo termina na iteração  $\nu$  quando se observa que

$$\frac{e^{(\nu-1)} - e^{(\nu)}}{e^{(\nu-1)}} < \varepsilon,$$

com

$$e^{(\nu)} = \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \left( \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)} \right)^2,$$

onde  $\varepsilon$  denota o patamar inferior para a diminuição relativa do erro  $e^{(\nu)}$ .

As equações do algoritmo (6.6.13) podem ser escritas em termos matriciais segundo

$$\begin{aligned}
 (\mathbf{G}_\alpha^{(\nu)})\hat{\alpha}^{(\nu+1)} &= (\mathbf{G}_\alpha^{(\nu)})\hat{\alpha}^{(\nu)} + \mathbf{L}_\alpha^{(\nu)} \\
 (\mathbf{G}_\kappa^{(\nu)})\hat{\kappa}^{(\nu+1)} &= (\mathbf{G}_\kappa^{(\nu)})\hat{\kappa}^{(\nu)} + \mathbf{L}_\kappa^{(\nu)} \\
 (\mathbf{G}_\beta^{(\nu)})\hat{\beta}^{(\nu+1)} &= (\mathbf{G}_\beta^{(\nu)})\hat{\beta}^{(\nu)} + \mathbf{L}_\beta^{(\nu)}
 \end{aligned} \tag{6.6.14}$$

onde  $\mathbf{G}_\alpha^{(\nu)}$ ,  $\mathbf{G}_\kappa^{(\nu)}$  e  $\mathbf{G}_\beta^{(\nu)}$  são matrizes diagonais de dimensão  $(\mathcal{N} \times \mathcal{N})$ ,  $(\mathcal{T} \times \mathcal{T})$  e  $(\mathcal{N} \times \mathcal{N})$ , respectivamente, e elemento representativo definido por

$$\begin{aligned}
 (\mathbf{G}_\alpha^{(\nu)})_{xx} &= t_{\max} - t_{\min} + 1 \\
 (\mathbf{G}_\kappa^{(\nu)})_{tt} &= \sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_x^{(\nu)})^2 \\
 (\mathbf{G}_\beta^{(\nu)})_{xx} &= \sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_t^{(\nu+1)})^2
 \end{aligned} \tag{6.6.15}$$

e  $\mathbf{L}_\alpha^{(\nu)}$ ,  $\mathbf{L}_\kappa^{(\nu)}$  e  $\mathbf{L}_\beta^{(\nu)}$  são vectores coluna de dimensão  $\mathcal{N}$ ,  $\mathcal{T}$  e  $\mathcal{N}$ , respectivamente, e elemento representativo definido por

$$\begin{aligned}
 (\mathbf{L}_\alpha^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)} \right] \\
 (\mathbf{L}_\kappa^{(\nu)})_t &= \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(\nu)} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu+1)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)} \right] \\
 (\mathbf{L}_\beta^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(\nu+1)} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu+1)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu+1)} \right].
 \end{aligned} \tag{6.6.16}$$

As estimativas produzidas pelo algoritmo (6.6.14) não cumprem naturalmente as restrições (6.6.2) e (6.6.3), pelo que é necessário introduzir a seguinte transformação

$$\begin{aligned}
 \hat{\alpha}_x &\leftarrow \hat{\alpha}_x + \hat{\beta}_x \bar{\kappa}_t \quad \text{com} \quad \bar{\kappa}_t = \frac{1}{\mathcal{T}} \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t \\
 \hat{\kappa}_t &\leftarrow (\hat{\kappa}_t - \bar{\kappa}_t) \hat{\beta}_\xi \quad \text{com} \quad \hat{\beta}_\xi = \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x \\
 \hat{\beta}_x &\leftarrow \frac{\hat{\beta}_x}{\hat{\beta}_\xi}
 \end{aligned}$$

O esquema iterativo agora descrito substitui as etapas 1 e 2 do procedimento SVD apresentado na secção anterior, mas não dispensa a terceira etapa que envolve, como vimos, a reestimação do índice temporal  $\kappa_t$  de maneira a assegurar que o método LC

reproduz de forma adequada o número de óbitos registados em cada ano.

#### 6.6.4 Estimação dos parâmetros por mínimos quadrados ponderados

O método dos mínimos quadrados ordinários descrito nas secções precedentes minimiza os desvios quadráticos sem destringar a contribuição de cada observação  $(x, t)$  para o número total de óbitos. Tendo em mente este aspecto, Wilmoth (1993) propõe a estimação dos parâmetros do método LC recorrendo ao método dos mínimos quadrados ponderados. Em termos formais, o método envolve a resolução do seguinte problema

$$\left( \hat{\alpha}_x, \hat{\beta}_x, \hat{\kappa}_t \right) = \arg \min_{\alpha_x, \beta_x, \kappa_t} \left\{ \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} w_{x,t} (\ln m_{x,t} - \alpha_x - \beta_x \kappa_t)^2 \right\} \quad (6.6.17)$$

com  $w_{x,t} = d_{x,t}$ .

A estimação dos parâmetros pode ser feita, à semelhança do que sucede na secção anterior, com o auxílio do esquema iterativo (6.6.14), bastando para tal a redefinição das matrizes em (6.6.15) e (6.6.16). Em termos mais precisos, as matrizes  $\mathbf{G}_\alpha^{(\nu)}$ ,  $\mathbf{G}_\kappa^{(\nu)}$  e  $\mathbf{G}_\beta^{(\nu)}$  são agora definidas por

$$\begin{aligned} (\mathbf{G}_\alpha^{(\nu)})_{xx} &= \sum_{t=t_{\min}}^{t_{\max}} d_{x,t} \\ (\mathbf{G}_\kappa^{(\nu)})_{tt} &= \sum_{x=x_{\min}}^{x_{\max}} d_{x,t} (\hat{\beta}_x^{(\nu)})^2 \\ (\mathbf{G}_\beta^{(\nu)})_{xx} &= \sum_{t=t_{\min}}^{t_{\max}} d_{x,t} (\hat{\kappa}_t^{(\nu+1)})^2. \end{aligned} \quad (6.6.18)$$

De igual modo, os vectores  $\mathbf{L}_\alpha^{(\nu)}$ ,  $\mathbf{L}_\kappa^{(\nu)}$  e  $\mathbf{L}_\beta^{(\nu)}$  são agora definidos por

$$\begin{aligned} (\mathbf{L}_\alpha^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} d_{x,t} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)} \right] \\ (\mathbf{L}_\kappa^{(\nu)})_t &= \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(\nu)} d_{x,t} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu+1)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)} \right] \\ (\mathbf{L}_\beta^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(\nu+1)} d_{x,t} \left[ \ln \hat{m}_{x,t} - \hat{\alpha}_x^{(\nu+1)} - \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu+1)} \right]. \end{aligned} \quad (6.6.19)$$

Nos demais aspectos (iniciação do algoritmo e critério de óptimo), o ajustamento do modelo segue os passos descritos na Secção 6.6.3. Wilmoth (1993) sugere igualmente um procedimento alternativo envolvendo a especificação de um modelo probabilístico, baseado na distribuição de Poisson, e a estimação dos parâmetros através do método

de máxima verosimilhança. Este procedimento está na origem do modelo de Poisson desenvolvido por Brouhns *et al.* (2002a), que analisaremos mais adiante na Secção 6.9.

### 6.6.5 Projecções de mortalidade

A obtenção de projecções de mortalidade no método LC decorre da aplicação da metodologia de Box-Jenkins à série temporal dos valores do parâmetro  $\kappa_t$ . Por exemplo, Lee e Carter (1992) preconizam a utilização de um modelo do tipo ARIMA(0, 1, 0), ou seja, de um modelo de passeio aleatório com tendência constante<sup>37</sup>

$$\kappa_t = \lambda + \kappa_{t-1} + u_t, \quad u_t \sim \mathcal{N}or(0, \sigma_u^2),$$

onde  $\lambda$  representa a tendência (*drift*), cujo estimador de máxima verosimilhança é dado simplesmente por  $\hat{\lambda} = (\hat{\kappa}_{t_{\max}} - \hat{\kappa}_{t_{\min}}) / (T - 1)$ .<sup>38</sup> Com base no modelo ARIMA seleccionado, é gerada uma previsão  $\{\hat{\kappa}_{t_{\max}+s} : s > 0\}$  para o índice temporal dada por  $\hat{\kappa}_{t_{\max}+s} = \mathbb{E}[\lambda s + \hat{\kappa}_{t_{\max}} + \sum_{t=1}^s u_t] = \hat{\kappa}_{t_{\max}} + \lambda s$ . Se recuperarmos a equação (6.6.1) e substituírmos  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_{t_{\max}+s}$ , logramos finalmente taxas de mortalidade projectadas  $\{\hat{m}_{x,t_{\max}+s} : s > 0\}$ .<sup>39</sup>

$$\hat{m}_{x,t_{\max}+s} = \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_{t_{\max}+s}), \quad s > 0. \quad (6.6.20)$$

<sup>37</sup>Nas projecções de população residente para Portugal e para o horizonte 2000-2050 o INE (2003) optou por uma solução alternativa ao considerar a modelação do índice temporal  $\kappa_t$  através de uma função logística:

$$\kappa_t = a_0 + a_1 \left( \frac{\exp(a_2 + a_3 t)}{1 + \exp(a_2 + a_3 t)} \right) + \epsilon_t,$$

onde  $t$  é a única variável explicativa e onde  $a_0$ ,  $a_1$ ,  $a_2$  e  $a_3$  são parâmetros da função. O termo  $\epsilon_t$  representa o erro associado ao ajustamento do modelo. A opção pela função logística é justificada, segundo o INE, pela sua maior flexibilidade em relação aos modelos ARIMA.

<sup>38</sup>A estimativa de máxima verosimilhança da variância  $\sigma_u^2$  é dada por:

$$\sigma_u^2 = \frac{1}{T-1} \sum_{t=1}^{h-1} (\hat{\kappa}_{t+1} - \hat{\kappa}_t - \hat{\lambda})^2,$$

com  $\text{Var}[\hat{\lambda}] = \sigma_u^2 / (T - 1)$ .

<sup>39</sup>Uma definição equivalente de (6.6.20) é dada por:

$$\begin{aligned} \hat{m}_{x,t_{\max}+s} &= \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_{t_{\max}+s}), \quad s > 0 \\ &= \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_{t_{\max}}) \exp \left[ \hat{\beta}_x (\hat{\kappa}_{t_{\max}+s} - \hat{\kappa}_{t_{\max}}) \right] \\ &= \left[ \exp \left( \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \ln m_{x,t} \right) \exp(\hat{\beta}_x \hat{\kappa}_{t_{\max}}) \right] \exp \left[ \hat{\beta}_x (\hat{\kappa}_{t_{\max}+s} - \hat{\kappa}_{t_{\max}}) \right]. \end{aligned}$$

Em alternativa a (6.6.20), Bell (1997) recomenda a substituição do parâmetro  $\alpha_x$  (estimado, como vimos, pela média de  $\ln m_{x,t}$ ) pelas taxas de mortalidade mais recentes, eventualmente graduadas em relação aos anos adjacentes de maneira a garantir uma transição suave. Esta opção visa assegurar que os valores estimados para o primeiro ano de previsão estão alinhados com as observações mais recentes. Em consequência desta alteração, as tábuas projectadas são calculadas mediante

$$\hat{m}_{x,t_{\max}+s} = m_{x,t_{\max}} \exp \left[ \hat{\beta}_x (\hat{\kappa}_{t_{\max}+s} - \hat{\kappa}_{t_{\max}}) \right]. \quad (6.6.21)$$

Note-se que o termo

$$RF(x, t_{\max} + s) = \exp \left[ \hat{\beta}_x (\hat{\kappa}_{t_{\max}+s} - \hat{\kappa}_{t_{\max}}) \right]$$

na equação (6.6.21) configura, na prática, um factor de redução da mortalidade.<sup>40</sup>

### 6.6.6 Apreciação do método e desenvolvimentos recentes

Uma das principais virtudes do método LC é, segundo os seus proponentes, o facto de este combinar um método demográfico que é simultaneamente rico mas parcimonioso, com um modelo de séries temporais, o que permite a obtenção de intervalos de confiança para as previsões (sob a hipótese de que os parâmetros são independentes). Por outro lado, a forma como o modelo demográfico subjacente está definido é suficiente para garantir que as taxas de mortalidade exibem um padrão de decréscimo exponencial, sem necessidade de impor qualquer limite assintótico arbitrário ou restrição que demarque os ganhos na esperança de vida. Outra vantagem do método refere-se à possibilidade de obter, de forma indirecta, estimativas nos períodos para os quais não existe informação disponível.<sup>41</sup>

<sup>40</sup> Por exemplo, se adoptarmos um modelo do tipo ARIMA(0, 1, 0) com tendência ditada pelo parâmetro  $\lambda$ , obtemos o seguinte factor de redução

$$RF(x, t_{\max} + s) = \exp(\hat{\beta}_x \hat{\lambda} s),$$

semelhante ao adoptado originalmente pelo CMIB.

<sup>41</sup> Com efeito, dadas duas tábuas de mortalidade para dois períodos distintos é possível gerar uma família de tábuas recorrendo ao seguinte procedimento rudimentar:

1. Atribui-se o valor  $\kappa_0 = 0$  a uma das tábuas e o valor  $\kappa_1 = 1$  à outra;
2. São usadas as seguintes estimativas para  $\alpha_x$  e  $\beta_x$

$$\hat{\alpha}_x = \ln m_{x,0} \quad \hat{\beta}_x = \ln m_{x,0} - \ln m_{x,1}$$

3. As previsões são obtidas atribuindo valores arbitrários a  $\kappa_t$ . Nos casos em que  $\kappa_t \in [0, 1]$ , o modelo interpola geometricamente entre as duas tábuas. Nos casos em que  $\kappa_t < 0$  ou  $\kappa_t > 1$ , o

O método LC apresenta ainda algumas vantagens face às técnicas de extrapolação dos perfis anteriormente analisadas. Em primeiro lugar, a modelação de um único índice temporal permite resumir a estrutura de covariâncias entre as  $m_x$  (ou  $q_x$ ) no tempo, uma tarefa bastante mais complexa nos casos em que as projecções por idade são feitas de forma independente. Em segundo, a modelação de cada série individual  $m_x$  por modelos de séries temporais aumentaria significativamente o número de parâmetros e a complexidade do problema. Por último, a modelação independente das  $m_x$  pode, num futuro longínquo, gerar perfis de mortalidade muito pouco plausíveis. Pelo contrário, no método LC a dependência dos perfis face ao parâmetro  $\kappa_t$  assegura um compromisso entre as tendências das taxas correspondentes a cada idade.

A um outro nível, o método LC permite contornar algumas das principais limitações dos métodos paramétricos de projecção. Em particular, o método LC reduz a complexidade do problema,<sup>42</sup> oferece uma solução simples para a obtenção de intervalos de confiança e contorna o problema da irregularidade dos parâmetros.<sup>43</sup>

O enorme entusiasmo que o método LC suscitou após a sua publicação e a sua aplicação em realidades distintas daquelas para a qual foi originalmente concebido, nomeadamente em países com perfis de diminuição da mortalidade distintos do registado nos EUA permitiram, por um lado, realçar as qualidades da abordagem e, por outro, pôr em destaque as suas limitações. Lee (2000), Lee e Miller (2001), Tuljapurkar e Boe (1998), Brouhns *et al.* (2002a), Wong-Fupuy e Haberman (2004), entre outros, procedem a uma avaliação detalhada das potencialidades e limitações do método LC, da sua performance em termos de previsão e dos melhoramentos introduzidos no modelo.

O método LC pertence, desde logo, à categoria dos métodos extrapolativos padecendo, por isso, da sua debilidade fundamental: assume que as tendências históricas se mantêm no futuro, não incorpora elementos de avaliação subjectiva e ignora as implicações de eventuais alterações estruturais (avanços na medicina ou no combate às epidemias, alterações nos hábitos de vida, etc.).

Por outro lado, o método admite um determinado padrão de distribuição etária da mortalidade, em que as taxas de decréscimo nas várias idades (ditadas por  $\beta_x(d\kappa_t/dt)$ )

---

modelo produz extrapolações a partir de uma das tábuas.

<sup>42</sup>Se é verdade que nos métodos paramétricos o número de parâmetros necessário para ajustar uma dada experiência é relativamente reduzido, não é menos verdade que o seu número é bem maior quando o objectivo é o de realizar extrapolações de tendências e derivar intervalos de confiança.

<sup>43</sup>Com efeito, a evidência empírica (McNown e Rogers, 1989; Benjamin e Soliman, 1993; Felipe *et al.*, 2001) mostra que as estimativas dos parâmetros apresentam um padrão bastante irregular, o que dificulta a projecção. Não é assim de estranhar que, por exemplo, McNown e Rogers (1989) tenham optado por realizar projecções a 15 anos, em contraste com 75 anos estipulados por Lee e Carter (1992).

mantém sempre a mesma relação de proporcionalidade no tempo, uma hipótese não confirmada nalguns estudos empíricos.<sup>44</sup> Por outras palavras, as tendências temporais de  $\ln(m_{x,t})$  estão deterministicamente relacionadas, apesar de poderem diferir entre idades.

A incerteza presente no método LC decorre, fundamentalmente, da incerteza na previsão do índice temporal  $\kappa_t$ . Lee e Carter (1992) demonstram que à medida que o horizonte de projecção aumenta, as restantes fontes de variabilidade, designadamente os erros na estimação de  $\beta_x$  e os erros de ajustamento do modelo demográfico  $\epsilon_{x,t}$ , perdem importância e são dominados pela incerteza relativa a  $\kappa_t$ .<sup>45</sup> Por outro lado, têm surgido algumas dúvidas relativamente aos intervalos de previsão reportados por Lee e Carter (1992), considerados demasiadamente estreitos, que levantam dúvidas sobre a capacidade do modelo para incorporar toda a incerteza em relação ao futuro.

Noutro domínio, a forma como o modelo foi ajustado originalmente conduzia a que, nalguns grupos etários, se registasse uma discrepância entre o valor observado para a taxa de mortalidade no último ano considerado na amostra e o primeiro ano de previsão, um aspecto que podia de algum modo distorcer a qualidade da projecção.<sup>46</sup> Este problema foi analisado por Bell (1997), que concluiu que a substituição do parâmetro  $\alpha_x$  pelas taxas de mortalidade mais recentes melhorava a qualidade do método LC.

Outra das críticas apontadas ao método LC refere-se ao comportamento assintótico de longo prazo dos factores de redução. Com efeito, a combinação típica de valores positivos para o parâmetro  $\beta_x$  com uma trajectória decrescente para  $\kappa_t$  conduz-nos a valores nulos para o factor de redução, um resultado dificilmente aceitável na realidade e que, como vimos acima, foi há muito assumido na metodologia do CMIB. Para contornar esta limitação, Renshaw e Haberman (2005) apresentam uma versão do método LC em que consideram que a mortalidade assintótica é positiva.

Nos últimos anos, foram propostas e implementadas inúmeras extensões do método LC, algumas das quais merecerão atenção especial nas secções seguintes deste capítulo.

---

<sup>44</sup>Estudos empíricos realizados em vários países (e.g., Horiuchi e Wilmoth (1998)) demonstraram que a taxa de variação das taxas de mortalidade pode oscilar no tempo evidenciando, em particular, que nas últimas décadas o decréscimo é mais rápido nas idades elevadas do que nas idades mais jovens.

<sup>45</sup>Outro dos problemas do método diz respeito ao facto do comportamento da mortalidade nos EUA nem sempre ter diminuído, no período 1900-1989, segundo a trajectória definida pelo parâmetro  $\kappa_t$ . Ou seja, a série de  $\kappa_t$  não representa uma propriedade fundamental do comportamento da mortalidade, o que coloca algumas interrogações sobre a capacidade do modelo para realizar projecções.

<sup>46</sup>Lee e Carter (1992) reconheceram este erro, que poderia assumir algum significado nos primeiros anos de previsão, mas preferiram despezá-lo argumentando, por um lado, que ele se manifestava sobretudo nas taxas correspondentes às idades mais jovens e exercia, por isso, pouca influência sobre a esperança de vida e, por outro, que forçar o modelo a passar pelo valor observado no último ano do período amostral deteriorava a qualidade do ajustamento.

Carter (1996) propõe a utilização de modelos estruturais de séries temporais de modo a permitir que os parâmetros da equação de estimação possam variar no tempo. Lee e Miller (2001) testam uma alteração na forma como o parâmetro  $\kappa_t$  é reestimado na etapa 3 do procedimento SVD. A alteração consiste em reestimar  $\kappa_t$  de modo a que o valor projectado para a esperança de vida à nascença seja igual ao respectivo valor observado num determinado ano.<sup>47</sup> Os autores reconhecem ainda que a maior limitação do método corresponde à hipótese de invariância no tempo da sua componente etária, uma assunção desmentida pela realidade empírica.<sup>48</sup>

Carter e Prskawetz (2001) incorporam no método LC a possibilidade de este contemplar alterações temporais nos parâmetros  $\alpha_x$  e  $\beta_x$ , ou seja, introduzem a possibilidade dos perfis etários da mortalidade se alterarem no tempo. Wilmoth e Valkonen (2002) desenvolvem uma extensão do método LC destinada a analisar os diferenciais de mortalidade entre idades e no tempo, onde incorporam um conjunto de variáveis explicativas.

Renshaw e Haberman (2003a) reformulam o método LC de modo torna-lo comparável com a abordagem pelos factores de redução descrita na Secção 6.8. Renshaw e Haberman (2003b) reinterpretem a estrutura subjacente ao método LC no contexto dos GLMs.<sup>49</sup>

Brouhns *et al.* (2002a) e Renshaw e Haberman (2003c) criticam a hipótese de homocedasticidade assumida pelo método LC e desenvolvem uma versão baseada no modelo probabilístico definido pela distribuição de Poisson.<sup>50</sup> Bell (1997), Booth *et al.* (2002) e Renshaw e Haberman (2003c, 2003d) sugerem a construção de projecções com base nos dois primeiros termos da decomposição em valores singulares, bem como a aplicação de uma estrutura multivariada na análise do comportamento das séries temporais relativas ao índice temporal  $\kappa_t$ .<sup>51</sup>

Renshaw e Haberman (2003d) desenvolvem uma extensão do método LC com vista à derivação de projecções em populações desagregadas (e.g., por sexo, por tipo de produto, duração, por montantes investidos, etc.). Outras extensões do método LC, sumariadas por Lee (2000), incluem a possibilidade de obter previsões de mortalidade desagregadas por causa de morte. Renshaw e Haberman (2005) analisam o problema da projecção dos factores de redução num contexto em que se conjugam o efeito idade, o efeito temporal e o efeito geracional. Girosi e King (2005) demonstram que o método LC é um caso particular dos convencionais modelos de passeio aleatório (com tendência) generalizados

---

<sup>47</sup> A principal vantagem desta solução é a de que ela dispensa o conhecimento sobre a distribuição etária da população em cada um dos anos do período em estudo.

<sup>48</sup> Para contornar este problema, Lee e Miller (2001) e Tuljapurkar *et al.* (2000) sugerem a utilização de dados apenas a partir de 1950, atenuando desta forma os efeitos das quebras estruturais.

<sup>49</sup> Esta versão do modelo é analisada em detalhe na Secção 6.7.

<sup>50</sup> Esta versão do modelo e os seus desenvolvimentos são analisados em detalhe na Secção 6.9.

<sup>51</sup> Para mais detalhes veja-se a Secção 6.8.4.

(RWD), em que a matriz de covariâncias depende do vector de parâmetros  $\beta_x$ , ao passo que em RWD esta não está restringida.<sup>52</sup>

De Jong e Tickle (2006) formulam o método em termos de um modelo estado-espço (*state space framework*). Hyndman e Ullah (2007) desenvolvem uma extensão do modelo em que assumem: (i) que a mortalidade é uma função suave da idade; (ii) que as taxas de mortalidade são estimadas através de métodos de gradação não-paramétricos; (iii) que a modelação do índice temporal é feita com recurso a métodos generalizados de tratamento de séries temporais; (iv) que a estimação é robusta mesmo na presença de observações anormais (e.g., associadas a epidemias ou guerras). Biffis e Denuit (2005) e Biffis *et al.* (2006) generalizam o modelo a um contexto estocástico, admitindo em particular que a dinâmica do índice  $\kappa_t$  que determina as tendências temporais da mortalidade pode ser descrita por equações diferenciais estocásticas.<sup>53</sup>

## 6.7 Modelos Lineares Generalizados com respeito à idade e ao tempo

Nas secções precedentes tivemos oportunidade de constatar que a abordagem tradicional à modelação da mortalidade num contexto dinâmico envolve duas etapas distintas. Uma primeira, onde são graduadas tábuas contemporâneas para um ou mais períodos passados. Uma segunda, onde os perfis de mortalidade (ou os parâmetros de uma função analítica) constituem um input de um qualquer método de extrapolação. Dito de outro modo, nestes métodos a construção de tábuas projectadas é realizada de forma separada com respeito à idade e ao tempo cronológico.

Em alternativa a esta abordagem, Renshaw *et al.* (1996) desenvolvem uma estrutura metodológica, fundada nos modelos lineares generalizados analisados no Capítulo 3, que incorpora simultaneamente a idade actuarial e o tempo cronológico na estimação das tendências futuras para a mortalidade. Em termos formais, o modelo estipula que a força de mortalidade correspondente à idade  $x$  e ano  $t$ ,  $\mu_{x,t}$ , é expressa em termos de uma função paramétrica do tipo

$$\mu_{x,t} = \exp \left[ \beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r \alpha_i t^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x') t^i \right], \quad (6.7.1)$$

---

<sup>52</sup>Os autores demonstram ainda que a incerteza no método LC provém de duas fontes: choques de mortalidade perfeitamente correlacionados entre as idades e choques não correlacionados e que, por isso, ao contrário do que acontece no modelo RWD, o método tem muita dificuldade em lidar com os choques correlacionados de forma imperfeita.

<sup>53</sup>Para mais detalhes sobre a modelação estocástica veja-se o Capítulo 9.

## 6.7. Modelos Lineares Generalizados com respeito à idade e ao tempo

---

com a convenção de que alguns parâmetros  $\gamma_{ij}$  podem ser nulos.  $L_j(x')$  denota os polinómios de Legendre e  $x'$  e  $t'$  representam, respectivamente, as idades e anos cronológicos transformados, onde

$$x' = \frac{x - u}{v} \quad \text{com} \quad \begin{cases} u = \frac{x^{\max} + x^{\min}}{2} \\ v = \frac{x^{\max} - x^{\min}}{2} \end{cases}$$

$$t' = \frac{t - \omega}{\xi} \quad \text{com} \quad \begin{cases} \omega = \frac{t^{\max} + t^{\min}}{2} \\ \xi = \frac{t^{\max} - t^{\min}}{2} \end{cases}.$$

Se reagruparmos os termos em (6.7.1), podemos escrever

$$\mu_{x,t} = \exp \left[ \sum_{j=0}^s \beta_j L_j(x') \right] \exp \left[ \sum_{i=1}^r \left( \alpha_i + \sum_{j=1}^s \gamma_{ij} L_j(x') \right) t'^i \right]. \quad (6.7.2)$$

Na equação (6.7.2), a força de mortalidade  $\mu_{x,t}$  é expressa em termos do produto de duas parcelas distintas. A primeira, corresponde a uma função Gompertz-Makeham do tipo  $GM(0, s + 1)$ . A segunda, traduz o ajustamento decorrente da inclusão do efeito calendário na análise, acerto esse que será específico para cada idade  $x$  se pelo menos um dos termos  $\gamma_{ij}$  for diferente de zero.

Para estimar os parâmetros  $\alpha_i$ ,  $\beta_j$  e  $\gamma_{ij}$  do modelo são assumidas algumas hipóteses quanto à distribuição da variável aleatória. Neste caso, admite-se que o número observado de óbitos  $d_{x,t}$  é modelado em termos de realizações independentes de variáveis aleatórias Poisson  $D_{x,t}$ , de um modelo GLM com média e variância dadas por

$$\mathbb{E}[D_{x,t}] = m_{x,t} = E_{x,t} \mu_{x,t}, \quad \text{Var}(m_{x,t}) = \phi m_{x,t},$$

onde  $E_{x,t}$  representa o número de vidas expostas centralmente ao risco,  $\phi$  é um parâmetro de sobredispersão (de modo a contemplar a existência de duplicidades na amostra) e as demais mantêm o significado atribuído na Secção 3.5. Supõe-se ainda que o predictor linear está ligado à média da distribuição através de um *link* logarítmico, i.e.,

$$\log m_{x,t} = \eta_{x,t} = \log E_{x,t} + \log \mu_{x,t},$$

donde resulta

$$\log m_{x,t} = \log E_{x,t} + \beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r \alpha_i t'^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x')_i t'^i \quad (6.7.3)$$

## 6.7. Modelos Lineares Generalizados com respeito à idade e ao tempo

A estimativa dos parâmetros do modelo é conseguida através de um procedimento de log-verosimilhança que envolve, neste caso em particular, a resolução do seguinte problema de optimização

$$(\hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}) = \arg \max_{\alpha_i, \beta_j, \gamma_{ij}} \left\{ \frac{1}{\phi} \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} (-m_{x,t} + d_{x,t} \log m_{x,t}) \right\}. \quad (6.7.4)$$

A qualidade do ajustamento é aferida de modo usual, i.e., mediante o cálculo da estatística *unscaled deviance* relativa aos valores graduados  $\hat{\mu}_{x,t}$ , cuja expressão correspondente é, neste caso, definida por

$$\mathcal{D}(\mathbf{y}; \mathbf{m}) = 2 \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \left\{ d_{x,t} \log \left( \frac{d_{x,t}}{\hat{m}_{x,t}} \right) - (d_{x,t} - \hat{m}_{x,t}) \right\}, \quad (6.7.5)$$

onde  $\hat{m}_{x,t} = E_{x,t} \hat{\mu}_{x,t}$ . A correspondente *scaled deviance* é definida por  $\mathcal{S}(\mathbf{y}; \mathbf{m}) = \mathcal{D}(\mathbf{y}; \mathbf{m}) / \phi$ , com  $\phi$  estimado pelo quociente entre o valor de  $\mathcal{D}(\mathbf{y}; \mathbf{m})$  referente ao modelo óptimo e os respectivos graus de liberdade.

O procedimento adoptado pelos autores para seleccionar o modelo óptimo desenrola-se num conjunto de etapas. Em primeiro lugar, é determinada a ordem  $(r, s)$  óptima do modelo.<sup>54</sup> Com base neste resultado, é analisado o impacto da introdução dos coeficientes  $\gamma_{ij}$  (que determinam a dependência do ajustamento temporal em relação à idade) na estatística deviance.<sup>55</sup> A validação final do modelo é feita analisando os resíduos e os resultados dos testes de ajustamento descritos no Capítulo 3.

Tratando-se de um método eminentemente extrapolativo, é necessário averiguar se a boa capacidade de ajustamento no período amostral tem correspondência em termos de projecção de mortalidade. Sithole *et al.* (2000) investigam este aspecto e concluem que o modelo que proporciona o melhor ajustamento aos dados não gera, necessariamente, projecções de mortalidade plausíveis. Por esta razão, recomendam a determinação do modelo por etapas, encontrando um equilíbrio entre qualidade do ajustamento e alisamento das projecções.<sup>56</sup>

<sup>54</sup> A escolha óptima é feita testando se a melhoria na estatística  $\mathcal{S}(\mathbf{y}; \mathbf{m})$  em resultado de acréscimos sucessivos no número de parâmetros do modelo é estatisticamente significativa (quando comparada com o valor de  $\chi_{gl=1}^2$ ). Os valores óptimos de  $r$  e  $s$  serão aqueles a partir dos quais os incrementos na deviance são insignificantes do ponto de vista estatístico.

<sup>55</sup> Nas duas etapas iniciais é imprescindível verificar se os parâmetros estimados são significativamente diferentes de zero, usando para tal o habitual teste  $t$ -student.

<sup>56</sup> Em particular, os autores reconhecem que alguma qualidade de ajustamento pode ser sacrificada com o intuito de obter projecções adequadas, concluindo ainda que a dependência da mortalidade em relação ao tempo cronológico é expressa por polinómios de baixa ordem. Seleccionado o modelo óptimo, os autores efectuem projecções com um horizonte temporal de 20 anos. Por fim, revêem o modelo com

## 6.8. Modelação dos factores de redução da mortalidade

---

Alcançado o compromisso entre a qualidade do ajustamento e a capacidade preditiva do modelo, é possível definir factores de redução na mortalidade para cada idade e ano recorrendo a

$$RF(x, t) = \frac{\hat{\mu}_{x,t}}{\mu_{x,0}}, \quad (6.7.6)$$

onde  $\mu_{x,0}$  representa a força de mortalidade para um indivíduo com idade  $x$  no ano base ( $t = 0$ ), extraída, por exemplo, de uma tábua contemporânea, e  $\hat{\mu}_{x,t}$  representa a força de mortalidade para uma vida que atinge a idade  $x$  no ano (futuro)  $t$ . Estes factores de redução são comparáveis com os produzidos pela metodologia do CMIB.

## 6.8 Modelação dos factores de redução da mortalidade

Nesta secção, analisamos a aplicação dos GLM's na modelação e projecção dos factores de redução da mortalidade proposta por Renshaw e Haberman (2000, 2003b,c,d).

### 6.8.1 Caracterização do problema

Seja  $t$  o tempo cronológico, com origem em  $t = 0$ , situado no centro do período base, e  $q_{x,t}$  a probabilidade de óbito entre as idades  $x$  e  $x + 1$  no momento  $t \geq 0$ . O factor de redução para  $q_{x,t}$  é definido pela equação

$$RF(x, t) = \frac{q_{x,t}}{q_{x,0}}, \quad \forall t \geq 0 \quad (6.8.1)$$

ou, de forma equivalente, por

$$q_{x,t} = q_{x,0}RF(x, t), \quad \forall t \geq 0$$

sujeito às restrições

$$\begin{cases} RF(x, 0) = 1 & \forall x \geq 0 \\ 0 < RF(x, t) \leq 1 & \forall x, t \geq 0, \end{cases} \quad (6.8.2)$$

onde  $q_{x,0}$  denota o valor no ano base.

Alternativamente, o factor de redução pode ser definido em termos de  $\mu_{x,t}$ , a força de mortalidade à idade  $x$  no momento  $t$  ( $t \geq 0$ )

$$\mu_{x,t} = \mu_{x,0}RF(x, t), \quad \forall t \geq 0 \quad (6.8.3)$$

---

base na informação extraída do ajustamento e das forças de mortalidade projectadas averiguando, em particular, se este proporciona projecções que apresentam um padrão de evolução suave com respeito à idade e ao tempo cronológico.

## 6.8. Modelação dos factores de redução da mortalidade

---

sujeito, de novo, às restrições (6.8.2), e com  $\mu_{x,0}$  determinado de forma análoga ao referido para  $q_{x,0}$ . Considere-se agora a transformação logarítmica de (6.8.3)<sup>57</sup> e respectivas restrições, i.e.,

$$\log \mu_{x,t} = \log \mu_{x,0} + \log RF(x,t) \text{ tal que } \log RF(x,0) = 0, \quad \forall x, t \geq 0. \quad (6.8.4)$$

Se adoptarmos o esquema básico do método Lee-Carter, podemos reescrever a equação (6.8.4) como

$$\log \mu_{x,t} = \alpha_{x,0} + \beta_x \kappa_t \text{ tal que } \kappa_0 = 0, \quad t \geq 0, \quad (6.8.5)$$

onde

$$\begin{cases} \alpha_{x,0} = \log \mu_{x,0} \Leftrightarrow \mu_{x,0} = \exp(\alpha_{x,0}) \\ \beta_x \kappa_t = \log RF(x,t) \Leftrightarrow RF(x,t) = \exp(\beta_x \kappa_t). \end{cases}$$

Ao contrário do método LC original, neste caso a origem  $t = 0$  é bem definida em termos temporais, razão pela qual o parâmetro  $\alpha_{x,0}$  adopta um sufixo adicional. A unicidade da solução requer, uma vez mais, a normalização das estimativas dos parâmetros de acordo com as condições

$$\kappa_0 = 0, \quad \sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1, \quad \alpha_{x,0} = \log \mu_{x,0}.$$

### 6.8.2 Factores de redução e modelos GLM

#### Definição do modelo

Consideremos em primeiro lugar a modelação de  $q_{x,t}$ . Denotemos por  $d_{x,t}$  e  $E_{x,t}$  o número de óbitos e de expostos inicialmente ao risco na idade  $x$  e momento  $t$ , respectivamente. Admitamos que  $d_{x,t} \sim \text{Bin}(E_{x,t}; q_{x,t})$ . Estas hipóteses definem um modelo GLM com variável de resposta

$$Y_{x,t} = \frac{d_{x,t}}{E_{x,t}} \text{ com } \mathbb{E}(Y_{x,t}) = q_{x,t}, \quad \text{Var}(Y_{x,t}) = \phi \frac{q_{x,t}(1-q_{x,t})}{E_{x,t}}$$

com predictor linear  $\eta_{x,t}$  e função *link*  $g$ , monótona e diferenciável, tal que<sup>58</sup>

$$\eta_{x,t} = g(q_{x,t}) \Leftrightarrow q_{x,t} = g^{-1}(\eta_{x,t}),$$

---

<sup>57</sup>Idêntico procedimento pode ser aplicado no caso de factores de redução definidos em termos do quociente de mortalidade  $q_x$ .

<sup>58</sup>A metodoclogia é compatível com todas as funções *link* consideradas no Capítulo 3.

## 6.8. Modelação dos factores de redução da mortalidade

onde  $\phi$  denota um parâmetro de sobredispersão. Se considerarmos a equação (6.8.1), o factor de redução é definido por

$$RF(x, t) = \frac{g^{-1}(\eta_{x,t})}{q_{x,0}} \quad \text{com} \quad RF(x, 0) = 1. \quad (6.8.6)$$

Para acomodar esta restrição, os autores sugerem uma decomposição aditiva do previsor linear

$$\eta_{x,t} = \eta_{x,0} + \eta'_{x,t} \quad \text{com} \quad \eta'_{x,0} = 0,$$

donde resulta

$$RF(x, t) = \frac{g^{-1}(\eta_{x,0} + \eta'_{x,t})}{q_{x,0}}. \quad (6.8.7)$$

A primeira condição em (6.8.2) implica que  $g^{-1}(\eta_{x,0}) = q_{x,0}$  ou ainda  $\eta_{x,0} = g(q_{x,0})$ . O termo  $\eta_{x,0}$  representa, assim, um “offset” conhecido que é adicionado ao previsor linear. O passo seguinte consiste em encontrar uma função paramétrica para  $\eta'_{x,t}$  que respeite a condição  $\eta'_{x,0} = 0$ , com base na qual é possível escrever<sup>59</sup>

$$RF(x, t) = \frac{g^{-1}(\eta_{x,0} + \eta'_{x,t})}{g^{-1}(\eta_{x,0})} \quad \text{tal que} \quad RF(x, 0) = 1. \quad (6.8.8)$$

### Estimação, previsão e intervalos de confiança

Considere-se, a título de exemplo, a função paramétrica  $\eta'_{x,t} = \beta_x t$ . O ajustamento do modelo baseia-se no previsor linear

$$\eta_{x,t} = g(q_{x,0}) + \beta_x t$$

com *offset*  $g(q_{x,0})$ , calculado aplicando a função *link* às probabilidades graduadas da tábua base.

<sup>59</sup>Entre as funções paramétricas de interesse consideradas pelos autores destacam-se as seguintes:

$$\eta'_{x,t} = \beta_x t, \quad \forall x, t$$

$$\eta'_{x,t} = \sum_{i=1}^r \sum_{j=1}^k \beta_{ij} (x - x_j)_+^i \times t, \quad \forall x, t.$$

No primeiro caso, trata-se de uma recta que passa pela origem e tem inclinação dependente da idade. No segundo, trata-se de uma função de splines envolvendo  $k+1$  nós  $x_j$  ( $x_0 < x_1 < \dots < x_k$ ). A utilização de splines cúbicos ( $r = 3$  e  $\beta_{ij} = 0, \forall i = 1, 2$ ) acrescenta a possibilidade de se conseguir algum alisamento nos factores de redução com respeito à idade, tendo-se revelado, na maioria dos casos, uma solução adequada.

A correspondente estimativa dos factores de redução é dada por

$$\widehat{RF}(x, t) = \frac{\hat{q}_{x,t}}{q_{x,0}} = \frac{g^{-1} \left\{ g(q_{x,0}) + \hat{\beta}_x t \right\}}{q_{x,0}} \quad (6.8.9)$$

ou, em escala logarítmica, por

$$\log \widehat{RF}(x, t) = \log \left[ g^{-1} \left\{ g(q_{x,0}) + \hat{\beta}_x t \right\} \right] - \log(q_{x,0}) \quad (6.8.10)$$

A derivação de intervalos de confiança (bilaterais) com nível de confiança  $(1 - \alpha) \times 100\%$  para as estimativas socorre-se da hipótese de aproximação por uma distribuição Normal estandardizada (McCullagh e Nelder, 1989), i.e.,

$$g(q_{x,t}) = \hat{\eta}_{x,t} \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\eta}_{x,t})} = g(q_{x,0}) + \left( \hat{\beta}_x \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\beta}_x)} \right) t, \quad (6.8.11)$$

onde  $z_{\alpha/2}$  representa o quantil  $\alpha/2$  da distribuição Normal standardizada. Os correspondentes intervalos de confiança para os factores de redução são dados por

$$g^{-1} \left[ g(q_{x,0}) + \left( \hat{\beta}_x \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\beta}_x)} \right) t \right] / q_{x,0} \quad (6.8.12)$$

ou, em escala logarítmica, por

$$\log \left\{ g^{-1} \left[ g(q_{x,0}) + \left( \hat{\beta}_x \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\beta}_x)} \right) t \right] \right\} - \log q_{x,0}. \quad (6.8.13)$$

A previsão no modelo consegue-se pela extrapolação do comportamento do previsor linear para além do intervalo amostral, ou seja, para um momento  $t_{\max} + s$ ,  $s > 0$ . Os valores projectados e os respectivos intervalos de confiança são obtidos substituindo  $t_{\max} + s$  nas expressões (6.8.9) e (6.8.12), respectivamente.

### Factores de redução baseados na força de mortalidade

A metodologia apresentada nas secções anteriores pode ser adaptada à modelação de factores de redução baseados na força de mortalidade  $\mu_{x,t}$  (Renshaw e Haberman, 2000, 2003b). Admitamos que  $D_{x,t} \sim \text{Poisson}(E_{x,t}\mu_{x,t})$ , onde  $E_{x,t}$  representa agora o número de expostos centralmente ao risco. O modelo GLM é definido por

$$\mathbb{E}(D_{x,t}) = E_{x,t}\mu_{x,t} = E_{x,t}\mu_{x,0}RF(x, t), \quad \text{Var}(Y_{x,t}) = \phi\mathbb{E}(D_{x,t}).$$

## 6.8. Modelação dos factores de redução da mortalidade

Considere-se ainda a seguinte relação entre predictor linear e função *link*

$$\eta_{x,t} = g(\mu_{x,t}) \Leftrightarrow \mu_{x,t} = g^{-1}(\eta_{x,t}).$$

Se usarmos de novo uma decomposição aditiva do predictor linear, i.e.,

$$\eta_{x,t} = \eta_{x,0} + \eta'_{x,t} \quad \text{com} \quad \eta'_{x,0} = 0$$

o factor de redução é definido por

$$RF(x, t) = \frac{\mu_{x,t}}{\mu_{x,0}} = \frac{g^{-1}(\eta_{x,0} + \eta'_{x,t})}{\mu_{x,0}} \quad (6.8.14)$$

condicionado pela restrição (6.8.2), que implica

$$\mu_{x,0} = g^{-1}(\eta_{x,0}) \Leftrightarrow \eta_{x,0} = g(\mu_{x,0}).$$

Nos demais aspectos, a metodologia apresentada na secção anterior aplica-se, com a diferença de que  $q_{x,t}$  é substituído por  $\mu_{x,t}$  e os expostos inicialmente ao risco pelos expostos centralmente.<sup>60</sup> No caso em que optamos por um *link* logarítmico, o predictor linear é dado por

$$\eta_{x,t} = \log E_{x,t} + \log \mu_{x,0} + \log RF(x, t) \quad \text{tal que} \quad \log RF(x, 0) = 0,$$

onde  $(\log E_{x,t} + \log \mu_{x,0})$  corresponde a um termo *offset*.

O ajustamento do modelo é feito, como habitualmente, através da optimização da função de verosimilhança. As projecções de mortalidade são calculadas com recurso a<sup>61</sup>

$$\hat{\mu}_{x,t_{\max}+s} = \hat{\mu}_{x,t_{\max}} \exp(\hat{\beta}_x s), \quad s > 0. \quad (6.8.15)$$

<sup>60</sup>Renshaw e Haberman (2003b) sugerem, por exemplo, as seguintes estruturas paramétricas para  $\log RF(x, t)$ :

1.  $\log RF(x, t) = \gamma_t t + \beta_x t$
2.  $\log RF(x, t) = \gamma_t t + \beta'_x (t - t_j)_- + \beta_x t$

onde

$$(t)_- = \begin{cases} t, & t \leq 0 \\ 0, & t > 0 \end{cases}, \quad t \in [t_{\min}, t_{\max}], \quad t_{\min} < t_j < t_0 < t_{\max}$$

onde  $t_0$  expressa o momento base. No segundo modelo introduz-se alguma flexibilidade ao incluir uma possível quebra estrutural, situada em  $t_j < t_0$ , de modo a dar mais ênfase às tendências mais recentes na mortalidade. Os termos  $\gamma_t$  são adicionados de modo a garantir a igualdade entre os números observado e esperado de óbitos, tal como no método LC.

<sup>61</sup>Considerando, a título de exemplo, a função paramétrica  $\log RF(x, t) = \gamma_t t + \beta_x t$ .

## 6.8. Modelação dos factores de redução da mortalidade

---

Renshaw e Haberman (2005) desenvolvem uma reformulação do método LC que contempla mortalidade assintótica positiva. A reformulação, inspirada nos modelos adoptados pelo CMIB, envolve a seguinte definição alternativa para os factores de redução

$$RF_{\alpha}(x, t) = \alpha(x) + \{1 - \alpha(x)\} \exp(\beta_x \kappa_t), \quad 0 \leq \alpha(x) \leq 1 \quad (6.8.16)$$

baseada na reparametrização de (6.5.12) com

$$\beta_x = \frac{1}{20} \log \{1 - f_m(x)\}, \quad \kappa_t = t$$

e valores assintóticos não nulos  $\alpha(x)$ . Os valores projectados são agora dados por

$$\dot{RF}_{\alpha}(x, t_{\max} + s) = \alpha(x) + \{1 - \alpha(x)\} \exp \left[ \hat{\beta}_x (\dot{\kappa}_{t_{\max}+s} - \dot{\kappa}_{t_{\max}}) \right], \quad s > 0. \quad (6.8.17)$$

Apesar das muitas semelhanças entre o método LC e abordagem GLM, existe uma diferença fundamental no tratamento do tempo cronológico. Com efeito, enquanto no método LC o tempo é modelado como um factor multiplicativo  $\kappa_t$ , conjuntamente com  $\beta_x$ , mediante uma SVD, na abordagem GLM o tempo é explicitamente considerado como variável explicativa, com uma origem bem definida, sendo que apenas  $\beta_x$  necessita de ser estimado.<sup>62</sup>

### 6.8.3 Método Lee-Carter modificado

Renshaw e Haberman (2003a) desenvolvem uma versão modificada do método Lee-Carter. As principais novidades referem-se à proposição de soluções alternativas para estimar  $\alpha_x$  e à gradação das séries brutas  $\{\hat{\alpha}_x\}$  e  $\{\hat{\beta}_x\}$ . Em termos formais, considere-se o seguinte intervalo rectangular de dados com  $x \in [x_{\min}, x_{\max}]$  e  $t \in [t_{\min}, t_{\max}]$ . Seja  $[t_m, t_{\max}]$  o subintervalo de dimensão  $k = (t_{\max} - t_m + 1)$  compreendendo os dados que servem de base à construção de uma tábua contemporânea, em cujo centro se fixa a origem  $t_0$ ,  $t_0 = t_{\max} - [k/2]$ , onde  $[k/2]$  simboliza a parte inteira de  $k/2$  e  $t_{\min} \leq t_m \leq t_0 \leq t_{\max}$ . Os autores preconizam a estimação do modelo segundo as seguintes etapas:

---

<sup>62</sup> Outra diferença diz respeito ao facto do ajustamento destinado a igualar o número de óbitos observado ao respectivo valor estimado ser feito, no método LC, sem a devida compensação no parâmetro  $\beta_x$ , enquanto que na abordagem GLM ele é tido em conta na estimação dos parâmetros  $\beta_x$ . Em contrapartida, como o tempo é modelado como variável explicativa e o termo  $\alpha_x$  como um offset, na abordagem GLM a única fonte de incerteza está associada aos erros na estimação do parâmetro  $\beta_x$ , um aspecto que limita, de certa forma, a sua capacidade para capturar a incerteza quanto ao futuro.

## 6.8. Modelação dos factores de redução da mortalidade

1. Estimação de  $\alpha_{x,0}$  por qualquer um dos seguintes métodos:
  - (a)  $\hat{\alpha}_{x,0} = \log \mu_{x,0}$ , onde  $\{\mu_{x,0}\}$  é extraída da tábua graduada com base na experiência observada no intervalo de tempo  $[t_m, t_{\max}]$ .
  - (b)  $\hat{\alpha}_{x,0} = \log \left( \sum_{t=t_m}^{t_{\max}} d_{x,t} / \sum_{t=t_m}^{t_{\max}} E_{x,t} \right)$ , com base nos dados brutos  $\{d_{x,t}; E_{x,t}\}$ .
  - (c)  $\hat{\alpha}_{x,0} = \log \left( \prod_{t=t_m}^{t_{\max}} m_{x,t}^{1/k} \right)$ , i.e., com base na média geométrica das  $m_{x,t}$ .<sup>63</sup>
2. Alisamento da série  $\{\hat{\alpha}_{x,0}\}$  obtida em 1.(b) e 1.(c), caso se considere apropriado.
3. Estimação de  $(\beta_x, \kappa_t)$  mediante a aplicação da SVD à matriz  $[z_{x,t}] = [\ln \hat{m}_{x,t} - \hat{\alpha}_{x,0}]$ , respeitando as restrições (6.6.2) e (6.6.3).
4. Eventual graduação da série  $\{\hat{\beta}_x\}$  usando métodos paramétricos ou não-paramétricos.<sup>64</sup>
5. Ajustamento das estimativas  $\hat{\kappa}_t$  de modo a igualar os totais de óbitos observados e estimados.
6. Conversão dos valores  $\hat{\kappa}_t$  em  $(\hat{\kappa}_t - \hat{\kappa}_0)$  e análise dos resíduos.

A previsão no modelo baseia-se, tal como no método LC original, na modelação da série  $\{\hat{\kappa}_t\}$  usando um modelo ARIMA. Denotando por  $\{\hat{\kappa}_{t_{\max}+s} : s > 0\}$  a correspondente previsão, o valor projectado do factor de redução é dado por<sup>65</sup>

$$\log RF(x, t_{\max} + s) = \hat{\beta}_x \hat{\kappa}_{t_{\max}+s}, \quad s > 0. \quad (6.8.18)$$

Renshaw e Haberman (2003a) estabelecem ainda o paralelismo entre a reformulação do método LC e a metodologia desenvolvida por Renshaw e Haberman (2000) no contexto dos modelos GLM. Em termos mais concretos,

$$RF(x, t) = \frac{g^{-1} \{g(\exp \alpha_{x,0}) + \beta_x \kappa_t\}}{\exp \alpha_{x,0}}, \quad \kappa_0 = 0, \quad (6.8.19)$$

<sup>63</sup>No caso em que  $k = 1$ , as alternativas (b) e (c) são naturalmente equivalentes.

<sup>64</sup>Para uma aplicação da técnica de  $P$ -Splines na graduação dos termos log-bilineares do método Lee-Carter e do modelo de Brouhns *et al.* (2002a) veja-se Durban *et al.* (2002) e Currie *et al.* (2003).

<sup>65</sup>No caso particular em que adoptamos um modelo ARIMA(0, 1, 0), obtemos

$$\hat{\kappa}_{t_n+s} = \hat{\kappa}_{t_n} + \hat{\lambda}s.$$

Neste caso, a equação (6.8.18) reduz-se a

$$\log RF(x, t_{\max} + s) = \hat{\beta}_x (\hat{\kappa}_{t_{\max}} + \hat{\lambda}s), \quad s > 0.$$

onde  $g$  é uma função *link* pré-definida no contexto de modelos do tipo Binomial ou Poisson descritos anteriormente.<sup>66</sup>

#### 6.8.4 Método Lee-Carter com duplo previsor log-bilinear

Booth *et al.* (2002) e Renshaw e Haberman (2003c,d) desenvolvem uma extensão do método LC onde consideram um segundo previsor log-bilinear e uma abordagem multi-variada. Neste caso, a estrutura base do método LC dá lugar a

$$\eta_{x,t} = \alpha_x + \sum_{i=1}^2 \beta_x^{(i)} \kappa_t^{(i)} \quad (6.8.20)$$

no contexto de um modelo Gaussiano ou, de forma equivalente, a

$$\eta_{x,t} = \log E_{x,t} + \alpha_x + \sum_{i=1}^2 \beta_x^{(i)} \kappa_t^{(i)} \quad (6.8.21)$$

no quadro do modelo de Poisson.<sup>67</sup>

As restrições que asseguram a identificação do modelo são agora definidas por

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_t^{(1)} = \sum_{t=t_{\min}}^{t_{\max}} \kappa_t^{(2)} = 0 \quad (6.8.22)$$

$$\sum_{x=x_{\min}}^{x_{\max}} \beta_x^{(1)} = \sum_{x=x_{\min}}^{x_{\max}} \beta_x^{(2)} = 1. \quad (6.8.23)$$

Em consequência, os valores de  $\alpha_x$  continuam a ser estimados pela média das  $\ln(m_{x,t})$ . Para estimar os parâmetros  $\beta_x^{(i)}$  e  $\kappa_t^{(i)}$ ,  $i = 1, 2$ , são sugeridas duas soluções. Uma primeira, que envolve a incorporação das segundas componentes da SVD (Booth *et al.*, 2002). Uma segunda, que implica a reformulação do procedimento iterativo (6.9.10) proposto por Brouhns *et al.* (2002a) nas parcelas relativas aos parâmetros  $\beta_x^{(i)}$  e  $\kappa_t^{(i)}$ .

<sup>66</sup>Em particular, a utilização do *link* logarítmico na modelação de  $RF(x, t)$  para  $\mu_{x,t}$  conduz-nos a

$$RF(x, t) = \exp(\beta_x \kappa_t),$$

ou seja, a uma estrutura idêntica à descrita em (6.8.5).

<sup>67</sup>Esta mesma estrutura e metodologia podem ser usadas para obter projecções de taxas de mortalidade em populações desagregadas (Renshaw e Haberman, 2003d). A única diferença refere-se à utilização de um previsor do tipo

$$\eta_{x,j,t} = \log E_{x,j,t} + \alpha_{x,j} + \sum_{i=1}^2 \beta_{x,j}^{(i)} \kappa_t^{(i)},$$

onde  $j$  ( $j = 1, 2, \dots$ ) identifica o subconjunto da população considerado.

## 6.8. Modelação dos factores de redução da mortalidade

No que respeita à previsão no modelo, podemos optar por gerar projecções modelando o comportamento das séries  $\{\kappa_t^{(i)} : i = 1, 2\}$  usando processos ARIMA separados (Renshaw e Haberman, 2003c) ou, em alternativa, recorrer a um vector autoregressivo (VAR) do tipo (Renshaw e Haberman, 2003d)<sup>68</sup>

$$\mathbf{y}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \Phi \mathbf{y}_{t-1} + \varepsilon_t \quad (6.8.24)$$

com

$$\mathbf{a}_0 = \begin{pmatrix} a_{10} \\ a_{20} \end{pmatrix}, \quad \mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix},$$

onde se assume que o vector de erros  $\varepsilon_t \sim \mathcal{N}or(0, \Omega)$  e  $\Omega$  é uma matriz simétrica definida positiva. O modelo VAR pode ser aplicado sobre variáveis em níveis (i.e.,  $\mathbf{y}_t = (\kappa_t^{(1)}, \kappa_t^{(2)})$ ) ou em diferenças (i.e.,  $\mathbf{y}_t = (\Delta\kappa_t^{(1)}, \Delta\kappa_t^{(2)})$ ), testando a existência de relações de cointegração. Denotemos por  $\{\hat{\kappa}_{t_{\max}+s}^{(i)} : s > 0, i = 1, 2\}$  os valores projectados do parâmetro  $\kappa_t$ . As taxas de mortalidade previstas pelo modelo são calculadas mediante

$$\hat{m}_{x, t_{\max}+s} = m_{x, t_{\max}} \exp \left[ \sum_{i=1}^2 \beta_x^{(i)} \left( \kappa_{t_{\max}+s}^{(i)} - \kappa_{t_{\max}}^{(i)} \right) \right], \quad s > 0, \quad (6.8.25)$$

onde

$$RF(x, t_{\max} + s) = \exp \left[ \sum_{i=1}^2 \beta_x^{(i)} \left( \kappa_{t_{\max}+s}^{(i)} - \kappa_{t_{\max}}^{(i)} \right) \right], \quad s > 0. \quad (6.8.26)$$

pode, uma vez mais, ser interpretado como um factor de redução da mortalidade.

### 6.8.5 Factores de redução com inclusão de efeitos geracionais

Num estudo recente sobre as tendências da mortalidade no Reino Unido e num conjunto de países desenvolvidos, Willets (2004) e Willets *et al.* (2004) encontraram evidência empírica suficiente para suportar a importância do efeito geração (*cohort*) no comportamento da mortalidade. Os estudos concluem que os efeitos temporais no comportamento da mortalidade acompanham a ocorrência de factores contemporâneos, como sejam o nível geral de saúde da população, a qualidade dos cuidados de saúde, a existência de condições climáticas especialmente adversas, entre outros. Em relação ao efeito geracional, os estudos indicam que existe uma relação próxima entre estes e alterações estruturais (e.g., II Guerra Mundial, mudanças nos hábitos alimentares e de tabagismo). Os estudos recomendam ainda a construção de cenários e a análise das causas de morte

<sup>68</sup>Para uma análise detalhada dos modelos VAR veja-se, por exemplo, Hamilton (1994).

## 6.8. Modelação dos factores de redução da mortalidade

para uma efectiva apreciação dos efeitos temporal e geracional.

A pensar neste fenómeno, foram propostos recentemente modelos que conjugam, simultaneamente, os efeitos idade, tempo cronológico e geração (e.g., Tabeau (2001), CMIB (2005)). O modelo base é definido por

$$\log \mu_{x,t} = s_a(x) + s_b(t) + s_c(\tau), \quad (6.8.27)$$

onde  $s_a$ ,  $s_b$  e  $s_c$  são funções da idade  $x$ , do ano  $t$  e da geração  $\tau = t - x$ , respectivamente.

Ainda neste domínio, Renshaw e Haberman (2005, 2006) desenvolvem uma adaptação do método LC com vista à projecção de factores de redução com inclusão de efeitos geracionais. Seja  $(d_{x,t}; E_{x,t})$  uma matriz rectangular onde  $d_{x,t}$  denota o número de óbitos e  $E_{x,t}$  simboliza o correspondente exposição ao risco. Supõe-se que os dados estão classificados por ano cronológico  $t \in [t_{\min}, t_{\max}]$  e por idade  $x \in [x_{\min}, x_{\max}]$ , agrupados por categorias, ou por ano cronológico individual, caso em que o ano de nascimento (ou ano da geração)  $\tau = t - x$ ,  $\tau \in [t_{\min} - x_{\max}, t_{\max} - x_{\min}]$  se encontra bem definido.

O objectivo da análise é agora o de captar as tendências temporais na mortalidade em cada idade, incorporando os efeitos geracionais quando a base de dados o permitir. Para o efeito, considera-se a seguinte definição geral do factor de redução

$$RF(x, t) = \exp \left\{ \beta_x^{(0)} \xi_{t-x} + \sum_{i=1}^c \beta_x^{(i)} \kappa_t^{(i)} \right\} \quad (6.8.28)$$

$$\mu_{x,t} = \exp(\alpha_x + \log RF(x, t)), \quad (6.8.29)$$

onde  $\xi_{t-x}$  identifica um índice que capta os efeitos geracionais,  $c$  denota a ordem da estrutura log-bilinear do modelo e as demais variáveis mantêm o significado habitual. O modelo definido em (6.8.28) incorpora, como casos particulares, algumas das estruturas já analisadas neste capítulo. De facto, no caso em que  $\beta_x^{(0)} = 0$  e  $c = 1$  obtemos a estrutura log-bilinear original do método LC. No caso especial em que  $c = 0$  ( $\beta_x^{(0)} \equiv \beta_x$ ), o modelo refere-se a uma estrutura em que se consideram apenas o efeito idade e geração.

Quando formulado no contexto dos GLMs, o modelo é definido por

$$Y_{x,t} = D_{x,t}, \quad \mathbb{E}(Y_{x,t}) = \mu_{x,t} E_{x,t}, \quad \text{Var}(Y_{x,t}) = \phi \frac{V(Y_{x,t})}{w_{x,t}} \quad (6.8.30)$$

com parâmetro de escala  $\phi$ , função de variância  $V(Y_{x,t}) = \mathbb{E}(Y_{x,t})$  e ponderadores fixos  $w_{x,t} = 1$ .

## 6.8. Modelação dos factores de redução da mortalidade

---

Usando um *link* logarítmico, o previsor não-linear  $\eta_{x,t}$  é definido por

$$\eta_{x,t} = \log E(Y_{x,t}) = \log E_{x,t} + \alpha_x + \log RF(x, t).$$

A estimação deste modelo é complexa uma vez que as variáveis estão restringidas pela relação geração=(período-idade). Para assegurar uma estimativa única para os parâmetros, os autores propõem um procedimento implementado em duas fases:

1. Estimação de  $\alpha_x$  segundo o método LC original (com média geométrica)

$$\hat{\alpha}_x = \log \left( \prod_{t=t_{\min}}^{t_{\max}} m_{x,t}^{1/T} \right), \quad T = t_{\max} - t_{\min} + 1.$$

2. Os restantes parâmetros são estimados adaptando o procedimento iterativo descrito em (6.9.10), considerando as variáveis de resposta  $y_{x,t} = d_{x,t}$  no modelo de Poisson, *link* logarítmico, *offset* ( $\log E_{x,t} + \alpha_x$ ) e restrições

$$\xi_{t_{\min}-x_{\max}} = 0, \quad \sum_{x=x_{\min}}^{x_{\max}} \beta_x^{(0)} = 1, \quad \kappa_{t_{\min}} = 0, \quad \sum_{x=x_{\min}}^{x_{\max}} \beta_x^{(1)} = 1.$$

As taxas de mortalidade projectadas, alinhadas com os últimos valores conhecidos  $\hat{m}_{x,t_{\max}}$ , são calculadas mediante

$$\hat{m}_{x,t_{\max}+s} = \hat{m}_{x,t_{\max}} \dot{RF}(x, t_{\max} + s), \quad (6.8.31)$$

onde

$$\dot{RF} = \exp \left\{ \hat{\beta}_x^{(0)} \left( \tilde{\xi}_{t_{\max}-x+s} - \hat{\xi}_{t_{\max}-x} \right) + \sum_{i=1}^c \hat{\beta}_x^{(i)} \left( \dot{\kappa}_{t_{\max}+s}^{(i)} - \hat{\kappa}_{t_{\max}}^{(i)} \right) \right\}, \quad s > 0, \quad (6.8.32)$$

e onde  $(\hat{\beta}_x^{(i)}, \hat{\xi}_{t-x}, \hat{\kappa}_t^{(i)})$  são vectores de parâmetros estimados. As séries  $(\tilde{\xi}_{t-x}, \dot{\kappa}_t^{(i)})$  de valores projectados são derivadas a partir de modelos ARIMA separados

$$\left\{ \hat{\kappa}_t^{(i)} : t \in [t_{\min}, t_{\max}], i = 1, \dots, c \right\} \mapsto \left\{ \dot{\kappa}_{t_{\max}+s}^{(i)} : s > 0, i = 1, \dots, c \right\}$$

$$\left\{ \hat{\xi}_\tau^{(i)} : \tau \in [t_{\min} - x_{\max}, t_{\max} - x_{\min}] \right\} \mapsto \left\{ \dot{\xi}_{t_{\max}-x_{\min}+s}^{(i)} : s > 0 \right\},$$

onde

$$\tilde{\xi}_{t_{\max}-x_{\min}+s} = \begin{cases} \hat{\xi}_{t_{\max}-x_{\min}+s}, & 0 \leq s \leq (x - x_{\min}) \\ \check{\xi}_{t_{\max}-x_{\min}+s}, & s > (x - x_{\min}) \end{cases}$$

## 6.9 Modelo de Poisson

### 6.9.1 Definição

Uma das principais críticas apontadas ao método LC refere-se a hipótese de que os erros são normalmente distribuídos com variância constante. A hipótese de homocedasticidade é claramente irrealista na medida em que se sabe que, devido à existência de um número reduzido de óbitos nas idades avançadas, o logaritmo de  $\mu_x$  é muito mais variável nestas idades do que nas idades jovens. Para contornar este problema, Brouhns *et al.* (2002a) desenvolvem um modelo, fundado no método LC, em que consideram que o número de óbitos verificados à idade  $x$  no ano  $t$  segue uma distribuição de Poisson

$$D_{x,t} \sim \text{Poisson}(\mu_{x,t} E_{x,t}) \quad (6.9.1)$$

com

$$\mu_{x,t} = \exp(\alpha_x + \beta_x \kappa_t), \quad (6.9.2)$$

onde os parâmetros respeitam as restrições<sup>69</sup>

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0 \quad \text{e} \quad \sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1. \quad (6.9.3)$$

O modelo definido por (6.9.1)-(6.9.2)-(6.9.3), que designaremos doravante por modelo de Poisson, requer ainda, no contexto da elaboração de tábuas de mortalidade prospectivas, uma generalização da hipótese (2.2.30) de constância da força de mortalidade entre idades, de modo a acomodar a dimensão temporal. Neste sentido, o modelo assume que  $\mu_{x,t}$  é constante no interior de cada quadrado do diagrama de Lexis, ou seja, admite que a mortalidade sofrida em cada instante por um indivíduo de idade  $x$  no momento  $t$  é constante em cada período sucessivo de um ano. Formalmente,

$$\mu_{x+\xi, t+\tau} = \mu_{x,t}, \quad 0 \leq \xi, \tau < 1 \quad \text{com } x \text{ e } t \text{ inteiros.} \quad (6.9.4)$$

<sup>69</sup> As semelhanças entre os métodos de Lee-Carter e de Poisson e a abordagem proposta por Renshaw *et al.* (1996), Renshaw e Haberman (2003a,b) e Sithole *et al.* (2000) são evidentes, diferenciando-se no entanto pelo já mencionado tratamento distinto do elemento temporal. De facto, enquanto que no primeiro caso o tempo é modelado como um factor, no segundo o elemento temporal é modelado como uma covariável conhecida.

Com base na hipótese (6.9.4), deduz-se igualmente que

$${}_{\xi}p_{x,t} = \exp(-\xi\mu_{x,t}) = \{p_{x,t}\}^{\xi}, \quad 0 \leq \xi < 1$$

e que a exposição ao risco à idade  $x$  e ano  $t$ ,  $E_{x,t}$ , tem a seguinte relação com  $q_{x,t}$

$$E_{x,t} = \frac{-l_{x,t}q_{x,t}}{\ln(1 - q_{x,t})},$$

onde  $l_{x,t}$  denota o número de sobreviventes à idade  $x$  e ano  $t$ .

Como se observa, o modelo preserva a estrutura log-bilinear para  $\mu_{x,t}$ , mas substitui as hipóteses relativa ao termo erro  $\epsilon_{x,t}$  pela lei de Poisson para  $d_{x,t}$ . Os parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  mantêm, na sua essência, o significado originalmente atribuído pelo método Lee-Carter.

Os proponentes do modelo de Poisson identificam, no entanto, algumas vantagens face ao método LC que o tornam especialmente atractivo em aplicações actuariais. Em primeiro lugar, o modelo reconhece explicitamente a natureza inteira de  $D_{x,t}$ , contrariamente ao que acontece no método LC. Em segundo, o modelo deixa cair a hipótese de homocedasticidade do termo erro e reconhece a maior variabilidade de  $\mu_{x,t}$  nas idades mais elevadas. Em terceiro, a especificação (6.9.1) permite o recurso ao método de máxima verosimilhança para estimar os parâmetros. O facto do ajustamento do modelo não depender de uma SVD significa que este não exige uma matriz rectangular de dados completa. Por fim, e não menos importante, o modelo dispensa o procedimento de reestimação dos  $\hat{\kappa}_t$ 's.

Uma das principais vantagens do modelo de Poisson é a sua flexibilidade. Com efeito, o modelo pode ser facilmente expandido de modo a considerar o impacto sobre a mortalidade de outras variáveis explicativas (covariáveis) que não apenas a idade  $x$  ou o tempo cronológico  $t$  (e.g., profissão, grau de incapacidade, nível de rendimento, etc.). Efectivamente, se adoptarmos, por exemplo, o esquema multiplicativo proposto por Cox (1972), a formulação (6.9.1) é substituída por

$$D_{x,t} \sim \text{Poisson} \left( E_{x,t}\mu_{x,t} \exp \left( \beta_0 + \sum_{j=1}^s \beta_j z_{x,t,j} \right) \right), \quad (6.9.5)$$

onde  $\mu_{x,t}$  representa a mortalidade base e  $z_{x,t,j}$  são covariáveis codificando a presença (ou a ausência) e a magnitude de características adicionais na população.

### 6.9.2 Fundamentação da utilização da distribuição de Poisson

Considere-se uma observação  $(x, t)$  qualquer. São registados  $D_{x,t}$  óbitos de um total de  $l_{x,t}$  indivíduos com idade  $x$  recenseados no período de observação anual. Admitamos que as suas durações de vida podem ser consideradas variáveis aleatórias independentes e identicamente distribuídas. Associe a cada um dos indivíduos uma variável  $\delta_{x,i}$ , que indica a ocorrência do óbito, i.e.,

$$\delta_{x,i} = \begin{cases} 1 & \text{se o indivíduo } i \text{ faleceu à idade } x \\ 0 & \text{caso contrário,} \end{cases}$$

$i = 1, 2, \dots, l_{x,t}$ . Representemos por  $\tau_{x,i}$  o tempo vivido pelo indivíduo  $i$ . Em termos mais precisos,  $\tau_{x,i} = 1$  no caso do indivíduo  $i$  alcançar a idade  $x + 1$ . No caso deste falecer à idade  $x$ ,  $\tau_{x,i}$  simboliza a fracção de ano vivida pelo indivíduo. Admitamos que dispomos, para cada um dos  $l_{x,t}$  indivíduos, das observações  $(\delta_{x,i}, \tau_{x,i})$ . Se considerarmos a hipótese (6.9.4), a contribuição do  $i^{\text{ésimo}}$  indivíduo para a verosimilhança pode ser definida da seguinte forma:

- no caso do indivíduo sobreviver, a verosimilhança correspondente é dada por

$$p_{x,t} = \exp(-\mu_{x,t})$$

- no caso do indivíduo falecer, a verosimilhança associada é a seguinte

$$\tau_{x,i} p_{x,t} \mu_{x,t}^{\tau_{x,i}} = \exp(-\tau_{x,i} \mu_{x,t}) \mu_{x,t}^{\tau_{x,i}}$$

Em alternativa, a contribuição do  $i^{\text{ésimo}}$  indivíduo pode ser rescrita segundo

$$\exp(-\tau_{x,i} \mu_{x,t}) \{\mu_{x,t}\}^{\delta_{x,i}}$$

Em resultado da hipótese de independência, é possível escrever a verosimilhança associada aos  $l_{x,t}$  indivíduos da seguinte forma

$$\begin{aligned} \mathcal{V}(\mu_{x,t}) &= \prod_{i=1}^{l_{x,t}} \exp(-\tau_{x,i} \mu_{x,t}) \{\mu_{x,t}\}^{\delta_{x,i}} \\ &= \exp(-\mu_{x,t} \tau_{\bullet}) \{\mu_{x,t}\}^{\delta_{\bullet}}, \end{aligned} \quad (6.9.6)$$

onde

$$\tau_{\bullet} = \sum_{i=1}^{l_{x,t}} \tau_{x,i} \text{ e } \delta_{\bullet} = \sum_{i=1}^{l_{x,t}} \delta_{x,i}.$$

O termo  $\delta_{\bullet}$  representa, assim, o número de óbitos observados à idade  $x$  no ano  $t$ ,  $D_{x,t}$ , enquanto  $\tau_{\bullet}$  denota o tempo total vivido pelos  $l_{x,t}$  indivíduos no decurso do ano  $t$ . A verosimilhança (6.9.6) difere da que decorre da hipótese (6.9.1) apenas por uma constante, pelo que no caso de recorrermos ao princípio de máxima verosimilhança para estimar os parâmetros do modelo, resulta indiferente trabalhar com uma ou outra.<sup>70</sup>

### 6.9.3 Estimação dos parâmetros por máxima verosimilhança

Uma vantagem importante deste modelo face ao método LC reside, como vimos, no facto de ele permitir a estimação dos parâmetros pelo método de máxima verosimilhança, sem necessidade de recorrer à SVD ou ao método dos mínimos quadrados ponderados. Em termos formais, designemos por  $\lambda_{x,t}$  o número esperado de óbitos na idade  $x$  e ano  $t$ , i.e.

$$\lambda_{x,t} = \mathbb{E}[D_{x,t}] = E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) \quad (6.9.7)$$

e por  $d_{x,t}$  o respectivo valor observado de  $D_{x,t}$ .

A função de log-verosimilhança  $\mathcal{L}(\alpha, \beta, \kappa) = \ln \mathcal{V}(\alpha, \beta, \kappa)$  é definida por

$$\begin{aligned} \mathcal{L}(\alpha, \beta, \kappa) &= \ln \left\{ \prod_{t=t_{\min}}^{t_{\max}} \prod_{x=x_{\min}}^{x_{\max}} \left( \frac{\lambda_{x,t}^{d_{x,t}} \exp(-\lambda_{x,t})}{(d_{x,t})!} \right) \right\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \{d_{x,t} \ln \lambda_{x,t} - \lambda_{x,t} - \ln [(d_{x,t})!]\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \{d_{x,t} \ln E_{x,t} + d_{x,t} (\alpha_x + \beta_x \kappa_t) \\ &\quad - E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) - \ln [(d_{x,t})!]\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \{d_{x,t} (\alpha_x + \beta_x \kappa_t) - E_{x,t} \exp(\alpha_x + \beta_x \kappa_t)\} + c, \end{aligned} \quad (6.9.8)$$

onde  $\alpha = (\alpha_{x_{\min}}, \dots, \alpha_{x_{\max}})$ ,  $\beta = (\beta_{x_{\min}}, \dots, \beta_{x_{\max}})$ ,  $\kappa = (\kappa_{t_{\min}}, \dots, \kappa_{t_{\max}})$  e  $c$  é uma constante.

<sup>70</sup>Brillinger (1986) demonstra, com base na formulação de hipóteses razoáveis quanto aos processos que governam os nascimentos e os óbitos, que a distribuição de Poisson surge naturalmente como uma solução particularmente adequada para modelar o número de óbitos observados nas diferentes idades.

Os estimadores de máxima verosimilhança dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  correspondem aos valores que maximizam  $\mathcal{L}(\alpha, \beta, \kappa)$ . A presença do termo bilinear  $\beta_x \kappa_t$  na especificação (6.9.2) de  $\mu_{x,t}$  impede a estimação dos parâmetros pelos *packages* informáticos tradicionais que incluem o modelo de Poisson. Por esta razão, os autores propõem a adopção do algoritmo iterativo desenvolvido por Goodman (1979), assente no método de Newton-Raphson. O procedimento assenta num princípio simples: em cada iteração  $v$ , um único conjunto de parâmetros  $\theta_j$  é actualizado, mantendo os restantes fixos à sua estimativa anterior, com a ajuda do seguinte esquema de actualização

$$\hat{\theta}_j^{(v+1)} = \hat{\theta}_j^{(v)} - \frac{\partial \mathcal{L}^{(v)} / d\theta_j}{\partial^2 \mathcal{L}^{(v)} / d\theta_j^2} \quad (6.9.9)$$

onde  $\mathcal{L}^{(v)} = \mathcal{L}(\hat{\theta}^{(v)})$ .

No nosso caso, existem três conjuntos distintos de parâmetros, correspondentes aos termos  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ . Neste sentido, dado um vector de valores iniciais para os parâmetros  $\hat{\alpha}_x^{(0)}$ ,  $\hat{\beta}_x^{(0)}$  e  $\hat{\kappa}_t^{(0)}$ , as estimativas são revistas segundo o algoritmo

$$\hat{\alpha}_x^{(v+1)} = \hat{\alpha}_x^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)})]}{\sum_{t=t_{\min}}^{t_{\max}} [E_{x,t} \exp(\hat{\alpha}_x^{(v)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)})]} \quad (6.9.10)$$

$$\hat{\kappa}_t^{(v+1)} = \hat{\kappa}_t^{(v)} - \frac{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(v)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)})]}{\sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_x^{(v)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)})]}$$

$$\hat{\beta}_x^{(v+1)} = \hat{\beta}_x^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(v+1)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v+1)})]}{\sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_t^{(v+1)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v+1)})]}$$

ou, em alternativa, recorrendo à formulação matricial de (6.9.10)

$$\begin{aligned} (\mathbf{G}_\alpha^{(\nu)}) \hat{\alpha}^{(\nu+1)} &= (\mathbf{G}_\alpha^{(\nu)}) \hat{\alpha}^{(\nu)} + \mathbf{L}_\alpha^{(\nu)} \\ (\mathbf{G}_\kappa^{(\nu)}) \hat{\kappa}^{(\nu+1)} &= (\mathbf{G}_\kappa^{(\nu)}) \hat{\kappa}^{(\nu)} + \mathbf{L}_\kappa^{(\nu)} \\ (\mathbf{G}_\beta^{(\nu)}) \hat{\beta}^{(\nu+1)} &= (\mathbf{G}_\beta^{(\nu)}) \hat{\beta}^{(\nu)} + \mathbf{L}_\beta^{(\nu)}, \end{aligned} \quad (6.9.11)$$

## 6.9. Modelo de Poisson

onde  $\mathbf{G}_\alpha^{(\nu)}$ ,  $\mathbf{G}_\kappa^{(\nu)}$  e  $\mathbf{G}_\beta^{(\nu)}$  são matrizes diagonais de dimensão  $(\mathcal{N} \times \mathcal{N})$ ,  $(\mathcal{T} \times \mathcal{T})$  e  $(\mathcal{N} \times \mathcal{N})$ , respectivamente, e elemento representativo definido por

$$\begin{aligned} (\mathbf{G}_\alpha^{(\nu)})_{xx} &= \sum_{t=t_{\min}}^{t_{\max}} E_{x,t} \exp(\hat{\alpha}_x^{(\nu)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)}) \\ (\mathbf{G}_\kappa^{(\nu)})_{tt} &= \sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_x^{(\nu)})^2 E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)}) \\ (\mathbf{G}_\beta^{(\nu)})_{xx} &= \sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_t^{(\nu+1)})^2 E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu+1)}) \end{aligned} \quad (6.9.12)$$

e  $\mathbf{L}_\alpha^{(\nu)}$ ,  $\mathbf{L}_\kappa^{(\nu)}$  e  $\mathbf{L}_\beta^{(\nu)}$  são vectores coluna de dimensão  $\mathcal{N}$ ,  $\mathcal{T}$  e  $\mathcal{N}$ , respectivamente, e elemento representativo definido por

$$\begin{aligned} (\mathbf{L}_\alpha^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} \left[ d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)}) \right] \\ (\mathbf{L}_\kappa^{(\nu)})_t &= \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(\nu)} \left[ d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu)}) \right] \\ (\mathbf{L}_\beta^{(\nu)})_x &= \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(\nu+1)} \left[ d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{\kappa}_t^{(\nu+1)}) \right], \end{aligned} \quad (6.9.13)$$

com  $\mathcal{N} = (x_{\max} - x_{\min} + 1)$  e  $\mathcal{T} = (t_{\max} - t_{\min} + 1)$ .

A iniciação do algoritmo e a fixação do critério de óptimo obedecem a procedimentos idênticos aos definidos para o método LC.<sup>71</sup> Ao contrário do que acontece no método LC, a maximização de  $\mathcal{L}(\alpha, \beta, \kappa)$  garante a igualdade entre os números observado e estimado de óbitos em cada idade ao longo do período de observação dispensando, por isso, o procedimento de reestimação dos  $\hat{\kappa}_t$ 's descrito na etapa 3 do método LC.

As estimativas  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  geradas pelo método de máxima verosimilhança não garantem, automaticamente, o cumprimento das restrições (6.9.3). Para o conseguir, é introduzida uma reparametrização que converte  $(\hat{\alpha}, \hat{\beta}, \hat{\kappa})$  em  $(\alpha^*, \beta^*, \kappa^*)$  com

$$\kappa_t^* \leftarrow (\hat{\kappa}_t - \bar{\kappa}) K \quad \text{e} \quad \beta_x^* \leftarrow \frac{\hat{\beta}_x}{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x}, \quad (6.9.14)$$

onde  $\bar{\kappa}$  corresponde à média aritmética de  $\hat{\kappa}_t$ , i.e.,

$$\bar{\kappa} = \frac{1}{\mathcal{T}} \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t$$

<sup>71</sup>Em alternativa, Renshaw e Haberman (2003b) propõem que o critério de óptimo envolva a optimização da função de verosimilhança monitorando o valor da estatística deviance.

e  $K$  é uma constante arbitrária. Os autores optam por seleccionar  $K$  segundo

$$K = \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x.$$

Por fim, da necessidade de respeitar a igualdade

$$\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t = \alpha_x^* + \beta_x^* \kappa_t^*$$

deduz-se que

$$\begin{aligned} \alpha_x^* &= \hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t - \beta_x^* \kappa_t^* \\ &= \hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t - \frac{\hat{\beta}_x}{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x} (\hat{\kappa}_t - \bar{\kappa}) K \\ &= \hat{\alpha}_x + \hat{\beta}_x \bar{\kappa} \end{aligned} \quad (6.9.15)$$

As novas estimativas  $\alpha_x^*$ ,  $\beta_x^*$  e  $\kappa_t^*$  asseguram o cumprimento das restrições (6.9.3) e fornecem a mesma estimativa do número de óbitos  $\hat{d}_{x,t}$  dado que  $\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t = \alpha_x^* + \beta_x^* \kappa_t^*$ . A qualidade do ajustamento é aferida pelo cálculo da deviance

$$\mathcal{D}(d_{x,t}, \hat{\lambda}_{x,t}) = 2 \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \omega_{x,t} \left\{ d_{x,t} \ln \left( \frac{d_{x,t}}{\hat{\lambda}_{x,t}} \right) - (d_{x,t} - \hat{\lambda}_{x,t}) \right\}, \quad (6.9.16)$$

onde

$$\omega_{x,t} = \begin{cases} 1, & E_{x,t} > 0 \\ 0, & E_{x,t} = 0 \end{cases}$$

e pela análise do comportamento dos resíduos

$$r_{x,t}^{\mathcal{D}} = \sqrt{2} \cdot \text{sign}(d_{x,t} - \hat{\lambda}_{x,t}) \sqrt{d_{x,t} \ln \frac{d_{x,t}}{\hat{\lambda}_{x,t}} - (d_{x,t} - \hat{\lambda}_{x,t})}. \quad (6.9.17)$$

A previsão no modelo baseia-se, à semelhança do que acontece no método LC, na aplicação da metodologia de Box-Jenkins à série temporal das estimativas  $\hat{\kappa}_t$ .

Um ponto delicado nos métodos de LC e de Poisson reside na não simultaneidade dos procedimentos de estimação dos parâmetros, uma vez que a modelação decorre em duas fases. Numa primeira, os parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  são estimados. Numa segunda, o parâmetro  $\kappa_t$  é extrapolado usando a metodologia de Box-Jenkins.

Para contornar esta limitação, Czado *et al.* (2005) desenvolvem uma formulação

Bayesiana do modelo de Poisson onde, após escolher a ordem do modelo ARIMA, procedem à estimação simultânea da totalidade dos parâmetros com o auxílio de técnicas de simulação de MCMC (*Markov Chain Monte Carlo*). A abordagem Bayesiana trata todos os parâmetros desconhecidos  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  como variáveis aleatórias e deriva a sua distribuição condicionada à informação conhecida  $(E_{x,t}; d_{x,t})$ . Esta abordagem apresenta ainda a vantagem de permitir estimar directamente intervalos de confiança para os indicadores de interesse.

### 6.9.4 Extensão do modelo com incorporação de P-Splines e estimação dos parâmetros por máxima verosimilhança penalizada

Os estudos empíricos realizados com base no modelo de Poisson (e no de Lee-Carter) revelam que as estimativas dos parâmetros apresentam, nalguns casos, alguma irregularidade entre idades consecutivas.<sup>72</sup> Se no caso da tendência temporal  $\hat{\kappa}_t$  este perfil é de certo modo esperado, na medida em que se admite que contemporaneamente ocorrem eventos com repercussões na mortalidade (e.g., avanços na medicina, conflitos militares, epidemias), nos casos das séries  $\hat{\alpha}_x$  e  $\hat{\beta}_x$  as descontinuidades devem merecer uma atenção especial.

O parâmetro  $\alpha_x$  representa, como vimos, o comportamento médio da mortalidade em cada idade e num intervalo de tempo relativamente longo. Por este motivo, não é expectável que as estimativas geradas pelo modelo apresentem anomalias significativas, uma hipótese que é confirmada pelos estudos empíricos.<sup>73</sup> Em relação às estimativas do parâmetro  $\hat{\beta}_x$ , o cenário é distinto na medida em que a generalidade dos estudos confirma a presença de alguma irregularidade. O parâmetro  $\beta_x$  condiciona, como vimos, todo o processo de extrapolação das taxas de mortalidade para o futuro. Neste sentido, qualquer irregularidade verificada nas suas estimativas traduzir-se-à, inevitavelmente, em tábuas de mortalidade projectadas com oscilações erráticas nos quocientes de mortalidade entre idades consecutivas, uma característica indesejada do ponto de vista actuarial.

De facto, os actuários baseiam o cálculo dos prémios e das responsabilidades financeiras das companhias de seguros em tábuas de mortalidade. Qualquer irregularidade verificada nestas propaga-se no *pricing* das apólices e nos balanços das empresas. Se exceptuarmos os casos em que as anomalias correspondem a características particulares

---

<sup>72</sup>No Capítulo 7 teremos oportunidade de verificar que o estudo realizado para a população portuguesa reforça esta afirmação.

<sup>73</sup>O facto de apresentar um perfil suave não significa, contudo, que a curva dos  $\hat{\alpha}_x$  deixa, em momento algum, de representar os fenómenos habituais, como sejam a mortalidade infantil e juvenil mais elevada, a tradicional “bossa” de acidentes, mais significativa nos homens, e a mortalidade crescente nas idades avançadas.

do risco coberto pelas seguradoras, a causa provável para este padrão reside em erros amostrais, razão já antes avançada para justificar o recurso a técnicas de graduação.

Uma solução simplista para este problema passaria, por exemplo, pelo alisamento das estimativas brutas dos parâmetros do modelo de Poisson recorrendo, por exemplo, a um dos métodos de graduação descritos no Capítulo 4. Uma outra envolveria o alisamento de alguns dos parâmetros (e.g.,  $\hat{\beta}_x$ ), preservando as estimativas iniciais dos restantes.<sup>74</sup> Embora simples, este procedimento tende, em potência, a produzir discrepâncias intoleráveis entre os valores observados para a mortalidade e os valores estimados pelo modelo. Uma alternativa possível envolve a graduação das tábuas de mortalidade produzidas pelo procedimento de estimação usando, por exemplo, métodos não-paramétricos. Contudo, esta solução implica, na prática, a perda de informação importante, em particular aquela que se refere à decomposição da projecção nas três séries de parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ .

Para resolver este problema, ensaiamos nesta secção uma nova técnica que incorpora as preocupações quanto ao alisamento das séries no procedimento de estimação dos parâmetros do modelo de Poisson.<sup>75</sup> A solução é inspirada na técnica de P-Splines desenvolvida por Eilers e Marx (1996) e analisada em detalhe no Capítulo 4. Em termos mais precisos, sugere-se a imposição de uma penalização (de ordem  $d$ ) incidindo sobre a diferença finita entre os coeficientes estimados do modelo de Poisson.

Denotemos por  $D_\alpha^{(d)}$ ,  $D_\beta^{(d)}$  e  $D_\kappa^{(d)}$  os operadores de diferenças  $\Delta^d$  em notação matricial, onde  $\Delta^d$  denota a  $d^{\text{ésima}}$  diferença finita, com dimensão  $(\mathcal{N} - d) \times \mathcal{N}$ ,  $(\mathcal{N} - d) \times \mathcal{N}$  e  $(\mathcal{T} - d) \times \mathcal{T}$ , respectivamente. Como vimos no Capítulo 4, a estrutura dos operadores de diferenças pode assumir inúmeras configurações possíveis. Uma matriz de diferenciação de primeira ordem ( $d = 1$ ),  $D_i^{(1)}$ ,  $i \in (\alpha, \beta, \kappa)$ , do tipo

$$D_i^{(1)} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

introduzirá uma penalização sobre as diferenças entre dois valores consecutivos dos parâmetros (e.g.,  $\Delta\beta_x = \beta_x - \beta_{x-1}$ ). De forma análoga, uma matriz de diferenciação de

<sup>74</sup>Esta solução não é virgem uma vez que, como vimos na Secção 6.8.3, foi sugerida por Renshaw e Haberman (2003a) no âmbito da versão modificada do método LC.

<sup>75</sup>Solução análoga pode ser desenvolvida para o método Lee-Carter.

segunda ordem ( $d = 2$ ),  $D_i^{(2)}$ ,  $i \in (\alpha, \beta, \kappa)$ , do tipo

$$D_i^{(2)} = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}$$

penalizará as diferenças de segunda ordem (e.g.,  $\Delta^2 \beta_x = \beta_x - 2\beta_{x-1} + \beta_{x-2}$ ), e proporcionará naturalmente séries estimadas mais alisadas.

Consideremos de novo o modelo definido por (6.9.1)-(6.9.2)-(6.9.3). A introdução de uma penalização no procedimento de estimação dos coeficientes do modelo de Poisson conduz-nos, no espírito de (4.3.17), a uma função de log-verosimilhança penalizada  $\mathcal{L}_P(\alpha, \beta, \kappa)$  definida por

$$\mathcal{L}_P(\alpha, \beta, \kappa) = \mathcal{L}(\alpha, \beta, \kappa) - \frac{1}{2} \lambda_\alpha \alpha' D_\alpha' D_\alpha \alpha - \frac{1}{2} \lambda_\beta \beta' D_\beta' D_\beta \beta - \frac{1}{2} \lambda_\kappa \kappa' D_\kappa' D_\kappa \kappa \quad (6.9.18)$$

onde  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$  são parâmetros de penalização ou alisamento e  $\mathcal{L}(\alpha, \beta, \kappa)$  representa a função de log-verosimilhança.

Como se observa, o primeiro termo da função objectivo (6.9.18) mede o grau de ajustamento face aos valores observados, enquanto os seguintes correspondem a penalizações directamente proporcionais às diferenças finitas (de ordem  $d$ ) nos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  do modelo. A função objectivo estabelece, deste modo, um compromisso entre a qualidade do ajustamento (primeiro termo) e o alisamento das séries  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  (restantes termos).

A importância relativa dos dois objectivos é controlada pelos parâmetros de penalização  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$ . No limite, quando  $\lambda_i \rightarrow \infty$   $i \in (\alpha, \beta, \kappa)$  o procedimento produz um ajustamento linear. No extremo oposto, i.e., no caso em que todos os parâmetros são nulos  $\lambda_\alpha = \lambda_\beta = \lambda_\kappa = 0$ , o procedimento devolve as estimativas iniciais geradas pelo método de máxima verosimilhança.

Para estimar os parâmetros, recorreremos a uma adaptação do esquema iterativo geral (4.3.18) desenvolvido em Eilers e Marx (1996), Marx e Eilers (1998) e Durban *et al.* (2002). A adaptação envolve a reformulação do procedimento iterativo (6.9.11) segundo

$$\begin{aligned} (G_\alpha^{(\nu)} + B_\alpha) \hat{\alpha}^{(\nu+1)} &= (G_\alpha^{(\nu)} + B_\alpha - P_\alpha) \hat{\alpha}^{(\nu)} + L_\alpha^{(\nu)} \\ (G_\kappa^{(\nu)} + B_\kappa) \hat{\kappa}^{(\nu+1)} &= (G_\kappa^{(\nu)} + B_\kappa - P_\kappa) \hat{\kappa}^{(\nu)} + L_\kappa^{(\nu)} \\ (G_\beta^{(\nu)} + B_\beta) \hat{\beta}^{(\nu+1)} &= (G_\beta^{(\nu)} + B_\beta - P_\beta) \hat{\beta}^{(\nu)} + L_\beta^{(\nu)}, \end{aligned} \quad (6.9.19)$$

onde  $\mathbf{P}_i = \lambda_i \mathbf{D}'_i \mathbf{D}_i$ ,  $i \in (\alpha, \beta, \kappa)$ , e  $\mathbf{B}_i = \text{diag}(\mathbf{P}_i)$  denota uma matriz composta exclusivamente pelos elementos da diagonal principal de  $\mathbf{P}_i$ .

Para implementar o método resta-nos, para além de iniciar o algoritmo e definir o critério de óptimo, escolher o valor dos parâmetros de penalização  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$ . Conforme referimos no Capítulo 4, o valor dos parâmetros de penalização pode ser escolhido de forma subjectiva ou, em alternativa, resultar de um procedimento automático. Eilers e Marx (1996) e Durban *et al.* (2002) consideram diferentes critérios (AIC, BIC, validação cruzada, validação cruzada generalizado).

Tomando como exemplo o critério de validação cruzada definido em (4.8.1), consideremos como erro de previsão  $\epsilon_{r_{\mathcal{D}}}(x, t, \lambda_i)$ ,  $i \in (\alpha, \beta, \kappa)$ , a definição de resíduos de deviance (6.9.17) usada no contexto dos GLMs, suprimindo a observação correspondente à idade  $x$  e ano de calendário  $t$ . O critério de validação cruzada selecciona os parâmetros de penalização  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$  minimizando a seguinte soma de quadrados

$$(\hat{\lambda}_\alpha, \hat{\lambda}_\beta, \hat{\lambda}_\kappa) = \arg \min_{\lambda_\alpha, \lambda_\beta, \lambda_\kappa} \left\{ \mathcal{E}_{r_{\mathcal{D}}} = \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \epsilon_{r_{\mathcal{D}}}(x, t, \lambda_i)^2 \right\}, \quad i \in (\alpha, \beta, \kappa). \quad (6.9.20)$$

Em alternativa, podemos recorrer aos critérios BIC (*Bayesian Information Criterion*) ou AIC (*Akaike Information Criterion*), definidos por

$$\mathcal{D}(\mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}) + \delta \dim(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}), \quad (6.9.21)$$

onde  $\mathcal{D}(\mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\lambda})$  denota a deviance,  $\dim(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}) = \text{trace}(\mathbf{H})$  a dimensão efectiva do vector de parâmetros,  $\mathbf{H}$  a matriz de projecção (*hat matrix*) e  $\delta = \ln(\mathcal{N} \times \mathcal{T})$  ou  $\delta = 2$  no caso dos critérios BIC e AIC, respectivamente.

Por último, como as estimativas  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  geradas pelo procedimento de estimação não asseguram automaticamente o respeito pelas restrições (6.9.3), recorre-se de novo a uma reparametrização do modelo nos termos do definido na Secção 6.9.3.

### 6.9.5 Extensão de Renshaw-Haberman com duplo predictor log-bilinear

Na sequência da sugestão apresentada por Renshaw e Haberman (2003c,d), desenvolvemos nesta secção a versão do modelo de Poisson com duplo predictor log-bilinear. Partindo de (6.9.2), a introdução de um segundo termo log-bilinear dá origem a um modelo definido por (6.9.1) e pela equação base

$$\mu_{x,t} = \exp(\alpha_x + \beta_{x,1} \kappa_{t,1} + \beta_{x,2} \kappa_{t,2}). \quad (6.9.22)$$

As restrições de identificação do modelo são agora descritas por

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_{t,1} = \sum_{t=t_{\min}}^{t_{\max}} \kappa_{t,2} = 0 \quad (6.9.23)$$

$$\sum_{x=x_{\min}}^{x_{\max}} \beta_{x,1} = \sum_{x=x_{\min}}^{x_{\max}} \beta_{x,2} = 1. \quad (6.9.24)$$

As hipóteses do modelo conduzem-nos, neste caso, à seguinte função de log-verosimilhança

$$\begin{aligned} \mathcal{L}(\alpha, \beta, \kappa) = & \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \{d_{x,t} (\alpha_x + \beta_{x,1}\kappa_{t,1} + \beta_{x,2}\kappa_{t,2}) \\ & - E_{x,t} \exp(\alpha_x + \beta_{x,1}\kappa_{t,1} + \beta_{x,2}\kappa_{t,2})\} + c, \end{aligned} \quad (6.9.25)$$

onde  $c$  é uma constante e as demais variáveis têm o significado habitual.

Para estimar os parâmetros, é reformulado o procedimento iterativo desenvolvido por Goodman (1979) de modo a contemplar os parâmetros adicionais  $\beta_{x,2}$  e  $\kappa_{t,2}$

$$\hat{\alpha}_x^{(v+1)} = \hat{\alpha}_x^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}{\sum_{t=t_{\min}}^{t_{\max}} [E_{x,t} \exp(\hat{\alpha}_x^{(v)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]} \quad (6.9.26)$$

$$\hat{\kappa}_{t,1}^{(v+1)} = \hat{\kappa}_{t,1}^{(v)} - \frac{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_{x,1}^{(v)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}{\sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_{x,1}^{(v)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}$$

$$\hat{\beta}_{x,1}^{(v+1)} = \hat{\beta}_{x,1}^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_{t,1}^{(v+1)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}{\sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_{t,1}^{(v+1)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}$$

$$\hat{\kappa}_{t,2}^{(v+1)} = \hat{\kappa}_{t,2}^{(v)} - \frac{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_{x,2}^{(v)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v+1)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}{\sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_{x,2}^{(v)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v+1)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v)})]}$$

$$\hat{\beta}_{x,2}^{(v+1)} = \hat{\beta}_{x,2}^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_{t,2}^{(v+1)} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v+1)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v+1)})]}{\sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_{t,2}^{(v+1)})^2 [E_{x,t} \exp(\hat{\alpha}_x^{(v+1)} + \hat{\beta}_{x,1}^{(v+1)} \hat{\kappa}_{t,1}^{(v+1)} + \hat{\beta}_{x,2}^{(v)} \hat{\kappa}_{t,2}^{(v+1)})]}$$

A iniciação do algoritmo requer um vector de valores iniciais para os parâmetros  $\hat{\alpha}_x^{(0)}$ ,  $\hat{\beta}_{x,i}^{(0)}$  ( $\hat{\beta}_{x,i}^{(0)} \neq 0$ ) e  $\hat{\kappa}_{t,i}^{(0)}$  ( $i = 1, 2$ ). As estimativas  $\hat{\alpha}_x$ ,  $\hat{\beta}_{x,i}$  e  $\hat{\kappa}_{t,i}$  ( $i = 1, 2$ ) geradas pelo algoritmo de estimação (6.9.26) não garantem necessariamente a satisfação das restrições (6.9.23) e (6.9.24). Para tal, é necessário aplicar uma vez mais a seguinte reparametrização

$$\begin{aligned}\hat{\alpha}_x &\leftarrow \hat{\alpha}_x + \sum_{i=1}^2 \hat{\beta}_{x,i} \bar{\kappa}_i \\ \hat{\kappa}_{t,1} &\leftarrow (\hat{\kappa}_{t,i} - \bar{\kappa}_1) \hat{\beta}_{\xi,i}, \quad (i = 1, 2) \\ \hat{\beta}_{x,i} &\leftarrow \frac{\hat{\beta}_{x,i}}{\hat{\beta}_{\xi,i}}, \quad (i = 1, 2)\end{aligned}\tag{6.9.27}$$

onde

$$\bar{\kappa}_i = \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_{t,i} \quad \text{e} \quad \hat{\beta}_{\xi,i} = \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_{x,i} \quad (i = 1, 2).$$

O aumento do número de parâmetros do modelo proporciona, como veremos no Capítulo 7, uma melhoria da sua capacidade de ajustamento. Em contrapartida, o aumento da complexidade do modelo condiciona, em certa medida, a sua aplicação prática em problemas de previsão. Com efeito, os dois índices temporais  $\hat{\kappa}_{t,1}$  e  $\hat{\kappa}_{t,2}$  são, na prática, duas séries temporais altamente correlacionadas. Este facto torna a sua extrapolação para o futuro uma tarefa bastante mais complexa do que no caso em que existe uma única série, uma vez que são exigidos modelos multivariados.

Por outro lado, no capítulo seguinte teremos oportunidade de verificar que as estimativas dos parâmetros tendem a apresentar uma irregularidade acrescida, que dificulta inclusive a sua interpretação. Acresce ainda que neste caso as séries  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  do modelo clássico e  $\hat{\beta}_{x,1}$  e  $\hat{\kappa}_{t,1}$  do modelo extendido podem divergir significativamente.

### 6.9.6 Extensão do modelo com inclusão de “Tábua Limite”

Conforme referimos anteriormente, o comportamento assintótico das taxas de mortalidade projectadas pelos modelos de Poisson e de Lee-Carter pode, nalguns casos, revelar-se pouco satisfatório. Os estudos empíricos realizados com base nestes modelos (onde se inclui o relativo à população portuguesa, apresentado no capítulo seguinte) revelam uma trajectória claramente decrescente para o índice temporal  $\hat{\kappa}_t$ , acompanhada de estimativas de  $\beta_x$  positivas, um resultado antecipado num contexto marcado pelo decréscimo da mortalidade.

Neste sentido, a aplicação de métodos de séries temporais na extrapolação de  $\hat{\kappa}_t$  para

um horizonte infinito conduz-nos, invariavelmente, a taxas de mortalidade assintóticas nulas, um cenário improvável, pelo menos com base no conhecimento actual. De facto, nos casos em que  $\hat{\beta}_x > 0$  e  $\hat{\alpha}_x$  é finito verifica-se que

$$\lim_{\hat{\kappa}_t \rightarrow -\infty} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t) = 0$$

As preocupações quanto ao comportamento assintótico dos modelos de projecção motivaram, como vimos nas secções anteriores, o desenvolvimento de soluções que impõem, à partida, um valor arbitrário positivo para as taxas de longo prazo. No caso do modelo exponencial implícito na metodologia do CMIB (1990) e no caso da reformulação do método LC proposta por Renshaw e Haberman (2005), este valor corresponde ao parâmetro positivo  $\alpha(x)$ .

A abordagem usada nestes estudos situa-se no campo dos denominados modelos de projecção com tábua limite (ou tábua objectivo). Desenvolvidos inicialmente por Bourgeois-Pichat (1952), estes modelos admitem a existência de uma tábua “óptima” para a qual convergem os ganhos de longevidade verificados ao longo do tempo.<sup>76</sup> Dito de outro modo, estes modelos admitem, explicitamente, que existem limites naturais para a longevidade humana, dos quais decorrem níveis de mortalidade mínimos abaixo dos quais se assume ser impossível descer no futuro.

Inspirados neste princípio, formulamos nesta secção uma extensão do modelo de Poisson em que admitimos que as projecções de mortalidade são balizadas por uma determinada tábua limite, para a qual tendem os ganhos de longevidade de uma população. Designemos por  $\mu_x^{\text{lim}}$  e por  $q_x^{\text{lim}}$  as taxas instantâneas e os quocientes de mortalidade correspondentes a essa tábua limite. A determinação deste limite requer, naturalmente, a formulação de um conjunto de juízos mais ou menos subjectivos sobre as tendências previsíveis nas diferentes causas de morte, sobre os avanços na medicina ou sobre o sucesso no combate às epidemias.

Na prática, na eventualidade de este exercício prospectivo se revelar difícil, uma forma diferente de interpretar o modelo passa por considerar como tábua limite uma tábua que reflecta o comportamento da mortalidade numa população mais avançada em termos condições de desenvolvimento económico e social (população alvo), para a qual tenderá a convergir a experiência em estudo. Outra solução passa por adoptar uma qualquer função paramétrica sobre a qual seja possível equacionar diferentes cenários quanto às grandes tendências demográficas na longevidade humana (e.g., rectangularização e expansão da curva de sobrevivência), incorporando informação estatística e avaliações de natureza

---

<sup>76</sup>Na secção seguinte voltaremos a este tema com mais detalhe.

## 6.9. Modelo de Poisson

subjectiva.<sup>77</sup>

Em termos formais, a incorporação de uma tábua limite no modelo de Poisson exige a substituição da hipótese (6.9.1) por

$$D_{x,t} \sim \text{Poisson}(E_{x,t}(\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}})) \quad (6.9.28)$$

com

$$\mu_{x,t}^{\text{ad}} = \exp(\alpha_x + \beta_x \kappa_t) \quad (6.9.29)$$

e as habituais restrições de identificação

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0 \quad \text{e} \quad \sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1 \quad (6.9.30)$$

Como se observa, o modelo estipula que o número de óbitos esperados à idade  $x$  e ano  $t$  é determinado pela exposição ao risco  $E_{x,t}$  e por uma força de mortalidade que resulta da soma do valor limite  $\mu_x^{\text{lim}}$  com o valor adicional (contemporâneo)  $\mu_{x,t}^{\text{ad}}$ . Esta solução simples para incorporar o valor limite é inspirada nas formulações (6.5.3) e (6.8.16) adoptadas no âmbito do modelo exponencial e na modelação dos factores de redução (Renshaw e Haberman, 2005). Seja  $\lambda_{x,t}$  o número de óbitos esperados à idade  $x$  e ano  $t$ , com

$$\lambda_{x,t} = \mathbb{E}[D_{x,t}] = E_{x,t} \left( \mu_x^{\text{lim}} + \exp(\alpha_x + \beta_x \kappa_t) \right) \quad (6.9.31)$$

A função de log-verosimilhança correspondente é definida por

$$\begin{aligned} \mathcal{L}(\alpha, \beta, \kappa) &= \ln \left\{ \prod_{t=t_{\min}}^{t_{\max}} \prod_{x=x_{\min}}^{x_{\max}} \left( \frac{\lambda_{x,t}^{d_{x,t}} \exp(-\lambda_{x,t})}{(d_{x,t})!} \right) \right\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \{d_{x,t} \ln \lambda_{x,t} - \lambda_{x,t} - \ln[(d_{x,t})!]\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \left\{ d_{x,t} \ln E_{x,t} + d_{x,t} \ln \left( \mu_x^{\text{lim}} + \exp(\alpha_x + \beta_x \kappa_t) \right) \right. \\ &\quad \left. - E_{x,t} \left( \mu_x^{\text{lim}} + \exp(\alpha_x + \beta_x \kappa_t) \right) - \ln[(d_{x,t})!] \right\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \left\{ d_{x,t} \ln \left( \mu_x^{\text{lim}} + \exp(\alpha_x + \beta_x \kappa_t) \right) - E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) \right\} + c \end{aligned} \quad (6.9.32)$$

<sup>77</sup> Por exemplo, Duchene e Wunsch (1988, 1991) construíram uma tábua de mortalidade unisexo com base na lei de Weibull (2.5.47), expressando um conjunto de quocientes  $q_x$  mínimos que, segundo os autores, constituem os limites naturais para a vida humana.

onde  $c$  é uma constante e as demais variáveis mantêm o seu significado habitual.

Em face das limitações dos *packages* tradicionais, ensaiamos de novo uma solução baseada no algoritmo de Goodman (6.9.9) para estimar os parâmetros. Seja  $\mathcal{L}(\theta_j) = \mathcal{L}(\alpha_j, \beta_j, \kappa_j)$ ,  $j \in (x, t)$ , a função objectivo. Considere um par  $(x, t)$  qualquer. A primeira derivada parcial da função objectivo com respeito a  $\theta_j$  é dada por

$$\frac{d\mathcal{L}(\theta_j)}{d\theta_j} = \frac{d_{x,t} \left( \frac{d\mu_{x,t}^{\text{ad}}}{d\theta_j} \right)}{(\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}})} - E_{x,t} \left( \frac{d\mu_{x,t}^{\text{ad}}}{d\theta_j} \right), \quad (6.9.33)$$

onde

$$\frac{d\mu_{x,t}^{\text{ad}}}{d\alpha_x} = \mu_{x,t}^{\text{ad}}, \quad \frac{d\mu_{x,t}^{\text{ad}}}{d\beta_x} = \kappa_t \mu_{x,t}^{\text{ad}}, \quad \frac{d\mu_{x,t}^{\text{ad}}}{d\kappa_t} = \beta_x \mu_{x,t}^{\text{ad}}.$$

A segunda derivada parcial da função objectivo com respeito a  $\theta_j$  é dada por

$$\frac{d^2\mathcal{L}(\theta_j)}{d\theta_j^2} = d_{x,t} \left\{ \frac{\left( \frac{d^2\mu_{x,t}^{\text{ad}}}{d\theta_j^2} \right) (\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}}) - \left( \frac{d\mu_{x,t}^{\text{ad}}}{d\theta_j} \right)^2}{(\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}})^2} \right\} - E_{x,t} \left( \frac{d^2\mu_{x,t}^{\text{ad}}}{d\theta_j^2} \right), \quad (6.9.34)$$

onde

$$\frac{d^2\mu_{x,t}^{\text{ad}}}{d\alpha_x^2} = \mu_{x,t}^{\text{ad}}, \quad \frac{d^2\mu_{x,t}^{\text{ad}}}{d\beta_x^2} = (\kappa_t)^2 \mu_{x,t}^{\text{ad}}, \quad \frac{d^2\mu_{x,t}^{\text{ad}}}{d\kappa_t^2} = (\beta_x)^2 \mu_{x,t}^{\text{ad}}.$$

Assim, por exemplo, as expressões (6.9.33) e (6.9.34) para  $\alpha_x$  são dadas por

$$\frac{d\mathcal{L}(\theta_j)}{d\alpha_x} = \frac{d_{x,t}\mu_{x,t}^{\text{ad}}}{(\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}})} - E_{x,t}\mu_{x,t}^{\text{ad}} \quad (6.9.35)$$

e por

$$\frac{d^2\mathcal{L}(\theta_j)}{d\alpha_x^2} = d_{x,t} \left\{ \frac{\mu_x^{\text{lim}} \mu_{x,t}^{\text{ad}}}{(\mu_x^{\text{lim}} + \mu_{x,t}^{\text{ad}})^2} \right\} - E_{x,t}\mu_{x,t}^{\text{ad}}, \quad (6.9.36)$$

respectivamente.

Expressões análogas podem ser derivadas para os parâmetros  $\beta_x$  e  $\kappa_t$ .

Seja  $\hat{\mu}_{x,t}^{\text{ad}(\nu_\alpha, \nu_\beta, \nu_\kappa)} = \exp(\hat{\alpha}_x^{(\nu_\alpha)} + \hat{\beta}_x^{(\nu_\beta)} \hat{\kappa}_t^{(\nu_\kappa)})$ . Dado um vector de valores iniciais  $(\hat{\alpha}_x^{(0)}, \hat{\beta}_x^{(0)}, \hat{\kappa}_t^{(0)})$  e as expressões (6.9.35) e (6.9.36), a adaptação do algoritmo (6.9.9)

conduz-nos a estimativas de  $\alpha_x$  revistas segundo

$$\begin{aligned}\hat{\alpha}_x^{(\nu+1,1)} &= \sum_{t=t_{\min}}^{t_{\max}} \left( d_{x,t} \frac{\hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)}}{\mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)}} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)} \right) \\ \hat{\alpha}_x^{(\nu+1,2)} &= \sum_{t=t_{\min}}^{t_{\max}} \left( d_{x,t} \frac{\mu_x^{\text{lim}} \hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)}}{\left( \mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)} \right)^2} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu,\nu,\nu)} \right) \\ \hat{\alpha}_x^{(\nu+1)} &= \hat{\alpha}_x^{(\nu)} - \frac{\hat{\alpha}_x^{(\nu+1,1)}}{\hat{\alpha}_x^{(\nu+1,2)}}.\end{aligned}$$

Usando um raciocínio análogo, as estimativas de  $\kappa_t$  e  $\beta_x$  são revistas segundo

$$\begin{aligned}\hat{\kappa}_t^{(\nu+1,1)} &= \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(\nu)} \left( d_{x,t} \frac{\hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)}}{\mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)}} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)} \right) \\ \hat{\kappa}_t^{(\nu+1,2)} &= \sum_{x=x_{\min}}^{x_{\max}} \left( \hat{\beta}_x^{(\nu)} \right)^2 \left( d_{x,t} \frac{\mu_x^{\text{lim}} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)}}{\left( \mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)} \right)^2} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu)} \right) \\ \hat{\kappa}_t^{(\nu+1)} &= \hat{\kappa}_t^{(\nu)} - \frac{\hat{\kappa}_t^{(\nu+1,1)}}{\hat{\kappa}_t^{(\nu+1,2)}} \\ \hat{\beta}_x^{(\nu+1,1)} &= \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(\nu+1)} \left( d_{x,t} \frac{\hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)}}{\mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)}} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)} \right) \\ \hat{\beta}_x^{(\nu+1,2)} &= \sum_{t=t_{\min}}^{t_{\max}} \left( \hat{\kappa}_t^{(\nu+1)} \right)^2 \left( d_{x,t} \frac{\mu_x^{\text{lim}} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)}}{\left( \mu_x^{\text{lim}} + \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)} \right)^2} - E_{x,t} \hat{\mu}_{x,t}^{\text{ad}(\nu+1,\nu,\nu+1)} \right) \\ \hat{\beta}_x^{(\nu+1)} &= \hat{\beta}_x^{(\nu)} - \frac{\hat{\beta}_x^{(\nu+1,1)}}{\hat{\beta}_x^{(\nu+1,2)}}.\end{aligned}$$

Uma solução simples para determinar a condição de paragem (critério de óptimo) deste algoritmo envolve a fixação de um valor mínimo para a variação absoluta na função objectivo  $\mathcal{L}(\alpha, \beta, \kappa)$ . Formalmente, o algoritmo termina na iteração  $\nu$  quando

$$\text{abs} \left( \mathcal{L}^{(\nu)}(\theta_j) - \mathcal{L}^{(\nu-1)}(\theta_j) \right) < \varepsilon,$$

onde  $\varepsilon$  denota um patamar inferior pré-definido.

Por fim, as estimativas iniciais dos parâmetros geradas pelo algoritmo são reajustadas de modo a respeitar as restrições de identificação (6.9.30). Usando argumentos idênticos

aos apresentados na Secção 6.9.3, o ajustamento envolve

$$\begin{aligned}\hat{\alpha}_x &\leftarrow \hat{\alpha}_x + \hat{\beta}_x \bar{\kappa}, & \bar{\kappa} &= \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t \\ \hat{\kappa}_t &\leftarrow (\hat{\kappa}_t - \bar{\kappa}) \hat{\beta}_\xi, & \hat{\beta}_\xi &= \sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x \\ \hat{\beta}_x &\leftarrow \frac{\hat{\beta}_x}{\hat{\beta}_\xi}.\end{aligned}$$

A avaliação da qualidade do ajustamento é aferida da forma habitual, com base na estatística deviance e na análise dos resíduos do modelo.

### 6.9.7 Construção de intervalos de confiança

O modelo de Poisson não permite a obtenção de soluções analíticas para os intervalos de confiança das estimativas das variáveis de interesse. A principal justificação prende-se com o facto do modelo combinar duas fontes de incerteza, relativas aos erros amostrais nos parâmetros do modelo de Poisson e referentes aos erros de previsão nos parâmetros do modelo ARIMA. Acresce ainda que as medidas de interesse (e.g,  $\mu_{x,t}$ ,  $e_{x,t}$ ) são funções não-lineares complexas dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  e dos parâmetros do modelo ARIMA.

Para contornar estas limitações e permitir a construção de intervalos de confiança, Brouhns *et al.* (2002b, 2005) desenvolvem três abordagens alternativas, todas elas assentes em métodos de simulação. Uma primeira, de natureza paramétrica ou por simulação de Monte Carlo, que assume que a especificação log-bilinear de Poisson (6.9.2) é correcta. Uma segunda, de natureza semi-paramétrica (*parametric bootstrap*), onde se admite que a hipótese relativa à distribuição de Poisson (6.9.1) é adequada para gerar matrizes aleatórias para o número de óbitos ou de resíduos. Uma terceira, de natureza não-paramétrica, onde apenas se impõem alguns pressupostos relativamente à distribuição Multinomial.

#### Método de Simulação de Monte Carlo

A abordagem paramétrica ou por Simulação de Monte Carlo (SMC), descrita em detalhe por Brouhns *et al.* (2002b), parte do pressuposto de que o modelo de Poisson é adequado. Por outro lado, como os parâmetros do modelo são estimados pelo método de máxima verosimilhança, sabemos que eles obedecem aproximadamente a uma lei Normal multivariada, de média  $(\alpha, \beta, \kappa)$  e matriz de variâncias-covariâncias igual à inversa da matriz de informação de Fisher  $\mathcal{I}$ . Os elementos desta matriz correspondem ao (nega-

tivo) do valor esperado das segundas derivadas parciais da função de log-verosimilhança  $\mathcal{L}(\alpha, \beta, \kappa)$  com respeito aos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ .<sup>78</sup>

A segunda fonte de variabilidade tem origem, como referimos, no modelo de previsão. No caso de um modelo ARIMA, toda a informação necessária está contida nas estimativas dos parâmetros e da variância  $\sigma^2$  do ruído branco.

Designemos por  $\alpha_x^m$ ,  $\beta_x^m$  e  $\kappa_t^m$  os valores dos parâmetros referentes à  $m^{\text{ésima}}$  simulação,  $m = 1, \dots, M$ , para  $t = t_{\max}$  e por  $\kappa_t^m$ ,  $t \geq t_{\max} + 1$ , os valores projectados do índice temporal. A primeira fase do método de simulação de Monte Carlo desenrola-se nas seguintes quatro etapas:

1. São geradas simulações de  $\alpha_x^m$ ,  $\beta_x^m$  e  $\kappa_t^m$  a partir da distribuição Normal multivariada com médias iguais às estimativas de máxima verosimilhança dos parâmetros e respectiva matriz de variâncias-covariâncias.
2. É estimado o modelo ARIMA usando os valores  $\kappa_t^m$ ,  $t = t_{\min}, \dots, t_{\max}$ , como observações.
3. Projecta-se o valor de  $\kappa_t^m$  usando o modelo ARIMA estimado em 2. Os erros de previsão  $\hat{\epsilon}_t^m$  são simulados a partir da distribuição Normal univariada com média 0 e desvio-padrão  $\sigma_\epsilon^m$ , i.e.,  $\mathcal{N}or(0, \sigma_\epsilon^m)$ .
4. Calculam-se as medidas de interesse para cada simulação realizada.

A primeira etapa reflecte a incerteza associada às estimativas dos parâmetros do modelo de Poisson. A segunda etapa toma em consideração o facto da incerteza nos parâmetros do modelo ARIMA depender da incerteza em relação aos parâmetros do modelo de Poisson. Na terceira etapa considera-se que a incerteza nos valores projectados de  $\kappa_t$  depende, não apenas do desvio-padrão do modelo ARIMA, mas também da incerteza em relação aos próprios parâmetros do modelo ARIMA. Tal é assim porquanto a cada simulação  $m$  correspondem parâmetros distintos no modelo ARIMA. Por fim, na quarta etapa todas as fontes de incerteza são incorporadas no cálculo das medidas de interesse.

A segunda fase do método SMC envolve a aplicação da abordagem pelo percentil para construir intervalos de confiança para os indicadores de interesse. Consideremos as  $M$  estimativas  $\{\theta_j(m), m = 1, \dots, M\}$  da variável  $\theta_j$  gerada pelo método SMC. O intervalo de confiança a  $(1 - 2\alpha)\%$  para a variável  $\theta_j$  é definido por  $[\hat{\theta}_j^{(\alpha)}, \hat{\theta}_j^{(1-\alpha)}]$ , onde  $\hat{\theta}_j^{(\alpha)}$  denota o  $(100 \times \alpha)^{\text{ésimo}}$  percentil empírico de  $\theta_j$ , correspondente ao  $(M \times \alpha)^{\text{ésimo}}$  valor na lista

<sup>78</sup>Para mais detalhes sobre a matriz de informação de Fisher, veja-se a Secção 7.8.1 do Capítulo 7.

ordenada de valores simulados  $\{\theta_j(1), \theta_j(2), \dots, \theta_j(M)\}$ . O  $100 \times (1 - \alpha)^{\text{ésimo}}$  percentil empírico  $\hat{\theta}_j^{(1-\alpha)}$  é obtido de forma análoga.

### Método de *bootstrap* paramétrico

O método de Simulação de Monte Carlo admite que o modelo de Poisson é adequado. Pressupõe ainda que a dimensão da amostra é suficientemente grande para justificar o recurso à lei assintótica (Normal) dos estimadores de máxima verosimilhança dos parâmetros. Desenvolvido por Brouhns *et al.* (2005), o método de *bootstrap* paramétrico é uma abordagem de natureza semi-paramétrica que retém apenas a hipótese (6.9.1), mas não impõe a estrutura log-bilinear (6.9.2) para  $\mu_{x,t}$ .

O método é implementado, uma vez mais, em duas fases. A primeira desagrega-se em cinco etapas, nos seguintes termos:

1. Partindo do conjunto de observações  $(E_{x,t}; d_{x,t})$ , são construídas  $B$  amostras aleatórias  $(E_{x,t}; d_{x,t}^b)$ ,  $b = 1, \dots, B$ , onde os valores  $d_{x,t}^b$  são gerados por simulação a partir da distribuição de Poisson com média  $E_{x,t} \hat{\mu}_{x,t} = d_{x,t}$ , i.e.,

$$d_{x,t}^b \sim \text{Poisson}(d_{x,t}).$$

2. Para cada simulação bootstrap  $(E_{x,t}; d_{x,t}^b)$ , os parâmetros do modelo de Poisson são estimados através do método de máxima verosimilhança.
3. Em cada simulação, são estimados os parâmetros do modelo ARIMA( $p, d, q$ ) seleccionado com base nas observações iniciais e recorrendo à série  $K^b = \{\kappa_t^b, t = t_{\min}, \dots, t_{\max}, b = 1, \dots, B\}$ . Note-se que a estrutura do modelo ARIMA é preservada, sendo apenas reestimados os seus parâmetros.
4. São geradas estimativas  $\kappa_t^b$  com base nos parâmetros do modelo ARIMA estimado na etapa 3. Os termos residuais  $\epsilon_t^b$  são simulados a partir da lei normal  $\mathcal{N}or(0, \sigma^b)$ .
5. Calculam-se os indicadores de interesse para cada simulação.

A segunda fase do método de bootstrap paramétrico decorre nos termos do anteriormente descrito para o método de SMC.

### Método de bootstrap paramétrico com base nos resíduos

Uma variante do método de bootstrap paramétrico é aquela que contempla a construção de intervalos de confiança a partir de um procedimento de bootstrap aplicado directa-

mente sobre os resíduos do modelo e não sobre os óbitos. O pressuposto base do modelo é o de que os resíduos são independentes e identicamente distribuídos. Em termos práticos, o método é implementado em duas fases, sendo que a primeira pode desagregar-se nas seguintes etapas:

1. Partindo do conjunto de observações  $(E_{x,t}; d_{x,t})$ , são estimados os parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  do modelo de Poisson por intermédio do método de máxima verosimilhança.
2. É calculada a matriz dos resíduos de deviance  $r_{\mathcal{D}}$  segundo (6.9.17). Em alternativa, o método pode ser aplicado com base nos resíduos de Pearson  $r_{x,t}^{Pearson} = (d_{x,t} - \hat{d}_{x,t}) / \sqrt{\hat{d}_{x,t}}$ .
3. São geradas  $B$  réplicas  $\{r_{\mathcal{D}}^{(b)}, b = 1, \dots, B\}$ , mediante um procedimento de amostragem com reposição dos elementos da matriz  $r_{\mathcal{D}}$ . A equação (6.9.17) para os resíduos de deviance é invertida de modo a obter as correspondentes matrizes para o número estimado de óbitos  $\{\hat{d}_{x,t}^{(b)}, b = 1, \dots, B\}$ .
4. Aplica-se o método de máxima verosimilhança às matrizes  $(E_{x,t}; \hat{d}_{x,t}^{(b)})$ , obtendo-se um total de  $B$  estimativas dos parâmetros  $\{\hat{\alpha}_x^b, \hat{\beta}_x^b, \hat{\kappa}_t^b, b = 1, \dots, B\}$ .
5. Em cada simulação, são estimados os parâmetros do modelo ARIMA( $p, d, q$ ) selecionado com base nas observações iniciais e recorrendo à série  $K^b = \{\kappa_t^b, t = t_{\min}, \dots, t_{\max}, b = 1, \dots, B\}$ .
6. São geradas estimativas  $\hat{\kappa}_t^b$  com base nos parâmetros do modelo ARIMA estimado na etapa 5. Os termos residuais  $\epsilon_t^b$  são simulados a partir da lei normal  $\mathcal{N}or(0, \sigma^b)$ .
7. Calculam-se os indicadores de interesse.

A segunda fase do método desenrola-se nos mesmos moldes que os métodos anteriores.

### Método de bootstrap não paramétrico

A aplicação do método de bootstrap não paramétrico na construção de intervalos de confiança para os indicadores de mortalidade é devida a Brouhns *et al.* (2005). Esta terceira abordagem, mais geral, implica o desenho de um esquema multinomial envolvendo a redistribuição do número de óbitos observado nas diferentes idades de cada pseudo-geração, segundo as probabilidades estimadas com base nas estatísticas de mortalidade disponíveis.

Em termos mais precisos, para cada par  $(x, t)$  é seguida a pseudo-geração  $(E_{x,t}; d_{x,t})$ ,  $(E_{x+1,t+1}; d_{x+1,t+1})$ ,  $(E_{x+2,t+2}; d_{x+2,t+2})$ , ... Registe-se, no entanto, que estamos na realidade perante uma geração fictícia, uma vez que ela combina, como vimos no quadrado de Lexis da Figura 6.3.1, indivíduos nascidos nos anos  $t - x$  e  $t - x - 1$ . Posto isto, a primeira fase do método de bootstrap não paramétrico pode decompor-se nas seguintes etapas:

1. Para cada uma das gerações fictícias, são geradas simulações do pseudo-número de óbitos  $d_{x,t}^b$ ,  $d_{x+1,t+1}^b, \dots$  com o auxílio de uma distribuição Multinomial com expoente

$$d_{\bullet} = \sum_{k \geq 0} d_{x+k,t+k}$$

e parâmetros

$$\begin{aligned} \hat{q}_{x,t} &= \frac{d_{x,t}}{d_{\bullet}} \\ \hat{q}_{x+1,t+1} &= \frac{d_{x+1,t+1}}{d_{\bullet}} \\ &\vdots \end{aligned}$$

2. A aplicação do método prossegue então segundo os procedimentos descritos nas etapas 2 a 5 do método de bootstrap paramétrico, considerando naturalmente os novos valores  $\{d_{x,t}^b, b = 1, \dots, B\}$  simulados em 1.

Saliente-se que a relação

$$d_{\bullet} = \sum_{k \geq 0} d_{x+k,t+k}$$

é respeitada em cada simulação bootstrap. Por outras palavras, o procedimento de bootstrap preserva o número total de óbitos de cada geração fictícia. A segunda fase do método de bootstrap não paramétrico decorre nos mesmos moldes do referido anteriormente para os métodos de SMC e bootstrap paramétrico.

## 6.10 Modelo Binomial bilinear

Cossette *et al.* (2005) desenvolvem uma variante do método Lee-Carter com recurso à distribuição Binomial, onde sustentam que a modelação de  $D_{x,t}$  através da distribuição de Poisson pode ser considerada uma aproximação ao “verdadeiro” processo binomial

## 6.10. Modelo Binomial bilinear

que gera o número de óbitos. Nesta secção, descrevemos os principais resultados obtidos pelos autores.

Admitamos que  $D_{x,t}$  segue uma distribuição Binomial com parâmetros  $E_{x,t}$  e  $q_{x,t}$ . A especificação completa do modelo é dada por

$$D_{x,t} \sim \text{Bin}(E_{x,t}; q_{x,t}) \quad (6.10.1)$$

com

$$q_{x,t} = 1 - \exp(-\exp(\alpha_x + \beta_x \kappa_t)), \quad (6.10.2)$$

onde, na equação (6.10.2), fizemos uso da definição de  $\mu_{x,t}$  apresentada em (6.9.2). A identificação do modelo requer de novo o cumprimento das restrições (6.9.3).

Com base na hipótese de independência entre os óbitos, é deduzida a seguinte função de log-verosimilhança

$$\begin{aligned} \mathcal{L}(\alpha, \beta, \kappa) &= \ln \left\{ \prod_{t=t_{\min}}^{t_{\max}} \prod_{x=x_{\min}}^{x_{\max}} \binom{E_{x,t}}{d_{x,t}} q_{x,t}^{d_{x,t}} (1 - q_{x,t})^{E_{x,t} - d_{x,t}} \right\} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} d_{x,t} \ln q_{x,t} + (E_{x,t} - d_{x,t}) \ln (1 - q_{x,t}) + \ln \binom{E_{x,t}}{d_{x,t}} \\ &= \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} d_{x,t} \ln [(1 - \exp(-\exp(\alpha_x + \beta_x \kappa_t)))] \quad (6.10.3) \\ &\quad - (E_{x,t} - d_{x,t}) \exp(\alpha_x + \beta_x \kappa_t) + c, \end{aligned}$$

onde  $c$  é uma constante e as demais variáveis assumem o seu significado habitual.

À semelhança do que acontece no modelo de Poisson, a presença do termo bilinear  $\beta_x \kappa_t$  força a estimação do modelo através do esquema iterativo sugerido por Goodman (1979). Seja

$$\mu_{x,t}^{(v)} = \exp(\hat{\alpha}_x^{(v)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)})$$

e

$$q_{x,t}^{(v)} = 1 - \exp(-\mu_{x,t}^{(v)}) = 1 - p_{x,t}^{(v)}.$$

O procedimento iterativo usado para derivar as estimativas de máxima-verosimilhança

é definido por

$$\begin{aligned}
 \hat{\alpha}_x^{(v+1)} &= \hat{\alpha}_x^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} \mu_{x,t}^{(v)} \left( \frac{d_{x,t}}{q_{x,t}^{(v)}} - E_{x,t} \right)}{\sum_{t=t_{\min}}^{t_{\max}} \mu_{x,t}^{(v)} \left( d_{x,t} \frac{q_{x,t}^{(v)} + \mu_{x,t}^{(v)} p_{x,t}^{(v)}}{(q_{x,t}^{(v)})^2} - E_{x,t} \right)} \\
 \hat{\beta}_x^{(v+1)} &= \hat{\beta}_x^{(v)} - \frac{\sum_{t=t_{\min}}^{t_{\max}} \hat{\kappa}_t^{(v)} \mu_{x,t}^{(v)} \left( \frac{d_{x,t}}{q_{x,t}^{(v)}} - E_{x,t} \right)}{\sum_{t=t_{\min}}^{t_{\max}} (\hat{\kappa}_t^{(v)})^2 \mu_{x,t}^{(v)} \left( d_{x,t} \frac{q_{x,t}^{(v)} + \mu_{x,t}^{(v)} p_{x,t}^{(v)}}{(q_{x,t}^{(v)})^2} - E_{x,t} \right)} \\
 \hat{\kappa}_t^{(v+1)} &= \hat{\kappa}_t^{(v)} - \frac{\sum_{x=x_{\min}}^{x_{\max}} \hat{\beta}_x^{(v+1)} \mu_{x,t}^{(v)} \left( \frac{d_{x,t}}{q_{x,t}^{(v)}} - E_{x,t} \right)}{\sum_{x=x_{\min}}^{x_{\max}} (\hat{\beta}_x^{(v+1)})^2 \mu_{x,t}^{(v)} \left( d_{x,t} \frac{q_{x,t}^{(v)} + \mu_{x,t}^{(v)} p_{x,t}^{(v)}}{(q_{x,t}^{(v)})^2} - E_{x,t} \right)}.
 \end{aligned} \tag{6.10.4}$$

A iniciação do algoritmo e a definição do critério de óptimo são feitas nos mesmos termos dos métodos anteriores. Os resíduos de deviance apropriados para monitorar a qualidade de ajustamento do modelo são agora calculados segundo

$$r_{x,t}^D = \sqrt{2} \cdot \text{sign} \left( d_{x,t} - \hat{\lambda}_{x,t} \right) \left\{ d_{x,t} \ln \left( \frac{1 - \frac{d_{x,t}}{E_{x,t}}}{1 - \hat{q}_{x,t}} \right) + d_{x,t} \ln \left( \frac{d_{x,t}}{E_{x,t} \hat{q}_{x,t}} \right) \right\}^{\frac{1}{2}}. \tag{6.10.5}$$

Cossette *et al.* (2005) comparam o desempenho dos modelos baseados nas distribuições de Poisson e Binomial numa população de pensionistas e concluem que o modelo de Poisson apresenta, em termos relativos, um desempenho ligeiramente superior.

## 6.11 Modelos de projecção com “Tábua Limite”

Os modelos de projecção analisados nas secções anteriores baseiam-se na extrapolação das tendências observadas na mortalidade. Apesar de esta ser a opção mais natural, existem abordagens distintas (algumas delas já analisadas) que merecem ser analisadas.

Uma das questões que desde há muito preocupa actuários e demógrafos refere-se à duração da vida humana. Com efeito, e como vimos no Capítulo 2, no século XVIII Abraham de Moivre representou a função de sobrevivência por uma equação matemática em que a idade máxima  $\omega$  era fixada em 86 anos. No século XIX, as investigações de

Gompertz e Makeham incidiram igualmente sobre o tema, embora Gompertz (1825) fosse da opinião de que “*The limit to the possible duration of life is a subject not likely ever to be determined, even should it exist.*” Um número considerável de leis de mortalidade foi entretanto desenvolvido, considerando distintos valores para  $\omega$  (e.g., De Graaf (1729) e Lambert (1772) assumiam que  $\omega = 92$  e  $\omega = 96$ , respectivamente).

Uma abordagem alternativa consiste, como vimos na Secção 6.9.6, em construir tábuas de mortalidade limite, tomando por base os dados relativos a países com elevada esperança de vida. Neste contexto, o modelo de projecção com relação a uma tábua limite (ou tábua objectivo), trabalho pioneiro atribuído a Bourgeois-Pichat (1952), admite a existência de uma tábua “óptima” para a qual convergem os melhoramentos na mortalidade verificados ao longo do tempo. A escolha deste limite compreende a formulação de um conjunto de hipóteses, a consideração de uma experiência alvo e/ou a elaboração de cenários para a evolução de indicadores característicos.

Em termos formais, seja  $q_x^{\text{lim}}$  a probabilidade limite para a idade  $x$  e  $q_{x,t_0}$  o seu respectivo valor actual. Admitamos que a mortalidade projectada é expressa da seguinte forma

$$q_x(t) = \mathcal{F}_\theta \left( q_x^{\text{lim}}; q_{x,t_0} \right), \quad (6.11.1)$$

onde  $\mathcal{F}_\theta(\cdot)$  denota uma qualquer função interpoladora. Em particular, se considerarmos uma função do tipo exponencial, i.e.,

$$q_{x,t} = q_x^{\text{lim}} + \left[ q_{x,t_0} - q_x^{\text{lim}} \right] e^{-\lambda_x(t-t_0)} \quad (6.11.2)$$

com  $\lambda_x > 0$ , deduz-se que  $\lim_{t \rightarrow \infty} q_{x,t} = q_x^{\text{lim}}$ . Para além da aplicação já demonstrada na Secção 6.9.6 e dos exemplos aí mencionados, a projecção com tábua limite pode ainda ser adoptada (embora com menor fiabilidade) no contexto dos métodos paramétricos analisados na Secção 6.4. A adopção do método vertical de projecção significa que o problema envolve agora a interpolação entre o valor dos parâmetros  $\theta$  que caracterizam a situação actual e o respectivo valor limite  $\theta^{\text{lim}}$ :

## 6.12 Modelos de projecção com “Tábua Modelo”

Em alternativa à utilização de modelos paramétricos, é possível representar a evolução da mortalidade através de um conjunto de indicadores resumo ou “*markers*” (Heligman, 1984; Pollard, 1987; Tabeau, 2001). O primeiro conjunto de tábuas modelo foi elaborado pelas Nações Unidas em 1955, com o objectivo de servir de base à elaboração de pro-

jecções nos países em desenvolvimento (United Nations, 1955).<sup>79</sup> Em termos genéricos, o procedimento de projecção, ilustrado na Figura 6.12.1, decompõe-se nas seguintes etapas:

1. É escolhido um conjunto de tábuas modelo, que representem a mortalidade de uma determinada população em diferentes épocas e que, se antevê, possam igualmente espelhar o seu comportamento futuro.
2. A trajectória expressa em cada tábua é resumida por um conjunto de “*markers*”.
3. São analisadas as tendências registadas nesses indicadores e extrapolado o seu valor futuro, por exemplo com recurso a um modelo matemático.<sup>80</sup>
4. Projectados os “*markers*”, são incorporadas hipóteses subjectivas destinadas a distribuir a mortalidade pelas idades que compõem a vida humana.<sup>81</sup>

As principais limitações deste método dizem respeito à dificuldade em seleccionar as tábuas modelo e à necessidade de introduzir hipóteses sobre a distribuição da mortalidade por idade. Com efeito, indicadores como  $\dot{e}_0$  são medidas altamente não lineares da mortalidade registada em cada idade. De facto, as tendências históricas demonstram que quando a mortalidade decresce de um nível elevado, a maior contribuição para o aumento de  $\dot{e}_0$  provém da diminuição na mortalidade infantil e juvenil. Pelo contrário, quando a mortalidade decresce de um nível baixo, como aquele que se regista actualmente nos países industrializados, os ganhos na esperança de vida são atribuídos, sobretudo, à redução da mortalidade nas idades mais avançadas.<sup>82</sup>

<sup>79</sup>A sua construção assentou em 158 tábuas de mortalidade distintas para cada sexo, e recorreu a técnicas estatísticas que relacionam a mortalidade de uma idade (ou grupos de idades) com a mortalidade da idade precedente através de uma função quadrática. Entre as contribuições importantes neste domínio, destaque para as tábuas modelo regionais de Coale e Demeny (1966) e para o sistema de tábuas modelo de Lederman e Breas (1959), que recorre à análise factorial para identificar as variáveis que explicam a variação entre as 154 tábuas usadas no estudo. Realce ainda para o modelo relacional (*logit*) de Brass (1971).

<sup>80</sup>Por exemplo, o Banco Mundial projecta o indicador  $\dot{e}_x$  usando uma função logística da taxa de variação da mortalidade no tempo. Outro exemplo interessante pode ser encontrado em Buettner (2002), que desenvolve uma variante do método Lee-Carter destinada a projectar o comportamento da mortalidade na base de tábuas modelo indexadas pela esperança de vida à nascença. Dado que as tábuas modelo não contêm uma referência temporal, a reformulação do método Lee-Carter envolve a substituição do índice temporal  $t$  em (6.6.1) por um índice que reflecta o nível de esperança de vida. A equação base do modelo é dada por

$$\ln m_{x,e} = \alpha_x + \beta_x \kappa_e + \epsilon_{x,e},$$

onde o parâmetro  $\kappa_e$  representa a tendência no nível de  $\dot{e}_0$ .

<sup>81</sup>A calibragem de tábuas projectadas pode envolver, por exemplo, a distribuição da mortalidade observada numa população tipo.

<sup>82</sup>Keyfitz (1985) demonstra, por exemplo, que mesmo nos casos em que se admite um decréscimo constante e simétrico  $\delta$  nas taxas de mortalidade de todas as idades a esperança de vida aumenta a uma taxa decrescente, numa proporção  $H\delta$  onde  $H$  denota a entropia da tábua.

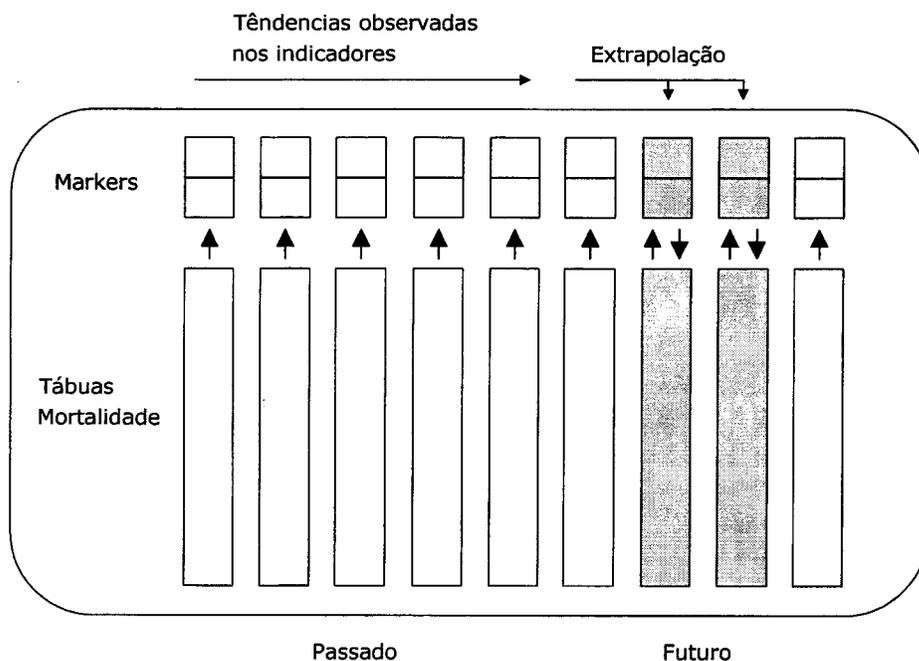


Figura 6.12.1: Projecção com Tábuas Modelo

## 6.13 Projecção via transformação de funções da tábua de mortalidade

Nas secções anteriores, enfatizámos que as tendências registadas na mortalidade podem ser analisadas com recurso a distintos indicadores:  $q_x$ ,  $\mu_x$ ,  $e_x$ ,  $q_x/p_x$ , entre outros. Uma solução alternativa para realizar projecções consiste em recorrer a transformações das funções base de uma tábua de mortalidade. Entre as múltiplas transformações possíveis, duas merecem destaque pela sua difusão e validade empírica: a transformação *logit* da função de sobrevivência, que dá corpo ao conhecido “método relacional”, e a transformação pela “função de resistência”.

### 6.13.1 Modelo relacional

Desenvolvido por Brass (1971, 1974), o “modelo relacional” presume que a relação entre a mortalidade de duas populações distintas pode ser devidamente representada por uma transformação *logit* da função de sobrevivência. Seja  $\Lambda_x$  a transformação *logit* de uma

tábua de mortalidade genérica  $\{q_x\}$

$$\Lambda_x = \frac{1}{2} \ln \left( \frac{{}_x q_0}{{}_x p_0} \right), \quad (6.13.1)$$

onde  $\frac{{}_x q_0}{{}_x p_0}$  representa o ganho máximo possível na função de sobrevivência à idade  $x$ , com  $S(x) = {}_x p_0$ . Note-se, em particular, que  $\lim_{x \rightarrow 0} \Lambda_x = -\infty$  e que  $\lim_{x \rightarrow +\infty} \Lambda_x = +\infty$ . Registe-se ainda que  $\Lambda_x = 0$  quando  ${}_x q_0 = {}_x p_0$ , ou seja, quando  ${}_x p_0 = \frac{1}{2}$ , o valor correspondente à mediana da vida aleatória residual, quando medida à nascença. O método relacional de projecção usa a função  $(1 - l_x)$ , com  $l_0 = 1$ , ou seja, considera a proporção dos falecidos à idade  $x$ . Substituindo em (6.13.1), obtemos

$$\Lambda_x = \frac{1}{2} \ln \left( \frac{1 - l_x}{l_x} \right). \quad (6.13.2)$$

Com base em estudos empíricos, Brass (1974) demonstra que a relação entre  $\Lambda_x$  e a correspondente transformação *logit* para uma população standard,  $\Lambda_x^{\text{stand}}$ , é aproximadamente linear<sup>83</sup>, i.e.,

$$\Lambda_x = \alpha + \beta \Lambda_x^{\text{stand}}, \quad (6.13.3)$$

onde os parâmetros  $\alpha$  e  $\beta$  são, nas palavras do autor, (quase) independentes da idade. O parâmetro  $\alpha$  denota o nível geral de mortalidade enquanto que o parâmetro  $\beta$  espelha a intensidade da sua relação com a mortalidade da população standard. A utilização do modelo relacional num contexto dinâmico exige, numa primeira fase, a obtenção de estimativas para os parâmetros  $\alpha_t$  e  $\beta_t$  em cada um dos anos que compõem a amostra. De seguida, são aplicados modelos ARIMA na extrapolação das tendências observadas em  $\alpha$  e  $\beta$ , visando a construção de tábuas projectadas.

O modelo relacional é particularmente interessante quando aplicado a dados geracionais, na medida em que os estudos empíricos realizados sugerem que os valores transformados pela função *logit* referentes a sucessivas gerações estão, aparentemente, linearmente relacionados (Pollard, 1987). Neste sentido, se designarmos por  $\Lambda_x(\tau)$  o *logit* da função de sobrevivência para a geração nascida no ano  $\tau$ ,  $l_x(\tau)$ , i.e.,

$$\Lambda_x(\tau) = \frac{1}{2} \ln \left( \frac{1 - l_x(\tau)}{l_x(\tau)} \right), \quad (6.13.4)$$

<sup>83</sup>Brass (1971) nota que a linearidade nem sempre existe e que, nalguns casos, ela pode ser obtida usando as diferenças entre os *logit* para diferentes momentos no tempo e a transformação *logit* da população standard, i.e.,

$$\Lambda_x - \Lambda_x^{\text{stand}} = \alpha + (\beta - 1) \Lambda_x^{\text{stand}}.$$

### 6.13. Projecção via transformação de funções da tábua de mortalidade

a relação entre a mortalidade de duas gerações nascidas em  $\tau_k$  e  $\tau_{k+1}$  é então definida por

$$\Lambda_x(\tau_{k+1}) = \alpha_k + \beta_k \Lambda_x(\tau_k). \quad (6.13.5)$$

As projecções de mortalidade envolvem, uma vez mais, a extrapolação das séries  $\{\alpha_k\}$  e  $\{\beta_k\}$ . Os valores projectados  $\{l_{x,t_{\max}+s}(\tau) : s > 0\}$  deduzem-se a partir da inversa da transformação logit<sup>84</sup>

$$l_x(\tau) = \frac{1}{1 + \exp[2\Lambda_x(\tau)]}. \quad (6.13.6)$$

Apesar da sua simplicidade e objectividade, o modelo relacional apresenta alguns problemas. Em particular, a transformação é efectivamente uma função da probabilidade de sobrevivência desde a nascença até uma dada idade futura  $x$  dependendo, por isso, de todas as taxas  $m_x$  para as idades anteriores a  $x$ . Significa isto que a função *logit* não é particularmente sensível à mortalidade registada numa idade particular, o que dificulta naturalmente a reconstituição adequada dos valores de  $q_x$ .

#### 6.13.2 Função de resistência

Petrioli e Berti (1975, 1979) sugerem uma transformação alternativa da função de sobrevivência  $S(x)$ , baseada na denominada “função de resistência”,  $r(x)$ , definida da seguinte forma

$$r(x) = \frac{\frac{S(x)}{\omega-x}}{\frac{1-S(x)}{x}}, \quad (6.13.7)$$

onde  $\omega$  denota a idade máxima na tábua. Como se observa, a transformação traduz, na prática, o rácio entre duas probabilidades médias anuais de óbito para um recém-nascido, respectivamente após a idade  $x$  (numerador) e antes da idade  $x$  (denominador).

<sup>84</sup>O modelo relacional pode ainda ser aplicado num contexto em que a mortalidade é representada por uma função paramétrica. Por exemplo, Petrioli (1975, 1996) admite que  $\mu_x$  segue a popular lei de Gompertz  $\mu_x = Bc^x$ , donde se deduz

$$\ln \mu_x = \ln B + x \ln c.$$

Considerando que a experiência de duas tábuas  $\{\mu_x^i\}_{i=1,2}$  é representada adequadamente pela lei de Gompertz, o modelo relacional é definido por

$$\ln(\mu_x^1) = \alpha + \beta \ln(\mu_x^2),$$

onde

$$\alpha = \frac{\ln B_1 - (\ln c_1 \ln B_2)}{\ln c_2} \quad \text{e} \quad \beta = \frac{\ln c_1}{\ln c_2}.$$

A função de resistência foi graduada usando a fórmula

$$r(x) = x^\alpha (\omega - x)^\beta e^{Ax^2+Bx+C} \quad (6.13.8)$$

ou, em particular, usando uma versão simplificada com apenas três parâmetros

$$r(x) = kx^\alpha (\omega - x)^\beta. \quad (6.13.9)$$

Num contexto dinâmico, as tendências na mortalidade são representadas assumindo que alguns dos parâmetros da função dependem do ano cronológico  $t$ . Assim, se considerarmos a equação (6.13.9), temos

$$r(x, t) = k(t) x^{\alpha(t)} (\omega - x)^{\beta(t)}$$

numa estrutura formal equiparável à dos modelos paramétricos analisados na Secção 6.4.<sup>85</sup> Registe-se, por fim, que a função de resistência pode ser incorporada num modelo relacional (Petrioli, 1996) com recurso a uma equação de ligação do tipo

$$\ln [r_1(x) \cdot \xi] = \alpha + \beta \ln [r_2(x) \cdot \xi], \quad (6.13.10)$$

onde  $\xi = (\omega - x) / x$ .

## 6.14 Métodos não paramétricos

### 6.14.1 P-Splines

Currie *et al.* (2003, 2004) demonstram que o método das P-Splines apresentado no Capítulo 4 pode ser utilizado para graduar e projectar tábuas num contexto bidimensional em que se consideram, simultaneamente, a idade actuarial e o tempo cronológico. Designemos por  $\mathbf{B}_a = \mathbf{B}(x_a)$ ,  $n_a \times c_a$ , a matriz de regressores contendo as B-Splines referentes à variável explicativa idade,  $x_a$ , onde  $n_a$  denota o número de idades consideradas e  $c_a$  representa o número de nós contemplado nesta dimensão. Denotemos por  $\mathbf{B}_t = \mathbf{B}(x_t)$ ,  $n_t \times c_t$ , a matriz de regressores contendo as B-Splines referentes à variável explicativa tempo cronológico,  $x_t$ , onde  $n_t$  define o número de anos incluídos na janela de observação e  $c_t$  simboliza o número de nós contemplado na dimensão temporal.

---

<sup>85</sup> Este modelo foi usado para projectar a mortalidade da população italiana e adoptado pela Associação Italiana de Actuários (ANIA) para construir tábuas projectadas para aplicação no mercado de rendas vitalícias.

A matriz de regressores para este modelo bidimensional é dada por

$$\mathbf{B} = \mathbf{B}_t \otimes \mathbf{B}_a, \quad (6.14.1)$$

onde  $\otimes$  simboliza o produto de Kronecker entre duas matrizes. Note-se que  $\mathbf{B}$  tem  $n_a n_t$  linhas e  $c_a c_t$  colunas. Associado a  $\mathbf{B}$  surge um vector de coeficientes da regressão  $\mathbf{a}$  de dimensão  $c_a c_t$ . Esses coeficientes são arrumados numa matriz  $\mathbf{A}$  com colunas e linhas dadas por  $\mathbf{A} = (a_1, \dots, a_{c_t})$  e  $\mathbf{A}' = (a_1^r, \dots, a_{c_a}^r)$ .

Seja  $\boldsymbol{\mu}$  o vector das  $\mu_{x,t}$  brutas, consideradas como realizações de variáveis Poisson com média  $\mu_{x,t} E_{x,t}$ . Quando especificado no contexto dos modelos GLM, o modelo é definido por

$$\log \boldsymbol{\mu} = \log \mathbf{E} + \mathbf{B} \mathbf{a}, \quad (6.14.2)$$

onde  $\log \mathbf{E}$  indica o tradicional *offset* do modelo.

Para evitar o problema de sobreparametrização do modelo os autores introduzem uma penalização sobre os coeficientes  $\mathbf{a}$ . A matriz de penalização  $\mathbf{P}$  é definida por

$$\mathbf{P} = \lambda_a \mathbf{I}_{c_a} \otimes \mathbf{D}'_a \mathbf{D}_a + \lambda_t \mathbf{D}'_t \mathbf{D}_t \otimes \mathbf{I}_{c_t}, \quad (6.14.3)$$

onde  $\mathbf{I}_{c_a}$  é uma matriz identidade de dimensão  $c_a$ ,  $\mathbf{D}_a$  é uma matriz de diferenças com dimensão  $(c_a - p_a) \times c_a$  onde  $p_a$  simboliza ordem da penalização sobre a idade.  $\mathbf{I}_{c_t}$  e  $\mathbf{D}_t$  têm interpretação análoga, com a diferença de que estão associados à penalização que incide sobre os coeficientes da dimensão temporal. Os parâmetros  $\lambda_a$  e  $\lambda_t$ , que determinam o nível de alisamento sobre a dimensão idade e tempo, respectivamente, são seleccionados através de um procedimento de optimização com respeito a um dado critério (e.g., BIC, AIC).

Os coeficientes da regressão são estimados maximizando a função de log-verosimilhança, com  $\mathbf{B}$  definida por (6.14.1), penalizada pela matriz  $\mathbf{P}$ , i.e.,

$$\mathcal{L}_p = \mathcal{L}(\mathbf{a}; \mathbf{y}) - \frac{1}{2} \mathbf{a}' \mathbf{P} \mathbf{a}. \quad (6.14.4)$$

A previsão no modelo assenta no princípio de que os valores futuros são desconhecidos, atribuindo valores arbitrários a  $d_{x,t}$  e  $E_{x,t}$ . A função de penalização possibilita então, de forma simultânea, o ajustamento aos dados observados e a projecção dos valores futuros, assegurando o alisamento da série projectada.

### 6.14.2 Métodos baseados em *smoothers* do tipo Kernel

Felipe *et al.* (2001) desenvolvem um método não-paramétrico de graduação com respeito à idade e ao tempo cronológico, onde combinam alguns princípios básicos de processos estocásticos com *smoothers* do tipo Kernel.<sup>86</sup> Em termos mais precisos, seja  $t$  o tempo cronológico,  $X(t)$  o processo seguido pela idade e  $n$  o número de indivíduos observados  $i = 1, \dots, n$ . Denotemos por  $N_i$  o processo de contagem do número de falhanços para o  $i^{\text{ésimo}}$  indivíduo no intervalo de tempo  $[t_1, t_2]$ . A intensidade aleatória do processo  $\lambda = (\lambda_1, \dots, \lambda_n)$  é definida por

$$\lambda_i(t) = \alpha \{X_i(t), t\} Y_i(t), \quad (6.14.5)$$

onde  $X_i(t)$  é um processo previsível unidimensional para a idade e  $\alpha(\cdot)$  é função sem forma funcional definida.  $Y_i(t)$  é um processo previsível que assume valores em  $\{0, 1\}$ , tomando o valor 1 quando o  $i^{\text{ésimo}}$  indivíduo está exposto ao risco e 0, caso contrário. Seja  $K$  uma função kernel unidimensional e  $b = (b_1, b_2)$  um vector bidimensional de *bandwidth*. O estimador de  $\alpha$  é dado por

$$\hat{\alpha}(x, t) = \frac{d_{x,t}^K}{E_{x,t}^K} \quad (6.14.6)$$

onde

$$\begin{aligned} d_{x,t}^K &= \sum_{i=1}^n \int_{t_1}^{t_2} K_{b_1} \{x - X_i(s)\} K_{b_2}(t - s) dN_i(s) \\ E_{x,t}^K &= \sum_{i=1}^n \int_{t_1}^{t_2} K_{b_1} \{x - X_i(s)\} K_{b_2}(t - s) Y_i(s) ds \end{aligned}$$

definem, respectivamente, o número de ocorrências e o número de expostos ao risco, devidamente suavizados. O estimador corresponde, assim, ao tradicional rácio entre o número de óbitos observados e o número de expostos ao risco. Para separar os efeitos idade e tempo cronológico, os autores analisam, em particular, o submodelo multiplicativo bidimensional, definido por

$$\alpha(x, t) = \alpha_1(x) \alpha_2(t).$$

---

<sup>86</sup>A metodologia é classificada, pelos próprios autores, como sendo uma técnica exploratória de dados, especialmente concebida para diagnosticar o comportamento da mortalidade nas dimensões etária e temporal, numa análise prévia à graduação mediante funções paramétricas.

O modelo pode ser interpretado como um ratio de mortalidade contínuo, cuja única hipótese se refere à interacção multiplicativa entre as duas componentes. Trata-se de uma hipótese algo restritiva, na medida em que pressupõe que a passagem do tempo influencia as taxas de mortalidade em todas as idades na mesma proporção. Os autores sustentam que a principal vantagem do método corresponde à separação das componentes idade e tempo cronológico. Reconhecem, todavia, as suas limitações, face às potencialidades oferecidas, por exemplo, pelos métodos paramétricos em termos de projecção, interpretação dos parâmetros e inferência estatística.

### 6.15 Métodos de projecção nas idades avançadas

Nos últimos dois séculos, a esperança de vida à nascença nos países desenvolvidos duplicou aproximadamente, atingindo os 72 e os 79 anos no caso de homens e mulheres, respectivamente (United Nations, 2002). Uma boa fatia deste acréscimo é atribuível à significativa redução alcançada na mortalidade infantil e juvenil, que assume hoje valores tão baixos que futuras diminuições pouco impacto terão no aumento da esperança de vida. Neste sentido, a manutenção da trajectória crescente da esperança de vida exigirá reduções importantes na mortalidade das idades mais avançadas.

O continuado aumento da esperança de vida têm implicações claras. Por um lado, um número crescente de pessoas sobrevive até às idades mais avançadas. Por outro, os óbitos registados nestas idades representam uma proporção cada vez maior do total de óbitos registados num ano. Durante muito tempo, a escassez e pouca fiabilidade da informação estatística sobre o número de óbitos nestas idades fez com que o perfil da mortalidade permanecesse relativamente desconhecido. Não admira, por isso, que no passado as estimativas resultassem, frequentemente, de extrapolações efectuadas com base nas idades mais jovens, usando modelos paramétricos fundados nas leis de Gompertz ou de Makeham.

Nos últimos anos, a melhoria da qualidade da informação estatística permitiu identificar um comportamento aparentemente inesperado na mortalidade das idades mais avançadas. Com efeito, os estudos empíricos realizados no campo da demografia (e.g., Horiuchi e Wilmoth (1998), Olshansky e Carnes (1997), Cheung e Robine (2007)) detetaram uma desaceleração na taxa de crescimento da mortalidade a partir de uma certa idade. Em particular, a curva  $x \mapsto m_x$  tende a apresentar, em escala logarítmica, uma forma côncava nas idades mais avançadas, formando uma espécie de “plateau”. Por outras palavras, o crescimento exponencial pressuposto, por exemplo, pela lei de Gompertz, não se confirma na realidade.

Esta incapacidade das leis de mortalidade clássicas para representar o perfil nas idades avançadas, com o progressivo afastamento face às hipóteses estipuladas pelas leis de Makeham ou Gompertz, motivou os investigadores na procura e desenvolvimento de modelos alternativos que captassem, de forma adequada, o comportamento do fenómeno. Nesta secção, apresentamos uma síntese dos principais contributos desenvolvidos nesta matéria, tendo em vista a selecção da melhor solução a adoptar na construção de tábuas prospectivas.<sup>87</sup>

### 6.15.1 Modelo de Heligman-Pollard

Uma das primeiras soluções propostas envolve a utilização de uma versão reduzida do modelo de Heligman e Pollard (1980), analisado em detalhe no Capítulo 3. O modelo incluía, na sua versão original, oito parâmetros e cobria todo o espectro de idades desde o nascimento até às idades mais avançadas. Quando aplicada apenas às idades mais avançadas, a segunda lei de HP (3.6.4) reduz-se a

$$q_x = \frac{GH^x}{1 + KGH^x}, \quad (6.15.1)$$

ou seja, a uma função com três parâmetros  $G$ ,  $H$  e  $K$ , cujo significado foi detalhado anteriormente. Na sua formulação original, o modelo destinava-se à graduação de tábuas contemporâneas. Como se observa, apesar da sua maior flexibilidade esta versão restrita do modelo de HP replica, em boa medida, o padrão de mortalidade implícito na lei de Gompertz.

### 6.15.2 Modelo Logit

O modelo relacional de Brass (1971), assente na parametrização da transformação *logit* de  $q_x$  por uma função linear, i.e.,

$$\text{logit}(q_x) = \ln\left(\frac{q_x}{1 - q_x}\right) = \alpha + \beta x \quad (6.15.2)$$

constitui uma alternativa simples para estimar a mortalidade nas idades mais avançadas.

As semelhanças entre o modelo *logit* e o modelo de Heligman-Pollard são evidentes. Com efeito, se aplicarmos a função inversa a (6.15.2) e resolvermos em ordem a  $q_x$ ,

---

<sup>87</sup> Para uma análise mais detalhada veja-se, por exemplo, Boleslawski e Tabeau (2001), Buettner (2002), Thatcher *et al.* (1998) e Pitacco (2004).

obtemos

$$q_x = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}} = \frac{e^{\alpha} e^{\beta x}}{1 + e^{\alpha} e^{\beta x}} = \frac{GH^x}{1 + GH^x} \quad (6.15.3)$$

onde  $G = e^{\alpha}$  e  $H = e^{\beta}$ .

### 6.15.3 Método de Coale-Kisker

Coale e Guo (1989) desenvolveram uma técnica de fecho de tábuas de mortalidade, que consiste na substituição das taxas de mortalidade observadas das idades mais avançadas (85 anos e seguintes) por uma sequência de valores extrapolados para os grupos de idades 85-89, 90-94, ..., 105-109 anos. Num estudo empírico sobre a experiência de seis países (Áustria, França, Japão, Noruega, Países-Baixos, Suécia), os autores detectaram uma diminuição relativamente constante na taxa de crescimento da mortalidade em intervalos de idades quinquenais a partir dos 80/85 anos.<sup>88</sup>

No ano seguinte, Coale e Kisker (1990) explanaram uma abordagem semelhante mas aplicada no fecho de tábuas completas. O método proposto por Coale e Guo (1989) e Coale e Kisker (1990) ficaria conhecido na literatura por método de Coale-Kisker. Em termos formais, designemos por  $k_x$  a taxa de variação da taxa de mortalidade  $m_x$

$$k_x = \ln \left( \frac{m_x}{m_{x-1}} \right) = \ln m_x - \ln m_{x-1}, \quad (6.15.4)$$

donde resulta

$$m_x = m_{x-1} \exp(k_x). \quad (6.15.5)$$

Se fixarmos, por exemplo,  $x = 85$  e aplicarmos recursivamente a equação (6.15.5), podemos escrever

$$m_x = m_{84} \exp \left( \sum_{y=85}^x k_y \right), \quad x = 85, \dots, x_{\max}, \quad (6.15.6)$$

onde  $x_{\max}$  representa a idade máxima (limite) considerada.<sup>89</sup>

Os autores assumem que  $k_x$  é linear acima de uma determinada idade, 85 anos neste

<sup>88</sup>O método proposto pelos autores foi aplicado no fecho das tábuas modelo de Coale-Demeny, tábuas abreviadas com idades agrupadas em intervalos quinquenais.

<sup>89</sup>Repare-se que no caso em que  $k_x$  é constante (i.e.,  $k_x = k$ ), a equação reduz-se ao clássico modelo de Gompertz

$$m_x = m_{84} \exp[(x - 84)k].$$

caso, ou seja

$$k_x = k_{85} + (x - 85) s, \quad x \in [85, x_{\max}]. \quad (6.15.7)$$

Para determinar a inclinação  $s$  da recta (6.15.7), Coale e Kisker (1990) impõem um valor pré-determinado para a taxa  $m_x$  correspondente à idade limite  $m_{x_{\max}}$ , que fixam em  $x_{\max} = 110$  anos a partir das observações efectuadas na população sueca

$$m_{110} = \begin{cases} 1.0, & \text{homens} \\ 0.8, & \text{mulheres.} \end{cases}$$

A fixação de valores diferentes para homens e mulheres é intencional e destina-se a evitar o cruzamento das  $m_x$  dos indivíduos de ambos os sexos nas idades mais avançadas. Se incorporarmos estas restrições em (6.15.6) obtemos

$$m_{110} = m_{84} \exp \left( \sum_{y=85}^{110} k_y \right) = m_{84} \exp \left[ \sum_{y=85}^{110} (k_{85} + (y - 85) s) \right],$$

donde se deriva a seguinte estimativa do parâmetro  $s$

$$s = \frac{\ln \left( \frac{m_{110}}{m_{84}} \right) - 26k_{85}}{325}. \quad (6.15.8)$$

Os valores de  $m_x$  relativos a cada idade são calculados mediante

$$m_x = m_{84} \exp \left\{ \sum_{y=85}^x [k_{85} + (y - 85) s] \right\}, \quad x = 85, \dots, x_{\max} \quad (6.15.9)$$

ou, se não desejarmos acumular os valores de  $k_x$  e  $s$ , através de

$$m_x = m_{x-1} \exp [k_{85} + (x - 85) s], \quad x = 85, \dots, x_{\max}. \quad (6.15.10)$$

O método parte da hipótese de que as taxas em torno da idade  $x = 85$  são fiáveis, garantindo que o parâmetro  $k_{85}$  pode ser calculado directamente a partir das estimativas brutas. Wilmoth (1995b) demonstra que o modelo de Coale-Kisker implica uma função exponencial de polinómios quadrática com respeito à idade, i.e.,

$$m_x = \exp (ax^2 + bx + c) \quad (6.15.11)$$

contrastando, assim, com a hipótese de Gompertz.

### 6.15.4 Método de Lindbergson

Lindbergson (2001) propõe uma técnica que consiste em modificar a fórmula de Makeham substituindo o crescimento exponencial por uma linha recta nas idades mais avançadas. Em termos formais, a força de mortalidade à idade  $x$  é definida por

$$\mu_x(a, b, c, d) = \begin{cases} a + be^{cx} & , x \leq \eta \\ a + be^{c\eta} + k(x - \eta) & , x > \eta, \end{cases} \quad (6.15.12)$$

onde  $a$ ,  $b$ ,  $c$  e  $k$  são parâmetros e  $\eta$  indica a idade a partir da qual a mortalidade tende a estabilizar. Dado um valor para  $\eta$ , os parâmetros podem ser estimados com base no método do mínimo Chi-Quadrado, i.e., pela minimização da seguinte equação

$$Q^2 = \sum_{x=x_{\min}}^{x_{\max}} \left\{ \frac{E_x}{\hat{\mu}_x} (\hat{\mu}_x - \mu_x(a, b, c, d))^2 \right\}, \quad (6.15.13)$$

onde  $\hat{\mu}_x$  e  $E_x$  assumem o seu significado habitual. Para determinar a idade de transição, é recomendada a repetição do procedimento de minimização (6.15.13) para diferentes valores de  $\eta$  no sentido de encontrar o valor que minimiza a estatística  $Q^2$ .<sup>90</sup> O autor testa o desempenho do modelo numa amostra relativa à população sueca e aos períodos 1988-1992 e 1993-1997, desagregada por sexo, e conclui que o ajustamento é satisfatório. O facto do modelo impor uma trajectória linear para  $\mu_x$  nas idades extremas contrasta com a evidência empírica que aponta, como vimos, para uma configuração côncava.

### 6.15.5 Método de Denuit e Goderniaux

O método desenvolvido por Denuit e Goderniaux (2005) trabalha com base nos quocientes de mortalidade  $q_x$  e prescreve uma condição de fecho para as tábuas. O método envolve o ajustamento, pelo método dos mínimos quadrados, do seguinte modelo log-quadrático

$$\ln \hat{q}_x = a + bx + cx^2 + \epsilon_x, \quad \epsilon_x \sim \mathcal{N}or(0, \sigma^2) \quad (6.15.14)$$

às observações relativas às idades mais elevadas, respeitando duas restrições

$$q_{x_{\max}} = 1 \quad (6.15.15)$$

$$q'_{x_{\max}} = 0, \quad (6.15.16)$$

<sup>90</sup> Assinale-se ainda que os cálculos são feitos considerando apenas as idades  $x$  a partir das quais se observam  $\mu_x$  crescentes.

onde  $q'_x$  designa a primeira derivada de  $q_x$  com respeito à idade  $x$ .

Se incorporarmos as restrições (6.15.15) e (6.15.16) em (6.15.14) demonstra-se, após alguma manipulação algébrica, que a equação base do modelo pode ser expressa segundo

$$\ln \hat{q}_x = (x_{\max}^2 - 2x(x_{\max}) + x^2) c + \epsilon_x, \quad \epsilon_x \sim \text{Nor}(0, \sigma^2). \quad (6.15.17)$$

As restrições (6.15.15) e (6.15.16) impõem à curva dos quocientes  $q_x$  uma configuração côncava nas idades elevadas e a existência de uma tangente horizontal no ponto  $x = x_{\max}$ . A restrição (6.15.16) visa impedir um eventual decréscimo dos quocientes nas idades avançadas.

Um dos aspectos críticos no modelo refere-se à determinação da idade a partir da qual as estimativas brutas  $\hat{q}_x$  são substituídas pelos valores ajustados provenientes do modelo (6.15.14). Em relação a este ponto, os autores recomendam a escolha da “idade de corte”,  $x_0$ , que maximiza o coeficiente de determinação da regressão  $R^2$ , repetindo o procedimento para diferentes valores de  $x_0$  no intervalo  $x \in [50, 85]$ . Alertam ainda para a provável necessidade de alisar a série  $\{\hat{q}_x\}$  em torno da idade de corte, recomendando para tal, por exemplo, a utilização de uma média geométrica dos  $\hat{q}_x$  nas idades  $x = x_0 - 5, \dots, x_0 + 5$ . Este procedimento evita o surgimento de descontinuidades no comportamento da série na vizinhança de  $x_0$ .

Os autores testam o desempenho empírico de diferentes métodos de fecho nos dados relativos às populações francesa, holandesa e sueca nos anos de 2002, 2001 e 1999, respectivamente, comparando os seus méritos relativos e analisando as consequências da sua utilização na tarifação de rendas vitalícias. Concluem que a escolha do método de fecho não é irrelevante e têm implicações no cálculo dos prémios.

### 6.15.6 Modelo logístico

O modelo logístico analisado no Capítulo 2 assenta numa curva em “S”, que cresce primeiro rapidamente para depois desacelerar essa progressão, apresentando um comportamento assintótico interessante na modelação das idades avançadas. Thatcher *et al.* (1998) testam o seu desempenho com base na seguinte formulação

$$\mu_x = \theta_1 + \frac{\theta_2 e^{\theta_3 x}}{1 + \theta_4 e^{\theta_3 x}}, \quad (6.15.18)$$

onde  $\theta_i$  ( $i = 1, \dots, 4$ ) são parâmetros a estimar.

No extremo inferior do intervalo de idades, as estimativas proporcionadas pelos modelos de Makeham e (6.15.18) não se afastam significativamente, na medida em que, para

valores de  $\theta_i$  reduzidos, o denominador do segundo termo em (6.15.18) se aproxima da unidade. Em contrapartida, no extremo superior os modelos divergem uma vez que, enquanto (6.15.18) nivela no patamar  $\theta_1 + \frac{\theta_2}{\theta_4}$ , a Lei de Makeham cresce sem limite. Recentemente, Bongaarts (2004) desenvolveu uma variante do modelo logístico, que apelidou de “*shifting logistic model*”, com base na seguinte função

$$\mu(x, t) = \frac{\alpha(t) e^{\beta x}}{1 + \alpha(t) e^{\beta x}} + \gamma(t), \quad (6.15.19)$$

onde  $\alpha(t)$  e  $\gamma(t)$  são parâmetros que dependem do tempo cronológico. O autor assume que o parâmetro  $\beta$  pode diferir entre populações e entre os sexos, mas é constante no tempo.

### 6.15.7 Modelo de Kannistö

O modelo de Kannistö (1992) constitui uma variante do modelo logístico em que a transformação *logit* de  $q_x$  é expressa em termos de uma função linear da idade. Formalmente, temos

$$q_x = 1 - \exp \left[ - \left( \frac{\theta_1 \exp(\theta_2 x)}{1 + \theta_1 (\exp(\theta_2 x) - 1)} \right) \right] \quad (6.15.20)$$

A equação (6.15.20) é um caso particular de (2.5.44) fixando  $\theta_1 = 0$  e  $\sigma^2 = \theta_3$ . A principal vantagem deste modelo prende-se com a sua facilidade de ajustamento, na medida em que apenas é necessário estimar dois parâmetros ( $\theta_1, \theta_2$ ).

### 6.15.8 Método das gerações extintas

Um dos principais problemas na análise da mortalidade nas idades avançadas refere-se à obtenção de estimativas de população. Atribuído a Vincent (1951), o método das gerações extintas constitui uma solução para derivar estimativas de população para as gerações sem membros sobreviventes no final do período de observação. Neste método, o efectivo populacional  $P_{x,t}$  à idade  $x$  no momento  $t$  é estimado somando todos os óbitos futuros desta geração, calculados da seguinte forma

$$P_{x,t} = \sum_{i=0}^{\infty} [d_{x+i,t+i,g}^U + d_{x+i+1,t+i,g}^L], \quad (6.15.21)$$

onde  $d_{\bullet}^L$  e  $d_{\bullet}^U$  denotam os óbitos registados nos triângulos inferior e superior, respectivamente.

O método assume que os óbitos têm dupla classificação (por idade e data de nascimento) e que não existem migrações internacionais após a idade  $x$  na geração  $g$  em análise, uma hipótese razoável apenas nas idades avançadas. Um ponto prévio na aplicação deste método prende-se com a determinação das gerações já extintas. Regra geral, assume-se que uma geração está extinta se alcançou a idade limite  $\omega$  (e.g., 120 anos) no final do período de observação.<sup>91</sup>

### 6.15.9 Método do rácio de sobrevivência

Desenvolvido por Thatcher *et al.* (2002), o método do rácio de sobrevivência (*survivor ratio method*) é usado igualmente para obter estimativas de população nas idades mais avançadas. O método das gerações extintas pressupõe que só após o falecimento de todos os membros de uma geração é possível reconstruir os números de sobreviventes nas idades anteriores, admitindo que nos confinamos a idades em que o fenómeno da migração é negligenciável.

Em alternativa, o método do rácio de sobrevivência não requer a extinção completa da geração, bastando para tal que os membros da geração atinjam uma dada idade (e.g. 100 anos). A justificação é a de que, nesta idade, se assume que apenas uma pequena proporção do efectivo original terá sobrevivido, ou seja, que o rácio entre o número de sobreviventes face aos elementos da geração que faleceram nos últimos  $k$  anos pode ser estimado com base na experiência das gerações precedentes.

Designemos por rácio de sobrevivência,  $RS_{x,t,g}$ , o rácio entre os sobreviventes da geração  $g$  com idade  $x$  em 1 de Janeiro do ano  $t$  e os sobreviventes da mesma geração  $k$  anos antes

$$RS_{x,t,g} = \frac{P_{x,t}}{P_{x-k,t-k}}. \quad (6.15.22)$$

Assumindo que não existe migração na geração no intervalo de tempo  $[(t-k), t]$ , este rácio pode ser definido por

$$RS_{x,t,g} = \frac{P_{x,t}}{P_{x,t} + D_{\bullet}}, \quad (6.15.23)$$

onde

$$D_{\bullet} = \sum_{i=1}^k [d_{x-i,t-i,g}^U + d_{x-i+1,t-i,g}^L].$$

Resolvendo (6.15.23) em ordem a  $P_{x,t}$ , obtemos

---

<sup>91</sup>Para mais detalhes sobre este tema veja Wilmoth *et al.* (2005).

$$P_{x,t} = \frac{RS_{x,t,g,k}}{(1 - RS_{x,t,g,k})} D_{\bullet}. \quad (6.15.24)$$

O rácio  $RS_{x,t,g}$  para a mais velha geração não extinta (com idade  $\omega - 1$  no momento de observação) é desconhecido, uma vez que não existe informação sobre a dimensão do efectivo  $P_{\omega-1,t}$ . No entanto, rácios comparáveis são conhecidos para todas as gerações precedentes, uma vez que é possível usar o método das gerações extintas para derivar estimativas de população. Admitamos que o rácio assume aproximadamente o mesmo valor na geração em análise face às  $m$  gerações precedentes, i.e.,

$$RS_{x,t,g,k} = \frac{P_{x,t}}{P_{x-k,t-k}} \approx \frac{P_{x,t-1}}{P_{x-k,t-k-1}} \approx \dots \approx \frac{P_{x,t-m}}{P_{x-k,t-k-m}}.$$

Uma forma de estimar  $RS_{x,t,g,k}$  consiste em calcular o rácio agregado para as  $m$  gerações precedentes, i.e.,

$$RS_{x,t,g,k}^* = \frac{\sum_{i=1}^m P_{x,t-i}}{\sum_{i=1}^m P_{x-k,t-k-i}}. \quad (6.15.25)$$

Conhecendo os valores de  $RS_{x,t,g,k}^*$  e  $D_{\bullet}$ , é possível estimar  $P_{x,t}$  mediante

$$\tilde{P}_{x,t} = \frac{RS_{x,t,g,k}^*}{(1 - RS_{x,t,g,k}^*)} D_{\bullet}. \quad (6.15.26)$$

Aplicando este método recursivamente desde a idade  $\omega - 1$  até à idade desejada (e.g., 90 anos), é possível derivar estimativas de população para as idades correspondentes.

Estudos empíricos realizados por Kannistö (1996) demonstraram, no entanto, que quando as taxas de mortalidade apresentam uma trajectória descendente, os rácios de sobrevivência não são constantes entre gerações, razão pela qual o autor propõe a introdução de um factor de correcção  $\alpha$  e a substituição da estimativa (6.15.26) por

$$\hat{P}_{x,t} = \alpha \frac{RS_{x,t,g,k}^*}{(1 - RS_{x,t,g,k}^*)} D_{\bullet}, \quad (6.15.27)$$

onde  $\alpha$  é um parâmetro a estimar. Quando a mortalidade é decrescente/crescente/constante, o valor de  $\alpha$  deve ser maior/menor/igual à unidade.<sup>92</sup>

<sup>92</sup>O autor defende ainda que o procedimento iterativo de reconstituição deve ser calibrado de modo a que as estimativas de população obtidas pelo método coincidam com os valores apurados pelos organismos oficiais para a idade máxima disponível (e.g., 85 anos em Portugal). O coeficiente  $\alpha$  desempenha, nesta matéria, um papel fundamental.

## Capítulo 7

# Tábuas de Mortalidade Prospectivas para a População Portuguesa

*“And in the end, it’s not the years in your life that count. It’s the life in your years.” (Abraham Lincoln)*

### 7.1 Introdução

Para responder aos desafios criados pelo aumento da longevidade humana, os actuários devem caminhar no sentido da utilização de tábuas de mortalidade prospectivas, que lhes permitam calcular os prémios e as reservas financeiras relativas a rendas vitalícias e outros produtos com benefícios de longo prazo em caso de vida, quer no contexto da actividade privada, quer no domínio dos sistemas públicos de segurança e protecção social. A construção de tábuas prospectivas envolve, como vimos no capítulo anterior, a adopção de métodos estatísticos que projectam as tendências históricas registadas na mortalidade para o futuro. Estes métodos permitem igualmente a derivação de medidas da incerteza subjacente às projecções, que se consubstanciam na construção de intervalos de confiança para as estimativas pontuais.

O objectivo deste capítulo é o de construir tábuas de mortalidade prospectivas para a população portuguesa visando, quer o estabelecimento de bases técnicas adequadas a uma correcta quantificação e controlo do risco de longevidade por parte das companhias de seguros, quer a avaliação da sustentabilidade financeira dos sistemas de segurança social. Os resultados obtidos permitem ainda analisar e projectar a evolução de indicadores

## 7.1. Introdução

---

como a esperança de vida e constituem, em conjunto com as componentes fecundidade e migração, um elemento chave na elaboração de projecções demográficas de longo prazo.

O Capítulo está organizado da seguinte forma. No que resta desta Secção, é discutida a importância das bases técnicas na actividade seguradora e a necessidade de tábuas prospectivas. Na Secção 7.2, são examinadas as grandes tendências de evolução da mortalidade em Portugal, considerando uma abordagem transversal assente em tábuas contemporâneas produzidas pelo INE. Esta análise transversal não tem naturalmente em conta o percurso longitudinal dos indivíduos no diagrama de Lexis, pelo que os indicadores analisados se referem a gerações fictícias. Nas Secções 7.3 e 7.4, testamos a aplicação do modelo de Poisson (nas suas múltiplas variantes) e do método Lee-Carter na obtenção de projecções de mortalidade para a população portuguesa.

Na Secção 7.5, investigamos a aplicação de métodos de séries temporais na projecção das tendências temporais. Na Secção 7.6, examinamos a utilização de métodos de projecção da mortalidade nas idades avançadas e de fecho de tábuas. Na Secção 7.7, analisamos a evolução da longevidade numa óptica transversal e numa óptica longitudinal ou prospectiva. Na Secção 7.8 procuramos quantificar a incerteza em torno das quantidades projectadas e o risco de longevidade. Para tal, são testadas duas metodologias alternativas de construção de intervalos de confiança empíricos, a saber o método de Simulação de Monte Carlo e o método *parametric bootstrap*. Por fim, na Secção 7.9 são resumidas as principais conclusões dos estudos efectuados.

### 7.1.1 Bases técnicas

A evolução da longevidade humana tem consequências óbvias no *pricing* e no cálculo das responsabilidades financeiras das empresas com produtos do ramo vida com benefícios em caso de vida, em particular com rendas vitalícias e pensões. A determinação do prémio puro de uma renda vitalícia assenta, à semelhança do que acontece noutros produtos do ramo vida, no princípio da equivalência. Segundo este princípio, o valor actual esperado das prestações devidas pela seguradora deve ser igual ao valor actual esperado dos prémios puros pagos pela pessoa segura. Ao prémio puro assim obtido, é habitual as companhias adicionarem um conjunto de cargas (margem de segurança, comissão de gestão, comissão de emissão, etc.), obtendo em resultado o denominado prémio comercial.

As bases técnicas de um contrato de seguro do ramo vida clássico compõem-se: (i) de uma taxa de juro técnica, usada para actualizar os fluxos financeiros futuros gerados pelo contrato, que é normalmente fixada obedecendo a regras de prudência sendo, nalguns

## 7.1. Introdução

---

casos, garantida durante todo o contrato; (ii) de uma lei de sobrevivência (descrita por uma tábua de mortalidade), e (iii) dos montantes (ou percentagens) das diferentes cargas adicionais que agravam o prémio puro.

Em Portugal, a legislação em vigor<sup>1</sup> estabelece que a taxa de juro técnica a usar no cálculo de prémios e provisões matemáticas do ramo vida deve ser escolhida de forma prudente, tendo em conta a natureza e a maturidade dos compromissos assumidos, bem como os activos em que a companhia de seguros se propõe investir os valores correspondentes às provisões.<sup>2</sup> Já no que se refere à componente demográfica, as entidades de supervisão recomendam a adopção de tábuas de mortalidade suficientemente prudentes e adequadas às classes de risco envolvidas, considerando margens de segurança razoáveis para fazer face a variações inesperadas desfavoráveis nos diferentes factores de risco.

Conforme referimos no Capítulo 5, não existem em Portugal tábuas de mortalidade regulamentares homologadas pelas entidades de supervisão, nem para a população portuguesa, nem para as subpopulações de pessoas seguras. Em consequência, as companhias a operar no mercado socorrem-se de tábuas adoptadas noutros países, seleccionando mormente aquelas que melhor aproximam o comportamento da mortalidade na sua carteira. Apesar de esta prática se enquadrar nas normas orientadoras da actividade seguradora, não deixa de constituir uma solução de recurso que envolve riscos de base substanciais e que deve, por isso, merecer uma atenção prioritária de companhias e entidades reguladoras no sentido de se colmatar esta lacuna. Neste sentido, centramos neste capítulo no desenvolvimento de tábuas de mortalidade prospectivas para a população portuguesa analisando, em particular, as suas implicações sobre a tarificação no mercado de rendas vitalícias.

### 7.1.2 Necessidade de tábuas prospectivas

Tradicionalmente, os actuários baseiam o cálculo dos prémios de seguro em tábuas contemporâneas ou de momento. Essas tábuas são, como vimos nos capítulos anteriores,

---

<sup>1</sup> Vidé o artigo 82 do Decreto-Lei 94-B/98 de 17 de Abril.

<sup>2</sup> Embora com carácter meramente prudencial, estabeleceram-se igualmente regras para a determinação da taxa de juro técnica a utilizar. Assim, para os contratos de taxa garantida, o Instituto de Seguros de Portugal (ISP) fixou por via da norma regulamentar n.º 19/94 uma taxa de juro técnica máxima de 4%, podendo ascender a 6% no caso dos contratos de rendas vitalícias. Não se encontram abrangidos por esta regra os contratos de seguros e operações ligados a fundos de investimento, contratos de prémio único com uma duração máxima de 8 anos e contratos sem participação de resultados. Em investimentos autónomos, recomenda-se que a taxa de juro técnica seja determinada em função da rentabilidade da aplicação, desde que cumpridas as margens e os requisitos estabelecidos pelo ISP. Caso não sejam cumpridos estes requisitos, poderá haver lugar à constituição de uma provisão adicional, designada por provisão para compromissos de taxa. Por fim, a taxa de juro utilizada nunca poderá ser superior ao rendimento dos activos.

## 7.2. Evolução da mortalidade em Portugal

---

construídas a partir de observações relativas a um período determinado (que vai, tipicamente, de 1 ano a um máximo de 5 anos). Esta forma de proceder afigura-se claramente desajustada nos casos em que a mortalidade evolui no tempo. Com efeito, suponhamos que o actuário utiliza as probabilidades anuais de morte (ou quocientes de mortalidade) obtidas a partir de um ano de observações da sua carteira. Nesse caso, os segurados de idade  $x$  servirão para estimar a probabilidade  $q_x$  de morte nesse ano para um indivíduo que alcança essa idade. As diferentes estimativas  $\hat{q}_x$  de  $q_x$  servirão para calcular os prémios dos diferentes produtos comercializados pela empresa.

Esta abordagem simples e pragmática sofre, no entanto, de uma debilidade fundamental. Com efeito, os prémios assim calculados não reflectirão o risco real de nenhuma das pessoas seguras da carteira, salvo na hipótese, pouco provável, de que a mortalidade permanece estática no tempo. De facto, na Secção 6.3 provámos que na realidade cada uma das  $\hat{q}_x$  é estimada com base em, pelo menos, uma geração diferente. Por outras palavras, se o objectivo for o de calcular o prémio a pagar por um indivíduo com  $x_0$  anos de idade no momento da subscrição do contrato,  $\hat{q}_{x_0}$  será estimado a partir dos indivíduos com essa idade na carteira (fazendo, por isso, parte da mesma geração que o segurado),  $\hat{q}_{x_0+1}$  a partir daqueles com idade  $x_0 + 1$  (ou seja, precedentes da geração imediatamente anterior à do novo segurado),  $\hat{q}_{x_0+2}$  com base nos indivíduos com idade  $x_0 + 2$ , e assim sucessivamente. Num cenário de diminuição da mortalidade, resulta evidente que as probabilidades adequadas a aplicar no cálculo do prémio serão inferiores a  $\hat{q}_{x_0+1}$  à idade  $x_0 + 1$ , inferiores a  $\hat{q}_{x_0+2}$  à idade  $x_0 + 2$ , e assim sucessivamente.

Com base neste raciocínio, conclui-se que, em contextos de aumento da longevidade, o *pricing* com base em tábuas contemporâneas conduz a uma subestimação sistemática dos prémios relativos a operações em caso de vida. Se é verdade que esta subestimação pode, pelo menos parcialmente, ser compensada pela aplicação de cargas adicionais sobre o prémio puro, num contexto em que a longevidade é dinâmica e os mercados são competitivos a adopção de bases técnicas adequadas passa, inevitavelmente, pela utilização de tábuas de mortalidade prospectivas.

## 7.2 Evolução da mortalidade em Portugal

Nesta secção, procedemos a uma caracterização da evolução da mortalidade em Portugal desde 1930 até ao presente, usando para o efeito as tábuas de mortalidade contemporâneas produzidas pelo INE e por Mexia e Corte-Real (1995).

## 7.2.1 Rectangularização da função de sobrevivência

A análise da distribuição do número de óbitos nas diferentes idades (i.e., do gráfico da função  $x \mapsto d_x/l_0$ ), representada na Figura 7.2.1, permite-nos constatar claramente uma crescente concentração da mortalidade em torno da moda da distribuição, localizada nas idades adultas. Em consequência, observa-se uma progressiva rectangularização da função de sobrevivência, sobretudo nos indivíduos do sexo feminino, um fenómeno bem patente nas Figuras 7.2.2 e 7.2.3 onde são representados os gráficos da função  $x \mapsto_x p_0 = l_x/l_0$ . Nestas figuras é representada, no eixo das ordenadas, a percentagem dos membros da população que alcançam a idade  $x$  indicada em abcissa.

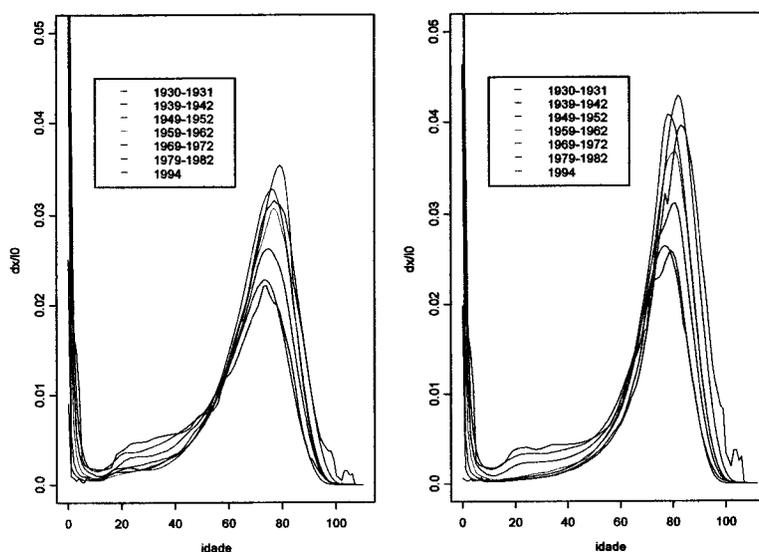


Figura 7.2.1: Densidade  $x \mapsto d_x/l_0$  (homens à esquerda e mulheres à direita)

Explicuemos em detalhe o fenómeno da rectangularização. O número de sobreviventes de uma geração decresce normalmente em função da idade, e desde os primeiros anos. A queda continuada dos quocientes  $q_x$  registada em Portugal, em quase todas as idades, tem como resultado que uma percentagem muito significativa dos elementos de cada geração sobrevive até às idades avançadas, por volta dos 70-75 anos. Esta evolução na mortalidade faz com que a função de sobrevivência se desloque verticalmente e se aproxime cada vez mais da forma rectangular, como se constata nas Figuras 7.2.2 e 7.2.3.

## 7.2. Evolução da mortalidade em Portugal

---

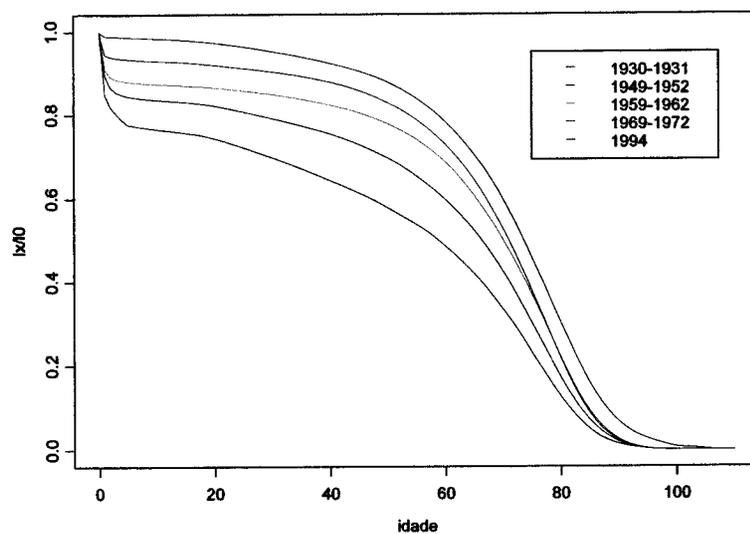


Figura 7.2.2: Função de sobrevivência  $x \mapsto l_x/l_0$ , Homens

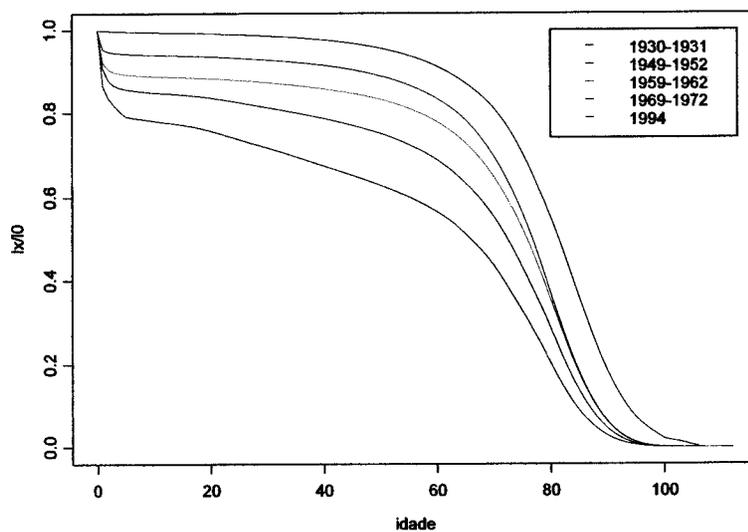


Figura 7.2.3: Função de sobrevivência  $x \mapsto l_x/l_0$ , Mulheres

## 7.2. Evolução da mortalidade em Portugal

Se observarmos atentamente a Figura 7.2.1, verificamos igualmente que a moda da distribuição tende a deslocar-se em direcção às idades mais elevadas (em torno dos 80 anos), naquilo que é conhecido na literatura como o fenómeno de expansão da função de sobrevivência.

No Quadro 7.2.1 verificamos que os fenómenos de rectangularização e de expansão registados em Portugal no período em análise têm conduzido, em ambos os sexos, a um significativo aumento da esperança de vida à nascença e a uma diminuição da variância (e do coeficiente de variação associado) da duração da vida humana ao longo do tempo. Como se observa no quadro, o coeficiente de variação diminuiu significativamente desde os anos 30 até ao presente passando, no caso dos homens, de um valor de 0.66 em 1930-1931 para um valor de 0.24 em 2002-2003, enquanto que no caso das mulheres se registou uma descida de um valor de 0.63 para 0.20 no início do século XXI.

| Período   | Esperança de vida<br>$e_0 = E [T]$ |          | Variância<br>$Var [T]$ |          | Coef. de variação<br>$\sqrt{Var [T]}/E [T]$ |          |
|-----------|------------------------------------|----------|------------------------|----------|---|----------|
|           | Homens                             | Mulheres | Homens                 | Mulheres | Homens                                      | Mulheres |
| 1930-1931 | 48.08                              | 52.12    | 994.81                 | 1066.40  | 0.66  | 0.63     |
| 1939-1942 | 48.58                              | 52.82    | 1005.11                | 1047.13  | 0.65  | 0.61     |
| 1949-1952 | 55.52                              | 60.50    | 880.57                 | 922.14   | 0.53  | 0.50     |
| 1959-1962 | 60.73                              | 66.35    | 778.36                 | 771.66   | 0.46  | 0.42     |
| 1969-1972 | 64.25                              | 70.48    | 578.83                 | 541.32   | 0.37  | 0.33     |
| 1979-1982 | 68.44                              | 75.35    | 459.48                 | 390.45   | 0.31  | 0.26     |
| 1990-1991 | 70.60                              | 77.52    | 393.60                 | 306.43   | 0.28  | 0.23     |
| 2002-2003 | 73.93                              | 80.23    | 327.03                 | 252.73   | 0.24  | 0.20     |

Quadro 7.2.1: Esperança de vida à nascença, variância e coeficiente de variação das tábuas de mortalidade da população portuguesa (I.N.E.) no período de 1930-31 a 2002-03

A análise do Quadro 7.2.1 permite-nos ainda concluir que o fenómeno de rectangularização da função de sobrevivência é ligeiramente mais acentuado nos indivíduos do sexo feminino do que nos indivíduos do sexo masculino.

Para quantificarmos o fenómeno da rectangularização, calculámos de seguida a denominada *entropia* da função de sobrevivência ou da tábua de mortalidade, definida como vimos no Capítulo 6 por

$$H = - \frac{\int_0^{\infty} l_x \ln(l_x) dx}{\int_0^{\infty} l_x dx},$$

onde  $l_x$  representa agora a função de sobrevivência (para  $l_0 = 1$ ).

## 7.2. Evolução da mortalidade em Portugal

Uma reformulação simples desta expressão conduz-nos a uma interpretação mais intuitiva do conceito de entropia.<sup>3</sup> Com efeito, se recordarmos que

$$\mu_x = -\frac{1}{l_x} \frac{d}{dx} l_x = -\frac{d}{dx} \ln(l_x)$$

obtemos, após alguma manipulação

$$\begin{aligned} H &= \frac{1}{l_0 \dot{e}_0} \int_0^\infty l_x \left( \int_0^x \mu_\xi d\xi \right) dx \\ &= \frac{1}{l_0 \dot{e}_0} \int_0^\infty \mu_\xi \left( \int_{x \geq \xi} l_x dx \right) d\xi \\ &= \frac{1}{l_0 \dot{e}_0} \int_0^\infty \mu_x l_x \dot{e}_x dx. \end{aligned} \quad (7.2.1)$$

Se considerarmos a última equação em (7.2.1), podemos interpretar o conceito de entropia da seguinte forma. O numerador de  $H$  representa o número médio de anos de vida restante “perdidos” em consequência dos óbitos registados. Tal é assim porquanto  $\mu_x l_x$  reproduz o número médio de óbitos à idade  $x$ , e  $\dot{e}_x$  denota o valor esperado da duração de vida restante de cada um deles nessa idade. O denominador  $l_0 \dot{e}_0$  traduz a esperança de vida à nascença de toda a geração. Uma diminuição na mortalidade traduz-se num aumento do número total de anos vividos por toda a geração. Neste sentido, a entropia pode ser entendida como um indicador dos ganhos relativos de esperança de vida conseguidos pela diminuição da mortalidade.

Dois casos extremos merecem, a este propósito, uma referência especial:

- Em primeiro lugar, no caso de todos os membros da população falecerem exactamente à mesma idade  $x^*$ , prova-se que  $H = 0$ . Com efeito, nesse caso  $\mu_x = 0$  para  $x \in [0, x^*[$  e  $l_x = 0$  para  $x \in ]x^*, \infty[$ , pelo que o integral em (7.2.1) é nulo.
- Em segundo, nos casos em que  $\mu_x$  é constante ao longo de todo o intervalo de idades, i.e.,  $\mu_x \equiv \mu \forall x$ , deduz-se que

$$\dot{e}_x = \frac{1}{\mu} \text{ e } l_x = l_0 \exp(-\mu x) \quad (7.2.2)$$

---

<sup>3</sup>O conceito de entropia é adoptado em domínios tão distintos como a termodinâmica, probabilidades e estatística, investigação operacional, economia, finanças, marketing, planeamento urbano e de redes de transportes, com definições e interpretações naturalmente distintas. Formulado originalmente no contexto da física, a entropia é adoptada como uma medida da desordem de um sistema físico. Já no domínio da teoria da informação, a sua introdução é atribuída a Shannon (1948), referindo-se à incerteza de uma distribuição de probabilidade.

## 7.2. Evolução da mortalidade em Portugal

---

donde se deriva que  $H = 1$ .

Em suma, o valor da entropia  $H$  de uma tábua de mortalidade posiciona-se entre estes dois extremos, ou seja,  $H \in [0, 1]$ . O fenómeno de rectangularização da função de sobrevivência traduz-se numa diminuição da entropia.

No Quadro 7.2.2 reportamos a evolução de  $H$  desde os anos 30 até ao presente. Como se observa, regista-se uma diminuição progressiva nos valores de  $H$  em ambos os sexos em todo o período analisado, reflectindo a diminuição da mortalidade infantil e a já mencionada rectangularização das curvas de sobrevivência. A comparação entre as trajetórias evidenciadas pelas séries de  $H$  relativas a ambos os sexos põe de novo em evidência a maior incidência do fenómeno da rectangularização no caso das mulheres.

| Período   | Entropia $H$ |          |
|-----------|--------------|----------|
|           | Homens       | Mulheres |
| 1930-1931 | 0.509        | 0.467    |
| 1939-1942 | 0.500        | 0.448    |
| 1949-1952 | 0.389        | 0.339    |
| 1959-1962 | 0.307        | 0.260    |
| 1969-1972 | 0.245        | 0.199    |
| 1979-1982 | 0.206        | 0.157    |
| 1990-1991 | 0.188        | 0.139    |
| 2002-2003 | 0.164        | 0.122    |

Quadro 7.2.2: Entropia das tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003

Uma forma alternativa de apreender o fenómeno da rectangularização consiste em analisar a evolução do indicador vida mediana (ou vida provável)  $vm_x$ , definida em (2.3.6) por  $vm_x p_x = 0.5$ . No Quadro 7.2.3 reportamos os valores da vida provável à nascença  $vm_0$  no período em análise. Como se observa pela leitura do quadro, a vida provável progrediu de forma assinalável em ambos os sexos entre 1930-31 e 2002-03, aumentando 18.68 e 17.80 anos nos homens e mulheres, respectivamente. Em resultado, um indivíduo nascido no período 2002-03 terá 50% de probabilidade de alcançar os 77.54 anos, se for do sexo masculino, e de perfazer 83.58 anos, se for do sexo feminino. Estes valores indicam que um número cada vez maior de elementos de cada geração sobrevive até mais tarde, até às idades avançadas, idades onde se registam a maioria dos óbitos e após as quais o número de sobreviventes decai de forma acelerada.

## 7.2. Evolução da mortalidade em Portugal

| Período   | Vida mediana $vm_0$ |          |
|-----------|---------------------|----------|
|           | Homens              | Mulheres |
| 1930-1931 | 58.86               | 65.78    |
| 1939-1942 | 59.79               | 66.52    |
| 1949-1952 | 66.33               | 72.31    |
| 1959-1962 | 70.06               | 75.75    |
| 1969-1972 | 70.99               | 76.71    |
| 1979-1982 | 73.62               | 79.67    |
| 1990-1991 | 75.04               | 81.21    |
| 2002-2003 | 77.54               | 83.58    |

Quadro 7.2.3: Vida mediana subjacente às tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003

Para concluir a análise sobre o fenómeno da rectangularização, analisamos no Quadro 7.2.4 a variabilidade da distribuição apresentando os valores do intervalo inter-quartis,  $IQR = x'' - x'$ , onde  $x'$  e  $x''$  são as idades para as quais  $l_{x'} = 0.75$  e  $l_{x''} = 0.25$  (com  $l_0 = 1$ ). Como se observa, a diminuição assinalável do  $IQR$  em ambos os sexos (especialmente no sexo feminino) verificada no período em análise acompanha a crescente concentração do número de óbitos em torno da moda, e confirma a cada vez menor dispersão da distribuição da vida residual da população portuguesa.

| Período   | Intervalo inter-quartis $IQR$ |                        |       |                        |                        |       |
|-----------|-------------------------------|------------------------|-------|------------------------|------------------------|-------|
|           | Homens                        |                        |       | Mulheres               |                        |       |
|           | Quartil <sub>inf</sub>        | Quartil <sub>sup</sub> | $IQR$ | Quartil <sub>inf</sub> | Quartil <sub>sup</sub> | $IQR$ |
| 1930-1931 | 18.17                         | 74.06                  | 55.90 | 22.01                  | 78.16                  | 56.15 |
| 1939-1942 | 19.55                         | 73.56                  | 54.01 | 23.97                  | 77.85                  | 53.89 |
| 1949-1952 | 41.03                         | 76.89                  | 35.86 | 50.33                  | 81.12                  | 30.79 |
| 1959-1962 | 54.00                         | 78.92                  | 24.93 | 63.16                  | 83.82                  | 20.65 |
| 1969-1972 | 58.42                         | 78.91                  | 20.49 | 67.19                  | 83.02                  | 15.83 |
| 1979-1982 | 62.09                         | 81.10                  | 19.01 | 71.25                  | 85.66                  | 14.41 |
| 1990-1991 | 64.10                         | 82.52                  | 18.42 | 73.33                  | ND                     | ND    |
| 2002-2003 | 67.53                         | ND                     | ND    | 76.20                  | ND                     | ND    |

Quadro 7.2.4: Intervalo inter-quartis  $IQR$  subjacente às tábuas de mortalidade da população portuguesa de 1930-1931 a 2002-2003

### 7.2.2 Evolução da esperança de vida

No Quadro 7.2.5 apresentamos a evolução da esperança de vida à nascença  $e_0(t)$  e aos 65 anos  $e_{65}(t)$ , extraída das tábuas de mortalidade para a população portuguesa publicadas pelo INE no período de 1930-31 a 2002-03. O facto mais destacado pela leitura

## 7.2. Evolução da mortalidade em Portugal

do quadro refere-se, naturalmente, aos ganhos notáveis nos indicadores de esperança de vida registados neste período em ambos os sexos. Com efeito, em apenas 70 anos a esperança de vida à nascença aumentou 25.85 anos (ou 54%), no caso dos homens, e 28.11 anos (ou 54%) no caso das mulheres, um progresso assinalável sob todos os pontos de vista. A evolução do indicador  $\dot{e}_{65}(t)$  é naturalmente menos expressiva em termos absolutos (4.08 e 5.59 anos, respectivamente), mas igualmente assinalável em termos relativos, com ganhos de 36% e 43% nos homens e mulheres, respectivamente. Registe-se, por fim, que os ganhos de esperança de vida ocorreram a um ritmo irregular, com as maiores conquistas a serem registadas na primeira metade do período analisado.

| Período   | $\dot{e}_0(t)$ |          | $\dot{e}_{65}(t)$ |          |
|-----------|----------------|----------|-------------------|----------|
|           | Homens         | Mulheres | Homens            | Mulheres |
| 1930-1931 | 48.08          | 52.12    | 11.49             | 13.09    |
| 1939-1942 | 48.58          | 52.82    | 10.82             | 12.70    |
| 1949-1952 | 55.52          | 60.50    | 11.90             | 13.96    |
| 1959-1962 | 60.73          | 66.35    | 12.42             | 14.64    |
| 1969-1972 | 64.25          | 70.48    | 12.02             | 14.49    |
| 1979-1982 | 68.44          | 75.35    | 13.41             | 16.36    |
| 1990-1991 | 70.60          | 77.52    | 14.12             | 17.13    |
| 2002-2003 | 73.93          | 80.23    | 15.57             | 18.68    |

Quadro 7.2.5: Evolução da esperança de vida da população portuguesa entre 1930-1931 e 2002-2003

A Figura 7.2.4 representa as funções  $x \mapsto \dot{e}_x(t)$  no período de 1930-31 a 1994. Como se observa, as curvas  $x \mapsto \dot{e}_x(t)$  apresentam, desde 1930-1931 até 1969-1972, um pico em torno dos 4-5 anos, ou seja, verifica-se que a esperança de vida começa por crescer até por volta dos 5 anos, para em seguida decrescer. Este comportamento, aparentemente inesperado, é característico de populações onde o nível de mortalidade infantil é todavia significativo. Para melhor compreendermos este fenómeno, notemos que

$$\begin{aligned} \frac{d}{dx} \dot{e}_x &= \frac{d}{dx} \left( \frac{1}{l_x} \int_0^\infty l_{x+\xi} d\xi \right) \\ &= -1 - \frac{1}{(l_x)^2} \left( \frac{d}{dx} l_x \right) \int_0^\infty l_{x+\xi} d\xi. \end{aligned}$$

Se recordarmos que  $\frac{d}{dx} l_x = -l_x \mu_x$  obtemos, após alguma manipulação matemática

$$\frac{d}{dx} \dot{e}_x = -1 + \mu_x \dot{e}_x. \quad (7.2.3)$$

## 7.2. Evolução da mortalidade em Portugal

O significado desta equação torna-se agora claro. Nas idades em que a força de mortalidade é muito baixa (i.e.,  $\mu_x \approx 0$ ), cada novo aniversário aproxima-nos um ano do fim da vida (i.e.,  $\frac{d}{dx}\hat{e}_x \approx -1$ ). Pelo contrário, nas idades em que a mortalidade é importante (i.e., em que o valor de  $\mu_x$  é significativo) e a esperança de vida  $\hat{e}_x$  é considerável, a derivada de  $\hat{e}_x$  em relação a  $x$  pode ser positiva. Este é, em particular, o caso das idades muito jovens em populações onde a mortalidade infantil é todavia muito elevada, nas quais é normal a esperança de vida remanescente de uma criança de 5 anos ser superior à de um recém-nascido.

A Figura 7.2.4 mostra igualmente que, num passado mais recente, a esperança de vida  $\hat{e}_x$  decresce regularmente com a idade, num sinal de que a mortalidade infantil em Portugal diminuiu para valores muito baixos. Com efeito, Portugal apresenta hoje indicadores de mortalidade infantil semelhantes aos reportados pelos países com melhor performance a nível mundial.

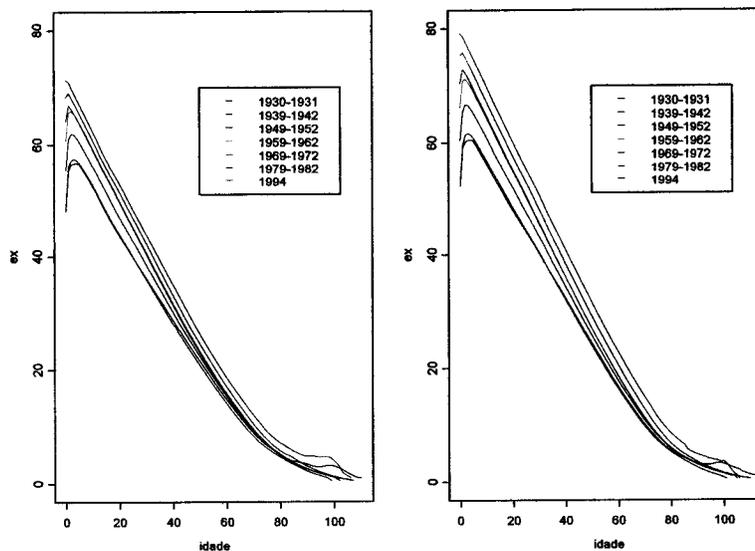


Figura 7.2.4: Esperança de vida  $x \mapsto \hat{e}_x(t)$  (homens à esquerda e mulheres à direita)

### 7.2.3 Bossa de acidentes

A denominada “bossa” de acidentes (assim comumente designada, embora ela represente outros fenómenos como o suicídio) representa um fenómeno de sobremortalidade observável nas idades juvenis da maioria das populações. Para confirmarmos a presença

## 7.2. Evolução da mortalidade em Portugal

---

deste fenómeno na população portuguesa, representamos nas Figuras 7.2.5 e 7.2.6 o logaritmo dos quocientes  $q_x$  para ambos os sexos no período em análise.

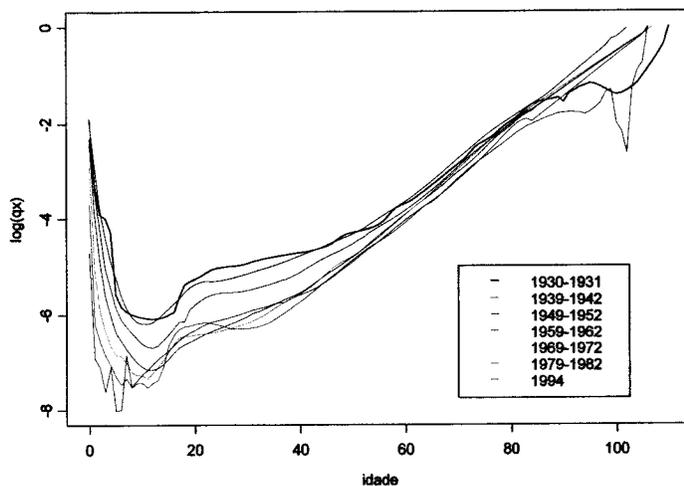


Figura 7.2.5: Logaritmo do quociente de mortalidade  $x \mapsto \ln(q_x)$ , Homens

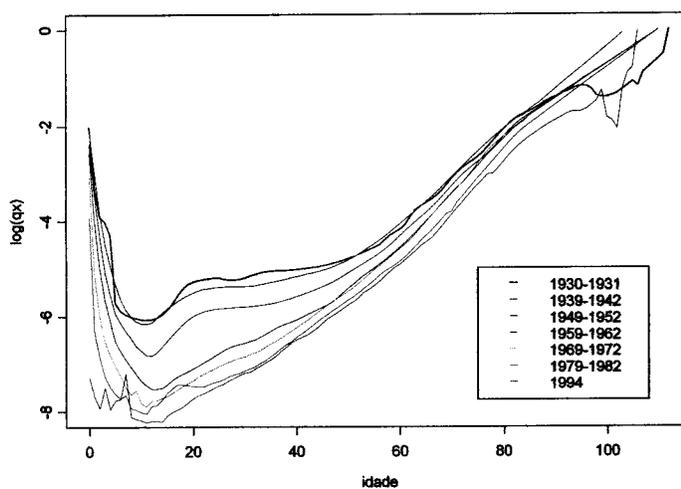


Figura 7.2.6: Logaritmo do quociente de mortalidade  $x \mapsto \ln(q_x)$ , Mulheres

Como se observa, a “bossa” de acidentes perdeu alguma da sua importância em ambos os sexos num passado mais recente, apresentando uma tendência para se dispersar, em especial no caso da população feminina. A bossa traduz-se por quocientes particularmente elevados no intervalo de idades entre os 15-30 anos, em consequência do risco acrescido da ocorrência de mortes violentas, em particular no caso da população masculina.

A observação das Figuras 7.2.5 e 7.2.6 permite ainda verificar que se regista um decréscimo nos quocientes  $q_x$  até por volta dos 15 anos, e que partir dos 25-30 anos o fenómeno segue a sua normal trajectória ascendente. Realce ainda para a irregularidade no comportamento dos quocientes nas idades elevadas, confirmando as suspeitas avançadas no capítulo anterior aquando da discussão dos métodos de fecho de tábuas.

### 7.2.4 Diferencial de mortalidade entre homens e mulheres

O debate em torno dos fundamentos que explicam as diferenças na esperança de vida entre homens e mulheres alimenta, desde há muito, acesas discussões na comunidade científica, em razão da dificuldade em decifrar as complexas interações entre condições biológicas, sociais e comportamentais que influenciam o fenómeno.<sup>4</sup>

As diferenças de mortalidade entre sexos têm naturalmente repercussões importantes sobre a proporção de mulheres na população idosa. Com efeito, dados do US Bureau of Census para os países desenvolvidos indicam que em 1998 existiam 69.7 homens por cada 100 mulheres com mais de 60 anos e 44.3 homens por cada 100 mulheres com mais de 80 anos. As projecções efectuadas por Kinsella e Gist (1998) para o horizonte de 2025 indicam que estas cifras se situarão nos 78 e 55.8 homens, respectivamente, ou seja, projectam uma ligeira redução na assimetria da distribuição por sexos da população idosa devido à (prevista) maior velocidade de crescimento da esperança de vida para os homens.

Neste contexto, mesmo que exista uma distribuição equilibrada entre sexos nos novos nascimentos, as diferenças na mortalidade entre homens e mulheres induzem uma alteração progressiva na composição da população adulta e idosa em favor do sexo feminino. Esta dominância feminina tem evidentemente consequências a vários níveis (e.g., sobre a gestão dos regimes de segurança social, sobre a procura de serviços de saúde). A importância deste fenómeno é analisada, entre outros, por Kinsella e Gist (1998). Os autores destacam a espectacular evolução da esperança de vida à nascença registada

---

<sup>4</sup>A análise detalhada deste tema vai muito para além do objectivo deste trabalho. Para uma contribuição recente veja-se, por exemplo, Lemaire (2001). Neste trabalho o autor analisa, para um total de 169 países, uma série de factores que explicam a maior longevidade feminina.

## 7.2. Evolução da mortalidade em Portugal

no século XX e evidenciam, em especial para os países desenvolvidos, que salvo raras exceções, os valores de  $\hat{e}_0$  são superiores nas mulheres. Concluem ainda que o ritmo de crescimento da longevidade está longe de ser linear, particularmente no caso dos homens, e que o diferencial de longevidade entre homens e mulheres cresceu com o tempo em favor das mulheres.

No que respeita à realidade portuguesa, os dados apresentados no Quadro 7.2.5 evidenciam claramente a maior longevidade das mulheres, que em 2002-03 registavam uma esperança de vida à nascença 6.3 anos superior à dos homens, contra um diferencial de 3.11 anos à idade 65. Este diferencial é comparável com o valor médio registado nos países desenvolvidos, estimado actualmente em cerca 7 anos.

Na Figura 7.2.7 representamos a evolução da esperança de vida à nascença  $\hat{e}_0$  e aos 65 anos  $\hat{e}_{65}$  por sexo, bem como o diferencial  $(\hat{e}_x^M - \hat{e}_x^H)_{x \in \{0,65\}}$ , onde  $M$  e  $H$  designam mulheres e homens, respectivamente. Como se observa, a amplitude (em termos absolutos) do diferencial entre as esperanças de vida de ambos os sexos aumentou entre 1930-31 e 2002-03, quer à nascença, quer aos 65 anos. Assinale-se, contudo, que durante a última década do século XX se observa uma ligeira diminuição no diferencial  $(\hat{e}_0^M - \hat{e}_0^H)$ .

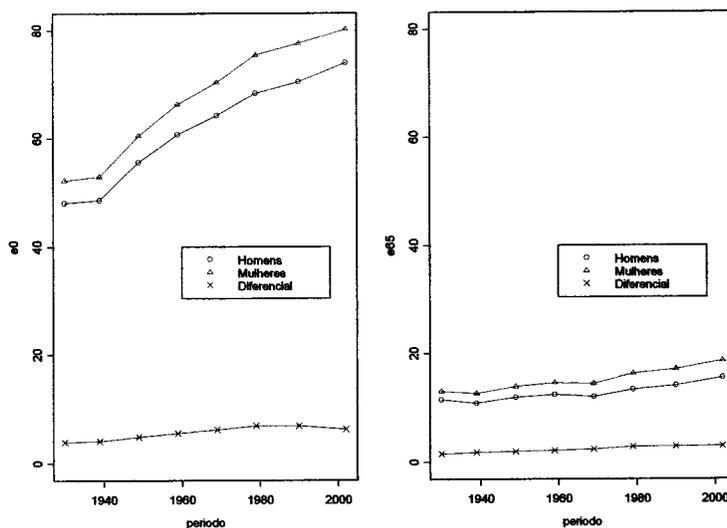


Figura 7.2.7: Evolução da esperança de vida à nascença  $\hat{e}_0$  e aos 65 anos  $\hat{e}_{65}$  por sexo

Esta conclusão é reforçada na Figura 7.2.8, onde se representa a diferença entre as esperanças de vida  $\hat{e}_x$  de homens e mulheres em função da idade  $x$ , nos vários momentos

## 7.2. Evolução da mortalidade em Portugal

considerados. Um valor positivo no gráfico indica uma esperança de vida (em anos) superior para as mulheres.

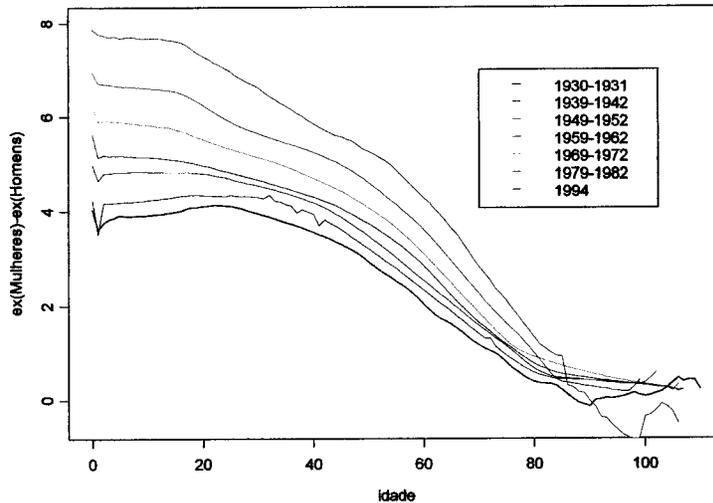


Figura 7.2.8: Evolução do diferencial entre as esperanças de vida ( $e_x^M - e_x^H$ ) em função da idade  $x$

Como se observa, o diferencial de longevidade esperada é favorável às mulheres em (praticamente) todas as idades e em todos os momentos analisados.<sup>5</sup> O aumento deste diferencial com o tempo é bem patente em todas as idades, mas sobretudo nas mais jovens onde atinge o seu valor máximo.

Na Figura 7.2.9 representamos o ratio entre os quocientes  $q_x$  de ambos os sexos  $x \mapsto q_x^H / q_x^M$  para as 7 tábuas de mortalidade consideradas. Um valor superior à unidade indica sobremortalidade dos homens face às mulheres na idade e período considerados.

A leitura do gráfico permite concluir, em primeiro lugar, que a mortalidade masculina é superior à feminina em praticamente todas as idades e períodos analisados. Em segundo, regista-se no decurso do tempo um aumento significativo na sobremortalidade masculina na generalidade das idades, em especial no intervalo 10-40, explicável em parte pela já mencionada “bossa” de acidentes, que afecta de forma mais premente os jovens do sexo masculino. Por fim, nas idades avançadas assiste-se a alguma convergência no

<sup>5</sup> Exceptua-se o caso da tábua de 1994 que apresenta um comportamento inexplicável e anormalmente errático nas idades elevadas.

### 7.3. Modelo de Poisson

---

comportamento da mortalidade entre os sexos.

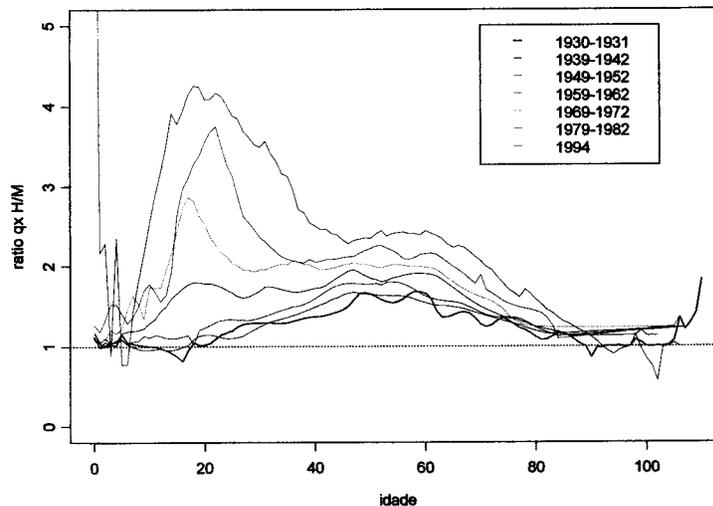


Figura 7.2.9: Ratio entre os quocientes de mortalidade  $q_x^H/q_x^M$  em função da idade  $x$

## 7.3 Modelo de Poisson

### 7.3.1 Estatísticas de mortalidade

O ponto de partida de qualquer análise sobre a evolução da mortalidade consiste em recolher um conjunto de dados fiáveis sobre a população estudada. Em particular, para os anos  $t_{\min}, \dots, t_{\max}$  e para as idades  $x_{\min}, \dots, x_{\max}$  é necessário compilar informação sobre o número de óbitos registados em cada idade e sobre a correspondente exposição ao risco. A este respeito, os dados usados neste estudo compreendem o número de óbitos  $d_{x,t}$ , discriminados por idade (mas não por data de nascimento, i.e., referem-se ao quadrado de Lexis) e sexo no período  $(t_{\min}, t_{\max}) = (1970, 2004)$ , com intervalo de idades igual a  $(x_{\min}, x_{\max}) = (0, 117)$ , no caso dos homens, e igual a  $(x_{\min}, x_{\max}) = (0, 124)$ , no caso das mulheres. Compreendem ainda as estimativas de população residente  $P_{x,t}$ , discriminadas por sexo e idade no período  $(t_{\min}, t_{\max}) = (1970, 2004)$ , com intervalo de idades igual a  $(x_{\min}, x_{\max}) = (0, 84)$  em ambos os sexos. Os dados foram fornecidos pelo Instituto Nacional de Estatística (I.N.E.) e referem-se ao conjunto da população portuguesa.

Conforme referimos no capítulo anterior, a estimação de alguns modelos (e.g., Lee-Carter) pressupõe a existência de uma matriz de dados completa na forma rectangular. Como se observa no caso em apreço, as restrições da amostra em termos de intervalo de idades coberto pelas estimativas de população residente condicional, invariavelmente, a análise do comportamento da mortalidade nas idades mais avançadas.

Já no que toca ao modelo de Poisson, salientámos oportunamente que o ajustamento não depende de uma decomposição em valores singulares, pelo que a sua implementação é possível independentemente da existência ou não de uma matriz rectangular de dados completa. Dito de outro modo,  $x_{\min}$  e  $x_{\max}$  não necessitam de ser idênticos em todos os anos de calendário  $t_{\min}, \dots, t_{\max}$  da amostra. Acresce ainda que o método acomoda perfeitamente casos em que alguma observação  $(x, t)$  não está disponível.

#### 7.3.2 Estimativa bruta da força de mortalidade

Nas Figuras 7.3.1 e 7.3.2 representamos a estimativa bruta de  $\mu_{x,t}$ ,  $x \in [0, 84]$ , no período considerado. A observação atenta dos gráficos permite detectar as mesmas tendências assinaladas na secção anterior no contexto da análise transversal, nomeadamente a diminuição progressiva da mortalidade ao longo do tempo em todas as idades, a queda acentuada da mortalidade infantil e a importância do fenómeno “bossa” de acidentes, sobretudo na população do sexo masculino.

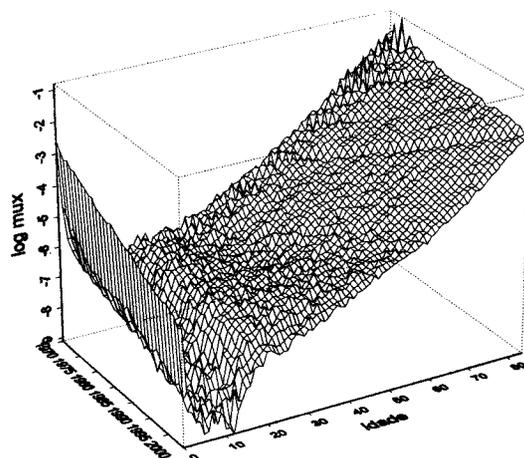


Figura 7.3.1: Estimativas brutas de  $\mu_{x,t}$  no período 1970-2004, Homens

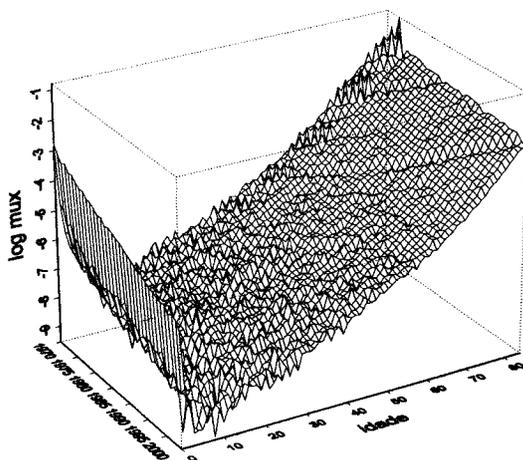


Figura 7.3.2: Estimativas brutas de  $\mu_{x,t}$  no período 1970-2004, Mulheres

### 7.3.3 Estimação dos parâmetros por máxima verosimilhança

Nesta secção testamos a versão base do modelo de Poisson descrito em detalhe no Capítulo 6, que envolve, como vimos, a estimação dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  através do método de máxima verosimilhança. A primeira etapa em qualquer procedimento de projecção consiste em reduzir a dimensão do problema. De facto, seria praticamente impossível tratar simultaneamente todas as séries cronológicas que descrevem a evolução ao longo do tempo dos quocientes  $q_x$  (ou das taxas instantâneas  $\mu_x$ ) nas diferentes idades.

O primeiro aspecto a ter em conta respeita à selecção do intervalo de idades a considerar pelos modelos na estimação dos parâmetros. Neste campo, se para estimar indicadores como  $\hat{e}_0(t)$  não nos resta outra hipótese que não a de contemplar todas as idades disponíveis, no caso da análise visar o cálculo de indicadores à idade 65 anos (e.g.,  $\hat{e}_{65}(t)$ ), podemos confinar o estudo e considerar apenas, por exemplo, a evolução da mortalidade nas idades superiores a 50 anos. Esta delimitação procura garantir que o modelo estimado é expurgado de um certo “ruído” gerado pela mortalidade infantil e juvenil, ou pela “bossa” de acidentes, fenómenos que, como veremos em seguida, são difíceis de modelar e podem influenciar a estimação (e naturalmente a projecção) do índice temporal  $\kappa_t$ .

Neste sentido, num primeiro momento testámos uma versão do modelo em que foram

### 7.3. Modelo de Poisson

---

considerados os dados disponíveis para todas as idades  $x \in [0, 84]$  e anos de calendário  $t \in [1970, 2004]$ . Posteriormente, testámos a sensibilidade dos resultados a uma alteração no intervalo de idades, restringindo a análise às idades  $x \in [50, 84]$ .

Para estimar os parâmetros do modelo, recorremos ao método de máxima verosimilhança adoptando o método iterativo desenvolvido por Goodman (1979) e materializado no algoritmo (6.9.10). Para iniciar o algoritmo, considerámos o seguinte vector de valores iniciais para os parâmetros:  $(\hat{\alpha}_x^{(0)}, \hat{\kappa}_t^{(0)}, \hat{\beta}_x^{(0)}) = (0, 1, 0.1)$ . A condição de paragem (critério de óptimo) da rotina envolveu a fixação de um valor mínimo  $\varepsilon = 10^{-4}$  para a variação absoluta na função de log-verosimilhança (função objectivo)  $\mathcal{L}(\alpha, \beta, \kappa)$ . As estimativas dos parâmetros do modelo e os resultados dos testes de qualidade do ajustamento foram conseguidos com o auxílio de uma rotina informática especialmente escrita para o efeito, implementada no *package* SAS (Release 9.1).

Nas Figuras 7.3.3 e 7.3.4 representamos as estimativas dos parâmetros  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  do modelo de Poisson aplicado, respectivamente, a todas as idades e apenas às idades  $x \geq 50$  anos. Os respectivos valores encontram-se discriminados no Anexo 7.A. O primeiro aspecto a realçar refere-se à influência exercida pela escolha do intervalo de idades nas estimativas de alguns dos parâmetros do modelo. Como se observa pela comparação das figuras, os parâmetros  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  diferem significativamente nalgumas idades, um resultado que confirma as diferentes trajectórias registadas na mortalidade das idades jovens e das idades avançadas.

A curva dos  $\hat{\alpha}_x$  assume, como se antevia, a sua configuração habitual, ou seja, apresentando taxas de mortalidade médias relativamente elevadas para os recém-nascidos e idades infantis, que decrescem muito rapidamente com a idade para atingir o seu mínimo por volta dos 12 anos, progredindo depois regularmente com a idade, com excepção da já referida “bossa” de acidentes em torno dos 20-25 anos (visivelmente mais acentuada no caso dos homens).

Por outro lado, é possível constatar que as maiores oscilações temporais nas taxas  $\mu_x$  (associadas a valores mais significativos do parâmetro  $\hat{\beta}_x$ ) estão situadas nas idades mais jovens, muito provavelmente em consequência dos progressos realizados pela medicina no combate à mortalidade infantil e juvenil. O padrão gráfico do parâmetro  $\hat{\beta}_x$  em torno dos 25 anos deixa ainda antever, por um lado a importância da mortalidade accidental no sexo masculino e, por outro, os progressos incipientes registados neste domínio no período em análise. Com efeito, os valores do parâmetro  $\hat{\beta}_x$  aproximam-se bastante de zero neste segmento, apresentando inclusive valores ligeiramente negativos nalgumas idades, sinónimo de que as taxas registadas nalgumas idades não acompanharam o decréscimo geral observado no período.

### 7.3. Modelo de Poisson

---

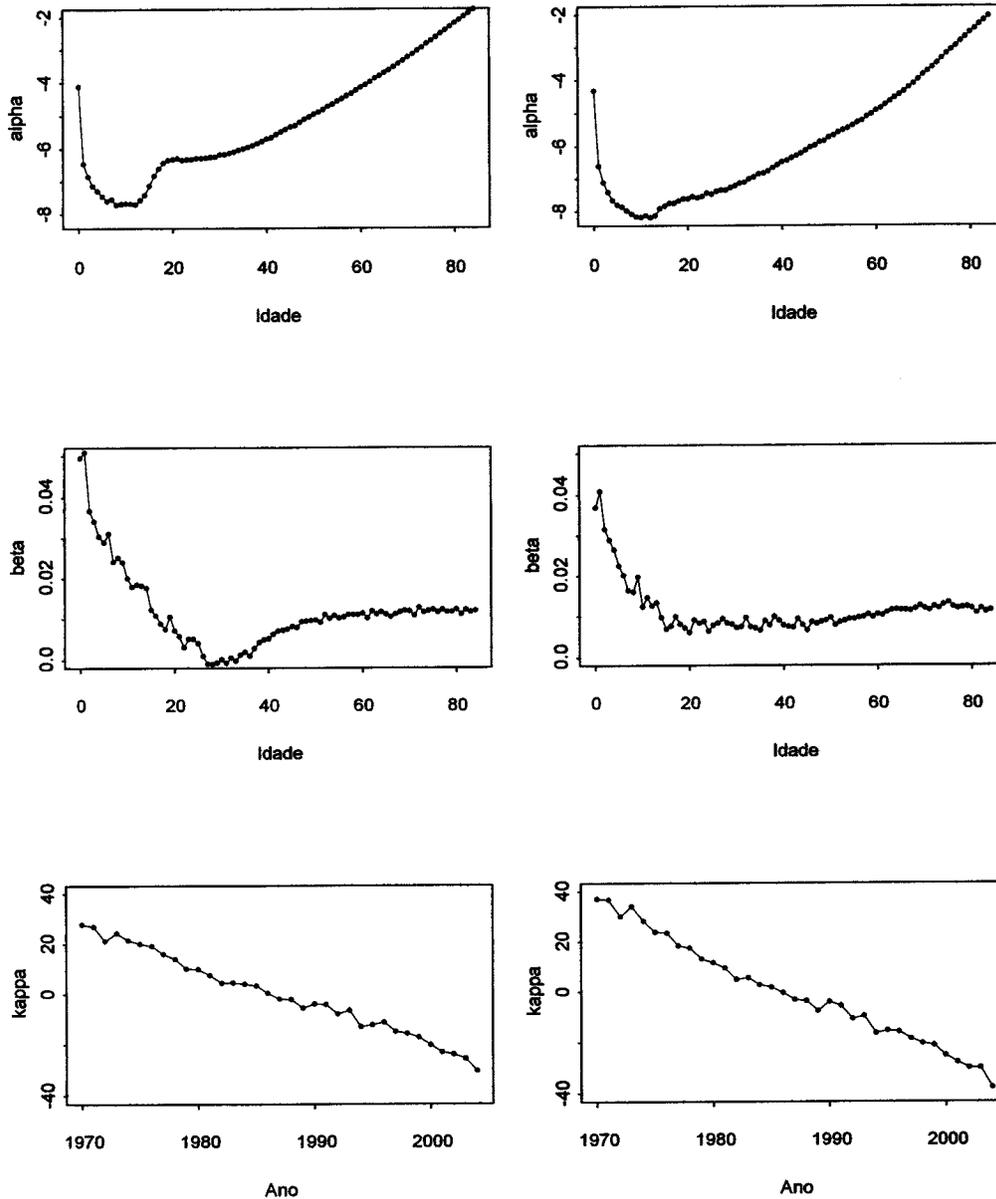


Figura 7.3.3: Estimativas dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  do modelo de Poisson aplicado a todas as idades (homens à esquerda e mulheres à direita)

### 7.3. Modelo de Poisson

---

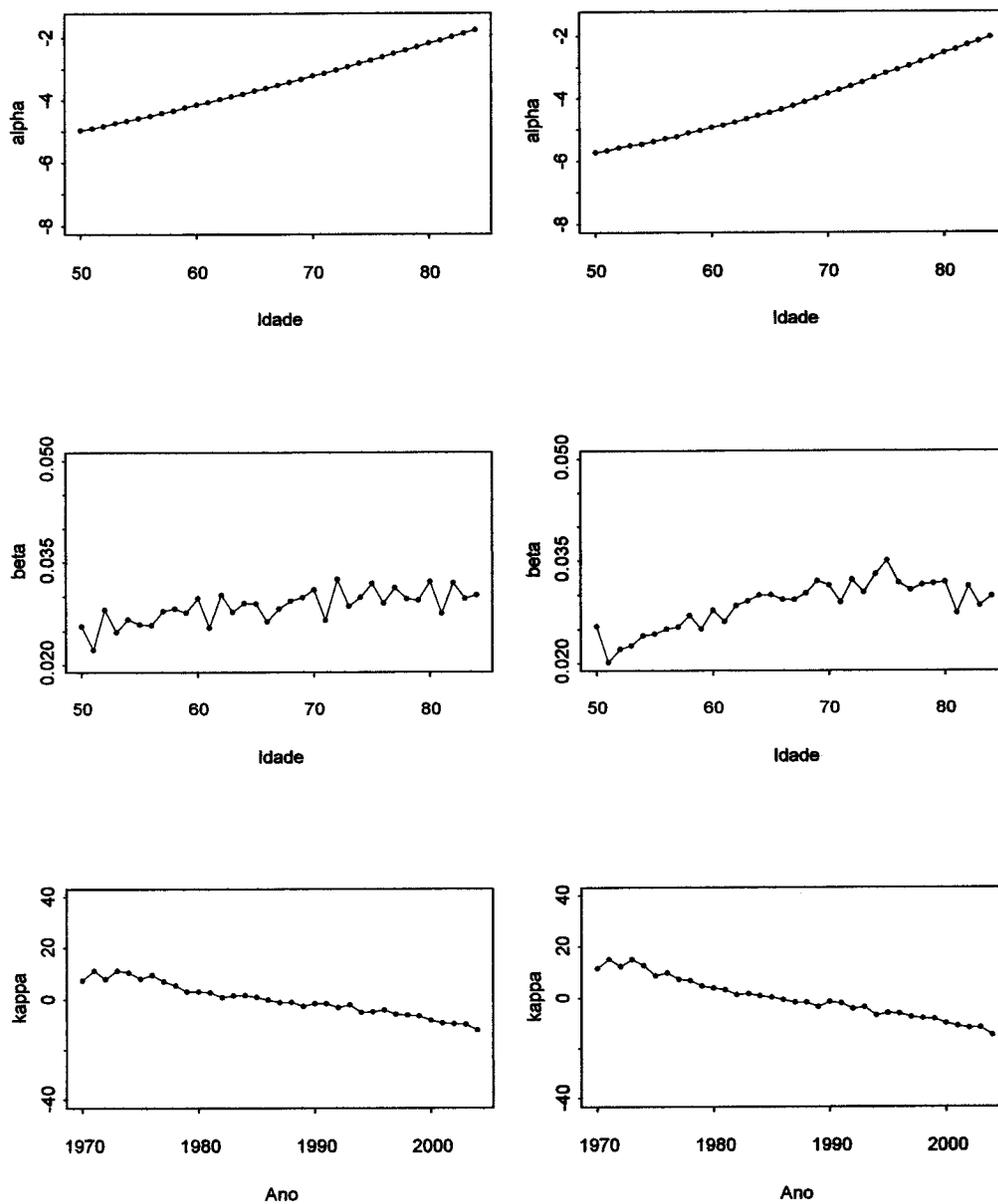


Figura 7.3.4: Estimativas dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  do modelo de Poisson aplicado apenas às idades  $x \geq 50$  (homens à esquerda e mulheres à direita)

Destaque ainda para a maior sensibilidade das taxas às variações em  $\kappa_t$  nas idades adultas do sexo masculino, por contraposição ao comportamento mais estável no sexo feminino.

Por último, a análise da trajectória do índice temporal  $\hat{\kappa}_t$  revela, em ambos os sexos, uma tendência decrescente bem definida (aproximadamente linear). Este resultado, conjugado com estimativas  $\hat{\beta}_x$  positivas para a esmagadora maioria das idades, traduz em ambos os sexos o decréscimo acentuado das taxas de mortalidade observado em Portugal no último terço do século XX.

Para aferir da qualidade do ajustamento, calculámos em primeiro lugar o indicador deviance definido em (6.9.16). Os resultados obtidos, no valor de 15667.51 no caso dos homens e de 15249.63 para as mulheres, põem em evidência o (ligeiro) melhor desempenho do modelo na população feminina, uma diferença que não é, todavia, estatisticamente significativa.

De seguida, e com o objectivo de obter um termo de comparação entre os desempenhos do modelo de Poisson e do método Lee-Carter, determinámos a percentagem da variância temporal explicada pelo modelo em cada idade  $x$ . Para tal recorreremos, em cada idade  $x$ , à seguinte expressão

$$1 - \frac{\frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} (\hat{\Delta}_{x,t})^2 - \left( \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \hat{\Delta}_{x,t} \right)^2}{\frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} (\hat{\mu}_{x,t})^2 - \left( \frac{1}{T} \sum_{t=t_{\min}}^{t_{\max}} \hat{\mu}_{x,t} \right)^2}, \quad (7.3.1)$$

onde  $T = (t_{\max} - t_{\min} + 1)$ ,  $\hat{\Delta}_{x,t} = \hat{\mu}_{x,t} - \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t)$  e  $\hat{\mu}_{x,t}$  denota a estimativa bruta de  $\mu_{x,t}$ .

O Quadro 7.3.1 reporta os resultados obtidos num conjunto de idades seleccionadas. A leitura do quadro permite concluir que a percentagem da variância explicada pelo modelo pode ser considerada boa na generalidade das idades, com excepção do intervalo 20-35 anos, no caso dos homens, onde a qualidade do ajustamento se deteriora fortemente. A explicação para este comportamento deficiente encontra-se, uma vez mais, na dificuldade do modelo em captar o fenómeno da mortalidade acidental, bastante irregular no tempo, que afecta sobretudo este segmento.

Por outro lado, verifica-se que o modelo aplicado no intervalo de idades 50-84 apresenta melhores resultados que o seu homólogo aplicado a todas as idades, dada a menor variabilidade observada nas idades avançadas no período em análise.

### 7.3. Modelo de Poisson

| Idade | $x \in [0, 84]$ |          | $x \in [50, 84]$ |          |
|-------|-----------------|----------|------------------|----------|
|       | Homens          | Mulheres | Homens           | Mulheres |
| 0     | 0.9828          | 0.9696   |                  |          |
| 5     | 0.9200          | 0.8553   |                  |          |
| 10    | 0.7401          | 0.6930   |                  |          |
| 15    | 0.5929          | 0.3325   |                  |          |
| 20    | 0.3055          | 0.4770   |                  |          |
| 25    | 0.0827          | 0.5975   |                  |          |
| 30    | 0.0003          | 0.5162   |                  |          |
| 35    | 0.0710          | 0.5580   |                  |          |
| 40    | 0.3895          | 0.7142   |                  |          |
| 45    | 0.6214          | 0.6637   |                  |          |
| 50    | 0.7226          | 0.8051   | 0.8188           | 0.8503   |
| 55    | 0.7980          | 0.8920   | 0.8241           | 0.9223   |
| 60    | 0.8035          | 0.8683   | 0.8828           | 0.9063   |
| 65    | 0.7950          | 0.8650   | 0.8775           | 0.9086   |
| 70    | 0.8030          | 0.8317   | 0.8957           | 0.8800   |
| 75    | 0.6165          | 0.7365   | 0.6904           | 0.7846   |
| 80    | 0.7342          | 0.7295   | 0.8174           | 0.8007   |
| 84    | 0.4978          | 0.4615   | 0.5284           | 0.5012   |

Quadro 7.3.1: Percentagem da variância explicada pelo modelo Poisson em idades seleccionadas

Para avaliar a capacidade de ajustamento do modelo, analisámos ainda o comportamento dos resíduos. O objectivo é o de averiguar se estes não apresentam qualquer padrão que indicie problemas na especificação do modelo e na sua capacidade para realizar previsão. Nesse sentido, calculámos num primeiro momento a diferença entre o número total de óbitos

$$D_{\bullet,t} = \sum_{x=x_{\min}}^{x_{\max}} d_{x,t}$$

observado em cada ano de calendário e a sua correspondente estimativa

$$\hat{D}_{\bullet,t} = \sum_{x=x_{\min}}^{x_{\max}} E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t).$$

Na Figura 7.3.5 apresentamos os resultados deste exercício. Como se constata, apenas no caso do modelo aplicado à população masculina nas idades 0-84 anos é possível descortinar um (ligeiro) padrão oscilatório. Em consequência do teorema do limite central,  $D_{\bullet,t}$  segue uma distribuição aproximadamente Normal, de média e variância igual a  $\sum_{x=x_{\min}}^{x_{\max}} E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t)$ . Esta aproximação permite-nos construir intervalos de confiança ao nível de 95%, representados na Figura 7.3.5 a tracejado.

### 7.3. Modelo de Poisson

Como se observa, todos os resíduos estão situados dentro dos intervalos correspondentes à distribuição Normal, um resultado conforme com o modelo.

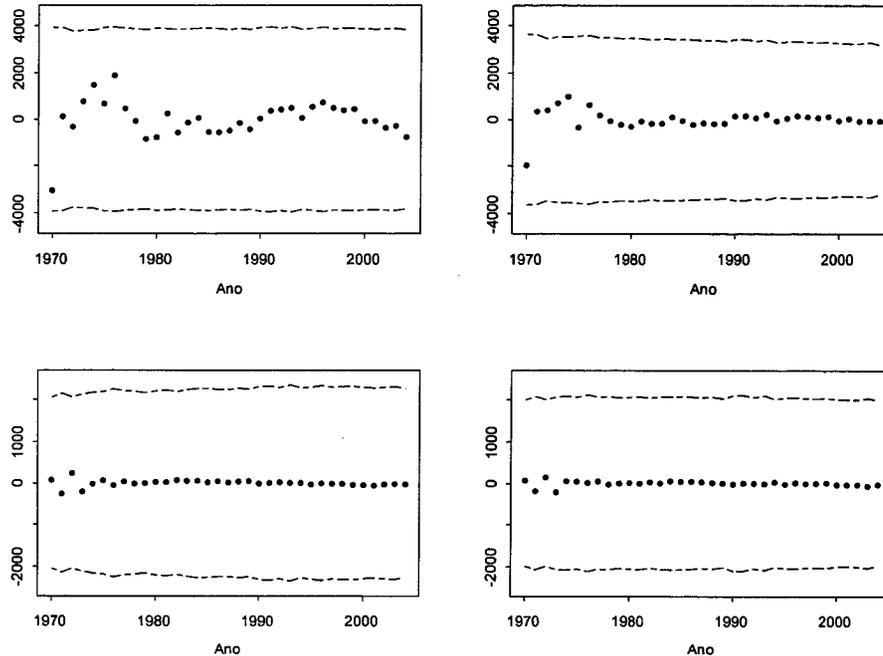


Figura 7.3.5: Diferença entre o número de óbitos observado em cada ano de calendário e o seu valor estimado pelo modelo aplicado a todas as idades (em cima) e apenas às idades  $x \geq 50$  (em baixo) e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita)

Num segundo momento, procedemos ao cálculo dos resíduos estandardizados (ou de deviance),  $r_{x,t}^D$ , definidos pela equação (6.9.17). A Figura 7.3.6 representa a evolução temporal dos resíduos  $r_{x,t}^D$  nalgumas idades seleccionadas. Tal como se antevia, é nas idades mais jovens 0-20 anos (sobretudo nos modelos para o sexo masculino) que se observam as estruturas de correlação mais pronunciadas. Nas restantes idades, verifica-se que a maioria dos resíduos não excede os valores de referência relativos aos intervalos de confiança da distribuição Normal. A qualidade do modelo é considerada adequada e a estrutura de correlação apresentada não aparenta ser suficiente para justificar a introdução de variáveis explicativas adicionais (e.g., de um segundo termo log-bilinear).

### 7.3. Modelo de Poisson

---

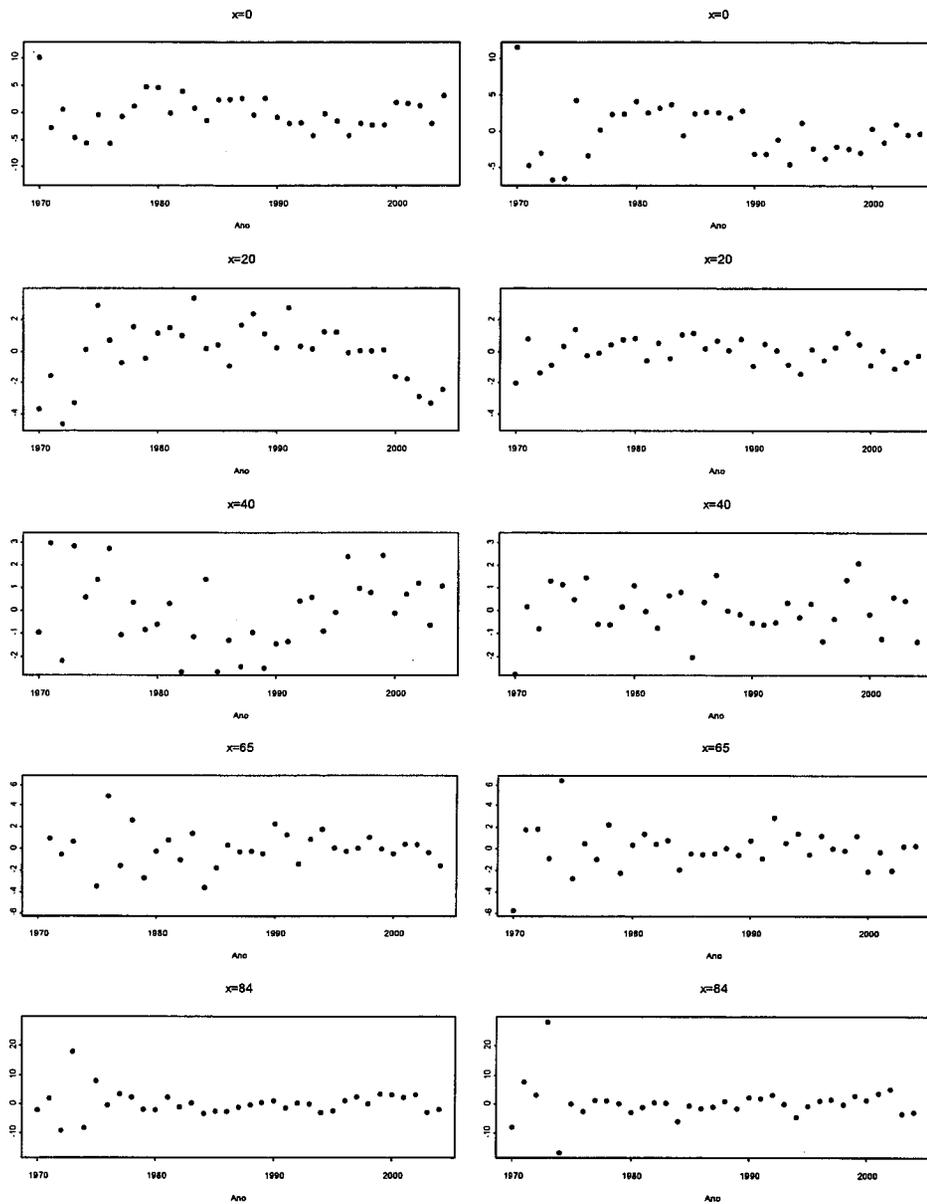


Figura 7.3.6: Resíduos de deviance padronizados do modelo de Poisson (homens à esquerda e mulheres à direita)

Ainda assim, na Secção 7.3.6 confirmamos esta suspeita e testamos uma extensão do modelo de Poisson com duplo preditor log-bilinear.

### 7.3.4 Modelo com incorporação de P-Splines e estimação dos parâmetros por máxima verosimilhança penalizada

Nesta secção, testamos a utilização do modelo de Poisson com inclusão de P-Splines e estimação dos parâmetros por máxima verosimilhança penalizada desenvolvido no capítulo anterior, considerando a amostra relativa à população portuguesa. Este estudo visa averiguar a capacidade do modelo para estabelecer um compromisso entre o grau de ajustamento face aos valores observados e a obtenção de estimativas alisadas para as séries de parâmetros.

Como vimos no capítulo anterior, o modelo impõe uma penalização na função objectivo que incide sobre a diferença finita entre os coeficientes estimados do modelo de Poisson clássico. Esta solução, inspirada na técnica de P-Splines, envolve a definição, num primeiro momento, da ordem da penalização a introduzir. No nosso caso, foram testadas diferentes versões do modelo, optando-se finalmente por considerar um valor  $d = 2$ , ou seja, por penalizar as diferenças de segunda ordem entre as estimativas dos parâmetros. Esta escolha revelou-se suficiente para obter séries graduadas.

A segunda etapa na aplicação do modelo envolveu a selecção do valor dos parâmetros de penalização  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$ . Conforme referimos no Capítulo 6, o valor destes parâmetros pode ser escolhido de forma subjectiva ou, em alternativa, resultar de um procedimento automático, nomeadamente do procedimento de validação cruzada com minimização do quadrado dos erros de previsão usando a definição de resíduos de deviance. O cálculo dos resíduos de deviance  $r_{x,t}^D$  para uso neste contexto implica a supressão da observação correspondente à idade  $x$  e ano de calendário  $t$ .

Em termos computacionais, o problema que é necessário resolver é complexo e extremamente exigente. Por exemplo, se considerarmos apenas o parâmetro  $\lambda_\beta$ , o procedimento de validação cruzada requer um total de  $(x_{\max} - x_{\min} + 1) \times (t_{\max} - t_{\min} + 1)$  ajustamentos, os quais não podem ser realizados pelo algoritmo (6.9.19) na medida em que em cada iteração um par  $(x, t)$  estará em falta.

Para contornar este problema, optámos por recorrer a uma solução simples baseada num procedimento de tentativa e erro, que envolve o cálculo do critério BIC para um número limitado de valores pré-definidos de  $\lambda_\alpha$ ,  $\lambda_\beta$  e  $\lambda_\kappa$ . Os estimadores  $\hat{\lambda}_\alpha$ ,  $\hat{\lambda}_\beta$  e  $\hat{\lambda}_\kappa$  corresponderão à combinação de parâmetros de penalização que minimiza este critério.

Para ilustrar a aplicação deste método, considerámos uma versão restrita do modelo em que a penalização incide unicamente sobre as diferenças finitas dos coeficientes  $\hat{\beta}_x$ , o que pressupõe que  $\lambda_\alpha = \lambda_\kappa = 0$ . Esta opção justifica-se, como vimos na secção anterior, pela maior irregularidade detectada neste parâmetro e pelo facto de  $\kappa_t$  ser objecto de

### 7.3. Modelo de Poisson

alisamento no âmbito do procedimento de extrapolação da tendência temporal.

Para iniciar o algoritmo, considerámos o seguinte vector de valores iniciais para os parâmetros:  $(\hat{\alpha}_x^{(0)}, \hat{\kappa}_t^{(0)}, \hat{\beta}_x^{(0)}) = (0, 1, 0.1)$ . A condição de paragem do algoritmo envolveu a fixação de um valor mínimo  $\varepsilon = 0.0001$  para a variação absoluta na função de log-verosimilhança penalizada  $\mathcal{L}_P(\alpha, \beta, \kappa)$ . A estimação dos parâmetros e o cálculo dos indicadores de qualidade do ajustamento foram obtidos com base numa rotina informática especialmente escrita para o efeito, implementada no *package* SAS (Release 9.1).

Nas Figuras 7.3.7 e 7.3.9 representamos as estimativas de  $\beta_x$  obtidas considerando um conjunto seleccionado de valores para o parâmetro de penalização  $\lambda_\beta$ . Como se observa, quanto maior for o valor do parâmetro de penalização maior é o grau de alisamento da série. O caso em que  $\lambda_\beta = 0$  traduz, naturalmente, as estimativas geradas pelo método de máxima verosimilhança clássico.

A introdução de uma penalização sobre as diferenças finitas dos parâmetros  $\beta_x$  tem obviamente consequências sobre os estimadores  $\hat{\alpha}_x$  e  $\hat{\kappa}_t$ , como se comprova pela observação das Figuras 7.3.8 e 7.3.10. Se no caso do estimador  $\hat{\alpha}_x$  a introdução de uma penalização na função objectivo exerce efeitos negligenciáveis, no caso do estimador do índice temporal  $\hat{\kappa}_t$  o impacto é um pouco mais notório e de certo modo expectável, atendendo à natureza multiplicativa deste em relação a  $\beta_x$ .

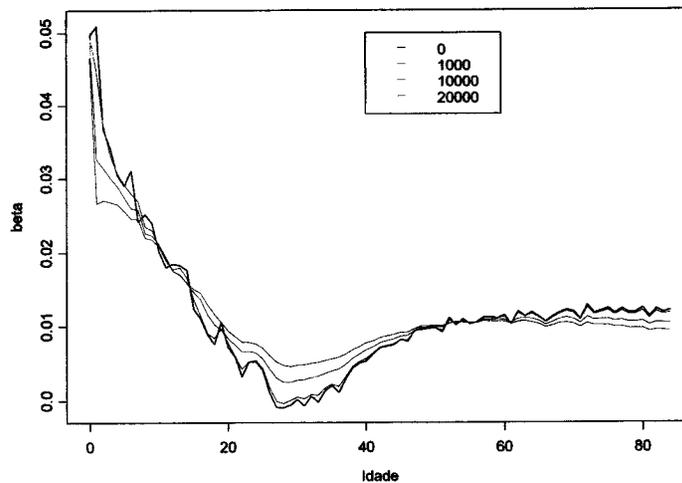


Figura 7.3.7: Estimativas do parâmetro  $\beta_x$  pelo método de máxima verosimilhança penalizada para valores seleccionados de  $\lambda_\beta$  (homens)

### 7.3. Modelo de Poisson

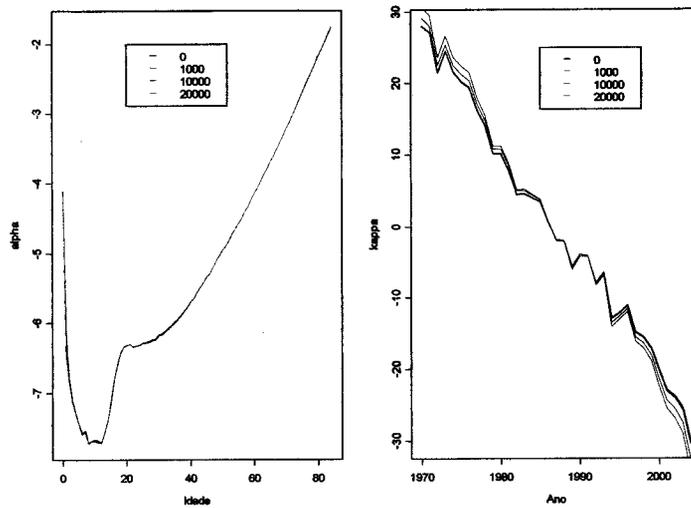


Figura 7.3.8: Estimativas dos parâmetros  $\alpha_x$  e  $\kappa_t$  pelo método de máxima verosimilhança penalizada para valores seleccionados de  $\lambda_\beta$  (homens)

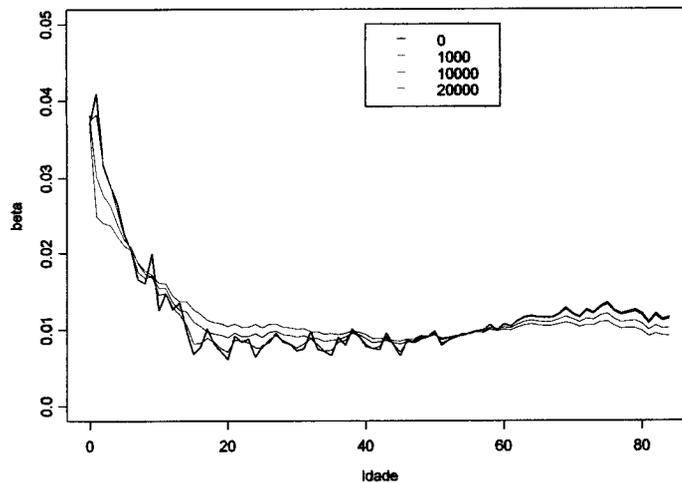


Figura 7.3.9: Estimativas do parâmetro  $\beta_x$  pelo método de máxima verosimilhança penalizada para valores seleccionados de  $\lambda_\beta$  (mulheres)

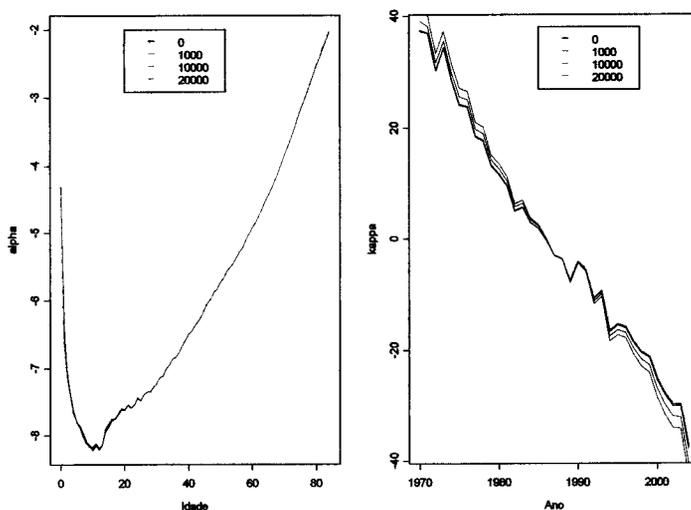


Figura 7.3.10: Estimativas dos parâmetros  $\alpha_x$  e  $\kappa_t$  pelo método de máxima verosimilhança penalizada para valores seleccionados de  $\lambda_\beta$  (mulheres)

Numa tentativa de determinar aproximadamente o valor óptimo da penalização  $\lambda_\beta$ , reportamos no Quadro 7.3.2 os valores do critério  $BIC(0, \lambda_\beta, 0)$  para um painel de valores arbitrários de  $\lambda_\beta$ . A leitura do quadro permite verificar que o valor óptimo do parâmetro de penalização é aproximado por  $\hat{\lambda}_\beta \approx 10000$ , no caso dos homens, e por  $\hat{\lambda}_\beta \approx 20000$  no caso das mulheres.

| Homens          |                            | Mulheres        |                            |
|-----------------|----------------------------|-----------------|----------------------------|
| $\lambda_\beta$ | $BIC(0, \lambda_\beta, 0)$ | $\lambda_\beta$ | $BIC(0, \lambda_\beta, 0)$ |
| 0               | 17307.10                   | 0               | 16889.22                   |
| 1               | 17307.10                   | 1               | 16889.19                   |
| 10              | 17306.99                   | 10              | 16888.95                   |
| 100             | 17303.65                   | 100             | 16881.20                   |
| 1000            | 17282.22                   | 10000           | 16838.48                   |
| 10000           | 16930.17                   | 20000           | 16786.69                   |
| 20000           | 17074.01                   | 40000           | 16880.16                   |
| 40000           | 17300.53                   | 80000           | 16895.81                   |
| 80000           | 17339.10                   | 100000          | 16975.63                   |
| 100000          | 17379.71                   | 200000          | 17237.64                   |

Quadro 7.3.2: Estimativa dos parâmetros de penalização óptimos

Investigações futuras deverão ser capazes de desenvolver um procedimento automático de determinação dos parâmetros óptimos de penalização que seja eficiente e computacionalmente viável, e de fácil incorporação no algoritmo (6.9.19).

#### 7.3.5 Extensão do modelo com inclusão de tábua limite

No Capítulo 6 desenvolvemos uma extensão do modelo de Poisson em que admitimos que as projecções de mortalidade são balizadas por uma determinada tábua limite, para a qual convergem os ganhos de longevidade. Nesta secção, testamos a aplicação desta versão do modelo considerando os dados relativos à população portuguesa.

Um dos aspectos críticos na aplicação de modelos de projecção com tábua limite refere-se à selecção da tábua limite, ou seja, à definição daquilo que são considerados os limites plausíveis para a longevidade humana. A fixação deste limite requer a formulação de hipóteses mais ou menos fundamentadas quanto à evolução futura de um conjunto muito significativo de variáveis biológicas, económicas e sociais. A este respeito, o autor desta dissertação não tem conhecimento de quaisquer estudos científicos que sustentem uma hipótese credível sobre a longevidade máxima da população portuguesa.

Uma solução para contornar o problema passa, como vimos no capítulo anterior, por considerar como tábua limite uma tábua que reflecta o comportamento da mortalidade numa população alvo, residindo a dificuldade neste caso na escolha da população alvo adequada, ou uma lei paramétrica que projecte as grandes tendências na longevidade.

Para ilustrar a aplicação deste método na população portuguesa, optámos por considerar como (hipotética) tábua limite a tábua unisexo proposta por Duchêne e Wunsch (1988, 1991). Esta tábua foi construída com base na hipótese de que a força de mortalidade é bem representada por uma lei de Weibull

$$\mu_x = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \quad (7.3.2)$$

com  $\beta = 95$  e  $\alpha = 14.40198275$ . Admite ainda que a idade limite  $\omega$  da população é igual a 115 anos, que a idade modal é de 95 anos e que, de um total de um milhão de nascimentos, apenas um indivíduo sobreviverá até aos 114 anos. Por fim, os autores admitem ainda, de forma algo arbitrária, que a mortalidade durante o primeiro ano de vida se traduz em dois óbitos por cada 1000 nascimentos.

Da equação (7.3.2) resulta a seguinte função de sobrevivência

$$l_x = l_0 \exp \left[ - \left( \frac{x}{\beta} \right)^\alpha \right] \quad (7.3.3)$$

A Figura 7.3.11 representa a função  $x \mapsto l_x/l_0$  resultante das hipóteses acima assumidas. Como se observa, a tábua limite estabelece uma elevada rectangularização da função de sobrevivência, uma consequência da quase ausência de mortalidade “prematura”.

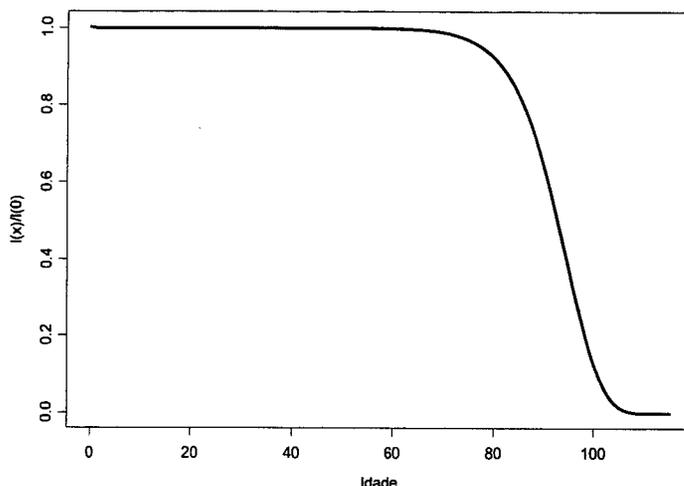


Figura 7.3.11: Tábua limite de Duchêne-Wunsch, função de sobrevivência  $x \mapsto l_x/l_0$

Na Figura 7.3.12 representamos as taxas de mortalidade ditadas pela tábua limite de Duchêne-Wunsch e as estimativas brutas de  $\mu_{x,t}$  para a população portuguesa em anos seleccionados  $t = 1970, 1985, 2004$ . Como se observa, se exceptuarmos a idade  $x = 0$ , a tábua limite estipula taxas mínimas nulas até por volta dos 40 anos, uma hipótese algo irrealista se considerarmos a probabilidade de ocorrerem acidentes.

As estimativas dos parâmetros do modelo (6.9.28)-(6.9.29)-(6.9.30) com inclusão da tábua limite de Duchêne-Wunsch foram obtidas com base no algoritmo (6.9.32) e através de uma rotina informática escrita e implementada no *package* SAS (Release 9.1). Para iniciar o algoritmo, considerámos o seguinte vector de valores iniciais para os parâmetros:  $(\hat{\alpha}_x^{(0)}, \hat{\kappa}_t^{(0)}, \hat{\beta}_x^{(0)}) = (0, 1, 0.1)$ . Os resultados da estimação são representados nas Figuras 7.3.14 e 7.3.13.

Como se observa, as principais diferenças em relação aos estimadores  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  gerados pelo modelo de Poisson clássico situam-se nas idades para as quais a tábua limite prevê patamares significativamente não nulos, ou seja, a partir dos 60 anos de idade. A influência do método no parâmetro  $\hat{\kappa}_t$  é pouco expressiva replicando, no essencial, a tendência decrescente na mortalidade registada no período em análise.

### 7.3. Modelo de Poisson

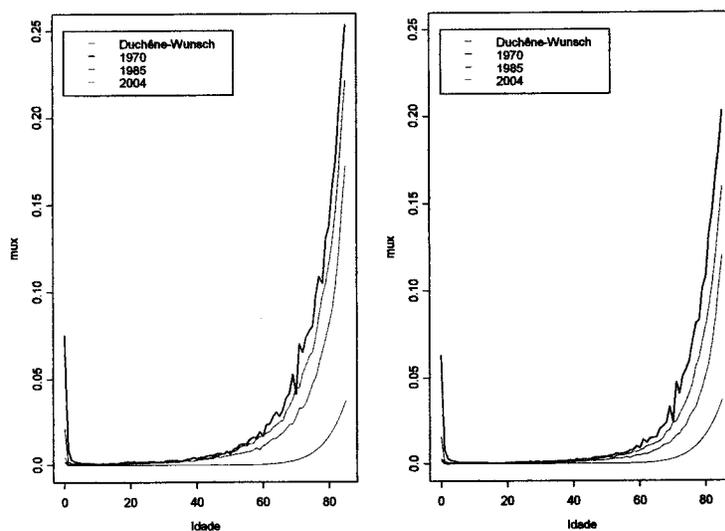


Figura 7.3.12: Comparação entre as taxas  $\mu_{x,t}$  da tábua limite de Duchêne-Wunsch e as estimativas brutas em 1970, 1985 e 2004

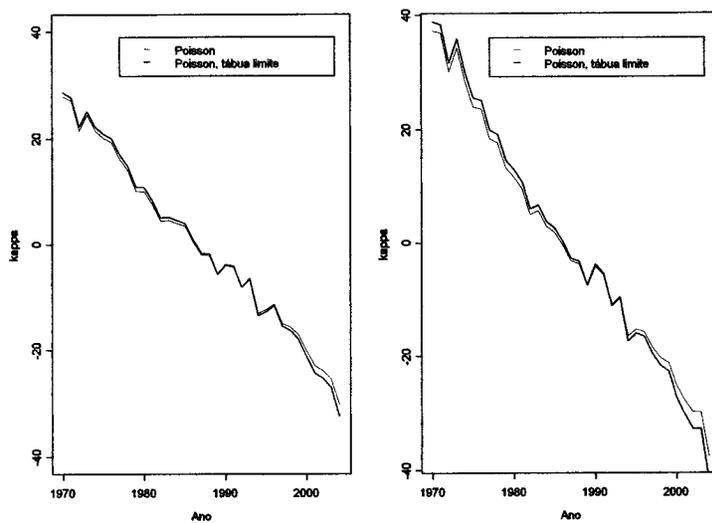


Figura 7.3.13: Estimativas dos parâmetros  $\kappa_t$  do modelo de Poisson com tábua limite de Duchêne-Wunsch (homens à esquerda e mulheres à direita)

### 7.3. Modelo de Poisson

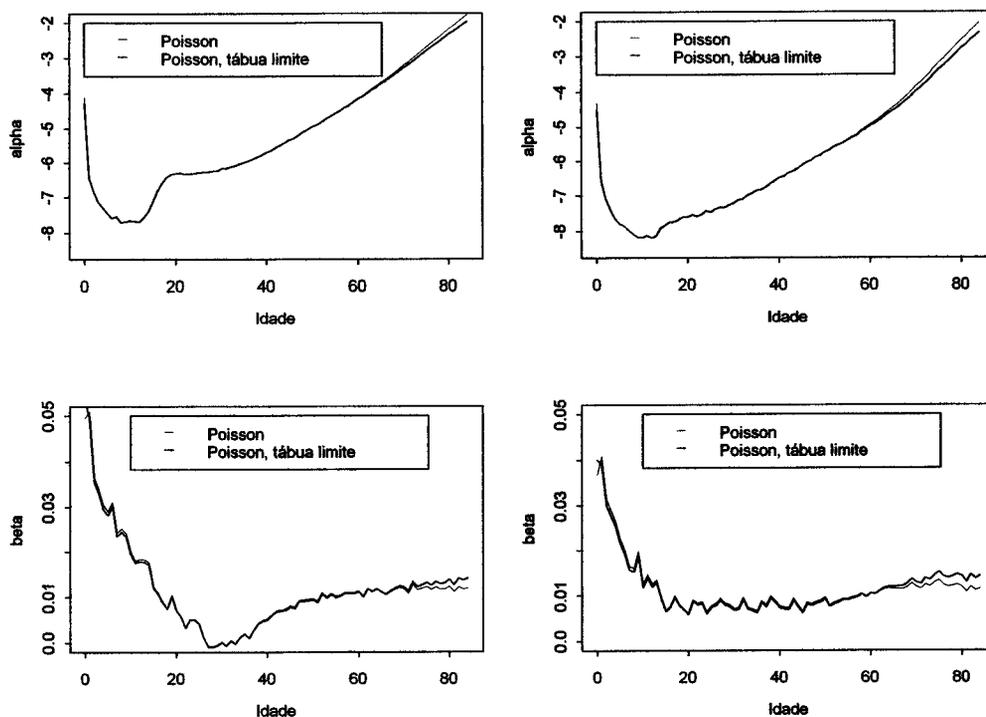


Figura 7.3.14: Estimativas dos parâmetros  $\alpha_x$  e  $\beta_x$  do modelo de Poisson com tábua limite de Duchêne-Wunsch (homens à esquerda e mulheres à direita)

Para termos uma melhor percepção sobre a influência exercida pela incorporação de uma tábua limite nas taxas projectadas pelo modelo de Poisson, representamos na Figura 7.3.15 as estimativas brutas de  $\hat{\mu}_{80,t}$  no intervalo  $t \in [1970, 2004]$  e a trajectória projectada pelos modelos de Poisson clássico e com inclusão de tábua limite. A linha horizontal a tracejado representa o valor limite  $\mu_{80}^{\text{lim}}$  estipulado pela tábua de Duchêne-Wunsch.

Como se constata, a convergência para um nível de mortalidade nulo pressuposta pelo modelo clássico é substituída por uma convergência para a taxa  $\mu_{80}^{\text{lim}}$  pressuposta pela tábua limite, balizando desta forma os potenciais ganhos na longevidade humana. Estudos futuros deverão ser capazes de desenvolver uma tábua limite para a população portuguesa assente, por exemplo, em hipóteses fundamentadas quanto à evolução de um indicador que sintetize a evolução da longevidade humana (e.g., esperança de vida).

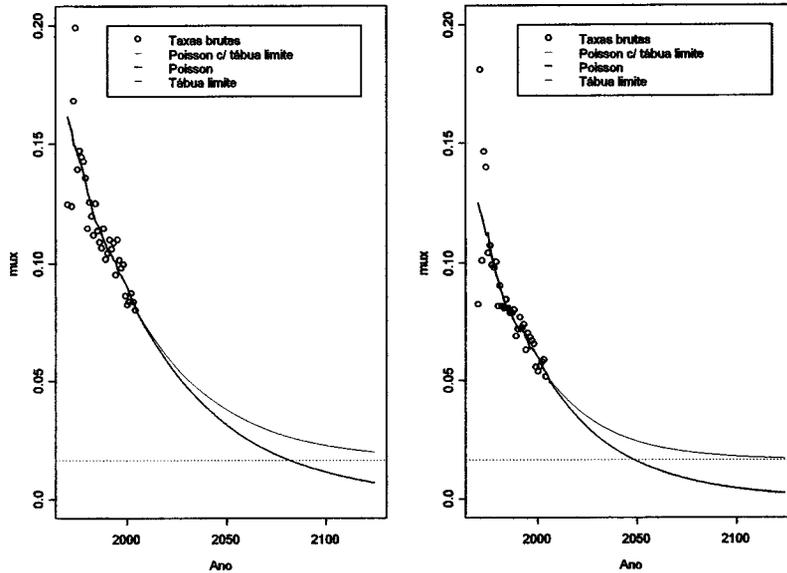


Figura 7.3.15: Trajectória projectada pelo modelo Poisson com tábua limite para a taxa  $\hat{\mu}_{80,t}$  (homens à esquerda e mulheres à direita)

### 7.3.6 Extensão do modelo com duplo previsor log-bilinear

Nesta secção testamos a aplicação da extensão do modelo de Poisson proposta por Renshaw e Haberman (2003c,d) na população portuguesa. Como vimos no capítulo anterior, a extensão materializa-se na inclusão de um segundo termo log-bilinear na especificação do modelo.

As estimativas dos parâmetros do modelo (6.9.22)-(6.9.23)-(6.9.24) foram obtidas com base no algoritmo (6.9.26) e através de uma rotina informática escrita e implementada no *package* SAS (Release 9.1). Para iniciar o algoritmo, considerámos o seguinte vector de valores iniciais:  $(\hat{\alpha}_x^{(0)}, \hat{\kappa}_{t,1}^{(0)}, \hat{\kappa}_{t,2}^{(0)}, \hat{\beta}_{x,1}^{(0)}, \hat{\beta}_{x,2}^{(0)}) = (0, 1, 1, 0.1, 0.1)$ .

Nas Figuras 7.3.16, 7.3.18, 7.3.17 e 7.3.19 representamos as estimativas dos parâmetros  $\hat{\alpha}_x$ ,  $\hat{\beta}_{x,1}$ ,  $\hat{\beta}_{x,2}$ ,  $\hat{\kappa}_{t,1}$  e  $\hat{\kappa}_{t,2}$  do modelo aplicado a todas as idades para ambos os sexos. Como se observa, a introdução de um segundo termo log-bilinear exerce um efeito negligenciável nas estimativas do parâmetro  $\hat{\alpha}_x$ . Pelo contrário, as estimativas dos parâmetros  $\hat{\beta}_{x,1}$  e  $\hat{\kappa}_{t,1}$  evidenciam, na amostra em estudo, algumas diferenças face aos valores  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  gerados pelo modelo de Poisson clássico.

### 7.3. Modelo de Poisson

---

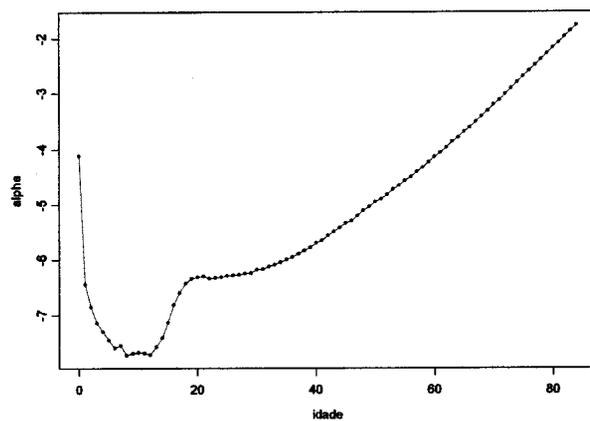


Figura 7.3.16: Estimativas de  $\alpha_x$  no modelo de Poisson com duplo predictor log-bilinear (Homens)

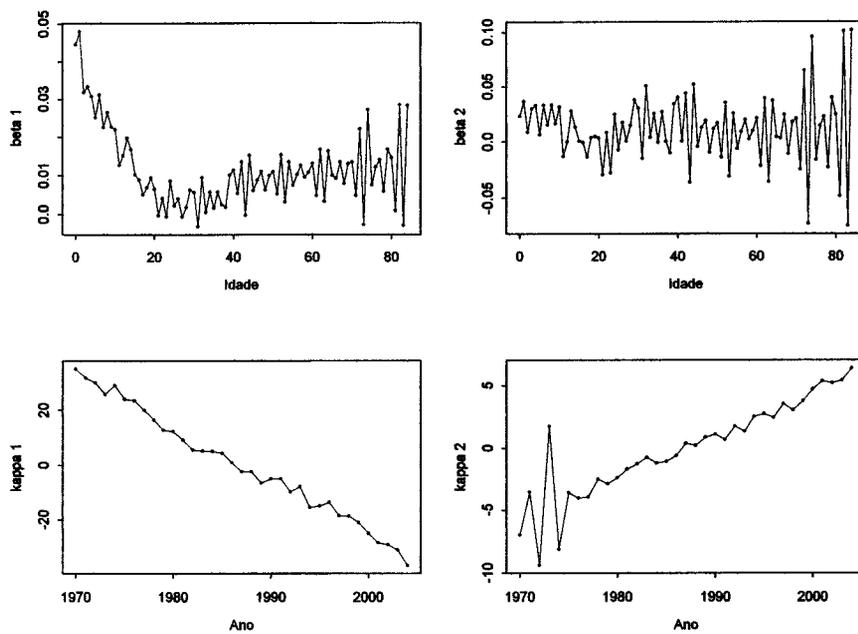


Figura 7.3.17: Estimativas de  $\beta_{x,1}$ ,  $\beta_{x,2}$ ,  $\kappa_{t,1}$  e  $\kappa_{t,2}$  do modelo de Poisson com duplo predictor log-bilinear (Homens)

### 7.3. Modelo de Poisson

---

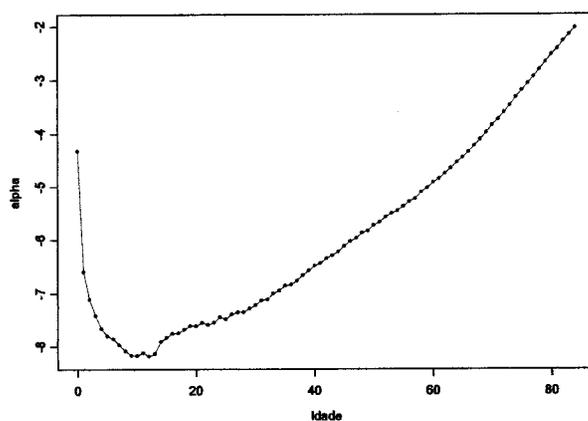


Figura 7.3.18: Estimativas de  $\alpha_x$  no modelo de Poisson com duplo previsor log-bilinear (Mulheres)

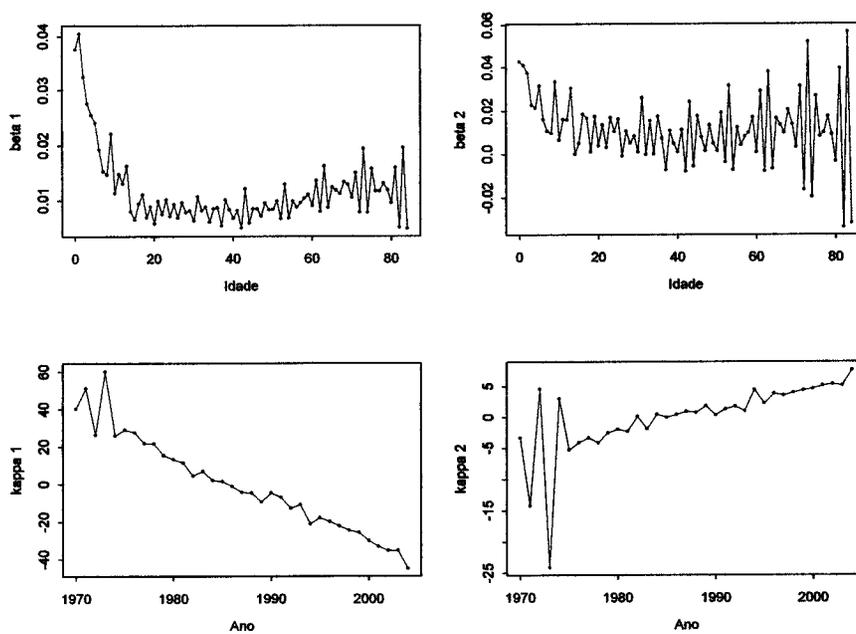


Figura 7.3.19: Estimativas de  $\beta_{x,1}$ ,  $\beta_{x,2}$ ,  $\kappa_{t,1}$  e  $\kappa_{t,2}$  do modelo de Poisson com duplo previsor log-bilinear (Mulheres)

Este resultado não surpreende na medida em que a estimação dos parâmetros no modelo (6.9.22) e no modelo de Poisson clássico é feita, ao contrário do que se passa no método Lee-Carter, de forma independente.<sup>6</sup> Verifica-se ainda que estimativas do segundo parâmetro de sensibilidade  $\hat{\beta}_{x,2}$  evidenciam, em ambos os sexos, um padrão gráfico bastante mais errático do que as correspondentes a  $\hat{\beta}_{x,1}$ , com a volatilidade a aumentar significativamente com a idade.

As estimativas de  $\hat{\kappa}_{t,1}$  replicam, no essencial, a trajetória decrescente detectada nas secções anteriores considerando o valor do índice temporal no modelo original. Um aspecto interessante refere-se, contudo, às tendências antagónicas seguidas por  $\hat{\kappa}_{t,1}$  e  $\hat{\kappa}_{t,2}$  em ambos os sexos, decrescentes no primeiro caso e ascendentes no segundo. Constata-se que o segundo índice temporal  $\hat{\kappa}_{t,2}$  revela, em ambos os sexos, uma tendência irregular e de mais difícil explicação.

Verifica-se, inclusive, que a partir de finais dos anos 80 o parâmetro assume valores positivos, um facto que não nos permite todavia tirar conclusões claras uma vez que o sinal de  $\hat{\beta}_{x,2}$  oscila significativamente entre idades. Como seria de prever, as séries  $\hat{\kappa}_{t,1}$  e  $\hat{\kappa}_{t,2}$  são altamente correlacionadas, como se depreende pelo elevado valor dos coeficientes de correlação (-0.92 e -0.83 para homens e mulheres, respectivamente).

Na Figura 7.3.20 representamos a percentagem da variância explicada pelo modelo de Poisson com um e dois termos log-bilineares, numa tentativa de analisar os ganhos obtidos em termos de qualidade de ajustamento. Como se constata, a introdução de um segundo termo log-bilinear influencia maioritariamente o comportamento do modelo nas idades adultas e mais avançadas, revelando-se incapaz de melhorar significativamente o ajustamento nas idades problemáticas, correspondentes à “bossa” de acidentes, mais importante como vimos no caso dos homens.

Idêntica conclusão pode ser retirada com base na análise da Figura 7.3.21, onde são representados os resíduos de deviance do modelo de Poisson com um e dois termos log-bilineares. Com efeito, observa-se que apenas nas idades adultas a estrutura dos resíduos do modelo com duplo previsor aparenta ser um pouco mais consentânea com o estipulado pela distribuição Normal, quando comparada com aquela proporcionada pelo modelo original.

---

<sup>6</sup> Com efeito, recordamos que na versão original do método Lee-Carter uma única decomposição em valores singulares permite estimar a totalidade dos parâmetros  $\alpha_x$ ,  $\beta_x = \beta_{x,1}$ ,  $\kappa_t = \kappa_{t,1}$ ,  $\beta_{x,2}$  e  $\kappa_{t,2}$ .

### 7.3. Modelo de Poisson

---

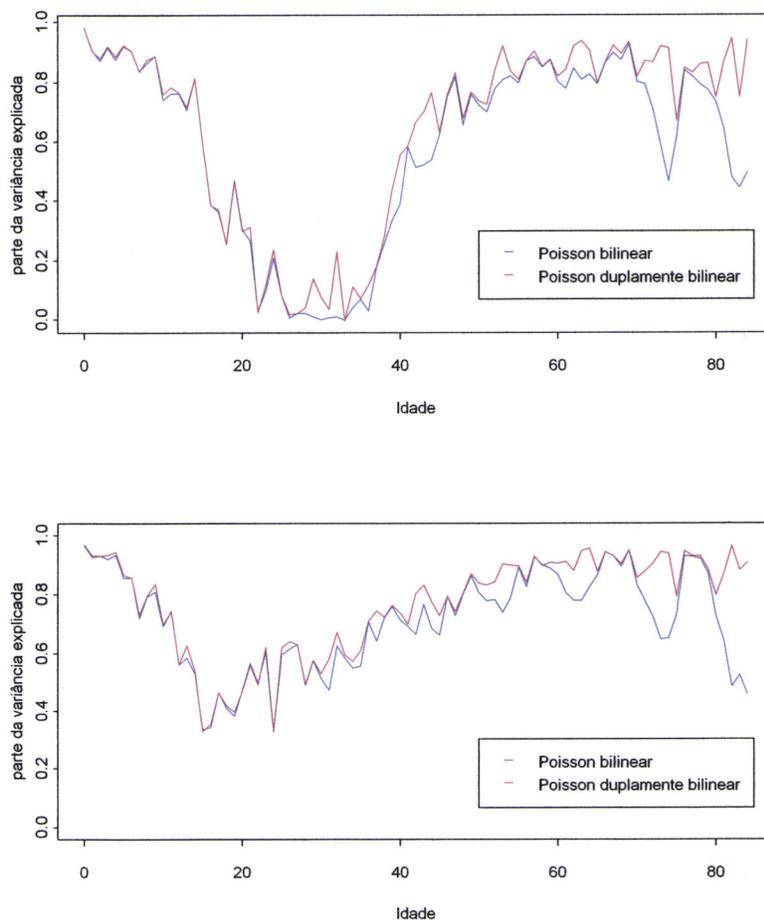


Figura 7.3.20: Percentagem da variância explicada pelo modelo de Poisson com um e dois termos log-bilineares (homens em cima e mulheres em baixo)

A análise dos resultados permite-nos concluir que, na amostra em estudo, o aumento do número de parâmetros decorrente da adição de um segundo termo log-bilinear no modelo de Poisson proporciona uma ligeira melhoria na qualidade do ajustamento. No entanto, a maior irregularidade nas estimativas dos parâmetros e o facto de  $\hat{\kappa}_{t,1}$  e  $\hat{\kappa}_{t,2}$  se constituírem como duas séries temporais fortemente correlacionadas dificulta a sua interpretação e extrapolação para o futuro, tornando bastante mais complexa a tarefa de projectar o comportamento da mortalidade.

### 7.3. Modelo de Poisson

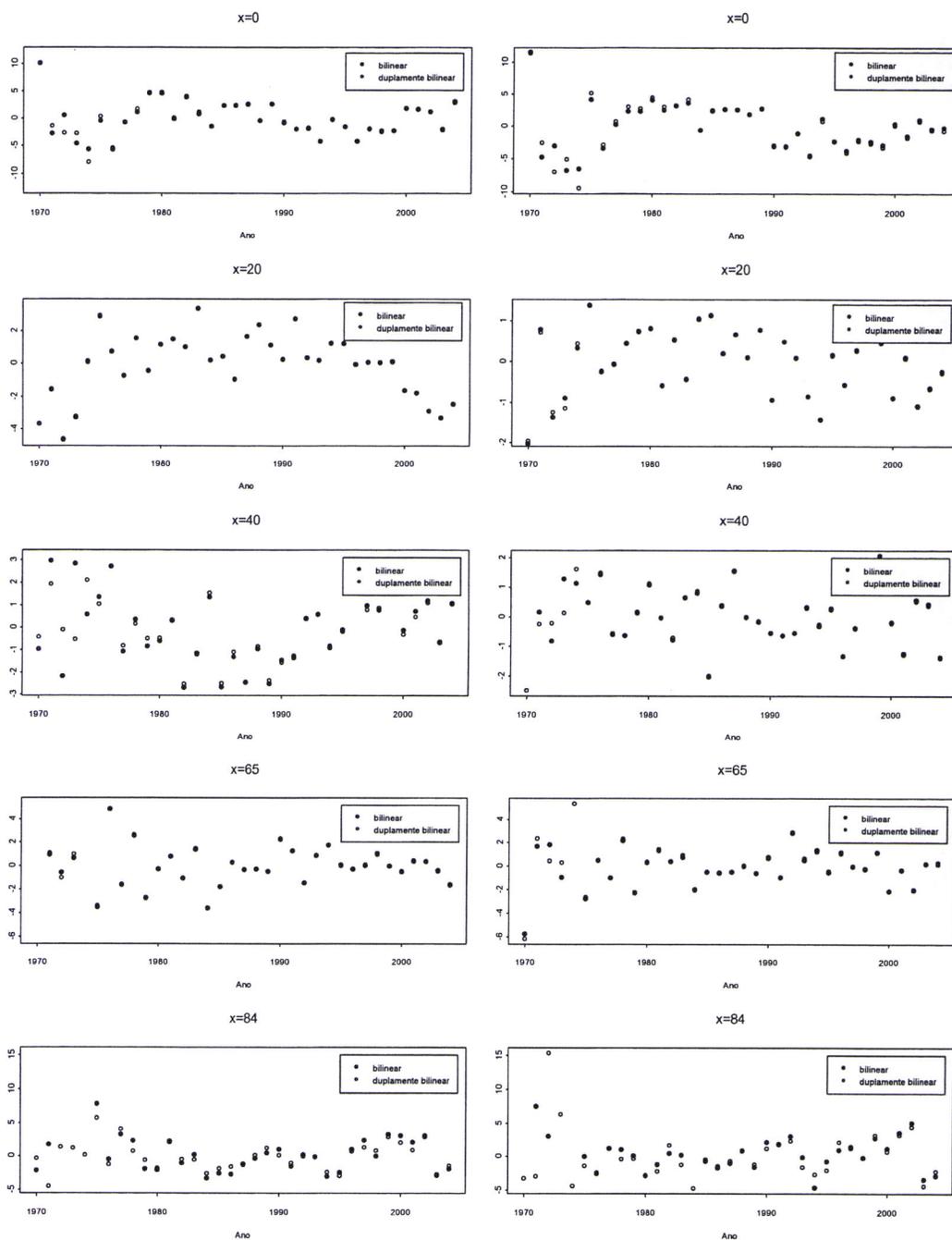


Figura 7.3.21: Resíduos do modelo de Poisson com um e dois termos log-bilineares (homens à esquerda e mulheres à direita)

## 7.4 Método Lee-Carter

Nesta secção, testamos a aplicação do método de Lee-Carter na população portuguesa e comparamos o seu desempenho com o proporcionado pelo modelo de Poisson. Os parâmetros do modelo 6.6.1 foram estimados pelo método dos mínimos quadrados ordinários, adoptando o método iterativo descrito na Secção 6.6.3.

A solução do problema de optimização 6.6.4 respeitou as três etapas preconizadas pelo método Lee-Carter, e assentou numa rotina informática especificamente concebida para o efeito e implementada no *package* SAS (Release 9.1). Na estimação dos parâmetros, considerámos os dados disponíveis para todas as idades e anos de calendário. Para iniciar o algoritmo 6.6.13, usámos o seguinte vector de valores iniciais  $(\hat{\alpha}_x^{(0)}, \hat{\kappa}_t^{(0)}, \hat{\beta}_x^{(0)}) = (0, 1, 0.1)$ .

Na Figura 7.4.1 efectuamos uma comparação entre as estimativas dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  obtidas pelos métodos Lee-Carter e de Poisson, aplicados a todas as idades. Saliente-se que os valores relativos ao parâmetro  $\kappa_t$  respeitam às estimativas iniciais, ou seja, antes da reparametrização destinada a respeitar as restrições do modelo.

Como se observa, as diferenças face ao modelo de Poisson são pouco significativas, em especial no que respeita ao parâmetro  $\alpha_x$ . De seguida, e em conformidade com a Etapa 3 do método Lee-Carter, reajustámos os  $\hat{\kappa}_t$  gerados pelo algoritmo de modo a garantir que o número total de óbitos observado em cada ano iguala o seu número previsto. Recorremos, para tal, a um algoritmo (6.6.12) baseado no método iterativo de Newton-Raphson. O impacto desta reestimação é visível na Figura 7.4.2.

A Figura 7.4.3 dá-nos uma ideia da capacidade de ajustamento de ambos os modelos. Como se observa, a percentagem da variância explicada pelos modelos é praticamente indistinguível na generalidade das idades, embora se constate que o modelo de Poisson apresenta um comportamento ligeiramente melhor nas idades mais jovens. Esta ténue superioridade do modelo de Poisson é confirmada pelos valores obtidos para a estatística deviance, que ascendem, no método Lee-Carter, a 16609.27 e 15650.48 para homens e mulheres, respectivamente, valores que contrastam com 15667.51 e 15249.63 calculados para o modelo de Poisson.

A análise do comportamento dos resíduos nalgumas idades seleccionadas não revela, como se observa pela leitura da Figura 7.4.4, diferenças substanciais em relação ao que afirmámos anteriormente a propósito do modelo de Poisson.

Em síntese, se somarmos as vantagens teóricas do modelo de Poisson clássico relativamente ao método Lee-Carter aos resultados obtidos em termos de grau de ajustamento e à maior facilidade de implementação, inclinamo-nos a favor da escolha do primeiro.

## 7.4. Método Lee-Carter

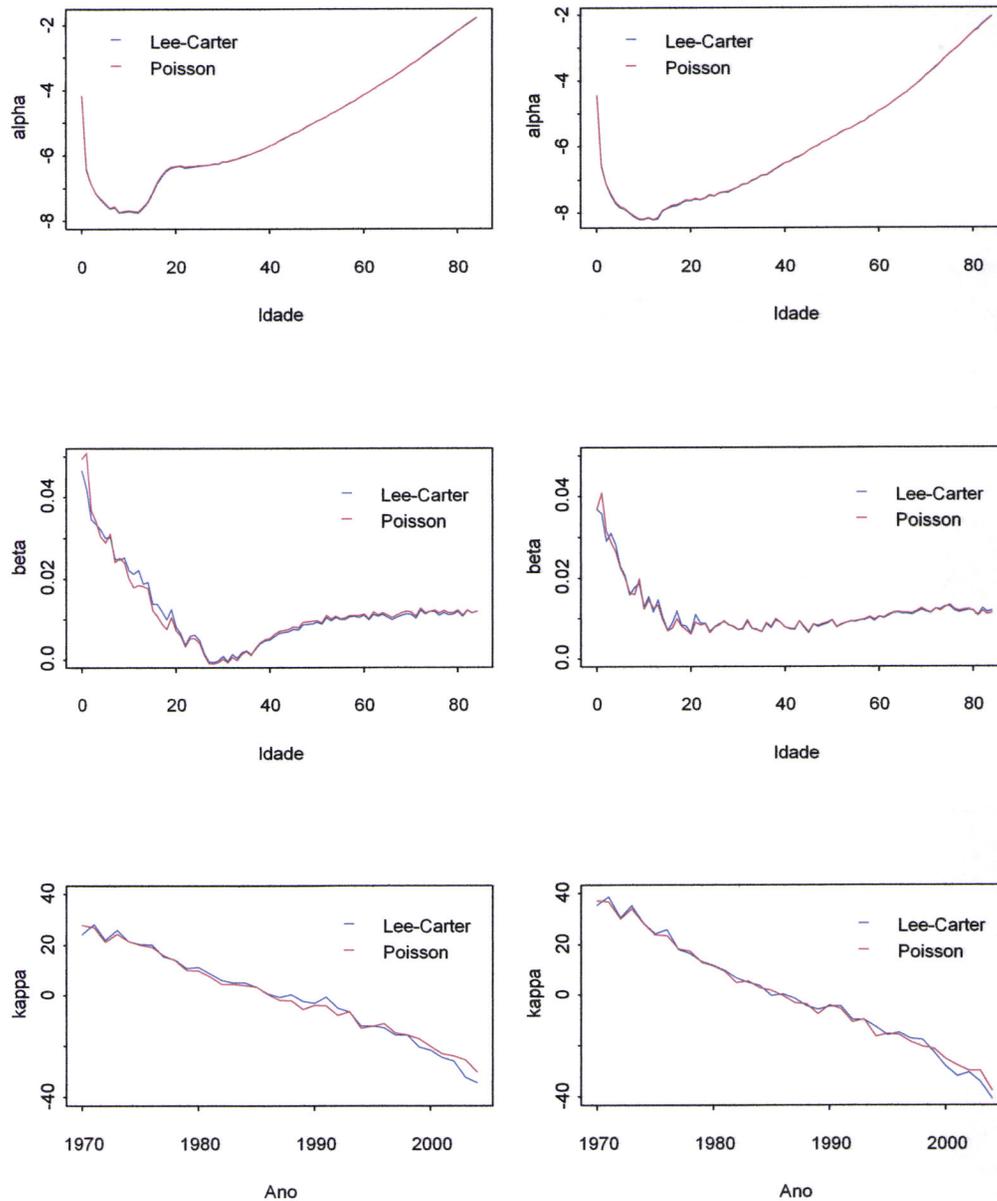


Figura 7.4.1: Estimativas dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  obtidas pelos métodos Lee-Carter e de Poisson aplicados a todas as idades (homens à esquerda e mulheres à direita)

## 7.4. Método Lee-Carter

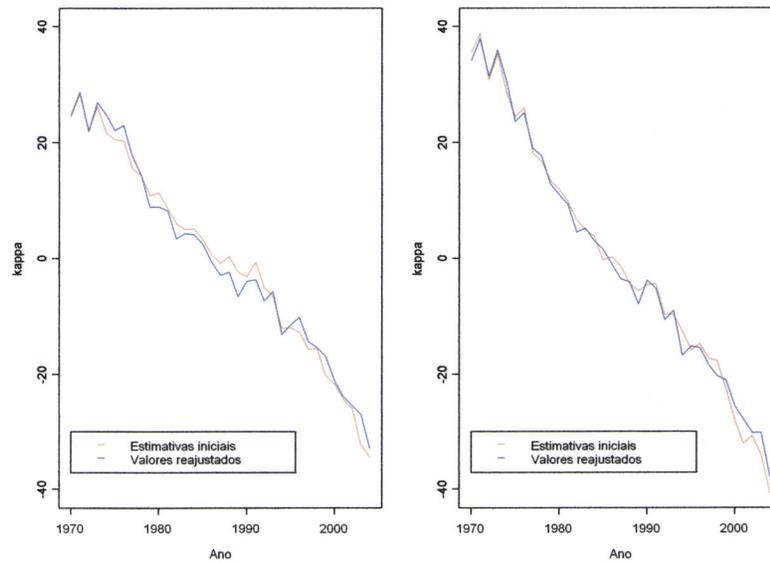


Figura 7.4.2: Reestimação do parâmetro  $\kappa_t$  de modo a satisfazer a restrição (6.6.11)

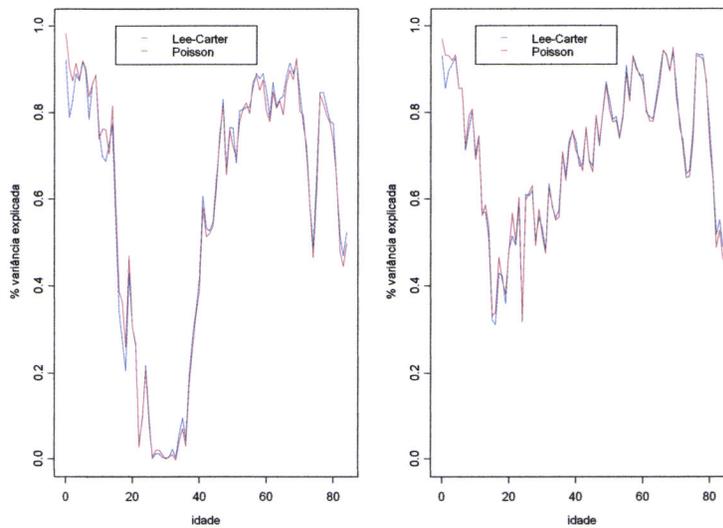


Figura 7.4.3: Percentagem da variância explicada pelos modelos de Lee-Carter e de Poisson (homens à esquerda e mulheres à direita)

## 7.4. Método Lee-Carter

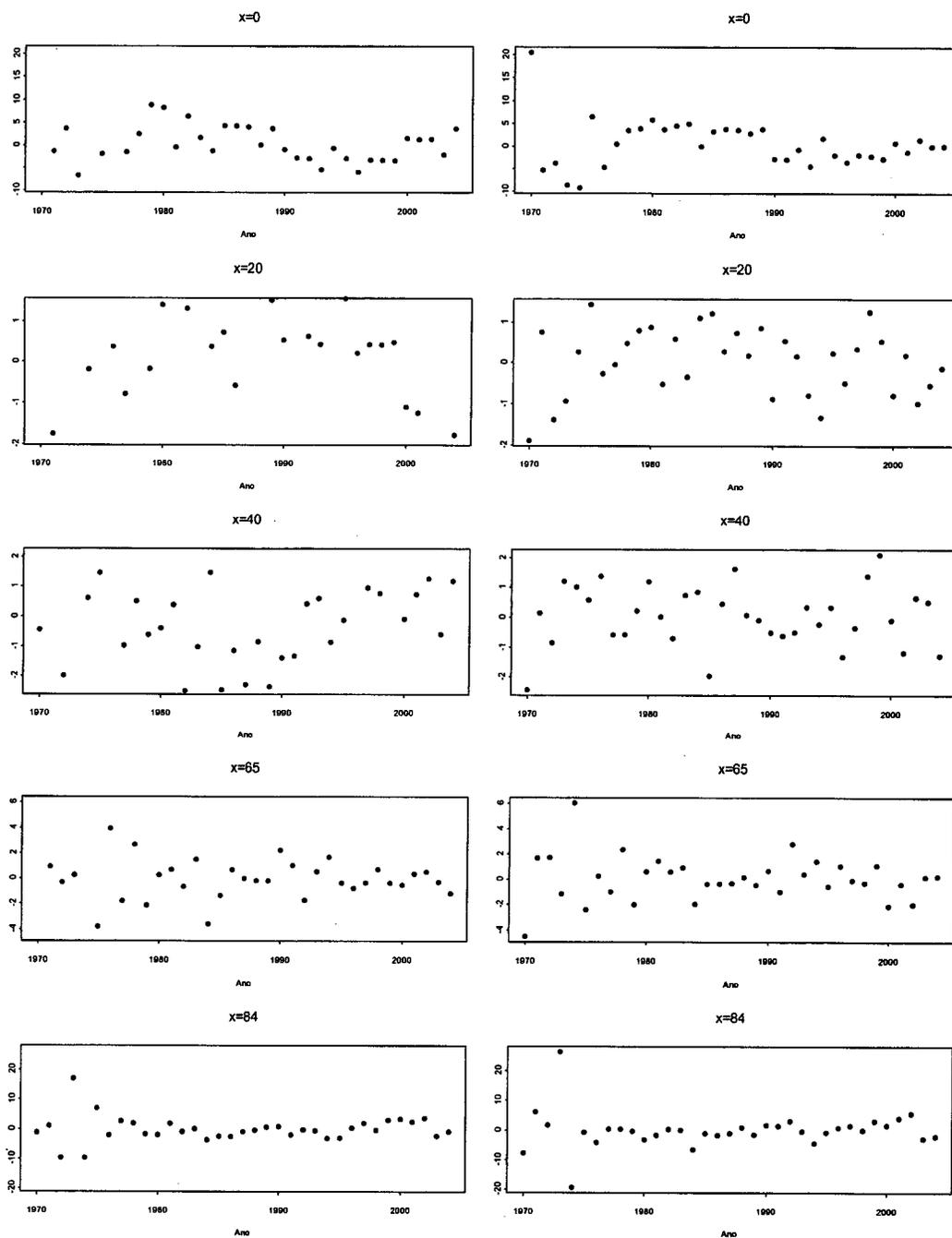


Figura 7.4.4: Resíduos de deviance estandardizados do método Lee-Carter (homens à esquerda e mulheres à direita)

## 7.5 Extrapolação da tendência temporal

Os modelos de Poisson e de Lee-Carter resumem toda a informação sobre a evolução da mortalidade no tempo no índice temporal  $\kappa_t$ . Neste sentido, a obtenção de projecções para os indicadores de mortalidade e, em particular, a construção de tábuas prospectivas passa, em primeiro lugar, pela modelação do comportamento desta série temporal com vista à sua extrapolação para o futuro (i.e., para  $t \geq t_{\max} + 1$ ).

O método de projecção por extrapolação possui, como é bem sabido, as vantagens e as desvantagens da objectividade. O seu pragmatismo reside no facto de procurar projectar no futuro um conjunto de tendências passadas detectadas na mortalidade, assumindo que os factores que ditaram essas tendências continuarão a marcar a sua evolução futura. Por esta razão, o método desvaloriza, em boa medida, a importância da opinião de especialistas sobre a evolução previsível de um conjunto de variáveis que condicionam, decisivamente, os ganhos futuros na longevidade (e.g., progressos da medicina, surgimento de novas doenças, evolução dos estilos de vida).

Esta abordagem é obviamente susceptível de crítica, entre outras razões pela sua incapacidade de prever choques estruturais (e.g., descoberta de um novo medicamento, terramoto). Pese embora estas limitações, o facto da mortalidade humana ser o resultado de uma interacção complexa de múltiplos factores sociais e biológicos, bastante difíceis de esmiuçar, faz com que a modelação por extrapolação constitua, ainda hoje, a técnica standard a nível internacional.

Seja  $\{\kappa_t, t = t_{\min}, \dots, t_{\max}\}$  uma realização de dimensão finita da série cronológica  $\mathcal{K} = \{\kappa_t, t \in \mathbb{N}\}$ . Para modelar o comportamento da série, recorreremos a modelos do tipo ARIMA e adoptamos a convencional metodologia de Box-Jenkins. Recordemos que a expressão geral de um modelo ARIMA é dada por<sup>7</sup>

$$(1 - B)^d \kappa_t = \mu + \frac{\Theta_q(B) \epsilon_t}{\Phi_q(B)},$$

onde  $B$  representa o operador de desfasamento (i.e.,  $B(\kappa_t) = \kappa_{t-1}$ ,  $B^2(\kappa_t) = \kappa_{t-2}, \dots$ ),  $1 - B$  é o operador de diferenças (i.e.,  $(1 - B)\kappa_t = \kappa_t - \kappa_{t-1}$ ,  $(1 - B)^2 \kappa_t = \kappa_t - 2\kappa_{t-1} + \kappa_{t-2}, \dots$ ),  $\Theta_q(B)$  é o polinómio relativo à componente média móvel, com coeficientes  $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ ,  $\Phi_q(B)$  é o polinómio relativo à componente autoregressiva, com coeficientes  $\phi = (\phi_1, \phi_2, \dots, \phi_p)$ , e  $\epsilon_t$  é um ruído branco com variância  $\sigma_\epsilon^2$ .

A modelação paramétrica de uma série temporal assenta essencialmente na estrutura de dependência existente entre os elementos que a compõem. Nesta secção apresentamos,

<sup>7</sup>Para uma descrição detalhada da metodologia de Box-Jenkins veja-se, por exemplo, Hamilton (1994).

de forma detalhada, os resultados obtidos pela aplicação da metodologia de Box-Jenkins à série de parâmetros  $\{\hat{\kappa}_t\}$  estimada pelo modelo de Poisson clássico aplicado a todas as idades. Um quadro recapitulativo final resume os modelos ARIMA seleccionados nas demais séries.

### 7.5.1 Estacionaridade da série

O método proposto por Box-Jenkins aplica-se a processos estocásticos estacionários em covariância (ou na forma fraca). Tal pressupõe, para além da existência dos dois primeiros momentos, a constância da média ao longo do tempo e funções de autocovariância e de auto-correlação que dependem apenas do desfasamento temporal (*lag*) entre as observações e não do momento (da data) da observação.

A primeira etapa da modelação consiste, assim, em assegurar que  $\mathcal{K}$  satisfaz adequadamente estas condições. Caso contrário, torna-se necessário introduzir alguma transformação (e.g., por diferenciação) que nos permita atingir a desejada estacionaridade. Existem dois métodos principais para apreciar a estacionaridade da série: (i) avaliação subjectiva com base na análise do gráfico da série temporal e do seu respectivo correlograma; (ii) testes estatísticos formais para a detecção de raízes unitárias.

A primeira condição necessária ao ajustamento de um modelo ARIMA é a estabilidade em variância da série. Tal exige que a variabilidade seja homogênea ao longo do período de observação. Na Figura 7.5.1 é possível constatar que a variabilidade é relativamente estável no período temporal em análise, pelo que à primeira vista não se afigura necessária qualquer transformação.

A metodologia de Box-Jenkins requer igualmente que as séries sejam estacionárias em média, ou seja, destituídas de tendência temporal. Na Figura 7.5.1 resulta claro que as séries anuais são claramente decrescentes, logo não estacionárias, mas não apresentam sinais de sazonalidade. Neste sentido, devemos aplicar uma primeira diferenciação, tanto no caso dos homens como no caso das mulheres. A aplicação desta técnica implica, naturalmente, a perda de uma observação. As séries diferenciadas  $\Delta\kappa_t$  assim obtidas são, como se observa na Figura 7.5.1, aparentemente estacionárias em média.

Para confirmar este diagnóstico preliminar, realizamos de seguida um conjunto de testes estatísticos formais (teste ADF e teste de Phillips-Perron) destinados a avaliar a presença de raízes unitárias e a ordem da estacionaridade da série. Recordamos que o teste ADF (*Augmented Dickey Fuller*) avalia a hipótese nula de que a série  $\Delta\kappa_t$  é  $I(0)$ , contra a alternativa de que é  $I(1)$ , assumindo que a dinâmica da série pode ser representada por uma estrutura do tipo ARMA( $p, q$ ) com ordem desconhecida. O teste

## 7.5. Extrapolação da tendência temporal

ADF baseia-se na seguinte equação de regressão

$$\Delta\kappa_t = \alpha + \gamma t + \pi\kappa_{t-1} + \sum_{j=1}^p \psi_j \Delta\kappa_{t-j} + \varepsilon_t, \quad (7.5.1)$$

onde  $\alpha$ ,  $\gamma$ ,  $\pi$  e  $\psi_j$  ( $j = 1, \dots, p$ ) são parâmetros a estimar. Os  $p$  termos (diferenças) desfasados,  $\Delta\kappa_{t-j}$ , são usados para aproximar a estrutura do tipo ARMA do termo erro, sendo que o valor de  $p$  é fixado de modo a garantir que o termo erro não exibe autocorrelação. Segundo a hipótese nula,  $\Delta\kappa_t$  é  $I(0)$ , o que implica que  $\pi = 0$ .

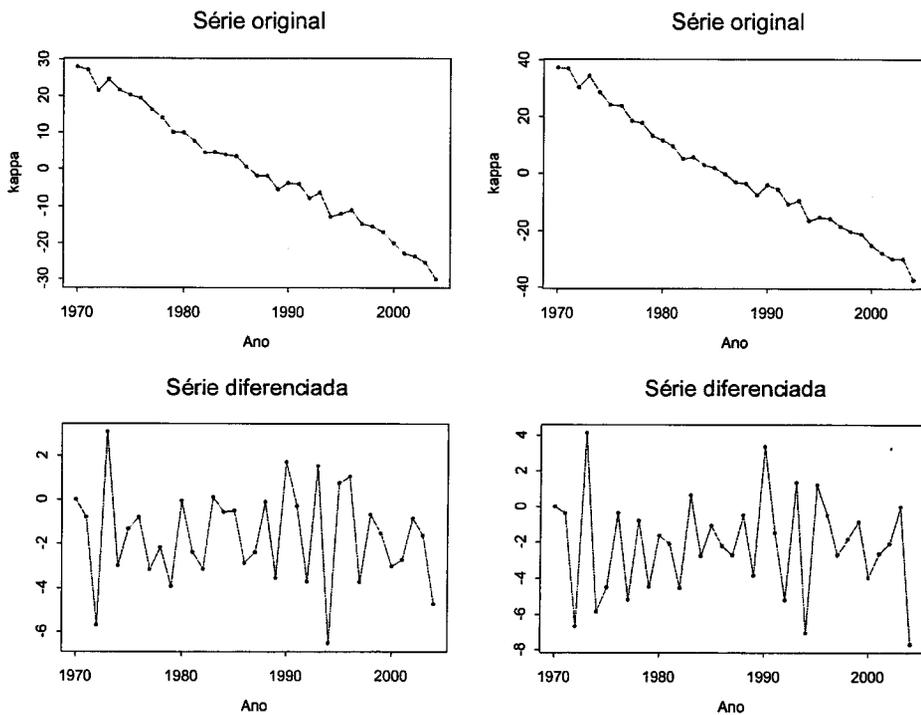


Figura 7.5.1: Séries iniciais  $\hat{\kappa}_t$  e primeiras diferenças  $(1 - B)\hat{\kappa}_t$  (homens à esquerda e mulheres à direita)

Uma solução alternativa para averiguar a presença de raízes unitárias consiste em usar o teste de Phillips-Perron (PP), que difere do teste ADF essencialmente pelo facto de ignorar a eventual presença de autocorrelação no termo erro da equação de regressão.

## 7.5. Extrapolação da tendência temporal

A equação base do teste é a seguinte

$$\Delta\kappa_t = \alpha + \gamma t + \pi\kappa_{t-1} + u_t, \quad (7.5.2)$$

onde  $u_t$  é  $I(0)$  e pode apresentar heterocedasticidade. O teste PP considera a eventual presença de autocorrelação e heterocedasticidade no termo erro  $u_t$  modificando diretamente a estatística de teste. O teste de PP é considerado mais robusto que o teste ADF porque admite formas mais gerais de heterocedasticidade no termo erro  $u_t$ . Outra vantagem é a de que neste caso não é necessário especificar a dimensão do desfasamento (parâmetro  $p$ ) na equação de regressão.

Os Quadros 7.5.1 e 7.5.2 resumem os resultados obtidos pela aplicação dos testes PP e ADF a ambas as séries. A análise dos valores das estatísticas de teste e dos valores de prova ( $p$ -values) permite-nos rejeitar a hipótese nula de que a raiz é unitária e as séries são  $I(0)$ , em favor da alternativa de que as séries  $\{\kappa_t\}$  são  $I(1)$ , i.e., estacionárias em (primeiras) diferenças.

| <b>Homens</b>  |           | Estatística teste | Valores críticos Dickey-Fuller |            |            |
|--|-----------|-------------------|--------------------------------|------------|------------|
|  |           |                   | 1%                             | 5%         | 10%        |
| $Z(\rho)$  |           | -28.294           | -23.652                        | -18.584    | -16.032    |
| $Z(\tau)$  |           | -4.192            | -4.297                         | -3.564     | -3.218     |
| * MacKinnon approximate $p$ -value for $Z(t) = 0.0046$ |           |                   |                                |            |            |
| Termo  | Parâmetro | Estimativa        | Std. error                     | $t$ -value | $p$ -value |
| constante  | $\alpha$  | 19.24462          | 5.212382                       | 3.69       | 0.001      |
| tendência  | $\gamma$  | -1.160177         | 0.288109                       | -4.03      | 0.000      |
| $\kappa_{t-1}$   | $\pi$     | 0.2653236         | 0.183177                       | 1.45       | 0.158      |

| <b>Mulheres</b>  |           | Estatística teste | Valores críticos Dickey-Fuller |            |            |
|--|-----------|-------------------|--------------------------------|------------|------------|
|  |           |                   | 1%                             | 5%         | 10%        |
| $Z(\rho)$  |           | -19.415           | -23.652                        | -18.584    | -16.032    |
| $Z(\tau)$  |           | -3.475            | -4.297                         | -3.564     | -3.218     |
| * MacKinnon approximate $p$ -value for $Z(t) = 0.0422$ |           |                   |                                |            |            |
| Termo  | Parâmetro | Estimativa        | Std. error                     | $t$ -value | $p$ -value |
| constante  | $\alpha$  | 17.91657          | 5.899336                       | 3.04       | 0.005      |
| tendência  | $\gamma$  | -1.11439          | 0.326156                       | -3.42      | 0.002      |
| $\kappa_{t-1}$   | $\pi$     | 0.4434802         | 0.161659                       | 2.74       | 0.010      |

Quadro 7.5.1: Testes de Phillips-Perron para a existência de raízes unitárias

## 7.5. Extrapolação da tendência temporal

| Homens                                     |           | Estatística |            | Valores críticos DF |            |        |
|--|-----------|-------------|------------|---------------------|------------|--------|
|  |           | teste       |            | 1%                  | 5%         | 10%    |
| $Z(t)$                                     |           | -4.011      |            | -4.297              | -3.564     | -3.218 |
| * MacKinnon $p$ -value for $Z(t) = 0.0042$ |           |             |            |                     |            |        |
| Termo                                      | Parâmetro | Estimativa  | Std. error | $t$ -value          | $p$ -value |        |
| constante                                  | $\alpha$  | 19.24462    | 5.212382   | 3.69                | 0.001      |        |
| tendência                                  | $\gamma$  | -1.160177   | 0.2881087  | -4.03               | 0.000      |        |
| $\kappa_{t-1}$                             | $\pi$     | -0.7346764  | 0.1831773  | -4.01               | 0.000      |        |

| Mulheres                                   |           | Estatística |            | Valores críticos DF |            |        |
|--|-----------|-------------|------------|---------------------|------------|--------|
|  |           | teste       |            | 1%                  | 5%         | 10%    |
| $Z(t)$                                     |           | -3.443      |            | -4.297              | -3.564     | -3.218 |
| * MacKinnon $p$ -value for $Z(t) = 0.0460$ |           |             |            |                     |            |        |
| Termo                                      | Parâmetro | Estimativa  | Std. error | $t$ -value          | $p$ -value |        |
| constante                                  | $\alpha$  | 17.91657    | 5.899335   | 3.04                | 0.005      |        |
| tendência                                  | $\gamma$  | -1.11439    | 0.326156   | -3.42               | 0.002      |        |
| $\kappa_{t-1}$                             | $\pi$     | -0.5565198  | 0.161659   | -3.44               | 0.002      |        |

Quadro 7.5.2: Testes ADF para a existência de raízes unitárias

Por fim, conduzimos um conjunto de testes à hipótese nula de ruído branco para as séries originais  $\{\kappa_t\}$  e em primeiras diferenças  $\{\Delta_1\kappa_t\}$ . Os resultados decorrentes da aplicação dos testes de Portmanteau e de Bartlett à hipótese nula de que os coeficientes de autocorrelação para diferentes níveis de defasamento são conjuntamente nulos, i.e.,  $H_0 : \rho_1 = \dots = \rho_l = 0$ , contra a alternativa  $H_1 : \rho_i \neq 0$  para algum  $i \in \{1, \dots, m\}$ , são apresentados no Quadro 7.5.3.

|          |                    | Teste Portmanteau |            | Teste Bartlett |            |
|----------|--------------------|-------------------|------------|----------------|------------|
|          |                    | $Q^*$ stat        | $p$ -value | $B$ stat       | $p$ -value |
| Homens   | $\kappa_t$         | 131.9774          | 0.0000     | 2.7040         | 0.0000     |
|          | $\Delta_1\kappa_t$ | 34.5482           | 0.0029     | 1.5596         | 0.0154     |
| Mulheres | $\kappa_t$         | 129.8582          | 0.0000     | 2.6771         | 0.0000     |
|          | $\Delta_1\kappa_t$ | 48.6481           | 0.0000     | 1.8542         | 0.0021     |

Quadro 7.5.3: Testes de Portmanteau e Bartlett à hipótese de ruído branco

A análise dos resultados obtidos nos testes permite concluir que a hipótese de ruído branco para as séries em nível e em primeiras diferenças deve ser rejeitada em ambos os sexos. Verifica-se, assim, que o modelo ARIMA(0, 1, 0), a escolha clássica do método Lee-Carter, não é adequado às séries em estudo, sendo por isso necessário o recurso aos procedimentos de identificação preconizados pela metodologia de Box-Jenkins.

### 7.5.2 Identificação do modelo

Para identificar o modelo adequado socorremo-nos, em primeiro lugar, da inspeção visual dos gráficos das funções de autocorrelação (ACF) e de autocorrelação parcial (PACF) das séries diferenciadas, representadas na Figura 7.5.2. Estas funções traduzem, por definição, a influência entre duas observações em função do desfaseamento que as separa e assumem um papel preponderante na análise da dependência temporal da série.

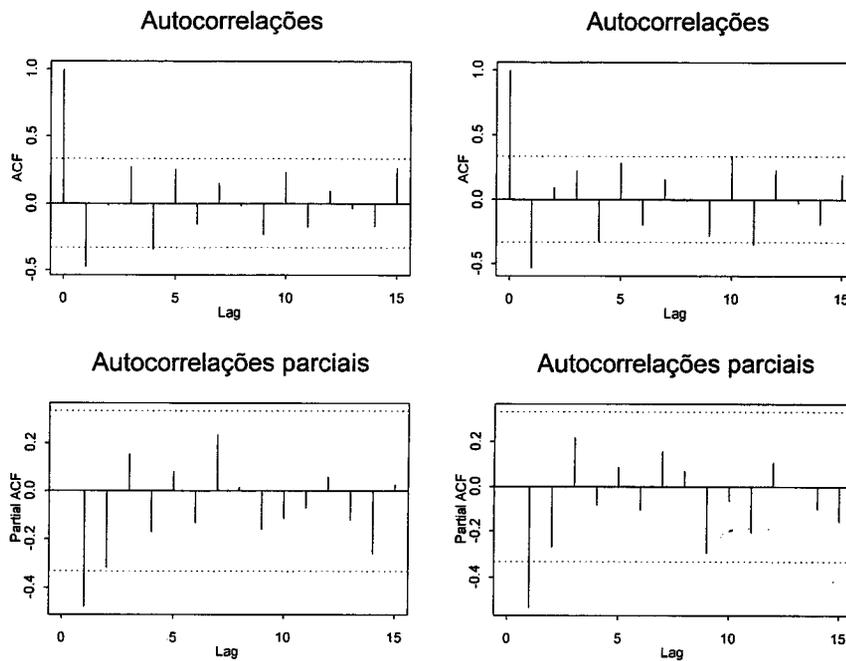


Figura 7.5.2: Coeficientes de autocorrelação e de autocorrelação parcial para as séries diferenciadas  $\Delta_1 \kappa_t$  (homens à esquerda e mulheres à direita)

Como se observa, o decréscimo relativamente rápido dos coeficientes de autocorrelação e de autocorrelação parcial para os limites estabelecidos pelos intervalos de confiança (a tracejado) confirmam a hipótese de estacionaridade da série diferenciada. Por outro lado, o comportamento dos coeficientes face aos respectivos intervalos de confiança deixa antever que um modelo ARIMA com uma estrutura relativamente simples deve ser suficiente para captar a estrutura de dependência da série. Ainda assim, é bem sabido que a ACF e a PACF não são particularmente elucidativas no que se refere à determinação da ordem adequada do modelo ARIMA.

## 7.5. Extrapolação da tendência temporal

Para esta finalidade, a nossa opção recaiu sobre a hierarquização dos modelos candidatos pelos critérios BIC (*Bayesian Information Criterion*) e AIC (*Akaike Information Criterion*). A informação fornecida por estes indicadores foi posteriormente complementada com a análise dos resultados proporcionados pela aplicação dos métodos **ESACF** (*Extended Sample Autocorrelation Function*) e **SCAN** (*Smallest CANonical correlation method*) desenvolvidos por Tsay and Tiao (1984, 1985), e pelo método **MINIC** (*MINimum Information Criterion*) proposto por Hannan and Rissanen (1982), disponíveis no *package* SAS. Por uma questão de parcimónia, restringimos a nossa apreciação a modelos com um número de parâmetros igual ou inferior a 4. Os resultados são apresentados no Quadro 7.5.4.

| Modelo         | Homens     |            | Mulheres   |            |
|----------------|------------|------------|------------|------------|
|                | <i>BIC</i> | <i>AIC</i> | <i>BIC</i> | <i>AIC</i> |
| ARIMA(0, 1, 0) | 150.74     | 149.21     | 168.44     | 166.91     |
| ARIMA(1, 1, 0) | 143.90     | 140.85     | 157.02     | 153.97     |
| ARIMA(0, 1, 1) | 141.49     | 138.43     | 158.46     | 155.40     |
| ARIMA(1, 1, 1) | 144.48     | 139.90     | 158.52     | 153.94     |
| ARIMA(2, 1, 0) | 143.44     | 138.86     | 156.95     | 152.37     |
| ARIMA(0, 1, 2) | 144.49     | 139.91     | 157.89     | 153.31     |
| ARIMA(2, 1, 1) | 144.99     | 138.88     | 158.67     | 152.57     |
| ARMA(1, 1, 2)  | 147.98     | 141.87     | 161.05     | 154.95     |
| ARIMA(3, 1, 0) | 146.41     | 140.30     | 158.92     | 152.82     |
| ARIMA(0, 1, 3) | 147.63     | 141.52     | 160.09     | 153.98     |

Quadro 7.5.4: Identificação da ordem do modelo ARIMA

A análise dos diferentes critérios levou-nos a identificação do modelo ARIMA(0, 1, 1) como o mais adequado para a população masculina, enquanto que no caso da população feminina a escolha recaiu sobre um modelo do tipo ARIMA(1, 1, 0) com estrutura autoregressiva. O Quadro 7.5.5 resume a ordem  $(p, d, q)$  dos modelos ARIMA seleccionados com o auxílio dos critérios BIC e AIC nos diferentes modelos testados, considerando distintos intervalos de idades. Assinale-se que, no caso dos homens, o intervalo de idades considerado exerce alguma influência sobre a ordem do modelo ARIMA.

| Modelo           | Poisson   | Poisson   | Lee-Carter |
|------------------|-----------|-----------|------------|
| Intervalo Idades | 0-84      | 50-84     | 0-84       |
| Homens           | (0, 1, 1) | (0, 1, 2) | (1, 1, 0)  |
| Mulheres         | (1, 1, 0) | (1, 1, 0) | (1, 1, 0)  |

Quadro 7.5.5: Ordem dos modelos ARIMA seleccionados

### 7.5.3 Estimação dos parâmetros do modelo ARIMA

Para estimar os modelos seleccionados na secção anterior, recorreremos ao método de máxima verosimilhança condicional. Consideremos os modelos ARIMA(0, 1, 1) para os homens

$$\Delta_1 \kappa_t^m = \rho^m + \theta^m \varepsilon_{t-1}^m + \varepsilon_t^m$$

e ARIMA(1, 1, 0) para as mulheres

$$\Delta_1 \kappa_t^w = \rho^w + \phi^w \kappa_{t-1}^w + \varepsilon_t^w$$

onde  $\varepsilon_t^m$  e  $\varepsilon_t^w$  são ruídos brancos com variância  $\sigma_m^2$  e  $\sigma_w^2$  e os índices  $m$  e  $w$  respeitam a homens e mulheres, respectivamente. As estimativas foram obtidas com recurso ao procedimento ARIMA, implementado no *package* SAS.

O Quadro 7.5.6 resume os principais resultados obtidos.

| Parâmetro  | Estimativa | Std error | t-value | p-value |
|------------|------------|-----------|---------|---------|
| $\rho^m$   | -1.64623   | 0.11663   | -14.11  | <.0001  |
| $\theta^m$ | 0.64315    | 0.14831   | 4.34    | 0.0001  |
| $\sigma_m$ | 1.800992   |           |         |         |
| $\rho^w$   | -2.14802   | 0.23969   | -8.96   | <.0001  |
| $\phi^w$   | -0.63606   | 0.15145   | -4.20   | 0.0002  |
| $\sigma_w$ | 2.263249   |           |         |         |

Quadro 7.5.6: Estimativas dos parâmetros do modelo ARMA

A análise dos valores de prova permite-nos concluir que todos os parâmetros são significativos ao nível de significância  $\alpha = 0.05$ .

### 7.5.4 Análise dos resíduos

Para completar a análise sobre a qualidade do ajustamento do modelo ARIMA, resta-nos verificar se os resíduos da estimação constituem realizações de um processo do tipo ruído branco. Para o efeito, dispomos de um teste global sobre a hipótese de correlação nula entre os resíduos, baseado na estatística  $Q(m)$  de Ljung e Box (1978), com  $Q(m) \sim \chi_m^2$ . Os resultados obtidos considerando diferentes níveis de desfasamento  $m = 6, 12, 18$  são reportados no Quadro 7.5.7.

A análise dos valores da estatística de teste e respectivos valores de prova não nos coloca, em todos os casos analisados, em condições de rejeitar a hipótese nula de ausência de correlação entre os resíduos dos modelos ARIMA.

## 7.5. Extrapolação da tendência temporal

| To lag | Homens   |    |         | Mulheres |    |         |
|--------|----------|----|---------|----------|----|---------|
|        | $\chi^2$ | df | p-value | $\chi^2$ | df | p-value |
| 6      | 4.39     | 5  | 0.4943  | 5.89     | 5  | 0.3170  |
| 12     | 9.96     | 11 | 0.5343  | 15.69    | 11 | 0.1530  |
| 18     | 18.05    | 17 | 0.3854  | 25.48    | 17 | 0.0846  |

Quadro 7.5.7: Teste de Ljung e Box para os resíduos do modelo ARIMA

Na Figura 7.5.3 avaliamos o grau de ajustamento dos resíduos dos modelos à distribuição normal estandardizada. Como se observa, os quartis empíricos e teóricos aproximam-se bastante, confirmando as conclusões retiradas com base no teste de Ljung-Box.

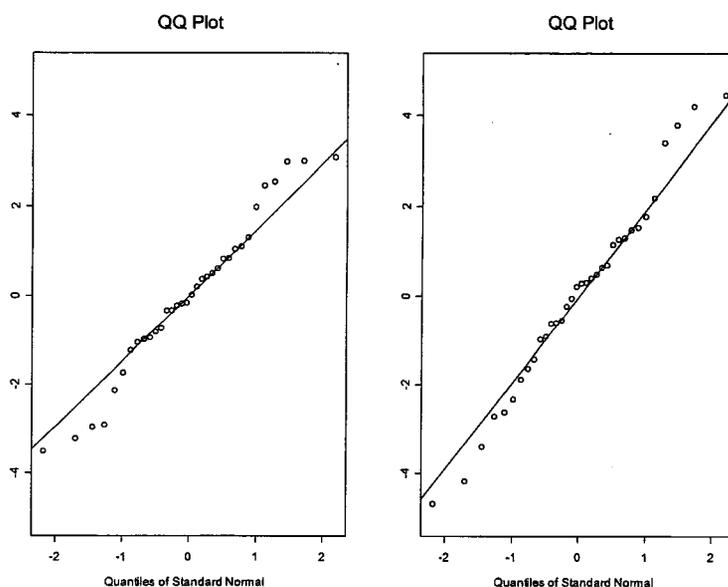


Figura 7.5.3: QQ-plots normal dos resíduos dos modelos ARIMA (homens à esquerda e mulheres à direita)

Esta conclusão foi igualmente confirmada pelos resultados obtidos nos testes de Shapiro-Wilk e Jarque-Bera para a hipótese nula de normalidade dos resíduos.

### 7.5.5 Previsão

Os modelos  $ARIMA(p, d, q)$  estimados na secção anterior colocam-nos em condições de projectar a série  $\{\hat{\kappa}_t\}$  para o futuro. Na Figura 7.5.4 representamos o valor projectado de

## 7.5. Extrapolação da tendência temporal

$\kappa_t \{ \hat{\kappa}_{t_{\max}+s} : s > 0 \}$  para o horizonte temporal 2060, bem como os respectivos intervalos de confiança a 95%. Em condições ideais, uma base de dados com apenas 35 anos não nos qualificaria à partida para efectuar previsões avisadas para um horizonte temporal tão longínquo.<sup>8</sup> Esta limitação é, neste caso, suplantada pela necessidade de calcular prémios relativos a rendas vitalícias e indicadores como a esperança de vida remanescente, cálculos que envolvem, inevitavelmente, a realização de projecções de longo prazo para  $\kappa_t$ . Registe-se, no entanto, que temos bem presente a incerteza inerente à realização de projecções para um horizonte tão longínquo.

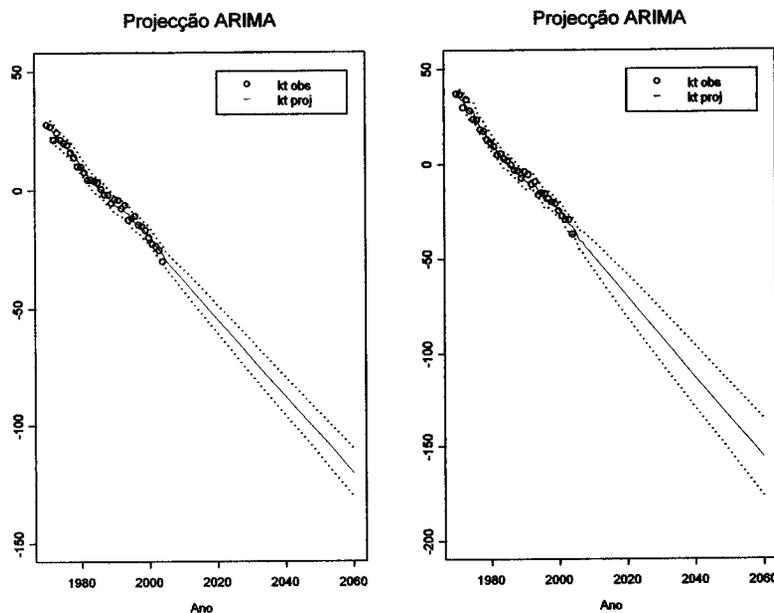


Figura 7.5.4: Projecção do parâmetro  $\kappa_t$  para o horizonte temporal 2060 e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita)

Feita esta ressalva, resta-nos usar os valores projectados de  $\kappa_t$  para obter tábuas de mortalidade prospectivas. Com efeito, tomando a projecção  $\{ \hat{\kappa}_{2004+s} : s = 1, 2, \dots \}$ , é possível deduzir

$$\hat{\mu}_{x,2004+s} = \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_{2004+s}), \quad s = 1, 2, \dots$$

<sup>8</sup>Esta afirmação decorre do princípio, muitas vezes seguido, de que o horizonte de projecção deve ser, no máximo, igual à dimensão do intervalo de estimação.

## 7.6 Projecção da mortalidade nas idades avançadas

Como se observa nas Figuras 7.3.1 e 7.3.2, as estimativas brutas de  $\mu_x$  calculadas de forma clássica sofrem de alguma instabilidade nas idades mais avançadas. Este comportamento, explicável em boa medida pela escassez de dados, justifica a introdução de algum alisamento nas séries.

Por outro lado, a idade máxima para a qual dispomos de observações (em especial no que toca às estimativas de população residente) é manifestamente insuficiente para permitir a modelação da extinção total de uma população. No nosso caso, 84 anos é certamente insuficiente quando sabemos que, no período em análise, existem óbitos registados aos 117 e aos 124 anos nas populações masculina e feminina, respectivamente.

Os continuados aumentos na esperança de vida registados nas últimas décadas traduzem-se, por um lado, num número crescente de sobreviventes até às idades mais avançadas e, por outro, num acréscimo da proporção dos óbitos registados nestas idades em relação ao total. Entende-se assim o crescente interesse pelo estudo do comportamento da mortalidade nas idades avançadas.

Nesta secção, são aplicadas técnicas de graduação com o objectivo de suavizar as estimativas brutas de  $\mu_x$  observadas nas idades mais avançadas. Em paralelo, é aplicado um método de fecho das tábuas que envolve a extrapolação condicionada dos quocientes observados nas idades avançadas para uma dada idade limite, fixada neste caso nos 120 anos.

### 7.6.1 Fixação da idade biológica máxima

O debate em torno da existência de uma idade limite  $\omega$  (ou idade biológica máxima) desperta desde há muito o interesse da comunidade científica. Trata-se, naturalmente, de um tema cuja discussão vai muito para além dos objectivos desta dissertação, sobre o qual não existe unanimidade e que não está isento de alguma controvérsia. Com efeito, o estado actual do conhecimento não nos coloca em condições de avançar com uma idade biológica máxima para a espécie humana.

Num estudo relativamente recente (Vaupel, 1997), elaborado a partir de uma importante base de dados, a saber 70 milhões de seres humanos de 14 países que atingiram os 80 anos de idade, entre os quais se contam 200.000 centenários, foi possível detectar uma desaceleração na taxa de crescimento dos  $q_x$  a partir dos 80 anos, segmento onde inclusive foi possível registar algum decréscimo no nível de mortalidade. Não sendo suficiente para abandonarmos formalmente a hipótese de uma idade máxima, este padrão de mortalidade alerta-nos para a incerteza em torno da duração da vida humana.

A maioria dos actuários, apoiados nas opiniões autorizadas de médicos ou proeminentes biólogos, adopta contudo uma posição pragmática e acredita na existência de um limite superior para a duração da vida humana. A consequência imediata é a necessidade de “fechar” as tábuas de mortalidade, ou seja, a necessidade de extrapolar as probabilidades anuais de óbito para além das idades observadas até uma idade limite. Uma alternativa simples passaria por admitir que as probabilidades  $q_x$  são constantes a partir do último valor observado, apesar dos estudos empíricos realizados apontarem para a formação de uma espécie de “plateau” nas idades avançadas.

No nosso caso, entendemos que a adopção de uma idade limite se apresenta como a hipótese mais razoável e consentânea com a realidade. Para tal, considerámos como idade limite para a população portuguesa os 120 anos, em ambos os sexos. Esta idade máxima pode, no entanto, ser entendida como um limite a partir do qual os eventuais sobreviventes serão negligenciados no cálculo de indicadores demográficos ou de prémios de contratos de seguros.

A idade de 120 anos proposta no nosso caso pode, à primeira vista, ser considerada elevada se atendermos unicamente ao reduzido número de centenários a residir actualmente em Portugal. No entanto, convém recordar que este mesmo limite será considerado em todo o horizonte temporal de projecção, bem assim como na construção de tábuas prospectivas para as gerações mais jovens. Este facto, combinado com a trajectória descendente da mortalidade registada nas últimas décadas pode, no futuro, transformar a idade limite de 120 anos num patamar perfeitamente trivial.

### 7.6.2 Método de fecho das tábuas

Fixada a idade limite, o passo seguinte consistia em seleccionar o método de fecho das tábuas mais adequado à realidade. A este respeito, testámos alguns dos modelos descritos na Secção 6.15, nomeadamente os métodos de Coale-Kisker e de Denuit e Goderniaux e os modelos logístico de Perks e de Kannisto analisados nos Capítulos 2 e 6.<sup>9</sup>

Os resultados obtidos permitiram concluir que os métodos de Denuit-Goderniaux (DG) e de Coale-Kisker (CK) adoptam trajectórias comparáveis para a evolução da mortalidade, embora se tenha constatado que no método CK a qualidade do ajustamento depende, em boa medida, dos valores pré-definidos para a taxa de mortalidade correspondente à idade limite  $m_{x_{\max}}$ . O melhor comportamento destes modelos não pode ser dissociado da imposição de restrições no processo de extrapolação, que permitem cal-

---

<sup>9</sup>Os resultados desta análise fazem parte de um estudo realizado pelo autor no âmbito de um grupo de trabalho criado pelo Instituto Nacional de Estatística para definir a nova metodologia de cálculo das tábuas de mortalidade contemporâneas portuguesas.

## 7.6. Projecção da mortalidade nas idades avançadas

---

ibrar os modelos às condições observadas em cada momento, atendendo nomeadamente à idade máxima para a qual são registados óbitos na população. O método DG apresenta, no entanto, uma vantagem importante face ao método CK, uma vez que permite fixar antecipadamente a idade de fecho da tábua, ou seja, à idade para a qual se estima um quociente de mortalidade unitário.

Pelo contrário, os modelos logísticos de Kannisto e Perks envolvem basicamente o ajustamento das estimativas brutas, seguido de extrapolação não condicionada da mortalidade. Em função disto, constatou-se que os modelos geram trajetórias bastante irregulares e inconsistentes com o comportamento normal da mortalidade nas idades avançadas, seja produzindo quocientes  $q_x$  estagnados com a idade, seja gerando trajetórias explosivas, que não acompanham o já referido fenómeno de desaceleração da taxa de crescimento. Acresce que os modelos não convergem automaticamente para a idade de fecho definida, apresentando ainda oscilações muito significativas de ano para ano.

Em face destes argumentos, os resultados do estudo recomendaram a utilização do método proposto por Denuit e Goderniaux (2005). Conforme vimos na Secção 6.15, o método trabalha directamente sobre os quocientes de mortalidade e introduz condições de fecho de modo a impor à curva dos  $q_x$  uma configuração côncava nas idades elevadas e a impedir um eventual decréscimo dos quocientes nas idades avançadas. A estimação do método envolveu o ajustamento, pelo método dos mínimos quadrados, do modelo log-quadrático (6.15.14), i.e., de

$$\ln \hat{q}_{x,t} = a(t) + b(t)x + c(t)x^2 + \epsilon_x(t), \quad \epsilon_{x,t} \sim \mathcal{N}or(0, \sigma^2(t))$$

com as seguintes restrições

$$q_{120} = 1 \quad \text{e} \quad q'_{120} = 0,$$

em cada um dos anos que integram a amostra, considerando as observações relativas às idades no intervalo  $65 \leq x \leq 84$ . Este procedimento foi efectuado com base numa rotina construída e implementada no *package* SAS (Release 9.1).

Um dos aspectos críticos no modelo refere-se à determinação da idade a partir da qual as estimativas brutas  $\hat{q}_x$  são substituídas pelos valores ajustados provenientes do modelo. Para responder a esta questão, usámos como critério de optimização a maximização do coeficiente de determinação  $R^2$  da regressão, fazendo variar a “idade de corte” num intervalo entre os 70 e os 84 anos. Os valores obtidos para a estatística  $R^2$  revelaram-se, em todos os casos, muito próximos e sempre superiores a 0.98. Em consequência desta evidência, e por uma questão de simplicidade, optámos por considerar um limite inferior único em todos os anos de calendário, e por substituir os quocientes brutos pelos valores

estimados pelo modelo a partir dos 85 anos, inclusive.

De seguida, combinámos os valores observados e projectados de  $q_x$  com os valores gerados pelo método de fecho. O resultado é uma nova matriz  $\{\hat{q}_{x,t} : x \in [0, 120], t \in [1970, 2124]\}$ , constituída pelos quocientes brutos para  $x \in [0, 84]$  e por valores graduados para  $x \in [85, 120]$ . Para evitar o surgimento de descontinuidades no comportamento da série na vizinhança da idade de corte, provocadas pela aplicação do procedimento de regressão, introduzimos algum alisamento na série recalculando os quocientes  $q_{x,t}$  nas idades  $x \in [80, 90]$  usando médias geométricas de 5 anos. Em termos formais, os quocientes suavizados são obtidos mediante

$$\hat{q}_x^* = (\hat{q}_{x-2} \cdot \hat{q}_{x-1} \cdot \hat{q}_x \cdot \hat{q}_{x+1} \cdot \hat{q}_{x+2})^{\frac{1}{5}}, \quad x = 80, \dots, 90. \quad (7.6.1)$$

Em síntese, dispomos agora de uma matriz de quocientes  $\{\tilde{q}_{x,t} : x = 0, \dots, 120; t = 1970, \dots, 2124\}$  onde  $\tilde{q}_{x,t} = \hat{q}_{x,t}^*$  para  $x \in [80, 90]$  e  $\tilde{q}_{x,t} = \hat{q}_{x,t}$  nas demais idades. Esta matriz torna possível a construção de tábuas projectadas e o cálculo de indicadores como a esperança de vida ou o preço de uma renda vitalícia, usando quer uma abordagem transversal, quer uma abordagem diagonal ou geracional.

Os quocientes obtidos pela aplicação do procedimento de fecho das tábuas são representados nas Figuras 7.6.1 e 7.6.2.

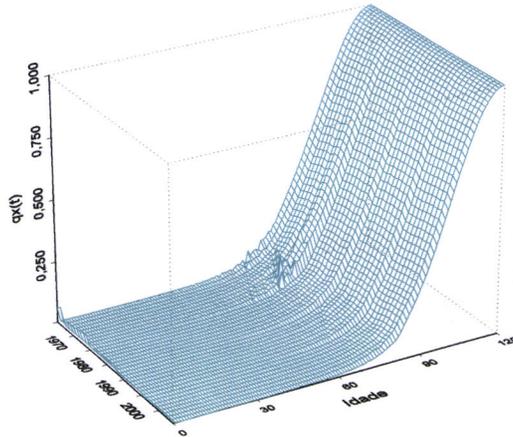


Figura 7.6.1: Quocientes  $q_{x,t}$  com tábuas fechadas (Homens)

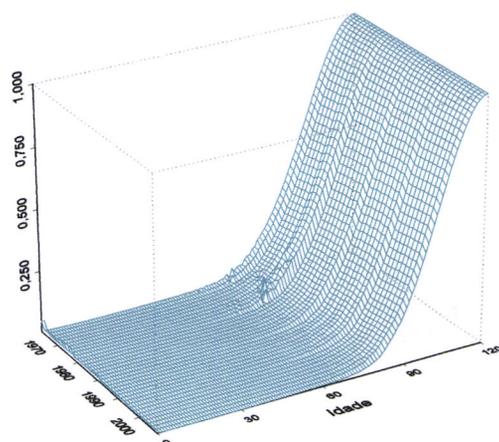


Figura 7.6.2: Quocientes  $q_{x,t}$  com tábuas fechadas (Mulheres)

## 7.7 Evolução da mortalidade

### 7.7.1 Por ano cronológico

As tábuas prospectivas derivadas na secção anterior colocam-nos em condições de analisar a evolução da mortalidade ao longo do tempo considerando, quer uma abordagem transversal ou de momento, quer uma abordagem geracional ou prospectiva. Todos os cálculos apresentados nesta secção são efectuados com base nas taxas ajustadas (para  $t \leq t_{\max}$ ) ou projectadas (para  $t \geq t_{\max}$ ) provenientes da estimação dos parâmetros do modelo de Poisson clássico considerando as estatísticas relativas à população portuguesa no período 1970-2004.

As Figuras 7.7.1 e 7.7.2 representam a evolução observada e estimada das taxas  $\mu_{x,t}$  em anos seleccionados entre 1970 e 2050, para os modelos aplicados a todas as idades ou apenas às idades no intervalo 50-84, numa óptica transversal. Na Figura 7.7.3 é possível observar a evolução dos quocientes de mortalidade entre 1970 e 2124 num conjunto de idades representativas. Como se observa, as figuras assinalam, de uma maneira geral, uma trajectória decrescente da mortalidade no tempo, com os valores projectados pelo modelo a replicarem as tendências observadas no período amostral. Assinala-se que os principais ganhos de longevidade são observados nas idades mais jovens, até por volta dos 15-20 anos, e em menor escala nas idades adultas a partir dos 40 anos.

## 7.7. Evolução da mortalidade

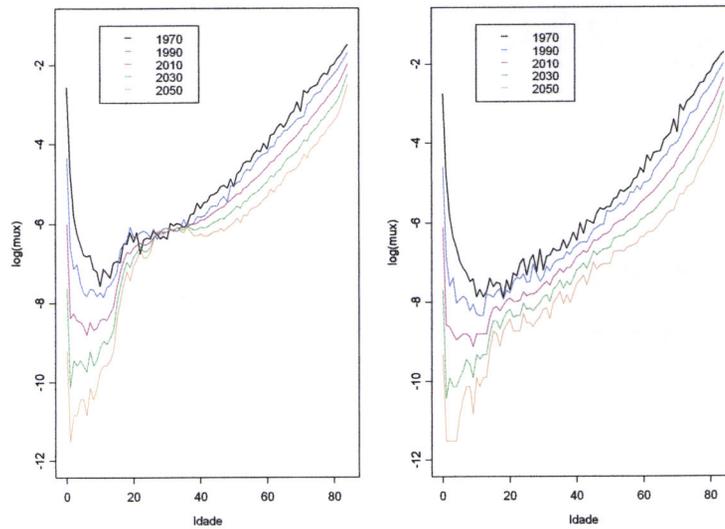


Figura 7.7.1: Evolução das  $\mu_{x,t}$  ajustadas pelo modelo de Poisson aplicado a todas as idades (homens à esquerda e mulheres à direita)

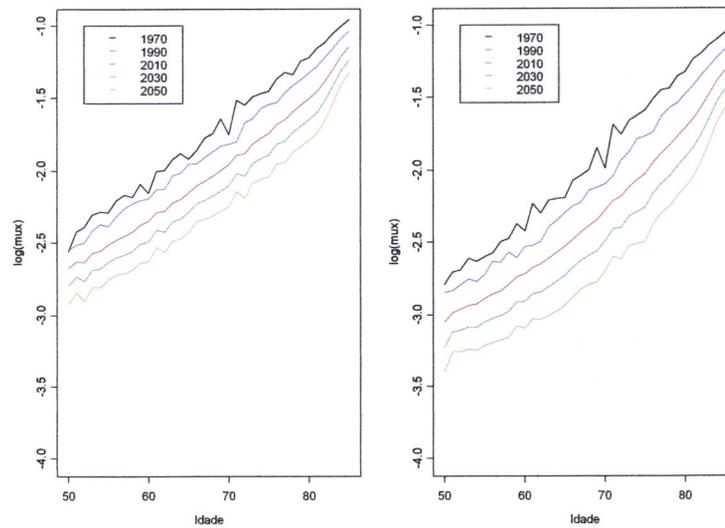


Figura 7.7.2: Evolução das  $\mu_{x,t}$  ajustadas pelo modelo de Poisson aplicado ao intervalo de idades  $x \in [50, 84]$  (homens à esquerda e mulheres à direita)

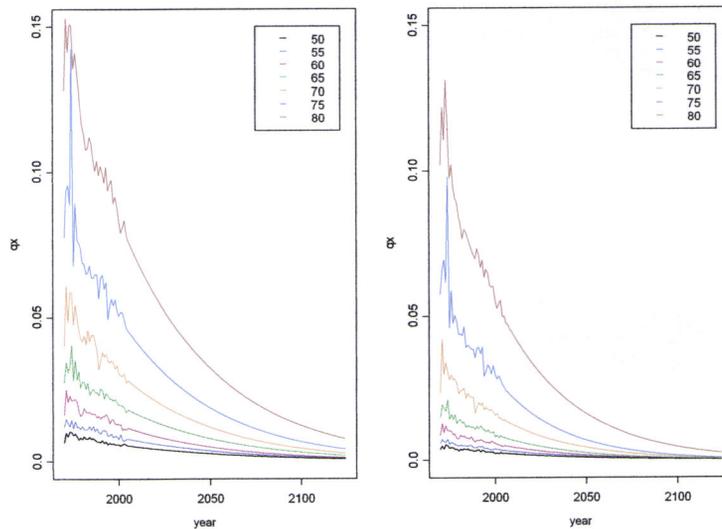


Figura 7.7.3: Evolução dos  $q_{x,t}$  entre 1970 e 2124 para algumas idades representativas (homens à esquerda e mulheres à direita)

O fenómeno “bossa” de acidentes registado no período amostral tende a acentuar-se com o tempo, em especial no caso da população masculina, onde se observa uma espécie de estagnação nas taxas  $\mu_{x,t}$  correspondentes a essas idades. Este resultado é reflexo dos ganhos incipientes ou nulos observados neste segmento no período 1970-2004, que projectando-se no futuro aumentam a sua importância relativa no conjunto do arco da vida humana. Nas idades avançadas, o padrão de decréscimo afigura-se consistente no tempo, a um ritmo ligeiramente mais acelerado no caso das mulheres.

### 7.7.2 Por geração ou ano de nascimento

A construção de tábuas prospectivas oferece-nos novas oportunidades em termos de análise do fenómeno da mortalidade, designadamente a possibilidade de investigar o seu comportamento não mais em termos de anos de calendário mas por geração ou data de nascimento. Se recapitularmos a ilustração do problema realizada na Figura 6.3.1, trata-se, no fundo, de passar de uma abordagem transversal (ou vertical) em que se analisam uma sequência de tábuas contemporâneas, uma por cada ano cronológico  $t$ , para uma abordagem longitudinal, em que o fenómeno é investigado considerado uma sequência de tábuas prospectivas, uma por cada ano de nascimento ou geração  $t$ .

## 7.7. Evolução da mortalidade

Nas Figuras 7.7.4 e 7.7.5 representamos a evolução das  $\mu_{x,t}$  ajustadas pelo modelo de Poisson para os indivíduos nascidos em gerações seleccionadas entre 1970 e 2004 e para os indivíduos que completam 50 anos de idade entre 1970 e 2004, respectivamente. Como se observa, as grandes características apontadas no âmbito da óptica transversal (decréscimo acentuado na mortalidade infantil, bossa de acidentes, etc.) são uma vez mais facilmente identificáveis agora numa óptica geracional. Constata-se, no entanto, que a evolução da mortalidade entre gerações sucessivas apresenta oscilações menos acentuadas do que entre anos de calendário, aparentando ser mais fiável e razoável.

Na Figura 7.7.6 efectuamos uma comparação entre as taxas  $\mu_{x,t}$  derivadas nas ópticas transversal e geracional em anos de calendário (e respectivas gerações) seleccionados. Como se observa, num cenário de diminuição generalizada da mortalidade os valores previstos segundo a óptica longitudinal são naturalmente menores que os registados segundo a abordagem transversal, com a diferença a aumentar com a idade do indivíduo e com o ano de nascimento da geração.

A excepção refere-se, uma vez mais, ao comportamento registado nas idades correspondentes à “bossa” de acidentes, nas quais é bem patente a estagnação (e até um ligeiro retrocesso previsto) na evolução da mortalidade.

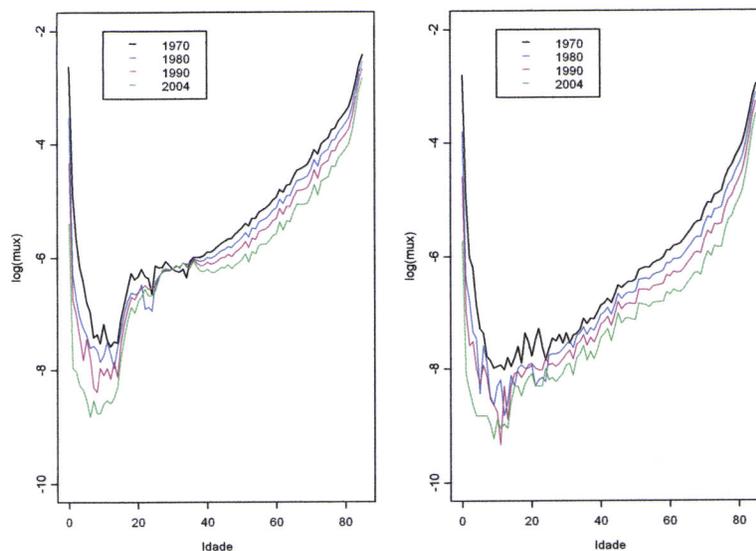


Figura 7.7.4: Evolução das  $\mu_{x,t}$  ajustadas pelo modelo de Poisson para os indivíduos nascidos nos anos de 1970, 1980, 1990 e 2004 (homens à esquerda e mulheres à direita)

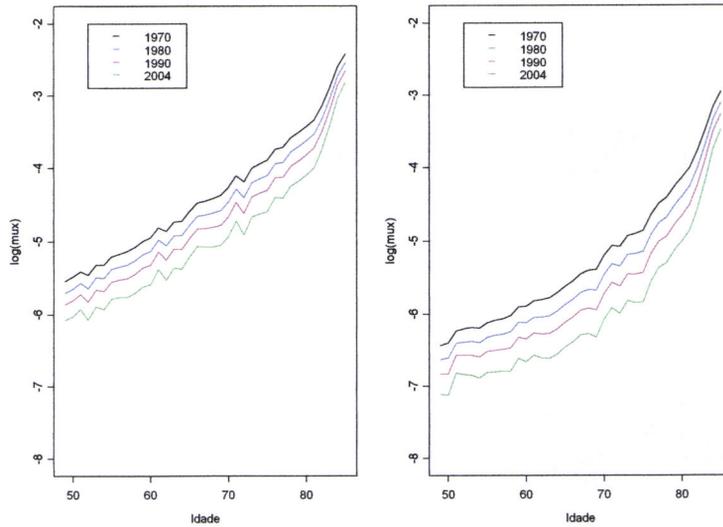


Figura 7.7.5: Evolução das  $\mu_{x,t}$  para as gerações de indivíduos que atingem os 50 anos em 1970, 1980, 1990 e 2004 (homens à esquerda e mulheres à direita)

### 7.7.3 Estimativas da esperança de vida

Nesta secção, analisamos a evolução das estimativas da esperança de vida completa  $\hat{e}_x(t)$  em função do ano de calendário  $t = 1970, \dots, 2004$  nas idades  $x = 0$  e 65 anos. Recordamos que, na abordagem transversal,  $\hat{e}_x(t)$  é calculado com base na mortalidade relativa ao ano  $t$  (i.e., com recurso aos quocientes  $q_{x+k,t}$ ,  $k = 0, 1, 2, \dots$ ), sendo por isso comparável com os valores reportados pelo I.N.E. para o ano  $t$ . Em termos formais, se assumirmos que  $\mu_{x+\xi,t} = \mu_{x,t}$  (para  $0 \leq \xi < 1$  e  $x$  inteiro), deduz-se que

$$\begin{aligned}
 \hat{e}_x(t) &= \int_0^1 \exp(-\xi \mu_{x,t}) d\xi \\
 &+ \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \exp(-\mu_{x+j,t}) \right\} \int_k^{k+1} \exp(-(\xi - k) \mu_{x+k,t}) d\xi \\
 &= \frac{1 - \exp(-\mu_{x,t})}{\mu_{x,t}} \\
 &+ \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \exp(-\mu_{x+j,t}) \right\} \frac{1 - \exp(-\mu_{x+k,t})}{\mu_{x+k,t}}. \tag{7.7.1}
 \end{aligned}$$

## 7.7. Evolução da mortalidade

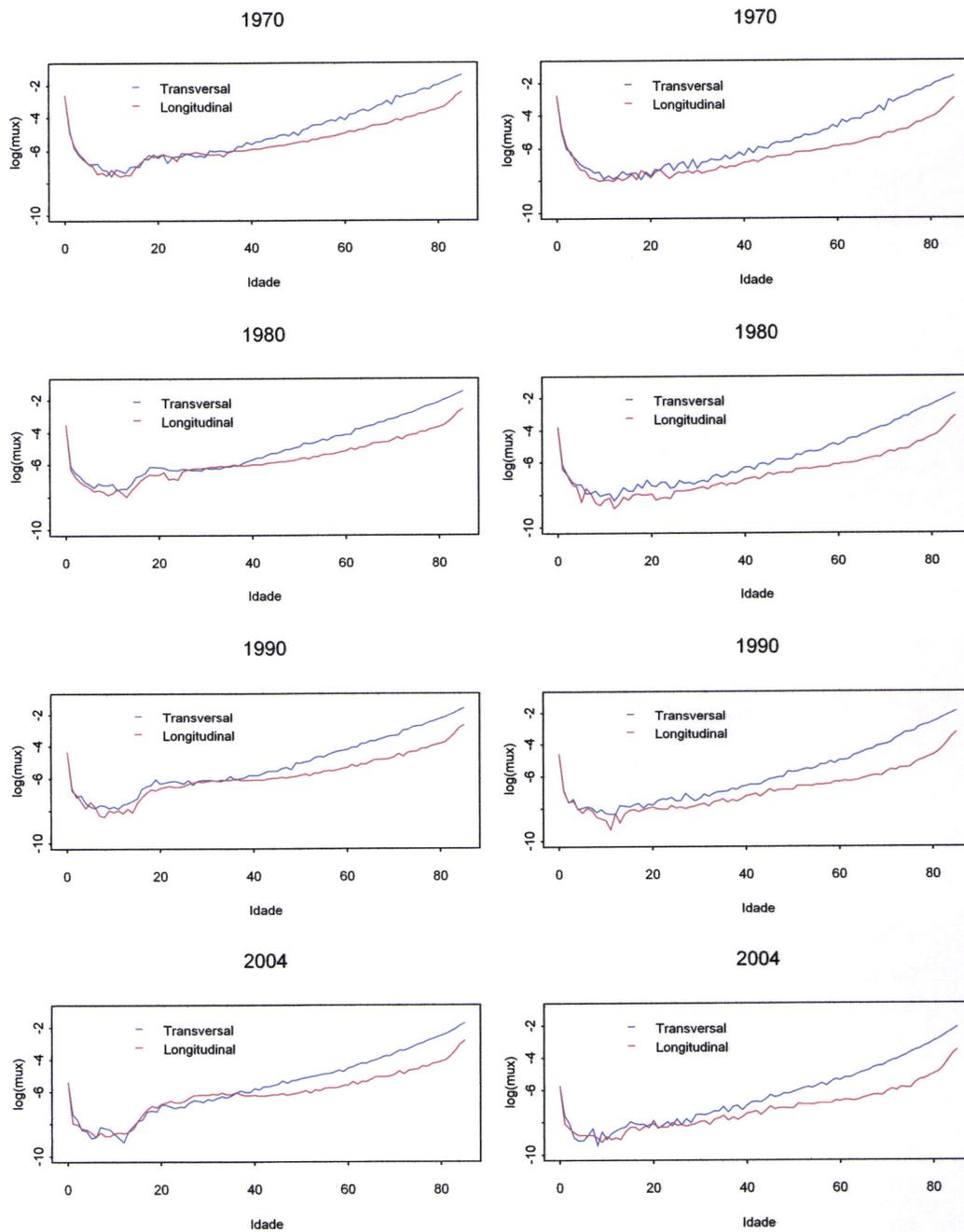


Figura 7.7.6: Comparação entre as taxas  $\mu_{x,t}$  obtidas nas ópticas transversal e geracional em anos seleccionados (homens à esquerda e mulheres à direita)

## 7.7. Evolução da mortalidade

Pelo contrário, na abordagem longitudinal  $\dot{e}_x(t)$  representa o valor esperado da duração de vida remanescente dos indivíduos que alcançam a idade  $x$  durante o ano  $t$ , sendo, por isso, calculada a partir dos quocientes  $q_{x+k,t+k}$  ( $k = 0, 1, 2, \dots$ ). Formalmente, prova-se que a aplicação da hipótese (6.9.4) nos conduz à seguinte expressão para  $\dot{e}_x(t)$

$$\begin{aligned} \dot{e}_x(t) &= \int_0^1 \exp(-\xi \mu_{x,t}) d\xi \\ &+ \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \exp(-\mu_{x+j,t+j}) \right\} \int_k^{k+1} \exp(-(\xi - k) \mu_{x+k,t+k}) d\xi \\ &= \frac{1 - \exp(-\mu_{x,t})}{\mu_{x,t}} \\ &+ \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \exp(-\mu_{x+j,t+j}) \right\} \frac{1 - \exp(-\mu_{x+k,t+k})}{\mu_{x+k,t+k}}. \end{aligned} \quad (7.7.2)$$

Dito isto, o Quadro 7.7.1 resume os valores estimados para a esperança de vida completa à nascença  $e_0(t)$  usando os modelos de Poisson e de Lee-Carter, segundo as ópticas longitudinal e transversal. Os resultados detalhados são apresentadas no Anexo 7.A. A coluna  $\Delta \dot{e}_0^{(P-LC)}$  reporta a diferença (em anos) entre os valores estimados pelo modelo de Poisson e pelo método Lee-Carter. A coluna  $\Delta \dot{e}_0$  indica os ganhos médios anuais (em dias) registados no indicador  $\dot{e}_0(t)$  entre 1970 e 2004 (tomando como referência os valores obtidos pelo modelo de Poisson).

| Homens         |              |            |                             |                    |             |            |                             |                    |
|----------------|--------------|------------|-----------------------------|--------------------|-------------|------------|-----------------------------|--------------------|
| $\dot{e}_0(t)$ | Longitudinal |            |                             |                    | Transversal |            |                             |                    |
| $t$            | Poisson      | Lee-Carter | $\Delta \dot{e}_0^{(P-LC)}$ | $\Delta \dot{e}_0$ | Poisson     | Lee-Carter | $\Delta \dot{e}_0^{(P-LC)}$ | $\Delta \dot{e}_0$ |
| 1970           | 71.99        | 72.09      | -0.10                       |                    | 63.21       | 63.29      | -0.08                       |                    |
| 2004           | 83.30        | 83.49      | -0.19                       | 117.9              | 74.55       | 74.66      | -0.11                       | 118.3              |

| Mulheres       |              |            |                             |                    |             |            |                             |                    |
|----------------|--------------|------------|-----------------------------|--------------------|-------------|------------|-----------------------------|--------------------|
| $\dot{e}_0(t)$ | Longitudinal |            |                             |                    | Transversal |            |                             |                    |
| $t$            | Poisson      | Lee-Carter | $\Delta \dot{e}_0^{(P-LC)}$ | $\Delta \dot{e}_0$ | Poisson     | Lee-Carter | $\Delta \dot{e}_0^{(P-LC)}$ | $\Delta \dot{e}_0$ |
| 1970           | 80.37        | 80.23      | 0.14                        |                    | 69.32       | 69.27      | 0.05                        |                    |
| 2004           | 90.63        | 90.43      | 0.20                        | 107.0              | 81.05       | 81.01      | 0.04                        | 122.4              |

Quadro 7.7.1: Evolução da esperança de vida à nascença  $e_0(t)$

O primeiro aspecto a salientar diz respeito aos ganhos notáveis de esperança de vida registados no período 1970-2004. Com efeito, segundo a visão transversal, por cada

ano decorrido a esperança de vida aumentou, em média, aproximadamente 4 meses em ambos os sexos (mais concretamente 118.3 e 122.4 dias para homens e mulheres, respectivamente). Estes valores são ligeiramente mais moderados (em especial para a população feminina) no caso da visão longitudinal, com ganhos de 117.9 e 107.0 dias para homens e mulheres, respectivamente.

O segundo dado importante refere-se à diferença significativa entre os valores calculados segundo as visões longitudinal e transversal. De facto, a utilização das tábuas prospectivas permite-nos estimar que a “verdadeira” esperança de vida à nascença para um indivíduo nascido em 2004 será, respectivamente, de 83.30 e 90.63 anos para homens e mulheres, e não de 74.55 e 81.05 anos, como prevê a tradicional visão transversal. Dito de outro modo, a projecção para o futuro das tendências de decréscimo da mortalidade verificadas entre 1970 e 2004 traduz-se numa subestimação da verdadeira esperança de vida em 8.75 e 9.58 anos para homens e mulheres, respectivamente, quando baseada na abordagem transversal. Esta conclusão, aparentemente surpreendente, é suficiente para reforçar a necessidade de dispor de tábuas de mortalidade prospectivas para uma correcta avaliação da longevidade humana.

Os valores de  $e_0(t)$  calculados com base nas estatísticas de mortalidade brutas publicadas pelo INE relativas ao período 2002-2003 ascendem a 73.93 e 80.23 anos para homens e mulheres, respectivamente. A mesma média, quando calculada com base no modelo de Poisson e relativa aos anos de 2002 e 2003, ascende a 73.74 e 80.33 anos. Ou seja, se compararmos as estimativas oficiais e as geradas pelo modelo de Poisson constatamos que as diferenças são negligenciáveis, o que atesta bem a qualidade do ajustamento proporcionada pelo modelo. Por outro lado, as diferenças entre os valores estimados pelos modelos de Poisson e de Lee-Carter são pouco significativas, em particular na abordagem transversal.

No que diz respeito às estimativas da esperança de vida restante aos 65 anos,  $e_{65}(t)$ , os valores reportados no Quadro 7.7.2 indicam que as diferenças entre as visões transversal e longitudinal continuam bem patentes, embora se apresentem, naturalmente, menos acentuadas (quer em termos absolutos quer em termos relativos). Ainda assim, e tomando como referência o modelo de Poisson aplicado no intervalo de idades  $x \in [0, 84]$ , a aplicação da visão transversal subestima a esperança de vida remanescente aos 65 anos em 1.07 e 1.65 anos para homens e mulheres, respectivamente. Estas diferenças podem, à primeira vista, ser consideradas de pouca monta, mas a verdade é que elas assumem, como veremos em seguida, uma expressão significativa no *pricing* de rendas vitalícias e de outros produtos com benefícios de longo prazo em caso de vida.

Os valores publicados pelo INE para o período 2002-2003 ascendem, como se observa

## 7.7. Evolução da mortalidade

no Quadro 7.2.5, a 15.57 e 18.68 anos para homens e mulheres, respectivamente, que comparam com 15.39 e 18.81 (média em 2002-2003) calculados segundo o modelo de Poisson.

A análise dos valores referentes aos ganhos de esperança de vida permite concluir que, em média, cada ano decorrido entre 1970 e 2004 significou um aumento de 39.2 (38.0) e 48.3 (49.0) dias no indicador  $\hat{e}_{65}(t)$  para homens e mulheres, respectivamente, de acordo com a visão longitudinal (transversal).

Em relação ao desempenho dos vários modelos testados, constata-se uma vez mais que a diferença entre os valores estimados é pouco expressiva. A consideração de diferentes intervalos de idades na estimação dos parâmetros do modelo de Poisson não se traduz, na prática, em diferenças importantes no indicador  $\hat{e}_{65}(t)$ . Ainda assim, verifica-se que os valores estimados pelo modelo aplicado no intervalo  $x \in [50, 84]$  são inferiores aos calculados pelo modelo aplicado a todas as idades. Esta diferença pode ser explicada pelo facto de o modelo restrito estar expurgado das idades infantis e juvenis onde o decréscimo na mortalidade foi, como vimos, mais acentuado.

| Homens            |                 |                  |                 |                      |                 |                  |                 |                      |
|-------------------|-----------------|------------------|-----------------|----------------------|-----------------|------------------|-----------------|----------------------|
| $\hat{e}_{65}(t)$ | Longitudinal    |                  |                 |                      | Transversal     |                  |                 |                      |
| $t$               | Poisson<br>0-84 | Poisson<br>50-84 | LCarter<br>0-84 | $\Delta\hat{e}_{65}$ | Poisson<br>0-84 | Poisson<br>50-84 | LCarter<br>0-84 | $\Delta\hat{e}_{65}$ |
| 1970              | 13.16           | 13.11            | 13.20           |                      | 12.21           | 12.17            | 12.27           |                      |
| 2004              | 16.91           | 16.78            | 17.04           | 39.2                 | 15.84           | 15.82            | 15.89           | 38.0                 |

| Mulheres          |                 |                  |                 |                      |                 |                  |                 |                      |
|-------------------|-----------------|------------------|-----------------|----------------------|-----------------|------------------|-----------------|----------------------|
| $\hat{e}_{65}(t)$ | Longitudinal    |                  |                 |                      | Transversal     |                  |                 |                      |
| $t$               | Poisson<br>0-84 | Poisson<br>50-84 | LCarter<br>0-84 | $\Delta\hat{e}_{65}$ | Poisson<br>0-84 | Poisson<br>50-84 | LCarter<br>0-84 | $\Delta\hat{e}_{65}$ |
| 1970              | 16.24           | 16.20            | 16.28           |                      | 14.53           | 14.48            | 14.56           |                      |
| 2004              | 20.87           | 20.73            | 20.83           | 48.3                 | 19.22           | 19.19            | 19.20           | 49.0                 |

Quadro 7.7.2: Evolução da esperança de vida aos 65 anos  $e_{65}(t)$

Na Figura 7.7.7 representamos a evolução da esperança de vida  $\hat{e}_x(t)$  segundo as ópticas transversal e longitudinal no período  $t \in [1970, 2004]$  e idades  $x = 0$  e 65 anos. Como se observa, os diferenciais entre as estimativas de  $\hat{e}_0(t)$  e  $\hat{e}_{65}(t)$  calculadas segundo as duas abordagens permanecem relativamente estáveis no período analisado, um resultado esperado atendendo à trajectória descendente (quase linear) das taxas  $\mu_{x,t}$ .

## 7.7. Evolução da mortalidade

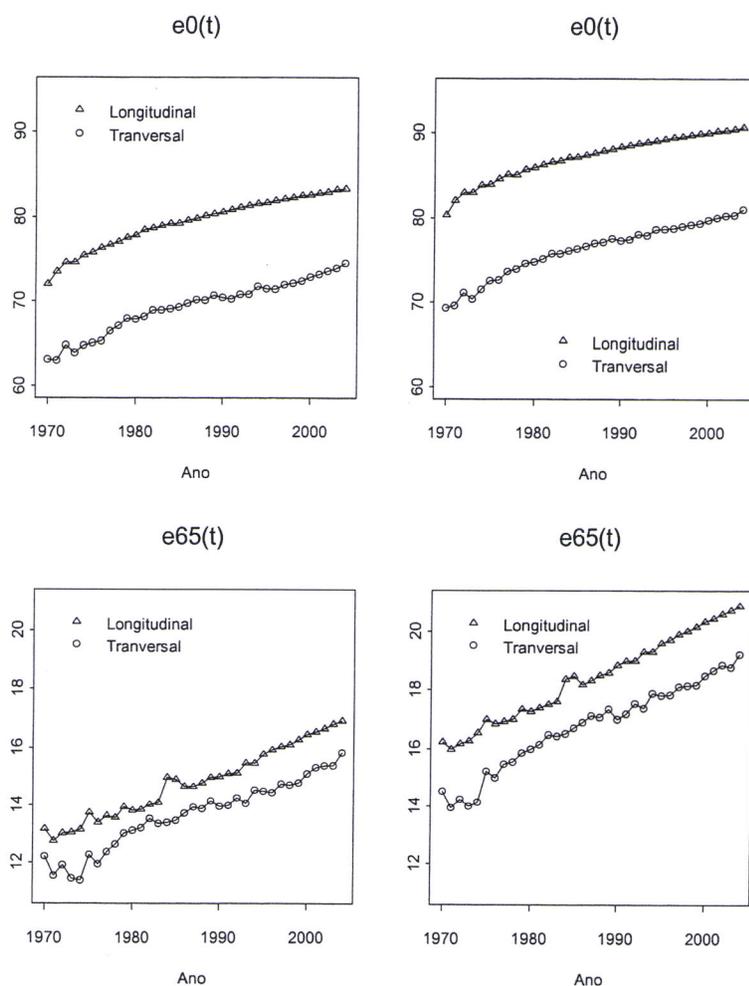


Figura 7.7.7: Evolução da esperança de vida  $e_x(t)$  segundo as ópticas transversal e longitudinal no período  $t \in [1970, 2004]$  e idades  $x = 0$  e  $65$  anos (homens à esquerda e mulheres à direita)

O Quadro 7.7.3 e a Figura 7.7.8 fornecem uma perspectiva de longo prazo para a evolução dos indicadores  $e_0(t)$  e  $e_{65}(t)$ , calculados segundo a visão transversal. Com a devida reserva que uma projecção para um horizonte tão longínquo deve merecer, o exercício de extrapolação permite perspectivar a manutenção da tendência de aumento da longevidade e de acréscimo da esperança de vida em ambos o sexos, embora seja bem patente alguma desaceleração no seu ritmo de crescimento.

## 7.7. Evolução da mortalidade

Registo ainda para a ligeira diminuição projectada para o diferencial de esperança de vida entre homens e mulheres no final do período de previsão.

| Ano  | $e_0(t)$ |          | $e_{65}(t)$ |          |
|------|----------|----------|-------------|----------|
|      | Homens   | Mulheres | Homens      | Mulheres |
| 2004 | 74.55    | 81.05    | 15.84       | 19.22    |
| 2025 | 77.54    | 84.60    | 18.05       | 21.81    |
| 2050 | 81.15    | 88.04    | 20.62       | 24.46    |
| 2075 | 84.15    | 90.68    | 22.84       | 26.60    |
| 2100 | 86.61    | 92.75    | 24.72       | 28.34    |
| 2124 | 88.56    | 94.35    | 26.26       | 29.74    |

Quadro 7.7.3: Evolução previsível da esperança de vida  $e_0(t)$  e  $e_{65}(t)$  segundo uma óptica transversal

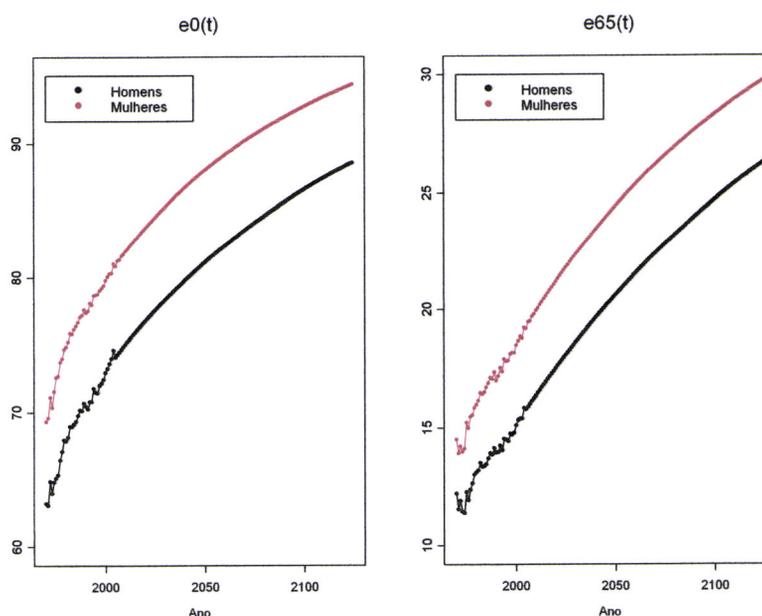


Figura 7.7.8: Evolução projectada da esperança de vida à nascença e aos 65 anos calculada numa óptica transversal

Por último, a nova Lei de Bases da Segurança Social (Lei nº 4/2007 de 16 de Janeiro) recentemente aprovada estipula, no número 1 do artigo 64º, que ao montante da pensão estatutária, calculada nos termos legais, é aplicável um *factor de sustentabilidade* relacionado com a evolução da esperança média de vida, tendo em vista a adequação do

sistema às modificações resultantes de alterações demográficas e económicas. O factor de sustentabilidade é definido pela relação entre a esperança média de vida aos 65 anos verificada em 2006 e a esperança média de vida registada no ano anterior ao do requerimento da pensão, calculada para o conjunto da população (i.e., igual para ambos os sexos). Esta definição implica que um aumento em  $\hat{e}_{65}(t)$  se traduzirá, inevitavelmente, num factor inferior à unidade, ou seja, numa redução do valor da pensão a receber.

Neste sentido, se considerarmos as estimativas produzidas por este estudo, o factor de sustentabilidade a aplicar, por exemplo, a um indivíduo que solicite a pensão em 2050 será de 0.775 (0.799) se for do sexo masculino (feminino), admitindo que o factor é diferenciado entre sexos.

#### 7.7.4 Evolução do preço de uma renda vitalícia

Nesta secção, centramos a nossa atenção na análise da evolução do preço de uma renda vitalícia imediata de valor unitário calculada para um indivíduo que atinge a idade  $x$  no ano  $t$ , i.e., segundo a óptica longitudinal. O seu valor actual,  $a_x(t)$ , é calculado mediante

$$a_x(t) = \sum_{k \geq 0} \left\{ \prod_{j=0}^k p_{x+j,t+j} \right\} \left( \frac{1}{1+r_{0,k+1}} \right)^{k+1}, \quad (7.7.3)$$

onde  $r_{0,k}$  representa a taxa de juro *spot* (anual nominal) aplicável a uma operação com maturidade em  $k$  (taxa de juro técnica).

Por uma questão de simplificação, assumiremos nos cálculos que a taxa de juro técnica é constante e igual a 3%, i.e.,  $r_{0,k} = i = 3\% \forall k$ . A consideração de um comportamento mais realista para a estrutura temporal de taxas de juro implicaria, naturalmente, o abandono desta hipótese e a necessidade de incorporar na análise modelos estocásticos capazes de projectar o comportamento de longo prazo das taxas de juro para as várias maturidades. Refira-se, contudo, que é sempre possível fazer oscilar a variável  $i$  num intervalo razoável de modo a testar a sensibilidade dos resultados a variações na taxa de juro. A este propósito, recordamos que, *ceteris paribus*, o impacto de variações na mortalidade sobre o preço de uma renda vitalícia é tanto maior quanto menor for a taxa de juro técnica adoptada.

A expressão (7.7.3) derivada numa óptica longitudinal contrasta com a sua equivalente no contexto da óptica transversal, definida por

$$a_x(t) = \sum_{k \geq 1} {}_k p_x(t) \cdot v^k, \quad (7.7.4)$$

## 7.7. Evolução da mortalidade

onde  $v = (1 + i)^{-1}$  denota o factor de desconto.

No Quadro 7.7.4 reportamos alguns valores que nos dão uma panorâmica sobre a evolução do preço de uma renda vitalícia à nascença  $a_0(t)$  e aos 65 anos  $a_{65}(t)$ , calculado segundo as ópticas transversal e longitudinal e considerando as estimativas geradas pelo modelo de Poisson. Os valores apresentados nas colunas  $\Delta a_0$  e  $\Delta a_{65}$  referem-se aos ganhos médios anuais no valor de  $a_x(t)$  registados entre 1970 e 2004.

| Homens |                          |                         |                         |                         |
|--------|--------------------------|-------------------------|-------------------------|-------------------------|
| $t$    | $a_0(t)$<br>Longitudinal | $\Delta a_0$<br>(anual) | $a_0(t)$<br>Transversal | $\Delta a_0$<br>(anual) |
| 1970   | 26.84                    |                         | 25.90                   |                         |
| 2004   | 29.82                    | 0.085                   | 29.02                   | 0.089                   |

| Mulheres |                          |                         |                         |                         |
|----------|--------------------------|-------------------------|-------------------------|-------------------------|
| $t$      | $a_0(t)$<br>Longitudinal | $\Delta a_0$<br>(anual) | $a_0(t)$<br>Transversal | $\Delta a_0$<br>(anual) |
| 1970     | 28.18                    |                         | 27.06                   |                         |
| 2004     | 30.72                    | 0.073                   | 29.89                   | 0.081                   |

| Homens |                             |                            |                            |                            |
|--------|-----------------------------|----------------------------|----------------------------|----------------------------|
| $t$    | $a_{65}(t)$<br>Longitudinal | $\Delta a_{65}$<br>(anual) | $a_{65}(t)$<br>Transversal | $\Delta a_{65}$<br>(anual) |
| 1970   | 9.88                        |                            | 9.29                       |                            |
| 2004   | 12.22                       | 0.0668                     | 11.64                      | 0.0671                     |

| Mulheres |                             |                            |                            |                            |
|----------|-----------------------------|----------------------------|----------------------------|----------------------------|
| $t$      | $a_{65}(t)$<br>Longitudinal | $\Delta a_{65}$<br>(anual) | $a_{65}(t)$<br>Transversal | $\Delta a_{65}$<br>(anual) |
| 1970     | 11.86                       |                            | 10.85                      |                            |
| 2004     | 14.56                       | 0.077                      | 13.70                      | 0.081                      |

Quadro 7.7.4: Evolução do preço de uma renda vitalícia à nascença  $a_0(t)$  e aos 65 anos  $a_{65}(t)$ , calculado segundo as ópticas transversal e longitudinal

Como se aprecia, a adopção de uma óptica transversal subavalia claramente o preço a pagar por uma renda vitalícia, quer no caso de um recém-nascido, quer no caso de um indivíduo com 65 anos. Esta subavaliação, bastante expressiva em ambos os sexos, chama a atenção para a importância que a utilização de tábuas prospectivas pelas companhias de seguros assume no presente contexto. Recordamos que num ambiente de taxas de juro historicamente baixas, as consequências de qualquer erro na avaliação da componente demográfica das bases técnicas são claramente mais gravosas do que num cenário em que o preço do dinheiro é mais elevado. Destaque ainda para os ganhos médios anuais

## 7.7. Evolução da mortalidade

---

no valor de  $a_x(t)$  registados entre 1970 e 2004.

No Quadro 7.7.5 admitimos que a evolução da mortalidade pode ser equiparada ao processo de desconto financeiro usando uma determinada taxa de juro. Nesse sentido, calculamos a Taxa Interna de Rentabilidade (TIR) (ou simplesmente taxa de desconto) necessária para igualar o valor de uma renda vitalícia imediata de valor unitário calculada no ano de 2004 segundo a óptica longitudinal ao valor dessa mesma renda calculado usando a correspondente tábua transversal. Os valores inscritos na coluna “ $\Delta(b.p.)$ ” traduzem, para cada idade, a diferença, medida em *basis points* (*b.p.*), entre a taxa de juro técnica de 3% e a TIR estimada.

| Idade | Homens |                | Mulheres |                |
|-------|--------|----------------|----------|----------------|
|       | TIR    | $\Delta(b.p.)$ | TIR      | $\Delta(b.p.)$ |
| 0     | 3,11%  | +11            | 3,10%    | +10            |
| 65    | 3,55%  | +55            | 3,60%    | +60            |

Quadro 7.7.5: Taxa interna de rentabilidade

A leitura do quadro permite verificar, por exemplo, que para um indivíduo do sexo masculino com 65 anos de idade a taxa de desconto requerida para igualar o valor de  $a_{65}(2004)$  calculado segundo as duas ópticas ascende a 3.55%, que compara com a taxa de juro técnica de 3%.

Este valor significa que o facto da mortalidade implícita na tábua prospectiva ser inferior à mortalidade implícita na tábua transversal equivale, na prática, à utilização de uma taxa de desconto superior em 55 *basis points* à taxa de juro base. Dito de outro modo, para que a companhia não apresente prejuízos financeiros pelo facto de subestimar o preço da renda em resultado da adopção de uma tábua transversal, é necessário que esta obtenha taxas de retorno nos mercados de capitais (ganhos financeiros) superiores em 55 *basis points* à taxa de juro técnica.

Estes valores constituem, uma vez mais, um sinal claro dos riscos subjacentes à utilização de tábuas de mortalidade desajustadas face ao comportamento da população em análise.

Por fim, na Figura 7.7.9 representamos a evolução do preço de uma renda vitalícia  $a_{65}(t)$  calculado segundo as ópticas transversal e longitudinal no período  $t \in [1970, 2004]$ , com  $i = 3\%$ . Como se observa, o diferencial entre os valores de  $a_{65}(t)$  calculados segundo as duas ópticas é relativamente estável, sobretudo a partir do final da década de 80.

## 7.8. Intervalos de confiança

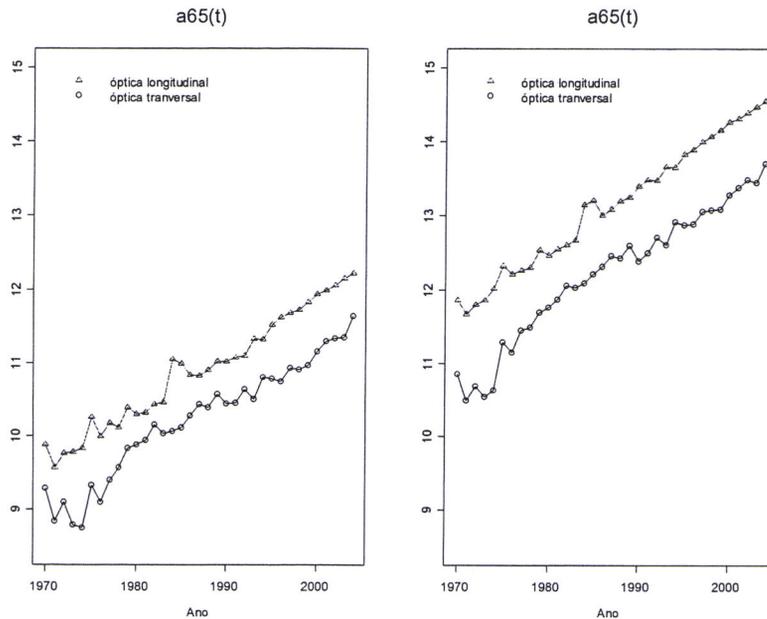


Figura 7.7.9: Evolução do preço de uma renda vitalícia  $a_{65}(t)$  calculado segundo as ópticas transversal e longitudinal no período  $t \in [1970, 2004]$ , com  $i = 3\%$  (homens à esquerda e mulheres à direita)

## 7.8 Intervalos de confiança

Em previsão, é indispensável fornecer informação sobre a variabilidade das quantidades projectadas. No caso em apreço, ela permitirá ao actuário avaliar o grau de confiança a atribuir às estimativas pontuais do prémio puro de uma renda vitalícia, facultando-lhe igualmente informação valiosa no momento de definir a margem de segurança adequada para efeitos de *pricing*, ou mesmo no momento de optar por uma solução de cobertura do risco de longevidade adequada (e.g., por um contrato de resseguro). A ferramenta mais indicada neste caso corresponde ao intervalo de confiança para a estimativa pontual. Neste contexto, o objectivo desta secção é o de construir intervalos de confiança para os indicadores de mortalidade (e.g.,  $e_x(t)$  e  $a_x(t)$ ) a partir do modelo de Poisson log-bilinear clássico.

Os modelos em análise não permitem a derivação analítica dos intervalos de confiança desejados. Esta impossibilidade é justificada por duas ordens de razões. Em primeiro lugar, na construção dos intervalos é necessário combinar duas fontes de variabilidade

bastante distintas: uma primeira, constituída pelos erros de estimação dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ ; uma segunda, associada aos erros de previsão do índice temporal  $\kappa_t$ . A segunda justificação prende-se com o facto dos indicadores que nos interessam serem funções não-lineares relativamente complexas, quer do modelo de Poisson, quer do modelo de previsão ARIMA adoptado. Em face destes constrangimentos, a derivação dos intervalos de confiança será feita com recurso a métodos de simulação.

### 7.8.1 Método de Simulação de Monte Carlo

Nesta secção, analisamos os resultados proporcionados pela aplicação do método de Simulação de Monte Carlo descrito na Secção 6.9.7 do Capítulo 6.

#### Matriz de informação de Fisher

De modo a simplificar a notação na caracterização dos elementos da matriz de informação de Fisher, o número esperado de óbitos de indivíduos com idade  $x$  no decurso do ano  $t$  de um total de  $E_{x,t}$  indivíduos expostos ao risco é definido de forma ligeiramente diferente. Formalmente, seja

$$\begin{aligned}\lambda_{x,t} &= E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) \\ &= E_{x,t} \exp \left[ \left( \sum_{y=x_{\min}}^{x_{\max}} a_{xy} \alpha_y \right) + \left( \sum_{y=x_{\min}}^{x_{\max}} b_{xy} \beta_y \right) + \left( \sum_{r=t_{\min}}^{t_{\max}-1} k_{tr} \kappa_r \right) \right],\end{aligned}$$

onde  $a_{xy}$ ,  $b_{xy}$  e  $k_{tr}$  são elementos de três matrizes de ponderação  $\mathbf{A}$  e  $\mathbf{B}$  de dimensão  $(\mathcal{N} \times \mathcal{N})$ ,  $\mathcal{N} = (x_{\max} - x_{\min} + 1)$ , e  $\mathbf{K}$  de dimensão  $\mathcal{T} \times (t_{\max} - t_{\min})$ ,  $\mathcal{T} = (t_{\max} - t_{\min} + 1)$ , cujos elementos estão associados aos parâmetros do modelo de Poisson. Para facilitar a derivação da matriz de informação de Fisher, introduzimos uma reparametrização do modelo de Poisson que passa, nomeadamente, pela substituição da restrição  $\sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1$  por

$$\beta_{x_{\min}} = 1. \tag{7.8.1}$$

Os novos estimadores são facilmente obtidos por transformação de variável

$$\begin{aligned}\hat{\kappa}_t &\leftarrow \hat{\kappa}_t \hat{\beta}_{x_{\min}} \\ \hat{\beta}_x &\leftarrow \frac{\hat{\beta}_x}{\hat{\beta}_{x_{\min}}}.\end{aligned}$$

Em resultado, a satisfação das restrições (6.9.3) e (7.8.1) impõe a seguinte definição

para as matrizes  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{K}$

$$\mathbf{A} = \mathbf{B} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \text{ e } \mathbf{K} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ -1 & \dots & -1 \end{pmatrix}$$

Note-se que fixar  $\kappa_{t_{\max}} = -1$  equivale a afirmar que  $\kappa_{t_{\max}} = -\sum_{r=t_{\min}}^{t_{\max}-1} \kappa_r$ , de maneira a respeitar as restrições de identificação. Esta é igualmente a razão pela qual a matriz  $\mathbf{K}$  tem apenas  $(t_{\max} - t_{\min})$  colunas em vez de  $\mathcal{T}$ .

Dito isto, se designarmos por  $\mathcal{V}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa})$  a função de verosimilhança, os elementos da matriz de informação de Fisher relativos aos parâmetros  $\alpha_{x_{\min}}, \dots, \alpha_{x_{\max}}, \beta_{x_{\min}+1}, \dots, \beta_{x_{\max}}$  e  $\kappa_{t_{\min}}, \dots, \kappa_{t_{\max}-1}$  são definidos por

$$\begin{aligned} -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\alpha_y d\alpha_{y'}} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} a_{xy} a_{xy'} \\ -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\beta_y d\beta_{y'}} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} (\kappa_t b_{xy}) (\kappa_t b_{xy'}) \\ -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\kappa_r d\kappa_{r'}} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} (\beta_x k_{tr}) (\beta_x k_{tr'}) \\ -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\alpha_y d\beta_{y'}} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} a_{xy} (\kappa_t b_{xy'}) \\ -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\alpha_y d\kappa_r} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} a_{xy} (\beta_x k_{tr}) \\ -\mathbb{E} \left( \frac{d^2 \ln \mathcal{V}}{d\beta_y d\kappa_r} \right) &= \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \lambda_{x,t} (t) (\kappa_t b_{xy}) (\beta_x k_{tr}) \end{aligned}$$

Se substituirmos  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$  pelas suas estimativas  $\tilde{\alpha}_x$ ,  $\tilde{\beta}_x$  e  $\tilde{\kappa}_t$  ajustadas de modo a respeitar (7.8.1), obtemos uma projecção da matriz de informação de Fisher  $\hat{\mathcal{I}}$ . Para retornar à parametrização original, procedemos à seguinte transformação

$$\begin{aligned} \hat{\kappa}_t &\leftarrow \tilde{\kappa}_t \tilde{\beta}_\xi, & \tilde{\beta}_\xi &= \sum_{x=x_{\min}}^{x_{\max}} \tilde{\beta}_x \\ \hat{\beta}_x &\leftarrow \frac{\tilde{\beta}_x}{\tilde{\beta}_\xi} \end{aligned}$$

## Resultados

A aplicação prática do método de Simulação de Monte Carlo (MSMC) consubstanciou-se na geração de  $M = 1000$  simulações aleatórias segundo os procedimentos descritos no Capítulo 6. Para tal, recorreremos a uma rotina informática construída e implementada no *package* SAS (Release 9.1).

No Quadro 7.8.1 reportamos os valores médios e os percentis empíricos para  $\hat{e}_x$  (2004) e  $a_x$  (2004),  $x = 0, 65$ . Os intervalos de confiança empíricos ao nível de 90% são definidos pelos intervalos situados entre os percentis  $q_{0.05}$  e  $q_{0.95}$ . Nas Figuras 7.8.1 e 7.8.2 representamos os histogramas da distribuição empírica dos indicadores no sentido de proporcionar uma percepção gráfica sobre os intervalos de confiança. Nestes, as linhas verticais a tracejado representam os percentis empíricos  $q_{0.05}$  e  $q_{0.95}$ , respectivamente.

| Homens                    |            |            |            |                            |            |            |            |
|---------------------------|------------|------------|------------|----------------------------|------------|------------|------------|
| $\hat{e}_0$ (2004)        | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_0$ (2004)               | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 83.63                     | 82.08      | 83.68      | 85.06      | 29.87                      | 29.71      | 29.88      | 30.03      |
| Margem erro: 1.49 (1.78%) |            |            |            | Margem erro: 0.16 (0.54%)  |            |            |            |
| $\hat{e}_{65}$ (2004)     | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}$ (2004)            | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 17.12                     | 16.54      | 17.12      | 17.64      | 12.34                      | 11.99      | 12.36      | 12.64      |
| Margem erro: 0.55 (3.21%) |            |            |            | Margem erro: 0.325 (2.64%) |            |            |            |

| Mulheres                   |            |            |            |                            |            |            |            |
|----------------------------|------------|------------|------------|----------------------------|------------|------------|------------|
| $\hat{e}_0$ (2004)         | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_0$ (2004)               | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 90.74                      | 90.24      | 90.76      | 91.19      | 30.73                      | 30.67      | 30.73      | 30.77      |
| Margem erro: 0.475 (0.52%) |            |            |            | Margem erro: 0.05 (0.16%)  |            |            |            |
| $\hat{e}_{65}$ (2004)      | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}$ (2004)            | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 20.98                      | 20.63      | 20.98      | 21.29      | 14.61                      | 14.42      | 14.61      | 14.79      |
| Margem erro: 0.28 (1.33%)  |            |            |            | Margem erro: 0.185 (1.27%) |            |            |            |

Quadro 7.8.1: Valores médios e percentis empíricos de  $e_x$ (2004) e  $a_x$ (2004),  $x = 0, 65$ , gerados pelo método de Simulação de Monte Carlo, onde  $Q_\alpha$  denota o percentil empírico de ordem  $\alpha$

Da análise dos valores destacamos, desde logo, o facto dos intervalos de confiança se revelarem mais estreitos para as mulheres do que para os homens. Este facto é explicado, em boa parte, pelo número mais elevado de termos residuais incluídos no modelo de previsão escolhido para projectar a tendência temporal  $\kappa_t$  na mortalidade masculina (modelo ARIMA(0, 1, 1)). Em consequência, as projecções  $\kappa_t^m$  geradas na etapa 3 da Simulação de Monte Carlo possuem um carácter aleatório mais pronunciado.

## 7.8. Intervalos de confiança

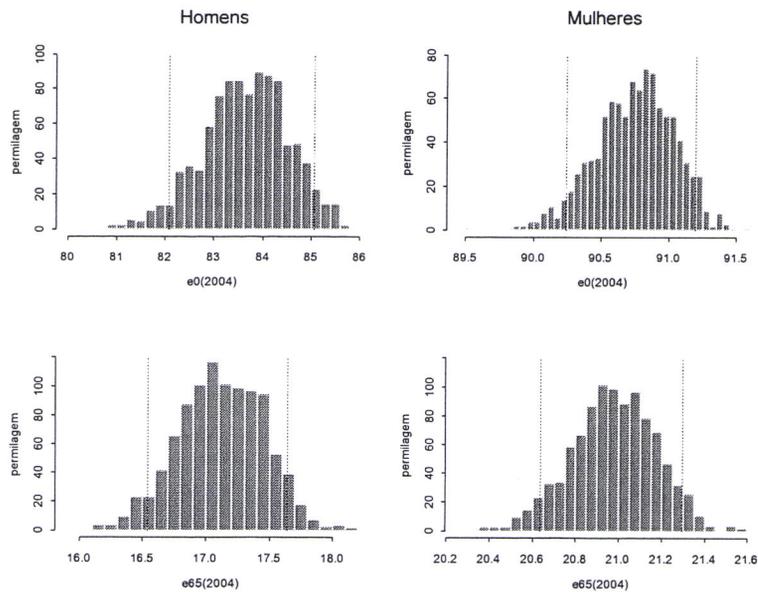


Figura 7.8.1: Histogramas dos valores empíricos de  $e_0(2004)$  e  $e_{65}(2004)$ , MSMC

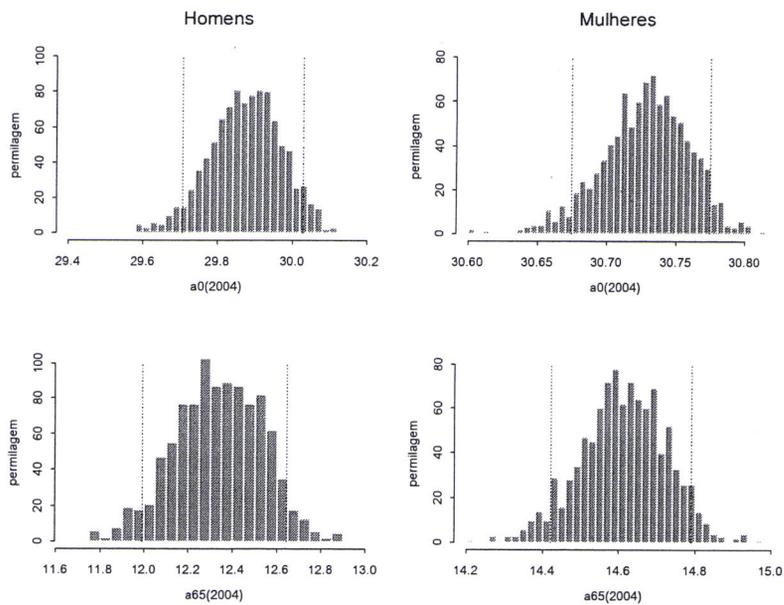


Figura 7.8.2: Histogramas dos valores empíricos de  $a_0(2004)$  e  $a_{65}(2004)$ , MSMC

Por outro lado, constata-se que as estimativas pontuais geradas pelo modelo de Poisson clássico e pelo método Lee-Carter se encontram claramente no interior dos intervalos de confiança a 90% estimados. Mais, as estimativas pontuais aproximam-se bastante dos valores médios da distribuição empírica gerada pela Simulação de Monte Carlo. Este facto prova que a diferença entre o ajustamento proporcionado pelos dois modelos não é muito significativa, pese embora a maior flexibilidade do modelo de Poisson.

Como se constata, os intervalos de confiança estimados (e as correspondentes margens de erro) podem ser considerados relativamente estreitos em ambos os sexos. Este facto é explicado, em boa medida, pelo comportamento assintótico assumido no modelo. Com efeito, na derivação dos intervalos considerámos como primeira fonte de variabilidade o risco subjacente à estimação dos parâmetros  $\alpha_x$ ,  $\beta_x$  e  $\kappa_t$ , admitindo em seguida que os estimadores de máxima verosimilhança seguem aproximadamente uma distribuição Normal Multivariada. Sendo esta uma escolha metodológica devidamente fundamentada, não podemos ignorar que ela restringe, de algum modo, o espaço para a ocorrência de valores extremos na distribuição.

As observação das Figuras 7.8.1 e 7.8.2 permite-nos verificar que o perfil gráfico da distribuição empírica dos indicadores é próximo do da tradicional distribuição gaussiana, com as estimativas pontuais do modelo de Poisson a situarem-se bem no centro dos intervalos de confiança empíricos.

### 7.8.2 Parametric bootstrap

Nesta secção, analisamos os resultados numéricos gerados pela aplicação do método *parametric bootstrap* (com base no número de óbitos), considerando um total de  $B = 1000$  realizações aleatórias. Para tal, recorreremos a uma rotina informática construída e implementada no *package* SAS (Release 9.1).

O Quadro 7.8.2 contém informação sobre os valores médios e sobre os percentis empíricos de  $\hat{e}_x(2004)$  e  $a_x(2004)$ ,  $x = 0, 65$ . As Figuras 7.8.3 e 7.8.4 permitem novamente visualizar os detalhes sobre a distribuição empírica dos indicadores analisados. Constata-se que os intervalos de confiança para a população do sexo feminino são indistinguíveis dos obtidos na secção anterior. Pelo contrário, os intervalos masculinos, sem apresentarem maior amplitude, são ligeiramente deslocados para a esquerda quando comparados com os gerados pelo método de Simulação de Monte Carlo.

Para termos uma ideia da variabilidade dos quocientes prospectivos simulados pelo método *parametric bootstrap*, representamos na Figura 7.8.5 100 realizações de  $q_{0+k}(2004+k)$  ( $k = 0, 1, 2, \dots$ ), que incidirão sobre os indivíduos nascidos em 2004.

7.8. Intervalos de confiança

| Homens                    |            |            |            |                            |            |            |            |
|---------------------------|------------|------------|------------|----------------------------|------------|------------|------------|
| $\hat{e}_0(2004)$         | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_0(2004)$                | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 83.60                     | 81.91      | 83.68      | 84.99      | 29.87                      | 29.68      | 29.88      | 30.02      |
| Margem erro: 1.54 (1.84%) |            |            |            | Margem erro: 0.17 (0.60%)  |            |            |            |
| $\hat{e}_{65}(2004)$      | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}(2004)$             | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 17.10                     | 16.50      | 17.14      | 17.66      | 12.33                      | 11.96      | 12.35      | 12.65      |
| Margem erro: 0.58 (3.39%) |            |            |            | Margem erro: 0.345 (2.80%) |            |            |            |

| Mulheres                   |            |            |            |                            |            |            |            |
|----------------------------|------------|------------|------------|----------------------------|------------|------------|------------|
| $\hat{e}_0(2004)$          | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_0(2004)$                | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 90.75                      | 90.26      | 90.77      | 91.23      | 30.73                      | 30.67      | 30.73      | 30.78      |
| Margem erro: 0.485 (0.53%) |            |            |            | Margem erro: 0.055 (0.18%) |            |            |            |
| $\hat{e}_{65}(2004)$       | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}(2004)$             | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 20.98                      | 20.64      | 20.99      | 21.28      | 14.61                      | 14.42      | 14.62      | 14.78      |
| Margem erro: 0.32 (1.53%)  |            |            |            | Margem erro: 0.18 (1.23%)  |            |            |            |

Quadro 7.8.2: Valores médios e percentis empíricos de  $e_x(2004)$  e  $a_x(2004)$ ,  $x = 0, 65$ , simulados pelo método parametric Bootstrap, onde  $Q_\alpha$  denota o percentil de ordem  $\alpha$

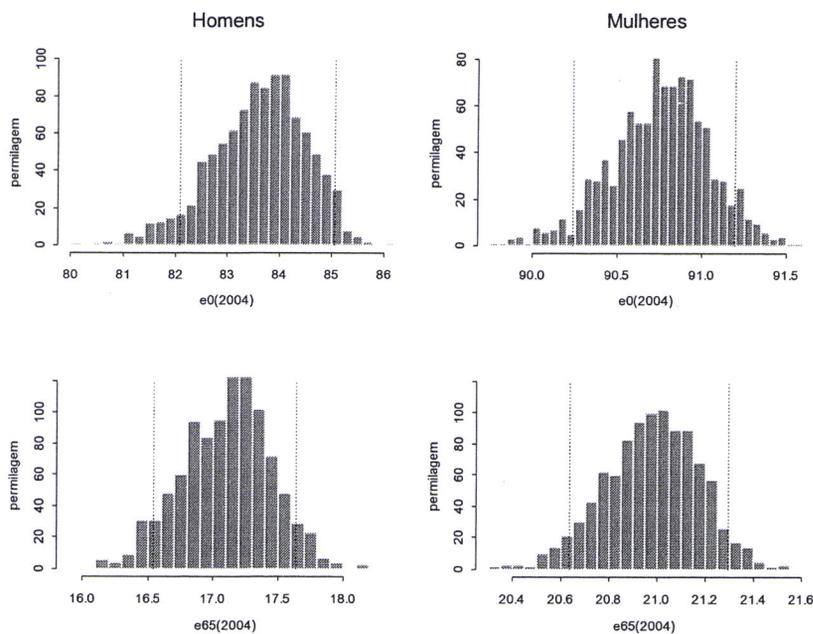


Figura 7.8.3: Histogramas dos valores de  $e_0(2004)$  e  $e_{65}(2004)$ , Parametric Bootstrap

## 7.8. Intervalos de confiança

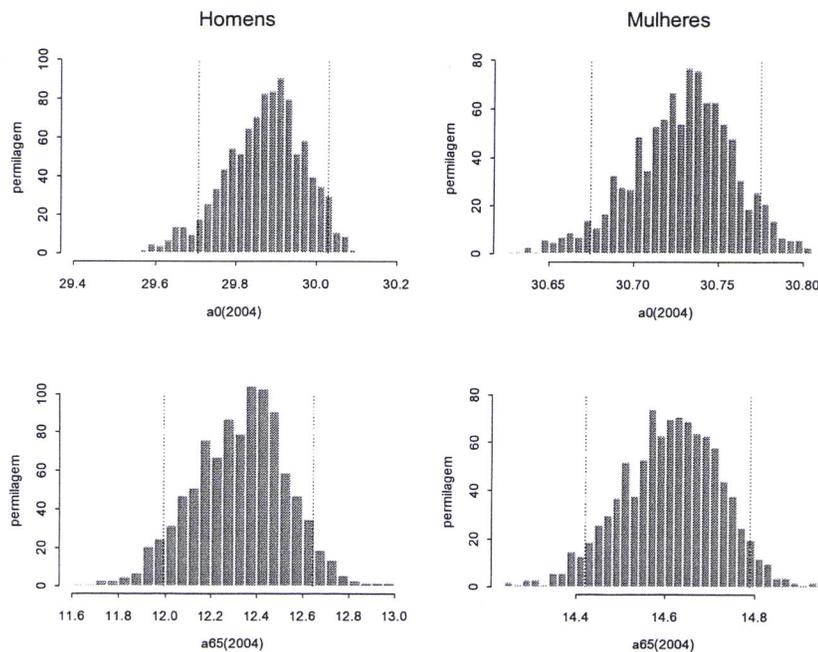


Figura 7.8.4: Histogramas dos valores de  $a_0(2004)$  e  $a_{65}(2004)$ , Parametric Bootstrap

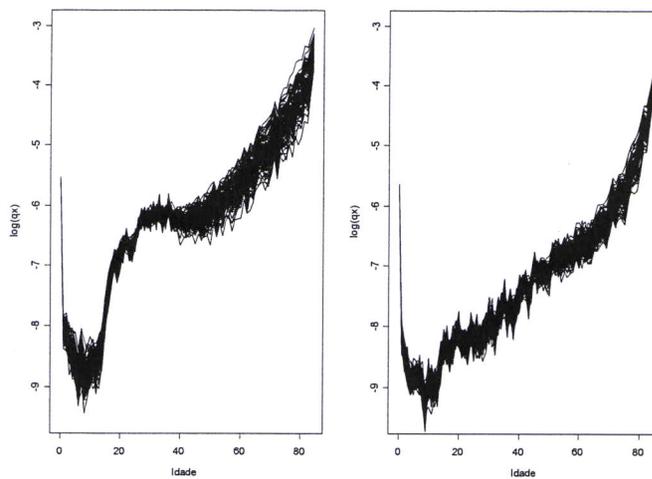


Figura 7.8.5: Quocientes de mortalidade  $q_{0+k}(2004+k)$  simulados pelo método parametric bootstrap (homens à esquerda e mulheres à direita)

Como se observa, a volatilidade das projecções é importante, ligeiramente mais significativa no caso dos homens em todas as idades, exceptuando o segmento relativo à bossa de acidentes onde a oscilação é menor em razão dos valores mais reduzidos do parâmetro  $\beta_x$ . Convém referir que uma boa parte desta variabilidade nos quocientes não se repercute depois no cálculo dos indicadores sintéticos inscritos no Quadro 7.8.2, em virtude do processo de actualização ou desconto.

## 7.9 Conclusão

O cálculo dos prémios puros nos contratos de seguros do ramo vida baseia-se, regra geral, no princípio da equivalência, segundo o qual o valor esperado das prestações a pagar pela seguradora deve ser igual ao valor esperado dos prémios puros pagos pela pessoa segura. Tradicionalmente, os actuários fundamentam o cálculo dos prémios em tábuas de mortalidade contemporâneas, relativas a gerações fictícias, residindo a principal vantagem desta abordagem no facto de ela apenas requerer a observação da população em causa durante um número reduzido de anos. O reverso da medalha é o de que uma tábua desta natureza mistura múltiplas gerações não descrevendo, na prática, o comportamento da mortalidade de nenhuma delas. Em particular, os valores característicos assim calculados (e.g., a esperança de vida) não têm uma interpretação concreta, nem correspondem a uma geração em particular.

Neste capítulo, examinámos num primeiro momento as grandes tendências na evolução da mortalidade da população portuguesa no período entre 1930-31 e 2002-03, adoptando uma abordagem transversal com base em tábuas contemporâneas publicadas pelo INE. Da análise dos diferentes indicadores destacamos, em ambos os sexos, a crescente concentração dos óbitos em torno da moda da distribuição, localizada nas idades adultas, uma progressiva rectangularização da função de sobrevivência, a expansão da longevidade humana com aumentos significativos na esperança de vida em todas as idades, em particular à nascença, a importância do fenómeno “bossa” de acidentes, em especial na população masculina, e o diferencial de longevidade favorável às mulheres, que ultrapassa actualmente os 6 anos de idade à nascença.

Num segundo momento, desenvolvemos as bases que sustentam a adopção de uma abordagem longitudinal no estudo da mortalidade. Em particular, testámos um conjunto de métodos de projecção e de elaboração de tábuas prospectivas, avaliando a sua capacidade de ajustamento aos dados, comparando em particular o desempenho do tradicional método Lee-Carter com as diferentes versões do modelo de Poisson. A modelação das séries temporais permitiu a realização de previsão nos modelos e a elaboração, pioneira,

de tábuas prospectivas para a população portuguesa.

A comparação entre os resultados proporcionados pelas abordagens transversal e longitudinal permitiu detectar diferenças significativas entre os valores da esperança de vida. Com efeito, a utilização das tábuas prospectivas permite-nos estimar em 83.30 e 90.63 anos para homens e mulheres, respectivamente, a “verdadeira” esperança de vida à nascença para um indivíduo nascido em 2004, valores que contrastam com os 74.55 e 81.05 anos calculados segundo a tradicional visão transversal. Significa isto que, a manterem-se as tendências de decréscimo da mortalidade verificadas entre 1970 e 2004, a abordagem tradicional subestima em 8.75 e 9.58 anos para homens e mulheres, respectivamente, a verdadeira esperança de vida à nascença. Idênticas conclusões são retiradas quando se investigam as estimativas da esperança de vida aos 65 anos e os prémios puros de rendas vitalícias segundo as duas abordagens.

Para quantificar a incerteza em torno das quantidades projectadas e estimar a importância do risco de longevidade, são derivados intervalos de confiança empíricos, usando os métodos de Simulação de Monte Carlo e *parametric bootstrap* (com base no número de óbitos). A aplicação destas técnicas permitiu a obtenção, nos vários indicadores actuariais (quocientes ou taxas de mortalidade, esperança de vida, prémio de rendas vitalícias), de medidas que quantificam a variabilidade da projecção.

Os resultados obtidos neste estudo confirmam que basear a tarificação numa tábua transversal equivale, implicitamente, a assumir que a mortalidade é constante no tempo, uma hipótese rejeitada pela evidência empírica observada na população portuguesa. Neste sentido, consideramos que esta realidade deve preocupar seriamente as companhias que operam o ramo vida e as entidades reguladoras do sector, constituindo a adopção de uma tábua prospectiva ou dinâmica uma solução indispensável para minimizar a importância dos riscos de longevidade.

## 7.A Anexos

## 7.A.1 Estimativas dos parâmetros do modelo de Poisson

| Idade | Homens     |           | Mulheres   |           | Idade | Homens     |           | Mulheres   |           |
|-------|------------|-----------|------------|-----------|-------|------------|-----------|------------|-----------|
|       | $\alpha_x$ | $\beta_x$ | $\alpha_x$ | $\beta_x$ |       | $\alpha_x$ | $\beta_x$ | $\alpha_x$ | $\beta_x$ |
| 0     | -4.10743   | 0.04955   | -4.31514   | 0.03692   | 43    | -5.48633   | 0.00728   | -6.29567   | 0.00954   |
| 1     | -6.44733   | 0.05092   | -6.59634   | 0.04093   | 44    | -5.41286   | 0.00748   | -6.22204   | 0.00804   |
| 2     | -6.85158   | 0.03671   | -7.10901   | 0.03158   | 45    | -5.33216   | 0.00820   | -6.11832   | 0.00660   |
| 3     | -7.13812   | 0.03412   | -7.40539   | 0.02898   | 46    | -5.28706   | 0.00794   | -6.02785   | 0.00865   |
| 4     | -7.29643   | 0.03046   | -7.65501   | 0.02667   | 47    | -5.19578   | 0.00938   | -5.96874   | 0.00828   |
| 5     | -7.44896   | 0.02907   | -7.79602   | 0.02259   | 48    | -5.10115   | 0.00949   | -5.86740   | 0.00878   |
| 6     | -7.59172   | 0.03113   | -7.85649   | 0.02027   | 49    | -5.02909   | 0.00963   | -5.82817   | 0.00907   |
| 7     | -7.54206   | 0.02424   | -7.97292   | 0.01653   | 50    | -4.94714   | 0.00976   | -5.72588   | 0.00977   |
| 8     | -7.71731   | 0.02527   | -8.08482   | 0.01602   | 51    | -4.89337   | 0.00916   | -5.66479   | 0.00796   |
| 9     | -7.68698   | 0.02405   | -8.16748   | 0.01989   | 52    | -4.81440   | 0.01106   | -5.57036   | 0.00864   |
| 10    | -7.66856   | 0.02014   | -8.17621   | 0.01250   | 53    | -4.71693   | 0.01011   | -5.50262   | 0.00900   |
| 11    | -7.68131   | 0.01802   | -8.12277   | 0.01473   | 54    | -4.64992   | 0.01091   | -5.45539   | 0.00946   |
| 12    | -7.70597   | 0.01848   | -8.18658   | 0.01265   | 55    | -4.56376   | 0.01024   | -5.36438   | 0.00941   |
| 13    | -7.56869   | 0.01828   | -8.13454   | 0.01348   | 56    | -4.49183   | 0.01052   | -5.28036   | 0.00978   |
| 14    | -7.40622   | 0.01768   | -7.90882   | 0.00976   | 57    | -4.39636   | 0.01113   | -5.22312   | 0.00994   |
| 15    | -7.13148   | 0.01228   | -7.82966   | 0.00686   | 58    | -4.32333   | 0.01111   | -5.09845   | 0.01058   |
| 16    | -6.81321   | 0.01094   | -7.74849   | 0.00772   | 59    | -4.22587   | 0.01105   | -5.01729   | 0.00991   |
| 17    | -6.59859   | 0.00894   | -7.74243   | 0.01008   | 60    | -4.13252   | 0.01140   | -4.91636   | 0.01068   |
| 18    | -6.42548   | 0.00755   | -7.67165   | 0.00813   | 61    | -4.05645   | 0.01025   | -4.84609   | 0.01042   |
| 19    | -6.34235   | 0.01053   | -7.60066   | 0.00719   | 62    | -3.96588   | 0.01194   | -4.74368   | 0.01109   |
| 20    | -6.31585   | 0.00720   | -7.60398   | 0.00614   | 63    | -3.86797   | 0.01128   | -4.64254   | 0.01158   |
| 21    | -6.29839   | 0.00584   | -7.53918   | 0.00918   | 64    | -3.79395   | 0.01183   | -4.53166   | 0.01175   |
| 22    | -6.34077   | 0.00318   | -7.58208   | 0.00842   | 65    | -3.68853   | 0.01121   | -4.43443   | 0.01160   |
| 23    | -6.32543   | 0.00513   | -7.53717   | 0.00878   | 66    | -3.61407   | 0.01056   | -4.33629   | 0.01159   |
| 24    | -6.31470   | 0.00516   | -7.43395   | 0.00646   | 67    | -3.51349   | 0.01123   | -4.22088   | 0.01155   |
| 25    | -6.28886   | 0.00408   | -7.47383   | 0.00788   | 68    | -3.41594   | 0.01177   | -4.10214   | 0.01198   |
| 26    | -6.28544   | 0.00107   | -7.38241   | 0.00841   | 69    | -3.32022   | 0.01215   | -3.97576   | 0.01279   |
| 27    | -6.27316   | -0.00090  | -7.34411   | 0.00950   | 70    | -3.20933   | 0.01192   | -3.83522   | 0.01204   |
| 28    | -6.24969   | -0.00093  | -7.34135   | 0.00841   | 71    | -3.12754   | 0.01089   | -3.72689   | 0.01162   |
| 29    | -6.24119   | -0.00059  | -7.27234   | 0.00811   | 72    | -3.01831   | 0.01283   | -3.59831   | 0.01259   |
| 30    | -6.17804   | 0.00015   | -7.21534   | 0.00726   | 73    | -2.91823   | 0.01168   | -3.47652   | 0.01214   |
| 31    | -6.16882   | -0.00067  | -7.12739   | 0.00752   | 74    | -2.80663   | 0.01203   | -3.31679   | 0.01299   |
| 32    | -6.12429   | 0.00060   | -7.10735   | 0.00977   | 75    | -2.70036   | 0.01240   | -3.18377   | 0.01349   |
| 33    | -6.08828   | -0.00018  | -7.00011   | 0.00749   | 76    | -2.59783   | 0.01171   | -3.06536   | 0.01248   |
| 34    | -6.04011   | 0.00133   | -6.94562   | 0.00718   | 77    | -2.49340   | 0.01231   | -2.94122   | 0.01205   |
| 35    | -5.99397   | 0.00202   | -6.85251   | 0.00666   | 78    | -2.39219   | 0.01179   | -2.80455   | 0.01229   |
| 36    | -5.94799   | 0.00108   | -6.83182   | 0.00900   | 79    | -2.28135   | 0.01183   | -2.66799   | 0.01251   |
| 37    | -5.89146   | 0.00291   | -6.75983   | 0.00799   | 80    | -2.17014   | 0.01243   | -2.52920   | 0.01207   |
| 38    | -5.83148   | 0.00432   | -6.65931   | 0.01003   | 81    | -2.07996   | 0.01121   | -2.42138   | 0.01088   |
| 39    | -5.77171   | 0.00500   | -6.57240   | 0.00908   | 82    | -1.97253   | 0.01232   | -2.27932   | 0.01206   |
| 40    | -5.69557   | 0.00526   | -6.48188   | 0.00784   | 83    | -1.86871   | 0.01183   | -2.16090   | 0.01126   |
| 41    | -5.64555   | 0.00626   | -6.43727   | 0.00755   | 84    | -1.76529   | 0.01207   | -2.03479   | 0.01158   |
| 42    | -5.55832   | 0.00709   | -6.35044   | 0.00740   |       |            |           |            |           |

Quadro 7.A.1: Estimativas dos parâmetros  $\alpha_x$  e  $\beta_x$  do modelo de Poisson

---

| Ano  | Homens   | Mulheres |
|------|----------|----------|
| 1970 | 27.9919  | 37.2707  |
| 1971 | 27.1893  | 36.8966  |
| 1972 | 21.4907  | 30.2154  |
| 1973 | 24.5586  | 34.3471  |
| 1974 | 21.5383  | 28.4966  |
| 1975 | 20.1871  | 24.0109  |
| 1976 | 19.3789  | 23.654   |
| 1977 | 16.1934  | 18.4823  |
| 1978 | 14.0073  | 17.7053  |
| 1979 | 10.0880  | 13.2522  |
| 1980 | 10.0183  | 11.6405  |
| 1981 | 7.6184   | 9.5543   |
| 1982 | 4.4546   | 5.0368   |
| 1983 | 4.5547   | 5.6693   |
| 1984 | 3.9801   | 2.9098   |
| 1985 | 3.4656   | 1.8482   |
| 1986 | 0.5712   | -0.3494  |
| 1987 | -1.8285  | -3.0761  |
| 1988 | -1.9484  | -3.5484  |
| 1989 | -5.5121  | -7.3859  |
| 1990 | -3.8161  | -4.015   |
| 1991 | -4.1225  | -5.4789  |
| 1992 | -7.8395  | -10.6824 |
| 1993 | -6.3165  | -9.3526  |
| 1994 | -12.8471 | -16.4262 |
| 1995 | -12.1167 | -15.2275 |
| 1996 | -11.0962 | -15.7167 |
| 1997 | -14.8403 | -18.4289 |
| 1998 | -15.5427 | -20.2696 |
| 1999 | -17.1130 | -21.1308 |
| 2000 | -20.1513 | -25.0894 |
| 2001 | -22.8932 | -27.7219 |
| 2002 | -23.7533 | -29.7903 |
| 2003 | -25.4111 | -29.8062 |
| 2004 | -30.1381 | -37.4941 |

Quadro 7.A.2: Estimativas do parâmetro  $\kappa_t$  do modelo de Poisson

## 7.A.2 Evolução da esperança de vida à nascença e aos 65 anos

| Ano  | Homens   |       |             |       | Mulheres |       |             |       |
|------|----------|-------|-------------|-------|----------|-------|-------------|-------|
|      | $e_0(t)$ |       | $e_{65}(t)$ |       | $e_0(t)$ |       | $e_{65}(t)$ |       |
|      | Long     | Trans | Long        | Trans | Long     | Trans | Long        | Trans |
| 1970 | 71.99    | 63.21 | 13.16       | 12.21 | 80.37    | 69.32 | 16.24       | 14.53 |
| 1971 | 73.52    | 63.07 | 12.75       | 11.54 | 82.05    | 69.59 | 15.99       | 13.94 |
| 1972 | 74.61    | 64.84 | 13.02       | 11.90 | 82.97    | 71.12 | 16.18       | 14.24 |
| 1973 | 74.60    | 63.96 | 13.05       | 11.45 | 82.97    | 70.37 | 16.28       | 14.00 |
| 1974 | 75.46    | 64.80 | 13.15       | 11.38 | 83.86    | 71.55 | 16.55       | 14.14 |
| 1975 | 75.77    | 65.09 | 13.73       | 12.27 | 83.93    | 72.56 | 16.99       | 15.23 |
| 1976 | 76.32    | 65.33 | 13.38       | 11.93 | 84.60    | 72.65 | 16.84       | 15.00 |
| 1977 | 76.73    | 66.47 | 13.63       | 12.37 | 85.07    | 73.67 | 16.92       | 15.48 |
| 1978 | 77.09    | 67.10 | 13.55       | 12.64 | 85.06    | 73.94 | 16.99       | 15.55 |
| 1979 | 77.55    | 67.95 | 13.94       | 13.02 | 85.68    | 74.62 | 17.34       | 15.87 |
| 1980 | 77.84    | 67.87 | 13.81       | 13.11 | 85.91    | 74.80 | 17.26       | 16.00 |
| 1981 | 78.47    | 68.16 | 13.84       | 13.20 | 86.23    | 75.17 | 17.38       | 16.16 |
| 1982 | 78.71    | 68.94 | 14.02       | 13.52 | 86.58    | 75.82 | 17.50       | 16.48 |
| 1983 | 78.96    | 68.93 | 14.09       | 13.35 | 86.68    | 75.76 | 17.60       | 16.43 |
| 1984 | 79.23    | 69.11 | 14.98       | 13.38 | 87.10    | 76.12 | 18.37       | 16.52 |
| 1985 | 79.26    | 69.30 | 14.91       | 13.46 | 87.15    | 76.36 | 18.47       | 16.71 |
| 1986 | 79.63    | 69.75 | 14.67       | 13.71 | 87.35    | 76.65 | 18.17       | 16.89 |
| 1987 | 79.85    | 70.18 | 14.66       | 13.93 | 87.59    | 77.02 | 18.31       | 17.11 |
| 1988 | 80.16    | 70.12 | 14.78       | 13.87 | 87.84    | 77.15 | 18.49       | 17.06 |
| 1989 | 80.41    | 70.68 | 14.97       | 14.15 | 88.05    | 77.59 | 18.58       | 17.34 |
| 1990 | 80.58    | 70.45 | 14.99       | 13.95 | 88.33    | 77.34 | 18.84       | 16.99 |
| 1991 | 80.84    | 70.25 | 15.10       | 13.98 | 88.50    | 77.48 | 18.98       | 17.17 |
| 1992 | 81.16    | 70.78 | 15.12       | 14.26 | 88.73    | 78.08 | 18.99       | 17.53 |
| 1993 | 81.36    | 70.80 | 15.48       | 14.05 | 88.90    | 77.95 | 19.30       | 17.36 |
| 1994 | 81.59    | 71.76 | 15.47       | 14.54 | 89.04    | 78.66 | 19.30       | 17.89 |
| 1995 | 81.71    | 71.47 | 15.78       | 14.50 | 89.25    | 78.71 | 19.60       | 17.80 |
| 1996 | 81.93    | 71.42 | 15.93       | 14.44 | 89.46    | 78.77 | 19.71       | 17.83 |
| 1997 | 82.11    | 71.97 | 16.04       | 14.75 | 89.55    | 79.03 | 19.90       | 18.11 |
| 1998 | 82.29    | 72.12 | 16.11       | 14.71 | 89.75    | 79.21 | 20.02       | 18.15 |
| 1999 | 82.50    | 72.40 | 16.28       | 14.79 | 89.91    | 79.39 | 20.16       | 18.16 |
| 2000 | 82.59    | 72.89 | 16.45       | 15.10 | 89.97    | 79.79 | 20.35       | 18.48 |
| 2001 | 82.80    | 73.19 | 16.54       | 15.32 | 90.20    | 80.08 | 20.45       | 18.66 |
| 2002 | 82.95    | 73.55 | 16.65       | 15.38 | 90.27    | 80.31 | 20.60       | 18.85 |
| 2003 | 83.18    | 73.92 | 16.80       | 15.39 | 90.41    | 80.35 | 20.73       | 18.77 |
| 2004 | 83.30    | 74.55 | 16.91       | 15.84 | 90.63    | 81.05 | 20.87       | 19.22 |

Quadro 7.A.3: Evolução da esperança de vida  $e_x(t)$  no período 1970-2004

## Capítulo 8

# Avaliação do Efeito Selecção Adversa no Mercado de Fundos de Pensões em Portugal

*“Nowadays, people know the price of everything and the value of nothing.”*  
(Oscar Wilde)

### 8.1 Introdução

A existência de selecção adversa é um fenómeno típico do mercado de seguros. A teoria económica prevê que no caso de não existirem quaisquer barreiras no acesso dos consumidores aos produtos de seguro (i.e., nos casos em que não existe qualquer filtragem devida a factores como o rendimento, nível de escolaridade, raça, sexo, etc.), a presença de selecção adversa manifesta-se através de taxas de mortalidade superiores para os indivíduos que adquirem seguros de vida e inferiores para aqueles que compram anuidades, relativamente à população em geral (McCarthy e Mitchell, 2003).

A importância da selecção adversa tende igualmente a ser maior nos casos em que os indivíduos compram seguros voluntariamente do que nos casos em que são compelidos a fazê-lo em razão de outras condicionantes (e.g., no caso de um seguro de vida de uma empresa). Nos países em desenvolvimento, em que é comum a existência de uma grande disparidade na distribuição do rendimento, espera-se analogamente que o fenómeno de selecção adversa seja mais significativo do que em sociedades onde a assimetria é menor.

Para avaliar a importância do efeito selecção adversa no contexto do problema de projecção da mortalidade, testamos neste capítulo a hipótese de que a população por-

tuguesa e a população de beneficiários de fundos de pensões, descrita no Capítulo 5, têm perfis de mortalidade semelhantes. Com efeito, constitui dúvida razoável admitir que a população coberta pelo universo dos fundos de pensões a operar no mercado português apresenta características específicas (e.g., condições sócio-económicas dos beneficiários, estado de saúde, níveis de educação e estilos de vida) que a distinguem da população em geral. A confirmar-se esta hipótese, as tábuas prospectivas desenvolvidas no capítulo anterior não podem ser consideradas adequadas para uso no mercado de fundos de pensões, devendo por isso ser ajustadas às características específicas de cada subpopulação.

O capítulo está organizado da seguinte forma. Na Secção 8.2, investigamos a aplicação de um modelo relacional do tipo sugerido por Brass (1971) no contexto de um modelo linear generalizado para estabelecer a ligação entre a mortalidade nas populações portuguesa e de pensionistas. Na Secção 8.3, são derivadas tábuas prospectivas para a população de pensionistas. Na Secção 8.4, estimamos a importância do efeito selecção adversa na esperança de vida e no preço de rendas vitalícias.

Na Secção 8.5, avaliamos a adequação de duas tábuas regulamentares usadas no mercado de seguros em Portugal ao comportamento da mortalidade nas populações portuguesa e de pensionistas. Na Secção 8.6, são derivados intervalos de confiança para as estimativas. Por fim, na Secção 8.7 são resumidas as principais conclusões deste estudo.

## 8.2 Relação entre a mortalidade das populações

O presente estudo sobre a relação entre a mortalidade dos pensionistas e da população portuguesa em geral (POP) assenta nos dados coligidos pelo Instituto de Seguros de Portugal (ISP) para a categoria beneficiários de fundos de pensões (BFP), relativos ao período temporal 1997-2004.

Na Figura 8.2.1 representamos as taxas brutas  $\mu_{x,2004}$  observadas em 2004 na população portuguesa e na carteira de beneficiários de fundos de pensões, em indivíduos com mais de 50 anos de ambos os sexos. Em paralelo, a Figura 8.2.2 põe em evidência a evolução temporal dessas mesmas taxas em idades seleccionadas.

A observação das figuras permite constatar, desde logo, a maior volatilidade do comportamento da mortalidade no caso dos BFP, um facto explicável pela menor dimensão (em termos relativos) da população exposta ao risco. Ressalvadas as diferenças na dimensão das amostras em estudo, detecta-se a existência de comportamentos distintos na mortalidade das duas populações, que não aparentando ter amplitude elevada, assumem algum significado no caso dos indivíduos do sexo masculino.

## 8.2. Relação entre a mortalidade das populações

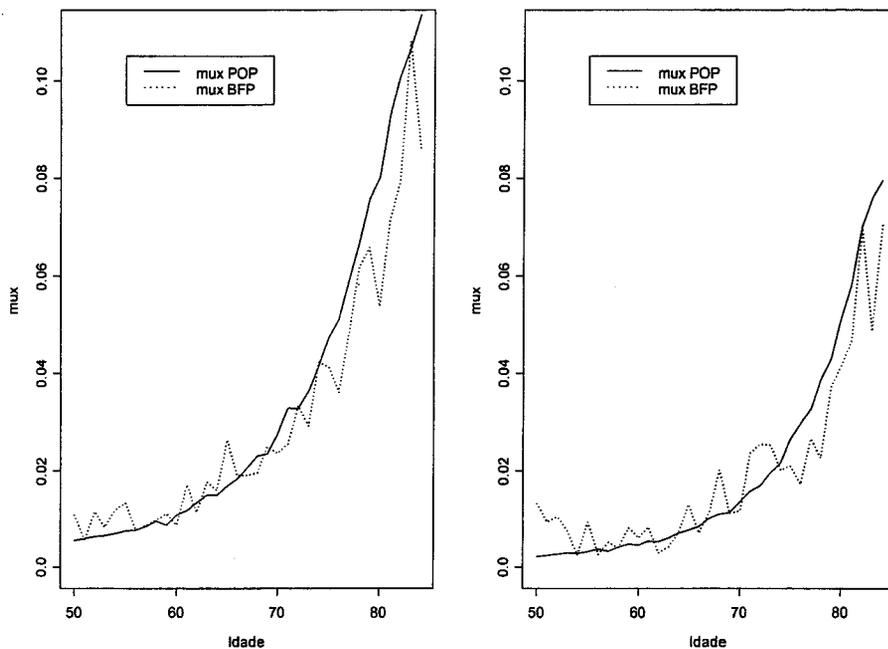


Figura 8.2.1: Taxas brutas  $\mu_{x,2004}$  ( $x \geq 50$ ) para a população portuguesa e para os beneficiários de fundos de pensões (homens à esquerda e mulheres à direita)

Em face das diferenças registadas, é importante adaptar as tábuas prospectivas construídas no capítulo anterior para a população portuguesa de modo a espelhar adequadamente a lei de sobrevivência na carteira de pensionistas.

À primeira vista, poderíamos tentar repetir a análise efectuada no capítulo anterior usando, por exemplo, o modelo de Poisson e os dados estatísticos relativos à população de pensionistas. Contudo, o facto de dispormos de uma série temporal de apenas 8 anos relativa ao período 1997-2004 (contra 35 anos no caso da população portuguesa) impede-nos, por uma questão de significância estatística, de reproduzir o estudo nos mesmos moldes e remete-nos, por isso, para novas soluções.

## 8.2. Relação entre a mortalidade das populações

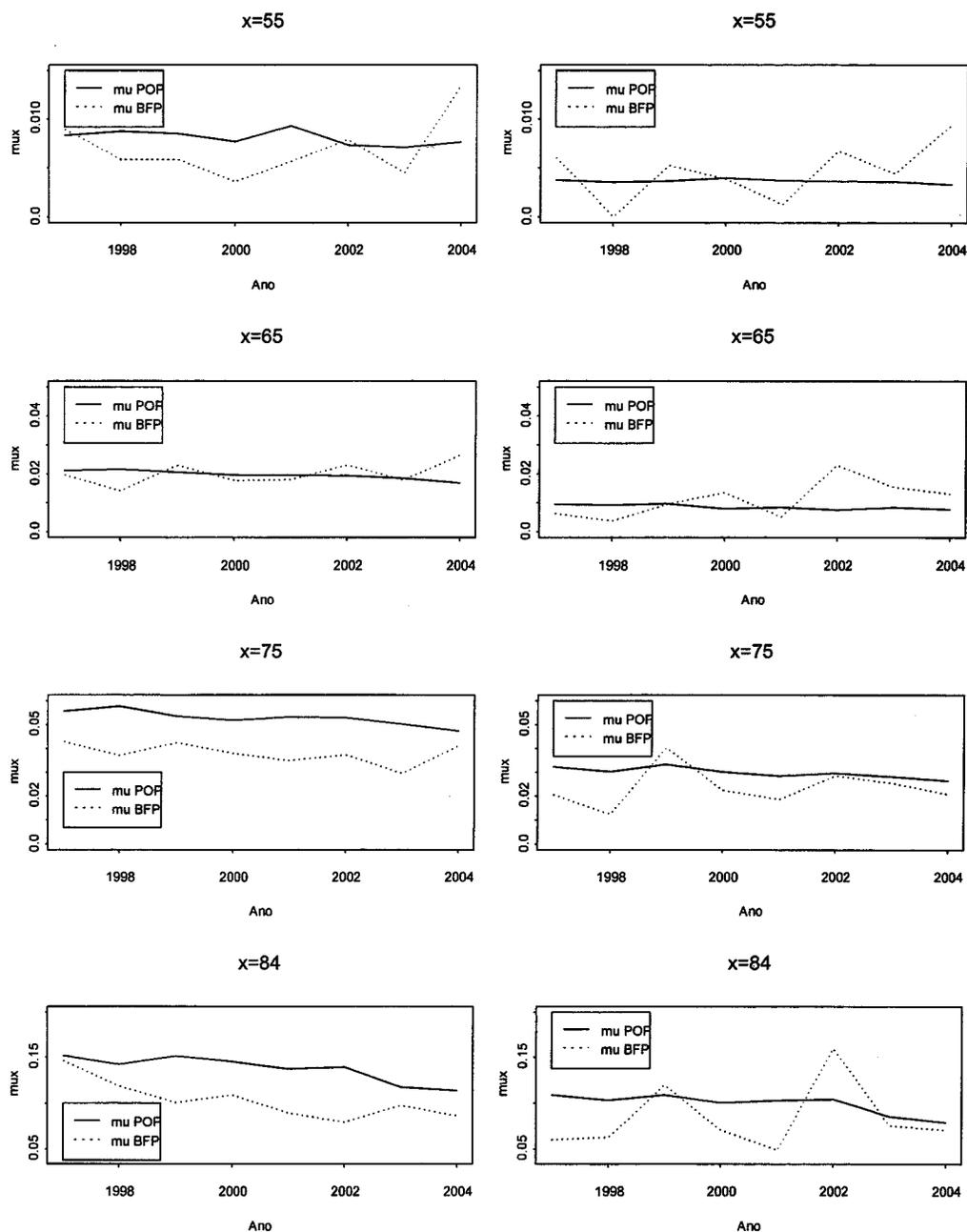


Figura 8.2.2: Taxas brutas  $\mu_{x,t}$  ( $x = 55, 65, 75, 85$ ) para a população portuguesa e para os beneficiários de fundos de pensões (homens à esquerda e mulheres à direita)

### 8.2.1 Modelo relacional de tipo Brass

Face à impossibilidade de aplicar directamente a metodologia usada no capítulo anterior para obter tábuas prospectivas para a população de BFP, a solução encontrada passou por adaptar um modelo relacional do tipo proposto por Brass (1971). A ideia subjacente é a de encontrar uma relação simples (idealmente linear) que associe as características da mortalidade das populações portuguesa e de BFP.

Encontrada essa relação, admitimos de seguida que ela é estável no tempo, o que nos permite derivar tábuas prospectivas para a população de BFP a partir das correspondentes tábuas elaboradas para a população portuguesa. Estas colocam-nos, por fim, em condições de estimar a importância do fenómeno de selecção adversa no mercado de fundos de pensões em Portugal. Formalmente, o modelo relacional é definido por

$$g(\mu_x) = \Psi\left(g(\mu_x^{\text{ref}})\right), \quad (8.2.1)$$

onde  $\Psi(\cdot)$  denota uma função qualquer,  $g(\mu_x)$  denota uma transformação de  $\mu_x$  (e.g., logarítmica, logit) e  $\mu_x^{\text{ref}}$  denota o comportamento de  $\mu_x$  numa população de referência.

### 8.2.2 Modelo log-linear com intercepto variável com a idade

Como ponto de partida para a nossa análise, representamos na Figura 8.2.3 os pares de valores

$$(\ln \hat{\mu}_{x,t}^{POP}, \ln \hat{\mu}_{x,t}^{BFP}), \quad x = 50, \dots, 84 \text{ e } t = 1997, \dots, 2004$$

referentes a observações para a população portuguesa (identificadas pelo índice superior “POP”) e de pensionistas (que reconheceremos pelo índice superior “BFP”), onde

$$\ln \hat{\mu}_{x,t}^{POP} = \hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t$$

denota o logaritmo da taxa  $\mu_{x,t}$  ajustada pelo modelo de Poisson e  $\hat{\mu}_{x,t}^{BFP}$  simboliza a estimativa bruta da taxa de mortalidade para os BFP de idade  $x$  no ano  $t$ . A Figura 8.2.3 deixa transparecer uma tendência relativamente linear entre  $\ln \hat{\mu}_{x,t}^{BFP}$  e  $\ln \hat{\mu}_{x,t}^{POP}$ , com uma inclinação que, numa parte significativas das idades aparenta ser independente da idade  $x$ . Ainda assim, detecta-se alguma variabilidade nas estimativas, sobretudo nas idades onde é menor a dimensão da amostra.

Neste sentido, testámos em primeiro lugar um modelo relacional log-linear com intercepto variável com a idade e inclinação constante, inserido na estrutura de um modelo linear generalizado com distribuição de Poisson.

## 8.2. Relação entre a mortalidade das populações

Em termos formais, seja

$$D_{x,t}^{BFP} \sim \text{Poisson}(E_{x,t}^{BFP} \mu_{x,t}^{BFP}), \quad (8.2.2)$$

onde  $E_{x,t}^{BFP}$  denota a população exposição ao risco (i.e., o número de pensionistas com direito ao recebimento de uma pensão) e

$$\begin{aligned} \ln \mu_{x,t}^{BFP} &= \eta_x + \theta \ln \hat{\mu}_x^{POP}(t) \\ &= \eta_x + \theta(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t), \quad x \geq 50, \end{aligned} \quad (8.2.3)$$

onde  $\eta_x$  e  $\theta$  são parâmetros a estimar (os valores  $\hat{\alpha}_x$ ,  $\hat{\beta}_x$  e  $\hat{\kappa}_t$  são iguais aos estimados para a população portuguesa).

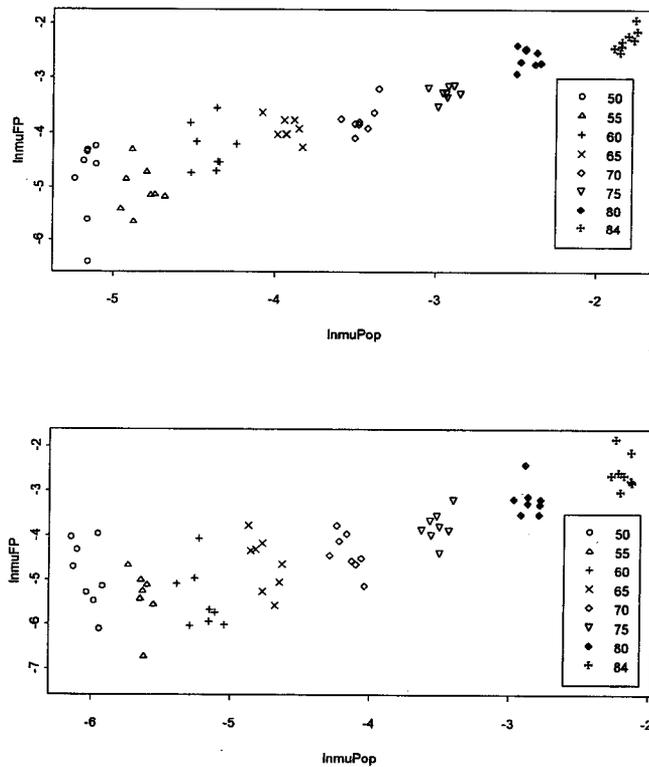


Figura 8.2.3: Relação entre as taxas  $\mu_{x,t}$  das populações portuguesa e de beneficiários de fundos de pensões em idades seleccionadas (homens em cima e mulheres em baixo)

## 8.2. Relação entre a mortalidade das populações

A especificação (8.2.3), que designaremos por modelo Poisson-Relacional  $\mathbf{A}$ ,  $\mathcal{PR}^A$ , admite uma inclinação constante mas estipula que a ordenada na origina varia em função da idade do indivíduo. Note-se, porém, que se assume que os parâmetros não são função do tempo cronológico. Como se trata de um modelo de Poisson linear, é possível recorrer às rotinas de estimação de GLMs implementadas na maioria dos *packages* informáticos. No nosso caso, recorreremos ao procedimento GENMOD<sup>1</sup> do *package* SAS (Release 9.1) para obter as estimativas dos parâmetros  $\eta_x$  e  $\theta$ . Em termos mais precisos, o procedimento GENMOD considera, neste caso, o método de máxima verosimilhança com distribuição de Poisson, *link* logarítmico, variável de resposta  $d_{x,t}^{BFP}$  e termo *offset*  $\ln E_{x,t}^{BFP}$ .

As estimativas de  $\theta$  são reportadas no Quadro 8.2.1.

| Sexo     | $\hat{\theta}$ | $\sigma(\hat{\theta})$ | $\chi^2$ | $p$ -value | IC 95%             |
|----------|----------------|------------------------|----------|------------|--------------------|
| Homens   | 0.1593         | 0.1396                 | 1.30     | 0.2538     | [-0.1143, 0.4328]  |
| Mulheres | -1.1793        | 0.1990                 | 35.13    | <0.0001    | [-1.5692, -0.7893] |

Quadro 8.2.1: Estimativas do parâmetro  $\theta$  do modelo log-linear com intercepto variável com a idade

Como se observa, no caso dos homens não podemos rejeitar a hipótese de que a estimativa de  $\theta$  seja nula.

Na Figura 8.2.4 representamos, por sua vez, as estimativas do parâmetro  $\eta_x$  do modelo (8.2.3) e respectivos intervalos de confiança a 95%. Como se observa, os valores de  $\hat{\eta}_x$  tendem a adoptar um comportamento tendencialmente linear e crescente com a idade, sobretudo no caso da população feminina. Perante esta evidência, e na procura de uma versão simplificada para o modelo em estudo, testámos a pertinência de uma relação linear entre  $\hat{\eta}_x$  e a idade  $x$ , i.e.,

$$\hat{\eta}_x = \phi_0 + \phi_1 x + \epsilon_x, \quad \epsilon_x \sim \text{Nor}(0, \sigma_\epsilon^2) \quad (8.2.4)$$

Os resultados da regressão (8.2.4) são reportados no Quadro 8.2.2. Como se observa, os valores obtidos para o coeficiente de determinação (0.9358 e 0.9628 para homens e mulheres, respectivamente), para a estatística RMSE e no teste  $F$  confirmam a significância estatística da relação linear.<sup>2</sup> A análise das estatísticas de teste e dos respectivos valores de prova permite-nos concluir que todos os parâmetros são significativos ao nível de significância  $\alpha = 0.01$ .

<sup>1</sup>Procedimento de estimação de modelos lineares generalizados (GLM's).

<sup>2</sup>Efectuámos ainda uma regressão linear ponderada pela inversa da variância associada a cada estimador  $\hat{\eta}_x$ . Os resultados obtidos foram sensivelmente iguais.

## 8.2. Relação entre a mortalidade das populações

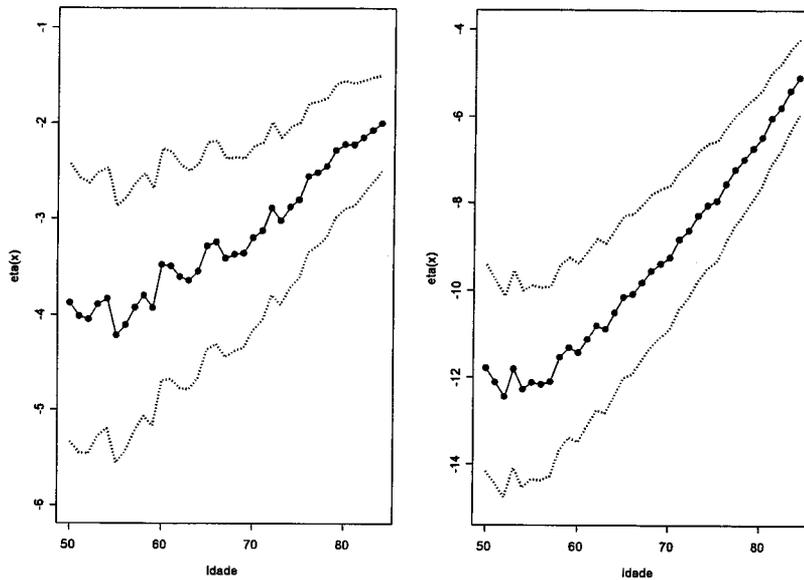


Figura 8.2.4: Estimativas do parâmetro  $\eta_x$  do modelo (8.2.3) e respectivos intervalos de confiança a 95% (homens à esquerda e mulheres à direita)

| HOMENS   |            |         |            |            |
|--|------------|---------|------------|------------|
| Parâmetro  | Estimativa | s.e.    | $t$ -value | $p$ -value |
| $\phi_0$   | -7.42923   | 0.19437 | -38.22     | <.0001     |
| $\phi_1$   | 0.06290    | 0.00287 | 21.93      | <.0001     |
| $R^2 = 0.9358$ ; $RMSE = 0.17140$ ; $F$ -value: 480.84 |            |         |            |            |

| MULHERES   |            |         |            |            |
|--|------------|---------|------------|------------|
| Parâmetro  | Estimativa | s.e.    | $t$ -value | $p$ -value |
| $\phi_0$   | -24.01177  | 0.50251 | -47.78     | <.0001     |
| $\phi_1$   | 0.21686    | 0.00742 | 29.24      | <.0001     |
| $R^2 = 0.9628$ ; $RMSE = 0.44312$ ; $F$ -value: 855.02 |            |         |            |            |

Quadro 8.2.2: Regressão linear do parâmetro  $\eta_x$  sobre a idade  $x$

A qualidade do ajustamento pode igualmente ser apreciada na Figura 8.2.5.

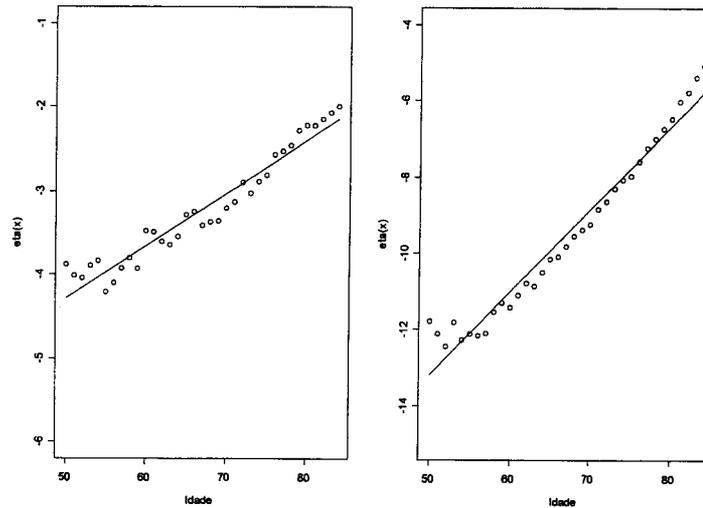


Figura 8.2.5: Resultado da regressão linear de  $\hat{\eta}_x$  sobre a idade  $x$  (homens à esquerda e mulheres à direita)

### 8.2.3 Modelo log-linear com intercepto linearizado

A boa aproximação conseguida na regressão linear entre  $\hat{\eta}_x$  e  $x$  incentivou-nos a testar uma versão simplificada do modelo (8.2.3) incorporando esta característica. O resultado é um modelo Poisson-Relacional do tipo **B**,  $\mathcal{PR}^B$ , definido por (8.2.2) e pela relação

$$\begin{aligned} \ln \mu_x^{BFP}(t) &= \phi_0 + \phi_1 x + \theta \ln \hat{\mu}_x^{POP}(t) \\ &= \phi_0 + \phi_1 x + \theta(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t), \quad x \geq 50, \end{aligned} \quad (8.2.5)$$

onde  $\phi_0$ ,  $\phi_1$  e  $\theta$  são parâmetros a estimar.

No Quadro 8.2.3 reportamos os resultados da estimação do modelo (8.2.2)-(8.2.5). A análise das estatísticas de teste e dos respectivos valores de prova permite-nos concluir que todos os parâmetros são significativos ao nível de significância  $\alpha = 0.05$ . O sinal positivo de  $\hat{\phi}_1$  observado em ambos os sexos sugere que o efeito selecção adversa decresce com a idade do indivíduo.

Embora descreva de forma adequada os dados do problema, o modelo (8.2.2)-(8.2.5) apresenta uma limitação importante. Com efeito, nas idades avançadas o termo linear na idade tende a dominar os restantes e a produzir taxas de mortalidade pouco razoáveis.

## 8.2. Relação entre a mortalidade das populações

| HOMENS                                     |            |        |                    |          |         |
|--|------------|--------|--------------------|----------|---------|
| Parâmetro                                  | Estimativa | s.e.   | Int confiança 95%  | $\chi^2$ | p-value |
| $\phi_0$                                   | -2.9455    | 1.0911 | [-5.0841, -0.8069] | 7.29     | 0.0069  |
| $\phi_1$                                   | 0.0219     | 0.0105 | [0.0013, 0.0425]   | 4.34     | 0.0373  |
| $\theta$                                   | 0.6373     | 0.1041 | [0.4332, 0.8413]   | 37.47    | <.0001  |
| Deviance: 701.41; Log Likelihood: 30034.07 |            |        |                    |          |         |

| MULHERES                                   |            |        |                    |          |         |
|--|------------|--------|--------------------|----------|---------|
| Parâmetro                                  | Estimativa | s.e.   | Int confiança 95%  | $\chi^2$ | p-value |
| $\phi_0$                                   | -6.0014    | 1.5269 | [-8.9941, -3.0087] | 15.45    | <.0001  |
| $\phi_1$                                   | 0.0519     | 0.0148 | [0.0229, 0.0809]   | 12.31    | 0.0005  |
| $\theta$                                   | 0.4221     | 0.1211 | [0.1847, 0.6595]   | 12.14    | 0.0005  |
| Deviance: 1075.44; Log Likelihood: 8641.93 |            |        |                    |          |         |

Quadro 8.2.3: Estimativas dos parâmetros do modelo log-linear com intercepto linearizado

### 8.2.4 Modelo log-linear com intercepto constante

Numa tentativa de simplificar o modelo relacional, ajustámos a seguinte regressão linear

$$\ln \mu_{x,t}^{BFP} = v_0 + v_1 \ln \hat{\mu}_{x,t}^{POP} + \epsilon_x, \quad \epsilon_x \sim \mathcal{N}(0, \sigma_\epsilon^2) \quad (8.2.6)$$

aos pares de pontos  $(\ln \mu_{x,t}^{BFP}, \ln \hat{\mu}_{x,t}^{POP})$ , usando o método dos mínimos quadrados ponderados, com cada observação a ser ponderada pela respectiva exposição ao risco.

Os resultados da estimação são apresentados no Quadro 8.2.4.

| HOMENS             |            |                |         |         |
|--------------------|------------|----------------|---------|---------|
| Parâmetro          | Estimativa | s.e.           | t-value | p-value |
| $v_0$              | -0.76410   | 0.09337        | -8.18   | <.0001  |
| $v_1$              | 0.83891    | 0.02381        | 35.23   | <.0001  |
| Adj $R^2 = 0.9264$ |            | RMSE = 9.39823 |         |         |

| MULHERES           |            |                 |         |         |
|--------------------|------------|-----------------|---------|---------|
| Parâmetro          | Estimativa | s.e.            | t-value | p-value |
| $v_0$              | -0.67332   | 0.18932         | -3.56   | 0.0004  |
| $v_1$              | 0.87805    | 0.04184         | 20.99   | <.0001  |
| Adj $R^2 = 0.8151$ |            | RMSE = 14.24706 |         |         |

Quadro 8.2.4: Estimativas dos parâmetros do modelo log-linear com intercepto constante pelo método dos mínimos quadrados ponderados

## 8.2. Relação entre a mortalidade das populações

Os valores do coeficiente de determinação ajustado, 0.9264 e 0.8151 para homens e mulheres respectivamente, atestam a qualidade da relação. Verifica-se igualmente que todos os parâmetros são significativos ao nível de significância  $\alpha = 0.01$ .

A Figura (8.2.6) complementa esta análise e dá-nos uma panorâmica quanto ao grau de ajustamento da regressão aos dados, em particular no caso dos homens onde a variabilidade em torno de uma tendência positiva bem definida é menor.

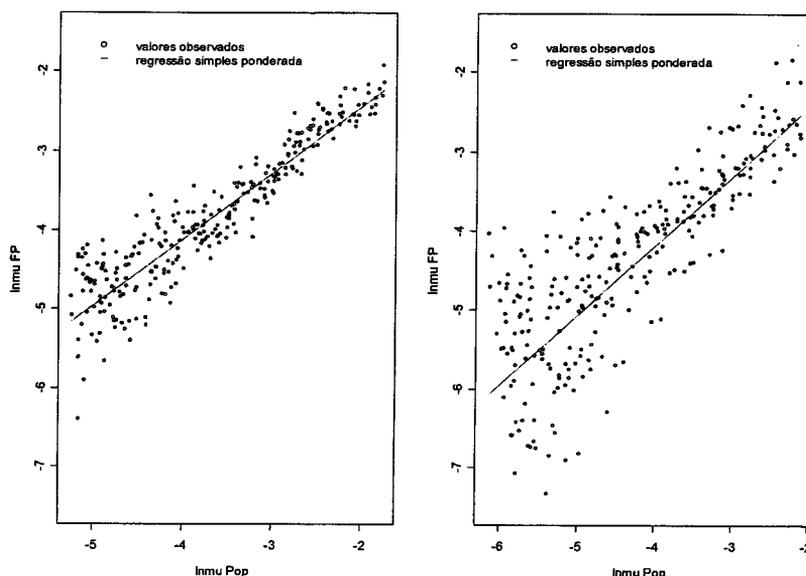


Figura 8.2.6: Estimação do modelo log-linear com intercepto constante pelo método dos mínimos quadrados ponderados (homens à esquerda e mulheres à direita)

Em face destes resultados, equacionámos uma versão simplificada do modelo (8.2.3)<sup>3</sup> considerando um intercepto independente da idade. A tradução deste princípio é um modelo Poisson-Relacional do tipo C,  $\mathcal{PR}^C$ , definido pela hipótese (8.2.2) e pela relação

$$\ln \mu_{x,t}^{BFP} = \eta + \theta \ln \hat{\mu}_{x,t}^{POP}, \quad x \geq 50 \quad (8.2.7)$$

com apenas dois parâmetros para estimar.

Os resultados da estimação dos parâmetros  $\eta$  e  $\theta$  do modelo (8.2.2)-(8.2.7) com recurso ao procedimento GENMOD são reportados no Quadro 8.2.5.

<sup>3</sup>Ou, equivalentemente, do modelo (8.2.5) considerando  $\phi_1 = 0$ .

## 8.2. Relação entre a mortalidade das populações

| HOMENS                                     |            |        |          |         |                    |
|--|------------|--------|----------|---------|--------------------|
| Parâmetro                                  | Estimativa | s.e.   | $\chi^2$ | p-value | IC 95%             |
| $\eta$                                     | -0.6748    | 0.0401 | 282.98   | <.0001  | [-0.7534, -0.5962] |
| $\theta$                                   | 0.8527     | 0.0119 | 5164.07  | <.0001  | [0.8294, 0.8760]   |
| Deviance: 705.76; Log Likelihood: 30031.89 |            |        |          |         |                    |

| MULHERES                                   |            |        |          |         |                    |
|--|------------|--------|----------|---------|--------------------|
| Parâmetro                                  | Estimativa | s.e.   | $\chi^2$ | p-value | IC 95%             |
| $\eta$                                     | -0.6462    | 0.0585 | 122.17   | <.0001  | [-0.7608, -0.5316] |
| $\theta$                                   | 0.8443     | 0.0158 | 2856.97  | <.0001  | [0.8133, 0.8752]   |
| Deviance: 1087.97; Log Likelihood: 8635.67 |            |        |          |         |                    |

Quadro 8.2.5: Estimativas dos parâmetros do modelo log-linear com intercepto constante no contexto dos modelos lineares generalizados

Na Figura 8.2.7 comparamos os valores estimados pelo modelo de regressão linear (estimado pelo método dos mínimos quadrados ponderados) e pelo modelo de Poisson (estimado pelo método de máxima verossimilhança).

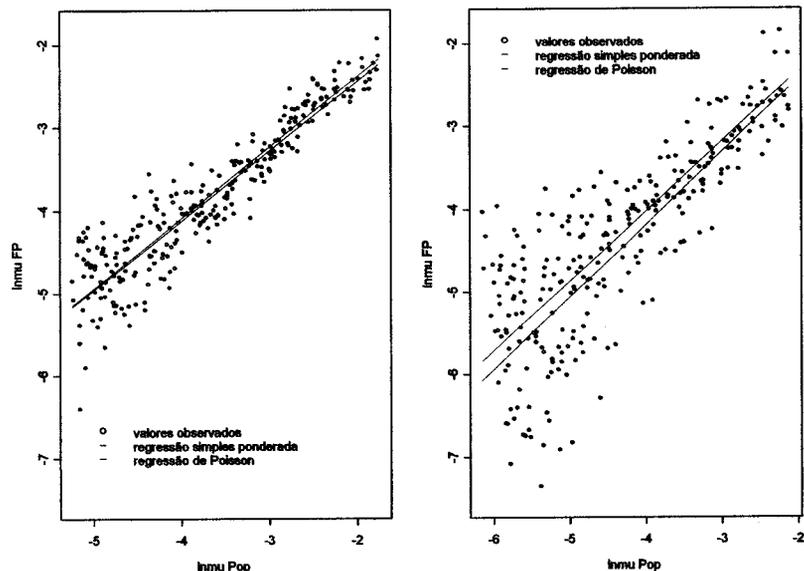


Figura 8.2.7: Estimativa do modelo relacional (8.2.7) ligando as taxas de mortalidade da população e dos BFP (homens à esquerda e mulheres à direita)

## 8.2. Relação entre a mortalidade das populações

Como se observa, o critério dos mínimos quadrados ponderados tende a sobrestimar o efeito de selecção adversa na população de beneficiários de fundos de pensões, em especial no caso de indivíduos do sexo feminino. Esta conclusão decorre do facto do critério dos mínimos quadrados ponderados estimar em todas as idades valores inferiores aos gerados pelo modelo de Poisson, o que significa que prevê taxas de mortalidade  $\mu_{x,t}$  também elas inferiores.

### 8.2.5 Comparação entre os modelos

Combinando as estimativas dos parâmetros dos três modelos relacionais estudados com as tábuas de mortalidade prospectivas calculadas para a população portuguesa, estamos agora em condições de construir tábuas prospectivas para a população de beneficiários de fundos de pensões e de apurar um conjunto de indicadores característicos, que sintetizem a presença ou não de significativos efeitos de selecção adversa.

Nesta secção, comparamos o desempenho dos três modelos analisados em termos dos indicadores  $\hat{e}_{65}(2004)$  e  $a_{65}(2004)$  ( $i = 3\%$ ), calculados numa óptica transversal (Trans) e numa óptica longitudinal (Long). Os resultados são sintetizados no Quadro 8.2.6.

| HOMENS |                      |        |                |        |
|--------|----------------------|--------|----------------|--------|
| Modelo | $\hat{e}_{65}(2004)$ |        | $a_{65}(2004)$ |        |
|        | Long                 | Trans  | Long           | Trans  |
| $PR^A$ | 16.866               | 16.698 | 12.232         | 12.140 |
| $PR^B$ | 17.598               | 16.925 | 12.631         | 12.267 |
| $PR^C$ | 17.996               | 17.087 | 12.858         | 12.370 |

| MULHERES |                      |        |                |        |
|----------|----------------------|--------|----------------|--------|
| Modelo   | $\hat{e}_{65}(2004)$ |        | $a_{65}(2004)$ |        |
|          | Long                 | Trans  | Long           | Trans  |
| $PR^A$   | 16.642               | 18.405 | 12.214         | 13.190 |
| $PR^B$   | 19.943               | 19.278 | 14.047         | 13.696 |
| $PR^C$   | 20.969               | 19.630 | 14.603         | 13.906 |

Quadro 8.2.6: Estimativas de  $e_{65}(2004)$  e  $a_{65}(2004)$  segundo as ópticas transversal e longitudinal calculadas com base nas tábuas geradas pelos modelos relacionais

A observação atenta do Quadro 8.2.6 evidencia a existência de diferenças importantes entre os três modelos testados. Estas são mais acentuadas nos indicadores calculados segundo a óptica longitudinal e sobretudo no caso da população feminina. De facto, verifica-se que neste caso os modelos  $PR^A$  e  $PR^B$  proporcionam valores considerados

muito baixos face aos correspondentes indicadores estimados para a população em geral, atendendo aos valores observados na Figura 8.2.1 para as taxas de mortalidade nas duas populações. Pelo contrário, o modelo log-linear com intercepto constante (modelo  $\mathcal{PR}^C$ ) apresenta, em ambos os sexos, estimativas consideradas mais razoáveis. Por esta razão, e pela sua simplicidade (o modelo inclui apenas dois parâmetros), retemos o modelo  $\mathcal{PR}^C$  para a análise que se segue.

### 8.3 Tábuas de mortalidade prospectivas para a população de pensionistas

Se admitirmos que a equação (8.2.7), que relaciona a mortalidade observada nas populações portuguesa em geral e de beneficiários de fundos de pensões, permanece válida no tempo, podemos recorrer ao modelo (8.2.2)-(8.2.7) para estimar as taxas  $\mu_{x,t}^{BFP}$  a partir das tábuas prospectivas elaboradas para a população portuguesa.

Na Figura 8.3.1 fornecemos uma primeira visão sobre a evolução temporal dos quocientes  $q_{x,t}^{BFP}$  ajustados pelo modelo Poisson-Relacional. Na Figura 8.3.2 comparamos as curvas de taxas  $\hat{\mu}_{x,t}^{POP}$  e  $\hat{\mu}_{x,t}^{BFP}$  brutas e ajustadas, respectivamente, com o auxílio dos modelos de Poisson e Poisson-Relacional (8.2.2)-(8.2.7).

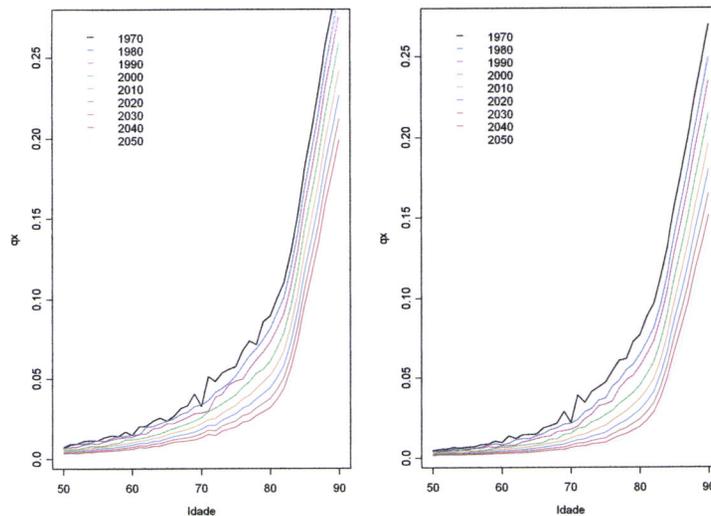


Figura 8.3.1: Evolução temporal dos quocientes  $q_{x,t}^{BFP}$  ajustados pelo modelo Poisson-Relacional (homens à esquerda e mulheres à direita)

### 8.3. Tábuas de mortalidade prospectivas para a população de pensionistas

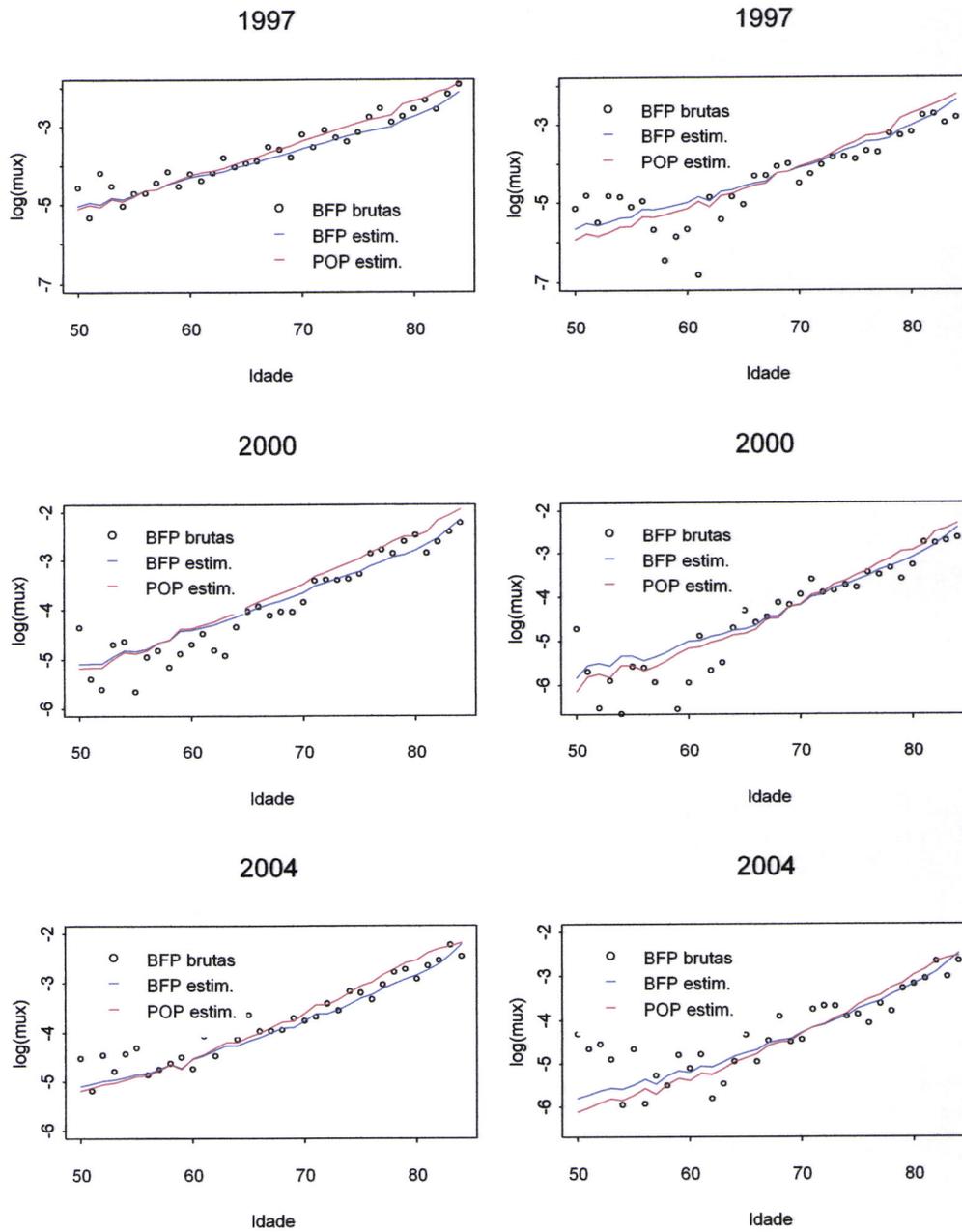


Figura 8.3.2: Comparação entre as taxas  $\mu_{x,t}$  brutas e estimadas para as populações portuguesa e de BFP em anos seleccionados (homens à esquerda e mulheres à direita)

### 8.3. Tábuas de mortalidade prospectivas para a população de pensionistas

Conforme se antevia, a trajectória descendente projectada para os  $q_{x,t}^{BFP}$  acompanha no tempo o comportamento previsto para a população portuguesa. O modelo reconstitui de forma adequada as taxas  $\hat{\mu}_{x,t}^{BFP}$  observadas, pese embora a maior variabilidade registada no caso da população de BFP, explicável pela menor exposição relativa ao risco quando comparada com a população portuguesa. O comportamento da mortalidade na população de BFP aproxima-se bastante do da população em geral, mas é evidente a trajectória mais favorável no caso particular dos homens, a partir dos 60 anos.

A informação contida nas tábuas prospectivas construídas para a população de BFP permite-nos igualmente projectar a trajectória dos quocientes  $q_{x,2004}^{POP}$  e  $q_{x,2004}^{BFP}$  relativos à geração que atinge a idade  $x = 50$  em 2004. Em termos mais precisos, estamos interessados na evolução dos quocientes de mortalidade  $q_{50+k}(2004+k)$  ( $k = 0, 1, \dots$ ) a aplicar aos indivíduos que alcançam os 50 anos de idade em 2004. As curvas projectadas são visíveis na Figura 8.3.3.

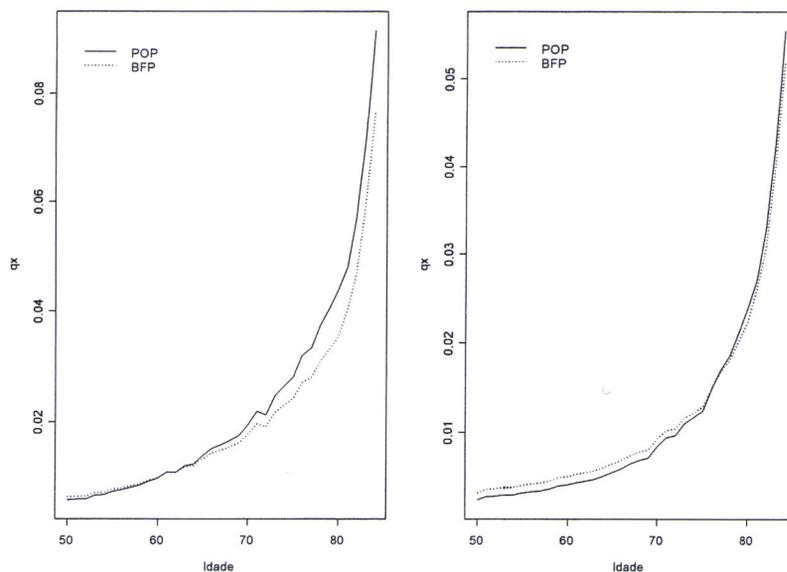


Figura 8.3.3: Quocientes de mortalidade  $q_{x,2004}^{POP}$  e  $q_{x,2004}^{BFP}$  relativos à geração que atinge a idade  $x = 50$  em 2004 (homens à esquerda e mulheres à direita)

A submortalidade da população masculina de BFP e o comportamento muito próximo da população feminina relativamente à população em geral são, uma vez mais, postos em evidência.

## 8.4 Estimativa do efeito selecção adversa

### 8.4.1 Na esperança de vida residual

A importância do efeito selecção adversa no mercado de fundos de pensões pode ser aferida através das diferenças na esperança de vida face à população em geral. Em termos teóricos, os indivíduos que têm uma esperança de vida mais elevada atribuem maior importância à necessidade de cobrir o risco de longevidade uma vez que a probabilidade de esgotarem a sua riqueza em vida é maior. Esta maior preocupação com a necessidade de garantir um nível de rendimento adequado durante todo o ciclo de vida pode traduzir-se, por exemplo, na aquisição de rendas vitalícias ou no aumento da poupança durante a sua vida activa.

O Quadro 8.4.1 projecta a evolução da esperança de vida completa remanescente aos 65 anos  $\{\hat{e}_{65}^{POP}(t), \hat{e}_{65}^{BFP}(t)\}$  entre 2004 e 2019, calculadas segundo as ópticas longitudinal e transversal. A importância do efeito selecção adversa pode ser aferida pela diferença absoluta entre as esperanças de vida,  $\{\hat{e}_{65}^{BFP}(t) - \hat{e}_{65}^{POP}(t)\}$ , ou pela sua diferença relativa  $\Delta_{\hat{e}_{65}}$  (%)

$$\Delta_{\hat{e}_{65}} (\%) = \frac{\hat{e}_{65}^{BFP}(t) - \hat{e}_{65}^{POP}(t)}{\hat{e}_{65}^{POP}(t)}.$$

Como se constata, a submortalidade da população masculina de BFP traduz-se em esperanças de vida  $\hat{e}_{65}^{BFP}(t)$  claramente superiores às suas homologas para a população portuguesa  $\hat{e}_{65}^{POP}(t)$ , com as diferenças a atingirem, no ano de 2004, os 6.40% (7.84%) segundo a visão longitudinal (transversal), embora com tendência para diminuir com o tempo. Os valores apresentados indicam, por exemplo, que um pensionista que atinge os 65 anos de idade em 2004 viverá mais 1.09 anos (1.25, segundo a visão transversal) que o seu homólogo da população portuguesa.

O efeito selecção adversa no caso da população feminina de BFP é bem menos evidente, podendo até ser considerado negligenciável se atendermos às diferenças absolutas e relativas registadas na esperança de vida, em particular na óptica longitudinal.

Assinale-se ainda que o diferencial de esperança de vida entre sexos é maior na população em geral do que na população de BFP (de aproximadamente 1 ano), com tendência para um ligeiro acréscimo.

8.4. Estimativa do efeito selecção adversa

| Visão Longitudinal |                         |                         |                             |                         |                         |                             |
|--------------------|-------------------------|-------------------------|-----------------------------|-------------------------|-------------------------|-----------------------------|
| Ano                | Homens                  |                         |                             | Mulheres                |                         |                             |
|                    | $\hat{e}_{65}^{POP}(t)$ | $\hat{e}_{65}^{BFP}(t)$ | $\Delta_{\hat{e}_{65}}(\%)$ | $\hat{e}_{65}^{POP}(t)$ | $\hat{e}_{65}^{BFP}(t)$ | $\Delta_{\hat{e}_{65}}(\%)$ |
| 2004               | 16.91                   | 18.00                   | 6.40                        | 20.87                   | 20.97                   | 0.46                        |
| 2005               | 17.01                   | 18.08                   | 6.25                        | 21.01                   | 21.08                   | 0.35                        |
| 2006               | 17.13                   | 18.17                   | 6.09                        | 21.14                   | 21.19                   | 0.24                        |
| 2007               | 17.25                   | 18.27                   | 5.92                        | 21.27                   | 21.30                   | 0.13                        |
| 2008               | 17.37                   | 18.37                   | 5.75                        | 21.41                   | 21.41                   | 0.03                        |
| 2009               | 17.49                   | 18.46                   | 5.59                        | 21.54                   | 21.52                   | -0.08                       |
| 2010               | 17.60                   | 18.56                   | 5.43                        | 21.67                   | 21.63                   | -0.18                       |
| 2011               | 17.72                   | 18.66                   | 5.27                        | 21.80                   | 21.74                   | -0.27                       |
| 2012               | 17.84                   | 18.75                   | 5.12                        | 21.92                   | 21.84                   | -0.37                       |
| 2013               | 17.95                   | 18.85                   | 4.97                        | 22.05                   | 21.95                   | -0.46                       |
| 2014               | 18.07                   | 18.94                   | 4.82                        | 22.18                   | 22.05                   | -0.55                       |
| 2015               | 18.19                   | 19.04                   | 4.67                        | 22.30                   | 22.16                   | -0.64                       |
| 2016               | 18.30                   | 19.13                   | 4.52                        | 22.42                   | 22.26                   | -0.73                       |
| 2017               | 18.42                   | 19.22                   | 4.38                        | 22.55                   | 22.36                   | -0.81                       |
| 2018               | 18.53                   | 19.32                   | 4.24                        | 22.67                   | 22.46                   | -0.90                       |
| 2019               | 18.64                   | 19.41                   | 4.10                        | 22.79                   | 22.56                   | -0.98                       |

| Visão Transversal |                         |                         |                             |                         |                         |                             |
|-------------------|-------------------------|-------------------------|-----------------------------|-------------------------|-------------------------|-----------------------------|
| Ano               | Homens                  |                         |                             | Mulheres                |                         |                             |
|                   | $\hat{e}_{65}^{POP}(t)$ | $\hat{e}_{65}^{BFP}(t)$ | $\Delta_{\hat{e}_{65}}(\%)$ | $\hat{e}_{65}^{POP}(t)$ | $\hat{e}_{65}^{BFP}(t)$ | $\Delta_{\hat{e}_{65}}(\%)$ |
| 2004              | 15.84                   | 17.09                   | 7.84                        | 19.22                   | 19.63                   | 2.11                        |
| 2005              | 15.79                   | 17.04                   | 7.92                        | 19.19                   | 19.60                   | 2.14                        |
| 2006              | 15.90                   | 17.13                   | 7.74                        | 19.47                   | 19.83                   | 1.88                        |
| 2007              | 16.02                   | 17.23                   | 7.56                        | 19.51                   | 19.87                   | 1.83                        |
| 2008              | 16.13                   | 17.33                   | 7.39                        | 19.70                   | 20.03                   | 1.65                        |
| 2009              | 16.25                   | 17.42                   | 7.22                        | 19.80                   | 20.11                   | 1.56                        |
| 2010              | 16.36                   | 17.52                   | 7.05                        | 19.95                   | 20.24                   | 1.42                        |
| 2011              | 16.48                   | 17.61                   | 6.88                        | 20.07                   | 20.34                   | 1.32                        |
| 2012              | 16.59                   | 17.71                   | 6.72                        | 20.21                   | 20.45                   | 1.19                        |
| 2013              | 16.71                   | 17.80                   | 6.56                        | 20.33                   | 20.55                   | 1.08                        |
| 2014              | 16.82                   | 17.90                   | 6.40                        | 20.47                   | 20.66                   | 0.97                        |
| 2015              | 16.93                   | 17.99                   | 6.24                        | 20.59                   | 20.77                   | 0.86                        |
| 2016              | 17.05                   | 18.08                   | 6.08                        | 20.72                   | 20.88                   | 0.76                        |
| 2017              | 17.16                   | 18.18                   | 5.93                        | 20.84                   | 20.98                   | 0.65                        |
| 2018              | 17.27                   | 18.27                   | 5.78                        | 20.97                   | 21.08                   | 0.55                        |
| 2019              | 17.38                   | 18.36                   | 5.63                        | 21.09                   | 21.19                   | 0.45                        |

Quadro 8.4.1: Comparação entre as esperanças de vida aos 65 anos para as populações portuguesa e de BFP, segundo as visões longitudinal e transversal

### 8.4.2 No preço das rendas vitalícias

Uma forma diferente de quantificar a importância financeira do fenómeno selecção adversa, consiste em estimar as diferenças nos prémios de rendas vitalícias em resultado da aplicação de bases demográficas distintas.

Neste sentido, calculámos os prémios únicos puros de rendas vitalícias de valor unitário, considerando as tábuas prospectivas deduzidas para as populações portuguesa e de BFP e uma taxa de juro técnica  $i = 3\%$ . Os resultados são resumidos no Quadro 8.5.1 e nas Figuras 8.5.1 e 8.5.2.

O diferencial relativo entre  $a_{65}^{BFP}(t)$  e  $a_{65}^{POP}(t)$ ,  $\Delta_{a_{65}}(\%)$ , é calculado mediante

$$\Delta_{a_{65}}(\%) = \frac{a_{65}^{BFP}(t) - a_{65}^{POP}(t)}{a_{65}^{POP}(t)}.$$

Como se observa, uma companhia que utilize as tábuas para a população portuguesa para fazer o *pricing* de uma renda vitalícia comercializada a um pensionista subestima em 6% o valor do prémio puro no caso dos homens. No caso das mulheres, as diferenças são novamente pouco expressivas (inferiores a 1%, na óptica longitudinal).

Se desprezarmos a importância das cargas adicionais sobre o prémio puro, conclui-se que a não aplicação das tábuas prospectivas específicas para a população de BFP pode conduzir, em especial no caso dos homens, a prejuízos significativos para companhias a operar o ramo vida. Este custo é o resultado da presença do efeito selecção adversa detectado neste estudo.

## 8.5 Tábuas prospectivas vs tábuas regulamentares

Para aquilatarmos da importância que o recurso a tábuas prospectivas assume no mercado de fundos de pensões em Portugal, representamos na Figura 8.5.3 a evolução projectada do prémio único puro  $a_{65}^{POP}(t)$  e  $a_{65}^{BFP}(t)$  calculado com base nas tábuas de mortalidade prospectivas (visão longitudinal) deduzidas para as populações portuguesa e de BFP, respectivamente, e o valor correspondente determinado a partir de duas tábuas regulamentares (tábuas TV 73/77 e TV 88/90), bastante usadas no mercado português.

Como se observa, a tábua TV 73/77 subestima largamente o valor actual das rendas, sobretudo no caso das mulheres. Se a aplicação desta tábua com inclusão de uma margem de segurança significativa relativamente às projecções realizadas para a população portuguesa a tornava aceitável no caso dos homens até ao início do século XXI, o prémio resultante é largamente insuficiente quando levamos em linha de conta o efeito

## 8.5. Tábuas prospectivas vs tábuas regulamentares

selecção adversa e o comportamento da mortalidade na população de BFP.

| Visão Longitudinal |                    |                    |                        |                    |                    |                        |
|--------------------|--------------------|--------------------|------------------------|--------------------|--------------------|------------------------|
| Ano                | Homens             |                    |                        | Mulheres           |                    |                        |
|                    | $a_{65}^{POP} (t)$ | $a_{65}^{BFP} (t)$ | $\Delta_{a_{65}} (\%)$ | $a_{65}^{POP} (t)$ | $a_{65}^{BFP} (t)$ | $\Delta_{a_{65}} (\%)$ |
| 2004               | 12.22              | 12.86              | 5.21                   | 14.56              | 14.60              | 0.31                   |
| 2005               | 12.28              | 12.90              | 5.09                   | 14.63              | 14.67              | 0.22                   |
| 2006               | 12.35              | 12.96              | 4.95                   | 14.71              | 14.73              | 0.13                   |
| 2007               | 12.42              | 13.02              | 4.81                   | 14.78              | 14.79              | 0.05                   |
| 2008               | 12.49              | 13.08              | 4.67                   | 14.86              | 14.85              | -0.03                  |
| 2009               | 12.57              | 13.14              | 4.54                   | 14.93              | 14.91              | -0.11                  |
| 2010               | 12.64              | 13.19              | 4.40                   | 15.00              | 14.97              | -0.19                  |
| 2011               | 12.71              | 13.25              | 4.27                   | 15.07              | 15.03              | -0.26                  |
| 2012               | 12.78              | 13.31              | 4.15                   | 15.14              | 15.09              | -0.34                  |
| 2013               | 12.85              | 13.36              | 4.02                   | 15.21              | 15.15              | -0.41                  |
| 2014               | 12.92              | 13.42              | 3.90                   | 15.28              | 15.21              | -0.48                  |
| 2015               | 12.98              | 13.47              | 3.77                   | 15.35              | 15.27              | -0.54                  |
| 2016               | 13.05              | 13.53              | 3.65                   | 15.42              | 15.32              | -0.61                  |
| 2017               | 13.12              | 13.58              | 3.53                   | 15.48              | 15.38              | -0.67                  |
| 2018               | 13.19              | 13.64              | 3.42                   | 15.55              | 15.44              | -0.73                  |
| 2019               | 13.25              | 13.69              | 3.31                   | 15.62              | 15.49              | -0.79                  |

| Visão Transversal |                    |                    |                        |                    |                    |                        |
|-------------------|--------------------|--------------------|------------------------|--------------------|--------------------|------------------------|
| Ano               | Homens             |                    |                        | Mulheres           |                    |                        |
|                   | $a_{65}^{POP} (t)$ | $a_{65}^{BFP} (t)$ | $\Delta_{a_{65}} (\%)$ | $a_{65}^{POP} (t)$ | $a_{65}^{BFP} (t)$ | $\Delta_{a_{65}} (\%)$ |
| 2004              | 11.64              | 12.37              | 6.30                   | 13.70              | 13.91              | 1.52                   |
| 2005              | 11.59              | 12.33              | 6.39                   | 13.68              | 13.89              | 1.54                   |
| 2006              | 11.66              | 12.39              | 6.24                   | 13.84              | 14.02              | 1.33                   |
| 2007              | 11.73              | 12.45              | 6.09                   | 13.87              | 14.05              | 1.30                   |
| 2008              | 11.81              | 12.51              | 5.94                   | 13.98              | 14.14              | 1.16                   |
| 2009              | 11.88              | 12.57              | 5.80                   | 14.03              | 14.18              | 1.09                   |
| 2010              | 11.95              | 12.62              | 5.66                   | 14.12              | 14.26              | 0.98                   |
| 2011              | 12.02              | 12.68              | 5.52                   | 14.19              | 14.31              | 0.90                   |
| 2012              | 12.09              | 12.74              | 5.38                   | 14.26              | 14.38              | 0.80                   |
| 2013              | 12.16              | 12.80              | 5.24                   | 14.33              | 14.44              | 0.72                   |
| 2014              | 12.23              | 12.85              | 5.11                   | 14.41              | 14.50              | 0.63                   |
| 2015              | 12.30              | 12.91              | 4.98                   | 14.48              | 14.56              | 0.55                   |
| 2016              | 12.37              | 12.97              | 4.85                   | 14.55              | 14.62              | 0.47                   |
| 2017              | 12.43              | 13.02              | 4.72                   | 14.62              | 14.68              | 0.39                   |
| 2018              | 12.50              | 13.08              | 4.60                   | 14.69              | 14.74              | 0.31                   |
| 2019              | 12.57              | 13.13              | 4.47                   | 14.76              | 14.80              | 0.23                   |

Quadro 8.5.1: Comparação entre os prémios de rendas vitalícias calculados aos 65 anos para as populações portuguesa e de BFP, segundo as visões longitudinal e transversal

## 8.5. Tábuas prospectivas vs tábuas regulamentares

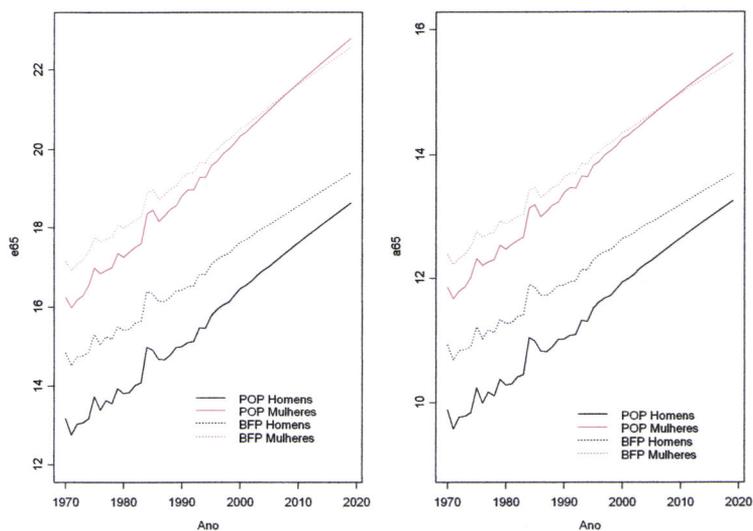


Figura 8.5.1: Estimativas de  $e_{65}^e(t)$  (à esquerda) e de  $a_{65}(t)$  (à direita) para as populações portuguesa e de BFP, segundo a óptica longitudinal

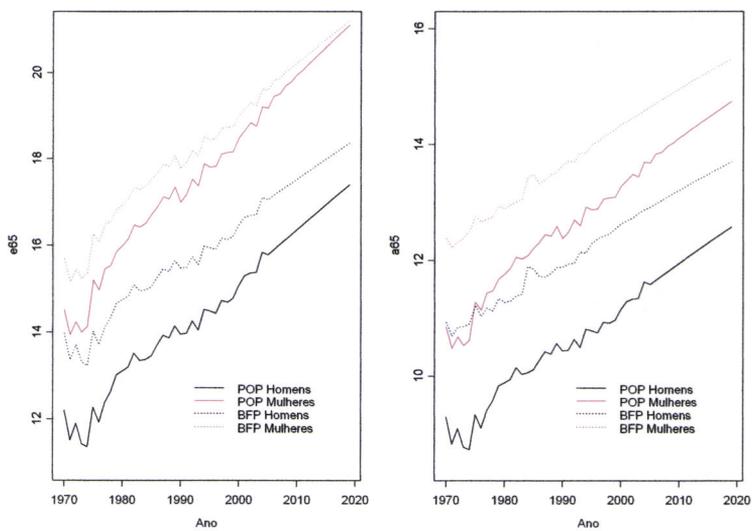


Figura 8.5.2: Estimativas de  $e_{65}^e(t)$  (à esquerda) e de  $a_{65}(t)$  (à direita) para as populações portuguesa e de BFP, segundo a óptica transversal

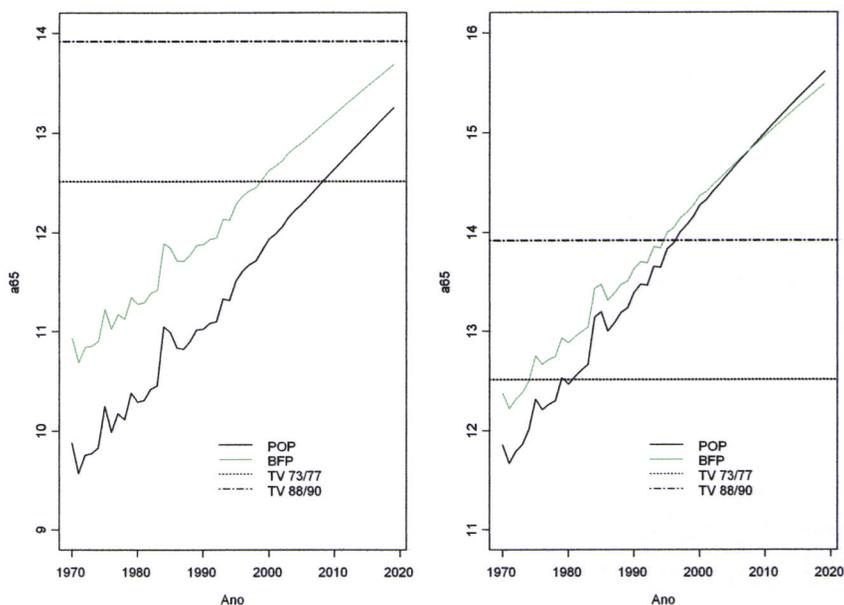


Figura 8.5.3: Prémio puro  $a_{65}^{POP}(t)$  e  $a_{65}^{BFP}(t)$  calculado com base nas tábuas prospectivas obtidas para as populações portuguesa e de BFP, respectivamente, e com base nas tábuas regulamentares TV 73/77 e TV 88/90 (homens à esquerda e mulheres à direita)

Neste contexto, as companhias que neste momento calculam os prémios com base na tábua TV 73/77 para indivíduos de ambos os sexos devem estar preparadas para assumir perdas técnicas importantes no caso da sua carteira se assemelhar à representada pela população de BFP.

Relativamente à tábua regulamentar TV 88/90, a Figura 8.5.3 mostra-nos duas realidades distintas. Se, por um lado, ela apresenta uma margem de segurança significativa relativamente às populações do sexo masculino, por outro ela revela-se desajustada face ao comportamento das populações do sexo feminino a partir de meados da década de 90. A sua utilização neste último caso comporta, de novo, riscos consideráveis para as companhias a operar no mercado.

Em face desta conclusão, podemos questionar-nos sobre qual será a taxa de juro técnica  $i^*$  necessária para igualar o prémio puro de uma renda vitalícia calculada em 2004 segundo as tábuas prospectivas (para as populações portuguesa e de BFP) ao seu valor deduzido a partir das tábuas TV 73/77 e TV 88/90. Os resultados deste exercício

## 8.5. Tábuas prospectivas vs tábuas regulamentares

são reportados no Quadro 8.5.2.

Um prémio único puro de 12.51 e 13.92, igual para ambos os sexos, calculado com base nas tábuas TV 73/77 e TV 88/90 (considerando uma taxa técnica de 3%), respectivamente, corresponde a uma taxa de juro de 4.57% e 3.47%, respectivamente, no caso da mortalidade evoluir conforme as tábuas prospectivas calculadas para a população feminina de BFP.

| HOMENS   |           |                |       |                |
|----------|-----------|----------------|-------|----------------|
| Tábua    | População |                | BFP   |                |
|          | $i^*$     | $\Delta(b.p.)$ | $i^*$ | $\Delta(b.p.)$ |
| TV 73/77 | 2.74%     | -26            | 3.29% | +29            |
| TV 88/90 | 1.62%     | -138           | 2.18% | -82            |

| MULHERES |           |                |       |                |
|----------|-----------|----------------|-------|----------------|
| Tábua    | População |                | BFP   |                |
|          | $i^*$     | $\Delta(b.p.)$ | $i^*$ | $\Delta(b.p.)$ |
| TV 73/77 | 4.54%     | +154           | 4.57% | +157           |
| TV 88/90 | 3.60%     | +60            | 3.47% | +47            |

Quadro 8.5.2: Taxa de juro técnica necessária para igualar o prémio puro de uma renda vitalícia calculada em 2004 segundo as tábuas prospectivas ao seu valor deduzido a partir das tábuas TV 73/77 e TV 88/90

Uma forma diferente de avaliar os resultados consiste em interpretar estes valores como taxas de retorno financeiro que as companhias devem realizar de maneira a compensar o risco de longevidade (resultante de uma mortalidade mais favorável do que a expressa nas tábuas TV 73/77 e TV 88/90). O diferencial em *basis points*  $\Delta(b.p.)$  dá-nos, quando negativo, uma ideia da margem de segurança (face à taxa de juro técnica  $i = 3\%$ ) associada ao uso das tábuas regulamentares. Quando positivo, dá informação sobre os ganhos financeiros a realizar para compensar as perdas na componente demográfica.

Na conjuntura actual, a obtenção de taxas de retorno da ordem das acima reportadas não se afigura fácil e expectável. A consequência imediata é a de que as tentativas de compensar o risco de longevidade mediante uma política de gestão financeira adequada, que rentabilize os patrimónios colocados à disposição dos fundos de pensões, podem revelar-se ilusórias e infrutíferas. Nalguns casos, a existência de determinados arranjos contratuais faz com que esta tarefa seja ainda mais difícil. Exemplo disso são as cláusulas que obrigam as companhias a distribuir uma percentagem (mais ou menos elevada) dos rendimentos financeiros gerados pelos activos que constituem os fundos (na parcela que excede o rendimento garantido contratualmente).

## 8.6 Intervalos de confiança

O objectivo desta secção é o de obter intervalos de confiança para os indicadores  $e_{65}^{BFP}(t)$  e  $a_{65}^{BFP}(t)$  associados à população de beneficiários de fundos de pensões. Nesse sentido, recorreremos a um algoritmo deduzido a partir do método *parametric bootstrap* descrito no Capítulo 6.

O procedimento, que designamos de *duplo-bootstrap* por compreender simulações das tábuas prospectivas para a população portuguesa e do modelo relacional, decompõe-se, neste caso, nas seguintes etapas:

1. Partindo dos valores observados, são geradas amostras aleatórias  $(E_{x,t}^{POP,b}, d_{x,t}^{POP,b}, E_{x,t}^{BFP,b}, d_{x,t}^{BFP,b})$ ,  $b = 1, \dots, B$ , a partir da população portuguesa em geral, por um lado, e da população de BFP por outro. Em termos mais precisos,  $d_{x,t}^{POP,b}$  é obtido através de uma realização da lei de Poisson de média  $d_{x,t}^{POP}$  e  $d_{x,t}^{BFP,b}$  é deduzido com base numa realização da lei de Poisson de média  $d_{x,t}^{BFP}$ :

$$\begin{aligned} d_{x,t}^{POP,b} &\sim \text{Poisson}(d_{x,t}^{POP}) \\ d_{x,t}^{BFP,b} &\sim \text{Poisson}(d_{x,t}^{BFP}) \end{aligned}$$

2. Em cada simulação  $b$   $(E_{x,t}^{POP,b}, d_{x,t}^{POP,b})$ , são estimados os parâmetros  $\hat{\alpha}_x^b$ ,  $\hat{\beta}_x^b$  e  $\hat{\kappa}_t^b$  do modelo de Poisson.
3. Em cada simulação, são estimados os parâmetros do modelo ARIMA seleccionado com base nas observações iniciais e recorrendo à série  $\mathcal{K}^b = \{\hat{\kappa}_t^b, t = t_{\min}, \dots, t_{\max}, b = 1, \dots, B\}$ . Note-se, uma vez mais, que a estrutura do modelo ARIMA é preservada enquanto que os parâmetros são reestimados.
4. São gerados valores projectados  $\hat{\kappa}_t^b$ , com base nos parâmetros do modelo ARIMA estimado na etapa 3. Os termos residuais  $\epsilon_t^b$  são simulados a partir da lei Normal  $\text{Nor}(0, \sigma^b)$ .
5. São construídas tábuas de mortalidade prospectivas para a população portuguesa.
6. Em cada simulação  $b$ , são estimados os parâmetros  $\hat{\eta}^b$  e  $\hat{\gamma}^b$  do modelo Poisson-Relacional (8.2.2)-(8.2.7).
7. São projectadas as taxas instantâneas de mortalidade da seguinte maneira

$$\ln \mu_{x,t}^{BFP,b} = \hat{\eta}^b + \hat{\theta}^b \ln \hat{\mu}_{x,t}^{POP,b}$$

## 8.6. Intervalos de confiança

onde

$$\ln \hat{\mu}_{x,t}^{POP,b} = \hat{\alpha}_x^b + \hat{\beta}_x^b \hat{\kappa}_t^b$$

e  $b = 1, \dots, B$ ,  $x_{\min} \leq x \leq x_{\max}$  e  $t \geq t_{\max}$ .

8. Em cada simulação, são calculados os indicadores de interesse  $\hat{e}_{65}^{BFP,b}(t)$  e  $a_{65}^{BFP,b}(t)$ .

A aplicação do método *parametric bootstrap* e o cálculo dos indicadores de interesse foram conseguidos com o auxílio de um algoritmo informático escrito e implementado no *package* SAS (Release 9.1). Os resultados numéricos considerando um total de  $B = 1000$  realizações aleatórias são resumidos no Quadro 8.6.1.

Registe-se que os valores apresentados referem-se apenas aos valores médios e percentis empíricos de  $\hat{e}_{65}(2004)$  e  $a_{65}(2004)$ , calculados segundo a óptica longitudinal. A Figura 8.6.1 fornece-nos uma panorâmica sobre a distribuição empírica dos indicadores.

| HOMENS                    |            |            |            |                           |            |            |            |
|---------------------------|------------|------------|------------|---------------------------|------------|------------|------------|
| $\hat{e}_{65}(2004)$      | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}(2004)$            | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 18.22                     | 17.40      | 18.23      | 19.04      | 12.99                     | 12.51      | 13.00      | 13.46      |
| Margem erro: 1.64 (8.98%) |            |            |            | Margem erro: 0.95 (7.30%) |            |            |            |

| MULHERES                  |            |            |            |                           |            |            |            |
|---------------------------|------------|------------|------------|---------------------------|------------|------------|------------|
| $\hat{e}_{65}(2004)$      | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ | $a_{65}(2004)$            | $Q_{0.05}$ | $Q_{0.50}$ | $Q_{0.95}$ |
| 21.14                     | 20.72      | 21.13      | 21.57      | 14.70                     | 14.46      | 14.69      | 14.94      |
| Margem erro: 0.85 (4.02%) |            |            |            | Margem erro: 0.48 (3.28%) |            |            |            |

Quadro 8.6.1: Valores médios e quartis empíricos de  $e_x(2004)$  e  $a_x(2004)$  para a população de BFP, simulados pelo método de *parametric bootstrap*, onde  $Q_\alpha$  denota o percentil empírico de ordem  $\alpha$

Como se observa, o procedimento de duplo *bootstrap* gera intervalos de confiança a 90% relativamente estreitos, em especial no caso das mulheres. Por exemplo, para o indicador  $a_{65}(2004)$  o seu valor é de 0.95 e 0.48 respectivamente para homens e mulheres, o que corresponde a margens de erro de 7.30% e 3.28% face ao valor médio do indicador. No caso dos homens, os intervalos são mais significativos, um resultado que se explica, uma vez mais, pelo maior número de termos residuais incluídos no modelo ARIMA usado na projecção da tendência temporal  $\kappa_t$ .

Como se atesta pela comparação dos valores inscritos nos Quadros 8.6.1 e 7.8.2, o procedimento de duplo *bootstrap* aplicado na construção de intervalos de confiança para a população de BFP proporciona, naturalmente, amplitudes de intervalo mais significativas

## 8.6. Intervalos de confiança

relativamente às obtidas para a população em geral, face ao maior número de parâmetros (incerteza) envolvido.

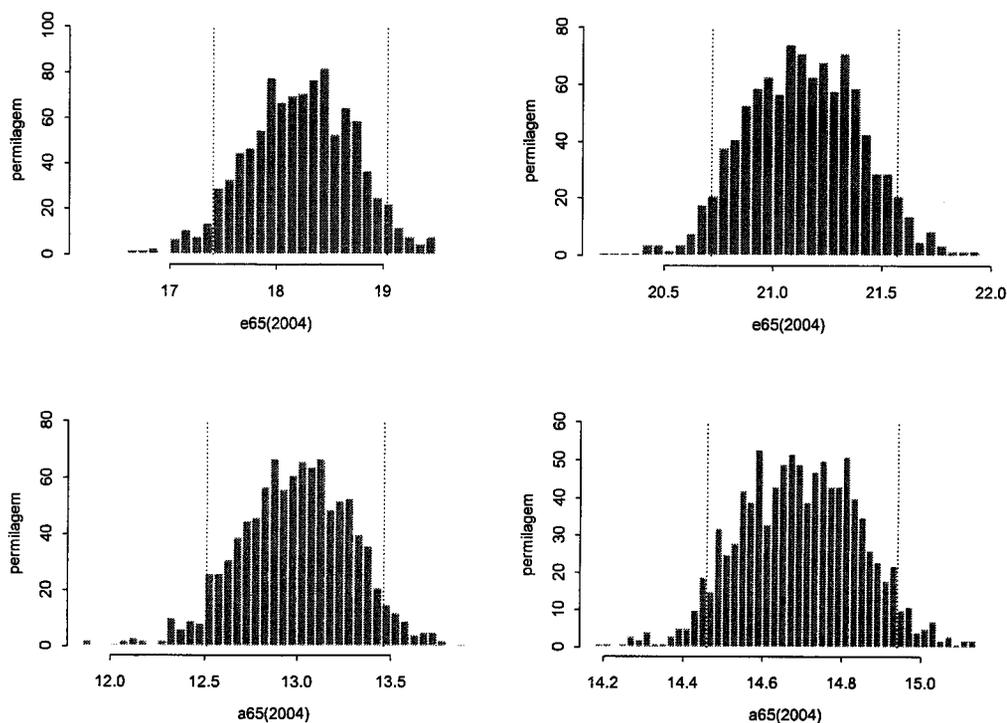


Figura 8.6.1: Histogramas dos valores de  $\hat{e}_{65}^{BFP}(2004)$  e  $\hat{a}_{65}^{BFP}(2004)$  gerados pelo método de Parametric Bootstrap (homens à esquerda e mulheres à direita)

As simulações efectuadas para obter os percentis empíricos inscritos no Quadro 8.6.1 proporcionam igualmente uma visão sobre a variabilidade das taxas de mortalidade prospectivas. Neste sentido, a Figura 8.6.2 representa 100 realizações das taxas  $\mu_{50+k}^{BFP}(2004+k)$  ( $k = 0, 1, 2, \dots$ ), que incidirão sobre os indivíduos que atingem os 50 anos em 2004.

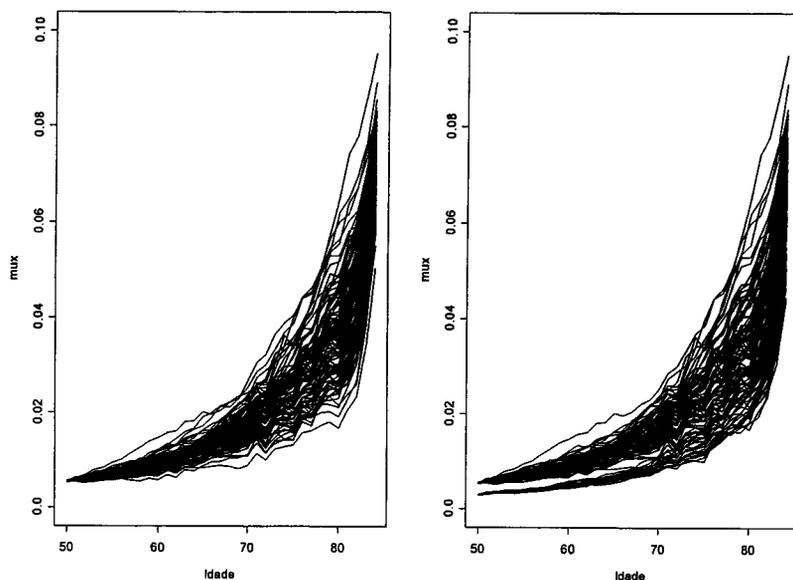


Figura 8.6.2: Taxas de mortalidade  $\mu_{50+k}^{BFP}$  (2004 +  $k$ ) simuladas pelo método parametric bootstrap (homens à esquerda e mulheres à direita)

Observa-se alguma variabilidade nas projecções, que é no entanto atenuada no cálculo dos valores característicos representados no Quadro 8.6.1, em resultado do processo de actualização. Tal variabilidade estará, no entanto, presente no cálculo das projecções de cash flows relativos a uma certa de rendas vitalícias.

## 8.7 Conclusão

Neste capítulo, investigámos a importância do efeito selecção adversa na carteira de beneficiários de fundos de pensões em Portugal. Para tal, adoptámos uma metodologia, assente num modelo relacional incorporado num modelo linear generalizado com distribuição de Poisson, que permite ajustar as projecções de mortalidade derivadas para a população portuguesa ao comportamento observado na população de pensionistas.

Foram investigadas diferentes formulações do modelo relacional, tendo a escolha recaído numa relação log-linear com intercepto constante para todas as idades. Os resultados obtidos com base neste modelo permitiram-nos derivar pela primeira vez tábuas de mortalidade prospectivas para a população de beneficiários de fundos de pensões em

Portugal, e avaliar a importância do efeito selecção adversa.

Da análise dos resultados destacamos a importância do efeito selecção adversa na população masculina, com uma clara submortalidade face à população portuguesa em geral. Em contrapartida, constatámos que no período em análise o fenómeno assume pouca expressão no caso da população feminina. Estes resultados sugerem que o desenvolvimento do mercado de fundos de pensões em Portugal deve ter em conta as características particulares da mortalidade nesta população. Caso contrário, as insuficientes provisões financeiras podem induzir perdas consideráveis nos resultados das empresas.

Os estudos realizados demonstraram igualmente a insuficiência dos prémios brutos puros calculados com base em duas tábuas regulamentares - tábuas TV 73/77 e TV 88/90 -, comumente usadas no mercado português, face ao comportamento previsional da mortalidade, em especial no caso das mulheres. Este facto chama a atenção para o problema da subestimação das tarifas e das responsabilidades financeiras neste mercado. As conclusões deste estudo suportam a ideia de que este problema deve merecer uma atenção especial por parte das entidades reguladoras, cuja resolução passa, inevitavelmente, pela adopção de tábuas prospectivas específicas e pela imposição de bases técnicas suficientemente prudentes.

Uma característica interessante da metodologia usada refere-se à possibilidade de derivar intervalos de confiança empíricos para os valores característicos calculados para as populações portuguesa e de beneficiários de fundos de pensões, elementos importantes na determinação das cargas de segurança a aplicar sobre os prémios puros e na quantificação do risco de longevidade a que estão expostas as carteiras.

## **Parte III**

# **Mortalidade Estocástica: Aplicações Actuariais e Cobertura do Risco de Longevidade**

## Capítulo 9

# Mortalidade Estocástica

*“Do not worry about your difficulties in mathematics. I can assure you mine are still grater.”* (Albert Einstein, 1879-1955)

### 9.1 Introdução

Nas companhias de seguros do ramo vida, a abordagem tradicional ao cálculo dos prêmios e das reservas matemáticas baseia-se, como vimos nos capítulos anteriores, na utilização de uma intensidade de mortalidade determinística, função apenas da idade do indivíduo (tábua contemporânea), e de uma taxa de juro técnica constante, usada para descontar os cash flows futuros dos produtos.

O comportamento observado nos mercados financeiros e as tendências demográficas recentes vieram demonstrar, à saciedade, que nem a taxa de juro nem a intensidade de mortalidade são determinísticas, uma evidência particularmente inquietante para as companhias que oferecem produtos com benefícios de longo prazo em caso de vida (e.g., pensões e rendas vitalícias). Neste contexto, as companhias estão, grosso modo, expostas a três tipos de riscos: riscos financeiros, riscos de mortalidade sistemáticos (riscos de longevidade), relativos à evolução futura da intensidade de mortalidade subjacente, e riscos de mortalidade não sistemáticos, relativos a uma (eventual) evolução aleatória adversa na mortalidade na carteira.

No passado, as companhias lidavam com o problema escolhendo taxas de juro e tábuas de mortalidade prudentes (do ponto de vista da empresa), o que fazia com que à medida que o retorno dos investimentos e o comportamento da mortalidade eram conhecidos se gerassem excedentes, posteriormente distribuídos (parcial ou totalmente) às pessoas seguras sob a forma de bónus. Como os contratos tem normalmente durações superiores

a 30-35 anos, o que no início se consideravam valores prudentes sobre a evolução das variáveis de interesse revela-se, hoje, perfeitamente desajustado. O risco de longevidade tem uma natureza distinta da dos riscos financeiros. Com efeito, enquanto que os preços dos activos financeiros tendem a exibir uma volatilidade elevada, o que suscita uma necessidade permanente de adoptar estratégias de cobertura adequadas, as alterações na intensidade de mortalidade tendem a ocorrer de forma lenta no tempo, razão pela qual a atenção dada à incerteza associada a este factor de risco tem sido menor.

No capítulo anterior vimos que uma das formas de lidar com o problema da longevidade passa pela utilização de tábuas de mortalidade dinâmicas ou prospectivas. Para a sua construção, foi apresentado um leque alargado e variado de metodologias que permitem, ora em tempo discreto, ora em tempo contínuo, a implementação de uma abordagem que, sendo dinâmica é, na sua essência, determinística.

Neste capítulo, adoptamos uma perspectiva mais ampla e consideramos que a intensidade de mortalidade pode ser modelada através de processos estocásticos. Esta abordagem, simultaneamente dinâmica e estocástica, permite-nos captar de forma mais realista duas características importantes da intensidade de mortalidade: dependência temporal e incerteza sobre a trajectória futura.

A introdução de uma intensidade estocástica comporta vantagens importantes. Em primeiro lugar, ela permite a derivação de prémios e reservas mais realistas face à natureza dos factores de risco subjacentes. Em segundo, ela vai de encontro às orientações recentemente propostas pelo International Accounting Standards Board (IASB) (IASB, 2001, 2004), nas quais se reconhece a necessidade de explicitar adequadamente todas as fontes de risco, incluindo os riscos não financeiros, quer estes sejam diversificáveis (e.g., desvios aleatórios da mortalidade face aos seus níveis esperados) ou não diversificáveis (risco de longevidade), e advogam a contabilização das responsabilidades das companhias de seguros pelo seu valor de mercado.<sup>1</sup>

A um outro nível, a adopção de uma estrutura conceptual mais rica para a intensidade de mortalidade permite equacionar, de forma mais rigorosa, um conjunto de estratégias de cobertura do risco de longevidade, que envolvem nomeadamente a sua transferência para outras contrapartes. Entre elas, consideramos em particular a possibilidade de transferir o risco de longevidade para a pessoa segura (e.g., através de contratos com

---

<sup>1</sup>Em termos mais concretos, o IASB define como valor justo (*fair value*) de uma carteira de responsabilidades de uma companhia de seguros o seu valor de troca numa hipotética transacção em mercado secundário. Apesar de não existirem mercados secundários organizados para a maioria destas responsabilidades, e em consequência não serem observados preços de troca, é defendida a utilização de argumentos de não-arbitragem e a avaliação das responsabilidades com base no desconto dos cash flows futuros esperados. A aplicação de uma abordagem *risk-neutral* esbarra, contudo, na incompletude dos mercados e requer, por isso, uma análise mais cuidada, que retomaremos no Capítulo 11.

prémios e/ou benefícios indexados à evolução da intensidade de mortalidade), para o mercado (e.g., mediante a emissão de obrigações com cupões e/ou principal indexado à evolução da longevidade), as tradicionais opções de resseguro e securitização, e a adopção de estratégias de *hedging* assentes em derivados financeiros (opções, futuros, swaps, etc.) contingentes à evolução da intensidade de mortalidade.

O capítulo está organizado da seguinte forma. Na Secção 9.2, introduzimos as principais ferramentas matemáticas usadas no capítulo. Na Secção 9.3, descrevemos o enquadramento teórico adoptado para modelar a mortalidade num contexto estocástico. Em particular, exploramos o paralelismo entre a modelação do incumprimento no âmbito da literatura sobre risco de crédito e a intensidade de mortalidade. É justificada a adopção de equações diferenciais estocásticas (com saltos) para a intensidade de mortalidade e a representação da probabilidade de sobrevivência mediante funções do tipo afim, aproveitando alguns resultados bem estabelecidos na literatura sobre taxas de juro. Nesta secção efectuamos ainda uma revisão crítica da literatura publicada até ao momento sobre esta matéria, realçando o facto da análise da mortalidade num contexto estocástico se encontrar ainda num estágio inicial de desenvolvimento.

Na Secção 9.4, estudamos a adequação de um conjunto de processos estocásticos do tipo afim para ajustar a trajectória da intensidade de mortalidade. Os modelos analisados, na sua maioria decalcados da literatura sobre taxas de juro, consideram uma única variável de estado ou fonte de incerteza, admitem soluções analíticas fechadas e distinguem-se, essencialmente, pelo facto de estipularem diferentes processos estocásticos para a intensidade de mortalidade.

Na Secção 9.5, desenvolvemos uma nova abordagem teórica ao problema da modelação da mortalidade num contexto estocástico onde consideramos, de forma simultânea, a evolução da intensidade em todas as idades que integram o arco da vida humana. Esta formulação assenta no princípio simples de que a intensidade de mortalidade pode ser representada por uma função do tipo afim de um conjunto de factores latentes, com dinâmica ditada por equações diferenciais estocásticas com saltos. São analisados três casos particulares, onde se catapulta para um contexto estocástico algumas das mais conhecidas leis de mortalidade determinísticas analisadas no Capítulo 2.

## 9.2 Ferramentas matemáticas

Nesta secção, introduzimos as principais ferramentas matemáticas adoptadas no capítulo para modelar a mortalidade num contexto estocástico e para caracterizar algumas aplicações actuariais. Para um tratamento mais detalhado deste tema, recomendamos

a leitura de Karatzas e Shreve (1991), Musiela e Rutkowski (1997), Braumann (2005) e das numerosas referências citadas no texto.

Consideremos o seguinte espaço de probabilidade filtrado (ou base estocástica)  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ . O universo ou espaço amostral  $\Omega$  representa o conjunto (suposto não-vazio) de todos os possíveis resultados de uma experiência ou fenómeno aleatório.  $\mathcal{F}$  é uma álgebra- $\sigma$  de eventos, i.e., uma classe não-vazia de subconjuntos de  $\Omega$  fechada para a complementação (se  $A \in \mathcal{F}$ , então o complemento  $A^c := \Omega - A \in \mathcal{F}$ ) e para uniões contáveis (se  $A_i \in \mathcal{F}$ ,  $i = 1, 2, \dots$ , então  $A := \cup_{i=1}^{\infty} A_i \in \mathcal{F}$ ). Os conjuntos  $A \in \mathcal{F}$  são habitualmente designados de acontecimentos ou conjuntos mensuráveis.

A probabilidade  $\mathbb{P}$  é uma função de  $\mathcal{F}$  em  $[0, 1]$ , normada ( $\mathbb{P}(\Omega) = 1$ ) e aditiva- $\sigma$  (a probabilidade de uma união de elementos disjuntos de  $\mathcal{F}$  é igual à soma das probabilidades individuais de cada elemento, i.e.,  $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$  para  $A_i \in \mathcal{F}$  e  $A_i \cap A_j = \emptyset \forall i \neq j$ ) que atribui probabilidades aos eventos de  $\mathcal{F}$ .

A filtragem  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  define uma família de sub-álgebras- $\sigma$  não decrescente  $\mathcal{F}_s \subset \mathcal{F}_t$ , quando  $s \leq t$ , representando o fluxo de informação sobre o processo estocástico ao longo do tempo, onde  $\mathcal{F}_t$  contém a informação revelada a todos os agentes até ao momento  $t$ . Admitamos ainda que todas as filtrações satisfazem, após o necessário alargamento se for o caso, as denominadas condições habituais, i.e., continuidade à direita e completude do espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Um processo estocástico no espaço de probabilidade filtrado  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  é uma colecção indexada de variáveis aleatórias  $\{X_t\}_{t \in I}$ , onde  $t$  denota o tempo e o conjunto de índices  $I$  define um intervalo de tempo (e.g., da forma  $[0, +\infty)$ ). Como cada variável aleatória  $X_t = X_t(\omega)$  é uma função do “acaso”  $\omega \in \Omega$ , um processo estocástico é simultaneamente função do tempo e do acaso. Um processo estocástico diz-se adaptado à filtragem  $(\mathcal{F}_t)_{t \in I}$  se, para cada  $t \in I$ ,  $X_t$  for  $\mathcal{F}_t$ -mensurável. Um processo estocástico é previsível se o seu valor em cada momento  $t$  depender apenas da informação disponível até esse momento (exclusive), contida na filtragem  $\mathcal{F}_t$ .

### 9.2.1 Processos pontuais e intensidades aleatórias

Neste secção introduzimos alguns resultados sobre processos pontuais, intensidades aleatórias e processos duplamente estocásticos. Todas as propriedades descritas abaixo respeitam a uma dada base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ .<sup>2</sup>

---

<sup>2</sup>Para um tratamento mais detalhado deste tema recomendamos a leitura de, por exemplo, Brémaud (1981), Duffie (1996), Fleming e Harrington (1991) e Rolski *et al.* (1999).

### Processos pontuais

Consideremos em primeiro lugar a definição de tempos de paragem, processos pontuais e respectivos processos de contagem. De uma maneira muito simples, os dois primeiros conceitos representarão as ocorrências do fenómeno em análise, enquanto o terceiro regista essas mesmas ocorrências.

**Definição 9.2.1 (Tempos de paragem, processos pontuais, processos de contagem)** *Uma variável aleatória  $T$  em  $\mathcal{T} = [0, \infty]$ , fazendo a correspondência entre  $\Omega$  e o horizonte temporal  $\mathcal{T}$ , diz-se um tempo de paragem (stopping time) com respeito à filtragem  $\mathcal{F}_t$  se o evento  $\{\omega : T(\omega) \leq t\} \in \mathcal{F}_t, \forall t \in \mathcal{T}$ . Um processo pontual (point process) é uma sequência não decrescente de tempos de paragem  $(T_n)_{n \geq 0}$  em  $[0, \infty]$  tal que  $T_0 = 0$  e  $T_n < T_{n+1}$  no conjunto  $\{T_n < \infty\}$ , para todo o  $n$ . O correspondente processo de contagem (counting process)  $N$  é definido segundo*

$$N_t = \sum_{n=1}^{\infty} \mathbb{I}_{\{T_n \leq t\}} = \begin{cases} n & \text{para } t \in [T_n, T_{n+1}), n \geq 0 \\ \infty & \text{se } t \geq T_{\infty}, \end{cases}$$

onde  $T_{\infty} = \lim_{n \rightarrow \infty} T_n$  designa o momento de explosão de  $N$ . Consequentemente, denotamos por  $N_{\infty} = \lim_{n \rightarrow \infty} N_t$  o número total de momentos de chegada.

O processo  $N$  tem trajectórias que são funções discretas começando em 0 e com saltos de amplitude 1. A variável  $T_n$  simboliza o momento do  $n^{\text{ésimo}}$  salto do processo  $N$ , enquanto  $N_t$  regista o número de saltos ocorridos até ao momento  $t$  (inclusive). A informação considerada pelos processos  $N$  e  $(T_n)$  é a mesma, uma vez que  $\{N_t \geq n\} = \{T_n \leq t\} \forall t \geq 0$  e  $n$  inteiro não-negativo. Tal implica que, dada uma filtragem  $\mathbb{F}$ ,  $N$  é adaptado a  $\mathbb{F}$  sse  $T_n$  é um tempo de paragem- $\mathbb{F}$  para todo o  $n \geq 0$ . Daqui resulta que a filtragem mínima com respeito à qual os  $T_n$ 's são tempos de paragem é a filtragem natural  $\mathbb{F}^N = (\mathcal{F}_t^N)_{t \geq 0}$  com  $\mathcal{F}_t^N = \sigma(N_s : 0 \leq s \leq t)$ . No nosso caso, estamos interessados em processos de contagem não explosivos.

**Definição 9.2.2 (Processos pontuais integráveis e não explosivos)** *Um processo pontual diz-se não explosivo se  $T_{\infty} = \infty$   $\mathbb{P}$ -q.c.<sup>3</sup> ou, de forma equivalente, se  $N_t < \infty$  para todo o  $t \geq 0$   $\mathbb{P}$ -q.c.. Diz-se integrável se, adicionalmente,  $\mathbb{E}[N_t] < \infty$  para todo o  $t \geq 0$ .*

---

<sup>3</sup>Usaremos a abreviatura q.c. para designar a expressão *quase certamente*.

A condição de não-explosão significa que existe um número finito de momentos de chegada em cada intervalo de tempo limitado.

### Intensidade aleatória

Considere-se um processo de contagem não explosivo. Como ele é crescente, segundo a decomposição de Doob-Meyer existe um único processo crescente previsível  $\Lambda$  tal que  $\Lambda_0 = 0$  e  $M = N - \Lambda$  é uma martingala local- $\mathbb{F}$  (Fleming e Harrington, 1991). O processo  $\Lambda$  é designado de *compensador previsível* (*predictable compensator*) de  $N$ , enquanto  $M$  é apelidado de *processo de contagem compensado*. O conceito de intensidade aleatória relaciona-se com  $\Lambda$  nos termos da seguinte definição:

**Definição 9.2.3 (Intensidade estocástica)** *Seja  $N$  um processo de contagem- $\mathbb{F}$  não explosivo e  $\lambda$  um processo não-negativo. Diz-se que  $\lambda$  é a intensidade- $\mathbb{F}$  de  $N$  se o compensador de  $N$ ,  $M_t = N_t - \Lambda_t$ , verificar  $\Lambda_t = \int_0^t \lambda(s) ds$ .*

A intensidade é invariante em relação a alargamentos na filtragem, desde que a informação adicional seja independente de  $\mathbb{F}$ . O objectivo subjacente à especificação de uma intensidade aleatória é o de descrever o número médio de acréscimo de saltos verificado num pequeno intervalo de tempo, como demonstramos em seguida no caso em que admitimos que  $\lambda$  é limitado e contínuo (neste caso  $M$  é uma martingala). De acordo com a definição 9.2.3, para qualquer  $\Delta t, t \geq 0$  temos

$$\mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t] = \mathbb{E}\left[\int_t^{t+\Delta t} \lambda(u) du \middle| \mathcal{F}_t\right]. \quad (9.2.1)$$

Com base no teorema do valor médio para a integração<sup>4</sup>, o lado direito da equação (9.2.1) pode ser escrito segundo

$$\mathbb{E}\left[\int_t^{t+\Delta t} \lambda(u) du \middle| \mathcal{F}_t\right] = \Delta t \mathbb{E}[\lambda_{\xi(t, \Delta t)} | \mathcal{F}_t],$$

onde  $t \leq \xi(t, \Delta t) \leq t + \Delta t$   $\mathbb{P}$ -q.c.. Se calcularmos os limites em ambos os lados da equação (9.2.1) e explorarmos as propriedades assintóticas de  $\lambda$ , podemos usar o teorema

---

<sup>4</sup>De acordo com o teorema do valor médio para a integração, existe um  $c$  em  $[x_1, x_1 + \Delta x]$  tal que:

$$\int_{x_1}^{x_1 + \Delta x} f(t) dt = f(c) \Delta x.$$

da convergência dominada para escrever<sup>5</sup>

$$\lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t]}{\Delta t} = \mathbb{E} \left[ \lim_{\Delta t \rightarrow 0} \lambda_{\xi(t, \Delta t)} \middle| \mathcal{F}_t \right] = \lambda(t)$$

donde se deduz

$$\mathbb{E}[N_{t+\Delta t} - N_t | \mathcal{F}_t] = \lambda(t) \Delta t + o(\Delta t), \quad (9.2.2)$$

onde  $o(\Delta t)$  é tal que  $\lim_{\Delta t \rightarrow 0} o(\Delta t) / \Delta t = 0$ .

A equação (9.2.2) põe em evidência a importância da intensidade  $\lambda$  na determinação do número médio (condicionado) de saltos do processo estocástico num pequeno intervalo de tempo. O condicionamento do valor esperado em  $\mathcal{F}_t$  significa que a informação disponível no momento  $t$  fornece uma indicação sobre o número esperado de saltos num futuro próximo ou, o que é o mesmo, sobre a verosimilhança de um salto no futuro imediato. Refira-se, no entanto, que esta informação não é suficiente para prever a ocorrência efectiva desse salto, que surge assim como uma “surpresa total” no sentido informacional.

### Processos duplamente estocásticos

Recordemos em primeiro lugar a definição e as propriedades básicas de um processo de Poisson  $N$  definido com respeito a uma base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  e uma intensidade determinística  $\lambda(\cdot)$ , onde  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  é uma função localmente integrável (i.e.,  $\int_0^t \lambda(s) ds < \infty$  para  $t \geq 0$ ).

**Definição 9.2.4 (Processo de Poisson)** *O processo de Poisson- $\mathbb{F}$  em  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  é um processo de contagem- $\mathbb{F}$   $N$  com intensidade determinística  $\lambda(\cdot)$  se os seus incrementos  $N_s - N_t$ , para  $0 \leq t \leq s$ , satisfazem as seguintes condições:*

- $N_s - N_t$  é independente de  $\mathcal{F}_t$
- $N_s - N_t$  segue uma lei de Poisson com parâmetro  $\Lambda(s) - \Lambda(t)$ , onde  $\Lambda(h) = \int_0^h \lambda(u) du$  para  $h \geq 0$ .

---

<sup>5</sup>O teorema da convergência dominada de Lebesgue estabelece que, se uma sequência  $\{f_n : n = 1, 2, 3, \dots\}$  de funções reais mensuráveis num espaço de medida  $S = (\Omega, \mathcal{F}, \mu)$  converge em todo o seu domínio, e é dominada por uma qualquer função não-negativa  $g$  em  $L^1$  (i.e.,  $|f_n(x)| \leq g(x)$ ,  $\forall n$ ), então

$$\int_S \left( \lim_{n \rightarrow \infty} f_n \right) d\mu = \lim_{n \rightarrow \infty} \int_S f_n d\mu.$$

Em termos mais específicos, para qualquer  $0 \leq t \leq s$  e parâmetro não-negativo  $k$ , podemos escrever

$$\mathbb{P}(N_s - N_t = k | \mathcal{F}_t) = \frac{(\Lambda_s - \Lambda_t)^k}{k!} e^{-(\Lambda_s - \Lambda_t)}. \quad (9.2.3)$$

Quando  $\mathbb{F} = \mathbb{F}^N$ , diz-se simplesmente que  $N$  é um processo de Poisson com intensidade  $\lambda(\cdot)$ . O caso em que  $\lambda(\cdot) \equiv \bar{\lambda}$ , com  $\bar{\lambda} > 0$ , é referido como um processo de Poisson *homogéneo*, uma vez que para qualquer  $t$  e  $s$  fixos (com  $s \geq t \geq 0$ ), os incrementos  $N_{t+s} - N_t$  têm todos a mesma lei de probabilidade, que é igual à da variável aleatória  $N_s$ . Quando  $\bar{\lambda} \equiv 1$ ,  $N$  corresponde a um processo de Poisson standard.

Um *processo duplamente estocástico* (também conhecido por *processo de Cox*) é uma generalização do processo de Poisson no sentido de que permite que a intensidade seja aleatória ou estocástica.

**Definição 9.2.5 (Processo duplamente estocástico)** *Seja  $N$  um processo de contagem- $\mathbb{F}$  não explosivo com intensidade  $\lambda$  definido em  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  e  $\mathbb{G}$  uma subfiltragem de  $\mathbb{F}$ , tal que  $\mathcal{G}_t \subset \mathcal{F}_t$  para  $t \geq 0$ . Dizemos que  $N$  é duplamente estocástico na filtragem  $\mathbb{G}$  se  $\lambda$  é  $\mathbb{G}$ -previsível e se, condicional em  $\mathcal{F}_t \vee \mathcal{G}_T$  (para  $T \geq t \geq 0$ ), os incrementos  $N_T - N_t$  seguem uma distribuição de Poisson com parâmetro  $\Lambda_T - \Lambda_t$ , onde  $\Lambda_h = \int_0^h \lambda(u) du$  para  $h \geq 0$ . Especificamente, para qualquer  $0 \leq t \leq T$  e parâmetro não-negativo  $k$  temos*

$$\begin{aligned} \mathbb{P}(N_T - N_t = k | \mathcal{F}_t \vee \mathcal{G}_T) &= \frac{(\Lambda_T - \Lambda_t)^k}{k!} e^{-(\Lambda_T - \Lambda_t)} \\ &= \exp\left(-\int_t^T \lambda(u) du\right) \frac{\left(\int_t^T \lambda(u) du\right)^k}{k!}. \end{aligned} \quad (9.2.4)$$

Em particular, se assumirmos que  $N_0 = 0$ , temos

$$\mathbb{P}(N_T = 0 | \mathcal{F}_0 \vee \mathcal{G}_T) = \exp\left(-\int_0^T \lambda(u) du\right).$$

A definição 9.2.5 estabelece que, para uma trajectória particular  $t \mapsto \lambda(t, \omega)$  de  $\lambda(t)$ , com  $\omega \in \Omega$  fixo, o processo de contagem  $N$  torna-se um processo de Poisson não-homogéneo com parâmetro  $\int_0^\cdot \lambda(s, \omega) ds$  (Biffis, 2005). A ideia subjacente à especificação de uma subfiltragem  $\mathbb{G}$  é a de que os tempos de salto de  $N$  são momentos de paragem com respeito a  $\mathbb{F}$ , mas fora do domínio de  $\mathbb{G}$ , que por sua vez comporta informação suficiente para revelar a intensidade  $\lambda$ , i.e., a probabilidade de esses saltos ocorrerem.

Um processo duplamente estocástico determinado por uma dada filtragem  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$  e com intensidade  $\lambda$  pode ser construído (e simulado) usando uma sequência de variáveis aleatórias  $E_n$  com distribuição exponencial standard <sup>6</sup>, independentes de  $\mathbb{G}$ , despoletando a ocorrência de saltos em  $N$ . Designemos por  $T_n$  os tempos de ocorrência dos saltos (*jump-arrival times*) de  $N$ , definidos por

$$T_n = \inf \left\{ t \geq T_{n-1} : \int_{T_{n-1}}^t \lambda(u) du = E_n \right\}, \quad n \geq 1, \quad (9.2.5)$$

com  $T_0 = 0$ . Com base em (9.2.5), podemos verificar que o  $n^{\text{ésimo}}$  salto de  $N$  está mais “próximo” quando as realizações de  $\lambda$  são maiores e o integral alcança o valor de  $E_n$  mais rapidamente.

Uma consequência imediata da natureza duplamente estocástica do processo é a de que, considerando a lei das esperanças iteradas e (9.2.4), a probabilidade de, no momento  $t$ , ocorrerem  $k$  saltos adicionais antes de  $T \geq t$ , pode ser expressa segundo

$$\begin{aligned} \mathbb{P}(N_T - N_t = k | \mathcal{F}_t) &= \mathbb{E} [\mathbb{I}_{\{N_T - N_t = k\}} | \mathcal{F}_t] \\ &= \mathbb{E} [\mathbb{P}(N_T - N_t = k | \mathcal{F}_t \vee \mathcal{G}_T) | \mathcal{F}_t] \\ &= \mathbb{E} \left[ \frac{(\Lambda_T - \Lambda_t)^k}{k!} e^{-(\Lambda_T - \Lambda_t)} \Big| \mathcal{F}_t \right] \end{aligned} \quad (9.2.6)$$

Diz-se que um tempo de paragem  $\tau$  tem uma intensidade  $\lambda$  se é o momento do primeiro salto (*first jump time*) de um processo de contagem não explosivo  $N$  com intensidade  $\lambda$ . Se  $N$  é duplamente estocástico e ditado por  $\mathbb{G} \subset \mathbb{F}$ , então  $\tau$  diz-se duplamente estocástico com intensidade  $\lambda$ .

Se fixarmos  $k = 0$  em (9.2.6), a probabilidade de  $\tau$  ocorrer depois de  $T \geq t$  (dado que  $\tau > t$ ) pode ser representada por

$$\mathbb{P}(N_T - N_t = 0 | \mathcal{F}_t) = \mathbb{E} \left[ e^{-\int_t^T \lambda(u) du} \Big| \mathcal{F}_t \right]. \quad (9.2.7)$$

Um outro resultado interessante refere-se à função de densidade de probabilidade de um tempo de paragem duplamente estocástico  $\tau$ . Seja  $p(\cdot)$  a “função de sobrevivência” de  $\tau$ , i.e., a função  $p : \mathbb{R}_+ \mapsto [0, 1]$  tal que  $p(t) = \mathbb{P}(\tau > t)$ . Assumindo que existe, a função de densidade de  $\tau$ ,  $f(t)$ , é dada por  $-p'(t)$ . A função de densidade de  $\tau$

<sup>6</sup>Ou seja, respeitando  $\mathbb{P}(E_n > \alpha) = e^{-\alpha}$ , para  $n$  inteiro.

condicional a  $\{\tau > t\}$ ,  $f_t(s)$  (para  $0 \leq t \leq s$ ), é dada por

$$f_t(s) = \frac{d}{ds} \mathbb{P}(\tau \leq s | \mathcal{F}_t) = \mathbb{E} \left[ \lambda(s) e^{-\int_t^s \lambda(u) du} \middle| \mathcal{F}_t \right]. \quad (9.2.8)$$

A importância destes resultados no contexto actuarial tornar-se-á evidente na secção seguinte, onde  $\tau$  é interpretado como o tempo de vida remanescente de um indivíduo com idade  $x$ .

### 9.2.2 Processos do tipo afim

Dada uma base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , um processo do tipo afim (*affine process*)  $X$  em  $\mathbb{R}^n$ , relativo a um determinado espaço de estados  $D \subset \mathbb{R}^n$ , é um processo de Markov cuja função característica condicional é do tipo

$$\mathbb{E} \left[ e^{iu \cdot X_T} \middle| \mathcal{F}_t \right] = e^{\alpha(t) + \beta(t) \cdot X_t}, \quad (9.2.9)$$

com  $u \in \mathbb{R}^n$  e  $T \geq t \geq 0$  e funções  $\alpha(\cdot) \doteq \alpha(\cdot; u; T) : [0, \infty) \rightarrow \mathbb{C}$  e  $\beta(\cdot) \doteq \beta(\cdot; u; T) : [0, \infty) \rightarrow \mathbb{C}^n$ . O processo diz-se regular se os coeficientes  $\alpha(\cdot)$  e  $\beta(\cdot)$  da função característica são diferenciáveis e as suas derivadas contínuas em  $t = 0$ .

As propriedades fundamentais dos processos do tipo afim são analisadas em detalhe por Duffie *et al.* (2003) e Filipovic (2005). Nesta secção adoptamos uma perspectiva restritiva mas mais intuitiva destes processos, habitualmente usada no contexto das aplicações financeiras. Tomando como referência Duffie (1996, 2005), Björk (1998), Duffie e Kan (1996), Duffie *et al.* (2000) e Piazzesi (2005), os processos do tipo afim são definidos em termos de soluções para determinadas equações diferenciais estocásticas (EDEs) numa base estocástica filtrada fixa.

Dada uma base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , um processo do tipo afim com saltos (*affine jump-diffusion process*)  $X$  em  $\mathbb{R}^n$  é um processo de Markov especificado como a solução  $X$  da seguinte EDE

$$dX_t = \delta(t, X_t)dt + \sigma(t, X_t)dW_t + \sum_{h=1}^m dJ_t^h, \quad (9.2.10)$$

onde  $W$  é um  $\mathbb{F}$ -movimento browniano standard em  $\mathbb{R}^n$  e cada  $J^h$  é um processo *pure-jump* em  $\mathbb{R}^n$  com intensidade de salto  $\{\eta^h(t, X_t) : t \geq 0\}$  e função de distribuição da dimensão dos saltos  $\nu_t^h$  em  $\mathbb{R}^n$ . Um requisito fundamental é o de que a tendência  $\delta$ , a matriz de variâncias-covariâncias  $\sigma\sigma^T$  e a intensidade de salto  $\eta^h$  devem apresentar uma dependência do tipo afim em relação a  $X$ . De forma inversa, processos estocásticos do

tipo (9.2.10) são processos do tipo afim no sentido de (9.2.9).

Em termos mais explícitos, a dependência de tipo afim dos coeficientes da equação (9.2.10) pode ser representada por

$$\begin{aligned} \delta(t, X) &= d_0(t) + d_1(t) \cdot X \\ (\sigma(t, X)\sigma(t, X))^T_{i,j} &= (V_0(t))_{i,j} + (V_1(t))_{i,j} \cdot X \quad i, j = 1, \dots, n \\ \eta^h(t, X) &= \eta_0^h(t) + \eta_1^h(t) \cdot X. \quad h = 1, \dots, m \end{aligned}$$

As funções  $d \doteq (d_0, d_1)$ ,  $V \doteq (V_0, V_1)$  e  $\eta^h \doteq (\eta_0^h, \eta_1^h)$  são definidas em  $[0, \infty)$ , tomam valores em  $\mathbb{R}^n \times \mathbb{R}^{n \times n}$ ,  $\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n \times n}$  e  $\mathbb{R} \times \mathbb{R}^n$ , respectivamente, e assume-se que são limitadas e contínuas. A distribuição da dimensão dos saltos  $\nu_t^h$  do  $h^{\text{ésimo}}$  processo de salto é determinada pela sua Transformada de Laplace, i.e.,

$$\mathcal{L}^h(t, c) = \int_{\mathbb{R}^n} e^{c \cdot z} d\nu_t^h(z), \quad (9.2.11)$$

definida em  $t \in [0, \infty)$ ,  $c \in \mathbb{C}^n$  e de modo a que o integral seja finito.

A escolha de uma função de distribuição dos saltos com transformada de Laplace conhecida (ou facilmente derivável) é obviamente importante para preservar a flexibilidade analítica do modelo. A transformada  $\mathcal{L} \doteq (\mathcal{L}^h)$  e as funções  $d$ ,  $V$  e  $\eta \doteq (\eta^h)$  determinam completamente a distribuição de  $X$ , uma vez fixada a condição inicial  $X_0$ .

Um resultado importante no contexto dos processos do tipo afim é o de que, para qualquer  $a \in \mathbb{C}^n$ , um dado  $T \geq t$  e uma função do tipo afim  $R$  definida por  $R(t, X) = \rho_0(t) + \rho_1(t) \cdot X$  (para uma dada função  $\rho \doteq (\rho_0, \rho_1)$  em  $\mathbb{R} \times \mathbb{R}^n$  contínua e limitada), a transformada  $\phi$  de  $X_T$ , definida por

$$\phi(a, X_t, t, T) = \mathbb{E} \left[ e^{-\int_t^T R(s, X_s) ds} e^{a \cdot X_T} \middle| \mathcal{F}_t \right] \quad (9.2.12)$$

admite, verificadas determinadas condições (Duffie *et al.* (2000), Prop. 1), a seguinte representação

$$\phi(a, X_t, t, T) = e^{\alpha(t) + \beta(t) \cdot X_t}. \quad (9.2.13)$$

Com base no teorema de Feynman-Kac, sabemos que  $\phi(a, X, t, T) \doteq v(X, t)$  é solução da seguinte equação diferencial parcial

$$\frac{\partial v}{\partial t} + \delta \frac{\partial v}{\partial X} + \frac{1}{2} \sigma^2 \frac{\partial^2 v}{\partial X^2} + \sum_{h=1}^m \eta^h \int_{\mathbb{R}^n} [v(X+z, t) - v(X, t)] d\nu_t^h(z) - Xv = 0 \quad (9.2.14)$$

ou, tomando em consideração a representação do tipo afim

$$\begin{aligned} \dot{\alpha}(t) + \dot{\beta}(t) \cdot x + \delta\beta(t) + \frac{1}{2}\sigma^2\beta^2(t) \\ + \sum_{h=1}^m \eta^h \int_{\mathbb{R}^n} [v(X+z, t) - v(X, t)] d\nu_t^h(z) - Xv = 0, \end{aligned} \quad (9.2.15)$$

onde  $\delta \doteq \delta(t, X)$ ,  $\sigma \doteq \sigma(t, X)$ ,  $\eta^h \doteq \eta^h(t, X)$  e onde  $\alpha(\cdot) \doteq \alpha(\cdot; a, T)$ ,  $\beta(\cdot) \doteq \beta(\cdot; a, T)$  respeitam as seguintes equações diferenciais ordinárias (EDO) do tipo Riccati

$$\dot{\beta}(t) = \rho_1(t) - d_1(t)^T \beta(t) - \frac{1}{2} \beta(t)^T V_1(t) \beta(t) - \sum_{h=1}^m \eta_1^h [\mathcal{L}^h(t, \beta) - 1] \quad (9.2.16)$$

$$\dot{\alpha}(t) = \rho_0(t) - d_0(t) \cdot \beta(t) - \frac{1}{2} \beta(t)^T V_0(t) \beta(t) - \sum_{h=1}^m \eta_0^h [\mathcal{L}^h(t, \beta) - 1] \quad (9.2.17)$$

com condições limite  $\alpha(T) = 0$  e  $\beta(T) = a$ , onde  $\eta_j^h \equiv \eta_j^h(t)$  ( $j = 1, 2$ ).

Consideremos agora, para qualquer  $a, b \in \mathbb{C}^n$ ,  $c \in \mathbb{C}$ , uma transformada estendida de  $X_T$ , definida para  $T \geq t$  por

$$\psi(a, b, c, X_t, t, T) = \mathbb{E} \left[ e^{-\int_t^T R(s, X_s) ds} e^{a \cdot X_T} (b \cdot X_T + c) \middle| \mathcal{F}_t \right]. \quad (9.2.18)$$

Verificadas determinadas condições (Duffie *et al.* (2000), Prop. 3), prova-se que

$$\psi(a, b, c, X_t, t, T) = e^{\alpha(t) + \beta(t) \cdot X_t} [\hat{\alpha}(t) + \hat{\beta}(t) \cdot X_t], \quad (9.2.19)$$

onde as funções  $\hat{\alpha}(\cdot) \doteq \hat{\alpha}(\cdot; a, b, c, T)$ ,  $\hat{\beta}(\cdot) \doteq \hat{\beta}(\cdot; a, b, c, T)$  constituem soluções das seguintes EDO

$$\dot{\hat{\beta}}(t) = -d_1(t)^T \hat{\beta}(t) - \beta(t)^T V_1(t) \hat{\beta}(t) - \sum_{h=1}^m \eta_1^h [\nabla^h(t, \beta) \cdot \hat{\beta}(t)] \quad (9.2.20)$$

$$\dot{\hat{\alpha}}(t) = -d_0(t) \cdot \hat{\beta}(t) - \beta(t)^T V_0(t) \hat{\beta}(t) - \sum_{h=1}^m \eta_0^h [\nabla^h(t, \beta) \cdot \hat{\beta}(t)], \quad (9.2.21)$$

com condições limite  $\hat{\alpha}(T) = c$  e  $\hat{\beta}(T) = b$ , onde  $\nabla^h(t, c)$  denota o gradiente de  $\mathcal{L}^h(t, c)$  com respeito a  $c \in \mathbb{C}^n$ , i.e.,  $\nabla^h(t, c) = \int_{\mathbb{R}^n} c \exp(c \cdot z) \nu_t^h(dz)$ .

No caso particular em que  $a = 0$ ,  $b = 0$ ,  $c = 1$ ,  $\rho_0 = 0$ ,  $\rho_1 = 1$ , verifica-se que  $\psi(\cdot) = \mathbb{E} \left[ e^{-\int_t^T X_s ds} \middle| \mathcal{F}_t \right]$ , donde se deriva que  $\psi(\cdot) = e^{\alpha(t) + \beta(t) \cdot X_t}$  com  $\hat{\alpha}(t) = \hat{\alpha}(T) = 1$ ,  $\hat{\beta}(t) = \hat{\beta}(T) = 0$ .

A importância destes resultados na derivação de soluções para um conjunto de aplicações financeiras e para a modelação da mortalidade é evidente. Por exemplo, uma boa parte das especificações do tipo afim usadas para modelar a estrutura temporal de taxas de juro é derivada como um caso especial de (9.2.18), considerando o valor esperado com respeito a uma medida de probabilidade neutra face ao risco e fixando  $r_s \doteq R(s, X_s)$  (com  $r_s$  a denotar uma taxa de juro de curto prazo) e  $a, b = 0$  e  $c = 1$ . Deste leque fazem parte, por exemplo, o modelo de Vasicek (1977), onde  $X$  é um processo de Ornstein-Uhlenbeck, o modelo de Cox, Ingersoll e Ross (1985) (CIR), onde  $X$  é uma equação de difusão de Feller, os modelos multivariados introduzidos por Duffie e Kan (1996) e várias outras extensões que analisaremos nas secções seguintes.<sup>7</sup>

Em termos mais gerais, a tradicional avaliação com base nos fluxos de caixa descontados (com taxa de desconto  $R(s, X_s)$ ) pode ser implementada em todos os casos em que os produtos financeiros têm cash flows abrangidos pelo termo  $e^{a \cdot X_T} (b \cdot X_T + c)$  na especificação (9.2.19).

Por outro lado, as equações (9.2.18) e (9.2.19) podem ser usadas na modelação do risco de crédito, onde o momento do incumprimento (*default*) é expresso em termos do primeiro tempo de salto de um processo de contagem duplamente estocástico  $N$  ditado por um processo do tipo afim  $X$  (veja-se, por exemplo, Lando (1998, 2004), Duffie e Singleton (1999), Schönbucher (1998, 2003), Bielecki e Rutkowski (2002), Ammann (2001) e Duffie (2005)).

A atractividade dos processos do tipo afim está relacionada com as características das EDO (9.2.16)-(9.2.21), que são essencialmente equações de Riccati generalizadas, resolúveis, regra geral, mediante métodos numéricos convencionais. Conforme teremos oportunidade de demonstrar nas secções seguintes, para algumas especificações dos coeficientes  $\delta$ ,  $\sigma$ ,  $\eta^h$  e  $\nu^h$  é possível derivar soluções fechadas para as variáveis de interesse. Tal é o caso, por exemplo, dos modelos de Vasicek, CIR sem saltos, Merton ou Feller. Nos casos em que o modelo inclui saltos, soluções explícitas podem igualmente ser derivadas quando o processo que regula os saltos é do tipo Poisson, dependendo da especificação de  $\nu^h$ .<sup>8</sup>

---

<sup>7</sup>Para uma lista detalhada dos modelos incluídos nesta categoria veja-se Dai e Singleton (2000) e referências nele citadas.

<sup>8</sup>Por exemplo, Das (1998) demonstra que no caso do modelo gaussiano de Vasicek, a escolha da distribuição  $\nu^h$  é relativamente livre. No caso do modelo de CIR, a escolha é mais restrita, devido à restrição de não-negatividade de  $X$ . Ainda assim, soluções fechadas estão disponíveis considerando as distribuições degenerada (dimensão do salto fixa), uniforme, exponencial e binomial.

### 9.3 Modelação da mortalidade num contexto estocástico

A modelação da mortalidade num contexto estocástico é muito semelhante à modelação do incumprimento (*default*) no contexto da literatura sobre risco de crédito. As abordagens adoptadas na avaliação de activos sujeitos a risco de crédito são, grosso modo, agrupadas em duas categorias principais:<sup>9</sup>

- A chamada *abordagem estrutural*, que procura descrever o evento de incumprimento modelando a dinâmica do valor dos activos da empresa e a sua estrutura de capitais, na qual o evento de incumprimento é despoletado quando o valor dos activos da empresa desce abaixo de um determinado patamar.
- A denominada *intensity based approach* (também conhecida por *reduced-form approach*), na qual a evolução do valor total dos activos da empresa não assume qualquer papel e onde o momento de incumprimento é modelado como um tempo de paragem imprevisível, que ocorre de forma totalmente inesperada no sentido informacional. Esta última abordagem baseia-se na especificação exógena da probabilidade condicional de incumprimento, regra geral mediante modelação da intensidade de incumprimento, admitindo-se que esta possa depender, quer de variáveis observáveis, quer de variáveis não observáveis.<sup>10</sup>

Um resultado-chave na abordagem pela intensidade é aquele onde se determina que o preço de activos sujeitos a risco de crédito pode ser feito com base na metodologia habitualmente adoptada na avaliação de activos financeiros sem risco. A única diferença reside no facto de, num contexto “*risk-neutral*”, o preço de um activo ser igual ao valor esperado dos seus cash flows futuros, descontados segundo uma taxa de juro ajustada de modo a ter em conta, quer a probabilidade, quer o momento do incumprimento. Em termos práticos, este resultado significa que é possível adoptar toda a formalização matemática desenvolvida na literatura sobre a estrutura temporal de taxas de juro, bastando para tal que se parametrize adequadamente a taxa de juro ajustada em vez da taxa de juro sem risco.

Transpondo este resultado para o contexto actuarial, podemos interpretar o incumprimento como correspondendo ao evento ditado pelo óbito da pessoa segura e a intensidade de incumprimento como a intensidade de mortalidade. Esta semelhança entre as intensidades de mortalidade e de incumprimento permite-nos aplicar os modelos desenvolvidos

---

<sup>9</sup>Para uma análise detalhada deste tema veja-se, por exemplo, Lando (1998, 2004), Duffie e Singleton (1999), Schönbucher (1998, 2003), Bielecki e Rutkowski (2002), Ammann (2001) e Duffie (2005).

<sup>10</sup>Entre os trabalhos pioneiros nesta abordagem, destacamos a contribuição de Artzner e Delbaen (1995), onde se estabelece o paralelismo com a avaliação de rendas vitalícias.

no contexto da literatura sobre risco de crédito na avaliação de activos contingentes à evolução da mortalidade.

Refira-se, no entanto, que existem algumas diferenças entre o risco de mortalidade e o risco de crédito que se devem reflectir na escolha do modelo a usar. Por exemplo, diferentes empresas no contexto do risco de crédito podem ser associadas a diferentes gerações num modelo de mortalidade. Porém, se na modelação da mortalidade existe uma ordenação natural das gerações (pela sua idade), com correlações elevadas entre gerações adjacentes, em contrapartida não existe uma ordenação das empresas individualmente, apesar do seu incumprimento poder estar, de alguma forma, correlacionado.

A um outro nível, a intensidade de incumprimento num modelo de risco de crédito é susceptível de ser modelada recorrendo a uma equação de difusão com mecanismo de reversão em relação à média, possivelmente homogénea em termos temporais. Em contraste, os modelos de mortalidade são certamente não homogéneos temporalmente e devem incorporar, necessariamente, componentes de não reversão em relação a uma média determinística qualquer. Com efeito, a inclusão de um mecanismo de reversão em relação à média significaria, desde logo, que no caso do comportamento da mortalidade se revelar mais favorável do que o antecipado no passado, o potencial para eventuais progressos estaria substancialmente reduzido no futuro. Por outro lado, em casos extremos os ganhos de longevidade registados no passado poderiam ser invertidos se a velocidade de reversão em direcção à média do modelo fosse significativa, um padrão dificilmente justificável se atendermos à evolução observada na mortalidade ou à percepção generalizada quanto à previsível evolução, por exemplo, dos progressos nos cuidados de saúde.

Noutro patamar, se na literatura sobre risco de crédito a ocorrência de incumprimento pode ser seguida da recuperação (parcial ou total) dos montantes em dívida, na modelação da mortalidade a “perda” é obviamente total. Por fim, a existência de correlação entre a intensidade de incumprimento e taxa de juro é muito mais plausível do que entre esta última e a intensidade de mortalidade.

#### 9.3.1 Estrutura de mortalidade do tipo afim

Admitamos que a incerteza é descrita por uma dada base estocástica fixa  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  e centremos a análise numa pessoa segura com idade  $x$  no momento 0. A duração (aleatória) da sua vida residual é modelada em termos de um  $\mathbb{F}$ -tempo de paragem  $\tau_x$  admitindo uma intensidade  $\mu_x$ . Em termos mais específicos, a variável  $\tau_x$  é entendida como o primeiro tempo de salto (*jump-time*) de um processo de contagem não explosivo  $N$

### 9.3. Modelação da mortalidade num contexto estocástico

com intensidade aleatória  $\mu_x$ , que regista em cada momento  $t \geq 0$  se o indivíduo faleceu ( $N_t \neq 0$ ) ou sobreviveu ( $N_t = 0$ ). O tempo de paragem  $\tau_x$  admite uma intensidade  $\mu_x$  se  $N$  também o faz, i.e., se  $\mu_x$  é um processo não-negativo previsível tal que  $\int_0^t \mu_x(s) ds < \infty$  para  $t \geq 0$  e tal que o processo compensado  $M_t = \left\{ N_t - \int_0^t \mu_x(s) ds : t \geq 0 \right\}$  é uma  $\mathbb{F}$ -martingala local.

A ideia subjacente à especificação de uma intensidade de mortalidade aleatória  $\mu_x$  é a de que, em cada momento  $t \geq 0$  e estado da natureza  $\omega \in \Omega$ , tal que  $\tau_x(\omega) > t$ , podemos escrever

$$\mathbb{P}(t < \tau_x \leq t + \Delta t | \mathcal{F}_t)(\omega) = \mu_x(t, \omega) \Delta t + o(\Delta t, \omega), \quad (9.3.1)$$

onde  $o(\Delta t, \omega)$  é tal que  $\lim_{\Delta t \rightarrow 0} o(\Delta t, \omega) / \Delta t = 0$ .

Esta equação equivale, num contexto estocástico, à definição de quociente de mortalidade instantânea avançada em (2.2.19), um conceito familiar no contexto actuarial. Conforme veremos em seguida, a definição (2.2.19) emerge no actual contexto como um caso particular em que a intensidade de mortalidade é determinística.

A abordagem adoptada nesta secção é, como se depreende, simultaneamente dinâmica e estocástica: em cada momento  $t \geq 0$ , a intensidade aleatória é definida como  $\mu_x(t, \omega)$ , onde  $t \geq 0$  simboliza a contrapartida em tempo contínuo da noção de “tempo de calendário” adoptada no Capítulo 6 no contexto dos modelos dinâmicos de mortalidade, e  $\omega \in \Omega$  traduz o “estado da natureza”, precisando uma trajectória particular de  $\mu_x$ . Em consequência, realçamos que a perspectiva da mortalidade adoptada neste contexto é claramente diagonal, relativa a uma única geração  $x$ , com a incerteza a depender do fluxo de informação contido na álgebra- $\sigma$   $\mathcal{F}_t$ . A este respeito, saliente-se que o modelo permite a actualização no tempo do fluxo de informação sobre a evolução da intensidade  $\mu_x$ .

Do ponto de vista actuarial, a expressão (9.3.1) enfatiza a variação estocástica da intensidade ao longo do tempo, à medida que nova informação sobre a mortalidade da população segura se torna disponível. Tal corresponde, na prática, a uma modificação na base demográfica, representada pela escolha (ou disponibilização) de uma nova intensidade (ou de uma nova tábua de mortalidade da qual a intensidade é extraída) em cada momento  $t$ . O tipo de base (realista ou prudente) não é especificado nesta secção, mantendo-se uma definição geral em relação à medida de probabilidade  $\mathbb{P}$  com base na qual são calculados os valores esperados.

Se admitirmos que a pessoa segura é representativa de um grupo de segurados (e.g., com a mesma idade e estado de saúde), com vidas residuais consideradas independentes

### 9.3. Modelação da mortalidade num contexto estocástico

e identicamente distribuídas, os resultados aqui obtidos podem ser estendidos a todos os indivíduos que integram a população homogénea. Por uma questão de simplificação, daqui em diante deixamos cair a referência à idade  $x$  na notação usada para o tempo de paragem  $\tau_x$ . Em particular, substituímos  $\tau_x$  por  $\tau$  e  $\mu_x(t, \omega)$  por  $\mu_x(t)$ , e referimo-nos a intensidades determinísticas (como as inferidas com base numa tábua de mortalidade) escrevendo  $\mu_{x+t}$ . Referimo-nos a intensidades aleatórias por  $\mu_x(t)$ , onde o índice inferior denota a idade alcançada pelo indivíduo no instante  $(t)$ .

Assumimos que  $N$  é um processo duplamente estocástico determinado por  $\mathbb{G} \subset \mathbb{F}$  e com intensidade aleatória  $\mathbb{G}$ -previsível  $\mu_x$ . Recordamos que, com base nesta hipótese e condicional a uma dada trajectória  $t \mapsto \mu_x(t, \omega)$  da intensidade  $\mu_x$  (para  $\omega \in \Omega$  fixo), o processo de contagem  $N$  associado a  $\tau$  torna-se um processo de Poisson não-homogéneo com parâmetro  $\int_0^\cdot \mu_x(s, \omega) ds$ . Assinale-se que o momento do óbito  $\tau$  é um  $\mathbb{F}$ -tempo de paragem, mas não necessariamente um  $\mathbb{G}$ -tempo de paragem: a ideia por detrás da especificação da intensidade  $\mathbb{G}$ -previsível  $\mu_x$  é a de que  $\mathbb{G}$  contém informação suficiente sobre a probabilidade de se verificar o óbito, mas não sobre a sua ocorrência efectiva, informação essa que está contida na filtragem mais geral  $\mathbb{F}$ . Em termos intuitivos, o facto de  $\tau$  ter uma intensidade significa que o óbito do indivíduo surge como uma “surpresa total” no sentido informacional.

Considere-se um indivíduo com idade  $x + t$  no momento  $t$ , i.e., com idade  $x$  no momento 0. Se fizermos uso da equação (9.2.7), a sua probabilidade de sobrevivência no horizonte  $(t, T]$ , com  $T \geq t \geq 0$  fixo, pode ser definida segundo

$$\mathbb{P}(\tau > T | \mathcal{F}_t) = \mathbb{E} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right], \quad (9.3.2)$$

supondo que  $\tau > t$ . Esta definição é comparável com a sua análoga num contexto determinístico  $T-t p_{x+t}$ , expressa pela fórmula (2.2.10). Assinale-se que o facto de não conhecermos de antemão a evolução futura da mortalidade implica a substituição da probabilidade de sobrevivência derivada num contexto determinístico por um valor esperado, condicionado neste caso pela evolução conhecida do fenómeno até ao momento  $t$ , i.e., pelo fluxo de informação  $\mathcal{F}_t$ .

Num contexto duplamente estocástico, a função de densidade condicional  $f_t(\cdot)$  é dada, no intervalo  $\{\tau > t\}$  (para  $0 \leq t \leq s$ ), pela expressão

$$f_t(s) = \frac{d}{ds} \mathbb{P}(\tau \leq s | \mathcal{F}_t) = \mathbb{E} \left[ \mu_{x+s}(s) e^{-\int_t^s \mu_{x+u}(u) du} \middle| \mathcal{F}_t \right]. \quad (9.3.3)$$

As expressões (9.3.2) e (9.3.3) são casos particulares cobertos pela transformada

(9.2.19), pelo que a utilização de uma estrutura do tipo afim para a mortalidade se revela conveniente do ponto de vista computacional. Neste sentido, considere-se o seguinte processo do tipo afim (com saltos) para  $\mu_x(t)$  nos termos da equação (9.2.10)<sup>11</sup>

$$d\mu_x(t) = \alpha^\mu(t, x, \mu_x(t))dt + \sigma^\mu(t, x, \mu_x(t))dW_t + \sum_{h=1}^m dJ_t^h, \quad (9.3.4)$$

onde  $W$  é um movimento browniano standard com respeito a uma filtragem  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$  e  $J_t^h$  é um processo de salto nos termos do definido na secção anterior.

Esta especificação pode naturalmente ser alargada a um contexto multivariado, admitindo que  $W$  é  $n$ -dimensional e  $\alpha^\mu(\cdot)$  e  $\sigma^\mu(\cdot)$  são vectores (de dimensão  $n \times n$ ), de modo a admitir correlações imperfeitas entre as variações nas taxas  $\mu_x(t)$  observadas em diferentes idades. Acresce que, como veremos mais adiante a propósito dos modelos com factores latentes, as funções  $\alpha^\mu(\cdot)$  e  $\sigma^\mu(\cdot)$  podem elas próprias depender de outros processos de difusão, também eles adaptados a  $\mathcal{F}_t$ .

O uso de uma abordagem descontínua na modelação da intensidade  $\mu_x(t)$ , decorrente da adopção de processos estocásticos com saltos, pode parecer estranho aqueles que, em demografia, se debruçam sobre este fenómeno. A este respeito, notamos que a dimensão e a frequência dos choques descontínuos representados por  $J_t$  não têm que ser necessariamente significativos, garantindo-se assim que o modelo preserva no essencial a natureza do fenómeno. Por outro lado, é nossa opinião que a inclusão desta fonte de risco na dinâmica de  $\mu_x(t)$  proporciona um aumento na flexibilidade do modelo e uma riqueza teórica que compensam, em certa medida, a utilização de uma abordagem com estas características.

A principal desvantagem da especificação Poisson-Gaussiana (9.3.4) é a de que ela admite a possibilidade de ocorrerem valores negativos na intensidade  $\mu_x(t)$ , violando desta forma a condição de não-negatividade. Trata-se de um problema bem identificado no contexto da modelação da estrutura temporal de taxas de juro (e.g., James e Webber (2000), Brigo e Mercurio (2001)), ou na literatura sobre risco de crédito assente em equações de difusão gaussianas (e.g., Lando (1998), Schönbucher (2003)). Na prática, podemos sempre fixar os parâmetros do modelo de modo a assegurar que a probabilidade de se registarem valores negativos em  $\mu_x(t)$  é muito reduzida.

---

<sup>11</sup> Refira-se que o processo  $X$  pode incluir quer variáveis observáveis, quer variáveis não observáveis que influenciem o comportamento da intensidade de mortalidade. No primeiro caso, podemos incluir indicadores (“*markers*”) que espelhem adequadamente a evolução da intensidade, como sejam a esperança de vida ou a taxa de mortalidade relativas a determinadas idades. No segundo caso, podemos pensar em factores genéricos não observáveis, cuja dinâmica possa ser estimada mediante calibragem em relação a indicadores objectivos.

### 9.3. Modelação da mortalidade num contexto estocástico

Como a dinâmica de  $\mu_x(t)$  em (9.3.4) depende apenas do estado corrente do processo, estamos claramente na presença de um processo de Markov. Seja  $S(t, x, \mu_{x+t}(t), T)$  a probabilidade de sobrevivência no horizonte  $(t, T]$  para um indivíduo com idade  $x + t$

$$S(t, x, \mu_{x+t}(t), T) = \mathbb{P}(\tau > T | \mathcal{F}_t) = \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right]. \quad (9.3.5)$$

Note-se que a intensidade de mortalidade é modelada com respeito a uma medida de probabilidade objectiva  $\mathbb{P}$ . Mais adiante veremos que a utilização desta metodologia em aplicações actuariais requer a adopção de uma medida de probabilidade “neutra face ao risco” (*risk-neutral*)  $\mathbb{Q}$ , incorporando as preferências em relação ao nível de risco dos cash flows.

Para uma dada idade  $x$ , designemos por  $M$  a martingala- $\mathbb{P}$  definida por

$$M_t = \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right]$$

interpretada neste contexto como a probabilidade, no momento  $t$ , de sobrevivência no intervalo  $[0, T]$ . Por definição, podemos escrever  $M_t$  segundo

$$\begin{aligned} M_t &= e^{-\int_0^t \mu_{x+s}(s) ds} \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right] \\ &= e^{-\int_0^t \mu_{x+s}(s) ds} S(t, x, \mu_{x+t}(t), T). \end{aligned}$$

Se admitirmos que  $S(\cdot) \in \mathbb{C}^{1,2}$ , da aplicação do Lemma de Itô, da regra do produto e de (9.3.4), resulta

$$\begin{aligned} dM_t &= e^{-\int_0^t \mu_{x+s}(s) ds} (dS_t - \mu_x S_t dt) \\ &= e^{-\int_0^t \mu_{x+s}(s) ds} \left( \partial_t S_t + \alpha_t^\mu \partial_\mu S_t + \frac{1}{2} (\sigma_t^\mu)^2 \partial_{\mu\mu} S_t - \mu_x S_t \right) dt + e^{-\int_0^t \mu_{x+s}(s) ds} \sigma_t^\mu \partial_\mu S_t dW_t \\ &\quad + e^{-\int_0^t \mu_{x+s}(s) ds} \sum_{h=1}^m \left( \eta^h \int_{\mathbb{R}^n} [S(t, \mu_x + z) - S(t, \mu_x)] d\nu_t^h(z) \right), \end{aligned}$$

onde adoptamos a notação simplificada  $S_t \doteq S(t, x, \mu_{x+t}(t), T)$ ,  $\partial_t S_t = \frac{d}{dt} S_t$ ,  $\partial_\mu S_t = \frac{d}{d\mu} S_t$ ,  $\partial_{\mu\mu} S_t = \frac{d^2}{d\mu^2} S_t$ .

Como  $M$  é uma martingala- $\mathbb{P}$ , o termo correspondente à tendência deve ser nulo, donde resulta a seguinte equação diferencial parcial para a probabilidade de sobrevivência

$$\partial_t S(t, \mu, T) + \alpha^\mu(t, \mu) \partial_\mu S(t, \mu, T) + \frac{1}{2} (\sigma^\mu(t, \mu))^2 \partial_{\mu\mu} S(t, \mu, T) - \mu S(t, \mu, T) = 0 \quad (9.3.6)$$



com condição limite  $S(T, x, \mu_x(t), T) = 1$ . A equação (9.3.6) é análoga à equação diferencial derivada para obrigações de cupão zero num contexto de taxas de juro estocásticas.

A opção por uma estrutura do tipo afim permite-nos representar a probabilidade de sobrevivência  $S(t, x, \mu_{x+t}(t), T)$  mediante uma expressão simples do tipo

$$S(t, x, \mu_{x+t}(t), T) = e^{\mathcal{A}(t, x, T) + \mathcal{B}(t, x, T) \cdot \mu_{x+t}(t)}, \quad (9.3.7)$$

onde as funções determinísticas  $\mathcal{A}(t, x, T)$  e  $\mathcal{B}(t, x, T)$  constituem, respectivamente, soluções das equações (9.2.16) e (9.2.17) com condições terminais  $\mathcal{A}(T, x, T) = 0$  e  $\mathcal{B}(T, x, T) = 0$ . Estas equações podem ser resolvidas recorrendo, regra geral, a métodos numéricos, sendo que para algumas especificações da intensidade  $\mu_x(t)$ , devidamente exploradas nas secções seguintes, é possível derivar soluções analíticas fechadas.

A definição (9.3.7) inclui, como caso particular, modelos com intensidade determinística. Com efeito, se escolhermos  $\mathcal{A}(\cdot)$  e  $\mathcal{B}(\cdot)$  segundo

$$\mathcal{A}(t, x, T) = - \int_t^T \mu_{x+s} ds \quad \text{e} \quad \mathcal{B}(t, x, T) = 0,$$

respectivamente, facilmente verificamos que a probabilidade de sobrevivência é dada por  $S(t, x, \mu_{x+t}(t), T) = e^{-\int_t^T \mu_{x+s} ds}$ .

Um outro aspecto interessante diz respeito à relação entre os processos estocásticos para a intensidade  $\mu_{x+t}(t)$  e a tradicional força de mortalidade determinística  $\mu_{x+t}$ , habitualmente usada pelos actuários. Recordamos que  $\mu_{x+t}$  é definida por

$$\mu_{x+t} = \lim_{\Delta t \rightarrow 0^+} \frac{\mathbb{P}[t < T_x \leq t + \Delta t | T_x > t]}{\Delta t} = - \frac{\frac{d}{dt} S(t, x, \mu_{x+t}(t), T)}{S(t, x, \mu_{x+t}(t), T)}.$$

Considerando que probabilidade de sobrevivência admite uma estrutura do tipo afim (9.3.7) e notando que

$$\frac{d}{dt} S(t, x, \mu_{x+t}(t), T) = \frac{d}{dt} \{ \mathcal{A}(t, x, T) + \mathcal{B}(t, x, T) \cdot \mu_{x+t}(t) \} \cdot S(t, x, \mu_{x+t}(t), T),$$

a equação acima pode ser escrita segundo

$$\mu_{x+t} = - \frac{d}{dt} \{ \mathcal{A}(t, x, T) + \mathcal{B}(t, x, T) \cdot \mu_{x+t}(t) \}. \quad (9.3.8)$$

Assinale-se ainda que a parametrização de  $\mu_{x+t}(t)$  por processos do tipo afim permite o cálculo de indicadores demográficos tradicionais como a esperança de vida residual à

idade  $x$  ou a entropia da função de sobrevivência.<sup>12</sup>

Por analogia com o conceito de taxas de juro *forward*, introduzimos o conceito de *intensidade de mortalidade forward* para a geração  $x$  em  $t$  para o momento  $T$ ,  $f^\mu(t, x, T)$ , definido por (Dahl, 2004)

$$f^\mu(t, x, T) = -\frac{d}{dT} \ln S(t, x, \mu_{x+t}(t), T). \quad (9.3.9)$$

De forma equivalente, podemos expressar a relação entre a intensidade de mortalidade *forward* e a probabilidade de sobrevivência mediante

$$S(t, x, \mu_{x+t}(t), T) = \exp\left(-\int_t^T f^\mu(t, x, s) ds\right). \quad (9.3.10)$$

A intensidade de mortalidade *forward*  $f^\mu(t, x, T)$  pode ser interpretada como a intensidade determinística no intervalo  $[s, s + ds)$ , tal que a probabilidade de sobrevivência em  $[s, s + ds)$ , usando esta intensidade, coincide com a probabilidade de sobrevivência determinada no momento  $t$  via (9.3.5). A função de intensidade de mortalidade *forward* para a geração  $x$ ,  $(f_\mu(t, x, s))_{t \leq s \leq T}$ , é uma função que iguala a probabilidade de sobrevivência no momento  $t$  para um indivíduo com idade  $x + t$  a  $e^{-\int_t^\tau f_\mu(t, x, s) ds}$  para  $t \leq \tau \leq T$ . No caso particular em que  $T = t$ , obtemos a *força de mortalidade spot*

$$\mu(t, x + t) = f^\mu(t, x + t, t).$$

### 9.3.2 Revisão da literatura sobre mortalidade estocástica

O interesse pela modelação da mortalidade num contexto estocástico é muito recente. Por este motivo, não é de estranhar a escassez de literatura sobre esta matéria e o estágio pouco consolidado do corpo teórico. A maioria dos modelos que encontramos na literatura resulta da associação que é feita entre os comportamentos da força de mortalidade e das taxas de juro (veja-se, por exemplo, Milevsky e Promislow (2001), Dahl (2004), Biffis (2005), Cairns *et al.* (2006a)). Em particular, ambos são processos positivos, possuem estruturas temporais e, não menos importante, são fundamentalmente estocásticos na sua natureza.

As abordagens possíveis à modelação da dinâmica da força de mortalidade são elencadas por Cairns *et al.* (2006a). Elas replicam, no essencial, o leque de alternativas habitualmente disponíveis na modelação da dinâmica da estrutura temporal de taxas de juro, e podem ser classificadas do seguinte modo:

---

<sup>12</sup>Veja-se, por exemplo, Biffis (2005).

- *short rate mortality models* para a dinâmica de  $\mu_x(t)$ , correspondentes aos modelos desenvolvidos para a taxa de juro *spot* (sem risco) de curto prazo, categoria onde se incluem, entre outros, os modelos de Merton (1973), Vasicek (1977), Cox, Ingersoll e Ross (1985), Hull e White (1990), e Black e Karasinski (1991).
- *forward mortality models* para a dinâmica de  $f^\mu(t, x, T)$ , correspondentes aos denominados modelos de não arbitragem para a estrutura temporal de taxas de juro, onde se destaca a abordagem de Heath, Jarrow e Morton (1992).
- *positive-mortality modelling framework* para as probabilidades de sobrevivência *spot*  $p_x(t)$ , correspondentes à abordagem desenvolvida por Flesaker e Hughson (1996), Rogers (1997) e Rutkowski (1997) para as taxas de juro.
- *modelos de mercado* para a probabilidade de sobrevivência *forward* ou para os preços *forward* de rendas vitalícias, correspondentes aos modelos de mercado (LIBOR e Swap Market Model) desenvolvidos por Brace, Gatarek e Musiela (1997) e Jamshidian (1997).

A estas abordagens acrescem naturalmente as já mencionadas formulações que estabelecem um paralelo entre a modelação da mortalidade e a modelação do incumprimento no contexto da literatura sobre risco de crédito.

A maioria dos modelos de mortalidade estocástica propostos até ao momento enquadra-se nos denominados *short rate mortality models* incidindo, ora sobre o quociente de mortalidade *spot*  $q_x(t)$ , ora sobre a força de mortalidade *spot*  $\mu_x(t)$  (Milevsky e Promislow, 2001; Dahl, 2004; Biffis, 2005; Biffis e Millossovich, 2004, 2006; Schrage, 2006; Di Lorenzo *et al.*, 2006). A par destes, encontramos igualmente um número significativo de formulações em tempo discreto, que partem de uma projecção determinística para os quocientes *spot*  $\hat{q}_x(t)$  (ou para a força de mortalidade  $\hat{\mu}_x(t)$ ) e aplicam um ajustamento que evolui no tempo de forma estocástica (Yang, 2001; Olivier e Jeffery, 2004; Smith, 2005; Biffis e Denuit, 2005; Lin e Cox, 2005b; Biffis *et al.*, 2006; Ballotta e Haberman, 2006; Cairns *et al.*, 2005, 2006b).

Em alternativa, Lin e Cox (2005a) propõem uma metodologia substancialmente diferente das anteriores. Em vez de adoptarem uma especificação concreta para a intensidade  $\mu_x(t)$ , os autores recorrem à Transformada de Wang (1996, 2000, 2002, 2003) para converter taxas projectadas num contexto determinístico naquilo que se poderia designar de «probabilidades ajustadas pelo risco», tendo em mente o *pricing* de activos financeiros contingentes à evolução da longevidade.

Milevsky e Promislow (2001) foram os primeiros a propor um modelo estocástico em tempo contínuo para a força de mortalidade (*hazard rate*), com o objectivo de calcular o preço justo de GAO's (*Guaranteed Annuity Options*) incluídas em anuidades de termos variáveis. Os autores demonstram, num primeiro momento em tempo discreto, que os riscos de taxa de juro e de mortalidade desta opção (*mortality option*) podem ser cobertos e que o seu preço (prémio) pode ser calculado com base no valor de uma carteira replicante constituída por dotes puros (*endowments*), rendas vitalícias e obrigações sem risco.<sup>13</sup> Num segundo momento, os autores calculam o preço da opção em tempo contínuo e num contexto de neutralidade face ao risco, assumindo que a taxa de juro de curto prazo e a intensidade de mortalidade evoluem no tempo de forma independente e segundo um modelo do tipo Cox-Ingersoll-Ross e uma equação de difusão simples com reversão em relação à média derivada da lei de Gompertz (denominada de *mean reverting Brownian Gompertz*), respectivamente.

Em termos formais, os autores consideram a seguinte especificação para a intensidade de mortalidade

$$\mu_x(t) = \mu_x(0) \exp(gt + \sigma Y_t), \quad g, \sigma, \mu_x(0) > 0, \quad (9.3.11)$$

com

$$dY_t = -bY_t dt + dW_t^\mu, \quad Y_0 = 0, \quad b \geq 0. \quad (9.3.12)$$

Como se observa, no caso particular em que  $b = 0$ , o processo estocástico converte-se num tradicional movimento browniano geométrico. Registe-se, por fim, que o modelo não admite solução fechada para a probabilidade de sobrevivência pelo que os resultados numéricos reportados pelos autores resultaram da aplicação de métodos de Simulação de Monte Carlo.

Dahl (2004) desenvolve uma abordagem mais abrangente e considera inicialmente que a intensidade  $\mu_x(t)$  tem um comportamento dinâmico ditado por uma equação de difusão do tipo

$$d\mu_x(t) = \alpha^\mu(t, x, \mu_x(t))dt + \sigma^\mu(t, x, \mu_x(t))dW_t, \quad (9.3.13)$$

onde  $W_t$  é um movimento browniano standard com respeito à filtragem  $\mathbb{F}$ , para se centrar posteriormente em processos estocásticos do tipo afim. O autor deriva equações diferenciais parciais para o preço justo de contratos de seguros e para contratos derivados (*mortality derivatives*) quando a mortalidade subjacente é estocástica. Adicionalmente, analisa um conjunto de mecanismos de cobertura do risco de longevidade, soluções que

---

<sup>13</sup> Refira-se, a propósito, que os argumentos usados dependem da capacidade que a companhia de seguros tem de vender contratos de seguros e rendas vitalícias a uma população com idades situadas num mesmo grupo, uma hipótese considerada algo irrealista.

envolvem a transferência do risco de mortalidade sistemática para outras contrapartes, seja para a pessoa segura (ligando os prémios e/ou os benefícios do contrato à evolução de  $\mu_x(t)$  de uma forma similar à adoptada nos contratos de seguro do tipo *unit-link*), seja para os mercados financeiros (mediante produtos derivados com activo subjacente dependente da evolução de  $\mu_x(t)$ ).

Biffis e Milossovich (2004) generalizam a equação (9.3.13) e trabalham com um processo bidimensional de modo a analisar a evolução no tempo de carteiras de produtos do ramo vida relativas a diferentes gerações (ou classes de risco) de pessoas seguras.

Dahl e Moller (2005) baseiam-se num modelo com estrutura do tipo afim para derivar estratégias de cobertura destinadas a minimizar o risco de mortalidade de carteiras de responsabilidades de companhias de seguros, num contexto particular em que o mercado não oferece a possibilidade de transaccionar derivados de mortalidade.

Biffis (2005) estende a abordagem de Dahl (2004) e emprega processos do tipo afim com saltos do tipo (9.2.10) para modelar, num contexto actuarial, quer os factores de risco financeiros, quer os factores de risco demográficos. O autor deriva fórmulas fechadas (condicionadas à solução de equações diferenciais ordinárias) para o valor teórico de numerosos contratos de seguro do ramo vida. Propõe ainda dois modelos específicos para a evolução de  $\mu_x(t)$ : um primeiro, assente num processo Poisson-Gaussiano definido por

$$\mu_t = m(t) + \eta_1 Y_{t-}, \quad (9.3.14)$$

onde  $m(t)$  traduz uma intensidade determinística, reflectindo hipóteses realistas sobre o comportamento da mortalidade,<sup>14</sup> e o processo  $\eta_1 Y_{t-}$  representa os desvios em relação à base inicial, capturando flutuações aleatórias e/ou sistemáticas com dinâmica da variável de estado  $Y$  descrita pela seguinte EDE

$$dY_t = \gamma(\bar{y}(t) - Y_t)dt + \sigma dW_t - dJ_t, \quad (9.3.15)$$

onde  $\gamma, \bar{y}(t), \sigma > 0$ ,  $\bar{y}(t)$  representa o nível (variável com o tempo) para o qual  $Y$  reverte a uma velocidade determinada pelo parâmetro  $\gamma$ , após a ocorrência de cada flutuação ou salto. A componente de salto é definida por um processo de Poisson independente de

---

<sup>14</sup>Esta função determinística pode representar, por exemplo: (i) uma tábua de mortalidade usada pelo mercado para realizar projecções sobre a mortalidade da população segura; (ii) uma dada base demográfica prudencial, usada para efeitos de *pricing*, constituição de reservas ou para avaliação da solvência das empresas; (iii) uma estimativa realista da mortalidade, incorporando as expectativas não enviesadas sobre a evolução futura do fenómeno, construída a partir da informação mais recente disponível no mercado.

$W$ , com intensidade de chegada do salto  $k > 0$  e dimensão dos saltos exponencialmente distribuída com média  $j > 0$ .

No segundo modelo, o autor considera um processo do tipo afim bidimensional  $Y = (\mu_x(t), \bar{\mu}_x(t))$  em  $\mathbb{R}_+^2$ , cuja primeira componente corresponde à intensidade de mortalidade propriamente dita  $\mu_x(t)$ , enquanto o segundo denota a tendência estocástica  $\bar{\mu}_x(t)$  dessa intensidade. Em particular, os processos  $\mu_x(t)$  e  $\bar{\mu}_x(t)$  têm dinâmica descrita pelas seguintes EDEs

$$d\mu_x(t) = \gamma_1 (\bar{\mu}_x(t) - \mu_x(t)) dt + \sigma_1 \sqrt{\mu_x(t)} dW_t^1 \quad (9.3.16)$$

$$d\bar{\mu}_x(t) = \gamma_2 (m(t) - \bar{\mu}_x(t)) dt + \sigma_2 \sqrt{\bar{\mu}_x(t) - m^{\text{lim}}(t)} dW_t^2, \quad (9.3.17)$$

onde  $W = (W^1, W^2)$  denota um movimento browniano standard em  $\mathbb{R}^2$ ,  $\gamma_1, \gamma_2 > 0$  são parâmetros que ditam a velocidade de reversão em relação à média de  $\mu_x(t)$  para  $\bar{\mu}_x(t)$  e de  $\bar{\mu}_x(t)$  para  $m(t)$ , respectivamente, após a ocorrência de flutuações devidas a  $W$ . As funções  $m(t)$  e  $m^{\text{lim}}(t)$  simbolizam, respectivamente, uma dada base demográfica e um patamar inferior (um limite assintótico biológico) para a tendência estocástica  $\bar{\mu}_x(t)$ , que restringe os ganhos potenciais na evolução da mortalidade, nos termos do definido no Capítulo 6.

Schräger (2006) desenvolve um modelo com estrutura do tipo afim no qual procura descrever a mortalidade em todas as idades que compõem o arco da vida humana de forma simultânea, adoptando para tal uma relação funcional entre a intensidade  $\mu_x(t)$  e a idade. Esta abordagem contrasta com as formulações anteriores onde se modela o comportamento da intensidade de uma única geração.<sup>15</sup>

Biffis e Denuit (2005) e Biffis *et al.* (2006) generalizam o modelo proposto por Lee e Carter (1992) a um contexto estocástico, admitindo em particular que a dinâmica do índice  $\kappa_t$ , que captura as tendências temporais da mortalidade, pode ser descrita por equações diferenciais estocásticas. Em termos formais, considera-se que o modelo de Lee-Carter definido em (6.6.1)-(6.6.2)-(6.6.3) pode ser interpretado como uma versão (em tempo discreto) de uma intensidade  $\mu_x(t)$  em tempo contínuo do tipo

$$\mu_x(t) = \exp(\alpha_x(x+t) + \beta_x(x+t) \cdot \kappa_t), \quad (9.3.18)$$

onde  $\alpha_x$  e  $\beta_x$  são estimativas pontuais das funções  $\alpha_x(\cdot)$  e  $\beta_x(\cdot)$  para cada idade, e onde

---

<sup>15</sup>O autor apresenta ainda algumas aplicações na avaliação de contratos contingentes à evolução da mortalidade e testa a qualidade de ajustamento do modelo usando dados para a população masculina holandesa no período 1950-2002, recorrendo ao filtro de Kalman para estimar os parâmetros.

$\kappa$  é um processo de difusão com dinâmica determinada por

$$d\kappa_t = \delta(t, \kappa_t) dt + \sigma(t, \kappa_t) dW_t, \quad (9.3.19)$$

onde  $W$  denota um movimento browniano standard (potencialmente multivariado).

Note-se que a dependência das funções  $\alpha_x(\cdot)$  e  $\beta_x(\cdot)$  em relação ao tempo se deve ao facto do processo  $\mu_x(t)$  descrever a evolução da intensidade para um indivíduo com idade  $x + t$  no momento  $t$ .<sup>16</sup>

Os modelos anteriores assumem, com alguma propriedade e conveniência, que num contexto de neutralidade face ao risco a dinâmica da estrutura temporal da mortalidade é independente da dinâmica da estrutura temporal das taxas de juro, i.e., assumem que os factores de risco demográficos e financeiros são independentes. A principal vantagem deste pressuposto é a de que ele nos permite separar o *pricing* do risco de mortalidade do *pricing* do risco de taxa de juro. Contrariando esta hipótese, Miltersen e Persson (2005) e Bauer e Russ (2006) consideram a possibilidade de existirem correlações entre os factores de risco. Tomando como referência a abordagem de Heath *et al.* (1992), os autores modelam a intensidade de mortalidade *forward* em vez da intensidade *spot*, e consideram toda a curva das  $f^\mu(t, x, T)$  como uma variável de estado de dimensão  $n$ . Usando argumentos similares aos adoptados por Heath *et al.* (1992), são derivadas condições de não arbitragem para a tendência do processo estocástico.

Olivier e Jeffery (2004) e Smith (2005) propõem uma abordagem alternativa à modelação das taxas de mortalidade *forward*. Os autores desenvolvem a sua formulação a partir das probabilidades de sobrevivência *forward* e exploram as propriedades analíticas da distribuição Gamma para derivar o comportamento dinâmico da mortalidade num contexto *risk-neutral*.

Cairns *et al.* (2006a) sugerem uma abordagem diferente, baseada na formulação de Pelsser (2003), onde modelam a denominada “*Survivor Credit Offer Rate (SCOR)*”, i.e., o retorno de um dote puro (*endowment*) anual em caso de sobrevivência, usando uma estrutura similar à dos chamados LIBOR *market models*. Os autores descrevem o modo como as variáveis *forward SCOR* podem ser calculadas e ilustram a sua aplicação no âmbito do *pricing* de obrigações com cupões e/ou principal contingentes à evolução da longevidade (*longevity bonds*), que analisaremos em detalhe no Capítulo 11.

Ballotta e Haberman (2006) desenvolvem um modelo semelhante ao proposto por Milevsky e Promislow (2001), onde combinam uma projecção determinística para a força

---

<sup>16</sup>Os autores demonstram ainda que é possível adoptar uma versão *risk-neutral* do modelo de Lee-Carter para avaliar contratos de seguros, nomeadamente pela introdução de mudanças adequadas na medida de probabilidade.

### 9.3. Modelação da mortalidade num contexto estocástico

de mortalidade, resultante da aplicação dos modelos lineares generalizados com respeito à idade actuarial e ao tempo cronológico descritos no Capítulo 6, com um ajustamento que evolui no tempo de forma estocástica e que representa as oscilações aleatórias na tendência projectada.

Formalmente, a evolução da força de mortalidade no tempo é determinada por

$$\mu(x+z, t+z) = \mu(x+z, 0) \exp \{(\alpha + \beta(x+t))(t+z) + \sigma_h Y_{t+z}\}, \quad (9.3.20)$$

onde

$$\begin{aligned} \mu(x+z, 0) &= a_1 + a_2 x' + \exp \left\{ b_1 + b_2 x' + b_3 \left[ 2(x')^2 - 1 \right] \right\} \\ x' &= \frac{(x+z) - 70}{50}, \quad x \geq 50, \end{aligned} \quad (9.3.21)$$

e onde  $(Y_t : t \geq 0)$  é um processo estocástico em  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , governado por uma equação do tipo Ornstein-Uhlenbeck respeitando a seguinte EDE

$$dY_t = -aY_t dt + dW_t^\mu, \quad Y_0 = 0, \quad a \geq 0, \quad (9.3.22)$$

com  $(W_t : t \geq 0)$  a denotar um movimento browniano unidimensional, independente dos factores de risco financeiros.

O modelo não admite uma solução fechada pelo que os autores recorrem a métodos de simulação de Monte Carlo para analisar as implicações da introdução da mortalidade estocástica na avaliação de opções de conversão em renda vitalícia com taxa de juro garantida (GAO's).

O modelo de Yang (2001) adopta uma filosofia semelhante ao de Ballotta e Haberman (2006), com a diferença de que a abordagem é desenvolvida em tempo discreto. Tomando como ponto de partida uma previsão para os quocientes de mortalidade  $\hat{q}_x(t)$  realizada no momento 0, os autores estimam o valor actual  $q_x(t)$  mediante<sup>17</sup>

$$q_x(t) = \hat{q}_x(t) \exp \left[ X(t) - \frac{1}{2} \sigma_Y^2 + \sigma_Y Z_Y(t) \right], \quad (9.3.23)$$

onde

$$X(t) = X(t-1) - \frac{1}{2} \sigma_X^2 + \sigma_X Z_X(t) \quad (9.3.24)$$

e  $Z_X(t)$  e  $Z_Y(t)$  são sequências independentes de variáveis i.i.d. que seguem uma dis-

<sup>17</sup>Note-se que  $q_x(t)$  representa, neste caso, a probabilidade de um indivíduo que atinge a idade  $x$  no momento  $t$  falecer antes do momento  $t+1$ .

tribuição normal standardizada. A variável  $X(t)$  determina a tendência estocástica da curva de mortalidade enquanto o termo  $-\frac{1}{2}\sigma_Y^2 + \sigma_Y Z_Y(t)$  modela choques únicos (saltos) associados, por exemplo, a epidemias ou catástrofes.

Cairns *et al.* (2005, 2006b) desenvolvem igualmente um modelo em tempo discreto onde consideram que os quocientes  $q_x(t)$  evoluem no tempo segundo a trajectória ditada por dois factores estocásticos e uma relação funcional derivada da lei de mortalidade de Perks (1932). Formalmente, temos

$$q_x(t) = \frac{e^{a(t)+b(t)x}}{1 + e^{a(t)+b(t)x}}, \quad (9.3.25)$$

onde  $(a(t), b(t))$  é um vector de parâmetros cuja dinâmica das respectivas séries temporais é governada por um modelo *ARIMA*(1, 1, 0). O factor  $a(t)$  afecta todas as idades de igual forma enquanto que o factor  $b(t)$  tem um efeito sobre a mortalidade proporcional à idade do indivíduo. O modelo é facilmente estimado notando que  $p_x(t) = 1 - q_x(t)$  e

$$\ln\left(\frac{q_x(t)}{p_x(t)}\right) = a(t) + b(t)x. \quad (9.3.26)$$

Os autores usam o modelo para determinar a importância do risco de longevidade em *longevity bonds* com diferentes maturidades e relativas a distintas gerações.

### 9.3.3 Requisitos importantes nos modelos

Admitindo que a incerteza subjacente à evolução futura de  $\mu_x(t)$  é bem captada por um processo estocástico, é razoável pensar-se que na modelação do seu comportamento devem ser tidas em conta um conjunto de características peculiares que distinguem a trajectória da mortalidade da seguida, por exemplo, pelos preços das acções, pelas taxas de câmbio ou até mesmo das taxas de juro.

Neste sentido, e à semelhança do que acontece na modelação das taxas de juro, um processo estocástico plausível para  $\mu_x(t)$  deve, segundo Cairns *et al.* (2006a):

- garantir que a intensidade  $\mu_x(t)$  é sempre (por definição) positiva
- ser consistente com a evolução histórica do fenómeno
- assegurar que as tendências de longo prazo do modelo são biologicamente admissíveis. Por exemplo, o modelo deve excluir à partida uma curva de mortalidade invertida, i.e., uma situação em que as taxas  $\mu_x(t)$  nas idades avançadas decrescem com a idade.

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

---

- incorporar necessariamente componentes de não reversão em relação a uma média pré-determinada qualquer, em especial no longo prazo. A inclusão de reversão em relação à média significaria, desde logo, que no caso da longevidade evoluir de forma mais favorável do que o antecipado no passado, o potencial para eventuais progressos estaria substancialmente reduzido no futuro.
- não ignorar que a volatilidade das taxas  $\mu_x(t)$  não é homogénea ao longo de todo o arco da vida humana, com as idades mais avançadas a evidenciarem, regra geral, uma maior volatilidade, devido sobretudo à maior importância do risco amostral neste segmento.
- ser intuitivo e flexível o suficiente para ser aplicável em problemas de *pricing* e *hedging*.
- permitir, desejavelmente, a avaliação de activos contingentes à evolução da longevidade usando soluções analíticas fechadas ou, caso tal não seja possível, métodos numéricos convencionais.

Em face da lista de critérios apresentada, não existe até ao momento uma abordagem ou um modelo que possa ser considerado, de forma consistente, superior aos restantes. Neste sentido, é nossa convicção que a escolha do modelo a aplicar deve basear-se, quer em fundamentos de natureza teórica, quer em argumentos de natureza empírica, que aconselhem a sua utilização em razão do seu ajustamento à realidade observada numa dada população.

### 9.4 Modelos do tipo afim para a intensidade de mortalidade

Nesta secção, investigamos a adaptação de um conjunto de processos estocásticos do tipo afim para a ajustar a trajectória da intensidade de mortalidade  $\mu_x(t)$ . Os modelos analisados, na sua maioria decalcados da literatura sobre taxas de juro (veja-se, por exemplo, Rebonato (1998), Björk (1998), James e Webber (2000), Brigo e Mercurio (2001)), são modelos de factor único, i.e., consideram uma única variável de estado ou fonte de incerteza, diferindo entre si essencialmente pela circunstância de prescreverem distintos processos estocásticos para a intensidade  $\mu_x(t)$ .

A opção por modelos de factor único não significa que ignoramos a possibilidade de a correlação entre intensidades  $\mu_x(t)$  correspondentes às diferentes idades ser imperfeita .

No entanto, não existindo ainda evidência empírica suficiente para sustentar a adopção de modelos multifactor, esta escolha representa um compromisso entre a adequação do modelo e sua complexidade analítica. Com efeito, uma característica importante dos modelos analisados é a de que todos eles admitem uma solução analítica fechada, facilitando sobremaneira a sua utilização no contexto actuarial.

Estudos empíricos recentes no domínio da abordagem *reduced-form approach* à modelação do risco de crédito sugerem que modelos tradicionais como os de Vasicek (1977) e Cox, Ingersoll e Ross (1985) são adequados para representar a trajectória da intensidade de incumprimento (veja-se, por exemplo, Lando (2004), Schönbucher (2003), Bielecki e Rutkowski (2002)). Apesar de estes modelos perfilarem, à partida, um conjunto de características teóricas desajustadas das trajectórias de  $\mu_x(t)$ , entendemos que o teste definitivo à sua capacidade de ajustamento deve ser ditado pelos resultados obtidos nos estudos empíricos, razão pela qual os incluímos na lista de processos a avaliar.

Os modelos analisados de seguida diferem entre si essencialmente pela:

1. inclusão (ou não) de componentes de reversão em relação a uma média de longo prazo;
2. inclusão (ou não) de saltos (*jumps*) no processo estocástico.

#### 9.4.1 Modelo de Vasicek

Consideremos em primeiro lugar a adaptação do modelo com reversão à média de Vasicek (1977). O modelo estipula que a intensidade  $\mu_{x+t}(t)$  evolui no tempo segundo um processo de Ornstein-Uhlenbeck com coeficientes constantes, i.e.,

$$d\mu_{x+t}(t) = a(\theta - \mu_{x+t}(t))dt + \sigma dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.1)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\theta > 0$ ,  $\sigma \geq 0$ , e as demais variáveis assumem o seu significado habitual.

O parâmetro  $a$  denota a velocidade de convergência de  $\mu_{x+t}(t)$  em direcção à sua média de longo prazo  $\theta$ , específica para cada idade. Integrando a equação (9.4.1), obtemos, para  $t \geq s$ ,

$$\mu_{x+t}(t) = \mu_{x+s}(s)e^{-a(t-s)} + \theta(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW(u), \quad (9.4.2)$$

donde se deduz que  $\mu_x(t)$  segue uma distribuição Normal com média e variância (condi-

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

cionadas por  $\mathcal{F}_s$ ) dadas, respectivamente, por

$$\begin{aligned}\mathbb{E} \{ \mu_{x+t}(t) | \mathcal{F}_s \} &= \mu_{x+s}(s)e^{-a(t-s)} + \theta \left( 1 - e^{-a(t-s)} \right) \\ \text{Var} \{ \mu_{x+t}(t) | \mathcal{F}_s \} &= \frac{\sigma^2}{2a} \left[ 1 - e^{-2a(t-s)} \right].\end{aligned}\quad (9.4.3)$$

No modelo de Vasicek, existe a probabilidade positiva de que  $\mu_{x+t}(t)$  possa assumir valores negativos. Essa probabilidade pode ser calculada segundo

$$\begin{aligned}\mathbb{P}(\mu_{x+t}(t) \leq 0) &= \mathbb{P} \left( \mu_{x+t}(0)e^{-at} + \theta(1 - e^{-at}) + \sigma \sqrt{\frac{1 - e^{-2at}}{2a}} N \leq 0 \right) \\ &= \mathbb{P} \left( N \leq -\frac{\mu_{x+t}(0)e^{-at} + \theta(1 - e^{-at})}{\sigma \sqrt{\frac{1 - e^{-2at}}{2a}}} \right) = \Phi(\xi(\sigma, a, \theta))\end{aligned}$$

com

$$\xi(\sigma, a, \theta) = -\frac{\mu_{x+t}(0)e^{-at} + \theta(1 - e^{-at})}{\sigma \sqrt{\frac{1 - e^{-2at}}{2a}}}$$

onde  $N \sim \mathcal{N}or(0, 1)$  e  $\Phi(\cdot)$  denota a respectiva função de distribuição.

Como se observa, a função  $\xi(\cdot)$  é uma função crescente de  $\sigma$  e uma função decrescente de  $\theta$ , sendo inconclusiva a sua dependência em relação ao parâmetro  $a$ . Significa isto que a maior ou menor probabilidade de o modelo proporcionar valores indesejados para a intensidade  $\mu_{x+t}(t)$  dependerá sempre das estimativas dos parâmetros obtidas em estudos empíricos concretos.

Representemos a probabilidade de sobrevivência para um indivíduo com idade  $x + t$  no momento  $t$  por uma função do tipo afim, i.e.,

$$\begin{aligned}{}_{T-t}p_{x+t}(t) &= \mathbb{E} \left[ \exp \left( - \int_t^T \mu_{x+s}(s) ds \right) \middle| \mathcal{F}_t \right] \\ &= e^{\mathcal{A}(t, x, T) + \mathcal{B}(t, x, T) \cdot \mu_{x+t}(t)}\end{aligned}\quad (9.4.4)$$

onde as funções determinísticas  $\mathcal{A}(t, x, T)$  e  $\mathcal{B}(t, x, T)$  constituem, respectivamente, soluções das equações (9.2.16) e (9.2.17) com condições terminais  $\mathcal{A}(T, x, T) = 0$  e  $\mathcal{B}(T, x, T) = 0$ .

Para simplificar a notação, introduzimos a seguinte mudança de variável  $\tau = T - t$ , com  $\mathcal{A}(\tau) \doteq \mathcal{A}(t, x, T)$  e  $\mathcal{B}(\tau) \doteq \mathcal{B}(t, x, T)$ . Neste caso, podemos reescrever a equação (9.4.4) segundo

$${}_{T-t}p_{x+t}(t) \doteq \psi(\tau, \mu_{x+t}(t)) = e^{\mathcal{A}(\tau) + \mathcal{B}(\tau) \cdot \mu_{x+t}(t)}\quad (9.4.5)$$

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

Aplicando o Lemma de Itô e calculando as respectivas derivadas parciais, prova-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(\tau, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau)\mu_{x+t}(t) + \mathcal{B}(\tau) [a(\theta - \mu_{x+t}(t))] + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - \mu_{x+t}(t) \right\} = 0 \quad (9.4.6)$$

onde chamamos a atenção para o facto de

$$\begin{aligned} \dot{\mathcal{A}}(\tau) &= \frac{d}{dt}\mathcal{A}(\tau) = \frac{d\mathcal{A}(\tau)}{d\tau} \underbrace{\frac{d\tau}{dt}}_{=-1} = -\frac{d\mathcal{A}(\tau)}{d\tau} \\ \dot{\mathcal{B}}(\tau) &= \frac{d}{dt}\mathcal{B}(\tau) = \frac{d\mathcal{B}(\tau)}{d\tau} \underbrace{\frac{d\tau}{dt}}_{=-1} = -\frac{d\mathcal{B}(\tau)}{d\tau}. \end{aligned}$$

Se dividirmos ambos os lados da equação (9.4.6) por  $\psi(\tau, \mu_{x+t}(t))$  e arrumarmos as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) - a\mathcal{B}(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + a\theta\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem, nos termos definidos em (9.2.16)-(9.2.17), soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = -a\mathcal{B}(\tau) - 1 \quad (9.4.7)$$

$$\dot{\mathcal{A}}(\tau) = a\theta\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) \quad (9.4.8)$$

com condições terminais (ou limite)

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.9)$$

Resolvendo o problema (9.4.7)-(9.4.8)-(9.4.9), obtemos as seguintes soluções fechadas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = \frac{e^{-a\tau} - 1}{a} \quad (9.4.10)$$

$$\mathcal{A}(\tau) = \frac{[\tau + \mathcal{B}(\tau)] \left( \frac{\sigma^2}{2} - a^2\theta \right)}{a^2} - \frac{\sigma^2\mathcal{B}^2(\tau)}{4a} \quad (9.4.11)$$

Prova: Para confirmarmos a solução do problema para  $\mathcal{B}(\tau)$ , notemos em primeiro

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

lugar que  $\dot{\mathcal{B}}(\tau) = -e^{-a\tau}$ . Substituindo na respectiva EDO em (9.4.7), obtemos

$$-\dot{\mathcal{B}}(\tau) - a\mathcal{B}(\tau) = e^{-a\tau} - a \left[ \frac{e^{-a\tau} - 1}{a} \right] = 1$$

Para derivar a expressão de  $\mathcal{A}(\tau)$ , resta-nos integrar a respectiva EDO em (9.4.8) no intervalo entre 0 e  $\tau$ , i.e.,

$$\int_0^\tau \left( -\dot{\mathcal{A}}(s) + a\theta\mathcal{B}(s) + \frac{\sigma^2}{2}\mathcal{B}^2(s) \right) ds = 0$$

ou, de forma equivalente,

$$\mathcal{A}(\tau) - \underbrace{\mathcal{A}(0)}_{=0} = a\theta \int_0^\tau \mathcal{B}(s)ds + \frac{\sigma^2}{2} \int_0^\tau \mathcal{B}^2(s)ds$$

Se substituirmos a solução derivada para  $\mathcal{B}(\tau)$  e resolvermos o integral, obtemos a seguinte expressão para  $\mathcal{A}(\tau)$

$$\begin{aligned} \mathcal{A}(\tau) &= a\theta \int_0^\tau \left( \frac{e^{-as} - 1}{a} \right) ds + \frac{\sigma^2}{2} \int_0^\tau \left( \frac{e^{-as} - 1}{a} \right)^2 ds \\ &= \theta \int_0^\tau (e^{-as} - 1) ds + \frac{\sigma^2}{2a^2} \int_0^\tau (1 - 2e^{-as} + e^{-2as}) ds \\ &= \theta \left[ - \left( \frac{e^{-a\tau} - 1}{a} \right) - \tau \right] + \frac{\sigma^2}{2a^2} \left[ \tau + 2 \left( \frac{e^{-a\tau} - 1}{a} \right) + \frac{1 - e^{-2a\tau}}{2a} \right] \\ &= -\theta [\mathcal{B}(\tau) + \tau] + \frac{\sigma^2}{2a^2} \left\{ \tau + 2\mathcal{B}(\tau) + \frac{1 - e^{-2a\tau}}{2a} \right\} \\ &= \frac{-a^2\theta\mathcal{B}(\tau) - a^2\theta\tau + \frac{\sigma^2}{2}\tau + \frac{\sigma^2}{2}\mathcal{B}(\tau)}{a^2} + \frac{\sigma^2}{2a^2}\mathcal{B}(\tau) + \frac{\sigma^2}{4a} \left( \frac{1 - e^{-2a\tau}}{a^2} \right) \\ &= \frac{[\tau + \mathcal{B}(\tau)] \left( \frac{\sigma^2}{2} - a^2\theta \right)}{a^2} + \frac{\sigma^2}{4a} \left( \frac{2(e^{-a\tau} - 1)}{a^2} + \frac{(1 - e^{-2a\tau})}{a^2} \right) \\ &= \frac{[\tau + \mathcal{B}(\tau)] \left( \frac{\sigma^2}{2} - a^2\theta \right)}{a^2} + \frac{\sigma^2}{4a} \left( \frac{-1 + 2e^{-a\tau} - e^{-2a\tau}}{a^2} \right) \\ &= \frac{[\tau + \mathcal{B}(\tau)] \left( \frac{\sigma^2}{2} - a^2\theta \right)}{a^2} - \frac{\sigma^2\mathcal{B}^2(\tau)}{4a} \end{aligned}$$

■

### 9.4.2 Modelo de Vasicek com incerteza ditada exclusivamente por saltos

Nesta secção, consideramos uma variante do modelo de Vasicek em que admitimos que o coeficiente de difusão (ou de volatilidade) é nulo e que a incerteza é determinada exclusivamente por saltos (*jumps*) aleatórios descontínuos, caracterizados por uma determinada distribuição de probabilidade. Formalmente, admitimos que a intensidade  $\mu_{x+t}(t)$  é governada pela seguinte equação diferencial estocástica

$$d\mu_{x+t}(t) = a(\theta - \mu_{x+t}(t))dt + dJ(t), \quad \mu_{x+t}(0) = \bar{\mu}_x \quad (9.4.12)$$

com

$$J(t) = \sum_{i=1}^{N_t} \varepsilon_i, \quad (9.4.13)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\theta > 0$ . O termo  $J(t)$  designa um processo de Poisson composto com intensidade constante  $\eta \geq 0$ , enquanto  $\{\varepsilon_i : i = 1, \dots, \infty\}$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.). Na sequência dos resultados de Kou (2002), Kou e Wang (2003), Chen e Kou (2005), entre outros, supomos que a dimensão dos saltos (*jump sizes*) segue uma distribuição assimétrica duplamente exponencial com densidade

$$f(z) = \pi_1 \left( \frac{1}{v_1} \right) e^{-\frac{z}{v_1}} \mathbb{I}_{\{z \geq 0\}} + \pi_2 \left( \frac{1}{v_2} \right) e^{\frac{z}{v_2}} \mathbb{I}_{\{z < 0\}} \quad (9.4.14)$$

onde  $\pi_1, \pi_2 \geq 0$  ( $\pi_1 + \pi_2 = 1$ ), representam as probabilidades de saltos positivos (com dimensão média  $v_1 > 0$ ) e negativos (com dimensão média  $v_2 > 0$ ), respectivamente. No caso particular em que consideramos apenas saltos negativos, i.e., no caso em que estamos apenas interessados nos efeitos do aumento da longevidade (e.g., Biffis (2005)),  $\pi_1 = 0$ . No caso particular em que  $\eta = 0$ , o modelo torna-se determinístico. No caso em que  $v_1 = v_2$  e  $\pi_1 = \pi_2 = \frac{1}{2}$  obtemos a chamada “primeira lei de Laplace”.

As investigações realizadas até ao momento considerando equações de difusão com saltos para mortalidade (e.g., Biffis (2005)), concentram-se exclusivamente na análise do impacto de saltos negativos na intensidade  $\mu_{x+t}(t)$ , ou seja, na importância do risco de longevidade. Não negligenciando a importância e o interesse deste risco em aplicações actuariais, a nossa opinião é a de que um processo estocástico mais rico deve igualmente contemplar a possibilidade de ocorrerem choques negativos (aumentos) na intensidade  $\mu_{x+t}(t)$  associados, por exemplo, a catástrofes naturais ou a novas epidemias (e.g., gripe das aves).

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

A opção pela distribuição duplamente exponencial em detrimento, por exemplo, da distribuição Normal ou da distribuição uniforme, permite-nos captar adequadamente este fenómeno, ao mesmo tempo que é preservada a intuição dos parâmetros e a flexibilidade analítica do modelo, indispensável à derivação de soluções analíticas fechadas. A distribuição exponencial é habitualmente adoptada para modelar processos de Poisson, em que uma variável transita de estado segundo uma probabilidade por unidade de tempo constante.

A distribuição duplamente exponencial apresenta ainda duas características interessantes no contexto da modelação da mortalidade. Em primeiro lugar, a distribuição é leptocúrtica, ou seja, regista com maior frequência valores em torno da média (nula no caso simétrico) do que a distribuição Normal, correspondendo à evidência empírica de que os saltos sobre a mortalidade são, maioritariamente, de reduzida dimensão. Em segundo, a distribuição apresenta caudas pesadas correspondentes, neste caso, a eventos extremos.

Da aplicação da transformada de Laplace resulta, neste caso

$$\begin{aligned}
 \mathcal{L}\{f(z)\}(c) &= \int_{-\infty}^{+\infty} e^{cz} f(z) dz \\
 &= \int_{-\infty}^{+\infty} e^{cz} \left\{ \pi_1 \left( \frac{1}{v_1} \right) e^{-\frac{z}{v_1}} \mathbb{I}_{\{z \geq 0\}} + \pi_2 \left( \frac{1}{v_2} \right) e^{\frac{z}{v_2}} \mathbb{I}_{\{z < 0\}} \right\} dz \\
 &= \pi_1 \int_0^{+\infty} e^{cz} \frac{1}{v_1} e^{-\frac{z}{v_1}} dz + \pi_2 \int_{-\infty}^0 e^{cz} \frac{1}{v_2} e^{\frac{z}{v_2}} dz \\
 &= \frac{\pi_1}{1 - v_1 c} + \frac{\pi_2}{1 + v_2 c} \tag{9.4.15}
 \end{aligned}$$

para

$$c - \frac{1}{v_1} < 0 \quad \text{e} \quad c + \frac{1}{v_2} > 0, \quad \text{i.e.,} \quad -\frac{1}{v_2} < c < \frac{1}{v_1}.$$

Consideremos agora, nos termos da definição (9.3.7), que a probabilidade de sobrevivência é representada por uma função do tipo afim (9.4.5). Demonstra-se, após alguma manipulação algébrica, que a solução do problema admite a seguinte representação de Feynman-Kac

$$\begin{aligned}
 \psi(\tau, \mu_{x+t}(t)) &\left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau) \mu_{x+t}(t) + \mathcal{B}(\tau) [a(\theta - \mu_{x+t}(t))] \right. \\
 &\left. + \eta \left( \frac{\pi_1}{1 - v_1 \mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2 \mathcal{B}(\tau)} - 1 \right) - \mu_{x+t}(t) \right\} = 0.
 \end{aligned}$$

Dividindo ambos os lados da equação por  $\psi(\tau, \mu_{x+t}(t))$  e arrumando as parcelas,

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) - a\mathcal{B}(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + a\theta\mathcal{B}(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = -a\mathcal{B}(\tau) - 1 \quad (9.4.16)$$

$$\dot{\mathcal{A}}(\tau) = a\theta\mathcal{B}(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \quad (9.4.17)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.18)$$

A solução do problema (9.4.16)-(9.4.17)-(9.4.18) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = \frac{e^{-a\tau} - 1}{a} \quad (9.4.19)$$

$$\begin{aligned} \mathcal{A}(\tau) = & \left[ -\theta + \eta \left( \frac{\pi_1 a}{a + v_1} + \frac{\pi_2 a}{a - v_2} - 1 \right) \right] \tau - \theta\mathcal{B}(\tau) \\ & + \eta \left\{ \frac{\pi_1 \ln(1 - v_1\mathcal{B}(\tau))}{a + v_1} + \frac{\pi_2 \ln(1 + v_2\mathcal{B}(\tau))}{a - v_2} \right\} \end{aligned} \quad (9.4.20)$$

definidas para

$$1 - v_1\mathcal{B}(\tau) > 0 \quad \text{e} \quad 1 + v_2\mathcal{B}(\tau) > 0$$

ou, de forma equivalente, para

$$-\frac{1}{v_2} < \mathcal{B}(\tau) < \frac{1}{v_1}$$

Seja  $\tau \geq 0$ . Como  $\mathcal{B}(\tau) \leq 0$ , basta verificar se  $\mathcal{B}(\tau) > -\frac{1}{v_2}$ , ou seja, se  $e^{-a\tau} > 1 - \frac{a}{v_2}$ , condição que se verifica caso  $a \geq v_2$ . No caso de  $a < v_2$ , a relação  $\mathcal{B}(\tau) > -\frac{1}{v_2}$  equivale a  $\tau < -\frac{1}{a} \ln \left( 1 - \frac{a}{v_2} \right)$ , pelo que a solução (9.4.20) só é válida para valores de  $\tau = T - t$  inferiores a  $-\frac{1}{a} \ln \left( 1 - \frac{a}{v_2} \right)$ .

Prova: A verificação da solução (9.4.19) é idêntica à apresentada no modelo de Vasicek. Em relação à expressão para  $\mathcal{A}(\tau)$ , o procedimento envolve novamente a integração da respectiva EDO no intervalo entre 0 e  $\tau$ , i.e.,

$$\mathcal{A}(\tau) = a\theta \int_0^\tau \mathcal{B}(s) ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 - v_1\mathcal{B}(s)} + \frac{\pi_2}{1 + v_2\mathcal{B}(s)} - 1 \right) ds$$

ou, de forma equivalente,

$$\begin{aligned}
 \mathcal{A}(\tau) &= a\theta \int_0^\tau \left( \frac{e^{-as} - 1}{a} \right) ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 - v_1 \left( \frac{e^{-as} - 1}{a} \right)} + \frac{\pi_2}{1 + v_2 \left( \frac{e^{-as} - 1}{a} \right)} - 1 \right) ds \\
 &= \theta \left[ - \left( \frac{e^{-a\tau} - 1}{a} \right) - \tau \right] + \eta \left\{ \frac{\pi_1 [a\tau - \ln(a) + \ln(a - v_1 (e^{-a\tau} - 1))]}{a + v_1} \right. \\
 &\quad \left. + \frac{\pi_2 [a\tau - \ln(a) + \ln(a + v_2 (e^{-a\tau} - 1))]}{a - v_2} - \tau \right\} \\
 &= -\theta [\mathcal{B}(\tau) + \tau] + \eta \left\{ \frac{\pi_1 [a\tau - \ln(a) + \ln(a (1 - \frac{v_1}{a} (e^{-a\tau} - 1)))]}{a + v_1} \right. \\
 &\quad \left. + \frac{\pi_2 [a\tau - \ln(a) + \ln(a (1 + \frac{v_2}{a} (e^{-a\tau} - 1)))]}{a - v_2} - \tau \right\} \\
 &= -\theta [\mathcal{B}(\tau) + \tau] + \eta \left\{ \frac{\pi_1 [a\tau + \ln(1 - v_1 \mathcal{B}(\tau))]}{a + v_1} + \frac{\pi_2 [a\tau + \ln(1 + v_2 \mathcal{B}(\tau))]}{a - v_2} - \tau \right\} \\
 &= \left[ -\theta + \eta \left( \frac{\pi_1 a}{a + v_1} + \frac{\pi_2 a}{a - v_2} - 1 \right) \right] \tau - \theta \mathcal{B}(\tau) \\
 &\quad + \eta \left\{ \frac{\pi_1 \ln(1 - v_1 \mathcal{B}(\tau))}{a + v_1} + \frac{\pi_2 \ln(1 + v_2 \mathcal{B}(\tau))}{a - v_2} \right\}
 \end{aligned}$$

■

### 9.4.3 Modelo de Cox-Ingersoll-Ross

O modelo de equilíbrio geral desenvolvido por Cox, Ingersoll e Ross (CIR) (1985) conduziu à introdução de modificações no coeficiente de difusão do modelo de Vasicek de modo a assegurar que a taxa de juro de curto prazo se mantém sempre positiva. Em consequência, e salvaguardadas as devidas adaptações, a formulação de CIR para a intensidade  $\mu_{x+t}(t)$  é definida por

$$d\mu_{x+t}(t) = a(\theta - \mu_{x+t}(t))dt + \sigma \sqrt{\mu_{x+t}(t)} dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.21)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\theta > 0$ ,  $\sigma \geq 0$  e as demais variáveis assumem o seu significado habitual. A condição  $2a\theta > \sigma^2$  é imposta de modo a garantir que  $\mu_{x+t}(t)$  é sempre positiva. As propriedades do modelo de CIR estão bem detalhadas na literatura, pelo que recordamos apenas que o valor esperado e a variância de  $\mu_x(t)$  são dadas, respectivamente, por

$$\mathbb{E} \{ \mu_{x+t}(t) | \mathcal{F}_s \} = \mu_{x+s}(s) e^{-a(t-s)} + \theta (1 - e^{-a(t-s)}) \quad (9.4.22)$$

e por

$$\text{Var} \{ \mu_{x+t}(t) | \mathcal{F}_s \} = \mu_{x+s}(s) \frac{\sigma^2}{a} \left[ e^{-a(t-s)} - e^{-2a(t-s)} \right] + \theta \frac{\sigma^2}{2a} \left( 1 - e^{-a(t-s)} \right)^2. \quad (9.4.23)$$

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(\tau, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau) \mu_{x+t}(t) + \mathcal{B}(\tau) [a(\theta - \mu_{x+t}(t))] + \frac{1}{2} \sigma^2 \mu_{x+t}(t) \mathcal{B}^2(\tau) - \mu_{x+t}(t) \right\} = 0.$$

Dividindo os termos da equação por  $\psi(\tau, \mu_{x+t}(t))$  e arrumando as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) - a\mathcal{B}(\tau) + \frac{1}{2} \sigma^2 \mathcal{B}^2(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + a\theta \mathcal{B}(\tau) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = -a\mathcal{B}(\tau) + \frac{1}{2} \sigma^2 \mathcal{B}^2(\tau) - 1 \quad (9.4.24)$$

$$\dot{\mathcal{A}}(\tau) = a\theta \mathcal{B}(\tau) \quad (9.4.25)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.26)$$

A solução do problema (9.4.24)-(9.4.25)-(9.4.26) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{A}(\tau) = \frac{a\theta}{\phi} \tau - \frac{a\theta(\phi + \xi)}{\kappa\phi\xi} \ln \left( \frac{\phi + \xi}{\phi + \xi e^{-\kappa\tau}} \right) \quad (9.4.27)$$

$$\mathcal{B}(\tau) = \frac{1 - e^{-\kappa\tau}}{\phi + \xi e^{-\kappa\tau}}, \quad \text{com} \quad \begin{cases} \kappa = \sqrt{a^2 + 2\sigma^2} \\ \phi = -\frac{(a+\kappa)}{2} \\ \xi = \frac{(a-\kappa)}{2}. \end{cases} \quad (9.4.28)$$

Refira-se, por fim, que para fazer sentido a solução (9.4.27) deve respeitar o seguinte requisito técnico

$$\frac{\phi + \xi}{\phi + \xi e^{-\kappa\tau}} > 0 \quad (9.4.29)$$

Note que  $\phi + \xi = -\kappa < 0$ . Logo a condição técnica é equivalente a  $\phi + \xi e^{-\kappa\tau} < 0$ , ou ainda, a  $e^{-\kappa\tau} > -\frac{\phi}{\xi}$ . Como  $-\phi > 0$  e  $\xi < 0$  vem  $-\frac{\phi}{\xi} < 0$ , pelo que a condição verifica-se sempre.

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

---

Prova: Consideremos em primeiro lugar a verificação da solução (9.4.28) para  $B(\tau)$ . A equação  $\dot{B}(\tau) = 0$  admite duas raízes ou pontos de equilíbrio, i.e.,

$$\dot{B}(\tau) = 0 \Leftrightarrow -aB(\tau) + \frac{1}{2}\sigma^2 B^2(\tau) - 1 = 0 \Leftrightarrow B(\tau) = \begin{cases} \frac{a + \sqrt{a^2 + 2\sigma^2}}{\sigma^2} \\ \frac{a - \sqrt{a^2 + 2\sigma^2}}{\sigma^2} \end{cases}.$$

Denotemos por  $\kappa$  a seguinte expressão

$$\kappa = \sqrt{a^2 + 2\sigma^2}$$

e consideremos a seguinte transformação de variável

$$u(\tau) = B(\tau) - \frac{\kappa + a}{\sigma^2}.$$

Substituindo na equação base, deduz-se a seguinte EDO

$$\begin{aligned} \dot{u}(\tau) &= -a \left( u(\tau) + \frac{\kappa + a}{\sigma^2} \right) + \frac{1}{2}\sigma^2 \left( u(\tau) + \frac{\kappa + a}{\sigma^2} \right)^2 - 1 \\ &= \frac{1}{2}\sigma^2 u(\tau)^2 + \kappa u(\tau) - \frac{a(\kappa + a)}{\sigma^2} - 1 + \frac{(\kappa + a)^2}{2\sigma^2} \\ &= \frac{1}{2}\sigma^2 u(\tau)^2 + \kappa u(\tau) + \frac{-2a(\kappa + a) - 2\sigma^2 + (\kappa^2 + 2\kappa a + a^2)}{2\sigma^2} \\ &= \frac{1}{2}\sigma^2 u(\tau)^2 + \kappa u(\tau) \end{aligned}$$

cuja solução vem dada por

$$u(\tau) = -\frac{2\kappa}{\sigma^2 - 2\kappa e^{-\kappa\tau} c_1},$$

onde  $c_1$  denota uma constante. Invertendo a substituição de variável, deriva-se que

$$\begin{aligned} B(\tau) &= u(\tau) + \frac{\kappa + a}{\sigma^2} = -\frac{2\kappa}{\sigma^2 - 2\kappa e^{-\kappa\tau} c_1} + \frac{\kappa + a}{\sigma^2} \\ &= \frac{-2\kappa\sigma^2 + (\kappa + a)(\sigma^2 - 2\kappa e^{-\kappa\tau} c_1)}{\sigma^2(\sigma^2 - 2\kappa e^{-\kappa\tau} c_1)}. \end{aligned}$$

Recuperando as condições limite para  $B(\tau)$ ,  $B(0) = 0$ , deduz-se que

$$c_1 = \frac{\sigma^2(a - \kappa)}{2\kappa(a + \kappa)}.$$

Substituindo a expressão de  $c_1$  em  $B(\tau)$ , obtém-se finalmente a solução desejada

$$\begin{aligned}
 B(\tau) &= \frac{-2\kappa\sigma^2 + (\kappa + a) \left( \sigma^2 - 2\kappa e^{-\kappa\tau} \left( \frac{\sigma^2(a-\kappa)}{2\kappa(a+\kappa)} \right) \right)}{\sigma^2 \left( \sigma^2 - 2\kappa e^{-\kappa\tau} \left( \frac{\sigma^2(a-\kappa)}{2\kappa(a+\kappa)} \right) \right)} \\
 &= \frac{-2\kappa\sigma^2 + \sigma^2(\kappa + a) - \sigma^2(a - \kappa)e^{-\kappa\tau}}{\sigma^2 \left( \sigma^2 - \frac{\sigma^2(a-\kappa)}{(a+\kappa)} e^{-\kappa\tau} \right)} \\
 &= \frac{\sigma^2(a + \kappa) \left[ (a - \kappa) - \frac{(a+\kappa)(a-\kappa)}{(a+\kappa)} e^{-\kappa\tau} \right]}{\sigma^2 [\sigma^2(a + \kappa) - \sigma^2(a - \kappa) e^{-\kappa\tau}]} \\
 &= \frac{(a + \kappa)(a - \kappa)(1 - e^{-\kappa\tau})}{\sigma^2 [(a + \kappa) - (a - \kappa) e^{-\kappa\tau}]} \\
 &= \frac{(a^2 - \kappa^2)(1 - e^{-\kappa\tau})}{\sigma^2 [(a + \kappa) - (a - \kappa) e^{-\kappa\tau}]} \\
 &= \frac{1 - e^{-\kappa\tau}}{-\frac{(a+\kappa)}{2} + \frac{(a-\kappa)}{2} e^{-\kappa\tau}} \\
 &= \frac{1 - e^{-\kappa\tau}}{\phi + \xi e^{-\kappa\tau}}, \quad \begin{cases} \kappa = \sqrt{a^2 + 2\sigma^2} \\ \phi = -\frac{(a+\kappa)}{2} \\ \xi = \frac{(a-\kappa)}{2} \end{cases}
 \end{aligned}$$

Para derivar a expressão analítica de  $A(\tau)$ , resta-nos substituir  $B(\tau)$  e integrar a respectiva EDO no intervalo entre  $t$  e  $T$  (ou, de forma equivalente, entre 0 e  $\tau$ ), i.e.,

$$\begin{aligned}
 A(\tau) &= a\theta \int_t^T \left( \frac{1 - e^{-\kappa(T-s)}}{\phi + \xi e^{-\kappa(T-s)}} \right) ds \\
 &= a\theta \left\{ \frac{\xi [\kappa(T-t) - \ln((\phi + \xi) e^{\kappa T}) + \ln(\phi e^{\kappa T} + \xi e^{\kappa t})]}{\kappa\phi\xi} \right. \\
 &\quad \left. + \frac{\phi [-\kappa(T-t) - \ln(\phi + \xi) + \ln(\xi + \phi e^{\kappa(T-t)})]}{\kappa\phi\xi} \right\} \\
 &= \frac{a\theta}{\kappa\phi\xi} \left\{ \xi \left[ \kappa(T-t) - \ln(\phi + \xi) - \kappa T + \ln \left( e^{\kappa T} (\phi + \xi e^{-\kappa(T-t)}) \right) \right] \right. \\
 &\quad \left. \phi \left[ -\kappa(T-t) - \ln(\phi + \xi) + \ln \left( e^{\kappa(T-t)} (\phi + \xi e^{-\kappa(T-t)}) \right) \right] \right\} \\
 &= \frac{a\theta}{\kappa\phi\xi} \left\{ \xi\kappa(T-t) - (\phi + \xi) \left[ \ln(\phi + \xi) - \ln(\phi + \xi e^{-\kappa(T-t)}) \right] \right\},
 \end{aligned}$$

donde se obtém finalmente a solução (9.4.27). ■

#### 9.4.4 Modelo de Merton

Consideremos agora um conjunto de processos que não contemplam a hipótese de reversão em relação à média. Em primeiro lugar, analisamos uma versão do modelo desenvolvido por Merton (1973) para a dinâmica das taxas de juro, assente na seguinte equação de difusão

$$d\mu_{x+t}(t) = a dt + \sigma dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.30)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\sigma \geq 0$  e  $W(t)$  denota um movimento browniano standard.

Integrando a equação (9.4.30) obtemos, para  $0 \leq s \leq t$ ,

$$\mu_{x+t}(t) = \mu_{x+s}(s) + a(t-s) + \sigma(W(t) - W(s)). \quad (9.4.31)$$

A distribuição de  $\mu_{x+t}(t)$ , condicionada à informação disponível no momento  $s$ , é normal com média e variância dadas, respectivamente, por

$$\mathbb{E} \{ \mu_{x+t}(t) | \mathcal{F}_s \} = \mu_{x+s}(s) + a(t-s) \quad (9.4.32)$$

$$\text{Var} \{ \mu_{x+t}(t) | \mathcal{F}_s \} = \sigma^2(t-s). \quad (9.4.33)$$

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(\tau, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau)\mu_{x+t}(t) + a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - \mu_{x+t}(t) \right\} = 0.$$

Dividindo ambos os lados da equação por  $\psi(\tau, \mu_{x+t}(t))$  e arrumando as parcelas, obtém-se

$$\left[ -\dot{\mathcal{B}}(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = -1 \quad (9.4.34)$$

$$\dot{\mathcal{A}}(\tau) = a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) \quad (9.4.35)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.36)$$

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

A solução do problema (9.4.34)-(9.4.35)-(9.4.36) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = -\tau \quad (9.4.37)$$

$$\mathcal{A}(\tau) = -\frac{a}{2}\tau^2 + \frac{\sigma^2}{6}\tau^3. \quad (9.4.38)$$

Prova: A verificação da solução para  $\mathcal{B}(\tau)$  é trivial. Para derivar a expressão analítica de  $\mathcal{A}(\tau)$ , resta-nos novamente substituir  $\mathcal{B}(\tau)$  e integrar a respectiva EDO no intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned} \mathcal{A}(\tau) &= a \int_0^\tau \mathcal{B}(s) ds + \frac{1}{2}\sigma^2 \int_0^\tau \mathcal{B}^2(s) ds \\ &= a \int_0^\tau (-s) ds + \frac{1}{2}\sigma^2 \int_0^\tau (-s)^2 ds \\ &= -\frac{a}{2}\tau^2 + \frac{\sigma^2}{6}\tau^3. \end{aligned}$$

■

Se substituirmos as equações (9.4.38) e (9.4.37) em (9.4.5), obtemos a seguinte expressão analítica simples para a probabilidade de sobrevivência

$${}_{T-t}p_{x+t}(t) = \exp \left\{ -\frac{a}{2}\tau^2 + \frac{\sigma^2}{6}\tau^3 - \tau\mu_{x+t}(t) \right\}.$$

#### 9.4.5 Modelo de Merton com saltos

Admitamos agora que os desvios na tendência do processo estocástico de intensidade  $\mu_{x+t}(t)$  são governados, não apenas por um coeficiente de difusão constante, mas também por saltos descontínuos. Nesse caso, a equação de difusão base do modelo de Merton dá lugar a

$$d\mu_{x+t}(t) = a dt + \sigma dW(t) + dJ(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.39)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\sigma \geq 0$  e onde  $J(t)$  designa um processo de Poisson composto, independente de  $W(t)$ , com intensidade de salto  $\eta > 0$  e dimensão dos saltos com função de densidade definida em (9.4.14).

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-

Kac

$$\psi(t, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau)\mu_{x+t}(t) + a\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) - \mu_{x+t}(t) \right\} = 0.$$

Dividindo ambos os lados da equação por  $\psi(t, \mu_{x+t}(t))$ , deduz-se que

$$\left[ -\dot{\mathcal{B}}(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + a\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = -1 \quad (9.4.40)$$

$$\dot{\mathcal{A}}(\tau) = a\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \quad (9.4.41)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.42)$$

A solução do problema (9.4.40)-(9.4.41)-(9.4.42) remete-nos, neste caso, para as seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = -\tau \quad (9.4.43)$$

$$\begin{aligned} \mathcal{A}(\tau) = & -\frac{a}{2}\tau^2 + \frac{\sigma^2}{6}\tau^3 + \eta\mathcal{B}(\tau) \\ & + \eta \left\{ \frac{\pi_1 \ln(1 - v_1\mathcal{B}(\tau))}{v_1} - \frac{\pi_2 \ln(1 + v_2\mathcal{B}(\tau))}{v_2} \right\} \end{aligned} \quad (9.4.44)$$

sendo que, no caso da solução (9.4.44), devem ainda ser respeitados os seguintes requisitos técnicos

$$1 - v_1\mathcal{B}(\tau) > 0 \quad \text{e} \quad 1 + v_2\mathcal{B}(\tau) > 0$$

ou, de forma equivalente,

$$-\frac{1}{v_2} < \mathcal{B}(\tau) < \frac{1}{v_1}. \quad (9.4.45)$$

Como  $\mathcal{B}(\tau) = -\tau$ , para confirmar este requisito basta verificar se  $-\frac{1}{v_2} < \mathcal{B}(\tau)$ , o que equivale a  $\tau < \frac{1}{v_2}$ .

Prova: A verificação da solução para  $\mathcal{B}(\tau)$  é, de novo, trivial. Para derivar a expressão analítica de  $\mathcal{A}(\tau)$ , seguimos o procedimento habitual e integramos a respectiva EDO no

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned}
 \mathcal{A}(\tau) &= a \int_0^\tau \mathcal{B}(s) ds + \frac{\sigma^2}{2} \int_0^\tau \mathcal{B}^2(s) ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 - v_1 \mathcal{B}(s)} + \frac{\pi_2}{1 + v_2 \mathcal{B}(s)} - 1 \right) ds \\
 &= a \int_0^\tau (-s) ds + \frac{\sigma^2}{2} \int_0^\tau (-s)^2 ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 + v_1 s} + \frac{\pi_2}{1 - v_2 s} - 1 \right) ds \\
 &= -\frac{a}{2} \tau^2 + \frac{\sigma^2}{6} \tau^3 + \eta \left\{ \frac{\pi_1 \ln(1 + v_1 \tau)}{v_1} - \frac{\pi_2 \ln(1 - v_2 \tau)}{v_2} - \tau \right\} \\
 &= -\frac{a}{2} \tau^2 + \frac{\sigma^2}{6} \tau^3 + \eta \left\{ \frac{\pi_1 \ln(1 - v_1 \mathcal{B}(\tau))}{v_1} - \frac{\pi_2 \ln(1 + v_2 \mathcal{B}(\tau))}{v_2} + \mathcal{B}(\tau) \right\}
 \end{aligned}$$

■

#### 9.4.6 Modelo de Dothan

Consideremos agora uma variante naïve do modelo de Dothan (1978), na qual admitimos que a dinâmica da intensidade  $\mu_{x+t}(t)$  é ditada pela seguinte equação diferencial estocástica<sup>18</sup>

$$d\mu_{x+t}(t) = \sigma \sqrt{\mu_{x+t}(t)} dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.46)$$

onde  $\bar{\mu}_x > 0$ ,  $\sigma > 0$  e as demais variáveis têm o significado habitual.

A equação (9.4.46) constitui, na prática, uma versão simplificada do modelo de CIR sem tendência. Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(t, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau) \mu_{x+t}(t) + \frac{1}{2} \sigma^2 \mu_{x+t}(t) \mathcal{B}^2(\tau) - \mu_x(t) \right\} = 0.$$

Dividindo ambos os termos por  $\psi(t, \mu_{x+t}(t))$  e arrumando as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) + \frac{1}{2} \sigma^2 \mathcal{B}^2(\tau) - 1 \right] \mu_{x+t}(t) - \dot{\mathcal{A}}(\tau) = 0,$$

<sup>18</sup>Na sua formulação original, o modelo de Dothan (1978) estipula o seguinte processo estocástico para a taxa de juro de curto prazo  $r(t)$

$$dr(t) = \sigma r(t) dW(t), \quad r(0) = r_0,$$

onde  $r_0, \sigma > 0$ .

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

---

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 \quad (9.4.47)$$

$$\dot{\mathcal{A}}(\tau) = 0 \quad (9.4.48)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.49)$$

A solução do problema (9.4.47)-(9.4.48)-(9.4.49) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{A}(\tau) = 0 \quad (9.4.50)$$

$$\mathcal{B}(\tau) = \phi \frac{(e^{\xi\tau} - 1)}{(1 + e^{\xi\tau})}, \quad \begin{cases} \phi = \frac{\sqrt{2}}{\sigma} \\ \xi = -\sigma\sqrt{2} \end{cases} \quad (9.4.51)$$

Prova: Verifiquemos a solução (9.4.51) para  $\mathcal{B}(\tau)$ . A equação  $\dot{\mathcal{B}}(\tau) = 0$  admite duas raízes, i.e.,

$$\dot{\mathcal{B}}(\tau) = 0 \Leftrightarrow \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 = 0 \Leftrightarrow \mathcal{B}(\tau) = \begin{cases} \frac{\sqrt{2}}{\sigma} \\ -\frac{\sqrt{2}}{\sigma} \end{cases}$$

Considere-se a seguinte transformação de variável  $u(\tau) = \mathcal{B}(\tau) - \frac{\sqrt{2}}{\sigma}$ . Substituindo, obtemos a seguinte EDO

$$\dot{u}(\tau) = \frac{1}{2}\sigma^2 \left( u(\tau) + \frac{\sqrt{2}}{\sigma} \right)^2 - 1 = \frac{1}{2}\sigma^2 u(\tau)^2 + \sqrt{2}\sigma u(\tau),$$

cujas soluções vem dada por

$$u(\tau) = -\frac{4}{\sigma\sqrt{2} - 4e^{(-\tau\sigma\sqrt{2})}c_1},$$

onde  $c_1$  é uma constante. Invertendo a substituição de variável, obtemos

$$\begin{aligned} \mathcal{B}(\tau) &= u(\tau) + \frac{\sqrt{2}}{\sigma} = -\frac{4}{\sigma\sqrt{2} - 4e^{(-\tau\sigma\sqrt{2})}c_1} + \frac{\sqrt{2}}{\sigma} \\ &= \frac{-4\sigma + \sqrt{2} [\sigma\sqrt{2} - 4e^{(-\tau\sigma\sqrt{2})}c_1]}{\sigma [\sigma\sqrt{2} - 4e^{(-\tau\sigma\sqrt{2})}c_1]}. \end{aligned}$$

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

---

obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) + a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 \right] \mu_{x+t}(t) - \dot{\mathcal{A}}(\tau) = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 \quad (9.4.53)$$

$$\dot{\mathcal{A}}(\tau) = 0 \quad (9.4.54)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.55)$$

A solução do problema (9.4.53)-(9.4.54)-(9.4.55) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = \frac{1 - e^{\kappa\tau}}{\alpha_0 + \alpha_1 e^{\kappa\tau}}, \quad \text{com} \quad \begin{cases} \kappa = \sqrt{a^2 + 2\sigma^2} \\ \alpha_0 = \frac{(a+\kappa)}{2} \\ \alpha_1 = \frac{(\kappa-a)}{2} \end{cases} \quad (9.4.56)$$

$$\mathcal{A}(\tau) = 0. \quad (9.4.57)$$

Prova: Consideremos em primeiro lugar a verificação da solução (9.4.56) para  $\mathcal{B}(\tau)$ . A equação  $\dot{\mathcal{B}}(\tau) = 0$  admite duas raízes, i.e.,

$$\dot{\mathcal{B}}(\tau) = 0 \Leftrightarrow a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 = 0 \Leftrightarrow \mathcal{B}(\tau) = \begin{cases} \frac{-a - \sqrt{a^2 + 2\sigma^2}}{\sigma^2} \\ \frac{-a + \sqrt{a^2 + 2\sigma^2}}{\sigma^2} \end{cases}.$$

Denotemos por  $\kappa$  a seguinte expressão

$$\kappa = \sqrt{a^2 + 2\sigma^2}$$

e consideremos a seguinte transformação de variável

$$u(\tau) = \mathcal{B}(\tau) + \frac{\kappa + a}{\sigma^2}.$$

Substituindo na equação base, deduz-se a seguinte EDO

$$\dot{u}(\tau) = \frac{1}{2}\sigma^2 u(\tau)^2 - \kappa u(\tau)$$

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

cuja solução vem dada por

$$u(\tau) = \frac{2\kappa}{\sigma^2 + 2\kappa e^{\kappa\tau} c_1},$$

onde  $c_1$  denota uma constante. Invertendo a substituição de variável e recuperando as condições limite para  $\mathcal{B}(\tau)$ ,  $\mathcal{B}(0) = 0$ , deduz-se que

$$c_1 = \frac{\sigma^2 (k - a)}{2\kappa(a + \kappa)}.$$

Substituindo a equação de  $c_1$  em  $\mathcal{B}(\tau)$  obtém-se, após alguma manipulação algébrica, a seguinte expressão

$$\begin{aligned} \mathcal{B}(\tau) &= -\frac{(a + \kappa) [(k - a) (e^{\kappa\tau} - 1)]}{\sigma^2 [(a + \kappa) + (k - a) e^{\kappa\tau}]} \\ &= -\frac{(\kappa^2 - a^2) (e^{\kappa\tau} - 1)}{\sigma^2 [(a + \kappa) + (k - a) e^{\kappa\tau}]} \\ &= \frac{(1 - e^{\kappa\tau})}{\frac{(a + \kappa)}{2} + \frac{(\kappa - a)}{2} e^{\kappa\tau}}, \end{aligned}$$

donde se deriva finalmente a solução (9.4.56). ■

#### 9.4.8 Modelo de Feller com saltos

Consideremos agora que os desvios face à tendência do processo estocástico preconizada pelo modelo de Feller são determinados igualmente por saltos descontínuos. Nesse caso, a equação de difusão base do modelo dá lugar a

$$d\mu_{x+t}(t) = a\mu_{x+t}(t)dt + \sigma\sqrt{\mu_{x+t}(t)}dW(t) + dJ(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.58)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\sigma \geq 0$  e onde  $J(t)$  designa, de novo, um processo de Poisson composto, independente de  $W(t)$ , com intensidade de salto  $\eta > 0$  e dimensão dos saltos com função de densidade definida em (9.4.14).

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\begin{aligned} \psi(t, \mu_{x+t}(t)) &\left\{ -\dot{A}(\tau) - \dot{B}(\tau)\mu_x(t) + a\mu_{x+t}(t)\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mu_{x+t}(t)\mathcal{B}^2(\tau) \right. \\ &\left. + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) - \mu_{x+t}(t) \right\} = 0. \end{aligned}$$

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

Dividindo ambos os lados da equação por  $\psi(t, \mu_{x+t}(t))$  e arrumando as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) + a\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = a\mathcal{B}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) - 1 \quad (9.4.59)$$

$$\dot{\mathcal{A}}(\tau) = \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \quad (9.4.60)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.61)$$

A solução do problema (9.4.59)-(9.4.60)-(9.4.61) conduz-nos às seguintes expressões analíticas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\begin{aligned} \mathcal{A}(\tau) = & \eta\pi_1 \left\{ \frac{\alpha_0\tau}{(\alpha_0 - v_1)} + \frac{v_1(\alpha_0 + \alpha_1) [\ln(\alpha_0 + \alpha_1) - \ln(\alpha_0 - v_1 + (\alpha_1 + v_1)e^{\kappa\tau})]}{\kappa(\alpha_0 - v_1)(\alpha_1 + v_1)} \right\} \\ & + \eta\pi_2 \left\{ \frac{\alpha_0\tau}{(\alpha_0 + v_2)} + \frac{v_2(\alpha_0 + \alpha_1)}{\kappa(\alpha_1 - v_2)(\alpha_0 + v_2)} [-\ln(\alpha_0 + \alpha_1) \right. \\ & \left. + \ln(\alpha_0 + v_2 + (\alpha_1 - v_2)e^{\kappa\tau})] \right\} - \eta\tau \end{aligned} \quad (9.4.62)$$

$$\mathcal{B}(\tau) = \frac{1 - e^{\kappa\tau}}{\alpha_0 + \alpha_1 e^{\kappa\tau}}, \quad \text{com} \quad \begin{cases} \kappa = \sqrt{a^2 + 2\sigma^2} \\ \alpha_0 = \frac{a+\kappa}{2} \\ \alpha_1 = \frac{\kappa-a}{2} \end{cases} \quad (9.4.63)$$

Para fazer sentido, a solução (9.4.62) deve ainda respeitar os seguintes requisitos técnicos

$$\alpha_0 - v_1 + (\alpha_1 + v_1)e^{\kappa\tau} > 0 \quad \text{e} \quad \alpha_0 + v_2 + (\alpha_1 - v_2)e^{\kappa\tau} > 0$$

ou, de forma equivalente

$$1 - v_1\mathcal{B}(\tau) > 0 \quad \text{e} \quad 1 + v_2\mathcal{B}(\tau) > 0, \quad \text{i.e.,} \quad -\frac{1}{v_2} < \mathcal{B}(\tau) < \frac{1}{v_1}$$

Como  $\mathcal{B}(\tau) \leq 0$ , basta verificar se  $\mathcal{B}(\tau) > -\frac{1}{v_2}$ , o que equivale a  $e^{\kappa\tau}(v_2 - \alpha_1) < \alpha_0 + v_2$ . Como  $\alpha_1 < v_2$ , esta condição equivale a  $\tau < \frac{1}{\kappa} \ln \left( \frac{\alpha_0 + v_2}{v_2 - \alpha_1} \right)$ . No caso em que  $\alpha_1 \geq v_2$ , a condição é respeitada para todo  $\tau \geq 0$ .

Prova: A verificação da solução (9.4.63) para  $\mathcal{B}(\tau)$  é idêntica à apresentada para

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

o modelo de Feller sem saltos, pelo que é omitida por uma questão de redundância. Resta-nos derivar a expressão de  $\mathcal{A}(\tau)$ , bastando para tal integrar a respectiva EDO em (9.4.62) no intervalo entre 0 e  $\tau$  ou, em alternativa, entre  $t$  e  $T$

$$\begin{aligned}
 \mathcal{A}(\tau) &= \int_t^T \eta \left( \frac{\pi_1}{1 - v_1 \mathcal{B}(T-s)} + \frac{\pi_2}{1 + v_2 \mathcal{B}(T-s)} - 1 \right) ds \\
 &= \frac{\eta \pi_1}{\kappa (\alpha_0 - v_1) (\alpha_1 + v_1)} \left\{ \kappa \alpha_1 (\alpha_0 - v_1) (T-t) + v_1 (\alpha_0 + \alpha_1) \left[ \ln \left( (\alpha_0 + \alpha_1) e^{\kappa T} \right) \right. \right. \\
 &\quad \left. \left. - \ln \left( e^{\kappa t} \left( \alpha_0 - v_1 + (\alpha_1 + v_1) e^{\kappa(T-t)} \right) \right) \right] \right\} \\
 &\quad + \frac{\eta \pi_2}{\kappa (\alpha_1 - v_2) (\alpha_0 + v_2)} \left\{ \kappa \alpha_1 (\alpha_0 + v_2) (T-t) + v_2 (\alpha_0 + \alpha_1) \left[ - \ln \left( (\alpha_0 + \alpha_1) e^{\kappa T} \right) \right. \right. \\
 &\quad \left. \left. + \ln \left( e^{\kappa t} \left( \alpha_0 + v_2 + (\alpha_1 - v_2) e^{\kappa(T-t)} \right) \right) \right] \right\} - \eta (T-t) \\
 &= \eta \pi_1 \left\{ \frac{\alpha_0 \tau}{(\alpha_0 - v_1)} + \frac{v_1 (\alpha_0 + \alpha_1) \left[ \ln (\alpha_0 + \alpha_1) - \ln \left( \alpha_0 - v_1 + (\alpha_1 + v_1) e^{\kappa \tau} \right) \right]}{\kappa (\alpha_0 - v_1) (\alpha_1 + v_1)} \right\} \\
 &\quad + \eta \pi_2 \left\{ \frac{\alpha_0 \tau}{(\alpha_0 + v_2)} + \frac{v_2 (\alpha_0 + \alpha_1) \left[ - \ln (\alpha_0 + \alpha_1) + \ln \left( \alpha_0 + v_2 + (\alpha_1 - v_2) e^{\kappa \tau} \right) \right]}{\kappa (\alpha_1 - v_2) (\alpha_0 + v_2)} \right\} \\
 &\quad - \eta \tau
 \end{aligned}$$

■

#### 9.4.9 Equação de Ornstein-Uhlenbeck com tendência positiva

Na categoria dos modelos sem reversão à média, consideramos nesta secção a tradicional equação de Ornstein-Uhlenbeck, ligeiramente adaptada de modo a contemplar um *drift* positivo, segundo a qual a intensidade de mortalidade evolui segundo

$$d\mu_{x+t}(t) = a\mu_{x+t}(t)dt + \sigma dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.64)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\sigma \geq 0$ , e as demais variáveis assumem o seu significado habitual. Note-se que no caso em que  $a < 0$  e  $\theta = 0$  esta formulação é um caso particular do modelo de Vasicek.

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(t, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau) \mu_{x+t}(t) + a \mu_{x+t}(t) \mathcal{B}(\tau) + \frac{1}{2} \sigma^2 \mathcal{B}^2(\tau) - \mu_{x+t}(t) \right\} = 0.$$

Dividindo ambos os lados da equação por  $\psi(t, \mu_{x+t}(t))$  e arrumando as parcelas,

#### 9.4. Modelos do tipo afim para a intensidade de mortalidade

---

obtemos

$$\left[-\dot{\mathcal{B}}(\tau) + a\mathcal{B}(\tau) - 1\right] \mu_x(t) + \left[-\dot{\mathcal{A}}(\tau) + \frac{1}{2}\sigma^2\mathcal{B}^2(\tau)\right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = a\mathcal{B}(\tau) - 1 \quad (9.4.65)$$

$$\dot{\mathcal{A}}(\tau) = \frac{1}{2}\sigma^2\mathcal{B}^2(\tau) \quad (9.4.66)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.67)$$

Resolvendo o sistema (9.4.65)-(9.4.66)-(9.4.67), obtemos as seguintes soluções fechadas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$ , com  $\tau = (T - t)$

$$\mathcal{B}(\tau) = \frac{1 - e^{a\tau}}{a} \quad (9.4.68)$$

$$\mathcal{A}(\tau) = \frac{\sigma^2}{2a^2} [\tau + \mathcal{B}(\tau)] + \frac{\sigma^2}{4a} \mathcal{B}^2(\tau). \quad (9.4.69)$$

Prova: A verificação da solução (9.4.68) para  $\mathcal{B}(\tau)$  é semelhante à apresentada para o modelo de Vasicek, pelo que é omitida por uma questão de redundância. Em relação à expressão de  $\mathcal{A}(\tau)$ , resta-nos integrar a respectiva EDO em (9.4.66) no intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned} \mathcal{A}(\tau) &= \frac{1}{2}\sigma^2 \int_0^\tau \left(\frac{1 - e^{as}}{a}\right)^2 ds \\ &= \frac{\sigma^2}{2a^2} \int_0^\tau (1 - 2e^{as} + e^{2as}) ds \\ &= \frac{\sigma^2}{2a^2} \left[ \tau + 2\left(\frac{1 - e^{a\tau}}{a}\right) + \frac{e^{2a\tau} - 1}{2a} \right] \\ &= \frac{\sigma^2}{2a^2} \left[ \tau + 2\mathcal{B}(\tau) + \frac{e^{2a\tau} - 1}{2a} \right] \\ &= \frac{\sigma^2}{2a^2} [\tau + \mathcal{B}(\tau)] + \frac{\sigma^2}{4a} \left[ \frac{1 - 2e^{a\tau} + e^{2a\tau}}{a^2} \right] \\ &= \frac{\sigma^2}{2a^2} [\tau + \mathcal{B}(\tau)] + \frac{\sigma^2}{4a} \mathcal{B}^2(\tau) \end{aligned}$$

■

### 9.4.10 Equação de Ornstein-Uhlenbeck com tendência positiva e saltos

Consideremos agora que os desvios na tendência do processo estocástico da intensidade  $\mu_{x+t}(t)$  são ditados, quer por um coeficiente de difusão constante, quer por saltos descontínuos. Nesse caso, a equação diferencial estocástica base do modelo anterior dá lugar a

$$d\mu_{x+t}(t) = a\mu_{x+t}(t)dt + \sigma dW(t) + dJ(t), \quad \mu_{x+t}(0) = \bar{\mu}_x, \quad (9.4.70)$$

onde  $\bar{\mu}_x > 0$ ,  $a > 0$ ,  $\sigma \geq 0$  e onde  $J(t)$  designa, de novo, um processo de Poisson composto, independente de  $W(t)$ , com intensidade de salto  $\eta > 0$  e dimensão dos saltos com função de densidade definida em (9.4.14).

Representando a probabilidade de sobrevivência por uma função do tipo afim (9.4.5), demonstra-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\psi(t, \mu_{x+t}(t)) \left\{ -\dot{\mathcal{A}}(\tau) - \dot{\mathcal{B}}(\tau)\mu_{x+t}(t) + a\mu_{x+t}(t)\mathcal{B}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) - \mu_{x+t}(t) \right\} = 0.$$

Dividindo ambos os lados da equação por  $\psi(t, \mu_x(t))$  e arrumando as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}(\tau) + a\mathcal{B}(\tau) - 1 \right] \mu_{x+t}(t) + \left[ -\dot{\mathcal{A}}(\tau) + \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \right] = 0,$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}(\tau)$  constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}(\tau) = a\mathcal{B}(\tau) - 1 \quad (9.4.71)$$

$$\dot{\mathcal{A}}(\tau) = \frac{\sigma^2}{2}\mathcal{B}^2(\tau) + \eta \left( \frac{\pi_1}{1 - v_1\mathcal{B}(\tau)} + \frac{\pi_2}{1 + v_2\mathcal{B}(\tau)} - 1 \right) \quad (9.4.72)$$

com condições terminais

$$\mathcal{B}(0) = 0, \quad \mathcal{A}(0) = 0. \quad (9.4.73)$$

Resolvendo o sistema (9.4.71)-(9.4.72)-(9.4.73), obtemos as seguintes soluções fechadas para  $\mathcal{B}(\tau)$

$$\mathcal{B}(\tau) = \frac{1 - e^{a\tau}}{a} \quad (9.4.74)$$

e para  $\mathcal{A}(\tau)$

$$\begin{aligned} \mathcal{A}(\tau) &= \left[ \frac{\sigma^2}{2a^2} + \eta \left( \frac{\pi_1 a}{a - v_1} + \frac{\pi_2 a}{a + v_2} - 1 \right) \right] \tau \\ &\quad + \left[ \frac{\sigma^2}{2a^2} \left( 1 + \frac{a\mathcal{B}(\tau)}{2} \right) \right] \mathcal{B}(\tau) - \eta \left\{ \frac{\pi_1 \ln(1 - v_1 \mathcal{B}(\tau))}{a - v_1} + \frac{\pi_2 \ln(1 + v_2 \mathcal{B}(\tau))}{a + v_2} \right\} \end{aligned} \quad (9.4.75)$$

com  $\tau = (T - t)$ . No caso da solução (9.4.75), devem ainda ser respeitados os seguintes requisitos técnicos

$$-\frac{1}{v_2} < \mathcal{B}(\tau) < \frac{1}{v_1}. \quad (9.4.76)$$

Como  $\mathcal{B}(\tau) \leq 0$ , basta verificar se  $\mathcal{B}(\tau) > -\frac{1}{v_2}$  o que equivale a  $\tau < \frac{1}{a} \ln \left( 1 + \frac{a}{v_2} \right)$ .

Prova: A verificação da solução (9.4.74) para  $\mathcal{B}(\tau)$  é idêntica à apresentada para o modelo de Vasicek, sendo por isso omitida. Em relação à expressão de  $\mathcal{A}(\tau)$ , ela é derivável integrando a respectiva EDO (9.4.72) no intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned} \mathcal{A}(\tau) &= \int_0^\tau \frac{\sigma^2}{2} \mathcal{B}^2(s) ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 - v_1 \mathcal{B}(s)} + \frac{\pi_2}{1 + v_2 \mathcal{B}(s)} - 1 \right) ds \\ &= \frac{\sigma^2}{2} \int_0^\tau \left( \frac{1 - e^{as}}{a} \right)^2 ds + \eta \int_0^\tau \left( \frac{\pi_1}{1 - v_1 \left( \frac{1 - e^{as}}{a} \right)} + \frac{\pi_2}{1 + v_2 \left( \frac{1 - e^{as}}{a} \right)} - 1 \right) ds \\ &= \frac{\sigma^2}{2a^2} \left[ \tau + 2 \left( \frac{1 - e^{a\tau}}{a} \right) + \frac{e^{2a\tau} - 1}{2a} \right] + \eta \left\{ \frac{\pi_1}{a - v_1} [a\tau + \ln(a) \right. \\ &\quad \left. - \ln(a - v_1(1 - e^{a\tau}))] + \frac{\pi_2 [a\tau + \ln(a) - \ln(a + v_2(1 - e^{a\tau}))]}{a + v_2} - \tau \right\} \\ &= \frac{\sigma^2}{2a^2} [\tau + \mathcal{B}(\tau)] + \frac{\sigma^2}{4a} \left[ \frac{1 - 2e^{a\tau} + e^{2a\tau}}{a^2} \right] + \eta \left\{ \frac{\pi_1}{a - v_1} [a\tau + \ln(a) \right. \\ &\quad \left. - \ln \left( a \left( 1 - \frac{v_1}{a} (1 - e^{a\tau}) \right) \right) \right] + \frac{\pi_2 [a\tau + \ln(a) - \ln \left( a \left( 1 + \frac{v_2}{a} (1 - e^{a\tau}) \right) \right)]}{a + v_2} - \tau \right\} \\ &= \frac{\sigma^2}{2a^2} [\tau + \mathcal{B}(\tau)] + \frac{\sigma^2}{4a} \mathcal{B}^2(\tau) + \eta \left\{ \frac{\pi_1 [a\tau - \ln(1 - v_1 \mathcal{B}(\tau))]}{a - v_1} \right. \\ &\quad \left. + \frac{\pi_2 [a\tau - \ln(1 + v_2 \mathcal{B}(\tau))]}{a + v_2} - \tau \right\} \\ &= \left[ \frac{\sigma^2}{2a^2} + \eta \left( \frac{\pi_1 a}{a - v_1} + \frac{\pi_2 a}{a + v_2} - 1 \right) \right] \tau + \left[ \frac{\sigma^2}{2a^2} \left( 1 + \frac{a\mathcal{B}(\tau)}{2} \right) \right] \mathcal{B}(\tau) \\ &\quad - \eta \left\{ \frac{\pi_1 \ln(1 - v_1 \mathcal{B}(\tau))}{a - v_1} + \frac{\pi_2 \ln(1 + v_2 \mathcal{B}(\tau))}{a + v_2} \right\} \end{aligned}$$

■

## 9.5 Modelos do tipo afim com factores latentes

Na secção anterior analisámos um conjunto de modelos com estrutura do tipo afim para a intensidade de mortalidade nos quais, verificadas determinadas condições, foi possível derivar soluções fechadas para a probabilidade de sobrevivência.

A esmagadora maioria dos modelos desenvolvidos até ao momento no campo da mortalidade estocástica apresenta três limitações fundamentais. Em primeiro lugar, a modelação de  $\mu_x(t)$  mediante (9.3.4) envolve um conjunto de hipóteses simplificadoras, entre as quais o pressuposto de que os choques são todos do mesmo tipo e afectam todas as idades/gerações que compõem o arco da vida humana de igual modo. Neste sentido, uma formulação mais realista deve ser capaz de reconhecer, não apenas que as intensidades  $\mu_x(t)$  são influenciadas por múltiplos factores de risco, mas também que esses mesmos factores exercem, potencialmente, efeitos diferenciados sobre a evolução de  $\mu_x(t)$  nas várias idades/gerações.

Em segundo lugar, os modelos desenvolvidos por Milevsky e Promislow (2001), Dahl (2004), Biffis (2005), entre outros, examinam a evolução da intensidade de uma única idade/geração de cada vez. Uma formulação mais geral do problema deve permitir uma análise conjunta da dinâmica da intensidade  $\mu_x(t)$  em todas as idades que integram o arco da vida humana.

Em terceiro, os modelos desenvolvidos, quer num contexto dinâmico mas determinístico (e.g., Lee e Carter (1992), Renshaw *et al.* (1996), Brouhns *et al.* (2002a), Renshaw e Haberman (2003c,d)), quer no contexto estocástico (e.g., Milevsky e Promislow (2001), Dahl (2004), Cairns *et al.* (2006a)), ignoram a importância dos saltos (*jumps*) na evolução da mortalidade. Em nossa opinião, a ocorrência de saltos na mortalidade não deve ser desvalorizada na medida em que uma boa parte da racionalidade económica subjacente à compra e venda de contratos de seguro do ramo vida, às operações de titularização ou, até mesmo, à tomada de posições em derivados de mortalidade, corresponde a estratégias de cobertura (*hedging*) ou aumento da exposição aos riscos de mortalidade catastrófica, i.e., aos riscos de mortalidade normalmente associados a saltos inesperados.

Em face destes argumentos, desenvolvemos nesta secção uma abordagem onde consideramos a evolução de  $\mu_x(t)$  em todas as idades que integram o arco da vida humana de forma simultânea. O princípio subjacente é o de que a intensidade  $\mu_x(t)$  pode ser representada por uma função do tipo afim de um conjunto de factores latentes, cujo comportamento dinâmico é determinado por equações de difusão do tipo (9.3.4).

Esta solução é inspirada nos modelos do tipo afim com factores latentes propostos na literatura sobre taxas de juro e risco de crédito e generaliza os modelos de Milevsky

e Promislow (2001), Yang (2001), Schräger (2006) e Ballotta e Haberman (2006).

Ao contrário do que acontece, por exemplo, nos métodos de Lee-Carter e de Poisson, onde a preocupação se centra fundamentalmente na análise das propriedades das séries temporais da mortalidade e na sua projecção (extrapolação), neste caso procuramos desenvolver uma formulação teórica que fundamente adequadamente o *pricing* dos produtos e as operações de cobertura de risco.

A opção por esta metodologia apresenta algumas vantagens importantes. Em primeiro lugar, ela catapulta para um contexto estocástico algumas das mais conhecidas relações de dependência entre a idade e a intensidade  $\mu_x(t)$  (leis de mortalidade), concedendo-lhes uma riqueza teórica impensável no passado.

Em segundo, os modelos assumem uma natureza multivariada e reconhecem, de forma explícita, que as intensidades  $\mu_x(t)$  são influenciadas por múltiplos factores de risco com impactos diferenciados sobre o comportamento do fenómeno nas várias idades. A flexibilidade dos modelos permite-nos inclusive contemplar uma estrutura de correlação entre os distintos factores de risco. Em terceiro, a opção por modelos do tipo afim preserva a riqueza e a flexibilidade analítica demonstrada na secção anterior e permite, nalguns casos, a derivação de soluções fechadas para as probabilidades de morte e sobrevivência. Em quarto, um dos principais atractivos desta abordagem está relacionado com o facto dos modelos capturarem, de forma clara, os principais tipos de risco de mortalidade com que se confrontam os gestores de carteira: risco de longevidade (i.e., de desvios sistemáticos face à mortalidade projectada), risco de processo ou volatilidade (i.e. discrepâncias aleatórias entre a mortalidade projectada e observada), risco de base (i.e., carteira vs população) e, não menos importante, risco de mortalidade catastrófica, menosprezado por exemplo por Schräger (2006).

Por fim, a abordagem aqui descrita é compatível com uma interpretação clara dos factores de risco e é consistente com a aplicação dos tradicionais modelos de avaliação de derivados financeiros no *pricing* e *hedging* dos riscos de longevidade. Se assumirmos que os factores de risco demográficos e financeiros são independentes, podemos facilmente combinar esta metodologia com processos estocásticos para a estrutura temporal de taxas de juro visando a avaliação de, por exemplo, opções incorporadas em contratos de seguros.

### 9.5.1 Formulação teórica

Considere-se o seguinte espaço de probabilidade filtrado  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , onde a filtração  $\mathbb{F}$  satisfaz as condições habituais e  $\mathbb{P}$  denota uma medida de probabilidade. Admitamos que

## 9.5. Modelos do tipo afim com factores latentes

a intensidade de mortalidade de um indivíduo com idade  $x + t$  no momento  $t$ ,  $\mu_{x+t}(t)$ , é representada pela seguinte equação paramétrica geral

$$\mu_{x+t}(t) = g_0(x, t) + \sum_{j=1}^M g_j(x, t) X_j(t), \quad (9.5.1)$$

onde  $g_j : x \rightarrow \mathbb{R}_+$  é uma função real qualquer, potencialmente dependente da idade  $x$  do indivíduo, e  $X_j(t)$  são factores latentes que determinam a trajectória da intensidade. A dinâmica multidimensional dos  $M$  factores é ditada pela equação diferencial estocástica

$$dX(t) = \delta(\theta - X(t))dt + \Sigma\sqrt{V_t}dW_t^{\mathbb{P}} + dJ_t, \quad X(0) = \bar{X}, \quad (9.5.2)$$

onde  $W_t^{\mathbb{P}}$  é um movimento browniano  $M$ -dimensional,  $J_t$  denota uma componente de salto,  $\delta$  e  $\Sigma$  são matrizes de dimensão  $M \times M$  e  $\theta$  é um vector de dimensão  $M$ . A matriz  $V_t$  é uma matriz diagonal contendo os coeficientes de difusão dos factores na diagonal, i.e.,

$$V_{t,(jj)} = \alpha_j + \beta_j' X(t), \quad j = 1, 2, \dots, M \quad (9.5.3)$$

onde os  $\beta_j$  são vectores de dimensão  $M$ .

Como se observa, a intensidade de mortalidade é uma função do tipo afim dos factores latentes. Acresce que a tendência do processo, a matriz de variâncias-covariâncias e a intensidade de salto apresentam, também elas, uma dependência do tipo afim em relação aos factores latentes. Ao mencionarmos explicitamente os valores iniciais dos factores,  $X(0)$ , podemos condicionar o modelo de modo a obter trajectórias plausíveis.

Ao contrário de formulações anteriores, que consideram uma única idade de cada vez, a equação (9.5.1) permite modelar a intensidade  $\mu_{x+t}(t)$  em todas as idades de forma simultânea. Por analogia com a formulação de Duffie e Kan (1996), admitimos que a intensidade  $\mu_{x+t}(t)$  pode ser representada por uma função (exponencial) afim dos factores latentes.

Formalmente, considere-se o modelo (9.5.1)-(9.5.2) e definamos por  $g(x + s)$  o vector

$$g(x + s) = [g_1(x + s), \dots, g_M(x + s)]'. \quad (9.5.4)$$

Neste caso, a probabilidade de sobrevivência de um indivíduo com idade  $x + t$  no horizonte  $(T - t)$  é dada por

$${}_{T-t}p_{x+t}(t) = \mathbb{E} \left[ \exp \left( - \int_t^T \mu_{x+s}(s) ds \right) \middle| \mathcal{F}_t \right].$$

## 9.5. Modelos do tipo afim com factores latentes

Com base no teorema de Feynman-Kac, sabemos que  $T-t p_{x+t}(t) \doteq \psi(t, X_t)$  é solução da seguinte equação diferencial parcial (abreviando a notação de  $X_j(t)$  para  $X_{tj}$  e  $\psi(t, X_t)$  para  $\psi$ )

$$\begin{aligned} \frac{d\psi}{dt} + (\theta'\delta' - X_t'\delta') \frac{d\psi}{dX_t} + \frac{1}{2} \sum_{k,j=1}^M (\Sigma V_t \Sigma')_{kj} \frac{d^2\psi}{dX_{tk}dX_{tj}} \\ + \sum_{h=1}^m \eta^h(X_t, t) \int_{\mathbb{R}^n} [\psi(X_t + z, t) - \psi(X_t, t)] d\nu_t^h(z) - [g_0(x+t) + X_t'g(x+t)] \psi = 0. \end{aligned}$$

Em  $J_t = \sum_{h=1}^m J_t^h$  o tipo de salto  $h$  tem uma função de distribuição  $\nu_t^h$  no momento  $t$ , dependendo apenas de  $t$ , e uma intensidade estocástica de salto  $\{\eta^h(t, X_t) : t \geq 0\}$  para  $h \in \{1, \dots, m\}$ , onde  $\eta^h$  é definido por  $\eta^h(t, X) = \eta_0^h(t) + \eta_1^h(t) \cdot X$ . A distribuição da dimensão dos saltos  $\nu_t^h$  do  $h^{\text{ésimo}}$  processo de salto é determinada pela sua Transformada de Laplace

$$\zeta^h(t, c) = \int_{\mathbb{R}^n} e^{c \cdot z} d\nu_t^h(z) \quad (9.5.5)$$

definida em  $t \in [0, \infty)$ ,  $c \in \mathbb{C}^n$  e de modo a que o integral seja finito.

Representemos  $\psi(t, X_t)$  por uma função do tipo afim

$$T-t p_{x+t}(t) = \exp \{ \mathcal{A}(x, t, T) + \mathcal{B}(x, t, T) X_t \}. \quad (9.5.6)$$

Substituindo na equação acima, obtemos

$$\begin{aligned} \left[ \dot{\mathcal{A}}_t + \dot{\mathcal{B}}_t' X_t + (\theta'\delta' - X_t'\delta') \mathcal{B}_t + \frac{1}{2} \sum_{k,j=1}^M \sum_{i=1}^M \{ \Sigma_{ki} (\alpha_i + \beta_i' X_t) \Sigma_{ji} \mathcal{B}_{tk} \mathcal{B}_{tj} \} \right. \\ \left. + \sum_{h=1}^m (\eta_0^h + \eta_1^h X_t) [\zeta^h(t, \mathcal{B}_t) - 1] - g_0(x+t) - X_t'g(x+t) \right] \psi = 0. \end{aligned} \quad (9.5.7)$$

Se notarmos que

$$\sum_{k,j=1}^M \sum_{i=1}^M \{ \Sigma_{ki} (\alpha_i + \beta_i' X_t) \Sigma_{ji} \mathcal{B}_{tk} \mathcal{B}_{tj} \} = \sum_{i=1}^M \left\{ (\alpha_i + \beta_i' X_t) \left( \sum_{k=1}^M \Sigma_{ki} \mathcal{B}_{tk} \right) \left( \sum_{j=1}^M \Sigma_{ji} \mathcal{B}_{tj} \right) \right\}$$

e que

$$\sum_{k=1}^M \Sigma_{ki} \mathcal{B}_{tk} = [\Sigma' \mathcal{B}(x, t, T)]_i,$$

a expressão (9.5.7) pode ser simplificada para dar lugar a

$$\begin{aligned} \dot{A}_t + \dot{B}'_t X_t + (\theta' \delta' - X'_t \delta') B_t + \frac{1}{2} \sum_{i=1}^M \left\{ [\Sigma' B_t]_i^2 (\alpha_i + \beta'_i X_t) \right\} \\ + \sum_{h=1}^m \left( \eta_0^h + \eta_1^h X_t \right) \left[ \zeta^h(t, B_t) - 1 \right] - g_0(x+t) - X'_t g(x+t) = 0, \end{aligned} \quad (9.5.8)$$

onde os coeficientes  $A_t \equiv \mathcal{A}(x, t, T)$  e  $B_t \equiv \mathcal{B}(x, t, T)$  constituem soluções do seguinte sistema de EDO

$$\dot{A}_t = -\theta' \delta' B_t - \frac{1}{2} \sum_{i=1}^M [\Sigma' B_t]_i^2 \alpha_i - \sum_{h=1}^m \eta_0^h \left[ \zeta^h(t, B_t) - 1 \right] + g_0(x+t) \quad (9.5.9)$$

$$\dot{B}_t = \delta' B_t - \frac{1}{2} \sum_{i=1}^M [\Sigma' B_t]_i^2 \beta_i - \sum_{h=1}^m \eta_1^h \left[ \zeta^h(t, B_t) - 1 \right] + g(x+t), \quad (9.5.10)$$

onde  $\dot{A}_t \equiv \frac{dA_t}{dt}$  e  $\dot{B}_t \equiv \frac{dB_t}{dt}$ .

Não obstante a sua flexibilidade e consistência como os modelos tradicionais de avaliação de activos contingentes, a formulação (9.5.1)-(9.5.2) do modelo do tipo afim com factores latentes afigura-se bastante geral. De forma a adaptá-la ao contexto actuarial, desenvolvemos nas secções seguintes parametrizações particulares, combinando um modelo demográfico tradicional (e.g., lei de Gompertz-Makeham) com equações de difusão para a dinâmica dos factores latentes.

### 9.5.2 Caso I: Modelo de Makeham

Para ilustrar a aplicação desta metodologia na derivação de soluções fechadas para a probabilidade de sobrevivência num contexto estocástico, consideramos em primeiro lugar a tradicional lei de Makeham (1867) caracterizada em detalhe no Capítulo 2. Na formulação original, a lei de Makeham estabelece que a relação de dependência entre a idade e a intensidade de mortalidade é dada pela seguinte equação determinística

$$\mu_{x+t} = X_1 + X_2 c^{x+t} \quad (9.5.11)$$

onde  $X_1 > 0$ ,  $X_2 > 0$  e  $c > 1$ .

A lei de Makeham pode ser enquadrada na formulação geral (9.5.1) do modelo com factores latentes como um caso particular notando que  $g_0(x) = 0$ ,  $g_1(x) = 1$  e  $g_2(x) = c^{x+t}$ .

## 9.5. Modelos do tipo afim com factores latentes

Se admitirmos que a equação (9.5.11) ajusta adequadamente o comportamento da mortalidade ao longo de todo o arco da vida humana, as modificações na intensidade  $\mu_{x+t}(t)$  podem ser expressas em termos de variações nos parâmetros que a caracterizam (factores latentes). Por outras palavras, neste modelo a incerteza reflecte-se no facto de desconhecermos as trajectórias futuras dos parâmetros da equação.

A escolha de uma forma funcional envolve naturalmente algum risco de modelo (ou de processo estocástico). A verdade é que, neste caso, podemos quantificar e controlar esse risco de forma sistemática, na medida em que é sempre possível escolher a função paramétrica mais apropriada, i.e., aquela que minimiza o erro de ajustamento.

Admitamos agora que todos (ou apenas alguns) os parâmetros da equação (9.5.11) seguem processos estocásticos nos termos do definido em (9.5.2). Por forma a obtermos soluções analíticas fechadas para a probabilidade de sobrevivência, restringimos num primeiro momento a nossa análise a modelos dinâmicos do tipo gaussiano para os factores latentes, i.e., supomos que os factores seguem um processo de Ornstein-Uhlenbeck multivariado com saltos. Por fim, e sem perda de generalidade, supomos que o parâmetro  $c$  é constante no tempo. O resultado é o seguinte modelo

$$\mu_{x+t}(t) = X_1(t) + X_2(t)c^{x+t}, \quad (9.5.12)$$

onde os parâmetros  $X_j$  ( $j = 1, 2$ ) têm um comportamento dinâmico ditado por

$$\begin{aligned} dX_j(t) &= a_j(\theta_j - X_j(t))dt + \sigma_j dW_{jt}^{\mathbb{P}} + dJ_j(t), \quad X_j(0) = \bar{X}_j \\ dW_{1t}^{\mathbb{P}} dW_{2t}^{\mathbb{P}} &= \rho dt, \end{aligned} \quad (9.5.13)$$

onde  $a_j > 0$ ,  $\theta_j > 0$ ,  $\sigma_j \geq 0$ ,  $W_1^{\mathbb{P}}$  e  $W_2^{\mathbb{P}}$  são movimentos brownianos correlacionados e  $J_j(t)$  designa um processo de Poisson composto, independente de  $W_t$ , com intensidade de salto  $\eta_j > 0$  e dimensão dos saltos ditada por uma distribuição assimétrica duplamente exponencial com função de densidade definida em (9.4.14).

Representemos a probabilidade de sobrevivência de um indivíduo com idade  $x+t$  no horizonte  $(T-t)$  por uma função do tipo afim

$$\begin{aligned} {}_{T-t}p_{x+t}(t) &= \mathbb{E} \left[ \exp \left( - \int_t^T \mu_x(s) ds \right) \middle| \mathcal{F}_t \right] = \psi(t, X_t) \\ &= \exp \left\{ \mathcal{A}(x, t, T) + \sum_{j=1}^2 \mathcal{B}_j(x, t, T) X_j \right\}. \end{aligned} \quad (9.5.14)$$

Introduzindo a habitual mudança de variável  $\tau = T-t$ , aplicando o Lemma de Itô e

## 9.5. Modelos do tipo afim com factores latentes

calculando as respectivas derivadas parciais, prova-se que a solução do problema admite, neste caso, a seguinte representação de Feynman-Kac

$$\psi(t, X_t) \left\{ -\dot{\mathcal{A}}(\tau) - \sum_{j=1}^2 \dot{\mathcal{B}}_j(\tau) X_j + \sum_{j=1}^2 a_j [\theta_j - X_j(t)] \mathcal{B}_j(\tau) + \frac{1}{2} \sum_{j=1}^2 \sigma_j^2 \mathcal{B}_j^2(\tau) \right. \\ \left. + \mathcal{B}_1(\tau) \mathcal{B}_2(\tau) \sigma_1 \sigma_2 \rho + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - v_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + v_{2j} \mathcal{B}_j(\tau)} - 1 \right) - (X_1 + X_2 e^{x+t}) \right\} = 0.$$

Dividindo ambos os lados da equação por  $\psi(t, X_t)$  e arrumando as parcelas, obtemos

$$\begin{aligned} & \left[ -\dot{\mathcal{B}}_1(\tau) - a_1 \mathcal{B}_1(\tau) - 1 \right] X_1 + \left[ -\dot{\mathcal{B}}_2(\tau) - a_2 \mathcal{B}_2(\tau) - e^{x+t} \right] X_2 - \dot{\mathcal{A}}(\tau) \\ & + \sum_{j=1}^2 a_j \theta_j \mathcal{B}_j(\tau) + \sum_{j=1}^2 \frac{\sigma_j^2}{2} \mathcal{B}_j^2(\tau) + \mathcal{B}_1(\tau) \mathcal{B}_2(\tau) \sigma_1 \sigma_2 \rho \\ & + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - v_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + v_{2j} \mathcal{B}_j(\tau)} - 1 \right) = 0, \end{aligned}$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}_j(\tau)$  ( $j = 1, 2$ ) constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}_1(\tau) = -a_1 \mathcal{B}_1(\tau) - 1 \quad (9.5.15)$$

$$\dot{\mathcal{B}}_2(\tau) = -a_2 \mathcal{B}_2(\tau) - e^{x+t} \quad (9.5.16)$$

$$\begin{aligned} \dot{\mathcal{A}}(\tau) = & \sum_{j=1}^2 a_j \theta_j \mathcal{B}_j(\tau) + \sum_{j=1}^2 \frac{\sigma_j^2}{2} \mathcal{B}_j^2(\tau) + \mathcal{B}_1(\tau) \mathcal{B}_2(\tau) \sigma_1 \sigma_2 \rho \\ & + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - v_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + v_{2j} \mathcal{B}_j(\tau)} - 1 \right) \end{aligned} \quad (9.5.17)$$

com condições terminais

$$\mathcal{A}(0) = 0, \quad \mathcal{B}_j(0) = 0, \quad j = 1, 2. \quad (9.5.18)$$

Admitamos, por simplificação, que  $\pi_{12} = 0$  (donde  $\pi_{22} = 1$ ), ou seja, que apenas são esperados choques negativos sobre o factor latente  $X_2$ .

Resolvendo o sistema de equações (9.5.15)-(9.5.16)-(9.5.17) e (9.5.18), obtemos as

seguintes soluções fechadas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}_j(\tau)$  ( $j = 1, 2$ )

$$\mathcal{B}_1(\tau) = \frac{e^{-a_1\tau} - 1}{a_1} \quad (9.5.19)$$

$$\mathcal{B}_2(\tau) = c^{x+t} \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right), \quad \xi = \ln c \quad (9.5.20)$$

$$\begin{aligned} \mathcal{A}(\tau) = & -\theta_1 [\mathcal{B}_1(\tau) + \tau] + \frac{a_2\theta_2}{\xi - a_2} [\tau c^{x+t} - \mathcal{B}_2(\tau)] \quad (9.5.21) \\ & + \frac{\sigma_1^2}{2a_1^2} [\tau + \mathcal{B}_1(\tau)] - \frac{\sigma_1^2}{4a_1} \mathcal{B}_1^2(\tau) \\ & + \frac{\sigma_2^2 c^{2(x+t)}}{2(a_2 - \xi)^3} \left[ (a_2 - \xi)\tau + 2e^{(\xi-a_2)\tau} - \frac{1}{2}e^{2(\xi-a_2)\tau} - \frac{3}{2} \right] \\ & + \frac{\rho\sigma_1\sigma_2 c^{x+t}}{a_1(a_2 - \eta)} \left\{ \frac{1 - e^{(\xi-a_2-a_1)\tau}}{a_1 + a_2 - \xi} + \mathcal{B}_1(\tau) + \frac{\mathcal{B}_2(\tau)}{c^{x+t}} + \tau \right\} \\ & + \eta_1 \left\{ \frac{\pi_{11} [a_1\tau + \ln(1 - v_{11}\mathcal{B}_1(\tau))]}{a_1 + v_{11}} + \frac{\pi_{21} [a_1\tau + \ln(1 + v_{21}\mathcal{B}_1(\tau))]}{a_1 - v_{21}} \right. \\ & \left. - \tau \right\} + \eta_2 \frac{\{v_{22}c^{x+t}\tau + \ln[1 + v_{22}\mathcal{B}_2(\tau)]\}}{a_2 - \xi - v_{22}c^{x+t}}. \end{aligned}$$

Prova: Para confirmarmos a solução do problema para  $\mathcal{B}_1(\tau)$ , notemos em primeiro lugar que  $\dot{\mathcal{B}}_1(\tau) = -e^{-a_1\tau}$ . Substituindo na respectiva EDO (9.5.15), obtém-se

$$-\dot{\mathcal{B}}_1(\tau) - a_1\mathcal{B}_1(\tau) = e^{-a_1\tau} - a_1 \left[ \frac{e^{-a_1\tau} - 1}{a_1} \right] = e^{-a_1\tau} + 1 - e^{-a_1\tau} = 1.$$

Para verificarmos a solução do problema para  $\mathcal{B}_2(\tau)$ , assinalamos desde logo que  $-\dot{\mathcal{B}}_2(\tau) = c^{x+t}\xi \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right) + c^{x+t}e^{(\xi-a_2)\tau}$ . Substituindo na respectiva EDO (9.5.16) obtém-se

$$\begin{aligned} & -\dot{\mathcal{B}}_2(\tau) - a_2\mathcal{B}_2(\tau) \\ = & c^{x+t}\xi \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right) + c^{x+t}e^{(\xi-a_2)\tau} - a_2c^{x+t} \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right) \\ = & -c^{x+t}(a_2 - \xi) \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right) + c^{x+t}e^{(\xi-a_2)\tau} = c^{x+t}. \end{aligned}$$

A derivação da expressão analítica de  $\mathcal{A}(\tau)$  procede da forma habitual, i.e., integrando a EDO (9.5.17) no intervalo entre 0 e  $\tau$ , como se demonstra no Anexo 9.A. ■

O modelo (9.5.12)-(9.5.13) apresenta algumas vantagens face aos modelos de factor único do tipo afim analisados na Secção 9.4. Em primeiro lugar, se negligenciarmos a importância dos saltos descontínuos em (9.5.13) e admitirmos que  $\bar{X}_j > \theta_j$ , os factores latentes apresentam uma tendência exponencial decrescente, compatível com uma trajectória descendente para as taxas de mortalidade.

Em segundo lugar, o modelo captura, de forma relativamente simples, os quatro principais tipos de risco de mortalidade com que se defrontam os gestores de carteira, a saber, risco de flutuações aleatórias, risco de longevidade, risco catastrófico e risco de base. Com efeito, o modelo está perfeitamente capacitado para isolar as características dos vários tipos de risco de mortalidade, usando para tal os parâmetros das equações de difusão dos factores latentes. A tendência geral da mortalidade e o risco de longevidade são captados pelos parâmetros  $a_j$ . O risco de flutuações aleatórias é identificado pela matriz de variâncias-covariâncias.

A direcção e a magnitude do risco de mortalidade catastrófica estão perfeitamente identificadas e correspondem às componentes de salto na equação de difusão dos factores. Por fim, o modelo é compatível com um modelo relacional do tipo sugerido por Brass (1971), onde se ligam as estimativas obtidas para a população em geral e para uma subpopulação (e.g., de pessoas seguras), proporcionando uma quantificação do risco de base (efeito selecção adversa).

Por último, não obstante a sua complexidade o modelo admite soluções analíticas fechadas e dispensa, por isso, a utilização de métodos numéricos ou de simulação para resolver as EDO. Por esta razão, ele pode revelar-se de grande utilidade para efeitos de *pricing* e de quantificação do risco de longevidade.

Em contrapartida, a utilização de uma dinâmica gaussiana para os factores latentes não exclui a possibilidade de taxas  $\mu_{x+t}(t)$  negativas, uma característica bem conhecida na literatura sobre taxas de juro. Esta possibilidade não é, no entanto, suficiente para rejeitarmos a dinâmica gaussiana na modelação da mortalidade estocástica. Desde logo porque a sua flexibilidade analítica facilita, como vimos, a aplicação em problemas de *pricing* e medição do risco de longevidade. Por outro lado, e não menos importante, os modelos apresentam uma boa capacidade de ajustamento aos dados empíricos, como teremos oportunidade de constatar no Capítulo 10.

Ainda assim, uma (das múltiplas) alternativa à especificação (9.5.13) poderia passar pela adopção de equações de Feller (com saltos) para os factores latentes, i.e.,

$$\begin{aligned} dX_j(t) &= a_j(\theta_j - X_j(t))dt + \sigma_j\sqrt{X_j(t)}dW_{jt}^P + dJ_j(t), \quad X_j(0) = \bar{X}_j(9.5.22) \\ dW_{1t}^P dW_{2t}^P &= \rho dt, \end{aligned}$$

## 9.5. Modelos do tipo afim com factores latentes

com todas as variáveis a assumirem o seu significado habitual.

Representando a probabilidade de sobrevivência por uma função do tipo afim, aplicando o Lemma de Itô e calculando as respectivas derivadas parciais, prova-se que a solução do problema admite, neste caso, a seguinte representação de Feynman-Kac

$$\begin{aligned} \psi(t, X_t) \left\{ -\dot{A}(\tau) - \sum_{j=1}^2 \dot{B}_j(\tau) X_j + \sum_{j=1}^2 a_j [\theta_j - X_j(t)] B_j(\tau) + \frac{1}{2} \sum_{j=1}^2 \sigma_j^2 B_j^2(\tau) X_j \right. \\ \left. + B_1(\tau) X_1 B_2(\tau) X_2 \sigma_1 \sigma_2 \rho + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - \nu_{1j} B_j(\tau)} + \frac{\pi_{2j}}{1 + \nu_{2j} B_j(\tau)} - 1 \right) \right. \\ \left. - (X_1 + X_2 c^{x+t}) \right\} = 0. \end{aligned}$$

Não existindo evidência empírica que suporte a correlação entre os parâmetros da equação de Makeham, admitimos que os processos de Wiener são não correlacionados, i.e.,  $\rho = 0$ .

Dividindo ambos os lados da equação por  $\psi(t, X_t)$  e arrumando as parcelas, obtemos

$$\begin{aligned} \left[ -\dot{B}_1(\tau) - a_1 B_1(\tau) + \frac{\sigma_1^2}{2} B_1^2(\tau) - 1 \right] X_1 + \left[ -\dot{B}_2(\tau) - a_2 B_2(\tau) + \frac{\sigma_2^2}{2} B_2^2(\tau) - c^{x+t} \right] X_2 \\ - \dot{A}(\tau) + \sum_{j=1}^2 a_j \theta_j B_j(\tau) + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - \nu_{1j} B_j(\tau)} + \frac{\pi_{2j}}{1 + \nu_{2j} B_j(\tau)} - 1 \right) = 0, \end{aligned}$$

onde  $A(\tau)$  e  $B_j(\tau)$  ( $j = 1, 2$ ) constituem soluções do seguinte sistema de EDO

$$\dot{B}_1(\tau) = -a_1 B_1(\tau) + \frac{1}{2} \sigma_1^2 B_1^2(\tau) - 1 \quad (9.5.23)$$

$$\dot{B}_2(\tau) = -a_2 B_2(\tau) + \frac{1}{2} \sigma_2^2 B_2^2(\tau) - c^{x+t} \quad (9.5.24)$$

$$\dot{A}(\tau) = \sum_{j=1}^2 a_j \theta_j B_j(\tau) + \sum_{j=1}^2 \eta_j \left( \frac{\pi_{1j}}{1 - \nu_{1j} B_j(\tau)} + \frac{\pi_{2j}}{1 + \nu_{2j} B_j(\tau)} - 1 \right) \quad (9.5.25)$$

e respectivas condições terminais

$$A(0) = 0, \quad B_j(0) = 0, \quad j = 1, 2. \quad (9.5.26)$$

Resolvendo o sistema de equações (9.5.23)-(9.5.24)-(9.5.25) e (9.5.26), obtemos a

seguinte solução fechada para  $\mathcal{B}_1(\tau)$

$$\mathcal{B}_1(\tau) = \frac{1 - e^{-\kappa_1 \tau}}{\phi_1 + \xi_1 e^{-\kappa_1 \tau}}, \quad \text{com} \quad \begin{cases} \kappa_1 = \sqrt{a_1^2 + 2\sigma_1^2} \\ \phi_1 = -\frac{(a_1 + \kappa_1)}{2} \\ \xi_1 = \frac{(a_1 - \kappa_1)}{2} \end{cases} \quad (9.5.27)$$

enquanto que  $\mathcal{B}_2(x, t, T)$  e  $\mathcal{A}(x, t, T)$  poderão ser obtidos por resolução numérica das EDO (9.5.24) e (9.5.25).

Como se comprova, a versão não gaussiana do modelo de Makeham não proporciona soluções analíticas fechadas para a probabilidade de sobrevivência, uma característica que dificulta a sua estimação e aplicação no contexto actuarial.

### 9.5.3 Caso II: Modelo de Siler

Nesta secção, analisamos a aplicação do modelo de Siler (1979, 1983) num contexto estocástico. Recordamos que o modelo, descrito no Capítulo 2, assenta na seguinte forma funcional para a intensidade de mortalidade

$$\mu_{x+t} = X_1 e^{-\beta_1(x+t)} + X_2 + X_3 e^{\beta_3(x+t)}, \quad (9.5.28)$$

onde todos os parâmetros são positivos.

Esta especificação permite a modelação do comportamento de  $\mu_{x+t}(t)$  nas diferentes faixas etárias. O primeiro termo, decrescente com a idade, corresponde à mortalidade nas idades jovens. O segundo termo, independente da idade, considera que existem outras influências sobre a mortalidade para além do envelhecimento, nomeadamente causas acidentais, cuja incidência é transversal a todas as idades. O terceiro termo corresponde à clássica lei de Gompertz e capta a evolução normal da mortalidade nas idades adultas (crescente com a idade).

O modelo de Siler pode ser enquadrado na formulação geral (9.5.1) como um caso particular, notando que  $g_0(x) = 0$ ,  $g_1(x) = e^{-\beta_1(x+t)}$ ,  $g_2(x) = 1$  e  $g_3(x) = e^{\beta_3(x+t)}$ .

Para introduzir incerteza no modelo, admitimos neste caso a hipótese de que os parâmetros (factores) em (9.5.28) seguem processos de Ornstein-Uhlenbeck com coeficientes constantes e sem saltos.

O resultado é o seguinte modelo

$$\mu_x(t) = X_1(t) e^{-\beta_1(x+t)} + X_2(t) + X_3(t) e^{\beta_3(x+t)}, \quad (9.5.29)$$

onde os parâmetros  $X_j$  ( $j = 1, 2, 3$ ) têm um comportamento dinâmico ditado por

$$dX_j(t) = a_j(\theta_j - X_j(t))dt + \sigma_j dW_{jt}^{\mathbb{P}}, \quad X_j(0) = \bar{X}_j \quad (9.5.30)$$

$$dW_{it}^{\mathbb{P}} dW_{jt}^{\mathbb{P}} \sim \Sigma dt = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{pmatrix} dt,$$

onde  $a_j > 0$ ,  $\theta_j > 0$ ,  $\sigma_j \geq 0$  e  $W_j^{\mathbb{P}}$  são movimentos brownianos correlacionados.

Representemos  $T-t p_{x+t}(t)$  por uma função do tipo afim, i.e., por

$$T-t p_{x+t}(t) = \exp \left\{ \mathcal{A}(x, t, T) + \sum_{j=1}^3 \mathcal{B}_j(x, t, T) X_j \right\}. \quad (9.5.31)$$

Introduzindo a habitual mudança de variável  $\tau = T - t$ , aplicando o Lemma de Itô e calculando as respectivas derivadas parciais, prova-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\begin{aligned} \psi(t, X_t) \left\{ -\dot{\mathcal{A}}(\tau) - \sum_{j=1}^3 \dot{\mathcal{B}}_j(\tau) X_j + \sum_{j=1}^3 \mathcal{B}_j(\tau) [a_j(\theta_j - X_j)] + \sum_{j=1}^3 \frac{1}{2} \sigma_j^2 \mathcal{B}_j^2(\tau) \right. \\ \left. + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} - (X_1 e^{-\beta_1(x+t)} + X_2 + X_3 e^{\beta_3(x+t)}) \right\} = 0 \end{aligned}$$

ou, de forma equivalente,

$$\begin{aligned} \left[ -\dot{\mathcal{B}}_1(\tau) - a_1 \mathcal{B}_1(\tau) - e^{-\beta_1(x+t)} \right] X_1 + \left[ -\dot{\mathcal{B}}_2(\tau) - a_2 \mathcal{B}_2(\tau) - 1 \right] X_2 \\ + \left[ -\dot{\mathcal{B}}_3(\tau) - a_3 \mathcal{B}_3(\tau) - e^{\beta_3(x+t)} \right] X_3 - \dot{\mathcal{A}}(\tau) + \sum_{j=1}^3 a_j \theta_j \mathcal{B}_j(\tau) \\ + \sum_{j=1}^3 \frac{1}{2} \sigma_j^2 \mathcal{B}_j^2(\tau) + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} = 0, \end{aligned}$$

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}_j(\tau)$  ( $j = 1, 2, 3$ ) constituem soluções do seguinte sistema de quatro EDO

$$\dot{\mathcal{B}}_1(\tau) = -a_1 \mathcal{B}_1(\tau) - e^{-\beta_1(x+t)} \quad (9.5.32)$$

$$\dot{\mathcal{B}}_2(\tau) = -a_2 \mathcal{B}_2(\tau) - 1 \quad (9.5.33)$$

## 9.5. Modelos do tipo afim com factores latentes

$$\dot{\mathcal{B}}_3(\tau) = -a_3 \mathcal{B}_3(\tau) - e^{\beta_3(x+t)} \quad (9.5.34)$$

$$\dot{\mathcal{A}}(\tau) = \sum_{j=1}^3 a_j \theta_j \mathcal{B}_j(\tau) + \sum_{j=1}^3 \frac{1}{2} \sigma_j^2 \mathcal{B}_j^2(\tau) + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} \quad (9.5.35)$$

com condições terminais

$$\mathcal{A}(0) = 0, \quad \mathcal{B}_j(0) = 0, \quad j = 1, 2, 3. \quad (9.5.36)$$

Resolvendo o sistema (9.5.32)-(9.5.35) e (9.5.36), obtemos as seguintes soluções fechadas para  $\mathcal{A}(\tau)$  e  $\mathcal{B}_j(\tau)$  ( $j = 1, 2, 3$ )

$$\mathcal{B}_1(\tau) = e^{-\beta_1(x+t)} \left( \frac{e^{-(\beta_1+a_1)\tau} - 1}{a_1 + \beta_1} \right) \quad (9.5.37)$$

$$\mathcal{B}_2(\tau) = \frac{e^{-a_2\tau} - 1}{a_2} \quad (9.5.38)$$

$$\mathcal{B}_3(\tau) = e^{\beta_3(x+t)} \left( \frac{e^{(\beta_3-a_3)\tau} - 1}{a_3 - \beta_3} \right) \quad (9.5.39)$$

$$\mathcal{A}(\tau) = \epsilon_0 \tau + \sum_{j=1}^3 \epsilon_j \mathcal{B}_j(\tau) + \epsilon_4 \quad (9.5.40)$$

onde

$$\begin{aligned} \epsilon_0 &= \left\{ -\frac{a_1 \theta_1 e^{-\beta_1(x+t)}}{a_1 + \beta_1} - \theta_2 - \frac{a_3 \theta_3 e^{\beta_3(x+t)}}{a_3 - \beta_3} + \frac{\sigma_1^2 e^{-2\beta_1(x+t)}}{2(a_1 + \beta_1)^2} + \frac{\sigma_2^2}{2a_2^2} + \frac{\sigma_3^2 e^{2\beta_3(x+t)}}{2(a_3 - \beta_3)^2} \right. \\ &\quad \left. + \frac{\sigma_1 \sigma_2 \rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1) a_2} + \frac{\sigma_1 \sigma_3 \rho_{13} e^{(\beta_3 - \beta_1)(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} + \frac{\sigma_2 \sigma_3 \rho_{23} e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \right\} \\ \epsilon_1 &= -\frac{a_1 \theta_1}{a_1 + \beta_1} + \frac{\sigma_1^2 e^{-\beta_1(x+t)}}{(a_1 + \beta_1)^2} + \frac{\sigma_1 \sigma_2 \rho_{12}}{(a_1 + \beta_1) a_2} + \frac{\sigma_1 \sigma_3 \rho_{13} e^{\beta_3(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} \\ \epsilon_2 &= -\theta_2 + \frac{\sigma_2^2}{a_2^2} + \frac{\sigma_1 \sigma_2 \rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1) a_2} + \frac{\sigma_2 \sigma_3 \rho_{23} e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \\ \epsilon_3 &= -\frac{a_3 \theta_3}{a_3 - \beta_3} + \frac{\sigma_3^2 e^{\beta_3(x+t)}}{(a_3 - \beta_3)^2} + \frac{\sigma_1 \sigma_3 \rho_{13} e^{-\beta_1(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} + \frac{\sigma_2 \sigma_3 \rho_{23}}{a_2(a_3 - \beta_3)} \end{aligned}$$

$$\begin{aligned}
 \epsilon_4 = & \frac{\sigma_1^2 e^{-2\beta_1(x+t)} (1 - e^{-2(\beta_1+a_1)\tau})}{4(a_1 + \beta_1)^3} + \frac{\sigma_2^2 (1 - e^{-2a_2\tau})}{4a_2^3} \\
 & + \frac{\sigma_3^2 e^{2\beta_3(x+t)} (1 - e^{2(\beta_3-a_3)\tau})}{4(a_3 - \beta_3)^3} + \frac{\sigma_1\sigma_2\rho_{12}e^{-\beta_1(x+t)}}{(a_1 + \beta_1)a_2} \left( \frac{1 - e^{-(\beta_1+a_1+a_2)\tau}}{a_1 + \beta_1 + a_2} \right) \\
 & + \frac{\sigma_1\sigma_3\rho_{13}e^{(\beta_3-\beta_1)(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} \left( \frac{1 - e^{[(\beta_3-a_3)-(\beta_1+a_1)]\tau}}{a_3 - \beta_3 + a_1 + \beta_1} \right) \\
 & + \frac{\sigma_2\sigma_3\rho_{23}e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \left( \frac{1 - e^{(\beta_3-a_3-a_2)\tau}}{a_3 - \beta_3 + a_2} \right)
 \end{aligned}$$

Prova: A verificação da solução do problema para  $\mathcal{B}_1(\tau)$  é simples, bastando para tal a substituição na respectiva EDO (9.5.32) e um pouco de álgebra

$$\begin{aligned}
 & -\dot{\mathcal{B}}_1(\tau) - a_1\mathcal{B}_1(\tau) \\
 = & -\beta_1\mathcal{B}_1(\tau) + e^{-\beta_1(x+t)}e^{-(\beta_1+a_1)\tau} - a_1\mathcal{B}_1(\tau) \\
 = & -(a_1 + \beta_1) \frac{e^{-\beta_1(x+t)}(e^{-(\beta_1+a_1)\tau} - 1)}{a_1 + \beta_1} + e^{-\beta_1(x+t)}e^{-(\beta_1+a_1)\tau} = e^{-\beta_1(x+t)}
 \end{aligned}$$

A verificação da solução para  $\mathcal{B}_2(\tau)$  é semelhante à apresentada para o modelo de Vasicek, pelo que é omitida por uma questão de redundância. Em relação à solução para  $\mathcal{B}_3(\tau)$ , substituindo na respectiva EDO (9.5.34) deduz-se

$$\begin{aligned}
 & -\dot{\mathcal{B}}_3(\tau) - a_3\mathcal{B}_3(\tau) \\
 = & \beta_3\mathcal{B}_3(\tau) + e^{\beta_3(x+t)}e^{(\beta_3-a_3)\tau} - a_3\mathcal{B}_3(\tau) \\
 = & -(a_3 - \beta_3) \frac{e^{\beta_3(x+t)}(e^{(\beta_3-a_3)\tau} - 1)}{a_3 - \beta_3} + e^{\beta_3(x+t)}e^{(\beta_3-a_3)\tau} = e^{\beta_3(x+t)}
 \end{aligned}$$

Por último, a expressão de  $\mathcal{A}(\tau)$  obtém-se da forma habitual integrando a respectiva EDO (9.5.35) no intervalo entre 0 e  $\tau$ , como se demonstra no Anexo 9.A. ■

### 9.5.4 Caso III: Modelo de Thiele não Gaussiano

Nesta secção, consideramos uma versão estocástica da parametrização de  $\mu_{x+t}$  proposta por Thiele (1872) para todas as idades que compõem o arco da vida humana. Esta lei, descrita em detalhe no Capítulo 2, baseia-se na seguinte função

$$\mu_{x+t} = X_1 e^{-\beta_1(x+t)} + X_2 e^{(-\frac{1}{2}\beta_2(x+t-\xi)^2)} + X_3 e^{\beta_3(x+t)}, \quad x+t \geq 0, \quad (9.5.41)$$

## 9.5. Modelos do tipo afim com factores latentes

onde  $\alpha_i, \beta_i > 0$  ( $i = 1, 2$ ) e  $\alpha_3, \beta_3, \xi \geq 0$ . Este modelo pode ser encaixado na formulação geral (9.5.1), notando que  $g_0(x) = 0$ ,  $g_1(x) = e^{-\beta_1(x+t)}$ ,  $g_2(x) = e^{(-\frac{1}{2}\beta_2(x+t-\xi)^2)}$  e  $g_3(x) = e^{\beta_3(x+t)}$ . Para garantir que a intensidade  $\mu_x(t)$  é sempre não-negativa, a nossa opção recaiu neste caso pela utilização de processos de Feller com saltos para os factores latentes  $X_j$  ( $j = 1, 2, 3$ ). O modelo resultante pode ser formulado da seguinte forma

$$\mu_{x+t}(t) = X_1(t)e^{-\beta_1(x+t)} + X_2(t)e^{(-\frac{1}{2}\beta_2(x+t-\xi)^2)} + X_3(t)e^{\beta_3(x+t)}, \quad (9.5.42)$$

onde os factores  $X_j$  ( $j = 1, 2, 3$ ) têm um comportamento dinâmico ditado por

$$\begin{aligned} dX_j(t) &= a_j(\theta_j - X_j(t))dt + \sigma_j\sqrt{X_j(t)}dW_{jt}^{\mathbf{P}} + dJ_j(t), \quad X_j(0) = \bar{X}_j \\ dW_{it}^{\mathbf{P}}dW_{jt}^{\mathbf{P}} &\sim \Sigma dt. \end{aligned} \quad (9.5.43)$$

Representemos  $T-t p_{x+t}(t)$  por uma função do tipo afim

$$T-t p_{x+t}(t) = \exp \left\{ \mathcal{A}(x, t, T) + \sum_{j=1}^3 \mathcal{B}_j(x, t, T) X_j \right\}. \quad (9.5.44)$$

Introduzindo a habitual mudança de variável  $\tau = T - t$ , aplicando o Lemma de Itô e calculando as respectivas derivadas parciais, prova-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\begin{aligned} \psi(t, X_t) \left\{ -\dot{\mathcal{A}}(\tau) - \sum_{j=1}^3 \dot{\mathcal{B}}_j(\tau) X_j + \sum_{j=1}^3 a_j [\theta_j - X_j(t)] \mathcal{B}_j(\tau) + \sum_{j=1}^3 \frac{\sigma_j^2}{2} \mathcal{B}_j^2(\tau) X_j \right. \\ \left. + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} + \sum_{j=1}^3 \eta_j \left( \frac{\pi_{1j}}{1 - \nu_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + \nu_{2j} \mathcal{B}_j(\tau)} - 1 \right) - \mu_x(t) \right\} = 0 \end{aligned}$$

ou, de forma equivalente

$$\begin{aligned} \left[ -\dot{\mathcal{B}}_1(\tau) - a_1 \mathcal{B}_1(\tau) + \frac{\sigma_1^2}{2} \mathcal{B}_1^2(\tau) - e^{-\beta_1(x+t)} \right] X_1 + \left[ -\dot{\mathcal{B}}_2(\tau) - a_2 \mathcal{B}_2(\tau) + \frac{\sigma_2^2}{2} \mathcal{B}_2^2(\tau) \right. \\ \left. - e^{-\beta_2((x+t)-\xi)^2} \right] X_2 + \left[ -\dot{\mathcal{B}}_3(\tau) - a_3 \mathcal{B}_3(\tau) + \frac{\sigma_3^2}{2} \mathcal{B}_3^2(\tau) - e^{\beta_3(x+t)} \right] X_3 + \sum_{j=1}^3 a_j \theta_j \mathcal{B}_j(\tau) \\ - \dot{\mathcal{A}}(\tau) + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} + \sum_{j=1}^3 \eta_j \left( \frac{\pi_{1j}}{1 - \nu_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + \nu_{2j} \mathcal{B}_j(\tau)} - 1 \right) = 0, \end{aligned}$$

## 9.5. Modelos do tipo afim com factores latentes

---

onde  $\mathcal{A}(\tau)$  e  $\mathcal{B}_j(\tau)$  ( $j = 1, 2, 3$ ) constituem soluções do seguinte sistema de EDO

$$\dot{\mathcal{B}}_1(\tau) = -a_1\mathcal{B}_1(\tau) + \frac{\sigma_1^2}{2}\mathcal{B}_1^2(\tau) - e^{-\beta_1(x+t)} \quad (9.5.45)$$

$$\dot{\mathcal{B}}_2(\tau) = -a_2\mathcal{B}_2(\tau) + \frac{\sigma_2^2}{2}\mathcal{B}_2^2(\tau) - e^{-\beta_2((x+t)-\xi)^2} \quad (9.5.46)$$

$$\dot{\mathcal{B}}_3(\tau) = -a_3\mathcal{B}_3(\tau) + \frac{\sigma_3^2}{2}\mathcal{B}_3^2(\tau) - e^{\beta_3(x+t)} \quad (9.5.47)$$

$$\begin{aligned} \dot{\mathcal{A}}(\tau) = & \sum_{j=1}^3 a_j \theta_j \mathcal{B}_j(\tau) + \sum_{j=1}^3 \sum_{\substack{i=1 \\ i \neq j}}^3 \mathcal{B}_j(\tau) \mathcal{B}_i(\tau) \sigma_j \sigma_i \rho_{ji} \\ & + \sum_{j=1}^3 \eta_j \left( \frac{\pi_{1j}}{1 - v_{1j} \mathcal{B}_j(\tau)} + \frac{\pi_{2j}}{1 + v_{2j} \mathcal{B}_j(\tau)} - 1 \right) \end{aligned} \quad (9.5.48)$$

com condições terminais

$$\mathcal{A}(0) = 0, \quad \mathcal{B}_j(0) = 0, \quad j = 1, 2, 3. \quad (9.5.49)$$

A resolução do sistema de equações (9.5.45)-(9.5.48) e (9.5.49) só é possível mediante a aplicação de métodos numéricos, mesmo no caso em que admitimos que os processos de Wiener são não correlacionados.

## 9.A Anexos

### Demonstração da equação (9.5.21)

A derivação da expressão analítica de  $\mathcal{A}(\tau)$  requer a integração da EDO (9.5.17) no intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned} \mathcal{A}(\tau) &= \sum_{j=1}^2 a_j \theta_j \int_0^\tau \mathcal{B}_j(s) ds + \sum_{j=1}^2 \frac{\sigma_j^2}{2} \int_0^\tau (\mathcal{B}_j(s))^2 ds \\ &\quad + \rho \sigma_1 \sigma_2 \int_0^\tau \mathcal{B}_1(s) \mathcal{B}_2(s) ds \\ &\quad + \sum_{j=1}^2 \eta_j \int_0^\tau \left( \frac{\pi_{1j}}{1 - \nu_{1j} \mathcal{B}_j(s)} + \frac{\pi_{2j}}{1 + \nu_{2j} \mathcal{B}_j(s)} - 1 \right) ds \end{aligned}$$

Substituindo as soluções analíticas de  $\mathcal{B}_1(\tau)$  e  $\mathcal{B}_2(\tau)$  derivadas em (9.5.19) e (9.5.20), obtemos

$$\begin{aligned} \mathcal{A}(\tau) &= a_1 \theta_1 \int_0^\tau \frac{e^{-a_1 s} - 1}{a_1} ds + a_2 \theta_2 \int_0^\tau c^{x+t} \frac{e^{(\xi - a_2)s} - 1}{a_2 - \xi} ds \\ &\quad + \frac{\sigma_1^2}{2} \int_0^\tau \left( \frac{e^{-a_1 s} - 1}{a_1} \right)^2 ds + \frac{\sigma_2^2}{2} \int_0^\tau \left( c^{x+t} \frac{e^{(\xi - a_2)s} - 1}{a_2 - \xi} \right)^2 ds \\ &\quad + \rho \sigma_1 \sigma_2 \int_0^\tau \left( \frac{e^{-a_1 s} - 1}{a_1} \right) \left( c^{x+t} \frac{e^{(\xi - a_2)s} - 1}{a_2 - \xi} \right) ds \\ &\quad + \eta_1 \int_0^\tau \left( \frac{\pi_{11}}{1 - \nu_{11} \left( \frac{e^{-a_1 s} - 1}{a_1} \right)} + \frac{\pi_{21}}{1 + \nu_{21} \left( \frac{e^{-a_1 s} - 1}{a_1} \right)} - 1 \right) ds \\ &\quad + \eta_2 \int_0^\tau \left( \frac{\pi_{22}}{1 + \nu_{22} \left( c^{x+t} \frac{e^{(\xi - a_2)s} - 1}{a_2 - \xi} \right)} - 1 \right) ds \end{aligned}$$

ou, de forma equivalente,

$$\begin{aligned}
 \mathcal{A}(\tau) = & \theta_1 \left[ - \left( \frac{e^{-a_1\tau} - 1}{a_1} \right) - \tau \right] + a_2\theta_2 c^{x+t} \left( \frac{1 - e^{(\xi-a_2)\tau + (\xi-a_2)\tau}}{(a_2 - \xi)^2} \right) \\
 & + \frac{\sigma_1^2}{2a_1^2} \left\{ \tau + 2 \left( \frac{e^{-a_1\tau} - 1}{a_1} \right) + \frac{1 - e^{-2a_1\tau}}{2a_1} \right\} \\
 & + \frac{\sigma_2^2}{2} \left\{ \frac{c^{2(x+t)}}{(a_2 - \xi)^2} \left[ \tau + 2 \left( \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} \right) + \frac{1 - e^{2(\xi-a_2)\tau}}{2(a_2 - \xi)} \right] \right\} \\
 & + \frac{\rho\sigma_1\sigma_2 c^{x+t}}{a_1(a_2 - \eta)} \left\{ - \frac{e^{(\xi-a_2-a_1)\tau} - 1}{a_1 + a_2 - \xi} + \frac{e^{-a_1\tau} - 1}{a_1} + \frac{e^{(\xi-a_2)\tau} - 1}{a_2 - \xi} + \tau \right\} \\
 & + \eta_1 \left\{ \frac{\pi_{11} [a_1\tau - \ln(a_1) + \ln(a_1 - v_{11}(e^{-a_1\tau} - 1))]}{a_1 + v_{11}} \right. \\
 & \left. + \frac{\pi_{21} [a_1\tau - \ln(a_1) + \ln(a_1 + v_{21}(e^{-a_1\tau} - 1))]}{a_1 - v_{21}} - \tau \right\} \\
 & + \eta_2 \frac{\left\{ v_{22}c^{x+t}\tau - \ln(a_2 - \xi) + \ln \left[ (a_2 - \xi) \left( 1 + v_{22} \frac{c^{x+t}(e^{(\xi-a_2)\tau} - 1)}{a_2 - \xi} \right) \right] \right\}}{a_2 - \xi - v_{22}c^{x+t}}
 \end{aligned}$$

Simplificando esta expressão, deduz-se finalmente a equação (9.5.21)

$$\begin{aligned}
 \mathcal{A}(\tau) = & -\theta_1 [\mathcal{B}_1(\tau) + \tau] + \frac{a_2\theta_2}{\xi - a_2} [\tau c^{x+t} - B_2(\tau)] \\
 & + \frac{\sigma_1^2}{2a_1^2} [\tau + \mathcal{B}_1(\tau)] - \frac{\sigma_1^2}{4a_1} \mathcal{B}_1^2(\tau) \\
 & + \frac{\sigma_2^2 c^{2(x+t)}}{2(a_2 - \xi)^3} \left[ (a_2 - \xi)\tau + 2e^{(\xi-a_2)\tau} - \frac{1}{2}e^{2(\xi-a_2)\tau} - \frac{3}{2} \right] \\
 & + \frac{\rho\sigma_1\sigma_2 c^{x+t}}{a_1(a_2 - \eta)} \left\{ \frac{1 - e^{(\xi-a_2-a_1)\tau}}{a_1 + a_2 - \xi} + \mathcal{B}_1(\tau) + \frac{B_2(\tau)}{c^{x+t}} + \tau \right\} \\
 & + \eta_1 \left\{ \frac{\pi_{11} [a_1\tau + \ln(1 - v_{11}\mathcal{B}_1(\tau))]}{a_1 + v_{11}} + \frac{\pi_{21} [a_1\tau + \ln(1 + v_{21}\mathcal{B}_1(\tau))]}{a_1 - v_{21}} - \tau \right\} \\
 & + \eta_2 \frac{\left\{ v_{22}c^{x+t}\tau + \ln[1 + v_{22}B_2(\tau)] \right\}}{a_2 - \xi - v_{22}c^{x+t}}
 \end{aligned}$$

**Demonstração da equação (9.5.40)**

A derivação da expressão analítica de  $\mathcal{A}(\tau)$  requer a integração da EDO (9.5.35) no intervalo entre 0 e  $\tau$ , i.e.,

$$\begin{aligned} \mathcal{A}(\tau) &= \sum_{j=1}^3 \left\{ a_j \theta_j \left( \int_0^\tau \mathcal{B}_j(s) ds \right) \right\} + \sum_{j=1}^3 \left\{ \frac{1}{2} \sigma_j^2 \left( \int_0^\tau \mathcal{B}_j(s)^2 ds \right) \right\} \\ &\quad + \sum_{j=1}^3 \sum_{\substack{i=1 \\ j \neq i}}^3 \int_0^\tau \{ \mathcal{B}_j(s) \mathcal{B}_i(s) \sigma_j \sigma_i \rho_{ji} \} ds \end{aligned}$$

Substituindo as soluções analíticas de  $\mathcal{B}_1(\tau)$ ,  $\mathcal{B}_2(\tau)$  e  $\mathcal{B}_3(\tau)$  derivadas em (9.5.37), (9.5.38) e (9.5.39), respectivamente, obtemos

$$\begin{aligned} \mathcal{A}(\tau) &= a_1 \theta_1 \int_0^\tau e^{-\beta_1(x+t)} \left( \frac{e^{-(\beta_1+a_1)s} - 1}{a_1 + \beta_1} \right) ds + a_2 \theta_2 \int_0^\tau \left( \frac{e^{-a_2s} - 1}{a_2} \right) ds \\ &\quad + a_3 \theta_3 \int_0^\tau e^{\beta_3(x+t)} \left( \frac{e^{(\beta_3-a_3)s} - 1}{a_3 - \beta_3} \right) ds \\ &\quad + \frac{\sigma_1^2}{2} \int_0^\tau \left( e^{-\beta_1(x+t)} \left( \frac{e^{-(\beta_1+a_1)s} - 1}{a_1 + \beta_1} \right) \right)^2 ds + \frac{\sigma_2^2}{2} \int_0^\tau \left( \frac{e^{-a_2s} - 1}{a_2} \right)^2 ds \\ &\quad + \frac{\sigma_3^2}{2} \int_0^\tau \left( e^{\beta_3(x+t)} \left( \frac{e^{(\beta_3-a_3)s} - 1}{a_3 - \beta_3} \right) \right)^2 ds \\ &\quad + \sigma_1 \sigma_2 \rho_{12} \int_0^\tau \left( e^{-\beta_1(x+t)} \left( \frac{e^{-(\beta_1+a_1)s} - 1}{a_1 + \beta_1} \right) \right) \left( \frac{e^{-a_2s} - 1}{a_2} \right) ds \\ &\quad + \sigma_1 \sigma_3 \rho_{13} \int_0^\tau \left( e^{-\beta_1(x+t)} \left( \frac{e^{-(\beta_1+a_1)s} - 1}{a_1 + \beta_1} \right) \right) \left( e^{\beta_3(x+t)} \left( \frac{e^{(\beta_3-a_3)s} - 1}{a_3 - \beta_3} \right) \right) ds \\ &\quad + \sigma_2 \sigma_3 \rho_{23} \int_0^\tau \left( \frac{e^{-a_2s} - 1}{a_2} \right) \left( e^{\beta_3(x+t)} \left( \frac{e^{(\beta_3-a_3)s} - 1}{a_3 - \beta_3} \right) \right) ds \end{aligned}$$

Integrando no intervalo entre 0 e  $\tau$ , obtemos

$$\begin{aligned} \mathcal{A}(\tau) &= \frac{a_1 \theta_1 e^{-\beta_1(x+t)}}{a_1 + \beta_1} \left[ -\frac{e^{-(\beta_1+a_1)\tau} - 1}{a_1 + \beta_1} - \tau \right] + \theta_2 \left[ -\frac{e^{-a_2\tau} - 1}{a_2} - \tau \right] \\ &\quad + \frac{a_3 \theta_3 e^{\beta_3(x+t)}}{a_3 - \beta_3} \left[ -\frac{e^{(\beta_3-a_3)\tau} - 1}{a_3 - \beta_3} - \tau \right] \end{aligned}$$

$$\begin{aligned}
 & + \frac{\sigma_1^2 e^{-2\beta_1(x+t)}}{2(a_1 + \beta_1)^2} \left[ -\frac{e^{-2(\beta_1+a_1)\tau} - 1}{2(a_1 + \beta_1)} + 2 \left( \frac{e^{-(\beta_1+a_1)\tau} - 1}{a_1 + \beta_1} \right) + \tau \right] \\
 & + \frac{\sigma_2^2}{2a_2^2} \left[ \tau + 2 \left( \frac{e^{-a_2\tau} - 1}{a_2} \right) + \frac{1 - e^{-2a_2\tau}}{2a_2} \right] \\
 & + \frac{\sigma_3^2 e^{2\beta_3(x+t)}}{2(a_3 - \beta_3)^2} \left[ -\frac{e^{2(\beta_3-a_3)\tau} - 1}{2(a_3 - \beta_3)} + 2 \left( \frac{e^{(\beta_3-a_3)\tau} - 1}{a_3 - \beta_3} \right) + \tau \right] \\
 & + \frac{\sigma_1 \sigma_2 \rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1) a_2} \left[ -\frac{e^{-(\beta_1+a_1+a_2)\tau} - 1}{a_1 + \beta_1 + a_2} + \frac{e^{-(\beta_1+a_1)\tau} - 1}{a_1 + \beta_1} + \tau + \frac{e^{-a_2\tau} - 1}{a_2} \right] \\
 & + \frac{\sigma_1 \sigma_3 \rho_{13} e^{(\beta_3-\beta_1)(x+t)}}{(a_1 + \beta_1) (a_3 - \beta_3)} \left[ -\frac{e^{[(\beta_3-a_3)-(\beta_1+a_1)]\tau} - 1}{a_3 - \beta_3 + a_1 + \beta_1} + \tau + \frac{e^{-(\beta_1+a_1)\tau} - 1}{a_1 + \beta_1} \right. \\
 & \left. + \frac{e^{(\beta_3-a_3)\tau} - 1}{a_3 - \beta_3} \right] + \frac{\sigma_2 \sigma_3 \rho_{23} e^{\beta_3(x+t)}}{a_2 (a_3 - \beta_3)} \left[ -\frac{e^{(\beta_3-a_3-a_2)\tau} - 1}{a_3 - \beta_3 + a_2} + \frac{e^{-a_2\tau} - 1}{a_2} \right. \\
 & \left. + \frac{e^{(\beta_3-a_3)\tau} - 1}{a_3 - \beta_3} + \tau \right]
 \end{aligned}$$

Simplificando,

$$\begin{aligned}
 \mathcal{A}(\tau) & = -\frac{a_1 \theta_1 e^{-\beta_1(x+t)}}{a_1 + \beta_1} \left[ e^{\beta_1(x+t)} \mathcal{B}_1(\tau) + \tau \right] - \theta_2 \left[ \mathcal{B}_2(\tau) + \tau \right] \\
 & - \frac{a_3 \theta_3 e^{\beta_3(x+t)}}{a_3 - \beta_3} \left[ e^{-\beta_3(x+t)} \mathcal{B}_3(\tau) + \tau \right] + \frac{\sigma_1^2 e^{-2\beta_1(x+t)}}{2(a_1 + \beta_1)^2} \left[ 2e^{\beta_1(x+t)} \mathcal{B}_1(\tau) + \tau \right. \\
 & \left. + \frac{1 - e^{-2(\beta_1+a_1)\tau}}{2(a_1 + \beta_1)} \right] + \frac{\sigma_2^2}{2a_2^2} \left[ \tau + 2\mathcal{B}_2(\tau) + \frac{1 - e^{-2a_2\tau}}{2a_2} \right] \\
 & + \frac{\sigma_3^2 e^{2\beta_3(x+t)}}{2(a_3 - \beta_3)^2} \left[ \frac{1 - e^{2(\beta_3-a_3)\tau}}{2(a_3 - \beta_3)} + 2e^{-\beta_3(x+t)} \mathcal{B}_3(\tau) + \tau \right] \\
 & + \frac{\sigma_1 \sigma_2 \rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1) a_2} \left[ \frac{1 - e^{-(\beta_1+a_1+a_2)\tau}}{a_1 + \beta_1 + a_2} + e^{\beta_1(x+t)} \mathcal{B}_1(\tau) + \mathcal{B}_2(\tau) + \tau \right] \\
 & + \frac{\sigma_1 \sigma_3 \rho_{13} e^{(\beta_3-\beta_1)(x+t)}}{(a_1 + \beta_1) (a_3 - \beta_3)} \left[ \frac{1 - e^{[(\beta_3-a_3)-(\beta_1+a_1)]\tau}}{a_3 - \beta_3 + a_1 + \beta_1} + e^{\beta_1(x+t)} \mathcal{B}_1(\tau) + \tau \right. \\
 & \left. + e^{-\beta_3(x+t)} \mathcal{B}_3(\tau) \right] + \frac{\sigma_2 \sigma_3 \rho_{23} e^{\beta_3(x+t)}}{a_2 (a_3 - \beta_3)} \left[ \frac{1 - e^{(\beta_3-a_3-a_2)\tau}}{a_3 - \beta_3 + a_2} + \mathcal{B}_2(\tau) \right. \\
 & \left. + e^{-\beta_3(x+t)} \mathcal{B}_3(\tau) + \tau \right]
 \end{aligned}$$

deduz-se finalmente a equação (9.5.40)

$$\begin{aligned}
 \mathcal{A}(\tau) = & \tau \left\{ -\frac{a_1\theta_1 e^{-\beta_1(x+t)}}{a_1 + \beta_1} - \theta_2 - \frac{a_3\theta_3 e^{\beta_3(x+t)}}{a_3 - \beta_3} + \frac{\sigma_1^2 e^{-2\beta_1(x+t)}}{2(a_1 + \beta_1)^2} + \frac{\sigma_2^2}{2a_2^2} \right. \\
 & \left. + \frac{\sigma_3^2 e^{2\beta_3(x+t)}}{2(a_3 - \beta_3)^2} + \frac{\sigma_1\sigma_2\rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1)a_2} + \frac{\sigma_1\sigma_3\rho_{13} e^{(\beta_3 - \beta_1)(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} + \frac{\sigma_2\sigma_3\rho_{23} e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \right\} \\
 & + \mathcal{B}_1(\tau) \left\{ -\frac{a_1\theta_1}{a_1 + \beta_1} + \frac{\sigma_1^2 e^{-\beta_1(x+t)}}{(a_1 + \beta_1)^2} + \frac{\sigma_1\sigma_2\rho_{12}}{(a_1 + \beta_1)a_2} + \frac{\sigma_1\sigma_3\rho_{13} e^{\beta_3(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} \right\} \\
 & + \mathcal{B}_2(\tau) \left\{ -\theta_2 + \frac{\sigma_2^2}{a_2^2} + \frac{\sigma_1\sigma_2\rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1)a_2} + \frac{\sigma_2\sigma_3\rho_{23} e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \right\} \\
 & + \mathcal{B}_3(\tau) \left\{ -\frac{a_3\theta_3}{a_3 - \beta_3} + \frac{\sigma_3^2 e^{\beta_3(x+t)}}{(a_3 - \beta_3)^2} + \frac{\sigma_1\sigma_3\rho_{13} e^{-\beta_1(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} + \frac{\sigma_2\sigma_3\rho_{23}}{a_2(a_3 - \beta_3)} \right\} \\
 & + \left\{ \frac{\sigma_1^2 e^{-2\beta_1(x+t)} (1 - e^{-2(\beta_1 + a_1)\tau})}{4(a_1 + \beta_1)^3} + \frac{\sigma_2^2 (1 - e^{-2a_2\tau})}{4a_2^3} \right. \\
 & + \frac{\sigma_3^2 e^{2\beta_3(x+t)} (1 - e^{2(\beta_3 - a_3)\tau})}{4(a_3 - \beta_3)^3} + \frac{\sigma_1\sigma_2\rho_{12} e^{-\beta_1(x+t)}}{(a_1 + \beta_1)a_2} \left( \frac{1 - e^{-(\beta_1 + a_1 + a_2)\tau}}{a_1 + \beta_1 + a_2} \right) \\
 & + \frac{\sigma_1\sigma_3\rho_{13} e^{(\beta_3 - \beta_1)(x+t)}}{(a_1 + \beta_1)(a_3 - \beta_3)} \left( \frac{1 - e^{[(\beta_3 - a_3) - (\beta_1 + a_1)]\tau}}{a_3 - \beta_3 + a_1 + \beta_1} \right) \\
 & \left. + \frac{\sigma_2\sigma_3\rho_{23} e^{\beta_3(x+t)}}{a_2(a_3 - \beta_3)} \left( \frac{1 - e^{(\beta_3 - a_3 - a_2)\tau}}{a_3 - \beta_3 + a_2} \right) \right\}
 \end{aligned}$$

## Capítulo 10

# Ensaio de Calibragem dos Modelos Estocásticos do Tipo Afim na População Portuguesa

*“If I find 10.000 ways something won't work, I haven't failed. I am not discouraged, because every wrong attempt discarded is another step forward.”*  
(Thomas Edison, 1847-1931)

### 10.1 Introdução

No capítulo anterior, argumentámos que a adopção de uma perspectiva mais ampla e realista sobre a evolução da longevidade humana no tempo e sobre a incerteza que encerra a sua trajectória futura pode ser conseguida modelando a intensidade de mortalidade mediante processos estocásticos.

Para tal, investigámos a adequação de um conjunto de processos estocásticos do tipo afim para ajustar a trajectória da intensidade  $\mu_x(t)$ , incluindo ou não mecanismos de reversão em direcção a uma média de longo prazo incorporando, nalguns casos, saltos (positivos e negativos) descontínuos, mas admitindo, em todos os modelos, soluções analíticas fechadas.

Neste capítulo, realizamos um primeiro ensaio exploratório quanto à capacidade destes modelos para descrever adequadamente a trajectória da intensidade  $\mu_x(t)$  na população portuguesa, considerando a realidade reproduzida pelas tábuas de mortalidade prospectivas derivadas no Capítulo 7. A utilização de processos estocásticos no *pricing* de contratos de seguros e na quantificação do risco de longevidade exige a cali-

bragem do modelo adequado e a estimação dos parâmetros que determinam a trajectória de  $\mu_x(t)$ .

O capítulo está organizado da seguinte forma. Na Secção 10.2, descrevemos brevemente as características da amostra usada no estudo e a metodologia adoptada na calibragem dos modelos. Na Secção 10.3, apresentamos os principais resultados do estudo. Por fim, na Secção 10.4 resumimos as principais conclusões dos ensaios efectuados.

## 10.2 Dados e metodologia de calibragem

### 10.2.1 Caracterização da amostra

Para testar a qualidade de ajustamento dos modelos do tipo afim para a intensidade de mortalidade na população portuguesa, consideramos as tábuas de mortalidade prospectivas derivadas no Capítulo 7. Recordamos que as tábuas foram estimadas com base no modelo de Poisson a partir dos dados brutos sobre o número de óbitos  $d_{x,t}$  e estimativas de população residente  $P_{x,t}$ , discriminados por idade e sexo no período  $(t_{\min}, t_{\max}) = (1970, 2004)$ , com intervalo de idades igual a  $(x_{\min}, x_{\max}) = (0, 84)$ , fornecidos pelo Instituto Nacional de Estatística para o conjunto da população portuguesa.

A aplicação da metodologia de Box-Jenkins na projecção do índice temporal, por um lado, e do método de fecho de tábuas proposto por Denuit e Goderniaux (2005), por outro, permitiu-nos finalmente derivar uma matriz  $\{q_{x,t} : x \in [0, 120], t \in [1970, 2124]\}$ , constituída pelos quocientes brutos para  $\{x \in [0, 84], t \in [1970, 2004]\}$  e por valores graduados para  $\{x \in [85, 120], t \in [1970, 2004]\}$  e por valores projectados para  $\{x \in [0, 120], t \in [2005, 2124]\}$ .

Conforme referimos no capítulo anterior, a perspectiva da mortalidade adoptada no contexto dos modelos de mortalidade estocástica é uma perspectiva diagonal, relativa a uma única geração  $x$ . Neste sentido, o procedimento de calibragem de cada um dos modelos realizado neste capítulo reporta-se a uma única geração.

A base de dados ideal para testar a capacidade dos modelos para descrever o comportamento da intensidade  $\mu_x(t)$  é uma base composta exclusivamente por taxas (ou quocientes) estimadas a partir de estatísticas demográficas observadas numa população. Esta tarefa requer o acompanhamento de uma geração desde o seu nascimento até à extinção. No caso da população portuguesa não existe, como vimos, informação suficiente para construir uma tábua prospectiva exclusivamente com base em dados observados, sendo necessário recorrer a um método de projecção. Trata-se de uma limitação que restringe, mas não invalida, a nossa análise. No entanto, devemos ter em conta que

uma parte da matriz usada para calibrar os modelos contém quocientes suavizados pela aplicação dos procedimentos de projecção e graduação.

### 10.2.2 Metodologia de calibragem

Para testar a qualidade dos modelos, realizámos um ensaio exploratório do procedimento de calibragem considerando a intensidade  $\mu_x(t)$  para uma única idade inicial e quatro gerações diferentes. Em termos mais precisos, a calibragem dos modelos incide sobre as probabilidades de sobrevivência das gerações que em 1970, 1980, 1990 e 2004 atingiram os 65 anos de idade.

A idade  $x = 65$  é uma idade chave, quer para os sistemas públicos e privados de segurança social, quer na avaliação de rendas vitalícias. A calibragem do modelo em quatro momentos diferentes permite aferir a sua consistência no tempo, verificando nomeadamente em que medida a qualidade do ajustamento é sensível às alterações na mortalidade observadas no tempo.

No total, são testados 10 modelos diferentes, a saber, o modelo de Vasicek, o modelo de Vasicek sem coeficiente de difusão e com saltos, o modelo de Cox-Ingersoll-Ross, o modelo de Dothan e os modelos de Merton, Feller e de Ornstein-Uhlenbeck com e sem saltos descontínuos. No caso dos modelos com saltos, adoptamos a hipótese avançada no capítulo anterior e supomos que a dimensão dos saltos segue uma distribuição assimétrica duplamente exponencial.

Na calibragem dos modelos, admitimos explicitamente a hipótese de que podem ocorrer, quer saltos positivos (com probabilidade  $\pi_1 \geq 0$  e dimensão média  $v_1 > 0$ ), quer saltos negativos (com probabilidade  $\pi_2 \geq 0$  e dimensão média  $v_2 > 0$ ) na intensidade  $\mu_x(t)$ . Para definir as probabilidades de choques negativos (subidas) e positivos (diminuições) sobre  $\mu_x(t)$ , simulámos o procedimento de calibragem considerando diferentes valores para  $\pi_1$  e  $\pi_2$  com o intuito de avaliar a sensibilidade dos resultados. Em função das conclusões retiradas deste exercício e das tendências decrescentes observadas na mortalidade, acabámos por fixar finalmente  $\pi_1 = 5\%$  e  $\pi_2 = 95\%$ .

A maior probabilidade atribuída a saltos negativos em  $\mu_x(t)$  é motivada pela expectativa de melhorias inesperadas na mortalidade. Estes ganhos de longevidade, correspondentes a pontos de descontinuidade no processo estocástico de  $\mu_x(t)$ , podem resultar, por exemplo, da descoberta de uma vacina contra uma doença letal. A frequência e a dimensão dos saltos são determinadas pelo procedimento de calibragem.

A estimação dos parâmetros é efectuada com recurso ao método dos mínimos quadrados ordinários, minimizando os desvios quadráticos entre as probabilidades de sobre-

vivência estimadas pelos diferentes modelos  ${}_{T-t}p_{65}^{\text{model}}(t)$  e as implícitas nas tábuas prospectivas  ${}_{T-t}p_{65}^{\text{TP}}(t)$ . Em termos formais, o vector de parâmetros  $\Theta$  de cada modelo é o resultado do seguinte problema de optimização:

$$\hat{\Theta} = \arg \min_{\Theta} \left\{ \mathcal{Q}^2 = \sum_{T=t+1}^{t+(x_{\max}-65)+1} \left( {}_{T-t}p_{65}^{\text{model}}(t) - {}_{T-t}p_{65}^{\text{TP}}(t) \right)^2 \right\} \quad (10.2.1)$$

onde  $x_{\max} = 120$  e  $t \in \{1970, 1980, 1990, 2004\}$ .

O procedimento de minimização é naturalmente condicionado pelas restrições impostas aos parâmetros na definição dos modelos. A resolução do problema de optimização e os resultados dos testes de qualidade do ajustamento foram obtidos com recurso a uma rotina informática especialmente escrita e implementada no *package* S-PLUS 2000 (Release 3).

## 10.3 Resultados

O Quadro 10.3.1 sintetiza, para ambos os sexos e momentos no tempo, o erro de estimação (ou de calibragem)  $\mathcal{Q}^2$  apurado nos 10 modelos do tipo afim para  $\mu_x(t)$  testados nesta amostra. Nos Quadros 10.3.2 a 10.3.11 reportamos, para cada um dos modelos, o valor óptimo dos parâmetros, o erro de calibragem e o valor inicial do processo estocástico,  $\mu_{65}(t)$ . O valor inicial  $\mu_{65}(t)$  foi estimado por  $-\ln(p_{65}(t))$ , ou seja, com base na hipótese de força de mortalidade constante avançada em (6.9.4).

Nas Figuras 10.3.1 a 10.3.10 representamos, para os diferentes modelos, a função de sobrevivência  ${}_{T-t}p_{65}(t)$  das gerações que em 1970 e 2004 alcançaram os 65 anos de idade e a estimada pelas tábuas prospectivas para a população portuguesa.

Como se observa, a qualidade de ajustamento varia substancialmente entre os modelos analisados. A variante naïve representada pelo modelo de Dothan proporciona, conforme se esperava, os piores resultados em ambos os sexos e em todos os períodos analisados. Com efeito, o modelo revela-se incapaz de captar o natural crescimento da intensidade  $\mu_x(t)$  com a idade.

Em face dos resultados obtidos nesta amostra podemos concluir, com alguma segurança, que os processos estocásticos sem componente determinística (tendência) se revelam, na amostra em estudo, desajustados para descrever a trajectória da intensidade de mortalidade.

### 10.3. Resultados

| Homens                              |             |            |            |            |
|-------------------------------------|-------------|------------|------------|------------|
| Modelo                              | $t = 1970$  | $t = 1980$ | $t = 1990$ | $t = 2004$ |
| Vasicek                             | 0.04235259  | 0.06089494 | 0.09792257 | 0.1598514  |
| Vasicek c/saltos                    | 0.0163585   | 0.02269441 | 0.03732777 | 0.0697703  |
| CIR                                 | 0.04271031  | 0.0624167  | 0.1063244  | 0.173152   |
| Merton                              | 0.04045165  | 0.05808783 | 0.09458064 | 0.1561025  |
| Merton c/saltos                     | 0.04021574  | 0.05717108 | 0.09409978 | 0.1559751  |
| Dothan                              | 5.219403    | 4.842206   | 5.095721   | 8.920934   |
| Feller                              | 0.000741098 | 0.00504437 | 0.01785774 | 0.0195093  |
| Feller c/saltos                     | 0.000483312 | 0.00113514 | 0.00423265 | 0.00743112 |
| OU <i>c/drift</i> positivo          | 0.000741105 | 0.00504440 | 0.0178578  | 0.0195094  |
| OU <i>c/drift</i> positivo e saltos | 0.000492032 | 0.00122845 | 0.00457461 | 0.00797233 |
| Mulheres                            |             |            |            |            |
| Modelo                              | $t = 1970$  | $t = 1980$ | $t = 1990$ | $t = 2004$ |
| Vasicek                             | 0.1136527   | 0.1580783  | 0.2455139  | 0.3832589  |
| Vasicek c/saltos                    | 0.0425649   | 0.0546209  | 0.08910286 | 0.1461823  |
| CIR                                 | 0.117125    | 0.1577336  | 0.2485609  | 0.3872163  |
| Merton                              | 0.1099472   | 0.1495434  | 0.2386794  | 0.3743746  |
| Merton c/saltos                     | 0.1090888   | 0.1490257  | 0.2381378  | 0.3739403  |
| Dothan                              | 11.16352    | 11.37778   | 12.48583   | 15.8944    |
| Feller                              | 0.00086477  | 0.00540914 | 0.02129198 | 0.03144563 |
| Feller c/saltos                     | 0.00035363  | 0.00071319 | 0.00348215 | 0.00615531 |
| OU <i>c/drift</i> positivo          | 0.00086480  | 0.00540923 | 0.02129223 | 0.03144625 |
| OU <i>c/drift</i> positivo e saltos | 0.00037507  | 0.00083302 | 0.00403972 | 0.00718258 |

Quadro 10.3.1: Erro de calibragem dos diferentes modelos estocásticos do tipo afim

Nos processos com tendência, verifica-se que os modelos que incorporam o mecanismo de reversão em direcção a uma média de longo prazo apresentam, regra geral, uma performance significativamente pior do que os modelos que não incluem essa característica. Com efeito, se exceptuarmos o desempenho do modelo de Vasicek sem termo de difusão e com saltos face ao modelo de Merton com e sem saltos, constata-se que os modelos com mecanismo de reversão, que tão bons resultados proporcionam na modelação das taxas de juro, se revelam menos apropriados para modelar a intensidade  $\mu_x(t)$  que os modelos sem esse traço distintivo.

Com efeito, verifica-se inclusive que o erro de calibragem aumenta significativamente quando passamos das gerações mais antigas (observadas em 1970) para as gerações mais jovens (observadas em 2004), um sintoma da incapacidade destes modelos para captar adequadamente os fenómenos de rectangularização e de expansão da função de sobrevivência registados neste período.

Esta conclusão é claramente confirmada nas Figuras 10.3.1, 10.3.2 e 10.3.3, onde é evidente que o ajustamento pouco satisfatório (em particular no casos dos homens) conseguido em 1970 se deteriora significativamente na geração que em 2004 alcança os 65 anos. A inabilidade dos modelos com mecanismo de reversão para captar o fenómeno de rectangularização manifesta-se ainda através de probabilidades de sobrevivência bastante inferiores às implícitas nas tábuas prospectivas até por volta dos 85-90 anos, e por uma clara subestimação da mortalidade nas idades avançadas.

Por outro lado, nos processos estocásticos com tendência determinística e sem mecanismo de reversão, observa-se que os modelos cuja componente determinística cresce de forma exponencial (modelos de Feller e Ornstein-Uhlenbeck com e sem saltos) apresentam um desempenho consistentemente mais adequado do que os seus homólogos onde o crescimento é linear (modelo de Merton). Acresce que a boa qualidade de ajustamento destes modelos é preservada nas gerações mais jovens, onde o fenómeno de rectangularização é mais significativo.

Assinale-se que este resultado é consistente com os tradicionais modelos determinísticos analisados nos capítulos precedentes, que estipulam, como vimos, um crescimento exponencial da mortalidade com a idade e não incluem mecanismo de reversão.

Outra conclusão interessante é a de que os modelos que incluem saltos na equação de difusão apresentam sistematicamente melhores resultados do que aqueles que não admitem choques descontínuos (positivos ou negativos) na tendência do processo estocástico. Como se comprova pela observação do Quadro 10.3.1 e das Figuras 10.3.8 e 10.3.10, a qualidade do ajustamento proporcionada por estes modelos é bastante boa, mesmo perante a acentuação do fenómeno de rectangularização nas gerações mais jovens.

Este resultado sugere que a inclusão de saltos descontínuos (em particular de saltos negativos) pode contribuir para, em conjunto com o termo de difusão, descrever as variações aleatórias na intensidade de mortalidade no tempo. Esta conclusão não surpreende de todo uma vez que a intensidade  $\mu_x(t)$  é uma variável aleatória cuja volatilidade de curto prazo é pouco significativa, mas com potencial para grandes oscilações a médio e longo prazo.

De entre os vários modelos analisados, o modelo de Feller com saltos apresenta, em ambos os sexos e em todos os momentos analisados, os melhores resultados em termos de ajustamento, seguido de perto pela equação de Ornstein-Uhlenbeck com saltos.

### 10.3. Resultados

| Homens        |             |             |             |             |
|---------------|-------------|-------------|-------------|-------------|
|               | $t = 1970$  | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901  | 0.02774125  | 0.02558451  | 0.01689187  |
| $a$           | 0.002602931 | 0.003064172 | 0.002597631 | 0.001811042 |
| $\theta$      | 1.99982     | 1.499783    | 1.410356    | 1.4018110   |
| $\sigma$      | 0.0000157   | 0.0000139   | 0.0000122   | 0.0000118   |
| $Q^2$         | 0.04235259  | 0.06089494  | 0.09792257  | 0.1598514   |

| Mulheres      |             |             |             |             |
|---------------|-------------|-------------|-------------|-------------|
|               | $t = 1970$  | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793  | 0.01375416  | 0.01163745  | 0.007780187 |
| $a$           | 0.002131906 | 0.003882015 | 0.002211804 | 0.00197203  |
| $\theta$      | 1.979812    | 1.89738     | 1.478399    | 1.460733    |
| $\sigma$      | 0.0000144   | 0.0000133   | 0.0000128   | 0.0000125   |
| $Q^2$         | 0.1136527   | 0.1580783   | 0.2455139   | 0.3832589   |

Quadro 10.3.2: Estimativas dos parâmetros do modelo de Vasicek

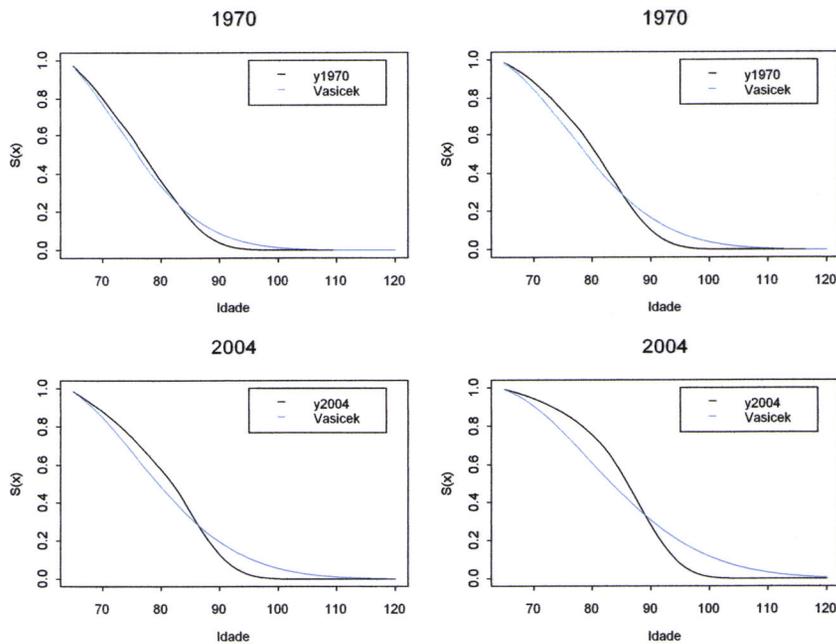


Figura 10.3.1: Calibragem da probabilidade de sobrevivência  ${}_{T-t}p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Vasicek (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |             |              |               |             |
|---------------|-------------|--------------|---------------|-------------|
|               | $t = 1970$  | $t = 1980$   | $t = 1990$    | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901  | 0.02774125   | 0.02558451    | 0.01689187  |
| $a$           | 0.007744907 | 0.007143821  | 0.006444377   | 0.006059281 |
| $\theta$      | 1.023741    | 0.9945932    | 0.9856125     | 0.9631203   |
| $\eta$        | 0.006723491 | 0.002601993  | 0.00001443609 | 0.02750041  |
| $v_1$         | 0.000013521 | 0.0000091582 | 0.04339722    | 0.00011292  |
| $v_2$         | 0.005939451 | 0.006452496  | 0.006440902   | 0.003046289 |
| $Q^2$         | 0.0163585   | 0.02269441   | 0.03732777    | 0.0697703   |

| Mulheres      |              |              |             |             |
|---------------|--------------|--------------|-------------|-------------|
|               | $t = 1970$   | $t = 1980$   | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793   | 0.01375416   | 0.01163745  | 0.007780187 |
| $a$           | 0.00681621   | 0.006364764  | 0.005853443 | 0.005433239 |
| $\theta$      | 1.058501     | 0.9975712    | 0.9896934   | 0.9820363   |
| $\eta$        | 0.00401142   | 0.004294324  | 0.00552713  | 0.01437855  |
| $v_1$         | 0.0009546517 | 0.0009506174 | 0.00014321  | 0.05048119  |
| $v_2$         | 0.005747897  | 0.005369336  | 0.004802932 | 0.003539567 |
| $Q^2$         | 0.0425649    | 0.0546209    | 0.08910286  | 0.1461823   |

Quadro 10.3.3: Estimativas dos parâmetros do modelo de Vasicek com incerteza ditada exclusivamente por saltos

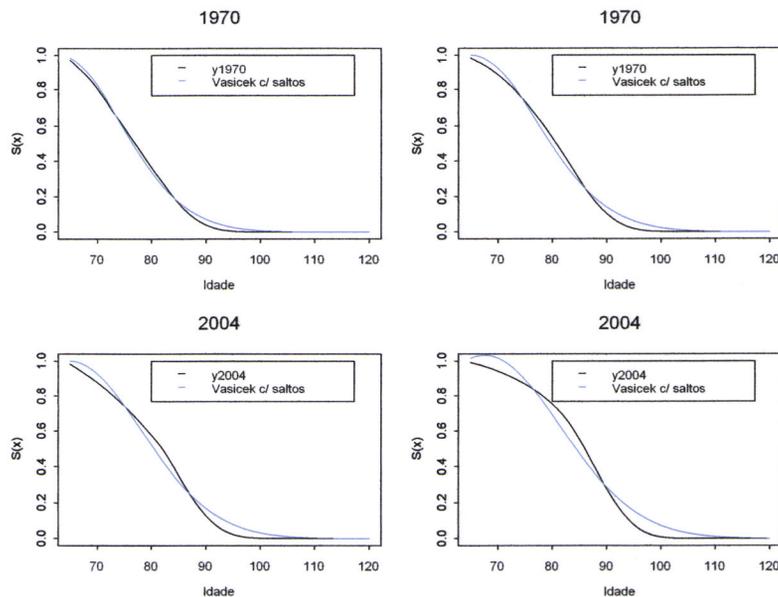


Figura 10.3.2: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Vasicek com saltos (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |               |             |              |              |
|---------------|---------------|-------------|--------------|--------------|
|               | $t = 1970$    | $t = 1980$  | $t = 1990$   | $t = 2004$   |
| $\mu_{65}(t)$ | 0.02765901    | 0.02774125  | 0.02558451   | 0.01689187   |
| $a$           | 0.003087752   | 0.004710139 | 0.009001458  | 0.008128622  |
| $\theta$      | 1.695073      | 0.9973636   | 0.4687738    | 0.4685706    |
| $\sigma$      | 0.00001192125 | 0.00001163  | 0.0000146337 | 0.0000103841 |
| $Q^2$         | 0.04271031    | 0.0624167   | 0.1063244    | 0.173152     |

| Mulheres      |             |               |               |              |
|---------------|-------------|---------------|---------------|--------------|
|               | $t = 1970$  | $t = 1980$    | $t = 1990$    | $t = 2004$   |
| $\mu_{65}(t)$ | 0.01472793  | 0.01375416    | 0.01163745    | 0.007780187  |
| $a$           | 0.004110858 | 0.003726601   | 0.003188738   | 0.002846031  |
| $\theta$      | 1.0059467   | 1.0039989     | 0.9963081     | 0.9858996    |
| $\sigma$      | 0.000009723 | 0.00001334821 | 0.00001036019 | 0.0000465832 |
| $Q^2$         | 0.117125    | 0.1577336     | 0.2485609     | 0.3872163    |

Quadro 10.3.4: Estimativas dos parâmetros do modelo de Cox-Ingersoll-Ross

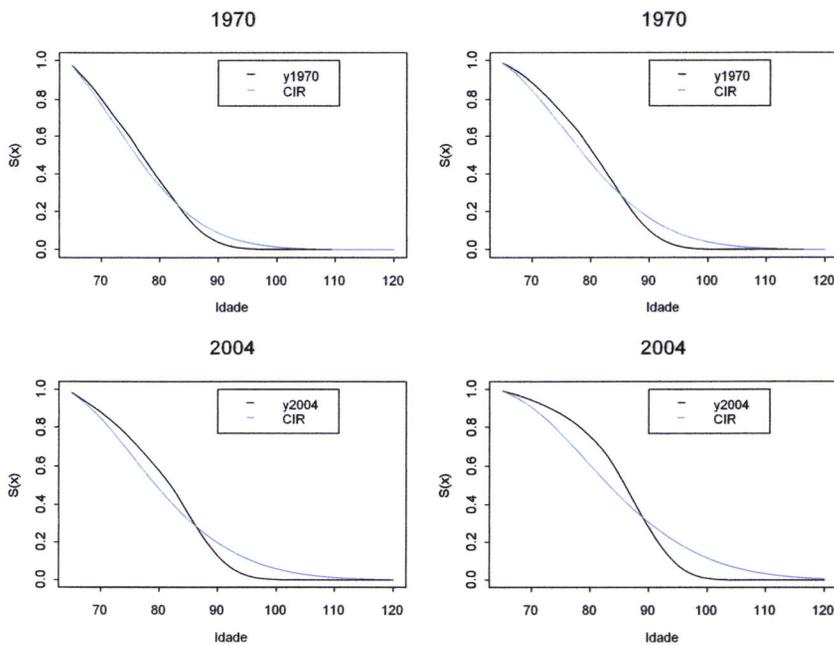


Figura 10.3.3: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Cox-Ingersoll-Ross (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |             |             |             |             |
|---------------|-------------|-------------|-------------|-------------|
|               | $t = 1970$  | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901  | 0.02774125  | 0.02558451  | 0.01689187  |
| $a$           | 0.005055519 | 0.004427015 | 0.003765256 | 0.003545631 |
| $\sigma$      | 0.00001192  | 0.00001163  | 0.00001146  | 0.00001091  |
| $Q^2$         | 0.04045165  | 0.05808783  | 0.09458064  | 0.1561025   |
| Mulheres      |             |             |             |             |
|               | $t = 1970$  | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793  | 0.01375416  | 0.01163745  | 0.007780187 |
| $a$           | 0.004171711 | 0.003725035 | 0.003237471 | 0.002896283 |
| $\sigma$      | 0.00001071  | 0.00001101  | 0.00001056  | 0.00001083  |
| $Q^2$         | 0.1099472   | 0.1495434   | 0.2386794   | 0.3743746   |

Quadro 10.3.5: Estimativas dos parâmetros do modelo de Merton

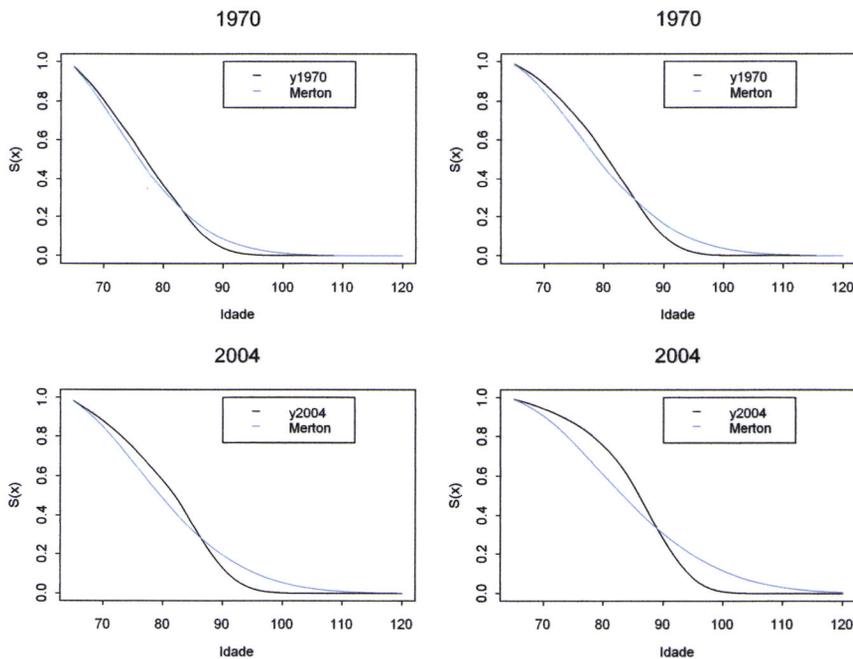


Figura 10.3.4: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Merton (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|
|               | $t = 1970$   | $t = 1980$   | $t = 1990$   | $t = 2004$   |
| $\mu_{65}(t)$ | 0.02765901   | 0.02774125   | 0.02558451   | 0.01689187   |
| $a$           | 0.005053242  | 0.004424883  | 0.003763314  | 0.003543845  |
| $\sigma$      | 0.000011071  | 0.000010393  | 0.000011229  | 0.000010184  |
| $\eta$        | 0.00001115   | 0.00001104   | 0.00001087   | 0.00001067   |
| $v_1$         | 0.0000451156 | 0.0001107311 | 0.0000100005 | 0.0001939105 |
| $v_2$         | 0.0000105816 | 0.000010499  | 0.0000132623 | 0.0000104505 |
| $Q^2$         | 0.04021574   | 0.05717108   | 0.09409978   | 0.1559751    |

| Mulheres      |             |              |             |             |
|---------------|-------------|--------------|-------------|-------------|
|               | $t = 1970$  | $t = 1980$   | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793  | 0.01375416   | 0.01163745  | 0.007780187 |
| $a$           | 0.004169785 | 0.00372323   | 0.003235819 | 0.002894764 |
| $\sigma$      | 0.00001045  | 0.00001074   | 0.00000983  | 0.00001161  |
| $\eta$        | 0.00009851  | 0.00010124   | 0.00009193  | 0.00011012  |
| $v_1$         | 0.000108623 | 0.0001013144 | 0.000106861 | 0.000084690 |
| $v_2$         | 0.000094152 | 0.0000879932 | 0.000087001 | 0.000097976 |
| $Q^2$         | 0.1090888   | 0.1490257    | 0.2381378   | 0.3739403   |

Quadro 10.3.6: Estimativas dos parâmetros do modelo de Merton com saltos

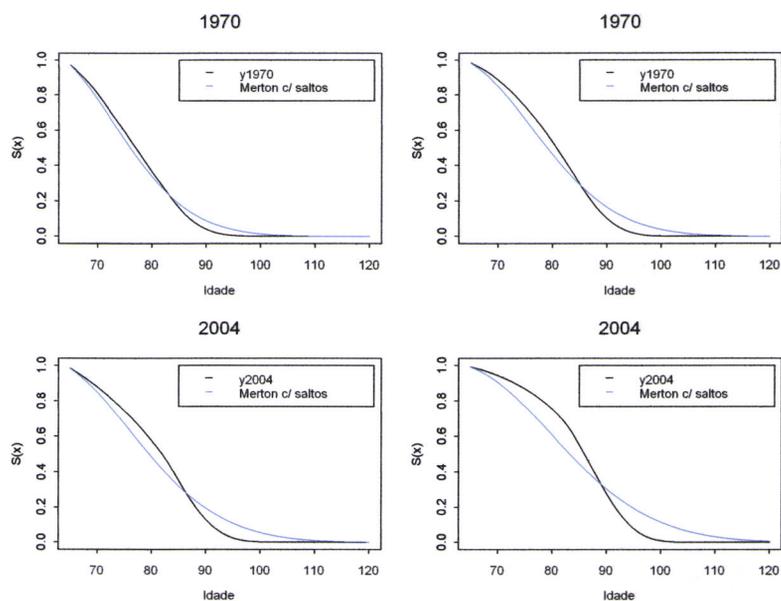


Figura 10.3.5: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Merton com saltos (homens à esquerda e mulheres à direita)

10.3. Resultados

| Homens        |            |            |            |             |
|---------------|------------|------------|------------|-------------|
|               | $t = 1970$ | $t = 1980$ | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901 | 0.02774125 | 0.02558451 | 0.01689187  |
| $\sigma$      | 0.00090181 | 0.00001011 | 0.00001817 | 0.00001911  |
| $Q^2$         | 5.219403   | 4.842206   | 5.095721   | 8.920934    |
| Mulheres      |            |            |            |             |
|               | $t = 1970$ | $t = 1980$ | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793 | 0.01375416 | 0.01163745 | 0.007780187 |
| $\sigma$      | 0.00013215 | 0.00001598 | 0.00006799 | 0.00000921  |
| $Q^2$         | 11.16352   | 11.37778   | 12.48583   | 15.8944     |

Quadro 10.3.7: Estimativas dos parâmetros do modelo de Dothan

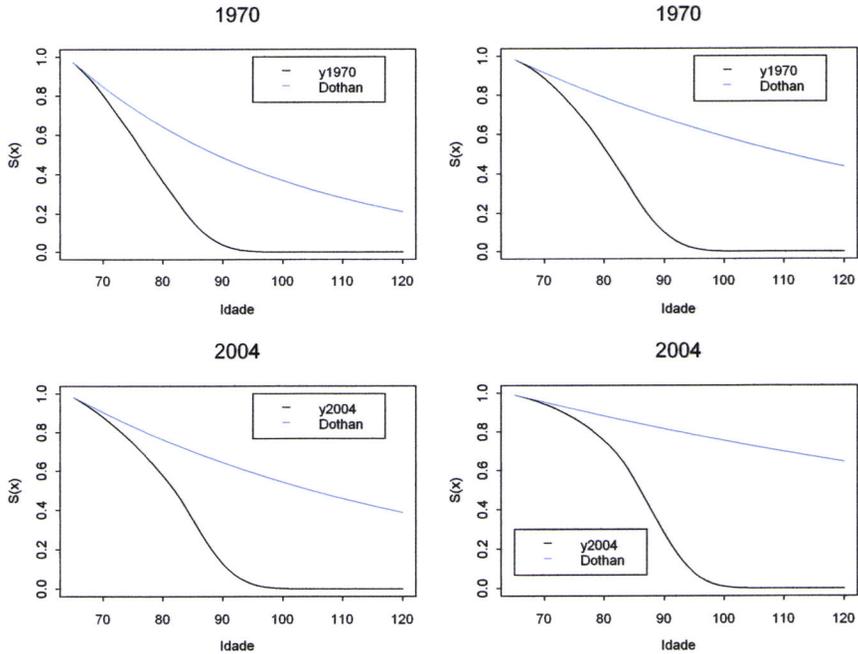


Figura 10.3.6: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Dothan (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |              |             |            |             |
|---------------|--------------|-------------|------------|-------------|
|               | $t = 1970$   | $t = 1980$  | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901   | 0.02774125  | 0.02558451 | 0.01689187  |
| $a$           | 0.09363077   | 0.08492182  | 0.07872597 | 0.09258513  |
| $\sigma$      | 0.000009721  | 0.000014591 | 0.00001071 | 0.00001931  |
| $Q^2$         | 0.0007410979 | 0.005044374 | 0.01785774 | 0.01950928  |
| Mulheres      |              |             |            |             |
|               | $t = 1970$   | $t = 1980$  | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793   | 0.01375416  | 0.01163745 | 0.007780187 |
| $a$           | 0.1103031    | 0.1051975   | 0.1028372  | 0.1123945   |
| $\sigma$      | 0.00001067   | 0.00001123  | 0.00001198 | 0.00001001  |
| $Q^2$         | 0.0008647697 | 0.005409142 | 0.02129198 | 0.03144563  |

Quadro 10.3.8: Estimativas dos parâmetros do modelo de Feller

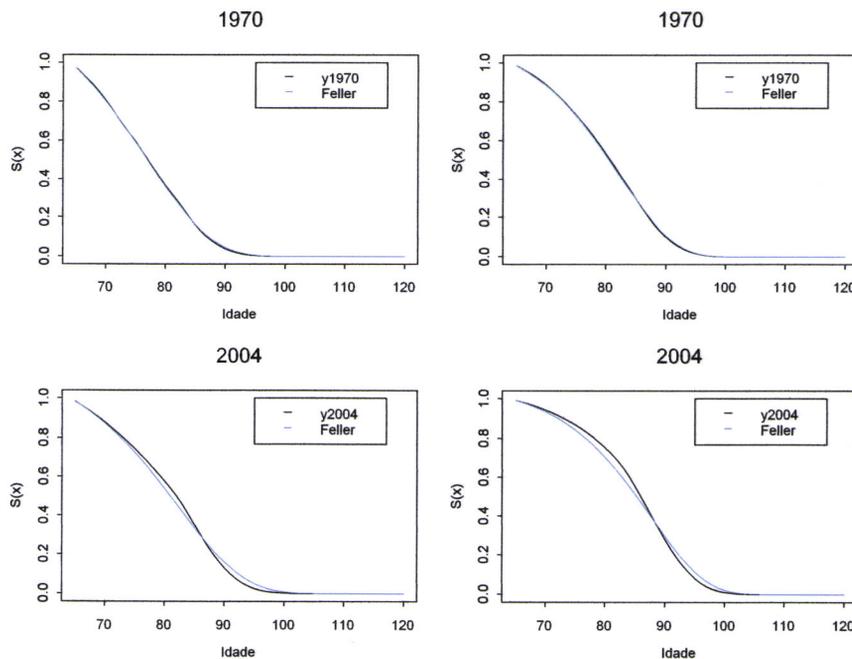


Figura 10.3.7: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Feller (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |             |             |            |               |
|---------------|-------------|-------------|------------|---------------|
|               | $t = 1970$  | $t = 1980$  | $t = 1990$ | $t = 2004$    |
| $\mu_{65}(t)$ | 0.02765901  | 0.02774125  | 0.02558451 | 0.01689187    |
| $a$           | 0.09516212  | 0.09033169  | 0.08739382 | 0.09949474    |
| $\sigma$      | 0.00001013  | 0.00001131  | 0.00000981 | 0.00000978    |
| $\eta$        | 0.0117887   | 0.03936915  | 0.06544481 | 0.05226689    |
| $v_1$         | 0.02654017  | 0.02876439  | 0.02726195 | 0.02757463    |
| $v_2$         | 0.001128449 | 0.0001023   | 0.0001102  | 0.00009724921 |
| $Q^2$         | 0.000483312 | 0.001135141 | 0.00423265 | 0.007431117   |

| Mulheres      |              |              |             |             |
|---------------|--------------|--------------|-------------|-------------|
|               | $t = 1970$   | $t = 1980$   | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793   | 0.01375416   | 0.01163745  | 0.007780187 |
| $a$           | 0.1119171    | 0.1096041    | 0.1101916   | 0.1199389   |
| $\sigma$      | 0.00001044   | 0.00001033   | 0.00001082  | 0.00001049  |
| $\eta$        | 0.01180289   | 0.03190069   | 0.05536174  | 0.05693019  |
| $v_1$         | 0.0284391    | 0.02890383   | 0.02727089  | 0.02644525  |
| $v_2$         | 0.0001189    | 0.0001098    | 0.0001072   | 0.0001066   |
| $Q^2$         | 0.0003536312 | 0.0007131984 | 0.003482145 | 0.006155311 |

Quadro 10.3.9: Estimativas dos parâmetros do modelo de Feller com saltos

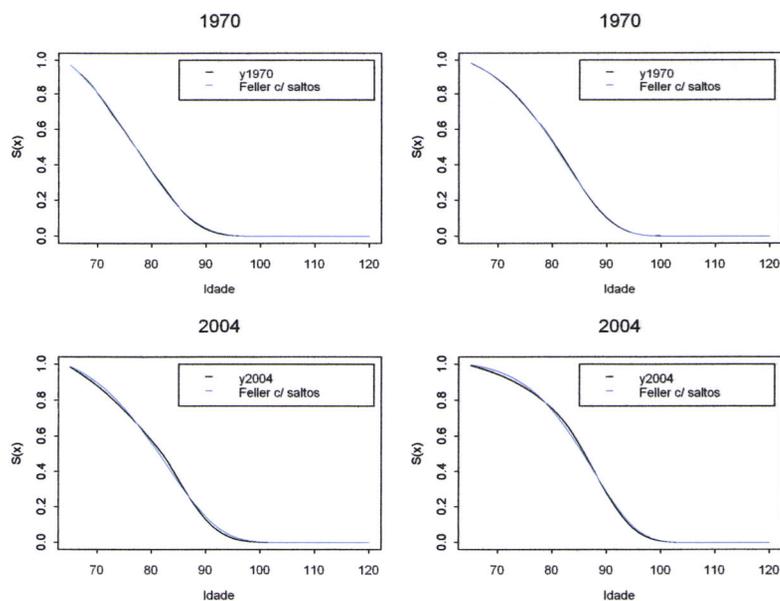


Figura 10.3.8: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pelo modelo de Feller com saltos (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |              |             |            |             |
|---------------|--------------|-------------|------------|-------------|
|               | $t = 1970$   | $t = 1980$  | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901   | 0.02774125  | 0.02558451 | 0.01689187  |
| $a$           | 0.0936308    | 0.08492186  | 0.078726   | 0.0925852   |
| $\sigma$      | 0.000012138  | 0.000009750 | 0.00009862 | 0.000011292 |
| $Q^2$         | 0.0007411049 | 0.005044397 | 0.0178578  | 0.0195094   |
| Mulheres      |              |             |            |             |
|               | $t = 1970$   | $t = 1980$  | $t = 1990$ | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793   | 0.01375416  | 0.01163745 | 0.007780187 |
| $a$           | 0.1103031    | 0.1051976   | 0.1028373  | 0.1123947   |
| $\sigma$      | 0.00001414   | 0.00001281  | 0.00001198 | 0.00001125  |
| $Q^2$         | 0.000864797  | 0.005409227 | 0.02129223 | 0.03144625  |

Quadro 10.3.10: Estimativas dos parâmetros da equação de Ornstein-Uhlenbeck com *drift* positivo

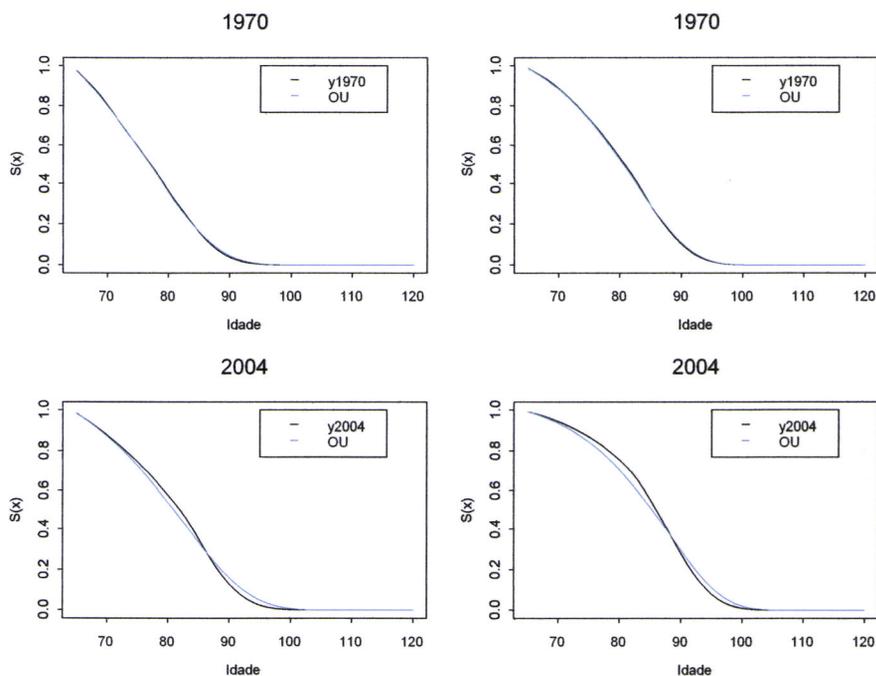


Figura 10.3.9: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pela equação de Ornstein-Uhlenbeck com *drift* positivo (homens à esquerda e mulheres à direita)

### 10.3. Resultados

| Homens        |              |             |             |             |
|---------------|--------------|-------------|-------------|-------------|
|               | $t = 1970$   | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.02765901   | 0.02774125  | 0.02558451  | 0.01689187  |
| $a$           | 0.09502302   | 0.08987547  | 0.08663877  | 0.09876713  |
| $\sigma$      | 0.00001756   | 0.00001279  | 0.00001022  | 0.00001018  |
| $\eta$        | 0.02164396   | 0.07929218  | 0.1301158   | 0.09708766  |
| $v_1$         | 0.00001095   | 0.00001039  | 0.000009122 | 0.00001198  |
| $v_2$         | 0.001401751  | 0.001471971 | 0.002069324 | 0.000119    |
| $Q^2$         | 0.0004920324 | 0.001228453 | 0.004574614 | 0.007972327 |

| Mulheres      |              |             |             |             |
|---------------|--------------|-------------|-------------|-------------|
|               | $t = 1970$   | $t = 1980$  | $t = 1990$  | $t = 2004$  |
| $\mu_{65}(t)$ | 0.01472793   | 0.01375416  | 0.01163745  | 0.007780187 |
| $a$           | 0.1117468    | 0.1091729   | 0.109424    | 0.1190731   |
| $\sigma$      | 0.000013215  | 0.000009158 | 0.00009679  | 0.000011292 |
| $\eta$        | 0.02105941   | 0.05881956  | 0.09900384  | 0.09759169  |
| $v_1$         | 0.000009723  | 0.000013348 | 0.00001036  | 0.00009658  |
| $v_2$         | 0.0000951    | 0.0001024   | 0.0000993   | 0.0001012   |
| $Q^2$         | 0.0003750678 | 0.00083302  | 0.004039717 | 0.007182579 |

Quadro 10.3.11: Estimativas dos parâmetros da equação de Ornstein-Uhlenbeck com *drift* positivo e saltos

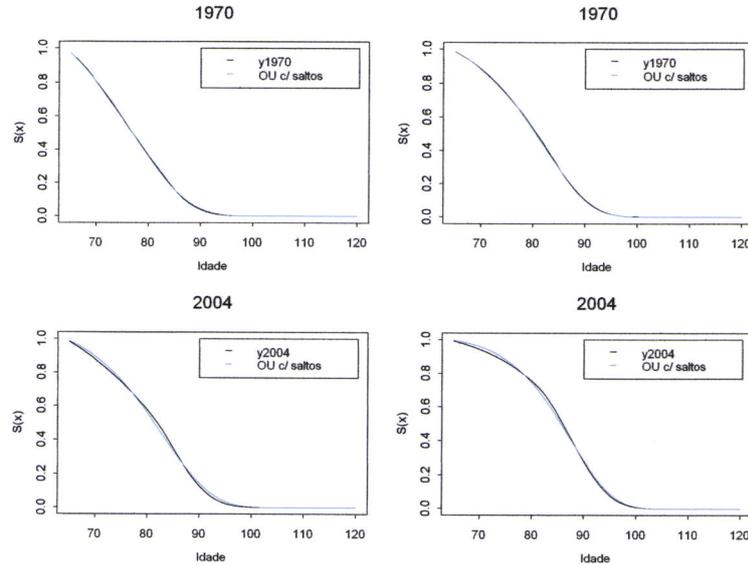


Figura 10.3.10: Calibragem da probabilidade de sobrevivência  $T-t p_{65}(t)$  em função da idade  $x + T - t$  para  $t = 1970$  e  $t = 2004$  pela equação de Ornstein-Uhlenbeck com *drift* positivo e saltos (homens à esquerda e mulheres à direita)

### 10.3. Resultados

Nas Figuras 10.3.11, 10.3.12, 10.3.13 e 10.3.14 oferecemos uma panorâmica geral sobre o desempenho dos modelos<sup>1</sup> representando os desvios entre as probabilidades  $T-t p_{65}^{\text{model}}(t)$  e  $T-t p_{65}^{\text{TP}}(t)$ . Assinalamos, em particular, os desvios significativos em torno de  $T-t p_{65}^{\text{TP}}(t)$  dos modelos com mecanismo de reversão à média.

Em relação às estimativas dos parâmetros, o primeiro aspecto a realçar diz respeito ao valor bastante baixo do coeficiente de difusão  $\sigma$  dos processos estocásticos. Em contrapartida, a dimensão dos saltos (positivos ou negativos), expressa nos parâmetros  $v_1$  e  $v_2$ , assume maior expressão, oscilando bastante entre um mínimo de 0.00001 e um máximo de 0.05, embora com uma frequência (dada pelo parâmetro  $\eta$ ) baixa.

Estes resultados constituem um sinal claro de que os saltos descontínuos assumem um papel importante no processo de calibragem e de explicação da componente estocástica da intensidade  $\mu_x(t)$ .

Nos modelos com mecanismo de reversão à média, observa-se que a velocidade de convergência dos processos estocásticos (parâmetro  $a$ ) é relativamente estável no tempo, em particular nas duas variantes do modelo de Vasicek. Nestes modelos, a média de longo prazo  $\theta$  apresenta, como se antevia, uma tendência decrescente no tempo, embora o valor tenda a estabilizar nas gerações mais jovens.

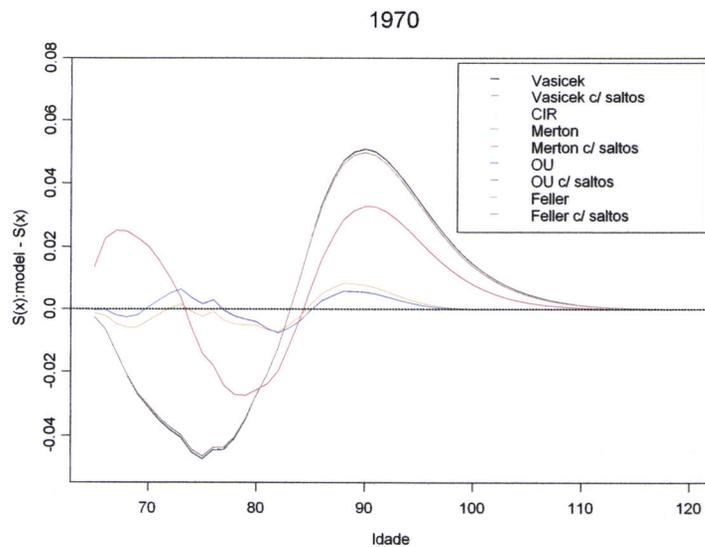


Figura 10.3.11: Desvios entre a função de sobrevivência  $T-t p_{65}(1970)$  estimada pelos modelos e a implícita nas tábuas prospectivas, Homens

<sup>1</sup>Excluindo o modelo de Dothan pela sua clara desadequação ao problema em análise.

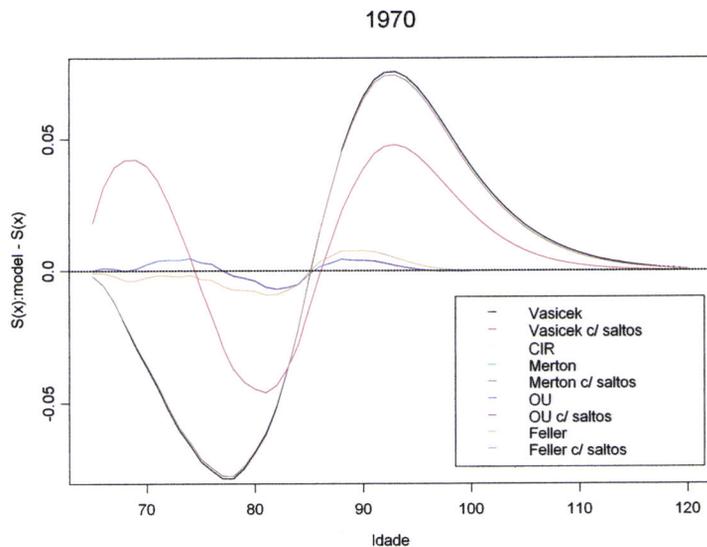


Figura 10.3.12: Desvios entre a função de sobrevivência  $T-tP_{65}(1970)$  estimada pelos modelos e a implícita nas tábuas prospectivas, Mulheres

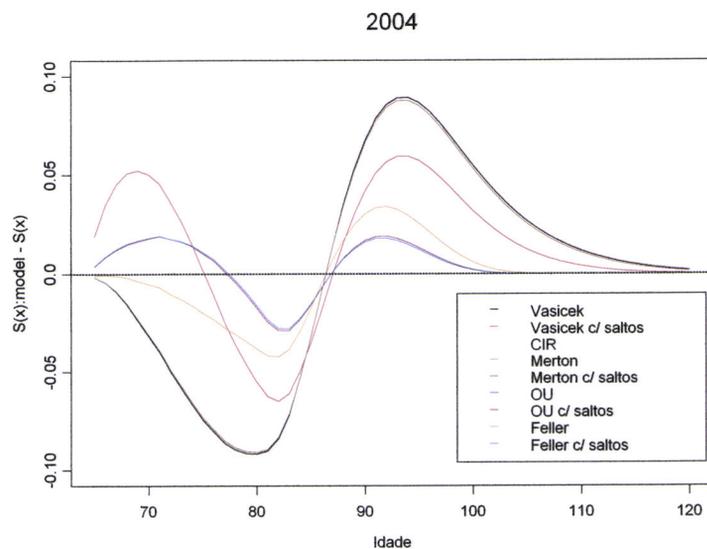


Figura 10.3.13: Desvios entre a função de sobrevivência  $T-tP_{65}(2004)$  estimada pelos modelos e a implícita nas tábuas prospectivas, Homens

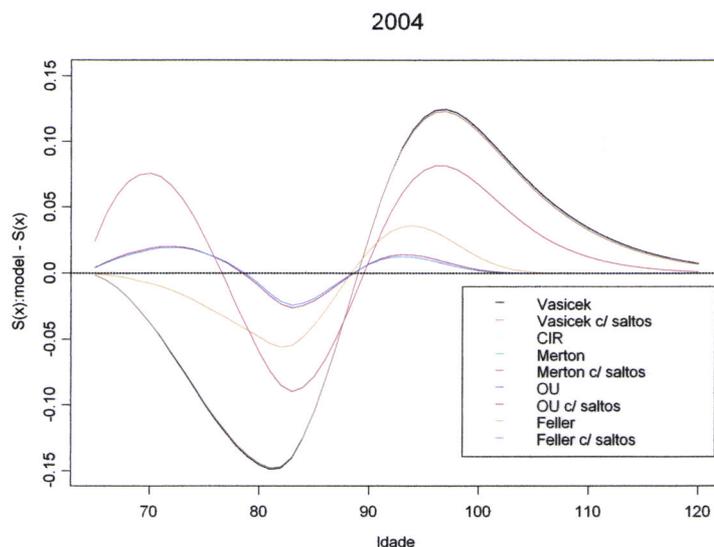


Figura 10.3.14: Desvios entre a função de sobrevivência  $T-t p_{65}(2004)$  estimada pelos modelos e a implícita nas tábuas prospectivas, Mulheres

## 10.4 Conclusão

Neste capítulo, investigámos a utilização de processos estocásticos do tipo afim para descrever a trajetória da intensidade  $\mu_x(t)$  na população portuguesa. Para tal, calibrámos as equações de difusão de um conjunto variado de modelos usando a informação contida nas tábuas prospectivas derivadas no Capítulo 7.

Os resultados obtidos no procedimento de calibragem sugerem que os processos estocásticos sem componente determinística (tendência) e os processos com mecanismo de reversão em direção a uma média de longo prazo não podem ser considerados adequados para caracterizar a trajetória da intensidade  $\mu_x(t)$ . Em contrapartida, os processos do tipo afim cuja componente determinística cresce de forma exponencial revelam um desempenho adequado e consistente no tempo, sobretudo quando comparados com processos cuja tendência determinística cresce de forma linear.

Os resultados dos testes de ajustamento indiciam igualmente que os modelos que incluem saltos na equação de difusão apresentam, sistematicamente, melhores resultados do que aqueles que não admitem choques descontínuos (positivos ou negativos) na tendência do processo estocástico. Esta conclusão é reforçada pelo bom desempenho

#### 10.4. Conclusão

---

dos modelos com estas características em contextos de maior incidência do fenómeno de rectangularização da função de sobrevivência. Em consequência, conclui-se que a inclusão de saltos descontínuos na componente estocástica dos processos pode contribuir para descrever as variações aleatórias na intensidade de mortalidade.

Investigações futuras deverão ser capazes de avaliar a sensibilidade destes resultados a alterações na amostra, na idade inicial do procedimento de calibragem e de verificar a importância de um eventual efeito geração numa base mais alargada.

## Capítulo 11

# Aplicações Actuarias e Cobertura do Risco de Longevidade

*“Longevity risk is the next big frontier for financial markets - unless of course someone ruins it all in the meantime and discovers the secret of eternal life.”*

(Blake *et al.*, 2006a)

### 11.1 Introdução

As alterações demográficas registadas nas últimas décadas tiveram impactos significativos no mercado de seguros e de fundos de pensões. Empresas e entidades reguladoras começam a tomar consciência da importância que a inclusão da incerteza em torno da evolução da mortalidade, em particular do risco de desvios sistemáticos em relação aos valores esperados, assume na correcta avaliação da capacidade das instituições para respeitarem os seus compromissos.

Este risco é acrescido pelos problemas de sustentabilidade financeira dos tradicionais sistemas de segurança social, nos quais a moderação dos benefícios concedidos pelos sistemas de repartição, a progressiva conversão dos planos de pensões de benefício definido em planos de contribuição definida e a tendência para a adopção de sistemas de segurança social assentes em mecanismos de capitalização, tenderão a reduzir a parcela do rendimento na reforma que é “garantida”, aumentando a procura de soluções de cobertura individual.

No mercado de seguros, a determinação dos prémios dos contratos segue uma prática convencional que envolve a aplicação do princípio da equivalência e o uso de bases técnicas (demográficas e financeiras) determinísticas prudentes. No passado, esta prática

## 11.1. Introdução

---

conduzia frequentemente a excedentes sistemáticos, que eram usados para adquirir cobertura adicional ou redistribuídos entre os beneficiários sob a forma de bónus. Independentemente da maior ou menor margem de segurança que a escolha de uma tábua de mortalidade conservadora pudesse representar no momento da celebração dos contratos, a verdade é que a evolução da mortalidade reforçou a ideia de que as companhias estão expostas ao risco de longevidade.

A natureza sistemática deste risco faz com que a adopção dos tradicionais mecanismos de agregação (“*pooling*”) dos contratos não seja, neste caso, eficaz, uma vez que afecta todas as apólices na mesma direcção e ao mesmo tempo. Em consequência, as companhias devem procurar mecanismos de cobertura do risco de longevidade diferentes, que envolvam soluções internas ou externas, usando técnicas clássicas no mercado de seguros ou clamando por soluções inovadoras, que passem pela titularização dos riscos ou pela criação de um novo mercado de derivados de mortalidade.

Os problemas no *pricing* e no cálculo das reservas criados pelas tendências demográficas levaram as entidades reguladoras (em particular o International Accounting Standards Board) a advogar a avaliação de activos e responsabilidades pelo seu valor justo, i.e., pelo seu valor de mercado. Em termos práticos, esta proposição determina que as avaliações devem considerar convenientemente todos os factores de risco (financeiros e demográficos, diversificáveis ou sistemáticos) e fundamentar as hipóteses usadas na sua medição.

Neste capítulo analisamos a aplicação de modelos de mortalidade estocástica na avaliação de contratos de seguro com cash flows contingentes à sobrevivência ou morte do indivíduo e discutimos variadíssimas técnicas de cobertura do risco de longevidade.

O capítulo está organizado da seguinte forma. Nas Secções 11.2 e 11.3 fazemos a analogia entre contratos de seguro do ramo vida e activos sujeitos a risco de crédito e descrevemos a estrutura do modelo de avaliação num contexto estocástico, considerando quer os factores de risco financeiros, quer os demográficos.

Na Secção 11.4 discutimos a aplicação desta metodologia na avaliação de apólices tradicionais do ramo vida e de produtos estruturados, dando especial atenção ao problema da determinação do preço de mercado do risco de longevidade. Em particular, desenvolvemos um novo método de estimação do preço do risco, recorrendo a um princípio clássico de determinação do prémio em contratos de seguro.

Por fim, na Secção 11.5 discutimos um conjunto extenso de soluções de cobertura do risco de longevidade que as companhias podem adoptar para gerir a exposição das suas carteiras.

## 11.2 Contratos de seguro e activos com risco de crédito

No capítulo anterior vimos que existem, grosso modo, duas grandes abordagens na avaliação de activos sujeitos a risco de crédito: a abordagem estrutural, que procura descrever o evento de incumprimento modelando a dinâmica do valor dos activos da empresa, e a abordagem pela intensidade, que assenta na fixação exógena da probabilidade condicional de incumprimento, materializada na especificação de equações diferenciais estocásticas para a intensidade de incumprimento.

Um resultado fundamental na abordagem pela intensidade é aquele que estabelece que a avaliação de activos com risco de crédito pode ser realizada usando a mesma construção teórica adoptada na avaliação de activos financeiros sem risco. A única diferença diz respeito à necessidade de, num contexto de neutralidade face ao risco, se efectuar a operação de desconto usando, não uma taxa de juro sem risco, mas uma taxa de desconto (ajustada ou fictícia) que leve em linha de conta quer o momento, quer a probabilidade de ocorrência do incumprimento.

Em particular, designemos por  $R$  o processo para a taxa de desconto ajustada e por  $B^d(0, T)$  o preço de uma obrigação de cupão zero com risco de crédito, maturidade em  $T$  e valor nominal unitário. Com base nos resultados de, por exemplo, Artzner e Delbaen (1995), Lando (1998), Duffie e Singleton (1999), podemos escrever

$$B^d(0, T) = \mathbb{E} \left[ e^{-\int_0^T R_s ds} \right] + \int_0^T w_s q_s ds, \quad (11.2.1)$$

onde o primeiro termo no lado direito corresponde ao preço de um activo que paga uma unidade monetária na maturidade em caso de respeito integral das obrigações por parte do emitente,  $w_s$  denota o montante recuperado em caso de incumprimento no momento  $s$  e  $q_s$  a função de densidade do preço do activo no intervalo  $(s, s + ds)$ . O valor esperado e a densidade são ambos especificados com respeito a uma medida de probabilidade “*risk-neutral*”.

Se interpretarmos o incumprimento como o evento correspondente ao óbito da pessoa segura (e a intensidade de incumprimento como a intensidade de mortalidade), podemos verificar que a equação (11.2.1) cobre um conjunto significativo de contratos de seguro tradicionais. Por exemplo, a equação contempla dotes puros (*endowments*) de maturidade arbitrária e outros benefícios de natureza análoga. Uma renda (temporária ou vitalícia) pode ser representada por combinações de cash flows abrangidos pelo termo valor esperado em (11.2.1), com diferentes maturidades (datas de pagamento). Recordamos que, da mesma forma que uma obrigação clássica pode ser decomposta (*stripped*)

### 11.3. Estrutura do modelo de avaliação

---

num conjunto de obrigações de cupão zero, uma renda pode ser decomposta numa sequência de dotes puros.

Apesar das semelhanças na análise, devemos ter em conta que a interpretação do evento de incumprimento deve, neste caso, ser invertida. Com efeito, enquanto que numa renda vitalícia os pagamentos cessam quando a pessoa segura (o beneficiário) morre, numa obrigação os cupões deixam de ser pagos quando o emitente (entidade responsável pela liquidação dos cupões e amortização do principal) entra em incumprimento.<sup>1</sup>

A analogia que é feita entre um dote puro (renda) e uma obrigação de cupão zero (obrigação clássica) com risco de crédito comporta algumas vantagens. Em particular, ela facilita a implementação prática de uma abordagem onde se conjugam, simultaneamente, factores de risco demográficos e financeiros. A metodologia de avaliação resultante concede maior riqueza teórica a algumas técnicas actuariais tradicionais, ao mesmo tempo que nos catapulta para a análise de problemas mais complexos, como sejam o *pricing* de garantias e opções incluídas em contratos de seguro. Por exemplo, a avaliação das denominadas *guaranteed annuity options* (GAO's), opções que concedem ao detentor da apólice o direito de converter no futuro um determinado valor acumulado numa renda vitalícia a uma dada taxa garantida, requer o exame conjunto das dinâmicas da taxa de juro e da mortalidade.

Assinale-se ainda que muitos dos conceitos adoptados na avaliação de activos financeiros já estão, de certo modo, incorporados na prática actuarial, o que pode facilitar a preparação da documentação financeira segundo as normas definidas pelos IASB.<sup>2</sup>

### 11.3 Estrutura do modelo de avaliação

Nesta secção descrevemos, separadamente, as fontes de incerteza que determinam o valor dos contratos de seguro: a incerteza financeira e a intensidade de mortalidade. O objectivo é combiná-las de modo a caracterizar adequadamente a dinâmica inerente a um contrato de seguros. A discussão que se segue é parcialmente baseada em Dahl (2004), Biffis e Millossovich (2004, 2006) e Biffis (2005).

#### 11.3.1 Mercados financeiros

Considere-se um mercado financeiro composto por apenas dois activos: um activo com risco  $S_t$  e um activo sem risco (*money market account*)  $B_t$ . De acordo com os postulados

---

<sup>1</sup>A hipótese de incumprimento por parte da companhia de seguros não é concebida nesta análise.

<sup>2</sup>Por exemplo, o conceito de mudança de medida de probabilidade está essencialmente associado à escolha de bases técnicas adequadas.

teóricos habituais, assumimos que o mercado se rege pelas seguintes condições: (i) o mercado é completamente líquido, i.e., é sempre possível comprar ou vender qualquer quantidade de ambos os activos, (ii) não existem custos de transacção, (iii) o *spread bid-ask* é nulo e (iv) são admitidas posições curtas em ambos os activos.<sup>3</sup>

Seja  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  um espaço de probabilidade fechado fixo. A dinâmica dos preços dos activos é definida por

$$dS_t = \alpha^s(t, S_t) S_t dt + \sigma^s(t, S_t) S_t dW_t, \quad S_0 > 0 \quad (11.3.1)$$

$$dB_t = r_t B_t dt, \quad B_0 = 1, \quad (11.3.2)$$

onde  $B_t$  simboliza o valor capitalizado no momento  $t$  de um depósito inicial ( $t = 0$ ) de uma unidade monetária, remunerado a uma taxa de juro sem risco de curto prazo  $r$ , e  $(W_t)_{0 \leq t \leq T}$  é um processo de Wiener standard.

Na ausência de oportunidades de arbitragem, existe uma medida de probabilidade equivalente (*equivalent martingale measure*)  $\mathbb{Q}$ , com base na qual o processo  $B^{-1}S$  é uma martingala (Duffie, 1996). Admitamos que o activo com risco gera dividendos e denotemos por  $D$  o seu processo de acumulação. Neste caso, a dinâmica de  $S_t$  é definida por

$$S_t = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T r_s ds} S_T + \int_t^T e^{-\int_t^u r_s ds} dD_u \middle| \mathcal{F}_t \right]. \quad (11.3.3)$$

Suponhamos agora que a dinâmica dos preços dos activos é ditada por um processo estocástico do tipo afim (com saltos)  $Y$  em  $\mathbb{R}^k$ . Em termos mais precisos, assumimos que  $r$  é expressa mediante  $r_t = r(Y_t)$ , onde  $Y_t$  satisfaz uma equação do tipo (9.2.10) e as condições aí expressas e  $r(\cdot)$  é uma função real do tipo afim  $r(x) = \rho_0 + \rho_1 \cdot x$ , com  $\rho_0 \in \mathbb{R}$  e  $\rho_1 \in \mathbb{R}^k$ . Em resultado, a estrutura temporal de taxas de juro pode ser determinada em termos dos preços de obrigações de cupão zero  $B(t, T)$ ,  $T \geq t \geq 0$ ,

$$B(t, T) = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T r(Y_s) ds} \middle| \mathcal{F}_t \right] = e^{\alpha(t, T) + \beta(t, T) \cdot Y_t} \quad (11.3.4)$$

e assume, como se observa, uma estrutura exponencial de uma função do tipo afim.

### 11.3.2 Mortalidade

Concentremo-nos na intensidade de mortalidade de uma pessoa segura pertencente a um grupo homogéneo (i.e., integrando um grupo de pessoas seguras com a mesma idade e estado de saúde, cuja vida residual se pode assumir que tem uma distribuição idêntica).

<sup>3</sup>Admite-se igualmente que os mercados financeiros podem ser incompletos.

Para uma dada base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , considere-se uma pessoa segura com idade  $x$  no momento 0 (data de realização do contrato). A duração aleatória da sua vida residual é modelada em termos de um tempo de paragem  $\tau_x$  admitindo uma intensidade  $\mu$ . À semelhança do que fizemos no Capítulo 9, assumimos que  $\mu = \mu(X_t)$ , onde  $X_t$  é um processo do tipo afim (com saltos) com dinâmica ditada por (9.2.10) e  $\mu(\cdot)$  é uma função do tipo afim. Na prática, a intensidade pode ser descrita por qualquer um dos modelos analisados na Secções 9.4 e 9.5 do Capítulo 9.

#### 11.3.3 Modelo combinado

Considere-se uma base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  suficientemente abrangente para suportar os processos  $Y$  e  $X$ . De forma natural, assumimos que as variáveis de estado que governam os preços dos activos financeiros ( $Y$  em  $\mathbb{R}^k$ ) são independentes das variáveis de estado que influenciam a dinâmica da intensidade de mortalidade ( $X$  em  $\mathbb{R}^d$ ). Trata-se, como mencionámos no capítulo anterior, de uma hipótese conveniente, habitualmente aceite na literatura, mas que é contestada por alguns autores (e.g., Miltersen e Persson (2005) e Bauer e Russ (2006)).

A nossa análise incide sobre uma pessoa segura (representativa) com idade  $x$  no momento 0, com duração de vida residual descrita por um tempo de paragem  $\tau_x$  admitindo uma intensidade estocástica  $\mu$ .

O fluxo de informação disponível no momento  $t \geq 0$  é representado pela filtragem  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ , que encerra elementos sobre a evolução de todas as variáveis de estado até ao momento  $t$  e informação sobre se o óbito da pessoa segura ocorreu entretanto. Seja  $Z = (X, Y)$  o processo conjunto das variáveis de estado em  $\mathbb{R}^{k+d}$ . Como  $X$  e  $Y$  são processos do tipo afim independentes, o processo  $Z$  também o é.

Para efeitos de avaliação das responsabilidades pelo seu valor justo (*fair value*, nos termos da definição e metodologia sugeridas pelo IASB), considera-se que existem mercados secundários onde as (res)seguradoras podem transaccionar carteiras de apólices, de modo a que seja possível tomar posições curtas ou longas em contratos de seguro.

Dependendo do tipo de contrato a avaliar, admite-se que é possível transaccionar continuamente em mercado secundário um conjunto de contratos de seguro básicos, que representam os activos “primitivos” usados no *pricing* segundo argumentos de não arbitragem.<sup>4</sup>

Nos casos em que a avaliação se destina a calcular o valor das reservas financeiras ou ao *pricing* de produtos em mercado primário a situação é mais complexa, na medida

---

<sup>4</sup>Por exemplo, na avaliação de rendas vitalícias admite-se, implicitamente, que os dotes puros de diferentes maturidades podem ser considerados activos primitivos.

## 11.4. Avaliação de contratos de seguro do ramo vida

---

em que nos referimos a mercados onde as companhias de seguros assumem, regra geral, posições curtas em contratos de seguro, ao passo que as pessoas seguras apenas tomam posições longas.<sup>5</sup>

A um outro nível, a transposição para um nível agregado (i.e., de uma carteira) dos argumentos de não arbitragem usados no *pricing* de contratos individuais só pode ser considerada como uma aproximação, uma vez que cada apólice se refere apenas a uma determinada pessoa segura. De facto, os argumentos habituais, envolvendo cobertura perfeita e carteiras replicantes, só se aplicam a cada uma das apólices consideradas de *per si*. Qualquer extensão dos resultados a carteiras de apólices deve ser entendida como uma aproximação, cujo grau de precisão dependerá naturalmente da dimensão da carteira e do grau de homogeneidade das pessoas que integram a população.

### 11.4 Avaliação de contratos de seguro do ramo vida

Nesta secção analisamos a avaliação de contratos de seguro com cash flows contingentes à sobrevivência ou morte do indivíduo. Para tal, exploramos a já mencionada analogia entre os conceitos de intensidade de incumprimento e de intensidade de mortalidade. Designemos por  $R_x(t)$  a taxa de desconto actuarial que resulta da soma da taxa de juro com a intensidade de mortalidade, i.e.,

$$R_x(t) = r_t + \mu_x(t), \quad (11.4.1)$$

onde notamos que  $R_x(t)$  é simultaneamente dependente da idade e do tempo cronológico.

Suponhamos que existe uma medida de probabilidade equivalente  $\mathbb{Q}$ , com base na qual todos os activos financeiros são considerados martingalas após desconto usando a *money market account* como numerário. Com base em  $\mathbb{Q}$ , assume-se que o tempo de paragem  $\tau_x$  é duplamente estocástico e condicionado por  $\mathbb{Q}^X \subset \mathbb{F}$  com intensidade  $\mu_x$ . Com base nos resultados de Lando (1998), Duffie e Singleton (1999) e Jamshidian (2004), entre outros, a avaliação de um activo contingente à sobrevivência do indivíduo reduz-se à avaliação de um activo sem risco de incumprimento. Tal significa que, usando argumentos de não arbitragem, os preços dos contratos de seguro são martingalas- $\mathbb{Q}$  após desconto usando a taxa de desconto actuarial (*money market account* ajustada pelo risco de mortalidade) como numerário.

---

<sup>5</sup> Assinale-se, todavia, que as restrições no lado das companhias de seguros podem ser contornadas se admitirmos que existe um mercado de resseguro ilimitado, permitindo assim uma tomada de posições longas nos contratos vendidos. Outra solução passa pela exploração das chamadas “coberturas naturais”, um tema a que regressaremos na Secção 11.5.

Conforme veremos em seguida, a avaliação de contratos de seguro num contexto estocástico assenta em dois activos base (ou primitivos). No primeiro caso, trata-se de um activo que paga uma determinada quantia fixa em caso de sobrevivência da pessoa segura num dado momento futuro. No segundo, trata-se de um activo que paga uma quantia fixa em caso de morte da pessoa segura num determinado período de tempo. A combinação destes dois activos base permite-nos posteriormente alargar os resultados em termos de avaliação a diferentes tipos de contratos de seguro (apólices tradicionais, do tipo *unit-link* ou indexadas, contratos com opções incorporadas, etc.). Um aspecto crucial na avaliação refere-se à determinação do *prémio de risco de mortalidade*, problema que nos remete para a discussão em torno da aplicação de medidas de probabilidade realistas ou *risk-neutral* no contexto actuarial. Os resultados apresentados nas secções seguintes são parcialmente baseados Biffis e Millossovich (2004, 2006), Biffis (2005) e Ballotta e Haberman (2006).

#### 11.4.1 Activos base

Consideremos em primeiro lugar a determinação do valor justo dos activos base envolvidos na avaliação de contratos de seguro. Estes dois activos têm uma matriz de *payoffs* simples, que se resume, no primeiro caso, a um determinado cash flow num dado momento futuro, em caso de sobrevivência e, no segundo, a um dado cash flow caso o óbito seja registado num determinado período de tempo.

O valor justo apresentado corresponde ao prémio puro estimado considerando apenas os factores de risco financeiro e demográfico. Não inclui, por isso, quaisquer comissões ou cargas adicionais cobradas pela seguradora para fazer face a encargos administrativos ou financeiros. A estrutura de avaliação usada foi descrita na secção anterior.

Designemos por  $SB_t(C_T; T)$  o valor justo, calculado no momento  $t$ , de um benefício conhecido de montante  $C_T$  a receber no momento  $T$  em caso de sobrevivência, com  $0 \leq t \leq T$ . O seu valor é dado por

$$\begin{aligned} SB_t(C_T; T) &= \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T r(Y_u) du} \mathbb{I}_{\{\tau_x > t\}} C_T \middle| \mathcal{F}_t \right] \\ &= \mathbb{I}_{\{\tau_x > t\}} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T (r(Y_u) + \mu(X_u)) du} C_T \middle| \mathcal{G}_t \right], \end{aligned} \quad (11.4.2)$$

onde o valor esperado é calculado com respeito a uma medida de probabilidade *risk-neutral*  $\mathbb{Q}$ .

Em particular, se assumirmos a independência entre os mercados financeiros e o comportamento da mortalidade e explorarmos a independência entre as filtragens  $\mathbb{G}^X =$

$(\mathcal{G}_t^X)_{t \geq 0}$  e  $\mathbb{G}^Y = (\mathcal{G}_t^Y)_{t \geq 0}$ , com  $\mathcal{G}_t^X = \sigma(X_s : 0 \leq s \leq t)$  e  $\mathcal{G}_t^Y = \sigma(Y_s : 0 \leq s \leq t)$ , podemos escrever

$$SB_t(C_T; T) = \mathbb{I}_{\{\tau_x > t\}} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T r(Y_u) du} C_T \middle| \mathcal{G}_t^Y \right] \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T \mu(X_u) du} \middle| \mathcal{G}_t^X \right]. \quad (11.4.3)$$

O valor do contrato  $SB_t(C_T; T)$  é nulo no intervalo  $\{\tau_x \leq t\}$ . Restringindo-nos ao intervalo  $\{\tau_x > t\}$ , a natureza duplamente estocástica do tempo de paragem permite-nos escrever (Biffis, 2005)

$$SB_t(C_T; T) = \mathbb{I}_{\{\tau_x > t\}} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T (r_u + \mu_u) du} C_T \middle| \mathcal{F}_t \right] \quad (11.4.4)$$

Representemos agora por  $DB_t(C_{\tau_x}; T)$  o valor justo, calculado no momento  $t$ , de um benefício em caso de morte de montante  $C_{\tau_x}$ , pagável caso o beneficiário da apólice faleça antes do momento  $T$ , com  $0 \leq t \leq T$ . O seu valor é dado por

$$\begin{aligned} DB_t(C_{\tau_x}; T) &= \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{\tau_x} r(Y_u) du} C_{\tau_x} \mathbb{I}_{\{t < \tau_x \leq T\}} \middle| \mathcal{F}_t \right] \\ &= \mathbb{I}_{\{\tau_x > t\}} \int_t^T \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^s (r(Y_u) + \mu(X_u)) du} \mu(X_s) C_s \middle| \mathcal{G}_t \right] ds. \end{aligned} \quad (11.4.5)$$

Se assumirmos, em particular, a independência entre os processos  $X$  e  $Y$  obtemos

$$DB_t(C_{\tau_x}; T) = \mathbb{I}_{\{\tau_x > t\}} \int_t^T \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^s r(Y_u) du} C_s \middle| \mathcal{G}_t^Y \right] \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^s \mu(X_u) du} \mu(X_s) \middle| \mathcal{G}_t^X \right] ds. \quad (11.4.6)$$

Admitamos agora que num contexto *risk-neutral*  $r$ ,  $\mu$ ,  $X$  e  $Y$  são processos do tipo afim com respeito a  $\mathbb{F}$ , nos termos do descrito nas secções anteriores. Considerando que  $X$  e  $Y$  são independentes, o processo conjunto  $Z = (X, Y)$  assume, também ele, uma estrutura do tipo afim. No caso de os benefícios implícitos nos contratos serem funções de  $Z$  cobertos pela transformada (9.2.18), as expressões para o valor dos contratos ganham uma flexibilidade analítica acrescida admitindo, nalguns casos, soluções fechadas.

### 11.4.2 Contratos de seguro do ramo vida

A combinação adequada dos activos base definidos na secção anterior permite-nos derivar resultados para um número significativo de contratos de seguro do tipo clássico (i.e., sem bónus), do tipo *unit-link* e indexados. No primeiro caso, referimo-nos a contratos em que o benefício  $C$  é determinístico. No segundo caso, estamos perante contratos

em que o benefício é aleatório, dependente da evolução de um fundo de referência ou de um índice. Refira-se ainda que os processos estocásticos que descrevem os benefícios podem depender do vector de variáveis de estado  $Z$ , o que possibilita lidar com contratos de seguro diferidos e com benefícios contingentes à evolução da mortalidade. Este é, por exemplo, o caso de seguros com retornos ligados, de alguma forma, à evolução de indicadores demográficos.

Apresentamos, de seguida, uma visão geral sobre os contratos cobertos pela metodologia de avaliação, referindo-nos em todos os casos a acordos efectuados por indivíduos com idade  $x$  no momento  $t = 0$ . Sempre que possível, usamos a notação actuarial clássica para facilitar a sua comparação com as correspondentes fórmulas derivadas num contexto determinístico.

### Seguros de vida

Num seguro de vida, a companhia compromete-se a pagar ao beneficiário o valor nominal da apólice em caso de morte da pessoa segura. Os seguros de vida são chamados de *temporários* se as responsabilidades da seguradora cessam ao fim de  $T$  anos, caso o segurado esteja vivo, ou de *vida inteira* se as responsabilidades da seguradora se extinguem apenas com a morte do segurado, independentemente da idade em que esta ocorra.

Tomando como referência o activo base definido em (11.4.6), o valor justo (prémio puro pago de uma só vez) de um *seguro de vida temporário* garantindo um benefício  $C$  em caso de morte da pessoa segura no período  $(0, T]$ ,  $A_{x:T|}$ , é dado por

$$A_{x:T|} = DB_0(C_{\tau_x}; T). \quad (11.4.7)$$

A equação (11.4.7) é comparável com a sua homóloga em contexto actuarial clássico (determinístico), definida por (e.g., Gerber (1995), Bowers *et al.* (1997))

$$A_{x:T|} = \sum_{k=0}^{T-1} v^{k+1} {}_k p_x \cdot q_{x+k}, \quad (11.4.8)$$

onde  $v = (1 + i)^{-1}$ ,  $i$  denota uma *taxa de juro técnica* fixa e se assume, por simplificação, que o montante seguro é unitário (i.e.,  $C = 1$ ).

No caso de um *seguro de vida inteira*, apenas precisamos de estender a maturidade do contrato até à idade máxima admissível  $\omega$ , i.e.,

$$A_x = DB_0(C_{\tau_x}; \omega - x), \quad (11.4.9)$$

equação que compara com a sua homóloga em contexto actuarial clássico

$$A_x = \sum_{k=0}^{\omega-x-1} v^{k+1} {}_k p_x \cdot q_{x+k}. \quad (11.4.10)$$

### Dotes puros e seguros dotais

Um dote puro (*pure endowment*) envolve o pagamento a uma pessoa segura de uma determinada quantia numa data futura, desde que essa pessoa esteja viva para o receber. O valor justo (prémio único puro) de um dote puro de montante  $C$  a receber no momento  $T$  em caso de sobrevivência,  ${}_T E_x$ , é calculado, nos termos da definição (11.4.4), mediante

$${}_T E_x = SB_0(C_T; T). \quad (11.4.11)$$

As vantagens decorrentes do recurso a modelos do tipo afim para a modelação do comportamento dinâmico dos factores de risco tornam-se evidentes quando constatamos que, no caso de  $r$ ,  $\mu$ ,  $X$  e  $Y$  seguirem processos do tipo afim, a equação (11.4.11) se simplifica para dar lugar a

$${}_T E_x = e^{\mathcal{A}_r(0,T) + \mathcal{B}_r(0,T) \cdot Y_0} e^{\mathcal{A}_\mu(0,x,T) + \mathcal{B}_\mu(0,x,T) \cdot X_0}, \quad (11.4.12)$$

onde  $\mathcal{A}_r(\cdot)$ ,  $\mathcal{B}_r(\cdot)$ ,  $\mathcal{A}_\mu(\cdot)$  e  $\mathcal{B}_\mu(\cdot)$  são soluções de um sistema de EDO's para os processos  $r$  e  $\mu$ .

A expressão (11.4.11) compara, no contexto actuarial clássico, com (para  $C = 1$ )

$${}_T E_x = v^T \cdot {}_T p_x. \quad (11.4.13)$$

Um dote genérico (*endowment*) é um contrato que combina um dote puro de montante  $C'$  a receber ao fim de  $T$  anos com um seguro de vida temporário de valor  $C''$  no intervalo  $(0, T]$ . Se conjugarmos as equações (11.4.4) e (11.4.6), o seu valor justo é dado por

$$A_{x:T} (C', C'') = SB_0(C'_T; T) + DB_0(C''_T; T). \quad (11.4.14)$$

Como se observa, esta definição genérica inclui os dotes puros ( $C'' = 0$ ) e os denominados *seguros dotais* ( $C' = C''$ ) como casos particulares, assim como combinações mais complexas. Por exemplo, o valor justo de um chamado *dote duplo* (*double endowment*), i.e., de um dote em que o benefício em caso de sobrevivência é o dobro do benefício em caso de morte é obtido a partir de (11.4.14) considerando  $C' = 2C''$ .

### Rendas vitalícias

Uma *renda vitalícia* (*life annuity*) é um contrato mediante o qual a companhia de seguros se compromete a pagar a um beneficiário com idade inicial  $x$  um conjunto de rendimentos (uma renda) de forma periódica (mensalmente, trimestralmente, anualmente, etc.) enquanto este estiver vivo, cessando as obrigações da companhia apenas após a sua morte. Em contrapartida, o beneficiário entrega à companhia um ou mais prêmios, conforme o acordo estabelecido. Nos casos em que o contrato estabelece que as obrigações da seguradora terminam ao fim de um determinado número de anos, mesmo que o beneficiário esteja vivo, falamos de *rendas vitalícias temporárias*.

Uma renda vitalícia pode ser decomposta numa sequência de dotes puros com diferentes maturidades. Considere-se, por exemplo, uma renda vitalícia antecipada que paga um benefício de montante  $C$  no início de cada período  $t = 0, 1, 2, \dots$  em caso de sobrevivência do beneficiário da apólice. O seu valor justo, assumindo pagamentos discretos,<sup>6</sup> é dado por

$$\ddot{a}_x(C) = \sum_{h=0}^{\omega-x-1} h E_x = \sum_{h=0}^{\omega-x-1} SB_0(C_h; h) \quad (11.4.15)$$

ou, admitindo uma estrutura do tipo afim para os processos  $r$  e  $\mu$ , por

$$\ddot{a}_x(C) = \sum_{h=0}^{\omega-x-1} C_h e^{A_r(0,h)+B_r(0,h) \cdot Y_0} e^{A_\mu(0,x,h)+B_\mu(0,x,h) \cdot X_0}. \quad (11.4.16)$$

Estas expressões comparam com a sua equivalente actuarial clássica (com  $C = 1$ ), definida por

$$\ddot{a}_x(C) = \sum_{h=0}^{\omega-x-1} v^h \cdot h p_x. \quad (11.4.17)$$

Consideremos agora o caso de uma *renda vitalícia diferida*  $t$  anos, que envolve o pagamento contínuo de um benefício indexado em caso de sobrevivência do beneficiário da apólice. Suponhamos que o benefício é composto por um montante fixo e por uma parcela variável igual a uma percentagem  $\eta$  do nível de taxa de juro de curto prazo observada em cada momento no tempo.<sup>7</sup>

<sup>6</sup>Se admitirmos que a renda paga continuamente um benefício a uma taxa instantânea  $C$  durante a vida do beneficiário da apólice, o valor justo do contrato é dado por:

$$\ddot{a}_x(C) = \int_0^{\omega-x} SB_0(C_s; s) ds$$

<sup>7</sup>Para uma análise detalhada de rendas vitalícias com diferentes mecanismos de indexação do benefício

O valor justo deste contrato é dado por

$$SB_0(C_t; t) = \sum_{h=t}^{\omega-x-1} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^h r_u du} (1 + \eta r_h) \right] \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^h \mu_u du} \right]. \quad (11.4.18)$$

### Contratos do tipo *Unit-Link*

Considere-se um dote puro com maturidade em  $T$  do tipo *unit-link*, que paga um benefício dependente da evolução de um activo financeiro com processo de preço  $S = \exp(Y^{(i)})$  e taxa de retorno  $\zeta(t, Y_t) = q_0(t) + q_1(t) \cdot Y_t$ . O seu valor justo é dado por (Biffis, 2005)

$$SB_0(e^{Y_T^i}; T) = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^T r_u du} e^{Y_T} \right] \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^h \mu_u du} \right].$$

### Contratos com garantias de retorno

A inclusão de garantias de retorno em contratos de seguro é hoje prática comum em muitos mercados, em especial no caso de rendas vitalícias diferidas dependentes da evolução de um fundo de referência durante o período de diferimento. As garantias assumidas estipulam benefícios mínimos em caso de morte (*Guaranteed Minimum Death Benefits*, GMDB) ou em caso de sobrevivência (*Guaranteed Minimum Living Benefits*, GMLB).

As garantias previstas em caso de morte da pessoa segura contemplam, por exemplo, a devolução do valor dos prémios pagos, do valor acumulado da conta no momento do óbito ou do valor acumulado à taxa garantida. Por seu turno, as garantias definidas em caso de sobrevivência podem ser classificadas em três grandes grupos. As denominadas GMAB (*Guaranteed Minimum Accumulation Benefits*) proporcionam um benefício mínimo em caso de sobrevivência num dado momento futuro  $G_T^A$  e destinam-se a proteger os beneficiários das apólices de performances negativas nos mercados bolsistas.

No caso das GMIB (*Guaranteed Minimum Income Benefits*), os produtos estabelecem um valor mínimo  $G_T^I$  no momento  $T$ , com a diferença de que a garantia apenas se aplica no caso do beneficiário exercer a sua opção de conversão do valor garantido numa renda vitalícia a determinadas taxas fixadas no início do contrato.<sup>8</sup>

veja-se, por exemplo, Hardy (2003).

<sup>8</sup>Ou seja, para além da possibilidade standard de receber o valor acumulado (sem garantia) ou de converter o valor de mercado do fundo numa renda vitalícia usando as taxas em vigor no mercado no momento  $T$ , as GMIB oferecem ao detentor da apólice uma terceira opção, nomeadamente a conversão de um determinado montante garantido numa renda vitalícia às taxas observadas aquando da celebração do contrato ( $t = 0$ ).

O terceiro tipo de garantias em caso de sobrevivência corresponde às denominadas GMWB (*Guaranteed Minimum Withdrawal Benefits*). Nestas, é assegurado um montante mínimo para levantamentos (resgates) durante a vida dos contratos, desde que o montante retirado em cada ano e durante a duração da apólice permaneça dentro de determinados limites.

O interesse pela avaliação deste tipo de contratos é relativamente recente (veja-se, e.g., Milevsky e Posner (2001), Milevsky e Salisbury (2002, 2006), JPMorgan (2004), Lehman Brothers (2005), Schräger (2006)), e são escassas as referências em que a mortalidade é tratada num contexto estocástico.

Para ilustrar a avaliação deste tipo de contratos no contexto da metodologia acima descrita, consideremos um contrato que garante, daqui a  $n$  anos, o maior de dois valores: o valor de um investimento de referência  $S_t$  ou um determinado valor garantido  $K$ , em caso de sobrevivência da pessoa segura. Para um beneficiário com idade  $x$ , o valor no momento  $t$  da garantia incorporada neste contrato,  $G_x(t)$ , é dado por

$$\begin{aligned} G_x(t) &= \mathbb{E}_t^{\mathbb{Q}} \left[ \exp \left( - \int_0^n R_{x+u}(t+u) du \right) (K - S_{t+n})^+ \right] \\ &= {}_n p_x^{\mathbb{Q}}(t) \cdot B(t, t+n) \cdot \mathbb{E}_t^{\mathbb{Q}^{t+n}} [(K - S_{t+n})^+], \end{aligned} \quad (11.4.19)$$

onde  $(x)^+ = \max(x, 0)$ ,  $\mathbb{Q}^{t+n}$  denota uma medida de probabilidade prospectiva (no momento  $t+n$ ) e as demais variáveis assumem o significado habitual.

Como se observa, num contexto *risk-neutral* o valor desta garantia obtém-se descontando para o momento da avaliação o valor intrínseco de uma opção de venda, usando uma taxa de desconto que considera simultaneamente os factores de risco demográficos e financeiros.

### Guaranteed Annuity Options

As *Guaranteed Annuity Options* (GAO's), também conhecidas nos E.U.A. por *Guaranteed Annuity Rates* (GAR), diferem de uma GMIB pelo facto de não assegurarem que o valor acumulado pelo fundo de referência no fim do período de diferimento da renda vitalícia é, no mínimo, igual a um determinado valor pré-estabelecido.

Uma GAO concede ao detentor da apólice o direito, mas não a obrigação, de converter, no vencimento do contrato (opção do tipo Europeu), um dado montante seguro (capital acumulado, não garantido) numa renda vitalícia a uma taxa de juro correspondente ao maior de dois valores: à taxa de juro em vigor no mercado nesse momento ou a uma determinada taxa de conversão definida no início do contrato.

A opção é exercida na maturidade, contingente à sobrevivência do indivíduo, se o valor dos benefícios gerados pela renda vitalícia calculada à taxa garantida no início do contrato for superior ao montante seguro na maturidade. Por outras palavras, a opção (*put option*) é exercida se a taxa garantida no momento da realização do contrato for, na data de exercício (maturidade), superior à taxa de juro em vigor no mercado para rendas vitalícias. Na prática, o valor da opção depende não apenas do preço das rendas vitalícias negociadas no mercado (determinado pelas taxas de juro, níveis de mortalidade e cargas aplicadas), mas também das preferências e expectativas dos detentores das apólices, elementos não captados pelos preços de mercado.<sup>9</sup>

O problema complexo da avaliação de opções de conversão em renda vitalícia a taxa garantida é analisado na literatura por inúmeros autores (veja-se, e.g., Milevsky e Promislow (2001), Ballota e Haberman (2003, 2006), Boyle e Hardy (2003), Olivieri e Pitacco (2003a), Pelsser (2003), Wilkie *et al.* (2003), Biffis e Millosovich (2006), Schräger (2006)).

Formalmente, considere-se um indivíduo com idade  $x$  e vida residual  $T_x$ , que entra no contrato no momento  $t = 0$ . O montante do benefício disponível na maturidade, i.e., no momento  $T > 0$ , é dado pelo valor de um fundo de referência com processo de preço  $S$ , enquanto que a taxa de conversão garantida  $0 < r_x^G < 1$  é fixada no início do contrato. Designemos por  $r_{x,T}(t) \equiv B(t, T) / a_x^{T-t}(t)$  a taxa de juro *forward* implícita no preço de uma renda vitalícia com início em  $T$  para um indivíduo com idade  $x$ .

Com base nos resultados de Pelsser (2003), o valor no momento  $t$  de uma opção de venda GAO sobre uma renda vitalícia com data de exercício  $T$ , é calculado segundo

$$\begin{aligned} GAO(t, T, x) &= \mathbb{E}_t^{\mathbb{Q}} \left[ \exp \left( - \int_0^{T-t} R_{x+u}(t+u) du \right) a_{x+T-t}^{T-t}(T) (r_x^G - r_{x+T-t,T}(T))^+ \right] \\ &= a_x^{T-t}(t) \cdot \mathbb{E}_t^{\mathbb{Q}} \left[ (r_x^G - r_{x+T-t,T}(T))^+ \right], \end{aligned} \quad (11.4.20)$$

onde  $a_x^{T-t}(t)$  denota o preço de uma renda vitalícia com primeiro pagamento diferido  $T - t$  anos. Demonstra-se que, no caso em que se assume uma estrutura gaussiana para as taxas de juro e para a intensidade de mortalidade estocástica, o preço de uma GAO pode ser aproximado por uma solução analítica (Schräger, 2006).

<sup>9</sup>As GAO's assumiram uma relevância muito grande na segunda metade dos anos 90, em consequência da redução das taxas de juro e da diminuição, não antecipada, das taxas de mortalidade nas idades mais avançadas. Em resultado do efeito combinado destes dois movimentos, muitas companhias de seguros no Reino Unido que haviam vendido maciçamente apólices com GAO's experimentaram graves problemas de solvência, que requereram a constituição de reservas adicionais e conduziram, por exemplo, a mais antiga companhia de seguros de mundo (Equitable Life Assurance Society) a cancelar a venda de novos contratos em 2000.

### 11.4.3 *Risk-neutral* valuation e determinação do preço de mercado do risco de mortalidade

Nas secções anteriores argumentámos que a metodologia de avaliação preconizada assentava no cálculo de um valor esperado com respeito a uma medida de probabilidade *risk-neutral*  $\mathbb{Q}$ . Esta medida deve, naturalmente, incorporar as preferências dos indivíduos em relação ao nível de risco dos benefícios gerados pelos contratos de seguro, possibilitando a utilização dos habituais argumentos de não arbitragem e o desconto dos cash flows usando uma taxa de juro sem risco.

A avaliação dos contratos de seguro é condicionada pela evolução estocástica dos factores de risco financeiros e demográficos. Em relação aos factores de risco financeiros, a prática habitual consiste em inferir a medida  $\mathbb{Q}$ , restringida a  $\mathbb{G}_{\infty}^Y$ , a partir dos preços observados nos mercados para os activos financeiros.<sup>10</sup> A especificação da medida  $\mathbb{Q}$ , restringida a  $\mathbb{G}_{\infty}^X$ , revela-se, pelo contrário, bem mais difícil e complexa, uma vez que não existem (por enquanto) nos mercados activos contingentes à evolução da intensidade de mortalidade, com base nos quais seria possível calibrar o modelo de modo a ter em conta o risco de longevidade. Com efeito, a incompletude dos mercados impede, neste caso, o *pricing* dos produtos com base no custo da carteira replicante.

Nos capítulos anteriores assinalámos que existem, grosso modo, quatro fontes de risco de mortalidade que afectam os contratos de seguro: risco de base, risco não sistemático, risco sistemático (risco de longevidade) e risco de mortalidade catastrófica. Uma questão natural que resulta desta classificação é a de saber se todas estas fontes de risco estão efectivamente incorporadas no *pricing* dos produtos. Por outras palavras, e usando um conceito chave na avaliação de activos financeiros, importa saber se existem ou não *preços de mercado para o risco de mortalidade* associado a cada uma destas fontes.

A primeira observação a fazer a este respeito é a de que o risco de base apresenta uma natureza manifestamente não sistemática, na medida em que depende fundamentalmente das características da população segura incluída na carteira. O mesmo se pode afirmar em relação ao risco de mortalidade catastrófica, associado a eventos extremos imprevisíveis e com uma periodicidade relativamente pouco frequente. Por esta razão, admitimos que quer o risco de base, quer o risco catastrófico são riscos de mortalidade não sistemáticos, pelo que centraremos a nossa discussão apenas na distinção entre riscos de mortalidade sistemáticos e não sistemáticos.

Em relação ao risco não sistemático, não existe um consenso na literatura sobre se este deve ou não ser considerado no *pricing* dos produtos. Tratando-se de um risco

---

<sup>10</sup>Para uma revisão da literatura sobre esta matéria veja-se, por exemplo, Duffie (1996).

claramente diversificável (“*pooling risk*”), a teoria económica sustenta que ele não deve ser recompensado em termos de retorno adicional. Por exemplo, o modelo CCAPM (*Consumption CAPM*) não considera o risco diversificável (idiossincrático) para efeitos de determinação do preço (Cochrane, 2005). Outro exemplo é o dos contratos do tipo *unit-link*, onde o risco financeiro assume um papel preponderante (veja-se, e.g., Bacinello e Persson (2002) e referência nele citadas). Nestes, assume-se que a medida de probabilidade *risk-neutral* coincide com a medida real, baseada na estimação estatística.

O argumento por detrás desta escolha é o de que a seguradora está exposta a um risco de mortalidade que decorre, sobretudo, de flutuações aleatórias nas taxas de mortalidade em torno dos seus valores esperados, um risco que pode, no limite, ser eliminado aumentando o número de contratos similares na carteira. Se o risco é completamente diversificável, num mercado eficiente o prémio de risco associado deve ser nulo.

Se, pelo contrário, se considerasse justificável a inclusão de uma compensação pelo risco de mortalidade não sistemática, esta dependeria naturalmente da dimensão da carteira. Nesse caso, o prémio seria tanto maior quanto menor fosse o número de apólices na carteira. Seguindo este raciocínio, as grandes seguradoras estariam em condições de oferecer os seus produtos a preços mais competitivos do que as pequenas companhias conduzindo, a prazo, a um mercado dominado apenas por grandes *players*. Na prática, verificamos que no mercado de seguros operam um grande número de companhias, de dimensão muito variada, o que significa que, ou não existe qualquer prémio pelo risco de mortalidade não sistemático, ou todas as companhias têm uma dimensão suficientemente grande para negligenciar a importância deste tipo de risco.

Por outro lado, a diversificação do risco é uma das funções centrais da actividade seguradora, nomeadamente pelo uso da lei dos grandes números e pelo recurso a instrumentos de cobertura sofisticados. Neste sentido, exigir um prémio pelo risco de mortalidade não sistemática seria entendido como uma “recompensa” pelo mau desempenho dos gestores de carteira na redução e/ou eliminação deste tipo de risco. Em face destes argumentos, assumiremos daqui em diante que as companhias não exigem qualquer prémio pela assunção do risco de mortalidade não sistemática.

Pelo contrário, existe uma (quase<sup>11</sup>) unanimidade relativamente à necessidade de exigir uma compensação pelo risco de longevidade no *pricing* de activos contingentes à evolução da mortalidade. De facto, o risco de mortalidade sistemática é um risco não diversificável que a seguradora não consegue eliminar aumentando a dimensão da

---

<sup>11</sup>Ballotta e Haberman (2006) assumem que o mercado é completamente *risk-neutral* com respeito ao risco de mortalidade. Refira-se, no entanto, que a sua escolha é motivada sobretudo pelo facto de não conseguirem desenvolver uma forma adequada de incluir este risco no modelo, e não tanto pela convicção de que o risco de mortalidade sistemática não requer uma compensação.

carteira ou apelando à lei dos grandes números. Por exemplo, no mercado de rendas vitalícias o risco de flutuações aleatórias na mortalidade assume um papel secundário, quando comparado com a importância que os desvios sistemáticos face aos valores projectados representam para a rentabilidade do negócio. Neste caso, aumentar a dimensão da carteira não só não elimina o risco de mortalidade como aumenta a exposição da companhia.

A um outro nível, a modelação da mortalidade mediante processos estocásticos fornece um novo enquadramento para o risco de longevidade. Com efeito, neste contexto o risco de mortalidade sistemática refere-se ao risco associado a oscilações na intensidade  $\mu_x(t)$ , enquanto que o risco de mortalidade não sistemática respeita ao risco associado à aleatoriedade dos óbitos numa carteira com intensidade  $\mu_x(t)$  conhecida.

O risco de mortalidade sistemática não é negociado eficientemente nos mercados financeiros ou no mercado de resseguro. Em consequência, a medida de probabilidade  $\mathbb{Q}$  não pode ser inferida a partir dos preços dos activos como acontece, por exemplo, no caso das obrigações e outros activos financeiros. A questão que se coloca é, pois, a de definir a estrutura do prémio de risco de mortalidade a aplicar e a forma como este deve ser incorporado na avaliação, quer de contratos de seguro tradicionais, quer de derivados financeiros contingentes à evolução da longevidade. Não existe, a este respeito, um consenso entre os investigadores. Analisamos, de seguida, os principais métodos recentemente sugeridos na literatura actuarial e financeira para determinar o prémio de risco. A juntar a estes, desenvolvemos um novo método de estimação do prémio de risco, inspirado num dos princípios clássicos de determinação do prémio em contratos de seguro.

#### **Abordagem *risk-neutral***

A adopção de uma abordagem *risk-neutral* no *pricing* de contratos de seguro e na determinação do preço de mercado do risco de longevidade é defendida na literatura por inúmeros autores (veja-se, e.g., Milevsky e Promislow (2001), Dahl (2004), Dahl e Möller (2005), Miltersen e Persson (2005), Cairns *et al.* (2006a,b), Schräger (2006), Bauer e Russ (2006)). Esta posição é sustentada em princípios teóricos consolidados na teoria financeira, nos quais se estabelece que, mesmo num mercado incompleto, se o mercado como um todo exclui oportunidades de arbitragem, existirá pelo menos uma medida de probabilidade *risk-neutral*  $\mathbb{Q}$  que pode ser usada para calcular os preços teóricos dos activos. O problema reside, neste sentido, na identificação desta medida de *pricing risk-neutral*, dada a ausência de activos de referência com base nos quais calibrar o modelo.

Em termos formais, para um dado espaço de probabilidade filtrado  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , admitamos que  $\{W_t^\mu\}_{t \geq 0}$  é um  $\mathbb{P}$ -movimento Browniano com filtragem natural  $\{\mathcal{F}_t\}_{t \geq 0}$  e que  $\{\theta_t\}_{t \geq 0}$  é um processo satisfazendo a condição de Novikov<sup>12</sup>, i.e.,

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \theta_t^2 dt \right) \right] < \infty.$$

Considere-se a martingala  $\xi_t$ , definida segundo

$$\xi_t = \exp \left( - \int_0^t \theta(s, \mu_{x+s}) dW_s - \frac{1}{2} \int_0^t \theta^2(s, \mu_{x+s}) ds \right). \quad (11.4.21)$$

A medida de probabilidade equivalente  $\mathbb{Q}$  é definida por

$$\mathbb{Q}(A) = \int_A \xi_t(\omega) \mathbb{P}(d\omega). \quad (11.4.22)$$

Neste caso, segundo a medida de probabilidade  $\mathbb{Q}$  o processo  $\{W_{0 \leq t \leq T}^\mathbb{Q}\}$  definido por

$$W_t^{\mu, \mathbb{Q}} = W_t^\mu + \int_0^t \theta(s, \mu_{x+s}) ds \quad (11.4.23)$$

é um movimento browniano standard.

O processo  $\theta$  representa o preço de mercado do risco de mortalidade.<sup>13</sup> Note-se que, como o mercado combinado (financeiro e de seguros) é incompleto, a medida de probabilidade equivalente não é única. Neste sentido, supomos que os cálculos são realizados com respeito a uma dada escolha de  $\mathbb{Q}$ .

Um exemplo desta abordagem é sugerido por Cairns *et al.* (2006b), onde se assume que o preço de mercado do risco de longevidade é constante. Os autores adoptam um modelo bifactorial em tempo discreto, onde o preço de mercado do risco de longevidade é determinado por um parâmetro de salto, calibrado a partir do prémio implícito no preço de subscrição da emissão de *longevity bonds* estruturada (embora finalmente não emitida) pelo consórcio EIB/BNP em Novembro de 2004.<sup>14</sup>

O modelo de Schräger (2006) fornece uma solução alternativa, permitindo nomeada-

<sup>12</sup>Para uma análise mais detalhada deste tema veja-se, por exemplo, Musiela e Rutkowski (1997), Karatzas e Shreve (1991) e Braumann (2005).

<sup>13</sup>Nos casos em que existe dependência entre os factores de risco financeiros e demográficos (e.g., Miltersen e Persson (2005)), o prémio de risco global pode ser decomposto em três parcelas: o prémio de risco de taxa de juro, associado ao factor de desconto, o prémio de risco resultante da correlação entre as taxas de juro e a mortalidade subjacente, e um prémio relativo ao risco de mortalidade específica.

<sup>14</sup>A hipótese de prémio de risco constante é, naturalmente, uma hipótese simplificadora, aceitável apenas num quadro em que não existem nos mercados activos contingentes à evolução da longevidade.

mente que os preços de mercado do risco de longevidade sejam escolhidos de maneira a que a intensidade  $\mu_x(t)$  (e respectivas probabilidades de sobrevivência), derivada com base numa medida *risk-neutral*, seja “prudente” quando comparada com a derivada num contexto real, i.e.,  ${}_{T-t}p_{x+t}^{\mathbb{Q}}(t) > {}_{T-t}p_{x+t}^{\mathbb{P}}(t)$ . Contudo, o autor não especifica um método para calibrar o modelo.

Em alternativa, Biffis (2005) sugere a utilização de taxas de desconto ajustadas no âmbito da metodologia *embedded value method*, habitualmente usada para avaliar companhias de seguros.<sup>15</sup> Refira-se que esta taxa ajustada incorpora, habitualmente, o custo do capital regulamentar, as responsabilidades com impostos e prémios de risco para fazer face a todos os riscos não diversificáveis subjacentes.

No futuro, o previsível desenvolvimento dos mercados de activos contingentes à evolução da longevidade proporcionará informação suficiente para calibrar os modelos e para a avaliar a adequação de hipóteses alternativas acerca do preço de mercado do risco de longevidade. Entretanto, a escolha de  $\mathbb{Q}$  continuará a depender de hipóteses subjetivas, condicionadas em grande parte por julgamentos em relação à sua razoabilidade económica.

A definição de uma nova medida de probabilidade tem consequências sobre a dinâmica da intensidade de mortalidade. Por exemplo, a versão *risk-neutral* do modelo de Cox-Ingersoll-Ross pode ser escrita como

$$d\mu_x(t) = \left( \beta^\mu(t) - \gamma^\mu(t) \mu_x(t) - \rho^\mu(t) \sqrt{\mu_x(t)} \theta(s, \mu_x(t)) \right) dt + \rho^\mu(t) \sqrt{\mu_x(t)} dW_t^{\mu, \mathbb{Q}}, \quad (11.4.24)$$

onde consideramos um preço de mercado do risco do tipo

$$\theta(s, \mu_x(t)) = \delta(t) \sqrt{\mu_x(t)} + \frac{\delta^*(t)}{\sqrt{\mu_x(t)}}, \quad (11.4.25)$$

onde  $\delta(t)$  e  $\delta^*(t)$  são funções determinísticas.

Substituindo na equação acima, a dinâmica de  $\mu$  num contexto *risk-neutral* é dada por

$$d\mu_x(t) = \left( \beta^{\mu, \mathbb{Q}}(t) - \gamma^{\mu, \mathbb{Q}}(t) \mu_x(t) \right) dt + \rho^{\mu, \mathbb{Q}}(t) \sqrt{\mu_x(t)} dW_t^{\mu, \mathbb{Q}}, \quad (11.4.26)$$

<sup>15</sup>O valor intrínseco (*embedded value*) de uma carteira de seguros corresponde à soma do valor dos activos líquidos que suportam a carteira com o valor esperado descontado dos lucros distribuíveis (ou *free cash flows*) provenientes das apólices já emitidas, calculados usando hipóteses realistas para as responsabilidades futuras.

onde

$$\begin{aligned}\beta^{\mu, \mathbb{Q}}(t) &= \beta^{\mu}(t) + \rho^{\mu}(t) \delta^*(t) \\ \gamma^{\mu, \mathbb{Q}}(t) &= \gamma^{\mu}(t) - \rho^{\mu}(t) \delta(t) \\ \rho^{\mu, \mathbb{Q}}(t) &= \rho^{\mu}(t).\end{aligned}$$

Numa abordagem distinta, Beelders e Colarossi (2004) avaliam um activo com fluxos de caixa contingentes à evolução da mortalidade (*mortality bond*) usando a teoria dos valores extremos. Os autores assumem, em particular, uma distribuição de Pareto para a mortalidade.

### Adopção de princípios clássicos de determinação do prémio

Nesta secção, desenvolvemos uma nova metodologia de estimação do prémio de risco de longevidade para uso na avaliação de contratos de seguro com benefícios de longo prazo em caso de vida. A solução é inspirada em princípios consagrados na teoria actuarial e na prática habitual das companhias de seguros.

A determinação do prémio comercial nos contratos de seguro baseia-se, regra geral, na adição de cargas adicionais (*loadings*) ao prémio puro da apólice, determinado com base na avaliação actuarial dos factores de risco subjacentes. Essas cargas destinam-se a fazer face a riscos não antecipados, despesas administrativas, impostos ou ao lucro dos accionistas. No problema em apreço, a aplicação directa desta regra prática exigiria a determinação da compensação (monetária ou outra) a exigir ao beneficiário da apólice pela transferência do risco de longevidade para a companhia. Uma solução *ad-hoc* consiste em aplicar um factor multiplicativo (superior à unidade) ao prémio puro para chegar ao prémio comercial.

No nosso caso, optámos por uma solução mais consistente, que envolve a estimativa das probabilidades de sobrevivência do beneficiário da apólice incluindo uma compensação (*risk loading*) pelo risco de longevidade, ou seja, contemplando um preço pelo risco de mortalidade sistemática. A solução encontrada para ajustar as probabilidades de sobrevivência assenta na utilização de um princípio clássico de determinação do prémio em contratos de seguro.

Os princípios teóricos de determinação do prémio (*premium principles*) estabelecem, grosso modo, regras explícitas para a fixação dos prémios associados a cada factor de risco, estipulando quer os montantes monetários a exigir ao segurado em contrapartida pela assunção dos riscos, quer os montantes adicionais impostos (*risk loadings*) para

precaver desvios imprevistos nas perdas esperadas.

De entre o vasto catálogo de *premium principles* existente, a nossa escolha recaiu sobre o tradicional *Standard Deviation Premium Principle* (SDPP).<sup>16</sup> Segundo este princípio, o prêmio é determinado adicionando ao valor esperado (esperança matemática) dos riscos cobertos pela apólice uma compensação proporcional ao desvio padrão desses mesmos riscos. Em termos formais,

$$\mathcal{H}[X] = \mathbb{E}^{\mathbb{P}}(X) + \beta \sqrt{\text{Var}_{\mathbb{P}}(X)}, \quad \beta > 0, \quad (11.4.27)$$

onde  $\mathcal{H}[X]$  representa a regra de determinação do prêmio,  $X$  denota o factor de risco subjacente e  $\beta$  define a carga adicional (*risk loading*) exigida. Os operadores valor esperado e variância são calculados com respeito à medida de probabilidade real  $\mathbb{P}$ .

O princípio do desvio-padrão apresenta algumas propriedades interessantes: independência,<sup>17</sup> *risk loading* (i.e.,  $\mathcal{H}[X] \geq \mathbb{E}^{\mathbb{P}}(X)$ ,  $\forall X$ ), invariância de escala (i.e.,  $\mathcal{H}[bX] = b\mathcal{H}[X]$ ,  $\forall X$ ,  $b \geq 0$ ), encontrando-se entre os mais usados na prática, em particular na tarifação de contratos de seguro do ramo não-vida.

Este método aproxima o prêmio obtido com base na aplicação do tradicional *princípio da utilidade equivalente*. Segundo este, o prêmio  $\mathcal{H}[X]$  deve ser fixado de modo a que a seguradora seja indiferente entre aceitar ou não cobrir o risco (i.e.,  $\mathcal{H}[X]$  constitui a solução da equação  $\dot{u}(\omega) = \mathbb{E}[u(\omega - X + \mathcal{H}[X])]$ , onde  $u(\cdot)$  denota a função utilidade e  $\omega$  a riqueza inicial da seguradora). O próprio princípio de Wang, que analisaremos mais adiante, reduz-se ao princípio do desvio-padrão em muitos casos (Young, 2004). Schweizer (2001) e Möller (2001a,c) discutem formas de adaptar o princípio da variância e do desvio-padrão no *pricing* de riscos financeiros em contexto estocástico.

Partindo de (9.2.7), a aplicação directa do princípio (11.4.27) no cálculo de probabilidades de sobrevivência ajustadas pela inclusão de um prêmio pelo risco de longevidade conduz-nos à seguinte expressão geral

$${}_{T-t}p_{x+t}^*(t) = \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right] + \beta \left\{ \text{Var}_{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right] \right\}^{\frac{1}{2}}. \quad (11.4.28)$$

A equação (11.4.28) estabelece que a inclusão do prêmio de risco de longevidade na avaliação de contratos de seguro é feita calibrando a probabilidade de sobrevivência

---

<sup>16</sup> Para uma revisão detalhada deste tema veja-se, por exemplo, Landsmana e Sherris (2001) e Young (2004).

<sup>17</sup> Segundo esta propriedade  $\mathcal{H}[X]$  depende apenas da distribuição de probabilidade de  $X$ , i.e., depende apenas da probabilidade de ocorrência do evento segurável e da perda monetária resultante, e não da causa da perda monetária.

por um factor  $\beta$  proporcional à própria incerteza da estimativa (desvio-padrão), e não adicionando um montante ao prémio puro calculado usando bases técnicas clássicas. Dito de outro modo, o prémio de risco é incorporado directamente na estimativa do prémio puro.

Para ilustrar a aplicação prática deste método, admitamos que a intensidade  $\mu_{x+t}(t)$  é modelada através de um processo do tipo Cox-Ingersoll-Ross, i.e.,

$$d\mu_{x+t}(t) = a(\theta - \mu_{x+t}(t))dt + \sigma\sqrt{\mu_{x+t}(t)}dW(t), \quad \mu_{x+t}(0) = \bar{\mu}_x \quad (11.4.29)$$

Representando a probabilidade de sobrevivência por uma função do tipo afim

$$\mathbb{E}^{\mathbb{P}} \left[ \exp \left( - \int_t^T \mu_{x+s}(s) ds \right) \middle| \mathcal{F}_t \right] = e^{\mathcal{A}_1(t,x,T) + \mathcal{B}_1(t,x,T) \cdot \mu_{x+t}(t)}, \quad (11.4.30)$$

onde  $\mathcal{A}_1(\cdot)$  e  $\mathcal{B}_1(\cdot)$  constituem soluções de um sistema de EDO's nos termos do definido em (9.2.16) e (9.2.17), respectivamente.

A variância de  $T-t p_{x+t}(t)$  é dada por

$$\text{Var}_{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_t^T 2\mu_{x+s}(s) ds} \right] - \left\{ \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_t^T \mu_{x+s}(s) ds} \right] \right\}^2. \quad (11.4.31)$$

Considere-se agora uma nova intensidade de mortalidade  $\tilde{\mu}_{x+t}(t) = 2\mu_{x+t}(t)$ . Com base em (11.4.29), a sua dinâmica é determinada por

$$d\tilde{\mu}_{x+t}(t) = a(2\theta - \tilde{\mu}_{x+t}(t))dt + \sqrt{2}\sigma\sqrt{\tilde{\mu}_{x+t}(t)}dW(t). \quad (11.4.32)$$

Representando a probabilidade de sobrevivência calculada com base nesta nova intensidade em termos de uma função do tipo afim, i.e., por

$$\mathbb{E}^{\mathbb{P}} \left[ \exp \left( - \int_t^T \tilde{\mu}_{x+s}(s) ds \right) \middle| \mathcal{F}_t \right] = e^{\mathcal{A}_2(t,x,T) + \mathcal{B}_2(t,x,T) \cdot \tilde{\mu}_{x+t}(t)}. \quad (11.4.33)$$

Aplicando o Lemma de Itô e calculando as respectivas derivadas parciais, prova-se que a solução do problema admite a seguinte representação de Feynman-Kac

$$\begin{aligned} & \psi(\tau, \tilde{\mu}_{x+t}(t)) \left\{ -\dot{\mathcal{A}}_2(\tau) - \dot{\mathcal{B}}_2(\tau)\tilde{\mu}_{x+t}(t) + \mathcal{B}_2(\tau) [a(2\theta - \tilde{\mu}_{x+t}(t))] \right. \\ & \left. + \sigma^2 \tilde{\mu}_{x+t}(t) \mathcal{B}_2^2(\tau) - \tilde{\mu}_{x+t}(t) \right\} = 0. \end{aligned}$$

Dividindo os termos da equação por  $\psi(\tau, \tilde{\mu}_{x+t}(t))$  e arrumando as parcelas, obtemos

$$\left[ -\dot{\mathcal{B}}_2(\tau) - a\mathcal{B}_2(\tau) + \sigma^2\mathcal{B}_2^2(\tau) - 1 \right] \tilde{\mu}_x(t) + \left[ -\dot{\mathcal{A}}_2(\tau) + 2a\theta\mathcal{B}_2(\tau) \right] = 0,$$

onde  $\mathcal{A}_2(\tau)$  e  $\mathcal{B}_2(\tau)$  constituem soluções do seguinte sistema de EDO's

$$\dot{\mathcal{B}}_2(\tau) = -a\mathcal{B}_2(\tau) + \sigma^2\mathcal{B}_2^2(\tau) - 1 \quad (11.4.34)$$

$$\dot{\mathcal{A}}_2(\tau) = 2a\theta\mathcal{B}_2(\tau) \quad (11.4.35)$$

com condições terminais

$$\mathcal{B}_2(0) = 0, \quad \mathcal{A}_2(0) = 0. \quad (11.4.36)$$

A solução do problema (11.4.34)-(11.4.35)-(11.4.36) remete-nos para as seguintes expressões analíticas para  $\mathcal{A}_2(\tau)$  e  $\mathcal{B}_2(\tau)$

$$\mathcal{A}_2(\tau) = 2 \left\{ \frac{a\theta}{\phi}(T-t) - \frac{a\theta(\phi+\xi)}{\kappa\phi\xi} \ln \left( \frac{\phi+\xi}{\phi+\xi e^{-\kappa(T-t)}} \right) \right\} \quad (11.4.37)$$

$$\mathcal{B}_2(\tau) = \frac{1 - e^{-\kappa(T-t)}}{\phi + \xi e^{-\kappa(T-t)}}, \quad \text{com} \quad \begin{cases} \kappa = \sqrt{a^2 + 4\sigma^2} \\ \phi = -\frac{(a+\kappa)}{2} \\ \xi = \frac{(a-\kappa)}{2}. \end{cases} \quad (11.4.38)$$

Prova: A verificação das soluções (11.4.37) e (11.4.38) para  $\mathcal{A}_2(\tau)$  e  $\mathcal{B}_2(\tau)$  é semelhante à apresentada no Capítulo 9 para o modelo de Cox-Ingersoll e Ross. ■

Se substituirmos as expressões (11.4.30) e (11.4.33) em (11.4.28), obtemos a seguinte expressão para a probabilidade de sobrevivência ajustada

$$\begin{aligned} {}_{T-t}P_{x+t}^*(t) &= e^{\mathcal{A}_1(t,x,T) + \mathcal{B}_1(t,x,T) \cdot \mu_x(t)} \\ &+ \beta \left\{ e^{\mathcal{A}_2(t,x,T) + 2\mathcal{B}_2(t,x,T) \cdot \mu_x(t)} - e^{2\mathcal{A}_1(t,x,T) + 2\mathcal{B}_1(t,x,T) \cdot \mu_x(t)} \right\}^{\frac{1}{2}}, \end{aligned} \quad (11.4.39)$$

onde  $\mathcal{A}_1$ ,  $\mathcal{B}_1$ ,  $\mathcal{A}_2$  e  $\mathcal{B}_2$  são funções definidas em (9.4.27), (9.4.28), (11.4.37) e (11.4.38), respectivamente.

Como se observa, a aplicação do princípio (11.4.27) fornece uma solução simples para incluir o prémio de risco de longevidade directamente no *pricing* de contratos de seguro. A incerteza em torno da probabilidade  ${}_{T-t}p_{x+t}(t)$  (e implicitamente em torno da trajectória da intensidade  $\mu_x(t)$ ) é incorporada directamente na avaliação admitindo-se, no entanto, que o parâmetro  $\beta$  possa incluir julgamentos subjectivos sobre a dimensão do risco envolvido.

### Transformada de Wang

Uma solução alternativa para determinar o preço de mercado do risco de longevidade, preconizada por Lin e Cox (2005a,b, 2006), Dowd *et al.* (2006), Blake *et al.* (2006b), Denuit *et al.* (2007), entre outros, consiste em ajustar a distribuição de probabilidade recorrendo à transformada de Wang (Wang, 1996, 2000, 2002, 2003). Esta abordagem “corrige” a distribuição do índice de sobrevivência para gerar valores esperados ajustados pelo risco de longevidade (ou equivalentes certos), que podem então ser usados para descontar os cash flows futuros segundo uma taxa de juro sem risco.

A abordagem pela transformada de Wang difere substancialmente das anteriores. Com efeito, em vez de usar modelos estocásticos para determinar o preço de activos contingentes à evolução da longevidade, a metodologia usa os preços observados no mercado de rendas vitalícias para estimar o preço de mercado do risco de longevidade implícito nas cotações. Esta estimativa, expressa em termos de um parâmetro de salto da transformada de Wang, é então usada para derivar probabilidades de sobrevivência ajustadas.

Formalmente, as probabilidades ajustadas pela transformada de Wang são dadas por

$${}_t\tilde{q}_x^W = \Phi \left[ \Phi^{-1}({}_t\tilde{q}_x^{\text{be}}) - \lambda \right], \quad (11.4.40)$$

onde  ${}_t\tilde{q}_x = 1 - {}_t\tilde{p}_x$ ,  $\Phi(\cdot)$  denota a função de distribuição de probabilidade Normal estandardizada e  $\lambda$  é o parâmetro de ajustamento da transformada de Wang. Este parâmetro simboliza, neste caso, o preço de mercado do risco de longevidade. Para distinguir as probabilidades estimadas com relação a diferentes medidas, denotamos por  ${}_t\tilde{q}_x^{\text{be}}$  os valores estimados com base na informação mais recente e por  ${}_t\tilde{q}_x^W$  os respectivos valores ajustados pela transformada de Wang.

Como se depreende, a cada valor do parâmetro  $\lambda$  corresponde uma transformada diferente. Uma solução possível para derivar o valor “óptimo” de  $\lambda$  consiste, como referimos, em usar a informação contida nas cotações de mercado das rendas vitalícias. Em termos formais, conhecida a estrutura temporal de taxas de juro, a estimativa do prémio de risco corresponde ao valor do parâmetro  $\lambda$  que iguala o valor hipotético de uma renda vitalícia  $a_x(\lambda)$  que paga  $K$  unidades monetárias anualmente ao seu valor observado no mercado, i.e., que respeita a igualdade

$$a_x(\lambda) = K \sum_{t=1}^{\infty} {}_t\tilde{p}_x^W B(0, t) = K \sum_{t=1}^{\infty} \left[ 1 - \Phi \left( \Phi^{-1}({}_t\tilde{q}_x^{\text{be}}) - \lambda \right) \right] B(0, t), \quad (11.4.41)$$

#### 11.4. Avaliação de contratos de seguro do ramo vida

---

onde  $B(0, t)$  denota o valor actual ( $t = 0$ ) de uma obrigação de cupão zero com maturidade em  $t$ .

Em suma, a ideia-base é a de que, após a transformação, o preço teórico de um activo corresponde ao valor esperado descontado dos seus cash flows futuros, segundo uma taxa de juro sem risco e uma distribuição de probabilidade ajustada.

A aplicação da transformada de Wang na obtenção de estimativas do prémio de risco e na avaliação de activos com fluxos de caixa contingentes à evolução da longevidade é criticada pelo facto de carecer de uma fundamentação teórica sólida. Como assinala Pelsser (2004), a transformada de Wang constitui um dos inúmeros mecanismos arbitrários que é possível usar para corrigir a distribuição da probabilidade de sobrevivência, não fornecendo, por isso, uma solução universal para o *pricing* dos riscos nos mercados financeiros e de seguros. O autor demonstra ainda, considerando processos estocásticos genéricos, que a transformada de Wang não gera preços consistentes com a ausência de oportunidades de arbitragem.

Outra limitação apontada a esta abordagem está relacionada com o facto de ela não clarificar o modo como os preços dos activos ou as probabilidades de sobrevivência evoluem no tempo. Um outro aspecto que merece ser questionado diz respeito à natureza da relação entre os preços de rendas vitalícias e os preços de outros activos financeiros contingentes à evolução da longevidade. Com efeito, não é garantido que as probabilidades implícitas nos preços das rendas constituam um bom ponto de partida na avaliação de outros activos, dada a natureza distinta das carteiras.

Por último, e como assinalam Cairns *et al.* (2006a), na abordagem pela transformada de Wang não é clara a forma como as diferentes transformadas relativas a gerações e maturidades distintas se relacionam entre si, formando um corpo teórico (uma estrutural temporal) coerente.

Recentemente, Cox *et al.* (2006) e Wang (2006) propuseram a utilização de um método denominado *multivariate exponential tilting* para avaliar activos contingentes à evolução da longevidade. Este método de avaliação, adoptado em contextos de mercados incompletos, equivale à aplicação da transformada de Wang da função de distribuição acumulada de uma variável aleatória.

#### Ratio de Sharpe

Milevsky *et al.* (2005, 2006) desenvolvem uma nova abordagem para avaliar activos contingentes à evolução da longevidade. A abordagem, decalcada de soluções análogas adoptadas na avaliação de activos sujeitos a riscos de mercado, assume que a incerteza

subjacente pode ser decomposta numa componente idiossincrática (diversificável) e numa componente sistemática (não diversificável).

Em consonância, os autores defendem que as companhias que comercializam apólices de seguro e rendas devem ser compensadas, na parcela correspondente ao risco de mortalidade sistemática, sob a forma de um determinado *Ratio de Sharpe Instantâneo*  $\alpha(r_t, \mu_x, t)$  definido, para uma dada carteira de cobertura, por

$$\alpha(r_t, \mu_x, t) = \frac{\lim_{h \rightarrow 0} \frac{1}{h} (\mathbb{E} [\Pi_{t+h} | \mathcal{F}_t] - r\Pi_t)}{\sqrt{\lim_{h \rightarrow 0} \frac{1}{h} (\text{Var} [\Pi_{t+h} | \mathcal{F}_t])}}, \quad (11.4.42)$$

onde  $\Pi$  denota o valor da carteira de cobertura,  $r$  a taxa de juro instantânea (sem risco) e as demais variáveis assumem o significado habitual.

Em termos intuitivos, os autores assumem que as companhias de seguros escolhem um determinado objectivo em termos de excesso de retorno por unidade de volatilidade (desvio-padrão), definido por  $\alpha(r_t, \mu_x, t)$ , com base no qual estipulam o preço dos contratos. Dito de outro modo, os preços dos contratos devem ser tais que as companhias consigam alcançar os seus objectivos globais em termos de valor pré-determinado para o ratio de Sharpe instantâneo da carteira. Os autores defendem, nomeadamente, que o preço de mercado do risco de longevidade deve ser fixado de modo a que o ratio de Sharpe seja consistente com o valor observado noutras classes de activos na economia.

Um resultado-chave nesta abordagem é aquele que estabelece que nos casos em que a intensidade de mortalidade é estocástica, o valor financeiro dos activos contingentes à evolução da longevidade é maior do que o valor actual (descontado) dos seus cash flows futuros num contexto de mortalidade determinística.

Bayraktar e Young (2007) ampliaram recentemente esta metodologia assumindo, em particular, que as companhias de seguros avaliam de forma simultânea uma carteira constituída por dotes puros e seguros de vida, através da especificação do ratio de Sharpe instantâneo para uma carteira de cobertura óptima.

### Consumption CAPM

Webb e Friedberg (2006) preconizam uma solução alternativa para determinar o preço de mercado do risco de longevidade num contexto estocástico, assente no modelo CCAPM (*Consumption Capital Asset pricing Model*).<sup>18</sup> Recordamos que no modelo CCAPM o retorno esperado de um activo com risco deve igualar a taxa de juro sem risco acrescida de um prémio ou compensação pelo risco sistemático.

---

<sup>18</sup>Para uma revisão detalhada do modelo CCAPM veja-se, por exemplo, Mehra e Prescott (2003).

#### 11.4. Avaliação de contratos de seguro do ramo vida

---

No modelo CCAPM, a compensação pelo risco sistemático é determinada pela relação entre o retorno esperado do activo e a utilidade marginal do consumo. Formalmente,

$$\mathbb{E}_t(R_{e,t+1}) = R_{f,t+1} + \frac{\text{Cov}_t\{-U'(c_{t+1}), R_{e,t+1}\}}{\mathbb{E}_t(U'(c_{t+1}))}, \quad (11.4.43)$$

onde  $U'(c)$  representa a utilidade marginal do consumo,  $R_e$  o retorno esperado dos activos e  $R_f$  a taxa de juro sem risco. Com base nesta relação, o prémio de risco  $\mathbb{E}_t(R_{e,t+1}) - R_{f,t+1}$  pode ser facilmente calculado.

O *rationale* por detrás da equação (11.4.43) é simples. Quando o consumo diminui num determinado período, a utilidade do consumo  $U(c)$  decresce, enquanto que a utilidade marginal  $U'(c)$  aumenta relativamente a períodos em que o nível de consumo é mais elevado. Neste sentido, o consumo permitido por qualquer rendimento extra gerado pelo investimento em períodos de baixo consumo é, em termos relativos, mais valorizado.

Em consequência, os investidores exigem um prémio de risco pela detenção de activos cujos retornos estejam positivamente correlacionados com choques sobre o nível global de consumo, uma vez que esses activos geram fluxos de caixa maiores nos estados da natureza em que o consumo e a utilidade são elevados, mas onde a utilidade marginal do consumo é baixa. Inversamente, os investidores tomarão posições significativas em activos que proporcionam retornos elevados quando o nível de consumo é baixo (e a correspondente utilidade marginal elevada), e estão dispostos a adquiri-los mesmo no caso em que o retorno esperado é inferior à taxa de juro sem risco.

Webb e Friedberg (2006) sustentam, com base em evidência empírica para o mercado norte-americano, que a correlação contemporânea entre os choques sobre o consumo agregado e os choques sobre a mortalidade é negativa. Em consequência, alegam que os investidores devem estar dispostos a aceitar um desconto (prémio de risco negativo) pela aquisição de obrigações contingentes à evolução da longevidade.

O argumento chave é o de que a maior volatilidade nas taxas de mortalidade ocorre nas idades avançadas, onde se regista o maior número de óbitos. Admitamos que um número inesperadamente baixo de idosos morre num determinado ano. Em resposta, o rendimento nacional não deve aumentar significativamente, uma vez que o stock de capital é maioritariamente fixo no curto prazo e um número muito reduzido de pessoas idosas trabalha. Em resultado, o rendimento *per-capita* diminuirá, dado que o produto é relativamente constante e deve agora ser partilhado por um número de indivíduos maior que o esperado.

Por outro lado, as reduções na mortalidade podem implicar aumentos nas despesas com saúde, reduzindo desta forma o orçamento disponível para consumo. Por fim, e na

medida em que o aumento da longevidade for global um país não poderá financiar-se externamente para atenuar os efeitos do choque sobre o consumo.

### Estratégias de minimização do risco

Os contratos de seguro e outros activos contingentes à evolução da longevidade estão simultaneamente expostos a riscos financeiros e riscos de mortalidade. Como os mercados são incompletos, não é possível fazer o *pricing* e *hedging* dos contratos com base na avaliação da carteira replicante. Um dos métodos sugeridos na literatura para lidar com este problema envolve a adopção das chamadas estratégias (de cobertura) de minimização do risco. Estas estratégias assentam, essencialmente, na minimização da variância da carteira de responsabilidades da seguradora com respeito a uma determinada medida de probabilidade equivalente (veja-se, e.g., Möller (1998, 2001a,c, 2002) e Dahl e Möller (2005)).

Considere-se uma dada base estocástica  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  e um mercado financeiro constituído por uma obrigação de cupão zero (ZCB) com maturidade em  $T$  e um activo sem risco (*money market account*). Designemos por  $(X_t)_{0 < t < T} = B(t, T)$  o processo seguido pelo preço da ZCB. Uma estratégia de negociação (*trading strategy*) é um processo  $\varphi_t = (\xi_t, \eta_t)$ , onde  $\xi_t$  denota o número de ZCB detidas e  $\eta_t$  o valor actualizado do activo sem risco no momento  $t$ . O par  $\varphi_t = (\xi_t, \eta_t)$  representa a carteira detida no momento  $t$ . O processo (descontado) associado à estratégia  $\varphi$  é definido por

$$V_t(\varphi) = \xi_t X_t + \eta_t, \quad 0 \leq t \leq T. \quad (11.4.44)$$

Considere-se um acordo entre duas partes (e.g., uma companhia de seguros e uma pessoa segura), especificando determinados pagamentos simbolizados pelo processo  $A = (A_t)_{0 < t < T}$ , ou seja,  $A$  estabelece o valor líquido (actualizado) dos pagamentos efectuados pela companhia ao comprador da apólice.

Considere-se uma companhia cujas responsabilidades num contrato são definidas por  $A$  e que aplica a estratégia  $\varphi$ . Imediatamente após a celebração do contrato, a empresa paga  $A_0$  ao comprador e faz um investimento inicial  $V_0(\varphi)$ . O custo inicial associado a  $(A, \varphi)$ ,  $C_0(\varphi)$ , é dado por  $C_0(\varphi) = V_0(\varphi) + A_0$ . No intervalo  $(0, T]$ , a companhia suportará custos adicionais pelo que, em qualquer momento  $t$ , os custos acumulados do processo em  $[0, t]$ ,  $C_t(\varphi)$ , são definidos por

$$C_t(\varphi) = V_t(\varphi) - \int_0^t \xi_u dX_u + A_t, \quad (11.4.45)$$

onde  $V_t(\varphi)$  deve ser interpretado como o valor da carteira  $\varphi_t = (\xi_t, \eta_t)$  em  $t$  depois de deduzidos os pagamentos  $A_t$ . Em particular,  $V_T(\varphi)$  denota o valor da carteira  $\varphi_T$  após liquidação de todas as responsabilidades, pelo que uma restrição natural consiste em considerar estratégias que satisfazem  $V_T(\varphi) = 0$ , i.e., carteiras que asseguram o *matching* perfeito entre activos e responsabilidades no horizonte  $T$ .

Como base na equação (11.4.45), verificamos que  $C_t(\varphi)$  é igual ao valor actualizado da carteira, subtraído dos ganhos de *trading* e somado dos pagamentos líquidos efectuados pela seguradora.

Uma estratégia diz-se minimizadora do risco se, para qualquer  $t \in [0, T]$ , minimiza a quantidade

$$R_t(\varphi) = \mathbb{E}^{\mathbb{Q}} \left[ (C_T(\varphi) - C_t(\varphi))^2 \middle| \mathcal{F}_t \right] \quad (11.4.46)$$

entre todas as estratégias com o mesmo valor no momento  $T$ ,  $V_T(\varphi)$ . Dito de outro modo, é uma estratégia que minimiza o valor esperado do quadrado dos seus custos futuros. O processo  $R_t(\varphi)$  pode ser entendido como uma medida do risco remanescente da companhia.

Se admitirmos que  $V_T(\varphi) = 0$ , um custo inicial  $C_0(\varphi)$  e a prossecução de uma estratégia de negociação  $\varphi$  conduz a um valor de carteira nulo após liquidação de todas as responsabilidades. Neste sentido, se a estratégia é admissível, o único preço que garante a ausência de oportunidades de arbitragem (*fair price*) é  $C_0(\varphi)$ . Não existindo processos admissíveis, não é possível determinar um preço justo único, tornando-se necessário o recurso a estratégias de minimização do risco.

### Estratégias de média-variância

A aplicação dos princípios de média-variância (*mean-variance indifference pricing*) ou de *pricing via equivalente* certo na avaliação de contratos de seguro do ramo vida é sugerida por Schweizer (2001), Möller (2001b, 2002) e Dahl e Möller (2005), entre outros. Estes princípios constituem um exemplo adicional da utilização de métodos desenvolvidos em contexto de mercados incompletos no tratamento conjunto dos factores de risco financeiros e demográficos. Nesta secção, enunciamos brevemente os principais resultados obtidos neste domínio.

Nas estratégias de média-variância, o objectivo consiste simplesmente em encontrar a estratégia autofinanciadora  $\hat{\varphi}_t = (\hat{\xi}_t, \hat{\eta}_t)$  que minimiza

$$\mathbb{E}^{\mathbb{P}} \left[ (H - V_T(\varphi))^2 \right], \quad (11.4.47)$$

onde  $H$  denota as responsabilidades da companhia de seguros.

O objectivo central da estratégia é simplesmente o de aproximar o mais possível o valor das responsabilidades do valor terminal da carteira de cobertura  $\varphi$ . Uma vez que trabalhamos apenas com estratégias autofinanciadoras, a carteira óptima é determinada pelo par  $(V_T(\hat{\varphi}), \hat{\vartheta})$ , onde  $V_T(\hat{\varphi})$  denota o preço aproximado para  $H$  (i.e.,  $V_T(\hat{\varphi}) = \mathbb{E}^Q[H]$ ) e  $\hat{\vartheta}$  denota a estratégia de cobertura de média-variância óptima.

## 11.5 Cobertura do risco de longevidade

O risco de longevidade, i.e., o risco de que os membros de uma dada população de referência sobrevivam, em média e em termos agregados, para além do estimado constitui uma das principais fontes de risco com que se defrontam as companhias de seguros e os fundos de pensões. Por exemplo, no caso da população de referência representar beneficiários de rendas vitalícias, o risco de longevidade corresponde ao risco de que os beneficiários vivam sistematicamente para além do estimado pelas tábuas de mortalidade usadas pelas companhias para determinar o prémio da renda.<sup>19</sup> O risco de longevidade representa um problema importante, quer pela incerteza em torno da evolução da longevidade humana, quer pelo valor financeiro das responsabilidades exposto ao risco.

Apesar deste facto, só muito recentemente a procura de soluções de cobertura (*hedging*) do risco de longevidade mereceu por parte da indústria seguradora uma atenção redobrada. Com efeito, no passado as companhias estavam habituadas a cobrir eventuais desvios demográficos desfavoráveis com lucros generosos nos seus investimentos financeiros.

A desregulação e liberalização dos mercados de seguros, com o conseqüente aumento da competição entre as empresas e redução das margens de lucro, e a diminuição das taxas de rentabilidade dos investimentos financeiros, em resultado da redução generalizada das taxas de juro, tornou cada mais difícil a camuflagem do problema por esta via.<sup>20</sup>

As companhias expostas ao risco de longevidade podem responder de forma diversa ao problema. Podem simplesmente aceitá-lo como um risco legítimo inerente ao negócio, que dominam quantitativa e qualitativamente e que estão dispostas e preparadas a assumir, ou podem entender que a natureza da incerteza subjacente e as perdas potenciais a ela associadas são demasiado importantes para dispensar soluções de minimização ou de cobertura do risco.

---

<sup>19</sup> Refira-se que as companhias de seguros e os fundos de pensões estão igualmente expostos ao risco de longevidade via pensões em pagamento e via coberturas de cuidados de saúde de longo prazo.

<sup>20</sup> Nalguns países, as próprias entidades reguladoras não permitem a cobertura de perdas técnicas com mais-valias financeiras.

## 11.5. Cobertura do risco de longevidade

---

Nesta secção, analisamos de forma sintética um conjunto de soluções que as companhias podem adoptar para gerir a sua exposição ao risco de longevidade. As ferramentas ao dispor das empresas compreendem:

1. O (re)desenho dos produtos e a adopção de estratégias de *pricing* distintas, que passem eventualmente pela inclusão de mecanismos de redistribuição de resultados entre os beneficiários;
2. A aplicação de regras de afectação do capital conducentes a margens de solvência adequadas;
3. A implementação de estratégias de *hedging* do risco de longevidade, entre as quais destacamos:
  - Estratégias de diversificação entre diferentes coberturas suportadas pela mesma companhia;
  - Mecanismos de transferência do risco para outras contrapartes,<sup>21</sup> nomeadamente através de contratos de resseguro e soluções de mercado de capitais;
  - O uso de ferramentas de gestão do risco assentes em contratos de derivados.

As soluções discutidas nesta secção incluem, como veremos em seguida, mecanismos de cobertura interna e externa, que fazem apelo a técnicas tradicionalmente usadas na indústria seguradora, mas que apontam igualmente para a necessidade de desenvolver soluções de *hedging* inovadoras, que envolvam a titularização dos riscos ou a criação de um mercado de derivados de mortalidade.

### 11.5.1 Desenho dos produtos e estratégias de *pricing*

Uma das primeiras técnicas ao dispor das companhias para mitigar o risco de longevidade passa por redesenhar as características dos produtos tradicionais por estas oferecidos. Uma solução possível, aplicável apenas aos novos contratos e não aos já existentes passa, por exemplo, pela aplicação de cargas (*safety loadings*) adicionais aos prémios puros determinados segundo bases técnicas determinísticas.

Outra solução envolve a oferta de rendas vitalícias com prémios variáveis, i.e., prémios que tomam em consideração a evolução da esperança de vida da geração a que pertence

---

<sup>21</sup>Por exemplo, os pequenos fundos de pensões estão habituados a comprar rendas vitalícias para os seus participantes no momento da reforma. Esta solução simples permite-lhes eliminar totalmente o risco de mortalidade (sistemático e idiosincrático) da carteira de pensionistas, mantendo apenas o risco de longevidade referente aos membros no activo.

## 11.5. Cobertura do risco de longevidade

---

a pessoa segura. Esta última solução permite transferir uma parte (ou a totalidade) do risco de longevidade para os beneficiários das apólices, mas não surge frequentemente como uma das estratégias preferidas. A justificação é simples: em mercados desenvolvidos, ela colide com o desejo de manter uma posição competitiva no mercado.

Uma solução alternativa envolve a substituição das tradicionais rendas vitalícias, que garantem um benefício certo, por contratos que incluam mecanismos de participação nos resultados (bónus) dos fundos autónomos, ou cláusulas que permitam a revisão dos parâmetros de cálculo durante a vida do contrato. Esta abordagem não elimina o risco de mortalidade em termos agregados, apenas transfere uma parte do risco para os detentores de apólices sobreviventes.

Autores como Wadsworth *et al.* (2001) e Blake *et al.* (2003) descrevem mecanismos mediante os quais as rendas pagam bónus (*survivor credits*) aos seus beneficiários em função das taxas de mortalidade observadas na subpopulação abrangida. Este método contrasta, naturalmente, com as tradicionais rendas em que os benefícios são baseados exclusivamente nas probabilidades de sobrevivência no momento da compra do produto. Com mercados perfeitos e consumidores bem informados, este tipo de produtos deve, em equilíbrio, oferecer um retorno acrescido face às rendas tradicionais, em compensação pela exposição ao risco de longevidade.

Uma outra solução passa por estabelecer estratégias de *pricing* que diferenciam os indivíduos de acordo com os seus factores de risco específicos (e.g., estado de saúde, estilos de vida, factores genéticos), ou seja, que envolvem a aplicação de cargas diferenciadas (*contingency loadings*).<sup>22</sup> Este tipo de prática, possível apenas em mercados com bastante profundidade e experiência, minimizaria a subsidiação cruzada que existe actualmente entre indivíduos com distintos estados de saúde, que beneficia aqueles em melhor condição, e atenuaria os fenómenos de selecção adversa.<sup>23</sup>

### 11.5.2 Conceito de solvência estocástica

O cumprimento integral das responsabilidades das companhias de seguros e dos fundos de pensões implica a constituição de reservas financeiras (i.e., de uma carteira de activos) adequadas e a afectação de uma determinada margem de solvência (*risk-based capital*).

---

<sup>22</sup>Entre os produtos que tipificam claramente esta abordagem, destaque para as denominadas *impaired life annuities*, rendas em que se antecipa que a pessoa segura vai experimentar taxas de mortalidade mais elevadas do que a média, seja devido ao seu estilo de vida (e.g., fumador, alcoólico), seja devido a doenças crónicas.

<sup>23</sup>No futuro, a desagregação da população por grupos e factores de risco deve permitir às companhias oferecer melhores preços aos indivíduos com menores perspectivas de sobrevivência.

A abordagem tradicionalmente usada no cálculo das reservas baseia-se na determinação do valor esperado das responsabilidades futuras (benefícios líquidos de prémios) adoptando bases técnicas (demográficas e financeiras) prudentes. As hipóteses clássicas nesta matéria contemplam taxas de juro determinísticas e níveis de mortalidade relativos a uma experiência recente, eventualmente projectados para o futuro no caso de produtos com benefícios de longo prazo.

A avaliação da solvabilidade das instituições assenta na comparação dos perfis aleatórios das carteiras de activos e de responsabilidades. Nos casos em que estão envolvidos benefícios de longo prazo, exigíveis enquanto a pessoa segura sobreviver, esta abordagem pode revelar-se insuficiente. Com efeito, nestes casos a incerteza em torno das tendências demográficas (sobretudo nas idades adultas e avançadas) e em relação ao desempenho dos mercados financeiros dificulta enormemente a avaliação da adequabilidade das reservas considerando apenas cenários determinísticos sobre o futuro. Significa isto que a comparação entre os activos e as reservas pode revestir-se, neste caso, de pouco significado, em particular quanto à capacidade dos activos para satisfazer as obrigações futuras em cenários realistas.

Uma forma diferente de enfrentar o risco de longevidade, defendida por exemplo em Faculty of Actuaries (1986) e Olivieri e Pitacco (2003b), consiste em ajustar o cálculo das reservas matemáticas e da margem de solvência desejada (ou imposta) de modo a tomar em consideração a natureza estocástica dos factores de risco.

O modelo baseia-se num novo conceito de solvência, denominado de *solvência estocástica*, no qual a aferição da solvabilidade das instituições é feita comparando o seu nível de activos com o valor (aleatório) actual das responsabilidades futuras, ou seja, sem referência explícita à noção de reservas, qualquer que seja a base de avaliação adoptada no cálculo destas. Na prática, trata-se de aplicar um procedimento do tipo *mark-to-market*, quer para a carteira de activos, quer para a carteira de responsabilidades. Neste contexto, uma empresa diz-se solvente se for capaz de, com uma determinada (elevada) probabilidade, cumprir as suas obrigações futuras, avaliadas segundo uma estrutura probabilística realista.

Designemos por  $\tau$  o momento em que a solvabilidade é avaliada e por  $T$  o horizonte temporal de análise.<sup>24</sup> Considerando valores anuais, o requisito de solvência em termos patrimoniais impõe que o valor dos activos totais da empresa no momento da avaliação

---

<sup>24</sup>Refira-se a este propósito que no contexto do Projecto Solvência II, a proposta aponta no sentido de considerar  $T = 1$ , i.e., de fazer a avaliação dos requisitos anualmente.

$\tau$ ,  $A_\tau$ , deve ser tal que

$$\mathbb{P} \left\{ \bigwedge_{t=1}^T (A_{\tau+t} - Y_{\tau+t}^{[II]}) \geq 0 \right\} = 1 - \varepsilon, \quad (11.5.1)$$

onde  $Y_{\tau+t}^{[II]}$  representa o valor actual esperado das responsabilidades futuras e  $\varepsilon$  simboliza uma dada probabilidade de ruína. Desta definição resulta que os requisitos de capital (*target capital*) a exigir aos accionistas no momento  $\tau$ ,  $M_\tau$ , são iguais a

$$M_\tau = A_\tau - Y_\tau^{[II]}. \quad (11.5.2)$$

A aplicação do conceito de solvência estocástica envolve a especificação de uma série de elementos (e.g., definição dos activos e das responsabilidades, respectivo horizonte temporal, probabilidade de incumprimento tolerada, natureza aberta ou fechada do portfolio, periodicidade da avaliação da solvabilidade, ponto de vista do regulador ou do accionista), cada um dos quais conduzindo a um nível de activos exigível distinto. Esta carteira de activos deve ser financiada quer com prémios (ou contribuições), quer com a afectação de capital. Numa perspectiva tradicional, o montante total dos activos assim obtido (que deve, no mínimo, igualar o montante requerido) pode ser decomposto num montante relativo às reservas (calculadas segundo uma dada base técnica de avaliação) e num montante respeitante à margem de solvência (componente residual).

A definição de solvência expressa em (11.5.1) é mais exigente do que a tradicional, na medida em que  $Y_\tau^{[II]}$  inclui não apenas as responsabilidades decorrentes do negócio (i.e., para com os beneficiários das apólices) mas também as obrigações corporativas (i.e., para com os accionistas). Para além de permitir uma melhor aferição da solvabilidade do ponto de vista dos accionistas, esta definição é consistente com as novas directrizes contabilísticas (e com o projecto solvência II), que preconizam a avaliação das responsabilidades pelo seu valor de mercado.

Como não existe um mercado secundário desenvolvido para as responsabilidades, o seu valor pode aqui ser entendido como traduzindo a melhor estimativa possível, derivada num contexto estocástico e considerando uma estrutura probabilística realista em relação aos factores de risco. Acresce ainda que o conceito de solvência estocástica preconiza o uso de bases técnicas consistentes na avaliação de activos e responsabilidades. A aplicação do conceito de solvência estocástica melhora a avaliação dos riscos mas envolve, naturalmente, custos para o negócio e potencialmente maiores requisitos de capital, razão pela qual não será facilmente aceite pelos accionistas.

### 11.5.3 Estratégia de diversificação e *natural hedging*

A diversificação das carteiras constitui uma das primeiras soluções a adoptar na cobertura do risco de longevidade, em particular na sua componente idiossincrática. A diversificação dos riscos de longevidade requer uma adequada gestão da carteira da empresa, balanceando a exposição aos diferentes produtos, a diferentes países ou, até mesmo, entre grupos sócio-económicos. Alguns autores (e.g., Dowd (2003)) argumentam que, pelo facto da correlação entre os choques de mortalidade de diferentes países não ser perfeita, as companhias de seguros podem aproveitar as vantagens da diversificação internacional e, desta forma, reduzir (pelo menos parcialmente) a exposição ao risco.

Uma das soluções mais simples passa, no entanto, por tentar estruturar a carteira de negócios de modo a explorar potenciais “coberturas naturais”, que advêm da exposição simultânea a produtos com reacções opostas à evolução da longevidade subjacente, nomeadamente seguros de vida em caso de morte (posição longa) e rendas vitalícias (posição curta). Com efeito, se a mortalidade decresce relativamente ao esperado, as responsabilidades da companhia de seguros diminuem devido ao diferimento do momento de pagamento dos benefícios em caso de morte da pessoa segura. Pelo contrário, as companhias que comercializam rendas vitalícias incorrem em perdas relativamente ao antecipado, na medida em que se vêm obrigadas a pagar benefícios durante um período superior ao estimado inicialmente. Se o comportamento da mortalidade se deteriorar, a situação inverte-se, i.e., as companhias enfrentam perdas nos seguros em caso de morte e ganhos potenciais no mercado de rendas.

O aproveitamento das coberturas naturais constitui, assim, uma solução interna eficiente para reduzir os riscos de longevidade, um mecanismo de compensação entre posições longas e curtas que oferece às companhias uma vantagem competitiva em termos de mercado.

A importância das coberturas naturais (*natural hedging*) na protecção face ao risco de longevidade é investigada, entre outros, por Frees *et al.* (1996), Marceau e Gaillardetz (1999), Milevsky e Promislow (2001), Wang *et al.* (2003), Cox e Lin (2004), Cairns *et al.* (2006a) e Gründl *et al.* (2006). A generalidade dos estudos realizados nesta matéria explora as implicações de alterações na mortalidade nos seguros de vida e nas rendas vitalícias de forma separada, ou analisa combinações simples entre seguros de vida e dotes puros (e.g., Frees *et al.* (1996), Marceau e Gaillardetz (1999), Milevsky e Promislow (2001, 2003), Cairns *et al.* (2006a)). O próprio foco da análise é muitas vezes distinto nestes estudos, sendo que no caso dos seguros de vida o interesse se centra sobretudo nas consequências dos choques “negativos” sobre a mortalidade (aumentos),

## 11.5. Cobertura do risco de longevidade

---

enquanto que no caso das rendas a atenção está virada para o impacto dos ganhos de longevidade.

Wang *et al.* (2003) analisam as implicações decorrentes de alterações nos factores que determinam o processo seguido pelas taxas de mortalidade. Os autores desenvolvem um modelo de cobertura do risco de longevidade que envolve a escolha da combinação óptima entre seguros de vida e rendas.

Marceau e Gaillardetz (1999), Milevsky e Promislow (2001, 2003) e Bayraktar e Young (2007) analisam o problema do *pricing* e da constituição de reservas num contexto em que os factores de risco demográficos e financeiros são estocásticos. Os autores desenvolvem um raciocínio de arbitragem em carteiras que combinam seguros de vida e dotes puros, concluindo que a diversificação da carteira possibilita ganhos em termos de redução do risco.

Gründl *et al.* (2006) analisam as implicações dos riscos demográficos na combinação óptima entre capital próprio, estratégia de investimento adoptada (activos com e sem risco) e carteira de produtos (seguros de vida/rendas vitalícias), no âmbito de um modelo de maximização do valor para os accionistas de uma seguradora. Os autores concluem que a utilização da estratégia de “cobertura natural” apenas pode ser considerada óptima nos casos em que o capital é escasso.<sup>25</sup>

Cox e Lin (2004) estudam a eficácia da cobertura natural na eliminação dos riscos de mortalidade e sugerem que as companhias que, pela sua natureza do seu negócio principal, tenham exposições assimétricas ao risco (e.g., maior peso do segmento de rendas) devem realizar contratos de *swap*, intercambiando fluxos financeiros de modo a conseguir o matching dos riscos. O objectivo é neste caso o de potenciar as vantagens da cobertura natural, não no seio de uma empresa, mas entre companhias especializadas em determinados segmentos. Esta solução pode revelar-se mais lógica (e exequível) e menos dispendiosa do que a de alterar radicalmente a composição do negócio de modo a alcançar uma cobertura natural completa.

A eficácia da estratégia de diversificação depende, em grande medida, da dimensão da empresa, da abrangência do seu portfolio de negócios, do seu grau de internacionalização e da homogeneidade das características das populações de beneficiários de seguros de vida e de rendas. Em relação a este último aspecto, não é crível que as populações das duas carteiras apresentem características semelhantes em termos de longevidade, pelo que uma empresa que opere simultaneamente as duas linhas de negócio enfrenta uma espécie

---

<sup>25</sup> Os autores demonstram ainda que a estratégia óptima depende simultaneamente do montante de capital disponível, do grau de aversão (à insolvência) dos compradores de seguros, das cargas adicionais definidas pela empresa e do modelo probabilístico usado para descrever a situação demográfica.

de “risco de base”. A cobertura completa do risco de longevidade pode, neste sentido, requerer o recurso a mecanismos alternativos, designadamente a sua transferência para os mercados financeiros.

### 11.5.4 Contratos de resseguro

Uma solução alternativa para realizar a cobertura do risco de longevidade envolve a cessão (parcial ou total) dos riscos assumidos por um segurador directo perante os segurados a uma resseguradora através de contratos de resseguro ou outras determinações legais. Refira-se, neste ponto, que o mecanismo do resseguro não conduz a uma redução global do risco de longevidade, devido à natureza sistemática deste, mas apenas a uma mudança na entidade que o suporta. Ainda assim, esta ferramenta constitui uma forma válida de ajustar a exposição ao risco.

O recurso ao resseguro é uma estratégia comum na cobertura dos riscos “convencionais” (e.g., riscos catastróficos, acidentes de viação, aéreos ou marítimos). Pelo contrário, são escassos os registos públicos sobre a realização de contratos de resseguro de risco de longevidade, pese embora comecem a surgir sinais claros do interesse de seguradoras e resseguradoras pelo desenvolvimento deste mercado (veja-se, e.g., Dowd (2003)).

Conforme referimos acima, apenas as grandes companhias estão em condições de maximizar os benefícios da estratégia de diversificação pelo que, em princípio, uma resseguradora, operando em vários mercados e áreas geográficas, estará melhor capacitada para suportar uma maior exposição ao risco e para eliminar os riscos idiossincráticos. Por outro lado, a cessão dos riscos visa limitar as consequências de eventuais movimentos adversos nos factores de risco, em particular os associados a catástrofes.

Existem, em tese, inúmeros acordos de resseguro que podem ser gizados para proporcionar às companhias de seguro uma cobertura adequada à sua exposição ao risco de longevidade. Neste sentido, a nossa contribuição neste domínio consubstancia-se na discussão das potencialidades e desvantagens de um leque restrito de propostas de resseguro.

Entre os acordos analisados, encontramos contratos do tipo proporcional (i.e., em que existe uma proporção fixa que estabelece antecipadamente a repartição dos prémios e responsabilidades entre segurador directo e o ressegurador) e contratos do tipo não-proporcional, contratos que visam uma intervenção ao nível individual e contratos que intervêm ao nível de uma carteira de apólices.

### Resseguro do tipo quota-parte

A proposta mais simples envolve um resseguro proporcional do tipo “quota-parte”. Segundo os termos deste contrato, o ressegurador assumiria a responsabilidade por uma percentagem fixa (quota) de todas as apólices de renda vitalícia subscritas pelo segurador directo. A quota negociada nos termos do acordo determinaria a repartição das reservas, dos prémios e dos benefícios pagos aos detentores das apólices entre segurador directo e ressegurador. A contrapartida para o ressegurador poderia constituir-se sob a forma de um prémio de resseguro único pago à cabeça, indexado à importância segura ou dependente do valor da quota-parte do risco assumida.

Segundo este acordo, existiria uma partilha dos riscos (financeiro e de longevidade) e dos lucros entre segurador e ressegurador, que exigiria um alinhamento de interesses muito grande entre as partes. Em particular, a celebração do contrato exige um acordo total quanto às hipóteses (comportamento da mortalidade e taxas de juro de referência) usadas no *pricing* e na avaliação das reservas matemáticas e, nalguns casos (e.g., casos em que as rendas envolvem benefícios indexados à evolução do nível de preços) em relação à evolução da inflação.

Os contratos de resseguro do tipo quota-parte apresentam, no entanto, algumas desvantagens. Desde logo, o facto de o ressegurador avaliar globalmente a carteira, não considerando suficientemente as diferentes necessidades de resseguro do segurador directo. Acresce que um contrato desta natureza não contribuiria para a homogeneidade da carteira uma vez que não limitaria a exposição aos riscos de ponta, nomeadamente aos eventos associados a uma longevidade superior ao estimado. O contrato limitar-se-ia a reduzir a escala da exposição ao risco, sem alterar as características da distribuição dos riscos. Nalgumas carteiras, o contrato ofereceria mesmo coberturas de resseguro desnecessárias, reduzindo inutilmente a margem de lucro do segurador directo.

Ainda assim, este tipo de contrato é indicado sobretudo para companhias em desenvolvimento, sem um conhecimento profundo do mercado de rendas vitalícias, que não dispõem de um capital de experiência na determinação dos prémios adequado e que procuram, por esta via, uma partilha dos riscos do negócio.

### Resseguro do tipo *Excess of Loss* em valor

Uma solução alternativa clássica envolveria um resseguro não-proporcional do tipo *Excess of Loss* (XL *reinsurance*) em valor. Neste tipo de acordo, o ressegurador não partilha todos os riscos como no caso do contrato do tipo quota-parte. Com efeito, no caso específico de uma carteira de rendas o segurador directo reteria todos os riscos

até uma determinada quantia máxima (prioridade), definida para cada apólice ou em termos agregados para a carteira (contrato do tipo *Aggregate Excess of Loss*), cedendo ao ressegurador a parte da cobertura que excede a prioridade. Um acordo tipo poderia ainda contemplar a existência de uma franquia, definida em termos de montante individual (ou em termos agregados), discriminada ou não por sexo e/ou por factores de risco. Adicionalmente, a responsabilidade do ressegurador pode ser balizada incluindo no acordo cláusulas de limitação de perdas (*Stop Loss*) definidas, por exemplo, através de um múltiplo da prioridade.

Da relação entre a prioridade e os valores cedidos em resseguro resultaria um indicador da proporção do risco que é ressegurado, com base na qual seria possível determinar a repartição dos prémios de seguro, das reservas e dos benefícios gerados pelas rendas entre seguradores directos e resseguradores.

Na prática, com este tipo de contrato apenas as rendas vitalícias de valor mais elevado seriam cedidas. A maior parte dos estudos empíricos realizados sugerem a existência de uma relação positiva entre valor das rendas e a esperança de vida (Mitchell *et al.* (1999)). Neste sentido, a principal vantagem de um contrato desta natureza residiria na possibilidade de reduzir os impactos financeiros resultantes da presença de selecção adversa na carteira. De facto, este contrato constituiria, para o segurador directo, uma forma de homogeneizar a exposição ao risco e de, ao mesmo tempo, restringir a exposição aos riscos mais significativos.

### **Resseguro do tipo *Excess of Loss* em duração de vida**

Uma solução interessante na cobertura do risco de longevidade é aquela que é proporcionada por um resseguro do tipo *Excess of Loss* (*XL reinsurance*) em duração de vida. Neste tipo de acordo, o segurador directo assumiria financeiramente (sozinho) todos os pagamentos decorrentes dos eventos cobertos pelo contrato de renda vitalícia até um determinado limite, independentemente da importância assegurada. Por sua vez, o ressegurador pagaria ao cedente a parcela “final” da renda, i.e., a parcela devida no caso do beneficiário sobreviver para além de uma determinada idade (prioridade)  $x_0 + \tau$  (e.g., 85 anos), onde  $x_0$  representa a idade do indivíduo no momento da realização do contrato e  $\tau$  denota o período máximo de pagamento da renda em que o cedente não recebe qualquer benefício do ressegurador. Se o beneficiário estiver vivo à idade  $x_0 + \tau$ , o ressegurador assume o pagamento da renda perante o cedente, um pagamento que perdurará enquanto este sobreviver (ou até a uma idade máxima definida no contrato, fixada através de um plafond de garantia ou medida em anos).

## 11.5. Cobertura do risco de longevidade

---

O valor da prioridade seria calculado isoladamente para cada apólice que integra a carteira, admitindo-se igualmente a existência de uma franquia, discriminada ou não por sexo e/ou outros factores de risco. Como se depreende, este tipo de resseguro estabelece uma intervenção de cobertura a um nível individual.

A principal vantagem deste contrato para o cedente consistiria na possibilidade de converter rendas vitalícias em rendas temporárias, sendo que apenas o risco de longevidade seria transferido (potencialmente na íntegra) para o ressegurador. Por seu turno, em razão deste contrato o ressegurador adquire uma posição curta numa renda vitalícia diferida  $\tau$  anos no futuro.

A contrapartida para o ressegurador envolveria o pagamento à cabeça de um prémio de resseguro único, considerando eventualmente cláusulas de revisão em momentos definidos no contrato, no caso da mortalidade evoluir de forma significativamente diferente da trajectória prevista no momento de celebração do acordo. Admite-se igualmente que o ressegurador possa financiar parcialmente o cedente permitindo que este reparta o pagamento do prémio num intervalo de tempo mais ou menos alargado. O prémio a pagar deve ainda contemplar uma estimativa do custo da longevidade adoptando, desejavelmente, uma perspectiva estocástica.

No que diz respeito às reservas matemáticas, como existe uma divisão do período de cobertura entre as partes admite-se como razoável que o cedente seja obrigado a constituir reservas apenas com referência aos benefícios a pagar até à prioridade, enquanto que o ressegurador deve constituir reservas apenas para fazer face aos pagamentos a efectuar a partir dessa idade. Dito de outro modo, o contrato proposto envolve a transferência de parte das reservas do cedente para o ressegurador.

A principal diferença deste tipo de acordo em relação aos contratos do tipo proporcional nos quais, como vimos, o ressegurador participa em todas as responsabilidades decorrentes das apólices, refere-se ao facto do ressegurador apenas assumir o pagamento das responsabilidades com as rendas após a prioridade.

Um contrato deste tipo constituiria uma solução eficaz de redução do risco de longevidade para o cedente. Pelo contrário, o acordo apresenta-se à partida como bastante arriscado para o ressegurador, cuja intervenção ocorre apenas no segmento do arco da vida humana em que a avaliação das perspectivas de longevidade é mais difícil, devido à escassez de informação sobre o comportamento da mortalidade nestas idades. Significa isto que é de prever que exista alguma reticência por parte das resseguradoras em participar neste tipo de contratos, pelo menos enquanto não estiverem disponíveis no mercado outras soluções de *hedging* do risco, que envolvam nomeadamente a sua transferência para o mercado de capitais.

Por fim, uma variante interessante deste contrato seria aquela que estipulasse uma dupla prioridade, que seja simultaneamente função da duração de vida do beneficiário da renda e do valor do contrato. Trata-se, no fundo, de combinar um contrato do tipo *Excess of Loss* em duração de vida com um contrato do tipo *Excess of Loss* em valor. Esta variante constituiria uma forma de ter em conta a eventual heterogeneidade da população segura em termos de níveis de renda e de longevidade.

### Resseguro do tipo *Stop-Loss* com base nos activos

Consideremos agora um tradicional contrato de resseguro do tipo *Stop-Loss* com base nos activos, definido num horizonte de médio e longo prazo, em que as intervenções do ressegurador teriam como objectivo prevenir uma situação de insolvência por parte do cedente em consequência de desvios significativos no comportamento da mortalidade.

O *rationale* por detrás deste contrato é simples. Assume-se que o efeito dos desvios na mortalidade sobre o negócio se torna perceptível comparando, num dado momento, o valor dos activos na carteira com o valor das reservas requeridas para fazer face às obrigações para com os beneficiários de rendas vitalícias sobreviventes nesse momento. Por outras palavras, a perda (*Loss*) materializa-se na insuficiência de activos.

Definido um horizonte temporal, se no final desse período o valor dos activos for inferior às reservas o ressegurador cobre uma parte (ou a totalidade, consoante o acordo) da diferença, i.e., o ressegurador intervém se no momento  $t$  se verificar que

$$A_t < (1 - \pi) V_t^{[II]},$$

onde  $V_t^{[II]}$  denota o valor das reservas e  $\pi V_t^{[II]}$  simboliza o valor da prioridade (em termos de perda, *Stop-Loss*) fixada no acordo.

Uma vez que o objectivo é o de proporcionar cobertura face ao risco de longevidade, o horizonte temporal a definir no contrato deve ser de longo prazo, de modo a expurgar potenciais oscilações no valor da carteira de activos devidas, não ao risco de longevidade, mas a flutuações aleatórias de curto prazo na mortalidade. Para compensar a significativa exposição ao risco a que o ressegurador fica sujeito em consequência deste acordo, é de esperar que o valor dos prémios de resseguro a cobrar (no momento da celebração do contrato ou periodicamente) sejam elevados.

A implementação de um contrato desta natureza exige, naturalmente, o acordo entre cedente e ressegurador em relação aos activos a comparar com as reservas, ao cálculo do valor das reservas, ao horizonte temporal do contrato e, no caso de cobertura parcial das perdas, da prioridade *Stop-Loss*. A obrigatoriedade de constituição das reservas

permanece no cedente, admitindo-se no entanto que o ressegurador possa exercer algum controlo sobre a política de investimentos deste, uma vez que a sua intervenção no contrato depende directamente evolução da carteira de activos do segurador directo.

A principal desvantagem deste tipo de acordo para o cedente respeita ao facto de as eventuais perdas registadas durante o período de cobertura não estarem abrangidas pelo acordo de resseguro, uma vez que o ressegurador intervém apenas no final do horizonte definido. Para o ressegurador, a principal limitação decorre da potencial falta de controlo sobre a forma como o cedente gere os seus activos durante o horizonte do acordo.

### **Resseguro do tipo *Stop-Loss* com base nos fluxos de caixa**

Uma alternativa ao resseguro do tipo *Stop-Loss* (SL) com base nos activos passa por considerar que a intervenção do ressegurador ocorre apenas quando o valor dos benefícios pagos pelas rendas vitalícias é superior ao valor esperado no momento da celebração do contrato. Por outras palavras, um contrato desta natureza determina que o ressegurador intervém apenas em caso de aumentos inesperados na longevidade das pessoas seguras, tomando a seu cargo as responsabilidades adicionais que daí advém. No limite, se o objectivo do contrato for apenas o de transferir a totalidade (ou pelo menos a maioria) do risco de longevidade do cedente para o ressegurador, a prioridade deve ser fixada acima do valor esperado dos fluxos de caixa da carteira, de modo a excluir a transferência do risco de flutuações aleatórias em torno da tendência de longo prazo.

Um acordo deste tipo deve ainda estipular, entre outros aspectos, o horizonte temporal da cobertura e a frequência da intervenção nesse intervalo (e.g., no final de cada ano), assim como a existência de eventuais limites para as perdas a assumir pelo ressegurador (cláusulas *stop-loss*). O *pricing* deste tipo de acordos afigura-se decisivo, tendo em conta o valor e a natureza (sistemática) do risco envolvidos.

Os contratos do tipo SL apresentam, ao contrário dos contratos do tipo XL, a vantagem de requerer o envolvimento do cedente e do ressegurador durante a totalidade do período de cobertura. Um acordo centrado nos fluxos de caixa afigura-se nos mais fácil de implementar do que aquele assente nos activos, devido à maior objectividade na definição e medição dos cash flows. Acresce ainda que um acordo baseado nos fluxos de caixa torna mais fácil a sua titularização, nomeadamente pela emissão de títulos de dívida com cash flows indexados à evolução da longevidade.

### 11.5.5 Titularização do risco: *Longevity/Survivor Bonds*

A oferta limitada de resseguro e os previsíveis custos elevados associados a alguns dos tradicionais acordos acima analisados, em conjunto com as crescentes preocupações quanto ao risco de crédito dos resseguradores, levaram muito recentemente as companhias de seguros a considerarem a titularização como uma solução alternativa na cobertura dos seus riscos de longevidade.

No passado, as companhias funcionavam como verdadeiros reservatórios de riscos, mais ou menos estanques, com muito pouca transferência para os mercados de capitais, se exceptuarmos naturalmente a captação de capital e a emissão de alguns instrumentos de dívida.

Nos últimos anos, as companhias de seguros transformaram-se progressivamente em intermediários financeiros, que colocam vários tipos de instrumentos junto dos mercados de capitais, transferindo desta forma os riscos directamente para os investidores.

Os principais aspectos envolvidos na titularização dos riscos de longevidade são analisados por Cowley e Cummins (2005), Krutov (2006) e Lin e Cox (2005a,b). Entre as principais finalidades da titularização incluem-se o financiamento dos custos associados à captação de novas apólices, a libertação de capital regulamentar “excessivo” em virtude de práticas conservadoras na constituição de reservas, a monetarização de um conjunto de cash flows gerados pela carteira de negócios, a redução ou eliminação da exposição a riscos (catastróficos e sistemáticos) demográficos ou de mercado, ou ainda a oportunidade de investir em novos activos e tipos de risco que melhorem a eficiência da carteira pela via da diversificação.

As operações de titularização no ramo vida adoptam uma estrutura similar à observada nos denominados *Catastrophe Bonds* (*CAT Bonds*). Nestas operações, a seguradora (ou resseguradora) transfere um conjunto de cash flows (e.g., prémios de rendas vitalícias) para um *Special Purpose Vehicle* (SPV), uma entidade passiva que é criada apenas para emitir títulos de dívida (em que os cash flows são usados como colateral) e para os colocar junto dos investidores. Estes contribuem com fundos para o SPV e assumem os riscos do investimento. Os fundos são por sua vez investidos e devem ser suficientes para pagar os cupões aos investidores e para assumir as responsabilidades com as rendas.

No actual contexto, o interesse pela titularização é justificado pelas potencialidades que a emissão de activos financeiros com cash flows contingentes à evolução da longevidade (em particular, dos chamados *longevity bonds*) oferece na cobertura do risco.

**Longevity (Survivor) Bonds**

Propostas pela primeira vez por Blake e Burrows (2001), as *Longevity Bonds* (LB's) ou *Survivor Bonds* (SB's) são instrumentos de dívida cujos cash flows dependem da realização de um determinado índice de sobrevivência  $\mathbb{I}_{S_x(t)}$  no momento  $t$ . Como o nome sugere, este índice representa a proporção de uma determinada população com idade  $x$  no momento  $t = 0$  que sobrevive até ao momento  $t$ . Na sua formulação original, a obrigação paga cupões até à extinção de todos os elementos de uma dada geração. Por outras palavras, trata-se na prática de uma obrigação clássica com maturidade estocástica, em que a variável estocástica corresponde à duração de vida do indivíduo com maior longevidade entre a geração de referência.<sup>26</sup>

Como se depreende, este tipo de obrigações de longo prazo é desenhado de forma a proteger o seu detentor contra quaisquer melhorias não antecipadas na mortalidade até à maturidade do título. Por esta razão, constitui-se como um excelente instrumento de cobertura para uma carteira dominada por rendas vitalícias, i.e., de uma carteira com uma exposição predominantemente curta ao risco de longevidade. Com efeito, na eventualidade dos beneficiários das rendas viverem mais tempo do que o esperado, a

<sup>26</sup> Não existem, até ao momento, registos públicos sobre a emissão de qualquer obrigação com as características de uma *Longevity Bond*. Existem, contudo, dados precisos sobre uma emissão anunciada em Novembro de 2004 pelo Banco Europeu de Investimento (BEI) e estruturada pelo Banco francês BNP Paribas (veja-se Blake *et al.* (2006a)). A emissão tinha um valor facial de 540 milhões de libras, um cupão inicial de 50 milhões de libras (£50m) e uma maturidade de 25 anos. Tratava-se de uma obrigação perpétua (*annuity or amortising bond*) com cupões variáveis, cuja característica inovadora consistia em ligar os cupões a um índice de sobrevivência de uma geração  $\mathbb{I}_{S_x(t)}$ , baseado nas taxas de mortalidade da população masculina de Inglaterra e do País de Gales com 65 anos em 2002. Em termos mais precisos, os fluxos de caixa gerados pela emissão eram dados por:

$$f_t(S_{65}(t)) = £50m \times \mathbb{I}_{S_{65}(t)}, \quad t = 1, 2, \dots, 25$$

O risco de longevidade seria coberto pela resseguradora Partner Re, que participava através de um contrato de *swap* onde se comprometia a fazer pagamentos anuais variáveis (no valor de  $£50m \times \mathbb{I}_{S_{65}(t)}$ ) ao BEI, calculados com base na mortalidade observada na população masculina de Inglaterra e do País de Gales com 65 anos, recebendo do BEI um conjunto de pagamentos anuais fixos determinados com base na mortalidade projectada para esta geração. Uma vez que o BEI desejava efectuar pagamentos em euros, esta operação era ainda complementada com um *swap* cambial (*fixed-sterling-for-floating-euro*) entre o BEI e o BNP. Como se depreende, a emissão estava estruturada para servir como activo de cobertura do risco de longevidade para um investidor com uma carteira dominada por rendas vitalícias. A obrigação de longo prazo estava desenhada para proteger o investidor contra diminuições não antecipadas na mortalidade até à maturidade do título. Por seu turno, o emitente realizaria lucros no caso de  $\mathbb{I}_{S_x(t)}$  ser inferior ao antecipado. Ao contrário do que se esperava, a emissão não teve grande sucesso junto dos investidores (foi apenas parcialmente subscrita), acabando por ser retirada nos finais de 2005 para reformulação. Entre as explicações para este insucesso contam-se o prazo da emissão (considerado curto), o valor do capital exigido face aos benefícios da redução da exposição ao risco, o risco de base subjacente à definição do índice de mortalidade (não devidamente compensado em termos de retorno) e o risco de modelação elevado na avaliação do título.

## 11.5. Cobertura do risco de longevidade

---

companhia de seguros é obrigada a efectuar pagamentos durante períodos mais longos do que o previsto recebendo, em contrapartida, cupões mais elevados nas suas posições em LB's.

Este tipo de obrigações constitui uma solução mais flexível e menos onerosa de cobertura do risco de longevidade do que os tradicionais contratos de resseguro. Com efeito, a solução permite a transferência do risco para um número mais alargado de contrapartes, dá origem a instrumentos financeiros que dependem apenas da evolução da mortalidade, que não necessitam de ser forçosamente ajustados às características da carteira de cada companhia e que são negociáveis em mercado secundário, o que permite uma gestão mais flexível das coberturas.

Para além do interesse previsível das companhias de seguros, outros investidores (e.g., bancos de investimento ou *hedge funds*) podem estar potencialmente interessados em adquirir uma exposição ao risco de longevidade, uma vez que este apresenta uma correlação baixa com os tradicionais factores de risco. Neste sentido, a combinação de um parâmetro beta baixo, um parâmetro alfa potencialmente positivo e uma taxa de retorno razoável transforma os activos com cash flows contingentes à longevidade num investimento atractivo do ponto de vista da diversificação da carteira.

Um mercado desenvolvido tenderá igualmente a atrair especuladores, que negociam com base em expectativas sobre a direcção dos preços dos activos (e da longevidade), e arbitragistas, que procuram explorar eventuais anomalias no *pricing* dos produtos. Saliente-se, no entanto, que a emissão deste tipo de obrigações não elimina o risco de longevidade da economia, apenas o transfere neste caso das seguradoras para as entidades emitentes dos títulos.

A emissão de LB's levanta um número significativo de questões, analisadas em detalhe por Blake e Burrows (2001), Dowd (2003), Blake (2003), Lin e Cox (2005a,b), Milevsky (2005), Cairns *et al.* (2005, 2006a,b), Blake *et al.* (2006a,b), Brown e Orzag (2006), Stallard (2006), Webb e Friedberg (2006), Bauer e Russ (2006), Cox *et al.* (2006) e Denuit *et al.* (2007).

A primeira questão importante refere-se à existência de um previsível desequilíbrio estrutural neste mercado. De facto, se do lado da procura parecem existir motivações naturais para a aquisição deste tipo de activos por parte de companhias de seguros, fundos de pensões e outros investidores, do lado da oferta não se vislumbra automaticamente o universo dos potenciais emitentes.

À partida, entre os candidatos naturais a emitentes encontram-se entidades (públicas ou privadas) com posições longas no risco de longevidade (e.g., empresas farmacêuticas, empresas do sector da distribuição, empresas direccionadas para o fornecimento de bens,

serviços ou alojamento a idosos, estados ou municípios que atraem reformados ricos (e.g., Estado Americano da Florida)), que desejem fazer o *hedging* da sua posição vendendo obrigações. Dowd (2003) sugere mesmo que entre o universo de potenciais emitentes podem surgir companhias de seguros em que o peso da carteira de seguros de vida seja relativamente superior ao da carteira de rendas vitalícias. Ainda assim, é de prever que este mercado seja caracterizado por um excesso natural de procura, que não deverá ser interpretado como uma falha de mercado, mas em que o preço das obrigações e as respectivas *yields* ajustarão (preços sobem e yields baixam) de modo a refrear a procura e a encorajar a oferta.

Em alternativa, Blake e Burrows (2001), Blake (2003) e Brown e Orzag (2006) propõem que as LB's sejam emitidas pelos Estados. Para fundamentar a sua proposta argumentam desde logo que o Estado, garante de última instância, deve suprimir a falha de mercado que representa a inexistência de mecanismos de cobertura do risco de longevidade pelas empresas privadas. Sustentam ainda que a dispersão do risco pela população (pelos contribuintes) elimina o prémio de risco associado e que o Estado contribui para o aumento da longevidade (e.g., através dos sistemas públicos de saúde), pelo que deve suportar as consequências dessas políticas. Por último, defendem que o Estado tem, ao contrário das empresas privadas, uma vantagem comparativa que lhe permite efectuar a partilha intergeracional do risco de forma mais eficiente, nomeadamente através da política fiscal (sistemas de protecção social, impostos, prestações sociais, dívida pública). Na prática, não é certo que da intervenção do Estado resultará uma partilha óptima do risco entre gerações, na medida em que os decisores políticos podem, por razões eleitorais, ser tentados a proteger excessivamente a geração actual à custa das gerações futuras.

Em oposição, alguns autores (e.g., Dowd (2003), Turner (2006)) questionam a transferência de mais risco de longevidade para o Estado, na medida em que este já está consideravelmente exposto, seja através dos sistemas de segurança social (assentes em mecanismos de repartição), seja através dos sistemas públicos de saúde. Estes autores argumentam ainda que os mercados de capitais estão melhor preparados para suportar e partilhar riscos que o Estado, que apenas diversifica os riscos entre um grupo circunscrito de indivíduos (contribuintes), enquanto que nos mercados é possível uma diversificação internacional.

O segundo aspecto chave na emissão de LB's prende-se com a escolha do índice de sobrevivência (Blake *et al.*, 2006a,b). O sucesso deste tipo de emissões depende do grau de ajustamento (*matching*) entre os cash flows da obrigação e as responsabilidades da seguradora. Por este motivo, o índice seleccionado deve, desejavelmente, ter em conta as características de mortalidade da população cujo risco se pretende cobrir de maneira

## 11.5. Cobertura do risco de longevidade

---

a minimizar o “risco de base”, abranger o segmento do arco da vida humana a que as empresas estão expostas e atender aos factores de risco da população (e.g., sexo).

A construção de um índice de sobrevivência enfrenta ainda outros problemas. Em primeiro lugar, os índices são construídos a partir de estatísticas pouco frequentes e obtidas com razoável desfasamento temporal. Em segundo, os dados de base são muitas vezes objecto de gradação, mediante métodos que podem mudar no tempo. Em terceiro, os resultados podem ser contaminados por erros nas observações. Em quarto, devemos ter em conta a importância do risco moral, que emerge neste contexto pela possibilidade de manipular as estatísticas de base ou pelo acesso antecipado à informação sobre o índice. Por último, os índices envolvem, em muitos casos, projecções de mortalidade que estão sujeitas a riscos de modelação.

Um terceiro aspecto-chave na emissão de LB's respeita à escolha do método de avaliação adequado quando os cash flows dos activos são contingentes à evolução da mortalidade. Com efeito, enquanto que as obrigações clássicas podem ser avaliadas usando a estrutura temporal de taxas de juro e os princípios de não arbitragem, tal não é possível neste contexto devido à natureza incompleta do mercado. Entre as soluções apresentadas na literatura encontram-se, como vimos na Secção 11.4.3, a adopção de uma abordagem *risk-neutral*, a utilização de princípios clássicos de determinação do prémio em contratos de seguro, o uso de abordagens baseadas na Transformada de Wang, no Rácio de Sharpe ou no modelo Consumption CAPM, estratégias de minimização do risco ou estratégias de média-variância.

Um último aspecto a ter em conta refere-se ao risco de crédito das emissões. A principal função das LB's (e de outros activos contingentes à evolução da longevidade que analisaremos no capítulo seguinte) é a de proporcionar às empresas uma oportunidade de fazer o *hedging* do risco de longevidade. A procura desta cobertura não significa que os investidores estão dispostos a aumentar a sua exposição ao risco de crédito ou a outros tipos de riscos de mercado. Neste sentido, na eventualidade de as entidades emitentes acarretarem risco de crédito devem ser consideradas soluções para mitigar este problema (e.g., SPV, *credit insurance arrangements*, derivados de crédito).

### ***Mortality-Linked Bonds***

Para além da obrigação acima apresentada, outras modalidades de *Longevity Bond* (LB) foram recentemente propostas e discutidas na literatura (veja-se, e.g., Blake *et al.* (2006a,b), Cairns *et al.* (2005, 2006a,b)). A LB “clássica” é uma obrigação perpétua com pagamentos dependentes da evolução de um índice de sobrevivência em que apenas

os cupões são contingentes à evolução da mortalidade. A natureza desta dependência pode, no entanto, ser distinta e ser especificada de modo a que o investidor esteja exposto ao risco de perder parte ou a totalidade dos cupões (e/ou do principal) se a mortalidade ultrapassar um determinado patamar.

Uma variante interessante das LB's são as chamadas *Zero-Coupon Longevity Bonds* (ZCLB), uma modalidade decalcada das tradicionais obrigações de cupão zero negociadas nos mercados de taxa de juro e que poderiam ser obtidas por decomposição (*stripping*) de uma LB clássica. No limite, podemos imaginar um completo espectro bidimensional de ZCLB's, em que a primeira dimensão está relacionada com a geração em causa e a segunda diz respeito à maturidade do título. A existência deste tipo de obrigações permitiria às companhias de seguros construir carteiras de cobertura adequadas ao seu perfil (demográfico e temporal) de responsabilidades.

A emissão de obrigações em que os fluxos de caixa dependem inversamente da evolução de um índice de sobrevivência (*Inverse Longevity Bonds*) constitui uma alternativa possível (e.g.,  $f_t(\mathbb{I}_{S_x(t)}) = \alpha(1 - \mathbb{I}_{S_x(t)})$ ,  $\alpha > 0$ ). Esta modalidade é inspirada nas convencionais *inverse floaters*, obrigações de taxa variável em que os cupões variam inversamente como os movimentos nas taxas de juro.

Outra das modalidades sugeridas refere-se às denominadas *Deferred Longevity Bonds* (DLB), obrigações em que os cash flows são diferidos no tempo em relação ao momento da emissão. Esta modalidade, que configura na prática uma espécie de contrato *forward* sobre a mortalidade, tem em conta o facto dos progressos na longevidade só assumirem maior expressão a médio e longo prazo.

Uma modalidade distinta é aquela representada pela primeira obrigação com cash flows contingentes à evolução da mortalidade efectivamente emitida e subscrita nos mercados de capitais. Esta emissão, lançada pela resseguradora Swiss Re em Dezembro de 2003, tinha uma maturidade de apenas três anos, um principal de 400 milhões USD (\$400m) e oferecia aos investidores um cupão variável indexado à LIBOR, adicionado de um spread de 135 *basis points*. A característica distintiva desta obrigação referia-se, no entanto, ao facto do reembolso do principal depender do valor observado para um índice de mortalidade,  $M_t$ , calculado com base na média ponderada das taxas de mortalidade registadas em cinco países, em ambos os sexos e num intervalo de idades.

Em termos mais precisos, o principal seria reembolsado na totalidade se o valor de  $M_t$  não excedesse 130% do valor base (observado em 2002) durante a vida da obrigação, caso contrário dependeria dos valores efectivamente registados para  $M_t$ , de acordo com a seguinte função  $f_t(\cdot)$

## 11.5. Cobertura do risco de longevidade

$$f_t(\cdot) = \begin{cases} LIBOR + spread & t = 1, \dots, T - 1 \\ LIBOR + spread + \max \left\{ 0, 100\% - \sum_t L_t \right\} & t = T, \end{cases} \quad (11.5.3)$$

onde  $L_t$  é uma função de perda definida segundo

$$L_t = \begin{cases} 0\% & \text{se } M_t < 1.3M_0 \\ [(M_t - 1.3M_0) / (0.2M_0)] \times 100\% & \text{se } 1.3M_0 \leq M_t \leq 1.5M_0 \quad \forall t \\ 100\% & \text{se } 1.5M_0 < M_t, \end{cases}$$

onde  $M_0$  e  $M_t$  denotam os valores do índice de mortalidade observados no ano base e no ano  $t$ , respectivamente. Como se constata, o valor dos cupões é imune à evolução do índice  $M_t$ , enquanto que o principal é uma função linear (por troços) de  $M_t$ .

Uma das características mais distintivas desta emissão é, no entanto, o facto da obrigação configurar uma espécie de *Catastrophe Mortality Bond*, estruturada, não para cobrir o risco de longevidade do investidor, mas para proteger o emitente do risco de mortalidade catastrófica. Com efeito, o emitente lucra numa situação em que  $M_t$  é extremamente elevado (e.g., numa catástrofe natural), ou seja, permite transferir para o mercado de capitais o risco de mortalidade a que está exposta uma carteira de apólices de seguro/resseguro de vida. Neste caso, o risco para o emitente é o de que as responsabilidades com as apólices se tornem exigíveis antes do previsto.

A obrigação foi emitida através de um SPV chamado Vita Capital, uma solução conveniente para o emitente uma vez lhe permitia manter as responsabilidades fora do balanço ao mesmo tempo que reduzia o risco de crédito suportado pelos investidores.<sup>27</sup>

<sup>27</sup> Após o sucesso da primeira emissão, a Swiss Re anunciou em Abril de 2005 que havia emitido, através de um novo SPV chamado Vita Capital II, uma segunda *Catastrophe Mortality Bond* com um principal de 362 milhões USD (\$362m), dividido em três tranches: Classe B (\$62m), Classe C (\$200m) e Classe D (\$100m), maturidade em 2010 e um índice ponderado de mortalidade calculado com base na população dos EUA, Alemanha, Japão e Canadá. A especificação da obrigação estipulava que o principal estaria exposto ao risco se, em dois anos consecutivos antes da maturidade, o índice combinado de mortalidade excedesse determinadas percentagens do nível de mortalidade esperada (a saber, 120% para a Classe B, 115% para a Classe C e 110% para a Classe D). Esta emissão foi, uma vez mais, totalmente subscrita e seria seguida, já em 2007, de uma nova emissão de \$390m e €240m com características semelhantes, emitida através de um SPV designado por Vita Capital III. Em 2006, as companhias Axa e Scottish Re emitiram, através de SPV's chamados de Osiris Capital e Tartan Capital, respectivamente, catastrophe mortality bonds no valor de \$350m + e50m e \$155m com base no comportamento da população de (França, Japão e EUA) e EUA, respectivamente.

### 11.5.6 Derivados de mortalidade

A hipótese, razoável, de que as principais entidades expostas ao risco de longevidade estão interessadas em precaver-se contra esta fonte de incerteza, permite-nos equacionar a criação e desenvolvimento no futuro de um mercado de contratos derivados de mortalidade. Neste mercado, seria possível negociar um leque de novos instrumentos financeiros com cash flows contingentes à evolução da longevidade, apresentando características similares às observadas nos tradicionais contratos de futuros, opções financeiras e de *swap*.

Em particular, podemos imaginar a tradicional segmentação entre produtos negociados em bolsas organizadas (e.g., futuros) e produtos transaccionados em mercados *over-the-counter* (OTC) (e.g., swaps). Os primeiros apresentam como principais vantagens a liquidez (que facilita a anulação das posições, curtas ou longas), a transparência (decorrente da centralização das operações em Bolsa), a eliminação do risco de crédito (pela intervenção de uma Câmara de Compensação, que se constitui como contraparte nos negócios) e a rapidez negocial, tendo no maior risco de base a sua maior desvantagem.

Por seu turno, os derivados de mortalidade negociados em mercados OTC têm como principal atractivo a possibilidade de serem estruturados à medida das necessidades dos intervenientes, de modo a minimizar o risco de base, mas a sua negociação posterior é limitada pela quase inexistência de mercados secundários, o que dificulta o rearranjo das posições.

A criação e utilização de derivados na gestão do risco de longevidade foi sugerida recentemente por Blake (2003), Dowd (2003), Cox e Lin (2004), Lin e Cox (2005a,b), Blake *et al.* (2006a,b) e Dowd *et al.* (2006). Nesta secção apresentamos uma breve descrição das principais características dos instrumentos financeiros (existentes ou hipotéticos) propostos na literatura.

#### Mortality Swaps

Um *swap* de mortalidade (*mortality swap*, MS) é um acordo no qual se estabelece o intercâmbio de um ou mais cash flows no futuro com base na evolução de um índice de mortalidade ou de sobrevivência.<sup>28</sup> Definidos desta forma, os *swaps* de mortalidade partilham muitas das características dos acordos de resseguro. Com efeito, ambos envolvem a troca de pagamentos com base em valores observados ou projectados para os factores de risco, podendo ser usados para o mesmo propósito. Refira-se, no entanto, que existem algumas diferenças importantes entre estes dois tipos de acordo. Em particular, os *swaps*

---

<sup>28</sup> Refira-se, a propósito, que a emissão de *longevity bonds* anunciada pelo BEI em 2004 incorporava um *swap* de mortalidade.

## 11.5. Cobertura do risco de longevidade

---

de mortalidade não são contratos de seguro no sentido legal do termo, não estando por isso sujeitos às disposições que regem este tipo de acordos, regulando-se por disposições geralmente menos limitativas do mercado de capitais.

Os MS's constituem igualmente uma alternativa, com vantagens, às LB's. Desde logo pelo facto de poderem ser estruturados de forma mais simples e ajustada às necessidades das partes, e com maior facilidade na extinção das posições. Por outro lado, não requerem mercados secundários líquidos, bastando apenas que existam vantagens comparativas para ambas as partes ou perspectivas distintas quanto ao desenvolvimento da mortalidade no futuro.

Num mercado de derivados de mortalidade desenvolvido, podemos equacionar diferentes modalidades de MS, envolvendo o intercâmbio de apenas um pagamento dependente da longevidade (por analogia com o que acontece num *forward Rate Agreement* (FRA)), ou múltiplos pagamentos (como no caso de um clássico *plain vanilla interest rate swap* (IRS)).<sup>29</sup> Neste último caso, existem contudo algumas diferenças que importa realçar. Com efeito, enquanto que a componente fixa de um IRS especifica pagamentos constantes no tempo, a componente fixa de um *vanilla mortality swap* (VMS) envolveria pagamentos pré-determinados que oscilam no tempo de acordo com a evolução prevista (no momento de realização do acordo) de um índice de sobrevivência. Por outro lado, a componente variável de um IRS está, regra geral, ligada à evolução de um indexante (e.g., LIBOR), enquanto que a parcela variável de um VMS dependeria da evolução observada no índice de sobrevivência.<sup>30</sup>

Os MS's podem ser usados com inúmeras finalidades. As companhias de seguros podem usar um MS para fazer a cobertura dos riscos de mortalidade a que estão expostas as suas carteiras, explorando as vantagens comparativas que resultam da detenção de exposições opostas no activo subjacente, i.e., longas (seguros de vida) ou curtas (rendas vitalícias). Podemos inclusive conceber, como o fazem Lin e Cox (2005a,b), que companhias de seguros que exploram simultaneamente estes dois segmentos de negócio possam

---

<sup>29</sup>Podemos ainda conceber formas mais elaboradas ou exóticas de MS: MS's do tipo *floating-for-floating, swaps* sobre o diferencial de mortalidade entre duas populações, *cross-currency* MS's, MS's em que um ou mais pagamentos dependem da evolução de variáveis não demográficas (e.g., taxas de juro, índices do mercado de acções,...) e MS com opções incorporadas (e.g., opções). A um outro nível, Charupat e Milevsky (2001) apelidam de *mortality swap* uma estratégia de investimento que consiste na aquisição de uma renda vitalícia imediata e na utilização dos recebimentos periódicos gerados por este contrato no financiamento da compra de seguros de vida. Reportando-se ao caso canadiano, os autores concluem que a forma como o sistema fiscal do país está desenhado permite a existência de oportunidades de arbitragem, notando em particular que a estruturação de um *swap* de mortalidade gera um retorno superior à taxa de juro sem risco.

<sup>30</sup>Refira-se, ainda, que a avaliação de VMS se afigura, à partida, mais difícil que a de um IRS na medida em que, devido à natureza incompleta dos mercados, não é possível recorrer aos tradicionais argumentos de não arbitragem.

recorrer a MS's para explorar as oportunidades de cobertura natural criadas pela forma como as posições respondem a variações na mortalidade.<sup>31</sup>

Os *swaps* de mortalidade oferecem às companhias outras formas de gerir a sua exposição ao risco. Como os acordos de *swap* se reportam a um dado horizonte temporal e ao comportamento da longevidade numa dada população, podemos conceber uma situação em que as companhias utilizam este instrumento para ajustar a sua exposição entre populações de referência distintas, ou até mesmo ao longo de uma *estrutura temporal da mortalidade*. Por exemplo, companhias de seguros sediadas em diferentes países e expostas ao comportamento da mortalidade nas respectivas populações podem considerar vantajosa a participação num acordo de *swap*. O fundamento para este acordo é simples. Como os factores de risco não estão, muito provavelmente, perfeitamente correlacionados, ambas as companhias beneficiam com a diversificação da sua exposição entre distintas populações de referência.

Por outro lado, as companhias poderiam usar MS's para reduzir a sua exposição ao comportamento da mortalidade num determinado segmento do arco da vida humana, transformando as responsabilidades da carteira de modo a conseguir uma distribuição mais equilibrada em toda a estrutura temporal.

A um outro nível, os MS's podem atrair a atenção de investidores que, na busca de uma melhor relação risco-retorno, procuram incessantemente por novas oportunidades de investimento em activos com correlação baixa (preferencialmente negativa) relativamente às suas carteiras, como é o caso dos instrumentos financeiros contingentes à longevidade.

Por fim, e não menos importante, num mercado de derivados desenvolvido os MS's podem servir de base à implementação de estratégias de especulação, assentes em projecções sobre a evolução da longevidade no futuro, ou até mesmo sobre o seu comportamento em determinados segmentos da sua estrutura temporal.

### *Mortality Futures*

A criação de contratos de futuros sobre a mortalidade (*mortality futures* (MF)) enquanto instrumento de cobertura do risco de longevidade é sugerida por Blake *et al.* (2006a). À semelhança do que acontece na generalidade dos contratos de futuros, a criação de um MF envolve a especificação de um activo subjacente e de uma data de vencimento do contrato. Segundo os autores, entre os diferentes factores que condicionam o sucesso de um mercado desta natureza (dimensão, liquidez e volatilidade do mercado *spot*, custos

---

<sup>31</sup>Neste caso em particular, um *swap* do tipo *floating-for-floating* seria o mais apropriado: uma das componentes estaria ligada às responsabilidades da companhia com a carteira de rendas, a outra dependeria dos pagamentos a realizar na carteira de seguros.

de transacção, etc.) o principal desafio coloca-se na escolha de um activo subjacente (índice de sobrevivência, *longevity bond*, etc.) adequado de modo a garantir o sucesso do mercado. Por exemplo, Cairns *et al.* (2006a) e Blake *et al.* (2006a) discutem a possibilidade de criar futuros sobre rendas vitalícias.

Num cenário hipotético em que o mercado de LB's floresce, podemos igualmente equacionar a possibilidade de oferecer contratos de futuros em que estas obrigações se constituem como activo subjacente, uma solução que não é virgem se considerarmos os contratos de futuros sobre obrigações clássicas transaccionados com sucesso em bolsas de derivados um pouco por todo o mundo. Uma solução alternativa passaria por considerar um índice de sobrevivência como activo subjacente, replicando de certa forma o que acontece nos futuros sobre índices do mercado accionista. Esta solução tem como principal vantagem o facto de não exigir um mercado líquido e eficiente para o activo subjacente.

Dito isto, os contratos de futuros sobre a mortalidade poderiam servir de base à implementação de estratégias de arbitragem, de especulação e de cobertura do risco de longevidade, de forma análoga à observada nos mercados de derivados para outros activos subjacentes.

### Mortality Options

A criação de um mercado de opções sobre mortalidade (*mortality options* (MO)) exibindo as características e a diversidade das opções financeiras clássicas foi proposta por Blake *et al.* (2006a). Segundo estes autores, as MO's são especialmente interessantes (por comparação com os MF's) para as companhias que desejam limitar as perdas decorrentes de movimentos adversos na mortalidade, ao mesmo tempo que beneficiam de uma eventual evolução favorável nos factores de risco.

Podem ainda revelar-se atractivas para especuladores que implementam estratégias baseadas na evolução da volatilidade e não no nível das taxas de mortalidade (ou produtos relacionados, e.g., rendas vitalícias). Uma questão crucial na avaliação de MO's e na gestão do risco das posições prende-se com a utilização de um modelo de mortalidade estocástica adequado (veja-se a discussão no Capítulo 9). À semelhança do mencionado acima para os MF's, um aspecto fundamental no desenvolvimento do mercado reside na escolha do activo subjacente.

Se considerarmos um índice de sobrevivência  $\mathbb{I}_{S_x}(t)$  como activo subjacente e designarmos por  $\mathbb{I}_{S_x^{cap}}(t)$  e  $\mathbb{I}_{S_x^{floor}}(t)$  a *cap rate* e a *floor rate* para exercício no momento  $t$ , respectivamente, podemos imaginar a criação de *longevity caplets* e *longevity floorlets* com

payoffs no momento  $t$  definidos por  $\max\{\mathbb{I}_{S_x(t)} - \mathbb{I}_{S_x^{cap}(t)}; 0\}$  e por  $\max\{\mathbb{I}_{S_x^{floor}(t)} - \mathbb{I}_{S_x(t)}; 0\}$ , respectivamente. A combinação de várias *longevity caplets* e *longevity floolets* com diferentes datas de exercício permite-nos igualmente engendrar a criação de *longevity caps* e *longevity floors*, de forma análoga ao que acontece nos mercados de opções sobre taxas de juro.

A utilização de um índice  $\mathbb{I}_{S_x(t)}$  em produtos financeiros apresenta, no entanto, uma debilidade fundamental. Com efeito, neste caso o activo subjacente seria observado com muito pouca frequência (e.g., uma vez por ano) e com erro apresentando, previsivelmente, pouca volatilidade.

Em alternativa ao uso de um índice  $\mathbb{I}_{S_x(t)}$ , podemos equacionar a utilização de outras variáveis como activo subjacente. Desde logo o recurso a *longevity bonds*. Outra solução passa por criar *mortality futures options* (MFO) usando os próprios preços de contratos de futuros sobre a mortalidade. Uma hipótese ainda mais exótica, considerada por Blake *et al.* (2006a), passa por especificar uma espécie de *mortality swaption* (MSW), um contrato em que o activo subjacente seria um *swap* de mortalidade, de um determinado tipo e maturidade.

Em suma, a criação de um mercado de opções colocaria ao dispor das companhias de seguros, dos especuladores e dos investidores em geral instrumentos poderosos na gestão do risco de longevidade.

### 11.5.7 Mercado de rendas vitalícias e cobertura individual do risco de longevidade

Nos tradicionais sistemas de segurança social do tipo *pay-as-you-go* (PAYGO), a cobertura do risco de longevidade é assegurada pela provisão de uma pensão pública de benefício definido, determinada em função de um conjunto de parâmetros do sistema (e.g., período contributivo do indivíduo, carreira salarial, idade de reforma,...). Esta promessa de benefícios futuros, mais ou menos credível consoante o grau de insustentabilidade económica e financeira do sistema, configura, na prática, uma renda vitalícia, regra geral de termos variáveis.

Os benefícios recebidos sob a forma de uma pensão desempenharam, historicamente, uma função importante ao garantir aos cidadãos um seguro base contra situações de ausência completa de rendimentos na velhice. As tendências demográficas registadas nas últimas décadas, caracterizadas por um envelhecimento acentuado, pela diminuição da taxa de fecundidade, pela diminuição da relação entre trabalhadores activos e dependentes e pelo significativo aumento da esperança média de vida à nascença e nas

## 11.5. Cobertura do risco de longevidade

---

idades adultas, ameaçam a sustentabilidade dos tradicionais sistemas de segurança social assentes no mecanismo de repartição contemporânea e tornam cada vez mais difícil a manutenção dos actuais níveis de vida das populações após o fim da sua vida activa.

Financiar o aumento da longevidade e garantir a liberdade de escolha dos cidadãos em relação aos seus rendimentos de ciclo de vida, em especial durante o período de desacumulação de poupança, criando-lhes condições para o exercício pleno dos seus direitos sociais, são alguns dos principais dilemas com que se deparam os actuais modelos sociais.

No domínio da Segurança Social, a tendência para a implementação de novos modelos, assentes numa estrutura multipilar que combina os tradicionais mecanismos de repartição intergeracional com sistemas complementares públicos e/ou privados de contribuição definida, baseados em contas individuais em regime de capitalização (real ou virtual), vai de encontro às doutrinas que sustentam que a actuação do Estado social deve ser pautada pelo respeito do princípio da subsidiariedade, neste caso no que toca à relação entre o Estado e o sector privado na oferta de protecção de rendimentos na velhice.

Na prática, o fim do papel monopolista do Estado neste domínio, com a criação de mecanismos de concorrência efectiva entre os vários subsistemas, assegurando a liberdade de escolha dos cidadãos e a sua co-responsabilização pelos rendimentos na reforma depende, em boa medida, da existência no mercado de instrumentos financeiros que permitam uma adequada cobertura do risco de longevidade.

Nos sistemas do tipo PAYGO esta função é, como vimos, assegurada pela provisão de uma pensão pública de benefício definido. A moderação dos benefícios concedidos pelos sistemas de repartição, a progressiva conversão dos planos de pensões de benefício definido em planos de contribuição definida e a tendência para a adopção de sistemas de segurança social assentes em mecanismos de capitalização, tenderão a reduzir a parcela do rendimento na reforma que é “garantida”. Neste cenário, a criação de um mercado privado de rendas vitalícias competitivo e eficiente emerge como um substituto natural da função de seguro social desempenhada pelas pensões públicas.<sup>32</sup>

Os ganhos (teóricos) de bem-estar decorrentes da possibilidade de converter a riqueza acumulada numa renda vitalícia são bem conhecidos desde o artigo pioneiro de Yaari (1965).<sup>33</sup> Com base numa série de hipóteses fortes (nomeadamente que os consumidores

---

<sup>32</sup>Saliente-se que o papel da família emerge, neste contexto, como um segundo substituto natural, sendo que esse papel pode ser descrito como um mercado de rendas incompleto, onde é possível implementar um conjunto de vínculos informais que proporcionam muitos dos benefícios associados à compra de uma renda. Refira-se, no entanto, que nas sociedades modernas, maioritariamente urbanas, o papel da família na provisão de um seguro de velhice diminui consideravelmente face às tradicionais sociedades rurais.

<sup>33</sup>Para uma discussão actualizada sobre a importância das rendas vitalícias enquanto mecanismo de cobertura do risco de longevidade e de desacumulação da poupança veja-se, por exemplo, Brown *et al.* (2001) e referências aí citadas.

## 11.5. Cobertura do risco de longevidade

---

maximizam a sua utilidade esperada, que as suas preferências são independentes do tempo, que o único risco que enfrentam é o risco de longevidade, que a cobertura total deste risco é possível através do mercado de rendas, que o preço destas é actuarialmente justo, que existe um único activo que remunera a uma dada taxa de juro, e que os consumidores podem emprestar e pedir emprestado a esta taxa), o modelo de Yaari conclui que a aquisição de rendas vitalícias é uma decisão óptima na ausência de motivações relacionadas com heranças.

O principal atractivo das rendas vitalícias é o de que estas oferecem aos indivíduos avessos ao risco a possibilidade de se protegerem contra o risco de esgotarem os seus activos em vida (ou simplesmente contra o risco de diminuição do rendimento na reforma), em troca do pagamento de um prémio (inicial ou periódico).

Num mundo sem incerteza, e na ausência de qualquer sistema formal de reforma, o modelo do ciclo de vida estipula que os indivíduos racionais tomam as suas decisões de consumo e poupança de forma óptima. Uma vez chegados à reforma, dividem a sua riqueza acumulada pelos anos de vida restantes, de modo a garantir um nível de consumo óptimo (e a salvaguardar eventuais preocupações para com os descendentes).

Na realidade, existem muitas fontes de incerteza que dificultam a escolha do nível óptimo de consumo na reforma. Um dos principais riscos é aquele que decorre de uma longevidade incerta: os reformados podem sobreviver para além dos horizontes esperados, esgotando a sua riqueza em vida. Neste cenário, os reformados podem sempre reduzir as suas despesas de consumo, mas tal estratégia aumenta a probabilidade de que se esteja a sacrificar “excessivamente” o nível de vida, não aproveitando de forma óptima a riqueza acumulada.

Ao proporcionar a um indivíduo avesso ao risco um rendimento certo enquanto este sobreviver, a compra de uma renda vitalícia elimina o risco de mortalidade individual enquanto fonte de incerteza financeira e aumenta as garantias quanto à manutenção do nível de vida e de um padrão de consumo estáveis durante todo o ciclo de vida.

Um mercado de rendas vitalícias competitivo deve ser capaz de oferecer aos cidadãos um leque de produtos financeiros variado, proporcionando fluxos de rendimento alternativos, estruturados segundo planos de prémios ajustados às necessidades e motivações de cada indivíduo, gerando benefícios em termos nominais ou reais, constantes ou variáveis, incluindo eventualmente garantias adicionais que cubram, por exemplo, as despesas de saúde na velhice.

Do ponto de vista colectivo, na ausência de um mercado privado de rendas vitalícias os indivíduos podem ter um incentivo para gastar a sua riqueza financeira e patrimonial muito rapidamente face aos horizontes de vida esperados, sabendo que em caso de

## 11.5. Cobertura do risco de longevidade

---

esgotarem prematuramente as suas reservas podem sempre reclamar dos governos benefícios adicionais (e.g., pensões de sobrevivência) para fazerem face a situações de pobreza extrema. Por outras palavras, na ausência de um mercado de rendas, os tradicionais fenómenos de risco moral e miopia geracional podem acentuar-se, limitando a efectiva liberdade de escolha dos cidadãos.

Apesar dos inquestionáveis ganhos de bem-estar associados à aquisição de rendas e do potencial de crescimento existente, a verdade é que nos países desenvolvidos os mercados privados de rendas vitalícias se encontram ainda num estágio de relativo subdesenvolvimento. As explicações para este facto abundam na literatura.<sup>34</sup> O facto de as rendas vitalícias terem sido, durante muito tempo, encaradas como um instrumento financeiro de acumulação de poupança e não como um mecanismo de cobertura contra o risco de longevidade na reforma ajuda, em parte, a explicar este fenómeno. A verdade, todavia, é que existem um conjunto de factores que, actuando ora ao nível da oferta, ora ao nível da procura, condicionam de forma decisiva a criação e desenvolvimento de um mercado de rendas líquido, competitivo e transparente.

Ao nível da procura, a inexistência de um mercado competitivo e a adopção de práticas conservadoras em termos de *pricing* faz com que os indivíduos percepcionem muitas vezes as rendas vitalícias como sendo “caras” face aos benefícios que proporcionam em termos de retorno e, sobretudo, de protecção contra o risco de longevidade.

Por outro lado, nos mercados competitivos modernos as decisões financeiras dos cidadãos são influenciadas pela sua percepção quanto à atractividade de diferentes alternativas de investimento, não apenas durante a vida activa, mas também após a reforma. Por exemplo, alguns indivíduos não encaram a compra de rendas vitalícias sob o argumento de que são capazes de gerir melhor os seus activos do que os gestores de fundos.

Alguns autores defendem que o *timing* de aquisição de uma renda deve ser postecipado para as idades mais avançadas, de modo a beneficiar do potencial de valorização das carteiras, enquanto outros argumentam que os reformados devem investir as suas poupanças em activos financeiros e não em rendas, uma vez que assim conseguem lidar melhor com a inflação.

Neste sentido, uma das formas de estimular a compra de seguros contra o risco de longevidade passa, por exemplo, pela criação de incentivos fiscais que sinalizem uma preocupação pública para com este tipo de risco e desequilibrem as escolhas em favor da compra de rendas vitalícias.

Outro dos motivos pelos quais os indivíduos podem ser levados a não comprar rendas

---

<sup>34</sup>Veja-se, por exemplo, Mitchell *et al.* (1999), Blake (1999), James e Vittas (2000), Mitchell e McCarthy (2002) e Davidoff *et al.* (2005).

prende-se com o desejo de reservar uma fracção significativa da sua riqueza em activos com liquidez elevada para fazer face, por exemplo, a despesas significativas inesperadas com cuidados de saúde, uma opção não contemplada nas rendas clássicas. Esta preocupação deve ser tida em conta pela indústria seguradora e pode ser mitigada, por exemplo, desenvolvendo produtos estruturados que combinem rendas vitalícias com coberturas de saúde e cláusulas de liquidez apropriadas.

Outro exemplo interessante é o das denominadas hipotecas reversíveis (*reverse mortgages*), em que o valor da habitação subjacente a um crédito hipotecário pode, em determinado momento, ser convertido numa determinada quantia monetária (*lump-sum*) ou numa renda vitalícia paga enquanto o indivíduo sobreviver (ou uma combinação destas duas opções), assegurando ao mesmo tempo o direito a habitar o imóvel.

Por último, e não menos importante, o desenvolvimento do mercado de rendas carece de uma política activa de educação dos cidadãos e do público em geral, quanto às características dos produtos e suas potencialidades em termos de cobertura do risco de longevidade. Tal pode ser conseguido, por exemplo, através de campanhas públicas de esclarecimento, de difusão de informação através das novas tecnologias ou pela via da actuação das associações de defesa do consumidor/investidor.

Ao nível da oferta, a principal razão apresentada na literatura para a exiguidade do mercado de rendas prende-se com a existência de problemas sérios de selecção adversa. Com efeito, existe uma ampla evidência (confirmada, como vimos no Capítulo 8, no caso português) que sustenta que os indivíduos que compram rendas vitalícias apresentam, em média, uma longevidade superior aos que não o fazem, o que leva as seguradoras a aumentar os preços de modo a compensar estes diferenciais de mortalidade. Nalguns países (e.g., Reino Unido) a solução encontrada para lidar com este fenómeno passou pela adopção de tábuas de mortalidade específicas para este tipo de populações, em detrimento das adoptadas para a população em geral.

A selecção adversa pode ser entendida como uma falha de mercado que pode, em princípio, ser regulada pela intervenção do Estado, nomeadamente impondo a todos os indivíduos a obrigatoriedade de converter uma determinada parcela da sua riqueza acumulada em renda. Esta solução apresenta algumas vantagens, nomeadamente o aumento da cobertura face aos riscos de longevidade, a redução dos fenómenos de selecção adversa e dos problemas derivados da existência de informação assimétrica entre seguradores e segurados, contribuindo igualmente para eliminar os riscos de mortalidade idiossincráticos. Adicionalmente, a conversão em renda (anuitização) promove a transferência dos riscos de mercado (investimento) dos cidadãos para entidades especializadas e o diferimento do momento de pagamento de impostos.

Em contrapartida, a conversão obrigatória em renda apresenta algumas desvantagens que não podem ser negligenciadas. Em particular, ela aumenta o risco de afectação sub-óptima da riqueza, pode promover uma redistribuição indesejada de rendimentos entre grupos da população, favorecendo nomeadamente os grupos de maior rendimento (e maior esperança de vida), expõe os indivíduos ao risco de insolvência das seguradoras, negligencia a importância de motivações pessoais em relação ao legado patrimonial a deixar aos descendentes e tende a aumentar os custos burocráticos e administrativos.

O desenvolvimento adequado da oferta no mercado de rendas carece ainda da existência de bases técnicas adequadas ao *pricing* dos produtos, não apenas em termos de modelos de longo prazo para o comportamento das taxas de juro, mas sobretudo em termos de tábuas de mortalidade que permitam estimar correctamente a longevidade de cada geração. A este propósito, nos capítulos anteriores argumentámos que as tábuas a adoptar devem ser prospectivas, atendendo à natureza dinâmica do fenómeno, e considerar desejavelmente as diferenças entre gerações. De facto, a incerteza quanto ao comportamento da mortalidade leva, na prática, as companhias de seguros a adoptarem uma posição conservadora no *pricing* das rendas, desincentivando por esta via o acesso ao mercado e agravando os já mencionados problemas de selecção adversa.

Por outro lado, a inexistência de estimativas credíveis quanto à longevidade humana leva muitas vezes os indivíduos a subestimar seriamente as suas perspectivas de sobrevivência levando-os, por seu turno, a subavaliar a importância da cobertura contra o risco de longevidade. Por exemplo, nalguns países (e.g., Reino Unido) as entidades reguladoras exigem (por razões constitucionais) a utilização de tábuas de mortalidade unisexo para calcular os benefícios associados aos planos de pensões, uma solução que gera uma espécie de subsídio cruzado entre reformados do sexo masculino e feminino e não favorece a transparência do mercado. Em suma, a disseminação de informação credível sobre a longevidade humana deve ser considerada uma matéria de interesse público e fazer parte da política de supervisão das autoridades competentes.

Ainda no lado da oferta, uma das principais limitações ao desenvolvimento do mercado refere-se à (quase) inexistência de activos que sirvam de base à adopção de estratégias de cobertura contra variações nos factores de risco de longo prazo subjacentes. Com efeito, as companhias que oferecem rendas no mercado enfrentam duas grandes fontes de risco: risco de taxa de juro e, principalmente, risco de longevidade.

Em relação ao risco de taxa de juro, a teoria da imunização financeira clássica sugere que, para proteger a carteira contra choques inesperados na estrutura temporal de taxas de juro, as companhias devem ajustar as carteiras de activos de modo a que as suas características em termos de risco (e.g., duração, *stochastic duration*, *M-Square*,...)

## 11.5. Cobertura do risco de longevidade

---

igualem as da carteira de responsabilidades com rendas, salvaguardadas as restrições em relação às condições de segunda ordem (convexidades).<sup>35</sup>

Na prática, a implementação desta estratégia é difícil, na medida em que requer a negociação nos mercados de obrigações com maturidade muito longa. Acresce ainda que a oferta de rendas vitalícias com garantia de rendimento real exige que no mercado estejam igualmente disponíveis obrigações com cupões indexados à inflação, algo que não se verifica na generalidade dos países. Em suma, a emissão de obrigações de dívida pública de longo prazo contribui para criar condições para o desenvolvimento do mercado de rendas, pese embora exemplos recentes (e.g., Rússia e Argentina) chamem a atenção para a necessidade de monitorizar a qualidade dos activos envolvidos face ao risco de incumprimento acrescido que esta solução representa.

Por outro lado, o risco de longevidade, i.e., a possibilidade de gerações inteiras viverem sistematicamente mais do que o projectado no passado, afigura-se como uma das principais preocupações com que se deparam as companhias que desejam expandir a sua área de negócio no mercado de rendas. Tal é assim porquanto aumentos inesperados na longevidade configuram riscos sistemáticos (não diversificáveis), que ameaçam seriamente a rentabilidade e a solvência das empresas.

Conforme mencionámos nas secções anteriores, as companhias de seguros dispõem de alguns mecanismos para minimizar a exposição a este risco, como sejam a venda simultânea de seguros de vida e rendas, a adopção de estratégias de diversificação entre grupos de idades, de diversificação internacional (e.g., através de contratos de resseguro), a possibilidade de partilha do risco com os beneficiários das rendas mas não conseguem, regra geral, eliminar completamente o risco de longevidade. Tal só se tornará possível, como salientámos anteriormente, com o desenvolvimento de um mercado de instrumentos financeiros contingentes à evolução da longevidade.

A um outro nível, o desenvolvimento do mercado de rendas exige a criação de uma política de regulação e de solvência adequadas, que ofereça garantias de segurança aos cidadãos que adquirem este tipo de produtos, que regule as condições de solvência das companhias de seguros que operam no mercado e que monitorize o risco moral das instituições. De facto, a necessidade de garantir o provisionamento e o pagamento de responsabilidades a muito longo prazo coloca novos desafios à regulação nestes mercados.

Em suma, as alterações demográficas e as modificações nos sistemas públicos e privados de segurança social permitem perspectivar um crescimento significativo na procura de rendas e de soluções que cubram o risco de uma longevidade acrescida. Em paralelo,

---

<sup>35</sup>Para uma discussão detalhada sobre as estratégias de imunização financeira considerando modelos unifactoriais e multifactor e modelos estocásticos veja-se, por exemplo, Bravo (2002).

### 11.5. Cobertura do risco de longevidade

---

é de prever igualmente um crescimento da oferta neste mercado, sendo que para tal os poderes públicos tem um papel fundamental, seja pela divulgação e disseminação de informação de qualidade sobre a mortalidade, pela criação de incentivos fiscais, pelo desenvolvimento do mercado de derivados de mortalidade ou pela sensibilização dos cidadãos em relação à necessidade de se protegerem em relação a uma longevidade acrescida.

## Capítulo 12

# Conclusão

“Se não receio o erro, é porque estou sempre disposto a corrigi-lo.” (Bento de Jesus Caraça, Matemático *Calipolense*, 1901-1948)

Os avanços na medicina, a melhoria nos níveis de vida, a tendência para a adopção de estilos de vida mais saudáveis e a ausência de conflitos militares globais e de surtos epidémicos graves explicam o significativo aumento da longevidade humana registado nas últimas décadas. Estas tendências, combinadas com níveis de fertilidade bastante baixos desafiam as sociedades, as empresas e os indivíduos a lidar com o risco de longevidade, refundando os tradicionais sistemas de segurança social do tipo *pay-as-you-go*, aumentando a importância do sector privado na oferta de soluções financeiras adequadas às fases de acumulação e desacumulação do ciclo de vida, adaptando os pressupostos actuariais usados no *pricing* e no *hedging* dos riscos e desenvolvendo mecanismos mais eficientes e transparentes de cobertura dos riscos de mortalidade e de longevidade.

Nesta dissertação desenvolvemos um conjunto de métodos, técnicas e algoritmos visando a medição e gestão dos riscos de mortalidade e de longevidade. As ferramentas discutidas ao longo do documento são especialmente úteis em aplicações de natureza actuarial mas podem ser adoptadas em múltiplos domínios, em particular em exercícios de projecção demográfica, em estudos de saúde pública, na biologia, entre outros.

Os resultados obtidos nos estudos de projecção da longevidade da população portuguesa confirmaram a existência de diferenças significativas entre os valores da esperança de vida calculados segundo as ópticas transversal e longitudinal. Em contextos de aumento da longevidade, a adopção de pressupostos actuariais clássicos conduz assim a uma sistemática subestimação dos prémios e das responsabilidades nos produtos com benefícios de longo prazo em caso de vida. Este facto torna cada vez mais urgente a necessidade de Portugal adoptar tábuas de mortalidade prospectivas regulamentares,

homologadas pelas respectivas autoridades de supervisão.

Os estudos empíricos sobre a importância do efeito selecção adversa na carteira de beneficiários de fundos de pensões em Portugal permitiram-nos derivar tábuas de mortalidade prospectivas para esta subpopulação e detectar a presença de uma clara submortalidade nos indivíduos do sexo masculino face à registada na população masculina portuguesa em geral. A curta dimensão da série recomenda a sua repetição no futuro para confirmação dos resultados agora obtidos.

A modelação da mortalidade através de processos estocásticos representa um salto importante na aplicação de metodologias condizentes com a verdadeira natureza do fenómeno: dinâmico e com trajectória incerta no futuro. Ela vai igualmente de encontro às orientações mais recentes das autoridades de supervisão na área dos seguros e fundos de pensões que recomendam a explicitação de todos os factores de risco (financeiros, demográficos, operacionais,...) a que estão expostas as carteiras e a avaliação das responsabilidades pelo seu justo valor, incorporando princípios consolidados na avaliação de activos financeiros. A progressiva introdução destes princípios na ciência actuarial constitui um marco importante e um sinal da sua crescente interacção com o corpo teórico produzido em economia financeira.

Por outro lado, a quantificação da componente sistemática do risco de mortalidade coloca-nos em condições de investigar novas soluções de cobertura do risco, que passem nomeadamente pela sua transferência para outras contrapartes, em especial para o mercado de capitais. A este nível, é previsível que num futuro próximo se desenvolva a nível mundial um mercado de produtos com cash flows contingentes à evolução da longevidade, correspondendo desta forma às necessidades dos intervenientes no mercado.

No decurso deste longo projecto, foram sendo identificadas um conjunto de novas linhas de investigação que pretendemos explorar num futuro próximo. Elas compreendem o desenvolvimento de novas variantes dos modelos de projecção da mortalidade analisados neste documento, o aprofundamento da modelação da mortalidade num contexto estocástico e o teste empírico dos modelos.

Chegado a este ponto, só me resta felicitar o leitor que ultrapassou o “Cabo das Tormentas” e “resistiu” até à última página. Espero, sinceramente, que todo o investimento realizado tenha, no futuro imediato, o retorno almejado.

# Bibliografia

- [1] Ammann, M. (2001). *Credit Risk Valuation: Methods, Models and Applications*. 2nd Edition, Springer, New York.
- [2] Anson, J. (1988). The parameters of death: a consideration of the quantity of information in a life table using a polynomial representation of the survivorship curve. *Statistics in Medicine*, 7, 895-912.
- [3] Artzner, P. e Delbaen, F. (1995). Default risk insurance and incomplete markets. *Mathematical Finance*, 5(3), 187-195.
- [4] Babbage, C. (1823). On the Tables of Single and Annual Assurance Premiums. *Journal of the Institute of Actuaries*, 6, 185.
- [5] Bacinello, A. e Persson, S. (2002). Design and pricing of equity-linked life insurance under stochastic interest rates. *The Journal of Risk and Insurance*, 3(2), 6-21.
- [6] Ballotta, L. e Haberman, S. (2003). Valuation of guaranteed annuity conversion options. *Insurance: Mathematics and Economics*, 33 (1), 87-108.
- [7] Ballotta, L. e Haberman, S. (2006). The fair valuation problem of guaranteed annuity options: The stochastic mortality environment case. *Insurance: Mathematics and Economics*, 38, 195-214.
- [8] Batten, R. W. (1978). *Mortality table construction*. Prentice-Hall Inc, Englewood Cliffs, New Jersey.
- [9] Bauer, D. e Russ, J. (2006). Pricing longevity bonds using implied survival probabilities. *Working Paper*, Ulm University and Institut für Finanz und Aktuarwissenschaften.
- [10] Bayraktar, E. e Young, V. (2007). Hedging Life Insurance with Pure Endowments. *Insurance Mathematics and Economics*, forthcoming.
- [11] Beard, R. (1959). Note on some mathematical mortality models. In: G. E. W. Wolstenholme and M.O'Connor (Eds.). *The Lifespan of Animals*. Little, Brown, Boston, 302-311.

- [12] Beard, R. (1971). Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In W. Brass (Ed.). *Biological Aspects of Demography*. Taylor & Francis, London.
- [13] Beelders, O. e Colarossi, D. (2004). Modelling mortality risk with extreme value theory: The case of Swiss Re's mortality-indexed bond. *Global Association of Risk Professionals*, 4, 26-30.
- [14] Bell, W. (1997). Comparing and assessing time series methods for forecasting age specific demographic rates. *Journal of Official Statistics*, 13, 279-303.
- [15] Benjamin, B. e Pollard, J. (1980). *The analysis of mortality and other actuarial statistics*. Heinemann, London.
- [16] Benjamin, B. e Soliman, A. (1993). *Mortality on the Move: methods of mortality projection*. City University, London.
- [17] Bielecki, T. e Rutkowski, M. (2002). *Credit Risk: Modelling, Valuation and Hedging*. Springer, New York.
- [18] Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37, 443-468.
- [19] Biffis, E. e Denuit, M. (2005). Lee Carter Goes Risk-Neutral: An Application to the Italian Annuity Market. *Actuarial Research Paper N.º 166*, Faculty of Actuarial Science and Statistics, Cass Business School, City University, London.
- [20] Biffis, E., Denuit, M. e Devolder, P. (2006). Stochastic mortality under measure changes. *Actuarial Research Paper*, Faculty of Actuarial Science and Statistics, Cass Business School, City University, London.
- [21] Biffis, E. e Millosovich, P. (2004). A Bidimensional Approach to Mortality Risk. *Working Paper*, Università Bocconi
- [22] Biffis, E. e Millosovich, P. (2006). The fair value of guaranteed annuity options. *Scandinavian Actuarial Journal*, 1, 23-41.
- [23] Björk, T. (1998). *Arbitrage Theory in Continuous Time*. Oxford University Press Inc., New York.
- [24] Black, F., e Karasinski, P. (1991). Bond and Option Pricing When Short Rates are Lognormal. *Financial Analyst Journal*, July-August, 52-59.
- [25] Blake, D. (1999). Annuity Markets: Problems and Solutions. *The Geneva Papers on Risk and Insurance*, 24 (3), 358-375.
- [26] Blake, D., (2003). Reply to "Survivor Bonds: A Comment on Blake and Burrows". *The Journal of Risk and Insurance*, 70, 349-351.

- [27] Blake, D. e Burrows, W. (2001). Survivor Bonds: Helping to Hedge Mortality Risk. *The Journal of Risk and Insurance*, 68, 339-348.
- [28] Blake, D., Cairns, A. e Dowd, K. (2003). Pensionmetrics II: Stochastic pension plan design during the distribution phase. *Insurance: Mathematics and Economics*, 33, 29-47.
- [29] Blake, D., Cairns, A. e Dowd, K. (2006a). Living with mortality: longevity bonds and other mortality-linked securities. *British Actuarial Journal*, forthcoming
- [30] Blake, D., Cairns, A., Dowd, K. e MacMinn, R. (2006b). Longevity Bonds: Financial Engineering, Valuation, and Hedging. *The Journal of Risk and Insurance*, 73(4), 647-672.
- [31] Blaschke, E. (1923). Sulle tavole di mortalità variabili col tempo. *Giornale di Matematica Finanziaria*, 5, 1-31.
- [32] Bloomfield, S. e Haberman, S. (1987). Graduation: some experiments with kernel methods. *Journal of the Institute of Actuaries*, 114, 339-369.
- [33] Boleslawski, L. e Tabeau, E. (2001). Comparing theoretical age patterns of mortality beyond the age of 80. In E. Tabeau *et al.* (Eds). *Forecasting Mortality in Developed Countries: insights from a statistical, demographical and epidemiological perspective*. Kluwer Academic Publishers, 127-155.
- [34] Bongaarts, J. (2004). Long-Range Trends in Adult Mortality: Models and Projection Methods. Policy Research Division Population Council, *Working Paper* 192.
- [35] Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22 (3), 547-581.
- [36] Booth, H., Maindonald, J. e Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56, 325-336.
- [37] Bourgeois-Pichat, J. (1952). Essai sur la mortalité biologique de l'homme. *Population*, 7(3), 1-31.
- [38] Bowers, N., Gerber, H., Hickman, J., Jones, D. e Nesbitt, C. (1997). *Actuarial Mathematics*. The Society of Actuaries, Schaumburg, Illinois.
- [39] Box, G. e Pierce, D. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 65, 1509-1526.
- [40] Boyle, P. e Hardy, M. (2003). Guaranteed annuity options. *ASTIN Bulletin*, 33 (2), 125-152.
- [41] Brace, A., Gatarek, D. e Musiela, M. (1997). The Market Model of Interest Rate Dynamics. *Mathematical Finance*, 7, N° 2, 127-55.

- [42] Brass, W. (1971). On the scale of mortality. In: *Biological Aspects of Demography*, London Taylor and Francis.
- [43] Brass, W. (1974). Mortality models and their uses in demography. *Transactions of the Faculty of Actuaries*, 33, 123-132.
- [44] Braumann, C. (2005). *Introdução às Equações Diferenciais Estocásticas e Aplicações*. Edições Sociedade Portuguesa de Estatística, Ericeira.
- [45] Bravo, J. M. (2002). *Modelos de Risco de Taxa de Juro: Estratégias de Cobertura e Imunização*. Dissertação de Mestrado em Economia Monetária e Financeira, Instituto Superior de Economia e Gestão – Universidade Técnica de Lisboa.
- [46] Brémaud, P. (1981). *Point Processes and Queues – Martingale Dynamics*. Springer-Verlag, New York.
- [47] Brigo, D. e Mercurio, F. (2001). *Interest Rate Models: Theory and Practice*. Springer-Verlag, Berlin.
- [48] Brillinger, D. (1961). A justification for some common laws of mortality. *Transaction of the Society of Actuaries*, 13, 116-119.
- [49] Brillinger, D. (1986). The natural variability of vital rates and associated statistics. *Biometrics*, 42, 693-734.
- [50] Brockett, P. (1991). Information Theoretic Approach to Actuarial Science. A Unification and Extension of Relevant Theory and Applications. *Transactions of the Society of Actuaries*, XLIII, 73-136.
- [51] Brouhns, N., Denuit, M. e Keilegom, V. (2005). Bootstrapping the Poisson log-bilinear model for mortality projection. *Scandinavian Actuarial Journal*, forthcoming.
- [52] Brouhns, N., Denuit, M. e Vermunt, J. (2002a). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31, 373-393.
- [53] Brouhns, N., Denuit, M. e Vermunt, J. (2002b). Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 105-130.
- [54] Brown, J. e Orzag, P. (2006). The Political Economy of Government-Issued Longevity Bonds. *The Journal of Risk and Insurance*, 73(4), 611-631.
- [55] Brown, J., Mitchell, O., Poterba, J. e Warshawsky, M (2001). *The role of annuity markets in financing retirement*. MIT Press.
- [56] Buettner, T. (2002). Approaches and experiences in projecting mortality patterns for the oldest-old. *North American Actuarial Journal*, 6(3), 14-25.

- [57] Butt, Z. e Haberman, S. (2004). Application of frailty-based mortality models using generalizad linear models. *ASTIN Bulletin*, 34(1), 175-197.
- [58] Cairns, A., Blake, D. e Dowd, K. (2005). Pricing the Risk on Longevity Bonds. *Life and Pensions*, October, 41-44.
- [59] Cairns, A., Blake, D. e Dowd, K. (2006a). Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk. *ASTIN Bulletin*, 36, 79-120.
- [60] Cairns, A., Blake, D. e Dowd, K. (2006b). A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. *The Journal of Risk and Insurance*, 73(4), 687-718.
- [61] Carriere, J. (1992). Parametric models for life tables. *Transactions of the Society of Actuaries*, 44, 77-99.
- [62] Carriere, J. (1994). An investigation of the Gompertz law of mortality. *ARCH*, 2, 161-177.
- [63] Carrilho, M. (1980). Tábuas Abreviadas de Mortalidade, 1941-1975. INE, *Série Estudos*, N.º 56.
- [64] Carrilho, M. e Patrício, L. (2004). Tábuas de mortalidade em Portugal. INE, *Revista de Estudos Demográficos*, N.º 36, 2º Semestre.
- [65] Carter, L. (1996). Forecasting U.S. mortality: a comparison of Box-Jenkins ARIMA and structural time series models. *The Sociological Quarterly*, 37(1), 127-144.
- [66] Carter, L. e Prskawetz, A. (2001). Examining structural shifts in mortality using the Lee-Carter method. *Max Planck Institute for Demographic Research* WP 2001-007, Germany.
- [67] Chan, F., Chan, L., Falkenberg, J. e Yu, M. (1986). Applications of linear and quadratic programmings to some cases of the Whittaker-Henderson graduation method. *Scandinavian Actuarial Journal*, 1, 141-153.
- [68] Charupat, N. e Milevsky, M. (2001). Mortality Swaps and Tax Arbitrage in the Canadian Insurance and Annuity Markets. *The Journal of Risk and Insurance*, 68(2), 277-302.
- [69] Chen, N. e Kou, S. (2005). Credit spread, optimal capital structure, and implied volatility with endogenous default and jump risk. *Working Paper*, Columbia University.
- [70] Cheung, S., e Robine, J. (2007). Increase in common longevity and the compression of mortality: The case of Japan. *Population Studies*, 61(1), 85-97.

- [71] Cleveland, W. (1979). Robust Locally-Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74, 829-836.
- [72] CMIB (1974). Considerations affecting the preparation of standard tables of mortality. *Journal of the Institute of Actuaries*, 101, 135-216.
- [73] CMIB (1976). *Mortality Investigation Reports*, CMIR 2. The Institute of Actuaries and Faculty of Actuaries, UK.
- [74] CMIB (1978). *Mortality Investigation Reports*, CMIR 3. The Institute of Actuaries and Faculty of Actuaries, UK.
- [75] CMIB (1990). *Mortality Investigation Reports*, CMIR 10. The Institute of Actuaries and Faculty of Actuaries, UK.
- [76] CMIB (1999). *Mortality Investigation Reports*, CMIR 17. The Institute of Actuaries and Faculty of Actuaries, UK.
- [77] CMIB (2005). Projecting Future Mortality: Towards a proposal for a stochastic methodology. *Working Paper N.º 15*, 2005.
- [78] Coale, A. e Demeny, P. (1966). *Regional model life tables and stable population*. Princeton University Press, Princeton.
- [79] Coale, A. e Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index*, 55, 613-643.
- [80] Coale, A. e Kisker, E. (1990). Defects in data on old age mortality in the United States: new procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. *Asian and Pacific Population Forum*, 4, 1-31.
- [81] Cochrane, J. (2005). *Asset Pricing*. Princeton University Press, Princeton.
- [82] Coelho, E. (2001). *O método Lee-Carter para a modelização e previsão da mortalidade*. Dissertação de Mestrado em Estatística e Gestão da Informação, Lisboa, ISEGI.
- [83] Congdon, P. (1993). Statistical graduation in local demographic analysis and projection. *Journal of the Royal statistical Society*, 156(2), 237-270.
- [84] Copas, J. e Haberman, S. (1983). Non-parametric graduation using kernel methods. *Journal of the Institute of Actuaries*, 110, 135-156.
- [85] Cossette, H. Delwarde, A., Denuit, M., Guillot, F. e Marceau, E. (2005). Pension plan valuation and dynamic mortality tables. Working paper 04-17, *Institut des Sciences Actuarielles*, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

- [86] Cowley, A., e Cummins, J. (2005). Securitization of Life Insurance Assets and Liabilities. *The Journal of Risk and Insurance*, 72, 193-226.
- [87] Cox, D. (1972). Regression Models and Life Tables (with discussion). *Journal of the Royal Statistical Society*, B34, 187-220.
- [88] Cox, J., Ingersoll, J. e Ross, S. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385-407.
- [89] Cox, S. e Lin, Y. (2004). Natural Hedging of Life and Annuity Mortality Risks. *Proceedings of the 14th AFIR Colloquium*, Boston, 483-487.
- [90] Cox, S., Lin, Y. e Wang, S. (2006). Multivariate exponential tilting and pricing implications for mortality securitization. *The Journal of Risk and Insurance*, 73(4), 719-736.
- [91] Cramér, H. e Wold, H. (1935). Mortality variations in Sweden: a study in graduation and forecasting. *Skandinavisk Aktuarietidskrift*, 18, 161-241.
- [92] Currie, I., Durban, M. e Eilers, P. (2003). Using P-Splines to extrapolate two-dimensional Poisson data. *Proceedings of the 18th International Workshop on Statistical Modelling*, Louvain, Belgium, 97-102.
- [93] Currie, I., Durban, M. e Eilers, P. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, 4, 279-298.
- [94] Cutler, D. e Meara, E. (2001). Changes in the age distribution of mortality over the 20th century. NBER *Working Paper* N.º 8556, NBER.
- [95] Czado, C., Delwarde, A. e Denuit, M. (2005). Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics*, 36, 260-284.
- [96] Dahl, M. (2004). Stochastic Mortality in Life Insurance: Market Reserves and Mortality-Linked Insurance Contracts. *Insurance: Mathematics and Economics*, 35, 113-136.
- [97] Dahl, M. e Moller, T. (2005). Valuation and hedging of life insurance liabilities with systematic mortality risk. In *Proceedings of the 15th International AFIR Colloquium*, Zurich.
- [98] Dai, Q. e Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55(5), 1943-1978.
- [99] Das, S. (1998). Poisson-Gaussian processes and the bond markets. *Working Paper* 6631, NBER.
- [100] Davidoff, T., Brown, J. e Diamond, P. (2005). Annuities and individual welfare. *The American Economic Review*, 95(5), 1573-1590.

- [101] Davidson, A. e Reid, A. (1927). On the calculation of rates of mortality. *Transactions of the Faculty of Actuaries*, 11(105), 183-232.
- [102] De Boor, C. (2001). *A Practical Guide to Splines. Revised Edition*. Applied Mathematical Sciences 27, Springer-Verlag, New York.
- [103] De Jong, P. e Tickle, L. (2006). Extending Lee-Carter mortality forecasting. *Mathematical Population Studies*, 13(1), 1-18.
- [104] De Moivre, A. (1725). *Annuities on lives or the valuation of annuities upon any number of lives as also of reversions*. William Person, London.
- [105] Dellaportas, P., Smith, A. e Stavropoulos, P. (2001). Bayesian analysis of mortality data. *Journal of the Royal statistical Society*, 164(2), 275-291.
- [106] Delwarde, A., Kachachidze, D., Olie, L. e Denuit, M. (2004). Modèles linéaires et additifs généralisés, maximum de vraisemblance local et méthodes relationnelles en assurance sur la vie. *Bulletin Français D'Actuariat*, 6(12), 77-102.
- [107] Denuit, M., Devolder, P. e Goderniaux, A. (2007). Securitization of Longevity Risk: Pricing Survivor Bonds With Wang Transform in the Lee-Carter Framework. *The Journal of Risk and Insurance*, 74(1), 87-113.
- [108] Denuit, M. e Goderniaux, A. (2005). Closing and projecting lifetables using log-linear models. *Bulletin de l'Association Suisse des Actuaries*, 1, 29-49.
- [109] Di Lorenzo, E., Sibillo, M. e Tessitore, G. (2006). A Stochastic proportional hazard model for the force of mortality. *Journal of Forecasting*, 25, 529-536.
- [110] Dormoy, E. (1878). Théorie mathématique des assurances sur la vie. In Dormoy, E (Eds). *Théorie mathématique des assurances sur la vie*, vol. 1. Paris: GauthierVillars, 1-47
- [111] Dothan, L. (1978). On the term structure of interest rates. *Journal of Financial Economics*, 6, 56-69.
- [112] Dowd, K. (2003). Survivor Bonds: A comment on Blake and Burrows. *The Journal of Risk and Insurance*, 70(2), 339-348.
- [113] Dowd, K., Blake, D., Cairns, A. e Dawson, P. (2006). Survivor Swaps. *The Journal of Risk and Insurance*, 73, 1-17.
- [114] Duchêne, J. e Wunsch, G. (1988). From the demographer's cauldron: single decrement life tables and the span of life. *Genus*, Vol. XLIV, n.º 3-4, 1-17.
- [115] Duchêne, J. e Wunsch, G. (1991). Population aging and the limits to human life. In Wolfgang Lutz (Eds). *Future demographic trends in Europe and North America: what can we assume today?*. Academic Press: San Diego, California/London, England; International Institute for Applied Systems Analysis, Laxenburg, Austria, 27-40.

- [116] Duffie, D. (1996). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, 2nd Edition.
- [117] Duffie, D. (2005). Credit risk modeling with affine processes. *Journal of Banking and Finance*, 29, 2751-2802
- [118] Duffie, D., Filipovic, D. e Schachermayer, W. (2003). Affine processes and their applications in finance. *Annals of Applied Probability*, 13(3), 984-1053.
- [119] Duffie, D. e Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4), 379-406.
- [120] Duffie, D., Pan, J. e Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68 (6), 1343-1376.
- [121] Duffie, D. e Singleton, K. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4), 687-720.
- [122] Durban, M., Currie, I. e Eilers, P. (2002). Using P-Splines to smooth two-dimensional Poisson data. *Proceedings of the 17th International Workshop on Statistical Modelling*, Crete, 207-214.
- [123] Eilers, P. e Marx, B. (1996). Flexible smoothing using B-Splines and penalized likelihood (with comments and Rejoinder). *Statistical Science*, 11(2), 89-121.
- [124] El Shaarawi, A., Prentice, R. e Fobes, W. (1974). The goodness of fit of certain aging models. *Journal of Chronic Diseases*, 27, 377-385.
- [125] Elandt-Johnson, R. e Johnson, N. (1980). *Survival models and data analysis*. Wiley Classics Library, John Wiley & Sons, New York.
- [126] Faculty of Actuaries' Solvency Working Party (1986) (Hardie, A. C., Limb, A. P., Loades, D. H., Lumsden, I. C., Mason, D. C., Pollock, G., Robertson, E. S., Scott, W. F. e Wilkie, A. D). The solvency of life assurance companies. *Transactions of the Faculty of Actuaries*, 39, 251-340.
- [127] Fan, J. e Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Chapman and Hall, London.
- [128] Felipe, A. e Guillén, M. (1999). Evolución y predicción de tablas de mortalidad dinámicas para la población española. *Fundación Mapfre Seguros, Cuadernos de la Fundación* N.º 46.
- [129] Felipe, A., Guillen, M. e Nielsen, J. (2001). Longevity studies based on kernel hazard estimation. *Insurance: Mathematics and Economics*, 28, 191-204.
- [130] Felipe, A., Guillén, M. e Perez-Marin, A. (2002). Recent mortality trends in the Spanish population. *British Actuarial Journal*, 8, 757-786.

- [131] Fígoli, M. (1998). Modelando e projectando a mortalidade no Brasil. *Revista Brasileira de Estudos de População*, 15 (1).
- [132] Filipovic, D. (2005). Time-inhomogeneous affine processes. *Stochastic Processes and their Applications*, 115, 639-659.
- [133] Fleming, T. e Harrington, D. (1991). *Counting Processes and Survival Analysis*. Wiley series in probability theory, John Wiley & Sons, New York.
- [134] Flesaker, B. e Hughson, L. (1996). *Positive interest*. *Risk*, 9(1), 46-49.
- [135] Forfar, D., McCutcheon, J. e Wilkie, A. (1988). On Graduation by Mathematical Formula. *Journal of the Institute of Actuaries*, 115, 1-149.
- [136] Forfar, D. e Smith, D. (1987). The changing shape of English life tables. *Transactions of the Faculty of Actuaries*, 40(1), 98-133.
- [137] Frees, E., Carriere, J. e Valdez, E. (1996). Annuity valuation with dependent mortality. *The Journal of Risk and Insurance*, 63(2), 229-261.
- [138] Friedman, J. e Silverman, B. (1989). Flexible parsimonious smoothing and additive modelling. *Technometrics*, 31, 3-39.
- [139] GAD – Government Actuary’s Department (2001). National Population Projections: Review of Methodology for Projecting Mortality. *Report N.º 8*, London.
- [140] Gavin, J., Haberman, S. e Verrall, R. (1993). Moving Weighted Average Graduation using Kernel Estimation. *Insurance: Mathematics and Economics*, 12, 113-126.
- [141] Gavin, J., Haberman, S. e Verrall, R. (1994). On the choice of bandwidth for kernel Graduation. *Journal of the Institute of Actuaries*, 121, 119-134.
- [142] Gavin, J., Haberman, S. e Verrall, R. (1995). Variable Kernel Graduation with a Boundary Correction. *Transaction of the Society of Actuaries*, XLVII, 173-209.
- [143] Gavrilov, L. e Gavrilova, N. (1991). *The biology of life span: a quantitative approach*. Harwood Academic Publishers, London.
- [144] Gerber, H. U. (1995). *Life insurance mathematics*. Second-Edition, Springer-Verlag, Berlin.
- [145] Girosi, F. e King, G. (2005). *Demographic Forecasting*. forthcoming.
- [146] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new method of determining the value of life contingencies. *Philosophical Transactions of the Royal Society*, 115, 513-585.
- [147] Gompertz, B. (1860). On one uniform law of mortality from birth to extreme old age and on the law of sickness. *Journal of the Institute of Actuaries*, 16, 329-344.

- [148] Goodman, L. (1979). Simple models for the analysis of association in cross classifications having ordered categories. *Journal of the American Statistical Association*, 74, 537-552.
- [149] Greville, T. (1969). *Theory and Applications of Spline Functions*. Academic Press, New York.
- [150] Greville, T. (1981a). Moving-weighted average smoothing extended to the extremities of the data I: Theory. *Scandinavian Actuarial Journal*, 1, 39-55.
- [151] Greville, T. (1981b). Moving-weighted average smoothing extended to the extremities of the data II: Methods. *Scandinavian Actuarial Journal*, 2, 65-81.
- [152] Gründl, H., Post, T. e Schulze, R. (2006). To Hedge or Not to Hedge: Managing Demographic Risk in Life Insurance Companies. *The Journal of Risk and Insurance*, 73(1), 19-41.
- [153] Gutterman, S. e Vanderhoof, I. (1998). Forecasting changes in mortality: a search for a law of causes and effects. *North American Actuarial Journal*, 2, 135-138.
- [154] Haberman, S. (1996). Landmarks in the history of actuarial science (up to 1919). *Actuarial Research Paper N.º 84*, Department of Actuarial Science and Statistics, City University, London.
- [155] Haberman, S. e Renshaw, A. (1996). Generalised linear models and actuarial science. *The Statistician*, 45, 407-436
- [156] Haberman, S. e Sibbett, T. editors (1995). *History of Actuarial Science*. 10 Volumes, Pickering and Chatto, London.
- [157] Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.
- [158] Hannan, E.J. and Rissanen, J. (1982). Recursive Estimation of Mixed Autoregressive Moving Average Order. *Biometrika*, 69 (1), 81-94.
- [159] Hannerz, H. (1999). *Methodology and application of a new law of mortality*. Department of Statistics, University of Lund, Sweden.
- [160] Hannerz, H. (2001). Presentation and derivation of a five-parameter survival function intended to model mortality in modern female populations. *Scandinavian Actuarial Journal*, 2, 176-187.
- [161] Hardy, M. (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. New Jersey: John Wiley & Sons.
- [162] Hartmann, M. (1987). Past and recent attempts to model mortality at all ages. *Journal of Official Statistics*, 3(1), 19-36.

- [163] Hastie, T. e Tibshirani, R. (1990). *Generalized Additive Models*. Monographs on Statistics and Applied Probability 43, Chapman & Hall, New York.
- [164] Heath, D., Jarrow, R. e Morton, A. (1992). Bond Pricing and Term Structure of Interest Rates. *Econometrica*, Vol. 60, 77-105.
- [165] Heligman, L. (1984). Construction of the new United Nations model life table system. In Vallin *et al.* (Eds). *Methodologies for the collection and analysis of mortality data*. Ordina Editions, 179-201.
- [166] Heligman, L. e Pollard, J. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, 107, 49-80.
- [167] Henderson, R. (1915). *Mortality laws and statistics*. Wiley and Sons, Inc. NY.
- [168] Henderson, R. (1924). A new method of graduation. *Transactions of the Actuarial Society of America*, Vol. 25, 29-40.
- [169] Hoem, J. e Linnemann, P. (1988). The tails in moving average graduations. *Scandinavian Actuarial Journal*, 4, 193-229.
- [170] Horiuchi, S. e Wilmoth, J. (1998). Deceleration in the age pattern of mortality at older ages. *Demography*, 35 (4), 391-412.
- [171] Hull, J. e White, A. (1990). Pricing Interest Rate Derivative Securities. *Review of Financial Studies*, 3, N.º. 4, 573-592.
- [172] Hürlimann, W. (1990). On life table applications of ordering risks. *Insurance: Mathematics and Economics*, 9, 277-279.
- [173] Hyndman, R. e Ullah, M. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis*, to appear.
- [174] IAA (2001). *Graduated Mortality Tables. Mortality Investigation IA 95-97 M and F*, Institute of Actuaries of Australia.
- [175] IASB (2001). *Draft Statement of Principles*. International Accounting Standards Board, London.
- [176] IASB (2004). *International Financial Reporting Standard N.º 4*. International Accounting Standards Board, London.
- [177] INE (2003). *Projeções de População Residente 2000-2050*, Lisboa.
- [178] ISP (2000). *Relatório do Sector Segurador e Fundos de Pensões 1999*. Instituto de Seguros de Portugal, Lisboa.
- [179] ISP (2003). *Relatório do Sector Segurador e Fundos de Pensões 2002*. Instituto de Seguros de Portugal, Lisboa.

- [180] James, E. e Vitas, D. (2000). Annuities Markets in Comparative Perspective: Do Consumers Get Their Money's Worth?. *Policy Research Working Paper WPS 2493*, The World Bank.
- [181] James, J. e Webber, N. (2000). *Interest Rate Modelling*. Wiley, Chichester.
- [182] Jamshidian, F. (1997). LIBOR and Swap Market Models and Measures. *Finance and Stochastics*, 1, 293-330.
- [183] Jamshidian, F. (2004). Valuation of credit default swaps and swaptions. *Finance and Stochastics*, 8, 343-371.
- [184] Jones, M. (1993). Simple Boundary correction for Kernel density estimation. *Statistics and Computing*, 135-146.
- [185] JPMorgan (2004). Variable annuity guarantees – increased volatility risk. *European Life Insurance*, 28, April 2004.
- [186] Kannistö, V. (1992). *Development of oldest-old mortality, 1950-1990: Evidence from 28 developed countries*. Odense University Press.
- [187] Kannistö, V. (1996). *The Advancing Frontier of Survival: Life Tables for Old Age*. Odense University Press, Odense, Denmark.
- [188] Karatzas, I. e Shreve, S. (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag, Berlin.
- [189] Keyfitz, N. (1982). Choice of function for mortality analysis: effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology*, 21, 329-352.
- [190] Keyfitz, N. (1985). *Applied Mathematical Demography*. Springer-Verlag, New York.
- [191] King, A. e Reid, A. (1936). Some notes on the double geometric formula. *Transactions of the Faculty of Actuaries*, 15, 229-234.
- [192] Kinsella, K. e Gist, Y. (1998). Gender and aging: mortality and health. *International Brief 98-02*, US Bureau of Census.
- [193] Kostaki, A. (1992). *Methodology and applications of the Helligman-Pollard formula*. University of Lund, Department of statistics, Sweden
- [194] Kou, S. (2002). A jump-diffusion model for option pricing. *Management Science*, 48, 1086-1101.
- [195] Kou, S. e Wang, H. (2003). First passage times of a jump diffusion process. *Advances in Applied Probability*, 35, 504-531.
- [196] Krutov, A. (2006). Insurance-Linked Securities: An Emerging Class of Financial Instruments. *Financial Engineering News*, 48, 7-16.

- [197] Lambert, J. (1776). *Dottrina degli azzardi*. Gaeta and Fontana, Milan.
- [198] Lancaster, P. e Salkaukas, K. (1986). *Curve and Surface Fitting: An Introduction*. Academic Press, London.
- [199] Lando, D. (1998). On Cox processes and credit derivative securities. *Review of Derivatives Research*, 2, 99-120.
- [200] Lando, D. (2004). *Credit Risk Modelling: Theory and Applications*. Princeton Series in Finance, Princeton University Press, Oxford.
- [201] Landsmana, Z. e Sherris, M. (2001). Risk measures and insurance premium principles. *Insurance: Mathematics and Economics*, 29 (1), 103-115.
- [202] Lazarus, W. (1862). The law of human mortality. *Journal of the Institute of Actuaries*, 10, 283-298.
- [203] Le Bras, H. (1976). Lois de mortalité et age limité. *Population*, 331, 655-692.
- [204] Lederman, S. e Breas, J. (1959). Les dimensions de la mortalité. *Population*, 14(4), 637-682.
- [205] Lee, R. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4(1), 80-93.
- [206] Lee, R. e Carter, L. (1992). Modelling and forecasting the time series of US mortality. *Journal of the American Statistical Association*, 87, 659-671.
- [207] Lee, R. e Miller, T. (2001). Evaluating the performance of the Lee-Carter approach to modelling and forecasting. *Demography*, 38, 537-549.
- [208] Lehman Brothers (2005). Variable annuity living benefit guarantees: over complex, over popular and over here?. *European Insurance*, 22, April 2005.
- [209] Lemaire, J. (2001). Why do females live longer than males?. *Belgian Actuarial Bulletin*, 1, 1-19.
- [210] Lidstone, G. (1935). The practical calculation of annuities for any number of joint lives on a mortality table following the double geometric law. *Journal of the Institute of Actuaries*, 66, 413-423.
- [211] Lin, Y. e Cox, S. (2005a). Securitization of mortality risks in life annuities. *The Journal of Risk and Insurance*, 72, 227-252.
- [212] Lin, Y. e Cox, S. (2005b). A mortality securitization model. *Working Paper*, Georgia State University. Atlanta.
- [213] Lin, Y. e Cox, S. (2006). Securitization of Catastrophe Mortality Risks. *Working Paper*, Georgia State University. Atlanta.

- [214] Lindbergson, M. (2001). Mortality Among the Elderly in Sweden 1988-1997. *Scandinavian Actuarial Journal*, 1, 79-94.
- [215] Lindsey, J. (1997). *Applying Generalized Linear Models*. Springer-Verlag, New York.
- [216] Littrow, J. (1852). *Über Lebensversicherungen und Andere Versorgungsanstalten. F. Beck'schen Universitäts-Buchhandlung*, Vienna, 52.
- [217] Ljung, G. e Box, G. (1978). On a Measure of Lack of Fit in Time Séries Models. *Biometrika*, 66, 67-72.
- [218] London, D. (1997). *Survival models and their estimation*. ACTEX Publications, Winstead, Connecticut.
- [219] López, J. (2001). *Cuantificación de riesgos y análisis global de la empresa aseguradora vida*. Editorial Dykinson, Madrid.
- [220] Lowrie, W. (1982). An extension of the Whittaker-Henderson method of graduation. *Transactions of the Society of Actuaries*, XXXIV, 329-337.
- [221] MacDonald, A. (1997). *The Second Actuarial Study of mortality in Europe*. Groupe Consultatif des Associations d'Actuaires des Pays des Communautés Europeennes, Oxford.
- [222] MacDonald, A., Cairns, A., Dwilt, P. e Millers, K. (1998). An international comparison of recent trends in population mortality. *British Actuarial Journal*, 4, 3-141.
- [223] Makeham, W. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries*, 8, 301-310.
- [224] Makeham, W. (1867). On the law of mortality. *Journal of the Institute of Actuaries*, 13, 325-358.
- [225] Makeham, W. (1889). On the further development of Gompertz's law. *Journal of the Institute of Actuaries*, 28, 152-159.
- [226] Marceau, E. e Gaillardetz, P. (1999). On life insurance reserve in a stochastic mortality and interest rates environment. *Insurance: Mathematics and Economics*, 25, 261-280.
- [227] Marocco, P. e Pitacco, E. (1998). Longevity risk and life annuity reinsurance. *Transactions of the 26th International Congress of Actuaries*, Birmingham, 453-479.
- [228] Marques, F. (1970). *O povoamento da Metrópole observado através dos Censos*. INE, Centro de Estudos Demográficos.

- [229] Martinelle, S. (1987). A generalized Perks formula for old-age mortality. *Research-Methods-Development Report N.º 38*, Statistiska Centralbyran: Stockholm, Sweden.
- [230] Marx, B. e Eilers, P. (1998). Direct generalized additive modeling with penalized likelihood. *Computational Statistics and Data Analysis*, 28, 193-209.
- [231] McCarthy, D. e Mitchell, O. (2003). International Adverse Selection in Life Insurance and Annuities. NBER *Working Paper* No. 9975, NBER.
- [232] McCullagh, P e Nelder, J. (1989). *Generalized Linear Models*. Second Edition, Chapman & Hall, Lond
- [233] McCutcheon, J. (1987). Experiments in graduating the data for the English Life Tables N.º 14. *Transactions of the Faculty of Actuaries*, 40, 135-147.
- [234] McNown, R. e Rogers, A. (1989). Forecasting mortality: A parameterized time series approach. *Demography*, 26 (4), 645-660.
- [235] Mehra, R. e Prescott, E. (2003). The equity premium in retrospect. In Constantinides, G., Harris, M. e Stulz, R. (Eds). *Handbook of the Economics of Finance*. London, Elsevier, 887-936.
- [236] Menthonnex, J e Wanner, P. (1998). *Tables de Mortalité Longitudinales pour la Suisse (Génération 1880-1980)*. Office Fédéral de la Statistique, Berne.
- [237] Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, 41, September, 867-886.
- [238] Mesrine, A. (2000). La surmortalité des chômeurs: un effet catalyseur du chômage?. *Economie et Statistique*, 334, 33-48.
- [239] Mexia, T. e Corte-Real, P. (1995). Tabelas de Mortalidade Auto-Correctivas: o caso Português. *Boletim do Instituto dos Actuários Portugueses*, 35, 1-120.
- [240] Milevsky, M. (2005). The Implied Longevity Yield: A Note on Developing an Index for Life Annuities. *The Journal of Risk and Insurance*, 72, 301-320.
- [241] Milevsky, M. e Posner, S. (2001). The Titanic option: valuation of the Guaranteed Minimum Death Benefit in variable annuities and mutual funds. *The Journal of Risk and Insurance*, 68 (1), 91-126.
- [242] Milevsky, M. e Promislow, S. (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29, 299-318.
- [243] Milevsky, M. e Promislow, S. (2003). Can Life Insurance be used to Hedge Payout Annuities? Part 1: Modeling Longevity Risk. IFID *Working Paper*, York University.

- [244] Milevsky, M., Promislow, S. e Young, V. (2005). Financial valuation of mortality risk via the instantaneous Sharpe ratio. *Working Paper*, IFIDC Center, Toronto, Canada.
- [245] Milevsky, M., Promislow, S. e Young, V. (2006). Killing the law of large numbers: mortality risk premiums and the Sharpe ratio. *The Journal of Risk and Insurance*, 73(4), 673-686.
- [246] Milevsky, M. e Salisbury, T. (2002). The real option to lapse and the valuation of death-protected investments. *Working Paper* York University and The Fields Institute, Toronto.
- [247] Milevsky, M. e Salisbury, T. (2006). Financial valuation of Guaranteed Minimum Withdrawal Benefits. *Insurance: Mathematics and Economics*, 38, 21-38.
- [248] Miltersen, K. e Persson, S. (2005). Is mortality dead? Stochastic force of mortality determined by no arbitrage. *Working paper*, University of Bergen.
- [249] Mitchell, O. e McCarthy, D. (2002). Annuities for an Ageing World. *CERP Working Paper* 21/02, Center for Research on Pension and Welfare Policies.
- [250] Mitchell, O., Poterba, J., Warshawsky, M. e Brown, J. (1999). New Evidence on the Money's Worth of Individual Annuities. *American Economic Review*, 89(5), 1299-1318.
- [251] Mode, C. e Busby, R. (1982). An eight-parameter model of human mortality – The single decrement case. *Bulletin of Mathematical Biology*, 44(5), 647-659.
- [252] Möller, T. (1998). Risk-Minimizing Hedging Strategies for Unit-Linked Life Insurance Contracts. *ASTIN Bulletin*, 28, 17-47.
- [253] Möller, T. (2001a). Hedging Equity-Linked Life Insurance Contracts. *North American Actuarial Journal*, 5(2), 79-95.
- [254] Möller, T. (2001b). On Transformations of Actuarial Valuation Principles. *Insurance: Mathematics and Economics*, 28(3), 281-303.
- [255] Möller, T. (2001c). Risk-Minimizing Hedging Strategies for Insurance Payment Processes. *Finance and Stochastics*, 5(4), 419-446.
- [256] Möller, T. (2002). On Valuation and Risk Management at the Interface of Insurance and Finance. *British Actuarial Journal*, 8(4), 787-828.
- [257] Morris, J., Cook, D. e Shaper, G. (1994). Loss of employment and mortality. *British Medical Journal*, 308, 1135-1139.
- [258] Moser, L. (1839). *Die Gesetze der Lebensdauer*. Verlag von Veit und Comp, Berlin, 315.

- [259] Musiela, M. e Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*. Springer-Verlag, Berlin.
- [260] Nadaraya, E. (1964). On estimating regression. *Theory of Probability and Its Applications*, 9, 141-42.
- [261] Nolfi, P. (1959). Die Berücksichtigung der Sterblichkeitsverbesserung in der Rentenversicherung nach der Optimalmethode der Spieltheorie. *Bulletin de la Association Suisse des Actuaires*, 1.
- [262] Olifiers, E. (1937). Graduation by double geometric laws supplementary to Makeham's basic curve, with an application to the graduation of the A 1924-29 (ultimate) table. *Journal of the Institute of Actuaries*, 68, 526-534.
- [263] Olivier, P. e Jeffery, T. (2004). Stochastic Mortality Models, *Presentation to the Society of Actuaries of Ireland* (disponível em <http://www.actuaries.ie>).
- [264] Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics and Economics*, 29, 231-245.
- [265] Olivieri, A. e Pitacco, E. (2003a). Annuity guarantee and uncertainty in mortality trends. *Working Paper N.º 30*, CERAP-Università Bocconi.
- [266] Olivieri, A. e Pitacco, E. (2003b). Solvency requirements for pension annuities. *Pension Economics and Finance*, 2(2), 127-157.
- [267] Olshansky, S. e Carnes, B. (1997). Ever since Gompertz. *Demography*, 34(1), 1-15.
- [268] Oppermann, L. (1870). On the Graduation of Life Tables. In Haberman, S. e Sibbett (1995), T. (Eds) (1995). *History of Actuarial Science*. Pickering & Chatto, Vol. 2, London.
- [269] Panjer, H. e Tan, K. (1995). Graduation of Canadian Individual Insurance Mortality Experience: 1986-1992. *Canadian Institute of Actuaries*, Toronto.
- [270] Pelsser, A. (2003). Pricing and hedging Guaranteed Annuity Options via static option replication. *Insurance: Mathematics and Economics*, 33, 283-296.
- [271] Pelsser, A. (2004). On the Applicability of the Wang Transform for Pricing Financial Risks. *Working Paper*, Erasmus University, Rotterdam.
- [272] Perks W. (1932). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63:12-57.
- [273] Petrioli, L. (1975). Connection among fertility distribution by the Gompertz function. *Quaderni n.19*, Istituto di Statistica della Facoltà di SEB, Università degli Studi di Siena.

- [274] Petrioli, L. (1996). Relational methods in demography. *Quaderni di Popolazione, Ambiente e Metodi Matematico-Statistici*, N.º 1, Novembre, Università degli Studi di Siena.
- [275] Petrioli, L. e Berti, M. (1975). Modelli di mortalità. *Quaderni n.17, Istituto di Statistica della Facoltà di SEB*, Università degli Studi di Siena.
- [276] Petrioli, L. e Berti, M. (1979). Modelli di mortalità e loro impiego. *Quaderni n.37, Istituto di Statistica della Facoltà di SEB*, Università degli Studi di Siena.
- [277] Piazzesi, M. (2005). Affine term structure models. In Y. Ait-Sahalia and L. P. Hansen (eds). *Handbook of Financial Econometrics*, Amsterdam: Elsevier-Science.
- [278] Pitacco, E. (2000). *Matematica e tecnica actuariale della assicurazioni sulla durata di vita*. Lint – Editoriale Associati, Trieste.
- [279] Pitacco, E. (2004). Survival models in a dynamic context: a survey. *Insurance: Mathematics and Economics*, 35, 279-298.
- [280] Pollard, A. (1949). Methods of forecasting mortality using Australian data. *Journal of the Institute of Actuaries*, 75, 151-182.
- [281] Pollard, J. (1987). Projection of age-specific mortality rates. *Population Bulletin of the United Nations*, 21/22, 55-69.
- [282] Pollard, J. e Valkovics, E. (1992). The Gompertz distribution and its applications. *Genus*, 48(3-4), 15-28.
- [283] Powell, M. (1981). *Approximation theory and methods*. Cambridge University Press, Cambridge.
- [284] Promislow, S. (2006). *Fundamentals of Actuarial Mathematics*. John Wiley & Sons, Chichester.
- [285] Quiquet, A. (1893). *Aperçu historique sur les formules d'interpolation des tables de survie et de mortalité*. Deuxième édition, Paris, Warnier et Compagnie.
- [286] Ramlaou-Hansen, H. (1983). The choice of kernel function in the graduation of counting process intensities. *Scandinavian Actuarial Journal*, 165-182.
- [287] Rebonato, R. (1998). *Interest-rate option models*. Second Edition. Wiley, Chichester.
- [288] Rempala, G. e Szatcschneider, K. (2004). Bootstrapping Parametric Models of Mortality. *Scandinavian Actuarial Journal*, 1, 53-78.
- [289] Renshaw, A. (1991). Actuarial graduation practice and generalized linear and non-linear models. *Journal of the Institute of Actuaries*, vol. 118, 295-312.

- [290] Renshaw, A. (1992). Comments on some results on the Gompertz and Heligman and Pollard laws of mortality. *Journal of the Institute of Actuaries*, 119, 107-113.
- [291] Renshaw, A. e Haberman, S. (1996). Dual Modelling and Select Mortality. *Insurance: Mathematics and Economics*, 18(2), 136-136.
- [292] Renshaw, A. e Haberman, S. (2000). Modelling for reduction factors. *Actuarial Research Paper N.º 127*, Cass Business School, City University, London.
- [293] Renshaw, A. e Haberman, S. (2003a). On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics*, 32, 379-401.
- [294] Renshaw, A. e Haberman, S. (2003b). Lee-Carter mortality forecasting, a parallel GLM approach: England and Wales mortality projections. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 52, 119-137.
- [295] Renshaw, A. e Haberman, S. (2003c). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.
- [296] Renshaw, A. e Haberman, S. (2003d). Lee-Carter mortality forecasting incorporating bivariate time series. *Actuarial Research Paper N.º 153*, Cass Business School, City University, London.
- [297] Renshaw, A. e Haberman, S. (2005). Mortality reduction factors incorporating cohort effects. *Actuarial Research Paper N.º 160*, Cass Business School, City University, London.
- [298] Renshaw, A. e Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38, 556-570.
- [299] Renshaw, A., Haberman, S. e Hatzopoulos, P. (1996). The modelling of recent mortality trends in United Kingdom male assured lives. *British Actuarial Journal*, 2 (II), 449-477.
- [300] Renshaw, A., Haberman, S. e Hatzopoulos, P. (1997). On the duality of assumptions underpinning the construction of life tables. *ASTIN Bulletin*, 27(I), 5-22.
- [301] Renshaw, A. e Hatzopoulos P. (1996). On the Graduation of Amounts. *British Actuarial Journal*, 2, 185-205.
- [302] Rogers, A. e Little, J. (1994). Parameterizing age patterns of demographic rates with the multiexponential model schedule. *Mathematical Population Studies*, 4, 175-195.
- [303] Rogers, A. e Planck, F. (1984). Parametrized multistage population projections. *Proceedings of the 1984 Annual Meeting of the Population Association of America*, Minnesota.

- [304] Rogers, L. (1997). The potential approach to the term-structure of interest rates and foreign exchange rates. *Mathematical Finance*, 7, 157-164.
- [305] Rolski, T., Schmidli, H., Schmidt, V. e Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*. Wiley series in probability and statistics, John Wiley & Sons, New York.
- [306] Rueff, F. (1955). Ableitung von Sterbetafeln für Rentenversicherung und sonstige Versicherungen mit Erlebensfallcharakter. Special Publication N.º 2 of the *Deutsche Gesellschaft für Versicherungsmathematik*, Konrad Tritsch Verlag, Würzburg.
- [307] Rutkowski, M. (1997). A note on the Flesaker & Hughson model of the term structure of interest rates. *Applied Mathematical Finance*, 4, 151-163.
- [308] Schönbucher, P. (1998). Term structure modelling of defaultable bonds. *Review of Derivatives Research*, 2, 161-192.
- [309] Schönbucher, P. (2003). *Credit Derivatives Pricing Models*. Wiley Finance, Chichester.
- [310] Schrage, D. (2006). Affine Stochastic Mortality. *Insurance: Mathematics and Economics*, 38, 81-97.
- [311] Schuette, D. (1978). A linear programming approach to graduation. *Transactions of the Society of Actuaries*, XXX, 407-431.
- [312] Schweizer, M. (2001). From Actuarial to Financial Valuation Principles. *Insurance: Mathematics and Economics*, 28, 31-47.
- [313] Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- [314] Siler, W. (1979). A competing-risk model for animal mortality. *Ecology*, 60, 750-757.
- [315] Siler, W. (1983). Parameters of mortality in human populations with widely varying life span. *Statistics in Medicine*, 2, 373-380.
- [316] Sithole, T., Haberman, S. e Verrall, R. (2000). An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data. *Insurance: Mathematics and Economics*, 27, 285-312.
- [317] Smith, A. (2005). Stochastic mortality modelling. *Presentation at the Workshop on the Interface between Quantitative Finance and Insurance*, Edinburgh, April 2005 (disponível em <http://www.icms.org.uk>).
- [318] Society of Actuaries Group Annuity Valuation Task Force (1995). 1994 Group Annuity Mortality Table and 1994 Group Annuity Reserving Table. *Transactions of the Society of Actuaries*, 47, 865-913.

- [319] Spencer, J. (1904). On the graduation of the rates of sickness and mortality presented by the experience of the Manchester Unity of Oddfellows during the period 1893-1897. *Journal of the Institute of Actuaries*, 38, 334-43.
- [320] Stallard, E. (2006). Demographic Issues in Longevity Risk Analysis. *The Journal of Risk and Insurance*, 73(4), 575-609.
- [321] Tabeau, E. (2001). A review of demographic forecasting models for mortality. In E. Tabeau *et al.* (Eds), *Forecasting Mortality in Developed Countries: insights from a statistical, demographical and epidemiological perspective*, Kluwer Academic Publishers, 1-32.
- [322] Taylor, G. (1992). A Bayesian interpretation of the Whittaker-Henderson graduation. *Insurance: Mathematics and Economics*, 11, 7-16.
- [323] Tenenbein, A. e Vanderhoof, I. (1980). New Mathematical Laws of Select and Ultimate Mortality. *Transactions of the Society of Actuaries*, XXXII, 119-183.
- [324] Thatcher, R., Kannistö, V. e Andreev, K. (2002). The Survivor Ratio Method for Estimating Numbers at high ages. *Demographic Research*, 6 (1), Max-Planck Institute for Demographic Research.
- [325] Thatcher, A, Kannistö, V. e Vaupel, J., (1998). *The force of mortality at ages 80 to 120*. Odense Monographs on Population Aging, Odense University Press, Odense.
- [326] Thiele, P. (1872). On a mathematical formula to express the rate of mortality throughout the whole life. *Journal of the Institute of Actuaries*, 16, 313-329.
- [327] Tsay, R.S. e Tiao, G.C. (1984). Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models. *JASA*, 79 (385), 84-96.
- [328] Tsay, R.S. e Tiao, G.C. (1985). Use of Canonical Analysis in Time Series Model Identification. *Biometrika*, 72 (2), 299-315.
- [329] Tuljapurkar, S. e Boe, C. (1998). Mortality change and forecasting: how much and how little do we know?. *North American Actuarial Journal*, 2(4), 13-47.
- [330] Tuljapurkar, J., Li, N. e Boe, C. (2000). A universal pattern of mortality decline in G7 countries. *Nature*, 405, 789-792.
- [331] Turner, A. (2006). Pensions, Risk, and Capital Markets. *The Journal of Risk and Insurance*, 73(4), 559-574.
- [332] United Nations (1955). Age and sex patterns of mortality: Model life tables for underdeveloped countries. *Population Studies*, 22, Department of Social Affairs, Sales N.º 1955.XIII.9.

- [333] United Nations (2002). *World Population Prospects 2002*. New York: United Nations.
- [334] Vallin, J. (1973). La mortalité par génération en France depuis 1889. *INED*. Paris.
- [335] Van den Berg Jeths, A., Hoogenveen, R., De Hollander, G. e Tabeau, E. (2001). A Review of Epidemiological Approaches to Forecasting Mortality and Morbidity. In E. Tabeau *et al.* (Eds), *Forecasting Mortality in Developed Countries: insights from a statistical, demographical and epidemiological perspective*, 33-56.
- [336] Vandeschrisck, C. (2001). The Lexis diagram, a misnomer. *Demographic Research*, 4(3), 97-124.
- [337] Vasicek, O. A. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, N° 5, 177-188.
- [338] Vaupel, J. (1997). Trajectory of mortality at advanced ages. In *Between Zeus and the Salmon: The Biodemography of longevity*, 17-37, National Academy of Science.
- [339] Vaupel, J., Manton, K. e Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439-454.
- [340] Venables, W. e Ripley, B. (2002). *Modern Applied Statistics with S*, 4th Edition, Springer-Verlag, New York.
- [341] Verrall, R. (1993a). A state space formulation of the Whittaker-Henderson graduation, with extensions. *Insurance: Mathematics and Economics*, 13, 7-14.
- [342] Verrall, R. (1993b). Graduation by dynamic regression methods. *Journal of the Institute of Actuaries*, 120, 153-170.
- [343] Verrall, R. (1994). Whittaker graduation and dynamic generalized linear models. *Insurance: Mathematics and Economics*, 13, 7-14.
- [344] Verrall, R. (1996). A unified framework for graduation. Faculty of Actuarial Science and Statistics *Actuarial Research Paper N.º 91*, Cass Business School, City University, London.
- [345] Vincent, P. (1951). La mortalité des vieillards. *Population*, 6, 181-204.
- [346] Wadsworth, M., Findlater, A. e Boardman, T. (2001). *Reinventing Annuities*. Staple Inn. Actuarial Society.
- [347] Wang, J. , Yang, M. , Pan, S. (2003). Hedging longevity risk in life insurance companies. *The Asia-Pacific Risk and Insurance Association Annual Meeting*.
- [348] Wang, S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26, 71-92.

- [349] Wang, S. (2000). A class of distortion operations for pricing financial and insurance risks. *The Journal of Risk and Insurance*, 67, 15-36.
- [350] Wang, S. (2002). A universal framework for pricing financial and insurance risks. *ASTIN Bulletin*, 32, 213-234.
- [351] Wang, S. (2003). Equilibrium pricing transforms: new results using the Buhlmann's 1980 economic model. *ASTIN Bulletin*, 33, 57-73.
- [352] Wang, S. (2006). Normalized exponential tilting: pricing and measuring multivariate risks. *Working Paper ERMII*.
- [353] Watson, G. (1964). Smooth regression analysis. *Sankhya*, A(26), 359-372.
- [354] Webb, T. e Friedberg, L. (2006). Life is Cheap: Using Mortality Bonds to Hedge Aggregate Mortality Risk. NBER *Working Paper* No. 11984, NBER.
- [355] Weibull, W. (1939). A statistical theory of the strength of materials. *Ing. Vetenskap Akad. Handl.*, 151 (1), 1-45.
- [356] Whittaker, E. (1923). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, Vol. 41, 63-75.
- [357] Wilkie, A. (1976). The Rate of Interest as a Stochastic Process -Theory and Applications, *Proceedings of the 20th International Congress of Actuaries*, Tokyo, 1, 1976, p.325-338.
- [358] Wilkie, A., Waters, H. e Yang, S. (2003). Reserving, pricing and hedging for policies with guaranteed annuity options. *British Actuarial Journal*, 9 (2), 263-291.
- [359] Willekens, F. (2001). Gompertz in context: the Gompertz and related distributions. In E. Tabeau *et al.* (Eds), *Forecasting Mortality in Developed Countries: insights from a statistical, demographical and epidemiological perspective*, 105-126.
- [360] Willemse, W. e Koppelaar, H. (2000). Knowledge elicitation of Gompertz's law of mortality. *Scandinavian Actuarial Journal*, 2, 168-179.
- [361] Willets, R. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, 10, 833-877.
- [362] Willets, R., Gallop, A., Leandro, P., MacDonald, A., Miller, K., Richards, S., Robjohns, N., Ryan, J. e Waters, H. (2004). Longevity in the 21st Century. *British Actuarial Journal*, 10, 685-832.
- [363] Wilmoth, J. (1993). Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. *Technical Report, Department of Demography*, University of California, Berkeley.

- [364] Wilmoth, J. (1995a). Are mortality projections always more pessimistic when disaggregated by cause of death?. *Mathematical Population Studies*, 5(4), 293-319.
- [365] Wilmoth, J. (1995b). Are mortality rates falling at extremely high ages? An investigation based on a model proposed by Coale and Kisker. *Population Studies*, 49(2), 281-295.
- [366] Wilmoth, J., Andreev, K., Jdanov, D. e Gleij, D. (2005). *Methods protocol for the Human Mortality Database*. Human Mortality Database.
- [367] Wilmoth, J. e Horiuchi, S. (1999). Rectangularization revisited: variability of age at death within human populations. *Demography*, 36(4), November, 475-495.
- [368] Wilmoth, J. e Valkonen, T. (2002). A parametric representation of mortality differentials over age and time. *Fifth seminar of the EAPS Working Group on Differentials in Health, Morbidity and Mortality in Europe*, Pontignano, Italy, April 2001.
- [369] Wittstein, T. (1883). The Mathematical Law of Mortality. *Journal of the Institute of Actuaries*, 24, 153-173.
- [370] Wong-Fillip, C. e Haberman, S. (2004). Projecting mortality trends: recent developments in the United Kingdom and the United States. *North American Actuarial Journal*, 8(2), 56-83.
- [371] Yaari, M. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2), 137-150.
- [372] Yang, S. (2001). *Reserving, pricing and hedging for guaranteed annuity options*. PhD Thesis, Heriot-Watt University, Edinburgh.
- [373] Yashin, A. (2001). Mortality models incorporating theoretical concepts of aging. In E. Tabeau *et al.* (Eds), *Forecasting Mortality in Developed Countries: insights from a statistical, demographical and epidemiological perspective*, 261-280.
- [374] Young, V. (2004). Premium Principles. In *Encyclopedia of Actuarial Science*. New York: John Wiley & Sons.