

UNIVERSIDADE DE ÉVORA

SÉRIES EMPARELHADAS DE ESTUDOS:
ESTRUTURA COMUM
INFERÊNCIA

Dissertação elaborada para obtenção do grau de Doutor em Matemática.

Não inclui as críticas e sugestões feitas pelo Júri.

Orientador: Maria Manuela Oliveira (Universidade de Évora)

Co-Orientador: João Tiago Praça Nunes Mexia (FCT/UNL)

Aníbal Manuel da Mota Areia

2009

UNIVERSIDADE DE ÉVORA

SÉRIES EMPARELHADAS DE ESTUDOS:
ESTRUTURA COMUM
INFERÊNCIA



170 181

Dissertação elaborada para obtenção do grau de Doutor em Matemática.

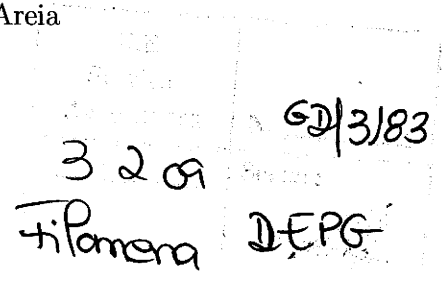
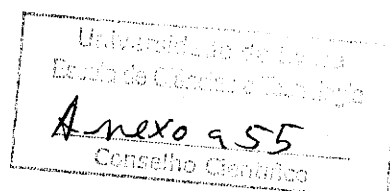
Não inclui as críticas e sugestões feitas pelo Júri.

Orientador: Maria Manuela Oliveira (Universidade de Évora)

Co-Orientador: João Tiago Praça Nunes Mexia (FCT/UNL)

Aníbal Manuel da Mota Areia

2009



Aos meus Filhos

À minha Mulher

Aos meus Pais

Agradecimentos

À Professora Doutora Maria Manuela Oliveira e ao Professor Doutor João Tiago Praça Nunes Mexia agradeço, muito reconhecido, a imprescindível orientação que sempre me concederam na preparação desta dissertação. O incentivo e apoio que me deram, assim como os seus conhecimentos científicos foram determinantes para mim na elaboração da dissertação.

RESUMO

Nesta dissertação procuramos desenvolver a teoria da inter-estrutura da metodologia STATIS para séries emparelhadas de estudos. As séries de estudos são emparelhadas quando existem bijecções entre os estudos das séries. Utilizando a abordagem de Escoufier (1973) consegue-se uma representação geométrica dos k estudos de uma série com pontos em R^k . Quando esses pontos se encontram próximos de um sub-espço com dimensão $h < k$ a série terá estrutura comum de grau h . Até agora (ver Oliveira e Mexia (1999a e 1999b), Oliveira (2002) e Ramos (2006)) apenas foram consideradas séries com estrutura comum de grau um.

A nossa abordagem assenta na teoria da "normalidade aproximada" que se apresenta no início da dissertação. Podemos assim realizar inferência com o objectivo de analisar séries individuais e séries emparelhadas de estudos. No caso das séries emparelhadas, admitimos que as mesmas correspondem aos tratamentos de um delineamento base. Mostramos como analisar a influência dos factores do delineamento base sobre as séries. Neste estudo consideramos dois tipos de análise:

- A análise longitudinal, em que se consideram medidas de evolução das séries;
- A análise transversal, em que se consideram simultaneamente estudos homólogos das várias séries.

A concluir, aplicamos os nossos resultados às eleições autárquicas de 1985 a 2001, em 16 concelhos de Portugal Continental. Estes concelhos pertencem a dois quar-

tetos de distritos, um situado ao longo da Costa e outro na fronteira com Espanha. Em cada distrito considerou-se, para além da capital, outro concelho. Teve-se assim em conta três factores: latitude, longitude e relevância administrativa.

ABSTRACT

In this dissertation, we seek to develop the theory of the STATIS methodology inter-structure for matched series of studies. The study series are matched when there are bijections between the studies of the series. Using the Escoufier (1973) approach a representation of the k studies of a series with points in R^k is obtained. When those points are close to a sub-space with an $h < k$ dimension the series will have an h degree common structure. Until now (see Oliveira and Mexia (1999a and 1999b), Oliveira (2002) and Ramos (2006)), only series of degree one common structures were considered.

Our approach rests on the theory of the “approximate normality” which is presented in the beginning of the dissertation. We can thus carry out inference with the objective of analysing individual series of studies as well as matched ones. In the case of matched series, we admit that the same correspond to the treatments of a base design. We show how to analyse the influence of the factors of the base design on the series. In this study, we consider two types of analyses:

- The longitudinal analysis, where evolution measures of the series are considered;
- The transversal analysis, where simultaneous homologous studies of the various series are considered.

Finally, we apply our results to the local elections (non-governmental elections) from 1985 to 2001 for 16 townships in Mainland Portugal. These townships belong

to two quartets of districts, one situated along the Coast and the other along the border with Spain. In each district, besides the capital, one other municipality was considered. Three factors were thus taken into account: latitude, longitude, and administrative relevance.

Lista de Símbolos e Notações

\in	pertence
\mathbf{v}^t	vector linha
\mathbf{A}^t	transposta da matriz \mathbf{A}
$vec(\mathbf{A})$	vector coluna formado pelo agrupamento dos vectores em coluna da matriz \mathbf{A}
$\mathbf{A} = [a_{ij}]$	matriz com elemento genérico a_{ij}
$tr(\mathbf{A})$	traço da matriz \mathbf{A}
$car(\mathbf{A})$	característica da matriz \mathbf{A}
\mathbf{A}^{-1}	matriz inversa de \mathbf{A}
\mathbf{A}^+	matriz inversa de Moore-Penrose de \mathbf{A}
\mathbf{I}_n	matriz identidade de ordem n
$\cdot \cdot$	produto escalar de Hilbert-Schmidt de matrizes
$\ \cdot\ $	norma Euclidiana de um vector ou de Hilbert-Schmidt para matrizes
$\xrightarrow[n \rightarrow \infty]{q.c.}$	converge quase certamente para, quando $n \rightarrow \infty$
$ \cdot $	função módulo
$E(\cdot)$	valor médio

$Var(\cdot)$	variância
$Cov(\cdot, \cdot)$	covariância
\sim	distribui-se como
$\overset{\circ}{\sim}$	distribui-se aproximadamente como
\approx	aproximadamente igual
\otimes	produto de Kronecker
$\Sigma(\cdot)$	matriz de covariância
$\mathbf{diag}(a_1, \dots, a_k)$	matriz diagonal com (a_1, \dots, a_k)
$N(\mu, \sigma^2)$	distribuição normal de valor médio μ e variância σ^2
χ_k^2	distribuição do qui-quadrado central com k graus de liberdade
$\chi_k^2(\delta)$	distribuição do qui-quadrado com k graus de liberdade e parâmetro de não centralidade δ
$F(\cdot g_1, g_2, \delta_i)$	distribuição F com g_1 e g_2 graus de liberdade e parâmetro de não centralidade δ_i

Lista de Tabelas

2.1	Estatísticas dos teste de Kolmorov-Smirnov para (P_2, L_2) .	18
2.2	Estatísticas do teste de Kolmorov-Smirnov para (P_2, L_2) .	19
2.3	Estatísticas do teste de Kolmorov-Smirnov para (P_0, L_0) .	21
2.4	Estatísticas do teste de Kolmorov-Smirnov para (P, L) e (P, L_0) .	22
2.5	Estatísticas do teste de Kolmorov-Smirnov para (P, L) e (P, L_0) .	23
2.6	Estatísticas do teste de Kolmorov-Smirnov para (P_S, L_S) e (P_D, L_D) .	25
2.7	Estatísticas do teste de Kolmorov-Smirnov para (P_S, L_S) e (P_D, L_D) .	26
2.8	Estatísticas do teste de Kolmorov-Smirnov para (P_S, L_S) e (P_D, L_D) .	27
2.9	Estatísticas do teste de Kolmorov-Smirnov para (P_S, L_S) e (P_D, L_D) .	28
3.1	Valores próprios e valores da preponderância.	46
3.2	Valores próprios e valores da preponderância.	46
3.3	Vectores de estrutura estimados.	47
3.4	Valores das estatísticas F_0 e F_1 - Componente aleatória em A_i .	48
3.5	Valores das estatísticas F_0 e F_1 - Componente aleatória em A_i .	48
3.6	Valores das estatísticas F_0 e F_1 - Componente aleatória em X_i .	54
3.7	Valores das estatísticas F_0 e F_1 - Componente aleatória em X_i .	54

4.1	Valores dos contrastes para os 16 concelhos.	77
4.2	Valores das estatísticas $F_{0i}(l)$.	78
4.3	Conjuntos associados aos factores.	81
4.4	Estatísticas $F(C)$ para efeitos do 1º, 2º e 3º factor.	82
4.5	Estatísticas $F(C)$ para efeitos da interacção entre os factores.	82
4.6	Estatísticas $F_i(C)$ para efeitos do 1º, 2º e 3º factor.	84
4.7	Estatísticas $F_i(C)$ para efeitos da interacção entre os factores.	85
4.8	Valores das estatísticas $F_{0i}(l)$.	87
4.9	Estatísticas $F(C)$ para efeitos do 1º, 2º e 3º factor.	88
4.10	Estatísticas $F(C)$ para efeitos da interacção entre os factores.	89
4.11	Estatísticas $F_i(C)$ para efeitos do 1º, 2º e 3º factor.	90
4.12	Estatísticas $F_i(C)$ para efeitos da interacção entre os factores.	91

ÍNDICE

Agradecimentos	iv
Resumo	v
Abstract	vii
Lista de Símbolos e Notações	ix
Lista de Tabelas	xi
1. Introdução	1
2. Normalidade aproximada	7
2.1 Considerações preliminares	7
2.2 Componente linear dominante	8
2.2.1 Convergência quase certa	8
2.3 Simulações	17
3. Modelos	30
3.1 Metodologia STATIS	30
3.1.1 Introdução	30
3.1.2 Inter-estrutura	31
3.2 Modelo com componente aleatória na matriz A_i	35
3.2.1 Formulação do modelo	35

3.2.2 Ajustamento do modelo	37
3.2.3 Validação do modelo	41
3.2.4 Aplicação: Eleições autárquicas em Portugal	44
3.2.4.1 Apresentação do problema	44
3.2.4.2 Ajustamento e validação do modelo	45
3.3 Modelo com componente aleatória na matriz X_i	50
3.3.1 Formulação, ajustamento e validação do modelo	50
3.3.2 Aplicação: Eleições autárquicas em Portugal	53
4. Inferência aplicada às séries de estudos	56
4.1 Modelo com componente aleatória em A_i	56
4.1.1 Considerações preliminares	56
4.1.2 Análise longitudinal para as séries individuais	59
4.1.3 Inferência para séries emparelhadas com estrutura de grau um	61
4.1.3.1 Análise longitudinal com um factor	61
4.1.3.2 Análise transversal com um factor	62
4.1.4 Inferência com vários factores e estrutura de grau um	65
4.1.4.1 Análise longitudinal	66
4.1.4.2 Análise transversal	69
4.1.5 Inferência com vários factores e estrutura com grau h	71
4.1.5.1 Análise longitudinal	71
4.1.5.2 Análise transversal	73
4.2 Aplicação: Eleições autárquicas em Portugal	76

4.2.1 Modelo com componente aleatória em A_i	76
4.2.1.1 Análise longitudinal para as séries individuais	76
4.2.1.2 Análise longitudinal e transversal sobre as séries emparelhadas	79
4.2.1.2.1 Análise longitudinal	81
4.2.1.2.2 Análise transversal	83
4.2.2 Modelo com componente aleatória em X_i	86
4.2.2.1 Análise longitudinal para as séries individuais	86
4.2.2.2 Análise longitudinal e transversal sobre as séries emparelhadas	87
4.2.2.2.1 Análise longitudinal	88
4.2.2.2.2 Análise transversal	90
5. Considerações finais	93
Bibliografia	96

CAPÍTULO 1

Introdução

Os polinómios de baixo grau em variáveis aleatórias normais e independentes com valores baixos dos coeficientes de variação têm distribuição aproximadamente normal, ver Ramos (2006) e Areia et al. (2008). Introduziu-se assim o conceito de "normalidade aproximada", que se afigurou de grande interesse, porque muitas estatísticas importantes são polinómios de baixo grau em variáveis independentes. Enquanto não se pôde admitir a normalidade dessas variáveis, a modelação com essas estatísticas era difícil, podendo agora ultrapassar-se essa dificuldade admitindo a normalidade das estatísticas, quando as observações têm valores baixos dos coeficientes de variação.

Nesta dissertação pretendemos aprofundar o estudo do conceito de "normalidade aproximada" e utilizá-la para realizar inferência na etapa inter-estrutura da metodologia STATIS ("*Structuration de Tableaux à Trois Indices de la Statistique*"). A inter-estrutura é a primeira fase do método STATIS (séries do primeiro tipo) e STATIS Dual (séries do segundo tipo). Estes métodos analisam séries de estudos em que se utilizam trios de matrizes $(\mathbf{X}_i, \mathbf{D}_n, \mathbf{D}_{p_i})$ que designamos por estudos.

Assim, um estudo é formado por uma matriz de dados, \mathbf{X}_i , com uma linha por objecto e uma coluna por variável e por duas matrizes, \mathbf{D}_n e \mathbf{D}_{p_i} , dos pesos dos objectos e das variáveis, respectivamente.

Escoufier (1973) desenvolveu uma interessante representação geométrica para os estudos destas séries. No caso das séries de estudos do primeiro tipo obteve os operadores $\mathbf{A}_i = \mathbf{X}_i \mathbf{D}_{p_i} \mathbf{X}_i^t \mathbf{D}_n$, construindo a matriz $\mathbf{S} = [s_{ij}]$, $i, j = 1, \dots, k$, obtida através dos produtos escalares de Hilbert-Schmidt, com $s_{ij} = \text{tr}(\mathbf{A}_i \mathbf{A}_j^t)$, $i, j = 1, \dots, k$. Para as séries do segundo tipo, tomou $s_{ij} = \text{tr}(\mathbf{B}_i \mathbf{B}_j^t)$, $i, j = 1, \dots, k$, com $\mathbf{B}_i = \mathbf{X}_i^t \mathbf{D}_{n_i} \mathbf{X}_i \mathbf{D}_p$, $i = 1, \dots, k$. A matriz \mathbf{S} é simétrica com valores próprios $\theta_1, \theta_2, \dots, \theta_k$, associados aos vectores próprios $\gamma_1, \gamma_2, \dots, \gamma_k$ ortonormados.

Lavit (1988) considerou que os estudos destas séries apresentam uma estrutura comum quando os pontos representativos dos estudos se encontram ao longo do primeiro eixo. Generalizaremos esta noção, dizendo que a série tem estrutura comum de grau h ($h < k$), se os pontos representativos dos respectivos estudos se encontram próximos de um sub-espço definido pela base ortonormada $\{\alpha_1, \alpha_2, \dots, \alpha_h\}$.

Oliveira e Mexia, (1999a) e (1999b), observaram que, nas séries com estrutura de grau um, o primeiro valor próprio da matriz \mathbf{S} predominava sobre os restantes, vindo

$$\mathbf{S} = \sum_{j=1}^k \theta_j \gamma_j \gamma_j^t \approx \theta_1 \gamma_1 \gamma_1^t,$$

tendo-se ainda

$$V = \|\mathbf{S} - \theta_1 \boldsymbol{\gamma}_1 \boldsymbol{\gamma}_1^t\|^2 = \sum_{j=2}^k \theta_j^2,$$

em que $\|\cdot\|$ representa a norma Euclidiana da matriz, dada pela raiz quadrada da soma dos quadrados dos seus elementos. Temos, então, o modelo da forma

$$\mathbf{S} = \boldsymbol{\lambda} \boldsymbol{\alpha} \boldsymbol{\alpha}^t + \mathbf{E},$$

onde \mathbf{E} é uma matriz aleatória tal que $\text{vec}(\mathbf{E})$ tem distribuição normal, com vector médio nulo e matriz de covariância $\sigma^2 \mathbf{I}_{k^2}$, $\text{vec}(\mathbf{E}) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{k^2})$. Este modelo mostrou-se bastante manejável, no entanto, a matriz \mathbf{S} seria apenas aproximadamente simétrica.

Posteriormente, Oliveira e Mexia (2004) tomaram $\mathbf{E} = \frac{1}{2} (\mathbf{E}^o + \mathbf{E}^{ot})$, com $\text{vec}(\mathbf{E}^o)$ normal, com vector médio nulo e matriz de covariância $\sigma^2 \mathbf{I}_{k^2}$. Conseguiu-se assim um modelo matemático mais correcto tendo, no entanto, o inconveniente de se admitir a normalidade dos elementos da matriz \mathbf{S} e não nos operadores e nem nas observações iniciais.

Ramos (2006) ultrapassou esta dificuldade, considerando a normalidade na componente aleatória das observações iniciais.

Nos modelos matemáticos desenvolvidos por Oliveira (2002) e Ramos (2006), a informação contida numa série de estudos é condensada num vector de estrutura estimado, $\tilde{\boldsymbol{\beta}} = \boldsymbol{\theta}_1 \boldsymbol{\gamma}_1$ de $\boldsymbol{\beta} = \boldsymbol{\lambda} \boldsymbol{\alpha}$ e numa soma de quadrados de resíduos.

Nesses modelos estudaram-se apenas séries de estudos com estrutura comum de grau um tendo-se realizado inferência para:

- ▶ uma série de estudos;
- ▶ séries emparelhadas, isto é, para conjuntos de séries entre cujos estudos existem correspondências biunívocas.

No presente trabalho desenvolvemos modelos matemáticos apenas para séries de estudos do primeiro tipo. Estes modelos podem ser aplicados de forma análoga às séries de estudos do segundo tipo.

Começamos por admitir que os elementos das matrizes $\mathbf{A}_i, i = 1, \dots, k$, são normais desenvolvendo, a partir daí, o modelo aplicável às séries de estudos. No desenvolvimento deste modelo consideramos:

- ▶ a "normalidade aproximada";
- ▶ a sua validação e determinação do grau da estrutura;
- ▶ que sendo $h < k$, o grau da estrutura, a informação é condensada nos vectores de estrutura estimados

$$\tilde{\beta}_j = \theta_j \gamma_j, j = 1, \dots, h,$$

e na soma dos quadrados dos resíduos

$$\left\| \mathbf{S} - \sum_{j=1}^h \theta_j \gamma_j \gamma_j^t \right\|^2.$$

Em seguida, realizamos um estudo análogo, admitindo agora que os elementos das matrizes de dados, $\mathbf{X}_i, i = 1, \dots, k$, são normais. Este modelo tem uma formulação matemática mais complexa, no entanto, a inferência baseada no mesmo revelou-se mais potente do que a anterior.

Concluído o desenvolvimento dos modelos e a sua aplicação aos resultados das

eleições, realizaremos inferência sobre as séries de estudos e sobre as séries emparelhadas de estudos a que correspondem aos tratamentos dum delineamento base. Interessará, ainda, estudar a acção dos factores que se consideram no delineamento base, sobre os vectores de estrutura. Para isso realizaram-se dois tipos de análise:

- ▶ Análise longitudinal, em que se estuda a acção dos factores sobre a evolução das séries de estudos. Para realizar esta análise, aplicamos vectores de contraste nos vectores de estrutura;

- ▶ Análise transversal, em que se estuda a acção dos factores sobre estudos homólogos, isto é, sobre estudos que se correspondem. Aos estudos homólogos corresponderão componentes com o mesmo índice dos vectores de estrutura.

Para completar o nosso trabalho incluímos uma aplicação dos modelos desenvolvidos aos resultados das eleições autárquicas no período de 1985 a 2001, em 16 concelhos de Portugal Continental.

Assim, no Capítulo 2, desenvolvemos o estudo do conceito de "normalidade aproximada" para mostrar que polinómios não lineares de baixo grau de variáveis aleatórias normais poderão ser tratadas como funções lineares de variáveis aleatórias, aprofundando os desenvolvimentos efectuados por Ramos (2006) e Areia et al. (2008), com recurso a várias simulações.

No Capítulo 3, apresentamos uma nova abordagem para a modelização da inter-estrutura de séries de estudos. Construiu-se um modelo matemático para séries de estudos de primeiro tipo, introduzindo a componente aleatória nos operadores desenvolvidos por Escoufier (1973) e nas matrizes de dados $\mathbf{X}_i, i = 1, \dots, k$. A concluir

este capítulo apresentamos uma aplicação aos resultados das eleições autárquicas de 16 concelhos de Portugal.

No Capítulo 4, desenvolvemos a inferência para estudar as séries de estudos individuais e para as séries de estudos emparelhadas associadas aos níveis dos factores, considerando as análises longitudinal e transversal. Efectuamos uma aplicação com os resultados das eleições autárquicas com a componente aleatória nos operadores de Escoufier e nas matrizes de dados $\mathbf{X}_i, i = 1, \dots, k$.

Utilizamos o Software R para as simulações sobre a "normalidade aproximada" e o Software Wolfram Research Mathematica 5.2, que permitiu efectuar os cálculos para o ajustamento dos modelos.

CAPÍTULO 2

Normalidade aproximada

2.1 Considerações preliminares

Neste capítulo vamos mostrar como polinómios de baixo grau em variáveis aleatórias normais independentes com baixos coeficientes de variação se distribuem muito aproximadamente como sendo variáveis normais.

Ramos (2006) mostrou que, quando os coeficientes de variação tendem para zero, os polinómios com coeficientes positivos têm uma componente dominante que é linear, logo normalmente distribuída.

No que se segue alargamos as condições sobre os sinais dos coeficientes quer no estudo teórico quer nas simulações que apresentaremos.

As simulações apresentadas mostram, ainda, que as distribuições de polinómios são aproximadamente normais para valores não excessivamente pequenos dos coeficientes de variação.

2.2 Componente linear dominante

No estudo da componente linear dominante utilizamos a convergência quase certa.

Apresentamos em seguida alguns resultados preliminares relativos à mesma.

2.2.1 Convergência quase certa

Dada uma sucessão $\{\mathbf{Z}_n\}$, $n \in \mathbb{N}$ de vectores aleatórios diz-se que \mathbf{Z}_n converge quase certamente para \mathbf{b} , $\left(\mathbf{Z}_n \xrightarrow[n \rightarrow \infty]{q.c.} \mathbf{b}\right)$, quando $\forall \varepsilon > 0$, se tem

$$P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} (\|\mathbf{Z}_n - \mathbf{b}\| < \varepsilon)\right) = 1,$$

isto é, sendo $A_n(\varepsilon)$ o acontecimento que se verifica quando $\|\mathbf{Z}_n - \mathbf{b}\| < \varepsilon$, o limite inferior de $\{A_n(\varepsilon)\}$, $n \in \mathbb{N}$ tem, $\forall \varepsilon > 0$, probabilidade 1. Por outras palavras, tem-se $\mathbf{Z}_n \xrightarrow[n \rightarrow \infty]{q.c.} \mathbf{b}$, se for 1 a probabilidade de se obter uma sequência de realizações dos vectores aleatórios da sequência $\{\mathbf{Z}_n\}$ com limite \mathbf{b} .

Teorema 2.1 Teorema de Slutsky

Sendo $\varphi(g)$ o conjunto de pontos de continuidade de uma função real g , se $\mathbf{b} \in \varphi(g)$

e se $\mathbf{Z}_n \xrightarrow[n \rightarrow \infty]{q.c.} \mathbf{b}$, tem-se $g(\mathbf{Z}_n) \xrightarrow[n \rightarrow \infty]{q.c.} g(\mathbf{b})$, isto é, $\forall \delta > 0$,

$$P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} (|g(\mathbf{Z}_n) - g(\mathbf{b})| < \delta)\right) = 1.$$

Nota: Este teorema dá-nos condições suficientes para $\forall \delta > 0$, se ter com probabilidade 1 uma sequência de realizações das variáveis da sequência $\{g(\mathbf{Z}_n)\}$ com limite $g(\mathbf{b})$.

Demonstração:

Representando por $A_n^o(\delta)$ o acontecimento que se verifica quando $|g(\mathbf{Z}_n) - g(\mathbf{b})| < \delta$.

Devido à continuidade de g , $\forall \delta > 0, \exists \varepsilon > 0$ tal que $A_n(\varepsilon) \implies A_n^o(\delta)$, vindo

$$\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n(\varepsilon) \implies \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^o(\delta).$$

Por hipótese $\mathbf{Z}_n \xrightarrow[n \rightarrow \infty]{q.c.} \mathbf{b}$, pelo que

$$P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n(\varepsilon)\right) = 1,$$

vindo

$$1 = P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n(\varepsilon)\right) \leq P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^o(\delta)\right) = 1,$$

o que estabelece a tese.

Lema 2.1 Qualquer que seja a variável aleatória X e $(a_n), n \in \mathbb{N}$, uma sucessão de valores reais tal que $a_n \rightarrow \infty$, tem-se $\frac{X}{a_n} \xrightarrow[n \rightarrow \infty]{q.c.} 0$.

Demonstração:

Sendo $F_X(x)$ a função distribuição da variável aleatória X , tem-se $\forall \varepsilon > 0$ e $b > 0$,

$$P\left(\left|\frac{X}{b}\right| < \varepsilon\right) = F(b\varepsilon) - F(-b\varepsilon^-).$$

Basta agora observar que, qualquer que seja $b > 0$ existe k tal que, para todo o $n > k$, se tem $a_n > b$ vindo

$$P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left|\frac{X}{a_n}\right| < \varepsilon\right) \geq P\left(\bigcap_{n=k}^{\infty} \left|\frac{X}{a_n}\right| < \varepsilon\right) \geq P\left(\left|\frac{X}{b}\right| < \varepsilon\right) = F(b\varepsilon) - F(-b\varepsilon^-),$$

e que, $F(b\varepsilon) - F(-b\varepsilon^-) \xrightarrow[b \rightarrow \infty]{} 1$, para se estabelecer a tese.

Proposição 2.1 Qualquer que seja o vector aleatório $\mathbf{X} = (X_1, X_2, \dots, X_m)$ e $\mu_{ni}, i = 1, \dots, m$, uma sucessão de valores reais, tal que $\mu_{ni} \xrightarrow[n \rightarrow \infty]{} +\infty, i = 1, \dots, m$, teremos,

$$\frac{\prod_{i=1}^m (X_i + \mu_{ni})^{r_i}}{\left(\prod_{i=1}^m \mu_{ni}^{r_i}\right) \left(1 + \sum_{i=1}^m r_i \frac{X_i}{\mu_{ni}}\right)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

qualquer que sejam os inteiros não negativos r_1, r_2, \dots, r_m .

Demonstração:

Seja

$$\begin{aligned} \prod_{i=1}^m (X_i + \mu_{ni})^{r_i} &= \prod_{i=1}^m \left[\sum_{j=1}^{r_i} \binom{r_i}{j} \mu_{ni}^{r_i-j} X_i^j \right] \\ &= \prod_{i=1}^m \mu_{ni}^{r_i} \prod_{i=1}^m \left(1 + \sum_{j=1}^{r_i} \binom{r_i}{j} \left(\frac{X_i}{\mu_{ni}} \right)^j \right) \end{aligned}$$

então

$$\frac{\prod_{i=1}^m (X_i + \mu_{ni})^{r_i}}{\prod_{i=1}^m \mu_{ni}^{r_i} \left(1 + \sum_{j=1}^m r_j \frac{X_j}{\mu_{nj}} \right)} = \frac{\prod_{i=1}^m \left(1 + \sum_{j=1}^{r_i} \binom{r_i}{j} \left(\frac{X_i}{\mu_{ni}} \right)^j \right)}{\left(1 + \sum_{j=1}^m r_j \frac{X_j}{\mu_{nj}} \right)},$$

bastando aplicar o Lema 2.1, em conjunto com o teorema de Slutsky para se estabelecer a tese.

Observe-se que se o vector aleatório \mathbf{X} for normal, sendo

$$\prod_{i=1}^m \mu_{ni}^{r_i} \left(1 + \sum_{j=1}^m r_j \frac{X_j}{\mu_{nj}} \right)$$

uma combinação linear de variáveis aleatórias com distribuição conjunta normal a mesma será normal, (ver Mexia, 1995).

Assim, quando os $\mu_{ni}, i = 1, \dots, m$, forem suficientemente grandes é de esperar que

$\prod_{i=1}^m (X_i + \mu_{ni})^{r_i}$ seja aproximadamente normal. As simulações que adiante apresentaremos confirmam esta situação.

Sejam os polinómios com os coeficientes $b(\mathbf{r}_i)$:

$$P(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{i=1}^u b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}$$

e

$$P_0(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma_0} b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}},$$

em que as m componentes dos vectores \mathbf{r}_i são inteiros não negativos. Sendo $\Gamma = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_u\}$, o conjunto desses vectores podemos definir no mesmo uma relação de ordem parcial entre os vectores, escrevendo $\mathbf{r}_i < \mathbf{r}_j$, se as componentes de \mathbf{r}_j não forem inferiores às componentes homólogas de \mathbf{r}_i e, pelo menos uma for superior. Se não existir \mathbf{r}_t tal que $\mathbf{r}_j < \mathbf{r}_t$, \mathbf{r}_j será maximal. Representamos por Γ_0 o conjunto dos vectores maximais de Γ .

Consideremos a seguinte proposição:

Proposição 2.2 Qualquer que seja o vector aleatório \mathbf{X} , se $b(\mathbf{r}_i) > 0, i = 1, \dots, u$, para todo o $\mathbf{r}_i \in \Gamma$ e se $\mu_{nj} \xrightarrow[n \rightarrow \infty]{} +\infty, j = 1, \dots, m$, tem-se

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{P_0(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Demonstração:

Tem-se $P(\mathbf{X}; \boldsymbol{\mu}_n) = P_0(\mathbf{X}; \boldsymbol{\mu}_n) + \sum_{\mathbf{r}_i \in \Gamma \setminus \Gamma_0} b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}$, ora qualquer que seja $\mathbf{r}_i \in \Gamma \setminus \Gamma_0$ existe $\mathbf{r}_{0i} \in \Gamma_0$ tal que $\mathbf{r}_i < \mathbf{r}_{0i}$ tendo-se

$$\frac{b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}}{b(\mathbf{r}_{0i}) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{0ij}}} \xrightarrow[n \rightarrow \infty]{q.c.} 0$$

e como os coeficientes são positivos

$$0 < b(\mathbf{r}_{0i}) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{0ij}} \leq P_0(\mathbf{X}; \boldsymbol{\mu}_n)$$

vindo

$$\frac{b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}}{P_0(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 0.$$

Basta agora aplicar o teorema de Slutsky para completar a demonstração.

Introduzimos agora as componentes lineares de $P(\mathbf{X}; \boldsymbol{\mu}_n)$

$$L(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{i=1}^u b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)$$

e de $P_0(\mathbf{X}; \boldsymbol{\mu}_n)$

$$L_0(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma_0} b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)$$

para estabelecer a Proposição 2.3 e o Teorema 2.2.

Proposição 2.3 Qualquer que seja o vector aleatório \mathbf{X} , se $b(\mathbf{r}_i) > 0$, $i = 1, \dots, u$,

para todo o $\mathbf{r}_i \in \Gamma$ e se $\mu_{nj} \xrightarrow[n \rightarrow \infty]{} +\infty$, $j = 1, \dots, m$, tem-se

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Demonstração:

Ora devido à Proposição 2.1, qualquer que seja $\mathbf{r}_i \in \Gamma$, tem-se

$$q_n(\mathbf{r}_i) = \frac{b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}}{b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

isto é, $P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} (|q_n(\mathbf{r}_i) - 1|) < \varepsilon\right) = 1$, $\forall \varepsilon > 0$. Sendo c o cardinal de Γ

($c = \#\Gamma$) e $Q(\mathbf{r}_i) = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} (|q_n(\mathbf{r}_i) - 1|)$ ter-se-á, atendendo às desigualdades

de Boole generalizadas

$$1 \geq P\left[\bigcap_{\mathbf{r}_i \in \Gamma} (Q(\mathbf{r}_i) < \varepsilon)\right] \geq \sum_{\mathbf{r}_i \in \Gamma} P(Q(\mathbf{r}_i) < \varepsilon) - (c - 1) = 1.$$

Basta agora observar que, devido aos coeficientes serem todos positivos, que

$$\bigcap_{\mathbf{r}_i \in \Gamma} (Q(\mathbf{r}_i) < \varepsilon)$$

implica

$$\left| \frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} - 1 \right| < \varepsilon,$$

tendo-se pois

$$P \left[\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left(\left| \frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} - 1 \right| < \varepsilon \right) \right] = 1$$

logo

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Esta Proposição corresponde ao resultado estabelecido em Ramos (2006). Observe-se ainda que se o vector aleatório \mathbf{X} é normal $L(\mathbf{X}; \boldsymbol{\mu}_n)$ será normal. As simulações, que adiante apresentaremos, mostram que quando os coeficientes de variação das componentes de $\mathbf{X} + \boldsymbol{\mu}_n$ são suficientemente pequenos, os polinómios $P(\mathbf{X}; \boldsymbol{\mu}_n)$ têm distribuição aproximadamente normal.

Para nos libertarmos da restrição quanto aos sinais dos coeficientes, vamos fraccionar $P(\mathbf{X}; \boldsymbol{\mu}_n)$ e $L(\mathbf{X}; \boldsymbol{\mu}_n)$ numa parte "positiva" e numa parte "negativa", sendo $P(\mathbf{X}; \boldsymbol{\mu}_n) = P^+(\mathbf{X}; \boldsymbol{\mu}_n) - P^-(\mathbf{X}; \boldsymbol{\mu}_n)$ e $L(\mathbf{X}; \boldsymbol{\mu}_n) = L^+(\mathbf{X}; \boldsymbol{\mu}_n) - L^-(\mathbf{X}; \boldsymbol{\mu}_n)$.

Tomemos

$$\Gamma^+ = \{\mathbf{r}_i : b(\mathbf{r}_i) > \mathbf{0}\}$$

e

$$\Gamma^- = \{\mathbf{r}_i : b(\mathbf{r}_i) < \mathbf{0}\},$$

bem como $\Gamma_0^+ = \Gamma^+ \cap \Gamma_0$ e $\Gamma_0^- = \Gamma^- \cap \Gamma_0$.

As partes "positivas" e "negativas" do polinómio $P(\mathbf{X}; \boldsymbol{\mu}_n)$ são:

$$P^+(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma^+} b(\mathbf{r}_i) \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}}$$

e

$$P^-(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma^-} |b(\mathbf{r}_i)| \prod_{j=1}^m (X_j + \mu_{nj})^{r_{ij}},$$

e da sua componente linear $L(\mathbf{X}; \boldsymbol{\mu}_n)$ são:

$$L^+(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma^+} b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)$$

e

$$L^-(\mathbf{X}; \boldsymbol{\mu}_n) = \sum_{\mathbf{r}_i \in \Gamma^-} |b(\mathbf{r}_i)| \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right).$$

Estamos agora em condições de estabelecer o seguinte teorema.

Teorema 2.2

Se $a_n = \frac{\sum_{\mathbf{r}_i \in \Gamma_0^-} |b(\mathbf{r}_i)| \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)}{\sum_{\mathbf{r}_i \in \Gamma_0^+} b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)} \xrightarrow[n \rightarrow \infty]{q.c.} 0$ e $\mu_{nj} \xrightarrow[n \rightarrow \infty]{} +\infty$, $j = 1, \dots, m$, qual-

quer que seja o vector aleatório \mathbf{X} tem-se

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Demonstração:

Começamos por admitir que o conjunto dos vectores maximais Γ_0^- não é vazio.

Atendendo à Proposição 2.3 temos que

$$\frac{P^+(\mathbf{X}; \boldsymbol{\mu}_n)}{L^+(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1$$

e

$$\frac{P^-(\mathbf{X}; \boldsymbol{\mu}_n)}{L^-(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Por outro lado, é fácil verificar, atendendo à Proposição 2.2, que

$$\frac{L^+(\mathbf{X}; \boldsymbol{\mu}_n)}{\sum_{\mathbf{r}_i \in \Gamma_0^+} b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

e

$$\frac{L^-(\mathbf{X}; \boldsymbol{\mu}_n)}{\sum_{\mathbf{r}_i \in \Gamma_0^-} |b(\mathbf{r}_i)| \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

logo, quando $a_n \rightarrow 0$, tem-se

$$\frac{L^-(\mathbf{X}; \boldsymbol{\mu}_n)}{L^+(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 0$$

e

$$\frac{P^-(\mathbf{X}; \boldsymbol{\mu}_n)}{P^+(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 0.$$

Como

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} = \frac{P^+(\mathbf{X}; \boldsymbol{\mu}_n)}{L^+(\mathbf{X}; \boldsymbol{\mu}_n)} \frac{1 - \frac{P^-(\mathbf{X}; \boldsymbol{\mu}_n)}{P^+(\mathbf{X}; \boldsymbol{\mu}_n)}}{1 - \frac{L^-(\mathbf{X}; \boldsymbol{\mu}_n)}{L^+(\mathbf{X}; \boldsymbol{\mu}_n)}}$$

e o segundo factor converge quase certamente para 1, quando $a_n \rightarrow 0$, a tese está estabelecida.

Se o conjunto dos vectores maximais $\Gamma_0^- = \emptyset$, temos $P_0^+ = P_0$ e $L_0^+ = L_0$, então

$$\frac{P_0^+(\mathbf{X}; \boldsymbol{\mu}_n)}{L_0^+(\mathbf{X}; \boldsymbol{\mu}_n)} = \frac{P_0(\mathbf{X}; \boldsymbol{\mu}_n)}{L_0(\mathbf{X}; \boldsymbol{\mu}_n)}.$$

Podemos ainda, utilizar as Proposições 2.2 e 2.3, para mostrar que:

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{P_0(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1$$

e que

$$\frac{P_0(\mathbf{X}; \boldsymbol{\mu}_n)}{L_0(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

o que implica

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1,$$

como se queria estabelecer.

Por simetria, se se tem

$$a_n = \frac{\sum_{\mathbf{r}_i \in \Gamma_0^+} b(\mathbf{r}_i) \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)}{\sum_{\mathbf{r}_i \in \Gamma_0^-} |b(\mathbf{r}_i)| \left(\prod_{j=1}^m \mu_{nj}^{r_{ij}} \right) \left(1 + \sum_{j=1}^m r_{ij} \frac{X_j}{\mu_{nj}} \right)} \xrightarrow[n \rightarrow \infty]{q.c.} 0$$

e $\mu_{nj} \xrightarrow[n \rightarrow \infty]{} +\infty$, $j = 1, \dots, m$, ter-se-à igualmente

$$\frac{P(\mathbf{X}; \boldsymbol{\mu}_n)}{L(\mathbf{X}; \boldsymbol{\mu}_n)} \xrightarrow[n \rightarrow \infty]{q.c.} 1.$$

Repetindo o que atrás se disse, observamos que quando \mathbf{X} é normal $L(\mathbf{X}; \boldsymbol{\mu}_n)$, será normal. No que se segue, começamos por considerar simulações para polinómios com coeficientes positivos, alargando de seguida o estudo a polinómios com coeficientes positivos e negativos.

2.3 Simulações

Para efectuar as simulações considerámos o polinómio $P(\mathbf{X}; \boldsymbol{\mu}_n)$ e a sua componente linear $L(\mathbf{X}; \boldsymbol{\mu}_n)$, com o vector aleatório $\mathbf{X} = (X_1, X_2)$. Através da utilização do Software R, foram geradas 1000 observações para cada variável $X_i, i = 1, 2$, tendo obtido amostras de $P(\mathbf{X}; \boldsymbol{\mu}_n)$ e de $L(\mathbf{X}; \boldsymbol{\mu}_n)$. A estas amostras aplicou-se o teste de Kolmogorov-Smirnov (K-S) para duas amostras, com um nível de significância de 0.05, cujo valor crítico é de 0.0608.

Os valores da estatística de teste que correspondem à não rejeição da hipótese de que as populações que originam as duas amostras têm a mesma distribuição encontram-se sublinhados.

■ 1.º Conjunto de simulações

Seja o polinómio

$$\blacktriangleright P_2 = (X_1 + \mu_1)^r (X_2 + \mu_2)^s$$

com as variáveis independentes $X_i \sim N(0, \sigma = 1)$, $i = 1, 2$, e a sua componente linear,

$$\blacktriangleright L_2 = \mu_1^r \mu_2^s + r \mu_1^{r-1} \mu_2^s X_1 + s \mu_1^r \mu_2^{s-1} X_2.$$

Nas tabelas 2.1 e 2.2 apresentam-se as estatísticas do teste de Kolmogorov-Smirnov (K-S) para as amostras de observações de (P_2, L_2) com $(r + s) \in \{2, 4, 6, 7\}$.

$n = 1000; (r = 1; s = 1)$								
$\mu_1 \backslash \mu_2$	1	2	5	10	20	30	50	70
1	0.117	0.088	<u>0.028</u>	<u>0.015</u>	<u>0.010</u>	<u>0.009</u>	<u>0.006</u>	<u>0.005</u>
2	0.088	0.078	<u>0.037</u>	<u>0.018</u>	<u>0.010</u>	<u>0.009</u>	<u>0.007</u>	<u>0.006</u>
5	<u>0.035</u>	<u>0.037</u>	<u>0.041</u>	<u>0.023</u>	<u>0.011</u>	<u>0.010</u>	<u>0.006</u>	<u>0.005</u>
10	<u>0.015</u>	<u>0.019</u>	<u>0.027</u>	<u>0.030</u>	<u>0.013</u>	<u>0.012</u>	<u>0.008</u>	<u>0.005</u>
20	<u>0.010</u>	<u>0.009</u>	<u>0.013</u>	<u>0.019</u>	<u>0.018</u>	<u>0.012</u>	<u>0.007</u>	<u>0.006</u>
30	<u>0.008</u>	<u>0.008</u>	<u>0.009</u>	<u>0.013</u>	<u>0.014</u>	<u>0.013</u>	<u>0.011</u>	<u>0.007</u>
50	<u>0.007</u>	<u>0.006</u>	<u>0.009</u>	<u>0.008</u>	<u>0.009</u>	<u>0.010</u>	<u>0.010</u>	<u>0.008</u>
70	<u>0.005</u>	<u>0.005</u>	<u>0.006</u>	<u>0.006</u>	<u>0.006</u>	<u>0.009</u>	<u>0.008</u>	<u>0.008</u>
$n = 1000; (r = 2; s = 2)$								
$\mu_1 \backslash \mu_2$	1	2	5	10	20	30	50	70
1	0.352	0.327	0.306	0.304	0.305	0.308	0.313	0.313
2	0.324	0.243	0.189	0.168	0.164	0.169	0.166	0.166
5	0.304	0.188	0.076	<u>0.058</u>	<u>0.055</u>	<u>0.053</u>	<u>0.053</u>	<u>0.053</u>
10	0.302	0.173	<u>0.050</u>	<u>0.037</u>	<u>0.027</u>	<u>0.027</u>	<u>0.033</u>	<u>0.031</u>
20	0.304	0.178	<u>0.046</u>	<u>0.029</u>	<u>0.021</u>	<u>0.019</u>	<u>0.019</u>	<u>0.015</u>
30	0.305	0.176	<u>0.045</u>	<u>0.026</u>	<u>0.021</u>	<u>0.016</u>	<u>0.014</u>	<u>0.013</u>
50	0.301	0.174	<u>0.047</u>	<u>0.025</u>	<u>0.019</u>	<u>0.016</u>	<u>0.012</u>	<u>0.010</u>
70	0.299	0.174	<u>0.049</u>	<u>0.024</u>	<u>0.014</u>	<u>0.012</u>	<u>0.012</u>	<u>0.009</u>

Tabela 2.1: Estatísticas do teste de Kolmogorov-Smirnov para (P_2, L_2) .

$n = 1000; (r = 3; s = 3)$								
$\mu_1 \backslash \mu_2$	1	2	5	10	20	30	50	70
1	0.268	0.304	0.280	0.273	0.271	0.273	0.272	0.272
2	0.303	0.312	0.267	0.256	0.253	0.254	0.251	0.251
5	0.286	0.258	0.130	0.107	0.098	0.094	0.089	0.090
10	0.283	0.244	0.094	<u>0.060</u>	<u>0.048</u>	<u>0.045</u>	<u>0.049</u>	<u>0.049</u>
20	0.279	0.248	0.085	<u>0.045</u>	<u>0.030</u>	<u>0.031</u>	<u>0.026</u>	<u>0.027</u>
30	0.277	0.245	0.084	<u>0.042</u>	<u>0.034</u>	<u>0.025</u>	<u>0.021</u>	<u>0.019</u>
50	0.275	0.245	0.087	<u>0.042</u>	<u>0.027</u>	<u>0.023</u>	<u>0.016</u>	<u>0.016</u>
70	0.275	0.244	0.086	<u>0.043</u>	<u>0.023</u>	<u>0.020</u>	<u>0.017</u>	<u>0.013</u>

$n = 1000; (r = 4; s = 3)$								
$\mu_1 \backslash \mu_2$	1	2	5	10	20	30	50	70
1	0.337	0.399	0.397	0.397	0.402	0.403	0.402	0.402
2	0.314	0.333	0.313	0.306	0.304	0.305	0.308	0.311
5	0.281	0.266	0.168	0.141	0.135	0.138	0.128	0.128
10	0.285	0.253	0.106	0.078	0.071	0.069	0.068	0.068
20	0.279	0.247	0.091	<u>0.047</u>	<u>0.041</u>	<u>0.035</u>	<u>0.036</u>	<u>0.039</u>
30	0.278	0.248	0.084	<u>0.048</u>	<u>0.029</u>	<u>0.028</u>	<u>0.027</u>	<u>0.024</u>
50	0.278	0.244	0.083	<u>0.041</u>	<u>0.027</u>	<u>0.023</u>	<u>0.021</u>	<u>0.019</u>
70	0.275	0.245	0.088	<u>0.042</u>	<u>0.026</u>	<u>0.022</u>	<u>0.017</u>	<u>0.016</u>

Tabela 2.2: Estatísticas do teste de Kolmogorov-Smirnov para (P_2, L_2) .

Da análise destas tabelas conclui-se que, quando:

- os coeficientes de variação diminuem, as estatísticas do teste de K-S também diminuem, podendo significar que para valores dos coeficientes de variação significativamente pequenos, a parte não linear do polinómio é escassa quando comparada com a parte linear, sugerindo que o polinómio P_2 terá distribuição aproximadamente normal;
- aumenta $(r + s)$, aumenta também a estatística do teste de K-S para os mesmos valores de μ_1 e μ_2 .

■ 2.º Conjunto de simulações

Seja o polinómio

$P = P_0 + (X_1 + \mu_1)^r (X_2 + \mu_2)^s + (X_1 + \mu_1)^s (X_2 + \mu_2)^r$, $r, s \in \{2, 4, 5\}$ e a sua componente linear,

$$L = L_0 + \mu_1^r \mu_2^s + r \mu_1^{r-1} \mu_2^s X_1 + s \mu_1^r \mu_2^{s-1} X_2 + \mu_1^s \mu_2^r + s \mu_1^{s-1} \mu_2^r X_1 + r \mu_1^s \mu_2^{r-1} X_2,$$

$r, s \in \{2, 4, 5\}$ com:

- $P_0 = (X_1 + \mu_1)^6 (X_2 + \mu_2)^4$, o termo dominante do polinómio P ;

e

- $L_0 = \mu_1^6 \mu_2^4 + 6 \mu_1^{6-1} \mu_2^4 X_1 + 4 \mu_1^6 \mu_2^{4-1} X_2$, a parte linear de P_0

Os resultados das estatísticas do teste de K-S são apresentados nas tabelas 2.3, 2.4 e 2.5, admitindo $X_i \sim N(0, \sigma = 1)$; $i = 1, 2$.

Consideram-se para estas simulações os pares (P_0, L_0) , (P, L) e (P, L_0) .

$n = 1000; (P_0, L_0)$								
$\mu_1 \backslash \mu_2$	10	20	30	40	50	60	80	100
10	0.130	0.119	0.116	0.113	0.115	0.116	0.116	0.115
20	0.083	0.064	<u>0.060</u>	<u>0.059</u>	<u>0.060</u>	<u>0.063</u>	<u>0.064</u>	<u>0.060</u>
30	0.080	<u>0.049</u>	<u>0.045</u>	<u>0.044</u>	<u>0.040</u>	<u>0.040</u>	<u>0.044</u>	<u>0.042</u>
40	0.073	<u>0.043</u>	<u>0.039</u>	<u>0.033</u>	<u>0.034</u>	<u>0.031</u>	<u>0.032</u>	<u>0.035</u>
50	0.066	<u>0.042</u>	<u>0.032</u>	<u>0.031</u>	<u>0.029</u>	<u>0.026</u>	<u>0.025</u>	<u>0.035</u>
60	0.063	<u>0.045</u>	<u>0.030</u>	<u>0.026</u>	<u>0.027</u>	<u>0.025</u>	<u>0.025</u>	<u>0.028</u>
80	<u>0.058</u>	<u>0.038</u>	<u>0.030</u>	<u>0.024</u>	<u>0.021</u>	<u>0.020</u>	<u>0.020</u>	<u>0.019</u>
100	<u>0.056</u>	<u>0.036</u>	<u>0.034</u>	<u>0.025</u>	<u>0.020</u>	<u>0.020</u>	<u>0.018</u>	<u>0.018</u>

Tabela 2.3: Estatísticas do teste de Kolmogorov-Smirnov para (P_0, L_0) .

$n = 1000; (P, L) (r = 2, s = 2)$								
$\mu_1 \backslash \mu_2$	10	20	30	40	50	60	80	100
10	0.130	0.119	0.116	0.113	0.115	0.116	0.116	0.115
20	0.083	0.064	<u>0.060</u>	<u>0.059</u>	<u>0.060</u>	<u>0.063</u>	<u>0.064</u>	<u>0.060</u>
30	0.080	<u>0.049</u>	<u>0.045</u>	<u>0.044</u>	<u>0.040</u>	<u>0.040</u>	<u>0.044</u>	<u>0.042</u>
40	0.073	<u>0.043</u>	<u>0.039</u>	<u>0.033</u>	<u>0.034</u>	<u>0.031</u>	<u>0.032</u>	<u>0.035</u>
50	0.066	<u>0.042</u>	<u>0.032</u>	<u>0.031</u>	<u>0.029</u>	<u>0.026</u>	<u>0.025</u>	<u>0.028</u>
60	0.063	<u>0.045</u>	<u>0.030</u>	<u>0.026</u>	<u>0.027</u>	<u>0.025</u>	<u>0.025</u>	<u>0.024</u>
80	<u>0.058</u>	<u>0.038</u>	<u>0.030</u>	<u>0.024</u>	<u>0.021</u>	<u>0.020</u>	<u>0.020</u>	<u>0.019</u>
100	<u>0.056</u>	<u>0.036</u>	<u>0.034</u>	<u>0.025</u>	<u>0.020</u>	<u>0.020</u>	<u>0.018</u>	<u>0.018</u>

$n = 1000; (P, L_0) (r = 2, s = 2)$								
$\mu_1 \backslash \mu_2$	10	20	30	40	50	60	80	100
10	0.130	0.119	0.116	0.113	0.115	0.116	0.116	0.115
20	0.083	0.064	<u>0.060</u>	<u>0.059</u>	<u>0.060</u>	<u>0.063</u>	<u>0.064</u>	<u>0.060</u>
30	0.080	<u>0.049</u>	<u>0.045</u>	<u>0.044</u>	<u>0.040</u>	<u>0.040</u>	<u>0.044</u>	<u>0.042</u>
40	0.073	<u>0.043</u>	<u>0.039</u>	<u>0.033</u>	<u>0.034</u>	<u>0.031</u>	<u>0.032</u>	<u>0.035</u>
50	0.066	<u>0.042</u>	<u>0.032</u>	<u>0.031</u>	<u>0.029</u>	<u>0.026</u>	<u>0.025</u>	<u>0.028</u>
60	0.063	<u>0.045</u>	<u>0.030</u>	<u>0.026</u>	<u>0.027</u>	<u>0.025</u>	<u>0.025</u>	<u>0.024</u>
80	<u>0.058</u>	<u>0.038</u>	<u>0.030</u>	<u>0.024</u>	<u>0.021</u>	<u>0.020</u>	<u>0.020</u>	<u>0.019</u>
100	<u>0.056</u>	<u>0.036</u>	<u>0.034</u>	<u>0.025</u>	<u>0.020</u>	<u>0.020</u>	<u>0.018</u>	<u>0.018</u>

Tabela 2.4: Estatísticas do teste de Kolmogorov-Smirnov para (P, L) e (P, L_0) .

$n = 1000; (P, L) (r = 5, s = 4)$								
$\mu_1 \backslash \mu_2$	10	20	30	40	50	60	80	100
10	0.127	0.111	0.108	0.105	0.105	0.102	0.099	0.099
20	0.084	0.063	<u>0.060</u>	<u>0.059</u>	<u>0.057</u>	<u>0.059</u>	<u>0.059</u>	<u>0.058</u>
30	0.080	<u>0.049</u>	<u>0.045</u>	<u>0.044</u>	<u>0.040</u>	<u>0.039</u>	<u>0.043</u>	<u>0.042</u>
40	0.072	<u>0.044</u>	<u>0.038</u>	<u>0.034</u>	<u>0.033</u>	<u>0.032</u>	<u>0.032</u>	<u>0.035</u>
50	0.066	<u>0.042</u>	<u>0.032</u>	<u>0.032</u>	<u>0.029</u>	<u>0.028</u>	<u>0.026</u>	<u>0.027</u>
60	0.063	<u>0.045</u>	<u>0.030</u>	<u>0.026</u>	<u>0.027</u>	<u>0.026</u>	<u>0.024</u>	<u>0.024</u>
80	<u>0.058</u>	<u>0.038</u>	<u>0.030</u>	<u>0.025</u>	<u>0.021</u>	<u>0.020</u>	<u>0.020</u>	<u>0.020</u>
100	<u>0.056</u>	<u>0.036</u>	<u>0.034</u>	<u>0.026</u>	<u>0.020</u>	<u>0.019</u>	<u>0.018</u>	<u>0.018</u>
$n = 1000; (P, L_0) (r = 5, s = 4)$								
$\mu_1 \backslash \mu_2$	10	20	30	40	50	60	80	100
10	0.147	0.182	0.240	0.293	0.349	0.395	0.472	0.539
20	0.101	0.120	0.157	0.187	0.221	0.255	0.307	0.364
30	0.089	0.090	0.117	0.144	0.162	0.188	0.237	0.277
40	0.078	0.082	0.098	0.117	0.135	0.151	0.189	0.228
50	0.073	0.066	0.089	0.102	0.118	0.130	0.159	0.189
60	0.071	0.064	0.082	0.091	0.105	0.118	0.142	0.165
80	0.063	0.051	0.058	0.083	0.090	0.100	0.119	0.135
100	<i>0.058</i>	0.049	0.050	0.061	0.084	0.091	0.104	0.119

Tabela 2.5: Estatísticas do teste de Kolmogorov-Smirnov para (P, L) e (P, L_0) .

Pela análise das tabelas 2.3, 2.4 e 2.5, verifica-se que:

- quando $r, s \in \{0, 1, 2, 3\}$ e P_0 é dominante, os valores das estatísticas do teste de K-S são iguais para os pares (P_0, L_0) , (P, L) e (P, L_0) . Assim, podemos utilizar a parte linear do termo dominante (L_0) como aproximação do polinómio P ;
- quando $r = 5$ e $s = 4$, os valores obtidos das estatísticas do teste de K-S para (P, L) e (P, L_0) são distintos. Isto deve-se a P_0 não ser dominante.

■ 3.º Conjunto de simulações

Consideremos os seguintes polinómios:

$$\blacktriangleright P_D(\mathbf{X}; \boldsymbol{\mu}) = P^+(\mathbf{X}; \boldsymbol{\mu}) - P^-(\mathbf{X}; \boldsymbol{\mu})$$

$$\blacktriangleright P_S(\mathbf{X}; \boldsymbol{\mu}) = P^+(\mathbf{X}; \boldsymbol{\mu}) + P^-(\mathbf{X}; \boldsymbol{\mu})$$

$$\text{com } P^+ = (X_1 + \mu_1)^4 (X_2 + \mu_2)^3 + (X_1 + \mu_1)^2 (X_2 + \mu_2)^2,$$

$$P^- = [(X_1 + \mu_1)^r (X_2 + \mu_2)^s + (X_1 + \mu_1)^1 (X_2 + \mu_2)^2], \quad r, s \in \{0, 1, 2, 3, 4, 5\},$$

e as suas componentes lineares:

$$\blacktriangleright L_D(\mathbf{X}; \boldsymbol{\mu}) = L^+(\mathbf{X}; \boldsymbol{\mu}) - L^-(\mathbf{X}; \boldsymbol{\mu})$$

$$\blacktriangleright L_S(\mathbf{X}; \boldsymbol{\mu}) = L^+(\mathbf{X}; \boldsymbol{\mu}) + L^-(\mathbf{X}; \boldsymbol{\mu})$$

$$\text{com } L^+(\mathbf{X}; \boldsymbol{\mu}) = \mu_1^4 \mu_2^3 + 4\mu_1^3 \mu_2^3 X_1 + 3\mu_1^4 \mu_2^2 X_2 + \mu_1^2 \mu_2^2 + 2\mu_1 \mu_2^2 X_1 + 2\mu_1^2 \mu_2 X_2 \text{ e}$$

$$L^-(\mathbf{X}; \boldsymbol{\mu}) = \mu_1^r \mu_2^s + r\mu_1^{r-1} \mu_2^s X_1 + s\mu_1^r \mu_2^{s-1} X_2 + \mu_1 \mu_2^2 + \mu_2^2 X_1 + 2\mu_1 \mu_2 X_2.$$

Apresentamos, nas tabelas 2.6, 2.7, 2.8 e 2.9, os resultados das estatísticas do teste de K-S para os pares (P_S, L_S) e (P_D, L_D) com $X_i \sim N(0, \sigma = 1)$, $i = 1, 2$.

$n = 1000; (P_S, L_S) \quad (r = 1; s = 1)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.314	0.310	0.297	0.290	0.291	0.293	0.295	0.294	0.295
5	0.278	0.172	0.137	0.128	0.127	0.121	0.122	0.124	0.126
10	0.265	0.112	0.078	0.071	0.064	<u>0.059</u>	<u>0.059</u>	<u>0.056</u>	<u>0.055</u>
15	0.262	0.092	0.064	<u>0.051</u>	<u>0.049</u>	<u>0.044</u>	<u>0.043</u>	<u>0.039</u>	<u>0.039</u>
20	0.260	0.087	<u>0.052</u>	<u>0.046</u>	<u>0.040</u>	<u>0.038</u>	<u>0.036</u>	<u>0.036</u>	<u>0.031</u>
30	0.256	0.088	<u>0.046</u>	<u>0.038</u>	<u>0.033</u>	<u>0.028</u>	<u>0.031</u>	<u>0.024</u>	<u>0.026</u>
40	0.254	0.085	<u>0.042</u>	<u>0.033</u>	<u>0.030</u>	<u>0.027</u>	<u>0.024</u>	<u>0.023</u>	<u>0.018</u>
60	0.254	0.083	<u>0.040</u>	<u>0.030</u>	<u>0.027</u>	<u>0.022</u>	<u>0.023</u>	<u>0.016</u>	<u>0.018</u>
80	0.256	0.084	<u>0.041</u>	<u>0.028</u>	<u>0.025</u>	<u>0.021</u>	<u>0.017</u>	<u>0.017</u>	<u>0.013</u>
$n = 1000; (P_D, L_D) \quad (r = 1; s = 1)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.329	0.312	0.297	0.293	0.292	0.293	0.295	0.294	0.295
5	0.279	0.172	0.137	0.128	0.127	0.121	0.122	0.124	0.126
10	0.265	0.112	0.078	0.071	0.064	<u>0.059</u>	<u>0.059</u>	<u>0.056</u>	<u>0.055</u>
15	0.262	0.092	0.064	<u>0.051</u>	<u>0.049</u>	<u>0.044</u>	<u>0.043</u>	<u>0.039</u>	<u>0.039</u>
20	0.260	0.087	<u>0.052</u>	<u>0.046</u>	<u>0.040</u>	<u>0.038</u>	<u>0.036</u>	<u>0.036</u>	<u>0.031</u>
30	0.256	0.088	<u>0.046</u>	<u>0.038</u>	<u>0.033</u>	<u>0.028</u>	<u>0.031</u>	<u>0.024</u>	<u>0.026</u>
40	0.254	0.085	<u>0.042</u>	<u>0.033</u>	<u>0.030</u>	<u>0.027</u>	<u>0.024</u>	<u>0.023</u>	<u>0.018</u>
60	0.254	0.083	<u>0.040</u>	<u>0.030</u>	<u>0.027</u>	<u>0.022</u>	<u>0.023</u>	<u>0.016</u>	<u>0.018</u>
80	0.256	0.084	<u>0.041</u>	<u>0.028</u>	<u>0.025</u>	<u>0.021</u>	<u>0.017</u>	<u>0.017</u>	<u>0.013</u>

Tabela 2.6: Estatísticas do teste de Kolmogorov-Smirnov para (P_S, L_S) e (P_D, L_D) .

$n = 1000; (P_S, L_S) (r = 2; s = 1)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.309	0.310	0.297	0.290	0.291	0.293	0.295	0.294	0.295
5	0.276	0.171	0.137	0.128	0.127	0.121	0.122	0.124	0.126
10	0.265	0.112	0.078	0.071	0.064	<u>0.059</u>	<u>0.059</u>	<u>0.056</u>	<u>0.055</u>
15	0.262	0.092	0.064	<u>0.051</u>	<u>0.049</u>	<u>0.044</u>	<u>0.043</u>	<u>0.039</u>	<u>0.039</u>
20	0.260	0.087	<u>0.052</u>	<u>0.046</u>	<u>0.040</u>	<u>0.038</u>	<u>0.036</u>	<u>0.036</u>	<u>0.031</u>
30	0.256	0.088	<u>0.046</u>	<u>0.038</u>	<u>0.033</u>	<u>0.028</u>	<u>0.031</u>	<u>0.024</u>	<u>0.026</u>
40	0.254	0.085	<u>0.042</u>	<u>0.033</u>	<u>0.030</u>	<u>0.027</u>	<u>0.024</u>	<u>0.023</u>	<u>0.018</u>
60	0.254	0.083	<u>0.040</u>	<u>0.030</u>	<u>0.027</u>	<u>0.022</u>	<u>0.023</u>	<u>0.016</u>	<u>0.018</u>
80	0.256	0.084	<u>0.041</u>	<u>0.028</u>	<u>0.025</u>	<u>0.021</u>	<u>0.017</u>	<u>0.017</u>	<u>0.013</u>

$n = 1000; (P_D, L_D) (r = 2; s = 1)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.333	0.312	0.297	0.293	0.292	0.293	0.295	0.294	0.295
5	0.280	0.172	0.137	0.128	0.127	0.121	0.122	0.124	0.126
10	0.266	0.112	0.078	0.071	0.064	<u>0.059</u>	<u>0.059</u>	<u>0.056</u>	<u>0.055</u>
15	0.262	0.092	0.064	<u>0.051</u>	<u>0.049</u>	<u>0.044</u>	<u>0.043</u>	<u>0.039</u>	<u>0.039</u>
20	0.261	0.087	<u>0.052</u>	<u>0.046</u>	<u>0.040</u>	<u>0.038</u>	<u>0.036</u>	<u>0.036</u>	<u>0.031</u>
30	0.256	0.088	<u>0.046</u>	<u>0.038</u>	<u>0.033</u>	<u>0.028</u>	<u>0.031</u>	<u>0.024</u>	<u>0.026</u>
40	0.254	0.085	<u>0.042</u>	<u>0.033</u>	<u>0.030</u>	<u>0.027</u>	<u>0.024</u>	<u>0.023</u>	<u>0.018</u>
60	0.254	0.083	<u>0.040</u>	<u>0.030</u>	<u>0.027</u>	<u>0.022</u>	<u>0.023</u>	<u>0.016</u>	<u>0.018</u>
80	0.256	0.084	<u>0.041</u>	<u>0.028</u>	<u>0.025</u>	<u>0.021</u>	<u>0.017</u>	<u>0.017</u>	<u>0.013</u>

Tabela 2.7: Estatísticas do teste de Kolmogorov-Smirnov para (P_S, L_S) e (P_D, L_D) .

$n = 1000; (P_S, L_S) \quad (r = 3; s = 5)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.325	0.285	0.260	0.249	0.240	0.236	0.234	0.238	0.238
5	0.289	0.182	0.121	0.103	0.090	0.088	0.086	0.085	0.082
10	0.273	0.154	0.094	0.078	0.068	<u>0.053</u>	<u>0.049</u>	<u>0.047</u>	<u>0.043</u>
15	0.262	0.140	0.079	0.064	<u>0.057</u>	<u>0.046</u>	<u>0.048</u>	<u>0.036</u>	<u>0.034</u>
20	0.256	0.132	0.074	<u>0.055</u>	<u>0.048</u>	<u>0.041</u>	<u>0.037</u>	<u>0.032</u>	<u>0.028</u>
30	0.250	0.123	0.072	<u>0.050</u>	<u>0.046</u>	<u>0.032</u>	<u>0.030</u>	<u>0.031</u>	<u>0.029</u>
40	0.244	0.118	0.068	<u>0.049</u>	<u>0.043</u>	<u>0.031</u>	<u>0.025</u>	<u>0.023</u>	<u>0.024</u>
60	0.241	0.106	0.065	<u>0.046</u>	<u>0.039</u>	<u>0.032</u>	<u>0.027</u>	<u>0.021</u>	<u>0.021</u>
80	0.240	0.102	0.063	<u>0.047</u>	<u>0.037</u>	<u>0.028</u>	<u>0.027</u>	<u>0.018</u>	<u>0.016</u>

$n = 1000; (P_D, L_D) \quad (r = 3; s = 5)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.276	0.282	0.257	0.247	0.240	0.236	0.234	0.238	0.238
5	0.270	0.198	0.124	0.102	0.089	0.087	0.085	0.085	0.082
10	0.253	0.193	0.093	0.078	0.065	<u>0.054</u>	<u>0.050</u>	<u>0.048</u>	<u>0.043</u>
15	0.148	0.213	0.083	0.062	<u>0.059</u>	<u>0.047</u>	<u>0.047</u>	<u>0.037</u>	<u>0.034</u>
20	0.190	0.232	0.083	<u>0.056</u>	<u>0.047</u>	<u>0.040</u>	<u>0.037</u>	<u>0.031</u>	<u>0.029</u>
30	0.215	0.250	0.084	<u>0.054</u>	<u>0.048</u>	<u>0.032</u>	<u>0.030</u>	<u>0.031</u>	<u>0.028</u>
40	0.222	0.382	0.092	<u>0.054</u>	<u>0.042</u>	<u>0.034</u>	<u>0.026</u>	<u>0.022</u>	<u>0.024</u>
60	0.224	0.168	0.100	<u>0.058</u>	<u>0.039</u>	<u>0.031</u>	<u>0.028</u>	<u>0.019</u>	<u>0.021</u>
80	0.227	0.067	0.120	<u>0.062</u>	<u>0.043</u>	<u>0.031</u>	<u>0.026</u>	<u>0.018</u>	<u>0.015</u>

Tabela 2.8: Estatísticas do teste de Kolmogorov-Smirnov para (P_S, L_S) e (P_D, L_D) .

$n = 1000; (P_S, L_S) \quad (r = 5; s = 5)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.377	0.345	0.335	0.330	0.333	0.330	0.328	0.326	0.325
5	0.358	0.250	0.198	0.179	0.176	0.173	0.172	0.168	0.166
10	0.346	0.199	0.123	0.098	0.090	0.084	0.078	0.075	0.077
15	0.349	0.189	0.100	0.080	0.068	0.063	<u>0.057</u>	<u>0.053</u>	<u>0.053</u>
20	0.347	0.181	0.086	0.068	<u>0.060</u>	<u>0.047</u>	<u>0.048</u>	<u>0.043</u>	<u>0.042</u>
30	0.348	0.177	0.082	<u>0.058</u>	<u>0.050</u>	<u>0.044</u>	<u>0.036</u>	<u>0.036</u>	<u>0.030</u>
40	0.345	0.179	0.077	<u>0.053</u>	<u>0.042</u>	<u>0.040</u>	<u>0.034</u>	<u>0.031</u>	<u>0.032</u>
60	0.345	0.181	0.077	<u>0.050</u>	<u>0.043</u>	<u>0.033</u>	<u>0.027</u>	<u>0.026</u>	<u>0.021</u>
80	0.344	0.179	0.077	<u>0.049</u>	<u>0.039</u>	<u>0.033</u>	<u>0.026</u>	<u>0.026</u>	<u>0.021</u>

$n = 1000; (P_D, L_D) \quad (r = 5; s = 5)$									
$\mu_1 \backslash \mu_2$	2	5	10	15	20	30	40	60	80
2	0.384	0.346	0.335	0.330	0.333	0.330	0.328	0.326	0.325
5	0.366	0.251	0.198	0.179	0.176	0.173	0.172	0.168	0.166
10	0.353	0.200	0.123	0.098	0.090	0.084	0.078	0.075	0.077
15	0.350	0.189	0.100	0.080	0.068	0.063	<u>0.057</u>	<u>0.053</u>	<u>0.053</u>
20	0.347	0.182	0.086	0.068	<u>0.060</u>	<u>0.047</u>	<u>0.048</u>	<u>0.043</u>	<u>0.042</u>
30	0.349	0.178	0.082	<u>0.058</u>	<u>0.050</u>	<u>0.044</u>	<u>0.036</u>	<u>0.036</u>	<u>0.030</u>
40	0.346	0.179	0.077	<u>0.053</u>	<u>0.042</u>	<u>0.040</u>	<u>0.034</u>	<u>0.031</u>	<u>0.032</u>
60	0.346	0.182	0.077	<u>0.050</u>	<u>0.043</u>	<u>0.033</u>	<u>0.027</u>	<u>0.026</u>	<u>0.021</u>
80	0.344	0.179	0.077	<u>0.049</u>	<u>0.039</u>	<u>0.033</u>	<u>0.026</u>	<u>0.026</u>	<u>0.021</u>

Tabela 2.9: Estatísticas do teste de Kolmogorov-Smirnov para (P_S, L_S) e (P_D, L_D) .

Com base nas tabelas, verifica-se, pois, concordância nas zonas de aceitação, quer os coeficientes dos termos dominantes sejam positivos, negativos ou positivos e negativos.

Existe, pois, possibilidade de aprofundar o estudo dos coeficientes dos termos dominantes dos polinómios.

CAPITULO 3

Modelos

3.1 Metodologia STATIS

3.1.1 Introdução

A metodologia STATIS foi inicialmente introduzida por Escoufier e L' Hermier des Plantes, no Laboratório de Probabilidades e Estatística da Universidade de Montpellier II, por volta de 1976 e posteriormente desenvolvida por Lavit em 1988. Este método permite analisar simultaneamente k estudos. Cada estudo é constituído por uma matriz de dados e duas matrizes de pesos: uma para objectos e outra para variáveis. Estes estudos podem incidir:

► **Sobre os mesmos objectos**, sendo os trios de matrizes da forma $(\mathbf{X}_i, \mathbf{D}_n, \mathbf{D}_{p_i})$, $i = 1, \dots, k$, com as matrizes de dados X_i , $i = 1, \dots, k$, do tipo $n \times p_i$, $i = 1, \dots, k$, onde constam n objectos sobre p_i variáveis quantitativas. Teremos:

$$X_i = \begin{bmatrix} x_{i11} & x_{i12} & \dots & x_{i1p_i} \\ x_{i21} & x_{i22} & \dots & x_{i2p_i} \\ \dots & \dots & \dots & \dots \\ x_{in1} & x_{in2} & \dots & x_{inp_i} \end{bmatrix}$$

em que os valores das variáveis para o i – *ésimo* objecto são

$$\begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{ip_i} \end{bmatrix},$$

sendo ainda

$$\begin{bmatrix} x_{i1j} \\ x_{i2j} \\ \dots \\ x_{inj} \end{bmatrix}$$

o vector dos valores da j – *ésima* variável.

► **Sobre as mesmas variáveis**, permitindo estudar a relação entre as variáveis (método STATIS dual) em que os objectos podem variar ao longo dos quadros. Neste caso o método é caracterizado por k estudos $(\mathbf{X}_i, \mathbf{D}_{n_i}, \mathbf{D}_p)$, $i = 1, \dots, k$.

A matriz de dados X_i é a matriz de dimensão $n_i \times p$, $i = 1, \dots, k$.

► **Sobre os mesmos objectos e sobre as mesmas variáveis**, situação que intersecta as duas situações anteriores.

3.1.2 Inter-estrutura

A utilização da metodologia STATIS permite efectuar uma comparação global das

k matrizes de dados utilizando distâncias euclidianas. Para tal é necessário definir um operador representativo para cada estudo, uma métrica entre os operadores representativos dos estudos e uma representação geométrica dos estudos.

Nas séries do primeiro tipo, um estudo terá a matriz

$$\mathbf{W}_i = \mathbf{X}_i \mathbf{D}_{p_i} \mathbf{X}_i^t, i = 1, \dots, k,$$

dos produtos escalares entre os objectos considerados nesse estudo, $i = 1, \dots, k$.

Então os $\mathbf{A}_i = \mathbf{W}_i \mathbf{D}_n, i = 1, \dots, k$, serão os operadores associados a esses estudos.

As distâncias entre os estudos podem ser medidas utilizando a norma euclidiana da diferença $\mathbf{A}_i - \mathbf{A}_j, i \neq j$, entre os operadores associados aos estudos.

Para representar graficamente os k estudos, utilizamos o produto escalar de Hilbert-Schmidt entre as matrizes:

$$\mathbf{A}_i | \mathbf{A}_j = \text{tr}(\mathbf{A}_i \mathbf{A}_j^t), i, j = 1, \dots, k,$$

introduzido inicialmente por Escoufier para definir uma distância euclidiana entre as matrizes \mathbf{A}_i e \mathbf{A}_j .

A matriz $\mathbf{S} = [s_{ij}], i, j = 1, \dots, k$, (matriz de inter-estrutura, simétrica, semidefinida positiva de ordem k) dos produtos escalares entre as matrizes \mathbf{A}_i e \mathbf{A}_j tem como termo geral:

$$s_{ij} = \mathbf{A}_i | \mathbf{A}_j = \text{tr}(\mathbf{A}_i \mathbf{A}_j^t), i, j = 1, \dots, k,$$

e terá a decomposição espectral

$$\mathbf{S} = \sum_{i=1}^k \theta_i \gamma_i \gamma_i^t = \sum_{i=1}^k \sqrt{\theta_i} \gamma_i \sqrt{\theta_i} \gamma_i^t,$$

com os valores próprios $\theta_1 > \theta_2 > \dots > \theta_k$ associados aos respectivos vectores próprios $\gamma_1, \gamma_2, \dots, \gamma_k$ permitindo uma representação geométrica dos estudos. As coordenadas do ponto representativo do j -ésimo estudo são as j -ésimas componentes dos vectores $\sqrt{\theta_1}\gamma_1, \sqrt{\theta_2}\gamma_2, \dots, \sqrt{\theta_k}\gamma_k$.

A matriz \mathbf{S} pode ainda ser considerada como uma matriz de covariâncias entre os k estudos e os seus elementos são não negativos. Esta matriz admite um vector próprio associado ao maior valor próprio cujas componentes têm todas o mesmo sinal, ver Lavit (1988).

Na prática, a imagem euclidiana restringe-se aos dois primeiros eixos, sendo representados todos no primeiro e quarto quadrante, porque as coordenadas do 1º vector próprio são positivas e $\theta_j = \|\sqrt{\theta_j}\gamma_j\|^2 \approx 0$, $j = 3, \dots, k$, obtendo-se uma imagem euclidiana plana aproximada dos k estudos.

Em particular a série de estudos terá uma estrutura comum de grau um quando os pontos representativos dos estudos se situam ao longo do primeiro eixo, isto é, $\theta_j = \|\sqrt{\theta_j}\gamma_j\|^2 \approx 0$, $j = 2, \dots, k$, ver Lavit (1988). Ter-se-á então

$$\mathbf{S} = \sum_{j=1}^k \theta_j \gamma_j \gamma_j^t \approx \theta_1 \gamma_1 \gamma_1^t,$$

com $\theta_1 \gamma_1 \gamma_1^t$ matriz de característica 1.

Consideremos o ponto S_i , cujas coordenadas são as componentes i dos vectores $\sqrt{\theta_j}\gamma_j$, $j = 1, \dots, k$.

Na figura 3.1, verificamos que os pontos não reflectem uma estrutura de objectos comum, uma vez que estão muito afastados do primeiro eixo.

Na figura 3.2, observamos que os pontos se situam ao longo do 1º eixo. Neste caso existe uma estrutura comum.

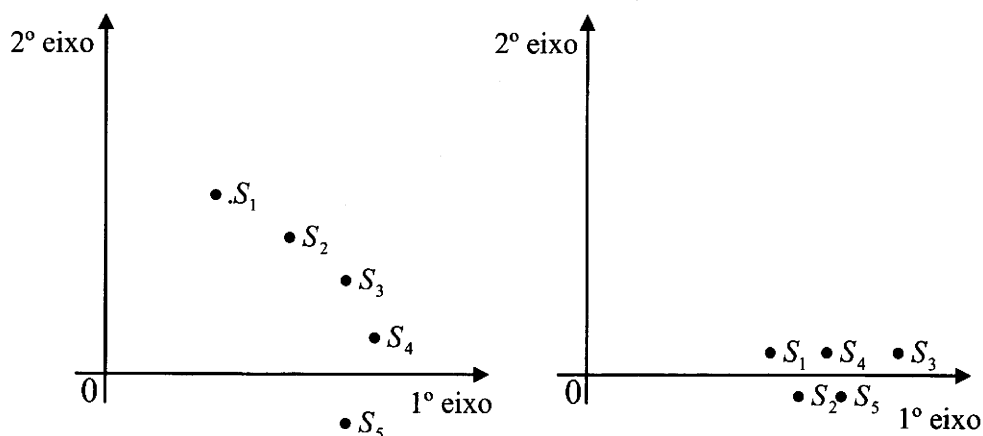


Figura 3.1: Não existe estrutura comum Figura 3.2: Existe estrutura comum

Até agora, (ver Oliveira e Mexia (1999a e 1999b), Oliveira (2002) e Ramos (2006)) apenas foram consideradas séries com estrutura comum de grau um. Nestas tem-se

$$\frac{\theta_j}{\theta_1} \approx 0, j = 2, \dots, k.$$

No caso da estrutura comum de grau h , apenas são desprezíveis os $\theta_j, j = h + 1, \dots, k$. Então os pontos representativos dos estudos estão próximos de um sub-espaco com dimensão h . Esse sub-espaco terá uma base ortonormada constituída por h vectores $\{\alpha_1, \alpha_2, \dots, \alpha_h\}$.

3.2 Modelo com componente aleatória na matriz \mathbf{A}_i

3.2.1 Formulação do modelo

Consideremos as matrizes de dados $\mathbf{X}_i, i = 1, \dots, k$, onde se consideram n objectos e p variáveis e as matrizes de pesos, $\mathbf{D}_n = \frac{\mathbf{I}_n}{n}$ e $\mathbf{D}_p = \mathbf{I}_p$, dos objectos e das variáveis, respectivamente. Associamos às matrizes de dados os operadores $\mathbf{A}_i = \mathbf{X}_i \mathbf{D}_p \mathbf{X}_i^t \mathbf{D}_n, i = 1, \dots, k$, que permitem determinar a matriz $\mathbf{S} = [s_{ij}]$ com $s_{ij} = \text{tr}(\mathbf{A}_i \mathbf{A}_j^t), i, j = 1, \dots, k$.

Na formulação matemática das séries do modelo, começamos por obter uma distribuição aproximada para $\mathbf{Z} = \text{vec}(\mathbf{S})$.

Admitamos então que

$$\mathbf{A}_i = \mathbf{A}_i^o + \mathbf{E}_i, i = 1, \dots, k,$$

com $\mathbf{A}_i^o = [a_{ilh}], l, h = 1, \dots, n$, uma matriz fixa e $\mathbf{E}_i = [e_{ilh}], l, h = 1, \dots, n$, com $\text{vec}(\mathbf{E}_i) \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{n^2}), i = 1, \dots, k$, sendo estes vectores independentes uns dos outros.

Assumiremos que os valores médios dos elementos das matrizes $\mathbf{A}_i, i = 1, \dots, k$, predominam sobre os correspondentes desvios padrão. Estes valores médios são os elementos das matrizes $\mathbf{A}_i^o, i = 1, \dots, k$.

Consideramos

$$\begin{aligned}
S_{ij} &= \text{tr}(\mathbf{A}_i \mathbf{A}_j^t) \\
&= \sum_{l=1}^n \sum_{h=1}^n (a_{ilh} + e_{ilh})(a_{jlh} + e_{jlh}) \\
&= \sum_{l=1}^n \sum_{h=1}^n (a_{ilh}a_{jlh} + a_{ilh}e_{jlh} + a_{jlh}e_{ilh} + e_{ilh}e_{jlh}) \\
&= S_{ij}^o + \sum_{l=1}^n \sum_{h=1}^n (e_{ilh}e_{jlh}), \quad i, j = 1, \dots, k
\end{aligned}$$

com $S_{ij}^o = \sum_{l=1}^n \sum_{h=1}^n (a_{ilh}a_{jlh} + a_{ilh}e_{jlh} + a_{jlh}e_{ilh})$.

Assim com $\mathbf{S}^o = [s_{ij}^o]$, $\text{vec}(\mathbf{S}^o)$ terá distribuição normal e \mathbf{Z} terá distribuição aproximadamente normal onde desprezamos os produtos dos erros $\sum_{l=1}^n \sum_{h=1}^n (e_{ilh}e_{jlh})$, $i = 1, \dots, k$, $j = 1, \dots, k$. No capítulo 2, apresentámos os resultados de vários conjuntos de simulações que mostram que polinómios de baixo grau de variáveis aleatórias normais independentes poderão ser aproximados por polinómios de grau um, normalmente distribuídos, quando as variáveis têm baixos coeficientes de variação. Neste pressuposto $S_{ij} \approx S_{ij}^o$, ($i, j = 1, \dots, k$).

Dada a independência e homocedasticidade dos $\text{vec}(\mathbf{E}_i)$, $i = 1, \dots, k$, virá

$$\text{Var} \left(\sum_{l=1}^n \sum_{h=1}^n a_{ilh}e_{jlh} + a_{jlh}e_{ilh} \right) = \sigma^2 \sum_{l=1}^n \sum_{h=1}^n (a_{ilh}^2 + a_{jlh}^2) = \sigma^2 (K_{ii} + K_{jj}); \quad i \neq j$$

e

$$\text{Var} \left(\sum_{l=1}^n \sum_{h=1}^n 2a_{ilh}e_{ilh} \right) = 4\sigma^2 \sum_{l=1}^n \sum_{h=1}^n a_{ilh}^2 = 4\sigma^2 K_{ii}. \quad (3.1)$$

Com

$$\begin{aligned}
K_{ij} &= \text{tr}(\mathbf{A}_i^o \mathbf{A}_j^{oT}), \\
&= \sum_{l=1}^n \sum_{h=1}^n (a_{ilh})(a_{jlh}), \quad (i, j = 1, \dots, k),
\end{aligned}$$

teremos

$$S_{ij} \overset{\circ}{\sim} N(K_{ij}; \sigma^2(K_{ii} + K_{jj})), i \neq j$$

e

$$S_{ii} \overset{\circ}{\sim} N(K_{ii}; 4\sigma^2 K_{ii}),$$

onde $\overset{\circ}{\sim}$ indica, distribui-se aproximadamente como...

De forma análoga, obteêm-se:

$$\begin{aligned} Cov(S_{ii}, S_{ij}) &= Cov(S_{ii}, S_{ji}) \approx 2\sigma^2 K_{ij}, i \neq j \\ Cov(S_{ii}, S_{ju}) &= 0, i \neq j \neq u \\ Cov(S_{ij}, S_{ji}) &= Var(S_{ij}) \approx \sigma^2(K_{ii} + K_{jj}), i \neq j \\ Cov(S_{ij}, S_{il}) &= Cov(S_{ij}, S_{li}) \approx \sigma^2(K_{jl}), i \neq j, j \neq l \\ Cov(S_{ij}, S_{lu}) &= 0, i \neq j, j \neq l, l \neq u. \end{aligned} \tag{3.2}$$

Assim com $\boldsymbol{\eta} = vec[(\mathbf{K})]$ e $\mathbf{K} = [k_{ij}]$, $(i, j = 1, \dots, k)$ teremos

$$\mathbf{Z} = vec(\mathbf{S}) \overset{\circ}{\sim} N(\boldsymbol{\eta}, \mathbf{L}),$$

sendo os elementos de \mathbf{L} dados pelas expressões (3.1) e (3.2).

3.2.2 Ajustamento do modelo

A discussão apresentada, na secção 3.1, leva-nos a considerar para a matriz \mathbf{S} , o modelo com estrutura comum de grau h , da forma

$$\mathbf{S} = \sum_{i=1}^h \lambda_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^t + \mathbf{E} = \sum_{i=1}^h \boldsymbol{\beta}_i \boldsymbol{\alpha}_i^t + \mathbf{E}, \tag{3.3}$$

com $\boldsymbol{\beta}_i = \lambda_i \boldsymbol{\alpha}_i$, $i = 1, \dots, h$, e \mathbf{E} uma matriz de erros.

Veremos como se:

- ▶ ajusta o modelo com estrutura comum de grau h ;
- ▶ valida o modelo ajustado ou se substituí h por $h + 1$.

Nesta secção consideramos o ajustamento, sendo a validação tratada na secção seguinte.

Sejam $\theta_1 > \theta_2 > \dots > \theta_k$ os valores próprios da matriz \mathbf{S} a que correspondem os vectores próprios $\gamma_1, \gamma_2, \dots, \gamma_k$. Caso $h = 1$ e $k \leq 20$, Oliveira e Mexia (2004) mostraram, através de estudos de simulação, que quando

$$\tau_1 = \frac{\theta_1^2}{\sum_{j=2}^k \theta_j^2} \geq 200$$

podemos usar $\tilde{\beta}_1 = \theta_1 \gamma_1$ para estimar β_1 .

Para o modelo proposto, utilizamos os estimadores $\tilde{\beta}_j = \theta_j \gamma_j$, $j = 1, \dots, h$ para estimar os $\beta_j = \lambda_j \alpha_j$, $j = 1, \dots, h$.

Ora, com $\mathbf{Z} = \text{vec}(\mathbf{S})$ e representando por \otimes o produto de Kronecker de matrizes, temos

$$\tilde{\beta}_j = \mathbf{S} \gamma_j = (\mathbf{I}_k \otimes \gamma_j^t) \mathbf{Z}, \quad j = 1, \dots, h$$

com vector médio

$$E(\tilde{\beta}_j) = (\mathbf{I}_k \otimes \gamma_j^t) \boldsymbol{\eta}, \quad j = 1, \dots, h,$$

e matriz de covariância

$$\sum(\tilde{\beta}_j) = (\mathbf{I}_k \otimes \gamma_j^t) \mathbf{L} (\mathbf{I}_k \otimes \gamma_j), \quad j = 1, \dots, h.$$

Caso a estrutura tenha grau h e os vectores de estrutura $\tilde{\beta}_1, \dots, \tilde{\beta}_h$ sejam bons estimadores dos vectores β_1, \dots, β_h ter-se-à

$$\mathbf{E} = \mathbf{S} - \sum_{i=1}^h \beta_i \alpha_i^t \approx \mathbf{S} - \sum_{i=1}^h \tilde{\beta}_i \gamma_i^t = \mathbf{S} - \sum_{i=1}^h (\mathbf{I}_k \otimes \gamma_i^t) \mathbf{Z} \gamma_i^t$$

e

$$\mathbf{R} = \text{vec}(\mathbf{E}) \approx \mathbf{Z} - \sum_{i=1}^h \text{vec}((\mathbf{I}_k \otimes \gamma_i^t) \mathbf{Z} \gamma_i^t).$$

De

$$\text{vec}((\mathbf{I}_k \otimes \gamma_i^t) \mathbf{Z} \gamma_i^t) = (\gamma_i \otimes \mathbf{I}_k \otimes \gamma_i^t) \mathbf{Z},$$

obtemos

$$\begin{aligned} \mathbf{R} &= \left[\mathbf{I}_{k^2} - \sum_{i=1}^h (\gamma_i \otimes \mathbf{I}_k \otimes \gamma_i^t) \right] \mathbf{Z} \\ &= [\mathbf{I}_{k^2} - \mathbf{W}] \mathbf{Z}, \end{aligned}$$

com $\mathbf{W} = \sum_{i=1}^h (\gamma_i \otimes \mathbf{I}_k \otimes \gamma_i^t)$. Assim a matriz de covariâncias de \mathbf{R} é da forma

$$\sum(\mathbf{R}) = (\mathbf{I}_{k^2} - \mathbf{W}) \mathbf{L} (\mathbf{I}_{k^2} - \mathbf{W})^t.$$

Enquanto \mathbf{R} pode ser considerado como um vector de resíduos $\tilde{\beta} = [\tilde{\beta}_1^t, \dots, \tilde{\beta}_h^t]^t$ será o vector global de estrutura comum ajustado. Estes dois vectores desempenham um papel central na inferência do modelo.

Em geral \mathbf{R} e $\tilde{\beta}$ não são independentes. Assim, somos levados a obter um vector de resíduos homocedástico e independente de $\tilde{\beta}$. Começamos por obter um vector \mathbf{AR} independente de $\tilde{\beta} = \mathbf{BZ}$, com

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_k \otimes \gamma_1^t \\ \dots \\ \mathbf{I}_k \otimes \gamma_h^t \end{bmatrix}.$$

Se \mathbf{Z} tem distribuição aproximadamente normal, \mathbf{AR} é independente de $\tilde{\boldsymbol{\beta}}$ se

$$\sum (\mathbf{AR}; \tilde{\boldsymbol{\beta}}) = \mathbf{A} (\mathbf{I}_{k^2} - \mathbf{W}) \mathbf{L} \mathbf{B}^t = \mathbf{0}.$$

Sendo c a característica da matriz $(\mathbf{I}_{k^2} - \mathbf{W}) \mathbf{L} \mathbf{B}^t$, aplicando o processo de ortogonalização de Gram-Schmidt aos vectores coluna de $(\mathbf{I}_{k^2} - \mathbf{W}) \mathbf{L} \mathbf{B}^t$, obtêm-se os vectores $\mathbf{z}_1, \dots, \mathbf{z}_c$, mutuamente ortogonais com norma 1. Sejam $\boldsymbol{\delta}_i$, os vectores com k^2 componentes todas nulas, com excepção da componente com índice i que é 1, $i = 1, \dots, k^2$. Aplicando agora o processo de ortonormalização de Gram-Schmidt aos vectores de $\mathbf{z}_1, \dots, \mathbf{z}_c, \boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_{k^2-c}$, obtemos de novo os vectores $\mathbf{z}_1, \dots, \mathbf{z}_c$, bem como os vectores linha da matriz \mathbf{A} : $\mathbf{a}_1, \dots, \mathbf{a}_{k^2-c}$.

Teremos então a matriz de covariância de \mathbf{AR} , dada por

$$\sum (\mathbf{AR}) = \mathbf{A} (\mathbf{I}_{k^2} - \mathbf{W}) \mathbf{L} (\mathbf{I}_{k^2} - \mathbf{W})^t \mathbf{A}^t,$$

com valores próprios não nulos ζ_1, \dots, ζ_g a que estão associados os vectores próprios $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_g$. Como as matrizes de covariância não têm valores próprios negativos, estará definida a matriz

$$\mathbf{G} = \text{diag} \left(\zeta_1^{-\frac{1}{2}}, \dots, \zeta_g^{-\frac{1}{2}} \right) \mathbf{G}',$$

onde \mathbf{G}' é a matriz com vectores linha $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_g$.

Verifica-se facilmente que

$$\sum (\mathbf{GAR}) = \mathbf{I}_g.$$

Assim, $\mathbf{T} = \mathbf{GAR}$ será o vector de resíduos homocedástico e independente de $\tilde{\boldsymbol{\beta}}$.

3.2.3 Validação do modelo

Na secção precedente, admitimos que:

- ▶ o modelo (ver 3.3, pág. 37) tinha estrutura comum de grau h ;
- ▶ o vector de resíduos $\mathbf{T} = \mathbf{GAR}$, cujas componentes T_1, \dots, T_g são variáveis aleatórias normais, independentes e identicamente distribuídas (i.i.d.), com valor médio nulo e variância σ^2 .

Para validar o modelo e os pressupostos anteriores haverá que testar as hipóteses H_{0h} e H_{0h}^0 .

No caso da 1.^a hipótese se verificar, verifica-se:

$$H_{0h} : T_1, \dots, T_g \text{ são variáveis aleatórias i.i.d. com distribuição } N(0, \sigma^2).$$

Observe-se que a normalidade de T_1, \dots, T_g é uma consequência testável da normalidade de \mathbf{Z} que se pretende verificar. Da normalidade de \mathbf{Z} resultaria a normalidade de $\tilde{\boldsymbol{\beta}} = \mathbf{BZ}$.

Para testar a hipótese H_{0h} , podemos, por exemplo, utilizar o teste bem conhecido de Shapiro-Wilk. No entanto, se suspeitarmos que a estrutura comum é de grau $h' > h$, seremos levados a previlgiar as hipóteses alternativas:

$$H_{1h} : T_1, \dots, T_g \text{ são variáveis aleatórias independentes com distribuição } N(\mu_i, \sigma^2),$$

$$i = 1, \dots, g.$$

A estatística

$$F_0 = \frac{(g-1) \left(\sum_{i=1}^g T_i \right)^2}{g \sum_{i=1}^g T_i^2 - \left(\sum_{i=1}^g T_i \right)^2}$$

tem, caso H_{0h} se verifique, distribuição F central com 1 e $g-1$ graus de liberdade.

Quando uma das alternativas H_{1h} se verifica, a estatística tem distribuição F duplamente não central com parâmetros de não centralidade

$$\delta_1 = \frac{1}{g\sigma^2} \left(\sum_{i=1}^g \mu_i \right)^2$$

para o numerador e

$$\delta_2 = \frac{1}{\sigma^2} \left(\sum_{i=1}^g \mu_i^2 - \frac{1}{g} \left(\sum_{i=1}^g \mu_i \right)^2 \right)$$

para o denominador, ver Mexia (1989).

Existem hipóteses alternativas, H_{1h} em que δ_1 predomina sobre δ_2 (δ_2 predomina sobre δ_1) em que F_0 tende a tomar valores superiores (inferiores) aos que tomaria, caso H_{0h} se verificasse. Assim sendo, $F_{p,1,g-1}$ o quantil para a probabilidade p da distribuição F central com 1 e $g-1$ graus de liberdade, para um teste de nível α , toma-se a região de aceitação $[F_{\frac{\alpha}{2},1,g-1}; F_{1-\frac{\alpha}{2},1,g-1}]$. Um estudo numérico deste teste mostra que a respectiva potência cresce rapidamente com δ_1 e δ_2 , ver Mexia (1989).

Observe-se que, quando a hipótese H_{0h} não é rejeitada, somos levados a admitir que o grau da estrutura comum não excede h . De facto, podia-se ter uma estrutura comum de grau superior. É-se assim levado a testar a significância da última

componente, representada por $\tilde{\beta}_h$. Quando $E(\tilde{\beta}_h) = \mathbf{0}$, a forma quadrática

$$\mathbf{Q}_h = \tilde{\beta}_h^t [(\mathbf{I}_k \otimes \gamma_h^t) \mathbf{L} (\mathbf{I}_k \otimes \gamma_h)]^+ \tilde{\beta}_h$$

distribui-se como o produto de σ^2 por um qui-quadrado com

$$c_h = \text{Car} [(\mathbf{I}_k \otimes \gamma_h^t) \mathbf{L} (\mathbf{I}_k \otimes \gamma_h)]^+$$

graus de liberdade, sendo \mathbf{M}^+ a inversa de Moore-Penrose de \mathbf{M} .

A estatística

$$F_h = \frac{g}{c_h} \frac{\mathbf{Q}_h}{\|\mathbf{T}\|^2},$$

caso $H_{0h}^0 : \beta_h^0 = \mathbf{0}$ se verificar, terá distribuição F central com c_h e g graus de liberdade, ver Mexia (1995).

Quando $\beta_h^0 = E(\tilde{\beta}_h) \neq \mathbf{0}$, o numerador de F_h tem parâmetro de não centralidade

$$\delta_h^0 = \frac{1}{\sigma^2} \beta_h^{0t} [(\mathbf{I}_k \otimes \gamma_h^t) \mathbf{L} (\mathbf{I}_k \otimes \gamma_h)]^+ \beta_h^0.$$

Observe-se que convém começar por testar H_{0h} , pois quando se verifica uma das alternativas H_{1h} , $\|\mathbf{T}\|^2$ será o produto de σ^2 por um qui-quadrado com g graus de liberdade e parâmetro de não centralidade $\delta_1 + \delta_2$. Resultará daqui uma perda de potência do teste F_h , ver Mexia (1989).

Na prática convirá ir testando valores crescentes para os graus de estrutura comum até não se rejeitar a hipótese H_{0h} e se rejeitar a hipótese H_{0h}^0 .

3.2.4 Aplicação: Eleições autárquicas em Portugal

3.2.4.1 Apresentação do problema

O nosso estudo incidiu sobre os resultados das eleições autárquicas realizadas em Portugal entre 1985 e 2001.

Assim, foram seleccionadas as regiões Norte, Centro Norte, Centro e Sul, tendo em cada região sido seleccionados dois distritos: um no Litoral e outro no Interior. Em cada distrito foram analisados dois concelhos, sendo um deles a capital de distrito.

Os Concelhos analisados foram: Viana do Castelo (V.Cas), Ponta da Barca (P.Bar), Bragança (Bra), Alfandega da Fé (A.Fé), Aveiro (Ave), Santa Maria da Feira (S.M.F.), Guarda (Gua), Pinhel (Pin), Leiria (Lei), Alcobaça (Alc), Portalegre (Por), Elvas (Elv), Faro (Far), Silves (Sil), Beja (Bej) e Mértola (Mer).

Em cada concelho foram seleccionadas 10 freguesias. No caso em que o concelho não tinha 10 freguesias, juntaram-se freguesias de concelhos adjacentes. Ao concelho de Faro juntámos cinco freguesias do concelho de Olhão, ao concelho de Silves, juntámos três de Albufeira e ao concelho de Mértola juntámos uma freguesia do concelho de Castro Verde.

Consideraram-se as seguintes opções políticas: Partido Social Democrata (PSD); Partido Socialista (PS); Partido do Centro Democrático Social/Partido Popular (CDS/PP); Coligação Democrática Unitária, na qual o Partido Comunista tem um papel preponderante (CDU); Outros partidos (OUT); Votos brancos/nulos

(B/N) e Abstenção (ABS).

Nalgumas eleições e concelhos, o PSD e o CDS apresentaram listas conjuntas. Os votos na coligação foram distribuídos pelos dois partidos, tendo nestes casos sido ponderadas as suas votações proporcionalmente aos seus resultados nas eleições autárquicas imediatamente anteriores ou posteriores.

Os dados dos resultados eleitorais foram obtidos através do site do STAPE (Secretariado Técnico dos Assuntos para o Processo Eleitoral).

3.2.4.2 Ajustamento e validação do modelo

Com base nos resultados eleitorais foi possível construir, para cada eleição e para cada concelho, uma matriz de dados \mathbf{X}_i , $i = 1, \dots, 5$, com as percentagens de votos. As linhas da matriz correspondem às freguesias e as colunas às opções políticas. Atribuímos as matrizes de pesos $\mathbf{D}_{10} = \frac{1}{10}\mathbf{I}_{10}$ e $\mathbf{D}_7 = \mathbf{I}_7$ para as freguesias e para as opções políticas, respectivamente.

Foi utilizado o Software Wolfram Research Mathematica 5.2, para efectuar cálculos no âmbito do ajustamento e validação do modelo. Os dois primeiros valores próprios, θ_1 e θ_2 , e os valores da preponderância τ para os vários concelhos encontram-se nas tabelas 3.1 e 3.2.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F
θ_1	86.888	126.264	145.213	105.563	120.169	102.488
θ_2	0.719	0.672	0.485	1.248	0.597	0.533
τ	11401.4	29599.8	54427.3	7039.4	36052.9	24746.5
	Gua	Pin	Lei	Alc.	Por	Elv
θ_1	144.113	111.733	115.010	124.364	104.169	144.970
θ_2	0.972	0.792	0.179	0.593	0.129	0.514
τ	19042.7	14960.0	131484.1	33739.7	98166.1	62641.7

Tabela 3.1: Valores próprios e valores da preponderância.

	Far	Sil	Bej	Mer
θ_1	141.857	97.299	129.321	143.055
θ_2	0.219	0.199	0.164	0.109
τ	121701.2	73105.6	392064.8	280434.8

Tabela 3.2: Valores próprios e valores da preponderância.

A análise das tabelas permite constatar que o segundo valor próprio é muito inferior ao primeiro valor próprio para todos os concelhos analisados. Verifica-se, ainda, que o valor da preponderância em todos os concelhos é superior a 7039, permitindo concluir que a estrutura comum tem grau um. Ajustamos, assim, modelos da forma $\mathbf{S} = \lambda_1 \boldsymbol{\alpha}_1 \boldsymbol{\alpha}_1^t + \mathbf{E}$, sendo \mathbf{E} uma matriz de erros.

Na tabela 3.3, apresentam-se os vectores de estrutura estimados para os 16 con-

celhos. Estes resultados permitem verificar que há menor variabilidade nos valores das componentes do vector de estrutura nos concelhos de Beja e Leiria e maior variabilidade nos concelhos de Alcobaça e Elvas.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F	Gua	Pin
1985	35.722	51.256	61.229	48.695	53.941	40.077	57.716	48.432
1989	37.095	56.419	63.300	40.778	56.815	43.042	66.348	41.315
1993	33.817	62.272	65.362	52.640	53.900	46.369	62.877	40.435
1997	42.156	56.821	65.989	46.876	49.211	49.900	71.605	59.374
2001	44.461	55.007	68.587	46.273	54.554	49.039	62.885	57.196
	Lei	Alc.	Por	Elv	Far	Sil	Bej	Mer
1985	50.119	72.298	41.815	51.721	57.963	41.714	57.182	65.941
1989	52.948	50.326	49.264	49.985	67.139	45.439	55.938	66.916
1993	48.262	52.563	47.862	47.763	56.466	42.323	58.677	61.514
1997	51.991	48.495	46.030	76.315	68.640	39.776	58.982	63.639
2001	53.662	50.912	47.604	87.963	66.001	47.844	58.338	61.683

Tabela 3.3: Vectores de estrutura estimados.

Os valores observados das estatísticas F_0 e F_1 (pág. 42 e 43) para os vários concelhos encontram-se nas tabelas 3.4 e 3.5. Estas estatísticas validam de forma clara e para todos os concelhos o modelo.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F
F_0	2.864	4.250	0.477	0.583	0.038	0.003
F_1	160.901	285.115	289.266	130.612	324.730	241.668
$\ \mathbf{T}\ ^2$	0.270	0.221	0.251	0.404	0.185	0.212
	Gua	Pin	Lei	Alc.	Por	Elv
F_0	0.060	0.046	0.258	0.668	2.056	3.985
F_1	179.691	148.563	611.702	221.258	801.233	276.642
$\ \mathbf{T}\ ^2$	0.401	0.376	0.094	0.281	0.065	0.262

Tabela 3.4: Valores das estatísticas F_0 e F_1 - Componente aleatória em \mathbf{A}_i .

	Far	Sil	Bej	Mer
F_0	1.595	1.138	0.069	1.238
F_1	779.280	446.696	762.419	1740.400
$\ \mathbf{T}\ ^2$	0.084	0.110	0.085	0.040

Tabela 3.5: Valores das estatísticas F_0 e F_1 - Componente aleatória em \mathbf{A}_j .

Assim, a hipótese de que os resíduos tinham valores médios nulos e eram homocedásticos não foi rejeitada para nenhum dos concelhos. Com efeito, para um nível de significância de 5%, a região de não rejeição para o teste com estatística F_0 é $[F_{0.025;1,9}, F_{0.975;1,9}] = [0.0010; 7.21]$ para todos os concelhos com exceção de Faro. Para este concelho, a região de não rejeição, para o mesmo nível de significância, é $[F_{0.025;1,8}, F_{0.975;1,8}] = [0.0010; 7.57]$.

Por outro lado os valores das estatísticas F_1 foram altamente significativos, podendo pois não se rejeitar, para todos os concelhos, a existência de uma estrutura comum de grau um.

3.3 Modelo com componente aleatória na matriz \mathbf{X}_i

3.3.1 Formulação, ajustamento e validação do modelo

Na secção 3.2, construímos o modelo introduzindo a componente aleatória nos operadores $\mathbf{A}_i = \mathbf{A}_i^0 + \mathbf{E}_i$, $i = 1, \dots, k$, com \mathbf{A}_i^0 matriz fixa e admitindo que $\text{vec}(\mathbf{E}_i) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{np})$.

Vamos agora introduzir a componente aleatória ao nível da matriz de dados \mathbf{X}_i . Assim, admitimos que

$$\mathbf{X}_i = \mathbf{U}_i + \mathbf{E}_i, i = 1, \dots, k,$$

com \mathbf{U}_i matriz fixa e $\text{vec}(\mathbf{E}_i) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{np})$, $i = 1, \dots, k$, sendo estes vectores independentes uns dos outros. Assumimos que nas observações que constituem as matrizes \mathbf{X}_i , $i = 1, \dots, k$, o valor médio predomina sobre o desvio padrão estandardizado. Os valores médios são os elementos das matrizes \mathbf{U}_i , $i = 1, \dots, k$, e os elementos das matrizes \mathbf{E}_i , $i = 1, \dots, k$, têm o desvio estandardizado. O coeficiente de variação será pequeno para essas variáveis.

Associamos aos estudos as matrizes \mathbf{A}_i , com $\mathbf{A}_i = \mathbf{X}_i \mathbf{D}_p \mathbf{X}_i^t \mathbf{D}_n$, $i = 1, \dots, k$, com as matrizes diagonais $\mathbf{D}_n = \frac{1}{n} \mathbf{I}_n$ e $\mathbf{D}_p = \mathbf{I}_p$ que representam os pesos dos objectos e das variáveis, respectivamente. Como $\mathbf{A}_i = [a_{ilh}]$, $i = 1, \dots, k$; $l, h = 1, \dots, n$, e $\mathbf{X}_i = [x_{ijl}] = [u_{ijl} + e_{ijl}]$, $i = 1, \dots, k$, $j = 1, \dots, n$, $l = 1, \dots, p$, ao desprezar o produto dos erros $e_{ijl} e_{ijl}$, com base na normalidade aproximada, mencionada no

Capítulo 2, obtemos os valores $[a_{ilh}] \approx [\eta_{ilh} + e_{ilh}^\circ]$, $i = 1, \dots, k$, $l, h = 1, \dots, n$, com

$$\eta_{ilh} = \frac{1}{n} \sum_{t=1}^p (u_{ilt} u_{iht}),$$

valor fixo e a variável

$$e_{ilh}^\circ = \frac{1}{n} \sum_{t=1}^p (u_{ilt} e_{iht} + u_{iht} e_{ilt}).$$

Dado ter-se admitido que $e_{ijh} \sim N(0, \sigma^2)$, $i = 1, \dots, k$, $j = 1, \dots, n$, $h = 1, \dots, p$ e que estas variáveis são independentes, e_{ilh}° terá distribuição normal com parâmetros:

$$\begin{aligned} E(e_{ilh}^\circ) &= 0, \\ \text{Var}(e_{ilh}^\circ) &= \text{Var} \left[\frac{1}{n} \sum_{t=1}^p (u_{ilt} \cdot e_{iht} + u_{iht} \cdot e_{ilt}) \right], \\ &= \sigma^2 \left(\frac{1}{n} \right)^2 \sum_{t=1}^p [(u_{ilt}^2) + (u_{iht}^2)], \quad l \neq h, \end{aligned}$$

e

$$\text{Var}(e_{iil}^\circ) = \text{Var} \left[\frac{1}{n} \sum_{t=1}^p 2(u_{ilt} \cdot e_{ilt}) \right] = \sigma^2 \left(\frac{2}{n} \right)^2 \sum_{t=1}^p (u_{ilt}^2).$$

Para construir a matriz $\mathbf{S} = [s_{ij}]$, $i, j = 1, \dots, k$, determinamos:

$$\begin{aligned} S_{ij} &= \text{tr}(\mathbf{A}_i \mathbf{A}_j^t) = \sum_{l=1}^n \sum_{h=1}^n a_{ilh} a_{jlh} \\ &= \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} + e_{ilh}^\circ) (\eta_{jlh} + e_{jlh}^\circ) \\ &= \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} \eta_{jlh}) + \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} e_{jlh}^\circ) + \sum_{l=1}^n \sum_{h=1}^n (\eta_{jlh} e_{ilh}^\circ) \\ &\quad + \sum_{l=1}^n \sum_{h=1}^n (e_{ilh}^\circ e_{jlh}^\circ). \end{aligned}$$

Admitimos que

$$\sum_{l=1}^n \sum_{h=1}^n (e_{ilh}^\circ e_{jlh}^\circ) \approx 0,$$

para poder aplicar a normalidade aproximada, vindo

$$S_{ij} \approx \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} \eta_{jlh}) + \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} e_{jlh}^{\circ}) + \sum_{l=1}^n \sum_{h=1}^n (\eta_{jlh} e_{ilh}^{\circ}), i \neq j$$

e

$$S_{ii} \approx \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh}^2) + 2 \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} e_{ilh}^{\circ}).$$

Assim, ter-se-á

$$S_{ij} \overset{\circ}{\sim} N \left(\sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} \eta_{jlh}); \sum_{l=1}^n \sum_{h=1}^n \eta_{ilh}^2 \text{Var}(e_{jlh}^{\circ}) + \sum_{l=1}^n \sum_{h=1}^n \eta_{jlh}^2 \text{Var}(e_{ilh}^{\circ}) \right), i \neq j$$

e

$$S_{ii} \overset{\circ}{\sim} N \left(\sum_{l=1}^n \sum_{h=1}^n \eta_{ilh}^2; 4 \sum_{l=1}^n \sum_{h=1}^n \eta_{ilh}^2 \text{Var}(e_{ilh}^{\circ}) \right).$$

Por outro lado

$$\begin{aligned} \text{Cov}(S_{ii}, S_{ii}) &= \text{Var}(S_{ii}) = 4 \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh}^2 \text{Var}(e_{ilh}^{\circ})), i = 1, \dots, k, \\ \text{Cov}(S_{ij}, S_{ij}) &= \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh}^2 \text{Var}(e_{jlh}^{\circ})) + \sum_{l=1}^n \sum_{h=1}^n (\eta_{jlh}^2 \text{Var}(e_{ilh}^{\circ})), \\ \text{Cov}(S_{ii}, S_{ij}) &= 2 \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} \eta_{jlh}) \text{Var}(e_{ilh}^{\circ}), i \neq j, \\ \text{Cov}(S_{ij}, S_{it}) &= \sum_{l=1}^n \sum_{h=1}^n (\eta_{jlh} \eta_{ilh}) \text{Var}(e_{ilh}^{\circ}), i \neq j, i \neq t, j \neq t, \\ \text{Cov}(S_{ij}, S_{tu}) &= 0, i \neq t, i \neq u, j \neq t, j \neq u, \end{aligned} \tag{3.4}$$

o que permitirá determinar a matriz de covariância do $\text{vec}(\mathbf{S})$.

Se a matriz $\mathbf{K} = [k_{ij}]$, com $k_{ij} = \sum_{l=1}^n \sum_{h=1}^n (\eta_{ilh} \eta_{jlh})$, então para $\boldsymbol{\eta} = \text{vec}(\mathbf{K})$ teremos

$$\mathbf{Z} = \text{vec}(\mathbf{S}) \overset{\circ}{\sim} \mathbf{N}(\boldsymbol{\eta}, \mathbf{L}),$$

com os elementos da matriz de covariância \mathbf{L} dados pelas expressões (3.4).

Para uma estrutura comum de grau h , com os vectores de estrutura $\beta_i = \lambda_i \alpha_i$, $i = 1, \dots, h$, ter-se-á o modelo da forma

$$\mathbf{S} = \sum_{i=1}^h \beta_i \alpha_i^t + \mathbf{E},$$

sendo $\tilde{\beta}_i = \theta_i \gamma_i$, $i = 1, \dots, h$, os estimadores dos correspondentes vectores de estrutura.

Com $\mathbf{R} = \text{vec}(\mathbf{E})$ determinamos o vector de resíduos homocedástico $\mathbf{T} = \mathbf{GAR}$ independente de $\tilde{\beta} = [\tilde{\beta}_1^t, \dots, \tilde{\beta}_h^t]^t$, conforme o desenvolvimento efectuado na sub-secção 3.2.2.

Para validarmos o modelo com vector de resíduos $\mathbf{T} = \mathbf{GAR}$ e com estrutura de grau h , testamos as hipóteses H_{0h} e H_{0h}^0 , mencionadas anteriormente (ver sub-secção 3.2.3, pág.41 e 43). Assim utilizamos as estatísticas F_0 e F_h (pág.42 e 43) para testar se as componentes de \mathbf{T} têm valor médio nulo e variância σ^2 e se a estrutura comum tem grau h , testando a significância da componente β_h .

3.3.2 Aplicação: Eleições autárquicas em Portugal

O nosso estudo incidiu sobre os mesmos resultados das eleições autárquicas dos 16 concelhos mencionados na sub-secção 3.2.4.

Foi utilizado o Software Wolfram Research Mathematica 5.2 para efectuar os cálculos no âmbito do ajustamento e validação do modelo que foi aplicado, considerando as matrizes de pesos dos objectos (freguesias) $\mathbf{D}_{10} = \frac{1}{10} \mathbf{I}_{10}$ e das variáveis (opções políticas) $\mathbf{D}_7 = \mathbf{I}_7$.

A matriz de covariância, obtida através das expressões 3.4 (pág.52), permitiu obter o vector de resíduos homocedástico para cada concelho. Para validarmos o modelo aplicamos as estatísticas F_0 e F_1 definidas nas pág. 42 e 43.

Os valores destas estatísticas encontram-se nas tabelas 3.6 e 3.7.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F
F_0	3.178	0.782	2.608	0.330	0.798	0.293
F_1	485.8	3229.5	3042.0	2962.1	18086.6	5562.1
$\ \mathbf{T}\ ^2$	0.86	0.147	0.189	0.143	0.028	0.076
	Gua	Pin	Lei	Alc.	Por	Elv
F_0	0.125	0.227	0.653	0.207	5.146	0.041
F_1	798.0	561.2	12691.9	3961.0	16810.0	315.4
$\ \mathbf{T}\ ^2$	0.652	0.820	0.037	0.127	0.026	2.615

Tabela 3.6: Valores das estatísticas F_0 e F_1 - Componente aleatória em \mathbf{X}_i .

	Far	Sil	Bej	Mer
F_0	0.571	0.193	0.077	0.043
F_1	157.2	4675.6	3874.4	35800.0
$\ \mathbf{T}\ ^2$	3.280	0.085	0.125	0.015

Tabela 3.7: Valores das estatísticas F_0 e F_1 - Componente aleatória em \mathbf{X}_i .

Para um nível de significância de 0.05, a região de não rejeição para o teste com estatística F_0 é $[F_{0.025;1,9}, F_{0.975;1,9}] = [0.0010; 7.21]$ para os concelhos de Viana do Castelo, Ponta da Barca, Bragança, Aveiro, Guarda, Pinhel, Alcobaça, e El-

vas. Para os restantes concelhos, a região de não rejeição para o mesmo nível de significância é $[F_{0.025;1,8}, F_{0.975;1,8}] = [0.0010; 7.57]$. Assim a hipótese de que os resíduos tinham valores médios nulos e eram homocedásticos não foi rejeitada para nenhum dos concelhos.

Por outro lado, da aplicação da estatística F_1 , podemos concluir também para todos os concelhos, que existe estrutura comum do primeiro grau altamente significativa ao nível de 0.01.

CAPÍTULO 4

Inferência aplicada às séries de estudos

4.1 Modelo com a componente aleatória em A_i

4.1.1 Considerações preliminares

Neste capítulo começamos a explorar as potencialidades dos modelos desenvolvidos no capítulo 3. Com efeito as aplicações usuais da metodologia STATIS restringem-se a uma série isolada de estudos, enquanto a nossa aproximação permite considerar m séries emparelhadas de estudos. Por séries emparelhadas de estudos entende-se séries em que há correspondência um a um entre os respectivos estudos.

Admitimos que as séries emparelhadas de estudos têm estrutura comum de grau h . Sendo k o número de estudos, a informação contida nas mesmas pode ser condensada, como vimos, em h vectores de estrutura ajustados e numa soma de quadrados de resíduos para cada série. As i -ésimas componentes dos vectores de estrutura correspondem às i -ésimas séries de estudos.

Podemos agora realizar dois tipos de análise: análise longitudinal e análise transversal.

A análise longitudinal estudará a influência dos factores sobre a evolução das séries de estudos. Este tipo de análise é sobretudo relevante, quando os estudos em cada série são realizados em tempos sucessivos, sendo simultâneos os estudos que se correspondem nas diferentes séries como a seguir se esquematizam:

Tempo	1.^a Série	2.^a Série	...	m.^a Série
t_1	1. ^o Estudo	1. ^o Estudo	...	1. ^o Estudo
t_2	2. ^o Estudo	2. ^o Estudo	...	2. ^o Estudo
...
t_k	k. ^o Estudo	k. ^o Estudo	...	k. ^o Estudo

Na análise transversal, estuda-se a acção dos níveis dos factores nas componentes homólogas dos vectores de estrutura. Tal como a análise longitudinal, a análise transversal é particularmente interessante, quando os estudos se dispõem como atrás se esquematizou.

Se o grau da estrutura for um, temos um único vector de estrutura ajustado por série e, se os estudos se dispõem como indicado, podemos apresentar os resultados dos ajustamentos dos modelos como a seguir se indica:

	1. ^a Série	2. ^a Série	...	m. ^a Série
1. ^o Estudo	$\tilde{\beta}_1(1)$	$\tilde{\beta}_1(2)$...	$\tilde{\beta}_1(m)$
2. ^o Estudo	$\tilde{\beta}_2(1)$	$\tilde{\beta}_2(2)$...	$\tilde{\beta}_2(m)$
...
k. ^o Estudo	$\tilde{\beta}_k(1)$	$\tilde{\beta}_k(2)$...	$\tilde{\beta}_k(m)$

onde $\tilde{\beta}_i(l), i = 1, \dots, k; l = 1, \dots, m$ são as componentes dos vectores de estrutura ajustados para as diferentes séries.

Representamos por $\tilde{\beta}(l)$ o único vector de estrutura ajustado para a série indexada por l , com $l = 1, 2, \dots, m$ e por $V(l)$ a correspondente soma de quadrados dos resíduos.

Admitimos que $\tilde{\beta}(l) \sim N(\beta(l), \sigma^2 \mathbf{W}(l))$ é independente de $V(l) \sim \sigma^2 \chi_{g(l)}^2, l = 1, 2, \dots, m$. Como se têm m séries de estudos, atendendo à reprodutividade dos qui-quadrados ter-se-à

$$V = \sum_{l=1}^m V(l) \sim \sigma^2 \chi_g^2,$$

com $g = \sum_{l=1}^m g(l)$, independente de $\tilde{\beta}(l), l = 1, 2, \dots, m$.

No âmbito do estudo da inferência dos modelos para as m séries de estudos emparelhadas, podemos ainda construir intervalos de confiança simultâneos e realizar comparações múltiplas. De entre os vários métodos possíveis destacam-se o método de Scheffé.

Este método baseia-se no seguinte teorema:

■ Teorema de Scheffé

Se $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)$ tem distribuição normal com vector médio $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$ e matriz de covariância regular $\sigma^2\mathbf{M}$, independente de V , que tem distribuição qui-quadrado central com g graus de liberdade e sendo $F_{(1-q, n, g)}$ o quantil para a probabilidade $1 - q$, da distribuição F central, com n e g graus de liberdade, tem-se

$$P \left[\bigcap_{\mathbf{c}} \left(|\mathbf{c}^t \boldsymbol{\mu} - \mathbf{c}^t \mathbf{Y}| \leq \sqrt{n \cdot F_{1-q, n, g}(\mathbf{c}^t \mathbf{M} \mathbf{c}) \frac{V}{g}} \right) \right] = 1 - q$$

indicando $\bigcap_{\mathbf{c}}$ que a intersecção é para todos os vectores $\mathbf{c} \in \mathbb{R}^n$.

Os intervalos de confiança simultâneos de nível $1 - q$ para as combinações lineares $\mathbf{c}^t \boldsymbol{\mu}$, das componentes de $\boldsymbol{\mu}$, são da forma

$$\left[\mathbf{c}^t \mathbf{Y} - \sqrt{n F_{1-q, n, g}(\mathbf{c}^t \mathbf{M} \mathbf{c}) \frac{V}{g}}; \mathbf{c}^t \mathbf{Y} + \sqrt{n F_{1-q, n, g}(\mathbf{c}^t \mathbf{M} \mathbf{c}) \frac{V}{g}} \right].$$

Este teorema foi posteriormente generalizado por Mexia (1989) em que a matriz de covariância de \mathbf{Y} é apenas semi-definida positiva.

4.1.2 Análise longitudinal para as séries individuais

Esta análise é efectuada com recurso a contrastes para estudar a evolução das séries de estudos. Um contraste entre as componentes do vector $\mathbf{v} = (v_1, v_2, \dots, v_k)^t$ pode ser definido como uma combinação linear $\sum_{i=1}^k c_i v_i$, onde as componentes têm soma nula, $\sum_{i=1}^k c_i = 0$.

Sendo $\mathbf{P} = [p_{ij}]$ uma matriz ortogonal estandardizada, $k \times k$, com vectores linha

$\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_k$, tais que $\mathbf{p}_1 = (p_{11}, p_{12}, \dots, p_{1k})$, com $p_{1j} = \frac{1}{\sqrt{k}}, j = 1, \dots, k$, serão nulas as somas das componentes dos vectores $\mathbf{p}_2, \dots, \mathbf{p}_k$. Como $\mathbf{p}_i^t \mathbf{p}_j = 0, i \neq j$, diz-se que os contrastes $\mathbf{p}_i^t \mathbf{v}, t = 2, \dots, k$, são ortogonais.

Observe-se que há uma grande flexibilidade na escolha dos contrastes. Muitas vezes são utilizados vectores com norma um de forma a descrever convenientemente a evolução das séries de estudos. Assim, utiliza-se $\mathbf{c}_1^t = \left(-\frac{1}{\sqrt{2}}, 0, \dots, 0, \frac{1}{\sqrt{2}}\right)$ para avaliar a evolução global da série, se houver um número ímpar de estudos e $\mathbf{c}_2^t = \left(\frac{1}{\sqrt{6}}, 0, \dots, 0, -\frac{2}{\sqrt{6}}, 0, \dots, 0, \frac{1}{\sqrt{6}}\right)$ para avaliar a curvatura da evolução.

Caso $\mathbf{c}_l^t \mathbf{v} > 0$ [< 0], $l = 1, 2$, teremos uma evolução global positiva [negativa] com a concavidade voltada para cima [baixo] o que indica aceleração [desaceleração] da evolução.

Admitimos que as séries de estudos correspondem aos níveis de um factor e que se têm vectores de estrutura ajustados $\tilde{\boldsymbol{\beta}}(l) \sim N(\boldsymbol{\beta}(l), \sigma^2 \mathbf{W}(l)), l = 1, \dots, m$, independentes entre si e de $V \sim \sigma^2 \chi_g^2$. Então teremos $\mathbf{c}^t \tilde{\boldsymbol{\beta}}(l) \sim N(\eta_l, \sigma^2 d_l), l = 1, \dots, m$, com $\eta_l = \mathbf{c}^t \boldsymbol{\beta}(l)$ e $d_l = \mathbf{c}^t \mathbf{W}(l) \mathbf{c}, l = 1, \dots, m$, sendo estas variáveis independentes entre si e de $V \sim \sigma^2 \chi_g^2$, ver Mexia (1995).

Assim para testar a hipótese

$$H_0 : \mathbf{c}^t \boldsymbol{\beta}(l) = \mathbf{c}^t \boldsymbol{\beta}_0(l), l = 1, \dots, m$$

utilizamos a estatística de teste

$$F_0 = \frac{g}{1} \frac{\left(\mathbf{c}^t \tilde{\boldsymbol{\beta}}(l) - \mathbf{c}^t \boldsymbol{\beta}_0(l)\right) \left(\mathbf{c}^t \mathbf{W}(l) \mathbf{c}\right)^{-1} \left(\mathbf{c}^t \tilde{\boldsymbol{\beta}}(l) - \mathbf{c}^t \boldsymbol{\beta}_0(l)\right)}{V} \sim F(\cdot | 1, g, \delta),$$

com distribuição F e com 1 e g graus de liberdade e parâmetro de não centralidade

$$\delta = \frac{1}{\sigma^2} (\mathbf{c}^t \boldsymbol{\beta}(l) - \mathbf{c}^t \boldsymbol{\beta}_0(l)) (\mathbf{c}^t \mathbf{W}(l) \mathbf{c})^{-1} (\mathbf{c}^t \boldsymbol{\beta}(l) - \mathbf{c}^t \boldsymbol{\beta}_0(l)).$$

Como o valor $\mathbf{c}^t \mathbf{W}(l) \mathbf{c}$ é positivo, $\delta = 0$ se e só se $\mathbf{c}^t \boldsymbol{\beta}(l) - \mathbf{c}^t \boldsymbol{\beta}_0(l) = 0$.

4.1.3 Inferência para séries emparelhadas com estrutura de grau um

4.1.3.1 Análise longitudinal com um factor

Considerando o vector $Y = (Y_1, Y_2, \dots, Y_m)$ com $Y_l = \mathbf{c}^t \tilde{\boldsymbol{\beta}}(l) \sim N(\eta_l, \sigma^2 d_l)$, $l = 1, \dots, m$, e $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m)$ ter-se-à $Y \sim N(\boldsymbol{\eta}, \sigma^2 \mathbf{D})$ com \mathbf{D} a matriz diagonal de elementos principais d_1, d_2, \dots, d_m .

Com \mathbf{A} , matriz do tipo axm , de característica a , tomando $\tilde{\boldsymbol{\psi}} = \mathbf{A}Y$ e $\boldsymbol{\psi} = \mathbf{A}\boldsymbol{\eta}$ ter-se-à

$$\tilde{\boldsymbol{\psi}} \sim N(\boldsymbol{\psi}, \sigma^2 (\mathbf{A}\mathbf{D}\mathbf{A}^t)),$$

sendo $\mathbf{A}\mathbf{D}\mathbf{A}^t$ definida positiva, visto os vectores linha da matriz \mathbf{A} serem linearmente independentes.

Como a característica da matriz $\mathbf{A}\mathbf{D}\mathbf{A}^t$ é a tem-se, ver Mexia (1995)

$$(\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}_0)^t (\mathbf{A}\mathbf{D}\mathbf{A}^t)^{-1} (\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}_0) \sim \sigma^2 \chi_{a,\delta}^2$$

com

$$\delta = \frac{1}{\sigma^2} ((\boldsymbol{\psi} - \boldsymbol{\psi}_0)^t (\mathbf{A}\mathbf{D}\mathbf{A}^t)^{-1} (\boldsymbol{\psi} - \boldsymbol{\psi}_0)).$$

Como \mathbf{Y} é independente de V , $\tilde{\boldsymbol{\psi}}$ será independente de V e consequentemente

$$F_0 = \frac{g}{a} \frac{(\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}_0)^t (\mathbf{A}\mathbf{D}\mathbf{A}^t)^{-1} (\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}_0)}{V}$$

terá distribuição F com a e g graus de liberdade e parâmetro de não centralidade δ . Podemos pois utilizar esta estatística para testar a hipótese

$$H_0 : \psi = \psi_0.$$

Como o parâmetro de não centralidade se anula quando e só quando a hipótese H_0 se verifica, o teste será estritamente não distorcido.

Por outro lado a estatística

$$F' = \frac{g}{a} \frac{(\psi - \tilde{\psi})^t (\mathbf{A} \mathbf{D} \mathbf{A}^t)^{-1} (\psi - \tilde{\psi})}{V}$$

terá distribuição F central com a e g graus de liberdade, ver Mexia (1995). Sendo $F_{(1-q;a,g)}$ o quantil dessa distribuição para a probabilidade $1 - q$, ter-se-á

$$P \left[(\psi - \tilde{\psi})^t (\mathbf{A} \mathbf{D} \mathbf{A}^t)^{-1} (\psi - \tilde{\psi}) \leq a F_{(1-q;a,g)} \frac{V}{g} \right] = 1 - q,$$

obtendo-se assim um elipsóide de confiança de nível $1 - q$ para ψ . Observe-se que o teste F de nível q rejeita H_0 , se e só se ψ_0 não estiver contido no elipsóide de confiança de nível $1 - q$. Assim o teste F goza de dualidade.

4.1.3.2 Análise transversal com um factor

Para efectuarmos a análise transversal, continuamos a admitir que as séries de estudos correspondem aos m níveis de um factor e que o único vector de estrutura ajustado por série é $\tilde{\beta}(l) \sim N(\beta(l), \sigma^2 \mathbf{W}(l))$, $l = 1, \dots, m$.

Sendo $\tilde{\beta}_i^o = [\tilde{\beta}_i(1), \dots, \tilde{\beta}_i(m)]^t$, $i = 1, \dots, k$, o vector formado pelas i -ésimas componentes dos vectores de estrutura ajustados o mesmo terá como vector médio

$\beta_i^o = [\beta_i(1), \dots, \beta_i(m)]^t$ constituído pelas i -ésimas componentes dos vectores de estrutura.

Como $\tilde{\beta}(l) \sim N(\beta(l), \sigma^2 \mathbf{W}(l)), l = 1, \dots, m$ e estes vectores são independentes entre si e de $V \sim \sigma^2 \chi_g^2$, teremos $\tilde{\beta}_i^o \sim N(\beta_i^o, \sigma^2 \mathbf{M}_i)$ independente de V com a matriz diagonal $\mathbf{M}_i = \text{diag}(W_i(1)W_i(2)\dots W_i(m))$, cujos elementos principais são os i -ésimos elementos principais das matrizes $\mathbf{W}(l), l = 1, \dots, m$.

Para o contraste $\mathbf{c}^t \tilde{\beta}_i^o$ teremos

$$\mathbf{c}^t \tilde{\beta}_i^o \sim N(\mathbf{c}^t \beta_i^o, \sigma^2 \mathbf{c}^t \mathbf{M}_i \mathbf{c}), i = 1, \dots, k$$

e a estatística

$$\left(\mathbf{c}^t \tilde{\beta}_i^o - \mathbf{c}^t \beta_{0i}^o \right) \left(\mathbf{c}^t \mathbf{M}_i \mathbf{c} \right)^{-1} \left(\mathbf{c}^t \tilde{\beta}_i^o - \mathbf{c}^t \beta_{0i}^o \right) \sim \sigma^2 \chi_{1, \delta_i}^2, i = 1, \dots, k$$

com distribuição qui-quadrado independente de $V \sim \sigma^2 \chi_g^2$ e parâmetro de não centralidade

$$\delta_i = \frac{1}{\sigma^2} \left(\mathbf{c}^t \beta_i^o - \mathbf{c}^t \beta_{0i}^o \right) \left(\mathbf{c}^t \mathbf{M}_i \mathbf{c} \right)^{-1} \left(\mathbf{c}^t \beta_i^o - \mathbf{c}^t \beta_{0i}^o \right), i = 1, \dots, k$$

que se anula quando $\mathbf{c}^t \beta_i^o = \mathbf{c}^t \beta_{0i}^o$.

Para testar a hipótese

$$H_{0i} : \mathbf{c}^t \beta_i^o = \mathbf{c}^t \beta_{0i}^o, i = 1, \dots, k$$

utilizamos a estatística

$$F_0 = \frac{g \left(\mathbf{c}^t \tilde{\beta}_i^o - \mathbf{c}^t \beta_{0i}^o \right) \left(\mathbf{c}^t \mathbf{M}_i \mathbf{c} \right)^{-1} \left(\mathbf{c}^t \tilde{\beta}_i^o - \mathbf{c}^t \beta_{0i}^o \right)}{1 \quad V} \sim F(.|1, g, \delta_i)$$

que tem distribuição F com 1 e g graus de liberdade e parâmetro de não centralidade δ_i , ver Mexia (1995).

Como $\mathbf{c}^t \mathbf{M}_i \mathbf{c} > 0$ tem-se $\delta_i = 0$, quando e só quando $\mathbf{c}^t \boldsymbol{\beta}_i^o = \mathbf{c}^t \boldsymbol{\beta}_{0i}^o, i = 1, \dots, k$.

Assim a hipótese H_{0i} verifica-se se e só se $\delta_i = 0$.

Dada a matriz \mathbf{A} do tipo axm com vectores linha linearmente independentes, as matrizes $\mathbf{A} \mathbf{M}_i \mathbf{A}^t, i = 1, \dots, k$, serão definidas positivas.

Tomando-se $\tilde{\boldsymbol{\varphi}}_i = \mathbf{A} \tilde{\boldsymbol{\beta}}_i^o, i = 1, \dots, k$, e $\boldsymbol{\varphi}_i = \mathbf{A} \boldsymbol{\beta}_i^o, i = 1, \dots, k$, ter-se-á

$$\tilde{\boldsymbol{\varphi}}_i \sim N(\boldsymbol{\varphi}_i, \sigma^2 \mathbf{A} \mathbf{M}_i \mathbf{A}^t), i = 1, \dots, k$$

pelo que

$$F_i = \frac{g (\tilde{\boldsymbol{\varphi}}_i - \boldsymbol{\varphi}_{0i})^t (\mathbf{A} \mathbf{M}_i \mathbf{A}^t)^{-1} (\tilde{\boldsymbol{\varphi}}_i - \boldsymbol{\varphi}_{0i})}{a}, i = 1, \dots, k$$

terá distribuição F com a e g graus de liberdade e parâmetro de não centralidade

$$\delta_i = \frac{1}{\sigma^2} (\boldsymbol{\varphi}_i - \boldsymbol{\varphi}_{0i})^t (\mathbf{A} \mathbf{M}_i \mathbf{A}^t)^{-1} (\boldsymbol{\varphi}_i - \boldsymbol{\varphi}_{0i}), i = 1, \dots, k$$

que só se anula quando

$$H_{0i} : \boldsymbol{\varphi}_i = \boldsymbol{\varphi}_{0i}, i = 1, \dots, k,$$

se verifica.

Assim o teste F para H_{0i} com estatística F_i , será estritamente não distorcido.

Podemos obter os elipsóides de confiança de nível $1 - q$ para $\boldsymbol{\varphi}_i$, dados por

$$P \left[(\boldsymbol{\varphi}_i - \tilde{\boldsymbol{\varphi}}_i)^t (\mathbf{A} \mathbf{M}_i \mathbf{A}^t)^{-1} (\boldsymbol{\varphi}_i - \tilde{\boldsymbol{\varphi}}_i) \leq a F_{(1-q; a, g)} \frac{V}{g} \right] = 1 - q, i = 1, \dots, k$$

e os intervalos de confiança simultâneos para $\mathbf{c}^t \boldsymbol{\varphi}_i$ resultantes de

$$P \left[\bigcap_{\mathbf{c}} |\mathbf{c}^t \boldsymbol{\varphi}_i - \mathbf{c}^t \tilde{\boldsymbol{\varphi}}_i| \leq \sqrt{a F_{(1-q; a, g)} \mathbf{c}^t (\mathbf{A} \mathbf{M}_i \mathbf{A}^t) \mathbf{c} \frac{V}{g}} \right] = 1 - q.$$

Os testes F gozam de dualidade, são não distorcidos e uniformemente mais potentes (UMP) para a família dos testes, cuja potência é função dos parâmetros de não centralidade, ver Lehmann (1959).

4.1.4 Inferência com vários factores e estrutura de grau um

Admitamos que para todas as combinações de níveis de L factores se têm séries emparelhadas de estudos com estrutura comum de grau um. Sendo k o número de estudos das séries, a informação contida nas mesmas pode ser condensada, como vimos, num vector de estrutura ajustado e numa soma de quadrados de resíduos para cada série.

As análises longitudinal e transversal são efectuadas, recorrendo às matrizes ortogonais estandardizadas. Caso os factores tenham J_1, \dots, J_L níveis as combinações de níveis dos factores podem ser indicadas pelos vectores \mathbf{j} com componentes $j_l = 1, \dots, J_L, l = 1, \dots, L$. Assim é possível ordenar as séries de estudos, utilizando os índices

$$i(\mathbf{j}) = j_1 + \sum_{l=2}^L (j_l - 1) \prod_{h=1}^{l-1} J_h. \quad (4.1)$$

Este índice variará de 1 a $m = \prod_{h=1}^L J_h$, havendo aliás m tratamentos definidos pelas combinações de níveis dos factores.

Então, com $l = i(\mathbf{j})$ ter-se-ão os vectores de estrutura ajustados

$$\tilde{\boldsymbol{\beta}}(l) \sim N(\boldsymbol{\beta}(l), \sigma^2 \mathbf{W}(l)), l = 1, \dots, m,$$

independentes entre si e de $V \sim \sigma^2 \chi_g^2$.

4.1.4.1 Análise longitudinal

Estamos num quadro linear do tipo semelhante ao considerado em modelos de efeitos fixos, quando se utiliza a análise de variância (ANOVA). Para realizar a ANOVA correspondente ao estudo dos efeitos e interações dos factores sobre os valores dos contrastes $\eta_l, l = 1, \dots, m$, pode-se utilizar um algoritmo baseado no produto de Kronecker de matrizes. Este algoritmo vem descrito com detalhe em Oliveira (2002).

Dada a matrizes $\mathbf{U} = [u_{lj}]$ do tipo $r \times s$ e a matriz \mathbf{B} , o respectivo produto de Kronecker é dado por

$$\mathbf{U} \otimes \mathbf{B} = \begin{bmatrix} u_{11}\mathbf{B} & u_{12}\mathbf{B} & \dots & u_{1s}\mathbf{B} \\ u_{21}\mathbf{B} & u_{22}\mathbf{B} & \dots & u_{2s}\mathbf{B} \\ \dots & \dots & \dots & \dots \\ u_{r1}\mathbf{B} & u_{r2}\mathbf{B} & \dots & u_{rs}\mathbf{B} \end{bmatrix}$$

Tomando agora matrizes $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_L$ de ordem J_1, J_2, \dots, J_L ortogonais estandarizadas, obtemos a matriz

$$\mathbf{P} = \mathbf{P}_L \otimes \dots \otimes \mathbf{P}_1$$

que é ortogonal estandarizada porque o produto de Kronecker de matrizes ortogonais estandarizadas dá matrizes ortogonais estandarizadas.

Toda a linha de \mathbf{P} é o produto de Kronecker sucessivo de linhas de $\mathbf{P}_1, \dots, \mathbf{P}_L$. Assim, pode associar-se a cada linha de \mathbf{P} , o conjunto C dos índices das matrizes $\mathbf{P}_1, \dots, \mathbf{P}_L$, cuja primeira linha não foi utilizada para a formação da linha de \mathbf{P} .

Seja $\mathbf{A}(C)$ a submatriz de P formada pelas linhas associadas a $C \subseteq \{1, \dots, L\}$. Em particular $\mathbf{A}(\emptyset)$ será formada pela primeira linha de \mathbf{P} .

Considere-se o vector $\mathbf{Y} = (Y_1, \dots, Y_m) \sim N(\boldsymbol{\eta}, \sigma^2 \mathbf{D})$, cujas componentes correspondem aos m tratamentos, com a matriz diagonal

$$\mathbf{D} = \text{diag}(\mathbf{c}^t \mathbf{W}(1) \mathbf{c}, \mathbf{c}^t \mathbf{W}(2) \mathbf{c}, \dots, \mathbf{c}^t \mathbf{W}(m) \mathbf{c}).$$

Sendo a hipótese de ausência de efeitos do factor ou de ausência de interacção entre os factores com índices em C , pode a mesma ser escrita como

$$H_0(C) : \mathbf{A}(C) \boldsymbol{\eta} = (0, \dots, 0)^t,$$

ver Mexia (1995) e Oliveira (2002).

Para $\tilde{\boldsymbol{\psi}}(C) = \mathbf{A}(C) \mathbf{Y}$ e $\boldsymbol{\psi}(C) = \mathbf{A}(C) \boldsymbol{\eta}$, teremos

$$\tilde{\boldsymbol{\psi}}(C) \sim N(\boldsymbol{\psi}(C), \sigma^2 \mathbf{A}(C) \mathbf{D} \mathbf{A}(C)^t); C \subseteq \{1, \dots, L\}.$$

A matriz $\mathbf{A}(C)$ tem $g(C) = \prod_{l \in C} (J_l - 1)$; $C \subseteq \{1, \dots, L\}$, vectores linha linearmente independentes pelo que $(\mathbf{A}(C) \mathbf{D} \mathbf{A}(C)^t)$ será definida positiva com característica $g(C)$.

Então,

$$\left(\tilde{\boldsymbol{\psi}}(C) - \mathbf{b} \right)^t (\mathbf{A}(C) \mathbf{D} \mathbf{A}(C)^t)^{-1} \left(\tilde{\boldsymbol{\psi}}(C) - \mathbf{b} \right) \sim \sigma^2 \chi_{g(C), \delta(C, \mathbf{b})}^2$$

com parâmetro de não centralidade

$$\delta(C, \mathbf{b}) = \frac{1}{\sigma^2} \left((\boldsymbol{\psi}(C) - \mathbf{b})^t (\mathbf{A}(C) \mathbf{D} \mathbf{A}(C)^t)^{-1} (\boldsymbol{\psi}(C) - \mathbf{b}) \right); C \subseteq \{1, \dots, L\},$$

ver Mexia (1995).

Como $\tilde{\psi}(C)$ é independente de $V \sim \sigma^2 \chi_g^2$, a estatística

$$F(C, b) = \frac{g}{g(C)} \frac{\left(\tilde{\psi}(C) - \mathbf{b}\right)^t (\mathbf{A}(C)\mathbf{D}\mathbf{A}(C)^t)^{-1} \left(\tilde{\psi}(C) - \mathbf{b}\right)}{V}; C \subseteq \{1, \dots, L\}$$

terá distribuição F com $g(c)$ e g graus de liberdade e parâmetro de não centralidade $\delta(C, b)$ que se anula se e só se a hipótese

$$H_0(C, \mathbf{b}) : \psi(C) = \mathbf{b}$$

se verifica. Assim o teste F com estatística $F(C, \mathbf{b})$ será estritamente não distorcido, ver Mexia (1995).

Observamos que a hipótese $H_0(C, \mathbf{b})$ representa uma generalização da hipótese $H_0(C)$ já que

$$H_0(C) = H_0(C, 0); C \subseteq \{1, \dots, L\}.$$

Acresce ainda que

$$F'(C) = \frac{g}{g(C)} \frac{\left(\psi(C) - \tilde{\psi}(C)\right)^t (\mathbf{A}(C)\mathbf{D}\mathbf{A}(C)^t)^{-1} \left(\psi(C) - \tilde{\psi}(C)\right)}{V}; C \subseteq \{1, \dots, L\}$$

tem distribuição F central com $g(c)$ e g graus de liberdade. Assim, sendo $F_{1-q, g(c), g}$ o quantil para a probabilidade $1 - q$ dessa distribuição, tem-se para $\psi(C)$ o elipsóide de confiança de nível $1 - q$, resultando de

$$P \left(\left(\psi(C) - \tilde{\psi}(C)\right)^t (\mathbf{A}(C)\mathbf{D}\mathbf{A}(C)^t)^{-1} \left(\psi(C) - \tilde{\psi}(C)\right) \leq g(C) F_{1-q, g(c), g} \frac{V}{g} \right) =$$

$1 - q$.

Observa-se que o teste de nível q com estatística $F(C, \mathbf{b})$ rejeita $H_0(C, \mathbf{b})$, se e só se \mathbf{b} não estiver contido no elipsóide de nível $1 - q$, pelo que os testes F atrás

construídos gozam de dualidade.

Assim, ao alargar-se a família das hipóteses testadas, surge nova propriedade - a dualidade para os testes F . Esta perspectiva encontra-se desenvolvida em Mexia (1989).

4.1.4.2 Análise transversal

Para considerarmos os i -ésimos estudos utilizamos o vector $\xi_i, i = 1, \dots, k$, com todas as componentes nulas salvo a i -ésima que é igual a 1.

Como $\tilde{\beta}(l) \sim N(\beta(l), \sigma^2 \mathbf{W}(l)), l = 1, \dots, m$, então $\xi_i^t \tilde{\beta}(l)$ tem valor médio $\xi_i^t \beta(l)$ e variância $\sigma^2 W_i(l)$, ou seja,

$$\xi_i^t \tilde{\beta}(l) \sim N(\xi_i^t \beta(l), \sigma^2 W_i(l)), i = 1, \dots, k, l = 1, \dots, m.$$

O vector

$$\tilde{\Upsilon}_i = [\xi_i^t \tilde{\beta}(1), \xi_i^t \tilde{\beta}(2), \dots, \xi_i^t \tilde{\beta}(m)]^t$$

tem o vector médio

$$\Upsilon_i = [\xi_i^t \beta(1), \xi_i^t \beta(2), \dots, \xi_i^t \beta(m)]^t$$

e a matriz de covariância $\sigma^2 \mathbf{M}_i$ com,

$$\mathbf{M}_i = \text{diag}[W_i(1), W_i(2), \dots, W_i(m)], i = 1, \dots, k.$$

Assim,

$$\tilde{\varphi}_i(C) = \mathbf{A}(C) \tilde{\Upsilon}_i \sim N(\varphi_i(C), \sigma^2 (\mathbf{A}(C) \mathbf{M}_i \mathbf{A}(C)^t)); C \subseteq \{1, \dots, L\}, i = 1, \dots, k,$$

com $\varphi_i(C) = \mathbf{A}(C) \Upsilon_i$ e dado \mathbf{M}_i ser uma matriz definida positiva

$$\begin{aligned}
g(C) &= \text{car} (\mathbf{A}(C)\mathbf{M}_i\mathbf{A}(C)^t) \\
&= \text{car} (\mathbf{A}(C)) \\
&= \prod_{l \in C} (J_l - 1); C \subseteq \{1, \dots, L\}.
\end{aligned}$$

Como os vectores $\tilde{\boldsymbol{\beta}}(l) \sim N(\boldsymbol{\beta}(l), \sigma^2 \mathbf{W}(l))$, $l = 1, \dots, m$, são independentes entre si e de $V \sim \sigma^2 \chi_g^2$, então $\tilde{\boldsymbol{\varphi}}_i(C)$ também é independente de V .

Para testar a hipótese

$$H_0(C) : \boldsymbol{\varphi}_i(C) = \boldsymbol{\varphi}_i^0(C), i = 1, \dots, k$$

utilizamos a seguinte estatística:

$$F(C, \boldsymbol{\varphi}_i(C)) = \frac{g}{g(C)} \frac{(\tilde{\boldsymbol{\varphi}}_i(C) - \boldsymbol{\varphi}_i^0(C))^t (\mathbf{A}(C)\mathbf{M}_i\mathbf{A}(C)^t)^{-1} (\tilde{\boldsymbol{\varphi}}_i(C) - \boldsymbol{\varphi}_i^0(C))}{V};$$

$C \subseteq \{1, \dots, L\}$, que tem distribuição F com $g(C)$ e g graus de liberdade, em que a variável

$$(\tilde{\boldsymbol{\varphi}}_i(C) - \boldsymbol{\varphi}_i^0(C))^t (\mathbf{A}(C)\mathbf{M}_i\mathbf{A}(C)^t)^{-1} (\tilde{\boldsymbol{\varphi}}_i(C) - \boldsymbol{\varphi}_i^0(C)) \sim \sigma^2 \chi_{g(C), \delta(C, \boldsymbol{\varphi}_i)}^2$$

é independente de V e com parâmetro de não centralidade

$$\delta(C, \boldsymbol{\varphi}_i) = \frac{1}{\sigma^2} \left[(\boldsymbol{\varphi}_i(C) - \boldsymbol{\varphi}_i^0(C))^t (\mathbf{A}(C)\mathbf{M}_i\mathbf{A}(C)^t)^{-1} (\boldsymbol{\varphi}_i(C) - \boldsymbol{\varphi}_i^0(C)) \right],$$

$C \subseteq \{1, \dots, L\}$, que se anula se e só se a hipótese $H_0(C)$ se verificar, ver Mexia (1995). Assim o teste F com estatística $F(C, \boldsymbol{\varphi}_i(C))$ será estritamente não distorcido e verifica a dualidade.

4.1.5 Inferência com vários factores e estrutura com grau h

Até agora considerámos nos vários modelos apenas estruturas comuns de grau um. Quando o grau da estrutura é h as hipóteses apresentadas nas secções anteriores podem ser definidas para os vários vectores de estrutura.

Considere $\tilde{\boldsymbol{\beta}}^h(l) = [\tilde{\boldsymbol{\beta}}(1, l), \tilde{\boldsymbol{\beta}}(2, l), \dots, \tilde{\boldsymbol{\beta}}(h, l)]$, $l = 1, \dots, m$, formada por h vectores de estrutura ajustados para a l -ésima série e $V(l)$ a soma de quadrados dos resíduos. Verifica-se que os vectores de estrutura ajustados

$$\tilde{\boldsymbol{\beta}}(j, l) \sim N(\boldsymbol{\beta}(j, l), \sigma^2 \mathbf{W}(j, l)), j = 1, 2, \dots, h; l = 1, \dots, m,$$

são independentes entre si e de $V \sim \sigma^2 \chi_g^2$, com $V = \sum_{l=1}^m V(l)$.

4.1.5.1 Análise longitudinal

O tratamento das hipóteses relativamente a qualquer vector da estrutura $\tilde{\boldsymbol{\beta}}(j, l)$, $j = 1, 2, \dots, h; l = 1, \dots, m$, é exactamente o mesmo que para as hipóteses relativas ao único vector de estrutura no caso do grau ser um. Teremos ainda as hipóteses $H_0(C | \mathbf{b}_1, \dots, \mathbf{b}_h) = \cap_{j=1}^h H_{0,j}(C, \mathbf{b}_j)$ que é a hipótese global relativa a todos os vectores de estrutura e que se verifica se e só se verificarem as hipóteses relativas ao j -ésimo vector de estrutura; $H_{0,j}(C, \mathbf{b}_j)$, $j = 1, \dots, h$.

Seja

$$\tilde{\boldsymbol{\psi}}_j(C) = \mathbf{A}(C) \mathbf{c}^t [\tilde{\boldsymbol{\beta}}(j, 1), \tilde{\boldsymbol{\beta}}(j, 2), \dots, \tilde{\boldsymbol{\beta}}(j, m)], j = 1, \dots, h$$

e

$$\boldsymbol{\psi}_j(C) = \mathbf{A}(C)\mathbf{c}^t [\boldsymbol{\beta}(j, 1), \boldsymbol{\beta}(j, 2), \dots, \boldsymbol{\beta}(j, m)], j = 1, \dots, h$$

então

$$\tilde{\boldsymbol{\psi}}_j(C) \sim N(\boldsymbol{\psi}_j(C), \sigma^2 (\mathbf{A}(C)\mathbf{D}_j\mathbf{A}(C)^t)), j = 1, \dots, h,$$

com \mathbf{D}_j a matriz diagonal com elementos principais

$$\mathbf{D}_j = \text{diag}(\mathbf{c}^t\mathbf{W}(j, 1)\mathbf{c}, \mathbf{c}^t\mathbf{W}(j, 2)\mathbf{c}, \dots, \mathbf{c}^t\mathbf{W}(j, m)\mathbf{c}), j = 1, \dots, h.$$

Dada a independência dos vários vectores de estrutura ajustados para os diferentes estudos, as formas quadráticas

$$\mathbf{U}_j(C, \mathbf{b}_j) = \left(\tilde{\boldsymbol{\psi}}_j(C) - \mathbf{b}_j \right)^t (\mathbf{A}(C)\mathbf{D}_j\mathbf{A}(C)^t)^{-1} \left(\tilde{\boldsymbol{\psi}}_j(C) - \mathbf{b}_j \right) \sim \sigma^2 \chi_{g(C), \delta_j(C, \mathbf{b}_j)}^2,$$

$j = 1, \dots, h, C \subseteq \{1, \dots, L\}$ serão independentes entre si. Atendendo à reprodutibilidade dos qui-quadrados ter-se-á

$$\sum_{j=1}^h \mathbf{U}_j(C, \mathbf{b}_j) \sim \sigma^2 \chi_{hg(C), \sum_{j=1}^h \delta_j(C, \mathbf{b}_j)}^2,$$

independente de $V \sim \sigma^2 \chi_g^2$ e

$$\delta_j(C, \mathbf{b}_j) = \frac{1}{\sigma^2} \left((\boldsymbol{\psi}_j(C) - \mathbf{b}_j)^t (\mathbf{A}(C)\mathbf{D}_j\mathbf{A}(C)^t)^{-1} (\boldsymbol{\psi}_j(C) - \mathbf{b}_j) \right), j = 1, \dots, h;$$

$C \subseteq \{1, \dots, L\}$.

O teste F para a hipótese

$$H_0(C|\mathbf{b}_1, \dots, \mathbf{b}_h) = \boldsymbol{\psi}_1(C) = \mathbf{b}_1, \boldsymbol{\psi}_2(C) = \mathbf{b}_2, \dots, \boldsymbol{\psi}_h(C) = \mathbf{b}_h$$

com estatística

$$F(C|\mathbf{b}_1, \dots, \mathbf{b}_h) = \frac{g}{hg(C)} \frac{\sum_{j=1}^h \mathbf{U}_j(C, \mathbf{b}_j)}{V}; C \subseteq \{1, \dots, L\}$$

terá $hg(c)$ e g graus de liberdade e parâmetro de não centralidade

$$\delta(C|\mathbf{b}_1, \dots, \mathbf{b}_h) = \sum_{j=1}^h \delta_j(C, \mathbf{b}_j)$$

que se anula quando e só quando $H_0(C|\mathbf{b}_1, \dots, \mathbf{b}_h)$ se verifica, ou seja, se e só se todos os $\delta_j(C, \mathbf{b}_j) = 0$.

Este teste F é não distorcido e uniformemente mais potente (UMP) para a família dos testes cuja potência é função dos parâmetros de não centralidade.

4.1.5.2 Análise transversal

No âmbito da análise transversal, se consideramos o vector $\boldsymbol{\xi}_i, i = 1, \dots, k$, com todas as componentes nulas, salvo a i -ésima que é igual a 1, obtemos

$$\boldsymbol{\xi}_i^t \tilde{\boldsymbol{\beta}}(j, l) \sim N(\boldsymbol{\xi}_i^t \boldsymbol{\beta}(j, l), \sigma^2 W_i(j, l)), j = 1, \dots, h, l = 1, \dots, m.$$

Seja

$$\tilde{\boldsymbol{\Upsilon}}_{i,j} = [\boldsymbol{\xi}_i^t \tilde{\boldsymbol{\beta}}(j, 1), \boldsymbol{\xi}_i^t \tilde{\boldsymbol{\beta}}(j, 2), \dots, \boldsymbol{\xi}_i^t \tilde{\boldsymbol{\beta}}(j, m)]^t$$

com vector médio

$$\boldsymbol{\Upsilon}_{i,j} = [\boldsymbol{\xi}_i^t \boldsymbol{\beta}(j, 1), \boldsymbol{\xi}_i^t \boldsymbol{\beta}(j, 2), \dots, \boldsymbol{\xi}_i^t \boldsymbol{\beta}(j, m)]^t$$

e

$$\tilde{\varphi}_i(C, j) = \mathbf{A}(C) \tilde{\boldsymbol{\Upsilon}}_{i,j}, i = 1, \dots, k, j = 1, \dots, h$$

com distribuição normal

$$\tilde{\varphi}_i(C, j) \sim N(\boldsymbol{\varphi}_i(C, j), \sigma^2 (\mathbf{A}(C) \mathbf{M}_i(j) \mathbf{A}(C)^t)), j = 1, \dots, h$$

em que

$$\varphi_i(C, j) = \mathbf{A}(C)\Upsilon_{i,j}$$

e a matriz diagonal

$$\mathbf{M}_i(j) = \mathbf{diag}(W_i(j, 1), W_i(j, 2), \dots, W_i(j, m)).$$

Dada a independência dos vários vectores de estrutura ajustados para os diferentes estudos, as formas quadráticas

$$\mathbf{T}_i(C, j) = \left((\tilde{\varphi}_i(C, j) - \varphi_i^0(C, j))^t (\mathbf{A}(C)\mathbf{M}_i(j)\mathbf{A}(C)^t)^{-1} (\tilde{\varphi}_i(C, j) - \varphi_i^0(C, j)) \right),$$

$j = 1, \dots, h, C \subseteq \{1, \dots, L\}$, têm a seguinte distribuição:

$$\mathbf{T}_i(C, j) \sim \sigma^2 \chi_{g(C), \delta(C, \varphi_i(j))}^2, j = 1, \dots, h; C \subseteq \{1, \dots, L\}$$

com

$$\delta(C, \varphi_i(j)) = \left((\varphi_i(C, j) - \varphi_i^0(C, j))^t (\mathbf{A}(C)\mathbf{M}_i(j)\mathbf{A}(C)^t)^{-1} (\varphi_i(C, j) - \varphi_i^0(C, j)) \right)$$

e

$$\sum_{j=1}^h \mathbf{T}_i(C, j) \sim \sigma^2 \chi_{hg(C), \sum_{j=1}^h \delta(C, \varphi_i(j))}^2,$$

devido à reprodutividade dos qui-quadrados.

Para testar a hipótese

$$H_{i0}(C, j) : \varphi_i(C, 1) = \varphi_i^0(C, 1), \varphi_i(C, 2) = \varphi_i^0(C, 2), \dots, \varphi_i(C, h) = \varphi_i^0(C, h); i = 1, \dots, k,$$

utilizamos a estatística

$$F_i(C, j) = \frac{g}{hg(C)} \frac{\sum_{j=1}^h \mathbf{T}_i(C, j)}{V}; C \subseteq \{1, \dots, L\}$$

que tem distribuição F com $hg(c)$ e g graus de liberdade, em que a variável $\sum_{j=1}^h \mathbf{T}_i(C, j)$ é independente de V e com parâmetro de não centralidade $\sum_{j=1}^h \delta(C, \varphi_i(j))$ que se anula se e só se a hipótese $H_{i0}(C, j)$ se verifica, ou seja, se e só se todos os $\delta(C, \varphi_i(j)) = 0$.

Tal como no caso anterior, este teste F é não distorcido e UMP para a família dos testes cuja potência é função dos parâmetros de não centralidade.

4.2 Aplicação: Eleições Autárquicas

4.2.1 Modelo com componente aleatória em A_i

Na sub-secção 3.2.4.2 analisámos os resultados eleitorais de cinco eleições autárquicas realizadas em Portugal entre 1985 e 2001, em 16 concelhos. O modelo ajustado admitia estrutura comum de grau um. Os resultados das estatísticas dos testes F de validação do modelo foram apresentados no cap. 3, tabelas 3.4 e 3.5, pág. 48. Pretendemos agora, utilizando os vectores de estrutura ajustados (tabela 3.3, pág. 47) e a soma de quadrados dos resíduos, realizar:

- ▶ a análise longitudinal para as séries individuais;
- ▶ a análise longitudinal e transversal sobre as séries emparelhadas de estudos.

4.2.1.1 Análise longitudinal para as séries individuais

Pretendemos verificar, através de uma análise longitudinal efectuada aos 16 concelhos, se houve diferenças nos resultados eleitorais nos vários anos, aplicando os seguintes vectores de contraste aos vectores de estrutura estimados:

$$\mathbf{c}_1^t = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

$$\mathbf{c}_2^t = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

e

$$\mathbf{c}_3^t = \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 & -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{6}} \end{bmatrix}.$$

Consideram-se assim definidos os contrastes:

► $L_{1l} = \mathbf{c}_1^t \tilde{\boldsymbol{\beta}}(l), l = 1, \dots, 16$ – para determinar a evolução entre os resultados do l -ésimo concelho entre 2001 e 1985;

► $L_{2l} = \mathbf{c}_2^t \tilde{\boldsymbol{\beta}}(l), l = 1, \dots, 16$ – para determinar a evolução entre os resultados do l -ésimo concelho entre 1997 e 1989;

► $L_{3l} = \mathbf{c}_3^t \tilde{\boldsymbol{\beta}}(l), l = 1, \dots, 16$ – para avaliar a curvatura da evolução das várias opções políticas.

Na tabela 4.1 apresentamos para os 16 concelhos os valores dos contrastes obtidos.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F	Gua	Pin
L_{1l}	6.179	2.652	5.203	-1.712	0.433	6.336	3.654	6.196
L_{2l}	3.578	0.284	1.901	4.312	-5.377	4.849	3.717	12.769
L_{3l}	5.122	-7.463	-0.370	-4.209	0.283	-1.479	-2.103	10.107
	Lei	Alc.	Por	Elv	Far	Sil	Bej	Mer
L_{1l}	2.505	-15.122	4.092	25.627	5.683	4.334	0.817	-3.010
L_{2l}	-0.676	-1.295	-2.286	18.618	1.061	-4.004	2.152	-2.316
L_{3l}	2.962	7.383	-2.574	18.027	4.503	2.005	-0.748	1.876

Tabela 4.1: Valores dos contrastes para os 16 concelhos.

Rejeita-se a hipótese

$$H_{0i}(l) : \mathbf{c}_i^t \boldsymbol{\beta}(l) = 0, i = 1, 2, 3; l = 1, \dots, 16$$

ao nível α , se

$$F_{0i}(l) = \frac{g \mathbf{c}_i^t \tilde{\boldsymbol{\beta}}(l) (\mathbf{c}_i^t \mathbf{W}(l) \mathbf{c}_i)^{-1} \mathbf{c}_i^t \tilde{\boldsymbol{\beta}}(l)}{1} > F_{(1-\alpha, 1, g)}.$$

Os resultados obtidos para as estatísticas encontram-se na tabela 4.2:

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F	Gua	Pin
$F_{01}(l)$	20.71	2.66	8.90	1.32	0.07	18.57	4.43	16.27
$F_{02}(l)$	7.03	0.03	1.19	8.41	11.48	10.95	4.59	67.44
$F_{03}(l)$	14.32	20.97	0.04	8.01	0.03	1.02	1.47	42.75
	Lei	Alc.	Por	Elv	Far	Sil	Bej	Mer
$F_{01}(l)$	2.62	84.44	7.69	198.63	12.04	8.81	0.24	3.04
$F_{02}(l)$	0.19	0.64	2.41	109.33	0.38	7.87	1.72	1.80
$F_{03}(l)$	3.66	20.81	3.05	102.88	7.07	2.01	0.21	1.18

Tabela 4.2: Valores das estatísticas $F_{0i}(l)$.

Como o valor crítico é $F_{(0.95; 1, 159)} = 3.90$ e, pela observação da tabela 4.2, conclui-se que:

- os resultados obtidos na 5.^a eleição autárquica não são significativamente diferentes dos obtidos na 1.^a eleição autárquica, nos concelhos de Ponta da Barca, Alfandega da Fé, Aveiro, Leiria, Beja e Mértola;
- os resultados obtidos na 4.^a eleição autárquica não são significativamente diferentes dos obtidos na 2.^a eleição, nos concelhos de Ponta da Barca, Bragança, Leiria, Alcobaça, Portalegre, Faro, Beja e Mértola;
- no período de 1985 a 2001, os resultados são significativos ao nível da curvatura,

nos concelhos de Viana do Castelo, Ponta da Barca, Alfândega da Fé, Pinhel, Alcobaça, Elvas e Faro.

4.2.1.2 Análise longitudinal e transversal sobre as séries emparelhadas

A análise longitudinal e transversal é obtida através da utilização de matrizes ortogonais estandardizadas para estudar os efeitos dos factores e suas interações nos resultados eleitorais. Assim indexamos as séries (ver 4.1, pág.65) de acordo com os níveis dos factores, obtendo-se: 1-Viana Castelo; 2-Ponta da Barca; 3-Bragança; 4-Alfandega da Fé; 5-Aveiro; 6-Santa Maria da Feira; 7-Guarda; 8-Pinhel; 9-Leiria; 10-Alcobaça; 11-Portalegre; 12-Elvas; 13-Faro; 14-Silves; 15-Beja e 16-Mértola.

Temos os seguintes factores com os seguintes níveis:

- ▶ 1.º factor com dois níveis: Capital de Distrito e Não Capital;
- ▶ 2.º factor com dois níveis: Litoral e Interior;
- ▶ 3.º factor com quatro níveis: Norte, Centro Norte, Centro e Sul.

A estes factores factores correspondem as matrizes ortogonais estandardizadas

$$\mathbf{P}_1 = \mathbf{P}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

e

$$\mathbf{P}_3 = \mathbf{P}_2 \otimes \mathbf{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

Conjunto	Linha	Factores
$\{\emptyset\}$	1	
$\{1\}$	2	1º
$\{2\}$	3	2º
$\{1, 2\}$	4	1º e 2º
$\{3\}$	5, 9 e 13	3º
$\{1, 3\}$	6, 10 e 14	1º e 3º
$\{2, 3\}$	7, 11 e 15	2º e 3º
$\{1, 2, 3\}$	8, 12 e 16	1º, 2º e 3º

Tabela 43: Conjuntos associados aos factores.

4.2.1.2.1 Análise longitudinal

Vamos agora testar os efeitos dos factores e suas interacções nos valores dos contrastes, apresentados na tabela 4.1, pág.77.

Rejeitamos a hipótese

$$H_0(C) : \psi(C)=0; C \subseteq \{1, 2, 3\}$$

ao nível α , quando a estatística

$$F(C) = \frac{g}{g(C)} \frac{\left(\tilde{\psi}(C)\right)^t \left(\mathbf{A}(C)\mathbf{D}\mathbf{A}(C)^t\right)^{-1} \left(\tilde{\psi}(C)\right)}{V} > F_{(1-\alpha;g(C),g)}.$$

■ Efeitos do 1º, 2º e 3º factor

Na tabela 4.4 encontram-se os valores das estatísticas $F(C)$ para os efeitos dos

factores nos valores dos contrastes.

	$F\{1\}$	$F\{2\}$	$F\{3\}$
L_{1l}	0.261	19.010	1.803
L_{2l}	20.846	40.140	7.123
L_{3l}	9.042	1.099	17.88

Tabela 4.4: Estatísticas $F(C)$ para efeitos do 1.º, 2.º e 3.º factor.

Como o valor crítico é $F_{(0.95;1,159)} = 3.90$, e pela observação da tabela, concluímos que relativamente ao:

- 1.º factor, apenas são significativos os valores dos testes para o 2.º e 3.º contraste;
- 2.º factor, apenas são significativos os valores dos testes para o 1.º e 2.º contrastes;

Para o valor crítico de $F_{(0.95;3,159)} = 2.66$, verificamos que os efeitos do 3.º factor apenas são significativos para o 2.º e 3.º contraste.

■ Interacção entre os factores

Na tabela 4.5 encontram-se os valores das estatísticas dos testes F para a interacção dos factores relativamente aos valores dos contrastes:

	$F\{1, 2\}$	$F\{1, 3\}$	$F\{2, 3\}$	$F\{1, 2, 3\}$
L_{1l}	21.934	7.327	56.740	42.674
L_{2l}	17.426	21.353	7.495	10.341
L_{3l}	47.687	30.354	4.360	2.426

Tabela 4.5: Estatísticas $F(C)$ para efeitos da interacção entre os factores.

Como o valor crítico é $F_{(0.95;1,159)} = 3.90$, concluímos que a interacção entre o 1º e 2º factor é significativa para todos os contrastes.

Para o valor crítico $F_{(0.95;3,159)} = 2.66$, verifica-se que a interacção entre o:

- 1º e 3º factor é significativa em todos os contrastes, mas a interacção é mais fraca para o 1º contraste;
- 2º e 3º factor é significativa em todos os contrastes, mas mais elevada relativamente ao 1º contraste;
- 1º, 2º e 3º factor, apenas é significativa relativamente ao 1º e 2º contraste.

4.2.1.2.2 Análise transversal

No âmbito desta análise, pretendemos verificar os efeitos e interacções dos vários factores nos resultados eleitorais obtidos nos 16 concelhos para os diferentes anos.

Rejeita-se a hipótese

$$H_0(C) : \varphi_i(C) = 0, i = 1, \dots, 5; C \subseteq \{1, 2, 3\}$$

ao nível α , se:

$$F_i(C) = \frac{g}{g(C)} \frac{(\tilde{\varphi}_i(C))^t (\mathbf{A}(C) \mathbf{M}_i \mathbf{A}(C)^t)^{-1} (\tilde{\varphi}_i(C))}{V} > F_{1-\alpha, g(c), g}.$$

■ Efeitos do 1º 2º e 3º factor

Os valores das estatísticas dos testes encontram-se na tabela 4.6.

Ano	$F_i\{1\}$	$F_i\{2\}$	$F_i\{3\}$
1985	0.3127	13.895	9.735
1989	47.0351	9.553	17.156
1993	7.2363	26.907	6.568
1997	2.7077	100.854	4.850
2001	0.0004	71.323	7.754

Tabela 4.6: Estatísticas $F_i(C)$ para efeitos do 1º, 2º e 3º factor.

Dado ter-se para o valor crítico de $F_{(0.95;1,159)} = 3.90$, conclui-se que:

- não há diferença significativa nos resultados eleitorais obtidos entre a capital de distrito e o outro concelho não capital de distrito, nos anos de 85, 97 e 2001;
- os efeitos do 1º factor são muito elevados em 1989, mas nas eleições seguintes esses efeitos vão diminuindo;
- existem efeitos do 2º factor em todos os anos, sendo muito elevados no ano de 1997.

Para o valor crítico $F_{(0.95;3,159)} = 2.66$ verifica-se que os efeitos do 3º factor são significativos em todos os anos sendo mais elevados no ano de 1989.

■ Interação entre os factores

Para verificar a existência ou não de interação entre os vários factores foram efectuados os testes de hipóteses, cujos valores das estatísticas estão na tabela 4.7.

Ano	$F_i\{1, 2\}$	$F_i\{1, 3\}$	$F_i\{2, 3\}$	$F_i\{1, 2, 3\}$
1985	1.829	33.973	37.829	32.687
1989	4.601	19.073	3.047	59.798
1993	30.070	25.174	11.729	36.229
1997	6.394	28.038	4.279	70.667
2001	14.916	36.374	6.073	63.153

Tabela 4.7: Estatísticas $F_i(C)$ para efeitos da interacção entre os factores.

Para o valor crítico de $F_{(0.95;1,159)} = 3.90$, concluímos que entre o 1º e 2º factor não existe interacção no ano de 85, sendo a interacção mais elevada no ano de 1993 do que nos outros anos.

Verificamos para $F_{(0.95;3,159)} = 2.66$, que:

- há uma forte interacção entre o 1º e 3º factor, sendo mais elevada no ano de 2001;
- no ano de 1985, a interacção é mais elevada entre o 2º e 3º factor e mais baixa no anos de 1989 e 1997;
- existe uma forte interacção entre o 1º, 2º e 3º factor, tendo o valor mais elevado no ano de 1997 e mais baixo no ano de 1985.

4.2.2 Modelo com componente aleatória em X_i

A análise que efectuamos incide sobre os resultados das eleições autárquicas entre 1985 e 2001, em 16 concelhos (mencionados na pág. 45). Na sub-secção 3.3.2, apresentámos os resultados dos testes de validação do modelo com a componente aleatória em X_i (tabelas 3.6 e 3.7, pág. 54) que permitiu concluir que a estrutura comum é de grau um.

Pretendemos efectuar a análise longitudinal para as séries individuais e as análises longitudinal e transversal, sobre as séries emparelhadas de estudos.

4.2.2.1 Análise longitudinal para as séries individuais

Para efectuar esta análise, utilizamos as estimativas dos valores dos contrastes que se encontram na tabela 4.1, pág.77. Pretende-se, assim, verificar se existem diferenças significativas entre os resultados eleitorais nos vários anos.

Rejeita-se a hipótese

$$H_{0i}(l) : c_i^t \beta(l) = 0, i = 1, 2, 3; l = 1, \dots, 16,$$

ao nível α , se

$$F_{0i}(l) = \frac{g}{1} \frac{\left(c_i^t \tilde{\beta}(l) \right) \left(c_i^t \mathbf{W}(l) c_i^t \right)^{-1} \left(c_i^t \tilde{\beta}(l) \right)}{V} > F_{(1-\alpha, 1, g)}.$$

Os valores das estatísticas $F_{0i}(l)$ encontram-se na tabela 4.8.

	V.Cas	P.Bar	Bra	A. Fé	Ave	S.M.F	Gua	Pin
$F_{01}(l)$	64.69	7.95	23.93	3.04	0.13	58.54	11.14	43.27
$F_{02}(l)$	23.36	0.09	3.25	14.10	20.90	34.17	10.87	192.36
$F_{03}(l)$	50.10	58.02	0.12	16.03	0.07	3.26	3.82	129.82
	Lei	Alc.	Por	Elv	Far	Sil	Bej	Mer
$F_{01}(l)$	7.95	209.95	25.13	470.54	30.48	27.74	0.71	8.62
$F_{02}(l)$	0.58	2.00	7.55	294.59	1.00	24.88	4.90	5.00
$F_{03}(l)$	11.67	59.25	9.65	309.71	20.46	6.19	0.59	3.44

Tabela 4.8: Valores das estatísticas $F_{0i}(l)$.

Para o valor crítico de $F_{(0.95;1,152)} = 3.84$, concluímos que:

- os resultados obtidos não são significativamente diferentes entre os anos de 2001 e 1985 para os concelhos de Alfandega da Fé, Aveiro e Beja;
- entre 1997 e 1989 os resultados não são significativamente diferentes nos concelhos de Ponta da Barca, Bragança, Leiria, Alcobaça e Faro;
- entre 1985 a 2001, os resultados não são significativos ao nível da curvatura, nos concelhos de Bragança, Aveiro, Santa Maria da Feira, Guarda, Beja e Mértola.

4.2.2.2 Análise longitudinal e transversal sobre séries emparelhadas

A análise longitudinal e transversal é efectuada, indexando as séries de acordo com os níveis dos factores apresentadas na sub-secção 4.2.1.2 (pág. 79). Para analisar os efeitos dos factores e suas interacções utilizamos as matrizes ortogonais

estandardizadas (pág. 79) para obter a matriz \mathbf{P} estandardizada (pág. 80).

4.2.2.2.1 Análise longitudinal

Nesta análise, pretendemos verificar os efeitos e interações dos factores nas diferenças verificadas entre as várias eleições e na análise da curvatura da evolução, utilizando os valores dos contrastes (tabela 4.1, pág.77).

Rejeitamos a hipótese

$$H_0(C) : \psi(C) = 0; C \subseteq \{1, 2, 3\}$$

ao nível α , se

$$F(C) = \frac{g}{g(C)} \frac{\left(\tilde{\psi}(C)\right)^t \left(\mathbf{A}(C)\mathbf{D}\mathbf{A}(C)^t\right)^{-1} \left(\tilde{\psi}(C)\right)}{V} > F_{(1-\alpha;g(C),g)}.$$

■ Efeitos do 1º 2º e 3º factor

O valor das estatísticas dos testes para os efeitos dos factores nos valores dos contrastes, estão discriminados na seguinte tabela 4.9:

	$F\{1\}$	$F\{2\}$	$F\{3\}$
L_{1l}	0.688	50.076	4.854
L_{2l}	55.237	106.361	19.563
L_{3l}	25.248	3.071	50.984

Tabela 4.9: Estatísticas $F(C)$ para efeitos do 1º, 2º e 3º factor.

Com o valor critico de $F_{(0.95;1,152)} = 3.84$ concluímos que para:

- o 1º factor, apenas são significativos os valores dos testes para o 2º e 3º contraste, não tendo este factor influência significativa nas diferenças entre os resultados eleitorais dos anos de 2001 e 1985;

- o 2º factor, apenas são significativos os valores dos testes para o 1º e 2º contraste.

Para $F_{(0.95;3,152)} = 2.66$, verificamos que para o 3º factor são significativos os valores dos testes para os vários contrastes.

Constatamos que os efeitos do 2º e 3º factor são menores nas diferenças entre os resultados eleitorais dos anos de 2001 e 1985 do que entre 1997 e 1989.

■ Interação entre factores

Os resultados das estatísticas dos testes F para efeitos da interacção entre os vários factores relativamente aos valores dos contrastes estão na tabela 4.10.

	$F\{1, 2\}$	$F\{1, 3\}$	$F\{2, 3\}$	$F\{1, 2, 3\}$
L_{1l}	57.577	18.863	153.905	113.951
L_{2l}	46.176	58.126	20.055	29.130
L_{3l}	133.151	85.310	12.048	7.036

Tabela 4.10: Estatísticas $F(C)$ para efeitos da interacção entre os factores.

Para o valor $F_{(0.95;1,152)} = 3.84$ e $F_{(0.95;3,152)} = 2.60$ verifica-se que a interacção é significativa entre todas as interacções possíveis dos vários factores nos contrastes.

Para o valor $F_{(0.95;3,152)} = 2.60$ concluímos que a interacção entre:

- o 1º e 3º factor para o 1º contraste é inferior à do 2º contraste;

- o 2º e 3º factor e entre 1º, 2º e 3º factor é mais forte relativamente ao 1º contraste do que no 2º contraste.

4.2.2.2 Análise transversal

Nesta análise, pretendemos verificar os efeitos e interacções dos vários factores nos resultados eleitorais obtidos nos 16 concelhos.

Rejeita-se

$$H_{0i}(C) : \varphi_i(C) = 0, i = 1, \dots, 5; C \subseteq \{1, 2, 3\}$$

ao nível α , se

$$F_i(C) = \frac{g}{g(C)} \frac{(\tilde{\varphi}_i(C))^t (\mathbf{A}(C)\mathbf{M}_i\mathbf{A}(C)^t)^{-1} (\tilde{\varphi}_i(C))}{V} > F_{1-\alpha, g(c), g}.$$

■ Efeitos do 1º, 2º e 3º factor

Na tabela 4.11, apresentamos os valores das estatísticas dos testes $F_i(C)$ para efeitos dos factores.

Ano	$F_i\{1\}$	$F_i\{2\}$	$F_i\{3\}$
1985	0.899	39.946	28.130
1989	137.880	28.005	49.132
1993	21.718	80.754	20.164
1997	7.593	282.810	13.884
2001	0.001	194.857	21.473

Tabela 4.11: Estatísticas $F_i(c)$ para efeitos do 1º, 2º e 3º factor.

Para o valor $F_{(0.95;1,152)} = 3.84$, podemos concluir que:

- o 1º factor não tem efeito nos resultados eleitorais obtidos nos anos de 1985 e 2001, mas em 1989 há uma forte influência desse factor que vai diminuindo ao longo das eleições seguintes;
- no 2º factor são significativos os valores dos testes em todos os anos eleitorais, mas são mais elevados nos anos de 1997 e 2001.

Para $F_{(0.95;3,152)} = 2.60$, verifica-se que o efeito do 3º factor é significativo em todos os anos eleitorais, mas em 1989 a sua influência foi mais acentuada.

■ Interação entre os factores

Os resultados das estatísticas dos testes $F_i(C)$, para a interação entre os factores, estão discriminadas na tabela 4.12

Ano	$F_i\{1, 2\}$	$F_i\{1, 3\}$	$F_i\{2, 3\}$	$F_i\{1, 2, 3\}$
1985	5.258	96.598	107.864	94.821
1989	13.488	55.557	9.262	174.050
1993	90.250	74.208	35.826	106.558
1997	17.932	78.876	11.744	201.931
2001	40.751	96.603	16.204	175.302

Tabela 4.12: Estatísticas $F_i(c)$ para efeitos da interação entre os factores.

e permitem concluir que para $F_{(0.95;1,152)} = 3.84$ e $F_{(0.95;3,152)} = 2.60$ há interação entre os vários factores em todos os anos eleitorais seleccionados.

Verificamos, ainda, que a interacção entre o:

- 1º e 2º factor é mais elevada em 1993 e mais baixa 1985;
- 1º e 3º factor é mais elevada em 1985 e 2001 e mais baixa em 1989;
- 2º e 3º factor tem o valor mais elevado em 1985 e mais baixo em 89;
- 1º, 2º e 3º factor tem valor mais elevado em 1997 e mais baixo em 1985.

CAPÍTULO 5

Considerações finais

Neste trabalho utilizamos a normalidade aproximada para demonstrar que polinómios não lineares de baixo grau, com variáveis aleatórias independentes, podem ser aproximados a funções lineares de variáveis aleatórias independentes com baixos coeficientes de variação. Mostramos também que os polinómios podem ser aproximados pela respectiva componente linear dominante.

Com a utilização da metodologia STATIS, apresentámos um modelo para séries de estudos com estrutura comum de grau h com a componente aleatória nas matrizes \mathbf{A}_i e \mathbf{X}_i que permitiu determinar, nas duas situações, o vector de resíduos homocedástico para cada série independente do vector de estrutura. Através da validação deste modelo verificamos que em ambos os casos:

- ▶ não se rejeitava a hipótese que os vectores de resíduos eram homocedásticos com vector médio nulo;
- ▶ se ajustava o modelo com estrutura comum de grau um altamente significativa.

O modelo obtido e os seus resultados permitiram-nos realizar inferência sobre as eleições autárquicas, com a componente aleatória em \mathbf{A}_i e \mathbf{X}_i . Assim nas séries individuais foram utilizados vectores de contraste e nas séries emparelhadas realizou-se a análise longitudinal para estudar a acção dos factores sobre a evolução das séries. Na análise transversal estudou-se a acção dos factores nos estudos simultâneos, isto é, nos estudos que se correspondem. Estas análises foram obtidas recorrendo a matrizes ortogonais estandardizadas e ao produto de kronecker de matrizes.

A comparação dos dois modelos deve ser feita em termos de precisão dos mesmos, expressos pelos resultados a que conduzem nas análises longitudinal e transversal.

Ao observar os valores das estatísticas dos testes F com a componente aleatória em:

- ▶ \mathbf{A}_i , tabelas: 4.2, 4.4, 4.5, 4.6 e 4.7;
- ▶ \mathbf{X}_i , tabelas: 4.8, 4.9, 4.10, 4.11 e 4.12;

constatamos que esses valores, com a componente aleatória em \mathbf{X}_i são aproximadamente o triplo das estatísticas obtidas com a componente aleatória nos operadores \mathbf{A}_i .

Esta comparação sugere que o modelo, com a componente aleatória nos operadores \mathbf{A}_i , deve ser considerado como uma aproximação. Caso os resultados a que a mesma conduza sejam suficientes, isto é, se rejeitarmos as hipóteses H_0 , não é necessário efectuar o teste com a componente em \mathbf{X}_i . Se se quiser refinar a análise, dever-se-à utilizar o método com a componente aleatória em \mathbf{X}_i .

Este trabalho incidiu sobre a primeira etapa "inter-estrutura, da metodologia STATIS". Pretendemos vir a estender o mesmo tratamento à etapa, "o compromisso".

BIBLIOGRAFIA

Areia A., Oliveira M. M., Mexia, J. T. (2008). *Models for a series of studies based on geometrical representation*. Statistical Methodology. Vol. 5, N.º 3, 277-288.

Escoufier Y. (1973). *Le Traitement des Variables Vectorielles*. Biometrics 29, N.º 4, 751-760.

Escoufier Y., L'Hermier H. (1978). *A propos de la Comparaison Graphique des Matrices de Variance*. Biom. J. Vol. 20, N.º 5, 477-483.

Lavit C. (1988). *Analyse Conjointe de Tableaux Quantitatifs*. Collection Méthods + Programmes, Masson, Paris.

Lavit C., Escoufier Y., Sabatier R., Traissac P, (1994). *The ACT (STATIS method)*. Computation Statistics & Data Analysis, 97-119.

Lehman L. (1959). *Testing Statistical Hypothesis*. John Willey & Sons, New York.

Mexia J. T. (1988). *Standardized Orthogonal Matrices and the Decomposition of the Sum of Squares for Treatments*. Trabalhos de Investigação, N.º2, Departamento de Matemática-Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Mexia J. T. (1989). *Controlled Heterocedasticity, in: Quotient Vector Spaces and F tests for Hypotheses on Mean Vectors*, Trabalhos de Investigação, N.º1, Departa-

mento de Matemática-Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Mexia J. T. (1989). *Simultaneous Confidence Intervals: Generalization of the Scheffé Theorem*. Trabalhos de Investigação, N.º2, Departamento de Matemática-Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Mexia J. T. (1990). *Best Linear Unbiased Estimates, Duality of F Tests and the Scheffé Multiple Comparison Method in Presence of Controlled Heterocedasticity*. Computation Statistics & Data Analysis - Vol. 10, N.º3.

Mexia J. T. (1995). *Introdução à Inferência Estatística Linear*. Centros de Estudos de Matemática Aplicada, Edições Universitárias Lusófonas, Lisboa.

Montgomery, Douglas C. (2005). *Design and Analysis of Experiments*. 6th Ed. John Wiley & Sons, New York.

Oliveira M. M., Mexia, J. T. (1998). *Tests for the rank of Hilbert-Schmidt product matrices*. *Advances in Data Science and Classifications*. Springer, 619-625.

Oliveira M. M., Mexia, J. T. (1999a). *F tests for Hypothesis on the Structure Vectors of Series*. *Discussiones Mathematicae*. Biometrical Letters. Vol. 19, N.º2, 345-353.

Oliveira M. M., Mexia, J. T. (1999b). *Multiple Comparisons for Rank one Common Structures*. Biometrical Letters. Vol. 36, N.º2, 159-167.

Oliveira M. M., Mexia, J. T. (2004). *AIDS in Portugal: endemic versus epidemic forecasting scenarios for mortality*. *International Journal of Forecasting*, N.º20, 131-137.

Oliveira M. M., Mexia, J. T. (2007). *ANOVA like analysis of matched series of studies with a common structure*. Journal of Statistical Planning and Inference, N.º137, 1862-1870.

Oliveira M. M., Mexia, J. T. (2007). *Modelling series of studies with a common structure*. Computation Statistics and Data Analysis, N.º51, 5876-5885.

Oliveira M. M. (2002). *Modelação de Séries Emparelhadas com Estrutura Comum*. PhD thesis. Universidade de Évora.

Ramos (2006). *Quase-Normalidade e Inferência para Séries de Estudos Emparelhadas*. PhD thesis. Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia.

Rohatgi, V. K. (1976). *An Introduction to Probability and Mathematical Statistics*. Wiley, New York.

Scheffé H. (1959). *The Analysis of Variance*. John Wiley & Sons, New York.

Schott J. R. (1997). *Matrix Analysis for Statistics*. Wiley Series in Probability and Statistics.

Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.