

Universidade de Évora
Instituto de Investigação e Formação Avançada

**Resultados Assintóticos para Famílias Estruturadas de
Modelos Colectivos. Aplicação aos Fogos Florestais em
Portugal Continental**

Dissertação apresentada para obtenção do grau de
Doutor em Matemática na especialidade de Es-
tatística, pela Universidade de Évora

Ana Isabel Guerra Cantarinha
Orientador: Maria Manuela Melo Oliveira
Co-orientadores: João Tiago Praça Nunes Mexia;
José Guilherme Calvão Borges

Évora
2011

nº de arquivo

copyright

Universidade de Évora
Instituto de Investigação e Formação Avançada

**Resultados Assintóticos para Famílias Estruturadas de
Modelos Colectivos. Aplicação aos Fogos Florestais em
Portugal Continental**

Dissertação apresentada para obtenção do grau de
Doutor em Matemática na especialidade de Es-
tatística, pela Universidade de Évora

Ana Isabel Guerra Cantarina

Orientador: Maria Manuela Melo Oliveira

Co-orientadores: João Tiago Praça Nunes Mexia;

José Guilherme Calvão Borges

Évora

2011

Ao João e à Matilde

Agradecimentos

À minha orientadora Professora Doutora Maria Manuela Oliveira, pelo apoio, pela orientação e pela amizade que me transmitiu ao longo deste percurso, tornando assim possível chegar até aqui.

Ao meu co-orientador Professor Doutor João Tiago Mexia, foi um privilégio trabalhar com uma pessoa com tamanha sabedoria, generosidade e humildade. Pela sua total disponibilidade para acompanhar de perto todo o processo de elaboração desta dissertação, sempre com a sua simpatia e entusiasmo.

Ao meu co-orientador Professor Doutor José Guilherme Borges, pelo apoio e acesso aos dados utilizados na aplicação desta dissertação, e foi através dele que todo este caminho começou.

Na preparação dos dados devo agradecer ao ISA, em particular à Susete pelo envio e explicação dos dados.

Ao CMA, em particular à Elsa, à Sónia e ao Miguel pela disponibilidade e simpatia que sempre manifestaram.

À Fundação Eugénio de Almeida pelo apoio financeiro através da bolsa de Doutoramento que me atribuiu.

À minha família e aos meus amigos pelo apoio, pela compreensão e por aquilo que sou hoje.

À memória do meu pai.

Ao João, que em todos os momentos o sinto ao meu lado, por tudo o que temos partilhado e por me ter dado uma filha linda, que nasceu no percurso desta dissertação, que com o seu sorriso me ajudou a ultrapassar os momentos mais difíceis.

Não poderia deixar de agradecer a Deus, pelo facto de ter criado condições para que tivesse tido a oportunidade de encontrar as pessoas certas nos momentos exactos.

A todos os que, directa ou indirectamente, estiveram comigo e me apoiaram o meu Bem Haja.

Resumo

Desenvolveu-se uma teoria baseada em resultados assintóticos para:

- modelos colectivos multivariados da forma

$$\mathbf{S} = \sum_{i=1}^N \mathbf{X}_i$$

onde os $\mathbf{X}_1, \dots, \mathbf{X}_i, \dots$ são i.i.d. e independentes de N que tem distribuição de Poisson;

- famílias estruturadas de modelos em que a cada tratamento dum modelo base Π corresponde um modelo colectivo multivariado. Estuda-se a acção dos factores de Π sobre os modelos de famílias.

Os resultados obtidos aplicam-se a situações em que as amostras são grandes, o que na prática sucede muitas vezes. Uma tal situação é a dos incêndios florestais em Portugal Continental. Para esse efeito o território encontra-se dividido em cinco regiões o que justifica a abordagem através de famílias de modelos. Os factores que se tinham eram, além da região, o ano já que se pretendia analisar a evolução temporal de

- \mathbf{N} - número de fogos por ano e região,
- \mathbf{T} - área total ardida por ano e região,

não se detectando qualquer evolução na gravidade do problema.

Abstract

Title: "Asymptotic Results for Structure Families of Collective Models. Application to forest fires in Mainland Portugal." We developed a theory based on asymptotic results for:

- multivariate collective models

$$\mathbf{S} = \sum_{i=1}^N \mathbf{X}_i$$

where the $\mathbf{X}_1, \dots, \mathbf{X}_i, \dots$ are i.i.d. independent from N that has a Poisson distribution;

- structured families in which, for each treatment of a base design, we have a multivariate collective model. The aim is to study the action of the factors in Π on the models of the family.

Our approach holds for situations in which we deal with large samples, which is quite often the case. One such situation is that of forest fires in Mainland Portugal. There are five regions for forest fires control so we were led to consider a structured family with a model per region. Besides regions in Π we had a time factor with a level per year. Since we were interested in studying the evolution:

- \mathbf{N} - number of fires per year and region,
- \mathbf{T} - burned area per year and region.

No significant evolution was found.

Simbologia e Notações

N número de fogos

T área total ardida

i.i.d. independentes e identicamente distribuídos

v vector (letra minúscula a negrito)

Y vector aleatório (letra maiúscula a negrito)

A matriz (letra maiúscula a negrito)

II modelo de efeitos fixos

\mathbb{R} conjunto dos números reais

\boxplus soma directa ortogonal de sub-espacos

| | |
|--------------------------------------------------------|------------------------------------------------------------------------------------|
| $\mathbf{D}(r_1, \dots, r_k)$ | matriz diagonal cujos elementos principais são as componentes de r_1, \dots, r_k |
| $\mathbf{A} = [\mathbf{a}_1^k, \dots, \mathbf{a}_s^k]$ | matriz com vectores coluna $\mathbf{a}_1^k, \dots, \mathbf{a}_s^k$ |
| $\mathbf{A} = [a_{ij}]$ | matriz com elemento genérico a_{ij} |
| $car(\cdot)$ | característica duma matriz |
| $det(\mathbf{A})$ | determinante da matriz \mathbf{A} |
| \mathbf{A}^t | transposta de \mathbf{A} |
| \mathbf{A}^{-1} | matriz inversa de \mathbf{A} |
| \mathbf{P}_r | matriz ortogonal padronizada $r \times r$ |
| \mathbf{M}^+ | inversa de Moore-Penrose da matriz \mathbf{M} |
| \mathbf{I}_k | matriz identidade de ordem k |

| | |
|--------------------|--------------------------------------------------------|
| $\mathbf{0}_{n,m}$ | matriz de ordem $n \times m$ com elementos todos nulos |
| $\mathbf{0}_k$ | vector com componentes nulas |
| $\mathbf{1}_k$ | vector cujas componentes são unitárias |
| ∇ | sub-espaço |
| ∇^\perp | complemento ortogonal |
| MPO | matriz de projecção ortogonal |
| f.g.m. | função geradora de momentos |
| AJCS | álgebra de Jordan comutativa de matrizes simétricas |
| AJC | álgebra de Jordan comutativa |
| \otimes | produto de Kronecker entre matrizes |
| $R(\mathbf{A})$ | espaço imagem da matriz \mathbf{A} |
| \mathcal{A} | álgebra |

| | |
|-----------------------------------------|-----------------------------------------------------------------------------------------------------|
| $\xrightarrow[n \rightarrow \infty]{F}$ | converge em distribuição para, quando n tende para $+\infty$ |
| $\xrightarrow[n \rightarrow \infty]{p}$ | converge em probabilidade para, quando n tende para $+\infty$ |
| $E(\cdot)$ | valor médio |
| \sim | distribui-se como |
| $\mathcal{V}(\cdot)$ | matriz de covariância |
| <u>$grad$</u> | vector gradiente |
| <u>$Hess$</u> | matriz Hessiana |
| $P(\lambda)$ | distribuição de Poisson de parâmetro λ |
| $N(\mu, \sigma^2)$ | distribuição normal de valor médio μ e variância σ^2 |
| χ_k^2 | distribuição do qui-quadrado central com k graus de liberdade |
| $\chi_{k,\delta}^2$ | distribuição do qui-quadrado com k graus de liberdade e parâmetro de não centralidade δ |

Índice

| | |
|---------------------------------------------------------------------|-----------|
| Agradecimentos | ii |
| Resumo | iv |
| Abstract | v |
| Simbologia e Notações | vi |
| 1 Introdução | 1 |
| 2 Matrizes | 4 |
| 2.1 Matrizes Uniformizadoras e de Projeção Ortogonal | 4 |
| 2.2 Produto de Kronecker | 8 |
| 2.3 Álgebras de Jordan Comutativas de Matrizes Simétricas | 10 |
| 3 Inferência | 15 |
| 3.1 Momentos | 15 |
| 3.2 Funções Geradoras de Momentos e de Cumulantes | 16 |
| 3.3 Vectores Normais | 19 |
| 3.4 Convergência em Distribuição | 23 |
| 3.5 Testes e Regiões de Confiança | 30 |
| 3.6 Linearidade Assintótica | 32 |
| 4 Modelos Colectivos Multivariados | 36 |
| 4.1 Modelos Isolados | 36 |
| 4.2 Famílias de Modelos | 42 |

| | | |
|----------|----------------------------------------------------------------|------------|
| 4.2.1 | Famílias Estruturadas | 42 |
| 4.2.2 | Famílias Estruturadas Ortogonais | 50 |
| 5 | Fogos Florestais em Portugal Continental | 53 |
| 5.1 | Apresentação do problema | 53 |
| 5.2 | Inferência e Resultados | 55 |
| 5.2.1 | Região a Região | 56 |
| 5.2.2 | Ano a Ano | 61 |
| 5.2.3 | Análise conjunta Região e Ano | 68 |
| 6 | Considerações Finais | 71 |
| A | Coordenadas Polares Generalizadas e Testes Qui-Quadrado | 72 |
| B | Teorema de Polya | 76 |
| C | Intervalos de Confiança para os Modelos Isolados | 80 |
| D | Método de Comparação Múltipla de Scheffé | 87 |
| | Bibliografia | 102 |

Lista de Tabelas

| | | |
|------|-------------------------------------------------------------------------------------|----|
| 5.1 | Estatísticas de Qui-Quadrado | 57 |
| 5.2 | Efeitos para \mathbf{N} | 57 |
| 5.3 | Efeitos para \mathbf{T} | 58 |
| 5.4 | Estatísticas de Qui-Quadrado para E_L | 59 |
| 5.5 | Estatísticas de Qui-Quadrado para E_Q | 60 |
| 5.6 | Totais de \mathbf{N} | 62 |
| 5.7 | Totais de \mathbf{T} | 63 |
| 5.8 | Estatísticas de Qui-Quadrado para \mathbf{N} e \mathbf{T} | 64 |
| 5.9 | Valores de $d_{i,l}$ para o ano de 1983 | 65 |
| 5.10 | Valores de $\dot{d}_{i,l}$ para o ano de 1983 | 65 |
| 5.11 | Pares de regiões sem diferenças significativas | 67 |
| 5.12 | Estatísticas de Qui-Quadrado | 69 |
| 5.13 | Relevância Relativa para \mathbf{N} | 69 |
| 5.14 | Relevância Relativa para \mathbf{T} | 70 |
| C.1 | Intervalos de confiança a 90% para $\boldsymbol{\lambda}$ | 81 |
| C.2 | Intervalos de confiança a 95% para $\boldsymbol{\lambda}$ | 82 |
| C.3 | Intervalos de confiança a 99% para $\boldsymbol{\lambda}$ | 83 |
| C.4 | Intervalos de confiança a 90% para $\boldsymbol{\lambda}\boldsymbol{\mu}$ | 84 |

| | | |
|------|-----------------------------------------------------------|----|
| C.5 | Intervalos de confiança a 95% para $\lambda\mu$ | 85 |
| C.6 | Intervalos de confiança a 99% para $\lambda\mu$ | 86 |
| D.1 | Valores de $d_{i,l}$ para o ano de 1984 | 87 |
| D.2 | Valores de $\dot{d}_{i,l}$ para o ano de 1984 | 88 |
| D.3 | Valores de $d_{i,l}$ para o ano de 1985 | 88 |
| D.4 | Valores de $\dot{d}_{i,l}$ para o ano de 1985 | 88 |
| D.5 | Valores de $d_{i,l}$ para o ano de 1986 | 89 |
| D.6 | Valores de $\dot{d}_{i,l}$ para o ano de 1986 | 89 |
| D.7 | Valores de $d_{i,l}$ para o ano de 1987 | 89 |
| D.8 | Valores de $\dot{d}_{i,l}$ para o ano de 1987 | 90 |
| D.9 | Valores de $d_{i,l}$ para o ano de 1988 | 90 |
| D.10 | Valores de $\dot{d}_{i,l}$ para o ano de 1988 | 90 |
| D.11 | Valores de $d_{i,l}$ para o ano de 1989 | 90 |
| D.12 | Valores de $\dot{d}_{i,l}$ para o ano de 1989 | 91 |
| D.13 | Valores de $d_{i,l}$ para o ano de 1990 | 91 |
| D.14 | Valores de $\dot{d}_{i,l}$ para o ano de 1990 | 91 |
| D.15 | Valores de $d_{i,l}$ para o ano de 1991 | 92 |
| D.16 | Valores de $\dot{d}_{i,l}$ para o ano de 1991 | 92 |
| D.17 | Valores de $d_{i,l}$ para o ano de 1992 | 92 |
| D.18 | Valores de $d_{i,l}$ para o ano de 1993 | 92 |
| D.19 | Valores de $\dot{d}_{i,l}$ para o ano de 1993 | 93 |
| D.20 | Valores de $d_{i,l}$ para o ano de 1994 | 93 |
| D.21 | Valores de $\dot{d}_{i,l}$ para o ano de 1994 | 93 |
| D.22 | Valores de $d_{i,l}$ para o ano de 1995 | 94 |

| | | |
|------|-----------------------------------------------|-----|
| D.23 | Valores de $\dot{d}_{i,l}$ para o ano de 1995 | 94 |
| D.24 | Valores de $d_{i,l}$ para o ano de 1996 | 94 |
| D.25 | Valores de $\dot{d}_{i,l}$ para o ano de 1996 | 94 |
| D.26 | Valores de $d_{i,l}$ para o ano de 1997 | 95 |
| D.27 | Valores de $\dot{d}_{i,l}$ para o ano de 1997 | 95 |
| D.28 | Valores de $d_{i,l}$ para o ano de 1998 | 95 |
| D.29 | Valores de $\dot{d}_{i,l}$ para o ano de 1998 | 96 |
| D.30 | Valores de $d_{i,l}$ para o ano de 1999 | 96 |
| D.31 | Valores de $\dot{d}_{i,l}$ para o ano de 1999 | 96 |
| D.32 | Valores de $d_{i,l}$ para o ano de 2000 | 96 |
| D.33 | Valores de $\dot{d}_{i,l}$ para o ano de 2000 | 97 |
| D.34 | Valores de $d_{i,l}$ para o ano de 2001 | 97 |
| D.35 | Valores de $\dot{d}_{i,l}$ para o ano de 2001 | 97 |
| D.36 | Valores de $d_{i,l}$ para o ano de 2002 | 98 |
| D.37 | Valores de $\dot{d}_{i,l}$ para o ano de 2002 | 98 |
| D.38 | Valores de $d_{i,l}$ para o ano de 2003 | 98 |
| D.39 | Valores de $\dot{d}_{i,l}$ para o ano de 2003 | 98 |
| D.40 | Valores de $d_{i,l}$ para o ano de 2004 | 99 |
| D.41 | Valores de $\dot{d}_{i,l}$ para o ano de 2004 | 99 |
| D.42 | Valores de $d_{i,l}$ para o ano de 2005 | 99 |
| D.43 | Valores de $\dot{d}_{i,l}$ para o ano de 2005 | 100 |
| D.44 | Valores de $d_{i,l}$ para o ano de 2006 | 100 |
| D.45 | Valores de $\dot{d}_{i,l}$ para o ano de 2006 | 100 |
| D.46 | Valores de $d_{i,l}$ para o ano de 2007 | 100 |
| D.47 | Valores de $\dot{d}_{i,l}$ para o ano de 2007 | 101 |

Capítulo 1

Introdução

Os modelos colectivos, ver por exemplo Bowers et al. (1986), Centeno (2003), Reis (1993) e Grandell (1992), têm desempenhado um papel central na Teoria do Risco. Esta teoria estuda fundamentalmente a solvabilidade das seguradoras. Como veremos na aplicação, estes modelos podem ser utilizados em muitos outros campos nomeadamente nos fogos florestais. Este alargamento do campo de aplicação levou a que se procurasse desenvolver a teoria destes modelos segundo duas linhas:

- Passagem dos modelos univariados a modelos multivariados,
- Passagem de modelos isolados a famílias estruturadas de modelos.

Nesta fase de ampliação da teoria obtiveram-se teoremas do limite central que permitem fundamentar a inferência, quer para modelos isolados quer para famílias estruturadas de modelos. Nestas os modelos correspondem aos tratamentos dum modelo de efeitos fixos II. O estudo de tais famílias estará centrado na acção dos factores de II sobre os modelos.

Esta abordagem a famílias estruturadas de modelos encontra-se dentro duma linha de trabalho que tem vindo a ser desenvolvida, ver Mexia (1987), Mexia (1990), Mexia (1992) e Oliveira & Mexia (2007a) (2007b).

Num primeiro capítulo apresentamos resultados sobre matrizes que nos serão úteis. Salienta-se a utilização de álgebras de Jordan comutativas. Estas álgebras foram introduzidas por Jordan et al. (1934) numa reformulação da Mecânica Quântica. Posteriormente foram redescobertas por Seely que iniciou uma linha de investigação extremamente profícua no campo da Inferência Estatística Linear. Ver Seely (1970a), (1970b), (1971), (1977), Seely & Zyskind (1971) e Vanleeuwen et al. (1998), (1999).

Segue-se um capítulo em que apresentamos os fundamentos para realizar inferência em modelos colectivos isolados ou integrados em famílias. Salienta-se os resultados sobre distribuições limite em que utilizamos uma abordagem que vem por exemplo apresentada em Lukacs & Laha (1964). A ênfase neste tipo de resultados permite utilizar o facto de, muitas vezes, se trabalhar com amostras grandes.

Utilizamos ainda outro tipo de resultados limite em que se considera o crescimento não da dimensão da amostra mas da sua não centralidade, ver Mexia & Oliveira (2010). Conseguimos assim interpretar situações em que as estatísticas dos testes usuais tomam grandes valores. De facto é frustrante concluir-se, como até pode suceder, que todos os efeitos e interacções são altamente significativos como até encontramos na aplicação que faremos. Podemos agora estimar a relevância dos mesmos.

Segue-se o capítulo chave da tese em que começamos por apresentar modelos isolados. Como referimos o nosso tratamento destes modelos centrar-se-á na utilização de distribuições limite. Para isso admitimos que os vectores de observações são dados por

$$\mathbf{S} = \sum_{i=1}^N \mathbf{X}_i$$

onde os \mathbf{X}_i, \dots são i.i.d. e independentes de N que terá distribuição de Poisson com parâmetro $\boldsymbol{\lambda}$. Em particular as propriedades de reprodutibilidade das distribuições de Poisson desempenharão um papel muito importante. Assim subjacente aos nossos modelos estará um processo estocástico de Poisson. No futuro tencionamos generalizar o nosso trabalho considerando outros processos de contagem, como se faz actualmente para o caso univariado, ver por exemplo Grandell (1992).

As distribuições limite obtidas permitem-nos aplicar as técnicas de Inferência Estatística Linear quando passamos a famílias estruturadas. Neste ponto seguimos a mesma orientação que nos trabalhos atrás referidos.

Aplicamos em seguida estes modelos ao estudo dos fogos florestais em Portugal Continental. Nesse estudo utilizamos modelos bivariados sendo as variáveis número de fogos (\mathbf{N}) e área total ardida (\mathbf{T}) consideradas ano a ano e para as cinco regiões Norte, Centro, Lisboa e Vale do Tejo, Alentejo e Algarve em que Portugal Continental está dividido. A nossa abordagem permitirá estudar a acção dos factores Ano e Região nas variáveis consideradas.

Capítulo 2

Matrizes

2.1 Matrizes Uniformizadoras e de Projeção Ortogonal

Seja \mathbf{A} uma matriz simétrica de ordem k , existe então uma matriz ortogonal \mathbf{P} tal que

$$\mathbf{PAP}^t = \mathbf{D}(r_1, \dots, r_k)$$

sendo $\mathbf{D}(r_1, \dots, r_k)$ uma matriz diagonal, cujos elementos principais r_1, \dots, r_k são os valores próprios da matriz \mathbf{A} . Se \mathbf{A} for definida positiva tem-se $r_j > 0$, $j = 1, \dots, k$, estando definida a matriz

$$\mathbf{G}_0 = \mathbf{D}(r_1^{-1/2}, \dots, r_k^{-1/2})\mathbf{P}$$

que é solução da equação matricial

$$\mathbf{MAM}^t = \mathbf{I}_k.$$

As soluções desta equação serão as matrizes uniformizadoras de \mathbf{A} . Em particular \mathbf{G}_0 será a uniformizadora directa de \mathbf{A} .

Dado um sub-espaco ∇ , seja o respectivo complemento ortogonal ∇^\perp tendo todo o vector \mathbf{z} uma partição única da forma

$$\mathbf{z} = \mathbf{z}_\nabla + \mathbf{z}_{\nabla^\perp}$$

onde $\mathbf{z}_\nabla \in \nabla$ [$\mathbf{z}_{\nabla^\perp} \in \nabla^\perp$] é a projecção ortogonal de \mathbf{z} sobre ∇ [∇^\perp]. Tem-se ainda

$$\mathbf{z}_\nabla = \mathbf{Q}(\nabla)\mathbf{z}$$

e

$$\mathbf{z}_{\nabla^\perp} = \mathbf{Q}(\nabla^\perp)\mathbf{z}$$

com $\mathbf{Q}(\nabla)$ e $\mathbf{Q}(\nabla^\perp)$ as matrizes de projecção ortogonal sobre ∇ e ∇^\perp . Como

$$\mathbf{Q}(\nabla)\mathbf{z}_\nabla = \mathbf{z}_\nabla$$

e

$$\mathbf{Q}(\nabla^\perp)\mathbf{z}_{\nabla^\perp} = \mathbf{z}_{\nabla^\perp}$$

vê-se que $\mathbf{Q}(\nabla)$ e $\mathbf{Q}(\nabla^\perp)$ são idempotentes. Pode ainda mostrar-se que uma matriz é MPO se e só se for simétrica e idempotente.

Dada $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m\}$ uma base para ∇ , tal que

$$\boldsymbol{\alpha}_i^t \boldsymbol{\alpha}_j = 0,$$

para $i \neq j$ e

$$\|\boldsymbol{\alpha}_i\| = 1, i, j = 1, \dots, m,$$

diremos que a base $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m\}$ é uma base ortonormada. Esta base pode ser expandida de forma a obter uma base ortonormada $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m, \boldsymbol{\alpha}_{m+1}, \dots, \boldsymbol{\alpha}_n\}$

para \mathbb{R}^n . Podemos ver que $\{\alpha_{m+1}, \dots, \alpha_n\}$ é uma base ortonormada para ∇^\perp e que

$$\mathbf{z} = \sum_{j=1}^n (\alpha_j^t \mathbf{z}) \alpha_j = \mathbf{z}_\nabla + \mathbf{z}_{\nabla^\perp}$$

com

$$\mathbf{z}_\nabla = \sum_{j=1}^m (\alpha_j^t \mathbf{z}) \alpha_j$$

e

$$\mathbf{z}_{\nabla^\perp} = \sum_{j=m+1}^n (\alpha_j^t \mathbf{z}) \alpha_j.$$

Sejam $\mathbf{A} = [\alpha_1, \dots, \alpha_m]^t$ e $\mathbf{A}^\perp = [\alpha_{m+1}, \dots, \alpha_n]^t$ então

$$\mathbf{A}\mathbf{A}^t = \mathbf{I}_m$$

$$\mathbf{A}^\perp(\mathbf{A}^\perp)^t = \mathbf{I}_{n-m}$$

$$\mathbf{A}(\mathbf{A}^\perp)^t = \mathbf{0}_{n \times (n-m)},$$

com $\mathbf{0}_{r \times s}$ a matriz nula $r \times s$. Assim, é fácil ver que $\mathbf{Q}_\nabla = \mathbf{A}^t \mathbf{A}$ e $\mathbf{Q}_{\nabla^\perp} = (\mathbf{A}^\perp)^t \mathbf{A}^\perp$ são matrizes simétricas e idempotentes, tendo-se aliás

$$\mathbf{Q}(\nabla) = \mathbf{Q}_\nabla$$

e

$$\mathbf{Q}(\nabla^\perp) = \mathbf{Q}_{\nabla^\perp}.$$

Como

$$\mathbf{Q}_{\nabla^\perp} = \mathbf{I}_n - \mathbf{Q}_\nabla$$

temos

$$\mathbf{Q}_\nabla \mathbf{Q}_{\nabla^\perp} = \mathbf{Q}_{\nabla^\perp} \mathbf{Q}_\nabla = \mathbf{0}_{n \times n}$$

onde \mathbf{I}_n é a matriz identidade $n \times n$.

Seja $R(\mathbf{U})$ o espaço imagem da matriz \mathbf{U} , assim temos

$$\nabla = R(\mathbf{Q}_\nabla) = R(\mathbf{A}^t),$$

pois tendo

$$\mathbf{Q}_\nabla \mathbf{A}^t = \mathbf{A}^t$$

e

$$\mathbf{A}^t \mathbf{A} = \mathbf{Q}_\nabla$$

temos

$$R(\mathbf{A}^t) \subset R(\mathbf{Q}_\nabla)$$

e

$$R(\mathbf{Q}_\nabla) \subset R(\mathbf{A}^t).$$

Um raciocínio similar é feito para ∇^\perp .

Dizemos que duas MPO \mathbf{Q}_1 e \mathbf{Q}_2 são ortogonais se os seus espaços imagem $R(\mathbf{Q}_1)$ e $R(\mathbf{Q}_2)$ forem mutuamente ortogonais.

Suponhamos que

$$\mathbf{Q}_l = \sum_{j=1}^{s_l} \alpha_{l,j} \alpha_{l,j}^t, \quad l = 1, 2$$

com $\{\alpha_{l,1}, \dots, \alpha_{l,s_l}\}$ uma base ortonormada para $R(\mathbf{Q}_l)$, $l = 1, 2$, vemos que \mathbf{Q}_1 e \mathbf{Q}_2 são ortogonais se e só se

$$\alpha_{1,j_1}^t \alpha_{2,j_2} = 0, \quad j_l = 1, \dots, s_l, \quad l = 1, 2,$$

ou seja, se e só se se verificar a igualdade

$$\mathbf{Q}_1 \mathbf{Q}_2 = \mathbf{0}_{n \times n}.$$

2.2 Produto de Kronecker

Dadas as matrizes $\mathbf{A} = [a_{i,j}]$ do tipo $r \times s$ e \mathbf{B} do tipo $m \times n$, o produto de Kronecker \otimes destas matrizes é dado pela matriz por blocos $rm \times sn$

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ \vdots & \dots & \vdots \\ a_{r1}\mathbf{B} & \dots & a_{rn}\mathbf{B} \end{bmatrix}.$$

Considerando as matrizes \mathbf{A}_j e \mathbf{B}_j , $j = 1, 2$ e os números reais α e β tem-se que

1. Em geral $\mathbf{A}_1 \otimes \mathbf{A}_2 \neq \mathbf{A}_2 \otimes \mathbf{A}_1$;

2. $(\alpha\mathbf{A}_1) \otimes (\beta\mathbf{A}_2) = \alpha\beta(\mathbf{A}_1 \otimes \mathbf{A}_2)$;

3. $(\mathbf{A}_1 \otimes \mathbf{A}_2)^t = \mathbf{A}_1^t \otimes \mathbf{A}_2^t$;

4. Estando os produtos usuais $\mathbf{A}_1\mathbf{A}_2$ e $\mathbf{B}_1\mathbf{B}_2$ definidos, então

$$(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2);$$

5. O produto \otimes é fechado para alguns tipos de matrizes. Assim, com

$$\mathbf{J}_n = \mathbf{1}_n\mathbf{1}_n$$

tem-se

$$\mathbf{1}_u \otimes \mathbf{1}_v = \mathbf{1}_{uv}$$

$$\mathbf{I}_u \otimes \mathbf{I}_v = \mathbf{I}_{uv}$$

e

$$\mathbf{J}_u \otimes \mathbf{J}_v = \mathbf{J}_{uv}.$$

Uma matriz ortogonal $\mathbf{P} = [p_{i,j}] \sim r \times r$ é ortogonal padronizada se

$$p_{1,1} = \dots = p_{1,r} = \frac{1}{\sqrt{r}}.$$

Representamos por \mathbf{P}_r as matrizes ortogonais padronizadas $r \times r$.

Mostra-se ainda que o produto \otimes de matrizes ortogonais [ortogonais padronizadas; simétricas; idempotentes] dá matrizes ortogonais [ortogonais padronizadas; simétricas; idempotentes] respectivamente. Por outro lado

$$\left(\sum_{i=1}^n a_i \mathbf{A}_i \right) \otimes \left(\sum_{j=1}^n b_j \mathbf{B}_j \right) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \mathbf{A}_i \otimes \mathbf{B}_j$$

e se as \mathbf{A}_j , $j = 1, 2$ são matrizes $n \times n$ invertíveis, então

$$(\mathbf{A}_1 \otimes \mathbf{A}_2)^{-1} = \mathbf{A}_1^{-1} \otimes \mathbf{A}_2^{-1}.$$

Dada $\mathbf{D}(\mathbf{r})$ uma matriz diagonal em que os elementos principais são as componentes de \mathbf{r} temos,

$$\mathbf{D}(\mathbf{r}_1) \otimes \mathbf{D}(\mathbf{r}_2) = \mathbf{D}(\mathbf{r}_1 \otimes \mathbf{r}_2).$$

Dadas as matrizes simétricas \mathbf{M}_1 e \mathbf{M}_2 , existem matrizes ortogonais \mathbf{P}_1 e \mathbf{P}_2 e matrizes $\mathbf{D}(\mathbf{r}_1)$ e $\mathbf{D}(\mathbf{r}_2)$ tal que

$$\mathbf{P}_j \mathbf{M}_j \mathbf{P}_j^t = \mathbf{D}(\mathbf{r}_j), \quad j = 1, 2.$$

Os elementos diagonais de $\mathbf{D}(\mathbf{r}_j)$ são os valores próprios de \mathbf{M}_j associados aos vectores próprios que são os vectores linha de \mathbf{P}_j , $j = 1, 2$. Levando em conta as

propriedades 3. e 4. teremos

$$\begin{aligned} (\mathbf{P}_1 \otimes \mathbf{P}_2)(\mathbf{M}_1 \otimes \mathbf{M}_2)(\mathbf{P}_1 \otimes \mathbf{P}_2)^t &= (\mathbf{P}_1 \mathbf{M}_1 \mathbf{P}_1^t) \otimes (\mathbf{P}_2 \mathbf{M}_2 \mathbf{P}_2^t) \\ &= \mathbf{D}(\mathbf{r}_1 \otimes \mathbf{r}_2). \end{aligned}$$

Assim, os valores próprios de $\mathbf{M}_1 \otimes \mathbf{M}_2$ são os produtos dos valores próprios de \mathbf{M}_1 e \mathbf{M}_2 e os vectores próprios corresponderão aos vectores linha de $(\mathbf{P}_1 \otimes \mathbf{P}_2)$. Como a característica de uma matriz simétrica é igual ao número dos seus valores próprios não nulos, temos

$$\text{car}(\mathbf{M}_1 \otimes \mathbf{M}_2) = \text{car}(\mathbf{M}_1) \otimes \text{car}(\mathbf{M}_2).$$

Dado o determinante de $\mathbf{M}_1[\mathbf{M}_2, \mathbf{M}_1 \otimes \mathbf{M}_2]$ ser o produto dos seus valores próprios, teremos

$$\det(\mathbf{M}_1 \otimes \mathbf{M}_2) = \det(\mathbf{M}_1) \otimes \det(\mathbf{M}_2).$$

Para além disso, ver Silvey (1975)

$$\text{car}(\mathbf{A}_j) = \text{car}(\mathbf{A}_j \mathbf{A}_j^t), \quad j = 1, 2$$

e como

$$(\mathbf{A}_1 \otimes \mathbf{A}_2)(\mathbf{A}_1 \otimes \mathbf{A}_2)^t = (\mathbf{A}_1 \mathbf{A}_1^t) \otimes (\mathbf{A}_2 \mathbf{A}_2^t)$$

temos ainda

$$\text{car}(\mathbf{A}_1 \otimes \mathbf{A}_2) = \prod_{j=1}^2 \text{car}(\mathbf{A}_j \mathbf{A}_j^t) = \prod_{j=1}^2 \text{car}(\mathbf{A}_j).$$

2.3 Álgebras de Jordan Comutativas de Matrizes Simétricas

Consideremos a seguinte definição

Definição 2.1 *Álgebras de Jordan Comutativas de Matrizes Simétricas (AJCS) são espaços lineares constituídos por matrizes simétricas $M_{n \times n}$ que comutam, e que contêm os quadrados das respectivas matrizes.*

Seely (1971) mostrou que toda a AJCS \mathcal{A} tem uma e só uma base, a sua base principal $bp(\mathcal{A}) = \{\mathbf{Q}_1, \dots, \mathbf{Q}_w\}$, constituída por matrizes de projecção ortogonal mutuamente ortogonais com $w = \dim(\mathcal{A})$. Sendo \mathbf{Q} uma MPO pertencente a \mathcal{A} tem-se

$$\mathbf{Q} = \sum_{j=1}^w q_j \mathbf{Q}_j.$$

Como \mathbf{Q} é idempotente e as $\mathbf{Q}_1, \dots, \mathbf{Q}_w$ são idempotentes e mutuamente ortogonais tem-se $q_j = 0$ ou $q_j = 1$, $j = 1, \dots, w$. Assim toda a MPO pertencente a \mathcal{A} será a soma de matrizes da $bp(\mathcal{A})$. Tendo-se

$$\mathbf{Q} = \sum_{j \in \mathcal{C}} \mathbf{Q}_j,$$

pelo que o espaço imagem de \mathbf{Q} será

$$R(\mathbf{Q}) = \boxplus_{j \in \mathcal{C}} R(\mathbf{Q}_j),$$

onde \boxplus representa a soma directa ortogonal de subespaços, logo com $r(\mathbf{U})$ a característica da matriz \mathbf{U} tendo

$$r(\mathbf{Q}) = \sum_{j \in \mathcal{C}} r(\mathbf{Q}_j).$$

Vê-se pois que, se $r(\mathbf{Q}) = 1$, \mathbf{Q} terá de pertencer à $bp(\mathcal{A})$.

Em particular se

$$\frac{1}{n} \mathbf{J}_n = \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^t \in \mathcal{A},$$

$\frac{1}{n} \mathbf{J}_n$ pertencerá a $bp(\mathcal{A})$ tomando-se $\mathbf{Q}_1 = \frac{1}{n} \mathbf{J}_n$ e dizendo-se que \mathcal{A} é regular.

Se $\sum_{j=1}^w \mathbf{Q}_j = \mathbf{I}_n \in \mathcal{A}$ diz-se que \mathcal{A} é completa. Se \mathcal{A} não for completa pode juntar-se $\mathbf{Q}^\perp = \mathbf{I}_n - \sum_{j=1}^w \mathbf{Q}_j$ à $bp(\mathcal{A})$, obtendo-se assim a base principal de uma nova AJCS $\bar{\mathcal{A}}$ que se diz a completada de \mathcal{A} .

Um exemplo simples de uma AJCS completa e regular é a $\mathcal{A}(n)$, que tem como base principal

$$\left\{ \frac{1}{n} \mathbf{J}_n, \mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right\}.$$

Sejam

$$\mathbf{M} = \sum_{j=1}^w a_j \mathbf{Q}_j,$$

$$\nabla_j = R(\mathbf{Q}_j), j = 1, \dots, w$$

e

$$\mathcal{C} = \{j; a_j \neq 0\}$$

tem-se

$$R(\mathbf{M}) = \boxplus_{j \in \mathcal{C}} \nabla_j.$$

Como os subespaços $\nabla_j, j = 1, \dots, w$ são ortogonais entre si, com

$$g_j = r(\mathbf{Q}_j), j = 1, \dots, w,$$

ter-se-á

$$car(\mathbf{M}) = \dim(R(\mathbf{M})) = \sum_{j \in \mathcal{C}} \dim(\nabla_j) = \sum_{j \in \mathcal{C}} car(\mathbf{Q}_j) = \sum_{j \in \mathcal{C}} g_j.$$

Dada $\mathbf{M} = \sum_{j=1}^w a_j \mathbf{Q}_j$, os a_1, \dots, a_w são os valores próprios de \mathbf{M} com multiplicidades g_1, \dots, g_w , logo se \mathcal{A} for completa

$$\det \left(\sum_{j=1}^w a_j \mathbf{Q}_j \right) = \prod_{j=1}^w a_j^{g_j},$$

2.3. ÁLGEBRAS DE JORDAN COMUTATIVAS DE MATRIZES SIMÉTRICAS 13

visto o determinante duma matriz diagonalizável ser o produto dos respectivos valores próprios.

A inversa de Moore-Penrose de uma matriz \mathbf{M} é a matriz \mathbf{M}^+ , que verifica as seguintes condições:

- $\mathbf{M}\mathbf{M}^+\mathbf{M} = \mathbf{M}$
- $\mathbf{M}^+\mathbf{M}\mathbf{M}^+ = \mathbf{M}^+$
- $(\mathbf{M}\mathbf{M}^+)^t = \mathbf{M}\mathbf{M}^+$
- $(\mathbf{M}^+\mathbf{M})^t = \mathbf{M}^+\mathbf{M}$

Prova-se que esta matriz existe sempre e é única, (ver Pollock (1979)).

Considere-se $\mathbf{M} \in \mathcal{A}$ com $\{\mathbf{Q}_1, \dots, \mathbf{Q}_w\}$ a base principal de \mathcal{A} . Como as $\mathbf{Q}_1, \dots, \mathbf{Q}_w$ são idempotentes e mutuamente ortogonais ter-se-á

$$\mathbf{M}^+ = \sum_{j=1}^w a_j^+ \mathbf{Q}_j$$

com $a_j^+ = 0$ se $a_j = 0$ e $a_j^+ = a_j^{-1}$ se $a_j \neq 0$, $j = 1, \dots, w$.

Assim as AJC contêm as inversas de Moore-Penrose das suas matrizes. Em particular se \mathbf{M} for regular ter-se-á

$$\mathbf{M}^{-1} = \sum_{j=1}^w a_j^{-1} \mathbf{Q}_j.$$

Estabeleçamos a

Proposição 2.1 *Uma AJC \mathcal{A} contém matrizes regulares se e só se for completa.*

Demonstração Seja $\mathbf{M} = \sum_{j=1}^w a_j \mathbf{Q}_j \in \mathcal{A}$ uma matriz regular então

$$\mathbf{M}^{-1} = \sum_{j=1}^w a_j^{-1} \mathbf{Q}_j \in \mathcal{A},$$

bem como

$$\mathbf{I}_n = \mathbf{M}\mathbf{M}^{-1} = \sum_{j=1}^w \mathbf{Q}_j \in \mathcal{A}$$

e \mathcal{A} será regular. Sendo \mathcal{A} regular com base principal $\{\mathbf{Q}_1, \dots, \mathbf{Q}_w\}$ as matrizes de \mathcal{A} com $\mathcal{C} = \{1, \dots, w\}$ serão regulares o que completa a prova.

□

Sendo \mathbf{A} simétrica $n \times n$ com $\text{car}(\mathbf{A}) = l$, pode-se ordenar os vectores linha de uma matriz \mathbf{P} que diagonaliza \mathbf{A} por forma a ter-se

$$\mathbf{P}\mathbf{M}\mathbf{P}^t = \mathbf{D}(r_1, \dots, r_l, 0, \dots, 0)$$

sendo r_1, \dots, r_l os valores próprios não nulos de \mathbf{M} , ter-se-á então

$$\mathbf{M} = \mathbf{P}^t \mathbf{D}(r_1, \dots, r_l, 0, \dots, 0) \mathbf{P}$$

bem como

$$\mathbf{M}^+ = \mathbf{P}^t \mathbf{D}(r_1^{-1}, \dots, r_l^{-1}, 0, \dots, 0) \mathbf{P}.$$

Capítulo 3

Inferência

3.1 Momentos

Nesta secção são apresentados alguns resultados que utilizaremos nos capítulos seguintes. As demonstrações podem ser consultadas em, por exemplo Mexia (1995). Dado que o operador E é linear temos para os vectores médios

$$E(\mathbf{A}\mathbf{X} + \mathbf{b}) = \mathbf{A}E(\mathbf{X}) + \mathbf{b}.$$

As matrizes de covariância e de covariância cruzada são dadas por

$$\mathbf{\Sigma}(\mathbf{X}) = E [(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^t]$$

e

$$\mathbf{\Sigma}(\mathbf{X}_1; \mathbf{X}_2) = E [(\mathbf{X}_1 - E(\mathbf{X}_1))(\mathbf{X}_2 - E(\mathbf{X}_2))^t],$$

com $\underline{\underline{\mu}} = E(\mathbf{X}\mathbf{X}^t)$, $\boldsymbol{\mu} = E(\mathbf{X})$ e $\boldsymbol{\mu}_l = E(\mathbf{X}_l)$, $l = 1, 2$, podemos reescrever estas expressões como

$$\mathbf{\Sigma}(\mathbf{X}) = \underline{\underline{\mu}} - \boldsymbol{\mu}\boldsymbol{\mu}^t$$

e

$$\mathfrak{Y}(\mathbf{X}_1; \mathbf{X}_2) = E(\mathbf{X}_1 \mathbf{X}_2^t) - \boldsymbol{\mu}_1 \boldsymbol{\mu}_2^t.$$

Por outro lado é fácil utilizar a linearidade do operador E para mostrar que

$$\mathfrak{Y}(\mathbf{A}\mathbf{X} + \mathbf{b}) = \mathbf{A}\mathfrak{Y}(\mathbf{X})\mathbf{A}^t$$

e que

$$\mathfrak{Y}(\mathbf{A}_1\mathbf{X}_1 + \mathbf{b}_1; \mathbf{A}_2\mathbf{X}_2 + \mathbf{b}_2) = \mathbf{A}_1\mathfrak{Y}(\mathbf{X}_1; \mathbf{X}_2)\mathbf{A}_2^t.$$

3.2 Funções Geradoras de Momentos e de Cumulantes

A função geradora de momentos do vector aleatório \mathbf{X}^n é dada por

$$\varphi_{\mathbf{X}}(\mathbf{u}) = E(e^{\mathbf{u}^t \mathbf{X}}).$$

Caso $\varphi(\mathbf{u}^n | \mathbf{X}^n)$ esteja definida num conjunto aberto contendo a origem é, (ver Lukacs & Laha (1964)), indefinidamente derivável na origem tendo-se

$$\left(\frac{\partial^r \varphi_{\mathbf{X}}(\mathbf{u})}{\prod_{l=1}^n \partial \mathbf{u}_l^{j_l}} \right)_{\mathbf{u}^n = \mathbf{0}^n} = E \left(\prod_{l=1}^n \mathbf{X}_l^{j_l} \right) = \boldsymbol{\mu}'_{j_1, \dots, j_n} = \boldsymbol{\mu}'_{\mathbf{j}^n}$$

onde $r = \sum_{l=1}^n j_l$ e \mathbf{j}^n o vector com componentes j_1, \dots, j_n .

Assim, sendo $\underline{grad}\varphi_{\mathbf{X}}(\mathbf{u})$ e $\underline{Hess}\varphi_{\mathbf{X}}(\mathbf{u})$ o gradiente e a matriz hessiana de $\varphi_{\mathbf{X}}(\mathbf{u})$, tem-se

$$(\underline{grad}\varphi_{\mathbf{X}}(\mathbf{u}))_{\mathbf{u}=\mathbf{0}} = \boldsymbol{\mu}$$

e

$$(\underline{Hess}\varphi_{\mathbf{X}}(\mathbf{u}))_{\mathbf{u}=\mathbf{0}} = \underline{\underline{\boldsymbol{\mu}}}.$$

Quando \mathbf{X}_1 e \mathbf{X}_2 são vectores aleatórios independentes $e^{\mathbf{u}_1^t \mathbf{X}_1}$ será independente de $e^{\mathbf{u}_2^t \mathbf{X}_2}$ logo, com

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$$

e

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

ter-se-á

$$\begin{aligned} \varphi_{\mathbf{X}}(\mathbf{u}) &= E\left(e^{\sum_{i=1}^2 \mathbf{u}_i^t \mathbf{X}_i}\right) = E(e^{\mathbf{u}_1^t \mathbf{X}_1} e^{\mathbf{u}_2^t \mathbf{X}_2}) \\ &= E(e^{\mathbf{u}_1^t \mathbf{X}_1}) E(e^{\mathbf{u}_2^t \mathbf{X}_2}) \\ &= \varphi_{\mathbf{X}_1}(\mathbf{u}_1) \varphi_{\mathbf{X}_2}(\mathbf{u}_2). \end{aligned}$$

Pode mostrar-se, (ver Wilks (1961)), que esta igualdade é condição necessária e suficiente para a independência de \mathbf{X}_1 e \mathbf{X}_2 .

Se \mathbf{X}_1 e \mathbf{X}_2 além de independentes tiverem o mesmo número de componentes obtém-se analogamente

$$\begin{aligned} \varphi_{\mathbf{X}_1 + \mathbf{X}_2}(\mathbf{u}) &= E\left(e^{\mathbf{u}^t \sum_{i=1}^2 \mathbf{X}_i}\right) \\ &= E\left(e^{\mathbf{u}^t \mathbf{X}_1} e^{\mathbf{u}^t \mathbf{X}_2}\right) \\ &= \prod_{i=1}^2 E(e^{\mathbf{u}^t \mathbf{X}_i}) = \prod_{i=1}^2 \varphi_{\mathbf{X}_i}(\mathbf{u}). \end{aligned}$$

Por outro lado (ver Lucaks & Laha (1964)) as funções geradoras de momentos identificam as distribuições e se

$$\varphi_{\mathbf{X}_n}(\mathbf{u}) \xrightarrow{n \rightarrow \infty} \varphi_{\mathbf{X}}(\mathbf{u})$$

teremos

$$F_{\mathbf{X}_n} \xrightarrow{n \rightarrow \infty} F_{\mathbf{X}},$$

com $F_{\mathbf{X}_n}$ e $F_{\mathbf{X}}$ as distribuições de \mathbf{X}_n e de \mathbf{X} . Esta convergência dá-se para os pontos de continuidade de $F_{\mathbf{X}}$. Além disso se $F_{\mathbf{X}}$ for contínua, temos devido ao Teorema de Polya,

$$F_{\mathbf{X}_n} \xrightarrow[n \rightarrow \infty]{u} F_{\mathbf{X}},$$

com \xrightarrow{u} indicando convergência uniforme. Este teorema encontra-se estabelecido no Apêndice B.

Observe-se ainda que

$$\varphi_{\mathbf{A}\mathbf{X}+\mathbf{b}}(\mathbf{u}) = E\left(e^{\mathbf{u}^t(\mathbf{A}\mathbf{X}+\mathbf{b})}\right) = E\left(e^{(\mathbf{A}^t\mathbf{u})^t\mathbf{X}+\mathbf{b}^t\mathbf{u}}\right) = e^{\mathbf{b}^t\mathbf{u}}\varphi_{\mathbf{X}}(\mathbf{A}^t\mathbf{u}).$$

Temos também a função geradora de cumulantes

$$\psi_{\mathbf{X}}(\mathbf{u}) = \log(\varphi_{\mathbf{X}}(\mathbf{u}))$$

sendo fácil de verificar que

$$(\underline{grad}\psi_{\mathbf{X}}(\mathbf{u}))_{\mathbf{u}=\mathbf{0}} = \boldsymbol{\mu}$$

e

$$(\underline{Hess}\psi_{\mathbf{X}}(\mathbf{u}))_{\mathbf{u}=\mathbf{0}} = \mathcal{V}(\mathbf{X}),$$

e que se \mathbf{X}_1 é independente de \mathbf{X}_2 , \mathbf{X}_1 (i) \mathbf{X}_2 ,

$$\psi_{\mathbf{X}_1+\mathbf{X}_2}(\mathbf{u}) = \psi_{\mathbf{X}_1}(\mathbf{u}) + \psi_{\mathbf{X}_2}(\mathbf{u}).$$

A distribuição $F(|\theta)$ é reprodutiva em relação a θ , se com \mathbf{X}_1 (i) \mathbf{X}_2 e tendo \mathbf{X}_l a distribuição $F(|\boldsymbol{\theta}_l)$, $l = 1, 2$, $\mathbf{X}_1 + \mathbf{X}_2$ tem distribuição $F(|\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2)$, vindo para as funções geradoras de momentos e cumulantes

$$\varphi(\mathbf{u}|\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2) = \prod_{l=1}^2 \varphi(\mathbf{u}|\boldsymbol{\theta}_l)$$

e

$$\psi(\mathbf{u}|\theta_1 + \theta_2) = \sum_{l=1}^2 \psi(\mathbf{u}|\theta_l).$$

Usando a equação aditiva funcional, (ver Parzen (1962)), teremos

$$\psi(\mathbf{u}|\boldsymbol{\theta}) = \boldsymbol{\theta}\psi(\mathbf{u}|1).$$

3.3 Vectores Normais

O vector \mathbf{X}^n com $E(\mathbf{X}^n) = \boldsymbol{\mu}^n$ e $\boldsymbol{\Sigma}(\mathbf{X}^n) = \mathbf{M}$ é um vector normal regular se tiver a densidade

$$n(\mathbf{x}^n|\boldsymbol{\mu}^n, \mathbf{M}) = \frac{e^{-1/2(\mathbf{x}-\boldsymbol{\mu})^t\mathbf{M}^{-1}(\mathbf{x}-\boldsymbol{\mu})}}{(2\pi)^{n/2}\sqrt{\det(\mathbf{M})}}$$

em que \mathbf{M} é regular. De seguida vamos obter a função geradora de momentos $\varphi_N(\mathbf{u}^n|\boldsymbol{\mu}^n, \mathbf{M})$ deste vector, começando primeiro pelo caso reduzido e depois passando ao caso geral.

No caso reduzido em que $\boldsymbol{\mu}^n = \mathbf{0}^n$ e $\mathbf{V} = \mathbf{I}_n$ tem-se:

$$\begin{aligned} \varphi_N(\mathbf{u}^n|\mathbf{0}^n, \mathbf{I}_n) &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{\mathbf{u}^t\mathbf{x}} \frac{e^{-1/2\mathbf{x}^t\mathbf{I}_n-1\mathbf{x}}}{(2\pi)^{n/2}} \prod_{i=1}^n dx_i \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{e^{\sum_{i=1}^n u_i x_i - 1/2 \sum_{i=1}^n x_i^2}}{(2\pi)^{n/2}} \prod_{i=1}^n dx_i \\ &= \prod_{i=1}^n \int_{-\infty}^{+\infty} \frac{e^{-\frac{x_i^2 - 2u_i x_i}{2}}}{\sqrt{2\pi}} dx_i \\ &= \prod_{i=1}^n e^{\frac{u_i^2}{2}} \int_{-\infty}^{+\infty} \frac{e^{-\frac{(x_i - u_i)^2}{2}}}{\sqrt{2\pi}} dx_i \\ &= e^{\frac{1}{2} \sum_{i=1}^n u_i^2} \\ &= e^{\frac{1}{2} \mathbf{u}^t \mathbf{I}_n \mathbf{u}}. \end{aligned}$$

pois os últimos integrais são iguais a 1 visto serem os integrais entre $-\infty$ e $+\infty$ de densidades.

De seguida vamos passar ao caso geral. Sendo \mathbf{M} a matriz de covariância regular a mesma é definida positiva. Seja \mathbf{G}_0 a uniformizadora directa de \mathbf{M} , como

$$\mathbf{M}^{-1} = \mathbf{G}_0^t \mathbf{G}_0 \quad (*)$$

e

$$\det(\mathbf{G}_0^t) = \det(\mathbf{G}_0),$$

tem-se

$$\det(\mathbf{G}_0) = \sqrt{\det(\mathbf{M}^{-1})}.$$

Fazendo a mudança de variável

$$\mathbf{x}^n = \boldsymbol{\mu}^n + \mathbf{G}_0^{-1} \mathbf{z}^n \Leftrightarrow \mathbf{x}^n - \boldsymbol{\mu}^n = \mathbf{G}_0^{-1} \mathbf{z}^n$$

com o jacobiano dado por

$$\det(\mathbf{G}_0^{-1}) = \sqrt{\det(\mathbf{M})}$$

obtém-se

$$\begin{aligned} \varphi_N(\mathbf{u}^n | \boldsymbol{\mu}^n, \mathbf{M}) &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{\mathbf{u}^t \mathbf{x}} \frac{e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t \mathbf{M}^{-1}(\mathbf{x}-\boldsymbol{\mu})}}{(2\pi)^{n/2} \sqrt{\det(\mathbf{M})}} \prod_{i=1}^n dx_i \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{\mathbf{u}^t(\boldsymbol{\mu} + \mathbf{G}_0^{-1} \mathbf{z})} \frac{e^{-\frac{1}{2} \mathbf{z}^t \mathbf{z}}}{(2\pi)^{n/2}} \prod_{i=1}^n dz_i, \end{aligned}$$

pois $(\mathbf{x}^n - \boldsymbol{\mu}^n)^t \mathbf{M}^{-1}(\mathbf{x}^n - \boldsymbol{\mu}^n) = \mathbf{z}^t (\mathbf{G}_0^{-1})^t \mathbf{G}_0^t \mathbf{G}_0 \mathbf{G}_0^{-1} \mathbf{z} = \mathbf{z}^t \mathbf{z}$.

Assim, prosseguindo, vem

$$\begin{aligned}
\varphi_N(\mathbf{u}^n | \boldsymbol{\mu}^n, \mathbf{M}) &= e^{\mathbf{u}^t \boldsymbol{\mu}} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{((\mathbf{G}_0^{-1})^t \mathbf{u})^t \mathbf{z}} e^{\frac{-1}{2} \mathbf{z}^t \mathbf{I}_n^{-1} \mathbf{z}} \prod_{i=1}^n dz_i \\
&= e^{\boldsymbol{\mu}^t \mathbf{u}} \varphi((\mathbf{G}_0^{-1})^t \mathbf{u} | \mathbf{0}^n, \mathbf{I}_n) \\
&= e^{\boldsymbol{\mu}^t \mathbf{u}} e^{\frac{1}{2} \mathbf{u}^t \mathbf{G}_0^{-1} (\mathbf{G}_0^{-1})^t \mathbf{u}} \\
&= e^{\boldsymbol{\mu}^t \mathbf{u} + \frac{1}{2} \mathbf{u}^t \mathbf{M} \mathbf{u}}
\end{aligned}$$

já que por (*) tem-se

$$\mathbf{G}_0^{-1} (\mathbf{G}_0^{-1})^t = \mathbf{G}_0^{-1} (\mathbf{G}_0^t)^{-1} = (\mathbf{G}_0^t \mathbf{G}_0)^{-1} = (\mathbf{M}^{-1})^{-1} = \mathbf{M}.$$

Vejam os de seguida como se comporta a função geradora de momentos no que diz respeito às transformações lineares e independência. A densidade $n(\mathbf{x}^n | \boldsymbol{\mu}^n, \mathbf{M})$ corresponde a f.g.m. $\varphi_N(\mathbf{u}^n | \boldsymbol{\mu}^n, \mathbf{M})$. Da mesma forma, se \mathbf{X}^n tiver a densidade $n(\mathbf{x}^n | \boldsymbol{\mu}^n, \mathbf{M})$, ter-se-á

$$\begin{aligned}
\varphi(\mathbf{u}^s | \mathbf{W} \mathbf{X}^n + \mathbf{b}^s) &= e^{\mathbf{b}^t \mathbf{u}} \varphi(\mathbf{W}^t \mathbf{u} | \boldsymbol{\mu}^n, \mathbf{M}) \\
&= e^{\boldsymbol{\mu}^t \mathbf{u}} \varphi((\mathbf{G}_0^{-1})^t \mathbf{u} | \mathbf{0}^n, \mathbf{I}_n) \\
&= e^{\mathbf{b}^t \mathbf{u}} e^{\boldsymbol{\mu}^t (\mathbf{W} \mathbf{u}) + \frac{1}{2} (\mathbf{W}^t \mathbf{u})^t \mathbf{M} (\mathbf{W}^t \mathbf{u})} \\
&= e^{(\mathbf{W} \boldsymbol{\mu} + \mathbf{b})^t \mathbf{u} + \frac{1}{2} \mathbf{u}^t (\mathbf{W} \mathbf{M} \mathbf{W}^t) \mathbf{u}} \\
&= \varphi(\mathbf{u}^s | \mathbf{W} \boldsymbol{\mu} + \mathbf{b}, \mathbf{W} \mathbf{M} \mathbf{W}^t),
\end{aligned}$$

visto

$$E(\mathbf{W} \mathbf{X}^n + \mathbf{b}^s) = \mathbf{W} \boldsymbol{\mu}^n + \mathbf{b}^s$$

e

$$\mathfrak{V}(\mathbf{W} \mathbf{X}^n + \mathbf{b}^s) = \mathbf{W} \mathbf{M} \mathbf{W}^t$$

vendo-se que $\varphi(\mathbf{u}^s | \mathbf{W}\mathbf{X}^n + \mathbf{b}^s)$ é da mesma forma que as f.g.m. que se obtiveram para vectores normais regulares. Assim, o conceito de vector normal torna-se mais amplo, dizendo-se que $\mathbf{X}^{/s}$ com

$$E(\mathbf{X}^{/s}) = \boldsymbol{\mu}^{/s}$$

e

$$\boldsymbol{\Sigma}(\mathbf{X}^{/s}) = \mathbf{M}'$$

é normal se

$$\varphi(\mathbf{u}^s | \mathbf{X}^{/s}) = \varphi_N(\mathbf{u}^s | \boldsymbol{\mu}^{/s}, \mathbf{M}') = e^{\boldsymbol{\mu}'^t \mathbf{u} + \frac{1}{2} \mathbf{u}^t \mathbf{M}' \mathbf{u}}.$$

Se \mathbf{M}' for regular, $\mathbf{X}^{/s}$ terá densidade $n(\mathbf{x}^s | \boldsymbol{\mu}^{/s}, \mathbf{M}')$ sendo um vector normal regular (ver Lukacs & Laha (1964)). Se \mathbf{M}' for singular, $\mathbf{X}^{/s}$ será um vector normal singular, não tendo assim densidade.

Suponhamos que \mathbf{X}_1 e \mathbf{X}_2 têm vectores médios $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$, matrizes de covariância $\boldsymbol{\Sigma}(\mathbf{X}_1)$ e $\boldsymbol{\Sigma}(\mathbf{X}_2)$ e matriz de covariância cruzada nula. Então

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$$

terá vector médio

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$$

e matriz de covariância

$$\boldsymbol{\Sigma}(\mathbf{X}) = \begin{bmatrix} \boldsymbol{\Sigma}(\mathbf{X}_1) & 0 \\ 0 & \boldsymbol{\Sigma}(\mathbf{X}_2) \end{bmatrix}.$$

Assim se \mathbf{X} for normal terá, com

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

função geradora de momentos

$$\begin{aligned}\varphi_{\mathbf{X}}(\mathbf{u}) &= e^{\boldsymbol{\mu}^t \mathbf{u} + \frac{1}{2} \mathbf{u}^t \boldsymbol{\Sigma}(\mathbf{X}) \mathbf{u}} \\ &= \prod_{l=1}^2 e^{\mu_l^t \mathbf{u}_l + \frac{1}{2} \mathbf{u}_l^t \boldsymbol{\Sigma}(\mathbf{X}_l) \mathbf{u}_l} \\ &= \prod_{l=1}^2 \varphi_{\mathbf{X}_l}(\mathbf{u}_l)\end{aligned}$$

e, conseqüentemente, \mathbf{X}_1 (i) \mathbf{X}_2 vê-se assim que a nulidade da matriz de covariância cruzada é condição necessária e suficiente da independência de vectores com distribuição conjunta normal. No caso geral é apenas condição necessária.

3.4 Convergência em Distribuição

Nesta secção usaremos funções características e integrais de Lebesgue-Stieljes apresentando a demonstração de alguns resultados, uma vez que são menos conhecidas do que as das secções anteriores. A função característica de \mathbf{X} será (assumindo que \mathbf{X} tem k componentes):

$$\begin{aligned}\xi_{\mathbf{X}}(\mathbf{u}) &= E(e^{i\mathbf{u}^t \mathbf{X}}) \\ &= \int \dots \int_{R^k} e^{i\mathbf{u}^t \mathbf{X}} dF_{\mathbf{X}} \\ &= \int \dots \int_{R^k} \cos(\mathbf{u}^t \mathbf{X}) dF_{\mathbf{X}} + i \int \dots \int_{R^k} \sin(\mathbf{u}^t \mathbf{X}) dF_{\mathbf{X}}\end{aligned}$$

Além disso, é sabido (ver Lukacs & Laha, (1964)) que se

$$\xi_{\mathbf{X}_n}(\mathbf{u}) \xrightarrow[n \rightarrow \infty]{} \xi_{\mathbf{X}}(\mathbf{u})$$

temos

$$\mathbf{X}_n \xrightarrow[n \rightarrow \infty]{F} \mathbf{X},$$

e se

$$\xi_{\mathbf{X}_\theta}(\mathbf{u}) \xrightarrow{\theta \rightarrow +\infty} \xi_{\mathbf{X}}(\mathbf{u})$$

temos

$$\mathbf{X}_\theta \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{X}.$$

Observe-se que se tem

$$\xi_{\mathbf{X}_\theta}(\mathbf{u}) \xrightarrow{\theta \rightarrow +\infty} \xi_{\mathbf{X}}(\mathbf{u})$$

se, qualquer que seja a sequência $\{\theta_n\}$ tal que

$$\theta_n \xrightarrow[n \rightarrow +\infty]{} +\infty,$$

se tiver

$$\xi_{\mathbf{X}_{\theta_n}}(\mathbf{u}) \xrightarrow[n \rightarrow +\infty]{} \xi_{\mathbf{X}}(\mathbf{u}).$$

Estabeleçamos, a

Proposição 3.1 *Se $\mathbf{X}_\theta \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{X}$ e $\mathbf{g}(\cdot)$ é uma função contínua, temos*

$$\mathbf{g}(\mathbf{X}_\theta) \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{g}(\mathbf{X}).$$

Demonstração Como $\cos(\mathbf{v}^t \mathbf{g}(\mathbf{x}))$ e $\sin(\mathbf{v}^t \mathbf{g}(\mathbf{x}))$ são funções contínuas e limitadas, teremos de acordo com o lema de Helley-Brey

$$\begin{aligned} \int \dots \int_{R^k} \cos(\mathbf{v}^t \mathbf{g}(\mathbf{x})) dF(\mathbf{x}|\theta_n) &\xrightarrow[n \rightarrow \infty]{} \int \dots \int_{R^k} \cos(\mathbf{v}^t \mathbf{g}(\mathbf{x})) dF(\mathbf{x}) \\ \int \dots \int_{R^k} \sin(\mathbf{v}^t \mathbf{g}(\mathbf{x})) dF(\mathbf{x}|\theta_n) &\xrightarrow[n \rightarrow \infty]{} \int \dots \int_{R^k} \sin(\mathbf{v}^t \mathbf{g}(\mathbf{x})) dF(\mathbf{x}), \end{aligned}$$

com $F(\mathbf{x}|\theta)$ e $F(\mathbf{x})$ as distribuições de \mathbf{X}_θ e de \mathbf{X} . Assim,

$$\xi_{\mathbf{g}(\mathbf{X}_\theta)}(\mathbf{v}) \xrightarrow[\theta \rightarrow \infty]{} \xi_{\mathbf{g}(\mathbf{X})}(\mathbf{v})$$

pelo que,

$$g(\mathbf{X}_\theta) \xrightarrow[\theta \rightarrow +\infty]{F} g(\mathbf{X}).$$

□

Corolário 1 Se $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{V})$ com \mathbf{V} regular, $\mathbf{X}_\theta \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{X}$ e $\mathbf{V}_\theta \xrightarrow[\theta \rightarrow +\infty]{} \mathbf{V}$,

$$\mathbf{Q}_\theta = (\mathbf{X}_\theta - \mathbf{b})^t \mathbf{V}^{-1} (\mathbf{X}_\theta - \mathbf{b}) \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{Q}$$

onde \mathbf{Q} é um qui-quadrado com $r = \text{car}(\mathbf{V})$ graus de liberdade e o parâmetro de não-centralidade $\delta = (\boldsymbol{\mu} - \mathbf{b})^t \mathbf{V}^{-1} (\boldsymbol{\mu} - \mathbf{b})$, pondo $\mathbf{Q} \sim \chi_{r,\delta}^2$.

Demonstração Temos que

$$(\mathbf{X} - \mathbf{b})^t \mathbf{V}^{-1} (\mathbf{X} - \mathbf{b})$$

se distribui como um qui-quadrado com r graus de liberdade e parâmetro de não centralidade δ , (ver por exemplo Mexia (1990)).

Os elementos de \mathbf{V}^{-1} são dados por funções contínuas dos elementos de \mathbf{V} , logo se

$$\mathbf{V}_{\theta_n} \xrightarrow[n \rightarrow +\infty]{} \mathbf{V}$$

sempre que $\theta_n \rightarrow +\infty$ temos

$$\mathbf{V}_{\theta_n}^{-1} \xrightarrow[n \rightarrow \infty]{} \mathbf{V}^{-1},$$

pelo que

$$\mathbf{V}_\theta^{-1} \xrightarrow[\theta \rightarrow +\infty]{} \mathbf{V}^{-1},$$

e assim

$$(\mathbf{X} - \mathbf{b})^t \mathbf{V}_\theta^{-1} (\mathbf{X} - \mathbf{b}) \xrightarrow[\theta \rightarrow +\infty]{} (\mathbf{X} - \mathbf{b})^t \mathbf{V}^{-1} (\mathbf{X} - \mathbf{b})$$

Para completar a demonstração basta aplicar a Proposição 3.1. \square

Estabeleçamos,

Lema 3.2 *Se N_θ tem distribuição $F_N(|\theta)$ reprodutiva para θ , com valor médio $\mu_N(|\theta)$ e variância $\sigma_N^2(|\theta)$, temos $\frac{N(|\theta)}{\theta} \xrightarrow[\theta \rightarrow +\infty]{p} \mu_N(|1)$.*

Demonstração Dado termos

$$\psi_N(u|\theta) = \theta\psi_N(u|1)$$

teremos

$$\mu(N|\theta) = \theta\mu(N|1)$$

e

$$\sigma^2(N|\theta) = \theta\sigma^2(N|1),$$

logo de acordo com a desigualdade de Bienaymé-Tchebycheff, temos qualquer que seja $\varepsilon > 0$,

$$pr \left(\left| \frac{N_\theta}{\theta} - \mu(|1) \right| \leq \varepsilon \right) \geq 1 - \frac{\sigma^2(N|1)}{\theta\varepsilon^2} \xrightarrow[\theta \rightarrow \infty]{} 1$$

o que estabelece a tese. \square

De seguida temos,

Lema 3.3 *Se $\mathbf{V}_n \xrightarrow[n \rightarrow \infty]{} \mathbf{V}$, com \mathbf{V} regular, e $F_N(|\theta)$ for reprodutiva para θ temos*

$$\mathbf{V}_N \xrightarrow[\theta \rightarrow +\infty]{p} \mathbf{V}$$

bem como

$$\mathbf{V}_N^{-1} \xrightarrow[\theta \rightarrow +\infty]{p} \mathbf{V}^{-1}.$$

Demonstração Raciocinando como atrás vemos que, qualquer que seja $\varepsilon > 0$, existe $n(\varepsilon)$ tal que, para $n > n(\varepsilon)$, o valor absoluto da diferença entre os elementos correspondentes de \mathbf{V}_n e \mathbf{V} [\mathbf{V}_n^{-1} e \mathbf{V}^{-1}] é inferior a ε . Basta agora recordar que, de acordo como Lema 3.1, temos

$$pr(N \geq n(\varepsilon)) \xrightarrow{\theta \rightarrow +\infty} 1,$$

para completar a demonstração. \square

Tomando

$$\mathbf{v}_N = \mathbf{vec}(\mathbf{V}_N)$$

$$\mathbf{v} = \mathbf{vec}(\mathbf{V}),$$

teremos

$$\mathbf{v}_N \xrightarrow[\theta \rightarrow +\infty]{p} \mathbf{v}.$$

Observe-se que

$$h(\mathbf{y}, \mathbf{v}) = \mathbf{y}^t \mathbf{V}^{-1} \mathbf{y}$$

é uma função contínua de \mathbf{y} e de $\mathbf{v} = \mathbf{vec}(\mathbf{V})$ tendo-se

$$\mathbf{Q} = h(\mathbf{X}, \mathbf{v}).$$

Ponhamos $\mathbf{V}_\theta = \mathbf{V}_{N_\theta}$, $\mathbf{v}_\theta = \mathbf{v}_{N_\theta}$ e estabeleçamos a

Proposição 3.4 *Se $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{V})$ e se N_θ tiver distribuição reprodutiva para θ , temos*

$$\mathbf{X}_\theta \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{X}$$

e

$$\mathbf{V}_\theta \xrightarrow[\theta \rightarrow \infty]{p} \mathbf{V}$$

com \mathbf{V} regular, bem como

$$\mathbf{Q}_\theta \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{Q}$$

com

$$\mathbf{Q}_\theta = h(\mathbf{X}_\theta, \mathbf{v}_\theta)$$

Demonstração A demonstração resulta directamente da Proposição 3.1 dado que

$$[\mathbf{X}_\theta^t, \mathbf{v}_\theta^t]^t \xrightarrow[\theta \rightarrow \infty]{F} [\mathbf{X}^t, \mathbf{v}^t]^t.$$

Observemos agora que, se $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$, $\mathbf{X}_\theta = [\mathbf{X}_{1,\theta_1}^t, \dots, \mathbf{X}_{L,\theta_L}^t]^t$ tem distribuição

$$F_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\theta}) = \prod_{l=1}^k F_{\mathbf{X}_i}(x_i, \theta_i)$$

quando os \mathbf{X}_{l,θ_l} , $l = 1, \dots, L$, são independentes. Assim, quando

$$\mathbf{X}_{l,\theta_l} \xrightarrow[\theta_l \rightarrow \infty]{F} \mathbf{X}_l, \quad l = 1, \dots, L$$

temos

$$\mathbf{X}_\theta \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{X}$$

com

$$\mathbf{X} = [\mathbf{X}_1^t, \dots, \mathbf{X}_L^t]^t,$$

sendo os $\mathbf{X}_1, \dots, \mathbf{X}_L$ assumidos como independentes. \square

Podemos notar que quando a sequência de vectores $\boldsymbol{\theta}$ tem componentes a tender para $+\infty$, podemos escrever $\{(\theta_{1,m_n,1}, \dots, \theta_{L,m_n,L})\} \rightarrow +\infty$. Usando esta notação, podemos raciocinar como acima para estabelecer a generalização da Proposição 3.1 dada por:

Proposição 3.5 Se $\mathbf{X}_\theta \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{X}$ e $g(\cdot)$ é uma função contínua então

$$g(\mathbf{X}_\theta) \xrightarrow[\theta \rightarrow +\infty]{F} g(\mathbf{X}).$$

e a:

Proposição 3.6 Se:

- (i) $\mathbf{X}_\theta \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{V})$ onde $\boldsymbol{\mu} = [\boldsymbol{\mu}_1^t \dots \boldsymbol{\mu}_L^t]^t$ e \mathbf{V} é matriz diagonal por blocos com blocos principais $\mathbf{V}_1, \dots, \mathbf{V}_L$ regulares,
- (ii) $\mathbf{V}_{l,\theta_l} \xrightarrow[\theta_l \rightarrow \infty]{} \mathbf{V}_l, l = 1, \dots, L$
- (iii) \mathbf{N}_{l,θ_l} tiver distribuição reprodutiva para $\theta_l, l = 1, \dots, L$.

então, com $\boldsymbol{\theta} = (\theta_1, \dots, \theta_L)$, tem-se

$$Q_{\theta,N} \xrightarrow[\theta \rightarrow +\infty]{F} Q$$

onde

$$Q_{\theta,N} = (\mathbf{X}_\theta - \mathbf{b})^t \mathbf{V}_N^{-1} (\mathbf{X}_\theta - \mathbf{b})$$

$$Q = (\mathbf{X}_\theta - \boldsymbol{\mu})^t \mathbf{V}^{-1} (\mathbf{X}_\theta - \boldsymbol{\mu}) \sim \chi_{r,\delta}^2,$$

com $r = \sum_{l=1}^L r_l$, sendo r_l o número de componentes de $\mathbf{X}_l, l = 1, \dots, L$ e $\delta = \sum_{l=1}^L \delta_l$,

com

$$\delta_l = (\boldsymbol{\mu}_l - \mathbf{b}_l)^t \mathbf{V}^{-1} (\boldsymbol{\mu}_l - \mathbf{b}_l), l = 1, \dots, L,$$

tendo-se $\mathbf{b} = [\mathbf{b}_1^t \dots \mathbf{b}_L^t]^t$.

Quando os vectores linha da matriz \mathbf{A} são linearmente independentes e \mathbf{V} é regular, a matriz $\mathbf{A}\mathbf{V}\mathbf{A}^t$ também é regular, (ver Mexia (1995)), tendo-se

$$\mathbf{A}\mathbf{X}_\theta \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{A}\mathbf{X} \sim N(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\mathbf{V}\mathbf{A}^t)$$

sendo ainda fácil mostrar que, assumidas as hipóteses (i), (ii) e (iii) da proposição 3.6 e tomando

$$\begin{aligned}\dot{Q}_{\theta,N} &= (\mathbf{X}_\theta - \mathbf{b})^t \mathbf{A}^t (\mathbf{A} \mathbf{V}_N \mathbf{A}^t)^{-1} \mathbf{A} (\mathbf{X}_\theta - \mathbf{b}) \\ \dot{Q}_\theta &= (\mathbf{X}_\theta - \boldsymbol{\mu})^t \mathbf{A}^t (\mathbf{A} \mathbf{V}_N \mathbf{A}^t)^{-1} \mathbf{A} (\mathbf{X}_\theta - \boldsymbol{\mu}),\end{aligned}$$

temos

$$\dot{Q}_{\theta,N} \xrightarrow[\theta \rightarrow +\infty]{F} \dot{Q} \sim \chi_{\dot{r}, \dot{\delta}}^2$$

com

$$\dot{r} = \text{car}(\mathbf{A} \mathbf{V}_N \mathbf{A}^t)$$

e

$$\dot{\delta} = (\boldsymbol{\mu} - \mathbf{b})^t \mathbf{A}^t (\mathbf{A} \mathbf{V}_N \mathbf{A}^t)^{-1} \mathbf{A} (\boldsymbol{\mu} - \mathbf{b}).$$

3.5 Testes e Regiões de Confiança

A distribuição limite, quando $\theta \rightarrow +\infty$, de

$$Q_{\theta,N}(\mathbf{A}, \mathbf{b}) = (\mathbf{X}_\theta - \mathbf{b})^t \mathbf{A}^t (\mathbf{A} \mathbf{V}_N \mathbf{A}^t)^{-1} \mathbf{A} (\mathbf{X}_\theta - \mathbf{b})$$

é, como vimos, $\chi_{\dot{r}, \dot{\delta}}^2$ pelo que podemos usar essa estatística para testar

$$H_0(\mathbf{A}, \mathbf{b}) : \mathbf{A} \boldsymbol{\mu} = \mathbf{A} \mathbf{b}$$

rejeitando esta hipótese quando

$$Q_{\theta,N}(\mathbf{A}, \mathbf{b}) > x_{\dot{r}, 1-q},$$

com $x_{\dot{r}, 1-q}$ o $(1 - q)$ -ésimo quantil para $\chi_{\dot{r}}^2 (= \chi_{\dot{r}, 0}^2)$. O nível limite para este teste será q . Além disso, temos a variável pivot

$$\dot{Q}_{\theta,N}(\mathbf{A}) = (\mathbf{X}_\theta - \boldsymbol{\mu})^t \mathbf{A}^t (\mathbf{A} \mathbf{V}_N \mathbf{A}^t)^{-1} \mathbf{A} (\mathbf{X}_\theta - \boldsymbol{\mu}),$$

com distribuição limite $\chi_{\hat{r}}^2$, quando $\boldsymbol{\theta} \rightarrow +\infty$. Assim o elipsoide de confiança, de nível limite $1 - q$, para $\boldsymbol{\psi} = \mathbf{A}\boldsymbol{\mu}$ é dado por

$$\dot{Q}_{\theta, N}(\mathbf{A}) \leq x_{\hat{r}, 1-q}.$$

Assim, o teste com nível limite q rejeita $H_0(\mathbf{A}, \mathbf{b})$ se e só se este elipsoide de confiança não contiver $\boldsymbol{\psi}_0 = \mathbf{A}\mathbf{b}$. Logo o teste goza, no limite, de dualidade. Por outro lado $\hat{\delta} = 0$ apenas quando $H_0(\mathbf{A}, \mathbf{b})$ é válida sendo, o teste, no limite, estritamente não-enviesado. Isto é o poder limite, para qualquer alternativa, excederá q .

Qualquer que seja o vector não-nulo \mathbf{d} com a mesma dimensão que $\boldsymbol{\psi}$, existem, (ver Scheffé (1959)), dois planos ortogonais a \mathbf{d} tangentes ao elipsoide de confiança, $\boldsymbol{\psi}$ estando entre esses planos se e só se

$$|\mathbf{d}^t \boldsymbol{\psi} - \mathbf{d}^t \tilde{\boldsymbol{\psi}}| \leq \sqrt{x_{\hat{r}, 1-q} \mathbf{d}^t (\mathbf{A} \mathbf{V} \mathbf{A}^t) \mathbf{d}},$$

onde $\tilde{\boldsymbol{\psi}} = \mathbf{A} \mathbf{X}_{\boldsymbol{\theta}}$.

Ora, um ponto pertence a um elipsoide se e só se estiver entre todos os pares de planos paralelos tangente ao elipsoide obtendo-se assim a versão do teorema de Scheffé dada por

$$pr \left[\bigcap_{\mathbf{d}} \left(|\mathbf{d}^t \boldsymbol{\psi} - \mathbf{d}^t \tilde{\boldsymbol{\psi}}| \leq \sqrt{r x_{\hat{r}, 1-q} \mathbf{d}^t (\mathbf{A} \mathbf{V} \mathbf{A}^t) \mathbf{d}} \right) \right] \xrightarrow{\boldsymbol{\theta} \rightarrow +\infty} 1 - q$$

onde $\bigcap_{\mathbf{d}}$ indica que todos os possíveis vectores \mathbf{d} são considerados. Logo teremos intervalos de confiança simultâneos para os $\mathbf{d}^t \boldsymbol{\psi}$. O nível de confiança conjunto limite para estes intervalos de confiança será $1 - q$.

Suponhamos agora que $\mathbf{X} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{V})$, com \mathbf{V} definida positiva, independente de $\mathbf{S} \sim \sigma^2 \chi_g^2$, então, (ver por exemplo Mexia (1990)), temos

$$(\mathbf{X} - \boldsymbol{\mu})^t \mathbf{V}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \sigma^2 \chi_r^2$$

com $r = \text{car}(\mathbf{V})$, independente de \mathbf{S} , logo

$$\mathbf{F} = \frac{g}{r} \frac{(\mathbf{X} - \boldsymbol{\mu})^t \mathbf{V}^{-1} (\mathbf{X} - \boldsymbol{\mu})}{\mathbf{S}}$$

tem distribuição F central com r e g graus de liberdade.

3.6 Linearidade Assintótica

Vamos agora considerar uma abordagem recente, (ver Mexia & Oliveira (2010)), em que se procura obter distribuições limite ligadas ao crescimento da não centralidade e não da dimensão da amostra. Aliás, o ênfase está na expressão analítica da estatística para a qual se pretende uma distribuição limite.

Dado $\mathbf{x} \in \mathbb{R}^k$ a função $g(\mathbf{x})$ terá, caso estejam definidas, gradiente $\underline{g}(\mathbf{x})$ e matriz Hessiana $\underline{\underline{g}}(\mathbf{x})$ com raio espectral $r(\mathbf{x})$. Com

$$\rho_d(\mathbf{x}) = \text{Sup}\{r(\mathbf{x} + \mathbf{u}); \|\mathbf{u}\| \leq d\}$$

a função $g(\cdot)$ será assintoticamente linear se, $\forall d > 0$

$$u_d(h) = \text{Max} \left\{ \frac{\rho_d(\mathbf{x})}{\|\underline{g}(\mathbf{x})\|}; \|\mathbf{x}\| > h \right\} \xrightarrow{h \rightarrow +\infty} 0.$$

Então, qualquer que seja a distribuição $F_{\mathbf{w}}$ de \mathbf{w} , com

$$\mathbf{Z} = \frac{g(\boldsymbol{\mu} + \mathbf{w}) - g(\boldsymbol{\mu})}{\|\underline{g}(\boldsymbol{\mu})\|}$$

$$\mathbf{Z}^0 = \frac{\underline{g}(\boldsymbol{\mu})^t \mathbf{w}}{\|\underline{g}(\boldsymbol{\mu})\|}$$

tem-se, (ver Mexia & Oliveira (2010)),

$$\text{Sup} \{|F_{\mathbf{Z}}(y) - F_{\mathbf{Z}^0}(y)|\} \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow +\infty} 0.$$

Assim se F^0 for a distribuição limite de \mathbf{Z}^0 , quando $\|\boldsymbol{\mu}\| \rightarrow \infty$, também F será a distribuição limite de \mathbf{Z} .

Tomemos

$$g(\mathbf{x}) = \|\mathbf{x}\|^2$$

vindo

$$\underline{g}(\mathbf{x}) = 2\mathbf{x},$$

$$\underline{\underline{g}}(\mathbf{x}) = 2\mathbf{I}_k$$

e, caso $\mathbf{Y} \sim \chi_{k,\delta}^2$, podemos sempre admitir que

$$\mathbf{Y} = g(\boldsymbol{\mu} + \mathbf{w}),$$

e

$$\delta = g(\boldsymbol{\mu})$$

com $\mathbf{w} \sim N(\mathbf{0}, \mathbf{I}_k)$.

Teremos então

$$\mathbf{Z}^0 = \frac{\boldsymbol{\mu}^t \mathbf{w}}{\|\boldsymbol{\mu}\|} \sim N(0, 1)$$

isto é, qualquer que seja $\boldsymbol{\mu}$, $F_{\mathbf{Z}^0}$ é a distribuição normal padronizada.

Assim a distribuição limite de

$$\mathbf{Z} = \frac{\mathbf{Y} - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|}$$

será a normal reduzida. Ao escrevermos a expressão de \mathbf{Z} utilizamos o facto de se ter $\delta = \|\boldsymbol{\mu}\|^2$. Sendo \dot{z}_{1-q} o valor que numa densidade normal reduzida delimita duas caudas com probabilidade $\frac{q}{2}$, temos com $\|\boldsymbol{\mu}\|$ suficientemente grande

$$pr \left(-\dot{z}_{1-q} \leq \frac{y - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|} \leq \dot{z}_{1-q} \right) = 1 - q,$$

o que com $\theta = \|\boldsymbol{\mu}\|^2$ pode ser escrito como

$$pr \left(\theta^2 - 2(y + 2\dot{z}_{1-q}^2)\theta + y^2 \leq 0 \right) = 1 - q$$

e como

$$pr \left(y + 2\dot{z}_{1-q}^2 - \sqrt{4\dot{z}_{1-q}^2 y + 4\dot{z}_{1-q}^4} \leq \theta \leq y + 2\dot{z}_{1-q}^2 + \sqrt{4\dot{z}_{1-q}^2 y + 4\dot{z}_{1-q}^4} \right) = 1 - q$$

o que permite escrever

$$pr \left(1 + \frac{2\dot{z}_{1-q}^2}{y} - 2\dot{z}_{1-q} \sqrt{\frac{1}{y} + \frac{\dot{z}_{1-q}^2}{y^2}} \leq \frac{\theta}{y} \leq 1 + \frac{2\dot{z}_{1-q}^2}{y} + 2\dot{z}_{1-q} \sqrt{\frac{1}{y} + \frac{\dot{z}_{1-q}^2}{y^2}} \right) = 1 - q.$$

Assim quando se utiliza \mathbf{y} como estimador de $\boldsymbol{\theta}$ o erro relativo é pequeno caso \mathbf{y} seja grande.

Analogamente ao utilizar-se $\dot{\mathbf{y}} = \frac{\mathbf{y}}{\mathbf{g}}$ para estimar $\dot{\boldsymbol{\theta}} = \frac{\boldsymbol{\theta}}{\mathbf{g}}$ o erro relativo, que é o mesmo que se tinha ao utilizar \mathbf{y} para estimar $\boldsymbol{\theta}$, será pequeno se \mathbf{y} for grande.

Observe-se ainda que esse erro relativo é da ordem $2\dot{z}_{1-q}y^{-\frac{1}{2}}$.

Suponhamos agora que $y_i \sim \chi_{g_i, \delta_i}^2$, $l = 1, \dots, m$, sendo os qui-quadrados independentes. Raciocinando-se como atrás obtém-se

$$pr \left(a_i \leq \frac{\delta_i}{y_i} \leq b_i \right) = 1 - q_i, \quad i = 1, \dots, m$$

com

$$a_i = 1 + \frac{2\dot{z}_{1-q_i}^2}{y_i} - 2\dot{z}_{1-q_i} \sqrt{\frac{1}{y_i} + \frac{\dot{z}_{1-q_i}^2}{y_i^2}}, \quad i = 1, \dots, m$$

$$b_i = 1 + \frac{2\dot{z}_{1-q_i}^2}{y_i} + 2\dot{z}_{1-q_i} \sqrt{\frac{1}{y_i} + \frac{\dot{z}_{1-q_i}^2}{y_i^2}}, \quad i = 1, \dots, m,$$

logo, atendendo à independência de y_1, \dots, y_m ,

$$pr \left[\bigcap_{i=1}^m \left(a_i \leq \frac{\delta_i}{y_i} \leq b_i \right) \right] = \prod_{i=1}^m (1 - q_i).$$

Quando os q_1, \dots, q_m são pequenos tem-se

$$\prod_{i=1}^m (1 - q_i) \approx 1 - \sum_{i=1}^m q_i.$$

Por outro lado, se os erros relativos resultantes de se utilizar $\dot{y}_i = \frac{y_i}{g_i}$ para estimar $\dot{\delta}_i = \frac{\delta_i}{g_i}$, $i = 1, \dots, m$ forem pequenos, pode-se utilizar

$$\tilde{\dot{\nu}} = \frac{\dot{y}_i}{\sum_{i=1}^m \dot{y}_m}$$

para estimar

$$\dot{\nu} = \frac{\dot{\delta}_i}{\sum_{i=1}^m \dot{\delta}_i}$$

medindo estes últimos as relevâncias relativas.

Capítulo 4

Modelos Colectivos Multivariados

4.1 Modelos Isolados

Dados \mathbf{X}_1, \dots i.i.d. com vector médio $\underline{\mu}$ e matriz de segundos momentos

$$\underline{\underline{\mu}} = [E(X_i X_l^t)],$$

onde X_i e X_l são componentes de \mathbf{X} com a mesma distribuição dos \mathbf{X}_1, \dots ,
pondo-se

$$\mathbf{S} = \sum_{i=1}^N \mathbf{X}_i$$

teremos

$$\begin{aligned}
\varphi_{\mathbf{S}}(\mathbf{u}) &= E(e^{\mathbf{u}^t \mathbf{S}}) = \sum_{n=0}^{\infty} pr((N = n)E(e^{\mathbf{u}^t \mathbf{S}}|N = n)) \\
&= \sum_{n=0}^{\infty} pr(N = n)E\left(e^{\sum_{i=1}^n \mathbf{u}^t \mathbf{X}_i}\right) \\
&= \sum_{n=0}^{\infty} pr(N = n)\varphi_{\mathbf{X}}^n(\mathbf{u}) \\
&= \sum_{n=0}^{\infty} pr(N = n)e^{n\psi_{\mathbf{X}}(\mathbf{u})} \\
&= \varphi_N(\psi_{\mathbf{X}}(\mathbf{u})),
\end{aligned}$$

já que

$$\varphi_N(\mathbf{v}) = \sum_{n=0}^{\infty} pr(N = n)e^{n\mathbf{v}}.$$

Assim

$$\underline{grad}\varphi_{\mathbf{S}}(\mathbf{u}) = \varphi'_N(\psi_{\mathbf{X}}(\mathbf{u}))\underline{grad}\psi_{\mathbf{X}}(\mathbf{u}),$$

e dado que $\psi_{\mathbf{X}}(\mathbf{0}) = 0$, com $\boldsymbol{\nu}$ e $\boldsymbol{\mu}$ os vectores médios de \mathbf{S} e \mathbf{X} , respectivamente temos

$$\boldsymbol{\nu} = \mu(N)\boldsymbol{\mu}.$$

Teremos também

$$\begin{aligned}
\underline{Hess}\psi_{\mathbf{S}}(\mathbf{u}) &= \varphi''_N(\psi_{\mathbf{X}}(\mathbf{u}))\underline{grad}\psi_{\mathbf{X}}(\mathbf{u})\underline{grad}\psi_{\mathbf{X}}(\mathbf{u})^t + \\
&+ \varphi'_N(\psi_{\mathbf{X}}(\mathbf{u}))\underline{Hess}\psi_{\mathbf{X}}(\mathbf{u}),
\end{aligned}$$

bem como

$$\underline{\underline{\mu}}(\mathbf{S}) = \mu'_2(N)\boldsymbol{\mu}\boldsymbol{\mu}^t + \mu(N)\underline{\underline{\mu}}(\mathbf{X}),$$

sendo $\underline{\underline{\mu}}(\mathbf{S})$ e $\underline{\underline{\mu}}(\mathbf{X})$ as matrizes de segundos momentos de \mathbf{S} e \mathbf{X} e $\mu'_2(N)$ o momento de segunda ordem de N .

Para além disso

$$\psi_{\mathbf{S}}(\mathbf{u}) = \psi_N(\psi_{\mathbf{X}}(\mathbf{u})),$$

assim, com $\mathbf{\Sigma}(\mathbf{S})$ e $\mathbf{\Sigma}$ as matrizes de covariância de \mathbf{S} e \mathbf{X} , respectivamente, e $\sigma^2(N)$ a variância de N , tem-se

$$\boldsymbol{\nu} = \mu(N)\boldsymbol{\mu}$$

e

$$\mathbf{\Sigma}(\mathbf{S}) = \sigma^2(N)\boldsymbol{\mu}\boldsymbol{\mu}^t + \mu(N)\mathbf{\Sigma}$$

Quando a distribuição de N_θ é reprodutiva para θ tem-se

$$\mu(N_\theta) = \theta\dot{\mu}$$

$$\sigma^2(N_\theta) = \theta\dot{\sigma}^2,$$

com

$$\dot{\mu} = \mu(N_\theta)$$

$$\dot{\sigma}^2 = \sigma^2(N_\theta),$$

vindo

$$\boldsymbol{\nu}(\theta) = \theta\dot{\mu}\boldsymbol{\mu}$$

e

$$\mathbf{\Sigma}(\mathbf{S}_\theta) = \theta (\sigma^2(N_1)\boldsymbol{\mu}\boldsymbol{\mu}^t + \mu(N_1)\mathbf{\Sigma}).$$

Assim, com k o número de componentes de \mathbf{X}_1, \dots , caso $\frac{N}{\theta} \xrightarrow{\theta \rightarrow \infty} 1$

$$\mathbf{Z}_\theta = \frac{1}{\sqrt{N}}(\mathbf{S}_\theta - \boldsymbol{\nu}(\theta)),$$

terá vector médio nulo e matriz de covariância \mathbf{I}_k . Este número de componentes será a dimensão do modelo. Por outro lado utilizando o desenvolvimento de Taylor na origem no cálculo de $\psi_{\mathbf{Z}_\theta}(\mathbf{u})$ obtém-se

$$\begin{aligned}\psi_{\mathbf{Z}_\theta}(\mathbf{u}) &= -\theta^{\frac{1}{2}}\dot{\mu}\boldsymbol{\mu}^t\mathbf{u} + \psi_{\mathbf{S}_\theta}(\theta^{-\frac{1}{2}}\mathbf{u}) \\ &= -\theta^{\frac{1}{2}}\dot{\mu}\boldsymbol{\mu}^t\mathbf{u} + \theta^{\frac{1}{2}}\dot{\mu}\boldsymbol{\mu}^t\mathbf{u} + \frac{1}{2}\mathbf{u}^t(\dot{\sigma}^2\boldsymbol{\mu}\boldsymbol{\mu}^t + \dot{\mu}\boldsymbol{\Sigma}) + R_2\left(\theta^{-\frac{1}{2}}\mathbf{u}\right)\end{aligned}$$

com $R_2(\cdot)$ o resíduo da expansão de Lagrange de segunda ordem de $\psi_{\mathbf{Z}_\theta}(\mathbf{u})$.

Dado ter-se

$$R_2(\theta^{-\frac{1}{2}}\mathbf{u}) \xrightarrow{\theta \rightarrow \infty} 0,$$

temos

$$\mathbf{Z}_\theta \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{Z} \sim N(\cdot | \mathbf{0}, \dot{\sigma}^2\boldsymbol{\mu}\boldsymbol{\mu}^t + \dot{\mu}\boldsymbol{\Sigma}).$$

Sendo

$$\mathbf{Z}_\theta(\mathbf{b}) = \frac{1}{\sqrt{N}}(\mathbf{S}_\theta - \mathbf{b}),$$

se

$$\theta^{\frac{1}{2}}(\mathbf{b}_\theta - \boldsymbol{\nu}) \xrightarrow{\theta \rightarrow \infty} \boldsymbol{\eta},$$

podemos, de forma análoga, mostrar que

$$\mathbf{Z}_\theta(\mathbf{b}_\theta) \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{Z}(\boldsymbol{\eta}) \sim N(\cdot | \boldsymbol{\eta}, \dot{\sigma}^2\boldsymbol{\mu}\boldsymbol{\mu}^t + \dot{\mu}\boldsymbol{\Sigma}).$$

Nomeadamente quando N_θ tem distribuição de Poisson com parâmetro $\lambda(= \theta)$, temos $\dot{\mu} = \dot{\sigma}^2 = \lambda$ e a matriz de covariância \mathbf{V} para a distribuição limite será $\underline{\underline{\mu}}$. Neste caso, dada a amostra $\mathbf{x}_1, \dots, \mathbf{x}_n$, temos o estimador

$$\tilde{\mathbf{V}}_n = \underline{\underline{\tilde{\mu}}}_n = \frac{1}{n} \sum_{l=1}^n \mathbf{x}_l \mathbf{x}_l^t.$$

De um modo geral temos

$$\tilde{\boldsymbol{\mu}}_n = \frac{1}{n} \sum_{l=1}^n \mathbf{x}_l$$

e

$$\tilde{\boldsymbol{\Sigma}}_n = \tilde{\boldsymbol{\mu}}_n - \tilde{\boldsymbol{\mu}}_n \tilde{\boldsymbol{\mu}}_n^t,$$

obtendo-se o estimador

$$\tilde{\mathbf{V}}_n = \hat{\sigma}^2 \tilde{\boldsymbol{\mu}}_n \tilde{\boldsymbol{\mu}}_n^t + \hat{\mu} \tilde{\boldsymbol{\Sigma}}_n,$$

assumindo que os parâmetros $\hat{\mu}$ e $\hat{\sigma}^2$ são conhecidos. Para realizarmos inferência sobre o modelo, tomamos $\mathbf{V}_N = \tilde{\mathbf{V}}_n$.

Considerando agora o caso particular, em que

$$\mathbf{S}_\lambda = \begin{bmatrix} \mathbf{N} \\ \mathbf{T} \end{bmatrix}$$

e

$$\mathbf{T} = \sum_{i=1}^n X_i,$$

temos então

$$\tilde{\mathbf{V}}_n = \frac{1}{N} \begin{bmatrix} \mathbf{N} & \mathbf{T} \\ \mathbf{T} & \sum_{i=1}^n X_i^2 \end{bmatrix}$$

e, assim

$$\tilde{\mathbf{V}}_n^{-1} = \frac{\mathbf{N}}{\mathbf{N} \sum_{i=1}^n X_i^2 - T^2} \begin{bmatrix} \sum_{i=1}^n X_i^2 & -\mathbf{T} \\ -\mathbf{T} & \mathbf{N} \end{bmatrix}.$$

Seja

$$\begin{aligned} \mathbf{z} &= \frac{1}{\sqrt{\lambda}} (\mathbf{S}_\lambda - \boldsymbol{\nu}(\lambda)) = \frac{1}{\sqrt{\mathbf{N}}} \left(\begin{bmatrix} \mathbf{N} \\ \mathbf{T} \end{bmatrix} - \begin{bmatrix} \lambda \\ \lambda \mu \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{\mathbf{N}}} \begin{bmatrix} \mathbf{N} - \lambda \\ \mathbf{T} - \lambda \mu \end{bmatrix}, \end{aligned}$$

tendo-se a forma quadrática

$$Q = \mathbf{Z}^t \tilde{\mathbf{V}}_n^{-1} \mathbf{Z} = \frac{\mathbf{N}}{\mathbf{N} \sum_{i=1}^n X_i^2 - \mathbf{T}^2} \left[\sum_{i=1}^n X_i^2 (\boldsymbol{\lambda} - \mathbf{N})^2 - 2\mathbf{T}(\boldsymbol{\lambda} - \mathbf{N})(\boldsymbol{\lambda}\boldsymbol{\mu} - \mathbf{T}) + \mathbf{N}(\boldsymbol{\lambda}\boldsymbol{\mu} - \mathbf{T})^2 \right].$$

De acordo com o Teorema de Sheffé temos

$$pr \left[\bigcap_{\mathbf{d}} \left(|d_1 \mathbf{N} + d_2 \mathbf{T} - (d_1 \boldsymbol{\lambda} + d_2 \boldsymbol{\lambda}\boldsymbol{\mu})| \leq \sqrt{2x_{1-p,2} \mathbf{d}^t \mathbf{V} \mathbf{d}} \right) \right] = 1 - p$$

obtendo-se assim a elipse de confiança de nível $1 - p$ centrada no ponto (N, T) .

Desta forma obteremos os intervalos de confiança simultâneos, com nível de confiança $1 - p$, para os parâmetros $\boldsymbol{\lambda}$ e $\boldsymbol{\lambda}\boldsymbol{\mu}$, respectivamente

$$IC(\boldsymbol{\lambda}) = \left[\mathbf{N} - x_{1,1-p} \sqrt{\frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i^2 - \frac{\mathbf{T}^2}{\mathbf{N}}}}; \mathbf{N} + x_{1,1-p} \sqrt{\frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i^2 - \frac{\mathbf{T}^2}{\mathbf{N}}}} \right]$$

e

$$IC(\boldsymbol{\lambda}\boldsymbol{\mu}) = \left[\mathbf{T} - x_{1,1-p} \sqrt{\frac{\mathbf{N}}{\sum_{i=1}^n X_i^2 - \frac{\mathbf{T}^2}{\mathbf{N}}}}; \mathbf{T} + x_{1,1-p} \sqrt{\frac{\mathbf{N}}{\sum_{i=1}^n X_i^2 - \frac{\mathbf{T}^2}{\mathbf{N}}}} \right].$$

No Apêndice C encontram-se estes intervalos de confiança calculados para o caso particular referido atrás.

Até agora temos centrado o nosso estudo em \mathbf{S}_θ realizando-se inferência em $\boldsymbol{\nu}$. Esta é a abordagem clássica da Teoria de Risco na qual o total de *claims* desempenha um papel central. No entanto, podemos estar interessados na inferência sobre $\boldsymbol{\mu}$. Assim, tomando

$$\mathbf{M}_\theta = \frac{1}{N_\theta} \mathbf{S}_\theta$$

temos, como anteriormente,

$$\mathbf{W}_\theta = \frac{N_\theta}{\sqrt{\theta}}(\mathbf{M}_\theta - \boldsymbol{\mu}) \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{W} \sim N(\cdot | \mathbf{0}, \dot{\boldsymbol{\mu}} \boldsymbol{\Sigma}).$$

Mais geralmente, se

$$\sqrt{\theta}(\mathbf{a}_\theta - \boldsymbol{\mu}) \xrightarrow[\theta \rightarrow \infty]{} \boldsymbol{\lambda},$$

teremos

$$\mathbf{W}_\theta(\mathbf{a}_\theta) = \frac{N_\theta}{\sqrt{\theta}}(\mathbf{M}_\theta - \mathbf{a}_\theta) \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{W}(\boldsymbol{\lambda}) \sim N(\cdot | \boldsymbol{\lambda}, \dot{\boldsymbol{\mu}} \boldsymbol{\Sigma}).$$

No caso particular em que

$$\mathbf{M} = \frac{1}{\bar{N}} \begin{bmatrix} \mathbf{N} \\ \mathbf{T} \end{bmatrix} = \begin{bmatrix} 1 \\ X. \end{bmatrix},$$

teremos apenas que considerar intervalos de confiança para $E[X.]$.

4.2 Famílias de Modelos

4.2.1 Famílias Estruturadas

Consideremos uma família de L modelos colectivos de dimensão k com pares aleatórios independentes $(N_{l,\theta_l}, \mathbf{S}_{l,\theta_l})$, $l = 1, \dots, L$. A distribuição de N_{l,θ_l} , $l = 1, \dots, L$ é reprodutiva para θ_l , $l = 1, \dots, L$, e os parâmetros $\dot{\boldsymbol{\mu}}$ e $\dot{\sigma}^2$ são os mesmos para os L modelos. O vector médio e a matriz de covariância de \mathbf{S}_{l,θ_l} serão então

$$\boldsymbol{\nu}_l = \theta_l \dot{\boldsymbol{\mu}} \boldsymbol{\mu}_l$$

e

$$\theta_l (\dot{\sigma}^2 \boldsymbol{\mu}_l \boldsymbol{\mu}_l^t + \dot{\boldsymbol{\mu}} \boldsymbol{\Sigma}_l), \quad l = 1, \dots, L,$$

respectivamente. Assim, com \mathbf{V} e \mathbf{V}_N as matrizes diagonais por blocos com blocos principais \mathbf{V}_l e \mathbf{V}_{l,N_l} , $l = 1, \dots, L$, respectivamente, e tomando

$$\mathbf{Z}_{l,\theta_l} = \frac{1}{\sqrt{N_l}}(\mathbf{S}_{l,\theta_l} - \boldsymbol{\nu}_l), \quad l = 1, \dots, L$$

temos

$$\mathbf{Z}_\theta = [\mathbf{Z}_{1,\theta_1}^t \dots \mathbf{Z}_{L,\theta_L}^t]^t \xrightarrow[\theta \rightarrow \infty]{F} \mathbf{Z} \sim N(\cdot | \mathbf{0}, \mathbf{V}).$$

Em particular, resulta daqui que,

$$\tilde{\nu}_{l,\theta_l} = \frac{S_{l,\theta_l}}{\theta_l} \xrightarrow[\theta_l \rightarrow \infty]{p} \nu_l, \quad l = 1, \dots, L.$$

Analogamente, com

$$\tilde{\boldsymbol{\nu}}_\theta = \begin{bmatrix} \tilde{\nu}_{1,\theta_1} \\ \dots \\ \tilde{\nu}_{L,\theta_L} \end{bmatrix}; \quad \boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \dots \\ \nu_L \end{bmatrix}$$

temos

$$\tilde{\boldsymbol{\nu}}_\theta \xrightarrow[\theta \rightarrow \infty]{p} \boldsymbol{\nu}.$$

A distribuição limite de

$$\mathbf{Q}_\theta = \sum_{l=1}^L \mathbf{Q}_{l,\theta_l},$$

com

$$Q_{l,\theta_l} = \mathbf{Z}_{l,\theta_l}^t V_{l,N_l}^{-1} \mathbf{Z}_{l,\theta_l}, \quad l = 1, \dots, L,$$

será $G(\cdot | Lk)$, distribuição do qui-quadrado central χ_{Lk}^2 com Lk graus de liberdade.

De um modo geral, se

$$\theta_l^{\frac{1}{2}}(\mathbf{b}_{l,\theta_l} - \boldsymbol{\nu}_l) \xrightarrow[\theta_l \rightarrow +\infty]{} \boldsymbol{\eta}_l, \quad l = 1, \dots, L,$$

temos

$$\mathbf{Z}_\theta(\mathbf{b}_\theta) = [\mathbf{Z}_{1,\theta_1}(\mathbf{b}_{1,\theta_1})^t \dots \mathbf{Z}_{L,\theta_L}(\mathbf{b}_{L,\theta_L})^t]^t \xrightarrow[\theta \rightarrow +\infty]{F} \mathbf{Z}(\boldsymbol{\eta}) \sim N(\cdot | \boldsymbol{\eta}, \mathbf{V}).$$

Com

$$Q_{l,\theta_l}(\mathbf{b}_{l,\theta_l}) = \mathbf{Z}_{l,\theta_l}(\mathbf{b}_{l,\theta_l})^t \mathbf{V}_{l,N_l}^{-1} \mathbf{Z}_{l,\theta_l}(\mathbf{b}_{l,\theta_l}), \quad l = 1, \dots, L$$

a distribuição limite de

$$\mathbf{Q}_\theta(\mathbf{b}) = \sum_{l=1}^L \mathbf{Q}_{l,\theta_l}(\mathbf{b}_{l,\theta_l}),$$

será

$$G\left(\cdot | Lk, \sum_{l=1}^L \delta_l(\mathbf{b}_{l,\theta_l})\right),$$

a distribuição de

$$\chi_{Lk, \sum_{l=1}^L \delta_l(\mathbf{b}_{l,\theta_l})}^2.$$

O parâmetro de não-centralidade será

$$\delta_l(\mathbf{b}_{l,\theta_l}) = \boldsymbol{\eta}_l^t \mathbf{V}_l^{-1} \boldsymbol{\eta}_l, \quad l = 1, \dots, L.$$

Por outro lado, as distribuições limite de

$$Q_\theta(\mathbf{A}) = (\mathbf{A}\mathbf{Z}_\theta)^t (\mathbf{A}\mathbf{V}_N \mathbf{A}^t)^{-1} (\mathbf{A}\mathbf{Z}_\theta)$$

e de

$$Q_\theta(\mathbf{A}, \mathbf{b}_\theta) = (\mathbf{A}\mathbf{Z}_\theta(\mathbf{b}))^t (\mathbf{A}\mathbf{V}_N \mathbf{A}^t)^{-1} (\mathbf{A}\mathbf{Z}_\theta(\mathbf{b})),$$

serão $G(\cdot | h)$ e $G(\cdot | h, \delta(\mathbf{A}))$ com

$$\delta(\mathbf{A}) = (\mathbf{A}\boldsymbol{\eta})^t (\mathbf{A}\mathbf{V}\mathbf{A}^t)^{-1} (\mathbf{A}\boldsymbol{\eta}),$$

com \mathbf{A} conhecida com vectores linha linearmente independentes.

Lembremos que os testes de qui-quadrado, no limite, gozam de dualidade. Assim o teste de nível limite p para

$$H_0(\mathbf{A}, \mathbf{b}) : \mathbf{A}\boldsymbol{\nu} = \mathbf{A}\mathbf{b},$$

não a rejeita quando e só quando $\tilde{\boldsymbol{\nu}}_{\boldsymbol{\theta}}$ estiver contido no elipsoide de confiança de nível limite $1 - p$ para $\mathbf{A}\boldsymbol{\nu}$.

Ora

$$\mathbf{A}\tilde{\boldsymbol{\nu}}_{\boldsymbol{\theta}} \xrightarrow[\boldsymbol{\theta} \rightarrow +\infty]{p} \mathbf{A}\boldsymbol{\nu}$$

pelo que, quaisquer que seja a alternativa e o nível de confiança $1 - q$, existirá $\dot{\boldsymbol{\theta}}_{1-q}$ tal que para $\boldsymbol{\theta} > \dot{\boldsymbol{\theta}}_{1-q}$, $\mathbf{A}\mathbf{b}$ não se encontrará no elipsoide de confiança e $H_0(\mathbf{A}, \mathbf{b})$ será rejeitada. Assim os testes de qui-quadrado serão fortemente consistentes, (ver por exemplo T. de Oliveira (1980)).

Para poder ter uma potência de teste no limite menor do que um teremos de considerar sucessivas hipóteses alternativas da forma

$$H_1(\mathbf{A}, \mathbf{b}_{\boldsymbol{\theta}}) : \mathbf{A}\boldsymbol{\nu} = \mathbf{A}\mathbf{b}_{\boldsymbol{\theta}}$$

tal que, com

$$\mathbf{b}_{\boldsymbol{\theta}} = [\mathbf{b}_{1,\boldsymbol{\theta}_1}^t, \dots, \mathbf{b}_{L,\boldsymbol{\theta}_L}^t]^t$$

e

$$\mathbf{b} = [\mathbf{b}_1^t, \dots, \mathbf{b}_L^t]^t$$

tivermos

$$\sqrt{\theta_l}(\mathbf{b}_{l,\boldsymbol{\theta}_l} - \mathbf{b}_l) \xrightarrow[\theta_l \rightarrow +\infty]{} \boldsymbol{\eta}_l, \quad l = 1, \dots, L,$$

pelo que a distribuição limite de $Q_{\boldsymbol{\theta}}(\mathbf{A}, \mathbf{b}_{\boldsymbol{\theta}})$ será $G(\cdot|h, \delta(\mathbf{A}))$. Assim, com $x_{h,1-q}$ o $(1 - q)$ -ésimo quantil de $G(\cdot|h)$, a potência limite será $1 - G(x_{h,1-q}|h, \delta(\mathbf{A}))$.

Quando os modelos correspondem aos tratamentos dum modelo base, a família será estruturada. Podemos então estudar a acção dos factores no modelo base, nas combinações lineares

$$\mathbf{c}^t \boldsymbol{\nu}_l = \mu_l \theta_l \boldsymbol{\mu}_l$$

das componentes dos vectores médios dos \mathbf{S}_{l, θ_l} , $l = 1, \dots, L$, tomando

$$\mathbf{A} = \dot{\mathbf{A}} \otimes \mathbf{c}^t$$

onde \otimes representa o produto de Kroneckere e a matriz $\dot{\mathbf{A}}$ é referente ao modelo base.

Em particular, se tomarmos $\mathbf{c} = \boldsymbol{\delta}_j$, sendo nulas todas as componentes de $\boldsymbol{\delta}_j$ exceptuando a j -ésima que é 1, estudando a acção dos factores do delineamento base nas j -ésimas componentes dos $\boldsymbol{\nu}_l$, $l = 1, \dots, L$.

Tendo-se um modelo base de efeitos fixos com vector médio $\boldsymbol{\vartheta}$, as hipóteses de ausência de efeitos e/ou interacção para os factores do modelo base podem ser escritas como

$$\dot{H}_0 : \dot{\mathbf{A}} \boldsymbol{\vartheta} = \mathbf{0}.$$

Então as hipóteses correspondentes para as famílias estruturadas de modelos colectivos serão da forma

$$H_0 : (\dot{\mathbf{A}} \otimes \mathbf{c}, \mathbf{0}).$$

Por exemplo, se no modelo base tivermos J factores, com a_1, \dots, a_J níveis que cruzam, e $C \subseteq \{1, \dots, J\}$, $\dot{H}_0(C)$ será a hipótese de ausência de efeitos para o único factor com índice em C se $\#C = 1$ ou de ausência de interacções entre os factores com índices em C se $\#C > 1$. Então, (ver Fonseca et al. (2003))

$$\dot{\mathbf{A}}(C) = \otimes_{l=1}^J \mathbf{A}_l(C); \quad C \subseteq \{1, \dots, J\}$$

com

$$\mathbf{A}_l(C) = \frac{1}{\sqrt{a_l}} \mathbf{1}_{a_l}^t; \quad l \notin C, \quad C \subseteq \{1, \dots, J\}$$

e

$$\mathbf{A}_l(C) = \mathbf{T}_{a_l}; \quad l \in C, \quad C \subseteq \{1, \dots, J\}$$

onde \mathbf{T}_{a_l} é a matriz obtida pela eliminação da primeira linha igual a $\frac{1}{\sqrt{a_l}} \mathbf{1}_{a_l}^t$ duma matriz ortogonal $a_l \times a_l$, $l = 1, \dots, J$.

Se pretendermos realizar inferência para $\boldsymbol{\mu} = [\boldsymbol{\mu}_1^t \dots \boldsymbol{\mu}_L^t]^t$ temos apenas de substituir \mathbf{Z}_θ por

$$\mathbf{W}_\theta = [\mathbf{W}_{1,\theta_1}^t \dots \mathbf{W}_{L,\theta_L}^t]^t,$$

com

$$\mathbf{W}_{L,\theta_L} = \sqrt{N_{l,\theta_l}} (\mathbf{M}_{l,\theta_l} - \boldsymbol{\mu}_l), \quad l = 1, \dots, L$$

onde $\mathbf{M}_{l,\theta_l} = \frac{1}{N_{l,\theta_l}} \mathbf{S}_{l,\theta_l}$, $l = 1, \dots, L$. Teremos ainda de substituir \mathbf{V}_l por $\boldsymbol{\Sigma}_l$ e $\mathbf{V}_{l,N_l} = \tilde{\mathbf{V}}_{l,n_l}$ por $\boldsymbol{\Sigma}_{l,N_l} = \tilde{\boldsymbol{\Sigma}}_{l,n_l}$, $l = 1, \dots, L$. As matrizes \mathbf{A} serão as mesmas de anteriormente.

Na inferência, quer para $\boldsymbol{\nu}$ quer para $\boldsymbol{\mu}$, podemos considerar testes selectivos em que se privilegiam algumas hipóteses alternativas.

Na construção desses testes utilizamos coordenadas polares generalizadas, encontrando-se no Apêndice A os resultados mais relevantes sobre essas coordenadas. As coordenadas polares generalizadas $r, \theta_1, \dots, \theta_{k-1}$ do ponto com coordenadas cartesianas x_1, \dots, x_k satisfazem as relações

$$x_1 = r c_1 \dots c_{k-1} = r l_1(\boldsymbol{\theta})$$

$$x_2 = r c_1 \dots s_{k-1} = r l_2(\boldsymbol{\theta})$$

$$\begin{aligned} & \vdots \\ x_i &= r c_1 \dots s_{k+1-i} = r l_i(\boldsymbol{\theta}) \\ & \vdots \\ x_k &= r s_1 = r l_k(\boldsymbol{\theta}) \end{aligned}$$

com $r = \|\mathbf{x}\|$, $c_i = \cos(\theta_i)$ e $s_i = \sin(\theta_i)$, $i = 1, \dots, k$. Têm-se os limites de variação $r \geq 0$,

$$-\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}, \quad i = 1, \dots, k-1$$

e

$$0 \leq \theta_k \leq 2\pi$$

definindo estes o domínio \mathcal{D} de variação do vector de ângulos ao centro Θ .

Para evitar duplicação vamos considerar testes para hipóteses relativas a $\boldsymbol{\nu}$. Admitamos poder trabalhar com uma distribuição limite $N(\boldsymbol{\nu}, \mathbf{V})$ para \mathbf{Z}_θ pretendendo-se privilegiar, relativamente a $H_0(\mathbf{A}, \mathbf{0})$ hipóteses em que

$$\boldsymbol{\theta}(\mathbf{A}\boldsymbol{\nu}) \in \dot{\mathcal{D}}$$

com $\dot{\mathcal{D}}$ um sub-domínio de \mathcal{D} . Então, a distribuição limite de \mathbf{AZ}_θ será $N(\mathbf{A}\boldsymbol{\nu}, \dot{\mathbf{V}})$ com $\dot{\mathbf{V}} = \mathbf{A}\mathbf{V}\mathbf{A}^t$.

O Corolário 2 da Proposição A.1 do Apêndice A diz-nos que, quando $H_0(\mathbf{A}, \mathbf{0})$ se verifica

- o vector Θ dos ângulos ao centro de \mathbf{AZ}_θ tem densidade

$$m_0(\mathbf{u}|\dot{\mathbf{V}}) = \frac{\Gamma\left(\frac{k}{2}\right) c_1^{k-2} \dots c_{k-2}}{2\pi^{k/2} \left(\mathbf{l}(\mathbf{u})^t \dot{\mathbf{V}}^{-1} \mathbf{l}(\mathbf{u})\right) \sqrt{\det(\dot{\mathbf{V}})}}$$

- Θ é independente de $\mathbf{R} = (\mathbf{AZ})^t \dot{\mathbf{V}}^{-1} (\mathbf{AZ})$ que se distribui como um qui-quadrado central com $\dot{r} = \text{car}(\dot{\mathbf{V}})$ graus de liberdade.

Assim pode-se utilizar o par (\mathbf{R}, Θ) para testar $H_0(\mathbf{A}, \mathbf{0})$, rejeitando esta hipótese quando $\Theta \in \dot{\mathcal{D}}$ e $\mathbf{R} > c$, com c um valor crítico. Para determinarmos o nível do teste temos

$$q = \int \dots \int_{\dot{\mathcal{D}}} m_0(\mathbf{u} | \dot{\mathbf{V}}) du_1 \dots du_{k-1} \int_c^{+\infty} g(x | \dot{r}) dx.$$

Por vezes há que completar a análise testando contrastes e obtendo intervalos de confiança simultâneos.

Admitamos que se pode trabalhar com a distribuição limite $N(\boldsymbol{\nu}, \mathbf{V})$ para \mathbf{Z}_θ . Então um contraste $\mathbf{c}^t \mathbf{Z}_\theta$ terá distribuição limite $N(\mathbf{c}^t \boldsymbol{\nu}, \mathbf{c}^t \mathbf{V} \mathbf{c})$ e, no limite

$$\frac{\mathbf{c}^t \mathbf{Z}_\theta - \mathbf{c}^t \boldsymbol{\nu}}{\sqrt{\mathbf{c}^t \mathbf{V} \mathbf{c}}} \sim N(0, 1)$$

obtendo-se os intervalos de confiança de nível $1 - q$ com extremos

$$\mathbf{c}^t \mathbf{Z}_\theta \pm \dot{z}_{1-q} \sqrt{\mathbf{c}^t \mathbf{V} \mathbf{c}}.$$

Quando um valor d_0 não se encontra coberto pelo intervalo de nível $1 - q$ podemos rejeitar

$$H_0(d_0) : \mathbf{c}^t \boldsymbol{\nu} = d_0$$

ao nível (limite) q .

Em alternativa pode utilizar-se o facto de que, quando $H_0(d_0)$ se verifica

$$\frac{(\mathbf{c}^t \mathbf{Z} - d_0)^2}{\mathbf{c}^t \mathbf{V} \mathbf{c}}$$

se distribui, no limite, como um qui-quadrado com 1 grau de liberdade.

4.2.2 Famílias Estruturadas Ortogonais

Admitamos agora que o modelo Π está associado a uma AJCS com base principal

$$bp(\mathcal{A}) = \{\mathbf{Q}_1, \dots, \mathbf{Q}_w\},$$

tendo-se

$$\mathbf{Q}_j = \mathbf{A}_j^t \mathbf{A}_j, \quad j = 1, \dots, w$$

onde \mathbf{A}_j tem vectores linha que constituem uma base ortonormada para

$$\nabla_j = R(\mathbf{Q}_j), \quad j = 1, \dots, w.$$

Então Π estará associado à partição ortogonal

$$R^L = \boxplus_{j=1}^J \nabla_j.$$

Tal como atrás L representará o número de tratamentos e \boxplus indicará a soma ortogonal de sub-espacos.

Agora as hipóteses a testar serão da forma

$$H_0(\mathbf{A}_j, \mathbf{b}) : \mathbf{A}_j \boldsymbol{\nu} = \mathbf{A}_j \mathbf{b}, \quad j = 1, \dots, J,$$

aplicando-se a cada uma delas os resultados da secção anterior.

Em geral estas hipóteses generalizam as $H_0(\mathbf{A}_j, \mathbf{0})$, $j = 1, \dots, J$ que serão hipóteses de ausência de efeitos e ou interacções. Para construir os testes bastanos ter as matrizes \mathbf{A}_j , $j = 1, \dots, J$. Um caso importante é aquele em que o modelo base corresponde ao cruzamento e aninhamento de L grupos com u_1, \dots, u_L factores. Se $u_l > 1$ cada um dos $a_l(1)$ níveis do primeiro factor do grupo l aninha $a_l(2)$ níveis do segundo factor do grupo e assim sucessivamente, $l = 1, \dots, L$.

Admitindo-se que factores que aninham não interagem, os efeitos e interacções que se testam correspondem aos vectores

$$\mathbf{h} \in \Gamma = \{\mathbf{k} : k_l = 0, \dots, u_l, l = 1, \dots, L\}.$$

Se \mathbf{h} tiver uma única componente não nula corresponderá aos efeitos do factor indicado por essa componente. Se \mathbf{h} tiver mais duma componente não nula corresponderá à interacção entre os factores indicados por essas componentes. As matrizes \mathbf{A} são, (ver Fonseca et al (2003)), dadas por

$$\mathbf{A}(\mathbf{h}) = \otimes_{l=1}^L \mathbf{A}_l(h_l), \quad \mathbf{h} \in \Gamma$$

com

$$\mathbf{A}_l(h) = \otimes_{k=1}^{h-1} \mathbf{I}_{a_l(k)} \otimes \mathbf{T}_{a_l(h)} \otimes_{k=h+1}^{u_l} \frac{1}{\sqrt{a_l(k)}} \mathbf{1}_{a_l(k)}^t, \quad h = 0, \dots, u_l, \quad l = 1, \dots, L$$

onde se convencionou ter

$$\begin{aligned} \otimes_{k=1}^0 \mathbf{I}_{a_l(k)} &= \mathbf{I}_1, \quad l = 1, \dots, L \\ \otimes_{k=u_l+1}^{u_l} \frac{1}{\sqrt{a_l(k)}} \mathbf{1}_{a_l(k)}^t &= \mathbf{1}_1^t, \quad l = 1, \dots, L. \end{aligned}$$

Estas matrizes têm características

$$g(\mathbf{h}) = \prod_{l=1}^L g_l(h_l), \quad \mathbf{h} \in \Gamma,$$

com

$$g_l(h) = \prod_{k=1}^{h-1} a_l(a_l(h) - 1), \quad h = 1, \dots, u_l, \quad l = 1, \dots, L,$$

tomando-se

$$\prod_{k=1}^0 a_l = 1, \quad l = 1, \dots, L.$$

Quando há apenas cruzamento de factores estas fórmulas simplificam-se. Temos agora

$$\Gamma = \{\mathbf{k} : k_l = 0, 1; l = 1, \dots, L\}$$

e

$$\begin{cases} \mathbf{A}_l(0) = \frac{1}{\sqrt{a_l}} \mathbf{1}_{a_l}, l = 1, \dots, L \\ \mathbf{A}_l(1) = \mathbf{T}_{a_l}, l = 1, \dots, L, \end{cases},$$

onde $a_l = a_l(1)$, $l = 1, \dots, L$. Tem-se ainda

$$g(\mathbf{h}) = \prod_{l=1}^L g_l(h_l), \mathbf{h} = 1, \dots, L$$

com

$$\begin{cases} g_l(0) = 1, l = 1, \dots, L \\ g_l(1) = a_l - 1, l = 1, \dots, L, \end{cases}.$$

Adiante encontramos um caso em que se tem $L = 2$ com $a_1 = 25$ e $a_2 = 5$,

logo

$$\Gamma = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

obtendo-se

$$\begin{cases} \mathbf{A}_1(0) = \frac{1}{\sqrt{25}} \mathbf{1}_{25}^t \\ \mathbf{A}_1(1) = \mathbf{T}_{25} \\ \mathbf{A}_2(0) = \frac{1}{\sqrt{5}} \mathbf{1}_5^t \\ \mathbf{A}_2(1) = \mathbf{T}_5 \end{cases},$$

Um outro caso muito interessante é o dos factoriais de base p estudado por Jesus et al. (2007 e 2009). Não entraremos no detalhe desse caso por estar fora do âmbito do nosso estudo, limitando-nos a referir a possibilidade de apenas se considerar uma fracção do conjunto dos tratamentos possíveis.

Capítulo 5

Fogos Florestais em Portugal Continental

5.1 Apresentação do problema

A incidência de incêndios florestais tem aumentado nas últimas décadas no Mediterrâneo (ver Pereira et al. (2006) e Pausas, (2008)). Em Portugal, a área ardida atingiu um total de cerca de $3,8 \times 10^6 ha$ no período de 1975 a 2007, correspondendo a cerca de 40% da área do país. Os incêndios têm um impacto substancial na economia, na sociedade e no ambiente. Os gestores dos ecossistemas florestais e os decisores políticos, têm de enfrentar o desafio de desenvolver planos de gestão e políticas que integrem informação relativa à distribuição dos incêndios no espaço e no tempo.

Portugal Continental (Figura 5.1) estende-se por cerca de $89.000 km^2$ situada entre $37^\circ N$ e $42^\circ N$ de latitude e entre $6^\circ W$ e 10° de longitude W. Faixas de altitude do nível do mar até cerca de 2000 metros, com as elevações principais

concentrada no centro e no norte de Portugal.

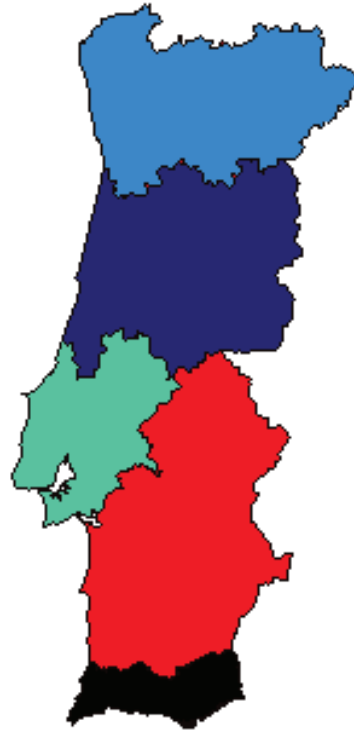


Figura 5.1: Mapa de Portugal Continental

A modelação da área ardida em Portugal foi baseada em informações históricas relativas aos incêndios no período de 33 anos (1975-2007). O mapeamento da área ardida foi obtido para cada ano neste período, pela classificação semi-automática das imagens de dados (Landsat Multi-Spectral Scanner, Landsat Thematic Mapper e Landsat TM + avançado).

Com base nestes dados foi realizado um primeiro trabalho Marques et al. (2011) utilizando modelos lineares generalizados. Utilizaremos agora a abordagem desenvolvida nesta dissertação.

5.2 Inferência e Resultados

Dispomos dos dados referentes a

- **N** - número de fogos,
- **T** - área total ardida.

para os anos de 1975 a 2007, mas em 1983 o problema dos fogos florestais agravou-se muito, pelo que resolvemos apenas trabalhar os dados posteriores, relativos às cinco regiões:

- Norte
- Centro
- Lisboa e Vale do Tejo
- Alentejo
- Algarve

em que Portugal Continental se encontra dividido.

Consideramos pois os factores região e ano com 5 e 25 níveis respectivamente.

Para analisar os dados consideramos três etapas:

- na primeira etapa realizou-se uma análise região a região pretendendo-se avaliar as respectivas evoluções no tempo.
- na segunda etapa trabalhou-se ano a ano, pretendendo-se para cada ano comparar as regiões.

- na terceira etapa estudou-se a acção dos dois factores considerando-se os efeitos dos mesmos e a respectiva interacção.

Estas análises foram efectuadas para o número de fogos (\mathbf{N}) e total da área ardida (\mathbf{T}).

5.2.1 Região a Região

Na primeira etapa o único factor que se considera tem 25 níveis, pelo que se deve trabalhar com uma matriz ortogonal padronizada

$$\mathbf{P}_{25} = \mathbf{P}_5 \otimes \mathbf{P}_5$$

onde, por exemplo se pode ter

$$P_5 = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & -\frac{3}{\sqrt{30}} & \frac{2}{\sqrt{30}} & -\frac{3}{\sqrt{30}} & \frac{2}{\sqrt{30}} \end{bmatrix}$$

Neste caso a matriz \mathbf{A} é formada pelas linhas de \mathbf{P}_{25} com excepção da primeira linha. Na Tabela 5.1 apresentam-se, para cada região, os valores das estatísticas de teste de qui-quadrado para a hipótese de não haver diferenças entre os anos.

Verificamos que todos os valores são altamente significativos, quer para o número de fogos quer para a área total ardida. Haverá pois uma evolução significativa para uma e outra variável em todas as regiões. As estatísticas de teste têm 24 graus de liberdade sendo dadas $\|\|\mathbf{AN}\|\|^2$ e $\|\|\mathbf{AT}\|\|^2$, respectivamente, para o número de fogos e a área total ardida, sendo as componentes de \mathbf{N} e \mathbf{T} os totais para os anos na região que se está a considerar e $\|\|\mathbf{W}\|\|^2 = \mathbf{W}^t \mathbf{\Sigma}(\mathbf{W})^{-1} \mathbf{W}$.

| Região | N | T |
|---------------|----------|----------|
| Norte | 9165 | 644018 |
| Centro | 2825 | 704250 |
| Alentejo | 1939 | 254109 |
| Lisboa | 776 | 24568 |
| Algarve | 427 | 142833 |

Tabela 5.1: Estatísticas de Qui-Quadrado

Para estudar essa evolução considerou-se, para uma e outra variável, um efeito linear

$$E_L = \mu_{25} - \mu_1$$

e um efeito quadrático

$$E_Q = \mu_{25} - 2\mu_{13} + \mu_1$$

onde os μ_j são os valores médios, duma ou doutra variável, no ano j . Nas Tabelas 5.2 e 5.3 encontram-se os valores destes efeitos para **N** e para **T**.

| Região | EL | EQ |
|---------------|-----------|-----------|
| Norte | 287 | -1505 |
| Centro | 29 | -979 |
| Alentejo | 62 | -26 |
| Lisboa | 24 | -6 |
| Algarve | 5 | -61 |

Tabela 5.2: Efeitos para **N**

| Região | EL | EQ |
|----------|--------|---------|
| Norte | 13579 | -64093 |
| Centro | -23308 | -104264 |
| Alentejo | 3501 | 1661 |
| Lisboa | 569 | -492 |
| Algarve | -9308 | -10003 |

Tabela 5.3: Efeitos para T

Na interpretação destes resultados observe-se que se têm os quatro casos possíveis (Figura 5.2).

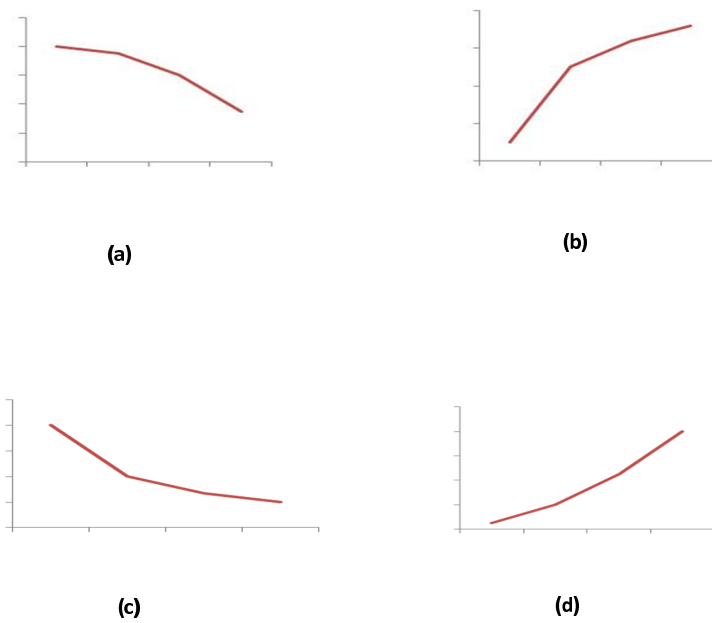


Figura 5.2: (a) $E_L < 0$, $E_Q < 0$ (b) $E_L > 0$, $E_Q < 0$ (c) $E_L < 0$, $E_Q > 0$
(d) $E_L > 0$, $E_Q > 0$

Assim para a variável **N** na região Norte temos um crescimento com a concavidade voltada para baixo. A situação piora, mas a uma taxa tendencialmente decrescente. Esta conclusão relativa a **N** verifica-se para as outras regiões. Para a variável **T** no Norte a situação mantém-se igual à da variável **N**. No Centro e no Algarve temos um decrescimento com a concavidade voltada para baixo, ou seja, a situação melhora a uma taxa crescente. Finalmente no Alentejo a situação piora a uma taxa tendencialmente crescente.

Por outro lado, as estatísticas para testar a significância destes efeitos são qui-quadrados com um grau de liberdade. Para os efeitos lineares são

$$Q = \frac{\frac{(N_{25} - N_1)^2}{2}}{\frac{N_1 + N_{25}}{2}} = \frac{(N_{25} - N_1)^2}{N_1 + N_{25}}$$

e

$$Q = \frac{(T_{25} - T_1)^2}{\sum_{j=1}^{25} x_{1,j}^2 + \sum_{j=1}^{25} x_{25,j}^2}$$

tendo-se obtido os valores da Tabela 5.4.

| Região | N | T |
|----------|----------|----------|
| Norte | 215 | 43 |
| Centro | 3 | 9 |
| Alentejo | 49 | 5 |
| Lisboa | 24 | 8 |
| Algarve | 3 | 1 |

Tabela 5.4: Estatísticas de Qui-Quadrado para E_L

Sendo o valor crítico de 3,84 conclui-se que se rejeita H_0 em todas as regiões à excepção do Centro e do Algarve no que diz respeito à variável **N**, enquanto que

relativamente à variável \mathbf{T} só não se rejeita H_0 no Algarve. Assim, na maioria dos casos evidencia-se uma evolução global.

As estatísticas para testar a significância dos efeitos quadráticos são

$$\mathbf{Q} = \frac{(N_1 - 2N_{13} + N_{25})^2}{N_1 + 4N_{13} + N_{25}}$$

e

$$\mathbf{Q} = \frac{(T_{25} - 2T_{13} + T_1)^2}{\sum_{j=1}^{25} x_{1,j}^2 + 4 \sum_{j=1}^{25} x_{13,j}^2 + \sum_{j=1}^{25} x_{25,j}^2}$$

tendo-se obtido os valores que constam na tabela

| Região | N | T |
|---------------|----------|----------|
| Norte | 545 | 52 |
| Centro | 347 | 15 |
| Alentejo | 2 | 1 |
| Lisboa | 0 | 1 |
| Algarve | 25 | 0 |

Tabela 5.5: Estatísticas de Qui-Quadrado para E_Q

que permitem concluir que a curvatura é significativa nas regiões Norte e Centro quer para a variável \mathbf{N} , quer para a variável \mathbf{T} . Esta significância também é observada para a variável \mathbf{N} no Algarve.

5.2.2 Ano a Ano

Na segunda etapa considera-se para cada ano o factor região com 5 níveis. Utilizou-se a mesma matriz \mathbf{P}_5 que atrás, e a matriz

$$\mathbf{A} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & -\frac{3}{\sqrt{30}} & \frac{2}{\sqrt{30}} & -\frac{3}{\sqrt{30}} & \frac{2}{\sqrt{30}} \end{bmatrix}$$

obtida retirando a primeira linha a \mathbf{P}_5 .

As estatísticas de teste qui-quadrado para a inexistência de diferenças entre regiões são $\|\mathbf{AN}\|^2$ e $\|\mathbf{AT}\|^2$, respectivamente, para o número de fogos e a área total ardida. As componentes de \mathbf{N} e \mathbf{T} são os totais de \mathbf{N} e \mathbf{T} para as diferentes regiões nos anos que se está a considerar. Esses totais encontram-se nas Tabelas 5.6 e 5.7.

Pela observação da Tabela 5.6 verifica-se que os anos onde ocorreu um maior número de fogos foram 1985, 1989, 2001 e 2002. As regiões mais afectadas foram o Norte e o Centro. No que diz respeito ao total da área ardida, pela Tabela 5.7 observa-se que os anos que se destacam foram 1985, 2003 e 2005, voltando o Norte e o Centro a serem as regiões mais afectadas.

| Ano | Norte | Centro | Lisboa | Alentejo | Algarve |
|------|-------|--------|--------|----------|---------|
| 1983 | 48 | 120 | 0 | 8 | 2 |
| 1984 | 1171 | 471 | 69 | 139 | 10 |
| 1985 | 1359 | 737 | 59 | 468 | 33 |
| 1986 | 753 | 493 | 64 | 219 | 19 |
| 1987 | 800 | 459 | 62 | 239 | 8 |
| 1988 | 338 | 186 | 31 | 101 | 11 |
| 1989 | 1481 | 653 | 47 | 86 | 3 |
| 1990 | 1089 | 440 | 26 | 116 | 20 |
| 1991 | 574 | 660 | 24 | 275 | 39 |
| 1992 | 258 | 330 | 29 | 84 | 18 |
| 1993 | 270 | 152 | 10 | 32 | 7 |
| 1994 | 753 | 511 | 4 | 50 | 9 |
| 1995 | 944 | 624 | 15 | 52 | 35 |
| 1996 | 1005 | 390 | 5 | 31 | 8 |
| 1997 | 456 | 298 | 7 | 7 | 7 |
| 1998 | 1311 | 417 | 39 | 88 | 12 |
| 1999 | 1012 | 399 | 17 | 40 | 11 |
| 2000 | 1019 | 578 | 36 | 118 | 10 |
| 2001 | 1131 | 510 | 33 | 156 | 50 |
| 2002 | 1067 | 498 | 74 | 220 | 25 |
| 2003 | 498 | 428 | 50 | 235 | 36 |
| 2004 | 400 | 191 | 21 | 89 | 29 |
| 2005 | 854 | 483 | 29 | 89 | 18 |
| 2006 | 196 | 89 | 1 | 27 | 0 |
| 2007 | 335 | 149 | 24 | 70 | 7 |

Tabela 5.6: Totais de N

| Ano | Norte | Centro | Lisboa | Alentejo | Algarve |
|------|--------|--------|--------|----------|---------|
| 1983 | 5213 | 36258 | 0 | 1304 | 10303 |
| 1984 | 68538 | 39379 | 2169 | 6349 | 275 |
| 1985 | 119706 | 134226 | 1568 | 31641 | 2583 |
| 1986 | 37643 | 61865 | 1155 | 10319 | 1317 |
| 1987 | 53965 | 67434 | 2132 | 11944 | 2199 |
| 1988 | 12169 | 13564 | 606 | 3930 | 1045 |
| 1989 | 112837 | 81752 | 1708 | 7390 | 89 |
| 1990 | 50540 | 57265 | 684 | 3781 | 761 |
| 1991 | 26116 | 115564 | 725 | 12828 | 5706 |
| 1992 | 8253 | 22012 | 813 | 2362 | 767 |
| 1993 | 15199 | 23319 | 794 | 1770 | 4363 |
| 1994 | 38346 | 47865 | 217 | 1666 | 723 |
| 1995 | 44049 | 76736 | 531 | 2224 | 10650 |
| 1996 | 65588 | 25743 | 145 | 1966 | 620 |
| 1997 | 12900 | 8703 | 59 | 144 | 123 |
| 1998 | 130661 | 79088 | 633 | 3703 | 774 |
| 1999 | 26854 | 36619 | 299 | 1255 | 924 |
| 2000 | 68137 | 67515 | 956 | 6947 | 327 |
| 2001 | 44952 | 44671 | 648 | 3979 | 3377 |
| 2002 | 67036 | 52503 | 1676 | 10513 | 1303 |
| 2003 | 33732 | 206204 | 4156 | 135954 | 60005 |
| 2004 | 35697 | 29893 | 1333 | 15933 | 32105 |
| 2005 | 147759 | 179718 | 945 | 6546 | 1499 |
| 2006 | 38501 | 19806 | 47 | 11372 | 0 |
| 2007 | 18792 | 12950 | 569 | 4805 | 995 |

Tabela 5.7: Totais de \mathbf{T}

| Ano | N | T |
|------|------|--------|
| 1983 | 178 | 53078 |
| 1984 | 1659 | 110889 |
| 1985 | 2170 | 266410 |
| 1986 | 1220 | 98133 |
| 1987 | 1400 | 113656 |
| 1988 | 490 | 23029 |
| 1989 | 2202 | 201689 |
| 1990 | 1442 | 104908 |
| 1991 | 1235 | 146053 |
| 1992 | 489 | 26204 |
| 1993 | 383 | 33883 |
| 1994 | 1262 | 85051 |
| 1995 | 1457 | 124038 |
| 1996 | 1370 | 91306 |
| 1997 | 717 | 21153 |
| 1998 | 1665 | 206953 |
| 1999 | 1339 | 61223 |
| 2000 | 1581 | 138038 |
| 2001 | 1460 | 85919 |
| 2002 | 1474 | 116289 |
| 2003 | 804 | 356476 |
| 2004 | 483 | 87354 |
| 2005 | 1234 | 323242 |
| 2006 | 313 | 69726 |
| 2007 | 465 | 30049 |

Tabela 5.8: Estatísticas de Qui-Quadrado para **N** e **T**

As estatísticas têm 4 graus de liberdade estando os respectivos valores na Tabela 5.8. Conclui-se que para todos os anos existem diferenças significativas ao nível de 5% entre regiões quer para **N** quer para **T**.

Para comparar as várias regiões, no que respeita às duas variáveis, utilizamos o método de comparação múltipla de Scheffé na variante atrás descrita. Como vimos, há diferenças para o par (i, l) no que respeita à variável **N** se

$$|N_i - N_l| > \sqrt{\chi_{1-p,4}^2(N_i + N_l)}$$

e para a variável **T** se

$$|T_i - T_l| > \sqrt{\chi_{1-p,4}^2(T_i + T_l)}.$$

Nas tabelas seguintes apresentamos, para o ano de 1983 os valores dos quocientes

$$d_{i,l} = \frac{|N_i - N_l|}{\sqrt{\chi_{0.95,4}^2(N_i + N_l)}}$$

e

$$\dot{d}_{i,l} = \frac{|T_i - T_l|}{\sqrt{\chi_{0.95,4}^2(T_i + T_l)}}$$

| Região | Centro | Alentejo | Lisboa | Algarve |
|----------|--------|----------|--------|---------|
| Norte | 2 | 2 | 2 | 2 |
| Centro | | 3 | 4 | 3 |
| Alentejo | | | 1 | 1 |
| Lisboa | | | | 0 |

Tabela 5.9: Valores de $d_{i,l}$ para o ano de 1983

| Região | Centro | Alentejo | Lisboa | Algarve |
|----------|--------|----------|--------|---------|
| Norte | 49 | 16 | 23 | 13 |
| Centro | | 59 | 62 | 39 |
| Alentejo | | | 12 | 27 |
| Lisboa | | | | 33 |

Tabela 5.10: Valores de $\dot{d}_{i,l}$ para o ano de 1983

As tabelas para os restantes anos encontram-se no Apêndice D. Do exame das tabelas para os vários anos a maior parte dos pares de regiões diferem significativamente ao nível de 5%. Na Tabela 5.11 realçamos pois, os pares de regiões que não diferem significativamente, tanto para a variável **N** como para a variável **T**. Observa-se que as regiões dividem-se em dois grupos, Norte e Centro por um lado, e Lisboa, Alentejo e Algarve por outro, para um nível de significância de 5%.

| Ano | N | T |
|------|------------------------------------------------------------|---------------------|
| 1983 | Lisboa vs Algarve; Alentejo vs Algarve; Alentejo vs Lisboa | |
| 1984 | | |
| 1985 | Lisboa vs Algarve | |
| 1986 | | |
| 1987 | | Lisboa vs Algarve |
| 1988 | | |
| 1989 | | |
| 1990 | Lisboa vs Algarve | Lisboa vs Algarve |
| 1991 | Lisboa vs Algarve; Norte vs Centro | |
| 1992 | Lisboa vs Algarve; Norte vs Centro | Lisboa vs Algarve |
| 1993 | Lisboa vs Algarve | |
| 1994 | Lisboa vs Algarve | |
| 1995 | Lisboa vs Algarve; Alentejo vs Algarve | |
| 1996 | Lisboa vs Algarve | |
| 1997 | Lisboa vs Algarve; Alentejo vs Lisboa; Alentejo vs Algarve | Alentejo vs Algarve |
| 1998 | | |
| 1999 | Lisboa vs Algarve; Alentejo vs Lisboa | |
| 2000 | | Norte vs Centro |
| 2001 | Lisboa vs Algarve | |
| 2002 | | |
| 2003 | Lisboa vs Algarve; Norte vs Centro | |
| 2004 | Lisboa vs Algarve | |
| 2005 | Lisboa vs Algarve | |
| 2006 | Lisboa vs Algarve | |
| 2007 | Lisboa vs Algarve | |

Tabela 5.11: Pares de regiões sem diferenças significativas

5.2.3 Análise conjunta Região e Ano

Vamos agora considerar a acção dos dois factores considerando os seus efeitos e a respectiva interacção. Temos um primeiro factor com 5 níveis, a região, e um segundo factor com 25 níveis, o ano. Tivemos pois de trabalhar com uma matriz

$$\mathbf{P}_{125} = \mathbf{P}_5 \otimes \mathbf{P}_5 \otimes \mathbf{P}_5$$

onde \mathbf{P}_5 é a mesma matriz que atrás. Então

- \mathbf{A}_1 , a matriz \mathbf{A} para o primeiro factor, é formada pelas linhas de \mathbf{P}_{125} com índices 26, 51, 76, 101;
- \mathbf{A}_2 , a matriz \mathbf{A} para o segundo factor, é formada pelas linhas de \mathbf{P}_{125} com índices 2 a 25;
- $\mathbf{A}_{1 \times 2}$, a matriz \mathbf{A} para a interacção, é formada pelas restantes linhas de \mathbf{P}_{125} com excepção da primeira.

Sendo \mathbf{Y} o vector dos totais de \mathbf{N} ou \mathbf{T} para os vários pares (região, ano) ordenados de acordo com

| Índice | Região | Ano |
|--------|---------|------|
| 1 | Norte | 1983 |
| ⋮ | ⋮ | ⋮ |
| 125 | Algarve | 2007 |

isto é, com o ano encaixado na região.

As estatísticas qui-quadrado são

$$\|\mathbf{A}_1 \mathbf{Y}\|^2,$$

$$||\mathbf{A}_2 \mathbf{Y}||^2$$

e

$$||\mathbf{A}_{1 \times 2} \mathbf{Y}||^2$$

com 4, 24 e 96 graus de liberdade respectivamente. Os valores obtidos para estas estatísticas constam na Tabela 5.12.

| | N | T |
|-----------|----------|----------|
| Região | 27546 | 2815301 |
| Anos | 11648 | 1363830 |
| Interação | 33630 | 3291861 |

Tabela 5.12: Estatísticas de Qui-Quadrado

Todas estas estatísticas foram altamente significativas.

Para avaliar a relevância relativa dos efeitos e interação utilizou-se a técnica baseada na linearidade assintótica descrita na secção 3.6. Os resultados vêm apresentados nas Tabelas 5.13 e 5.14.

| | SQ | GL | MQ | Relevância Relativa |
|-----------|-----------|-----------|-----------|----------------------------|
| Ano | 1921168 | 24 | 80049 | 0,03 |
| Região | 10194651 | 4 | 2548663 | 0,96 |
| Interação | 3078827 | 96 | 32071 | 0,01 |
| Total | | | 2660783 | 1 |

Tabela 5.13: Relevância Relativa para **N**

| | SQ | GL | MQ | Relevância Relativa |
|------------|-------------|-----------|-------------|----------------------------|
| Ano | 48592201010 | 24 | 2024675042 | 0,09 |
| Região | 78571106828 | 4 | 19642776707 | 0,88 |
| Interacção | 70402582681 | 96 | 733360236 | 0,03 |
| Total | | | 22400811985 | 1 |

Tabela 5.14: Relevância Relativa para **T**

De acordo com as tabelas conclui-se que a relevância relativa da região é dominante e que a relevância relativa da interacção é mínima para as duas variáveis.

Podemos concluir deste estudo que a metodologia seguida permite analisar a acção dos factores ano e região sobre

- número de fogos,
- área total ardida.

Infelizmente a gravidade destes sinistros não tem sido controlada. A Estatística permite apenas apontar para as conclusões que decorrem dos dados. Fica no entanto o alerta para a inoperância das políticas seguidas. Esta conclusão também é apoiada pelo trabalho de Marques et al. (2011) em que se concluiu que de um conjunto de variáveis as mais relevantes eram a Densidade Populacional e a Proximidade às Estradas o que sugere falta de controlo sobre a acção humana.

Capítulo 6

Considerações Finais

Ao fazer o balanço desta dissertação concluímos que os resultados assintóticos para os modelos colectivos multivariados permitem um tratamento adequado para

- Modelos Isolados,
- Famílias Estruturadas de Modelos

quando as amostras não são pequenas. Conseguiu-se de facto uma teoria homogénea que permite aplicar a Análise de Variância a modelos de efeitos fixos e técnicas associadas. Saliente-se o facto de termos encontrado situações em que as técnicas usuais referidas se revelam "como que demasiado potentes". Foi no entanto possível analisar esses casos utilizando a abordagem recente da linearidade assintótica.

Observe-se que os resultados assintóticos se "impõem", quando aplicáveis, pela sua beleza e facilidade de manejo.

Apêndice A

Coordenadas Polares Generalizadas e Testes Qui-Quadrado

As coordenadas polares generalizadas $r, \theta_1, \dots, \theta_{k-1}$ do ponto com coordenadas cartesianas x_1, \dots, x_k satisfazem as relações

$$x_1 = r c_1 \dots c_{k-1} = r l_1(\boldsymbol{\theta})$$

$$x_2 = r c_1 \dots s_{k-1} = r l_2(\boldsymbol{\theta})$$

\vdots

$$x_i = r c_1 \dots s_{k+1-i} = r l_i(\boldsymbol{\theta})$$

\vdots

$$x_k = r s_1 = r l_k(\boldsymbol{\theta})$$

com $r = \|\mathbf{x}\|$, $c_i = \cos(\theta_i)$ e $s_i = \sin(\theta_i)$, $i = 1, \dots, k$. Têm-se os limites de variação $r \geq 0$,

$$-\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}, i = 1, \dots, k-1$$

e

$$0 \leq \theta_k \leq 2\pi$$

definindo estes o domínio \mathcal{D} de variação do vector de ângulos ao centro Θ .

O Jacobiano da transformação de coordenadas cartesianas para coordenadas polares generalizadas é, (ver Kendal (1961))

$$J(r, \boldsymbol{\theta}) = r^{k-1} h(\boldsymbol{\theta})$$

com $h(\boldsymbol{\theta}) = \prod_{i=1}^{k-1} c_i^{k-1-i}$. A partir dos limites dados para os ângulos ao centro vemos que $J(r, \boldsymbol{\theta}) \geq 0$.

Seja $\boldsymbol{\theta}(\mathbf{x})$ o vector de ângulos centrais para (\mathbf{x}) , com \mathbf{V} uma matriz definida positiva, então

$$\mathbf{v} = \mathbf{v}^t \mathbf{V}^{-1} \mathbf{v},$$

assim, para $\mathbf{Z} \sim N(\cdot | \mathbf{d}, \mathbf{V})$ temos a densidade

$$\begin{aligned} n(\mathbf{z} | \mathbf{d}, \mathbf{V}) &= \frac{e^{-\frac{1}{2}(\mathbf{z}-\mathbf{d})^t \mathbf{V}^{-1}(\mathbf{z}-\mathbf{d})}}{(2\pi)^{k/2} \sqrt{\det(\mathbf{V})}} \\ &= \frac{e^{-\frac{1}{2}(\mathbf{z}-\mathbf{d})^t \mathbf{V}^{-1}(\mathbf{z}-\mathbf{d})}}{(2\pi)^{k/2} \sqrt{\det(\mathbf{V})}} \\ &= \frac{e^{-\frac{1}{2}(\mathbf{z}-\mathbf{d})^t \mathbf{V}^{-1}(\mathbf{z}-\mathbf{d})}}{(2\pi)^{k/2} \sqrt{\det(\mathbf{V})}} \sum_{i=0}^{\infty} \frac{(-d^t \mathbf{V}^{-1} \mathbf{z})^i}{i!}. \end{aligned}$$

Tomando

$$k(\boldsymbol{\theta}) = \mathbf{l}(\boldsymbol{\theta})$$

e

$$a(\boldsymbol{\theta}) = \mathbf{d}^t \mathbf{V}^{-1} \mathbf{l}(\boldsymbol{\theta})$$

temos

$$\mathbf{z} (= r^2 \mathbf{k}(\boldsymbol{\theta}))$$

e

$$\mathbf{d}^t \mathbf{V}^{-1} \mathbf{z} = r a(\boldsymbol{\theta}),$$

a densidade conjunta de $R = \|\mathbf{z}\|$ e $\boldsymbol{\Theta} = \boldsymbol{\theta}(\mathbf{z})$, será

$$f(r, \boldsymbol{\theta} | \mathbf{d}, \mathbf{V}) = j(r) \dot{j}(\boldsymbol{\theta}) r^{k-1} h(\boldsymbol{\theta}) \frac{e^{-\frac{(r^2 \mathbf{k}(\boldsymbol{\theta}) + \mathbf{d}) \mathbf{d}}{2}}}{(2\pi)^{k/2} \sqrt{\det(\mathbf{V})}} \sum_{i=0}^{\infty} \frac{ia(\boldsymbol{\theta})^i}{i!}$$

com $j(r)$ e $\dot{j}(\boldsymbol{\theta})$ as funções características de $[0; +\infty[$ e de \mathcal{D} .

Com $Y = \|\mathbf{Z}\|^2$, temos

$$R = \sqrt{\frac{Y}{K(\boldsymbol{\Theta})}}$$

pelo que a relação entre os valores destas variáveis são

$$r = \sqrt{\frac{y}{k(\boldsymbol{\theta})}}$$

logo a transformação de $(r, \boldsymbol{\theta})$ por $(y, \boldsymbol{\theta})$ terá o Jacobiano

$$\frac{\partial r}{\partial y} = \frac{1}{2\sqrt{y k(\boldsymbol{\theta})}},$$

assim a nova densidade conjunta virá dada pela expressão

$$\begin{aligned} f(y, \boldsymbol{\theta} | \mathbf{d}, \mathbf{V}) &= \frac{j(y) \dot{j}(\boldsymbol{\theta})}{2\sqrt{y k(\boldsymbol{\theta})}} \left(\frac{y}{k(\boldsymbol{\theta})}\right)^{\frac{k-1}{2}} h(\boldsymbol{\theta}) \frac{e^{-\frac{(y + \mathbf{d}) \mathbf{d}}{2}}}{(2\pi)^{\frac{k}{2}} \sqrt{\det(\mathbf{V})}} \sum_{i=0}^{\infty} \left(\frac{y}{k(\boldsymbol{\theta})}\right)^{\frac{i}{2}} \frac{a(\boldsymbol{\theta})^i}{i!} \\ &= e^{-\frac{\mathbf{d} \mathbf{d}}{2}} \sum_{i=0}^{\infty} \dot{j}(\boldsymbol{\theta}) \frac{2^{\frac{i}{2}}}{i!} \frac{\Gamma(\frac{k+i}{2}) a(\boldsymbol{\theta})^i h(\boldsymbol{\theta})}{2\pi^{k/2} k(\boldsymbol{\theta})^{\frac{k+i}{2}} \sqrt{\det(\mathbf{V})}} \left(\frac{j(y)}{2\Gamma(\frac{k+i}{2})} \left(\frac{y}{2}\right)^{\frac{k+i}{2}-1} e^{-\frac{y}{2}}\right). \end{aligned}$$

Assim, temos a

Proposição A.1 A função densidade conjunta de \mathbf{Y} e Θ é

$$\dot{f}(y, \boldsymbol{\theta} | \mathbf{d}, \mathbf{V}) = e^{\frac{-y}{2}} \sum_{i=0}^{\infty} m_i(\boldsymbol{\theta} | \mathbf{V}) g(y | k + i),$$

onde

$$m_i(\boldsymbol{\theta} | \mathbf{V}) = j(\boldsymbol{\theta}) \frac{2^{i/2}}{i!} \frac{\Gamma(\frac{k+i}{2}) a(\boldsymbol{\theta})^i h(\boldsymbol{\theta})}{2\pi^{k/2} k(\boldsymbol{\theta})^{\frac{k+i}{2}} \sqrt{\det(\mathbf{V})}}, \quad i = 0, \dots$$

sendo $g(y | k + i)$ a densidade de χ_{k+i}^2 .

Demonstração Temos apenas de observar que

$$g(y | k + i) = \frac{j(y)}{2\Gamma(\frac{k+i}{2})} \left(\frac{y}{2}\right)^{\frac{k+i}{2}} e^{-y/2}$$

para completar a prova. \square

Corolário 2 Se $\mathbf{d} = \mathbf{0}$, \mathbf{Y} e Θ são independentes com densidades $g(\cdot | k)$ e $m_0(\boldsymbol{\theta} | \mathbf{V})$.

Demonstração Se $\mathbf{d} = \mathbf{0}$, $a(\boldsymbol{\theta}) = 0$ e assim $\dot{f}(y, \boldsymbol{\theta} | \mathbf{0}, \mathbf{V}) = m_0(\boldsymbol{\theta} | \mathbf{V}) g(y | k)$ o que estabelece a tese. \square

Corolário 3 Se $\mathbf{Z}_\delta \xrightarrow{F} \mathbf{Z} \sim N(\cdot | \mathbf{d}, \mathbf{V})$, a densidade conjunta limite de $\mathbf{Y}_\delta = \mathbf{Z}_\delta$ e $\mathbf{U}_\delta = \mathbf{u}(\mathbf{Z}_\delta)$ será $\dot{f}(y, \mathbf{u} | \mathbf{d}, \mathbf{V})$.

Por outro lado com $\mathbf{V} = \mathbf{I}_k$ temos

$$\begin{cases} a(\boldsymbol{\theta}) = \mathbf{d}^t \mathbf{l}(\boldsymbol{\theta}) \\ k(\boldsymbol{\theta}) = \mathbf{l}(\boldsymbol{\theta})^t \mathbf{l}(\boldsymbol{\theta}) = 1 \end{cases},$$

onde $\|\mathbf{l}(\boldsymbol{\theta})\| = 1$, e

$$\begin{cases} m_0(\boldsymbol{\theta} | \mathbf{I}_k) = \frac{j(\boldsymbol{\theta}) \Gamma(k/2) h(\boldsymbol{\theta})}{2\pi^{k/2}} \\ m_l(\boldsymbol{\theta} | \mathbf{I}_k) = \frac{j(\boldsymbol{\theta}) \Gamma(k + l/2) h(\boldsymbol{\theta}) 2^{l/2}}{2\pi^{k/2} l!} (\mathbf{d}^t \mathbf{l}(\boldsymbol{\theta}))^l; \quad l = 1, \dots \end{cases}$$

Apêndice B

Teorema de Polya

Este teorema estabelece que: "Dada uma sucessão $\{F_n\}$ de funções distribuição e F uma função distribuição contínua, se

$$F_n \xrightarrow[n \rightarrow \infty]{} F$$

temos

$$\sup\{|F_n(\mathbf{y}) - F(\mathbf{y})|\} \xrightarrow[n \rightarrow \infty]{} 0."$$

A convergência pontual para uma distribuição contínua implica pois a convergência uniforme. Como foi referido este resultado é importante, uma vez que as distribuições limite obtidas são contínuas.

Demonstração Vamos usar duas medidas de probabilidade

$$\lambda(C) = \int \dots \int_C dF$$

$$\lambda_n(C) = \int \dots \int_C dF_n$$

onde C é um conjunto mensurável.

Seja X o produto cartesiano, tomando

$$P(\mathbf{A}) = X_{i=1}^k[-\mathbf{A}; \mathbf{A}]$$

e

$$[\mathbf{y}] = X_{i=1}^k[-\infty; y_i]$$

temos, com $\mathbf{y} \in \mathbb{R}^k$

$$\left\{ \begin{array}{l} F(\mathbf{y}) = \lambda([\mathbf{y}]) \\ F_n(\mathbf{y}) = \lambda_n([\mathbf{y}]) \end{array} \right. ,$$

Assumindo que F é contínua, qualquer que seja $\epsilon > 0$, existirá $A(\epsilon)$ tal que

$$\lambda(P(A(\epsilon))) = 1 - \epsilon.$$

Sejam \mathbf{y}_j , $j = 1, \dots, 2^k$ os vértices de

$$P = X_{i=1}^k[a_i, b_i]$$

então

$$(*) \left\{ \begin{array}{l} \lambda(P) = \sum_{j=1}^{2^k} (-1)^{s(\mathbf{y}_j)} F(\mathbf{y}_j) \\ \lambda_n(P) = \sum_{j=1}^{2^k} (-1)^{s(\mathbf{y}_j)} F_n(\mathbf{y}_j) \end{array} \right. ,$$

com $s(\mathbf{y}_j)$ o número de componentes de \mathbf{y}_j igual a a_i .

Assim, qualquer que seja \mathbf{A} ,

$$\lambda_n(\mathbf{A}) \xrightarrow{n \rightarrow \infty} \lambda(\mathbf{A}),$$

temos

$$\lambda_n(P(\mathbf{A}(\epsilon))) \xrightarrow{n \rightarrow \infty} 1 - \epsilon$$

havendo $n(\epsilon)$ tal que, para $n > n(\epsilon)$,

$$\lambda_n(P(\mathbf{A}(\epsilon))) > 1 - 2\epsilon.$$

Além disso, de acordo com o teorema de Heine, a restrição de $F(\cdot)$ a $P(\mathbf{A}(\epsilon))$ será uniformemente contínua. Observemos que se tem

$$P(\mathbf{A}(\epsilon)) = \bigcup_{j=1}^L P_j$$

onde os P_j são paralelepípedos com lados paralelos aos eixos coordenados tais que

$$F(\mathbf{y}_j'') - F(\mathbf{y}_j') \leq \epsilon, \quad j = 1, \dots, L$$

com \mathbf{y}_j' e \mathbf{y}_j'' os vértices de P_j com as menores e as maiores coordenadas. Temos então $n^+(\epsilon)$ tal que, para $n > n^+(\epsilon)$, temos

$$|F_n(\mathbf{y}_j') - F(\mathbf{y}_j')| < \epsilon; \quad j = 1, \dots, L$$

bem como,

$$|F_n(\mathbf{y}_j'') - F(\mathbf{y}_j'')| < \epsilon; \quad j = 1, \dots, L.$$

Se $\mathbf{y} \in P_j$, temos

$$\begin{cases} F(\mathbf{y}_j') \leq F(\mathbf{y}) \leq F(\mathbf{y}_j'') \\ F_n(\mathbf{y}_j') \leq F_n(\mathbf{y}) \leq F(\mathbf{y}_j'') \end{cases},$$

$j = 1, \dots, L$.

Assim, se $\mathbf{y} \in P(\mathbf{A}(\epsilon))$ e $n > n^+(\epsilon)$,

$$|F_n(\mathbf{y}) - F(\mathbf{y})| \leq 3\epsilon.$$

Logo, qualquer que seja $\mathbf{y} \in \mathbb{R}^k$, temos de acordo com (*)

$$\left| \lambda_n([\mathbf{y}] \cap P(\mathbf{A}(\epsilon))) - \lambda([\mathbf{y}] \cap P(\mathbf{A}(\epsilon))) \right| \leq 32^k \epsilon$$

visto os 2^k vértices de $[\mathbf{y}] \cap P(\mathbf{A}(\epsilon))$ pertencerem a $P(\mathbf{A}(\epsilon))$.

Para completar a demonstração temos apenas de salientar que

$$\begin{cases} F(\mathbf{y}) = \lambda([\mathbf{y}] \cap H(\mathbf{A})) + \lambda([\mathbf{y}] \setminus H(\mathbf{A})) \\ F_n(\mathbf{y}) = \lambda_n([\mathbf{y}] \cap H(\mathbf{A})) + \lambda_n([\mathbf{y}] \setminus H(\mathbf{A})) \end{cases}$$

e que

$$\begin{aligned} |F_n(\mathbf{y}) - F(\mathbf{y})| &\leq \left| \lambda_n([\mathbf{y}] \cap H(\mathbf{A})) - \lambda([\mathbf{y}] \cap H(\mathbf{A})) \right| + \\ &\quad + \left| \lambda_n([\mathbf{y}] \setminus H(\mathbf{A})) - \lambda([\mathbf{y}] \setminus H(\mathbf{A})) \right| \end{aligned}$$

onde

$$\begin{cases} 0 \leq \lambda([\mathbf{y}] \setminus H(\mathbf{A})) \leq \epsilon \\ 0 \leq \lambda_n([\mathbf{y}] \setminus H(\mathbf{A})) \leq \epsilon \end{cases} .$$

□

Apêndice C

Intervalos de Confiança para os Modelos Isolados

Apresentam-se de seguida os intervalos de confiança a 90%, 95% e 99% para os parâmetros λ e $\lambda\mu$ para cada uma das regiões nos anos em estudo.

| Ano | Norte Inf | Norte Sup | Centro Inf | Centro Sup | Alentejo Inf | Alentejo Sup | Lisboa Inf | Lisboa Sup | Algarve Inf | Algarve Sup |
|------|--------------|--------------|---------------|---------------|-----------------|-----------------|---------------|---------------|----------------|----------------|
| 1983 | 44 | 52 | 117 | 123 | 4 | 12 | 0 | 0 | -2 | 6 |
| 1984 | 1168 | 1174 | 468 | 474 | 136 | 142 | 66 | 72 | 5 | 15 |
| 1985 | 1356 | 1362 | 734 | 740 | 465 | 471 | 55 | 63 | 30 | 36 |
| 1986 | 750 | 756 | 490 | 496 | 216 | 222 | 60 | 68 | 16 | 22 |
| 1987 | 797 | 803 | 456 | 462 | 236 | 242 | 58 | 66 | 5 | 11 |
| 1988 | 335 | 341 | 183 | 189 | 97 | 105 | 26 | 36 | 7 | 15 |
| 1989 | 1478 | 1484 | 650 | 656 | 83 | 89 | 44 | 50 | -3 | 9 |
| 1990 | 1086 | 1092 | 437 | 443 | 113 | 119 | 23 | 29 | 16 | 24 |
| 1991 | 571 | 577 | 657 | 663 | 272 | 278 | 21 | 27 | 36 | 42 |
| 1992 | 255 | 261 | 327 | 333 | 81 | 87 | 26 | 32 | 14 | 22 |
| 1993 | 267 | 273 | 149 | 155 | 29 | 35 | 7 | 13 | 4 | 10 |
| 1994 | 750 | 756 | 508 | 514 | 47 | 53 | 0 | 8 | 6 | 12 |
| 1995 | 941 | 947 | 621 | 627 | 49 | 55 | 12 | 18 | 32 | 38 |
| 1996 | 1002 | 1008 | 387 | 393 | 27 | 35 | -1 | 11 | 5 | 11 |
| 1997 | 453 | 459 | 295 | 301 | 0 | 14 | -9 | 23 | 3 | 11 |
| 1998 | 1308 | 1314 | 414 | 420 | 85 | 91 | 35 | 43 | 9 | 15 |
| 1999 | 1009 | 1015 | 396 | 402 | 37 | 43 | 14 | 20 | 8 | 14 |
| 2000 | 1016 | 1022 | 575 | 581 | 115 | 121 | 33 | 39 | 6 | 14 |
| 2001 | 1128 | 1134 | 507 | 513 | 153 | 159 | 30 | 36 | 47 | 53 |
| 2002 | 1064 | 1070 | 495 | 501 | 217 | 223 | 70 | 78 | 22 | 28 |
| 2003 | 495 | 501 | 425 | 431 | 232 | 238 | 47 | 53 | 33 | 39 |
| 2004 | 397 | 403 | 188 | 194 | 86 | 92 | 18 | 24 | 26 | 32 |
| 2005 | 851 | 857 | 480 | 486 | 86 | 92 | 26 | 32 | 15 | 21 |
| 2006 | 193 | 199 | 86 | 92 | 24 | 30 | -19 | 21 | 0 | 0 |
| 2007 | 332 | 338 | 146 | 152 | 67 | 73 | 21 | 27 | 4 | 10 |

Tabela C.1: Intervalos de confiança a 90% para λ

82APÊNDICE C. INTERVALOS DE CONFIANÇA PARA OS MODELOS ISOLADOS

| | Norte | | Centro | | Alentejo | | Lisboa | | Algarve | |
|------|-------|------|--------|-----|----------|-----|--------|-----|---------|-----|
| Ano | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup |
| 1983 | 43 | 53 | 116 | 124 | 2 | 14 | 0 | 0 | -3 | 7 |
| 1984 | 1167 | 1175 | 467 | 475 | 134 | 144 | 65 | 73 | 3 | 17 |
| 1985 | 1355 | 1363 | 733 | 741 | 464 | 472 | 53 | 65 | 28 | 38 |
| 1986 | 749 | 757 | 489 | 497 | 215 | 223 | 59 | 69 | 15 | 23 |
| 1987 | 796 | 804 | 455 | 463 | 234 | 244 | 57 | 67 | 4 | 12 |
| 1988 | 333 | 343 | 182 | 190 | 96 | 106 | 24 | 38 | 5 | 17 |
| 1989 | 1477 | 1485 | 649 | 657 | 82 | 90 | 43 | 51 | -5 | 11 |
| 1990 | 1085 | 1093 | 436 | 444 | 111 | 121 | 21 | 31 | 14 | 26 |
| 1991 | 570 | 578 | 656 | 664 | 271 | 279 | 19 | 29 | 35 | 43 |
| 1992 | 254 | 262 | 326 | 334 | 79 | 89 | 25 | 33 | 12 | 24 |
| 1993 | 266 | 274 | 148 | 156 | 27 | 37 | 6 | 14 | 2 | 12 |
| 1994 | 749 | 757 | 507 | 515 | 45 | 55 | -2 | 10 | 4 | 14 |
| 1995 | 940 | 948 | 620 | 628 | 47 | 57 | 10 | 20 | 31 | 39 |
| 1996 | 1001 | 1009 | 386 | 394 | 26 | 36 | -4 | 14 | 4 | 12 |
| 1997 | 452 | 460 | 294 | 302 | -2 | 16 | -16 | 30 | 1 | 13 |
| 1998 | 1307 | 1315 | 413 | 421 | 84 | 92 | 33 | 45 | 7 | 17 |
| 1999 | 1008 | 1016 | 395 | 403 | 35 | 45 | 12 | 22 | 7 | 15 |
| 2000 | 1015 | 1023 | 574 | 582 | 114 | 122 | 32 | 40 | 4 | 16 |
| 2001 | 1127 | 1135 | 506 | 514 | 151 | 161 | 29 | 37 | 46 | 54 |
| 2002 | 1063 | 1071 | 494 | 502 | 216 | 224 | 69 | 79 | 21 | 29 |
| 2003 | 494 | 502 | 424 | 432 | 231 | 239 | 46 | 54 | 32 | 40 |
| 2004 | 396 | 404 | 187 | 195 | 85 | 93 | 17 | 25 | 25 | 33 |
| 2005 | 850 | 858 | 479 | 487 | 85 | 93 | 24 | 34 | 13 | 23 |
| 2006 | 191 | 201 | 85 | 93 | 23 | 31 | -27 | 29 | 0 | 0 |
| 2007 | 330 | 340 | 145 | 153 | 66 | 74 | 19 | 29 | 3 | 11 |

Tabela C.2: Intervalos de confiança a 95% para λ

| | Norte | | Centro | | Alentejo | | Lisboa | | Algarve | |
|------|-------|------|--------|-----|----------|-----|--------|-----|---------|-----|
| Ano | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup |
| 1983 | 39 | 57 | 113 | 127 | -3 | 19 | 0 | 0 | -7 | 11 |
| 1984 | 1164 | 1178 | 463 | 479 | 131 | 147 | 61 | 77 | -1 | 21 |
| 1985 | 1352 | 1366 | 730 | 744 | 461 | 475 | 48 | 70 | 24 | 42 |
| 1986 | 746 | 760 | 486 | 500 | 212 | 226 | 55 | 73 | 12 | 26 |
| 1987 | 793 | 807 | 452 | 466 | 231 | 247 | 53 | 71 | 1 | 15 |
| 1988 | 330 | 346 | 178 | 194 | 92 | 110 | 20 | 42 | 1 | 21 |
| 1989 | 1474 | 1488 | 646 | 660 | 79 | 93 | 40 | 54 | -11 | 17 |
| 1990 | 1082 | 1096 | 433 | 447 | 108 | 124 | 18 | 34 | 10 | 30 |
| 1991 | 567 | 581 | 653 | 667 | 268 | 282 | 16 | 32 | 32 | 46 |
| 1992 | 250 | 266 | 323 | 337 | 76 | 92 | 21 | 37 | 8 | 28 |
| 1993 | 263 | 277 | 145 | 159 | 24 | 40 | 2 | 18 | -1 | 15 |
| 1994 | 746 | 760 | 503 | 519 | 42 | 58 | -7 | 15 | 1 | 17 |
| 1995 | 937 | 951 | 617 | 631 | 44 | 60 | 7 | 23 | 28 | 42 |
| 1996 | 998 | 1012 | 383 | 397 | 22 | 40 | -11 | 21 | 1 | 15 |
| 1997 | 448 | 464 | 290 | 306 | -9 | 23 | -32 | 46 | -3 | 17 |
| 1998 | 1304 | 1318 | 410 | 424 | 80 | 96 | 29 | 49 | 4 | 20 |
| 1999 | 1005 | 1019 | 392 | 406 | 32 | 48 | 9 | 25 | 3 | 19 |
| 2000 | 1012 | 1026 | 571 | 585 | 111 | 125 | 29 | 43 | -1 | 21 |
| 2001 | 1124 | 1138 | 503 | 517 | 147 | 165 | 25 | 41 | 43 | 57 |
| 2002 | 1060 | 1074 | 491 | 505 | 212 | 228 | 65 | 83 | 17 | 33 |
| 2003 | 491 | 505 | 421 | 435 | 228 | 242 | 43 | 57 | 29 | 43 |
| 2004 | 393 | 407 | 184 | 198 | 82 | 96 | 13 | 29 | 22 | 36 |
| 2005 | 847 | 861 | 476 | 490 | 82 | 96 | 21 | 37 | 10 | 26 |
| 2006 | 188 | 204 | 81 | 97 | 20 | 34 | -47 | 49 | 0 | 0 |
| 2007 | 327 | 343 | 142 | 156 | 63 | 77 | 16 | 32 | 0 | 14 |

Tabela C.3: Intervalos de confiança a 99% para λ

84 APÊNDICE C. INTERVALOS DE CONFIANÇA PARA OS MODELOS ISOLADOS

| | Norte | Norte | Centro | Centro | Alentejo | Alentejo | Lisboa | Lisboa | Algarve | Algarve |
|------|--------|--------|--------|--------|----------|----------|--------|--------|---------|---------|
| Ano | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup |
| 1983 | 5213 | 5213 | 36258 | 36258 | 1304 | 1304 | 0 | 0 | 10303 | 10303 |
| 1984 | 68538 | 68538 | 39379 | 39379 | 6349 | 6349 | 2169 | 2169 | 275 | 275 |
| 1985 | 119706 | 119706 | 134226 | 134226 | 31641 | 31641 | 1568 | 1568 | 2583 | 2583 |
| 1986 | 37643 | 37643 | 61865 | 61865 | 10319 | 10319 | 1155 | 1155 | 1317 | 1317 |
| 1987 | 53965 | 53965 | 67434 | 67434 | 11944 | 11944 | 2132 | 2132 | 2199 | 2199 |
| 1988 | 12169 | 12169 | 13564 | 13564 | 3930 | 3930 | 606 | 606 | 1045 | 1045 |
| 1989 | 112837 | 112837 | 81752 | 81752 | 7390 | 7390 | 1708 | 1708 | 89 | 89 |
| 1990 | 50540 | 50540 | 57265 | 57265 | 3781 | 3781 | 684 | 684 | 761 | 761 |
| 1991 | 26116 | 26116 | 115564 | 115564 | 12828 | 12828 | 725 | 725 | 5706 | 5706 |
| 1992 | 8253 | 8253 | 22012 | 22012 | 2362 | 2362 | 813 | 813 | 767 | 767 |
| 1993 | 15199 | 15199 | 23319 | 23319 | 1770 | 1770 | 794 | 794 | 4363 | 4363 |
| 1994 | 38346 | 38346 | 47865 | 47865 | 1666 | 1666 | 217 | 217 | 723 | 723 |
| 1995 | 44049 | 44049 | 76736 | 76736 | 2224 | 2224 | 531 | 531 | 10650 | 10650 |
| 1996 | 65588 | 65588 | 25743 | 25743 | 1966 | 1966 | 145 | 145 | 620 | 620 |
| 1997 | 12900 | 12900 | 8703 | 8703 | 144 | 144 | 57 | 61 | 123 | 123 |
| 1998 | 130661 | 130661 | 79088 | 79088 | 3703 | 3703 | 633 | 633 | 774 | 774 |
| 1999 | 26854 | 26854 | 36619 | 36619 | 1255 | 1255 | 299 | 299 | 924 | 924 |
| 2000 | 68137 | 68137 | 67515 | 67515 | 6947 | 6947 | 956 | 956 | 327 | 327 |
| 2001 | 44952 | 44952 | 44671 | 44671 | 3979 | 3979 | 648 | 648 | 3377 | 3377 |
| 2002 | 67036 | 67036 | 52503 | 52503 | 10513 | 10513 | 1676 | 1676 | 1303 | 1303 |
| 2003 | 33732 | 33732 | 206204 | 206204 | 135954 | 135954 | 4156 | 4156 | 60005 | 60005 |
| 2004 | 35697 | 35697 | 29893 | 29893 | 15933 | 15933 | 1333 | 1333 | 32105 | 32105 |
| 2005 | 147759 | 147759 | 179718 | 179718 | 6546 | 6546 | 945 | 945 | 1499 | 1499 |
| 2006 | 38501 | 38501 | 19806 | 19806 | 11372 | 11372 | 47 | 47 | 0 | 0 |
| 2007 | 18792 | 18792 | 12950 | 12950 | 4805 | 4805 | 569 | 569 | 995 | 995 |

Tabela C.4: Intervalos de confiança a 90% para $\lambda\mu$

| | Norte | | Centro | | Alentejo | | Lisboa | | Algarve | |
|------|--------|--------|--------|--------|----------|--------|--------|------|---------|-------|
| Ano | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup |
| 1983 | 5213 | 5213 | 36258 | 36258 | 1304 | 1304 | 0 | 0 | 10303 | 10303 |
| 1984 | 68538 | 68538 | 39379 | 39379 | 6349 | 6349 | 2169 | 2169 | 275 | 275 |
| 1985 | 119706 | 119706 | 134226 | 134226 | 31641 | 31641 | 1568 | 1568 | 2583 | 2583 |
| 1986 | 37643 | 37643 | 61865 | 61865 | 10319 | 10319 | 1155 | 1155 | 1317 | 1317 |
| 1987 | 53965 | 53965 | 67434 | 67434 | 11944 | 11944 | 2132 | 2132 | 2199 | 2199 |
| 1988 | 12169 | 12169 | 13564 | 13564 | 3930 | 3930 | 606 | 606 | 1045 | 1045 |
| 1989 | 112837 | 112837 | 81752 | 81752 | 7390 | 7390 | 1708 | 1708 | 89 | 89 |
| 1990 | 50540 | 50540 | 57265 | 57265 | 3781 | 3781 | 684 | 684 | 761 | 761 |
| 1991 | 26116 | 26116 | 115564 | 115564 | 12828 | 12828 | 725 | 725 | 5706 | 5706 |
| 1992 | 8253 | 8253 | 22012 | 22012 | 2362 | 2362 | 813 | 813 | 767 | 767 |
| 1993 | 15199 | 15199 | 23319 | 23319 | 1770 | 1770 | 794 | 794 | 4363 | 4363 |
| 1994 | 38346 | 38346 | 47865 | 47865 | 1666 | 1666 | 217 | 217 | 723 | 723 |
| 1995 | 44049 | 44049 | 76736 | 76736 | 2224 | 2224 | 531 | 531 | 10650 | 10650 |
| 1996 | 65588 | 65588 | 25743 | 25743 | 1966 | 1966 | 145 | 145 | 620 | 620 |
| 1997 | 12900 | 12900 | 8703 | 8703 | 144 | 144 | 56 | 62 | 123 | 123 |
| 1998 | 130661 | 130661 | 79088 | 79088 | 3703 | 3703 | 633 | 633 | 774 | 774 |
| 1999 | 26854 | 26854 | 36619 | 36619 | 1255 | 1255 | 299 | 299 | 924 | 924 |
| 2000 | 68137 | 68137 | 67515 | 67515 | 6947 | 6947 | 956 | 956 | 327 | 327 |
| 2001 | 44952 | 44952 | 44671 | 44671 | 3979 | 3979 | 648 | 648 | 3377 | 3377 |
| 2002 | 67036 | 67036 | 52503 | 52503 | 10513 | 10513 | 1676 | 1676 | 1303 | 1303 |
| 2003 | 33732 | 33732 | 206204 | 206204 | 135954 | 135954 | 4156 | 4156 | 60005 | 60005 |
| 2004 | 35697 | 35697 | 29893 | 29893 | 15933 | 15933 | 1333 | 1333 | 32105 | 32105 |
| 2005 | 147759 | 147759 | 179718 | 179718 | 6546 | 6546 | 945 | 945 | 1499 | 1499 |
| 2006 | 38501 | 38501 | 19806 | 19806 | 11372 | 11372 | 46 | 48 | 0 | 0 |
| 2007 | 18792 | 18792 | 12950 | 12950 | 4805 | 4805 | 569 | 569 | 995 | 995 |

Tabela C.5: Intervalos de confiança a 95% para $\lambda\mu$

86 APÊNDICE C. INTERVALOS DE CONFIANÇA PARA OS MODELOS ISOLADOS

| | Norte | Norte | Centro | Centro | Alentejo | Alentejo | Lisboa | Lisboa | Algarve | Algarve |
|------|--------|--------|--------|--------|----------|----------|--------|--------|---------|---------|
| Ano | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup | Inf | Sup |
| 1983 | 5213 | 5213 | 36258 | 36258 | 1304 | 1304 | 0 | 0 | 10303 | 10303 |
| 1984 | 68538 | 68538 | 39379 | 39379 | 6349 | 6349 | 2169 | 2169 | 275 | 275 |
| 1985 | 119706 | 119706 | 134226 | 134226 | 31641 | 31641 | 1568 | 1568 | 2583 | 2583 |
| 1986 | 37643 | 37643 | 61865 | 61865 | 10319 | 10319 | 1155 | 1155 | 1317 | 1317 |
| 1987 | 53965 | 53965 | 67434 | 67434 | 11944 | 11944 | 2132 | 2132 | 2199 | 2199 |
| 1988 | 12169 | 12169 | 13564 | 13564 | 3930 | 3930 | 606 | 606 | 1045 | 1045 |
| 1989 | 112837 | 112837 | 81752 | 81752 | 7390 | 7390 | 1708 | 1708 | 89 | 89 |
| 1990 | 50540 | 50540 | 57265 | 57265 | 3781 | 3781 | 684 | 684 | 761 | 761 |
| 1991 | 26116 | 26116 | 115564 | 115564 | 12828 | 12828 | 725 | 725 | 5706 | 5706 |
| 1992 | 8253 | 8253 | 22012 | 22012 | 2362 | 2362 | 813 | 813 | 767 | 767 |
| 1993 | 15199 | 15199 | 23319 | 23319 | 1770 | 1770 | 794 | 794 | 4363 | 4363 |
| 1994 | 38346 | 38346 | 47865 | 47865 | 1666 | 1666 | 217 | 217 | 723 | 723 |
| 1995 | 44049 | 44049 | 76736 | 76736 | 2224 | 2224 | 531 | 531 | 10650 | 10650 |
| 1996 | 65588 | 65588 | 25743 | 25743 | 1966 | 1966 | 145 | 145 | 620 | 620 |
| 1997 | 12900 | 12900 | 8703 | 8703 | 143 | 145 | 54 | 64 | 123 | 123 |
| 1998 | 130661 | 130661 | 79088 | 79088 | 3703 | 3703 | 633 | 633 | 774 | 774 |
| 1999 | 26854 | 26854 | 36619 | 36619 | 1255 | 1255 | 299 | 299 | 924 | 924 |
| 2000 | 68137 | 68137 | 67515 | 67515 | 6947 | 6947 | 956 | 956 | 327 | 327 |
| 2001 | 44952 | 44952 | 44671 | 44671 | 3979 | 3979 | 648 | 648 | 3377 | 3377 |
| 2002 | 67036 | 67036 | 52503 | 52503 | 10513 | 10513 | 1676 | 1676 | 1303 | 1303 |
| 2003 | 33732 | 33732 | 206204 | 206204 | 135954 | 135954 | 4156 | 4156 | 60005 | 60005 |
| 2004 | 35697 | 35697 | 29893 | 29893 | 15933 | 15933 | 1333 | 1333 | 32105 | 32105 |
| 2005 | 147759 | 147759 | 179718 | 179718 | 6546 | 6546 | 945 | 945 | 1499 | 1499 |
| 2006 | 38501 | 38501 | 19806 | 19806 | 11372 | 11372 | 46 | 48 | 0 | 0 |
| 2007 | 18792 | 18792 | 12950 | 12950 | 4805 | 4805 | 569 | 569 | 995 | 995 |

Tabela C.6: Intervalos de confiança a 99% para $\lambda\mu$

Apêndice D

Método de Comparação Múltipla de Scheffé

Apresentam-se de seguida as tabelas correspondentes aos quocientes do método de comparação múltipla de Scheffé das regiões para cada um dos anos considerados.

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 6 | 9 | 10 | 11 |
| Centro | | 4 | 6 | 7 |
| Alent | | | 2 | 3 |
| Lisboa | | | | 2 |

Tabela D.1: Valores de $d_{i,l}$ para o ano de 1984

88APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 29 | 74 | 81 | 84 |
| Centro | | 50 | 59 | 64 |
| Alent | | | 15 | 24 |
| Lisboa | | | | 12 |

Tabela D.2: Valores de $\hat{d}_{i,l}$ para o ano de 1984

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 4 | 7 | 11 | 12 |
| Centro | | 3 | 8 | 8 |
| Alent | | | 6 | 6 |
| Lisboa | | | | 1 |

Tabela D.3: Valores de $d_{i,l}$ para o ano de 1985

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 9 | 73 | 110 | 109 |
| Centro | | 82 | 117 | 116 |
| Alent | | | 54 | 51 |
| Lisboa | | | | 5 |

Tabela D.4: Valores de $\hat{d}_{i,l}$ para o ano de 1985

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 6 | 8 | 9 |
| Centro | | 3 | 6 | 7 |
| Alent | | | 3 | 4 |
| Lisboa | | | | 2 |

Tabela D.5: Valores de $d_{i,l}$ para o ano de 1986

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 25 | 41 | 60 | 60 |
| Centro | | 62 | 79 | 78 |
| Alent | | | 28 | 27 |
| Lisboa | | | | 1 |

Tabela D.6: Valores de $\dot{d}_{i,l}$ para o ano de 1986

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 6 | 8 | 9 |
| Centro | | 3 | 6 | 7 |
| Alent | | | 3 | 5 |
| Lisboa | | | | 2 |

Tabela D.7: Valores de $d_{i,l}$ para o ano de 1987

90 APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 13 | 53 | 71 | 71 |
| Centro | | 64 | 80 | 80 |
| Alent | | | 27 | 27 |
| Lisboa | | | | 0 |

Tabela D.8: Valores de $\dot{d}_{i,l}$ para o ano de 1987

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 4 | 5 | 6 |
| Centro | | 2 | 3 | 4 |
| Alent | | | 2 | 3 |
| Lisboa | | | | 1 |

Tabela D.9: Valores de $d_{i,l}$ para o ano de 1988

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 21 | 33 | 31 |
| Centro | | 24 | 35 | 34 |
| Alent | | | 16 | 13 |
| Lisboa | | | | 4 |

Tabela D.10: Valores de $\dot{d}_{i,l}$ para o ano de 1988

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 6 | 11 | 12 | 12 |
| Centro | | 7 | 7 | 8 |
| Alent | | | 1 | 3 |
| Lisboa | | | | 2 |

Tabela D.11: Valores de $d_{i,l}$ para o ano de 1989

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 23 | 99 | 107 | 109 |
| Centro | | 81 | 90 | 93 |
| Alent | | | 19 | 27 |
| Lisboa | | | | 12 |

Tabela D.12: Valores de $\dot{d}_{i,l}$ para o ano de 1989

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 5 | 9 | 10 | 10 |
| Centro | | 4 | 6 | 6 |
| Alent | | | 2 | 3 |
| Lisboa | | | | 0 |

Tabela D.13: Valores de $d_{i,l}$ para o ano de 1990

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 7 | 65 | 72 | 71 |
| Centro | | 70 | 76 | 76 |
| Alent | | | 15 | 15 |
| Lisboa | | | | 1 |

Tabela D.14: Valores de $\dot{d}_{i,l}$ para o ano de 1990

92 APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 1 | 3 | 7 | 7 |
| Centro | | 4 | 8 | 8 |
| Alent | | | 5 | 4 |
| Lisboa | | | | 1 |

Tabela D.15: Valores de $d_{i,l}$ para o ano de 1991

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 77 | 22 | 50 | 37 |
| Centro | | 93 | 109 | 102 |
| Alent | | | 34 | 17 |
| Lisboa | | | | 20 |

Tabela D.16: Valores de $\hat{d}_{i,l}$ para o ano de 1991

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 1 | 3 | 4 | 5 |
| Centro | | 4 | 5 | 5 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 1 |

Tabela D.17: Valores de $d_{i,l}$ para o ano de 1992

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 4 | 5 | 5 |
| Centro | | 3 | 4 | 4 |
| Alent | | | 1 | 1 |
| Lisboa | | | | 0 |

Tabela D.18: Valores de $d_{i,l}$ para o ano de 1993

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 13 | 33 | 37 | 25 |
| Centro | | 44 | 47 | 37 |
| Alent | | | 6 | 11 |
| Lisboa | | | | 16 |

Tabela D.19: Valores de $\dot{d}_{i,l}$ para o ano de 1993

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 8 | 9 | 9 |
| Centro | | 6 | 7 | 7 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 0 |

Tabela D.20: Valores de $d_{i,l}$ para o ano de 1994

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 11 | 60 | 63 | 62 |
| Centro | | 67 | 71 | 69 |
| Alent | | | 11 | 6 |
| Lisboa | | | | 5 |

Tabela D.21: Valores de $\dot{d}_{i,l}$ para o ano de 1994

94 APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 9 | 10 | 9 |
| Centro | | 7 | 8 | 7 |
| Alent | | | 1 | 1 |
| Lisboa | | | | 1 |

Tabela D.22: Valores de $d_{i,l}$ para o ano de 1995

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 31 | 63 | 67 | 46 |
| Centro | | 86 | 89 | 73 |
| Alent | | | 10 | 24 |
| Lisboa | | | | 31 |

Tabela D.23: Valores de $\dot{d}_{i,l}$ para o ano de 1995

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 5 | 10 | 10 | 10 |
| Centro | | 6 | 6 | 6 |
| Alent | | | 1 | 1 |
| Lisboa | | | | 0 |

Tabela D.24: Valores de $d_{i,l}$ para o ano de 1996

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 43 | 79 | 83 | 82 |
| Centro | | 46 | 52 | 50 |
| Alent | | | 13 | 9 |
| Lisboa | | | | 6 |

Tabela D.25: Valores de $\dot{d}_{i,l}$ para o ano de 1996

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 7 | 7 | 7 |
| Centro | | 5 | 5 | 5 |
| Alent | | | 0 | 0 |
| Lisboa | | | | 0 |

Tabela D.26: Valores de $d_{i,l}$ para o ano de 1997

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 9 | 36 | 37 | 36 |
| Centro | | 30 | 30 | 30 |
| Alent | | | 2 | 0 |
| Lisboa | | | | 2 |

Tabela D.27: Valores de $\hat{d}_{i,l}$ para o ano de 1997

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 7 | 11 | 11 | 12 |
| Centro | | 5 | 6 | 6 |
| Alent | | | 1 | 2 |
| Lisboa | | | | 1 |

Tabela D.28: Valores de $d_{i,l}$ para o ano de 1998

96 APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 37 | 112 | 116 | 116 |
| Centro | | 85 | 90 | 90 |
| Alent | | | 15 | 14 |
| Lisboa | | | | 1 |

Tabela D.29: Valores de $\dot{d}_{i,l}$ para o ano de 1998

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 5 | 10 | 10 | 10 |
| Centro | | 6 | 6 | 6 |
| Alent | | | 1 | 1 |
| Lisboa | | | | 0 |

Tabela D.30: Valores de $d_{i,l}$ para o ano de 1999

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 13 | 50 | 52 | 51 |
| Centro | | 59 | 61 | 60 |
| Alent | | | 8 | 2 |
| Lisboa | | | | 6 |

Tabela D.31: Valores de $\dot{d}_{i,l}$ para o ano de 1999

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 4 | 9 | 10 | 10 |
| Centro | | 6 | 7 | 8 |
| Alent | | | 2 | 3 |
| Lisboa | | | | 1 |

Tabela D.32: Valores de $d_{i,l}$ para o ano de 2000

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 1 | 72 | 83 | 84 |
| Centro | | 72 | 83 | 84 |
| Alent | | | 22 | 25 |
| Lisboa | | | | 6 |

Tabela D.33: Valores de $\dot{d}_{i,l}$ para o ano de 2000

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 5 | 9 | 10 | 10 |
| Centro | | 4 | 7 | 6 |
| Alent | | | 3 | 2 |
| Lisboa | | | | 1 |

Tabela D.34: Valores de $d_{i,l}$ para o ano de 2001

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 0 | 60 | 67 | 61 |
| Centro | | 60 | 67 | 61 |
| Alent | | | 16 | 2 |
| Lisboa | | | | 14 |

Tabela D.35: Valores de $\dot{d}_{i,l}$ para o ano de 2001

98APÊNDICE D. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE SCHEFFÉ

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 5 | 8 | 10 | 10 |
| Centro | | 3 | 6 | 7 |
| Alent | | | 3 | 4 |
| Lisboa | | | | 2 |

Tabela D.36: Valores de $d_{i,l}$ para o ano de 2002

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 14 | 66 | 81 | 82 |
| Centro | | 54 | 71 | 72 |
| Alent | | | 26 | 28 |
| Lisboa | | | | 2 |

Tabela D.37: Valores de $\dot{d}_{i,l}$ para o ano de 2002

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 1 | 3 | 6 | 6 |
| Centro | | 2 | 6 | 6 |
| Alent | | | 4 | 4 |
| Lisboa | | | | 0 |

Tabela D.38: Valores de $d_{i,l}$ para o ano de 2003

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 114 | 81 | 49 | 28 |
| Centro | | 39 | 143 | 92 |
| Alent | | | 114 | 56 |
| Lisboa | | | | 72 |

Tabela D.39: Valores de $\dot{d}_{i,l}$ para o ano de 2003

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 5 | 6 | 6 |
| Centro | | 2 | 4 | 4 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 0 |

Tabela D.40: Valores de $d_{i,l}$ para o ano de 2004

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 7 | 28 | 58 | 4 |
| Centro | | 21 | 52 | 3 |
| Alent | | | 36 | 24 |
| Lisboa | | | | 55 |

Tabela D.41: Valores de $\hat{d}_{i,l}$ para o ano de 2004

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 8 | 9 | 9 |
| Centro | | 5 | 7 | 7 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 1 |

Tabela D.42: Valores de $d_{i,l}$ para o ano de 2005

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 18 | 117 | 124 | 123 |
| Centro | | 130 | 137 | 136 |
| Alent | | | 21 | 18 |
| Lisboa | | | | 4 |

Tabela D.43: Valores de $\dot{d}_{i,l}$ para o ano de 2005

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 2 | 4 | 5 | 5 |
| Centro | | 2 | 3 | 3 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 0 |

Tabela D.44: Valores de $d_{i,l}$ para o ano de 2006

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 25 | 39 | 64 | 64 |
| Centro | | 16 | 46 | 46 |
| Alent | | | 34 | 35 |
| Lisboa | | | | 2 |

Tabela D.45: Valores de $\dot{d}_{i,l}$ para o ano de 2006

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 3 | 4 | 5 | 6 |
| Centro | | 2 | 3 | 4 |
| Alent | | | 2 | 2 |
| Lisboa | | | | 1 |

Tabela D.46: Valores de $d_{i,l}$ para o ano de 2007

| Região | Centro | Alent | Lisboa | Algarve |
|--------|--------|-------|--------|---------|
| Norte | 11 | 30 | 43 | 41 |
| Centro | | 20 | 35 | 33 |
| Alent | | | 19 | 16 |
| Lisboa | | | | 3 |

Tabela D.47: Valores de $\dot{d}_{i,l}$ para o ano de 2007

Bibliografia

Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., Nesbet, C. J., (1986). *Actuarial Mathematics*. The Society of Actuaries.

Centeno, M. L., (2003). *Teoria do Risco na Actividade Seguradora*. Celta Editora - Colecção Económicas. Oeiras.

Reis, E. A. D., (1993). How long is the surplus below zero? *Insurance: Mathematics and Economics*, 12, 23 – 38.

Fonseca, M., Mexia, J. T. and Zmyslosny R., (2003). Estimators and tests for variance components in cross-nested orthogonal models. *Discussiones Mathematicae. Probability and Statistics*, 23(2) : 175 – 201.

Fonseca, M., Mexia, J. T. and Zmyslosny R., (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra and its Applications*, 417 : 75 – 86.

Fonseca, M.; Mexia, J. T. and Zmyslosny R., (2007). Jordan algebras. Generating pivot variables and orthogonal normal models. *Journal of Interdisciplinary Mathematics*, Vol.1, 305 – 326.

Fonseca, M., Mexia, J. T. and Zmyslosny R., (2008). Inference in normal models with commutative orthogonal block structure. *Acta et Commentationes Universitatis Tartunensis de Mathematica*, 3 – 16.

Fonseca, M., Jesus, V.; Mexia, J. T. and Zmyslosny R., (2009). Binary operations and canonical forms for factorial and related models. *Linear Algebra and its Applications*, 430, 2781 – 2797.

Grandell J., (1992). *Aspects of Risk Theory*. Springer Series in Statistics.

Jesus, V., Mexia, J. T., Rodrigues P. C., (2007). Inference for random effects in prime basis factorials using commutative jordan algebras *Discussiones Mathematicae. Probability and Statistics*, 27, 15 – 25.

Jesus, V., Mexia, J. T., Oliveira M. M., (2009). Binary operations and canonical forms for fractional and related models *Linear Algebra Appl.*, 430, 2781 – 2797.

Jordan, P., Von Neumann J., Wigner E., (2009). On the algebraic generalization of the quantum mechanical formalism *Ann. of Math.*, 36, 26 – 64.

Lukacs E.; Laha R. G., (1964). *Applications of characteristic functions*. Charles Griffin & Co. limited. London. Kendall, G., M., (1961). *A Course in the Geometry of n Dimensions*. Charles Griffin & Co. limited. London.

Mexia, J. T., (1987). *Multitreatment regression designs*. Trabalhos de Investigação N°1. 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 Scheffé's multiple comparison method in the presence of controlled heterocedasticity. *Journal Comp. Stat. Data Analysis*, 10(3) : 271 – 281.

Mexia, J. T., (1990). Variance free models. *Trabalhos de Investigação N°2*. Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Mexia, J. T., (1992). Assymptotic chi-squared tests, design and log-linear models. *Trabalhos de Investigação N°1*. Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

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

Mexia, J. T., Oliveira M. M., (2010). Asymptotic linearity and limit distributions, approximations. *Journal of Statistical Planning and Inference*, 140, (2), 353–357.

Oliveira, T., (1980). Statistical Choice of Univariate Extreme Models. *Statistical Distribution in Scientific Work*, vol. 6, C. Taillie et al. (eds.), 367 – 382. Reichel Dordrecht.

Oliveira M. M., Mexia, J. T., (2007a). ANOVA like analysis of matched series of studies with a common stucture. *Journal of Statistical Planning and Inference*, 137, 1862 – 1870.

Oliveira M. M., Mexia, J. T., (2007b). *Modelling series of studies with a common structure*. *Computational Statistics & Data Analysis*, 51, 5876 – 5885.

Parzen, E., (1962). *Stochastic Processes*. Holden-Day. San Francisco.

Pausas JP, Llovet J, Rodrigo A, Vallejo R (2008) Are wildfires a disaster in the Mediterranean basin: a review. *International Journal of Wildland Fire* 17(6): 713-723.

Pereira, J. M. C., Carreiras, J.M.B., Silva, J. M. N., Vasconcelos, M. J., (2006). Alguns conceitos básicos sobre fogos rurais em Portugal, in: Eds: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N., Silva, T. P., Incêndios Florestais em Portugal, ISAPress, Lisboa, 133-161.

Pollock, D., (1979). *The Algebra of Econometrics*. John Wiley & Sons. New York.

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

Seely, J., (1971a). Quadratic subspaces and completeness *The Annals of Mathematical Statistics*, 42(2) : 710 – 721.

Seely, J., Zyskind, (1971b). Linear spaces and minimum variance estimators *The Annals of Mathematical Statistics*, 42, 691 – 703.

Seely, J., (1970a). Linear spaces and unbiased estimators *The Annals of Mathematical Statistics*, 41, 1735 – 1745.

Seely, J., (1970b). Linear spaces and unbiased estimators. Application to a mixed linear model *The Annals of Mathematical Statistics*, 41, 1735 – 1745.

Seely, J., (1977a). Minimal sufficient statistics and completeness for multivariate normal families *Sankhya: The Indian Journal of Statistics, Series A*, 39(2), 170 – 185.

Silvey, S. D., (1977b). *Statistical Inference*. Chapman & Hall, London.

Marques, S., Borges, J. G., Garcia-Gonzalo, J., Moreira, F., Carreiras, J.M.B., Oliveira, M.M., Cantarina, A., Botequim, B., Pereira, J. M. C., (2011). Characterization of wildfires in Portugal *European Journal of Forest Research*, Vol.130, Issue 5 775 – 784.

Vanleeuwen, D. M., Birks, D. S., Seely, J., Mills, J., Greenwood, J. A., Jones, C. W., (1998). Sufficient conditions for orthogonal designs in mixed linear models *Journal of Statistical Planning and Inference*, 73, 373 – 389.

Vanleeuwen, D. M., Birks, D. S., Seely, J., Mills, J., Greenwood, J. A., Jones, C. W., (1999). Balance and orthogonality in designs for mixed classification models *The Annals of Mathematical Statistics*, 27(6), 1927 – 1947.

Wilks, S., (1961). *Mathematical Statistics*. John Wiley & Sons. New York.