



UNIVERSIDADE DE ÉVORA

Honesto Estudo com Longa Experiência Misturado

**PROCESSOS AUTO-REGRESSIVOS DE
COEFICIENTES ALEATÓRIOS NA MODELAÇÃO
DE DADOS DE CONTAGEM**

DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM MATEMÁTICA

Esta tese não inclui as críticas e sugestões feitas pelo júri



152 299

Dulce Maria de Oliveira Gomes

ORIENTADA POR: Professora Doutora Luísa Canto e Castro de Loura

JANEIRO DE 2005

Acção co-financiada pelo Fundo Social Europeu e Estado Português



À minha avó,

por ter plantado em mim o germe da leitura

Resumo

Modelar séries temporais de valores inteiros não-negativos pareceu-nos um desafio bastante aliciante, não só devido à sua importância, como também devido ao facto de ser um tema ainda pouco explorado, contrariamente à modelação de séries temporais com suporte nos reais, através de modelos lineares e não-lineares, que tem sido assunto de vários estudos.

Este trabalho centra-se, portanto, no estudo de modelos de séries temporais de valores inteiros não-negativos, também designadas de séries de contagem. Mais concretamente, centra-se no estudo de uma classe de modelos auto-regressivos de ordem 1, de coeficientes aleatórios e baseados numa generalização da operação *thinning*, proposta por Steutel e Van Harn em 1979. Dentro desta classe de modelos foram estudados dois casos particulares. O primeiro caso, que vai ser estudado ao longo capítulo 2, tem por base o facto dos coeficientes do modelo serem eles próprios também um processo estocástico. No segundo caso, considera-se que os coeficientes constituem uma sucessão de variáveis aleatórias independentes e identicamente distribuídas — sendo o modelo assim definido um caso particular do primeiro. O capítulo 3 desenvolve-se em torno deste caso particular.

Estes modelos foram designados, respectivamente, por DSINAR(1) generalizado (as iniciais da denominação em inglês, *Doubly Stochastic INteger AutoRegressive*) e RCINAR(1) generalizado (sigla de *Random Coefficient INteger AutoRegressive*).

Abstract

Modelling non-negative integer-valued time series seemed an interesting challenge, not only because of their importance but also because it is a subject which has not been very explored; unlike modelling time series with support on \mathbb{R} , through linear and non-linear models, which has been the subject of several studies.

This essay is, therefore, centred in the study of non-negative integer-valued time series models, also designated as time series of counts. More specifically, it is centred in the study of a class of autoregressive models of order 1, with random coefficients and based in a generalisation of the operation *thinning*, proposed by Steutel and Van Harn in 1979. Within this class of models we have studied two particular cases. The first, which is studied in chapter 2, is based in the fact that the model coefficients are themselves a stochastic process. In the second case, it is considered that the coefficient are an independent and identically distributed sequence of random variables. The model thus defined is a particular case of the previous. Chapter 3 is dedicated to such case.

These models were designated, respectively, by generalized DSINAR(1) (standing for *Doubly Stochastic INteger AutoRegressive*) and generalized RCINAR(1) (*Random Coefficient INteger AutoRegressive*).

Agradecimentos

Bem sei que estas breves palavras não vão ser suficientes para poder expressar a dívida de gratidão que me une àqueles que, de uma forma ou de outra, contribuíram para a elaboração deste trabalho. Mas quero deixar testemunho dos seus nomes e do meu afecto:

Obrigada à Prof^a Doutora Luísa Canto e Castro de Loura, orientadora e amiga, não só pela sua disponibilidade mas sobretudo pela sua grande generosidade e reconhecida paciência. Obrigada também pela confiança que, ao longo destes quase 12 anos, sempre depositou em mim, a qual me deu força para empreender mais este desafio. E obrigada pelo valioso contributo que tem tido na minha formação.

Obrigada à Dra. M^a de Fátima Espírito Santo, do Instituto de Meteorologia, pela sua enorme disponibilidade e simpatia.

Obrigada à Ana Coelho pelas correcções, pelas traduções e, sobretudo, por ter partilhado comigo algumas das minhas mais belas viagens, exposições, teatros...

À Júlia Teles e à Luzia Gonçalves pela partilha de programas em Latex. Ao Nuno Sousa, que me corrigiu os erros de português e à Isabel Sousa pelos seus belos cozinhados. Ao José António Ferreira, pela ajuda na procura de alguns artigos. Aos amigos citados, e aos não citados, obrigada por me terem oferecido a sua amizade mesmo nos dias mais ingratos em que surgia algum desânimo.

Aos colegas do Departamento de Matemática da Universidade de Évora, por todo o apoio que manifestaram. Em particular, ao Fernando Carapau pela ajuda informática.

À minha família, por sempre me terem permitido escolher o meu caminho.

Índice

<i>Lista de Figuras</i>	<i>xiv</i>
<i>Lista de Tabelas</i>	<i>xvi</i>
<i>Abreviaturas e Notações</i>	<i>xvii</i>
1 Introdução	1
1.1 Classe de modelos auto-regressivos médias móveis de valores inteiros	1
1.2 Classe de modelos de representação em espaço de estados, condicionalmente Poisson	8
1.2.1 Representação em espaço de estados — caso univariado	9
1.2.2 Modelos condicionalmente Poisson para séries temporais de contagem	13
2 Processo auto-regressivo de coeficientes estocásticos	19
2.1 Modelo DSINAR(1) generalizado	19
2.2 Existência e estacionariedade	24
2.3 Solução causal	29
2.4 Momentos e parâmetros	34
2.5 Condições necessárias e suficientes para a estacionariedade do processo	36
2.6 Especificação do processo DSINAR(1) generalizado em termos da sua distribuição marginal	38

2.7	DSINAR(1) generalizado: um modelo <i>observation driven</i> . . .	39
2.8	Estimação dos parâmetros do processo DSINAR(1) generalizado	41
2.8.1	Estimadores de máxima verosimilhança condicional . .	41
2.9	Teste de ajustamento do modelo DSINAR(1) generalizado . .	45
3	Processo auto-regressivo de coeficientes i.i.d.	51
3.1	Modelo RCINAR(1) generalizado	51
3.2	Estacionariedade do processo	54
3.3	Solução causal do processo RCINAR(1) generalizado	55
3.4	Momentos e parâmetros	57
3.5	Média e funções de autocovariância e autocorrelação amostrais	59
3.5.1	Comportamento assintótico da média amostral do processo	60
3.5.2	Comportamento assintótico das funções de autocovariância e autocorrelação amostrais do processo	62
3.6	Estimação dos parâmetros do processo RCINAR(1) generalizado	64
3.6.1	Estimadores de Yule-Walker	64
3.6.2	Estimadores dos mínimos quadrados condicionais . . .	66
3.6.3	Propriedades assintóticas dos estimadores	68
3.7	Teste de ajustamento do modelo RCINAR(1) estacionário . .	75
3.8	Previsão de p.e.'s estacionários a partir do modelo RCINAR(1)	77
4	Processo DSINAR(1) generalizado: um estudo por simulação	81
4.1	O operador Binomial <i>thinning</i>	83
4.1.1	Série explicativa estacionária	83
4.1.2	Série explicativa não-estacionária	93
4.2	O operador Poisson <i>thinning</i>	101
4.2.1	Série explicativa estacionária	101
4.2.2	Série explicativa não-estacionária	112

4.3	Séries temporais de valores inteiros com lacunas	118
5	Aplicação	121
5.1	Modelação do número de óbitos registado diariamente em Évora entre 1980 e 1997	124
5.1.1	Os dados	124
5.1.2	Modelação da série de óbitos registados em Évora	133
5.2	Outros modelos condicionalmente Poisson	142
5.2.1	Modelo estudado por Zeger e Qaqish, em 1988	143
5.2.2	Um modelo análogo ao modelo PAR(1)	144
6	Conclusões e sugestões para trabalho futuro	153
A	Processo DSAR(1)	159
B	Método de Newton-Raphson na estimação dos parâmetros do processo DSINAR(1)	163
	Referências bibliográficas	171

Lista de Figuras

4.1.1	Realização do processo DSINAR(1): Binomial <i>thinning</i> . . .	85
4.1.2	Histograma do processo DSINAR(1): Binomial <i>thinning</i> . . .	85
4.1.3	Funções de autocorrelação e autocorrelação parcial amostrais do processo DSINAR(1): Binomial <i>thinning</i>	86
4.1.4	Funções de autocorrelação e autocorrelação parcial amostrais do processo explicativo $\{X_t\}_{t \in \mathbb{Z}}$: Binomial <i>thinning</i>	86
4.1.5	Estimativas de máxima verosimilhança condicional de (ω, λ) e intervalos de confiança <i>bootstrap</i> a 95%: Binomial <i>thinning</i>	89
4.1.6	Resíduos do modelo DSINAR(1) ajustado: Binomial <i>thinning</i>	90
4.1.7	Histograma dos resíduos do modelo DSINAR(1) ajustado: Binomial <i>thinning</i>	91
4.1.8	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança <i>bootstrap</i> a 95%: Binomial <i>thinning</i>	91
4.1.9	Modelação de uma realização do processo DSINAR(1): Binomial <i>thinning</i>	93
4.1.10	Realização do processo DSINAR(1): Binomial <i>thinning</i> . . .	94
4.1.11	Histograma do processo DSINAR(1): Binomial <i>thinning</i> . . .	95
4.1.12	Funções de autocorrelação e autocorrelação parcial amostrais do processo DSINAR(1): Binomial <i>thinning</i>	95
4.1.13	Funções de autocorrelação e autocorrelação parcial amostrais do processo explicativo $\{X_t\}_{t \in \mathbb{Z}}$: Binomial <i>thinning</i>	96

4.1.14	Resíduos do modelo DSINAR(1) ajustado: Binomial <i>thinning</i>	98
4.1.15	Histograma dos resíduos do modelo DSINAR(1) ajustado: Binomial <i>thinning</i>	98
4.1.16	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo DSINAR(1) ajustado: Binomial <i>thinning</i>	99
4.1.17	Modelação de uma realização do processo DSINAR(1): Binomial <i>thinning</i>	101
4.2.1	Realização do processo DSINAR(1) generalizado: Poisson <i>thinning</i>	103
4.2.2	Histograma do processo DSINAR(1) generalizado: Poisson <i>thinning</i>	103
4.2.3	Funções de autocorrelação e autocorrelação parcial amostrais do processo DSINAR(1): Poisson <i>thinning</i>	104
4.2.4	Funções de autocorrelação e autocorrelação parcial amostrais do processo explicativo $\{X_t\}_{t \in \mathbb{Z}}$: Poisson <i>thinning</i> .	104
4.2.5	Estimativas de máxima verosimilhança condicional de (ω, λ) e intervalos de confiança <i>bootstrap</i> a 95%: Binomial <i>thinning</i>	108
4.2.6	Resíduos do modelo DSINAR(1) ajustado: Poisson <i>thinning</i>	109
4.2.7	Histograma dos resíduos modelo DSINAR(1) ajustado: Poisson <i>thinning</i>	109
4.2.8	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança <i>bootstrap</i> a 95%: Poisson <i>thinning</i>	110
4.2.9	Modelação de uma realização do processo DSINAR(1): Poisson <i>thinning</i>	111
4.2.10	Modelação de uma realização do processo DSINAR(1): Poisson <i>thinning</i>	112
4.2.11	Realização do processo DSINAR(1) generalizado: Poisson <i>thinning</i>	114
4.2.12	Histograma do processo DSINAR(1): Poisson <i>thinning</i> . . .	114

4.2.13	Funções de autocorrelação e autocorrelação parcial amostrais do processo DSINAR(1): Poisson <i>thinning</i>	115
4.2.14	Funções de autocorrelação e autocorrelação parcial amostrais do processo explicativo $\{X_t\}_{t \in \mathbb{Z}}$: Poisson <i>thinning</i>	115
4.2.15	Resíduos do modelo DSINAR(1) ajustado: Poisson <i>thinning</i>	116
4.2.16	Histograma dos resíduos modelo DSINAR(1) ajustado: Poisson <i>thinning</i>	117
4.2.17	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança <i>bootstrap</i> a 95%: Poisson <i>thinning</i>	117
4.2.18	Modelação de uma realização do processo DSINAR(1): Poisson <i>thinning</i>	118
5.1.1	Óbitos registados em Évora entre 1980 e 1992	125
5.1.2	Óbitos registados em Évora entre 1993 e 2002	125
5.1.3	Temperaturas máximas e mínimas registadas em Évora entre 1980 e 1992	126
5.1.4	Temperaturas máximas e mínimas registadas em Évora entre 1993 e 1997	126
5.1.5	Histograma dos óbitos registados em Évora entre 1980 e 1997	127
5.1.6	Funções de autocorrelação e autocorrelação parcial amostrais dos óbitos registados em Évora entre 1980 e 1997	127
5.1.7	Caixas-de-bigodes dos óbitos registados em Évora de 1980-2002	128
5.1.8	Caixas-de-bigodes dos óbitos registados em Évora nos meses de Junho, Julho e Agosto, entre 1980-1997	129
5.1.9	Caixas-de-bigodes dos óbitos registados em Évora nos meses de Dezembro, Janeiro e Fevereiro, entre 1980-1997	130
5.1.10	Número de óbitos antes, durante e após a onda de calor de Julho de 1991	131
5.1.11	Número de óbitos antes, durante e após a onda de frio de Fevereiro de 1983	131

5.1.12	Resíduos do modelo DSINAR(1) generalizado ajustado . .	136
5.1.13	Resíduos do modelo DSINAR(1) generalizado ajustado . .	137
5.1.14	Histograma dos resíduos do modelo DSINAR(1) generalizado ajustado	137
5.1.15	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo DSINAR(1) generalizado ajustado e respectivos intervalos de confiança <i>bootstrap</i> a 95%	138
5.1.16	Modelação da série dos órbitos	139
5.1.17	Modelação da série dos órbitos	139
5.2.1	Modelação da série dos órbitos	143
5.2.2	Modelação da série dos órbitos	144
5.2.3	Resíduos do modelo ajustado	144
5.2.4	Resíduos do modelo ajustado	145
5.2.5	Resíduos do modelo ajustado	147
5.2.6	Resíduos do modelo ajustado	147
5.2.7	Histograma dos resíduos do modelo ajustado	148
5.2.8	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo ajustado	148
5.2.9	Resíduos do modelo ajustado	149
5.2.10	Resíduos do modelo ajustado	149
5.2.11	Histograma dos resíduos do modelo ajustado	150
5.2.12	Funções de autocorrelação e autocorrelação parcial amostrais dos resíduos do modelo ajustado	150
5.2.13	Modelação da série dos órbitos	151
5.2.14	Modelação da série dos órbitos	151

Lista de Tabelas

4.1.1	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	87
4.1.2	Testes de normalidade. <i>p-values</i> dos testes de K-S e de S-W: caso estacionário	87
4.1.3	Probabilidades de cobertura para realizações de dimensão $N = 1000$ do processo DSINAR(1)	89
4.1.4	Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos	93
4.1.5	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	96
4.1.6	Testes de normalidade. <i>p-values</i> dos testes de K-S e de S-W: caso não-estacionário	97
4.1.7	Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos	100
4.1.8	Coefficiente de ajustabilidade do modelo.	100
4.2.1	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	105
4.2.2	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	105
4.2.3	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	105
4.2.4	Testes de normalidade. <i>p-values</i> dos testes de K-S e de S-W: caso estacionário	106

4.2.5	Testes de normalidade. <i>p-values</i> dos testes de K-S e de S-W: caso estacionário	106
4.2.6	Probabilidades de cobertura para realizações de dimensão $N = 1000$ do processo DSINAR(1) generalizado	107
4.2.7	Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos	111
4.2.8	Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos	113
4.2.9	Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos	116
4.2.10	Testes de normalidade. <i>p-values</i> dos testes de K-S e de S-W: caso não-estacionário	116
4.2.11	Coeficiente de ajustabilidade do modelo	118
5.1.1	Excessos durante as Ondas de Calor e de Frio	133
5.1.2	Estimativas de $(\omega_1, \omega_2, \lambda)$, por EMVC	135
5.1.3	Valores extremos causados por uma onda de frio e suas estimativas	140
5.1.4	Valores extremos causados por uma onda de frio e suas estimativas	140
5.2.1	Estimativas de (α, ω_1) , por EMVC	146

Glossário de abreviaturas

AR	auto-regressivo com suporte em \mathbb{R} .
ARMA	auto-regressivo média móvel com suporte em \mathbb{R} .
INAR	auto-regressivo com suporte em \mathbb{Z} .
INMA	média móvel com suporte em \mathbb{Z} .
RCINAR	auto-regressivo com coeficientes aleatórios e suporte em \mathbb{Z} .
DSINAR	auto-regressivo com coeficientes estocásticos e suporte em \mathbb{Z} .
EMQC	estimação por mínimos quadrados condicionais.
EMVC	estimação de máxima verosimilhança condicional.
EQM	erro quadrático médio.
EYW	estimação de Yule-Walker.
FAC	função de autocorrelação.
FACP	função de autocorrelação parcial.
f.d.	função densidade de probabilidade.
F.d.	função de distribuição.
f.g.m.	função geradora de momentos.
f.m.p.	função massa de probabilidade.
i.i.d.	independente(s) e identicamente distribuída(os).
K-S	Kolmogorov-Smirnov.
MA	média móvel com suporte em \mathbb{R} .
p.e.	processo estocástico.
S-W	Shapiro-Wilk.
v.a.	variável aleatória.
WN	ruído branco.

Notações

$\text{Cov}(Y_t, Y_s)$	covariância entre Y_t e Y_s .
$\text{Corr}(Y_t, Y_s)$	correlação entre Y_t e Y_s .
\xrightarrow{m}	convergência em média.
\xrightarrow{d}	convergência em distribuição.
$\xrightarrow{q.c.}$	convergência quase certa.
$Y_t \sim \text{Ber}(p)$	Y_t v.a. com distribuição Bernoulli de média p .
$Y_t \sim B(n, p)$	Y_t v.a. com distribuição Binomial de média np .
$Y_t \sim \text{Po}(\lambda)$	Y_t v.a. com distribuição Poisson de média λ .
$Y_t \sim \text{Geom}(p)$	Y_t v.a. com distribuição Geométrica de média p^{-1} .
\wedge sobre uma letra	estimador.
\sim sobre uma letra	estimativa.
\circ^G	operador <i>thinning</i> generalizado.
$\circ^B(\equiv *)$	operador Binomial <i>thinning</i> .
$P[Y = y]$	probabilidade de Y ser igual a y .
\vec{Y}_t	p.e. multivariado.
$E[Y_t] = \mu_Y$	média de Y_t .
$\text{Var}[Y_t] = \sigma_Y^2$	variância de Y_t .
\mathcal{A}^Y	álgebra- σ gerada por $\{Y_t : t \in \mathbb{Z}\}$.
\mathcal{A}_{t+}^Y	álgebra- σ gerada por $\{Y_s : s \geq t\}$.
\mathcal{A}_{t-}^Y	álgebra- σ gerada por $\{Y_s : s \leq t\}$.
$X _{Y=y}$	v.a. X condicionada a $Y = y$.

Capítulo 1

Introdução

Apesar da necessidade de encontrar modelos adequados para descrever (modelar, simular e/ou prever) séries temporais de valores inteiros, são muito poucos os modelos definidos até 1970 (McKenzie (2003), p. 574). A partir desta altura têm sido desenvolvidos, e aprofundados, diversos tipos de modelos de séries temporais de valores inteiros, também designadas de séries de contagem. De entre os vários modelos que, mais ou menos recentemente, têm surgido na literatura, iremos, neste capítulo, fazer um levantamento daqueles que, em nosso entender, se podem incluir em duas grandes classes: a classe de modelos auto-regressivos médias móveis de valores inteiros não-negativos e a classe de modelos de representação em espaço de estados condicionalmente Poisson. Estas duas classes não são, contudo, classes separadas, dado que é possível encontrar uma representação em espaço de estados para alguns dos modelos pertencentes à primeira classe.

1.1 Classe de modelos auto-regressivos médias móveis de valores inteiros

Jacobs e Lewis (1978) definiram uma classe de processos para séries temporais discretas, designados por processos DARMA (do inglês *Discrete AutoRegressive Moving Average*). Estes processos são baseados na família dos processos auto-regressivos médias móveis, com suporte nos reais, os designados processos ARMA (do inglês *AutoRegressive Moving Average*), definidos em 1970 por Box e Jenkins. Os processos DARMA *representam a primeira tentativa real de definir uma classe de modelos exclusivamente*

para séries temporais discretas (McKenzie (2003), p. 574). Desde então, têm surgido diversas abordagens ao estudo dos processos com suporte nos inteiros não-negativos e, por conseguinte, são vários os modelos já existentes. Não se pretende fazer aqui um estudo exaustivo de todos os modelos propostos mas apenas abordar os modelos baseados no operador Binomial *thinning* sugerido por Steutel e Van Harn em 1979. Este operador, representado por $*$, foi, na sua forma original, definido do seguinte modo:

Definição 1.1.1. Seja $\{U_i\}_{i \in \mathbb{N}}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas com distribuição de Bernoulli de valor esperado α , independente da variável aleatória Y inteira e não-negativa. Define-se a operação $*$ entre o parâmetro α e a variável Y como

$$\alpha * Y = \begin{cases} \sum_{i=1}^Y U_i, & Y > 0 \\ 0, & Y = 0. \end{cases}$$

(A) Modelos auto-regressivos de valores inteiros não-negativos e baseados no operador *thinning*:

McKenzie (1985) definiu um modelo para os inteiros não-negativos baseado no operador Binomial *thinning* com o objectivo de modelar séries temporais com determinadas distribuições marginais. No essencial, este tipo de modelo baseia-se nos modelos clássicos auto-regressivos de ordem 1 (denotados AR(1)). Se pensarmos num modelo AR(1) ($N_t = \phi N_{t-1} + \epsilon_t$, $\phi \in (0, 1)$) para modelar processos de contagem cujas observações tenham uma grandeza relativamente pequena (talvez agora, passados quase vinte anos, achemos impensável, mas o certo é que muitas vezes estes modelos baseados em variáveis contínuas foram utilizados para a modelação de variáveis discretas) podemos interpretá-lo do seguinte modo: o número de indivíduos no sistema no instante t , N_t , é igual à proporção ϕ (proporção esta que se mantém constante ao longo do tempo) de indivíduos que estavam no sistema no instante $t - 1$ mais os elementos que entram no sistema no intervalo $(t - 1, t]$, ϵ_t . Podemos ainda dizer que, dada a história passada, o número de indivíduos que permanecem no sistema (ou o número de sobreviventes da variável aleatória N_{t-1}) é uma variável determinística.

Ao substituir no modelo AR(1) a operação multiplicação pelo operador Binomial *thinning* ($N_t = \alpha * N_{t-1} + \epsilon_t$, $\alpha \in (0, 1)$), para além de garantir que as variáveis apenas tomam valores nos inteiros não-negativos,

McKenzie considera (dada a história passada do processo) os indivíduos que permanecem no sistema como uma variável aleatória. Ou seja, em cada instante de tempo t , cada indivíduo tem uma probabilidade α de permanecer no sistema e uma probabilidade $1 - \alpha$ de o abandonar. Cada indivíduo permanece ou sai do sistema de um modo independente. Deste modo, o número de indivíduos no sistema no instante t depende dos indivíduos que permanecem no mesmo (com probabilidade α) mais os “novos” elementos que entram no sistema. O parâmetro α neste modelo de valores inteiros é, portanto, análogo ao parâmetro ϕ ($\phi \in (0, 1)$) no modelo AR(1). Ou seja, enquanto que no modelo AR(1) um dos parâmetros a estimar é a proporção ϕ de indivíduos que permanecem no sistema, no modelo proposto por McKenzie esse parâmetro corresponde à probabilidade de um indivíduo, ao acaso, permanecer no sistema.

Suposições do modelo proposto por McKenzie:

- (i) cada indivíduo tem uma probabilidade α de permanecer no sistema,
- (ii) α constante ao longo do tempo,
- (iii) cada indivíduo permanece ou não no sistema no instante t , independentemente do número de indivíduos que estão no sistema nesse instante,
- (iv) em cada instante de tempo, cada “novo” indivíduo entra no sistema segundo uma determinada lei e independentemente dos indivíduos que saíram ou ainda permanecem no sistema,
- (v) $\alpha * N_{t-1} = \sum_{i=1}^{N_{t-1}} U_i(\alpha)$, onde $\{U_i(\alpha)\}_{i \in \mathbb{N}}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas, binárias, independentes de N_{t-1} e tais que $P[U_i(\alpha) = 1] = \alpha, \forall i \in \mathbb{N}$,
- (vi) $\{N_t\}_{t \in \mathbb{Z}}$ é uma sequência de variáveis aleatórias, estacionárias e com uma dada distribuição marginal.

McKenzie demonstra ainda que o processo assim definido e o processo AR(1) possuem a mesma estrutura de correlação.

McKenzie (1986) estudou este modelo no caso da distribuição marginal do processo ser geométrica (designado por processo AR(1) Geométrico) e no caso da distribuição marginal ser Binomial Negativa.

Al-Osh e Alzaid (1987) estudaram as propriedades do modelo proposto por McKenzie, o qual designaram por INAR(1) (sigla da denominação inglesa *first-order INteger AutoRegressive* — de ora avante utilizada para designar este tipo de modelos), mas sem considerarem margens com uma dada distribuição específica.

Para este modelo, e considerando que os erros têm distribuição Poisson (denominado modelo INAR(1) Poisson), propõem estimadores para os parâmetros do modelo, nomeadamente estimadores de Yule-Walker, estimadores dos mínimos quadrados condicionais e estimadores de máxima verosimilhança. Relativamente aos estimadores dos mínimos quadrados condicionais provam, recorrendo a resultados de Klimko e Nelson (1978), que são fortemente consistentes. Alzaid e Al-Osh (1988) prosseguem o estudo das propriedades dos modelos INAR(1) e obtêm a distribuição dos erros de modo a que o processo tenha margens com distribuição Geométrica (designado modelo GINAR(1)).

Muitos outros autores têm-se debruçado sobre o modelo auto-regressivo de ordem 1 de valores inteiros (igualmente com margens não especificadas). Por exemplo, Franke e Seligmann (1993), que para além de introduzirem um novo modelo, designado por SINAR(1), e que é uma generalização do modelo INAR(1) com inovações Poisson, provam que o processo INAR(1) é estritamente estacionário. Demonstram ainda algumas propriedades assintóticas dos estimadores de máxima verosimilhança condicional dos parâmetros do modelo INAR(1). Brännäs (1994), para além de obter os estimadores para os parâmetros do modelo INAR(1) Poisson através dos métodos já mencionados, aplica, também, o método dos momentos generalizados na obtenção dos mesmos.

Alzaid e Al-Osh (1990) e Du e Li (1991) generalizaram o modelo INAR para ordens superiores (INAR(p)). Alzaid e Al-Osh (1990) estudam as propriedades do modelo e provam, sob algumas condições, a existência de distribuição limite. Du e Li (1991) vão mais longe demonstrando que, quando as raízes do polinómio auto-regressivo de ordem p estão fora do círculo unitário, então existe solução estacionária (estacionariedade de 2^a ordem). Ou seja, demonstram que as condições de estacionariedade dos modelos INAR(p) e AR(p) são as mesmas. Demonstraram ainda que o processo além de estacionário também é ergódico. Estudam a estimação dos parâmetros do modelo através da estimação de Yule-Walker e dos mínimos quadrados condicionais bem como as suas propriedades. Estudam ainda a previsão com este tipo de modelos. A estimação por máxima verosimilhança condicional

dos parâmetros do modelo $\text{INAR}(p)$, que se saiba, ainda não foi efectuada.

Latour (1998) apresenta uma generalização do modelo $\text{INAR}(p)$, o qual designa por $\text{GINAR}(p)$ (que não se trata de uma generalização do processo proposto por Alzaid e Al-Osh (1988), igualmente designado por $\text{GINAR}(1)$). Esta generalização consiste apenas na substituição da variável aleatória de Bernoulli envolvida na operação *thinning* por qualquer outra variável aleatória inteira e não-negativa de valor esperado e variância finitos. O autor prova em que condição o processo $\text{GINAR}(p)$ (e, por conseguinte, o processo $\text{INAR}(p)$) é estacionário. Prova não só que os processos $\text{GINAR}(p)$ e $\text{AR}(p)$ têm uma mesma estrutura de correlação, como também que o processo GINAR é um processo AR (com inovações ruído branco). Obtém ainda os estimadores de Yule-Walker e dos mínimos quadrados condicionais dos parâmetros do modelo e estuda as suas propriedades assintóticas.

(B) Modelos médias móveis de valores inteiros e baseados no operador *thinning*:

Al-Osh e Alzaid (1988) introduziram o processo $\text{INMA}(q)$ (sigla da denominação inglesa *INteger-valued Moving Average*) estudando as suas propriedades básicas, nomeadamente as expressões do valor esperado, da variância e das funções autocovariância e autocorrelação. Estudaram ainda alguns aspectos do modelo $\text{INMA}(1)$ com margens Poisson. Contudo, sobre a estimação dos parâmetros do modelo e da validação deste nada foi mencionado.

McKenzie (1986) estuda igualmente o processo com estrutura média móvel de ordem 1, $\text{MA}(1)$, de margens Geométricas (modelo baseado no modelo $\text{EMA}(1)$ proposto por Lawrance e Lewis (1979) (referência de McKenzie) para modelar séries temporais de valores reais de margens Exponencial Negativa) e McKenzie (1988) estuda o processo $\text{MA}(1)$ com margens Poisson. Em ambos os artigos também não aparece qualquer referência à estimação dos parâmetros do modelo.

Brännäs e Hall (1998) estudam o modelo $\text{INMA}(q)$ apresentando novas caracterizações deste. Estudam as propriedades dos modelos e a estimação dos respectivos parâmetros, recorrendo à estimação de Yule-Walker, dos mínimos quadrados condicionais e através do método dos momentos generalizados baseado nas funções geradoras de probabilidade. Apresentam ainda um estudo sobre o desempenho destes métodos para amostras de

pequena dimensão.

(C) Modelos mistos auto-regressivos médias móveis de valores inteiros e baseados no operador *thinning*:

Em McKenzie (1986) podemos encontrar o estudo dos modelos mistos auto-regressivos médias móveis de valores inteiros não-negativos (de ordens p e q , denominado ARMA(p,q)) e com margens Geométricas e em McKenzie (1988) o estudo dos modelos ARMA(p,q) com margens Poisson.

Alzaid e Al-Osh (1993) também estudaram este tipo de modelos, considerando margens com distribuição Poisson generalizada.

A estimação dos parâmetros do modelo ARMA(p,q) (exceptuando os casos em que $q = 0$ (INAR(p)) ou $p = 0$ (INMA(q)) para valores inteiros ainda não se encontram estudados.

(D) Modelos auto-regressivos de ordem 1 de valores inteiros, baseados no operador *thinning*, com coeficientes não constantes:

(D1) Coeficientes aleatórios com uma dada distribuição:

Lewis (1982) propõe a modelação de um processo estocástico, com suporte em \mathbb{R} , através de um modelo linear, com estrutura auto-regressiva de ordem 1, com coeficientes aleatórios e com margens com distribuição Gama. McKenzie(1986) considera este modelo e volta a substituir a operação multiplicação pelo operador Binomial *thinning* com o intuito de modelar um processo estocástico com suporte nos inteiros não-negativos e com margens com distribuição Binomial Negativa. Para tal considera que os coeficientes (variáveis aleatórias) seguem uma distribuição Beta. O modelo assim definido *tem a mesma estrutura de autocorrelação e partilha muitas das propriedades do modelo desenvolvido por Lewis* (McKenzie (1986), p. 690).

(D2) Coeficientes a variar ao longo do tempo de forma determinística:

Mais tarde, Brännäs (1995) propôs um modelo, igualmente baseado no operador Binomial *thinning*, em que o parâmetro α varia ao longo do tempo através de um vector de covariáveis fixas.

Brännäs e Hellström (1999) consideraram ainda um modelo do tipo INAR(1), mas estabelecendo novas suposições. Consideraram, por exemplo, que cada indivíduo permanece no sistema com probabilidade α ou sai do sistema com probabilidade $1 - \alpha$, mas não de um modo independente.

(D3) Coeficientes a variar ao longo do tempo de forma aleatória:

Os modelos por nós desenvolvidos enquadram-se nesta sub-subclasse. Numa primeira fase do estudo tomou-se a ideia de Brännäs (1995) e substituiu-se o operador Binomial *thinning* pelo operador *thinning* generalizado^[1] e o vector de covariáveis fixas por uma sucessão de variáveis aleatórias independentes e identicamente distribuídas. Obteve-se, assim, um modelo de coeficientes aleatórios que se designou por RCINAR(1) generalizado (do inglês *Random Coefficient INteger AutoRegressive*) (capítulo 3, secção 3.1).

O modelo RCINAR(1) generalizado assim definido pode ser visto como um modelo RCA(1) (da denominação inglesa *Random Coefficient Autoregressive model of order 1* — estudados por Andel (1976), por Nicholls e Quinn (1980), por Quinn e Nicholls (1981, 1982b)) — no qual, seguindo a abordagem de McKenzie ao estudo das séries temporais de valores inteiros, se substitui a operação multiplicação pela operação *thinning* generalizada. Por sua vez, o processo RCINAR(1) generalizado generaliza o processo INAR(1), no sentido em que se tem um “parâmetro” que varia aleatoriamente ao longo do tempo. Contudo, neste tipo de modelos (RCINAR(1) generalizado ou RCA(1)) para além de se exigir que os parâmetros variem aleatoriamente ao longo do tempo exige-se ainda que sejam sequências de variáveis aleatórias independentes e identicamente distribuídas.

No capítulo 3, para além de se estabelecer o referido modelo auto-regressivo, de coeficientes i.i.d., de séries temporais de valores inteiros não-negativos, estuda-se a existência e estacionariedade do processo estocástico que o caracteriza (secção 3.2). Na secção 3.3 demonstra-se que

^[1]O operador *thinning* generalizado, que introduziremos no capítulo 2, é baseado na generalização introduzida por Latour (1998).

o processo RCINAR(1) generalizado admite uma representação causal e, na secção 3.4, são obtidas as expressões para os momentos, até à 2^a ordem, do processo. Nas restantes secções que constituem este capítulo são estudados o comportamento assintótico da média, da função de autocovariância e autocorrelação amostrais do processo (secção 3.5), os estimadores dos parâmetros do processo, bem como as suas propriedades assintóticas (secção 3.6). Por fim, são abordadas nas duas últimas secções a avaliação da qualidade do ajustamento deste tipo de modelos e a previsão de p.e.'s estacionários a partir dos mesmos.

Após este estudo, e de modo a poder incorporar no modelo a informação proveniente de outras séries temporais de valores reais, observadas nos mesmos instantes de tempo (séries explicativas), tomou-se a ideia anterior e substituiu-se a sequência de variáveis aleatórias independentes e identicamente distribuídas por um processo estocástico. Obteve-se, assim, um modelo duplamente estocástico, ou seja, um modelo em que os parâmetros são eles próprios também um processo estocástico que se designou por DSINAR(1) generalizado (do inglês *Doubly Stochastic INteger AutoRegressive*) (capítulo 2, secção 2.1).

O capítulo 2 apresenta-se estruturado de uma forma muito análoga à do capítulo 3. O motivo por que se apresenta no capítulo 2 o processo DSINAR(1) generalizado e, no capítulo 3, o processo RCINAR(1) generalizado é porque este último constitui um caso particular do primeiro, daí fazer mais sentido apresentá-lo depois.

1.2 Classe de modelos de representação em espaço de estados, condicionalmente Poisson

O modelo que surge de imediato para a modelação de dados de contagem é o modelo Poisson. Se se considerar que os dados seguem uma distribuição Poisson, em que o parâmetro é uma função de um vector de covariáveis, então tem-se o conhecido modelo de regressão de Poisson. Contudo, para além de assumir que os acontecimentos são independentes este modelo ainda assume que são equi-dispersos (a média é igual à variância), pelo que não é adequado quando se está a modelar séries temporais de contagens^[2].

^[2]Em geral, quando os dados de contagem se encontram correlacionados ao longo de tempo verifica-se quase sempre a sua sobre-dispersão (ver, por exemplo, Zeger (1988), p. 1.)

David Cox, em 1981, introduziu duas classes de modelos para séries temporais, que designou por modelos *observation driven* e modelos *parameter driven*, de modo a captar a autocorrelação e a sobre-dispersão dos dados. Estes modelos são dois tipos importantes de modelos em espaço de estados, que iremos introduzir de seguida.

1.2.1 Representação em espaço de estados — caso univariado

Um modelo em espaço de estados para uma série temporal, $\{Y_t\}_{t \in \mathbb{Z}}$, é um modelo linear que consiste em duas equações: a equação de observação (que descreve o processo que gera as observações) e a equação de estado (que descreve uma sucessão de variáveis que “controlam” o processo observável). A equação de observação é definida por

$$Y_t = g_t X_t + W_t, \quad t \in \mathbb{Z}, \quad (1.2.1)$$

onde $W_t \sim WN(0, \sigma_W^2)$ e g_t um escalar, possivelmente dependente de t .

Por sua vez, a equação de estado é definida por

$$X_{t+1} = f_t X_t + V_t, \quad t \in \mathbb{Z}, \quad (1.2.2)$$

onde $V_t \sim WN(0, \sigma_V^2)$, f_t um escalar (também possivelmente dependente de t) e $\{V_t\}_{t \in \mathbb{Z}}$ e $\{W_t\}_{t \in \mathbb{Z}}$ são dois processos não-correlacionados.

A equação de observação estabelece, assim, que a observação Y_t é uma função linear da variável estado, X_t , mais um ruído branco.

Diz-se, então, que uma série temporal $\{Y_t\}_{t \in \mathbb{Z}}$ tem uma representação em espaço de estados se existir um modelo em espaço de estados para $\{Y_t\}_{t \in \mathbb{Z}}$, isto é, um modelo especificado pelas equações (1.2.1) e (1.2.2)^[3]. Uma das vantagens deste tipo de representação é o facto de ser apropriada tanto para processos estacionários, como para processos não-estacionários.

Brockwell e Davis (1996) mostraram (exemplo 8.1.1., pp. 253-54) que é possível encontrar uma representação em espaço de estados para muitos dos modelos tradicionais, tais como os modelos ARMA e os modelos auto-regressivos médias móveis integrados (abreviadamente ARIMA).

Davis, Dunsmuir e Wang (1997) designaram os modelos desenvolvidos por Cox por modelos em espaço de estados generalizados. Estes modelos — *observation driven* e *parameter driven* — diferem entre si através da

^[3]Definição de Brockwell e Davis (1996), p. 253.

equação de estado, dado que a equação de observação é a mesma para ambos os modelos.

Os modelos designados *observation driven* (por vezes referidos como *modelos de transição na literatura sobre análise de dados longitudinais* (Davis, Dunsmuir e Streett (2001), p. 1)) assumem que o processo é explicado a partir dos próprios dados, de tal modo que o processo no instante t é modelado a partir da sua história passada.

Os modelos designados *parameter driven* assumem que o processo é independente da sua história passada e que é função de um processo latente, digamos ϵ_t , não observável.

Como facilmente se depreende, no modelo *observation driven* a autocorrelação dos dados está implícita na sua especificação. A sobre-dispersão dos dados é introduzida através da equação de estado. No modelo *parameter driven* a autocorrelação e a sobre-dispersão são introduzidas através do processo latente.

De modo a estabelecer, de um modo mais compreensível, a diferença entre estes dois tipos, vamos apresentar dois modelos não-lineares usuais, com suporte nos reais e pertencentes à classe dos modelos em espaço de estados generalizados, nomeadamente o modelo com *heteroscedasticidade condicional auto-regressiva de ordem 1*, introduzido por Engle (1982), — abreviadamente ARCH(1) — e o modelo com volatilidade estocástica de ordem 1, *first-order stochastic volatility* — abreviadamente SVAR(1) — popularizado por Taylor (ver, *e.g.*, Taylor (1994)).

Diz-se que um processo $\{Y_t\}_{t \in \mathbb{Z}}$ é um ARCH(1) se satisfizer as seguintes equações

$$Y_t = \sigma_t V_t \quad \text{e} \quad \sigma_t^2 = \omega_1 + \omega_2 Y_{t-1}^2,$$

onde $\omega_1 > 0$, $\omega_2 \geq 0$ e $\{V_t\}_{t \in \mathbb{Z}}$ é uma sucessão de variáveis aleatórias independentes e identicamente distribuídas. No caso em que $V_t \sim N(0, 1)$, tem-se que

$$Y_t | Y_{t-1} \sim N(0, \omega_1 + \omega_2 Y_{t-1}^2),$$

isto é, o processo é explicado a partir dos próprios dados, pelo que o modelo ARCH(1) é um *observation driven*^[4].

^[4]Por vezes, na literatura, o processo ARCH(1) também aparece definido como $Y_t = \sqrt{\sigma_t} V_t$, onde $\sqrt{\sigma_t} = \omega_1 + \omega_2 Y_{t-1}^2$.

Um exemplo de um modelo *parameter driven* é o modelo SVAR(1) — modelo alternativo ao modelo GARCH, modelo ARCH generalizado introduzido por Bollerslev (1986) — definido por

$$Y_t = \sigma_t W_t \quad \text{e} \quad \log(\sigma_t^2) = \mu + \phi(\log(\sigma_{t-1}^2) - \mu) + \theta \epsilon_t,$$

onde μ , ϕ e θ são constantes e $\{W_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são sucessões de variáveis aleatórias independentes e identicamente distribuídas, em que $\log(\sigma_t^2)$ é um processo AR(1). Se $W_t \sim N(0, 1)$ e $\epsilon_t \sim N(0, 1)$, então

$$Y_t | \sigma_t \sim N(0, \exp\{\mu + \phi(\log(\sigma_{t-1}^2) - \mu) + \theta \epsilon_t\}),$$

ou seja, o processo é independente da sua história passada e é função de um processo latente, digamos $\{\sigma_t\}_{t \in \mathbb{Z}}$, não observável.

Seguidamente vamos introduzir dois outros exemplos deste tipo de modelos, propostos por Davis *et al.* (1997), para séries temporais de contagens cujas observações se podem admitir como provenientes de uma distribuição Poisson em que a média é uma função do processo de estado. Os autores consideram, à semelhança do modelo de regressão de Poisson proposto por Zeger (1988), um vector de covariáveis, designado por \vec{X}_t . O modelo *parameter driven* proposto para séries temporais de contagem é da seguinte forma:

$$Y_t | \epsilon_t, \vec{X}_t = \vec{x}_t \sim Po(\epsilon_t \exp(\vec{x}_t \vec{\beta})),$$

onde ϵ_t é o processo latente e $\vec{\beta}$ o vector parâmetro a estimar. Os autores consideram que o processo latente, ϵ_t , é não-negativo e estritamente estacionário de média unitária e variância, σ_ϵ^2 , finita, e obtêm, respectivamente, as seguintes expressões para a média e para a variância do processo $\{Y_t\}_{t \in \mathbb{Z}}$, condicionado ao vector de covariáveis, $\vec{X}_t = \vec{x}_t$,

$$E[Y_t | \vec{X}_t = \vec{x}_t] = \exp(\vec{x}_t \vec{\beta})$$

e

$$Var[Y_t | \vec{X}_t = \vec{x}_t] = \exp(\vec{x}_t \vec{\beta})(1 + \sigma_\epsilon^2).$$

A partir destas expressões fica bem patente que a sobre-dispersão do processo $\{Y_t\}_{t \in \mathbb{Z}}$ é introduzida através do processo latente.

O modelo *observation driven* proposto pelos referidos autores é definido do seguinte modo:

$$Y_t | \mu_t \sim Po(\mu_t)$$

em que

$$W_t := \log(\mu_t) = \vec{x}_t \vec{\beta} + \sum_{i=1}^q \theta_i e_{t-i},$$

e $e_t = (Y_t - \mu_t)/\mu_t^\lambda$, $\lambda \geq 0$, fixo.

A partir destes exemplos, para processos com suporte em \mathbb{R} e em \mathbb{Z}^+ , vamos assim resumir o que se entende por modelos em espaço de estados generalizados.

Um processo $\{Y_t\}_{t \in \mathbb{Z}}$, com suporte em Ω , diz-se um *observation driven* se for da seguinte forma:

$$Y_t = f(A_{t-1}^Y, V_t), \quad (1.2.3)$$

onde \mathcal{A}_{t-1}^Y é a álgebra- σ gerada por $\{Y_{t-k}, k \geq 1\}$, $\{V_t\}_{t \in \mathbb{Z}}$ é um ruído branco e $f(\cdot)$ é uma função mensurável em Ω , e diz-se um *parameter driven* se for da seguinte forma:

$$Y_t = g(\epsilon_t, W_t) \quad \text{e} \quad \epsilon_t = h(A_{t-1}^\epsilon, Z_t), \quad (1.2.4)$$

onde $\mathcal{A}_{t-1}^\epsilon$ é a álgebra- σ gerada por $\{\epsilon_{t-k}, k \geq 1\}$ ^[5], $\{W_t\}_{t \in \mathbb{Z}}$ e $\{Z_t\}_{t \in \mathbb{Z}}$ são ruídos brancos e $g(\cdot)$ e $h(\cdot)$ são funções mensuráveis em Ω e Ω' , respectivamente.

Em ambos os casos pode ainda considerar-se a intervenção de covariáveis para a explicação do modelo, como se viu no segundo conjunto de exemplos. Neste caso, basta substituir na equação (1.2.3), V_t pelo par (\vec{X}_t, V'_t) , onde V'_t é o ruído branco, e na equação (1.2.4), ϵ_t pelo par (\vec{X}_t, ϵ_t) .

Tal como foi anteriormente referido, uma das vantagens da representação em espaço de estados é o facto de esta ser apropriada tanto para processos estacionários, como para processos não-estacionários. Contudo, existe, como é sobejamente sabido, vantagem em estabelecer propriedades como a estacionariedade e a ergodicidade, para o processo Y_t .

Relativamente aos modelos em espaço de estados generalizados a estacionariedade e a ergodicidade do processo Y_t é estabelecida à custa do processo de estado. Ou seja, se o processo de estado é estacionário e ergódico, o processo observado, Y_t , também o é.

^[5]No caso do modelo proposto por Davis *et al.* (1987) não se exige uma estrutura Markoviana para o processo latente.

Para os modelos *parameter driven*, estas propriedades são, em alguns casos, fáceis de estabelecer, uma vez que o processo de estado é, em geral, Markoviano e não depende do processo Y_t . No que diz respeito aos modelos *observation driven*, estas propriedades já não são tão fáceis de estabelecer, dado que o processo de estado depende do processo observado.^[6]

No capítulo 2, secção 2.7, prova-se que é possível encontrar uma representação em espaço de estados para os processos DSINAR(1) generalizados. E, por conseguinte, o mesmo acontece para os processos RCINAR(1) generalizados (secção 3.4).

1.2.2 Modelos condicionalmente Poisson para séries temporais de contagem

De entre as várias distribuições pertencentes à família exponencial, a distribuição Poisson é a que se vem impondo como a mais adequada para a modelação de séries temporais de valores inteiros não-negativos, ou séries de contagem. Assim, a grande parte dos modelos encontrados na literatura assumem que — num determinado instante de tempo — os dados de contagem, condicionados à sua história passada e, possivelmente, à história passada e presente de um vector de covariáveis, seguem uma distribuição Poisson, em que o valor esperado, μ_t , é uma função dessa informação (dos dados e das covariáveis, caso se estejam a considerar). Designaremos este tipo de modelos (que não são mais do que modelos *observation driven*) simplesmente por modelos Poisson, abreviadamente, ModP^[7].

A primeira característica que distingue os modelos para dados de contagem encontrados na literatura é a forma linear ou exponencial de μ_t . Outra característica encontrada é o facto de, no instante de tempo t , Y_t

^[6]Davis *et al.* (2001) tomam o modelo *observation driven*, proposto em 1987, e estudam as condições de estacionariedade forte e ergodicidade do processo Y_t , condicionado a $\vec{X}_t = \vec{x}_t$, para $q = 1$ e $\vec{x}_t \vec{\beta} = \beta$. Ou seja, sob estas duas condições determinaram quais os valores de λ que transformam o processo Y_t condicionado em um processo fortemente estacionário e ergódico.

^[7]Davis *et al.* (2001), p. 1, propõem o uso dos modelos *observation driven*, baseados na distribuição Poisson, para a análise temporal de séries de contagem, dado que, *em geral*, a estimação dos parâmetros dos modelos *parameter driven* exige um esforço computacional considerável, tal como a previsão de valores futuros (devido à existência do processo latente), apesar de a interpretação do efeito das covariáveis sobre o processo ser mais directa, como se pode verificar através do exemplo anteriormente apresentado, em que o valor esperado de Y_t , condicionado ao vector de covariáveis, é uma função apenas dessas covariáveis.

dependem apenas da sua história passada e do vector de covariáveis nesse instante de tempo, \vec{X}_t .

Com base nesta segunda característica, e de modo a sintetizar os vários modelos encontrados, defina-se a seguinte expressão

$$f(\vec{X}_t) \left\{ \sum_{i=1}^p \alpha_i Y_{t-i} \mu_{t-i}^{-\lambda} + \sum_{j=1}^q \beta_j \mu_{t-j}^{1-\lambda} \right\} + g_{\alpha, \beta}(\vec{X}_t),$$

onde α_i e β_j são números reais (a especificar consoante o suporte do processo), $\lambda \geq 0$ (fixo) e $f(\cdot)$ e $g_{\alpha, \beta}(\cdot)$ são funções mensuráveis, com $g(\cdot)$ possivelmente dependente de α_i e/ou de β_j , $i = 1, \dots, p$ e $j = 1, \dots, q$.

Podemos, assim, definir duas classes de modelos Poisson, de ordens p e q , tais que

$$Y_t | \mu_t \sim Po(\mu_t)$$

em que $\mu_t = E[Y_t | \mathcal{A}_{t-1}^Y, \vec{X}_t]$, \mathcal{A}_{t-1}^Y é a álgebra- σ gerada por $\{Y_{t-k}, k=1, \dots, p\}$ e \vec{X}_t é um p.e. multivariado. Ambas as classes são apropriadas para acontecimentos cujo valor esperado é inferior à variância.

Na primeira classe, que designaremos por $ModP_1(p, q)$, a equação de estado, μ_t , é definida como uma função linear, do seguinte modo:

$$\mu_t = f(\vec{X}_t) \left\{ \sum_{i=1}^p \alpha_i Y_{t-i} \mu_{t-i}^{-\lambda} + \sum_{j=1}^q \beta_j \mu_{t-j}^{1-\lambda} \right\} + g_{\alpha, \beta}(\vec{X}_t).$$

Na segunda classe, designada $ModP_2(p, q)$, a equação de estado é definida como

$$\log(\mu_t) = f(\vec{X}_t) \left\{ \sum_{i=1}^p \alpha_i Y_{t-i} \mu_{t-i}^{-\lambda} + \sum_{j=1}^q \beta_j \mu_{t-j}^{1-\lambda} \right\} + g_{\alpha, \beta}(\vec{X}_t).$$

Como exemplo de um modelo pertencente à primeira classe tem-se o modelo PAR(p) — *Poisson AutoRegressive model of order p* —, introduzido por Brandt *et al.* (2000).

O modelo PAR(p) é um modelo $ModP_1(p, 0)$, com $\lambda = 0$, cuja equação de estado é da seguinte forma

$$\mu_t = \sum_{i=1}^p \alpha_i (Y_{t-i} - \exp(\omega \vec{X}_t)) + \exp(\omega \vec{X}_t),$$

i.e., $f(\vec{X}_t) = 1$ e $g(\vec{X}_t) = \left(1 - \sum_{i=1}^p \alpha_i\right) \exp(\omega \vec{X}_t)$. Brandt *et al.* (2000) consideram \vec{X}_t um vector de v.a.'s i.i.d.

Demonstraremos na secção 2.7 que o processo DSINAR(1) Poisson — processo DSINAR(1) generalizado com a operação *thinning* especificada através da distribuição Poisson e inovações igualmente Poisson — é um $ModP_1(1, 0)$ com $\lambda = 0$, $\alpha_1 = 1$, $f(X_t) = \exp(\omega X_t)$ e $g(X_t) = \mu_Z$, $\forall t \in \mathbb{Z}$.

Como modelos pertencentes à segunda classe, apresentaremos três modelos. O primeiro modelo é uma das várias propostas estudadas por Zeger e Qaqish (1988)^[8]. Os segundo e terceiro modelos foram propostos por Davis *et al.* (1997), posteriormente desenvolvidos em 2001.

No primeiro modelo, a equação de estado, $\log(\mu_t)$, é definida como

$$\log(\mu_t) = \sum_{i=1}^p \alpha_i Y_{t-i} + \vec{X}_t \vec{\beta}. \quad (1.2.5)$$

Este modelo não é mais do que um $ModP_2(p, 0)$ com $\lambda = 0$, $f(\vec{X}_t) = 1$ e $g(\vec{X}_t) = \vec{X}_t \vec{\beta}$, onde $\vec{\beta}$ é o vector parâmetro a estimar.

A estacionariedade do modelo (1.2.5) só se verifica no caso em que $p = 1$ e $\alpha_i \leq 0$, excluindo-se, portanto, a possibilidade de uma dependência positiva (observação de Zeger e Qaqish, de 1988, citado em Davis *et al.* (1997)). Por este motivo, em 1997, Davis *et al.* propõem uma correcção no valor esperado de modo a obter um processo estacionário e estabelecem a seguinte equação de estado:

$$\log(\mu_t) = \sum_{i=1}^p \gamma_i e_{t-i} + \vec{X}_t \vec{\beta},$$

onde $e_t = (Y_t - \mu_t) / \mu_t^\lambda$. Ou seja, este modelo é um $ModP_2(p, p)$ com $f(\vec{X}_t) = 1$, $g(\vec{X}_t) = \vec{X}_t \vec{\beta}$ e $\beta_i = -\alpha_i$, $\forall i = 1, \dots, p$, cuja equação de estado se pode escrever como:

$$\log(\mu_t) = \sum_{i=1}^p \alpha_i Y_{t-i} \mu_{t-i}^{-\lambda} + \sum_{i=1}^p \beta_i \mu_{t-i}^{1-\lambda} + \vec{X}_t \vec{\beta}.$$

Contudo, apesar de pretenderem com esta modificação obter um processo estacionário, apenas encontram valores de λ para os quais o processo

^[8]Referência de Davis *et al.* (1997). Igualmente neste artigo estão apresentadas outras equações de estado, propostas em 1988 por Zeger e Qaqish.

$\{\log(\mu_t)\}_{t \in \mathbb{Z}}$ é estacionário, considerando $p = 1$ e $\vec{X}_t \vec{\beta} = \beta$ (Davis *et al.* (2001)).

Por último, vamos apresentar um modelo estabelecido por Davis *et al.* (1997) (posteriormente desenvolvido em 2001), designado por modelo GLARMA(p,q) — do inglês *Generalized Linear ARMA(p,q)* —, cuja equação de estado incorpora componentes auto-regressivas médias móveis, do seguinte modo:

$$\log(\mu_t) = \exp(\vec{X}_t \vec{\beta}) + \sum_{i=1}^{\infty} \tau_i e_{t-i},$$

em que $e_t = (Y_t - \mu_t)/\mu_t^\lambda$, $\lambda \geq 0$, fixo, e onde os coeficientes $\{\tau_i\}$ são determinados através da relação

$$\sum_{i=1}^{\infty} \tau_i z^i = \frac{1 + \sum_{i=1}^q \theta_i z^i}{1 - \sum_{i=1}^p \phi_i z^i} - 1 = \frac{\theta(z)}{\phi(z)} - 1,$$

onde $\phi(z)$ e $\theta(z)$ são, respectivamente, os polinómios auto-regressivo de ordem p e médias móveis de ordem q .

Podemos afirmar que o modelo GLARMA(p,q) é um $ModP_2(\infty, \infty)$, com $f(\vec{X}_t) = 1$, $g(\vec{X}_t) = \vec{X}_t \vec{\beta}$ e $\beta_i = -\alpha_i = \tau_i$, $\forall i$.

Streett (2000) estudou condições de estacionariedade (forte) e ergodicidade para o modelo GLARMA(0,q), considerando $\lambda = 1$ e, tal como no caso anterior, $\vec{X}_t \vec{\beta} = \beta$. Estas condições foram estabelecidas recorrendo a técnicas de cadeias de Markov. Provou ainda que o modelo GLARMA(0,1), com $\vec{X}_t \vec{\beta} = \beta = 0$ e $\theta_1 > 0$, admite pelo menos uma solução estacionária para $\lambda = 0.5$.

No que diz respeito à estruturação deste trabalho, temos ainda de fazer referência aos capítulos 4 e 5. No capítulo 4 faz-se um estudo por simulação do comportamento dos estimadores dos parâmetros do modelo DSINAR(1) generalizado e da adaptação deste. Nomeadamente, através do estudo do ajustamento do modelo proposto. É também abordada a metodologia *bootstrap* para a obtenção de intervalos de confiança para os parâmetros do modelo e para as funções de autocorrelação e autocorrelação parcial dos resíduos. Para estes fins foram elaborados programas em linguagem FORTRAN.

O capítulo 5 surge da necessidade de averiguar a aplicabilidade dos modelos, por nós propostos, a dados de um problema concreto. Assim,

motivadas por um trabalho levado a cabo pelo Instituto Nacional de Saúde Dr. Ricardo Jorge, sobre o impacto das ondas de calor na mortalidade, efectuámos um breve estudo sobre o impacto, não só das ondas de calor, mas também das ondas de frio na mortalidade no distrito de Évora. Com este objectivo, procedemos à modelação do número de óbitos registado diariamente no distrito de Évora, entre os anos de 1980 e 1997, através do modelo DSINAR(1) generalizado, usando como covariáveis as temperaturas máximas e mínimas diárias registadas, de igual modo, no distrito de Évora e durante esse período.

De modo a confrontar o modelo DSINAR(1) generalizado com a classe de modelos de representação em espaço de estados, condicionalmente Poisson, procedemos, ainda, à modelação da série dos óbitos diários (registados no distrito de Évora entre 1980 e 1997) através de dois modelos pertencentes a esta classe. Nomeadamente, através de um modelo similar ao modelo PAR(1) — que no contexto desta aplicação nos pareceu mais adequado do que este — e do modelo estudado por Zeger e Qaqish em 1988, cuja equação de estado é dada em (1.2.5).

Por último, no capítulo 6, apresentam-se as conclusões mais relevantes, em nosso entender, sobre o trabalho desenvolvido ao longo dos vários capítulos, bem como algumas sugestões para trabalho futuro. No final do trabalho encontram-se, ainda, dois apêndices. No primeiro, apresenta-se um breve resumo das características mais importantes dos processos auto-regressivos de ordem 1, com valores nos reais e de coeficientes estocásticos, os quais, adoptando a designação de Pourahamadi (1986), se denominam como *Doubly Stochastic AutoRegressive process of order one*, abreviadamente DSAR(1). No segundo apêndice descreve-se a aplicação do método de Newton-Raphson na estimação dos parâmetros de um processo pertencente à família dos processos DSINAR(1) generalizados.

Capítulo 2

Processo auto-regressivo de coeficientes estocásticos

Neste capítulo vamos propor um modelo — designado por DSINAR(1) generalizado — adequado para modelar processos de parâmetro discreto e com suporte nos inteiros não-negativos, para o qual se admite a existência de um processo explicativo subjacente.

Apesar de a estacionariedade fraca (ou estacionariedade de 2ª ordem) para este tipo de processos (em que o processo explicativo é estocástico e com uma determinada estrutura de covariância) ser relativamente fácil de estabelecer, o efeito dessas covariáveis sobre os dados não é, contudo, fácil de interpretar. O que constitui, de certo modo, uma desvantagem deste tipo de modelos.

Ao longo dos vários capítulos, e sempre que nos estivermos a referir à estacionariedade de um dado processo, significa que nos estamos a referir à estacionariedade de 2ª ordem.

2.1 Modelo DSINAR(1) generalizado

Vamos começar por definir o operador *thinning* generalizado (representado por \circ^G) entre as variáveis aleatórias α e Y .

Definição 2.1.1. Seja Y uma variável aleatória inteira e não-negativa e α uma variável aleatória com suporte em R^+ . Define-se a operação \circ^G entre

as variáveis aleatórias α e Y como

$$(\alpha \circ^G Y | \alpha, Y) \sim G(\mu, \sigma^2),$$

onde $G(\mu, \sigma^2)$ é uma dada distribuição do tipo discreto (associada à operação *thinning* generalizada) de média $\mu = \alpha Y$ e variância $\sigma^2 = \delta Y$ finita, com δ uma função de α e de Y , mensurável em \mathbb{R}^+ .

Definição 2.1.2. Um processo estocástico $\{Y_t\}_{t \in \mathbb{Z}}$, de parâmetro discreto e com suporte nos inteiros não-negativos, diz-se um DSINAR(1) generalizado (*Doubly Stochastic INteger AutoRegressive*) se existirem

- (i) um processo estocástico $\{\alpha_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{R}^+ , tal que $E[\alpha_t^2] < \infty$, e
- (ii) uma sucessão de variáveis aleatórias $\{Z_t\}_{t \in \mathbb{Z}}$ inteiras, não-negativas e não-correlacionadas, com uma dada distribuição discreta, de média μ_Z e variância σ_Z^2 finita,

tais que $\{Y_t\}_{t \in \mathbb{Z}}$ satisfaz a seguinte equação às diferenças estocástica (mais concretamente, e utilizando a designação de Tjøstheim (1986), equação às diferenças duplamente estocástica)

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z}, \quad (2.1.1)$$

onde:

- (a) $\{Z_t\}_{t \in \mathbb{Z}}$ é independente, quer da operação *thinning* generalizada, quer de $\{\alpha_t\}_{t \in \mathbb{Z}}$,
- (b) $Cov(Y_s, Z_t) = 0, \quad s < t$,
- (c) $\alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1} \sim G(\mu_t, \sigma_t^2)$, onde $G(\mu_t, \sigma_t^2)$ é uma dada distribuição do tipo discreto de média $\mu_t = \alpha_t Y_{t-1}$ e variância σ_t^2 finita,
- (d) $\alpha_t \circ^G Y_{t-1} | \mathcal{A}_{t+}^\alpha, \mathcal{A}^Y \stackrel{d}{=} \alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1}$, onde \mathcal{A}_{t+}^α é a álgebra- σ gerada por $\{\alpha_s : s \geq t\}$ e \mathcal{A}^Y é a álgebra- σ gerada por $\{Y_s : s \in \mathbb{Z}\}$.

Esta definição permite também incorporar a informação proveniente de processos estocásticos com suporte em \mathbb{R} . Mais concretamente, se se considerar que $\alpha_t = f(X_t)$, onde $\{X_t\}_{t \in \mathbb{Z}}$ é um p.e. com suporte em \mathbb{R} e

$f(\cdot)$ é tal que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é um p.e. com suporte em \mathbb{R}^+ . Por exemplo, quando se considera que $\alpha_t \in (0, 1)$ é usual admitir-se a seguinte representação

$$\alpha_t = \frac{\exp(\omega X_t)}{1 + \exp(\omega X_t)}, \quad \omega \in \mathbb{R} \quad (\omega : \text{parâmetro desconhecido}).$$

Vamos considerar $\sigma_t^2 = \delta_t Y_{t-1}$, com $\delta_t = g(\alpha_t, Y_{t-1})$ ($g(\cdot)$ uma função mensurável em \mathbb{R}^+) tal que $E[\delta_t^2] < \infty$. A título exemplificativo, apresentamos seguidamente a expressão de δ_t para as distribuições de tipo discreto usualmente utilizadas. Nos exemplos (1)-(3), $\{\delta_t\}_{t \in \mathbb{Z}}$ é apenas função de $\{\alpha_t\}_{t \in \mathbb{Z}}$, enquanto no exemplo (4) é função de $\{\alpha_t\}_{t \in \mathbb{Z}}$ e de Y_{t-1} .

$$(1) \quad G(\mu_t, \sigma_t^2) \equiv B(Y_{t-1}, \alpha_t) \implies \delta_t = \alpha_t(1 - \alpha_t).$$

$$(2) \quad G(\mu_t, \sigma_t^2) \equiv BN(Y_{t-1}, \frac{1}{1+\alpha_t}) \implies \delta_t = \alpha_t(1 + \alpha_t).$$

$$(3) \quad G(\mu_t, \sigma_t^2) \equiv Po(\alpha_t Y_{t-1}) \implies \delta_t = \alpha_t.$$

$$(4) \quad G(\mu_t, \sigma_t^2) \equiv Geom(\frac{1}{\alpha_t Y_{t-1}}) \implies \delta_t = \alpha_t(\alpha_t Y_{t-1} - 1).$$

Como o nome indica, o operador *thinning* generalizado generaliza o operador *thinning* ou Binomial *thinning* para o caso aleatório. Tal como foi referido na introdução, o operador *thinning* foi proposto por Steutel e Van Harn (1979), para o caso determinístico, e definido como $\alpha * Y = \sum_{i=1}^Y U_i$, com $\{U_i\}_{i \in \mathbb{N}}$ uma sequência de v.a.'s i.i.d. com distribuição de Bernoulli de valor esperado α , independente da v.a. Y , inteira não-negativa.

Por este facto, vamos particularizar os resultados obtidos considerando apenas a distribuição Binomial na especificação da operação *thinning* generalizada, ou seja,

$$(\alpha_t \circ^B Y_{t-1} | \alpha_t, Y_{t-1}) \sim B(Y_{t-1}, \alpha_t),$$

onde $\mu_t = \alpha_t Y_{t-1}$ e $\sigma_t^2 = \alpha_t(1 - \alpha_t)Y_{t-1}$, com $\{\alpha_t\}_{t \in \mathbb{Z}}$ um p.e. com suporte em $(0, 1)$. O processo assim definido designa-se simplesmente como processo DSINAR(1). No apêndice A apresentar-se-à um estudo comparativo entre

os processos DSINAR(1) e os processos auto-regressivos de ordem 1, com suporte nos reais. A importância (ou a pertinência) desta comparação reside no facto de os processos DSINAR(1) não serem mais do que uma versão discreta destes últimos.

De modo não só a simplificar a notação como também a estabelecer um paralelismo com o modelo INAR(1), vamos utilizar o símbolo $*$ em vez de \circ^B para representar a operação *thinning* ou Binomial *thinning*.

O estudo da existência e estacionariedade do processo $\{Y_t\}_{t \in \mathbb{Z}}$ será feito tendo apenas em atenção as seguintes expressões dos seus momentos, condicionados à observação do processo no instante anterior e à observação da covariável no instante presente,

$$E[Y_t | Y_{t-1}, \alpha_t] = \mu_t + \mu_Z$$

e

$$\text{Var}[Y_t | Y_{t-1}, \alpha_t] = \sigma_t^2 + \sigma_Z^2,$$

as quais se obtêm a partir das propriedades (a)-(d). Nenhuma suposição será feita relativamente à sua densidade condicional.

De modo a demonstrar a existência e a estacionariedade do processo definido em (2.1.1), comecemos por estabelecer o seguinte lema:

Lema 2.1.1. $\forall t, s \in \mathbb{Z}$ tem-se que:

- (1) $E[\alpha_t \circ^G Y_{t-1}] = E[\mu_t]$,
- (2) $\text{Var}[\alpha_t \circ^G Y_{t-1}] = \text{Var}[\mu_t] + E[\sigma_t^2]$,
- (3) $E[(\alpha_s \circ^G Y_{s-1})(\alpha_t \circ^G Y_{t-1})] = E[\mu_s \mu_t]$,
- (4) $E[\delta_t(\alpha_{t-1} \circ^G Y_{t-2})] = E[\delta_t \mu_{t-1}]$,
- (5) $E[\delta_t(\alpha_{t-1} \circ^G Y_{t-2})^2] = E[\delta_t \mu_{t-1}^2] + E[\delta_t \sigma_{t-1}^2]$.

Demonstração. Demonstremos (1) e (2), a partir da definição de operação *thinning* generalizada, aplicando a fórmula da dupla esperança^[1]. Note-se que ao demonstrar-se (2) está também a demonstrar-se (1).

$$\text{Var}[\alpha_t \circ^G Y_{t-1}] = E[(\alpha_t \circ^G Y_{t-1})^2] - (E[\alpha_t \circ^G Y_{t-1}])^2 =$$

^[1]Para duas v.a.'s X e Y , conjuntamente distribuídas, com $E[X] < \infty$, tem-se que $E[Y] = E[E[Y|X]]$.

$$\begin{aligned}
&= E[E[E[(\alpha_t \circ^G Y_{t-1})^2 | \alpha_t, Y_{t-1}]]] - (E[E[E[\alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1}]]])^2 = \\
&= E[E[\text{Var}[\alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1}] + (E[\alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1}])^2]] - \\
&\quad - (E[E[\alpha_t Y_{t-1} | Y_{t-1}]])^2 = \\
&= E[\delta_t Y_{t-1}] + E[(\alpha_t Y_{t-1})^2] - (E[\alpha_t Y_{t-1}])^2 = E[\sigma_t^2] + \text{Var}[\mu_t].
\end{aligned}$$

A partir destas definições e da propriedade (d) demonstram-se (3)-(5). A título exemplificativo, vamos demonstrar (3). Seja $W_s = \alpha_s \circ^G Y_{s-1}$. Então, para um instante de tempo s , com $s < t$,

$$\begin{aligned}
&E[(\alpha_s \circ^G Y_{s-1})(\alpha_t \circ^G Y_{t-1})] = \\
&= E[E[E[E[(\alpha_s \circ^G Y_{s-1})(\alpha_t \circ^G Y_{t-1}) | W_s, \alpha_t, Y_{t-1}]]]] = \\
&= E[E[E[W_s E[\alpha_t \circ^G Y_{t-1} | W_s, \alpha_t, Y_{t-1}]]]] = E[E[E[W_s(\alpha_t Y_{t-1}) | W_s, Y_{t-1}]]].
\end{aligned}$$

Continuando o mesmo raciocínio, demonstra-se que

$$E[E[E[W_s(\alpha_t Y_{t-1}) | \alpha_t, Y_{t-1}]]] = E[(\alpha_s Y_{s-1})(\alpha_t Y_{t-1})] = E[\mu_s \mu_t].$$

Para $s > t$, a demonstração é análoga. \square

Seguidamente, seguiremos um método análogo ao usado em Latour (1998), considerando, para cada t , a seguinte sucessão de v.a.'s

$$Y_t^{(n)} = \begin{cases} 0, & n < 0 \\ Z_t, & n = 0 \\ \alpha_t \circ^G Y_{t-1}^{(n-1)} + Z_t, & n > 0, \end{cases} \quad (2.1.2)$$

que aproximam Y_t quando $t \rightarrow \infty$.

Defina-se, tal como em Pourahmadi (1988), o processo $\{\beta_{t,i} : i \geq 0\}_{t \in \mathbb{Z}}$ do seguinte modo

$$\beta_{t,i} = \begin{cases} \prod_{j=0}^{i-1} \alpha_{t-j}, & i \geq 1 \\ 1, & i = 0 \end{cases} \quad (2.1.3)$$

e, de modo a simplificar a notação, defina-se igualmente o processo $\{\beta_{t,i}^{(j)} : i, j \geq 0\}_{t \in \mathbb{Z}}$ como

$$\beta_{t,i}^{(j)} = \begin{cases} \prod_{k=0}^{i-1} \alpha_{t-k}, & i \geq 1 \\ 1, & i = 0. \end{cases} \quad (2.1.4)$$

2.2 Existência e estacionariedade

Teorema 2.2.1. *Sejam $\{\alpha_t\}_{t \in \mathbb{Z}}$ um p.e. estacionário de \mathcal{Z}^{a} ordem com $\alpha_t \in \mathbb{R}^+$, $\{Z_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s não-correlacionadas, inteiras, não-negativas e independentes de $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\delta_t^{(n)}$ a função apresentada na definição 2.1.1 para um par $(\alpha_t, Y_{t-1}^{(n-1)})$, onde $\{Y_t^{(n)}\}_{t \in \mathbb{Z}}$ é a sucessão definida em (2.1.2). Defina-se $\beta_{t,i}$ e $\beta_{t,i}^{(j)}$ de acordo com (2.1.3) e (2.1.4), respectivamente. Se, $\forall t \in \mathbb{Z}$ e $k = 0, 1, \dots$,*

$$(1) \sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty \text{ e}$$

$$(2) \lim_{n \rightarrow \infty} \sum_{i=0}^n \sum_{j=i+1}^n E[\delta_{t-i}^{(n-i)} \beta_{t+k, i+k} \beta_{t,j}^{(i)}] < \infty,$$

então, existe um p.e. $\{Y_t\}_{t \in \mathbb{Z}}$, com suporte nos inteiros não-negativos e estacionário de \mathcal{Z}^{a} ordem tal que

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z}. \quad (2.2.1)$$

Demonstração. À semelhança das demonstrações de Du e Li (1991) ou de Latour (1998), de existência e estacionariedade dos processo INAR(p), vamos construir uma sucessão de v.a.'s $\{Y_t^{(n)}\}_{n \in \mathbb{N}}$ que podem igualmente ser consideradas como aproximações sucessivas de Y_t ,

$$Y_t^{(n)} = \begin{cases} 0, & n < 0 \\ Z_t, & n = 0 \\ \alpha_t \circ^G Y_{t-1}^{(n-1)} + Z_t, & n > 0. \end{cases}$$

Queremos demonstrar que a sucessão $\{Y_t^{(n)}\}_{n \in \mathbb{N}}$ converge. Para tal, vamos provar que $\{Y_t^{(n)}\}_{n \in \mathbb{N}}$ é uma sucessão de Cauchy. Mas antes vamos demonstrar que $\{Y_t^{(n)}\} \in L^2, \forall n \in \mathbb{N}$ e $\forall t \in \mathbb{Z}$ (com $L^2 = L^2(\Omega, \mathcal{A}, P)$ o espaço das v.a.'s com segundos momentos finitos). Defina-se ainda o produto escalar em L^2 como $\langle X, Y \rangle = E[XY]$ (L^2 é o espaço de Hilbert). Obtém-se recursivamente que

$$\begin{aligned} E[Y_t^{(n)}] &= E[\alpha_t Y_{t-1}^{(n-1)}] + \mu_Z = \\ &= E[\alpha_t \alpha_{t-1} Y_{t-2}^{(n-2)}] + \mu_Z(1 + E[\alpha_t]) = \\ &\quad \vdots \\ &= \mu_Z \left\{ 1 + \sum_{i=1}^n E \left[\prod_{j=0}^{i-1} \alpha_{t-j} \right] \right\} = \mu_Z \sum_{i=0}^n E[\beta_{t,i}]. \end{aligned}$$

Uma vez que a condição (1) implica que

$$\sum_{i=0}^{\infty} E[\beta_{t,i}] < \infty, \quad \forall t \in \mathbb{Z},$$

tem-se que $E[Y_t^{(n)}] < \infty$. Tendo em conta, para além disso, que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é um p.e. estacionário de 2ª ordem então, não só $E[Y_t^{(n)}] < \infty$ como é independente de t .

Obtém-se, após alguns cálculos simples, mas morosos (daí a sua omissão), a seguinte expressão para a covariância de $Y_t^{(n)}$,

$$\begin{aligned} \text{Cov}(Y_t^{(n)}, Y_{t+k}^{(n+k)}) &= E[\beta_{t+k,k} (Y_t^{(n)})^2] + \mu_Z \sum_{i=0}^{k-1} E[\beta_{t+k,i} Y_t^{(n)}] - \\ &\quad - \mu_Z^2 \sum_{i=0}^n E[\beta_{t,i}] \sum_{j=0}^n E[\beta_{t+k,j}], \quad k \geq 0. \end{aligned}$$

Quando $k = 0$

$$\sum_{i=0}^{k-1} E[\beta_{t+k,i} Y_t^{(n)}] = 0.$$



Desenvolvendo recursivamente os valores esperados que aparecem na expressão da covariância, prova-se que

$$\begin{aligned}
Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)}) &= \mu_Z \sum_{i=0}^{n-1} \sum_{j=i+1}^n E[\delta_{t-i}^{(n-i)} \beta_{t+k, i+k} \beta_{t,j}^{(i)}] + \\
&+ 2\mu_Z^2 \sum_{i=0}^{n-1} \sum_{j=i+k}^n E[\beta_{t,i} \beta_{t+k,j}] + (\sigma_Z^2 - \mu_Z^2) \sum_{i=0}^n E[\beta_{t,i} \beta_{t+k, i+k}] - \\
&- \mu_Z^2 \sum_{i=0}^n E[\beta_{t,i}] \sum_{j=0}^n E[\beta_{t+k,j}] + \mu_Z^2 \sum_{i=0}^{k-1} \sum_{j=0}^n E[\beta_{t+k,i} \beta_{t,j}], \quad k \geq 0.
\end{aligned}$$

A partir desta expressão tem-se que

$$\begin{aligned}
Var[Y_t^{(n)}] &= \mu_Z \sum_{i=0}^{n-1} \sum_{j=i+1}^n E[\delta_{t-i}^{(n-i)} \beta_{t,i} \beta_{t,j}^{(i)}] + 2\mu_Z^2 \sum_{i=0}^{n-1} \sum_{j=i}^n E[\beta_{t,i} \beta_{t,j}] + \\
&+ (\sigma_Z^2 - \mu_Z^2) \sum_{i=0}^n E[\beta_{t,i}^2] - \mu_Z^2 \left(\sum_{i=0}^n E[\beta_{t,i}] \right)^2.
\end{aligned}$$

Dado que, $\forall t \in \mathbb{Z}$,

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$$

e

$$\lim_{n \rightarrow \infty} \sum_{i=0}^n \sum_{j=i+1}^n E[\delta_{t-i}^{(n-i)} \beta_{t+k, i+k} \beta_{t,j}^{(i)}] < \infty$$

então $Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)}) < \infty, \forall k \geq 0$. Se, para além disso, $\{\alpha_t\}_{t \in \mathbb{Z}}$ for um p.e. estacionário de 2^a ordem segue-se igualmente que não só $Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)}) < \infty, \forall k \geq 0$, como também é independente de t . Deste modo, $\{Y_t^{(n)}\} \in L^2, \forall n \in \mathbb{N}$ e $\forall t \in \mathbb{Z}$.

Vamos agora demonstrar que $\|Y_t^{(m)} - Y_t^{(n)}\| \rightarrow 0$, quando $m, n \rightarrow \infty$. Prova-se facilmente (considerando, por exemplo, $m > n$) que

$$\|Y_t^{(m)} - Y_t^{(n)}\| \leq \mu_Z E[\beta_{t,m}] + \mu_Z \sum_{i=n}^m E[\beta_{t,i}].$$

Pelo que, sob a condição (1), $\|Y_t^{(m)} - Y_t^{(n)}\| \rightarrow 0$, quando $m, n \rightarrow \infty$. Segue-se deste facto que existe uma v.a. $Y_t \in L^2$ tal que $Y_t^{(n)} \xrightarrow{m.q.} Y_t$, quando $n \rightarrow \infty$ (Brockwell e Davis (1991), p. 68).

Prova-se ainda que o facto de $Y_t^{(n)}$ convergir, em média quadrática, para Y_t implica que $\alpha_{t+1} \circ^G Y_t^{(n)}$ também converge em média quadrática para $\alpha_{t+1} \circ^G Y_t$, quando $n \rightarrow \infty$. Daqui se conclui que existe uma v.a. $Y_t \in L^2$ que verifica (2.2.1).

De modo a provar a unicidade, vamos supor que existe outro p.e. estacionário, digamos $\{W_t\}_{t \in \mathbb{Z}}$, tal que $W_t = \alpha_t \circ^G W_{t-1} + Z_t, \forall t \in \mathbb{Z}$. Então

$$P(|Y_t^{(n)} - W_t| < \epsilon) \geq 1 - \frac{\mu_Z E[|\beta_{t,n}|] + E[|\prod_{j=0}^{n-1} \alpha_{t-j} W_{t-n}|]}{\epsilon},$$

para qualquer $\epsilon > 0$. Pelo facto de o processo $\{W_t\}_{t \in \mathbb{Z}}$ ser um processo estacionário resulta que $E[|\prod_{j=0}^{n-1} \alpha_{t-j} W_{t-n}|] \rightarrow 0$, quando $n \rightarrow \infty$ e, portanto,

$$P(|Y_t^{(n)} - W_t| < \epsilon) \rightarrow 1,$$

ou seja, $P(Y_t = W_t) = 1$.

□

Estudemos agora a existência e estacionariedade do processo, para o caso particular de se estar perante a operação Binomial *thinning*.

Corolário 2.2.1. *Sejam $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{Z_t\}_{t \in \mathbb{Z}}$ dois p.e.'s verificando as condições do teorema 2.2.1 e $\circ^G \equiv *$. Se*

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty, \quad \forall t \in \mathbb{Z},$$

então existe um p.e. $\{Y_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{Z}^+ e estacionário de 2^{a} ordem tal que

$$Y_t = \alpha_t * Y_{t-1} + Z_t, \quad t \in \mathbb{Z}. \quad (2.2.2)$$

Demonstração. Tendo em linha de conta a demonstração do teorema 2.2.1, comecemos por olhar para a expressão do valor esperado de $Y_t^{(n)}$. É sabido que no caso em que $G(\mu_t, \sigma_t^2) \equiv B(Y_{t-1}, \alpha_t)$, $\mu_t = \alpha_t Y_{t-1}$ e $\sigma_t^2 = \alpha_t(1 - \alpha_t)Y_{t-1}$. Por este motivo, a expressão de $E[Y_t^{(n)}]$ não se altera. Assim, tendo em conta a condição (1) do referido teorema mais o facto de $\{\alpha_t\}_{t \in \mathbb{Z}}$ ser um p.e. estacionário de 2^{a} ordem, conclui-se que $E[Y_t^{(n)}]$ é finito e independente de t .

Vamos agora demonstrar que as condições (1) e (2) do teorema anterior são equivalentes. Para tal, vamos substituir δ_t por $\alpha_t(1 - \alpha_t)$ na expressão da $Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)})$. Após alguns cálculos, verifica-se que

$$\begin{aligned} Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)}) &= \mu_Z \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} \left\{ E[\beta_{t+k, i+k} \beta_{t, j}] - E[\beta_{t+k, i+k+1} \beta_{t, j}] \right\} + \\ &+ 2\mu_Z^2 \sum_{i=0}^{n-k-1} \sum_{j=i+k}^{n-1} E[\beta_{t, i} \beta_{t+k, j}] + (\sigma_Z^2 - \mu_Z^2) \sum_{i=0}^{n-1} E[\beta_{t, i} \beta_{t+k, i+k}] - \\ &- \mu_Z^2 \sum_{i=0}^{n-1} E[\beta_{t, i}] \sum_{j=0}^{n-1} E[\beta_{t+k, j}] + \mu_Z^2 \sum_{i=0}^{k-1} \sum_{j=0}^{n-1} E[\beta_{t+k, i} \beta_{t, j}], \quad \forall k \geq 0. \end{aligned}$$

Assim, a condição

$$\sum_{i=0}^{\infty} E[\beta_{t, i}^2] < \infty, \quad \forall t \in \mathbb{Z},$$

permite concluir que $Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)}) < \infty$, $\forall k \geq 0$. Além disso, pelo facto de $\{\alpha_t\}_{t \in \mathbb{Z}}$ ser um p.e. estacionário de 2^{a} ordem, conclui-se ainda que $Cov(Y_t^{(n)}, Y_{t+k}^{(n+k)})$, $\forall k \geq 0$, é independente de t .

O restante da demonstração é idêntico à demonstração do teorema 2.2.1. \square

2.3 Solução causal

Proposição 2.3.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo DSINAR(1) generalizado, estacionário de 2ª ordem. Então, existe uma sucessão de v.a.'s $\{\epsilon_t\}_{t \in \mathbb{Z}}$, não-correlacionadas, tal que o processo $\{Y_t\}_{t \in \mathbb{Z}}$ admite como solução estacionária a seguinte representação causal*

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad (2.3.1)$$

onde $\epsilon_t = \alpha_t \circ^G Y_{t-1} - \alpha_t Y_{t-1} + Z_t$.

Antes de passarmos à demonstração desta proposição, vamos estabelecer o seguinte lema de modo a facilitar esta tarefa.

Lema 2.3.1. *Sejam $\{\alpha_t\}_{t \in \mathbb{Z}}$ um p.e. independente de $\{Z_t\}_{t \in \mathbb{Z}}$, com $\{Z_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s não-correlacionadas e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ definidas em (2.3.1). Tem-se, $\forall i \geq 1$, que*

$$(1) \text{Cov}(\beta_{t,i}, \epsilon_{t-i}) = 0$$

e

$$(2) \text{Cov}(\epsilon_t, \beta_{t,i} \epsilon_{t-i}) = 0.$$

Demonstração. Este lema demonstra-se, à semelhança do lema 2.1.1, recorrendo à definição de operação *thinning*. \square

Demonstração da proposição 2.3.1. De modo a demonstrar que o processo $\{Y_t\}_{t \in \mathbb{Z}}$, estacionário de 2ª ordem, admite uma representação causal vamos escrevê-lo na seguinte forma:

$$Y_t = \alpha_t Y_{t-1} + \epsilon_t, \quad t \in \mathbb{Z},$$

onde $\epsilon_t = \alpha_t \circ^G Y_{t-1} - \alpha_t Y_{t-1} + Z_t$.

Iterando Y_t regressivamente obtém-se

$$Y_t = \epsilon_t + \sum_{i=1}^{n-1} \left(\prod_{j=0}^{i-1} \alpha_{t-j} \right) \epsilon_{t-i} + \left(\prod_{j=0}^{n-1} \alpha_{t-j} \right) Y_{t-n} =$$

$$= \sum_{i=0}^{n-1} \beta_{t,i} \epsilon_{t-i} + \beta_{t,n} Y_{t-n}, \quad n > 0.$$

Seja

$$Y_t^{(n)} = \begin{cases} \sum_{i=0}^{n-1} \beta_{t,i} \epsilon_{t-i}^{(n-i)}, & n > 0 \\ 0, & n < 0 \\ Z_t, & n = 0 \end{cases}$$

e

$$\epsilon_t^{(n)} = \begin{cases} \alpha_t \circ^G Y_{t-1}^{(n-1)} - \alpha_t Y_{t-1}^{(n-1)} + Z_t, & n > 0 \\ 0, & n < 0 \\ Z_t, & n = 0. \end{cases}$$

Quando $n > 0$, $\epsilon_t^{(n)}$ verifica as seguintes propriedades:

$$E[\epsilon_t^{(n)}] = \mu_Z \neq 0, \quad \forall n \geq 0,$$

$$\begin{aligned} V[\epsilon_t^{(n)}] &= E[\delta_t^{(n)} Y_{t-1}^{(n-1)}] + \sigma_Z^2 = \\ &= \mu_Z \sum_{j=1}^{n-2} E[\delta_t^{(n)} \beta_{t,j}^{(0)}] + \sigma_Z^2 \end{aligned}$$

e

$$Cov(\epsilon_t^{(n)}, \epsilon_{t+k}^{(n+k)}) = 0, \quad k > 0.$$

Tal como na demonstração do teorema 2.2.1 vamos igualmente demonstrar que $Y_t^{(n)}$ é uma sequência de Cauchy em L^2 .

Com esse objectivo, comecemos por notar que

$$E[Y_t^{(n)}] = \mu_Z \sum_{i=0}^{n-1} E[\beta_{t,i}] < \infty,$$

dado que

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty.$$

Seguidamente vamos demonstrar que $Var[Y_t^{(n)}] < \infty$. A $Var[Y_t^{(n)}]$ pode escrever-se do seguinte modo:

$$\begin{aligned} Var[Y_t^{(n)}] &= Var\left[\epsilon_t^{(n)} + \sum_{i=1}^{n-1} \left(\prod_{j=0}^{i-1} \alpha_{t-j}\right) \epsilon_{t-i}^{(n-i)}\right] = \\ &= Var[\epsilon_t^{(n)}] + Var\left[\sum_{i=1}^{n-1} \beta_{t,i} \epsilon_{t-i}^{(n-i)}\right] + 2 \sum_{i=1}^{n-1} Cov\left(\epsilon_t^{(n)}, \beta_{t,i} \epsilon_{t-i}^{(n-i)}\right). \end{aligned}$$

Pelo lema 2.3.1, alínea (2), a terceira parcela do lado direito da equação anterior é nula, pelo que a expressão da variância se escreve como

$$Var[Y_t^{(n)}] = Var[\epsilon_t^{(n)}] + \sum_{i=1}^{n-1} Var[\beta_{t,i} \epsilon_{t-i}^{(n-i)}] + 2\mu_Z^2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} Cov(\beta_{t,i}, \beta_{t,j}).$$

Substituindo $\epsilon_{t-1}^{(n-1)} = \alpha_{t-1} \circ^G Y_{t-2}^{(n-2)} - \alpha_{t-1} Y_{t-2}^{(n-2)} + Z_{t-1}$ na primeira parcela do primeiro somatório obtém-se

$$Var[\beta_{t,1} \epsilon_{t-1}^{(n-1)}] = \mu_Z \sum_{j=2}^{n-1} E[\delta_{t-1}^{(n-1)} \beta_{t,1} \beta_{t,1}^{(1)}] + \sigma_Z^2 E[\beta_{t,1}^2] + \mu_Z^2 Var[\beta_{t,1}].$$

Seguindo o mesmo raciocínio tem-se que

$$\begin{aligned} Var[Y_t^{(n)}] &= \mu_Z \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} E[\delta_{t-i}^{(n-i)} \beta_{t,i} \beta_{t,j}^{(i)}] + \sigma_Z^2 \sum_{i=0}^{n-1} E[\beta_{t,i}^2] + \\ &+ \mu_Z^2 \sum_{i=0}^{n-1} Var[\beta_{t,i}] + 2\mu_Z^2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} Cov(\beta_{t,i}, \beta_{t,j}) = \\ &= \mu_Z \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} E[\delta_{t-i}^{(n-i)} \beta_{t,i} \beta_{t,j}^{(i)}] + \sigma_Z^2 \sum_{i=0}^{n-1} E[\beta_{t,i}^2] + \\ &+ \mu_Z^2 Var\left[\sum_{i=0}^{n-1} \beta_{t,i}\right]. \end{aligned}$$

Sob as condições (1) e (2) do teorema 2.2.1, $\{Y_t^{(n)}\} \in L^2, \forall t \in \mathbb{Z}$ e $\forall n \in \mathbb{N}$, e

$$\|Y_t^{(m)} - Y_t^{(n)}\| \leq \mu_Z \sum_{i=n}^{m-1} E[\beta_{t,i}] \rightarrow 0,$$

quando $m, n \rightarrow \infty$, pois a série $\sum_{i=0}^{\infty} E[\beta_{t,i}]$ é convergente.

Então, existe uma v.a. $Y_t \in L^2$ tal que $Y_t^{(n)} \xrightarrow{\text{m.q.}} Y_t$ (e, por conseguinte, $\epsilon_t^{(n)} \xrightarrow{\text{m.q.}} \epsilon_t$), quando $n \rightarrow \infty$, ou seja,

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad t \in \mathbb{Z}$$

é uma v.a. bem definida (em média quadrática).

Por último, prova-se que se $\{Y_t\}_{t \in \mathbb{Z}}$ é um processo estacionário e

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$$

então

$$\|\beta_{t,n} Y_{t-n}\|^2 \rightarrow 0,$$

quando $n \rightarrow \infty$, assegurando assim a unicidade da solução. \square

Portanto, o processo $\{Y_t\}_{t \in \mathbb{Z}}$, de média μ_Y , pode ser escrito na forma

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad t \in \mathbb{Z}, \quad (2.3.2)$$

onde ϵ_t verifica as seguintes propriedades:

(p1) $\mu_\epsilon = \mu_Z \neq 0$,

(p2) $\sigma_\epsilon^2 = \mu_\epsilon \sum_{j=1}^{\infty} E[\delta_t \beta_{t,j}^{(0)}] + \sigma_Z^2$,

(p3) $Cov(\epsilon_t, \epsilon_{t+k}) = 0, \quad k > 0$,

$$(p4) \text{ Cov}(Y_s, \epsilon_t) = 0, \quad s < t,$$

$$(p5) \text{ Cov}(\beta_{t,i}, \epsilon_{t-i}) = 0, \quad \forall i \geq 1.$$

O processo $\{Y_t\}_{t \in \mathbb{Z}}$, admitindo que tem média nula, pode também ser escrito da forma

$$Y_t = \mu_Z \sum_{i=0}^{\infty} \beta_{t,i} + \sum_{i=0}^{\infty} \beta_{t,i} \xi_{t-i}, \quad t \in \mathbb{Z}$$

onde ξ_t é ruído branco de média nula e variância σ_ϵ^2 , abreviadamente, $\xi_t \sim WN(0, \sigma_\epsilon^2)$.

Particularizando este resultado para o caso da operação Binomial *thinning*, prova-se facilmente que a variância do processo $\{\epsilon_t\}_{t \in \mathbb{Z}}$, ou do processo $\{\xi_t\}_{t \in \mathbb{Z}}$, (com $\epsilon_t = \alpha_t * Y_{t-1} - \alpha_t Y_{t-1} + Z_t$) é dada por

$$(p2') \quad \sigma_\epsilon^2 = \mu_\epsilon \sum_{i=1}^{\infty} E[\beta_{t,i}(1 - \beta_{t,i})] + \sigma_Z^2.$$

De modo a não sobrecarregar o texto, cometemos o abuso de usar as mesmas letras para designar diferentes processos.

Observação 2.3.1. Caso $\{\alpha_t\}_{t \in \mathbb{Z}}$ seja um p.e. não-estacionário e independente do processo de ruído branco $\{Z_t\}_{t \in \mathbb{Z}}$, o processo $\{\xi_t\}_{t \in \mathbb{Z}}$ verifica as seguintes propriedades:

$$(pi) \quad \mu_\xi = 0,$$

$$(pii) \quad \sigma_\xi^2 = E[\delta_t Y_{t-1}] + \sigma_Z^2,$$

$$(piii) \quad \text{Cov}(\xi_t, \xi_{t+k}) = 0, \quad k > 0.$$

Se $E[\delta_t Y_{t-1}]$ for finito, então σ_ξ^2 também é finita.

Observação 2.3.2. Dos resultados obtidos nestas duas últimas secções, podemos concluir que as condições para a existência de uma solução estacionária para os modelos DSINAR(1) e DSAR(1) (*Doubly Stochastic AutoRegressive process of order one*) — processo com suporte em \mathbb{R} , duplamente estocástico e com estrutura auto-regressiva de ordem 1 — apresentado no apêndice A, coincidem.

2.4 Momentos e parâmetros

A partir da representação médias móveis do processo (ou do facto de, quando $n \rightarrow \infty$, $Y_t^{(n)} \xrightarrow{m.q.} Y_t$) obtêm-se, respectivamente, as seguintes expressões para a média, variância e autocovariância do processo $\{Y_t\}_{t \in \mathbb{Z}}$:

$$E[Y_t] = \mu_Y = \mu_Z \sum_{i=0}^{\infty} E[\beta_{t,i}], \quad (2.4.1)$$

$$\begin{aligned} Var[Y_t] = \sigma_Y^2 = & \mu_Z^2 Var \left[\sum_{i=1}^{\infty} \beta_{t,i} \right] + \sigma_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i}^2] + \\ & + \mu_Z \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} E[\delta_{t-i} \beta_{t,i} \beta_{t,j}^{(i)}]. \end{aligned} \quad (2.4.2)$$

$$\begin{aligned} Cov(Y_t, Y_{t+k}) = \gamma_Y(k) = & \mu_Z \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} E[\delta_{t-i} \beta_{t+k,i+k} \beta_{t,j}^{(i)}] + \\ & + 2\mu_Z^2 \sum_{i=0}^{\infty} \sum_{j=i+k}^{\infty} E[\beta_{t,i} \beta_{t+k,j}] + (\sigma_Z^2 - \mu_Z^2) \sum_{i=0}^{\infty} E[\beta_{t,i} \beta_{t+k,i+k}] - \\ & - \mu_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i}] \sum_{j=0}^{\infty} E[\beta_{t+k,j}] + \mu_Z^2 \sum_{i=0}^{k-1} \sum_{j=0}^{\infty} E[\beta_{t+k,i} \beta_{t,j}], \quad k \geq 0. \end{aligned}$$

Particularizando, mais uma vez, para o caso do processo DSINAR(1) — processo DSINAR(1) generalizado cuja operação *thinning* generalizada é especificada através da distribuição Binomial — prova-se, após breves cálculos, que a expressão da média não se altera e que as expressões da variância e da autocovariância se podem escrever, respectivamente, da seguinte forma

$$\sigma_Y^2 = \mu_Z^2 Var \left[\sum_{i=0}^{\infty} \beta_{t,i} \right] + \sigma_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i}^2] +$$

$$+\mu_Z \sum_{i=1}^{\infty} E[\beta_{t,i}(1 - \beta_{t,i})],$$

ou seja,

$$\sigma_Y^2 = \mu_Z^2 \text{Var} \left[\sum_{i=0}^{\infty} \beta_{t,i} \right] + (\sigma_Z^2 - \mu_Z) \sum_{i=0}^{\infty} E[\beta_{t,i}^2] + \mu_Y \quad (2.4.3)$$

e

$$\begin{aligned} \gamma_Y(k) &= \sigma_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i} \beta_{t+k,i+k}] + \mu_Z^2 \sum_{i=0}^{k-1} \sum_{j=0}^{\infty} E[\beta_{t+k,i} \beta_{t,j}] + \\ &+ 2\mu_Z^2 \sum_{i=0}^{\infty} \sum_{j=i+k}^{\infty} E[\beta_{t,i} \beta_{t+k,j}] - \mu_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i} \beta_{t+k,i+k}] - \\ &- \mu_Z^2 \sum_{i=0}^{\infty} E[\beta_{t,i}] \sum_{j=0}^{\infty} E[\beta_{t+k,j}] + \mu_Z \sum_{i=1}^{\infty} E[\beta_{t,i} (\beta_{t+k,k} - \beta_{t+k,k+i})], \quad (2.4.4) \end{aligned}$$

para $k \geq 0$.

Note-se (equação (2.4.3)) que se $\sigma_Z^2 \geq \mu_Z$ então $\sigma_Y^2 > \mu_Y$. No caso bastante usual em que se tem ruídos poissonianos de parâmetro λ , *i.e.*, $Z_t \sim Po(\lambda)$, resulta que

$$\sigma_Y^2 = \lambda^2 \text{Var} \left[\sum_{i=0}^{\infty} \beta_{t,i} \right] + \mu_Y.$$

Como seria de esperar, e contrariamente aos processos INAR(1), o facto de os ruídos serem poissonianos não implica que o processo $\{Y_t\}_{t \in \mathbb{Z}}$ possua a mesma distribuição marginal. Pode assim afirmar-se que o processo DSINAR(1) não só é adequado para modelar séries temporais estacionárias, em que a evolução ao longo do tempo é influenciada por uma série explicativa estocástica, X_t ($\alpha_t = f(X_t)$), como também para modelar séries temporais estacionárias e sobre-dispersas, cuja sobre-dispersão é causada pela influência dessa série explicativa.

Embora no caso de os processos DSINAR(1) generalizados esta conclusão não seja fácil de retirar, os resultados obtidos a partir de estudos por simulação parecem, contudo, não contrariá-la.

Observação 2.4.1. Se $\{Y_t\}_{t \in \mathbb{Z}}$ é um processo DSINAR(1) generalizado, estacionário de 2^a ordem, cuja função de autocovariância é dada por

$$\gamma_Y(k) = Cov(Y_t, Y_{t+k}) = E[\beta_{t+k,k} Y_t^2] - \mu_Y E[\beta_{t+k,k} Y_t], \quad k \geq 0,$$

então,

$$Cov(Y_t, Y_{t+k} | \mathcal{A}_{t+1+}^\alpha) = \beta_{t+k,k} Var[Y_t | \mathcal{A}_{t+1+}^\alpha], \quad k \geq 0,$$

onde \mathcal{A}_{t+}^α é a álgebra- σ gerada por $\{\alpha_s : s \geq t\}$. Portanto, $Corr(Y_t, Y_{t+k} | \mathcal{A}_{t+1+}^\alpha) = \beta_{t+k,k}$.

Observação 2.4.2. Se compararmos as expressões (2.4.1), (2.4.3) e (2.4.4) com as correspondentes expressões para o processo com suporte em \mathbb{R} , duplamente estocástico e com estrutura auto-regressiva de ordem 1, DSAR(1), (expressões (A.0.3)-(A.0.5), apresentadas no apêndice A) chegamos à conclusão, como seria de esperar, da igualdade das médias e da forte semelhança entre as variâncias e as covariâncias destes dois processos. O resto da diferença entre as variâncias (covariâncias) destes processos é originado pelo facto de $E[(\alpha_t * Y_{t-1})^2] = E[(\alpha_t Y_{t-1})^2] + E[\alpha_t(1 - \alpha_t)Y_{t-1}]$.

2.5 Condições necessárias e suficientes para a estacionariedade do processo

A proposição seguinte (baseada no teorema 2.5 de Pourahmadi (1986)) estabelece uma condição necessária de existência de solução estacionária para que o processo DSINAR(1) generalizado.

Proposição 2.5.1. Se, $\forall t \in \mathbb{Z}$, $E[\log(\delta_t)] < \infty$, $\mu^{(1)} = E[\log(\alpha_t)] < \infty$ e $\mu^{(2)} = E[\log(\alpha_t) | \delta_{t-i}] < \infty$, $\forall i \geq 0$, então,

$$(1) \sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty \implies \mu^{(1)} < 0$$

$$(2) \lim_{n \rightarrow \infty} \sum_{i=0}^n \sum_{j=i+1}^n E[\delta_{t-i}^{(n-i)} \beta_{t+k,i+k} \beta_{t,j}^{(i)}] < \infty, \quad \forall k \geq 0, \implies \mu^{(2)} < 0.$$

Demonstração. (1) Se $\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$, então $E[\beta_{t,n}^2] \rightarrow 0$, quando $n \rightarrow \infty$. Pelo teorema da desigualdade condicional de Jensen (teorema 15-1I, Burrill

(1972)) tem-se que

$$E[\beta_{t,n}^2] \geq \exp \{2nE[\log(\alpha_t)]\}$$

o que implica, pelo critério da comparação, que $\mu^{(1)} < 0$.

(2) De igual modo, se, $\forall i \geq 0$, $\lim_{n \rightarrow \infty} \sum_{j=0}^{n-i} E[\delta_{t-i}^{(n-i)} \beta_{t,i} \beta_{t,j}^{(i)}] < \infty$, então $E[\delta_{t-i} \beta_{t,i} \beta_{t,n}^{(i)}] \rightarrow 0$, quando $n \rightarrow \infty$. Igualmente pela desigualdade condicional de Jensen e pelo critério de comparação, o facto de

$$E[\delta_{t-i}^{(n-i)} \beta_{t,i} \beta_{t,n}^{(i)}] = E \left[\exp \{ \log(\delta_{t-i}^{(n-i)}) + \sum_{j=0}^{n-1} \log(\alpha_{t-j}) \} | \delta_{t-i}^{(n-i)} \right]$$

convergir para zero, quando $n \rightarrow \infty$, implica que $\mu^{(2)} < 0$, $\forall i \geq 0$. □

No que diz respeito ao processo DSINAR(1), $\mu^{(1)} = E[\log(\alpha_t)] < \infty$ é uma condição necessária para que $\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$.

Nos teoremas seguintes estabelecem-se, respectivamente, uma condição necessária e uma condição suficiente de estacionariedade para o processo DSINAR(1).

Teorema 2.5.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo DSINAR(1), estacionário de 2ª ordem, e $S_n = \sum_{j=0}^{n-1} V_j$, com $V_j = \log(\alpha_{t-j})$. Então,*

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty \iff \sum_{i=0}^{\infty} M_{S_i}(2) < \infty, \quad \forall t \in \mathbb{Z},$$

onde $M_{S_n}(\cdot)$ é a função geradora de momentos de S_n .

Demonstração. Por definição,

$$M_{S_n}(2) = E[\exp(2S_n)].$$

Após simples cálculos, prova-se que

$$M_{S_n}(2) = E \left[\prod_{j=0}^{n-1} \alpha_{t-j}^2 \right].$$

Desta forma,

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$$

é equivalente à condição

$$\sum_{i=0}^{\infty} M_{S_i}(2) < \infty.$$

□

O teorema seguinte foi estabelecido por Karlsen, em 1990, para p.e.'s duplamente estocásticos, com suporte em \mathbb{R} . Contudo, este pode igualmente ser adoptado para os processos DSINAR(1) estacionários.

Teorema 2.5.2. *Seja $\{X_t\}_{t \in \mathbb{Z}}$ um processo Gaussiano e $E[Z_t^2] < \infty$. Então,*

$$(E[\alpha_t^{2\rho}])^{1/\rho} < 1$$

é uma condição suficiente para que o processo DSINAR(1) seja estacionário de 2ª ordem, onde $\rho = \sum_{k=-\infty}^{\infty} |\rho_X(k)|$.

Demonstração. Ver a demonstração deste teorema em Karlsen (1990) (teorema 4.2), conjuntamente com a segunda nota (p. 137), do mesmo artigo. □

2.6 Especificação do processo DSINAR(1) generalizado em termos da sua distribuição marginal

A simulação é de uma enorme utilidade no estudo dos modelos de séries temporais de valores inteiros, dado que a necessidade de gerar sequências de v.a.'s dependentes com determinadas distribuição marginal e estrutura de correlação é bastante comum (McKenzie (2003), p. 16).

Assim, com este propósito em mente e de modo a gerar, concretamente, realizações do processo DSINAR(1) generalizado, estacionário de 2ª ordem e com uma dada distribuição marginal do tipo discreto, vamos estabelecer a seguinte proposição:

Proposição 2.6.1. *Sejam $\{Y_t\}_{t \in \mathbb{Z}}$ um processo DSINAR(1) generalizado, estacionário de 2ª ordem e $\{Z_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s i.i.d. Então, admitindo que $\alpha_t = f(X_t)$, com $\{X_t\}_{t \in \mathbb{Z}}$ um p.e., e designando por $M_{Y_t}(s)$ a f.g.m. de $\{Y_t\}_{t \in \mathbb{Z}}$, tem-se que*

$$M_{Y_t|X_t=x_t}(s) = M_{Y_{t-1}|X_t=x_t}(\log(\alpha(x_t)e^s + 1 - \alpha(x_t))) \times M_{Z_t}(s),$$

ou seja,

$$M_{Y_t}(s) = E[M_{Y_{t-1}}(\log(\alpha_t e^s + 1 - \alpha_t))] \times M_{Z_t}(s).$$

Verifica-se, assim, que a geração de realizações do processo DSINAR(1) generalizado, estacionário de 2ª ordem e com uma dada distribuição marginal do tipo discreto, implica (tal como no caso dos processos INAR(1)) a especificação da distribuição do processo $\{Z_t\}_{t \in \mathbb{Z}}$, igualmente do tipo discreto, e a resolução da seguinte equação

$$M_{Z_t}(s) = \frac{M_{Y_t}(s)}{E[M_{Y_{t-1}}(\log(\alpha_t e^s + 1 - \alpha_t))]}.$$

2.7 DSINAR(1) generalizado: um modelo *observation driven*

Seja $\{Y_t\}_{t \in \mathbb{Z}}$ o processo DSINAR(1) generalizado, estacionário de 2ª ordem, satisfazendo a equação às diferenças

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z},$$

e \mathcal{H}_t a álgebra- σ gerada por $\{X_t, (Y_{t-1}, X_{t-1}), (Y_{t-2}, X_{t-2}), \dots\}$, com X_t um p.e. univariado com suporte em \mathbb{R} e $\alpha_t = f(X_t)$.

Este processo pode ser escrito, tal como já foi visto, da seguinte forma

$$Y_t - \mu_Z = \alpha_t Y_{t-1} + \xi_t, \quad t \in \mathbb{Z},$$

onde ξ_t é ruído branco de média nula e variância σ_ϵ^2 , finita.

Se definirmos a equação de estado, S_t , como

$$S_{t+1} = \alpha_{t+1} S_t + V_t, \quad t \in \mathbb{Z},$$

onde

$$S_1 = Y_1 = \mu_Z \sum_{i=0}^{\infty} \beta_{1,i} + \sum_{i=0}^{\infty} \beta_{1,i} \xi_{1-i}, \quad t \in \mathbb{Z}$$

e $V_t = \xi_{t+1}$, verifica-se facilmente que

$$S_2 = \mu_Z \{\alpha_2 + \alpha_2 \alpha_1 + \dots\} + \xi_2 + \alpha_2 \xi_1 + \dots,$$

tal que $S_2 = Y_2 - \mu_Z$. De igual modo, prova-se então que o processo $\{Y_t\}_{t \in \mathbb{Z}}$ satisfaz a seguinte equação de observação:

$$Y_t = S_t + \mu_Z,$$

com $g_t = 1$, $f_t = \alpha_{t+1}$ e $W_t = \mu_Z$ ^[2].

Se admitirmos ainda que Z_t segue uma distribuição Poisson de parâmetro μ_Z , $\mu_Z \geq 0$, e que a distribuição que especifica a operação *thinning* é igualmente Poisson de parâmetro $\alpha_t Y_{t-1}$ então, tem-se que

$$Y_t | \mathcal{H}_t \sim Po(\mu_t),$$

onde $\mu_t = \alpha_t Y_{t-1} + \mu_Z$.

Como $Y_t \in \mathbb{Z}^+$ escreve-se, em geral, $\mu_t = \exp(\omega X_t) Y_{t-1}$ — de modo a garantir que $\mu_t \in \mathbb{Z}^+$ —, com $\omega \in \mathbb{R}$ um parâmetro a estimar. Deste modo

$$\mu_t = \exp(\omega X_t) Y_{t-1} + \mu_Z.$$

^[2]Recorde-se que as equações de observação e de estado escrevem-se, respectivamente, como

$$Y_t = g_t X_t + W_t, \quad t \in \mathbb{Z},$$

onde $W_t \sim WN(0, \sigma_W^2)$ e g_t é um escalar (possivelmente dependente de t) e

$$X_{t+1} = f_t X_t + V_t, \quad t \in \mathbb{Z},$$

onde $V_t \sim WN(0, \sigma_V^2)$, f_t é um escalar (também possivelmente dependente de t) e $\{V_t\}_{t \in \mathbb{Z}}$ e $\{W_t\}_{t \in \mathbb{Z}}$ são dois processos não-correlacionados.

Ou seja, segundo a notação introduzida no capítulo 1, o modelo DSINAR(1) Poisson não é mais do que um $ModP_1(1, 0)$ com $\lambda = 0$, $\alpha_1 = 1$, $f(X_t) = \exp(\omega X_t)$ e $g(X_t) = \mu_Z$, $\forall t \in \mathbb{Z}$.

2.8 Estimação dos parâmetros do processo DSINAR(1) generalizado

2.8.1 Estimadores de máxima verosimilhança condicional

Nesta secção vamos proceder à estimação dos parâmetros do processo DSINAR(1) generalizado, através da estimação de máxima verosimilhança condicional. Para tal, considere-se $\alpha_t = f(X_t)$, com $\{X_t\}_{t \in \mathbb{Z}}$, um processo explicativo com suporte em \mathbb{R} e com uma determinada estrutura de covariância.

Sejam $\{(X_t, Y_t), t = 1, \dots, n\}$ as observações dos processos $\{X_t\}_{t \in \mathbb{Z}}$ e $\{Y_t\}_{t \in \mathbb{Z}}$, onde $\{X_t\}_{t \in \mathbb{Z}}$ é um p.e. com função densidade conjunta, $h(\vec{\theta}_3)$, sendo $\vec{\theta}_3$ o vector de parâmetros desconhecido de $\{X_t\}_{t \in \mathbb{Z}}$.

Caso a distribuição de $\{Z_t\}_{t \in \mathbb{Z}}$ seja conhecida (ou razoavelmente admitida) a menos de um vector θ_2 , de parâmetros desconhecidos, é possível escrever a função de verosimilhança do processo DSINAR(1) generalizado de uma forma explícita. Seja então, $L(\vec{\theta}) = L(\vec{y}_n, \vec{x}_n; \vec{\theta})$, a função de verosimilhança do processo, onde $\vec{\theta} = (\vec{\theta}_1, \vec{\theta}_2, \vec{\theta}_3)$ é o vector de parâmetros desconhecidos ($\vec{\theta}_1$ é o vector de parâmetros associado à função $\alpha_t = f(X_t)$). Após alguns cálculos, verifica-se que

$$L(\vec{\theta}) = P(Y_1 = y_1 | X_1 = x_1) \prod_{t=2}^n P(Y_t = y_t | Y_{t-1} = y_{t-1}, X_t = x_t) h(x_1, \dots, x_n),$$

com

$$P(Y_t = y_t | Y_{t-1} = y_{t-1}, X_t = x_t) =$$

$$P(\alpha_t \circ^G Y_{t-1} + Z_t = y_t | Y_{t-1} = y_{t-1}, X_t = x_t) = P(W_t + Z_t = y_t)$$

onde $W_t \sim G(\cdot)$ e Z_t é uma v.a. com uma dada distribuição discreta e independente de W_t . Então

$$P(W_t + Z_t = y_t) = \sum_{j=k_1(t)}^{k_2(t)} P(W_t = y_t - j) P(Z_t = j),$$

em que os limites inferior e superior do somatório, $k_1(t)$ e $k_2(t)$, respectivamente, são especificados consoante as distribuições de W_t e de Z_t .

Conhecida, portanto, a densidade de probabilidade da v.a. Y_t dado X_t e a densidade de probabilidade da v.a. Z_t o cálculo das estimativas de máxima verosimilhança dos parâmetros do modelo torna-se relativamente fácil de efectuar, mais que não seja recorrendo a métodos numéricos.

Desconhecendo-se a distribuição de Y_t condicionada a X_t pode considerar-se Y_1 como determinístico e proceder-se à maximização da verosimilhança condicionada à primeira observação, ou seja,

$$L(\vec{\theta}) = \prod_{i=2}^n g_i(y_{i-1}, y_i, x_i; \vec{\theta}_1, \vec{\theta}_2) L(x_1, \dots, x_n; \vec{\theta}_3), \quad (2.8.1)$$

onde $g_t(y_{t-1}, y_t, x_t; \vec{\theta}_1, \vec{\theta}_2) = \sum_{j=k_1(t)}^{k_2(t)} P(W_t = y_t - j) P(Z_t = j)$, $\forall t = 2, \dots, n$, a qual não depende de $\vec{\theta}_3$.

Considerando o caso em que a distribuição que especifica a operação *thinning* é a distribuição Binomial (processo DSINAR(1)) tem-se que $W_t \sim B(y_{t-1}, \alpha(x_t))$ (com $\alpha(x_t) := \alpha_t | X_t = x_t$, tal que $\alpha(x_t) \in (0, 1)$) e

$$g_t(y_{t-1}, y_t, x_t; \vec{\theta}) = \sum_{m=0}^{k_2(t)} \binom{y_{t-1}}{m} (\alpha(x_t))^m (1 - \alpha(x_t))^{y_{t-1}-m} P(Z_t = y_t - m),$$

onde $k_2(t) = \min(y_{t-1}, y_t)$.

Em geral, e por uma questão de interpretação dos parâmetros (em termos do *Odds ratio*), é usual assumir-se que

$$\alpha_t = \frac{\exp(\omega X_t)}{1 + \exp(\omega X_t)}, \quad \omega \in \mathbb{R} \quad (\omega : \text{parâmetro desconhecido}),$$

vindo, por isso, $\vec{\theta} = (\omega, \vec{\theta}_2, \vec{\theta}_3)$.

Se o tamanho da amostra for suficientemente grande a primeira observação não contribui de modo significativo para a verosimilhança total pelo que, em geral, pode considerar-se como função de verosimilhança a função definida em (2.8.1).

Os estimadores de máxima verosimilhança condicional obtêm-se através da maximização do logaritmo da função de verosimilhança, com respeito

ao vector $\vec{\theta}$, recorrendo, por exemplo, a métodos numéricos (*e.g.*, método iterativo de Newton-Raphson), uma vez que as equações de verosimilhança conduzem-nos a um sistema de equações não-lineares. Dado que não existem pacotes informáticos que, uma vez identificadas as distribuições de W_t e de Z_t , nos forneçam as estimativas dos parâmetros, torna-se necessário proceder à sua programação^[3].

Tal como é sobejamente sabido, se recorrermos aos métodos numéricos para a obtenção das estimativas de máxima verosimilhança condicional e se dispuser de valores iniciais para $\vec{\theta}$, razoavelmente próximos do máximo do logaritmo da função de verosimilhança, o processo iterativo acelera consideravelmente. Contudo, no caso do processo DSINAR(1) generalizado a obtenção de estimativas iniciais dos parâmetros não parece muito evidente.

Relativamente ao modelo DSINAR(1) e, para o caso bastante estudado, em que Z_t segue uma distribuição de Poisson de parâmetro λ ($\lambda \equiv \mu_Z$), os estudos por simulação evidenciaram ser possível obter *boas* estimativas de máxima verosimilhança condicional (através do método iterativo de Newton-Raphson) dos parâmetros do modelo, (λ, ω) , utilizando-se como valores iniciais ao método iterativo de maximização, $(\hat{\lambda}, 0)$, onde $\hat{\lambda} = \hat{\mu}_Z$ se obtém a partir de uma amostra de dimensão n , (Y_1, \dots, Y_n) , do processo Y_t do seguinte modo:

$$\hat{\lambda}_Z = \hat{\mu}_Z = \frac{1}{n-1} \sum_{t=2}^n Y_t - \frac{\hat{k}}{n-1} \sum_{t=2}^n Y_{t-1},$$

com

$$\hat{k} = \frac{\sum_{t=2}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^n (Y_{t-1} - \bar{Y})^2}.$$

Este estimador de λ não é mais do que o *estimador dos mínimos quadrados condicionais (EMQC)* considerando $E[Y_t | \mathcal{A}_{t-1}^Y] = \text{constante}$, $\forall t \in \mathbb{Z}$ (aqui $\mathcal{A}_{t-1}^Y = \sigma(Y_{t-1}, Y_{t-2}, \dots, Y_1)$ é a álgebra- σ gerada por $\{Y_{t-k}, k \geq 1\}$). Ou seja, $\hat{\mu}_Z$ é o EMQC de λ quando se assume que $\{\alpha_t\}_{t \in \mathbb{Z}}$ constitui uma sucessão de v.a.'s i.i.d. em vez de um p.e. estacionário de 2^a ordem.

^[3]Quer os estudos por simulação destes processos (em particular, o estudo do desempenho dos diversos métodos de estimação), quer a modelação do número de óbitos, foram efectuados a partir de programas elaborados em Fortran. No apêndice B apresentam-se, para uma especificação do modelo DSINAR(1) generalizado, as expressões das equações de verosimilhança e das suas derivadas, em ordem aos parâmetros do modelo.

A estimação dos parâmetros do modelo a partir do método dos mínimos quadrados condicionais irá ser estudada detalhadamente na secção 3.6.2.

Refira-se, a título de exemplo, que no caso de se estar a utilizar a operação Poisson *thinning*, o método iterativo de Newton-Raphson não converge quando se utiliza $(\tilde{\lambda}, 0)$, como valores iniciais ao método iterativo, apesar de os estudos por simulação efectuados no capítulo 4 revelarem que a média amostral (obtida em 100 realizações simuladas) das estimativas ser bastante próxima do verdadeiro valor de λ . Esta falta de convergência fica a dever-se ao facto de, para estes valores iniciais de λ e ω , a derivada em ordem a ω do logaritmo da função de verosimilhança tomar um valor demasiado pequeno e o programa o assumir como sendo menos infinito; o que não acontece quando se usa $(1, 0)$ como valores iniciais ao método iterativo. Para estes valores iniciais e para amostras de dimensão $n \geq 100$ obtiveram-se *boas* estimativas de máxima verosimilhança condicional dos parâmetros do modelo.

Como alternativa, e igualmente considerando que Z_t segue uma distribuição de Poisson de parâmetro λ , os estudos por simulação evidenciaram ainda ser possível obter *boas* estimativas de máxima verosimilhança condicional (através do método iterativo de Newton-Raphson) dos parâmetros do modelo, (λ, ω) , utilizando-se como valores iniciais ao método iterativo de maximização aqueles que conduzem a estimativas de máxima verosimilhança condicional que minimizam a soma dos quadrados dos resíduos. Mais concretamente, no caso em que $\{X_t\}_{t \in \mathbb{Z}}$ é um p.e. estacionário 2^a ordem obteve-se na secção 2.4 a seguinte expressão para a média do processo DSINAR(1) generalizado

$$\begin{aligned} \mu_Y &= \mu_Z \sum_{i=0}^{\infty} E[\beta_{t,i}] = \\ &= \mu_Z \{1 + E[\alpha_t] + E[\alpha_t \alpha_{t-1}] + \dots\}. \end{aligned}$$

Assim, dada uma a.a. de dimensão n dos processos $\{X_t\}_{t \in \mathbb{Z}}$ e $\{Y_t\}_{t \in \mathbb{Z}}$, $\{(X_t, Y_t), t = 1, \dots, n\}$, e admitindo que $\alpha_t = \exp(\omega X_t)$, escolha-se um valor de $\omega \in \mathbb{R}$, digamos ω_1 . Com esta informação é possível obter-se uma estimação aproximada de μ_Z do seguinte modo:

$$\hat{\mu}_Z = \frac{\bar{Y}}{1 + \sum_{i=1}^K \frac{1}{n-i+1} \sum_{j=i}^n \beta_{j,i}},$$

onde

$$\beta_{j,i} = \begin{cases} \prod_{m=0}^{j-1} \alpha_{j-m}, & i \geq 1 \\ 1, & i = 0, \end{cases}$$

e $K \in \mathbb{Z}$ (quanto maior for o valor de K , melhor serão as estimativas de μ_Z). Como iremos ver na secção 2.9, os resíduos, digamos ξ_t , são estimados do seguinte modo:

$$\begin{aligned}\hat{\xi}_t &= Y_t - E[Y_t | Y_{t-1}, X_t, \hat{\omega}_1, \hat{\lambda}] = \\ &= Y_t - \exp(\hat{\omega}_1 X_t) Y_{t-1} - \hat{\lambda}, \quad t = 2, \dots, n.\end{aligned}$$

Através dos estudos por simulação constatou-se, utilizando-se um valor de $K = 5$ e fazendo variar o valor de ω , com incrementos de uma unidade, que se obtêm igualmente *boas* estimativas de máxima verosimilhança condicional dos parâmetros do modelo para todos os valores de ω utilizados.

Apesar de esta alternativa à escolha dos valores iniciais ao método iterativo de maximização se ter revelado adequada, parece-nos mais apropriado recorrer às propostas anteriores, uma vez que em qualquer dos casos citados a obtenção das estimativas de máxima verosimilhança se faz de um modo muito mais célere do que neste caso. Para além disso, qualquer das soluções adoptadas conduz a semelhantes erros quadráticos médios amostrais das estimativas dos parâmetros do modelo.

Embora não se tenha conseguido obter as propriedades assintóticas dos estimadores de máxima verosimilhança condicional, os resultados por simulação obtidos no capítulo 4 revelam que deve ser possível provar que os estimadores de máxima verosimilhança condicional dos parâmetros do processo DSINAR(1) generalizado são fracamente consistentes, no sentido em que convergem em média e são assintoticamente centrados e normais.

2.9 Teste de ajustamento do modelo DSINAR(1) generalizado

Após a estimação dos parâmetros do modelo iremos, nesta secção, avaliar a qualidade do ajustamento do modelo estimado. Esta avaliação é feita, em geral, através da confrontação entre os valores observados e as correspondentes predições, obtidas a partir do modelo ajustado.

Define-se como “melhor” preditor aquele que minimiza o erro quadrático médio entre os valores observados (Y_t) e os valores estimados (\hat{Y}_t). Por outras palavras, é aquele que, segundo um dado critério (neste

caso a minimização do erro quadrático médio), *melhor compreende* a estrutura dos dados. Assim sendo, dadas as observações dos processos $\{X_t\}_{t \in \mathbb{Z}}$ e $\{Y_t\}_{t \in \mathbb{Z}}$, $\mathcal{H}_t = \{(X^{(t)}, Y^{(t-1)})\}$ (onde $X^{(t)} = (X_1, X_2, \dots, X_t)$ e $Y^{(t-1)} = (Y_1, Y_2, \dots, Y_{t-1})$, para $t = 2, \dots, n$), o “melhor” preditor define-se como $E[Y_t | \mathcal{H}_t]$, pelo que os resíduos, $\xi_t = Y_t - \hat{Y}_t$, são definidos através da diferença

$$\begin{aligned} \xi_t &= Y_t - E[Y_t | \mathcal{H}_t] \\ &= Y_t - \alpha_t Y_{t-1} - \mu_Z \\ &= \alpha_t \circ^G Y_{t-1} - \alpha_t Y_{t-1} + Z_t - \mu_Z, \quad t = 2, \dots, n. \end{aligned}$$

Assim, testar que o modelo estimado se ajusta bem ao conjunto de dados, isto é, que descreve bem o conjunto de dados em causa, consiste em estudar o comportamento dos resíduos.

Durante o estudo do modelo DSINAR(1) generalizado (secções 2.2 e 2.3) foram verificados vários pressupostos relativamente aos resíduos; nomeadamente o facto de estes constituírem uma sucessão de v.a.’s não-correlacionadas, de média nula e variância constante. No caso de o modelo ter sido correctamente estabelecido, os resíduos deverão, se não confirmar, pelo menos não negar estes pressupostos. Assim, se a análise dos resíduos permitir “concluir” que estes se comportam como um ruído branco, então pode concluir-se que o modelo estimado se ajusta bem à série temporal em causa.

Uma vez que o ruído branco é caracterizado essencialmente através das funções de autocorrelação (e de autocorrelação parcial) vamos definir o seguinte teste para a hipótese nula de que as primeiras m autocorrelações da série dos resíduos são nulas, ou seja, definir o teste estatístico

$$H_0 : \rho_\xi(1) = \rho_\xi(2) = \dots = \rho_\xi(m) = 0 \quad vs \quad H_1 : \exists_{k, k=1, \dots, m} : \rho_\xi(k) \neq 0.$$

Dado que o processo estacionário DSINAR(1) generalizado é um processo não-linear não nos parece muito simples a obtenção da distribuição assintótica do vector estimador das m funções de autocorrelação, $(\hat{\rho}_\xi(1), \hat{\rho}_\xi(2), \dots, \hat{\rho}_\xi(m))$, sob a hipótese de que ξ_t é ruído branco.

De modo a ultrapassar esta dificuldade, propõe-se o uso da metodologia *bootstrap* para a obtenção de intervalos de confiança para as funções de

autocorrelação e autocorrelação parcial dos resíduos escritas, de modo abreviado, como FAC e FACP, respectivamente. Como no caso do modelo DSINAR(1) generalizado os resíduos não constituem uma sequência de v.a.'s independentes não parece, portanto, adequado recorrer a uma metodologia que gere amostras *bootstrap* de v.a.'s independentes, mas sim que gere amostras *bootstrap* que reproduzam a estrutura de correlação da amostra original. Deste modo, vamos recorrer a metodologias de reamostragem para dados dependentes, ou seja, a metodologias de reamostragem por blocos (propostas, *e.g.*, por Künsch (1989) e Politis e Romano (1994)).

Na reamostragem por blocos salientam-se apenas duas abordagens: a reamostragem por blocos móveis (*moving blocks bootstrap*), proposta, *e.g.*, por Künsch em 1989, e a reamostragem estacionária por blocos, proposta por Politis e Romano em 1994. No essencial, o que separa estas duas metodologias é o facto de, com base numa dada amostra aleatória, se obter uma pseudo-série não-estacionária no primeiro caso e uma pseudo-série estacionária no segundo caso^[4].

Ambas as metodologias são metodologias de reamostragem por blocos, no sentido em que são construídos blocos de v.a.'s consecutivas, de modo a preservar a estrutura de autocorrelação da série. Na reamostragem por blocos móveis são construídos blocos sobrepostos de igual dimensão b , com b uma variável determinística. Na reamostragem estacionária por blocos, são construídos blocos de v.a.'s, sobrepostos ou não e de diferentes dimensões V_i , com V_i uma v.a. com distribuição Geométrica de parâmetro p . A primeira variável de cada bloco é também neste caso determinada aleatoriamente.

De entre o que é possível estudar a partir de estas duas abordagens, nesta secção vamos centrar-nos na construção de intervalos de confiança (IC) *bootstrap* para as FAC e FACP dos resíduos.

Seja, então, $\{\xi_t, t = 1, \dots, n\}$ uma a.a. do processo ruído branco $\{\xi_t\}_{t \in \mathbb{Z}}$, com função de distribuição $F_\xi(\cdot)$, desconhecida. Sejam $\{\hat{\rho}_\xi(j)\}_{j=1,2,\dots}$ e $\{\hat{\phi}_\xi(j)\}_{j=1,2,\dots}$ os estimadores usuais da FAC e FACP, respectivamente.

Os IC por blocos móveis obtêm-se de acordo com o seguinte algoritmo:

1. Com base na a.a. $\{\xi_t, t = 1, \dots, n\}$, construa-se o vector de blocos sobrepostos de dimensão b , $(B_1, B_2, \dots, B_{n-b+1})$, onde

^[4]**Proposição 1 (Politis e Romano, (1994)).** Condicional a $(\xi_1, \xi_2, \dots, \xi_n)$, a amostra *bootstrap*, $(\xi_1^*, \xi_2^* \dots, \xi_n^*)$, é estacionária.

$$B_i = (\xi_i, \dots, \xi_{i+b-1}).$$

2. Selecciona-se aleatoriamente (com reposição) $k = \lfloor \frac{n}{b} \rfloor$ blocos (onde cada bloco tem igual probabilidade de ser seleccionado) através dos quais se obtém uma nova pseudo-série $(\xi_1^*, \dots, \xi_m^*)$, com $m = kb \simeq n$.
3. Com esta nova pseudo-série estimem-se, através dos estimadores usuais e para cada *lag* j , as correspondentes FAC e a FACP, $\hat{\rho}_\xi^*(j)$ e $\hat{\phi}_\xi^*(j)$, respectivamente.
3. Repitam-se N vezes os passos 2 e 3.

Após a obtenção das N FAC e a FACP amostrais, tome-se um valor $\beta \in (0, 1)$ e construam-se, então, os respectivos intervalos de confiança $(1 - \beta) \times 100\%$. As FAC e FACP amostrais escrevem-se, no *lag* j , respectivamente, como $\tilde{\rho}_\xi^*(j)_i$ e $\tilde{\phi}_\xi^*(j)_i$, com $i = 1, \dots, N$.

O intervalo de confiança (bilateral) *bootstrap* aproximado, $[\hat{\rho}_\xi(j) - \chi_B(1 - \beta/2), \hat{\rho}_\xi(j) - \chi_B(\beta/2)]$, para $\rho_\xi(j)$ é construído do seguinte modo:

- (a) Ordene-se, por ordem crescente, os valores de $\tilde{\rho}_\xi^*(j)_i - \hat{\rho}_\xi(j)$, para $i = 1, \dots, N$;
- (b) Calcule-se $a = \lfloor \frac{\beta N}{2} + 1 \rfloor$ ($\lfloor x \rfloor$ representa o menor inteiro inferior a x) e $b = N - (a - 1)$;
- (c) Determinem-se $\chi_B^*(\beta/2)$ e $\chi_B^*(1 - \beta/2)$ como sendo, respectivamente, os elementos de ordem a e b da amostra ordenada obtida em (a).

Note-se que os IC propostos são totalmente não-paramétricos, no sentido em que não se admitem quaisquer modelos para séries temporais nem quaisquer distribuições assintóticas. Estabelece-se apenas que

$$(\hat{\rho}_\xi^*(j) - \hat{\rho}_\xi(j)) \stackrel{d}{=} (\hat{\rho}_\xi(j) - \rho_\xi(j)), \quad j = 1, \dots, m.$$

Os quantis $\chi_B(1 - \beta/2)$ e $\chi_B(\beta/2)$, tais que,

$$P(\rho_\xi(j) < \hat{\rho}_\xi(j) - \chi_B(1 - \beta/2)) \approx P(\rho_\xi(j) > \hat{\rho}_\xi(j) - \chi_B(\beta/2)) \approx \beta/2,$$

podem ser, portanto, aproximados por $\chi_B^*(1 - \beta/2)$ e $\chi_B^*(\beta/2)$, respectivamente. Os intervalos de confiança aproximados para $\phi_\xi(j)$ obtêm-se de igual modo.

De modo a expor a reamostragem estacionária por blocos tome-se igualmente uma a.a. de dimensão n do processo $\{\xi_t\}_{t \in \mathbb{Z}}$. Com base nesta amostra, construa-se um vector de blocos, $(B_{U_1, V_1}, B_{U_2, V_2}, \dots, B_{U_K, V_K})$ (com K aleatório), onde $\{U_i, i = 1, \dots, K\}$ é uma sequência de v.a.'s com distribuição Uniforme em $[1, n]$ e B_{U_i, V_i} é um bloco de dimensão aleatória, V_i , governada por uma distribuição Geométrica de parâmetro p , tal que $B_{U_i, V_i} = (\xi_{U_i}, \dots, \xi_{U_i+V_i-1})$. Neste método de reamostragem surge, contudo, o problema da escolha do parâmetro p . No capítulo 4 iremos apresentar uma proposta para a escolha deste parâmetro.

O algoritmo seguinte permite a construção de IC por reamostragem estacionária por blocos para as FAC e FACP dos resíduos:

1. Selecciona-se, aleatoriamente, uma observação, digamos ξ_{U_1} , tal que $\xi_1^* = \xi_{U_1}$. As primeiras V_1 observações da amostra *bootstrap*, $(\xi_1^*, \xi_2^*, \dots, \xi_{V_1}^*)$, são, assim, determinadas a partir do bloco $B_{U_1, V_1} = (\xi_{U_1}, \dots, \xi_{U_1+V_1-1})$.
2. Repita-se este procedimento até $\sum_{i=1}^K V_i = n$
3. Com a pseudo-série, $(\xi_1^*, \xi_2^*, \dots, \xi_n^*)$, estimem-se, através dos estimadores usuais e para cada *lag* j , as correspondentes FAC e a FACP, $\hat{\rho}_\xi^*(j)$ e $\hat{\phi}_\xi^*(j)$, respectivamente.
3. Repitam-se N vezes os passos 1 a 3.

Os intervalos de confiança (bilaterais) *bootstrap* aproximados para $\rho_\xi(j)$ e $\phi_\xi(j)$ constroem-se do mesmo modo que no caso da reamostragem por blocos móveis.

Ao optarmos por uma destas metodologias de reamostragem (consoante se pretenda gerar amostras *bootstrap* estacionárias ou não-estacionárias) deparamo-nos com o problema da escolha do tamanho dos blocos, ou da escolha do parâmetro da distribuição Geométrica.

Künsch (1989) e Politis e Romano (1994) estabelecem (sob algumas condições, nomeadamente condições sobre os momentos e condições de dependência) que a dimensão dos blocos, no caso da reamostragem por blocos móveis, — ou a dimensão média dos blocos, no caso da reamostragem estacionária por blocos, — deverá ser proporcional à dimensão da amostra, de modo a que com qualquer dos métodos de reamostragem a probabilidade

de cobertura assintótica dos IC seja $1 - \beta$. Nomeadamente, b deverá ser um valor tal que, quando $n \rightarrow \infty$, $b \rightarrow \infty$ e $b/n \rightarrow \infty$. Do mesmo modo, $p \rightarrow 0$ e $np \rightarrow \infty$, quando $n \rightarrow \infty$.

Contudo, o desempenho destes métodos de reamostragem, por exemplo, na construção de IC tem sido estudado, pelos vários autores que se têm debruçado sobre este assunto, sob a hipótese de $\{\xi_t\}_{t \in \mathbb{Z}}$ constituir uma sucessão de v.a.'s estritamente estacionárias.

Uma vez que a área em questão é extremamente vasta, no âmbito dos objectivos deste trabalho, para poder ser intensamente explorada, nomeadamente no que diz respeito ao desempenho destes métodos no caso da sucessão $\{\xi_t\}_{t \in \mathbb{Z}}$ constituir uma sucessão de v.a.'s estacionárias de 2^a ordem, optou-se, agora, por estudar o seu desempenho recorrendo a estudos por simulação, que se apresentarão no capítulo 4.

Após a verificação da validade do modelo, vamos sugerir uma quantificação da sua ajustabilidade. Designemos a esperança matemática do erro absoluto entre a série real e a série estimada por EA_t , *i.e.*, $EA_t = E[|Y_t - \hat{Y}_t|]$, com \hat{Y}_t o estimador de Y_t . Assim, de modo a avaliar a grandeza do erro absoluto, definiu-se o seguinte coeficiente de ajustabilidade

$$p_{E_t} = \frac{E[Y_t] - EA_t}{E[Y_t]}, \quad (2.9.1)$$

com $E[Y_t]$ a esperança matemática da série real. Este coeficiente será estimado por

$$\hat{p}_{E_t} = \frac{\frac{1}{n} \sum_{j=1}^n Y_j - \frac{1}{n-1} \sum_{j=2}^n |Y_j - \hat{Y}_j|}{\frac{1}{n} \sum_{j=1}^n Y_j},$$

com (Y_1, Y_2, \dots, Y_n) uma a.a., de dimensão n , do processo $\{Y_t\}_{t \in \mathbb{Z}}$.

A ajustabilidade do modelo seria perfeita se $EA_t = 0$, o que significaria que os valores da série real e os valores da série estimada coincidiriam em média. Por este motivo, se $p_{E_t}^2$ estiver próximo de 1 isto fornece-nos alguma indicação de que o modelo ajustado é adequado para descrever o conjunto de dados em causa. Se $p_{E_t}^2$ estiver próximo de 0 isto implica que os valores da série estimada se encontram muito próximos da média da série real. Quer isto dizer, que o modelo ajustado não consegue captar a variabilidade dos dados e, portanto, não será muito adequado para explicar o conjunto de dados em causa.

Capítulo 3

Processo auto-regressivo de coeficientes i.i.d.

Vamos agora considerar — à semelhança dos modelos RCA (*Random Coefficient Autoregressive*) estudados, por exemplo, por Nicholls e Quinn (1982) — que os coeficientes do modelo DSINAR(1) generalizado, $\{\alpha_t\}_{t \in \mathbb{Z}}$, em vez de admitirem uma determinada estrutura de correlação, constituem uma sucessão de v.a.'s i.i.d. Obtém-se, assim, um modelo de coeficientes aleatórios que iremos designar por RCINAR(1) generalizado (*Random Coefficient INteger AutoRegressive*).

Ao longo deste capítulo iremos fazer um estudo análogo ao que foi realizado para os processos DSINAR(1) generalizados, particularizando, no final de cada secção, os resultados obtidos (pelos motivos descritos no capítulo 2) considerando a operação *thinning* especificada através da distribuição Binomial. Neste caso, define-se igualmente o processo RCINAR(1) generalizado simplesmente como RCINAR(1), representando-se, abreviadamente, \circ^B por $*$.

3.1 Modelo RCINAR(1) generalizado

Tendo em conta a definição de operador *thinning* generalizado, introduzida no capítulo 2 (secção 2.1), vamos definir o que se entende por processo RCINAR(1) generalizado.

Definição 3.1.1. Um processo $\{Y_t\}_{t \in \mathbb{Z}}$, com suporte nos inteiros não-negativos, diz-se um RCINAR(1) generalizado se existirem

- (i) uma sucessão de v.a.'s i.i.d. $\{\alpha_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{R}^+ , tal que $E[\alpha_t^2] < \infty$ e
- (ii) uma sucessão de v.a.'s $\{Z_t\}_{t \in \mathbb{Z}}$ inteiras, não-negativas e não-correlacionadas, com uma dada distribuição discreta, de média μ_Z e variância σ_Z^2 finita,

tais que $\{Y_t\}_{t \in \mathbb{Z}}$ satisfaz a seguinte equação às diferenças estocástica

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z},$$

onde

- (a) $\{Z_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s, independente quer da operação *thinning* generalizada, quer de $\{\alpha_t\}_{t \in \mathbb{Z}}$, e definida no mesmo espaço de probabilidade que $\{\alpha_t\}_{t \in \mathbb{Z}}$,
- (b) $Cov(Y_s, Z_t) = 0, \quad s < t,$
- (c) $\alpha_t \circ^G Y_{t-1} | \alpha_t, Y_{t-1} \sim G(\mu_t, \sigma_t^2)$, onde $G(\mu_t, \sigma_t^2)$ é uma dada distribuição do tipo discreto de média $\mu_t = \alpha_t Y_{t-1}$ e variância $\sigma_t^2 = \delta_t Y_{t-1}$, finita e
- (d) $E[\delta_t^2] < \infty.$

O caso que estamos a considerar, ou seja, em que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d., reduz o lema 2.1.1 ao seguinte:

Lema 3.1.1. $\forall t, s \in \mathbb{Z},$

- (1) $E[\alpha_t \circ^G Y_{t-1}] = E[\alpha_t]E[Y_{t-1}],$
- (2) $E[(\alpha_t \circ^G Y_{t-1})^2] = E[\delta_t]E[Y_{t-1}] + E[\alpha_t^2]E[Y_{t-1}^2],$
- (3) $E[(\alpha_s \circ^G Y_{s-1})(\alpha_t \circ^G Y_{t-1})] = E[\alpha_s Y_{s-1} \alpha_t Y_{t-1}] =$

$$= \begin{cases} E^2[\alpha_t]E[Y_{t-1}Y_{s-1}], & |s - t| > 1, \\ E[\alpha_t]E[\alpha_{s-1}Y_{s-1}Y_{t-1}], & |s - t| = 1. \end{cases}$$

A partir do lema 3.1.1 e do facto da $Cov(Y_s, Z_t) = 0$, $s < t$, obtêm-se, respectivamente, as seguintes expressões para a média e variância de Y_t :

$$\begin{aligned} E[Y_t] &= E[\alpha_t \circ^G Y_{t-1} + Z_t] \\ &= E[\alpha_t]E[Y_{t-1}] + \mu_Z \end{aligned} \quad (3.1.1)$$

e

$$\begin{aligned} Var[Y_t] &= Var[\alpha_t \circ^G Y_{t-1} + Z_t] \\ &= E[\delta_t]E[Y_{t-1}] + Var[\alpha_t]E[Y_{t-1}^2] + E^2[\alpha_t]Var[Y_{t-1}] + \sigma_Z^2. \end{aligned}$$

Verifica-se ainda que a estrutura de covariância do processo RCINAR(1) generalizado é análoga à estrutura de covariância dos processos INAR(1) e AR(1) (ambos de coeficiente $E[\alpha_t]$), isto é,

$$Cov(Y_t, Y_{t+k}) = (E[\alpha_t])^k Var[Y_t], \quad k \geq 0.$$

Estudemos agora quais as diferenças entre os momentos condicionais até à 2ª ordem dos processos RCINAR(1) generalizados e dos processos INAR(1) e AR(1). Designemos por \mathcal{A}_{t-1}^Y a álgebra- σ gerada por $\{Y_{t-k}, k \geq 1\}$. Assim, a partir da definição de operador *thinning* podemos constatar que o processo RCINAR(1) generalizado possui as seguintes propriedades:

(1)

$$E[Y_t | \mathcal{A}_{t-1}^Y] = E[\alpha_t]Y_{t-1} + \mu_Z,$$

quer isto dizer, que a média de Y_t , dada a história passada do processo em instantes $t-1 > t-2 > \dots$, é uma função linear de Y_{t-1} (propriedade partilhada pelos processos INAR(1) e AR(1)).

(2)

$$Var[Y_t | \mathcal{A}_{t-1}^Y] = E[\delta_t]Y_{t-1} + Var[\alpha_t]Y_{t-1}^2 + \sigma_Z^2,$$

isto é, a variância de Y_t , dada a história passada do processo em instantes $t-1 > t-2 > \dots$, é uma função quadrática de Y_{t-1} (contrariamente ao processo INAR(1), que é uma função linear, e ao processo AR(1), que é uma constante).

Dos resultados a que se chegou vamos particularizar, para o caso do processo RCINAR(1), aqueles que consideramos mais relevantes. No que diz respeito aos primeiro e segundo momentos, verifica-se facilmente que as expressões para a média e para a correlação não se alteram — coincidindo com as respectivas expressões de um processo INAR(1) de coeficiente $E[\alpha_t]$ —, enquanto que a variância pode escrever-se da seguinte forma:

$$\begin{aligned} \text{Var}[Y_t] &= E[\alpha_t](1 - E[\alpha_t])E[Y_{t-1}] + E^2[\alpha_t]\text{Var}[Y_{t-1}] + \\ &+ \text{Var}[\alpha_t]E[Y_{t-1}(Y_{t-1} - 1)] + \sigma_Z^2. \end{aligned} \quad (3.1.2)$$

Comparando esta expressão com a expressão da variância de um processo INAR(1) de coeficiente $E[\alpha_t]$ ^[1], constata-se que a primeira é superior (inferior) à segunda sempre que $E[Y_{t-1}^2] > E[Y_{t-1}]$ ($E[Y_{t-1}^2] < E[Y_{t-1}]$).

Para finalizar, podemos ainda concluir de (3.1.1) e de (3.1.2) que, para o processo RCINAR(1), $\alpha_t * Y_{t-1} | Y_{t-1} \not\sim \text{Binomial}$. Tal já não sucederia se, em vez de definirmos a operação *thinning* como $\alpha_t * Y_{t-1} | \alpha_t, Y_{t-1} \sim B(Y_{t-1}, \alpha_t)$, a tivéssemos definido à semelhança do que foi proposto por Steutel e Van Harn. Ou seja, se a tivéssemos definido como

$$\alpha_t * Y_{t-1} = \sum_{i=1}^{Y_{t-1}} U_i,$$

sendo $\{U_i\}$ uma sucessão de v.a.'s i.i.d. e independentes de Y_{t-1} , tal que $U_i | \alpha_t \sim \text{Ber}(\alpha_t)$ (com $\alpha_t = f(X_t)$), o que nos conduziria a um modelo INAR(1) de parâmetro $E[\alpha_t]$, pelo que $\alpha_t * Y_{t-1} | Y_{t-1} \sim B(Y_{t-1}, E[\alpha_t])$.

3.2 Estacionariedade do processo

Teorema 3.2.1. *Seja $\{\alpha_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s i.i.d. com suporte em \mathbb{R}^+ e $\{Z_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s não-correlacionadas com suporte em \mathbb{Z}^+ , com uma dada distribuição discreta de média μ_Z e variância σ_Z^2 finita, e independentes de $\{\alpha_t\}_{t \in \mathbb{Z}}$. Se*

$$E[\alpha_t^2] < 1, \quad \forall t \in \mathbb{Z},$$

^[1]Seja $\{Y_t\}$ um processo INAR(1) de coeficiente $E[\alpha_t]$. Prova-se que $\text{Var}[Y_t] = E[\alpha_t](1 - E[\alpha_t])E[Y_{t-1}] + E^2[\alpha_t]\text{Var}[Y_{t-1}] + \sigma_Z^2$.

então, existe um p.e. $\{Y_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{Z}^+ e estacionário de \mathcal{Q}^e ordem tal que

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z}.$$

Demonstração. A partir do teorema 2.2.1 prova-se, facilmente, que as condições (1) e (2), de estacionariedade dos processos DSINAR(1) generalizados, se traduzem na condição $E[\alpha_t^2] < 1$, no caso em que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d. com suporte em \mathbb{R}^+ . \square

3.3 Solução causal do processo RCINAR(1) generalizado

Proposição 3.3.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1) generalizado, estacionário de \mathcal{Q}^e ordem. Então, existe uma sucessão de v.a.'s $\{\epsilon_t\}_{t \in \mathbb{Z}}$, não-correlacionadas, tal que o processo $\{Y_t\}_{t \in \mathbb{Z}}$ admite como solução estacionária a seguinte representação causal*

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad (3.3.1)$$

onde $\epsilon_t = \alpha_t \circ^G Y_{t-1} - \alpha_t Y_{t-1} + Z_t$. $\{\epsilon_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s que satisfaz as seguintes propriedades:

$$\mu_\epsilon = E[\epsilon_t] = \mu_Z, \quad \forall t,$$

$$Cov(\epsilon_t, \epsilon_s) = E[\epsilon_t \epsilon_s] = 0 \quad \forall t, s \quad \text{com } t \neq s,$$

$$Cov(Y_s, \epsilon_t) = 0, \quad s < t$$

e

$$\sigma_\epsilon^2 = E[\delta_t] \mu_Y + \sigma_Z^2. \quad (3.3.2)$$

Demonstração. Esta demonstração é imediata a partir da proposição 2.3.1. \square

O processo $\{Y_t\}_{t \in \mathbb{Z}}$ pode igualmente ser escrito, admitindo que não tem média nula, da forma

$$Y_t = \mu_Y + \sum_{i=0}^{\infty} \beta_{t,i} \xi_{t-i}, \quad t \in \mathbb{Z}$$

onde ξ_t é ruído branco de média nula e variância σ_ϵ^2 .

Teorema 3.3.1. *Para qualquer função de α_t , digamos $\phi_t = f(\alpha_t)$, que verifique a condição*

$$E[\phi_t] = E[\alpha_t], \quad \forall t \in \mathbb{Z},$$

e, se para além disso, tiver $\sigma_\epsilon^2 > 0$, onde $\epsilon_t = \phi_t Y_{t-1} - \alpha_t \circ^G Y_{t-1} + Z_t$, e se $E[\phi_t^2] < 1, \forall t \in \mathbb{Z}$, então o processo RCINAR(1) generalizado admite como única solução estacionária e causal

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad t \in \mathbb{Z},$$

onde

$$\beta_{t,i} = \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j}, & i \geq 1 \\ 1, & i = 0. \end{cases}$$

Quando $\phi_t = \alpha_t$ temos a representação apresentada em (3.3.1). Outra solução estacionária e causal surge quando $\phi_t = E[\alpha_t]$. Neste caso, a solução causal escreve-se como

$$Y_t = \mu_Y + \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i}, \quad t \in \mathbb{Z},$$

se $E[\alpha_t] < 1$, onde $\xi_t \sim WN(0, \sigma_\xi^2)$, com

$$\sigma_\xi^2 = E[\delta_t] \mu_Y + Var[\alpha_t] E[Y_t^2] + \sigma_Z^2. \quad (3.3.3)$$

O processo RCINAR(1) generalizado é, portanto, um processo AR(1) de coeficiente $E[\alpha_t]$ e com inovações ruído branco^[2].

Assim, de entre todas as representações causais do processo RCINAR(1) generalizado, e com base apenas na informação fornecida pelo processo $\{Y_t\}_{t \in \mathbb{Z}}$, a representação com $\phi_t = E[\alpha_t]$ é a única cujas inovações correspondem aos erros de predição linear. Este facto permite o estudo do processo RCINAR(1) generalizado com base apenas na informação fornecida por Y_t .

Fazendo agora um paralelismo entre os processos RCA(1) e RCINAR(1) generalizados, pode afirmar-se que ambos são estacionários de 2^a ordem se $E[\alpha_t^2] < 1$ e são processos AR(1) de coeficiente $E[\alpha_t]$ e inovações, $\{\xi_t\}_{t \in \mathbb{Z}}$, ruído branco de média nula e variância σ_ξ^2 , finita. Para o processo RCA(1) tem-se $\sigma_\xi^2 = Var[\alpha_t]E[Y_t^2] + \sigma_\epsilon^2$, enquanto que para o processo RCINAR(1) generalizado, se tem $\sigma_\xi^2 = Var[\alpha_t]E[Y_t^2] + E[\delta_t]\mu_Y + \sigma_\epsilon^2$.

Outro resultado importante refere-se aos processos RCINAR(1) generalizados, nos quais $\{\alpha_t\}_{t \in \mathbb{Z}}$ representa uma sucessão de v.a.'s i.i.d. com suporte em $(0, 1)$. Para estes casos, $\alpha_t \in (0, 1)$ é uma condição suficiente para que os processos RCINAR(1) generalizados sejam estacionários de 2^a ordem. Como caso particular de processos RCINAR(1) generalizados com esta característica salienta-se o caso do processo RCINAR(1) — processo cuja operação *thinning* generalizada é especificada através da distribuição Binomial.

3.4 Momentos e parâmetros

Em geral, os momentos até à 2^a ordem podem obter-se de diversas formas, nomeadamente através da solução causal. A expressão seguinte permite, a partir desta, obter de um modo bastante fácil os parâmetros do processo

^[2]Este resultado surge de imediato a partir do exercício 3.19 de Brockwell e Davis (1991), p. 112. Considere-se um processo AR(1), estacionário e de coeficiente $E[\alpha_t]$ e um processo RCINAR(1) generalizado, igualmente estacionário. Sem perda de generalidade, ambos os processos podem considerar-se como estacionários de média nula e com a mesma estrutura de covariância. Então, pelo referido exercício, o processo RCINAR(1) generalizado é também um processo AR(1).

RCINAR(1) generalizado:

$$Y_{t+k} = \sum_{i=0}^{k-1} \beta_{t+k,i} (\mu_Z + \xi_{t+k-i}) + \beta_{t+k,k} Y_t, \quad k \geq 1.$$

Desta expressão e a partir das propriedades de $\{\xi_t\}_{t \in \mathbb{Z}}$ obtêm-se, respectivamente, as seguintes expressões para a média, variância, autocovariância e autocorrelação do processo:

$$E[Y_t] = \frac{\mu_Z}{1 - E[\alpha_t]}, \quad (3.4.1)$$

$$\begin{aligned} \text{Var}[Y_t] &= \frac{E[\delta_t] \mu_Y + \text{Var}[\alpha_t] \mu_Y^2 + \sigma_Z^2}{1 - (\text{Var}[\alpha_t] + E^2[\alpha_t])} = \\ &= \frac{\sigma_\epsilon^2}{1 - (\text{Var}[\alpha_t] + E^2[\alpha_t])} + \text{Var}[\alpha_t] \frac{\mu_Y^2}{1 - (\text{Var}[\alpha_t] + E^2[\alpha_t])}, \end{aligned} \quad (3.4.2)$$

$$\gamma_Y(k) = (E[\alpha_t])^k \text{Var}[Y_t], \quad k \geq 0, \quad (3.4.3)$$

e

$$\rho_Y(k) = (E[\alpha_t])^k, \quad k \geq 0. \quad (3.4.4)$$

Na proposição seguinte são particularizados estes resultados para o caso do processo RCINAR(1).

Proposição 3.4.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1), estacionário de 2ª ordem e $Z_t \sim \text{Po}(\lambda)$. Então, prova-se que*

$$E[Y_t] = \frac{\lambda}{1 - E[\alpha_t]}$$

e

$$\begin{aligned} \text{Var}[Y_t] &= \frac{\lambda}{1 - E[\alpha_t]} + \lambda^2 \frac{\text{Var}[\alpha_t]}{(1 - E[\alpha_t^2])(1 - E[\alpha_t])^2} \\ &= E[Y_t] + (E[Y_t])^2 \frac{\text{Var}[\alpha_t]}{1 - E[\alpha_t^2]}. \end{aligned}$$

Demonstração. A demonstração é imediata bastando para tal substituir nas equações (3.4.1) e (3.4.2) μ_Z e $Var[Z_t]$ por λ e δ_t por $\alpha_t(1 - \alpha_t)$. \square

Através da proposição 3.4.1 verifica-se, contrariamente ao processo INAR(1) em que $\mu_Y = \sigma_Y^2$, que o facto de as inovações serem distribuídas de acordo com um modelo Poisson não implica que a distribuição marginal de Y_t seja Poisson, pois $\sigma_Y^2 > \mu_Y$.

Observação 3.4.1. Tal como para os processos DSINAR(1) generalizados, também para os processos RCINAR(1) generalizados é possível uma representação em espaço de estados. Seja, então, $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1) generalizado, estacionário de 2^a ordem e escrito na forma

$$Y_t - \mu_Y = E[\alpha_t]Y_{t-1} + \xi_t, \quad t \in \mathbb{Z},$$

tal que $\xi_t \sim WN(0, \sigma_\xi^2)$. Definindo a equação de estado como

$$S_{t+1} = E[\alpha_t]S_t + V_t, \quad t \in \mathbb{Z},$$

onde

$$S_1 = Y_1 - \mu_Y = \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i}$$

e $V_t = \xi_{t+1}$, verifica-se, facilmente, que o processo $\{Y_t\}_{t \in \mathbb{Z}}$ satisfaz a equação de observação $Y_t = S_t$, onde $g_t = 1$ e $W_t = \mu_Y$, $\forall t \in \mathbb{Z}$.

3.5 Média e funções de autocovariância e autocorrelação amostrais

Como vimos na secção anterior, o processo RCINAR(1) generalizado, estacionário de 2^a ordem, é um processo AR(1) de coeficiente $E[\alpha_t]$ e, como tal, pode ser representado na seguinte forma

$$Y_t = \mu_Y + \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i},$$

onde

$$E[\xi_t] = 0, \quad E[\xi_t \xi_s] = 0, \quad t \neq s,$$

e ξ_t são os erros de predição linear, cada um com variância

$$\sigma_\xi^2 = E[\delta_t] \mu_Y + \text{Var}[\alpha_t] E[Y_t^2] + \sigma_Z^2.$$

Seguidamente vamos utilizar este resultado para estudar o comportamento assintótico da média e das funções de autocovariância e autocorrelação amostrais do processo.

3.5.1 Comportamento assintótico da média amostral do processo

Proposição 3.5.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ o processo RCINAR(1) generalizado, fracamente estacionário, de média μ_Y e função de autocovariância $\gamma_Y(k) = (E[\alpha_t])^k \text{Var}[Y_t]$. A média amostral*

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$$

converge q.c. para μ_Y .

Demonstração. Como

$$\sum_{k=-\infty}^{\infty} |\gamma_Y(k)| = \sum_{k=-\infty}^{\infty} |(E[\alpha_t])^k \text{Var}[Y_t]| = \text{Var}[Y_t] \sum_{k=-\infty}^{\infty} |(E[\alpha_t])^k| < \infty,$$

conclui-se que a função de covariância é absolutamente somável, o que implica que a função de densidade espectral seja contínua. Como $\{Y_t\}_{t \in \mathbb{Z}}$ é um processo estacionário com espectro absolutamente contínuo tem-se, pelo teorema da ergodicidade, que $\bar{Y} \xrightarrow{\text{q.c.}} \mu_Y$ quando $n \rightarrow \infty$ (teorema 6.6 de Hall e Heyde (1980), p.184). \square

Proposição 3.5.2. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ o processo RCINAR(1) generalizado, fracamente estacionário com função de autocovariância $\gamma_Y(k)$ absolutamente somável. Então, o processo $\{Y_t\}_{t \in \mathbb{Z}}$ é ergódico para a média e tem função densidade espectral, $f(\theta)$, absolutamente contínua dada por*

$$f(\theta) = \frac{\sigma_\xi^2}{2\pi(1 - E[\alpha_t]e^{i\theta})(1 - E[\alpha_t]e^{-i\theta})}, \quad -\pi \leq \theta \leq \pi. \quad (3.5.1)$$

Demonstração. Pelo corolário 6.1.1.2 de Fuller (1976), p. 232, prova-se que se um processo estacionário tem função de autocovariância absolutamente somável então ele é ergódico para a média. Pela definição de espectro tem-se que

$$f(\theta) = \frac{\sigma_Y^2}{2\pi} \sum_{k=-\infty}^{\infty} (E[\alpha_t])^k e^{-i\theta k}.$$

Após alguns cálculos muito simples prova-se (3.5.1). □

Note-se que, pelo facto de se ter $E[\alpha_t] > 0$, serão sempre as baixas frequências que mais contribuirão para a variância do processo.

Vamos agora estabelecer a convergência em distribuição da média amostral, recorrendo ao corolário 5.2 de Hall e Heyde (1980), p. 135, que foi estabelecido para inovações independentes e identicamente distribuídas de média nula e variância finita. Aliás, é sobre esta condição que a teoria clássica de estimação de modelos auto-regressivos está assente. Contudo, Hall e Heyde demonstraram que a teoria clássica de estimação (e em particular o corolário 5.2) continua a ser aplicável, substituindo-se esta condição pela seguinte condição mais fraca

$$E[\xi_t | \mathcal{A}_{t-1}^\xi] = 0 \quad q.c. \quad \forall t, \quad (3.5.2)$$

onde $\mathcal{A}_{t-1}^\xi = \sigma(\xi_{t-1}, \xi_{t-2}, \dots)$ é a álgebra- σ gerada por ξ_{t-k} , $k \geq 1$.

Como os $\{\xi_t\}_{t \in \mathbb{Z}}$ são os erros de predição linear, o melhor preditor linear é o melhor preditor (ambos no sentido de média quadrática) para o processo $\{Y_t\}_{t \in \mathbb{Z}}$. Este facto implica a condição (3.5.2). Ou seja, $\{\xi_t\}_{t \in \mathbb{Z}}$ é uma

seqüência de diferenças de martingalas^[3].

Proposição 3.5.3. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ o processo RCINAR(1) generalizado, fracamente estacionário e ergódico, com função densidade espectral, $f(\theta)$, dada por (3.5.1). Se $f(\theta)$ é contínua em $\theta = 0$ então $n^{-1/2} \sum_{t=1}^n (Y_t - \mu_Y)$ converge em distribuição para uma Normal de média nula e variância $2\pi f(0)$, ou seja*

$$\bar{Y} \xrightarrow{d} N(\mu_Y, n^{-1} 2\pi f(0))$$

onde

$$2\pi f(0) = \frac{\sigma_\xi^2}{(1 - E[\alpha_t])^2}.$$

Demonstração. A demonstração da proposição pode ver-se recorrendo ao corolário 5.2 de Hall e Heyde (1980), p. 135. Pela definição de densidade espectral, $2\pi f(0) = \sigma_Y^2 \sum_{k=-\infty}^{\infty} (E[\alpha_t])^k$. A partir da equação (3.5.1) prova-se que $2\pi f(0) = \frac{\sigma_\xi^2}{(1 - E[\alpha_t])^2}$. □

3.5.2 Comportamento assintótico das funções de autocovariância e autocorrelação amostrais do processo

Tal como na estimação da média do processo, também na estimação das funções de autocovariância e autocorrelação, a inferência clássica assume que as inovações são i.i.d. De modo a estudar as propriedades dos estimadores destas funções num contexto mais vasto, em que não se requiere a estacionariedade do processo das inovações, Hall e Heyde (1980) substituíram esta condição pelas seguintes condições:

^[3]Seja (Ω, \mathcal{A}, P) um espaço de probabilidade e $\{\mathcal{A}_t\}_{t \in \mathbb{Z}}$ uma seqüência crescente de álgebras- σ . A seqüência $\{\epsilon_t\}_{t \in \mathbb{Z}}$ definida em (Ω, \mathcal{A}, P) , com $E[\epsilon_t] < \infty, \forall t \in \mathbb{Z}$, diz-se uma seqüência de martingalas se $E[\epsilon_t | \mathcal{A}_s] = \epsilon_s, s \leq t (s, t \in \mathbb{Z})$, onde $\mathcal{A}_s = \sigma(\epsilon_s, \epsilon_{s-1}, \dots)$. Uma seqüência $\{\xi_t\}_{t \in \mathbb{Z}}$ (com $E[\xi_t] < \infty, \forall t \in \mathbb{Z}$) diz-se uma seqüência de diferenças de martingalas se $E[\xi_t | \mathcal{A}_s] = 0, s \leq t (s, t \in \mathbb{Z})$, onde $\mathcal{A}_s = \sigma(\xi_s, \xi_{s-1}, \dots)$.

(i)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n E[\xi_t^2 | \mathcal{A}_{t-1}^\xi] = \sigma_\xi^2 > 0 \quad q.c.$$

(ii) Existe uma variável aleatória V com $E[V^2] < \infty$ tal que

$$P[|\xi_t| > u] \leq cP[|V| > u]$$

para algum $0 < c < \infty$ e para todo o $t \geq 1$ e $u \geq 0$.

Para processos estacionários (i) é substituída pela seguinte condição

(iii)

$$E[\xi_t^2 | \mathcal{A}_{t-1}^\xi] = \sigma_\xi^2 > 0 \quad q.c.$$

e a condição (ii) torna-se redundante.

Pelo teorema 6.6 de Hall e Heyde (1980), p. 184, prova-se que para o processo RCINAR(1) generalizado, $\{Y_t\}_{t \in \mathbb{Z}}$, estacionário de 2^a ordem, para o qual as condições (3.5.2) e (i) se verificam, o estimador usual da função de autocovariância

$$\hat{\gamma}_Y(k) = \frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}), \quad 0 \leq k \leq n-1,$$

converge *q.c.* para a função de autocovariância do processo, $\gamma_Y(k)$, e, se $\gamma_Y(0) > 0$,

$$\hat{\rho}_Y(k) = \hat{\gamma}_Y(k) / \hat{\gamma}_Y(0) \xrightarrow{q.c.} \rho_Y(k),$$

com $\rho_Y(k)$ a função de autocorrelação do processo.

De modo a estabelecer-se o teorema do limite central para a função de autocorrelação do processo, os referidos autores introduziram a seguinte condição:

(iv)

$$\sum_{i=1}^{\infty} i^{1/2} \{E[\alpha_i]\}^{2i} < \infty,$$

a qual é verificada pelo processo RCINAR(1) generalizado pois $E[\alpha_t^2] \in (0, 1)$. Além disso, provaram (teorema 6.7, p. 188) que se $\{Y_t\}_{t \in \mathbb{Z}}$ é um processo estacionário que satisfaz as condições (3.3.2), (iii) e (iv), então a distribuição conjunta de $\sqrt{n}(\hat{\rho}_Y(k) - \rho_Y(k))$, $1 \leq k \leq s$ converge para uma Normal multivariada de dimensão s com vector de médias nulo e matriz de covariâncias $\Sigma = [W_{ij}]$ onde

$$W_{ij} = \sum_{r=-\infty}^{\infty} \{\rho_Y(r)\rho_Y(r+i-j) + \rho_Y(r)\rho_Y(r+i+j) + 2\rho_Y(r)^2\rho_Y(i)\rho_Y(j) - 2\rho_Y(r)\rho_Y(i)\rho_Y(r+j) - 2\rho_Y(r)\rho_Y(j)\rho_Y(r+i)\}.$$

3.6 Estimação dos parâmetros do processo RCINAR(1) generalizado

3.6.1 Estimadores de Yule-Walker

Como se viu no sub-capítulo 3.3, o processo RCINAR(1) generalizado, de média nula e estacionário de 2^a ordem, é um processo AR(1) de coeficiente $E[\alpha_t]$ e, como tal, admite a seguinte representação causal

$$Y_t = \mu_Y + \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i}, \quad (3.6.1)$$

onde $\{\xi_t\}_{t \in \mathbb{Z}}$ ($\xi_t = \alpha_t \circ^G Y_{t-1} - E[\alpha_t]Y_{t-1} + Z_t - \mu_Z$) é um ruído branco de média nula e variância

$$\sigma_{\xi}^2 = E[\delta_t]\mu_Y + Var[\alpha_t]E[Y_t^2] + \sigma_Z^2. \quad (3.6.2)$$

Deste modo, podemos proceder à estimação de Yule-Walker usual (EYW) do vector de parâmetros $\bar{\theta}_1 = (E[\alpha_t], \mu_Z, \sigma_{\xi}^2)$. As propriedades assintóticas dos estimadores dos parâmetros serão facilmente obtidas pelo facto do processo ser estacionário.

Assim, dada uma amostra de dimensão n , (Y_1, \dots, Y_n) usa-se, em geral, como estimadores das funções de autocovariância e autocorrelação,

respectivamente $\gamma_Y(k)$ e $\rho_Y(k)$, as funções de autocovariância e autocorrelação amostrais definidas como

$$\hat{\gamma}_Y(k) = \frac{1}{n-k} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})$$

e

$$\hat{\rho}_Y(k) = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, \quad 0 \leq k \leq n-1.$$

Consequentemente, e tendo em consideração as equações (3.4.1) e (3.4.4) e também o facto da variância do processo $\{Y_t\}_{t \in \mathbb{Z}}$ se poder escrever na forma

$$Var[Y_t] = \frac{\sigma_\xi^2}{1 - (E[\alpha_t])^2}$$

tem-se que

$$\hat{E}[\alpha_t] = \hat{\rho}_Y(1),$$

$$\hat{\mu}_Z = (1 - \hat{\rho}_Y(1))\bar{Y},$$

onde \bar{Y} é a média amostral, e

$$\hat{\sigma}_Z^2 = \hat{\gamma}_Y(0) - \hat{\rho}_Y(1) \times \hat{\gamma}_Y(1).$$

Tal como vimos na secção 2.1, se a especificação da operação *thinning* for feita recorrendo às distribuições do tipo discreto usualmente usadas (por exemplo, Binomial, Poisson, Binomial Negativa e Geométrica), a estimação de $E[\delta_t]$ resume-se a estimar $E[\alpha_t]$ e/ou $E[\alpha_t^2]$. Nesta sequência, vamos considerar que $\circ^G \equiv *$ (os outros exemplos mencionados são muito análogos em termos de estimação). Assim, obtido $\hat{\sigma}_\xi^2$, o estimador de Yule-Walker da variância de Z_t escreve-se na forma

$$\hat{\sigma}_Z^2 = \hat{\sigma}_\xi^2 - \hat{\rho}_Y(1)(1 - \hat{\rho}_Y(1))\bar{Y} - \frac{\widehat{Var}[\alpha_t]}{n} \sum_{t=1}^n (Y_t^2 - Y_t). \quad (3.6.3)$$

Contudo, a partir das equações de Yule-Walker não é possível estimar $Var[\alpha_t]$ e, consequentemente, não é possível estimar σ_Z^2 .

3.6.2 Estimadores dos mínimos quadrados condicionais

Tal como foi referido na secção anterior, caso a especificação da operação *thinning* seja feita recorrendo às distribuições do tipo discreto usualmente usadas, a estimação de $E[\delta_t]$ resume-se a estimar $E[\alpha_t]$ e/ou $E[\alpha_t^2]$. Deste modo, vamos proceder à estimação dos mínimos quadrados condicionais (EMQC) dos parâmetros do processo RCINAR(1), ou seja, vamos considerar que $\circ^G \equiv *$.^[4]

Seja, então, $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1), estacionário de 2^a ordem^[5], cuja distribuição depende de um vector de parâmetros desconhecidos $\vec{\vartheta} = (\vec{\vartheta}_1, \vec{\vartheta}_2)$, sendo $\vec{\vartheta}_1 = (E[\alpha_t], \mu_Z)$ e $\vec{\vartheta}_2 = (Var[\alpha_t], \sigma_Z^2)$. $\vec{\vartheta}_1$ e $\vec{\vartheta}_2$, constituem, respectivamente, o vector das médias e das variâncias dos processos $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{Z_t\}_{t \in \mathbb{Z}}$.

Dada uma amostra de dimensão n , (Y_1, \dots, Y_n) , a estimação de $\vec{\vartheta}_1 = (E[\alpha_t], \mu_Z)$, através dos mínimos quadrados condicionais, consiste em minimizar a soma dos quadrados

$$Q_n(\vec{\vartheta}_1) = \sum_{t=2}^n (Y_t - E[Y_t | \mathcal{A}_{t-1}^Y])^2,$$

em ordem a $\vec{\vartheta}_1$, em que $\mathcal{A}_{t-1}^Y = \sigma(Y_{t-1}, Y_{t-2}, \dots)$ é a álgebra- σ gerada por $\{Y_{t-k}, k \geq 1\}$. Por outras palavras, consiste em minimizar a soma

$$Q_n(E[\alpha_t], \mu_Z) = \sum_{t=2}^n (Y_t - E[\alpha_t]Y_{t-1} - \mu_Z)^2.$$

Desta minimização resultam, respectivamente, os seguintes estimadores para $E[\alpha_t]$ e μ_Z ,

$$\hat{E}[\alpha_t] = \frac{\sum_{t=2}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^n (Y_{t-1} - \bar{Y})^2}$$

$$\hat{\mu}_Z = \frac{1}{n-1} \sum_{t=2}^n Y_t - \hat{E}[\alpha_t] \left\{ \frac{1}{n-1} \sum_{t=2}^n Y_{t-1} \right\}.$$

^[4]Caso a estimação de $E[\delta_t]$ não possa ser feita a partir de $E[\alpha_t]$ e/ou de $E[\alpha_t^2]$, esta pode ser feita através da minimização de $Q_n(\vec{\vartheta}_2)$ (introduzida na página seguinte), como se percebe facilmente ao longo desta secção.

^[5]Tal como no método anterior, as propriedades assintóticas dos estimadores dos mínimos quadrados obtêm-se a partir do facto de o processo ser estacionário.

A partir da equação (3.6.1) e da estimação de $\vec{\vartheta}_1$ pode ainda estimar-se σ_ξ^2 , através do estimador usual da variância,

$$\hat{\sigma}_\xi^2 = \frac{1}{n-2} \sum_{t=2}^n (Y_t - \hat{E}[\alpha_t]Y_{t-1} - \hat{\mu}_Z)^2.$$

Seguidamente procederemos à estimação do vector $\vec{\vartheta}_2$. A partir da equação (3.1.2) do capítulo 3.1 tem-se que

$$Var[Y_t|Y_{t-1}] = E[\alpha_t](1 - E[\alpha_t])Y_{t-1} + Var[\alpha_t]Y_{t-1}(Y_{t-1} - 1) + \sigma_Z^2.$$

Seguindo a proposta de Hwang e Basawa (1996), conhecido $\vec{\vartheta}_1$, a EMQC de $\vec{\vartheta}_2 = (Var[\alpha_t], \sigma_Z^2)$ consiste em minimizar a soma de quadrados

$$\begin{aligned} Q_n(\vec{\vartheta}_2) &= \sum_{t=2}^n (R_t(\vec{\vartheta}_1)^2 - E[R_t(\vec{\vartheta}_1)^2|\mathcal{A}_{t-1}^Y])^2 = \\ &= \sum_{t=2}^n (R_t(\vec{\vartheta}_1)^2 - E[R_t(\vec{\vartheta}_1)^2|Y_{t-1}])^2, \end{aligned} \quad (3.6.4)$$

em ordem a $\vec{\vartheta}_2$, onde $R_t(\vec{\vartheta}_1) = Y_t - E[Y_t|\mathcal{A}_{t-1}^Y]$. Note-se que $Var[Y_t|Y_{t-1}] = E[R_t(\vec{\vartheta}_1)^2|Y_{t-1}]$.

Seja $D_t = Y_{t-1}^2 - Y_{t-1}$ e $\hat{R}_t(\vec{\vartheta}_1) = Y_t - \hat{E}[\alpha_t]Y_{t-1} - \hat{\mu}_Z$. Minimizando a equação (3.6.4) em ordem a $\vec{\vartheta}_2$ resultam, respectivamente, os seguintes estimadores para $Var[\alpha_t]$ e σ_Z^2 ,

$$\hat{V}ar[\alpha_t] = \frac{\sum_{t=2}^n (\hat{R}_t(\vec{\vartheta}_1))^2 D_t + \hat{E}[\alpha_t](1 - \hat{E}[\alpha_t])(\bar{Y} - 1) \sum_{t=2}^n D_t - \hat{\sigma}_\xi^2 \sum_{t=2}^n D_t^2}{\sum_{t=2}^n D_t^2 - (\sum_{t=2}^n D_t)^2}$$

e

$$\hat{\sigma}_Z^2 = \hat{\sigma}_\xi^2 - \hat{E}[\alpha_t](1 - \hat{E}[\alpha_t])\bar{Y} - \frac{\hat{V}ar[\alpha_t]}{n-1} \sum_{t=2}^n D_t,$$

onde \bar{Y} é o estimador usual da média.

Por último, obtido o estimador de $Var[\alpha_t]$, o estimador de σ_Z^2 resulta de imediato através da expressão de σ_ξ^2 , isto é, de

$$\sigma_\xi^2 = E[\alpha_t(1 - \alpha_t)]\mu_Y + Var[\alpha_t]E[Y_t^2] + \sigma_Z^2.$$

3.6.3 Propriedades assintóticas dos estimadores

Nesta secção iremos estudar as propriedades assintóticas dos estimadores dos parâmetros obtidos nas secções anteriores. Mais concretamente, e recorrendo às notações aí introduzidas, estudaremos as propriedades assintóticas dos estimadores de Yule-Walker e dos mínimos quadrados condicionais de $(E[\alpha_t], \mu_Z)$ e de $(Var[\alpha_t], \sigma_Z^2)$. Os resultados obtidos nesta secção têm como pressuposto o facto de o processo $\{Y_t\}_{t \in \mathbb{Z}}$ ser estacionário de 2ª ordem.

Proposição 3.6.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1) generalizado, estacionário de 2ª ordem. Os estimadores de Yule-Walker de $(E[\alpha_t], \mu_Z)$ e de σ_ξ^2 são fortemente consistentes.*

Demonstração. Como o processo RCINAR(1) generalizado é um processo AR(1) de coeficiente $E[\alpha_t]$ e inovações ruído branco de média nula e variância σ_ξ^2 , finita, prova-se (Hall e Heyde (1980), p. 194) que $\hat{\gamma}_Y(k) \xrightarrow{\text{q.c.}} \gamma_Y(k)$ e $\hat{\rho}_Y(k) \xrightarrow{\text{q.c.}} \rho_Y(k)$, para $k = 0, 1$, e, portanto, $\hat{E}[\alpha_t] \xrightarrow{\text{q.c.}} E[\alpha_t]$. Por outras palavras, o estimador de Yule-Walker de $E[\alpha_t]$ é fortemente consistente.

As equações de Yule-Walker permitem estabelecer as seguintes relações entre as funções de autocovariância e os parâmetros

$$\gamma_Y(0) = E[\alpha_t]\gamma_Y(1) + \sigma_\xi^2$$

e

$$\gamma_Y(1) = E[\alpha_t]\gamma_Y(0),$$

isto é, σ_ξ^2 é uma função linear de $\gamma_Y(0)$. Deste modo, o estimador de σ_ξ^2 pode escrever-se como $\hat{\sigma}_\xi^2 = \hat{\gamma}_Y(0)[1 - (\hat{E}[\alpha_t])^2]$.

Como $\hat{E}[\alpha_t] \xrightarrow{\text{q.c.}} E[\alpha_t]$ e $\hat{\gamma}_Y(k) \xrightarrow{\text{q.c.}} \gamma_Y(k)$, $k = 0, 1$, resulta que $\hat{\sigma}_\xi^2 \xrightarrow{\text{q.c.}} \sigma_\xi^2$. A mesma conclusão pode ser retirada para $\hat{\mu}_Z$, i.e., $\hat{\mu}_Z \xrightarrow{\text{q.c.}} \mu_Z$. \square

Consideremos agora propriedades assintóticas dos estimadores dos mínimos quadrados condicionais. Para tal, vamos estabelecer o seguinte teorema.

Teorema 3.6.1. *Seja $\{Y_t\}_{t \in \mathbb{Z}}$ um processo RCINAR(1), estacionário de 2ª ordem e ergódico, com parâmetros $\vec{\vartheta}_1 = (E[\alpha_t], \mu_Z)$ e $\vec{\vartheta}_2 = (Var[\alpha_t], \sigma_Z^2)$.*

(a) *Se $E[Y_t^4] < \infty$, o vector estimador dos mínimos quadrados condicionais de $\vec{\vartheta}_1$ é fortemente consistente e assintoticamente Normal, isto é,*

$$\sqrt{n} \begin{bmatrix} \hat{E}[\alpha_t] \\ \hat{\mu}_Z \end{bmatrix} \xrightarrow{d} N_2(\vec{0}, \Sigma_1)$$

em que $N_2(\vec{0}, \Sigma_1)$ denota a Normal Bidimensional com vector média nulo e matriz de covariâncias $\Sigma_1 = [\sigma_{ij}]_{i,j=1,2} = V^{-1}WV^{-1}$, onde

$$V = \begin{bmatrix} \sigma_Y^2 + \mu_Y^2 & \mu_Y \\ \mu_Y & 1 \end{bmatrix}$$

e

$$W = \begin{bmatrix} E[\xi_t^2 Y_{t-1}^2] & E[\xi_t^2 Y_{t-1}] \\ E[\xi_t^2 Y_{t-1}] & E[\xi_t^2] \end{bmatrix}.$$

(b) *Se, além disso, $E[Y_t^8] < \infty$, então o vector estimador dos mínimos quadrados condicionais de $\vec{\vartheta}_2$ é também fortemente consistente e assintoticamente Normal, ou, por outras palavras,*

$$\sqrt{n} \begin{bmatrix} \hat{V}\hat{a}r[\alpha_t] \\ \hat{\sigma}_Z^2 \end{bmatrix} \xrightarrow{d} N_2(\vec{0}, \Sigma_2)$$

onde $\Sigma_2 = [\sigma_{ij}]_{i,j=1,2} = L^{-1}\Sigma L^{-1}$, em que

$$L = \begin{bmatrix} E[Y_t^2 - Y_t] & \sigma_Y^2 + \mu_Y^2 - \mu_Y \\ \sigma_Y^2 + \mu_Y^2 - \mu_Y & 1 \end{bmatrix}$$

e

$$\Sigma = \begin{bmatrix} E[u_t^2(\vec{\vartheta}_1)(Y_t^2 - Y_t)^2] & E[u_t^2(\vec{\vartheta}_1)(Y_t^2 - Y_t)] \\ E[u_t^2(\vec{\vartheta}_1)(Y_t^2 - Y_t)] & E[u_t^2(\vec{\vartheta}_1)] \end{bmatrix},$$

com $u_t(\vec{\vartheta}_1) = R_t(\vec{\vartheta}_1)^2 - E[R_t(\vec{\vartheta}_1)^2|Y_{t-1}]$.

Demonstração. (a) Começemos por demonstrar a consistência forte do vector estimador dos mínimos quadrados condicionais de $\vec{\vartheta}_1$. Para tal, iremos recorrer à notação de Klimko e Nelson (1980), pp. 630-631.

Seja $\{Y_t, t = 1, 2, \dots\}$ um processo estocástico definido no espaço de probabilidade $(\Omega, \mathcal{A}, P_{\vec{\theta}})$, cuja distribuição depende de um vector (coluna) $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$, de parâmetros desconhecidos, com $\vec{\theta}$ pertencente a um dado conjunto aberto do espaço Euclideano. Designe-se o valor esperado e o valor esperado condicional, relativos à medida de probabilidade $P_{\vec{\theta}}$, por $E_{\vec{\theta}}(\cdot)$ e $E_{\vec{\theta}}(\cdot|\cdot)$, respectivamente. Denote-se o “verdadeiro” valor de $\vec{\theta}$ por $\vec{\theta}^o$. Sejam $\{\mathfrak{S}_t\}_{t=0}^{\infty}$ a álgebra- σ gerada por $\{Y_k, 1 \leq k \leq n\}$ e \mathfrak{S}_0 a álgebra- σ trivial de Ω (i.e., $\mathfrak{S}_0 = \{\emptyset, \Omega\}$). Assuma-se ainda que $E[Y_t] < \infty, t = 1, 2, \dots$

Consideremos agora as funções $g(\cdot, \cdot)$ e $Q_n(\cdot)$ definidas, respectivamente, do seguinte modo

$$g(\vec{\theta}, \mathfrak{S}_t) = E_{\vec{\theta}}(Y_{t+1}|\mathfrak{S}_t)$$

e

$$Q_n(\vec{\theta}) = \sum_{t=0}^{n-1} [Y_{t+1} - g(\vec{\theta}, \mathfrak{S}_t)]^2,$$

para um dado conjunto de observações $Y_t, t = 1, \dots, n$.

Como é sabido, os estimadores dos mínimos quadrados são solução da equação

$$\partial Q_n(\vec{\theta})/\partial \theta_i = 0, \quad i = 1, 2, \dots, p,$$

pelo que se assume que a função $g(\vec{\theta}, \mathfrak{S}_t)$ é duas vezes continuamente diferenciável em $\vec{\theta}$, numa vizinhança do ponto $\vec{\theta}^o$.

Assim sendo, $Q_n(\cdot)$ admite a expansão

$$\begin{aligned} Q_n(\vec{\theta}) &= Q_n(\vec{\theta}^o) + (\vec{\theta} - \vec{\theta}^o)' \partial Q_n(\vec{\theta}^o)/\partial \vec{\theta} + \frac{1}{2} (\vec{\theta} - \vec{\theta}^o)' \partial^2 Q_n(\vec{\theta}^o)/\partial \vec{\theta}^2 = \\ &= Q_n(\vec{\theta}^o) + (\vec{\theta} - \vec{\theta}^o)' \partial Q_n(\vec{\theta}^o)/\partial \vec{\theta} + \frac{1}{2} (\vec{\theta} - \vec{\theta}^o)' V_n(\vec{\theta} - \vec{\theta}^o) + \end{aligned}$$

$$+\frac{1}{2}(\vec{\theta} - \vec{\theta}^o)'T_n(\vec{\theta}^*)(\vec{\theta} - \vec{\theta}^o),$$

onde $\vec{\theta}^*$ é um ponto intermédio entre $\vec{\theta}$ e $\vec{\theta}^o$,

$$V_n = (\partial^2 Q_n(\vec{\theta}^o)/\partial\vec{\theta}_i\partial\vec{\theta}_j)_{p \times p}$$

e

$$T_n(\vec{\theta}^*) = (\partial^2 Q_n(\vec{\theta}^*)/\partial\vec{\theta}^2 - V_n)_{p \times p}.$$

Teorema 2.1 (e corolário 2.1) (de Klimko e Nelson) Admitindo que:

- (i) $\lim_{n \rightarrow \infty} \sup_{\delta \rightarrow 0} (|T_n(\vec{\theta}^*)|_{ij}/n\delta) < \infty$, q.c., $i \leq p, j \leq p$, com $\vec{\theta}^*$ numa vizinhança δ de $\vec{\theta}^o$,
- (ii) $(2n)^{-1}V_n$ converge q.c. para uma matriz de constantes, $V_{p \times p}$, simétrica e definida positiva,
- (iii) $n^{-1}\partial Q_n(\vec{\theta}^o)/\partial\theta_i \rightarrow 0$, q.c., $i \leq p$,

então, existe uma sequência de estimadores $\{\hat{\theta}_n\}$ ($\hat{\theta}_n = (\hat{\theta}_{n1}, \hat{\theta}_{n2}, \dots, \hat{\theta}_{np})$) tal que $\hat{\theta}_n \xrightarrow{\text{q.c.}} \vec{\theta}^o$ e, para $\epsilon > 0$, existe um acontecimento E, com $P(E) > 1 - \epsilon$, e um n_0 tal que, em E e para $n > n_0$, $\hat{\theta}_n$ satisfaz o sistema de equações $(\partial Q_n(\vec{\theta})/\partial\theta_i = 0, i = 1, 2, \dots, p)$ e Q_n admite um mínimo relativo em $\hat{\theta}_n$.

Vamos agora demonstrar (usando a notação apropriada) que para os processos RCINAR(1) estacionários a condição (iii) do teorema 2.1 se verifica. Designemos ainda as componentes do vector de parâmetros $\vec{\vartheta}_1$ por $\vec{\vartheta}_1 = (\vartheta_{11}, \vartheta_{12}) = (E[\alpha_t], \mu_Z)$. Então, a derivada em ordem a ϑ_{11} será

$$\frac{1}{n} \frac{\partial Q_n(\vec{\vartheta}_1)}{\partial \vartheta_{11}} \Big|_{\vec{\vartheta}_1 = \vec{\vartheta}_1^o} = \frac{1}{n} \{2(n-1)\mu_Z - 2(n-1)\bar{Y} + 2(n-1)E[\alpha_t]\bar{Y},$$

com $\vec{\vartheta}_1^o$ o vector dos “verdadeiros valores” de $\vec{\vartheta}_1$.

Quando $n \rightarrow \infty$ resulta que

$$\frac{1}{n} \frac{\partial Q_n(\vec{\vartheta}_1)}{\partial \vartheta_{11}} \Big|_{\vec{\vartheta}_1 = \vec{\vartheta}_1^o} \xrightarrow{\text{q.c.}} 0,$$

uma vez que, por um lado, sendo Y_t um processo estacionário, $\bar{Y} \xrightarrow{\text{q.c.}} \mu_Y$ e, por outro, $\mu_Z = \mu_Y(1 - E[\alpha_t])$.

Tal como a derivada em ordem a ϑ_{11} , também para a derivada em ordem a ϑ_{12} , convenientemente normalizada, se tem o mesmo resultado, ou seja,

$$\begin{aligned} \frac{1}{n} \frac{\partial Q_n(\vec{\vartheta}_1)}{\partial \vartheta_{12}} \Big|_{\vec{\vartheta}_1 = \vec{\vartheta}_1^o} &= 2E[\alpha_t] \frac{1}{n} \sum_{t=2}^n (Y_{t-1} - \bar{Y})^2 - \frac{2(n-1)}{n} \bar{Y}^2 (1 - E[\alpha_t]) - \\ &\quad - \frac{2}{n} \sum_{t=2}^n (Y_t - \bar{Y})(Y_{t-1} \bar{Y}) + \frac{2\mu_Z}{n} \sum_{t=2}^n Y_{t-1}. \end{aligned}$$

Como $E[\alpha_t] = \rho_Y(1)$ e o processo é estacionário, então

$$\frac{1}{n} \frac{\partial Q_n(\vec{\vartheta}_1)}{\partial \vartheta_{12}} \Big|_{\vec{\vartheta}_1 = \vec{\vartheta}_1^o} \xrightarrow{\text{q.c.}} 0.$$

Relativamente à condição (ii), Klimko e Nelson, p. 631, referem que esta pode ser verificada mostrando que

$$V'_n = \left(\frac{1}{n} \sum_{t=2}^n \frac{\partial^2 g(\vartheta_1^o, \mathfrak{S}_{t-1})}{\partial \vartheta_i \partial \vartheta_j} u_t(\vartheta_1^o) \right)_{p \times p} \xrightarrow{\text{q.c.}} 0,$$

onde $u_t(\vartheta_1^o) = Y_t - E[Y_t | \mathfrak{S}_{t-1}]$, e

$$V''_n = \left(\frac{1}{n} \sum_{t=2}^n \frac{\partial g(\vartheta_1^o, \mathfrak{S}_{t-1})}{\partial \vartheta_i} \times \frac{\partial g(\vartheta_1^o, \mathfrak{S}_{t-1})}{\partial \vartheta_j} \right)_{p \times p} \xrightarrow{\text{q.c.}} V_{p \times p},$$

com V matriz definida positiva. Então, e dado que $g(\vartheta_1^o, \mathfrak{S}_{t-1}) = E[\alpha_t] Y_{t-1} + \mu_Z$, tem-se que $V'_n = (0)_{2 \times 2}$ e

$$V''_n = \begin{bmatrix} n^{-1} \sum_{t=2}^n Y_{t-1}^2 & n^{-1} \sum_{t=2}^n Y_{t-1} \\ n^{-1} \sum_{t=2}^n Y_{t-1} & 1 \end{bmatrix}$$

converge q.c. para a matriz definida positiva

$$\begin{bmatrix} \sigma_Y^2 + \mu_Y^2 & \mu_Y \\ \mu_Y & 1 \end{bmatrix}.$$

A condição (i) retira-se facilmente. Pelo teorema 2.1 prova-se que para os processos RCINAR(1), estacionários de 2^a ordem e com $\{\alpha_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s i.i.d., os estimadores dos mínimos quadrados condicionais de $\vec{\vartheta}_1$ são fortemente consistentes.

(b) Seguidamente, vamos proceder à demonstração da normalidade assintótica dos estimadores dos mínimos quadrados condicionais. Para tal, vamos utilizar de novo a notação proposta por Klimko e Nelson (1980).

Seja, então, $\{Y_t\}_{t=0}^\infty$ como uma sequência de v.a.'s estacionárias, ergódicas e integráveis e $\mathfrak{S}_t = \sigma(Y_t, Y_{t-1}, \dots, Y_{t-m+1})$, $t = m-1, m, \dots$, com m inteiro positivo arbitrário. Defina-se

$$Q_n(\vec{\theta}) = \sum_{t=m}^n [Y_{t+1} - g(\vec{\theta}, \mathfrak{S}_t)]^2,$$

sendo $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$.

Teorema 3.2 (de Klimko e Nelson) Admitindo que:

$$E[Y_t | Y_{t-1}, \dots, Y_0] = E[Y_t | Y_{t-1}, \dots, Y_{t-m}], \quad q. c., \quad t \geq m$$

e

$$E[u_m(\vec{\theta}^o)^2 | \partial g(\vec{\theta}^o, \mathfrak{S}_{m-1}) / \partial \theta_i \times \partial g(\vec{\theta}^o, \mathfrak{S}_{m-1}) / \partial \theta_j] < \infty, \quad i, j \leq p,$$

a sequência $\{\hat{\theta}_n\}$ de estimadores converge em distribuição para uma v.a. com distribuição Normal Multivariada de valor médio $\vec{\theta}^o$ e matriz de covariâncias $\Sigma = V^{-1} W V^{-1}$, onde

$$W = \left(E[u_m(\vec{\theta}^o)^2 \partial g(\vec{\theta}^o, \mathfrak{S}_{m-1}) / \partial \theta_i \times \partial g(\vec{\theta}^o, \mathfrak{S}_{m-1}) / \partial \theta_j] \right)_{p \times p}.$$

Ou seja,

$$\sqrt{n}(\hat{\theta}_n - \vec{\theta}^o) \xrightarrow{d} N_M(\vec{0}, V^{-1} W V^{-1}),$$

onde $N_M(\vec{0}, \Sigma)$ denota a Normal Multivariada.

De modo a demonstrar a normalidade assintótica do estimador dos mínimos quadrados condicionais de $\vec{\vartheta}_1$, vamos averiguar se para um processo

RCINAR(1), estacionário de 2ª ordem, se verificam as condições do teorema anterior.

Sendo o processo RCINAR(1) Markoviano, retira-se de imediato a validade da primeira condição do teorema. No que diz respeito à segunda condição, vamos considerar que $\vec{\vartheta}_1 = (\vartheta_{11}, \vartheta_{12}) = (\mu_Z, E[\alpha_t])$ e demonstrar que

$$(i) E[\xi_t^2 \left(\frac{\partial g(\vec{\vartheta}_1, \mathfrak{S}_{m-1})}{\partial \vartheta_{11}} \right)^2 | \vec{\vartheta}_1 = \vartheta_1^o] = \sigma_\xi^2 < \infty,$$

$$(ii) E[\xi_t^2 \frac{\partial g(\vartheta_{11}, \mathfrak{S}_{m-1})}{\partial \vartheta_{11}} | \vartheta_{11} = \vartheta_1^o \times \frac{\partial g(\vartheta_{11}, \mathfrak{S}_{m-1})}{\partial \vartheta_{12}} | \vartheta_{12} = \vartheta_1^o] = E[\xi_t^2 Y_{t-1}] < \infty \text{ e}$$

$$(iii) E[\xi_t^2 \left(\frac{\partial g(\vartheta_{11}, \mathfrak{S}_{m-1})}{\partial \vartheta_{11}} \right)^2 | \vartheta_{11} = \vartheta_1^o] = E[\xi_t^2 Y_{t-1}^2] < \infty.$$

Começemos por demonstrar a condição (ii), visto a (i) ser imediata. Após alguns cálculos prova-se que $E[\xi_t^2 Y_{t-1}] = E[\alpha_t(1 - \alpha_t)]E[Y_t^2] + Var[\alpha_t]E[Y_t^3] + E[Y_t]\sigma_Z^2 < \infty$ desde que $E[Y_t^3] < \infty$. De igual modo, tem-se que $E[\xi_t^2 Y_{t-1}^2] = E[\alpha_t(1 - \alpha_t)]E[Y_t^3] + Var[\alpha_t]E[Y_t^4] + E[Y_t^2]\sigma_Z^2 < \infty$ é válida desde que $E[Y_t^4] < \infty$, provando-se assim a condição (iii).

Deste modo, prova-se que os estimadores dos mínimos quadrados condicionais dos parâmetros do modelo RCINAR(1) (estacionário de 2ª ordem e com $\{\alpha_t\}_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s i.i.d.) são fortemente consistentes e assintoticamente normais, desde que $E[Y_t^4] < \infty$.

Para demonstrar que $\vec{\vartheta}_2$ também é um estimador fortemente consistente e assintoticamente normal com vector média nulo e matriz de covariâncias $\Sigma_2 = [\sigma_{ij}] = L^{-1}\Sigma L^{-1}$, $i, j = 1, 2$, basta considerar

$$g(\vartheta_2^o, \mathfrak{S}_{m-1}) = E[R_t(\vec{\vartheta}_1)^2 | Y_{t-1}]$$

e

$$u_m(\vartheta_2^o) = R_t(\vec{\vartheta}_1)^2 - E[R_t(\vec{\vartheta}_1)^2 | Y_{t-1}].$$

□

Observação 3.6.1. De modo a aumentar a eficiência dos estimadores dos mínimos quadrados condicionais do vector de parâmetros $\vec{\vartheta}_1 = (E[\alpha_t], \mu_Z)$, Hwang e Basawa (1996) propõem a sua ponderação pela variância

condicional do processo no instante t , Y_t , dado o seu passado e assumindo que $\vec{\vartheta}_2$ é conhecido. Isto é, propõem a estimação de $\vec{\vartheta}_1$ através da minimização da soma

$$Qp_n(\vec{\vartheta}_1) = \sum_{t=2}^n \left[\frac{(Y_t - E[Y_t | \mathcal{A}_{t-1}^Y])^2}{\text{Var}[Y_t | \mathcal{A}_{t-1}^Y]} \right].$$

Demonstram ainda, sob algumas condições, que os estimadores dos mínimos quadrados condicionais ponderados são fortemente consistentes e assintoticamente normais.

3.7 Teste de ajustamento do modelo RCINAR(1) estacionário

Depois de identificado o modelo e tendo-se procedido à estimação dos respectivos parâmetros (através dos métodos expostos na secção 3.6), inicia-se a avaliação da qualidade do ajustamento, testando-se o comportamento dos respectivos resíduos. Os resíduos, ξ_t , são definidos por

$$\begin{aligned} \xi_t &= Y_t - E[Y_t | Y_{t-1}] \\ &= Y_t - (E[\alpha_t]Y_{t-1} + \mu_Z). \end{aligned}$$

Tal como já foi referido, aquando da avaliação da qualidade do ajustamento dos modelos DSINAR(1) generalizados, se a análise dos resíduos não levar à rejeição de que estes se comportam como um ruído branco, então pode concluir-se que o modelo estimado se ajusta bem à série temporal em causa.

Comecemos por definir o seguinte teste estatístico

$$H_0 : \mu_\xi = 0 \quad \text{vs} \quad H_1 : \mu_\xi \neq 0.$$

No caso em que $\{\xi_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d. é possível, recorrendo ao conhecido teorema de Lindeberg-Levy, ensaiar facilmente a hipótese H_0 . Hall e Heyde (1980, p. 10) demonstram, sob algumas condições, que também se pode enunciar o teorema do limite central para

sucessões de diferenças de martingalas (e, portanto, para sucessões de v.a.'s não-independentes, verificando determinadas características).

Mais concretamente, seja $\{\xi_t\}_{t=1,\dots,n}$ uma sucessão de diferenças de martingalas, tal que $S_n^2 = E[V_n^2]$ e

$$V_n^2 = \sum_{i=1}^n E[\xi_i^2 | \mathcal{A}_{i-1}^\xi].$$

Se

- (1) $S_n^{-2} V_n^2 \xrightarrow{P} 1$ e
- (2) $S_n^{-2} \sum_{i=1}^n E[\xi_i^2 I(|\xi_i| \geq \delta S_n)] \rightarrow 0$ ($I(\cdot)$ designa a função indicatriz), quando $n \rightarrow \infty$, para todo o $\delta > 0$,

então

$$S_n^{-1} \sum_{i=1}^n \xi_i \xrightarrow{d} N(0, 1).$$

Sendo $\{\xi_t\}_{t \in \mathbb{Z}}$ um processo estacionário de 2^a ordem, a condição (2) encontra-se desde logo satisfeita. Tendo ainda em conta esta característica prova-se, recorrendo ao teorema 2.19 de Hall e Heyde (1980), p. 36, a validade da condição (1).

Um teste de diagnóstico usual, de avaliação da qualidade do ajustamento do modelo estimado, consiste em construir intervalos de confiança, por exemplo, para as funções de autocovariância dos resíduos, $\gamma_\xi(m)$, e verificar se, para os lags $m = 1, \dots, k$, $\hat{\gamma}_\xi(m)$ pertencem a estes intervalos.

Defindo-se $S_j = \xi_j \xi_{j-m}$, tem-se que $\{\sum_{j=1}^n \xi_j \xi_{j-m}, n \geq 1\}$ é uma sucessão de diferenças de martingalas. Então, sob as condições (1) e (2) anteriores prova-se, através do corolário 3.1 de Hall e Heyde (1980), p. 58, que

$$n^{-1/2} \sum_{i=1}^n \xi_i \xi_{i-k} \xrightarrow{d} N(0, \sigma_\xi^2), \quad \forall k > 0.$$

Outra técnica de diagnóstico consiste em testar a hipótese conjunta de que as primeiras m funções de autocorrelação dos resíduos são nulas. Ou seja, em definir o seguinte teste estatístico

$$H_0 : \rho_\xi(1) = \rho_\xi(2) = \dots = \rho_\xi(m) = 0 \quad vs \quad H_1 : \exists_{k, k=1, \dots, m} : \rho_\xi(k) \neq 0.$$

De modo a testar a hipótese H_0 , Box e Pierce propõem, em 1970, a seguinte estatística de teste (modificada posteriormente por Ljung e Box, em 1978)

$$Q_\xi = n \sum_{i=1}^m [\hat{\rho}_\xi(i)]^2.$$

Sob a hipótese de $\{\xi_t\}_{t \in \mathbb{Z}}$ constituir uma sucessão de v.a.'s i.i.d., a matriz de covariâncias assintótica do vector estimador das m autocorrelações é igual ao inverso da dimensão da amostra, a multiplicar pela matriz identidade, provando-se, portanto, que Q_ξ segue, aproximadamente, uma distribuição Qui-Quadrado com m graus de liberdade.

Deixa-se em aberto, para trabalho futuro, a aplicação deste teste de hipóteses ao caso em que $\{\xi_t\}_{t \in \mathbb{Z}}$ constitui, não uma sucessão de v.a.'s i.i.d., mas sim uma sucessão de diferenças de martingalas.

Tal como vimos na secção 2.9, após a verificação da validade do modelo, podemos ainda quantificar a ajustabilidade do modelo estimado através do coeficiente de ajustabilidade, p_{E_t} (ver equação (2.9.1)).

3.8 Previsão de p.e.'s estacionários a partir do modelo RCINAR(1)

Considerere-se, sem perda de generalidade, o processo $\{Y_t\}_{t \in \mathbb{Z}}$, estacionário de 2ª ordem e de média nula, representado na sua forma causal

$$Y_t = \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i}, \quad -\infty < t < \infty. \quad (3.8.1)$$

Como o processo RCINAR(1) generalizado pode ser escrito como um processo AR(1) de coeficiente $E[\alpha_t]$, a previsão a k passos faz-se de modo usual, isto é, através do valor esperado condicional.

O “melhor” preditor, *i.e.*, o preditor com menor erro quadrático médio, de Y_{t+1} (representado por $Y_t(1)$) é obtido do seguinte modo (Priestley (1981), p. 76)

$$Y_t(1) = E[Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1]$$

$$\begin{aligned}
&= E[E[\alpha_{t+1}]Y_t + \mu_Z + \epsilon_{t+1} | Y_t, Y_{t-1}, \dots, Y_1] \\
&= E[\alpha_{t+1}]Y_t + \mu_Z.
\end{aligned}$$

De modo análogo obtém-se o preditor de Y_{t+2}

$$\begin{aligned}
Y_t(2) &= E[Y_{t+2} | Y_t, Y_{t-1}, \dots, Y_1] \\
&= E[\alpha_{t+2}]E[Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1] + \mu_Z \\
&= E[\alpha_{t+2}]E[\alpha_{t+1}]Y_t + \mu_Z(1 + E[\alpha_{t+2}]),
\end{aligned}$$

ou seja

$$Y_t(2) = (E[\alpha_t])^2 Y_t + \mu_Z(1 + E[\alpha_t]).$$

Pelo que, para $k > 0$

$$Y_t(k) = (E[\alpha_t])^k Y_t + \mu_Z \sum_{j=0}^{k-1} (E[\alpha_t])^j.$$

Fazendo $k \rightarrow \infty$ verifica-se que

$$Y_t(k) \rightarrow \frac{\mu_Z}{1 - E[\alpha_t]}.$$

Por outras palavras, quando o horizonte da previsão aumenta indefinidamente, as previsões tendem, como seria de esperar para a média do processo Y_t .

Vamos agora proceder ao cálculo do erro de previsão. Seja $e(k) = Y_{t+k} - Y_t(k)$ o erro de previsão a k passos. Então, após cálculos simples, resulta

$$e(k) = \sum_{i=0}^{k-1} (E[\alpha_t])^i \xi_{t+k-i}$$

onde

$$E[e(k)] = 0$$

e

$$\text{Var}[e(k)] = \sigma_{\xi}^2 \sum_{i=0}^{k-1} (E[\alpha_t])^{2i}.$$

Fazendo igualmente $k \rightarrow \infty$ verifica-se que

$$\text{Var}[e(k)] \rightarrow \frac{\sigma_{\xi}^2}{1 - (E[\alpha_t])^2},$$

isto é, o preditor $Y_t(k)$ de Y_{t+k} é centrado e a variância tende, quando $k \rightarrow \infty$, para a variância do processo Y_t .

Capítulo 4

Processo DSINAR(1) generalizado: um estudo por simulação

De modo a estudar-se o desempenho do método de máxima verosimilhança, na estimação dos parâmetros do processo DSINAR(1) generalizado, bem como a sua adequação (através do estudo do ajustamento do modelo), simularam-se, a partir de um programa elaborado em linguagem FORTRAN, realizações do processo

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t = 1, \dots, N,$$

onde \circ^G representa o operador *thinning* generalizado (tal como definido no capítulo 2), $\alpha_t = f(X_t)$ é uma função de um p.e. $\{X_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{R} e $\{Z_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d. inteiras não-negativas com distribuição Poisson de parâmetro λ . Neste estudo optou-se apenas pela escolha de ruídos poissonianos, por uma racionalização do estudo e porque esta escolha permite definir um modelo condicionalmente Poisson para séries temporais de contagem.

Para este estudo consideraram-se os casos em que $\{X_t\}_{t \in \mathbb{Z}}$ é estacionário e não-estacionário. Para estes dois casos escolheram-se apenas duas distribuições na especificação da operação *thinning*: a distribuição Binomial e a distribuição Poisson. Para além da frequência com que estas duas distribuições discretas surgem em exemplos e/ou aplicações práticas, a escolha da distribuição Binomial ficou também a dever-se à própria definição

original do operador *thinning*. A escolha da distribuição Poisson (uma das distribuições discretas mais usuais), conjuntamente com a escolha de ruídos poissonianos, permite definir um modelo condicionalmente Poisson para séries temporais de contagem. Mais concretamente, permite definir um modelo *observation driven*, condicionalmente Poisson, bastante análogo à maior parte dos modelos encontrados na literatura (ver capítulo 1).

Para tal, foram simuladas realizações dos seguintes processos:

- (1) $\{X_t\}_{t \in \mathbb{Z}}$ um processo AR(1) gaussiano de média μ_X e de variância σ_X^2 , da forma

$$X_t - \mu_X = 0.6X_{t-1} + \epsilon_t$$

e

- (2) $\{X_t\}_{t \in \mathbb{Z}}$ um processo gaussiano, em que a média e a variância variam ao longo do tempo do seguinte modo:

$$\mu_t = \begin{cases} \mu_{t-1} + 0.04, & t \geq 1 \\ 2, & t = 0 \end{cases}$$

e

$$\sigma_t^2 = \begin{cases} \sigma_{t-1}^2 + 0.01, & t \geq 1 \\ 2, & t = 0. \end{cases}$$

Como valor inicial, necessário à simulação de $\{Y_t\}, t = 1, 2, \dots, N$, utilizou-se o valor médio do processo. Ou seja, de acordo com o explicitado na secção 2.4, considerou-se

$$Y_1 = \lambda \sum_{i=0}^K \bar{\beta}_{t,i},$$

onde $\bar{\beta}_{t,i} = \frac{1}{N-i+1} \sum_{t=i}^N \prod_{j=0}^{i-1} \alpha_{t-j}$, $i \geq 1$ (com $\bar{\beta}_{t,0} = 1$). Tal como foi referido na sub-secção 2.8.1, verificou-se, a partir dos estudos realizados que escolhendo-se um $K = 5$, $\hat{\lambda} = \bar{Y} / \sum_{i=0}^K \bar{\beta}_{t,i}$ permite obter estimativas do parâmetro λ muito próximas do verdadeiro valor e, por conseguinte, estimativas do parâmetro μ_Y também muito próximas do verdadeiro valor.

As simulações do processo foram obtidas para vários tamanhos de amostras, nomeadamente amostras de dimensão $N = 50, 100, 250, 500, 1000$.

Foi ainda considerado um período de aquecimento (*burn-in period*) de comprimento 500 na simulação das realizações do processo.

Para amostras de dimensão $N = 1000$ estudou-se, por simulação, a aplicação dos processos DSINAR(1) generalizado na modelação de séries de valores inteiros e não-negativos, usando como covariável um p.e. bivariado, $\{\vec{X}_t\}_{t \in \mathbb{Z}}$, com suporte em \mathbb{R}^2 . Como as conclusões a que chegámos não se alteram, quer usemos uma ou mais covariáveis, apresentamos apenas o estudo para o caso univariado.

4.1 O operador Binomial *thinning*

Tal como foi adoptado ao longo deste trabalho, vamos escrever o processo DSINAR(1) (que é o processo DSINAR(1) generalizado com a distribuição Binomial na especificação da operação *thinning*) da seguinte forma

$$Y_t = \alpha_t * Y_{t-1} + Z_t, \quad t = 1, \dots, N,$$

com $\alpha_t = \frac{\exp(\omega X_t)}{1 + \exp(\omega X_t)}$ e $\omega \in \mathbb{R}$.

A operação *thinning* define-se neste contexto do seguinte modo:

$$\alpha_t * Y_{t-1} | \alpha_t, Y_{t-1} \sim B(\mu_t, \sigma_t^2),$$

onde $B(\mu_t, \sigma_t^2)$ é a distribuição Binomial de valor esperado $\mu_t = \alpha_t Y_{t-1}$ e variância $\sigma_t^2 = \alpha_t(1 - \alpha_t)Y_{t-1}$.

4.1.1 Série explicativa estacionária

Comportamento assintótico dos estimadores dos parâmetros do modelo

De modo a estudar-se o comportamento assintótico dos estimadores de máxima verosimilhança condicional dos parâmetros do modelo, simularam-se realizações do processo $\{X_t\}_{t \in \mathbb{Z}}$, com $\mu_X = 4$ e $\sigma_X^2 = 2$ e, a partir destas, realizações do processo $\{Y_t\}_{t \in \mathbb{Z}}$ com $\omega = 0.2$ e $\lambda = 2$.

Outros valores de ω e λ foram utilizados, quer no presente estudo, quer no estudo que se apresenta na secção 4.2. Para estas escolhas, apenas se teve em conta o facto de se pretender simular séries temporais de valores inteiros, cujos valores das observações fossem relativamente reduzidos.

Após a simulação das várias realizações do processo procedeu-se à estimação dos parâmetros do modelo, (ω, λ) , através da estimação de máxima verosimilhança condicional (EMVC) e recorrendo ao método iterativo de Newton-Raphson, com o seguinte erro de paragem: $|\omega_i - \omega_{i-1}| \leq 10^{-5}$ e $|\lambda_i - \lambda_{i-1}| \leq 10^{-5}$ — onde o índice i designa a iteração. Como valores iniciais para o método iterativo, considerou-se $\omega^{(0)} = 0$ e $\lambda^{(0)} = \hat{\lambda}_{mq}$, onde $\hat{\lambda}_{mq}$ designa o estimador dos mínimos quadrados condicionais, quando se admite que o processo $\{X_t\}_{t \in \mathbb{Z}}$ é i.i.d.^[1] No apêndice B apresentam-se as expressões das equações de verosimilhança e das suas derivadas, em ordem aos parâmetros do modelo.

As figuras 4.1.1-4.1.4 que se apresentam de seguida dizem respeito a uma realização do processo DSINAR(1), para uma amostra de dimensão $N = 1000$. As funções de autocorrelação e autocorrelação parcial amostrais designam-se abreviadamente por FAC e FACP, respectivamente. Observando a figura 4.1.1 não se verificam quaisquer tendências, *i.e.*, nada parece contrariar a hipótese de estacionariedade da série. Podemos ainda constatar que as FAC e FACP amostrais são, para qualquer *lag*, bastante coincidentes com as FAC e FACP teóricas de um processo AR(1).

Seguidamente, na tabela 4.1.1 são apresentadas as médias, as variâncias e os erros quadráticos médios (EQM) amostrais (obtidos em 100 séries simuladas) das estimativas dos parâmetros do modelo, obtidas através da EMVC, para as várias dimensões da amostra.

A partir da sua análise pode concluir-se que os estimadores de máxima verosimilhança condicional conduzem a *boas* estimativas para o vector parâmetro, (ω, λ) , mesmo para amostras de menor dimensão ($N = 50$). Como seria de esperar, os EQM amostrais aproximam-se de zero à medida que a dimensão da amostra aumenta.

Verificou-se ainda (tabela 4.1.2), recorrendo aos testes de Kolmogorov-Smirnov (K-S) e Shapiro-Wilk (S-W), que as estimativas do vector parâmetro se podem considerar como provenientes de uma v.a. com distribuição Normal de valor esperado igual ao respectivo parâmetro. Note-se que esta conclusão é válida para os níveis de significância usuais e

^[1]Só para dar uma ideia, se escolhermos quaisquer valores de ω e λ para as simulações e, posteriormente, procedermos à sua estimação, utilizando como valores iniciais ao método iterativo de Newton-Raphson as estimativas propostas, o número médio de iterações necessárias até se atingir o erro de paragem é de 7, contra cerca de 50, no caso de se adoptar $\lambda^{(0)} = 1$. Em qualquer dos casos, o erro quadrático médio das estimativas é muito idêntico, qualquer que seja a dimensão da amostra.

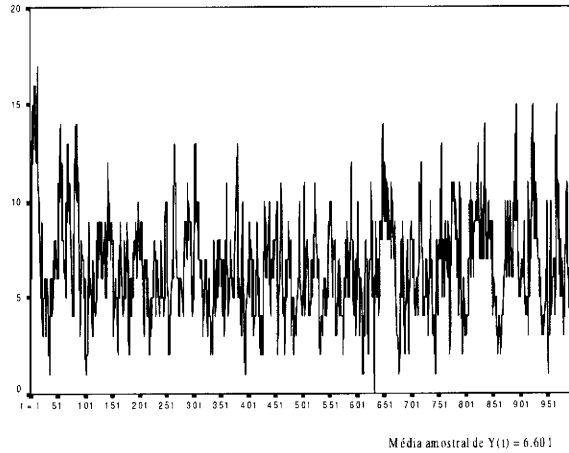


Figura 4.1.1: Realização, de dimensão $N = 1000$, do processo DSINAR(1): Binomial *thinning*.

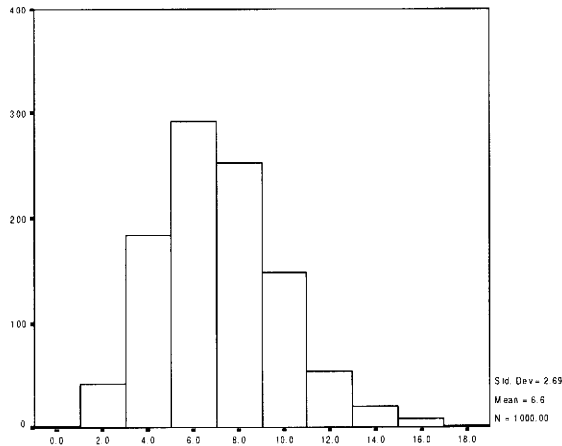


Figura 4.1.2: Histograma do processo DSINAR(1): Binomial *thinning*.

para amostras de dimensão reduzida ($N = 50$).

Embora a hipótese da normalidade assintótica dos estimadores de máxima verosimilhança dos parâmetros do modelo não seja de rejeitar, com base nos estudos por simulação, a obtenção dos respectivos intervalos de confiança não nos parece viável, dada a dificuldade em encontrar estimadores consistentes

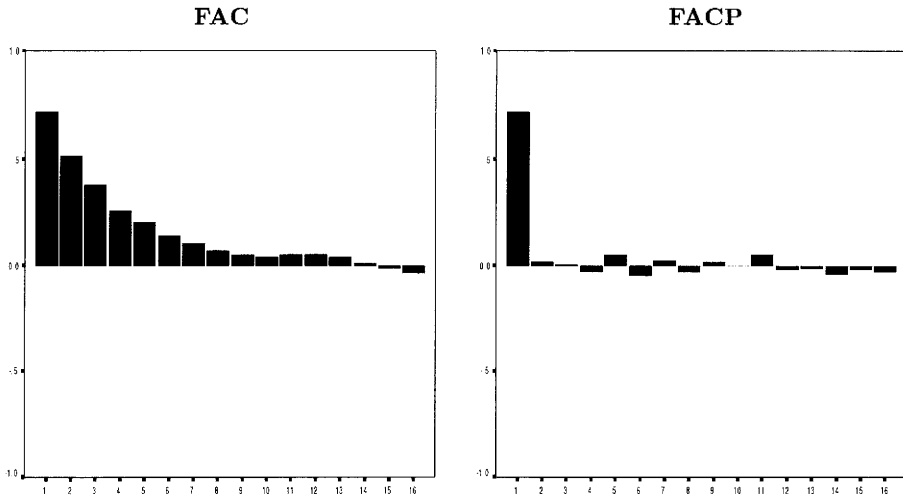


Figura 4.1.3: FAC e FACP amostrais do processo DSINAR(1): Binomial thinning.

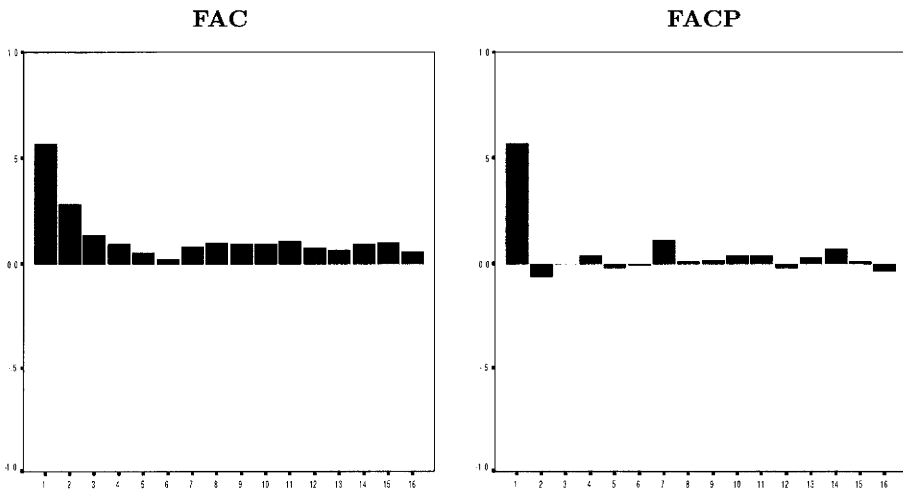


Figura 4.1.4: FAC e FACP amostrais do processo $\{X_t\}_{t \in \mathbb{Z}}$: Binomial thinning.

para a variância destes parâmetros. Assim, propõe-se o recurso aos

Tabela 4.1.1: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = 0.2$			$\lambda = 2$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
50	0.1836	0.003377	0.003613	2.1309	0.1343	0.1501
100	0.1812	0.001718	0.002072	2.1244	0.05989	0.07537
250	0.1962	0.0007873	0.0008017	1.9989	0.02758	0.02758
500	0.1995	0.0004160	0.0004163	1.9921	0.01309	0.01316
1000	0.1992	0.0002132	0.0002138	2.0067	0.009024	0.009069

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Binomial *thinning*.

Tabela 4.1.2: Testes de normalidade. *p-values* dos testes de K-S e de S-W: caso estacionário.

N	$\hat{\omega}$		$\hat{\lambda}$	
	K-S	S-W	K-S	S-W
50	> 0.200	0.508	> 0.200	0.741
100	> 0.200	0.811	0.190	0.522
250	> 0.200	0.544	> 0.200	0.072
500	> 0.200	0.595	> 0.200	0.972
1000	> 0.200	0.685	0.176	0.470

Nota: As estimativas de ω e de λ foram obtidas em 100 séries simuladas, utilizando-se o operador Binomial *thinning* para $\omega = 0.2$ e $\lambda = 2$.

métodos de reamostragem para séries temporais, propostos no capítulo 2.9, nomeadamente à reamostragem por blocos móveis e à reamostragem estacionária por blocos.

Considerando o segundo método de reamostragem, foram escolhidos vários valores de p , suficientemente pequenos, como parâmetros da distribuição Geométrica e estimadas as probabilidades de cobertura. Como critério para a escolha do valor de p mais adequado foi escolhido aquele que minimiza o erro da probabilidade de cobertura e o comprimento dos intervalos resultantes. De acordo com estes critérios, e para os valores dos parâmetros do modelo DSINAR(1) adoptados, obteve-se um valor de

$p = 0.02$, usando 500 repetições *bootstrap*.

Relativamente ao método de reamostragem por blocos móveis, de acordo com os mesmos critérios e igualmente para os valores dos parâmetros do modelo adoptados, foram testados diferentes tamanhos de bloco (o inverso do valor de p , devido à semelhança entre os dois métodos). Para cada um dos casos testados, as probabilidades de cobertura apresentam valores bastante afastados da probabilidade nominal. Isto deve-se ao facto de as estimativas de máxima verosimilhança condicional dos parâmetros do modelo, obtidas através destas amostras *bootstrap*, serem bastante próximas dos verdadeiros valores dos parâmetros. Daí, a maior parte dos intervalos de confiança não conterem os verdadeiros valores dos parâmetros.

Estes resultados, conjuntamente com os obtidos por reamostragem estacionária, são apresentados na tabela 4.1.3.

Nas figuras 4.1.5 apresentam-se os intervalos de confiança assintóticos a 95% para os parâmetros ω e λ , obtidos através da reamostragem estacionária por blocos com $p = 0.02$ e em 100 séries simuladas de dimensão $N = 1000$. Estas figuras contêm ainda os verdadeiros valores dos parâmetros (a vermelho) bem como as respectivas estimativas de máxima verosimilhança condicional (a tracejado).

Se, informalmente, pensarmos na amostra aleatória como sendo a população da qual se retiram, aleatoriamente, várias amostras (amostras *bootstrap*), então, deverá tomar-se como critério para a escolha do valor de p aquele que minimize o erro quadrático médio dos parâmetros ω e λ ; construindo-se, deste modo, outro critério aplicável ao caso usual em que se dispõe apenas de uma dada série temporal.

Assim, para os parâmetros $\omega = 0.2$, $\lambda = 2$ e $N = 1000$ obteve-se igualmente um valor de $p = 0.02$ como sendo aquele que minimiza os respectivos erros quadráticos médios.

Avaliação da qualidade do ajustamento do modelo

Depois de se ter procedido à estimação dos respectivos parâmetros inicia-se a avaliação da qualidade do ajustamento. Para tal, comecemos por definir o resíduo, ξ_t , no instante t , do seguinte modo:

$$\xi_t = Y_t - E[Y_t | Y_{t-1}, X_t] = Y_t - \alpha_t Y_{t-1} - \lambda.$$

Tabela 4.1.3: Probabilidades de cobertura (obtidas por reamostragem por blocos móveis (RBM) e por reamostragem estacionária por blocos (REB)) para realizações de dimensão $N = 1000$ do processo DSINAR(1).

RBM			REB		
b	ω	λ	p	ω	λ
6	0	0	0.167	0.01	0.01
16	0.02	0.01	0.0625	0.37	0.40
50	0.04	0.05	0.02	0.98	1.00
100	0.10	0.17	0.01	0.99	0.97
150	0.07	0.14	0.007	0.91	0.94

Nota: Estimativas obtidas em 500 amostras *bootstrap*.

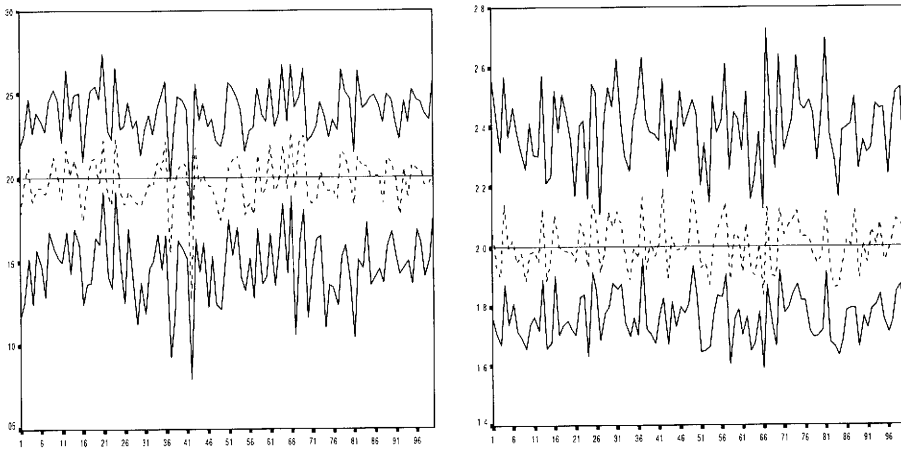


Figura 4.1.5: Estimativas de máxima verossimilhança condicional de ω (à esquerda, com o verdadeiro valor a vermelho), de λ (à direita, com o verdadeiro valor a vermelho) (ambas a tracejado) e intervalos de confiança *bootstrap* a 95% (a cheio), em 100 séries simuladas: Binomial *thinning*.

Testar que o modelo estimado se ajusta bem ao conjunto de dados, ou seja, que descreve bem o conjunto de dados em causa, consiste em testar o comportamento dos resíduos. De acordo com a proposição 2.3.1, se a análise dos resíduos permitir *concluir* que estes se comportam como um ruído branco de média nula, então pode concluir-se que o modelo estimado se ajusta bem

à série temporal em causa.

No caso do modelo DSINAR(1) ser apropriado para descrever a série temporal em estudo, o gráfico dos resíduos não deverá exibir valores discrepantes nem qualquer tendência (crescente ou decrescente no tempo), o histograma dos resíduos deverá apresentar uma média próxima de zero e as funções de autocorrelação e autocorrelação parcial amostrais desses mesmos resíduos deverão assemelhar-se às FAC e FACP de uma sucessão de v.a.'s não-correlacionadas.

Seguidamente apresenta-se, a título de exemplo, o estudo dos resíduos (para uma das amostras de dimensão $N = 1000$) obtidos a partir da estimação de máxima verosimilhança dos parâmetros do modelo, figuras 4.1.6-4.1.8. As FAC e FACP são apresentadas conjuntamente com os respectivos intervalos de confiança (a tracejado), obtidos através da reamostragem estacionária por blocos, com $p = 0.02$.

Através da análise das figuras, nada parece contrariar a hipótese de estacionariedade. As FAC e FACP apresentam, para qualquer *lag*, valores muito próximos de zero, o que é característico do processo de ruído branco. As figuras parecem assim evidenciar a não rejeição da hipótese de que os resíduos se comportam como um ruído branco, podendo então concluir-se que o modelo estimado se ajusta bem à série temporal em causa.

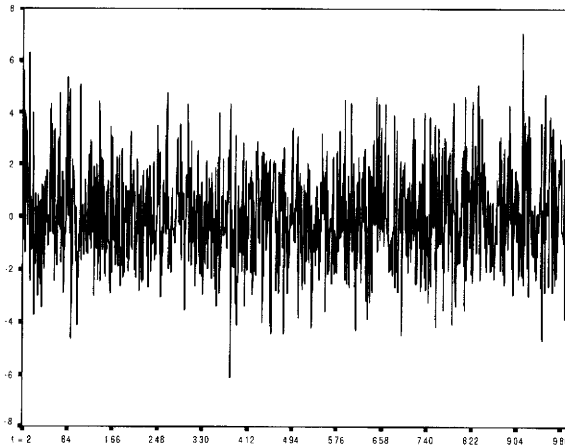


Figura 4.1.6: Resíduos do modelo DSINAR(1) ajustado: Binomial *thinning*.

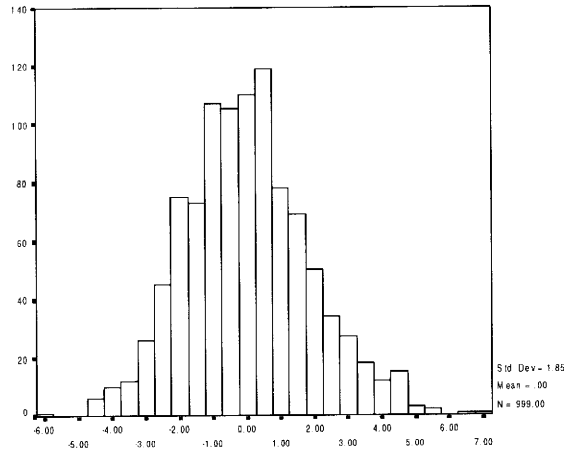


Figura 4.1.7: Histograma dos resíduos do modelo DSINAR(1) ajustado: Binomial *thinning*.

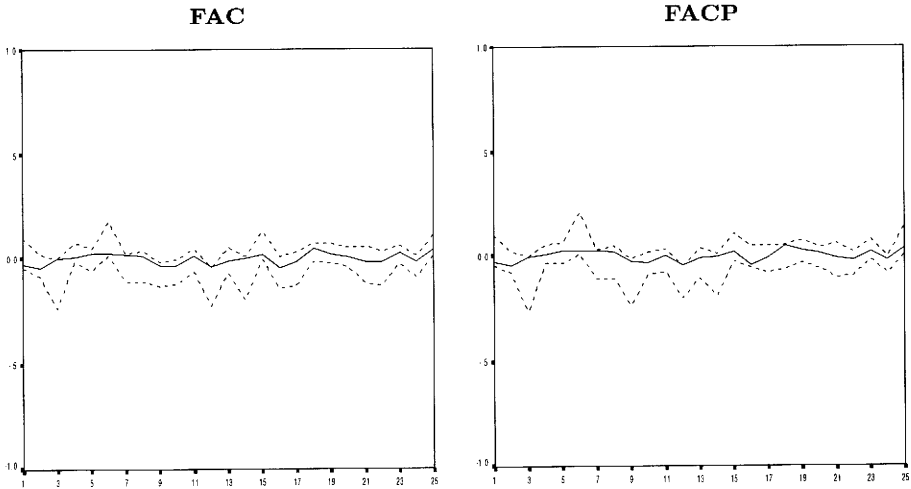


Figura 4.1.8: FAC e FACP amostrais (a cheio) dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança *bootstrap* a 95% (a tracejado): Binomial *thinning*.

Recorrendo-se à expressão de σ_{ξ}^2 , obtida na secção 2.3, obteve-se ainda a média e o EQM das estimativas da variância do processo $\{Z_t\}_{t \in \mathbb{Z}}$, σ_Z^2 . O

estimador de σ_Z^2 escreve-se como

$$\hat{\sigma}_Z^2 = \hat{\sigma}_\xi^2 - \hat{\mu}_Z \sum_{i=1}^K \overline{\beta_{t,i}(1 - \beta_{t,1})}, \quad (4.1.1)$$

onde $\overline{\beta_{t,i}(1 - \beta_{t,1})} = \frac{1}{N-i+1} \sum_{t=i}^N \prod_{j=0}^{i-1} \hat{\alpha}_{t-j}(1 - \hat{\alpha}_t)$, $i \geq 1$, e $\hat{\alpha}_t = \exp(\hat{\omega}X_t)/(1 + \exp(\hat{\omega}X_t))$. Note que ao usar-se a notação $\hat{\alpha}_t$ está a cometer-se um abuso de notação, uma vez que α_t é uma v.a.

Estudos por simulação revelaram que um K igual a 10% da dimensão da amostra conduz a boas estimativas da variância do processo $\{Z_t\}_{t \in \mathbb{Z}}$.

A tabela 4.1.4 contém a informação média acerca das médias e das variâncias amostrais dos resíduos, bem como acerca das variâncias e dos EQM amostrais da variância do processo $\{Z_t\}_{t \in \mathbb{Z}}$. Pode observar-se que as médias amostrais dos resíduos (em 100 séries simuladas) se encontram muito próximas de zero e que as variâncias amostrais de σ_Z^2 estão muito próximas do verdadeiro valor, com os EQM amostrais a aproximarem-se de zero à medida que a dimensão da amostra aumenta.

Por último, na figura 4.1.9 ilustra-se a modelação do conjunto de dados ($N = 1000$), a partir de $(\hat{\omega}, \hat{\lambda})^{[2]}$. Obtém-se ainda a média amostral do erro absoluto entre a série real e a série estimada que representaremos abreviadamente por \overline{EA} . Tal como vimos na secção 2.9, um modo de avaliar a grandeza do erro absoluto é através do coeficiente de ajustabilidade

$$\hat{p}_{E_t} = \frac{\overline{Y} - \overline{EA}_t}{\overline{Y}},$$

com \overline{Y} o estimador usual da média da série real.

A partir desta modelação (figura 4.1.9) obteve-se uma média do erro absoluto entre a série real e a série estimada de $\overline{EA}_t = 1.45522$, o que corresponde a um $\hat{p}_{E_t} = 0.7795$. Pode, portanto, afirmar-se que a ajustabilidade do modelo DSINAR(1) aos dados é bastante satisfatória.

^[2] $\hat{Y}_t = \exp(\hat{\omega}X_t)/(1 + \exp(\hat{\omega}X_t))Y_{t-1} + \hat{\lambda}$

Tabela 4.1.4: Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos.

N	$\hat{\mu}_\xi$	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_Z^2$	$EQM(\hat{\sigma}_Z^2)$
50	-0.0043878	3.3909	2.2391	0.43445
100	-0.0019120	3.3719	2.0621	0.24139
250	-0.0029829	3.3456	1.9842	0.068590
500	-0.0013999	3.3774	2.0189	0.042315
1000	-0.0015039	3.3651	2.0062	0.018016

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Binomial *thinning* para $\omega = 0.2$ e $\lambda = 2$.

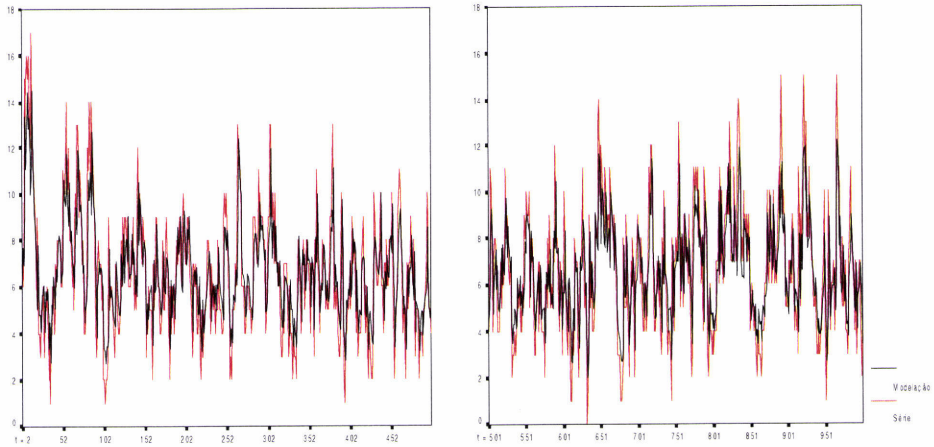


Figura 4.1.9: Modelação de uma realização do processo DSINAR(1): Binomial *thinning*.

4.1.2 Série explicativa não-estacionária

Nesta secção vamos proceder a um estudo análogo ao da secção anterior, substituindo o processo estacionário, $\{X_t\}_{t \in \mathbb{Z}}$, por um processo não-estacionário.

De modo a gerarem-se amostras do processo DSINAR(1) não-estacionário, simularam-se realizações do processo $\{X_t\}_{t \in \mathbb{Z}}$ de acordo

com (2), para as várias dimensões da amostra anteriormente propostas. Foram ainda escolhidos valores de $\omega = 0.05$ e $\lambda = 2$ para a simulação do processo $\{Y_t\}_{t \in \mathbb{Z}}$, para a qual se segue o mesmo procedimento que no caso estacionário.

Comportamento assintótico dos estimadores dos parâmetros do modelo

As figuras 4.1.10-4.1.13 que se apresentam de seguida dizem respeito a uma realização do processo DSINAR(1), considerando $Z_t \sim Po(2)$, $\omega = 0.05$ e uma amostra de dimensão $N = 1000$. Observando a figura 4.1.10 verifica-se uma tendência crescente, pelo que a realização é claramente não-estacionária, também evidente pelo decrescimento lento da FAC, figura 4.1.12. Note-se que, apesar da função de autocorrelação de um processo não-estacionário não existir, ela pode sempre ser estimada a partir de uma realização finita desse processo.

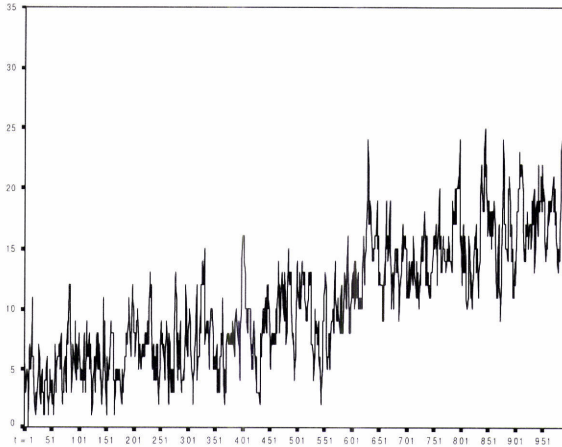


Figura 4.1.10: Realização, de dimensão $N = 1000$, do processo DSINAR(1): Binomial *thinning*.

Na tabela 4.1.5 apresentam-se, uma vez mais, as médias, as variâncias e os erros quadráticos médios amostrais (obtidos em 100 séries simuladas) das estimativas dos parâmetros do modelo, obtidas através da estimação de máxima verosimilhança condicional, para as várias dimensões da amostra. Pode concluir-se, a partir da sua análise, que estes estimadores conduzem a

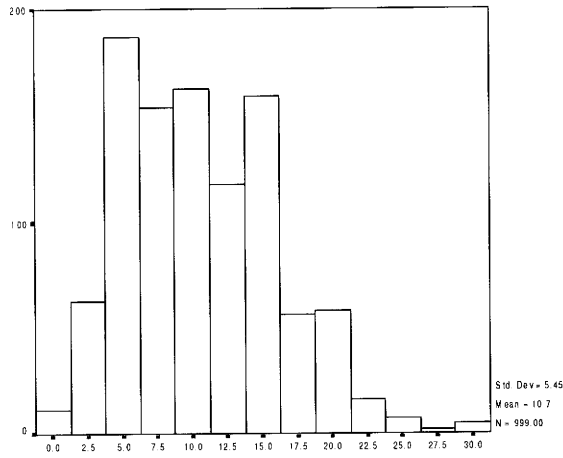


Figura 4.1.11: Histograma do processo DSINAR(1): Binomial *thinning*.

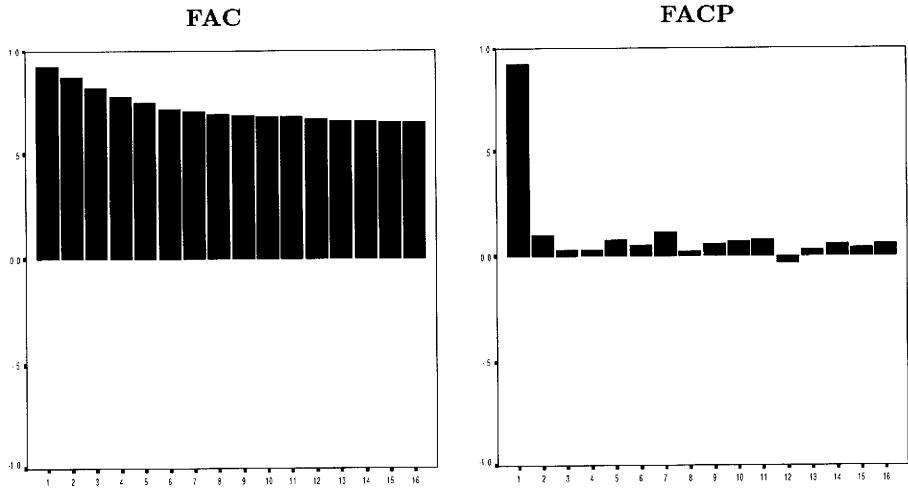


Figura 4.1.12: FAC e FACP amostrais do processo DSINAR(1): Binomial *thinning*.

boas estimativas para o vector parâmetro, (ω, λ) , mesmo para amostras de menor dimensão ($N = 50$) e que os EQM amostrais aproximam-se de zero à medida que a dimensão da amostra aumenta.

Efectuando os testes de Kolmogorov-Smirnov e Shapiro-Wilk verifica-se

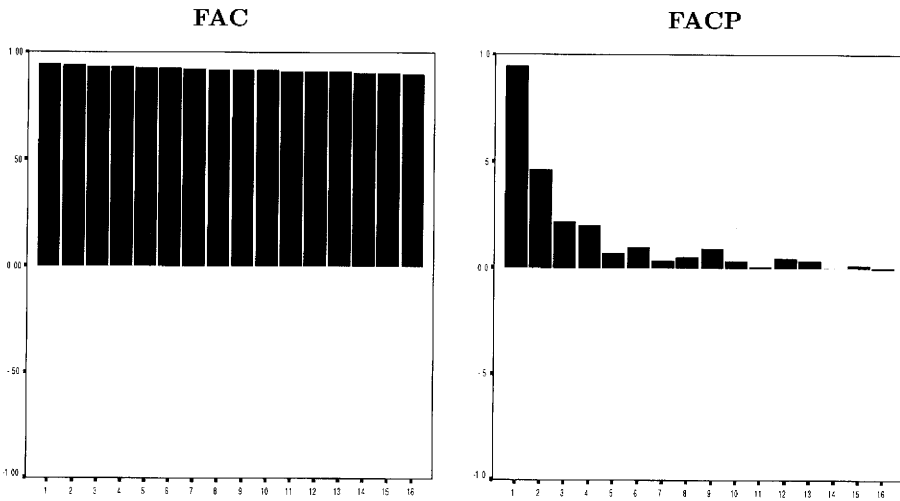


Figura 4.1.13: FAC e FACP amostrais do processo $\{X_t\}_{t \in \mathbb{Z}}$: Binomial thinning.

(tabela 4.1.6) que as estimativas de máxima verossimilhança condicional dos parâmetros do modelo se podem considerar, ao nível de significância de 5%, como provenientes de uma v.a. com distribuição Normal de valor esperado igual ao respectivo parâmetro. Esta conclusão não é, contudo, válida para o parâmetro λ considerando uma amostra de dimensão $N = 50$.

Tabela 4.1.5: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = 0.05$			$\lambda = 2$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
50	0.04993	0.001982	0.001962	1.9918	0.18279	0.18103
100	0.05132	0.000854	0.000855	2.0074	0.088147	0.088202
250	0.04953	0.000114	0.000114	2.0126	0.030199	0.030357
500	0.05012	0.000025	0.000025	1.9991	0.015394	0.015395
1000	0.05004	0.0000034	0.0000031	1.9942	0.006393	0.006393

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Binomial thinning.

Tabela 4.1.6: Testes de normalidade. *p-values* dos testes de K-S e de S-W: caso não-estacionário.

N	$\hat{\omega}$		$\hat{\lambda}$	
	K-S	S-W	K-S	S-W
50	> 0.200	0.396	0.011	0.0
100	> 0.200	0.586	> 0.200	0.412
250	> 0.200	0.611	> 0.200	0.162
500	> 0.200	0.056	> 0.200	0.213
1000	> 0.200	0.183	> 0.200	0.607

Nota: As estimativas de ω e de λ foram obtidas em 100 séries simuladas, utilizando-se o operador Binomial thinning para $\omega = 0.05$ e $\lambda = 2$.

Avaliação da qualidade do ajustamento do modelo

O resíduo, ξ_t , no instante t , define-se tal como no caso estacionário, ou seja,

$$\xi_t = Y_t - \alpha_t Y_{t-1} - \lambda, \quad t = 1, \dots, N.$$

Como os resíduos ξ_t não são outra coisa que a diferença entre os valores observados e os valores estimados através do modelo estocástico ajustado, espera-se — caso o modelo ajustado seja adequado para descrever os dados em causa — que esta diferença seja, em média, quase nula e que seja aproximadamente não-correlacionada. Ou seja, espera-se que a informação que o modelo ajustado não consegue explicar (ou captar) não apresente qualquer estrutura de correlação. Estas condições coincidem com as propriedades características das inovações de um processo DSINAR(1) generalizado, não-estacionário. Isto é, as inovações dadas por $\xi_t = \alpha_t \circ^G Y_{t-1} - \alpha_t Y_{t-1} + Z_t - \lambda$, $\forall t \in \mathbb{Z}$ e para qualquer distribuição $G(\cdot)$ têm média nula e autocovariância nula, para qualquer lag (ver observação 2.3.1). Sob a hipótese de que $Var[\xi_t] < \infty$, as inovações são não-correlacionadas. No que diz respeito à variância das inovações, não se espera (sob a hipótese de que $Var[\xi_t] < \infty$) que seja constante. Contudo, quanto menor esta for, em cada instante de tempo, relativamente à variância dos valores observados (nesse mesmo instante de tempo) melhor será a adequação do modelo ajustado. Quer isto dizer que é pouca a variabilidade da série que não é explicada pelo modelo.

Nas figuras 4.1.14-4.1.16, apresenta-se o estudo dos resíduos, para uma

das séries de dimensão $N = 1000$. A partir da observação das figuras podemos uma vez mais concluir que nada parece contrariar as hipóteses postuladas.

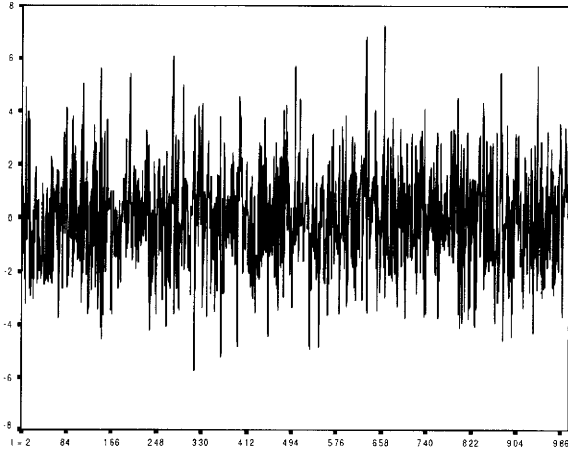


Figura 4.1.14: Resíduos do modelo DSINAR(1) ajustado: Binomial *thinning*.

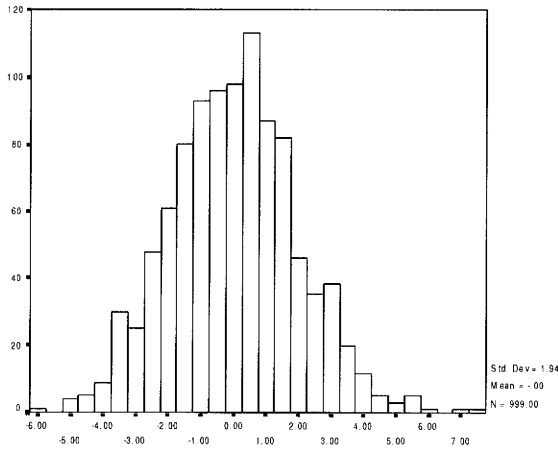


Figura 4.1.15: Histograma dos resíduos do modelo DSINAR(1) ajustado: Binomial *thinning*.

Na tabela 4.1.7 apresenta-se, para as várias dimensões da amostra,

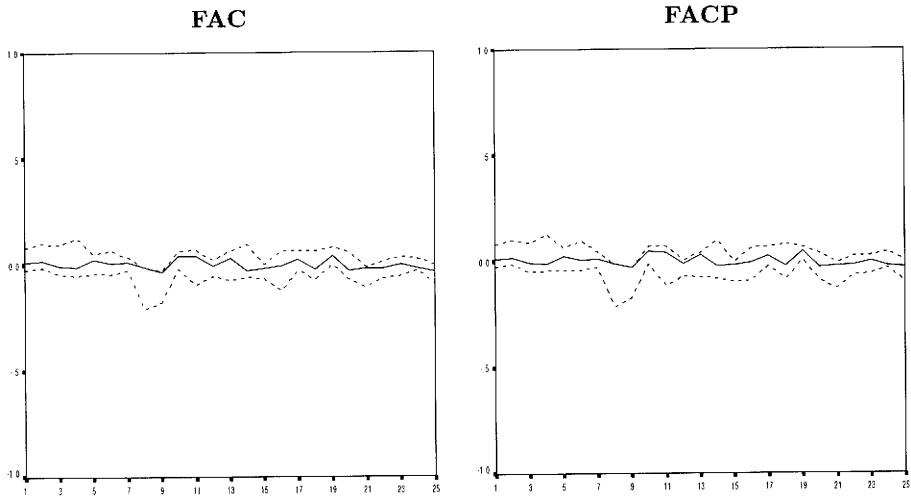


Figura 4.1.16: FAC e FACP amostrais (a cheio) dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança *bootstrap* a 95% (a tracejado): Binomial *thinning*.

a informação média acerca das médias e das variâncias amostrais dos resíduos. Apresentam-se ainda as estimativas da variância do processo $\{Z_t\}_{t \in \mathbb{Z}}$ (obtidas como se de um processo estacionário se tratasse), bem como os seus respectivos erros quadráticos médios amostrais.

Tal como no caso estacionário, verifica-se que as médias amostrais dos resíduos se encontram muito próximas de zero e que as estimativas médias de σ_Z^2 se encontram muito próximas do verdadeiro valor.

A modelação do conjunto de dados, escolhido para exemplificação, é ilustrada na figura 4.1.17. De modo a estudar-se a ajustabilidade do modelo à série temporal em causa e visto o processo não ser estacionário em média, procedeu-se à divisão das séries real e estimada em quatro troços ($[2, 250]$, $[251, 500]$, $[501, 750]$ e $[751, 1000]$). Esta divisão da série foi feita de modo simplesmente empírico, através da sua observação. Para cada um dos quatro troços foram calculadas a média amostral da série real, a média do erro absoluto entre a série real e a série estimada e estimado o coeficiente p_{E_t} (tabela 4.1.8).

Apenas para concluir, parece-nos, quer através do estudo dos resíduos quer através das estimativas (muito próximas de 1) do coeficiente de

Tabela 4.1.7: Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos.

N	$\hat{\mu}_\xi$	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_Z^2$	$EQM(\hat{\sigma}_Z^2)$
50	0.002469	3.1852	2.1230	0.343991
100	0.002347	3.1888	2.0189	0.167280
250	0.001875	3.2528	1.9979	0.075320
500	0.003583	3.3476	1.9872	0.039995
1000	0.001612	3.5117	1.9973	0.020196

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Binomial *thinning* para $\omega = 0.05$ e $\lambda = 2$.

Tabela 4.1.8: Coeficiente de ajustabilidade do modelo, p_{E_t} .

Troço	\bar{Y}	$\hat{p}_{E_t} = \frac{\bar{Y} - \bar{E}A_t}{\bar{Y}}$
250	5.54	0.7335
500	7.88	0.8037
750	12.04	0.8770
1000	17.10	0.9042

Nota: Estimativas obtidas utilizando-se o operador Binomial *thinning*, para $\omega = 0.05$ e $\lambda = 2$.

ajustabilidade que se obteve um bom ajustamento do modelo DSINAR(1) à série temporal não-estacionária.

Note-se que a ajustabilidade do modelo à série observada melhora, por um lado, à medida que a média amostral aumenta, o que fica a dever-se à própria definição de resíduo. Ou seja, sendo o ajustamento do modelo em cada instante de tempo t , \hat{Y}_t , obtido a partir da expressão linear

$$\hat{Y}_t = \frac{\exp(\hat{\omega}X_t)}{1 + \exp(\hat{\omega}X_t)}Y_{t-1} + \hat{\lambda},$$

resulta que $\hat{\lambda} < \hat{Y}_t < Y_{t-1} + \hat{\lambda}$, não captando, assim, os valores da série observada que se encontram abaixo de $\hat{\lambda}$. Por outro lado, e como seria de esperar, esta ajustabilidade também melhora à medida que a variância da série observada é comparativamente superior à variância da série residual.

Verificou-se, por simulação, que quanto menor é a média amostral de $\{Y_t\}_{t \in \mathbb{Z}}$ (ou, no caso de ω ser positivo (negativo), quanto menor (maior) é a média amostral de $\{X_t\}_{t \in \mathbb{Z}}$) maior é a quantidade de observações de grandeza inferior a λ . Deste modo, para séries temporais com médias amostrais muito pequenas a ajustabilidade do modelo só pode revelar-se pouco satisfatória. Uma série temporal nestas condições será apresentada na secção seguinte, para o caso da Poisson *thinning*.

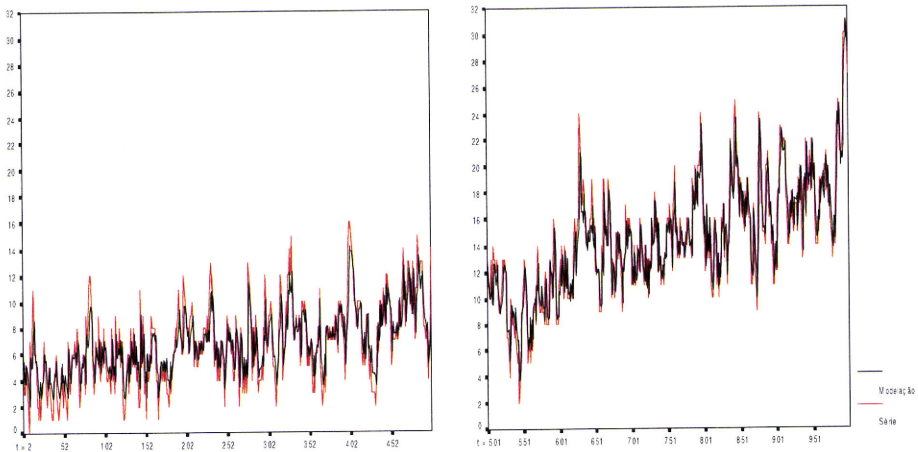


Figura 4.1.17: Modelação de uma realização do processo DSINAR(1): Binomial *thinning*.

4.2 O operador Poisson *thinning*

4.2.1 Série explicativa estacionária

Comportamento assintótico dos estimadores dos parâmetros do modelo

Seguidamente vamos proceder ao estudo por simulação do modelo DSINAR(1), considerando a distribuição Poisson na especificação da operação *thinning*. Mais precisamente, seja

$$Y_t = \alpha_t \circ^{Po} Y_{t-1} + Z_t, \quad t = 1, \dots, N,$$



com $\alpha_t = \exp(\omega X_t)$ e $\omega \in \mathbb{R}$, cuja operação *thinning* se define do seguinte modo:

$$\alpha_t \circ^{Po} Y_{t-1} | \alpha_t, Y_{t-1} \sim Po(\alpha_t Y_{t-1}),$$

onde $Po(\alpha_t Y_{t-1})$ é a distribuição Poisson de valor esperado $\mu_t = \alpha_t Y_{t-1}$.

Para a simulação de realizações do processo consideraram-se igualmente $\{X_t\}_{t \in \mathbb{Z}}$ como um processo AR(1) estacionário, da forma $X_t - \mu_X = 0.6X_{t-1} + \epsilon_t$, com $X_t \sim N(10, \sigma_X^2 = 4)$, $Z_t \sim Po(3)$ e $\omega = -0.1$. As simulações foram igualmente obtidas para amostras de dimensão $N = 50, 100, 250, 500, 1000$, tomando como valor inicial o valor médio do processo.

Como valores iniciais para o método iterativo considerou-se $\omega^{(0)} = 0$ e $\lambda^{(0)} = 1$ e o mesmo critério de paragem que anteriormente. Ao considerarmos como valor inicial para λ , $\lambda^{(0)} = \bar{Y} \sum_{i=0}^k \beta_{t,i}$, o método iterativo de Newton-Raphson, na maior parte dos casos estudados, não converge, apesar de o valor das estimativas ser bastante próximo do verdadeiro valor de λ . Esta falta de convergência fica a dever-se ao facto de, para estes valores iniciais de ω e λ , a derivada em ordem a ω do logaritmo da função de verosimilhança tomar um valor demasiado pequeno e o programa assumi-lo como menos infinito. O mesmo não acontece, no entanto, quando se usa, por exemplo, o valor inteiro mais pequeno que λ pode tomar, ou seja, $\lambda^{(0)} = 1$.

Para um valor de $\lambda = 3$, o número médio de iterações necessárias até se atingir o erro de paragem, para as várias dimensões das amostras e utilizando-se como valores iniciais $\lambda^{(0)} = 1$ e $\omega^{(0)} = 0$, é de cerca de 100.

As figuras 4.2.1-4.2.4 que se apresentam de seguida dizem respeito a uma das realizações do processo DSINAR(1) generalizado, considerando uma amostra de dimensão $N = 1000$.

Da análise da tabela 4.2.1 conclui-se (tal como no caso de se utilizar a distribuição Binomial para especificar o operador *thinning*) que a EMVC conduz a boas estimativas para o vector parâmetro, (ω, λ) . Relativamente ao parâmetro λ tem-se apenas a salientar que os EQM se aproximam de zero de um modo ligeiramente mais lento do que no caso anterior. A partir de estudos por simulação constatou-se que este facto parece ficar a dever-se à escolha do valor inicial de λ , para o método iterativo de Newton-Raphson^[3].

^[3]Quando se utiliza a distribuição Binomial para especificar o operador *thinning* e se escolhe como valor inicial para o método iterativo $\lambda^{(0)} = 1$, este caso não ocorre. Ou seja, para quaisquer dimensões da amostra e valores de λ que se escolham, os EQM amostrais são sempre aproximadamente zero.

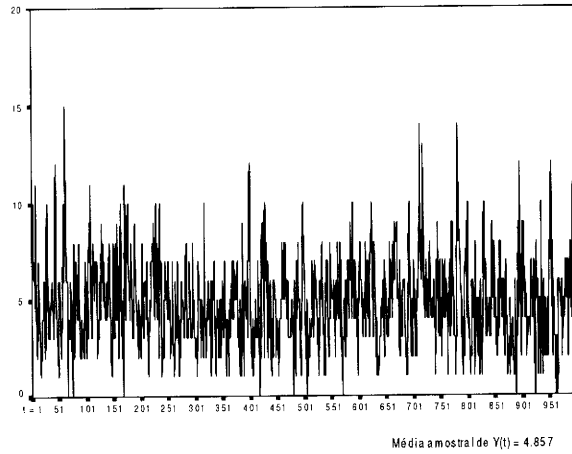


Figura 4.2.1: Realização, de dimensão $N = 1000$, do processo DSINAR(1) generalizado: Poisson *thinning*.

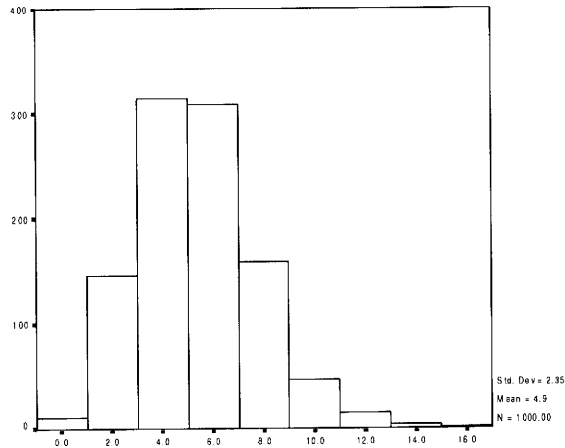


Figura 4.2.2: Histograma do processo DSINAR(1) generalizado: Poisson *thinning*.

Nas tabelas 4.2.2 e 4.2.3 (apenas, a título exemplificativo, para o caso de $N = 250$) apresentam-se dois estudos que confirmam esta explicação, nos quais o valor escolhido para λ e o valor inicial para o método iterativo pouco diferem.

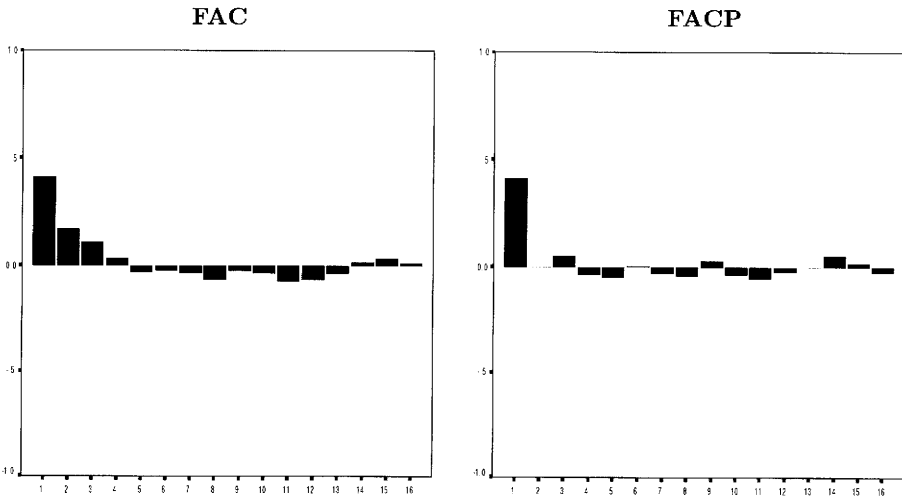


Figura 4.2.3: FAC e FACP amostrais do processo DSINAR(1) generalizado: Poisson *thinning*.

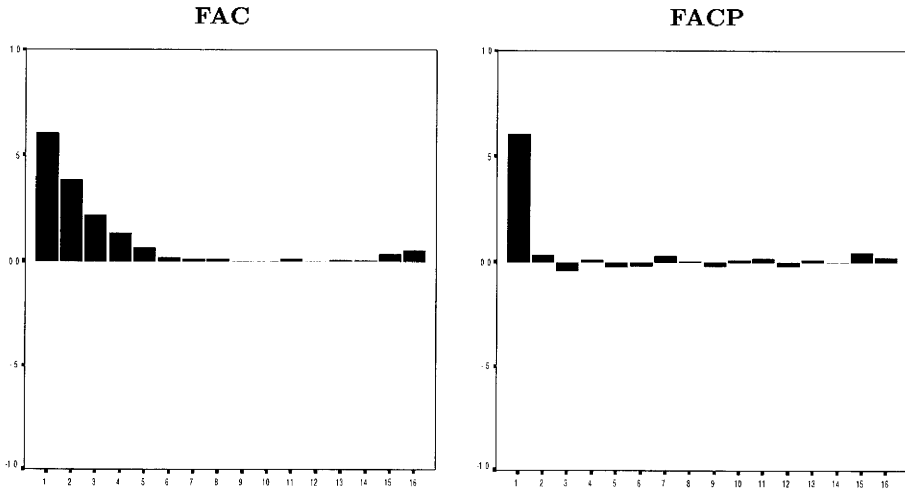


Figura 4.2.4: FAC e FACP amostrais do processo $\{X_t\}_{t \in \mathbb{Z}}$: Poisson *thinning*.

Tabela 4.2.1: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = -0.1$			$\lambda = 3$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
50	-0.1346	0.002807	0.004007	3.3555	0.35063	0.47703
100	-0.1138	0.000768	0.000960	3.1277	0.18344	0.19975
250	-0.1117	0.000449	0.000587	3.1602	0.088095	0.11378
500	-0.1029	0.000112	0.000120	3.0457	0.032217	0.034306
1000	-0.1040	0.000067	0.000083	3.0504	0.019107	0.021647

Nota: Estimativas obtidas em 100 séries simuladas, com $\lambda^{(0)} = 1$ e utilizando-se o operador Poisson *thinning*.

Tabela 4.2.2: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = -0.1$			$\lambda = 1.5$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
50	-0.1377	0.005909	0.007329	1.6214	0.15879	0.17353
100	-0.1133	0.001164	0.001339	1.5085	0.077474	0.077546
250	-0.1043	0.000538	0.000557	1.4912	0.036862	0.036941
500	-0.0988	0.000140	0.000141	1.4814	0.011929	0.012276
1000	-0.0983	0.000071	0.000074	1.4773	0.005947	0.006468

Nota: Estimativas obtidas em 100 séries simuladas, com $\lambda^{(0)} = 1$ e utilizando-se o operador Poisson *thinning*.

Tabela 4.2.3: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = -0.1$			$\lambda = 3$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
250	-0.1079	0.000370	0.000432	3.0678	0.066256	0.070847

Nota: Estimativas obtidas em 100 séries simuladas, com $\lambda^{(0)} = 2$ e utilizando-se o operador Poisson *thinning*.

Verificou-se ainda (tabela 4.2.4), recorrendo aos testes clássicos de Kolmogorov-Smirnov e de Shapiro-Wilk, que as estimativas de máxima verosimilhança de λ , obtidas em várias séries simuladas, se podem considerar (aos níveis de significância usuais) como provenientes de uma v.a. com distribuição Normal de valor médio igual ao respectivo parâmetro, inclusive, para amostras de menores dimensões ($N = 50$). Relativamente ao parâmetro ω , esta propriedade verifica-se apenas para amostras de dimensão $N \geq 500$. Quando se escolhem outros valores de λ e/ou $\lambda^{(0)}$ pouco ou nada se altera a esta conclusão.

Tabela 4.2.4: Testes de normalidade. *p-values* dos testes de K-S e de S-W: caso estacionário.

N	$\hat{\omega}$		$\hat{\lambda}$	
	K-S	S-W	K-S	S-W
50	0.0	0.0	> 0.200	0.560
100	0.003	0.009	> 0.200	0.475
250	0.029	0.0	> 0.200	0.697
500	> 0.200	0.438	> 0.200	0.812
1000	> 0.200	0.061	> 0.200	0.973

Nota: As estimativas de ω e de λ foram obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning* para $\omega = -0.1$ e $\lambda = 3$ ($\lambda^{(0)} = 1$).

Tabela 4.2.5: Testes de normalidade. *p-values* dos testes de K-S e de S-W: caso estacionário.

N	$\hat{\omega}$		$\hat{\lambda}$	
	K-S	S-W	K-S	S-W
50	0.0	0.0	> 0.200	0.073
100	0.145	0.017	> 0.200	0.270
250	0.0	0.0	0.031	0.125
500	0.173	0.139	> 0.200	0.445
1000	> 0.200	0.332	> 0.200	0.999

Nota: As estimativas de ω e de λ foram obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning* para $\omega = -0.1$ e $\lambda = 1.5$ ($\lambda^{(0)} = 1$).

Apesar de estes resultados revelarem ser possível provar que os estimadores de máxima verosimilhança condicional dos parâmetros do modelo são estimadores fracamente consistentes (no sentido em que convergem em média) e são assintoticamente centrados e normais, não nos parece viável encontrar estimadores consistentes (ou fracamente consistentes) para a variância desses estimadores. Assim, à semelhança do que foi feito na secção 4.1, vamos recorrer aos métodos de reamostragem para séries temporais para a obtenção dos pretendidos intervalos de confiança.

Na figura 4.2.5 apresentam-se os intervalos de confiança assintóticos a 95% (a cheio) para os parâmetros ω e λ . Estes intervalos foram obtidos através da reamostragem estacionária por blocos (usando-se 500 repetições *bootstrap*) para um valor de p de 0.02 e para 100 séries simuladas de dimensão $N = 1000$. Estas figuras contêm ainda os verdadeiros valores dos parâmetros (a vermelho) bem como as respectivas estimativas de máxima verosimilhança condicional (a tracejado). A tabela 4.2.6 contém as estimativas das probabilidades de cobertura associadas aos intervalos de confiança a 95% para os parâmetros do modelo, obtidos através da reamostragem por blocos móveis (a título de comparação) e da reamostragem estacionária por blocos.

Tabela 4.2.6: Probabilidades de cobertura (obtidas por reamostragem por blocos móveis (RBM) e por reamostragem estacionária por blocos (REB)) para realizações de dimensão $N = 1000$ do processo DSINAR(1) generalizado.

RBM			REB		
b	ω	λ	p	ω	λ
50	0.14	0.38	0.02	1.00	1.00

Nota: Estimativas obtidas em 500 amostras *bootstrap*.

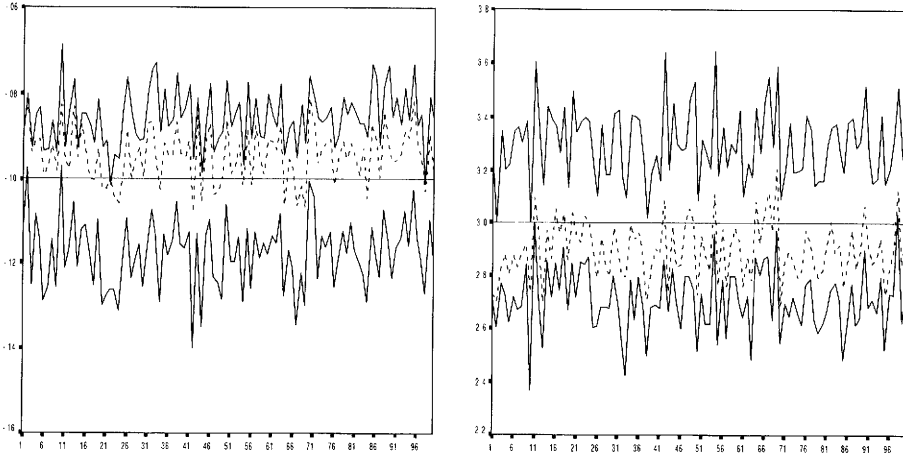


Figura 4.2.5: Estimativas de máxima verosimilhança condicional de ω (à esquerda, com o verdadeiro valor a vermelho), de λ (à direita, com o verdadeiro valor a vermelho) e intervalos de confiança *bootstrap* a 95% (a cheio), em 100 séries simuladas: Poisson *thinning*.

Avaliação da qualidade do ajustamento do modelo

Definindo o resíduo, ξ_t , no instante t , como

$$\xi_t = Y_t - \alpha_t Y_{t-1} - \lambda, \quad t = 1, \dots, N,$$

vamos proceder ao seu estudo para a amostra de dimensão $N = 1000$, com $\lambda = 3$ e $\omega = -0.1$, usada para exemplificação, figuras 4.2.6-4.2.8.

Apesar das FAC e FACP apresentarem, para qualquer *lag*, valores muito próximos de zero, o que é característico do processo de ruído branco, o gráfico dos resíduos exhibe alguns valores discrepantes, o que parece evidenciar que o modelo estimado se ajusta razoavelmente à série temporal em causa (veja-se a figura 4.2.9).

Tal como se pode observar na tabela 4.2.7, verifica-se uma vez mais que as médias amostrais dos resíduos se encontram muito próximas de zero e que as respectivas variâncias amostrais se mantêm praticamente inalteradas, à medida que se aumenta o tamanho da amostra. Contudo, as estimativas de σ_Z^2 apresentam EQM não tão próximos de zero, como quando se usa a Binomial *thinning*. O estimador da variância de Z_t escreve-se, tendo em

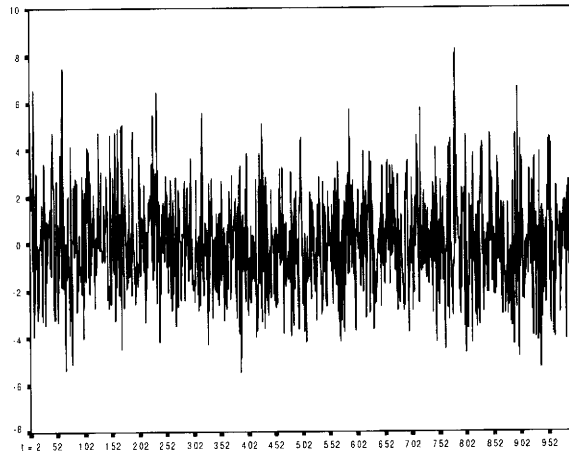


Figura 4.2.6: Resíduos do modelo DSINAR(1) ajustado: Poisson *thinning*.

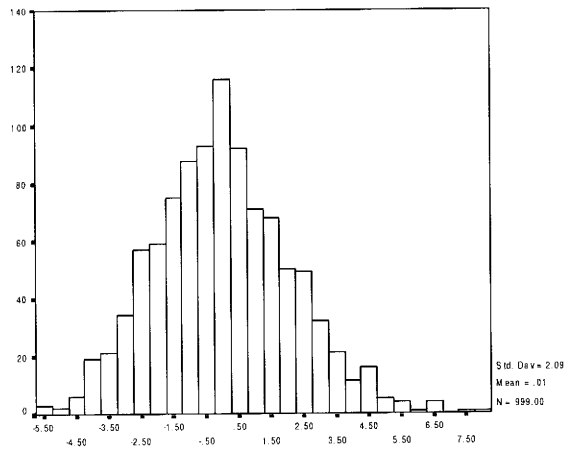


Figura 4.2.7: Histograma dos resíduos modelo DSINAR(1) ajustado: Poisson *thinning*.

conta a expressão da variância dos resíduos obtida na secção 2.3, como

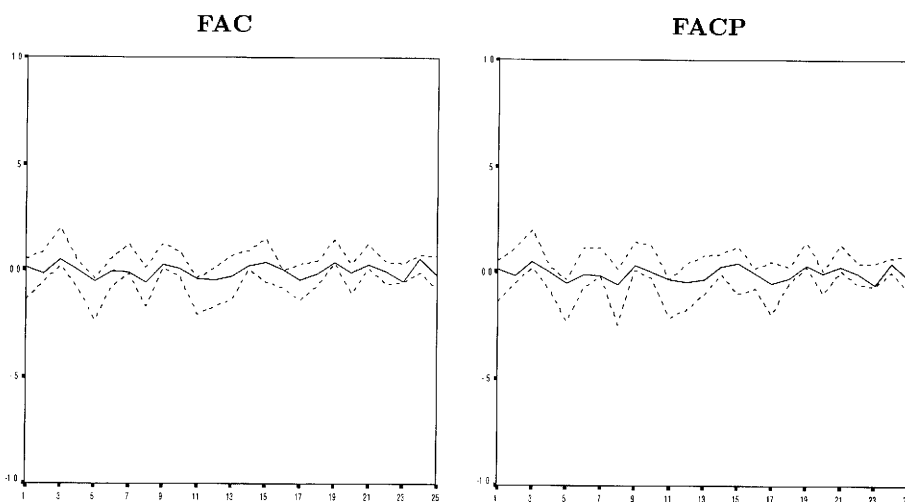


Figura 4.2.8: FAC e FACP amostrais (a cheio) dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança *bootstrap* a 95% (a tracejado): Poisson *thinning*.

$$\begin{aligned} \hat{\sigma}_Z^2 &= \hat{\sigma}_\xi^2 - \hat{\mu}_Z \sum_{i=1}^K \overline{\delta_t \beta_{t,i}^{(0)}} = \hat{\sigma}_\xi^2 - \hat{\mu}_Z \sum_{i=1}^K \overline{\alpha_t \beta_{t,i}^{(0)}} = \\ &= \hat{\sigma}_\xi^2 - \hat{\mu}_Z \left\{ \frac{1}{N} \sum_{t=1}^N \hat{\alpha}_t + \sum_{i=2}^K \frac{1}{N-i+1} \sum_{t=i}^N \hat{\beta}_{t,i} \right\}. \end{aligned} \quad (4.2.1)$$

De novo está a cometer-se um abuso de notação ao usar-se a notação $\hat{\beta}_{t,i} = \prod_{j=0}^{i-1} \hat{\alpha}_{t-j}$, uma vez que α_t é uma v.a. Tal como no caso anteriormente estudado, da Binomial *thinning*, os estudos por simulação revelaram que para um valor de K igual a 10% do tamanho da amostra o estimador conduz a boas estimativas do parâmetro.

Atráves da modelação da série (figura 4.2.9) obteve-se uma média do erro absoluto entre a série real e a série estimada de $\overline{EA}_t = 1.64225$, que corresponde a um coeficiente de ajustabilidade $\tilde{p}_{E_t} = 0.66188$. Pode ainda constatar-se que a ajustabilidade do modelo DSINAR(1) aos dados é razoavelmente satisfatória, no sentido em que os dados ajustados

Tabela 4.2.7: Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos.

N	$\hat{\mu}_\xi$	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_Z^2$	$EQM(\hat{\sigma}_Z^2)$
50	0.010676	4.8326	3.1891	1.55843
100	0.003607	4.7368	3.1254	0.66208
250	-0.000962	4.8269	3.1043	0.24669
500	-0.000001	4.7835	3.0153	0.15410
1000	-0.000206	4.8444	3.0479	0.075576

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning* para $\omega = -0.1$ e $\lambda = 3$.

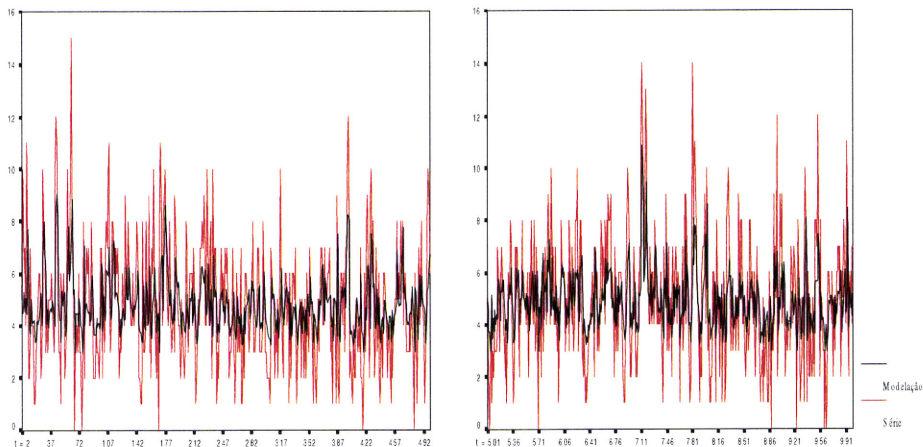


Figura 4.2.9: Modelação de uma realização do processo DSINAR(1): Poisson *thinning*.

acompanham a trajectória da série real, mas oscilando em torno do seu valor médio. Tal fica a dever-se, como já foi referido na secção anterior, ao facto de a média amostral da série temporal ser pequena ($\bar{Y} = 4.857$).

De modo a comprovar esta afirmação simularam-se outras realizações do processo $\{Y_t\}_{t \in \mathbb{Z}}$, com médias amostrais superiores às anteriores. Para tal, considerou-se $\{X_t\}_{t \in \mathbb{Z}}$ como sendo um processo AR(1) da forma $X_t - \mu_X = 0.6X_{t-1} + \epsilon_t$ (com $X_t \sim N(4, \sigma_X^2 = 2)$), $Z_t \sim Po(2)$ e $\omega = -0.1$.

A série temporal que se apresenta na figura 4.2.10 (a título de exemplo) possui uma média amostral de 6.341.

Como se pode agora verificar a partir da figura 4.2.10, a ajustabilidade do modelo aos dados melhora, obtendo-se uma média do erro absoluto entre a série real e a série estimada de 1.94394, que corresponde a um coeficiente de ajustabilidade de $\tilde{p}_{E_t} = 0.69343$.

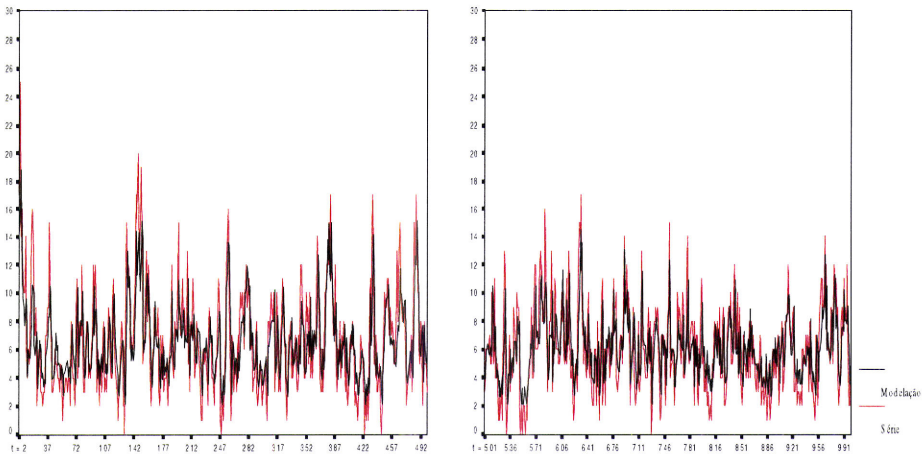


Figura 4.2.10: Modelação de uma realização do processo DSINAR(1): Poisson *thinning*.

Uma vez que com a escolha destes novos valores para os parâmetros se conseguiu melhorar a ajustabilidade do modelo, era de supor que as estimativas da variância de $\{Z_t\}_{t \in \mathbb{Z}}$ também melhorassem. Contudo, em termos dos EQM amostrais, não foi isto que se verificou. A título de exemplo, apresentam-se na tabela 4.2.8 as estimativas de $\sigma_{\mathbb{Z}}^2$, para amostras de dimensão $N = 500$ e $N = 1000$ ^[4].

4.2.2 Série explicativa não-estacionária

Por último, vai apresentar-se o estudo por simulação do processo DSINAR(1) generalizado, utilizando igualmente a distribuição Poisson na

^[4]Com outros valores de K (ver equação (4.2.1)), para além dos 10% do tamanho da amostra, obtiveram-se resultados análogos.

Tabela 4.2.8: Média variância e/ou EQM das estimativas de ξ_t e de σ_Z^2 , por amostras de vários tamanhos.

N	$\hat{\mu}_\xi$	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_Z^2$	$EQM(\hat{\sigma}_Z^2)$
500	-0.005417	6.2475	1.9244	0.23360
1000	-0.006237	6.3816	1.9871	0.12457

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning* para $\omega = -0.1$ e $\lambda = 2$.

especificação da operação *thinning*, mas considerando a série explicativa como não-estacionária. Para o efeito simularam-se realizações do processo $\{X_t\}_{t \in \mathbb{Z}}$ de acordo com (2), para as várias dimensões da amostra. A simulação de realizações do processo $\{Y_t\}_{t \in \mathbb{Z}}$ segue o procedimento já anteriormente descrito, escolhendo-se os valores de $\omega = -0.05$ e de $\lambda = 4$ para os parâmetros do modelo.

Nos gráficos 4.2.11-4.2.14 estão representadas, respectivamente, a série inteira e os respectivos histograma e funções de autocorrelação e autocorrelação parcial amostrais. Observa-se que a série apresenta uma média não constante, com uma forte tendência decrescente nos primeiros instantes de tempo. A não-estacionariedade da série pode ainda comprovar-se a partir da FAC amostral.

Devido à semelhança entre o caso estacionário e o caso não-estacionário (a diferença mais substancial está no aumento do número de iterações necessárias até o método de maximização de Newton-Raphson atingir o erro de paragem proposto) vai apresentar-se o estudo do comportamento assintótico dos estimadores de máxima verosimilhança condicional, apenas para amostras de dimensão $N = 1000$.

A partir destas estimativas de máxima verosimilhança obtiveram-se (para as 100 séries simuladas) uma média e uma variância amostral dos resíduos de $-4.87871E - 03$ e 6.41765 , respectivamente, e uma média amostral da variância do processo $\{Z_t\}_{t \in \mathbb{Z}}$, σ_Z^2 , de 3.97447 , com um EQM de 0.10618 .

A série dos resíduos e os respectivos histogramas e FAC e FACP amostrais, para uma das realizações de dimensão $N = 1000$, são apresentados nos gráficos 4.2.15-4.2.17. Através da sua análise, nada parece contrariar os pressupostos admitidos, podendo assim concluir-se que o modelo estimado

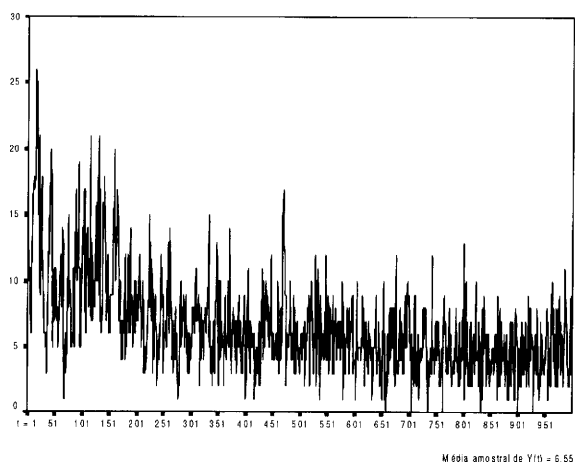


Figura 4.2.11: Realização, de dimensão $N = 1000$, do processo DSINAR(1) generalizado: Poisson *thinning*.

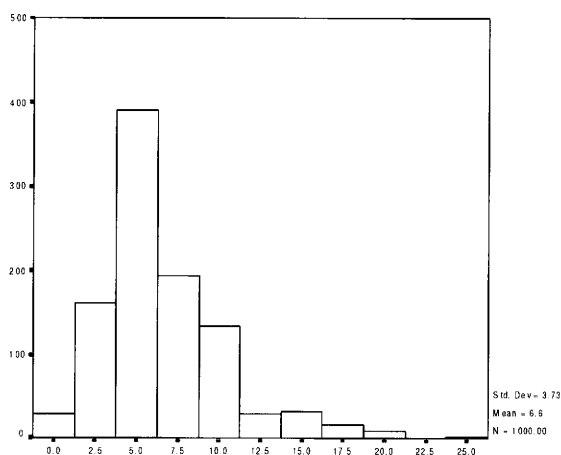


Figura 4.2.12: Histograma do processo DSINAR(1) generalizado: Poisson *thinning*.

se ajusta bem à série temporal em causa, embora o gráfico dos resíduos pareça mostrar maior variabilidade no início da série.

Por último, apresenta-se na figura 4.2.18 a modelação da série. Tal como para o caso da Binomial *thinning*, dividiram-se, de um modo puramente

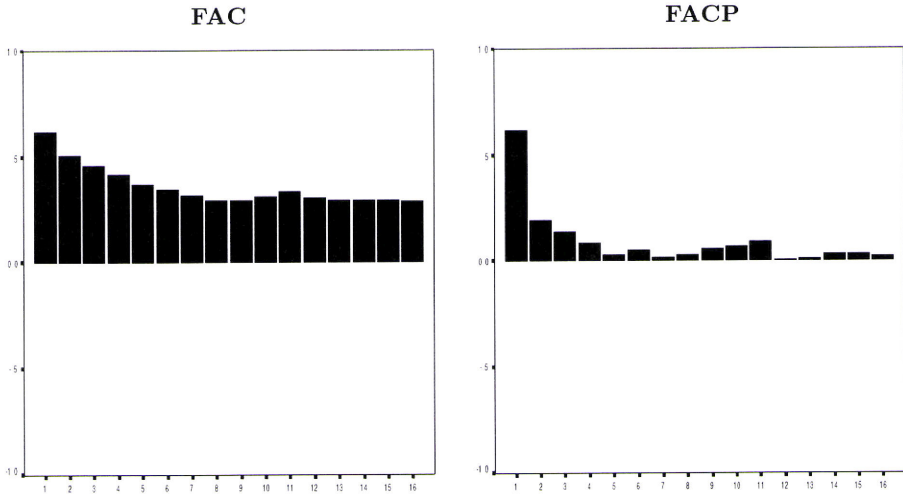


Figura 4.2.13: FAC e FACP amostrais do processo DSINAR(1) generalizado: Poisson *thinning*.

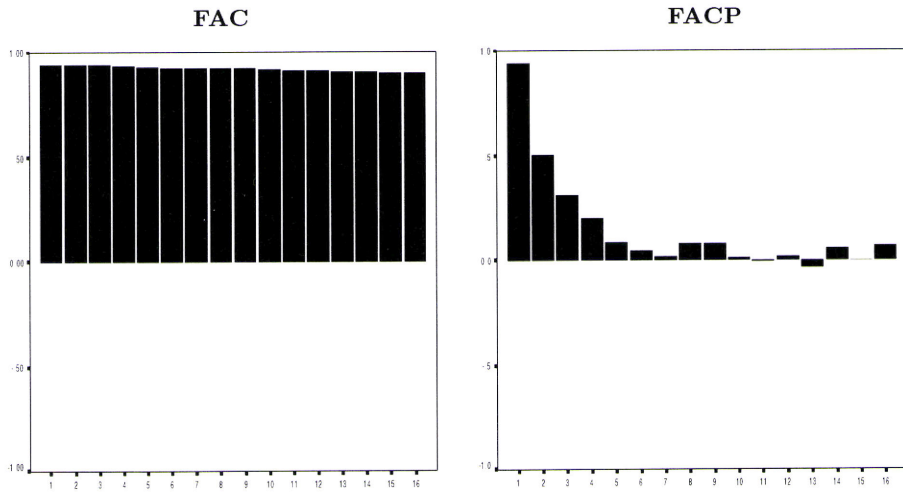


Figura 4.2.14: FAC e FACP amostrais do processo $\{X_t\}_{t \in \mathbb{Z}}$: Poisson *thinning*.

empírico, as séries real e estimada em quatro troços. A tabela 4.2.11 contém

Tabela 4.2.9: Média, variância e EQM das estimativas de (ω, λ) , por EMVC e amostras de vários tamanhos.

N	$\omega = -0.05$			$\lambda = 4$		
	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}^2$	$EQM(\hat{\omega})$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}^2$	$EQM(\hat{\lambda})$
1000	-0.0511	0.000014	0.000015	4.0148	0.023743	0.023726

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning*.

Tabela 4.2.10: Testes de normalidade. p -values dos testes de K-S e de S-W: caso não-estacionário.

N	$\hat{\omega}$		$\hat{\lambda}$	
	K-S	S-W	K-S	S-W
1000	> 0.200	0.896	> 0.200	0.548

Nota: Estimativas obtidas em 100 séries simuladas, utilizando-se o operador Poisson *thinning* para $\omega = -0.05$ e $\lambda = 4$.

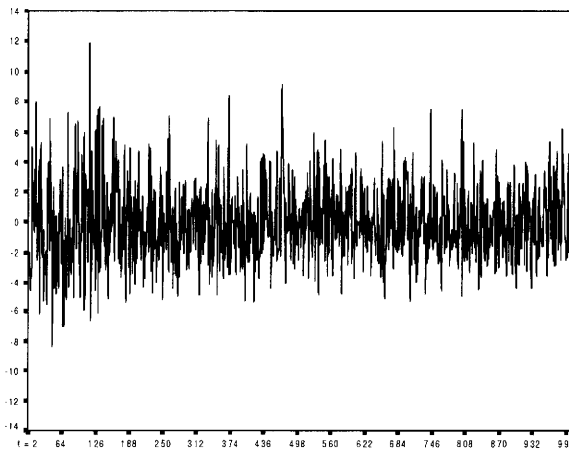


Figura 4.2.15: Resíduos do modelo DSINAR(1) ajustado: Poisson *thinning*.

a informação, por troço, relativa à média do erro absoluto entre a série real

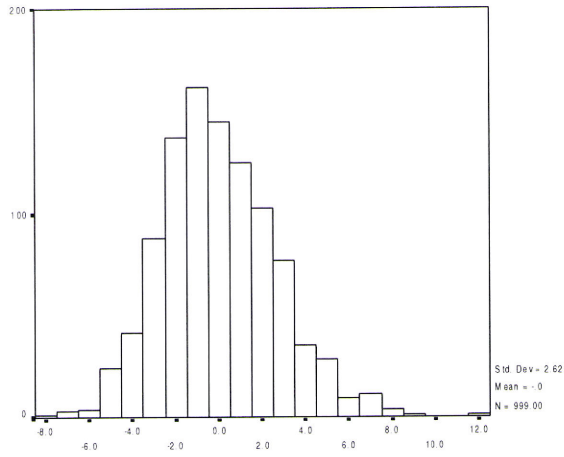


Figura 4.2.16: Histograma dos resíduos modelo DSINAR(1) ajustado: Poisson *thinning*.

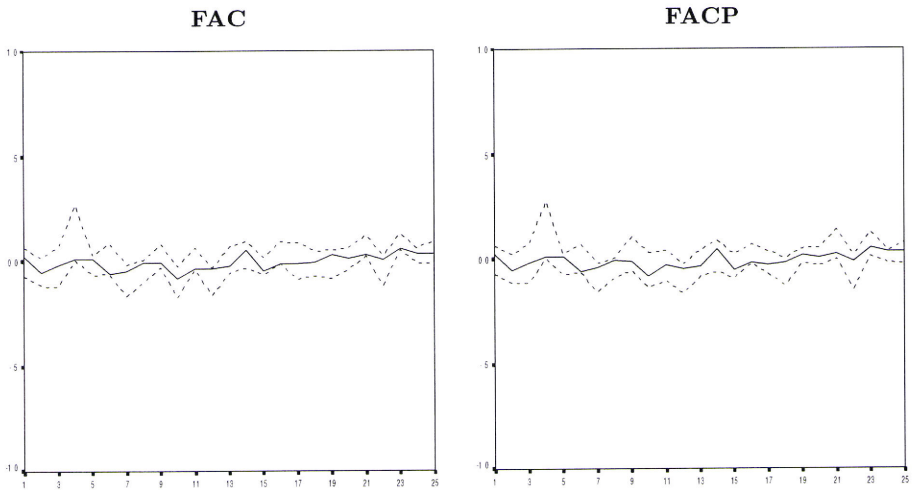


Figura 4.2.17: FAC e FACP amostrais (a cheio) dos resíduos do modelo DSINAR(1) ajustado e respectivos intervalos de confiança *bootstrap* a 95% (a tracejado): Poisson *thinning*.

e a série estimada e ao coeficiente p_{E_t} .

Uma vez mais verifica-se que o ajustamento do modelo à série inteira

Tabela 4.2.11: Coeficiente de ajustabilidade do modelo.

Troço	\bar{Y}	$\hat{p}_{E_t} = \frac{\bar{Y} - EA}{\bar{Y}}$
250	9.90	0.7385
500	6.26	0.6855
750	5.43	0.6651
1000	4.62	0.5926

Nota: Estimativas obtidas utilizando-se o operador Poisson *thinning*, para $\omega = -0.05$ e $\lambda = 4$.

é tanto melhor quanto maior for a média amostral do processo $\{Y_t\}_{t \in \mathbb{Z}}$ e quanto maior for a sua variância em relação à variância da série residual.

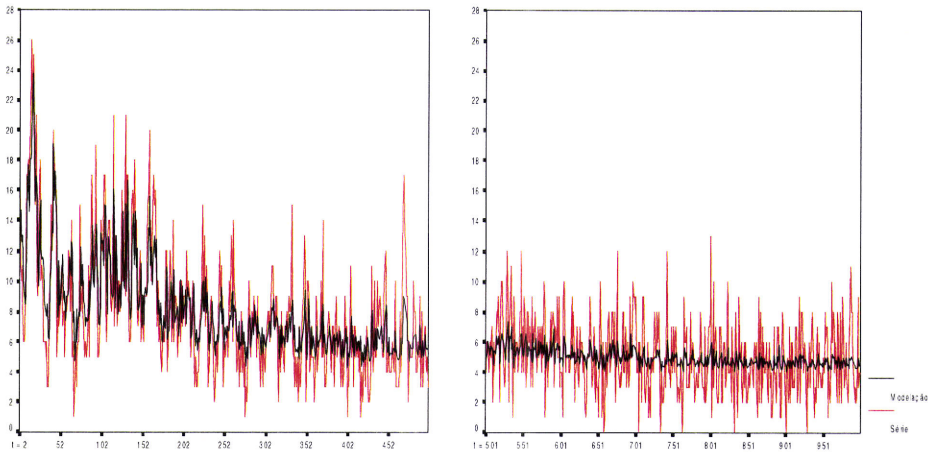


Figura 4.2.18: Modelação de uma realização do processo DSINAR(1): Poisson *thinning*.

4.3 Séries temporais de valores inteiros com lacunas

Na modelação DSINAR(1) generalizado (ou o caso particular da modelação RCINAR(1) generalizado) de séries temporais de valores inteiros

não-negativos, para além da dificuldade em prever valores no mesmo suporte, ainda nos deparamos com o problema destas poderem apresentar lacunas.

No caso da modelação ARMA de séries temporais com suporte em \mathbb{R} , qualquer destas situações não apresenta dificuldade de maior.

No que diz respeito aos processos de contagem, nomeadamente no contexto dos processos DSINAR(1) generalizados ou, no caso mais simples, dos processos RCINAR(1) — os quais até admitem uma representação AR(1) — como preencher lacunas de uma dada série temporal estacionária e de valores inteiros não-negativos?

Por exemplo, pensar no processo RCINAR(1) como um processo AR(1) de parâmetro $E[\alpha_t]$ não nos permite obter estimativas com suporte em \mathbb{Z}^+ para as observações em falta, quer através dos métodos usuais quer recorrendo à representação em espaço de estados^[5].

Será que faz sentido, tal como no caso do preenchimento de lacunas em séries temporais de valores reais^[6], pensar-se numa simples interpolação linear? Para dar uma resposta a esta questão, procedeu-se à simulação de realizações de um processo DSINAR(1) generalizado e à estimação dos seus parâmetros, através da estimação de máxima verosimilhança condicional. Seguidamente, de forma aleatória, criaram-se algumas lacunas (até 10% do total dos dados simulados) e utilizou-se a interpolação linear para o preenchimento das mesmas. Como as observações são inteiras considerou-se não a interpolação linear em si, mas sim o seu valor inteiro mais próximo.

No que diz respeito à modelação de realizações de um processo DSINAR(1) generalizado, o estudo por simulação, antes e após a substituição de algumas das observações simuladas pelas respectivas interpolações lineares, revelou-se muito semelhante. Podemos assim concluir que o preenchimento de lacunas em séries temporais de contagens, através da interpolação linear, não deve ser à partida rejeitado. Esta maneira, aparentemente muito simples, mas que permite obter uma *boa* explicação do modelo, deverá, contudo, ser confrontada com outras técnicas de preenchimento de lacunas, de modo a poder sustentar-se melhor esta afirmação. Para tal é necessário, como trabalho futuro, desenvolverem-se técnicas de previsão com suporte em \mathbb{Z}^+ .

^[5]Ver, *e.g.*, Brockwell e Davis (1987), pp. 482-88.

^[6]Pode ver-se um estudo bastante completo de modelação ARMA de séries temporais de valores reais com lacunas em Ferreira (1993).

Capítulo 5

Aplicação

Todos nós, ao longo do tempo, temo-nos vindo a aperceber de que o clima global está a sofrer grandes alterações, em particular, no que diz respeito ao aumento da temperatura média da superfície terrestre.

Segundo um relatório da UKCIP (do inglês *United Kingdom Climate Impact Programme*) — organismo criado em 1997 e sediado na Universidade de Oxford — desde o início do século XX, a temperatura da superfície terrestre aumentou cerca de 0.6°C, dos quais 0.4°C ocorreram após 1970. 1998 foi considerado o ano mais quente dos últimos 142 anos. Para dar um exemplo, só na Índia morreram pelo menos 2000 pessoas na primeira semana de Junho de 1998, com as temperaturas máximas a atingirem os 49°C^[1].

Mas o caso é mais grave do que aquilo que um leigo na matéria pode suspeitar. Por exemplo, o último relatório da Agência Europeia do Ambiente (AEA), *Sinais Ambientais 2004*^[2], mostra que a Europa está a aquecer mais rapidamente do que o resto do mundo e prevê que esta tendência se vá manter. De acordo com este relatório, estima-se que até 2100 as temperaturas possam subir entre 1.4 e 5.8 graus Celcius relativamente ao que se observava em 1990, incluindo um maior número de ondas de calor e mais intensas também. Já os Invernos correm o risco de desaparecerem até 2080. Estas variações climáticas não ficam apenas a dever-se à variação natural do sistema climático, mas também (ou sobretudo) a factores humanos, nomeadamente às emissões de gases responsáveis pelo efeito de estufa, entre os quais o dióxido de carbono, resultantes da queima de combustíveis fósseis.

^[1]Dados fornecidos pelo *India Today*, em 15 de Junho de 1998.

^[2]Ver em www.eea.eu.int.

No Verão de 2003, em que as temperaturas máximas na Europa atingiram valores bastante elevados, e durante um número bastante elevado de dias, o relatório aponta para a perda de 35 000 vidas, sobretudo no sul da Europa. A acrescentar a este enorme número de óbitos registado, tem-se ainda o impacto económico que estas condições climatéricas provocam. Segundo a AEA, as alterações climáticas e as condições climatéricas provocam, na Europa, um prejuízo anual médio de 10 mil milhões de euros, com tendência para aumentar. Só em Portugal, os incêndios florestais do Verão de 2003 custaram cerca de 925 milhões de euros.

Por tudo isto, muitos têm sido os estudos sobre o impacto das alterações climáticas e das condições climatéricas em variadíssimas áreas, nomeadamente na saúde, área sobre a qual iremos debruçar-nos.

Em Portugal, o Instituto Nacional de Saúde Dr. Ricardo Jorge desenvolveu um sistema de vigilância (sistema Ícaro, a funcionar desde 1999) para, com base na evolução das condições climatéricas na região de Lisboa, nomeadamente nas temperaturas máximas diárias, entre Maio e Setembro, prever e extrapolar a nível nacional a ocorrência de ondas de calor, dada a relação, em algumas regiões, que estas têm revelado com o aumento do número de óbitos (segundo estudos estatísticos efectuados por este Instituto).

Devido à motivação dada por este estudo, apresenta-se neste capítulo a modelação do número de óbitos registado diariamente na Conservatória do Registo Civil de Évora entre 1980 e 1997 (e que traduzem, portanto, o número de óbitos ocorridos no distrito de Évora), através do modelo DSINAR(1) generalizado, usando como covariável não só a série das temperaturas máximas diárias, mas também a das temperaturas mínimas. A série do número de óbitos diários foi fornecida pelo Instituto Nacional de Estatística e as séries das temperaturas máximas e mínimas diárias pelo Instituto de Meteorologia.

A opção por utilizar-se este tipo de modelos de séries temporais de valores inteiros deve-se ao facto de, a não ser que os valores das observações sejam suficientemente grandes, se tornar inadequada a utilização dos métodos baseados em ruídos gaussianos. No que diz respeito, em particular, ao distrito de Évora, a série do número de óbitos diários apresenta valores que oscilam entre 0 e 19, pelo que se justifica a opção tomada.

A utilização tanto da série das temperaturas máximas diárias, como das temperaturas mínimas diárias, deve-se ao facto de existirem regiões do país em que as ondas de calor têm um menor impacto no aumento do número

de mortes do que em outras. Nomeadamente, em regiões onde os Verões são, em geral, bastante quentes e os Invernos relativamente amenos. Estas condições permitem que a população se encontre fisiologicamente adaptada a condições climáticas que, em outras regiões, seriam consideradas adversas. Por exemplo, durante a onda de calor que ocorreu em Portugal, em Junho de 1981, registou-se, no distrito de Lisboa, um pico bastante significativo na série do número de óbitos (o número de óbitos diários atingiu um valor muito próximo dos 160. Fora dos períodos de ocorrência de ondas de calor, este valor não ultrapassa a barreira dos 90 óbitos (Nunes (1997))). Já no distrito de Évora, onde também se registou a ocorrência de uma onda de calor, observou-se um aumento do número de óbitos, mas muito pouco significativo (em geral, durante o período de Verão, o número máximo diário de óbitos, no distrito de Évora, é de 13. Durante o período em que decorreu a onda de calor, esse valor chegou aos 14).

Em resumo, pensamos que as extrapolações para outras regiões, a serem feitas, devem sê-lo com alguma parcimónia. Isto porque uma onda de calor não tem o mesmo impacto em todas as regiões. Além disso, parece-nos razoável admitir que, em algumas regiões, as ondas de frio também possam ter alguma influência no número de óbitos.

De seguida, vamos apresentar as definições, por nós adoptadas, de *onda de calor* e de *onda de frio*.

Segundo a definição do Instituto de Meteorologia, ocorre uma *onda de calor* sempre que a temperatura máxima seja, durante pelo menos 6 dias consecutivos, superior ao percentil 90, *i.e.*, seja superior ao valor da temperatura que ocorre em 90% do tempo ou que é susceptível de ser excedido em 10% do tempo.

Contrariamente ao que diz respeito às ondas de calor, não existe, em Portugal, uma definição oficial de onda de frio^[3]. Por este motivo, vamos dizer que ocorre uma *onda de frio* sempre as temperaturas mínimas, durante pelo menos 5 dias consecutivos, estejam abaixo dos 2°C e, destes 5 dias, em pelo menos 3 dias consecutivos, as temperaturas mínimas tenham sido

^[3]Por exemplo, também na Holanda esta definição não existe. Huynen *et al.* (2001) adoptaram como definição de onda de frio, para este país, a ocorrência de temperaturas mínimas abaixo dos -5°C (inclusive) durante pelo menos 9 dias dos quais, em pelo menos 6, a temperatura mínima foi de -10°C ou menos. A única definição oficial que encontramos diz respeito aos E.U.A. Segundo o *Weather Bureau* dos E.U.A., uma onda de frio caracteriza-se por uma diminuição na temperatura de 16°F (aprox. -8.9°C) em 24 horas, ou de 20°F (aprox. -6.7°C) em 36 horas, independentemente da hora do dia.

abaixo dos 0°C . Esta definição de onda de frio, que estamos adoptar para o distrito de Évora, baseou-se num estudo descritivo das séries dos óbitos e das temperaturas mínimas ocorridas neste distrito, ao longo dos 18 anos disponíveis. Talvez fosse mais correcto encontrar uma definição de onda de frio para todo o país. Optámos, contudo, por o não fazer uma vez que só dispomos da informação relativa ao distrito de Évora.

Será, portanto, com base nestas definições de onda de calor e de onda de frio que se irá averiguar a associação do efeito destas temperaturas “extremas” com o aumento do número de óbitos registado diariamente em Évora.

5.1 Modelação do número de óbitos registado diariamente em Évora entre 1980 e 1997

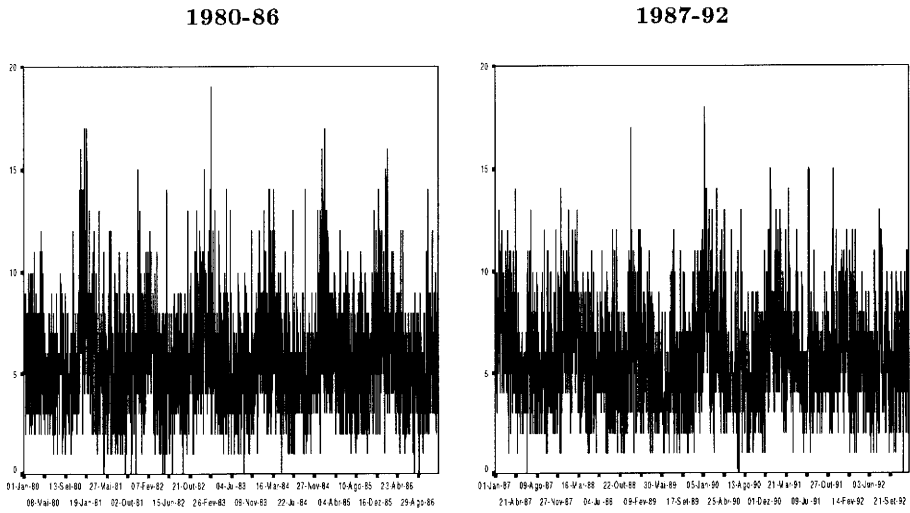
5.1.1 Os dados

Tal como já foi referido, dispomos para este estudo de três séries temporais relativas ao número de óbitos registado diariamente no distrito de Évora e às temperaturas máximas e mínimas diárias, referentes ao mesmo distrito. As temperaturas máximas e mínimas designam-se abreviadamente por T_{\max} e T_{\min} , respectivamente.

Apesar de dispormos da série dos óbitos entre 1980 (1 de Janeiro) e 2002 (31 de Dezembro) apenas dispomos das séries das temperaturas entre 1980 e 31 de Maio de 1998. Por este motivo, optámos por modelar a série dos óbitos durante o período entre 1980 e 1997. Contudo, a título de exemplo, as figuras 5.1.1 e 5.1.2 exibem os gráficos da série dos óbitos diários entre 1980 e 2002. As figuras 5.1.3 e 5.1.4 exibem os gráficos das séries das temperaturas máximas e mínimas diárias entre 1980 e 1997.

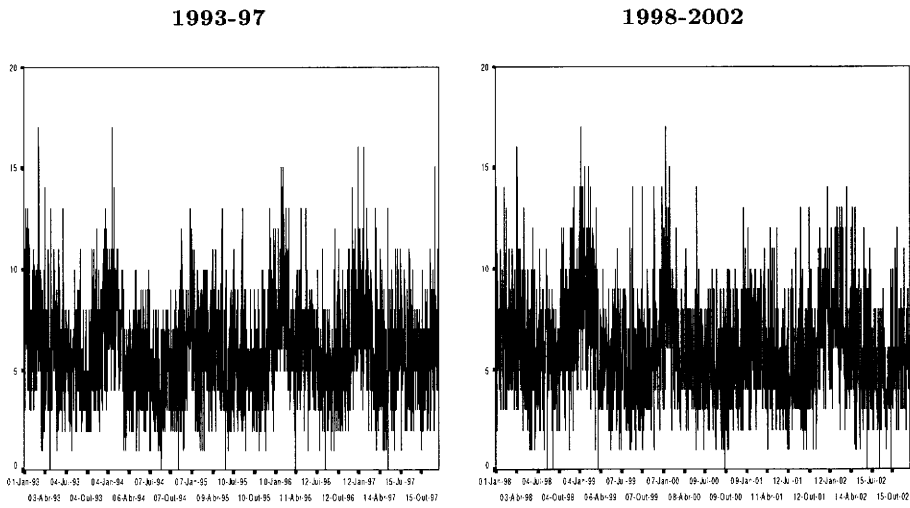
Da observação dos gráficos da série do número de óbitos constata-se que os dados apresentam uma sazonalidade anual, com os meses de Inverno a apresentarem, em geral, valores mais elevados que os restantes meses. Os meses de Primavera, de Verão e de Outono parecem, em alguns anos, exibir um comportamento análogo entre si.

A figura 5.1.7 exhibe uma representação gráfica, em forma de caixa-de-bigodes, do número de óbitos por ano (por uma questão de curiosidade apresentam-se os dados até 2002). À excepção de 1989, todos os



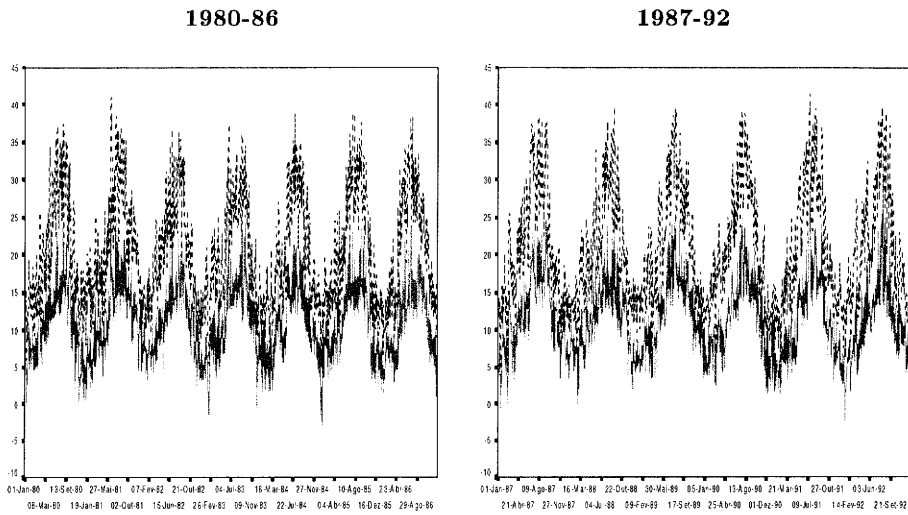
Fonte: Instituto Nacional de Estatística.

Figura 5.1.1: Óbitos registados em Évora entre 1980 e 1992.



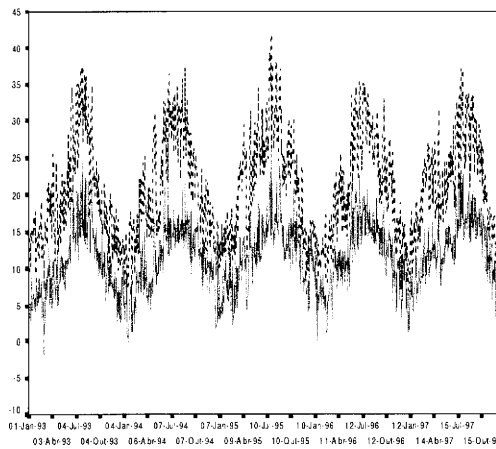
Fonte: Instituto Nacional de Estatística.

Figura 5.1.2: Óbitos registados em Évora entre 1993 e 2002.



Fonte: Instituto de Meteorologia.

Figura 5.1.3: Temperaturas máximas (a tracejado) e mínimas (a cheio) registadas em Évora entre 1980 e 1992.



Fonte: Instituto de Meteorologia.

Figura 5.1.4: Temperaturas máximas (a tracejado) e mínimas (a cheio) registadas em Évora entre 1993 e 1997.

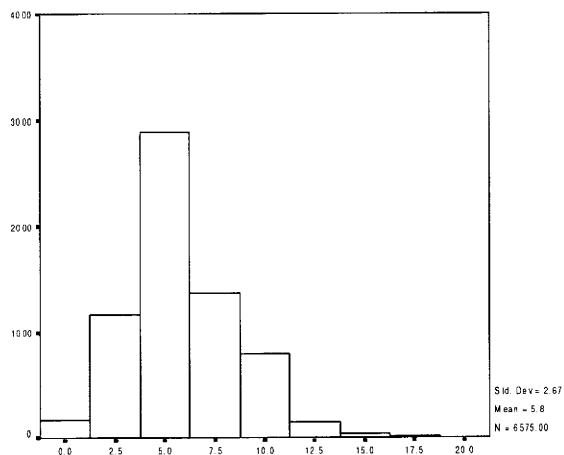


Figura 5.1.5: Histograma dos óbitos registados em Évora entre 1980 e 1997.

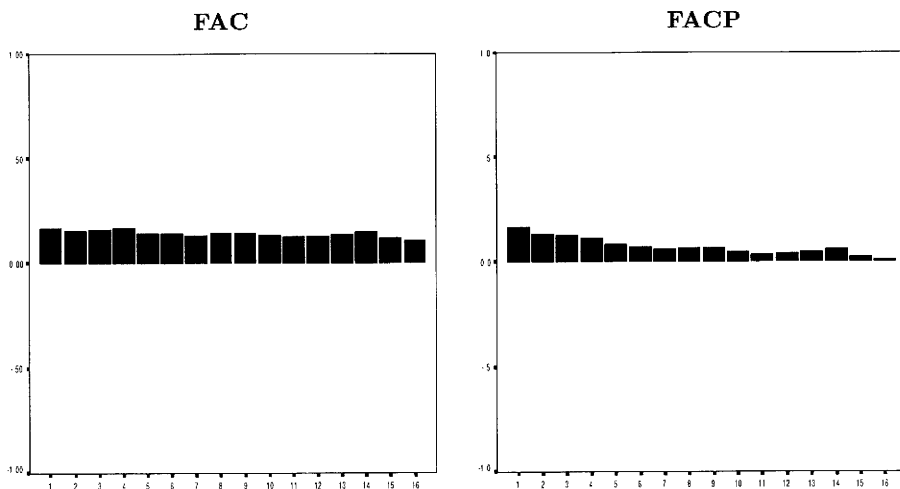


Figura 5.1.6: FAC e FACP amostrais dos óbitos registados em Évora entre 1980 e 1997.

anos exibem pelo menos dois *outliers* superiores, moderados e/ou severos. Através do teste não-paramétrico de Kruskal-Wallis rejeita-se a hipótese nula da igualdade das medianas dos 18 anos em estudo (80-97). Contudo, esta hipótese não se rejeita (ao nível de significância de 5%) para os anos de 80-88

($p\text{-value}=0.07$) e 90-97 ($p\text{-value}=0.073$).

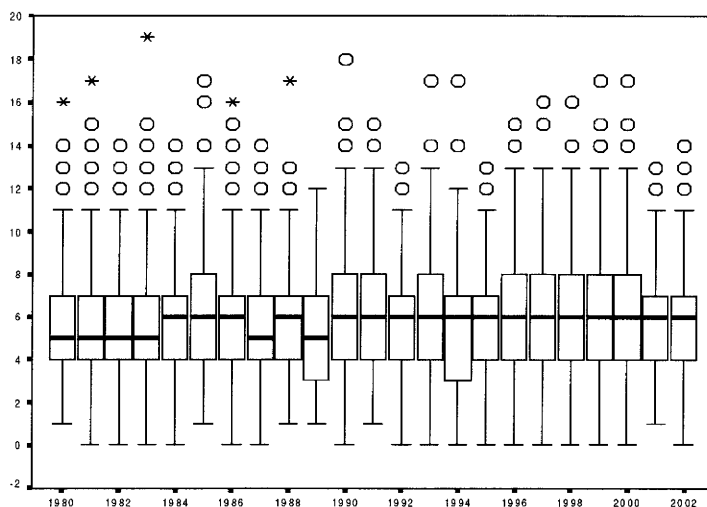


Figura 5.1.7: Caixas-de-bigodes dos óbitos registados em Évora de 1980-2002.

Passando agora a uma análise mais pormenorizada dos dados, vamos analisar a série do número de óbitos diários segundo as estações do ano. Para esta análise foi definido um período de Inverno (que vai desde 1 de Dezembro até 28/29 de Fevereiro) e um período de Verão (que vai desde 1 de Junho até 31 de Agosto)^[4]. Esta escolha deve-se ao facto de serem nestes meses que ocorrem as ondas de calor e de frio.

Centrando a análise na série do número de óbitos diários nos meses de Verão, verifica-se que o processo se mostra relativamente estacionário. Ou seja, os valores variam em torno de um valor fixo (aproximadamente 5) e esta variabilidade parece ser aproximadamente constante. Para os meses de Inverno, os valores variam em torno de um valor fixo (aproximadamente 7), mas a variabilidade já não parece ser constante. Da leitura da função de autocorrelação (figura 5.1.6) já havia indícios para suspeitar que a série dos óbitos, para além de apresentar um comportamento sazonal, também apresenta uma certa tendência.

^[4]O Inverno de 1980 apenas contém os meses de Janeiro e Fevereiro e o Inverno de 1997 está completo, ou seja, contém os meses de Dezembro de 1997 e os meses de Janeiro e Fevereiro de 1998, dado que até estes meses se dispõe das temperaturas máximas e mínimas diárias.

As figuras 5.1.8 e 5.1.9 exibem duas representações gráficas, em forma de caixa-de-bigodes, para os períodos de Verão e Inverno. Relativamente ao Verão, e recorrendo uma vez mais ao teste não-paramétrico de Kruskal-Wallis, não se rejeita, ao nível de significância de 1%, a hipótese nula da igualdade das medianas dos 18 anos. Para os períodos de 80-88 e de 90-97, não se rejeita igualmente esta hipótese obtendo-se, respectivamente, um p -value de 0.075 e um p -value de 0.312.

No que diz respeito ao Inverno rejeita-se, para qualquer nível de significância, a hipótese nula da igualdade das medianas, quer dos 18 anos quer dos períodos de 80-88 e de 90-97.

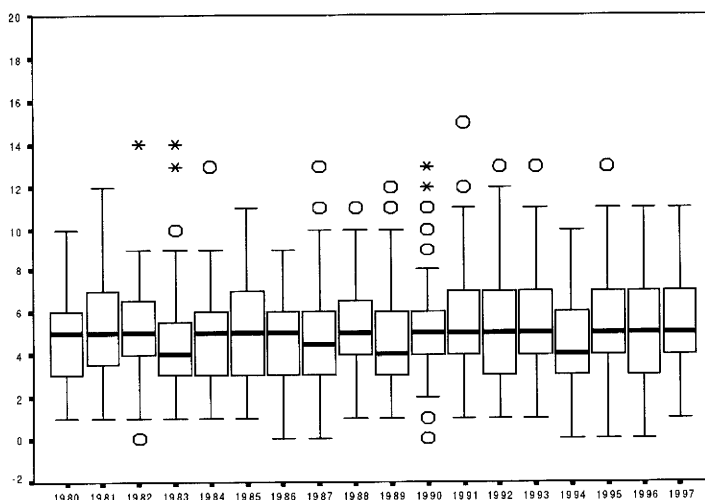


Figura 5.1.8: Caixas-de-bigodes dos óbitos registados em Évora nos meses de Junho, Julho e Agosto, entre 1980-1997.

Considerando ainda esta divisão entre Inverno e Verão, e após o cálculo das correspondentes médias e variâncias amostrais ($\bar{x}_{\text{Inverno}} = 7.20$, $S^2_{\text{Inverno}} = 8.036$) e ($\bar{x}_{\text{Verão}} = 5.08$, $S^2_{\text{Verão}} = 5.643$), fomos averiguar se estas duas colecções de dados podem ser consideradas amostras provenientes de v.a.'s com distribuição de Poisson. O teste de Kolmogorov-Smirnov leva-nos a não rejeitar as hipóteses postuladas (p -value (Inverno) = 0.626 e p -value (Verão) = 0.801). Contudo, esta hipótese é de rejeitar quando consideramos a série dos óbitos na sua totalidade.

Estudemos agora os efeitos das ondas de frio e de calor na mortalidade.

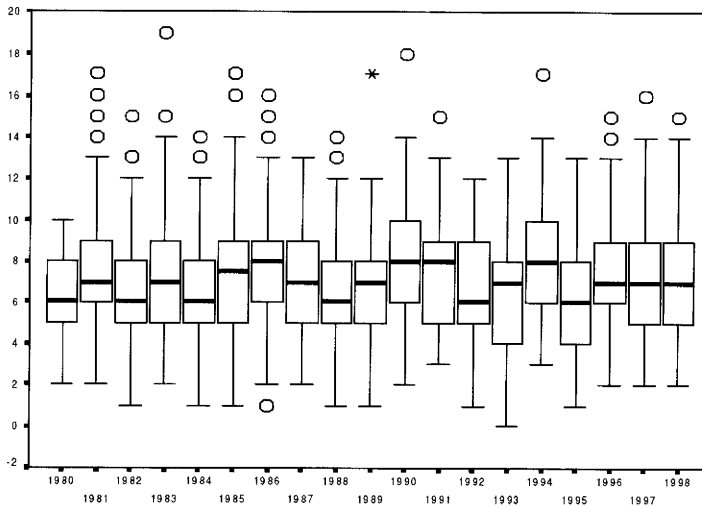


Figura 5.1.9: Caixas-de-bigodes dos óbitos registados em Évora nos meses de Dezembro, Janeiro e Fevereiro, entre 1980-1997.

Nota. Para o inverno de 1980 apenas constam os meses de Janeiro e Fevereiro.

Através do estudo das três séries temporais constatou-se que o impacto de uma onda de calor no aumento do número de óbitos é quase imediata e que, após os 3 dias para além do final de uma onda de calor, o número de óbitos volta ao seu comportamento anterior. Este comportamento parece sugerir que o impacto da onda de calor é sentido essencialmente por pessoas que na altura já se encontravam bastante débeis.

Nos períodos em que ocorreram ondas de frio verificou-se, em alguns casos, um certo acréscimo do número de óbitos. Contudo, em todas as ondas de frio que ocorreram entre 1980 e 1997, o maior impacto surge logo após estas terem terminado e prevalece durante 8 a 11 dias. Após estes 8 a 11 dias, para além do final da onda de frio, surge um valor mais extremo. Só para dar um exemplo, nos 10 dias após a onda de frio de Fevereiro de 1983 (na qual se registou a maior sucessão de dias frios, com temperaturas abaixo dos 0°C) ocorreram 19 óbitos, o maior valor observado nos 18 anos em estudo. Este impacto a longo prazo parece sugerir que o aumento do número de óbitos é, talvez, consequência de gripes ou pneumonias originadas devido a esse arrefecimento. Estas observações carecem, contudo, de uma análise mais aprofundada, nomeadamente de se conhecerem as causas que deram origem a esses óbitos. Os impactos de uma onda de calor e de uma onda

frio na mortalidade, no distrito de Évora, são descritos, respectivamente, na figuras 5.1.10 e 5.1.11.

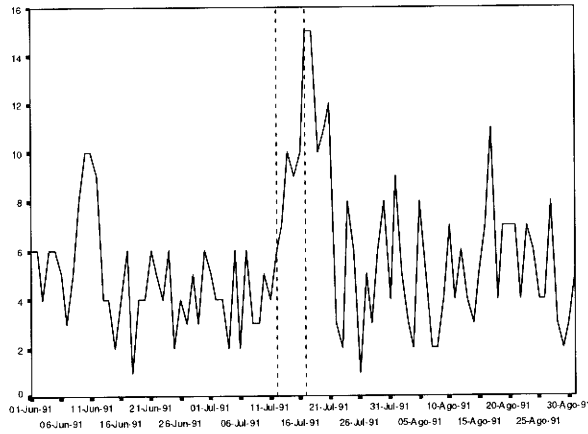


Figura 5.1.10: Número de óbitos antes, durante e após a onda de calor de Julho de 1991 (assinalada através das linhas a tracejado).

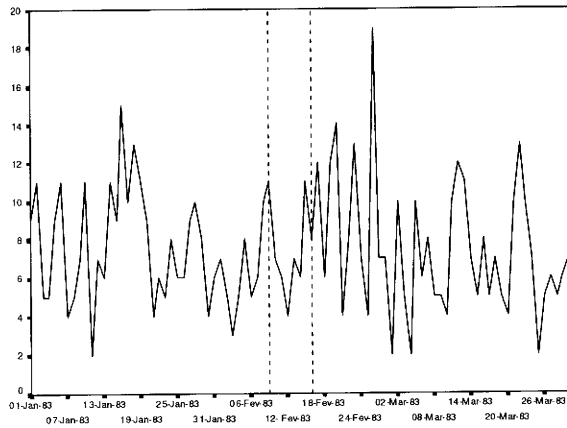


Figura 5.1.11: Número de óbitos antes, durante e após a onda de frio de Fevereiro de 1983 (assinalada através das linhas a tracejado).

Na tabela seguinte apresentam-se os excessos durante as ondas de calor e de frio registadas em Évora entre 1980 e 1997 (tabela 5.1.1).

Para se definir excessos numa onda de calor, ou de frio, começamos por definir *período alargado*. Define-se por *período alargado de uma onda de*

calor, o número de dias em que durou a onda de calor, acrescido de 3 dias após esta ter terminado. E por *período alargado de uma onda de frio*, o número de dias em que durou a onda de frio, acrescido de 11 dias após esta ter terminado.

Deste modo, definiu-se como *excessos numa onda de calor* a diferença entre o número de órbitos observados durante o seu período alargado e o número de órbitos observados durante o período alargado que a precedeu. A comparação com os dias que imediatamente precederam a onda de calor é uma das comparações propostas pelo Instituto Dr. Ricardo Jorge e que nos pareceu adequada, uma vez que nos dá alguma garantia de não ter havido grande alteração da dimensão da população, durante esse período de tempo.

Um outro modo de definir excessos numa onda de calor seria através da comparação do número de órbitos observados durante o respectivo período alargado com o número de órbitos observados em períodos homólogos (pertencentes aos anos anteriores e posteriores mais próximos do ano em que ocorreu a onda de calor), cuja temperatura média se encontrasse próxima da temperatura média que se verificou durante o período em que durou a onda de calor. Através desta comparação estar-se-ia a comparar o número de órbitos em igual número de dias e em condições climatéricas análogas, mas em que num caso a temperatura máxima se manteve acima de um determinado nível durante um certo número de dias consecutivos e no outro caso a temperatura máxima oscilou em torno desse nível. Esta comparação levanta, contudo, algumas dificuldades, nomeadamente na definição de temperaturas próximas.

Como *excessos numa onda de frio*, e seguindo a linha anterior, adoptou-se como definição a diferença entre o número de órbitos observados durante o correspondente período alargado e o número de órbitos observados durante o mesmo número de dias que precederam a onda de frio.

Não se tendo rejeitado a hipótese de que o número de órbitos diários, segundo as estações do ano, é proveniente de uma população com distribuição Poisson, estimámos a probabilidade de cauda associada ao ponto $x' + y'$, onde x' é o número total de órbitos observados durante as ondas de frio, ou de calor, e y' é o correspondente número de dias após estas terem terminado. Para tal, considerámos como estimativa do valor esperado da distribuição Poisson o número total de órbitos observados durante o mesmo número de dias que precederam estas ondas de frio, ou de calor.

Para as duas ondas de frio obtivemos valores muito próximos de zero, para

Tabela 5.1.1: Excessos durante as Ondas de Calor (OC) e de Frio (OF).

Data	Duração(x)	Total de Óbitos na(os)			Excesso
		OF	11 dias(y) após a OF	x+y dias antes da OF	
09/Fev/83	8	60	106	121	45
08/Jan/85	10	91	109	149	51
Total de Óbitos na(os)					
		OC	3 dias(y) após a OC	x+y dias antes da OC	
11/Jun/81	7	58	20	57	21
07/Ago/82	6	38	19	47	10
15/Jul/86	7	42	12	46	8
04/Set/88	6	29	26	41	14
12/Jul/91	6	66	33	37	62

as estimativas das probabilidades atrás referidas. Isto indica que os valores observados se afastam significativamente do padrão existente em situação de não ocorrência de ondas de frio. Relativamente às ondas de calor, apenas os valores observados nas ondas de calor de 1982 e de 1986 não se afastam significativamente do padrão existente, em situação de não ocorrência de ondas de calor.

5.1.2 Modelação da série de óbitos registados em Évora

Seguidamente apresenta-se uma aplicação dos processos DSINAR(1) generalizados na modelação da série de óbitos registados diariamente na Conservatória do Registo Civil de Évora entre 1980 e 1997 (correspondentes, portanto, ao distrito de Évora). Tal como foi atrás referido, nesta modelação serão utilizadas como covariáveis as temperaturas máximas e mínimas diárias. A não-estacionariedade da série dos óbitos não constitui um problema na obtenção das estimativas dos parâmetros do modelo, as quais são obtidas através da estimação de máxima verosimilhança condicional (EMVC). Os estudos por simulação efectuados no capítulo 4 evidenciaram que, para amostras de grande dimensão, os estimadores de máxima verosimilhança condicional conduzem a boas estimativas para o vector parâmetro do modelo, no sentido em que os erros quadráticos médios se encontram muito próximos de zero.

Tendo em atenção as considerações anteriores sobre o impacto das ondas de calor e de frio na mortalidade foram calculados os coeficientes de

correlação de Spearman entre a série dos óbitos e as séries das temperaturas máximas e mínimas (com e sem desfasamentos) durante as ondas de calor e de frio, respectivamente. Falar de séries das temperaturas (máximas e/ou mínimas) sem desfasamento significa que se está a considerar que, num determinado dia, o número de óbitos é influenciado pelas temperaturas desse mesmo dia. Contrariamente, falar em séries das temperaturas com desfasamento significa que está a considerar-se que o número de óbitos, num determinado dia, é influenciado pelas temperaturas registadas em dias anteriores.

Durante as ondas de calor, o maior coeficiente de correlação foi encontrado entre as variáveis *número de óbitos e temperatura máxima diária desfasada de 1 dia* (0.370) e, durante as ondas de frio, entre as variáveis *número diário de óbitos e temperatura mínima diária desfasada de 8 dias* (-0.208). Em ambos os casos a correlação entre as variáveis é estatisticamente significativa.

A modelação mais satisfatória (em termos da análise dos resíduos e do coeficiente de ajustabilidade, p_{E_t}) foi obtida usando o seguinte caso particular do modelo DSINAR(1) generalizado:

$$Y_t = \alpha_t \circ^{P_o} Y_{t-1} + Z_t,$$

onde $Z_t \sim \text{Po}(\lambda)$,

$$Y_t | Y_{t-1}, \text{Tmin}_{t-8}^*, \text{Tmax}_{t-1} \sim \text{Po}(\alpha_t Y_{t-1} + \lambda)$$

e $\alpha_t = \exp(\omega_1 \text{Tmin}_{t-8}^* + \omega_2 \text{Tmax}_{t-1})$. Por outras palavras, utilizou-se como covariáveis as temperaturas máximas desfasadas de 1 dia e a seguinte função das temperaturas mínimas desfasadas de 8 dias e que designaremos por Tmin_t^* :

$$\text{Tmin}_t^* = \begin{cases} \text{Tmin}_t - \overline{\text{Tmin}}, & 0^\circ\text{C} < \text{Tmin}_t < 5^\circ\text{C} \\ \text{Tmin}_t, & \text{c. c.} \end{cases}$$

onde $\overline{\text{Tmin}}$ é a média amostral das temperaturas mínimas durante o período em que estas estiveram abaixo dos 5°C .

A opção por utilizar-se esta função das temperaturas mínimas ficou a dever-se ao facto de em alguns períodos de dias consecutivos em que as temperaturas mínimas se situaram entre os 0°C e os 5°C (não constituindo, segundo a definição por nós adoptada, uma onda de frio) se ter observado um certo acréscimo do número de óbitos. Deste modo, com esta função pretendeu dar-se mais peso a estas temperaturas.

Note-se que os critérios de selecção da ordem do modelo mais apropriado na modelação de um dado conjunto de dados, nomeadamente o critério AIC (critério sugerido por Akaike em 1973) ou suas derivações, não foram aqui utilizados. Isto deve-se ao facto de o critério AIC, e as extensões deste terem sido construídos de modo a serem interpretados como uma estimativa (aproximada) centrada da medida de discrepância de Kullback-Leibler entre o modelo ajustado e o verdadeiro modelo, o qual se assume como linear e gaussiano.

A tabela 5.1.2 contém as estimativas de máxima verosimilhança condicional dos parâmetros do modelo, obtidas a partir de programas elaborados em linguagem FORTRAN.

Tabela 5.1.2: Estimativas de $(\omega_1, \omega_2, \lambda)$, por EMVC.

$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\lambda}$
-0.16219	-0.03875	5.05507

A partir destes resultados podemos constatar uma maior contribuição da variável T_{t-8}^* para a explicação de Y_t , do que da variável T_{t-1} . Face ao valor da estimativa de ω_2 , interessava agora saber se esse parâmetro é ou não nulo. Por outras palavras, interessa testar a seguinte hipótese

$$H_0 : Y_t \sim f_{\vec{\theta}}, \vec{\theta} = (\omega_1, \lambda) \quad vs \quad H_1 : Y_t \sim f_{\vec{\theta}}, \vec{\theta} = (\omega_1, \omega_2, \lambda).$$

Efectuando-se o teste da razão de verosimilhanças (baseado na verosimilhança do processo DSINAR(1) Poisson), o valor observado de menos duas vezes o logaritmo da estatística de teste foi de 766.36. Considerar a aproximação da distribuição amostral desta estatística de teste pela distribuição Qui-Quadrado, com 1 grau de liberdade, leva-nos a rejeitar a hipótese nula (o quantil de probabilidade 0.95 da distribuição Qui-Quadrado com 1 grau de liberdade é de 3.8415). Desta constatação decorre, então, que de entre os vários modelos estudados (nos quais se incorpora como única série explicativa, a série das temperaturas), o modelo postulado é aquele que melhor descreve a série temporal do número de óbitos registado diariamente no distrito de Évora.

Olhemos agora para os resíduos do modelo ajustado. Tal como foi definido no capítulo 2, o resíduo, ξ_t , define-se, no instante t , do seguinte modo:

$$\xi_t = Y_t - E[Y_t | Y_{t-1}, X_t] = Y_t - \alpha_t Y_{t-1} - \lambda.$$

Sendo $\{\alpha_t\}_{t \in \mathbb{Z}}$ um p.e. não-estacionário tem-se, para qualquer distribuição $G(\cdot)$ (com $G(\cdot)$ a distribuição discreta que especifica a operação *thinning* generalizada) que $\{\xi_t\}_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s de média nula e $Cov(\xi_t, \xi_{t+h}) = 0, h > 0$. Sob a hipótese de que a variância dos resíduos existe, prova-se que os resíduos são não-correlacionados (ver observação 2.3.1). Será, pois, sob esta hipótese que iremos estudar a adequação do modelo postulado.

As figuras 5.1.12-5.1.15 exibem, respectivamente, o gráfico dos resíduos do modelo DSINAR(1) generalizado ajustado, o respectivo histograma e as correspondentes funções de autocorrelação (FAC) e autocorrelação parcial (FACP) amostrais. A partir da sua observação consideramos que as hipóteses postuladas são razoavelmente confirmadas. Ou seja, os gráficos dos resíduos não exibem tendências. As FAC e FACP amostrais desses mesmos resíduos assemelham-se às FAC e FACP de uma sucessão de v.a.'s não-correlacionadas e o histograma apresenta-se razoavelmente simétrico em torno de zero.

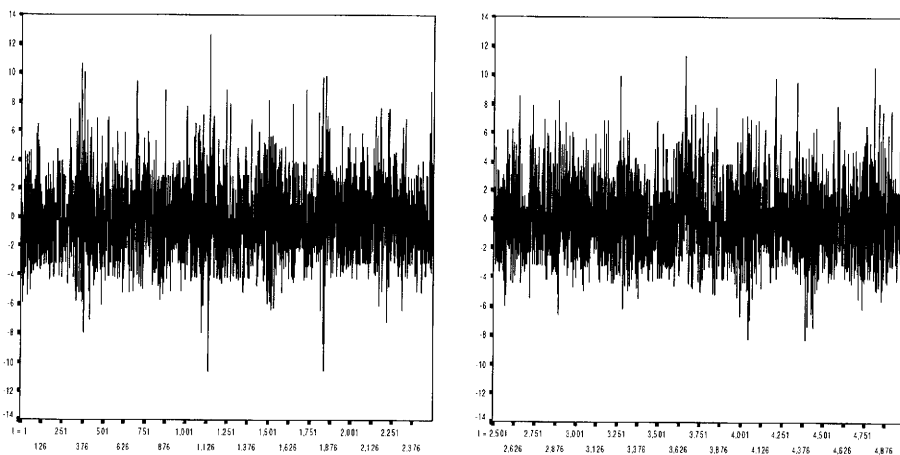


Figura 5.1.12: Resíduos do modelo DSINAR(1) generalizado ajustado.

Os intervalos de confiança (a 95%) para as FAC e FACP amostrais dos resíduos do modelo ajustado foram obtidos através de 500 amostras *bootstrap* obtidas, por sua vez, através da reamostragem estacionária por blocos, com um valor de p suficientemente pequeno ($p = 0.007$).

As figuras 5.1.16 e 5.1.17 exibem os gráficos da séries dos óbitos e da respectiva série estimada. De modo a estudar-se a ajustabilidade do

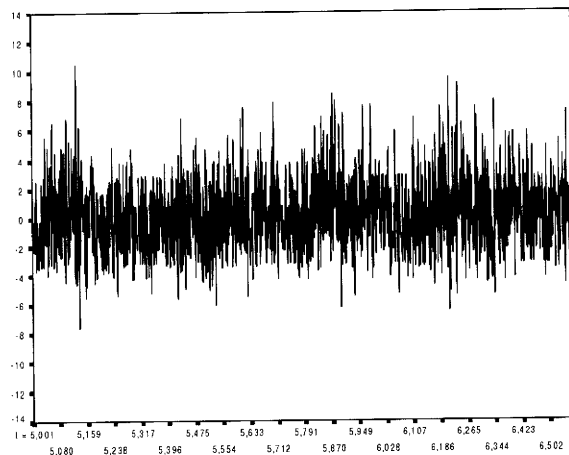


Figura 5.1.13: Resíduos do modelo DSINAR(1) generalizado ajustado.

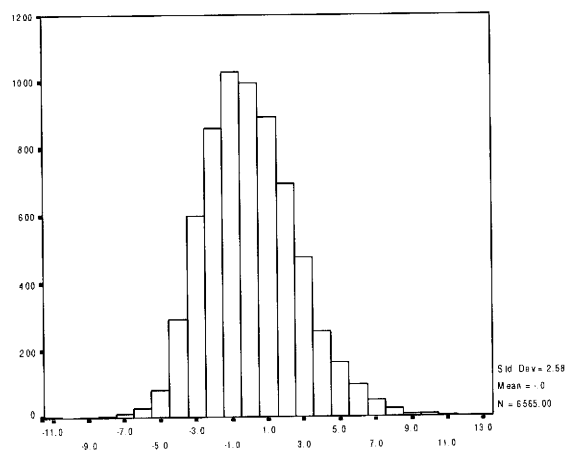


Figura 5.1.14: Histograma dos resíduos do modelo DSINAR(1) generalizado ajustado.

modelo à série temporal calculou-se, ainda, o quadrado do coeficiente de ajustabilidade, $p_{E_t}^2$. Para tal, dividiram-se as séries real e estimada em períodos de Inverno (apenas os meses de Dezembro, Janeiro e Fevereiro) e de Verão (meses de Junho, Julho e Agosto), calcularam-se a média amostral da

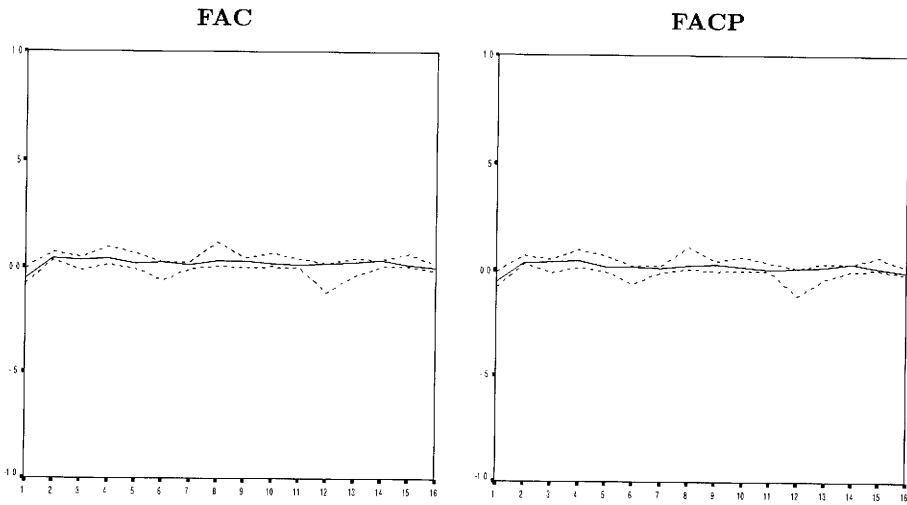


Figura 5.1.15: FAC e FACP amostrais (a cheio) dos resíduos do modelo DSINAR(1) generalizado ajustado e respectivos intervalos de confiança *bootstrap* a 95% (a tracejado).

série real e a média do erro absoluto entre a série real e a série estimada^[5]. Para o Inverno, obteve-se um $p_{E_t}^2$ médio de 0.552425 e para o Verão, um $p_{E_t}^2$ médio de 0.405104. Apesar de se terem obtido valores de $p_{E_t}^2$ algo afastados de 1, o modelo ajustado parece bastante promissor, uma vez que a temperatura é a única variável explicativa que se está de momento a incorporar no modelo. Daí este não conseguir captar adequadamente a variabilidade “total” dos dados. Embora o modelo não consiga captar adequadamente a variabilidade “total” dos dados, este consegue, contudo, captar adequadamente as principais características evidenciadas pelos dados.

Com base nas características da série dos óbitos diários, nomeadamente no aparecimento de um valor extremo após uma onda de frio, vamos fazer a previsão desses valores extremos, com base na informação até à ocorrência dessas mesmas ondas de frio. Verificou-se que estes valores extremos são tanto maiores, quanto maior é a duração da onda. Embora não seja possível tomar como segura esta constatação (por apenas dispormos de duas ondas de frio, com base na nossa definição), ela parece-nos bastante plausível, pelo que

^[5]A média amostral do média do erro absoluto entre a série real e a série estimada, considerando a série no seu todo, foi de 1.797295

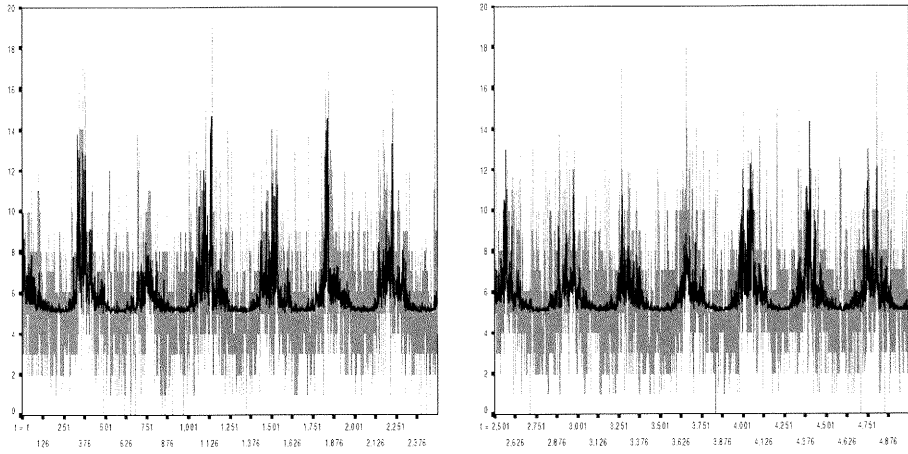


Figura 5.1.16: Série dos órbitos (a claro) e série estimada (a escuro).

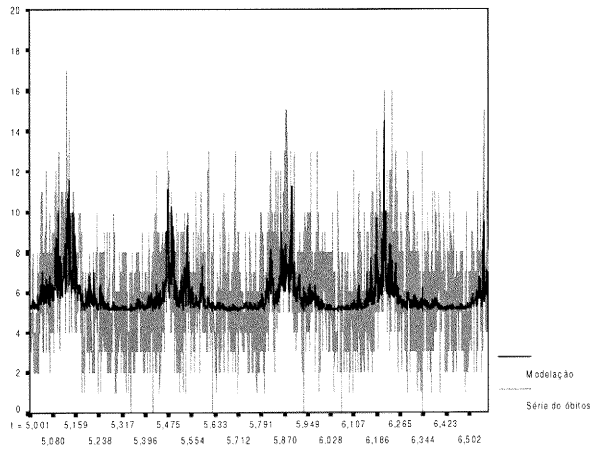


Figura 5.1.17: Série dos órbitos (a claro) e série estimada (a escuro).

as estimativas desses valores extremos serão obtidas tendo em consideração o efeito cumulativo das temperaturas mínimas, durante o período em que decorreu a onda de frio. As estimativas são, assim, obtidas através da

seguinte expressão:

$$\exp\left(\tilde{\omega}_1 \sum_{i=1}^m (T_{\min_i} - \overline{T_{\min}})\right) Y_{m+2} + \tilde{\lambda},$$

sendo m a duração da onda. As estimativas dos valores extremos ocorridos durante as ondas de frio registadas entre 1980 e 1997 são apresentadas na tabela seguinte:

Tabela 5.1.3: Valores extremos causados por uma onda de frio e suas estimativas.

Data	Nº de óbitos	Estimativa pontual do número de óbitos
26/Fev/83	19	20.89864
28/Jan/85	17	17.92792

Se admitirmos uma definição de onda de frio menos rigorosa, por exemplo, a ocorrência de pelo menos 5 dias consecutivos com $T_{\min} < 5^{\circ}\text{C}$ (o que corresponde a ter quatro ondas de frio), é possível, através da expressão anterior e com idêntica aproximação, estimar a ocorrência de valores extremos após essas ondas de frio. Na tabela seguinte apresentam-se as estimativas dos valores extremos ocorridos após as quatro ondas de frio.

Tabela 5.1.4: Valores extremos causados por uma onda de frio e suas estimativas.

Data	Número de óbitos	Estimativa pontual do número de óbitos
26/Fev/83	19	20.89864
28/Jan/85	17	17.92792
07/Mar/93	17	18.59962
23/Jan/94	17	16.24399

Observação 5.1.1. Em 1997, Baltazar Nunes (à altura, estagiário no Instituto de Saúde Dr. Ricardo Jorge) estudou o impacto das temperaturas máximas na mortalidade, no distrito de Lisboa. Para a modelação do número de óbitos diários, recorreu aos modelos ARIMA (*AutoRegressive Integreted Moving Average*) e aos modelos de Regressão Dinâmica. A escolha destes

processos, baseados em ruídos gaussianos, para a modelação de uma série de valores inteiros não-negativos, ficou a dever-se ao facto de os valores das observações no distrito de Lisboa serem bastante elevados. O autor chegou a resultados bastante satisfatórios com o modelo de Regressão Dinâmica, no sentido em que conseguiu captar a estrutura da série e ajustar o pico de mortalidade provocado pela onda de calor de 1981.

De modo a incorporar a informação proveniente da série das temperaturas máximas, o autor criou funções destas temperaturas.

- (1) Uma função que quantificasse o excesso de calor,

$$TM32_t = \begin{cases} T_{\max} - 32, & T_{\max} > 32^\circ\text{C} \\ 0, & \text{c. c.;} \end{cases}$$

- (2) Uma outra que representasse o número de dias (até ao instante de tempo t) em que a temperatura máxima se encontrou acima dos 32°C . Esta função foi designada por C_{ONDA32_t} ;

- (3) e, por último, uma função que representasse a sobrecarga térmica acumulada até ao instante de tempo t . Esta função foi definida à custa das anteriores do seguinte modo,

$$STRESS32_t = C_{ONDA32_t} \times TM32_t.$$

Estas variáveis foram também utilizadas como variáveis explicativas a incorporar no modelo DSINAR(1) generalizado. Contudo, com estas covariáveis, o método iterativo de Newton-Raphson (utilizado na obtenção das estimativas de máxima verosimilhança condicional) não converge. Mesmo modificando-se a variável $TM32_t$ (talvez seja razoável admitir, para o distrito de Évora, uma barreira acima dos 32°C) não conseguiu ultrapassar-se esta dificuldade.

5.2 Outros modelos condicionalmente Poisson

Dada a semelhança do modelo DSINAR(1) generalizado usado na modelação dos dados dos óbitos diários, e que tem a forma

$$Y_t = \alpha_t \circ^{Po} Y_{t-1} + Z_t,$$

onde $Z_t \sim \text{Po}(\lambda)$ e $\alpha_t = f(\vec{X}_t)$, isto é,

$$Y_t | Y_{t-1}, \vec{X}_t \sim \text{Po}(\alpha_t Y_{t-1} + \lambda),$$

com os modelos condicionalmente Poisson pertencentes às classes $\text{Mod}P_1(p, q)$ e $\text{Mod}P_2(p, q)$ (apresentadas no capítulo 1), pareceu-nos interessante apresentar uma análise dos dados dos óbitos (envolvendo igualmente as temperaturas mínimas e máximas diárias como covariáveis), através de dois modelos pertencentes às classes $\text{Mod}P_1(p, q)$ e $\text{Mod}P_2(p, q)$. Não se pretende, com isto, fazer um estudo da selecção do modelo mais apropriado para os dados em causa, mas apenas confrontar o modelo DSINAR(1) generalizado com dois modelos condicionalmente Poisson com duas definições diferentes de equação de estado.

Por este motivo, optámos por apresentar uma análise dos dados dos óbitos registados diariamente no distrito de Évora usando o modelo estudado por Zeger e Qaqish, em 1988, e de um modelo análogo ao modelo PAR(1), proposto por Brandt *et al.* (2000a).

A opção por utilizar o modelo estudado por Zeger e Qaqish, em 1988, em vez das alterações a este propostas por Davis *et al.* (1997, 2001), deveu-se não só ao facto de este primeiro ser mais simples de implementar, como também ao facto de as referidas alterações incidirem apenas numa correcção da equação de estado de modo a ser possível obter alguma solução estacionária para todo o valor de γ — um dos parâmetros da equação de estado (5.2.1). Como os dados com que estamos a trabalhar apresentam claramente uma não-estacionariedade não nos pareceu importante efectuar, por agora, a implementação destas alterações.

5.2.1 Modelo estudado por Zeger e Qaqish, em 1988

O modelo estudado por Zeger e Qaqish (1988) é da forma

$$Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, \vec{X}_t \sim Po(\mu_t),$$

cuja equação de estado é dada por

$$\log(\mu_t) = \sum_{i=1}^p \gamma_i Y_{t-i} + \vec{\beta} \vec{X}_t. \quad (5.2.1)$$

Para a análise dos dados dos órbitos considerou-se, à semelhança do modelo DSINAR(1) generalizado, que $p = 1$ e que $\vec{X}_t = (Tmax_{t-1}, Tmin_{t-8}^*)$.

O ajustamento deste modelo parece ser bastante pior que o do modelo DSINAR(1) generalizado. Nomeadamente, a série ajustada não acompanha as flutuações sazonais da série inicial (como se pode comprovar pelas figuras 5.2.1 e 5.2.2) e, como consequência, os gráficos dos resíduos do modelo ajustado (figuras 5.2.3 e 5.2.4) exibem igualmente um padrão sazonal. Como a série ajustada não consegue captar as principais características evidenciadas pelos dados, causa ainda o aparecimento de valores demasiado elevados nos resíduos.

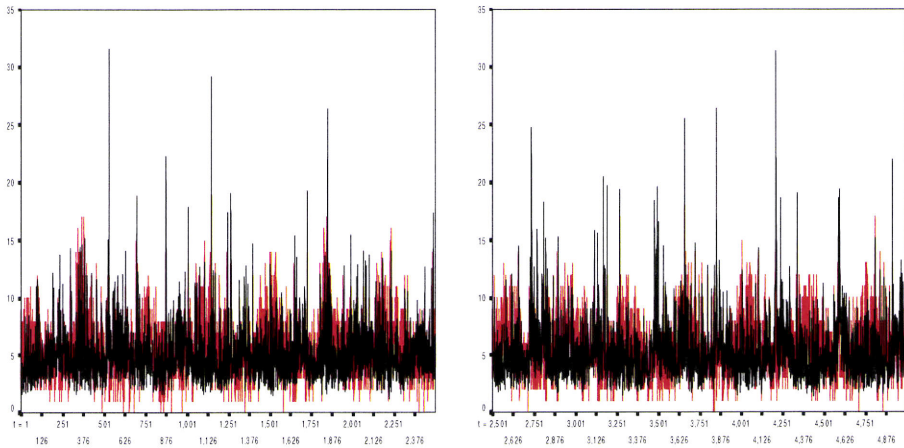


Figura 5.2.1: Série dos órbitos (a vermelho) e série estimada (a preto).

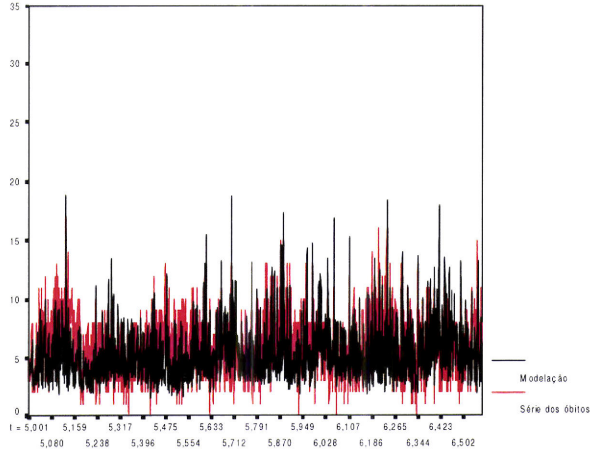


Figura 5.2.2: Série dos órbitos (a vermelho) e série estimada (a preto).

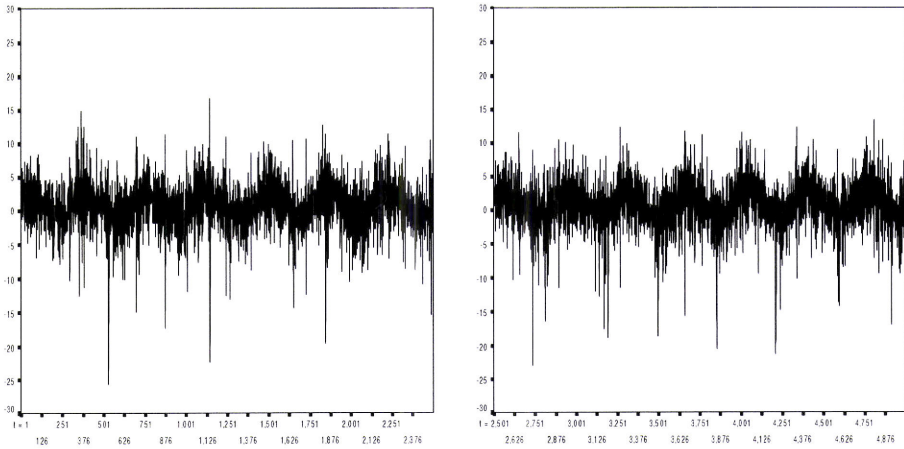


Figura 5.2.3: Resíduos do modelo ajustado.

5.2.2 Um modelo análogo ao modelo PAR(1)

Tal como foi referido anteriormente, uma segunda proposta consiste numa alteração ao modelo PAR(1), proposto por Brandt *et al.* (2000a). Uma primeira diferença entre este modelo e o modelo que vamos propor está na

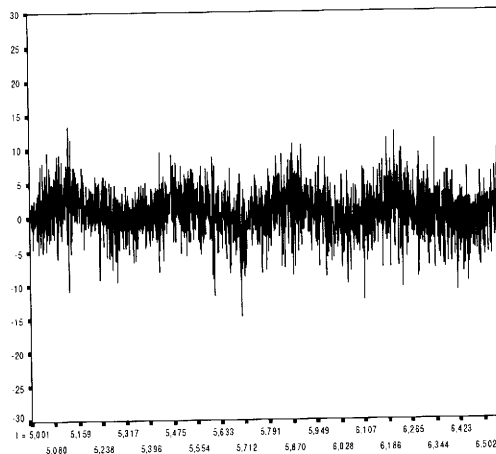


Figura 5.2.4: Resíduos do modelo ajustado.

forma de definir a equação de estado. Ou seja, o modelo por nós proposto é um modelo condicionalmente Poisson da forma

$$Y_t | Y_{t-1}, \vec{X}_t \sim Po(\mu_t),$$

cuja equação de estado é dada por

$$\mu_t = \alpha Y_{t-1} + \lambda_t,$$

sendo $\lambda_t = \exp(\vec{\omega} \vec{X}_t)$.

A segunda diferença está no facto de considerarmos que $\{\vec{X}_t\}_{t \in \mathbb{Z}}$ é um p.e. multivariado, com uma determinada estrutura de correlação, em vez de uma sucessão de v.a.'s i.i.d., tal como proposto por Brandt *et al.* (2000a).

A razão por termos optado por esta definição de equação de estado (em vez da proposta por Brandt *et al.* (2000a), *i.e.*, $\mu_t = \sum_{i=1}^p \alpha_i (Y_{t-i} - \exp(\omega \vec{X}_t)) + \exp(\omega \vec{X}_t)$) ficou a dever-se ao facto de nos ter parecido razoável admitir que o número médio de óbitos num determinado dia t fosse proporcional (com uma proporção α que se mantém constante ao longo do tempo) ao número de óbitos registado no dia anterior ($t-1$) mais o número médio de óbitos devidos à influência de determinados factores, $\{\vec{X}_t\}_{t \in \mathbb{Z}}$, observados no dia t .

Considerou-se como séries explicativas as séries das temperaturas máximas e das temperaturas mínimas, tal como no modelo DSINAR(1)

generalizado. A modelação mais satisfatória (em termos da análise dos resíduos e do coeficiente de ajustabilidade, p_{E_t}) foi obtida usando o seguinte modelo

$$Y_t | Y_{t-1}, T_{\min_{t-8}}^* \sim \text{Po}(\mu_t),$$

cuja equação de estado é dada por

$$\mu_t = \alpha Y_{t-1} + \lambda_t,$$

sendo $\lambda_t = \exp(\omega_1 T_{\min_{t-8}}^*)$. Isto é, usa-se como covariável apenas as temperaturas mínimas desfasadas de 8 dias.

A tabela seguinte contém as estimativas de máxima verosimilhança condicional dos parâmetros do modelo, obtidas a partir de um programa elaborado em linguagem FORTRAN.

Tabela 5.2.1: Estimativas de (α, ω_1) , por EMVC.

$\hat{\alpha}$	$\hat{\omega}_1$
0.92328	-0.17079

As figuras 5.2.5-5.2.8 exibem, respectivamente, o gráfico dos resíduos do modelo ajustado, o respectivo histograma e as correspondentes funções de autocorrelação e autocorrelação parcial amostrais.

A partir da sua observação constata-se que as FAC e FACP amostrais dos resíduos se assemelham às FAC e FACP de um modelo MA(1) (modelo de médias móveis de ordem 1). Como estas funções parecem sugerir que a série dos resíduos apresenta uma estrutura que poderá ser descrita por um modelo MA(1) procedemos, então, à sua modelação através do referido modelo. Nas figuras 5.2.9-5.2.12 estão representados os “novos” resíduos do modelo ajustado, o seu histograma e as correspondentes FAC e FACP amostrais. Graficamente, nada parece contrariar a hipótese de que os resíduos constituem uma sucessão de v.a.'s não-correlacionadas e de média nula.

Nas figuras 5.2.13 e 5.2.14 estão representadas a série real e a série estimada. A fim de se estudar a ajustabilidade do modelo à série temporal calculou-se ainda o quadrado do coeficiente de ajustabilidade, $p_{E_t}^2$, seguindo o mesmo procedimento que na secção anterior. Isto é, dividiram-se as séries real e estimada em períodos de Inverno e de Verão, calcularam-se a média

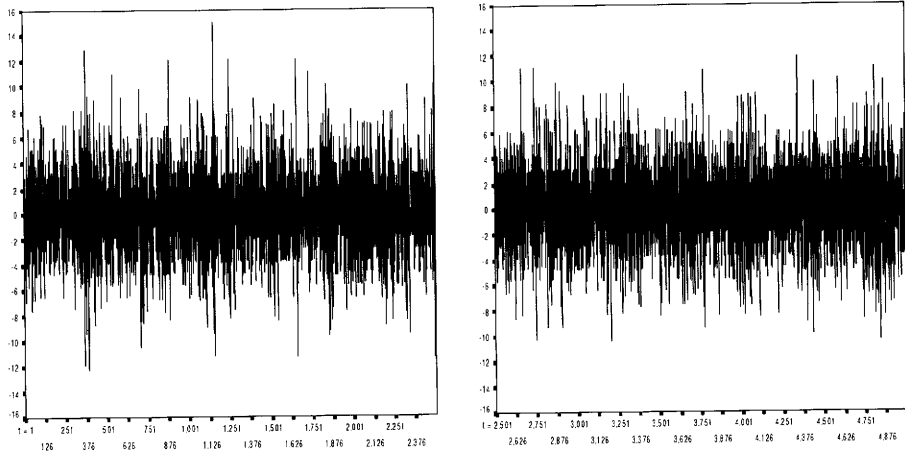


Figura 5.2.5: Resíduos do modelo ajustado.

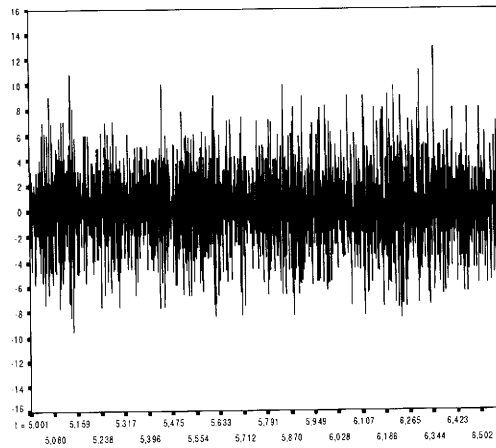


Figura 5.2.6: Resíduos do modelo ajustado.

amostral da série real e a média do erro absoluto entre a série real e a série estimada. Obteve-se, para o Inverno, um $p_{E_t}^2$ médio de 0.33495 e para o Verão, um $p_{E_t}^2$ médio de 0.27341. Estes resultados mostram que o modelo DSINAR(1) generalizado, anteriormente aplicado, parece (à luz deste coeficiente) explicar melhor a variabilidade dos dados.

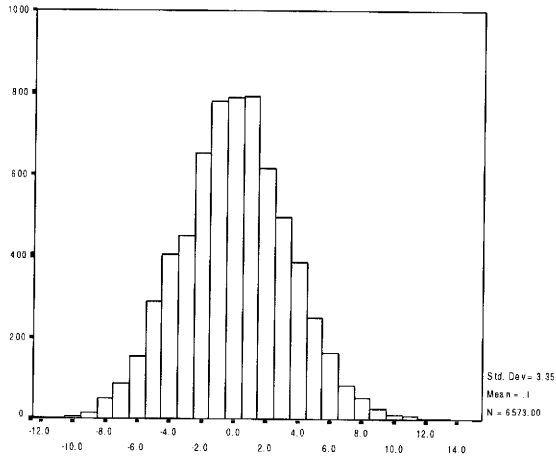


Figura 5.2.7: Histograma dos resíduos do modelo ajustado.

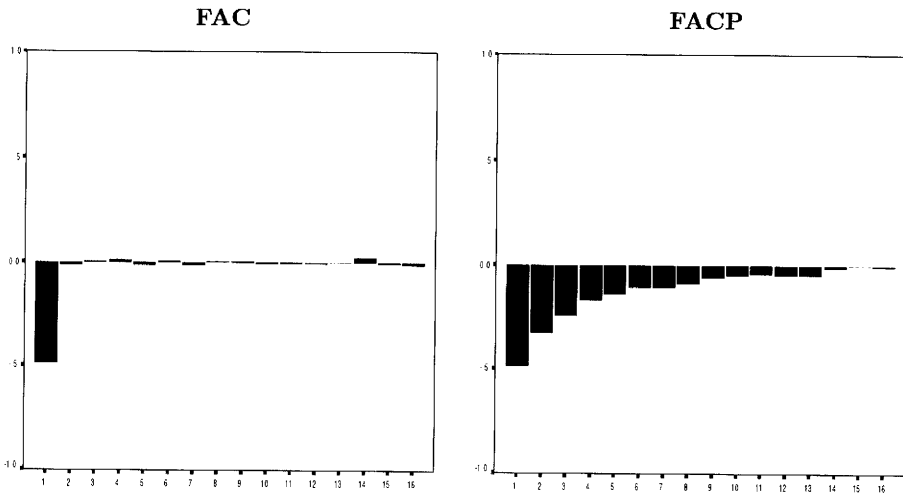


Figura 5.2.8: FAC e FACP amostrais dos resíduos do modelo ajustado.

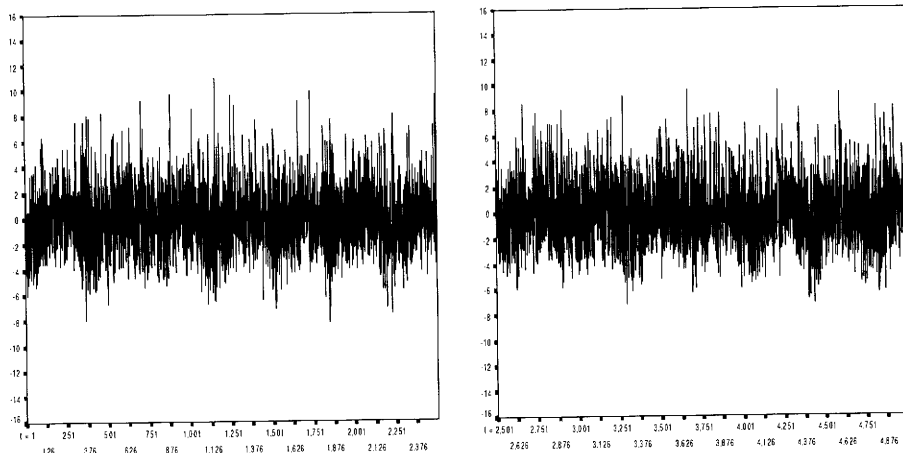


Figura 5.2.9: “Novos” resíduos do modelo ajustado.

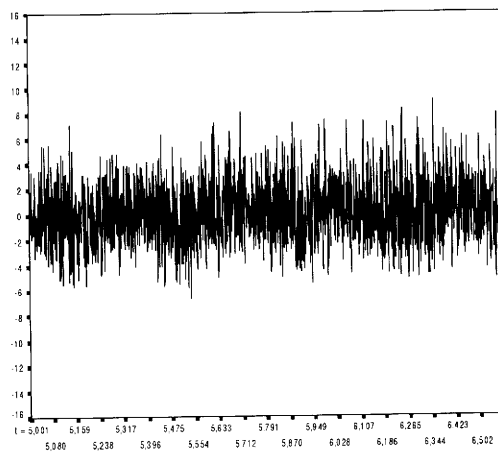


Figura 5.2.10: “Novos” resíduos do modelo ajustado.

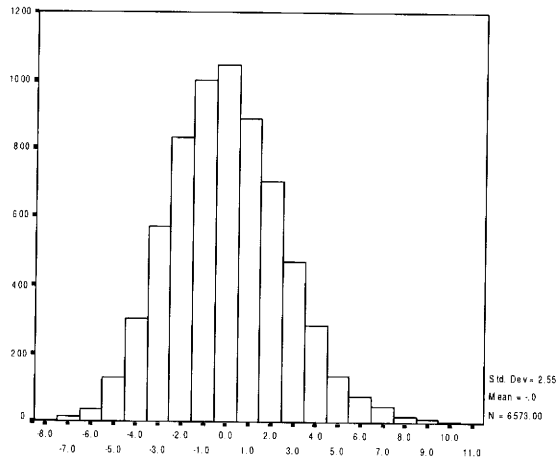


Figura 5.2.11: Histograma dos “novos” resíduos do modelo ajustado.

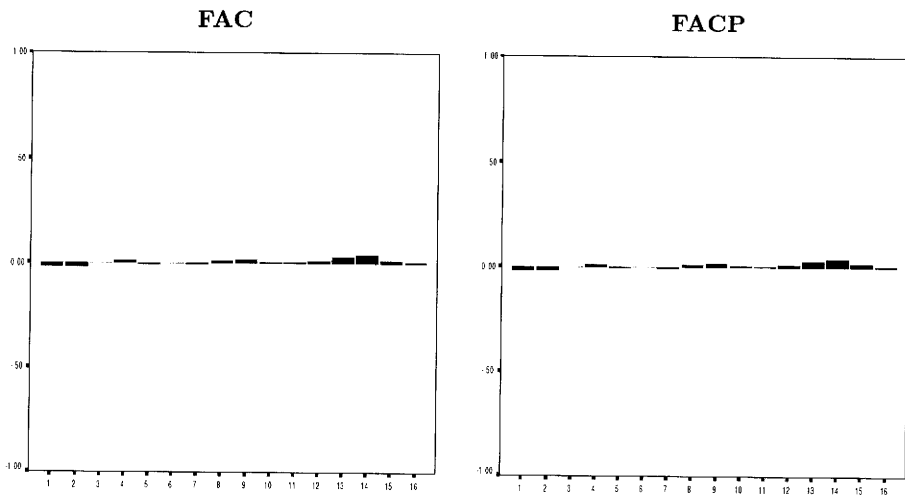


Figura 5.2.12: FAC e FACP amostrais dos “novos” resíduos do modelo ajustado.

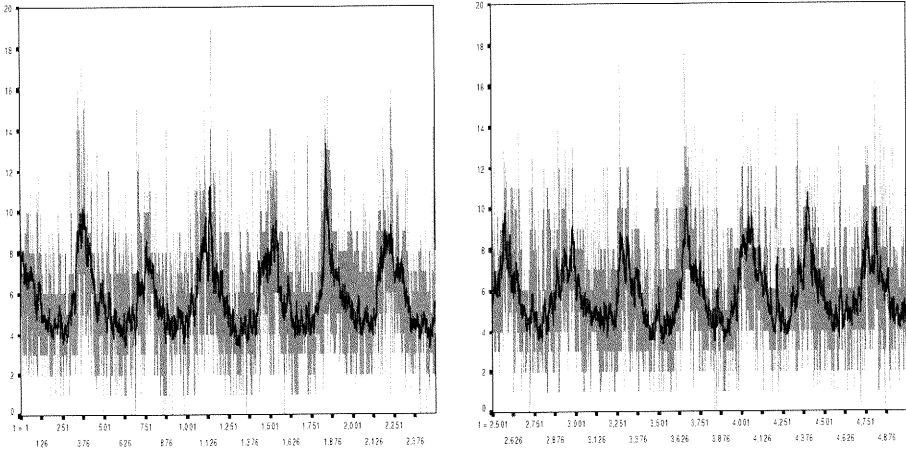


Figura 5.2.13: Série dos órbitos (a claro) e série estimada (a escuro).

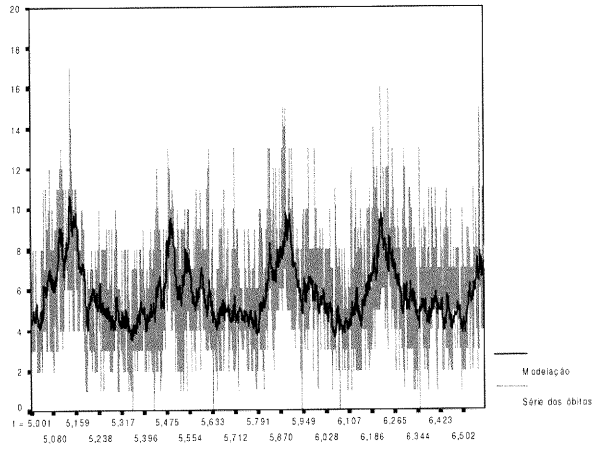


Figura 5.2.14: Série dos órbitos (a claro) e série estimada (a escuro).

Capítulo 6

Conclusões e sugestões para trabalho futuro

Neste trabalho foram estudados modelos de séries temporais de valores inteiros não-negativos que satisfazem a equação às diferenças duplamente estocástica

$$Y_t = \alpha_t \circ^G Y_{t-1} + Z_t, \quad t \in \mathbb{Z},$$

onde $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{Z_t\}_{t \in \mathbb{Z}}$ são duas sucessões de v.a.'s mutuamente independentes e \circ^G é um operador aleatório, denominado operador *thinning* generalizado.

Numa primeira fase considerou-se $\{\alpha_t\}_{t \in \mathbb{Z}}$ como um p.e. com segundos momentos finitos. O processo $\{Y_t\}_{t \in \mathbb{Z}}$ com esta característica, e admitindo a representação anterior, foi denominado DSINAR(1) generalizado (*Doubly Stochastic INteger AutoRegressive*). Foram estudadas as condições de existência e estacionariedade fraca (ou de 2ª ordem) de processos admitindo tal representação. Provou-se, também, que um processo estacionário de 2ª ordem solução da equação às diferenças duplamente estocástica tem, necessariamente, uma forma causal.

Da família dos modelos DSINAR(1) generalizados estudou-se o caso particular em que o operador *thinning* generalizado é especificado através da distribuição Binomial (denominado simplesmente processo DSINAR(1)). Provou-se que as condições necessárias e suficientes para a estacionariedade fraca dos processos auto-regressivos de ordem 1, de coeficientes estocásticos, são as mesmas quer tenham suporte em \mathbb{Z}^+ (processos DSINAR(1)) ou em

\mathbb{R} (*Doubly Stochastic AutoRegressive process of order one*, abreviadamente DSAR(1)).

Numa segunda fase, e à semelhança dos modelos auto-regressivos de coeficientes aleatórios e com suporte em \mathbb{R} (RCA), estudados, por exemplo, por Nicholls e Quinn (1982), considerou-se $\{\alpha_t\}_{t \in \mathbb{Z}}$ como uma sucessão de v.a.'s i.i.d. com segundos momentos finitos. O processo $\{Y_t\}_{t \in \mathbb{Z}}$ com esta característica e admitindo a representação anterior foi denominado RCINAR(1) generalizado (*Random Coefficient INteger AutoRegressive*).

Tal como para o processo anterior, foram igualmente estudadas as condições de existência e estacionariedade fraca de processos admitindo tal representação. Uma conclusão importante a que se chegou é que um processo estacionário de 2ª ordem solução da equação $Y_t = \phi_t Y_{t-1} + \epsilon_t$ — para qualquer função de α_t , digamos $\phi_t = f(\alpha_t)$, que verifique a condição $E[\phi_t] = E[\alpha_t]$ — é necessariamente da forma

$$Y_t = \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad t \in \mathbb{Z},$$

onde $\epsilon_t = \phi_t Y_{t-1} - \alpha_t \circ^G Y_{t-1} + Z_t$ e

$$\beta_{t,i} = \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j}, & i \geq 1 \\ 1, & i = 0. \end{cases}$$

Quando $\phi_t = \alpha_t$ temos a anterior representação causal, com $\epsilon_t = \alpha_t Y_{t-1} - \alpha_t \circ^G Y_{t-1} + Z_t$. Outra solução estacionária e causal surge quando $\phi_t = E[\alpha_t]$. Neste caso, a solução causal pode escrever-se como

$$Y_t = \mu_Y + \sum_{i=0}^{\infty} (E[\alpha_t])^i \xi_{t-i}, \quad t \in \mathbb{Z},$$

onde $\{\xi_t\}_{t \in \mathbb{Z}}$ é um processo ruído branco de média nula e variância $\sigma_\xi^2 > 0$.

Por outras palavras, o processo RCINAR(1) generalizado é um processo AR(1) de coeficiente $E[\alpha_t]$ e com inovações ruído branco. Assim, de entre todas as representações causais do processo RCINAR(1) generalizado, e com base apenas na informação fornecida pelo processo $\{Y_t\}_{t \in \mathbb{Z}}$, a representação com $\phi_t = E[\alpha_t]$ é a única cujas inovações correspondem aos erros de predição linear. Este facto permite o estudo do processo RCINAR(1) generalizado com base apenas na informação fornecida por Y_t .

A partir do estudo dos modelos RCINAR(1) generalizados torna-se possível estabelecer um paralelismo entre estes e os processos RCA(1). Ou seja, pode afirmar-se que ambos são estacionários de 2ª ordem se $E[\alpha_t^2] < 1$ e ambos são processos AR(1) de coeficiente $E[\alpha_t]$ e inovações, $\{\xi_t\}_{t \in \mathbb{Z}}$, ruído branco de média nula e variância σ_ξ^2 , finita.

Outro resultado importante refere-se aos processos RCINAR(1) generalizados, nos quais $\{\alpha_t\}_{t \in \mathbb{Z}}$ representa uma sucessão de v.a.'s i.i.d. com suporte em $(0, 1)$. Para estes casos, $\alpha_t \in (0, 1)$ é uma condição suficiente para que os processos RCINAR(1) generalizados sejam estacionários de 2ª ordem. Como caso particular de processos RCINAR(1) generalizados com esta característica salienta-se o caso do processo RCINAR(1) — processo cuja operação *thinning* generalizada é especificada através da distribuição Binomial.

Após termos resumido os resultados e as conclusões que, a nosso ver, nos pareceram mais importantes relativamente aos processos DSINAR(1) e RCINAR(1) generalizados, vamos agora abordar três questões, em três tópicos fundamentais na modelação de séries temporais, a desenvolver em trabalho futuro. Nomeadamente, a identificação do modelo, a estimação dos parâmetros desconhecidos e a validação do modelo ajustado.

No que diz respeito ao primeiro tópico, estudou-se, por simulação, o comportamento das funções de autocorrelação e de autocorrelação parcial do processo DSINAR(1) generalizado. Tal ficou a dever-se ao facto de estas poderem servir de indicador sobre a presença de um processo estacionário, ou não-estacionário, de ordem 1. Para realizações de processos DSINAR(1) generalizados estacionários observou-se que as funções de autocorrelação amostrais decaem para zero de um modo relativamente rápido (não tão rápido quanto nos modelos AR(p) estacionários) e as funções de autocorrelação parcial amostrais apresentam um decaimento brusco para zero a partir do *lag* 2. No caso não-estacionário, as funções de autocorrelação amostrais decaem para zero muito lentamente e as funções de autocorrelação parcial amostrais apresentam, ainda, um decaimento para zero a partir do *lag* 2, mas já não tão brusco como no caso anterior.

Tal como foi referido, estas características podem servir-nos como indicador sobre a apropriação, ou não, de um dado modelo pertencente à família dos modelos DSINAR(1) generalizados, para a modelação de uma dada série temporal. Através dos estudos por simulação efectuados verificaram-se ainda ligeiras diferenças entre as funções de autocorrelação e autocorrelação parcial amostrais dos dois modelos (pertencentes à família

dos modelos DSINAR(1) generalizados) estudados. Como trabalho futuro seria importante estudar-se o comportamento das funções de autocorrelação e autocorrelação parcial do processo DSINAR(1) generalizado de modo a perceber se as correspondentes funções empíricas poderão servir como uma indicação inicial da apropriação de um dado modelo, pertencente a essa família.

Relativamente à estimação dos parâmetros desconhecidos dos modelos DSINAR(1) e RCINAR(1) generalizados deixa-se, como proposta de trabalho futuro, o estudo das propriedades assintóticas dos estimadores de máxima verosimilhança condicional destes parâmetros. Apesar da dificuldade em estabelecer-se estas propriedades, os estudos por simulação levados a cabo revelaram que deve ser possível provar que estes estimadores dos parâmetros dos processos são fracamente consistentes, no sentido em que convergem em média e são assintoticamente centrados e normais.

Por último, e relativamente à análise da qualidade do ajustamento do modelo, muito trabalho ainda se encontra por desenvolver — por exemplo, no estudo de critérios adequados para a selecção da ordem do modelo DSINAR(p) generalizado, mais concretamente no estudo de um estimador da informação (ou distância) de Kullback-Leibler para este tipo de modelos, à semelhança dos estudos de Akaike, de 1973, para processos lineares gaussianos.

Mais vasto do que isto, pensamos que futuramente seria também interessante estudarem-se métodos de selecção do *melhor* modelo auto-projectivo explicativo de valores inteiros não-negativos (entre os quais se encontra o modelo DSINAR(1) generalizado), para um dado conjunto de dados em causa. Isto é, modelos da forma

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, \vec{X}_t, \vec{X}_{t-1}, \dots, \vec{X}_{t-q}, \epsilon_t),$$

em que se assume que o processo em causa no instante t é função dos seus valores passados, até ao instante $t - p$, de um processo explicativo multivariado $(\vec{X}_t, \vec{X}_{t-1}, \dots, \vec{X}_{t-q})$ (com p e q valores inteiros não-negativos) e de uma perturbação aleatória ϵ_t .

A última parte deste trabalho foi dedicada à modelação do número de óbitos registado diariamente na Conservatória do Registo Civil de Évora entre 1980 e 1997, através do modelo DSINAR(1) generalizado, usando como covariáveis as séries das temperaturas máximas e mínimas diárias registadas nesse distrito. Esta aplicação surgiu na sequência de estudos, nacionais e internacionais, sobre o impacto das alterações climáticas e das

condições climatéricas na saúde (em particular, nos estudos levados a cabo pelo Instituto Nacional de Saúde Dr. Ricardo Jorge).

O modelo adoptado revelou-se bastante satisfatório para os dados em causa, uma vez que consegue captar adequadamente as principais características evidenciadas por estes. Pensamos, contudo, que seria importante introduzir no modelo novas variáveis atmosféricas como, por exemplo, a humidade, a poluição do ar ou, até mesmo, a velocidade do vento. O conhecimento das causas de morte, por sexo e escalão etário, seria outro aspecto importante a introduzir neste estudo, nomeadamente na explicação do impacto na mortalidade das ondas de calor e de frio. Mais concretamente, no impacto na mortalidade quase imediato das primeiras e no impacto, passados alguns dias, das segundas.

Por último, pensamos que seria de todo importante alargar este estudo mais aprofundado a outras áreas geográficas nacionais, bem como comparar os seus resultados com os estudos efectuados noutros países. Só através de uma boa compreensão sobre o impacto das condições atmosféricas na mortalidade é possível adoptarem-se boas políticas de prevenção.

Apêndice A

Processo DSAR(1)

Seja $\{(\epsilon_t, \alpha_t)\}_{t \in \mathbb{Z}}$ um processo estocástico (p.e.) definido sobre um espaço de probabilidade (Ω, \mathcal{A}, P) e com valores em \mathbb{R}^2 . Defina-se o p.e. $\{Y_t\}_{t \in \mathbb{Z}}$, pela equação às diferenças estocástica

$$Y_t = \alpha_t Y_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}. \quad (\text{A.0.1})$$

Este processo foi estudado por diversos autores como, por exemplo, Tjøstheim (1986), Brandt (1986), Guyton (1986), Pourahamadi (1986, 1988) e Karlsen (1990). Pourahamadi (1986) denominou-o como *Doubly Stochastic AutoRegressive process of order one*, abreviadamente DSAR(1). Estes autores utilizaram abordagens diferentes no seu estudo, nomeadamente no estudo da existência de solução estacionária. Estas diferenças referem-se essencialmente à estacionariedade forte ou fraca do processo $\{\alpha_t\}_{t \in \mathbb{Z}}$ e às relações de independência entre este processo e o processo $\{\epsilon_t\}_{t \in \mathbb{Z}}$. Contudo, todos estes autores consideram $\{\epsilon_t\}_{t \in \mathbb{Z}}$ um p.e. estritamente estacionário.

Relativamente à segunda diferença, Guyton (1986) e Pourahamadi (1986, 1988) consideraram que $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são dois p.e.'s independentes, Tjøstheim (1986) considerou $\{\epsilon_{t+s}\}_{t,s \in \mathbb{Z}}$ independente de $\{\alpha_t, \alpha_{t-1}, \dots\}_{t \in \mathbb{Z}}$, para $s > 0$, enquanto que Brandt (1986) e Karlsen (1990) não supuseram qualquer condição de independência.

O processo DSAR(1) constitui uma generalização dos processos auto-regressivos de ordem 1 de coeficientes aleatórios (*RCA*) (estudados, por exemplo, por Āndel (1976)^[1], Vervaat (1979), Nicholls e Quinn (1980, 1982)),

^[1]Citado em Tjøstheim (1986)

onde $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são consideradas sucessões de v.a.'s i.i.d. mutuamente independentes.

Considerando que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é um p.e. estacionário de 2ª ordem, $\{\epsilon_t\}_{t \in \mathbb{Z}}$ é um p.e. estritamente estacionário e que $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são dois p.e.'s mutuamente independentes, os vários autores provam que se

$$\epsilon_t + \sum_{i=1}^{n-1} \left(\prod_{j=0}^{i-1} \alpha_{t-j} \right) \epsilon_{t-i}$$

converge em L^2 , quando $n \rightarrow \infty$, então

$$\begin{aligned} Y_t &= \epsilon_t + \sum_{i=1}^{\infty} \left(\prod_{j=0}^{i-1} \alpha_{t-j} \right) \epsilon_{t-i} \\ &= \sum_{i=0}^{\infty} \beta_{t,i} \epsilon_{t-i}, \quad t \in \mathbb{Z}, \end{aligned} \tag{A.0.2}$$

é uma solução estacionária de 2ª ordem de (A.0.1). Isto é, se $E[\epsilon_t^2] < \infty$ e

$$\sum_{i=0}^{\infty} E[\beta_{t,i}^2] < \infty$$

o processo (A.0.1) tem (A.0.2) como solução estacionária.

Brandt (1986) e Karlsen (1990) deduzem condições de estacionariedade mais gerais ao não suporem quaisquer condições de independência entre os processo $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$.

Por outro lado, por exemplo, Karlsen (1990) prova que quando $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são p.e.'s estritamente estacionários (como já foi referido, nenhuma suposição de independência é estabelecida entre estes dois p.e.'s), se

$$\epsilon_t + \sum_{i=1}^{n-1} \left(\prod_{j=0}^{i-1} \alpha_{t-j} \right) \epsilon_{t-i}$$

converge em L^p , $p \in [0, \infty]$, quando $n \rightarrow \infty$, então (A.0.2) é uma solução estritamente estacionária de (A.0.1), ou seja, se

$$\left(E[|\epsilon_t|^p] \right)^{1/p} \sum_{i=0}^{\infty} (E[|\beta_{t,i}|^p])^{k_p/p} < \infty,$$

em que $k_p = \min(1, p)$, o processo (A.0.1) tem (A.0.2) como solução estritamente estacionária.

Conclui-se, assim, que o processo DSAR(1) é estritamente estacionário ou estacionário de 2ª ordem consoante o tipo de convergência de $\{\beta_{t,i}; i \geq 0\}_{t \in \mathbb{Z}}$.

A partir da solução do processo DSAR(1), $\{Y_t\}_{t \in \mathbb{Z}}$, — admitindo que $\{\alpha_t\}_{t \in \mathbb{Z}}$ é um p.e. estacionário de 2ª ordem e que $\{\alpha_t\}_{t \in \mathbb{Z}}$ e $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são dois p.e.'s mutuamente independentes — obtivemos, respectivamente, as seguintes expressões para a média, variância e covariância do processo $\{Y_t\}_{t \in \mathbb{Z}}$:

$$E[Y_t] = \mu_\epsilon \sum_{i=0}^{\infty} E[\beta_{t,i}], \quad (\text{A.0.3})$$

$$Var[Y_t] = \mu_\epsilon^2 Var \left[\sum_{i=0}^{\infty} \beta_{t,i} \right] + \sigma_\epsilon^2 \sum_{i=0}^{\infty} E[\beta_{t,i}^2] \quad (\text{A.0.4})$$

$$\gamma_Y(k) = \mu_\epsilon^2 Cov \left(\sum_{i=0}^{\infty} \beta_{t,i}; \sum_{j=0}^{\infty} \beta_{t+k,j} \right) + \sigma_\epsilon^2 \sum_{i=0}^{\infty} E[\beta_{t,i} \beta_{t+k,i+k}], \quad k \geq 0 \quad (\text{A.0.5})$$

A partir destas expressões obtêm-se facilmente as expressões formuladas por Pourahmadi (1988), ao considerar $\{\alpha_t\}_{t \in \mathbb{Z}}$ como um p.e. estritamente estacionário.

Apêndice B

Método de Newton-Raphson na estimação dos parâmetros do processo DSINAR(1)

A título exemplificativo, vamos descrever como foram obtidas as soluções das equações de verosimilhança do processo DSINAR(1) — *i.e.*, quando no processo DSINAR(1) generalizado se considera a distribuição Binomial na especificação da operação *thinning* generalizada — com ruídos Poissonianos de parâmetro λ . Para quaisquer outras distribuições do tipo discreto, o procedimento é muito análogo.

Considere-se que $\{\vec{X}_t\}_{t \in \mathbb{Z}}$ é um p.e. multivariado com função densidade conjunta, $h(\vec{\theta}_3)$, sendo $\vec{\theta}_3$ o vector de parâmetros de $\{\vec{X}_t\}_{t \in \mathbb{Z}}$.

Assim, sendo $\{(\vec{X}_t, Y_t), t = 1, \dots, n\}$ uma a.a. dos processos $\{\vec{X}_t\}_{t \in \mathbb{Z}}$ e $\{Y_t\}_{t \in \mathbb{Z}}$, tem-se a seguinte expressão para a função de verosimilhança condicional do processo DSINAR(1), $L(\vec{\theta}) = L(\vec{Y}_n, \vec{X}_n; \vec{\theta})$ (com $\vec{\theta} = (\vec{\theta}_1, \lambda, \vec{\theta}_3)$ e $\vec{\theta}_1$ o vector de parâmetros associado à função $\alpha_t = f(\vec{X}_t)$),

$$L(\vec{\theta}) = \prod_{t=1}^{n-1} g_{t+1}(Y_t, Y_{t+1}, \vec{X}_{t+1}; \omega, \lambda) h(\vec{\theta}_3),$$

onde

$$g_{t+1}(Y_t, Y_{t+1}, \vec{X}_{t+1}; \omega, \lambda) = \sum_{m=0}^{k_2(t)} \binom{Y_t}{m} (\alpha_{t+1})^m (1 - \alpha_{t+1})^{Y_t - m} \frac{e^{-\lambda} \lambda^{Y_{t+1} - m}}{(Y_{t+1} - m)!},$$

$k_2(t) = \min(Y_{t+1}, Y_t)$ e $\alpha_{t+1} := \alpha_{t+1} | \vec{X}_{t+1}$, tal que $\alpha_{t+1} \in (0, 1)$. Mais concretamente, considera-se

$$\alpha_{t+1} = \frac{\exp(\vec{\theta}_1 \vec{X}_{t+1})}{1 + \exp(\vec{\theta}_1 \vec{X}_{t+1})}.$$

Após breves cálculos chega-se à seguinte expressão para o logaritmo da função de verosimilhança

$$\begin{aligned} \text{Log}(L(\vec{\theta})) &= \text{Log}(L_{\vec{X}_t}(\vec{\theta}_3)) - \lambda(n-1) + \\ &+ \sum_{t=1}^{n-1} \text{Log} \left[(Y_t)! \sum_{m=0}^{k_2(t)} \frac{1}{m!(Y_t-m)!} (\alpha_{t+1})^m (1-\alpha_{t+1})^{Y_t-m} \frac{\lambda^{Y_{t+1}-m}}{(Y_{t+1}-m)!} \right]. \end{aligned}$$

Os estimadores de máxima verosimilhança dos parâmetros do processo DSINAR(1), $(\vec{\theta}_1, \lambda)$, são, portanto, solução das seguintes equações de verosimilhança

$$\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \vec{\theta}_1} = 0 \quad \text{e} \quad \frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \lambda} = 0.$$

Começemos por considerar o caso em que $\{X_t\}_{t \in \mathbb{Z}}$ é um p.e. univariado, com $\vec{\theta}_1 = \omega$, o parâmetro a estimar.

Assim, as correspondentes soluções das equações de verosimilhança vão ser obtidas, tal como foi referido nos capítulos 2 e 4, através de métodos numéricos, nomeadamente através do método de Newton-Raphson, em que os valores de (ω, λ) são obtidos de um modo recursivo através da seguinte relação

$$(\omega_{i+1}, \lambda_{i+1}) = (\omega_i, \lambda_i) - J^{-1}(\omega_0, \lambda_0) \left(\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \omega}(\omega_i, \lambda_i), \frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \lambda}(\omega_i, \lambda_i) \right),$$

$i = 0, 1, \dots$, onde

$$J = \begin{bmatrix} \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega^2} & \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega \partial \lambda} \\ \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda \partial \omega} & \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda^2} \end{bmatrix}.$$

Os valores iniciais, (ω_0, λ_0) , para o método iterativo de maximização são os expostos nos capítulos 2 e 4. Considerou-se, ainda, o seguinte erro de paragem: $|\omega_i - \omega_{i+1}| \leq 10^{-5}$ e $|\lambda_i - \lambda_{i+1}| \leq 10^{-5}$, onde o índice i designa a iteração.

Com o intuito de enunciarmos, de um modo mais simplificado, as expressões das derivadas do logaritmo da função de verosimilhança, em ordem a cada um dos parâmetros, vamos definir algumas funções auxiliares. Designemos, então, por $S_1(t)$ a soma

$$S_1(t) = \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m (1 - \alpha_{t+1})^{Y_t - m} \lambda^{Y_{t+1} - m}}{m!(Y_t - m)!(Y_{t+1} - m)!}$$

e por $S_2(t)$ a soma

$$S_2(t) = \sum_{m=0}^{k_2(t)} \frac{m(\alpha_{t+1})^{m-1} (1 - \alpha_{t+1})^{Y_t - m} \lambda^{Y_{t+1} - m}}{m!(Y_t - m)!(Y_{t+1} - m)!} - \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m (1 - \alpha_{t+1})^{Y_t - m - 1} (Y_t - m) \lambda^{Y_{t+1} - m}}{m!(Y_t - m)!(Y_{t+1} - m)!}.$$

De modo a obterem-se as componentes da matriz jacobiana da função $f(\vec{\theta}) = \left(\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \omega}, \frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \lambda} \right)$, vamos proceder ao cálculo das derivadas de $S_1(t)$ e $S_2(t)$, em ordem aos parâmetros ω e λ . Começando pela soma $S_1(t)$, prova-se, facilmente, que

$$\frac{\partial S_1(t)}{\partial \omega} = S_2(t) \times \frac{\partial \alpha_{t+1}}{\partial \omega},$$

onde

$$\frac{\partial \alpha_{t+1}}{\partial \omega} = \frac{X_{t+1} \exp(\omega X_{t+1})}{[1 + \exp(\omega X_{t+1})]^2},$$

e que

$$\frac{\partial S_1(t)}{\partial \lambda} = \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m (1 - \alpha_{t+1})^{Y_t - m} (Y_{t+1} - m) \lambda^{Y_{t+1} - m - 1}}{m!(Y_t - m)!(Y_{t+1} - m)!},$$

a qual designaremos por $S_3(t)$. Relativamente às derivadas parciais da soma $S_2(t)$, prova-se que

$$\begin{aligned} \frac{\partial S_2(t)}{\partial \omega} &= \left(\frac{\partial \alpha_{t+1}}{\partial \omega} \right) \left\{ \sum_{m=0}^{k_2(t)} \frac{m(m-1)(\alpha_{t+1})^{m-2}(1-\alpha_{t+1})^{Y_t-m}\lambda^{Y_{t+1}-m}}{m!(Y_t-m)!(Y_{t+1}-m)!} - \right. \\ &\quad - 2 \sum_{m=0}^{k_2(t)} \frac{m(\alpha_{t+1})^{m-1}(1-\alpha_{t+1})^{Y_t-m-1}(Y_t-m)\lambda^{Y_{t+1}-m}}{m!(Y_t-m)!(Y_{t+1}-m)!} + \\ &\quad \left. + \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m(1-\alpha_{t+1})^{Y_t-m-2}(Y_t-m)(Y_t-m-1)\lambda^{Y_{t+1}-m}}{m!(Y_t-m)!(Y_{t+1}-m)!} \right\} = \\ &= \left(\frac{\partial \alpha_{t+1}}{\partial \omega} \right) S_4(t) \end{aligned}$$

e que

$$\begin{aligned} \frac{\partial S_2(t)}{\partial \lambda} &= \left(\frac{\partial \alpha_{t+1}}{\partial \omega} \right) \left\{ \sum_{m=0}^{k_2(t)} \frac{m(\alpha_{t+1})^{m-1}(1-\alpha_{t+1})^{Y_t-m}(Y_{t+1}-m)\lambda^{Y_{t+1}-m-1}}{m!(Y_t-m)!(Y_{t+1}-m)!} - \right. \\ &\quad \left. - \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m(1-\alpha_{t+1})^{Y_t-m-1}(Y_t-m)(Y_{t+1}-m)\lambda^{Y_{t+1}-m-1}}{m!(Y_t-m)!(Y_{t+1}-m)!} \right\}. \end{aligned}$$

Por último, vamos ainda enunciar as expressões das derivadas parciais da soma $S_3(t)$. Após cálculos simples, prova-se que

$$\frac{\partial S_3(t)}{\partial \omega} = \frac{\partial S_2(t)}{\partial \lambda} \frac{\partial \alpha_{t+1}}{\partial \omega}$$

e que

$$\frac{\partial S_3(t)}{\partial \lambda} = \sum_{m=0}^{k_2(t)} \frac{(\alpha_{t+1})^m(1-\alpha_{t+1})^{Y_t-m}(Y_{t+1}-m)(Y_{t+1}-m-1)\lambda^{Y_{t+1}-m-2}}{m!(Y_t-m)!(Y_{t+1}-m)!}.$$

As componentes da matriz jacobiana da função

$$f(\vec{\theta}) = \left(\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \omega}, \frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \lambda} \right) =$$

$$= \left(\sum_{t=1}^{n-1} \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega}}{S_1(t)}, -(n-1) + \sum_{t=1}^{n-1} \frac{S_3(t)}{S_1(t)} \right)$$

podem, assim, ser escritas do seguinte modo:

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega^2} = \sum_{t=1}^{n-1} \left\{ \frac{S_4(t) \left(\frac{\partial \alpha_{t+1}}{\partial \omega} \right)^2 + S_2(t) \frac{\partial^2 \alpha_{t+1}}{\partial \omega^2}}{S_1(t)} - \left[\frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega}}{S_1(t)} \right]^2 \right\},$$

onde

$$\frac{\partial^2 \alpha_{t+1}}{\partial \omega^2} = \frac{X_{t+1}^2 \exp(\omega X_{t+1}) [1 - \exp(\omega X_{t+1})]}{[1 + \exp(\omega X_{t+1})]^3},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda \partial \omega} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_2(t)}{\partial \lambda} \frac{\partial \alpha_{t+1}}{\partial \omega}}{S_1(t)} - \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega} S_3(t)}{[S_1(t)]^2} \right\},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega \partial \lambda} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_3(t)}{\partial \omega}}{S_1(t)} - \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega} S_3(t)}{[S_1(t)]^2} \right\} = \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda \partial \omega}$$

e

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda^2} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_3(t)}{\partial \lambda}}{S_1(t)} - \left[\frac{S_3(t)}{S_1(t)} \right]^2 \right\}.$$

Vamos agora considerar $\{\vec{X}_t\}_{t \in \mathbb{Z}}$ como um processo bivariado, de componentes $\{X_{1,t}\}_{t \in \mathbb{Z}}$ e $\{X_{2,t}\}_{t \in \mathbb{Z}}$. Por comodidade de exposição vamos, igualmente, centrar-nos na estimação de máxima verosimilhança condicional

dos parâmetros do modelo DSINAR(1), ou seja, na estimação de $(\omega_1, \omega_2, \lambda)$, em que

$$\alpha_{t+1} = \frac{\exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})}{1 + \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})}.$$

Recorrendo às notações anteriores, chega-se facilmente às seguintes equações de verosimilhança

$$\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \omega_1} = \sum_{t=1}^{n-1} \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_1}}{S_1(t)},$$

$$\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \omega_2} = \sum_{t=1}^{n-1} \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_2}}{S_1(t)}$$

e

$$\frac{\partial \text{Log}(L(\vec{\theta}))}{\partial \lambda} = -(n-1) + \sum_{t=1}^{n-1} \frac{S_3(t)}{S_1(t)},$$

onde

$$\frac{\partial \alpha_{t+1}}{\partial \omega_i} = \frac{X_{i,t+1} \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})}{[1 + \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})]^2}, \quad i = 1, 2.$$

As componentes da matriz jacobiana escrevem-se de um modo imediato, uma vez que a diferença entre este caso e o caso anterior reside apenas na expressão de α_{t+1} . Assim, tem-se que

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_1^2} = \sum_{t=1}^{n-1} \left\{ \frac{S_4(t) \left(\frac{\partial \alpha_{t+1}}{\partial \omega_1} \right)^2 + S_2(t) \frac{\partial^2 \alpha_{t+1}}{\partial \omega_1^2}}{S_1(t)} - \left[\frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_1}}{S_1(t)} \right]^2 \right\},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_2 \partial \omega_1} = \sum_{t=1}^{n-1} \left\{ \frac{S_4(t) \left(\frac{\partial \alpha_{t+1}}{\partial \omega_1} \right) \left(\frac{\partial \alpha_{t+1}}{\partial \omega_2} \right) + S_2(t) \frac{\partial^2 \alpha_{t+1}}{\partial \omega_2 \partial \omega_1}}{S_1(t)} - \right.$$

$$\left. - \left[\frac{S_2(t)}{S_1(t)} \right]^2 \left(\frac{\partial \alpha_{t+1}}{\partial \omega_1} \right) \left(\frac{\partial \alpha_{t+1}}{\partial \omega_2} \right) \right\} = \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_2 \partial \omega_1},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda \partial \omega_1} = \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_1 \partial \lambda} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_2(t)}{\partial \lambda} \frac{\partial \alpha_{t+1}}{\partial \omega_1}}{S_1(t)} - \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_1} S_3(t)}{[S_1(t)]^2} \right\},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_2^2} = \sum_{t=1}^{n-1} \left\{ \frac{S_4(t) \left(\frac{\partial \alpha_{t+1}}{\partial \omega_2} \right)^2 + S_2(t) \frac{\partial^2 \alpha_{t+1}}{\partial \omega_2^2}}{S_1(t)} - \left[\frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_2}}{S_1(t)} \right]^2 \right\},$$

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda \partial \omega_2} = \frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \omega_2 \partial \lambda} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_2(t)}{\partial \lambda} \frac{\partial \alpha_{t+1}}{\partial \omega_2}}{S_1(t)} - \frac{S_2(t) \frac{\partial \alpha_{t+1}}{\partial \omega_2} S_3(t)}{[S_1(t)]^2} \right\}$$

e, por último,

$$\frac{\partial^2 \text{Log}(L(\vec{\theta}))}{\partial \lambda^2} = \sum_{t=1}^{n-1} \left\{ \frac{\frac{\partial S_3(t)}{\partial \lambda}}{S_1(t)} - \left[\frac{S_3(t)}{S_1(t)} \right]^2 \right\}.$$

De modo a completar estas expressões, enunciam-se ainda as segundas derivadas parciais de $\alpha_{t+1} = f(\vec{X}_{t+1})$

$$\frac{\partial^2 \alpha_{t+1}}{\partial \omega_i \partial \omega_j} = \frac{X_{i,t+1} X_{j,t+1} \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})}{[1 + \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})]^3} - \frac{X_{i,t+1} X_{j,t+1} \exp(2\omega_1 X_{1,t+1} + 2\omega_2 X_{2,t+1})}{[1 + \exp(\omega_1 X_{1,t+1} + \omega_2 X_{2,t+1})]^3}, \quad i, j = 1, 2.$$

Referências Bibliográficas

- (2001). IPCC third assessment report - climate change 2001. Technical report, IPCC. (obtido em Maio de 2004 em: <http://www.ipcc.ch>).
- Al-Osh, M. A. e Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* 8(3), 261–275.
- Al-Osh, M. A. e Alzaid, A. A. (1988). Integer-valued moving average (INMA) process. *Statistische Hefte* 29, 281–300.
- Alzaid, A. A. e Al-Osh, M. A. (1988). First-order integer-valued autoregressive processes: Distributional and regression properties. *Statistica Neerlandica* 42, 53–61.
- Alzaid, A. A. e Al-Osh, M. A. (1990). An integer-valued p th-order autoregressive structure (INAR(p)) process. *Journal of Time Series Analysis* 27, 314–324.
- Alzaid, A. A. e Al-Osh, M. A. (1993). Some autoregressive moving average processes with generalized poisson marginal distribution. *Annals of the Institute of Statistical Mathematics* 45(2), 223–232.
- Andel, J. (1976). Autoregressive series with random parameters. *Math. Operationsforsch.u.Statistik* 7, 735–741.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307–27.
- Botelho, J. e Nogueira, P. (2004). Onda de calor de Agosto de 2003: os seus efeitos sobre a mortalidade da população Portuguesa. Technical report, Direcção de Serviços de Informação e Análise and Observatório Nacional de Saúde.
- Brandt, A. (1986). The stochastic equation with stationary coefficients. *Advances in Applied Probability* 18, 211–220.

- Brandt, P. e Williams, J. (2000). A linear poisson autoregressive model: The poisson (AR(p)). (obtido em Dezembro de 2003 em: <http://polmeth.wustl.edu/pa/PA92-164-184.pdf>).
- Brandt, P., Williams, J., Fordham, B. e Pollins, B. (2000). Dynamic modeling for persistent event count time series. (obtido em Dezembro de 2003 em: <http://psweb.sbs.ohio-state.edu/faculty/bpollins/papers/pewma3.pdf>).
- Brannas, K. (1994). Estimation and testing in integer-valued AR(1) models. *Umeå Economic Studies* 335.
- Brannas, K. (1995). Explanatory variables in the AR(1) count data model. *Umeå Economic Studies* 381.
- Brannas, K. e Hall, A. (1998). Estimation in integer-valued moving average models. *Umeå Economic studies* 477, 1–17.
- Brockwell, J. e Davis, R. A. (1996). *Introduction to Time Series and Forecasting*.
- Brockwell, P. e Davis, R. (1991). *Time Series: Theory and Methods*. (2^a ed.). Springer-Verlag: New York.
- Burrill, C. (1972). *Measure, Integration and Probability*. McGraw-Hill: New York.
- Davis, R., Dunsmuir, W. e Streett, S. (2001). Observation driven models for poisson counts. (obtido em Dezembro de 2003 em: <http://www.maths.unsw.edu.au/statistics>).
- Davis, R., Dunsmuir, W. e Wang, Y. (1997). Modelling time series of count data. Technical report, Colorado State University. (Enviado por correio pelos autores em 2001).
- Du, J.-G. e Yuan, L. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis* 11(2), 129–142.
- Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U. K. inflation. *Econometrica* 50, 987–1007.
- Feigin, P. D. e Tweedie, R. L. (1986). Random coefficient autoregressive processes: A markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis* 6(1), 1–14.
- Ferreira, A. (1993). Modelação estocástica da altura significativa de ondas do mar do norte. Trabalho Final de Curso, Instituto Superior Técnico, Univ. Técnica de Lisboa.

- Franke, J. e Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In S. Rao (Ed.), *Developments in Time Series Analysis*, pp. 310–330. T. Chapman and Hall, London.
- Freeland, K. e McCabe, B. (1998). Forecasting discrete valued low count time series.
- Fuller, W. A. (1976). *Introduction to Statistical Time Series*. John Wiley & Sons, New York.
- Guyton, D. (1986). A random parameter process for modeling and forecasting time series. *Journal of Time Series Analysis* 7(2), 105–115.
- Hall, A. (1998). *Extremos de Sucessões de Contagens*. Ph. D. thesis, Faculdade de Ciências da Universidade de Lisboa.
- Hall, P. e Heyde, C. C. (1980). *Martingale Limit Theory and its Application*. Academic Press: New York.
- Heinen, A. (2000). Modeling time series count data: An autoregressive conditional poisson model. (obtido em Dezembro de 2003 em: <http://www.cerge-ei.cz/pdf/events/papers/010222t.pdf>).
- Huynen, M., Martens, P., Schram, D., Weijenberg, P. e Junst, E. (2001). The impact of heat waves and cold spells on mortality rates in the Dutch population. (obtido em Setembro de 2004 em: <http://ehpnet1.niehs.nih.gov/docs/2001/109p463-47huynen>).
- Hwang, S. e Basawa, I. (1996). Parameter estimation for generalized random coefficient autoregressive processes. (obtido em Agosto de 2003 em: <http://citeseer.nj.nec.com/289444.html>).
- Jacobs, P. A. e Lewis, P. A. W. (1978). Discrete time series generated by mixtures. I: Correlational and runs properties. *Journal of the Royal Statistical Society, Series B* 40(1), 94–105.
- Karlsen, H. (1990). Existence of moments in a stationary stochastic difference equation. *Advances in Applied Probability* 22, 129–146.
- Klimko, L. e Nelson, P. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* 6(3), 629–642.
- Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis* 19(4), 439–455.

- Lawrance, A. J. e Lewis, P. A. W. (1977). An exponential moving-average sequence and point process (EMA1). *Journal of Applied Probability* 14, 98–113.
- McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* 21, 645–650.
- McKenzie, E. (1986). Autoregressive-moving-average processes with negative binomial and geometric marginal distributions. *Advances in Applied Probability* 18, 679–705.
- McKenzie, E. (1988). Some ARMA models for dependent sequences of poisson counts. *Advances in Applied Probability* 20, 822–835.
- McKenzie, E. (2003). Discrete variate time series. In D. N. Shanbhag e C. R. Rao (Eds.), *Handbook of Statistics*, Volume 21, pp. 573–606. Elsevier Science B. V.
- Nicholls, D. F. e Quinn, B. G. (1980). The estimation of random coefficient autoregressive models. I. *Journal of Time Series Analysis* 1(1), 37–46.
- Nicholls, D. F. e Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*. Springer-Verlag: New York.
- Nunes, B. (1997). Mortalidade em portugal: Modelação estocástica e estudo comparativo. Relatório de estágio, Faculdade de Ciências da Universidade de Lisboa e Instituto Nacional de Saúde Dr. Ricardo Jorge.
- Politis, D. e Romano, J. (1994). The stationary bootstrap. *Journal of the American Statistical Association* 89(428), 1303–1313.
- Pourahmadi, M. (1986). On stationarity of the solution of a doubly stochastic model. *Journal of Time Series Analysis* 7(2), 123–131.
- Pourahmadi, M. (1988). Stationarity of the solution of $x_t = a_t x_{t-1} + \epsilon_t$ and analysis of non-gaussian dependent random variables. *Journal of Time Series Analysis* 9(3), 225–239.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*, Volume 1. Academic Press: New York.
- Quinn, B. G. e Nicholls, D. F. (1981). The estimation of random coefficient autoregressive models. II. *Journal of Time Series Analysis* 2(3), 185–203.
- Quinn, B. G. e Nicholls, D. F. (1982). Testing for the randomness of autoregressive coefficients. *Journal of Time Series Analysis* 3(2), 123–135.

- Steutel, F. e Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability* 7, 893–99.
- Streett, S. (2000). *Some Observation Driven Models for Time Series*. Ph. D. thesis, Colorado State University. (obtido em Dezembro de 2003 em: <http://www.cgd.ucar.edu/streett/dis1.pdf>).
- Taylor, S. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* 4(2), 183–204.
- Tjøstheim, D. (1986). Some doubly stochastic time series models. *Journal of Time Series Analysis* 7(1), 51–72.
- Tjøstheim, D. (1994). Non-linear time series: A selective review. *Scandinavian Journal of Statistics* 21, 97–130.
- Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability* 11, 750–783.
- Zeger, S. (1988). A regression model for time series of counts. *Biometrika* 75, 621–629.
- Zeger, S. L. e Qaqish, B. (1988). Markov regression models for time series: a quasi-likelihood approach. *Biometrics* 44, 1019–1031.

