Famílias Estruturadas e Subfamílias de matrizes Estocásticas Quase Escalares

Uma Aplicação às Eleições Autárquicas em Portugal Continental de 1976 a 1997

António Afonso Delgado

Tese apresentada à Universidade de Évora
para obtenção do Grau de Doutor em Matemática.
Especialidade: Matemática e Aplicações

ORIENTAÇÃO: Maria Manuela Oliveira
CO-ORIENTAÇÕES: João Praça Nunes Mexia e António St. Aubyn

ÉVORA, NOVEMBRO DE 2013

INSTITUTO DE INVESTIGAÇÃO E FORMAÇÃO AVANÇADA

nº de arquivo

copyright

António Afonso Delgado

Famílias Estruturadas e Sub-Famílias de Matrizes Estocásticas Quase Escalares. Uma Aplicação às Eleições Autárquicas em Portugal Continental de 1976 a 1997

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

Évora
2013

Á memória do meu pai

Á minha mãe

Á memória de Viriato Marta e da sua esposa Irene Marta

Á minha Mulher Vanise da Costa Lima

Aos meus filhos Christian e Christinny

Agradecimentos

A realização desta Dissertação de Doutoramento é o resultado do forte contributo de algumas pessoas que muito me ajudaram na sua realização. Obviamente, que não poderei falar em todas, mas deixo aqui uma palavra para aquelas que mais me ajudaram. Em primeiro lugar, gostaria de endereçar os meus agradecimentos aos Professores Doutora Maria Manuela Melo Oliveira, Doutor João Tiago Nunes Mexia e Doutor António St. Aubyn, por terem, aceitado orientar-me, pela disponibilidade, paciência, dedicação e entusiasmo demonstrados ao longo da elaboração da dissertação, que motivaram e enriqueceram o trabalho.

À minha mulher e aos meus filhos pelo tempo de ausência.

À minha irmã Alice Gomes, meu cunhado João Gomes e família, agradeço pelo ambiente propício de instalação que me proporcionaram para frequência deste curso.

Aos meus pais e irmãos em geral, um especial agradecimento pelo apoio moral e incentivos que muito me ajudaram nos momentos mais difíceis.

À colega Dra. Sónia Inácio, pela partilha de conhecimentos durante o processo de recolha e análise de dados.

Aos professores Doutor Jorge Brito e José Tiago de Oliveira pelos seus amáveis conselhos e sugestões.

À Universidade de Cabo Verde, manifestamos os nossos agradecimentos pelas sucessivas licenças concedidas para a frequência do curso de Doutoramento.

À Fundação Calouste Gulbenkian pela atribuição da bolsa de estudo,fundamental para realização deste trabalho.

Às Dras. Margarida Abecassis e Cláudia Leitão, muito obrigado pela atenção dispensada.

Resumo

Desenvolvemos modelos para matrizes estocásticas simétricas com matriz média com característica um. Estas matrizes desempenham um papel relevante em muitas situações interessantes, por exemplo em problemas de componentes principais com primeira componente dominante.

Mostramos como ajustar e validar tais modelos e como medir o respectivo grau de ajustamento através de um estudo de simulações para evidenciar o bom comportamento dos estimadores obtidos. Esse bom comportamento verifica-se quando o primeiro valor próprio da matriz média é suficientemente preponderante. Utilizando a linearidade assimptótica foi possível testar se esta condição se verifica.

Além de modelos para matrizes quase escalares estudamos famílias estruturadas de modelos, cujas matrizes correspondem aos tratamentos de um delineamento base. Consideramos ainda famílias de modelos divididas em subfamílias que correspondem a esses tratamentos.

Estamos sobretudo interessados em modelos base com estrutura ortogonal. Apresentamos essa estrutura e mostramos como aplicar esses modelos no estudo de famílias estruturadas.

Damos uma aplicação dos nossos resultados à metodologia Statis, concretamente aos resultados de eleições autárquicas em Portugal Continental. Nesta aplicação utilizamos o facto destes modelos serem aplicáveis à primeira e segunda etapas desta metodologia. Assim, consideramos um par de famílias decompostas, em subfamílias que correspondem aos tratamentos do mesmo delineamento base. A primeira família corresponde à primeira etapa da metodologia Statis, e a segunda familia corresponde à segunda etapa da mesma metodologia. Estudámos a heterogeneidade das subfamílias de ambas as famílias de modelos. Concluímos que a ordenação das subfamílias a partir da respectiva heterogeneidade se mantém da primeira para a segunda etapa.

Abstract

Structured Families and Sub-Families Almost Stochastic Scalar.

An Application to Elections in Portugal from 1976 to 1997.

We developed models for stochastic matrices with rank one mean matrix. These matrices are relevant in many interesting situations, for instance in principal components with a dominant first one. It is shown how to adjust and validate models and how to measure their adjustment degree. Simulations were carried out to show the good behavior of our estimators whenever the predominance of the first eigenvalue was sufficiently high. Inference on the predominance of the first eigenvalue was carried out. To this study we used the recent technique of asymptotic linearity, Mexia e Oliveira (2011), which enables us to test if that predominance is sufficiently high. Besides models for single matrices we will consider structured families whose matrices correspond to the treatments of a base design. We will also consider families divided in subfamilies that correspond to the treatments of the base design. We are mainly interested in models with orthogonal structure. After presenting these models we show how to use them in the study of structured families namely to the results of election in mainland Portugal. We will give an application of our results to the Statis methodology. In this application we avail ourselves of our models being of use in the first and second phases of that methodology. We thus consider a pair of families decomposed in subfamilies that correspond to the treatments of the same base design. The first family corresponds to the first phase and the second to the second phase of the Statis methodology. We estimate the heterogeneity of the subfamilies of both families. We conclude that the ranking of subfamilies according to heterogeneity is carried out from the first to the second phase.

Conteúdo

Lista de Tabelas

Simbologia e Notações

$i.i.d.$	independentes e identicamente distribuídos
$\boldsymbol{v}$	vector (letra minúscula a negrito)
$\mathbf{Y}$	vector aleatório (letra maiúscula a negrito)
$\mathbf{A}$	matriz (letra maiúscula a negrito)
Π	modelo de efeitos fixos
$\mathbb{R}$	conjunto dos números reais
$\boxplus$	soma directa ortogonal de subespaços
$\mathbf{D}(r_1, \ldots, r_k)$	matriz diagonal cujos elementos principais são as componentes de $r_1, \ldots, r_k$
$\mathbf{A} = [\mathbf{a}_1^k, \cdots, \mathbf{a}_s^k]$	matriz com vectores coluna $\mathbf{a}_1^k, \cdots, \mathbf{a}_s^k$
$\mathbf{A} = [a_{ij}]$	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 1
∇	subespaço
$\nabla^{\perp}$	complemento ortogonal de ∇
MPO	matriz de projecção ortogonal
f.g.m.	função geradora de momentos
$\otimes$	produto de Kronecker de matrizes
$R(\mathbf{A})$	espaço imagem da matriz $\mathbf{A}$
$\aleph(\mathbf{A})$	espaço de nulidade da matriz $\mathbf{A}$
$\xrightarrow[n\to\infty]{u_F}$	converge uniformemente em distribuição para, quando n tende para $+\infty$
$\xrightarrow[n\to\infty]{p}$	converge em probabilidade para, quando n tende para $+\infty$
$E(\cdot)$	valor médio
$\sim$	distribui-se como
$\Sigma(\cdot)$	matriz de covariância
$\underline{grad}$	vector gradiente
$\underline{\underline{Hess}}$	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^2_k	qui-quadrado central com k graus de liberdade
$\chi^2_{k,\delta}$	qui-quadrado com k graus de liberdade e parâmetro de não centralidade δ

Capítulo 1

Introdução

A Inferência Estatística tem uma natureza dialéctica.

Dum lado tem-se um modelo que contém o que se sabe sobre o problema em estudo antes da colheita das observações e do outro as observações. É do jogo entre modelo e observações que se tiram as conclusões. Na formulação dos modelos usa-se a linguagem das probabilidades; no entanto, é no jogo entre modelos e observações que está o essencial do trabalho estatístico.

Na utilização do modelo, pode-se considerar uma primeira fase em que se trabalhava com modelos isolados, por exemplo, para interpretar uma dada experiência. No entanto, actualmente considera-se muito a interpretação de redes de ensaios e outras situações em que em vez dum modelo isolado se consideram vários. Nos casos melhor definidos, os vários modelos:

- têm a mesma estrutura

- correspondem a combinações de níveis dum dado conjunto de factores, por exemplo, aos pares local e ano, em que ensaios de comparação de variedades se realizam. Essas combinações serão os tratamentos do delineamento base. Neste caso os factores seriam a localização e o ano.

Diremos então que se tem uma família estruturada de ensaios em que, para além do estudo individual dos modelos, se procura analisar a acção dos factores do delineamento base sobre os resultados que se obtêm para os vários ensaios. Para isso, começa-se por condensar a informação transportada por cada ensaio num vector paramétrico ajustado e numa soma de quadrados de resíduos para o erro. Por exemplo, no primeiro exemplo destas famílias, a dos modelos regressionais múltiplos, os modelos individuais eram regressões lineares nas mesmas variáveis, Mexia (1987). A condensação conduziu aqui ao vector de coeficiente ajustado pelo método dos mínimos quadrados e à soma dos quadrados dos resíduos. O que distingue esta abordagem é o estudo da acção dos factores sobre os coeficientes. Aliás, à semelhança dos casos posteriores, esse estudo pode ser feito utilizando a Análise de Variância e técnicas relacionadas.

A esta primeira aplicação outras se seguiram, destacando-se as centradas em modelos para a metodologia Statis, Oliveira (2001), Ramos (2006) e Areia (2009). Esta metodologia trata séries de estudos. Cada estudo é constituído por um trio de matrizes; uma matriz de dados e duas matrizes de pesos, uma para objectos e outra para variáveis. Por exemplo, para cada eleição autárquica podemos, por concelho, ter um estudo em que as freguesias são os objectos e as opções partidárias são as variáveis. Os pesos das freguesias podem ser proporcionais aos respectivos números de eleitores, havendo várias possibilidades para os pesos das opções de voto.

Entre as aplicações mais recentes podemos referir a desenvolvida por Cantarinha (2012), para fogos florestais em Portugal Continental. Para efeito de combate a esses fogos o país encontra-se dividido em regiões: Norte, Centro, Vale do Tejo, Alentejo e Algarve. Acima do factor localização tem-se o factor ano. Para cada par (local, ano) temos um modelo. Os modelos utilizados enquadram-se nos modelos colectivos das Matemáticas Actuariais, ver por exemplo Bowers et al. (1986). Nestes modelos é o conjunto das apólices duma carteira que "gera"os sinistros $X_1, \ldots, X_N$, os quais podem ser considerados como realizações de variáveis independentes e identicamente distribuídas independente ainda do seu número N, sendo o total do sinistros

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

caso se considere um período de tempo fixo.

Agora o conjunto das apólices é substituído pelo conjunto dos povoamentos florestais de uma dada região e os sinistros, por exemplo pelas áreas ardidas nos vários fogos. Em muitos casos, por exemplo nas aplicações à metodologia Statis, os resultados são expressos em matrizes estocásticas simétricas, tendo-se verificado que, muitas vezes o seu primeiro valor próprio domina, fortemente os restantes. Então é-se levado a admitir que a matriz média de $\mathbf{M}$ tem característica um, dizendo-se que $\mathbf{M}$ é uma matriz quase escalar. O nosso estudo incidirá sobre famílias estruturadas destas matrizes, admitindo-se para as mesmas, modelos da forma

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

onde $\lambda \boldsymbol{\alpha}\boldsymbol{\alpha}^t$ é a matriz média e $\bar{\mathbf{E}}$ é uma matriz de erros simétrica. Como veremos a informação contida numa destas matrizes condensa-se num vector de estrutura,

$$\boldsymbol{\beta} = \lambda \boldsymbol{\alpha},$$

e numa soma de quadrados dos resíduos. O estudo conjunto destas famílias estruturadas centrar-se-á sobre a acção dos factores do delineamento base sobre os vectores de estruturas.

Na metodologia Statis considera-se três etapas; inter-estrutura, compromisso e trajectórias. Na primeira destas associa-se a cada estudo, um operador matricial. Sendo

$\mathbf{A}_1, \ldots, \mathbf{A}_k$ os operadores correspondentes aos estudos obtém-se a matriz

$$\mathbf{S} = [\mathbf{S}_{i,j}]$$

em que

$$\mathbf{S}_{i,j} = tr(\mathbf{A}_i \mathbf{A}_j^t), \quad i = 1, \ldots, k \quad i = 1, \ldots, k$$

é o produto de Hilbert -Schmidt das matrizes $\mathbf{A}_i$ e $\mathbf{A}_j$, $\quad i = 1, \ldots, k \quad i = 1, \ldots, k$. Até agora as aplicações tem-se reduzido a esta primeira etapa. No entanto, na segundo etapa, considera-se a matriz compromisso

$$\mathbf{C} = \frac{1}{k} \sum_{i=1}^{k} \mathbf{A}_i,$$

podendo os modelos apresentados serem aplicados às duas primeiras etapas da metodologia Statis, desde que os $\mathbf{A}_1, \ldots, \mathbf{A}_k$ sejam simétricas.

Convém referir que os delineamentos base que consideramos estão associados a partições ortogonais de $\Re^n$ sendo n o número de tratamentos.

Por exemplo, se tivermos n factores com $a_1, \ldots, a_n$ níveis e se considerarmos as combinações possíveis temos

$$n = \pi_{i=1}^{n} a_i,$$

tratamentos. Temos então de considerar, além de valores médios, os efeitos dos níveis dos n factores e as interações entre níveis de todos ou parte dos factores.

Quanto à articulação do trabalho, no Capítulo 2 consideramos os modelos para matrizes simétricas quase escalares. A formulação apresentada baseia-se na dada por Oliveira (2001) para a matriz $\mathbf{S}$ da inter-estrutura. Refez-se o desenvolvimento, mostrando que o modelo se aplica a qualquer matriz estocástica simétrica quase escalar. Como veremos ao longo da apresentação do modelo, para o mesmo ser aplicável, o primeiro valor próprio tem de ter uma preponderância suficientemente grande em relação aos restantes.

Esta condição pode ser formulada como uma hipótese sobre

$$\tau = \frac{\lambda}{\sigma^2},$$

desde que se admita que

$$\bar{\mathbf{E}} = \frac{1}{2}(\bar{E} + \bar{E}^t),$$

e que $\boldsymbol{vec}(\mathbf{E})$ seja $\aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2})$.

Recorde-se que

$$\boldsymbol{vec}\Big([\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k]\Big) = [\boldsymbol{a}_1^t, \ldots, \boldsymbol{a}_k^t]^t.$$

No entanto, como veremos, para testar hipóteses da forma

$$H_o : \tau = \tau_c$$

contra

$$H_1 : \tau > \tau_c,$$

isto é quando, nos interessa verificar se a preponderância excede um valor elevado τ_c, será necessário construir intervalos de confiança para τ utilizando a técnica da linearidade assimptótica, Mexia e Oliveira (2011).

Esta técnica, é apresentada com algum detalhe no Capítulo 3, com a construção dos intervalos de confiança atrás referidos. No Capítulo 2, apresentam-se resultados de simulações que facilitam a escolha dos valores τ_c.

No Capítulo 4 apresentaremos a estrutura algébrica dos delineamentos ortogonais. Cada modelo destes está associado a uma partição directa ortogonal

$$R^n = \boxplus_{j=1}^{m} \bar{w}_j,$$

sendo n o número de tratamentos. Mostraremos como aplicar a análise de variância quando se tem um vector

$$\boldsymbol{y} \sim \aleph(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n),$$

isto é, um vector normal com vector médio $\boldsymbol{\mu}$ e a matriz de covariância $\sigma^2\mathbf{I}_n$, independente de $\mathbf{S}$ em que $\mathbf{S} \sim \sigma^2\chi^2_g$. Esta formulação ajusta-se bem à condensação da informação contida em cada modelo da família, num par constituído por um vector de estrutura ajustado e numa soma de quadrados dos resíduos, testando-se as hipóteses

$$H_{0,i} : \boldsymbol{\mu} \in w_i = \bar{w}_i^{\perp}, \quad i = 1, \ldots, m,$$

com $\bar{w}_i^{\perp}$ o complemento ortogonal de $\bar{w}_i$, $\quad i = 1, \ldots, m$.

Por exemplo, se estivermos a considerar o cruzamento de m factores, os pares de subespaços $(w_i, \bar{w}_i)$, $\quad i = 1, \ldots, m$, corresponderão aos conjuntos de factores. Se se tiver um único factor n, um conjunto de hipóteses correspondente será de ausência de efeitos para os níveis desse factor. Se houver mais que um factor será de ausência de interacções entre as combinações de níveis dos factores de conjunto.

Mostraremos em seguida como aplicar os resultados apresentados às famílias estruturadas de matrizes quase escalares. Em particular consideram-se famílias divididas em subfamílias as quais, e não as matrizes individuais, correspondem aos tratamentos do delineamento base e veremos como medir a heterogeneidade de tais subfamílias.

No Capítulo 5 apresenta-se a metodologia Statis, introduzida por l'Hermier des Plantes (1976), Escoufier (1980) e Lavit (1989), aplicando-se os modelos desenvolvidos às:

- matriz $\mathbf{S}$ da inter-estrutura;

- matriz compromisso **C** da segunda etapa.

Se tivermos estudos agrupados em subfamilias que correspondem aos tratamentos dum delineamento base, temos duas familias de matrizes, uma para a primeira e outra para a segunda etapa divididas em subfamilias correspondentes aos tratamentos do delineamento base. Poderá então estudar-se a medida em que a heterogeneidade das subfamilias se transporta da primeira para a segunda etapa.
No Capítulo 6, aplicaremos os nossos modelos aos resultados das eleições autárquicas em Portugal Continental de 1976 a 1997. Para cobrir o país, considerou-se uma quadrícula com 3 distritos litorais e 3 distritos interiores.
Para cada concelho dum desses distritos e cada eleição tem-se um estudo, surgindo assim as subfamilias das familias de matrizes correspondentes à inter-estrututura e para o compromisso. Veremos que a heterogeneidade destas subfamilias se transporta da primeira para a segunda etapa.
Refira-se que os aspectos mais originais desta dissertação se prendem à:

- aplicação da linearidade assimptótica para construir intervalos de confiança para a preponderância;

- apresentação da versão multivariada da linearidade assimptótica;

- introdução das subfamílias das matrizes estruturadas e ao estudo da heterogeneidade das mesmas.

Capítulo 2

Modelos

2.1 Considerações prévias

Vamos agora, como referimos, apresentar um modelo para matrizes quase escalares da forma

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

onde $\lambda \boldsymbol{\alpha}\boldsymbol{\alpha}^t$ é a matriz média e se tem uma matriz de erros

$$\bar{\mathbf{E}} = \frac{1}{2}(\mathbf{E} + \mathbf{E}^t),$$

com $\boldsymbol{vec}(\mathbf{E}) \sim \aleph(\mathbf{0}^2, \sigma^2 \mathbf{I}_k)$. Este modelo foi apresentado por Oliveira(2001) para as matrizes de produto de Hilbert- Schmiidt da primeira etapa da metodologia Statis. Observe-se que já Lavit (1988), salientava tais matrizes terem muitas vezes um primeiro valor próprio dominante. Na secção seguinte estudamos a validação do modelo e a condensação da informação transportada no vector de estrutura e numa soma de quadrados de resíduos. Esta operação de condensação é muitas vezes utilizada em inferência. Assim, no caso das regressões, a informação é condensada no vector dos coeficientes ajustados e numa soma de quadrados dos resíduos. O vector de estrutura desempenha o papel que, nas regressões, cabia ao vector dos coeficientes.

Sendo θ e $\boldsymbol{\gamma}$ o primeiro valor e vector próprio de $\mathbf{M}$, ao longo desta secção admitimos que $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$. Adiante mostramos que essa suposição se justifica quando a preponderância

$$\tau = \frac{\lambda}{\sigma^2}$$

é suficiente grande, através de simulações que podem ainda ajudar a fixar o limiar de τ_c a partir do qual é de aceitar a suposição feita .

Na secção 2.3 mostra-se que sendo

$$\tilde{\mathbf{M}} = \theta\boldsymbol{\gamma}\boldsymbol{\gamma}^t,$$

a matriz ajustada, dado ter-se

$$\|\mathbf{M}\|^2 = \|\tilde{\mathbf{M}}\|^2 + \|\mathbf{M} - \tilde{\mathbf{M}}\|^2,$$

e $\|\tilde{\mathbf{M}}\|^2 = \theta^2$, se pode medir o valor de ajustamento por

$$R^2 = \frac{\theta^2}{\mathbf{M}^2}.$$

As simulações atrás referidas são apresentadas na secção 2.4. Segue-se uma secção dedicada à preponderância. Obtém-se um estimador centrado e um intervalo de confiança para τ. No entanto na prática esses intervalos são difíceis de construir para valores grandes de τ.

2.2 Validação e condensação

Sendo θ e $\boldsymbol{\gamma}$ o primeiro valor e vector próprio de $\mathbf{M}$, temos

$$\mathbf{M}\boldsymbol{\gamma} = \theta\boldsymbol{\gamma},$$

vindo

$$\|\mathbf{M}\boldsymbol{\gamma}\|^2 = \theta^2.$$

Dada uma matriz simétrica $\mathbf{A} = [a_{ij}]$ do tipo $k \times k$, seja $\boldsymbol{w}(\mathbf{A})$ o vector, com componentes

$$a_{22}, \ldots, a_{kk}, \sqrt{2}a_{23}, \ldots, \sqrt{2}a_{2k}, \ldots, \sqrt{2}a_{(k-1,k)}.$$

Com

$$\|\mathbf{A}\|^2 = \sum_{i=2}^{k}\sum_{j=2}^{k} a_{ij}^2,$$

e $\boldsymbol{v}(\mathbf{A})$, o primeiro vector linha de $\mathbf{A}$ temos

$$\begin{aligned} (2.1) \quad)\mathbf{A}(&= \sum_{i=2}^{k} a_{ii}^2 + 2\sum_{i=2}^{k}\sum_{j=i+1}^{k} a_{ij}^2 \\ &= \|\mathbf{A}\|^2 - 2\|\boldsymbol{v}(\mathbf{A})\|^2 + a_{11}^2 \end{aligned}$$

bem como

$$)\mathbf{A}(= \|\boldsymbol{w}(\mathbf{A})\|^2.$$

Interessa-nos agora estabelecer o **Lema 2.1**

Lema 2.1 *Sendo* $\mathbf{P}$ *uma matriz ortogonal tem-se* $\boldsymbol{vec}(\mathbf{PE}) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2})$.

Demonstração De $\boldsymbol{vec}(\mathbf{E}) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2})$, resulta ter-se

$$\mathbf{E} = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k],$$

com

$$\boldsymbol{e}_1^k, \ldots, \boldsymbol{e}_k^2 \quad i.i.d \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2}),$$

vindo

$$\mathbf{PE} = [\mathbf{P}\boldsymbol{e}_1, \ldots, \mathbf{P}\boldsymbol{e}_k],$$

com

$$\mathbf{P}\boldsymbol{e}_1^k, \ldots, \mathbf{P}\boldsymbol{e}_k^k \quad i.i.d \sim \aleph(\mathbf{0}_k, \sigma^2 \mathbf{I}_k)$$

donde resulta a tese. $\square$

Podemos agora estabelecer a Proposição seguinte

Proposição 2.2 *Com* $\boldsymbol{a} = \boldsymbol{v}(\mathrm{P})$ *temos*

$$\boldsymbol{v}(\mathbf{PE}) \sim \aleph(\mathbf{0}, \frac{\sigma^2}{2}(\mathbf{I}_k + \boldsymbol{a}\boldsymbol{a}^t)),$$

bem como

$$\boldsymbol{w}(\mathbf{P\bar{E}P}^t) \sim \aleph(\mathbf{0}, \sigma^2 \mathbf{I}_\ell)$$

com $\ell = \dfrac{k(k-1)}{2}$, *sendo* $\boldsymbol{v}(\mathbf{PE})$ *independente de* $\boldsymbol{w}(\mathbf{PE})$ *e de* $)\mathbf{P\bar{E}P}^t($, *tendo-se ainda*

$$)\mathbf{P\bar{E}P}^t(\sim \sigma^2 \chi^2_\ell.$$

Demonstração Sejam $a_1, \ldots, a_k$ as componentes de $\boldsymbol{a}$. Com $\mathbf{E} = [e_{ij}]$ as componentes de $\mathbf{E}\boldsymbol{a}$, $\mathbf{E}^t\boldsymbol{a}$ e de

$$\mathbf{E}\boldsymbol{a} = \frac{1}{2}(\mathbf{E}\boldsymbol{a} + \mathbf{E}^t\boldsymbol{a})$$

serão $\sum_{j=1}^k a_j e_{ij}$, $\sum_{j=1}^k a_j e_{ji}$ e $\frac{1}{2}\sum_{j=1}^k a_j(e_{ij} + e_{ji}), i = 1, \ldots, k$, respectivamente. Como $\boldsymbol{a} = \boldsymbol{v}(\mathbf{P})$ e $\bar{\mathbf{E}}$ é simétrica as componentes de

$$\boldsymbol{v}(\mathbf{P\bar{E}}) = \bar{\mathbf{E}}\boldsymbol{a},$$

serão

$$\frac{1}{2}\sum_{j=1}^k a_j(e_{ij} + e_{ji}),\ i = 1, \ldots, k$$

vendo-se que as componentes de $\boldsymbol{v}(\mathbf{P\bar{E}})$ têm valor médio nulo, e que

$$\begin{aligned} Var\left(\frac{1}{2}\sum_{j=1}^{k} a_j(e_{ij}+e_{ji})\right) &= \sigma^2 a_i^2 + \frac{\sigma^2}{2}\sum_{j\neq i}^{k} a_j^2 \\ &= \frac{\sigma^2}{2}(1+a_i^2),\ i=1,\ldots,k \end{aligned}$$

bem como

$$Cov\left(\frac{1}{2}\sum_{j=1}^{k} a_j(e_{ij}+e_{ji}); \frac{1}{2}\sum_{j=1}^{k} a_j(e_{hj}+e_{jh})\right) = \frac{\sigma^2}{2}a_i a_l;\ l\neq h.$$

Logo

$$\boldsymbol{vec}(\mathbf{P\bar{E}}) \sim \aleph\left(\mathbf{0}_k, \frac{\sigma^2}{2}(\mathbf{I}_k + \boldsymbol{a}\boldsymbol{a}^t)\right).$$

De acordo com o **Lema 2.1** tem-se $\boldsymbol{vec}(\mathbf{PE}) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2})$ e, dado $\boldsymbol{vec}(\mathbf{E}^t\mathbf{P}^t)$ ter as mesmas componentes que $\boldsymbol{vec}(\mathbf{PE})$ também temos

$$\boldsymbol{vec}(\mathbf{E}^t\mathbf{P}^t) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2}).$$

Podemos agora voltar a aplicar o **Lema 2.1** para mostrar que

$$\boldsymbol{vec}(\mathbf{PEP}^t) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2}).$$

Dado ter-se

$$\begin{aligned} \mathbf{P\bar{E}P}^t &= \frac{1}{2}(\mathbf{PEE}^t + \mathbf{PE}^t\mathbf{P}^t) \\ &= \frac{1}{2}\left(\mathbf{PEP}^t + (\mathbf{PEP}^t)^t\right), \end{aligned}$$

com, $\mathbf{P\bar{E}P}^t = [e^\circ_{ij}]$, as componentes de $\boldsymbol{w}(\mathbf{P\bar{E}P}^t)$ serão,

$$e^\circ_{22}, \ldots, e^\circ_{kk}, \frac{1}{\sqrt{2}}(e^\circ_{23}+e^\circ_{32})+, \ldots, \frac{1}{\sqrt{2}}(e^\circ_{(k-1),k}+e^\circ_{k(k-1)}),$$

independentes das componentes

$$e^\circ_{11}, \frac{1}{2}(e^\circ_{12} + e^\circ_{21}), \ldots, \frac{1}{2}(e^\circ_{1k} + e^\circ_{k1}),$$

de $\boldsymbol{vec}(\mathbf{P\bar{E}P}^t)$. Assim, os vectores

$$\boldsymbol{vec}(\mathbf{PE}) = \mathbf{P}^t \boldsymbol{v}(\mathbf{PEP}^t)$$

e $\boldsymbol{vec}(\boldsymbol{vec}(\mathbf{P\bar{E}P}^t)$ serão independentes de

$$\boldsymbol{w}(\mathbf{PE\bar{P}}^t)$$

e de

$$)\boldsymbol{w}(\mathbf{P\bar{E}P}^t)(.$$

Finalmente, de

$$\boldsymbol{vec}(\mathbf{PEP}^t) \sim \aleph(0_{k^2}, \sigma^2 \mathbf{I}_{k^2})$$

e das expressões das componentes de

$$\boldsymbol{w}(\mathbf{P\bar{E}P}^t)$$

resulta ter-se

$$\boldsymbol{w}(\mathbf{P\bar{E}P}^t) \sim \aleph(\mathbf{0}, \sigma^2 \mathbf{I}_\ell)$$

e, consequentemente

$$)\mathbf{P\bar{E}P}^t(= \|\boldsymbol{w}(\mathbf{P\bar{E}P}^t)\|^2 \sim \sigma^2 \chi^2_\ell,$$

o que completa a demonstração. $\square$

Admitindo que $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$ e $\theta \approx \lambda$, pode então escolher-se matrizes ortognais $\mathbf{P}^\circ$ e $\mathbf{P}^+$ tais que

$$\boldsymbol{v}(\mathbf{P}^\circ) = \boldsymbol{\gamma}$$

e, que

$$\boldsymbol{v}(\mathbf{P}^{+}) = \boldsymbol{\alpha}$$

e que $\mathbf{P}^{\circ} \approx \mathbf{P}^{+}$, vindo

$$\boldsymbol{w}(\mathbf{P}^{\circ}\mathbf{S}\mathbf{P}^{\circ t}) \approx \boldsymbol{w}(\mathbf{P}^{+}\mathbf{S}\mathbf{P}^{+t}),$$

Vaquinhas e Mexia (1992).

Sendo $\boldsymbol{\delta}_i$ o vector com i-éssimo componente 1 e as restantes nulas, temos

$$\begin{aligned}
\mathbf{P}^{+}\mathbf{M}\mathbf{P}^{+t} &= \mathbf{P}^{+}(\boldsymbol{\alpha}\boldsymbol{\alpha}^t + \bar{\mathbf{E}})\mathbf{P}^{+t} \\
&= \lambda\mathbf{P}^{+}\boldsymbol{\alpha}\boldsymbol{\alpha}^t\mathbf{P}^{+t} + \mathbf{P}^{+}\bar{\mathbf{E}}\mathbf{P}^{+t} \\
&= \lambda\boldsymbol{\delta}_1\boldsymbol{\delta}_1^t + \mathbf{P}^{+}\bar{\mathbf{E}}\mathbf{P}^{+t},
\end{aligned}$$

bem como

$$\boldsymbol{w}(\mathbf{P}^{+}\mathbf{M}\mathbf{P}^{+t}) = \boldsymbol{w}(\mathbf{P}^{+}\bar{\mathbf{E}}\mathbf{P}^{+t}) \sim \aleph(\mathbf{0}, \sigma^2\mathbf{I}_{\ell}).$$

Vamos agora ver como validar este modelo. Como

$$\bar{\mathbf{E}} = \frac{1}{2}(\mathbf{E} + \mathbf{E}^t)$$

com $\bar{\mathbf{E}} = [\bar{e}_{ij}]$ e $\mathbf{E} = [e_{ij}]$, temos

$$\bar{e}_{ii} = e_{ii}, i = 1, \ldots, k$$

bem como

$$\bar{e}_{ij} = \frac{1}{2}(e_{ij} + e_{ji}), i \neq j$$

dado $\boldsymbol{vec}(\mathbf{E}) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2\mathbf{I}_{k^2})$ vê-se que as variavéis

$$\bar{e}_{11}, \ldots, \bar{e}_{kk}, \sqrt{2}\bar{e}_{12}, \ldots, \sqrt{2}\bar{e}_{1k}, \ldots, \sqrt{2}\bar{e}_{23}, \ldots, \sqrt{2}\bar{e}_{(k-1)k},$$

são i.i.d. $\aleph(0, \sigma^2)$. Continuando a admitir que $\theta \approx \lambda$ e que $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$ temos

$$\theta \boldsymbol{\gamma}\boldsymbol{\gamma}^t \approx \lambda \boldsymbol{\alpha}\boldsymbol{\alpha}^t$$

pelo que

$$[z_{ij}] = \mathbf{Z} = \mathbf{M} - \theta \boldsymbol{\gamma}\boldsymbol{\gamma}^t \approx \mathbf{M} - \lambda \boldsymbol{\alpha}\boldsymbol{\alpha}^t = \bar{\mathbf{E}};$$

logo, com

$$g = \frac{k(k+1)}{2}$$

as variáveis aleatórias

$$y_1 = z_{11}, \ldots, y_k = z_{kk}, \quad y_{k+1} = \sqrt{2} z_{12}, \ldots, y_g = \sqrt{2} z_{(k-1)k}$$

serão (aproximadamente) i.i.d. $\aleph(0, \sigma^2)$.

Assim, é-se levado a testar a hipótese

$$H_0 : Y_1, \ldots, Y_g \quad i.i.d. \quad \aleph(0, \sigma^2).$$

A estatística

$$\Im = \frac{(g-1)\big(\sum_{i=1}^{g} y_i\big)^2}{g \sum_{i=1}^{g} y_i^2 - \big(\sum_{i=1}^{g} y_i\big)^2}$$

tem, caso H_0 se verifique, a distribuição $\bar{F}(|1, g-1)$ do quociente de dois qui-quadrados centrais com 1 e $g-1$ graus de liberdade, $\Im \approx \bar{F}(|1, g-1)$, Mexia (1989, pg. 41). Esta estatística foi construída para testar H_0 contra alternativas em que as variáveis $Y_1, \ldots, Y_g$ são independentes, com variância σ^2 e valores médios $\mu_1, \ldots, \mu_g$. Com

$$\begin{cases} \mu_\circ = \frac{1}{g}\sum_{i=1}^{g}\mu_i \\ \\ \qquad\qquad \delta_1 = \frac{1}{\sigma^2}g\mu_0^2 \\ \\ \delta_2 = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{g}(\mu_i-\mu_0)^2 = \dfrac{1}{\sigma^2}\Big(\displaystyle\sum_{i=1}^{g}\mu_i^2 - g\mu_0^2\Big) \end{cases} ,$$

A hipótese H_0 verifica-se quando

$$\delta_1 = \delta_2 = 0.$$

Sob as alternativas, $\Im$ distribui-se como o quociente de qui-quadrados independentes com 1 e $g-1$ graus de liberdade e parâmetro de não centralidade δ_1 e δ_2 pondo-se $\Im \approx \bar{F}(|1, g-1, \delta_1, \delta_2)$, Mexia(1989 a).

Existem alternativas em que δ_1 predomina sobre δ_2 e outras em que δ_2 predomina sobre δ_1. Nas primeiras $\Im$ tende a tomar valores maiores do que quando a hipótese H_0 se verifica. No caso das alternativas do segundo tipo passa-se o contrário. Assim, seguindo Mexia (1989 a), vamos considerar testes bilaterais. Sendo o p-éssimo quantil de $\bar{F}(.|1, g-1)$ a região de aceitação para o teste de nível q será

$$\Big[\bar{f}_{\frac{q}{2}, 1, g-1}; \bar{f}_{1-\frac{q}{2}, 1, g-1}\Big].$$

Mostra-se que, para os níveis usuais de significância, a potência destes testes cresce rapidamente com δ_1 e δ_2 Mexia(1989 a, pgs 43 a 63).

No caso que se está a considerar a rejeição da hipótese H_0 pode resultar de

- não adequação do modelo;

- θ e $\boldsymbol{\gamma}$ não serem bons estimadores de λ e $\boldsymbol{\alpha}$, respectivamente.

Logo, quando H_0 não é rejeitada somos levados a admitir que:

- o modelo se ajusta;
- θ e $\boldsymbol{\gamma}$ são bons estimadores de λ e $\boldsymbol{\alpha}$, respectivamente.

Observe-se que se poderia ter $\lambda = 0$ reduzindo-se o modelo a

$$\mathbf{M} = \bar{\mathbf{E}}$$

o que pouco interesse teria para as aplicações. Interessa-nos pois testar a hipótese

$$H_0^\circ : \lambda = 0$$

contra

$$H_1^\circ : \lambda > 0$$

Estabeleçamos então a Proposição seguinte

Proposição 2.3 *Quando a hipótese H_0° se verifica,*

$$F = \frac{(k-1)\theta^2}{2(\|\mathbf{M}\|^2 - \theta^2)}$$

tem distribuição F central com k e ℓ graus de liberdade, $F \sim F(.|k, \ell)$

Demonstração

Como

$$\boldsymbol{v}(\mathbf{P}^\circ \mathbf{M} \mathbf{P}^{\circ t}) = \theta \boldsymbol{\delta}_1$$

temos, atendendo á expressão 2.1

$$)\mathbf{P}^\circ \mathbf{M} \mathbf{P}^{\circ t}(= \|\mathbf{M}\|^2 - \theta^2$$

o que nos permite reescrever a estatistica F como

$$F = \frac{(k-1)\theta^2}{2)\mathbf{P}^\circ\mathbf{M}\mathbf{P}^{\circ t}(}.$$

Ora, quando H_0° se verifica, $\mathbf{M} = \bar{\mathbf{E}}$ e, de acordo com a **Proposição 2.2**, temos

$$\boldsymbol{v}(\mathbf{PM}) = \boldsymbol{v}(P\bar{\mathbf{E}}) \sim \aleph(\mathbf{0}_k, \frac{\sigma^2}{2}(\mathbf{I}_k + \boldsymbol{a}\boldsymbol{a}^t))$$

independente de

$$)\mathbf{PMP}^t(\sim \aleph(\mathbf{0}, \sigma^2\mathbf{I}_\ell).$$

Assim, tem -se

$$U(\mathbf{P}) = 2\boldsymbol{v}(\mathbf{PM})^t(\mathbf{I}_k + \boldsymbol{a}.\boldsymbol{a}^t)^{-1}\boldsymbol{v}(\mathbf{PM}) \sim \sigma^2\chi_k^2,$$

bem como

$$F(\mathbf{P}) = \frac{(k-1)U(\mathbf{P})}{2)\mathbf{PMP}^t(} \sim F(.|k,\ell),$$

Mexia (1990). Sendo, $\boldsymbol{\gamma}$ vector próprio de $\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t$, com valor próprio $2\boldsymbol{\gamma}$, é vector próprio de $(\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t)^{-1}$ com valor próprio $\frac{1}{2}$, logo

$$\boldsymbol{\gamma}^t(\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t)^{-1}\boldsymbol{\gamma} = \frac{1}{2},$$

e, como $\boldsymbol{\gamma}^t\mathbf{M} = \theta\boldsymbol{\gamma}^t$, temos

$$\boldsymbol{v}(\mathbf{P}^\circ\mathbf{M}) = \theta\boldsymbol{\gamma},$$

vindo

$$\begin{aligned}
U(\mathbf{P}^\circ) &= 2\boldsymbol{v}(\mathbf{P}^\circ\mathbf{M})^t(\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t)^{-1}\boldsymbol{v}(\mathbf{P}^\circ\mathbf{M}) \\
&= 2(\theta\boldsymbol{\gamma})^t(\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t)^{-1}(\theta\boldsymbol{\gamma}) \\
&= 2(\theta\boldsymbol{\gamma})^t(\frac{1}{2}\theta\boldsymbol{\gamma}) \\
&= \theta^2,
\end{aligned}$$

pelo que $F = F(\mathbf{P}^\circ)$.

Para completar a demonstração observamos que quando $H_\circ^\circ$ se verificar ao procurarmos a distribuição de F recondicionamos em ordem a $\mathbf{P}^\circ$, obtemos $F(.|k,\ell)$, visto todas as estatísticas $F(\mathbf{P})$ terem essa distribuição.

A fim de estudar o comportamento deste teste debaixo das alternativas H_1° estabelecemos a

Proposição 2.4 $F(\mathbf{P}^+)$ *tem distribuição* F *com* k *e* l *graus de liberdade e parâmetro de não centralidade*

$$\tau = \frac{\lambda^2}{\sigma^2},$$

pondo-se $F(\mathbf{P}^+) \sim F(.|k,l,\tau)$

Demonstração

Como

$$\mathbf{P}^+\mathbf{M} = \lambda \mathbf{P}^+ \boldsymbol{\alpha}\boldsymbol{\alpha}^t + \mathbf{P}^t\bar{\mathbf{E}},$$

tem-se

$$\boldsymbol{v}(\mathbf{P}^+\mathbf{M}) = \lambda\boldsymbol{\alpha} + \boldsymbol{v}(\mathbf{P}^t\bar{\mathbf{E}}).$$

Igualmente se vê que

$$\begin{aligned} \mathbf{P}^+\mathbf{M}\mathbf{P}^{+t} &= \mathbf{P}^+(\lambda\boldsymbol{\alpha}\boldsymbol{\alpha}^t + \bar{\mathbf{E}})\mathbf{P}^{+t} \\ &= \lambda\boldsymbol{\delta}_1\boldsymbol{\delta}_1^t + \mathbf{P}^+\bar{\mathbf{E}}\mathbf{P}^{+t}, \end{aligned}$$

já que $P^+\boldsymbol{\alpha} = \boldsymbol{\alpha}$, tendo-se pois

$$)\mathbf{P}^+\mathbf{M}\mathbf{P}^{+t}(\quad = \quad)\mathbf{P}^+\bar{\mathbf{E}}\mathbf{P}^+(,$$

logo atendendo à **Proposição 2.2**, temos

$$)\mathbf{P}^{+}\mathbf{M}\mathbf{P}^{+t} \sim \sigma^2\chi^2_{\ell},$$

independente de

$$\boldsymbol{v}(\mathbf{P}^{+}\bar{\mathbf{E}}) \sim \aleph(\mathbf{0}_k, \frac{\sigma^2}{2}(\mathbf{I}_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^t)),$$

e de

$$\boldsymbol{v}(\mathbf{P}^{+}\mathbf{M}) \sim \aleph(\lambda\boldsymbol{\alpha}, \frac{\sigma^2}{2}(\mathbf{I}_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^t)).$$

Tem-se então,

$$U(\mathbf{P}^{+}) = 2\boldsymbol{v}(\mathbf{P}^{+}\mathbf{M})^t(\mathbf{I}_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^t)^{-1}\boldsymbol{v}(\mathbf{P}^{+}\mathbf{M}) \sim \sigma^2\chi^2_{k,\delta}$$

com

$$\begin{aligned} \delta &= \frac{2}{\sigma^2}(\lambda\boldsymbol{\alpha})^t(\mathbf{I}_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^t)(\lambda\boldsymbol{\alpha}) \\ &= \frac{\lambda^2}{\sigma^2} \\ &= \tau. \end{aligned}$$

A tese resulta pois de $\boldsymbol{v}(\mathbf{P}^{+}\mathbf{M})$ e, consequentemente, $U(\mathbf{P}^{+})$ serem independentes de $)\mathbf{P}^{+}\mathbf{M}\mathbf{P}^{+t}($.

Ora quando $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$, temos

$$F = F(P^{\circ}) \approx F(P^{+}).$$

Pelo que F, terá aproximadamente, distribuição $F(.|k, \ell, \tau)$. Este facto permite estender ao teste com estatística F, o bom comportamento dos testes F para modelos de efeitos fixos. Mexia (1989 a) mostra que esses testes são UMP nas famílias dos testes:

- cuja potência depende apenas do parâmetro de não centralidade;
- invariantes para translações, rotações e homotetias.

Observe-se agora que, qualquer que seja a matriz ortogonal $\mathbf{P}$,

$$)\mathbf{PSP}^t(\sim \sigma^2\chi^2_\ell.$$

Logo, descondicionando em ordem a P°, obtém-se

$$)\mathbf{P}^\circ\mathbf{SP}^{\circ^t}(\sim \sigma^2\chi^2_\ell.$$

Obtendo-se assim o estimador centrado

$$\tilde{\sigma}^2 = \frac{V}{\ell}$$

com

$$\begin{aligned} V &=)\mathbf{P}^\circ\mathbf{SP}^{\circ^t}(\\ &= \|M\|^2 - \theta^2. \end{aligned}$$

A informação contida em $\mathbf{M}$ fica, quando este modelo se ajusta , condensada no par $(\tilde{\boldsymbol{\beta}}, V)$, observando-se, que quando

$$\boldsymbol{\gamma} \approx \boldsymbol{\alpha},$$

se tem

$$\tilde{\boldsymbol{\beta}} \approx \mathbf{M}\boldsymbol{\alpha}.$$

Por outro lado, com $\not\Sigma$, indicando matriz de covariância, tem-se também

$$\not\Sigma(\mathbf{M}\boldsymbol{\alpha}) = \not\Sigma(\bar{\mathbf{E}}\boldsymbol{\alpha}),$$

quando o modelo se ajusta. Assim, raciocinando como para estabelecer a **Proposição 2.2**, mostra-se que, quando $\boldsymbol{\gamma} \simeq \boldsymbol{\alpha}$

$$\not{\Sigma}(\tilde{\boldsymbol{\beta}}) \approx \not{\Sigma}(\mathbf{M}\boldsymbol{\alpha}) = \sigma^2(\mathbf{I}_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^t) \approx \sigma^2(\mathbf{I}_k + \boldsymbol{\gamma}\boldsymbol{\gamma}^t).$$

No entanto esta heterocedasticidade é pouco acentuada já que tem-se

$$\|\boldsymbol{\alpha}\|^2 = 1,$$

os elementos de $\boldsymbol{\alpha}\boldsymbol{\alpha}^t$ são em geral pequenos. No Capítulo 4 voltamos a considerar o problema de heterocedasticidade.

2.3 Valor de ajustamento

A estatística do teste F para a hipótese H_o°

$$\mathcal{F} = \frac{k-1}{2}\frac{\theta^2}{\|\mathbf{M}\|^2 - \theta^2}$$

tem, aproximadamente, distribuição $F(|k, g, \tau)$ sendo

$$\tau = \frac{\lambda^2}{\sigma^2}$$

a predominância.

Por outro lado, tendo em conta a matriz ajustada

$$\tilde{\mathbf{M}} = \theta_1 \boldsymbol{\gamma}_1^t \boldsymbol{\gamma}_1$$

e como $\mathbf{M}$ tem decomposição espectral

$$\mathbf{M} = \sum_{j=1}^{k} \theta_j \boldsymbol{\gamma}_j^t \boldsymbol{\gamma}_j$$

temos

$$\begin{cases} \|\tilde{\mathbf{M}}\|^2 = \theta_1^2 \\ \\ \|\mathbf{M}-\tilde{\mathbf{M}}\|^2 = \sum_{j=2}^{k} \theta_j^2 \\ \\ \|\mathbf{M}\|^2 = \sum_{j=2}^{k} \theta_j^2 = \|\tilde{\mathbf{M}}\|^2 + \|\mathbf{M}-\tilde{\mathbf{M}}\|^2, \end{cases}$$

podendo-se medir o valor de ajustamento por

$$\begin{aligned} R^2 &= 1 - \frac{\|\mathbf{M}-\tilde{\mathbf{M}}\|^2}{\|\mathbf{M}\|^2} \\ &= \frac{\|\tilde{\mathbf{M}}\|^2}{\|\tilde{\mathbf{M}}\|^2 + \|\mathbf{M}-\tilde{\mathbf{M}}\|^2} \\ &= \frac{\theta_1^2}{\theta_1^2 + (\|\mathbf{M}\|^2 - \theta_1^2)} \\ &= \frac{\theta^2}{\theta^2 + (\|\mathbf{M}\|^2 - \theta^2)} \\ &= \frac{\theta^2}{\|\mathbf{M}\|^2}. \end{aligned}$$

Observa-se agora que,

$$\mathcal{F} = \frac{\theta^2}{\|\mathbf{M}\|^2 - \theta^2} = \frac{2}{k-1} F$$

tem distribuição $\bar{F}(|k, g, \tau)$ do quociente $\chi^2_{k,\tau}/\chi^2_g$ de qui-quadrados independentes, e que

$$R^2 = \frac{\mathcal{F}}{\mathcal{F}+1} = \frac{\frac{2}{k-1}F}{\frac{2}{k-1}F + 1} = \frac{F}{F + \frac{k-1}{2}}$$

Relacionando-se assim R^2 com as estatística $\mathcal{F}$ e F. Esta medida R^2 do valor do ajustamento foi obtida em Oliveira (2001). No entanto, o tratamento apresentado no âmbito desta dissertação é diferente baseando-se em $\tilde{\mathbf{M}}$ poder ser considerada como projecção ortogonal da matriz $\mathbf{M}$ sobre o espaço das matrizes $K \times k$ com característica um.

2.4 Simulações

Vamos agora realizar um conjunto de simulações para mostrar que, quando τ é suficientemente grande,

$$\boldsymbol{\gamma} \simeq \boldsymbol{\alpha}.$$

Os valores considerados para k foram: 2, 4, 6, 8, 10, 12, 15 e 20 e os valores de τ foram $12,5$, 50, 100, 200 e 400. Na construção das matrizes $\mathbf{S}$ aproveitou-se a simetria do problema tomando-se

$$\boldsymbol{\alpha} = \boldsymbol{\delta}_1.$$

Para cada par (k, τ) geraram-se 1000 matrizes $\mathbf{E}$ com

$$\boldsymbol{vec}(\mathbf{E}) \sim \aleph(\mathbf{0}_{k^2}, \mathbf{I}_{k^2})$$

a partir dos quais se obtiveram as matrizes

$$\mathbf{S} = \lambda \boldsymbol{\delta}_1 \boldsymbol{\delta}_1^t + \bar{\mathbf{E}}$$

com

$$\lambda = \sqrt{2\tau}$$

e $\boldsymbol{\alpha} = \boldsymbol{\delta}_1$. Sendo $\boldsymbol{\gamma}_1$ o primeiro vector próprio de $\mathbf{S}$, como

$$\|\boldsymbol{\delta}_1\|^2 = \|\boldsymbol{\gamma}_1\|^2 = 1,$$

tem-se

$$
\begin{aligned}
\|\boldsymbol{\delta}_1 - \boldsymbol{\gamma}_1\|^2 &= (\boldsymbol{\delta}_1 - \boldsymbol{\gamma}_1)^t(\boldsymbol{\delta}_1 - \boldsymbol{\gamma}) \\
&= \boldsymbol{\delta}_1^t\boldsymbol{\delta}_1 - 2\boldsymbol{\delta}_1^t\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_1^t\boldsymbol{\gamma}_1 \\
&= 2(1-\gamma_{11}),
\end{aligned}
$$

visto ter-se

$$\gamma_{11} = \boldsymbol{\delta}_1^t\boldsymbol{\gamma}_1$$

com γ_{11} a primeira componente de $\boldsymbol{\gamma}_1$.
Como $\|\boldsymbol{\gamma}_1\|^2 = 1$ temos

$$-1 \leq \gamma_{11} \leq 1$$

e

$$\|\boldsymbol{\alpha} - \boldsymbol{\gamma}_1\|^2 = \|\boldsymbol{\delta}_1 - \boldsymbol{\gamma}_1\|^2$$

é tanto menor quanto mais próximo de 1 estiver γ_{11}. Assim, podemos considerar γ_{11} como uma medida de eficiência de $\boldsymbol{\gamma}$, para estimar $\boldsymbol{\alpha}$.

Para cada par (k, τ) apresentamos na Tabela 1 o valor médio (V. médio) e o desvio padrão (d.p.) para a amostra de valores de γ_{11}.

Como veremos muitas vezes os valores estimados para τ caem na zona em que somos levados a admitir ter $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$ e consequentemente, $\theta_k \approx \lambda$.

Tabela 2.1: Valores médios e desvio padrão da variável γ_{11}

	k=2		k=4		k=6		k=8	
τ	V. médio	d.p	V. médio	d.p	V. médio	d.p	V. médio	d.p
12.5	0.912	0.181	0.784	0.258	0.673	0.283	0.594	0.291
50	0.988	0.020	0.965	0.053	0.940	0.060	0.909	0.117
100	0.995	0.008	0.983	0.015	0.972	0.021	0.963	0.023
200	0.996	0.004	0.992	0.007	0.987	0.009	0.982	0.010
400	0.999	0.000	0.998	0.002	0.997	0.002	0.996	0.002
	k=10		k=12		k=15		k=20	
τ	V. médio	d.p	V. médio	d.p	V. médio	d.p	V. médio	d.p
12.5	0.499	0.294	0.446	0.284	0.379	0.272	0.311	0.241
50	0.877	0.129	0.842	0.158	0.801	0.173	0.707	0.234
100	0.949	0.031	0.939	0.032	0.921	0.052	0.888	0.072
200	0.977	0.012	0.971	0.014	0.964	0.015	0.951	0.018
400	0.994	0.003	0.993	0.003	0.991	0.003	0.988	0.004

2.5 Inferência para a preponderância

Os resultados das simulações mostram que caso τ seja suficientemente grande podemos

- usar θ e $\boldsymbol{\gamma}$ como estimadores de λ e $\boldsymbol{\alpha}$;
- validar o modelo e testar a exitência de componente deterministica ($\lambda > 0$).

Assim, nesta secção vamos

- determinar um estimador pontual para τ;
- obter intervalos de confiança para τ.

Observe-se que, por dualidade, se podem utilizar estes intervalos para construir testes de hipóteses para

$$H_0(\tau_0) : \tau = \tau_0.$$

A hipótese testada é rejeitada pelo teste de nível q, se o intervalo de confiança de nível $1-q$ não contiver τ_0. A partir de intervalos limitados dos dois lados [limitados à esquerda; limitados à direita], constroem-se testes bilaterais [unilaterais esquerdos; unilaterais direito]. Para além destes testes, interessa o teste para $H_0(\tau_0)$ contra

$$H_1(\tau_0) : \tau > \tau_0$$

já que ao rejeitar-se a hipótese $H_0(\tau_0)$ se está a optar por $\tau > \tau_0$, podendo τ_0 ser o limite inferior, para a preponderância à posteriori do qual se pode acrescentar que $\boldsymbol{\gamma}_1 \approx \alpha_1$, e que $\theta_1 \approx \lambda_1$. Este teste é constituído por dualidade a partir desses intervalos de confiança de nível $1 - q$ limitado à direita.

2.5.1 Estimação pontual

Os resultados desta sub-secção encontram-se aprofundados com maior detalhe em Oliveira (2002). Sendo

$$F = F(\mathbf{P}^{+})$$

a estatística atrás utilizada para testar

$$H_0^{\circ} : \tau = 0.$$

vamos substitui-la por

$$\mathcal{F} = \frac{2}{k-1}F$$

mais manejável. Observa-se que $\mathcal{F}$ se distribui como o quociente $\chi^2_{k,\tau}/\chi^2_\ell$ dos qui-quadrados independentes, tendo densidade

$$\bar{f}(z|k,\ell,\tau) = e^{\frac{-\tau}{2}} \sum_{j=0}^{\infty} \frac{(\frac{\tau}{2})^j}{j!} \bar{f}(z|k+2j,l)$$

onde, com $h = \ell + 2j$, se tem

$$\bar{f}(z|h,\ell) = \frac{\Gamma(\dfrac{h+l}{2})}{\Gamma(\dfrac{h}{2})\Gamma(\dfrac{l}{2})} \frac{z^{\dfrac{h}{2}-1}}{(1+z)^{\dfrac{h+l}{2}}}; \quad para \quad z > o$$

(anulando-se para $z < 0$).

Como $\ell > 2r$ o r-ésimo momento de $\bar{f}(z|h,\ell)$ relativa á origem é

$$\mu'_r(|h,\ell) = \frac{\Gamma(\dfrac{h+2r}{2})}{\Gamma(\frac{h}{2})\Gamma(\frac{\ell}{2})} \frac{\Gamma(\frac{\ell-2r}{2})}{\Gamma(\frac{\ell}{2})}$$

em particular se $\ell > 2$

$$\mu(|h,\ell) = \frac{h}{\ell-2}$$

e se $\ell > 4$, tem-se

$$\mu'_2(|h,\ell) = \frac{(h+2)h}{(\ell-4)(\ell-2)}.$$

Se $\ell > 2$ temos

$$\mu(|k,\ell,\tau) = e^{-\dfrac{\tau}{2}} \sum_{j=0}^{\infty} \frac{(\frac{\tau}{2})^j}{j!} \frac{h+2j}{\ell-2} = \frac{k}{\ell-2} + \frac{\tau}{\ell-2}.$$

Obtém-se o estimador centrado

$$\widetilde{\tau} = (l-2)\top^\circ - k.$$

Analogamente se, $\ell > 4$, tem-se

$$\mu'_2(k,\ell,\tau) = \frac{k^2 + 2k + 2(k+2)\tau + \tau^2}{(\ell-4)(\ell-2)}$$

vindo

$$var(\mathcal{F}) = \frac{2(k+\tau)^2 + 2(\ell-2)(k+1)\tau}{(l-4)(l-2)^2},$$

pelo que a variância do estimador será

$$var(\bar{\tau}) = \frac{2(k+\tau)^2 + 2(\ell-2)(k+1)\tau}{\ell-4}.$$

2.5.2 Intervalos de confiança

A função $\bar{F}(z|k,\ell,\tau)$ descresce estritamente com τ, pelo que podemos utilizar o método geral para obter intervalos de confiança para τ (Mood et al. (1987, pg. 389 a 391)). Assim, sendo $\tau_{\frac{q}{2}}$ e $\tau_{1-\frac{q}{2}}$ as soluções das equações em τ

$$\begin{cases} \bar{F}(\top^\circ|k,\ell,\tau) = 1 - \dfrac{q}{2} \\ \\ \bar{F}(\top^\circ|k,\ell,\tau) = \dfrac{q}{2} \end{cases},$$

teremos, para τ, o intervalo de confiança de nivel $1-q$

$$[\tau_{\frac{q}{2}}; \tau_{1-\frac{q}{2}}].$$

Analogamente os intervalos de confiança de nivel $1-q$ limitada à esquerda e à direita serão

$$[\tau_q; +\infty[$$

e

$$[0; \tau_{1-q}[.$$

Observe-se que para resolver as equações acima referidas se pode utilizar o método da bissecção, ver Hammig (1962, pg.352). Como

$$\bar{F}(z|k,\ell,\tau) = e^{-\frac{\tau}{2}} \sum_{j=0}^{\infty} \frac{(\frac{\tau}{2})^j}{j!} \bar{F}(\top^\circ|k+2j,\ell),$$

tomando-se

$$\bar{F}_N(z|k,\ell,\tau) = e^{-\frac{\tau}{2}} \sum_{j=0}^{N} \frac{(\dfrac{\tau}{2})^j}{j!} \bar{F}(\top^\circ|k+2j,\ell),$$

tem-se

$$\bar{F}_N(z|k,\ell,\tau) < \bar{F}(z|k,\ell,\tau) < \bar{F}_N(z|k,\ell,\tau) + e^{\dfrac{-\tau}{2}} \sum_{j=N+1}^{\infty} \frac{(\dfrac{\tau}{2})^j}{j!} =, \bar{F}_N^+(z|k,\ell,\tau),$$

já que

$$0 < \bar{F}(z|k+2j,\ell) < 1, \quad j = 0,1,\ldots,$$

tomando-se então a média das soluções obtidas usando $\bar{F}_N(z|k,\ell,\tau)$ e $\bar{F}_N^+(z|k,\ell,\tau)$. Podemos pois controlar o erro da truncatura que se tem quando se substitui $\bar{F}$ por $\bar{F}_N$. Em particular, dado $\varepsilon > 0$ torna-se fácil obter o número mínimo de termos necessários para que o erro de truncatura não exceda ε. Como a derivada de

$$g_N(\tau) = e^{-\dfrac{\tau}{2}} \sum_{j=N+1}^{\infty} \frac{(\dfrac{\tau}{2})^j}{j!},$$

é

$$\begin{aligned} g'_N(\tau) &= \frac{1}{2} e^{-\dfrac{\tau}{2}} \sum_{j=N+1}^{\infty} \frac{(\dfrac{\tau}{2})^j}{j!} + \frac{1}{2} e^{-\frac{\tau}{2}} \sum_{j=N+1}^{\infty} \frac{(\frac{\tau}{2})^{j-1}}{(j-1)!} \\ &= e^{-\dfrac{\tau}{2}} \sum_{j=N+1}^{\infty} \frac{(\dfrac{\tau}{2})^N}{N!} > 0, \end{aligned}$$

$g_N(\tau)$ e, consequentemente, o número mínimo de termos a tomar cresce com τ. Assim, temos de fixar um limite superior $\bar{\tau}$ para τ e determinar o número N' mínimo de termos para garantir um erro de truncatura quando muito igual a ε, quando $\tau \leq \bar{\tau}$. Tem-se aliás

$$N' = Min\{N : g_N(\widetilde{\tau}) \leq \varepsilon\}.$$

Tabela 2.2: Número de termos

	ε				
τ	10^{-2}	10^{-3}	10^{-4}	10^{-6}	
5	7	9	10	13	
10	11	13	15	19	
20	18	21	24	28	
50	37	42	46	52	
100	67	73	78	87	

Na Tabela 2.2, encontram-se indicados, em função de $\bar{\tau}$ e ε, os números mínimos de termos a tomar.

Verifica-se no entanto, que quando T° é grande a solução numérica das equações em τ

$$\bar{F}(T^{\circ}|k, \ell, \tau) = p$$

é difícil. Ora, são precisamente as situações em que T° toma valores elevados, indicando elevadas predominâncias, que nos interessa.

Capítulo 3

Linearidade assimptótica

3.1 Considerações prévias

Vimos atrás que τ é o parâmetro de não centralidade duma distribuição F. Dada a dificuldade de resolver, em ordem a τ, a equação

$$F(z|r, s, \tau) = p$$

quando z é grande, resulta daqui ter de se substituir a abordagem clássica na construção de intervalos de confiança para τ por outra, baseada na linearidade assimptótica. A noção chave desta abordagem é a noção de função assimptoticamente linear que introduziremos na secção 3.2.

Sendo $g(.)$ uma função assimptoticamente linear mostraremos, debaixo de condições bastante gerais, que, com

$$\begin{cases} Y^\circ & = \frac{g(\boldsymbol{a}+\boldsymbol{e})-g(\boldsymbol{a})}{\|g(\boldsymbol{a})\|} \\ \\ Z^\circ & = \frac{\boldsymbol{g}(\boldsymbol{a})^t\boldsymbol{e}}{\|\boldsymbol{g}(\boldsymbol{a})\|}, \end{cases}$$

se tem

$$Sup\{|F_{Y^\circ} - F_{Z^\circ}|\} \underset{\|\boldsymbol{a}\|\to\infty}{\to} 0.$$

Em particular se Z^0 tiver distribuição que não depende de $\boldsymbol{a}$ temos

$$F_{Y^\circ} \underset{\|\boldsymbol{a}\|\to\infty}{\overset{u}{\to}} F_{Z^\circ},$$

onde $\overset{u}{\to}$ indica convergência uniforme. Se $\boldsymbol{e}$ for normal com valor médio nulo e matriz de covariância $\sigma^2\mathbf{I}_k, \boldsymbol{e} \sim \aleph(\mathbf{0}, \sigma^2\mathbf{I}_k)$, tem-se $Z^\circ \sim \aleph(0, \sigma^2)$, Aplicamos estes resultados aos

- qui-quadrados não centrais;
- parâmetros de não centralidade das distribuições F.

Com vista a desenvolvimentos futuros, incluímos ainda uma secção com a versão multivariada da linearidade assimptótica.

3.2 Distribuições limite

Sejam $\boldsymbol{g}(.)$ e $\boldsymbol{g}^*(.)$ o gradiente e a matriz Hessiana duma função com derivadas parciais de segunda ordem $\boldsymbol{g} : \Re^k \longrightarrow \Re$. Representamos por $r_d(\boldsymbol{x})$ o supremo do raio espectral $g^k(\boldsymbol{y})$ quando

$$\|\boldsymbol{y} - \boldsymbol{x}\| \leq d$$

e pomos

$$k_d(u) = sup\{\frac{r_d(\boldsymbol{x})}{\|\boldsymbol{g}(\boldsymbol{x})\|}; \|\boldsymbol{x}\| \geq u\}$$

A função $g(.)$ será assimptoticamente linear se

$$\forall_d, k_d(u) \underset{\boldsymbol{u}\to\infty}{\to} 0,$$

já que então ao afastarmo-nos de origem, a norma do gradiente domina o raio espectral da matriz Hessiana.

Vamos agora considerar as estatísticas

$$y = g(\boldsymbol{a} + \boldsymbol{e})$$

, onde $g(.)$ é assimptoticamente linear, obtendo distribuição limite quando $\|a\| \longrightarrow \infty$. Como $\boldsymbol{e}$ é sempre o mesmo a dimensão da amostra dada pelo número de componentes de $\boldsymbol{a} + \boldsymbol{e}$ é fixa. Sendo $g(.)$ assimptoticamente linear mostra-se, ver Mexia e Oliveira (2010) que se a densidade de Z^0 é limitada por C

$$Sup\{\mid F_{Y^0}(z) - F_{Z^0}(z) \mid\} \leq 2d(\epsilon),$$

onde,

$$d(\epsilon) = C\epsilon + 1 - p(\epsilon),$$

com

$$p(\epsilon) = pr\{|Y^0 - Z^0|\} < \epsilon),$$

tendo-se ainda

$$Sup\{|F_{Y^0}(z) - F_{Z^0}(z)|\} \underset{\|\boldsymbol{a}\| \to \infty}{\to} 0.$$

Até agora temos considerado a aproximação a F_{Y° dada por F_{Z°. Pode-se no entanto procurar melhorar a mesma, com $E(U)$ o vector médio de U,

$$\triangle(\boldsymbol{a}) = E'(Y^\circ - Z^\circ),$$

tomamos, em vez de Z°,

$$Z^{\circ\circ} = Z^\circ + \triangle(\boldsymbol{a}).$$

Na secção seguinte mostramos um exemplo desta nova aproximação. Por outro lado, com y_p° e z_p° os p-ésimos quantis para as distribuições, F_{Y° e de F_{Z°, temos agora a,

Proposição 3.1 *Sendo g assimptoticamente linear, tendo-se*

$$Sup\{f_{Z^\circ}(z)\} \leq C,$$

bem como

$$f_{Z^\circ}(z) \geq m(q) > 0$$

quando

$$z_{q-d(\varepsilon)} \leq z \leq z^\circ_{1-q+d(\varepsilon)},$$

teremos

$$Sup\{\mid y_p^\circ - z_p^\circ \mid ; q < p < 1-q\} \leq 2\left(\varepsilon + \frac{d(\varepsilon)}{m(q)}\right).$$

Demonstração

De acordo com o **Corolário 1** da **Proposição 3.3** tem-se

$$\begin{aligned} F_{Z^\circ}(y_p^\circ - \varepsilon) - d(\varepsilon) &\leq F_{Y^\circ}(y_p^\circ) \\ &= p \leq F_{Z^\circ}(y_p^\circ + \varepsilon) + d(\varepsilon). \end{aligned}$$

Logo

$$\begin{cases} F_{Z^\circ}(y_p^\circ - \varepsilon) & \leq p + d(\varepsilon) \\ F_{Z^\circ}(y_p^\circ + \varepsilon) & \geq p - d(\varepsilon), \end{cases}$$

pelo que

$$\begin{cases} y_p^\circ - \varepsilon & \leq z^c_{p+d(\varepsilon)} irc \\ y_p^\circ + \varepsilon & \leq z^c_{p-d(\varepsilon)} irc, \end{cases}$$

vindo

$$z^{\circ}_{p-d(\varepsilon)} - \varepsilon \leq y^{\circ}_{p} \leq z^{\circ}_{p+d(\varepsilon)} + \varepsilon.$$

Para completar a demonstação basta observar que

$$z^{\circ}_{p+d(\varepsilon)} - z^{\circ}_{p-d(\varepsilon)} \leq \frac{2d(\varepsilon)}{m(q)}$$

sempre que

$$q \leq p - d(\varepsilon) < p + d(\varepsilon) \leq 1 - q.$$

$\square$

Quando F_{Z^0} não depende de $\boldsymbol{u}$ temos

$$F_{Y_0} \xrightarrow{u} F_{Z^0}$$

onde $\xrightarrow{u}$ representa a convergência uniforme. Quando $\boldsymbol{e}$ é normal com vectores médio nulo e matriz de covariância $\not\Sigma$

$$\boldsymbol{e} \sim \aleph(\mathbf{0}, \not\Sigma);$$

mostra-se que, quando $\not\Sigma$ é invertível

$$Sup\{F_{Z_0}(z)\} \leq \frac{1}{\sqrt{2\pi}\theta_k},$$

com θ_k o menor valor próprio de $\not\Sigma$ o que se permite tomar

$$C = \frac{1}{\sqrt{2\pi}\theta_k},$$

para obter um majorante para

$$Sup\{F_{Z^0}(z) - F_{Y^0}(z)\},$$

Mexia e Oliveira (2010). Por outro lado, sendo as caudas com probabilidade $\frac{q}{2}$ da normal reduzida limitadas por $-z_q$ e z_q temos

$$m_q = \frac{e^{\frac{-z^2}{q/2}}}{\sqrt{2\pi}}.$$

Observa-se que no caso em que $\boldsymbol{e} \sim \aleph(\mathbf{0}, \not\Sigma)$,

$$\boldsymbol{b} = \frac{1}{\|\boldsymbol{a}\|}\boldsymbol{a},$$

se tem

$$Z^0 \sim \aleph(0, \boldsymbol{b}^t \not\Sigma \boldsymbol{b}).$$

pelo que

$$F_{Y^0}(.) \underset{\|\boldsymbol{a}\|\to\infty}{\overset{u}{\to}} \aleph(0, \boldsymbol{b}^t \not\Sigma \boldsymbol{b})$$

caso $\boldsymbol{b}$ seja fixo.

3.3 Distribuição limite dos qui-quadrados

No presente trabalho os qui-quadrados desempenham um papel importante o que justifica a inclusão desta aplicação aos mesmos. Assim, vamos obter distribuição limite normal para

$$Y^\circ = \frac{\chi^2_{k,\delta} - \delta}{2\sqrt{\delta}},$$

quando $\delta \to \infty$. Teremos agora

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

pelo que

$$\begin{cases} \boldsymbol{g}(\boldsymbol{x}) & = 2\boldsymbol{x} \\ \boldsymbol{g}^\star(\boldsymbol{x}) & = 2I_k. \end{cases}$$

Logo, com

$$Y = g(\boldsymbol{\mu} + \boldsymbol{e})$$

onde $\boldsymbol{e} \sim \aleph(\mathbf{0}, I_k)$, teremos

$$Y \sim \chi^2_{k,\delta},$$

com $\delta = \|\boldsymbol{\mu}\|^2 = g(\boldsymbol{\mu})$. Como $\boldsymbol{g}(\boldsymbol{\mu}) = 2\boldsymbol{\mu}$ temos,

$$\|\boldsymbol{g}(\boldsymbol{\mu})\| = 2\|\boldsymbol{\mu}\| = 2\sqrt{\delta},$$

vindo

$$Z^\circ = \frac{2\boldsymbol{\mu}^t\boldsymbol{e}}{2\|\boldsymbol{\mu}\|} \sim \aleph(0,1).$$

Por outro lado,

$$g(\boldsymbol{\mu} + \boldsymbol{e}) = \|\boldsymbol{\mu} + \boldsymbol{e}\|^2 = \|\boldsymbol{\mu}\|^2 + 2\boldsymbol{\mu}^t\boldsymbol{e} + \|\boldsymbol{e}\|^2,$$

vindo,

$$Y^\circ - Z^\circ = \frac{\|\boldsymbol{e}\|^2}{2\sqrt{\delta}} = \frac{\chi^2_k}{2\sqrt{\delta}}.$$

Logo

$$\bigtriangledown = \bigtriangledown(\boldsymbol{\mu}) = \frac{k}{2\sqrt{\delta}},$$

tendo-se pois,

$$Z^{\circ\circ} = Z^0 + \bigtriangledown \sim \aleph(\bigtriangledown, 1)$$

visto que $Z^0 \sim \aleph(0,1)$.

Vamos agora utilizar a desigualdade de Berry (1941) e Essen(1945) para obter um majorante para $\mid F_{Y^\circ} - F_{Z^{\circ\circ}} \mid$. Assim, tem-se

$$Sup\{\mid F_1(z) - F_2(z) \mid\} \leq \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mid \frac{\varphi_1(t) - \varphi_2(t)}{t} \mid dt,$$

onde φ_1 e φ_2 são funções caracteristicas de F_1 e F_2.

Ora,

$$\begin{cases} \varphi_{Y^\circ}(t) = \varphi_{(\frac{\chi^2_{k,\delta}-\delta}{\sqrt{\delta}})}(t) \quad = e^{-i\sqrt{\delta}t}\varphi_{\chi^2_{k,\delta}}(\frac{t}{\sqrt{\delta}}) = \dfrac{e^{-\dfrac{t^2}{2(1-\frac{t}{\sqrt{\delta}})}}}{(1-\frac{it}{\sqrt{\delta}})^{k/2}} \\ \\ \varphi_{2^{\circ\circ}}(t) = e^{\dfrac{ikt}{2\sqrt{\delta}} - \dfrac{t^2}{2}} \quad . \end{cases}$$

Utilizando o MAT-LAB, obtiveram-se os valores apresentados na Tabela 3.1 para

$$\triangle(k,\delta) = Sup\{\mid F_{Y^\circ}(z) - F_{Z^{\circ\circ}}(z) \mid\}.$$

Tabela 3.1: Limites de Berry- Essen

	k			
δ	10	20	30	50
10	0.081	0.116	0.149	0.201
20	0.053	0.073	0.054	0.131
50	0.031	0.039	0.048	0.068
100	0.021	0.024	0.029	0.04
1000	0.006	0.007	0.007	0.008
10000	0.002	0.002	0.007	0.002

A análise desta Tabela exibe resultados perfeitamente compatíveis com ter-se

$$\triangle(k,\delta) \xrightarrow[\delta\to\infty]{} 0.$$

3.4 Parâmetro de não centralidade para estatísticas F

Esta aplicação é de certa forma uma continuação da anterior e é de interesse para inferência sobre preponderância em situações de forte domínio da mesma. Voltamos a tomar

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

voltando a ter

$$\begin{cases} g_1(\boldsymbol{x}) & = 2\boldsymbol{x} \\ g_2(\boldsymbol{x}) & = 2I_k. \end{cases}$$

Então, se

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

onde $\boldsymbol{e} \sim \aleph(\boldsymbol{0}, \sigma^2 I_k)$, ter-se-á $Y \sim \sigma^2 \chi^2_{k,\delta}$, com

$$\delta = \frac{1}{\sigma^2}\|\boldsymbol{\mu}\|^2 = \frac{1}{\sigma^2}g(\boldsymbol{\mu}).$$

Temos então

$$Z^\circ = \frac{\boldsymbol{\mu}^t \boldsymbol{e}}{\|\boldsymbol{\mu}\|^2} \sim \aleph(0, \sigma^2).$$

Logo F_{Z° não depende de $\boldsymbol{\mu}$ e a distribuição limite de

$$Y^\circ = \frac{Y - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|} = \frac{\|\boldsymbol{\mu} + \boldsymbol{e}\|^2 - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|},$$

quando $\|\boldsymbol{\mu}\| \to \infty$, será $\aleph(0, \sigma^2)$. Com z_q o limite para caudas com probabilidade $\frac{q}{2}$ da normal reduzida tomando

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{e},$$

temos

$$pr\Big(- z_q\sigma \leq \frac{\|\boldsymbol{y}\|^2 - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|} \leq z_q\sigma\Big) = 1 - q,$$

o que é equivalente a ter-se

$$pr\Big(-2z_q\sigma\|\boldsymbol{\mu}\| \leq \|\boldsymbol{y}\|^2 - \|\boldsymbol{\mu}\|^2 \leq 2z_q\sigma\|\boldsymbol{\mu}\|\Big) = 1 - q.$$

Ora, podemos utilizar este par de desigualdades para obter limites para

$$\Upsilon = \|\boldsymbol{\mu}\|^2.$$

Observe-se ainda que o par de desigualdades pode ser reescrito como

$$pr\Big((\|\boldsymbol{y}\|^2 - \Upsilon)^2 \leq 4z_q^2\sigma^2 x\Big) = 1 - q.$$

Como

$$pr\big(\|\boldsymbol{y}\|^4 - 2\|\boldsymbol{y}\|^2\Upsilon - 4z_q^2\sigma^2\Upsilon + \Upsilon^2 \leq 0\big) = 1 - q,$$

e como

$$\begin{aligned} pr\Big[\Big(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2 - \sqrt{(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2)^2 - \|\boldsymbol{y}\|^4}\Big) &\leq \Upsilon \\ &\leq \Big(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2 + \sqrt{(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2)^2 - \|\boldsymbol{y}\|^4}\Big)\Big] \\ &= 1 - q. \end{aligned}$$

Este último par de desigualdades pode ser simplificado para

$$\begin{aligned} pr\Big[\Big(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2 - 2z_q\sigma^2\sqrt{\|\boldsymbol{y}\|^2 + z_q^2\sigma^2}\Big) &\leq \|\boldsymbol{\mu}\|^2 \\ &\leq \Big(\|\boldsymbol{y}\|^2 + 2z_q^2\sigma^2 + 2z_q\sigma^2\sqrt{\|\boldsymbol{\mu}\|^2 + z_q^2\sigma^4}\Big)\Big] \\ &= 1 - q. \end{aligned}$$

Se $\sigma^2 = 1$ ter-se-ia $||\mu||^2 = \delta$, obtendo-se os limites pretendidos para $\|\boldsymbol{\mu}\|^2$. Analogamente, para $\|\boldsymbol{\mu}\|^2/\|\boldsymbol{y}\|^2$, tem-se

$$
\begin{aligned}
pr\Bigg[1+2z_q^2\frac{\sigma^2}{\|\boldsymbol{y}\|^2}-2z_q\frac{\sigma^2}{\|\boldsymbol{y}\|^2}\sqrt{1+\frac{z_q^2\sigma^2}{\|\boldsymbol{y}\|^2}} &\leq \frac{\|\boldsymbol{\mu}\|^2}{\|\boldsymbol{y}\|^2}\\
&\leq 1+2z_q^2\frac{\sigma^2}{\|\boldsymbol{y}\|^2}+2z_q\frac{\sigma^2}{\|\boldsymbol{y}\|^2}\sqrt{1+\frac{z_q^2\sigma^2}{\|\boldsymbol{y}\|^2}}\Bigg]\\
&= 1-q.
\end{aligned}
$$

Ora, em muitos casos, $\sigma/\|\boldsymbol{y}\|$ é suficientemente pequeno para que o erro relativo ε_r, resultante de se utilizar $\|\boldsymbol{y}\|^2$ como estimador de $\|\boldsymbol{\mu}\|^2$ poder ser desprezado. Observe-se que os limites obtidos para $\|\boldsymbol{\mu}\|^2$ e para $\|\boldsymbol{\mu}\|^2/\|\boldsymbol{y}\|^2$ dependem de σ^2 pelo que para os utilizarmos temos de ter informação adicional sobre σ^2. Admitimos agora que $\boldsymbol{y}$ é independente de $S \sim \sigma^2\chi_g^2$.

Como

$$
\frac{\|\boldsymbol{y}\|}{\|\boldsymbol{\mu}\|} \xrightarrow[u\to\infty]{p} 1,
$$

e como a distribuição limite de

$$
Z^\circ = \frac{\|\boldsymbol{y}\|^2}{2\|\boldsymbol{\mu}\|} - \frac{\|\boldsymbol{\mu}\|}{2},
$$

quando $\|\boldsymbol{\mu}\| \to \infty$, é $\aleph(0,\sigma^2)$ a distribuição limite de

$$
Z^{\circ\circ} = \frac{\|\boldsymbol{y}\| - \|\boldsymbol{\mu}\|}{2},
$$

quando $\|\boldsymbol{\mu}\| \to \infty$, também é $\aleph(0,\sigma^2)$. Então a distribuição limite de

$$
t^{\circ\circ} = \frac{\|\boldsymbol{y}\| - \|\boldsymbol{\mu}\|}{2\sqrt{\frac{s}{2}}},
$$

quando $\|\boldsymbol{\mu}\| \to \infty$, é a distribuição t central com g graus de liberdade, Mexia (1992). Seja $t_{g,q}$, o limite para as caudas com probabilidade q' dessa distribuição. Temos

$$
pr\Bigg(-t_{g,q} \leq \frac{\|\boldsymbol{y}\| - \|\boldsymbol{\mu}\|}{2\sqrt{\frac{s}{g}}} \leq t_{g,q}\Bigg) \xrightarrow[u\to\infty]{} 1-q,
$$

logo,

$$pr\left[-2t_{g,q}\sqrt{\frac{s}{g}} \leq \|\boldsymbol{y}\| - \|\boldsymbol{\mu}\| \leq 2t_{g,q}\sqrt{\frac{s}{g}}\right] \xrightarrow[u\to\infty]{} 1-q$$

vindo

$$pr\left(-\|\boldsymbol{y}\| - 2t_{g,q}\sqrt{\frac{s}{g}} \leq \|\boldsymbol{\mu}\| \leq -\|\boldsymbol{y}\| + 2t_{g,q}\sqrt{\frac{s}{g}}\right) \xrightarrow[u\to\infty]{} 1-q,$$

bem como

$$pr\left[\left(\|\boldsymbol{y}\| - 2t_{g,q}\sqrt{\frac{s}{g}}\right)^2 \leq \|\boldsymbol{\mu}\|^2 \leq \left(\|\boldsymbol{y}\| + 2t_{g,q}\sqrt{\frac{s}{g}}\right)^2\right] \xrightarrow[u\to\infty]{} 1-q$$

e como

$$pr\left[\left(1 - 2t_{g,q}\sqrt{\frac{s}{g\|\boldsymbol{y}\|^2}}\right)^2 \leq \frac{\|\boldsymbol{\mu}\|^2}{\|\boldsymbol{y}\|^2} \leq \big(1 + 2t_{g,q}\sqrt{\frac{s}{g\|\boldsymbol{y}\|^2}}\right)^2\right] \xrightarrow[u\to\infty]{} 1-q.$$

Estes intervalos de confiança não dependem de σ^2. Vê-se ainda que quando $\frac{s}{\|\boldsymbol{y}\|^2}$ é pequeno o erro relativo resultante de se tomar $\|\boldsymbol{y}\|^2$ como estimador de $\|\boldsymbol{\lambda}\|^2$ é desprezível.

Chegando ao ponto central desta alínea, observamos que para σ^2 e $1/\sigma^2$ temos os intervalos de confiança de nível $1-q'$ dados por

$$\begin{cases} \left[\dfrac{S}{\chi_{g,1-\frac{q'}{2}}};\quad \dfrac{S}{\S_{g,\frac{q'}{2}}}\right] \\ \\ \left[\dfrac{\chi_{g,\frac{q'}{2}}}{S};\quad \dfrac{\chi_{g,1-\frac{q'}{2}}}{S}\right]. \end{cases}$$

Ora, quando ambos os pares de desigualdades

$$\begin{cases} \left(\|\boldsymbol{y}\| - 2t_{g,q}\sqrt{\frac{S}{g}}\right)^2 \leq \ \|\boldsymbol{\mu}\|^2 \leq \left(\|\boldsymbol{y} + 2t_{g,q}\sqrt{\frac{s}{g}}\right)^2 \\ \\ \qquad\qquad \frac{\chi}{g,\frac{q'}{2}}S \leq \qquad \frac{1}{\sigma^2} \leq \frac{\chi}{g,1-\frac{q'}{2}S} \end{cases}$$

se verificam, temos

$$\left(\|\boldsymbol{y}\| - 2t_{g,q}\sqrt{\frac{S}{g}}\right)^2 \frac{\chi_{g,\frac{q'}{2}}}{S} \leq \delta \leq \left(\|\boldsymbol{y}\| + 2t_{g,q}\sqrt{\frac{S}{g}}\right)^2 \frac{\chi_{g,1-\frac{q'}{2}}}{S}.$$

Logo atendendo à desigualdade de Boole, temos

$$\begin{aligned} pr\left[\left(\|\boldsymbol{y}\| - 2t_{g,q}\sqrt{\frac{S}{2}}\right)\frac{\chi_{g,\frac{q'}{2}}}{S}\right. &\leq \delta \\ &\leq \left.\left(\|\boldsymbol{y}\| + 2t_{g,q}\sqrt{\frac{S}{g}}\right)^2 \frac{\chi_{g,1-\frac{q'}{2}}}{S}\right] \\ &\geq 1 - q - q'. \end{aligned}$$

Na prática torna-se $q < q\prime$ visto que a precisão da estimação de $\|\boldsymbol{\mu}\|^2$ usando $\|\boldsymbol{y}\|^2$ ser muito maior que a de estimar σ^2 usando s/g.

Uma aplicação à preponderância

Recordamos agora que, atendendo á **Proposição 2.4**, $F(\mathbf{P}^+)$ tem distribuição F com k e ℓ graus de liberdade e parâmetro de não centralidade

$$\delta = \tau.$$

Assim, o intervalo de confiança para δ pode ser reinterpretado como intervalo de confiança para τ. Através da dualidade podemos a partir destes intervalos testar a hipótese

$$H_0 : \tau = \tau_0.$$

3.5 Caso multivariado

Apresentamos de seguida os resultados para a linearidade assimptótica no caso multivariado.

Trabalhamos com estatísticas vectoriais com n componentes

$$Y_i = g_i(\boldsymbol{\mu} + \boldsymbol{e}),\ i = 1, \ldots, n.$$

A passagem do caso univariado para o caso multivariado é facilitada pela

Proposição 3.2se $F(\cdot)$ tiver distribuições marginais $F_i(\cdot),\ i = 1, \ldots, n$, com $\boldsymbol{a} \leq \boldsymbol{b}$, isto é se $a_i \leq b_i,\ i = 1, \ldots, n$ tem-se

$$F(\boldsymbol{b}) - F(\boldsymbol{a}) \leq \sum_{i=1}^{n} (F_i(b_i) - F_i(a_i)).$$

Demonstração: Com $\boldsymbol{X}$ com componentes $X_i, i = 1, \ldots, n$, e distribuição F, termos $\boldsymbol{X} \leq \boldsymbol{a}$ implicará $\boldsymbol{X} \leq \boldsymbol{b}$, vindo

$$\begin{aligned} F(\boldsymbol{b}) - F(\boldsymbol{a}) &= pr(\boldsymbol{X} \leq \boldsymbol{b}) - pr(\boldsymbol{X} \leq \boldsymbol{a}) \\ &= pr[(\boldsymbol{X} \leq \boldsymbol{b}) \setminus (\boldsymbol{X} \leq \boldsymbol{a})] \\ &= pr[(\boldsymbol{X} \leq \boldsymbol{b}) \cap (\boldsymbol{X} \leq \boldsymbol{a})^c] \end{aligned}$$

onde $\setminus$ representa a diferença de acontecimentos e $(\boldsymbol{X} \leq \boldsymbol{a}_i)^c$, o acontecimento complementar de $(\boldsymbol{X} \leq \boldsymbol{a})$. Ora,

$$(\boldsymbol{X} \leq \boldsymbol{a})^c = (\cap_{i=1}^{n} (X_i \leq a_i))^c = \cup_{i=1}^{n} (X_i > a_1)$$

e

$$(\boldsymbol{X} \leq \boldsymbol{b}) \cap (\cup_{i=1}^{n} (X_i > a_i)) = \cup_{i=1}^{n} ((\boldsymbol{X} \leq b_i) \cap (X_i > a_i))$$

$$= \cup_{i=1}^{n} (\cap_{i' \neq i} (X_{i'} \leq b_{i'}) \cap (a_i < X_i \leq b_i)),$$

vindo

$$\begin{aligned} F(\boldsymbol{b}) - F(\boldsymbol{a}) &\leq \sum_{i=1}^{n} pr\Big(\cap_{i' \neq i} (X_{i'} \leq b_{i'}) \cap (a_i < X_i \leq b_i) \Big) \\ &\leq \sum_{i=1}^{n} pr(a_i < X_i \leq b_i) = \sum_{i=1}^{n} \Big(F_i(b_i) - F_i(a_i) \Big), \end{aligned}$$

estando pois a tese estabelecida. Ainda temos uma generalização dos resultados da secção 3.2 que passamos a considerar. Para isso começamos por observar que, caso

$$\|\boldsymbol{u}-\boldsymbol{v}\| \leq d,$$

se tem

$$(1) \quad \boldsymbol{v}-\frac{d}{\sqrt{n}}\mathbf{1}_n \leq \boldsymbol{u} \leq \boldsymbol{v}+\frac{d}{\sqrt{n}}\mathbf{1}_n$$

e que, quando

$$Sup\{\mid \frac{1}{\sqrt{n}}\mathbf{1}_n^t \underline{g}_i(\boldsymbol{x}) \mid\} \leq \bar{\underline{g}}_i, \quad i=1,\ldots,n$$

se tem,

$$(2) \quad \mid g_i(\boldsymbol{v}+\frac{d}{\sqrt{n}}1_k)-g_i(\boldsymbol{v}-\frac{d}{\sqrt{n}}1_k) \mid \leq 2\bar{g}_i d, \quad i=1,\ldots,n.$$

Seja $A(\varepsilon)$ o acontecimento que se verifica quando

$$\mid \boldsymbol{Y}^\circ - \boldsymbol{Z}^\circ \mid \leq \varepsilon,$$

com $\boldsymbol{Y}^\circ$ e $\boldsymbol{Z}^\circ$ com componentes

$$\begin{cases} Y_i^\circ & = \frac{g_i(\boldsymbol{\mu}+\boldsymbol{e})-g_i(\boldsymbol{\mu})}{\|g_i(\boldsymbol{\mu})\|}, \quad i=1,\ldots,n. \\ Z_i^\circ & = \boldsymbol{g}_i(\boldsymbol{\mu})^t\boldsymbol{e}, \quad i=1,\ldots,n, \end{cases}$$

e tal, como atrás

$$p(\varepsilon)=pr(A(\varepsilon)).$$

Observe-se que, quando $A(\varepsilon)$ se verifica, se tem com $\Rightarrow$ indicando implicação

$$(\boldsymbol{Z}^\circ \leq \boldsymbol{z}-\frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n) \Rightarrow (\boldsymbol{Y}^\circ \leq \boldsymbol{z}) \Rightarrow (\boldsymbol{Z}^\circ \leq \boldsymbol{z}+\frac{\varepsilon}{\sqrt{n}}1_n),$$

tendo-se pois,

$$\begin{aligned} pr\Big(\boldsymbol{Z}^\circ \leq \boldsymbol{z}-\frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n \setminus A(\varepsilon)\Big) &\leq pr\Big(\boldsymbol{Y}^\circ \leq \boldsymbol{z} \setminus A(\varepsilon)\Big) \\ &\leq pr(\boldsymbol{Z}^\circ \leq \boldsymbol{z}+\frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n \setminus A(\varepsilon)). \end{aligned}$$

Logo

$$
\begin{aligned}
F_{\boldsymbol{Z}^\circ}\Big(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}1_n) - \Big(1 - p(\varepsilon)\Big) &= p(\varepsilon)pr\Big(\boldsymbol{Z}^\circ \le \boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}1_n/A(\varepsilon)\Big) \\
&+ \Big(1 - p(\varepsilon)pr\Big(\boldsymbol{Z}^\circ \le \boldsymbol{z} - \frac{\varepsilon}{\sqrt{k}}1_k/A^c(\varepsilon)\Big) - (1 - p(\varepsilon)\Big) \\
&\le p(\varepsilon)pr\Big(\boldsymbol{Z}^\circ \le \boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}1_n/A(\varepsilon)\Big) \\
&\le p(\varepsilon)pr\Big(\boldsymbol{Y}^\circ \le \boldsymbol{z}/A(\varepsilon)\Big) \\
&\le pr(\varepsilon)pr\Big(\boldsymbol{Y}^\circ \le \boldsymbol{z}/A(\varepsilon)\Big) + \Big(1 - p(\varepsilon\Big)pr\Big(\boldsymbol{Y}^\circ \le \boldsymbol{z}/A^c(\varepsilon)\Big) \\
&= F \quad {}_{\boldsymbol{z}^\circ}(\boldsymbol{z}).
\end{aligned}
$$

Da mesma maneira se mostra que

$$F_{\boldsymbol{Y}^\circ}(\boldsymbol{z}) \le F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{n}}1_n) + p(\varepsilon)$$

estabelendo-se assim a

Proposição 3.3 *Tem-se*

$$F_{\boldsymbol{Z}^\circ}\Big(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n\Big) - (1-p)(\varepsilon) \le F_{\boldsymbol{Y}^\circ}(\boldsymbol{z}) \le F_{\boldsymbol{Z}^\circ}\Big(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n\Big) + \Big(1 - p(\varepsilon)\Big).$$

Corolário 1

Tendo $F_{\boldsymbol{Z}^\circ}$ gradiente $\underline{F}_{\boldsymbol{Z}^\circ}$ tal que

$$\bar{\underline{F}}_{\boldsymbol{Z}^\circ} = Sup\{\mid \frac{1}{\sqrt{n}}\mathbf{1}_n^t \underline{F}_{\boldsymbol{Z}^\circ}(\boldsymbol{z}) \mid\} < +\infty,$$

tem-se

$$Sup\{\mid F_{\boldsymbol{Y}^\circ}(\boldsymbol{z}) - F_{\boldsymbol{Z}^\circ}(\boldsymbol{z}) \mid\} \le d(\varepsilon)$$

com

$$d(\varepsilon) = 2\Big(\bar{\underline{F}}_{\boldsymbol{Z}^\circ}\varepsilon + 1 - p(\varepsilon)\Big).$$

Demonstração:

Do mesmo modo como se obteve(2), mostra-se que

$$\begin{aligned}0 &\leq F_{\boldsymbol{Z}^\circ}\big(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n\big) - F_{\boldsymbol{Z}^\circ}\big(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n\big)\\ &\leq 2\underline{\bar{F}}_{\boldsymbol{Z}^\circ}\varepsilon.\end{aligned}$$

Pelo que

$$\left(F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n) + 1 - p(\varepsilon)\right) - (F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}\mathbf{1}_n) - (1 - p(\varepsilon)) \leq d(\varepsilon).$$

Para completar a demonstração basta observar que

$$\begin{aligned}F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{n}}1_n) &\leq F_{\boldsymbol{Z}^\circ}(\boldsymbol{z})\\ &\leq F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{n}}1_n).\end{aligned}$$

Regressando à estatística $\boldsymbol{Y}$, com componentes $Y_i,\ i = 1, \ldots, n$, admitamos que as funções $g_i(\cdot)$, são assimptoticamente lineares tendo-se, qualquer que seja $d > 0$,

$$k_{d,i}(u) \xrightarrow[u\to\infty]{} 0, \quad i = 1, \ldots, n.$$

De uma forma análoga vê-se que, quaisquer que sejam $d > 0$ e $\varepsilon > 0$,

$$u_i(d, \frac{\varepsilon}{\sqrt{k}}) = Inf\{u : k_{d,i}(u)d^2 \leq \frac{\varepsilon}{\sqrt{k}}\} < +\infty, \quad i = 1, \ldots, n.$$

Seja

$$p(\varepsilon) = pr(A(\varepsilon)),$$

com $A(\varepsilon)$ o acontecimento que se verifica quando

$$\cap_{i=1}^{n}(\mid Y_i^\circ - Z_i^\circ \mid \leq \frac{\varepsilon}{k}).$$

Quando

$$\|\boldsymbol{\mu_i}\| \geq Max\{u_i(\ell_{1-q}, \frac{\varepsilon}{\sqrt{k}}), \quad i = 1, \dots, n\},$$

tem-se $p(\varepsilon) \geq 1 - q$ e como, atendendo á **Proposição 3.3**,

$$\begin{aligned} F_{\boldsymbol{Z}^\circ}(\boldsymbol{z} - \frac{\varepsilon}{\sqrt{k}} 1_k) - q(\varepsilon) &\leq F_{\boldsymbol{Y}^\circ}(\boldsymbol{z}) \\ &\leq F_{\boldsymbol{Y}^\circ}(\boldsymbol{z} + \frac{\varepsilon}{\sqrt{k}} 1_k) + q. \end{aligned}$$

Devido ao **Corolário 1** da **Proposição 3.3** temos igualmente

$$Sup\{\mid F_{\boldsymbol{Y}^\circ}(\boldsymbol{y}) - F_{\boldsymbol{Z}^\circ}(\boldsymbol{z}) \mid\} \leq 2(\underline{\bar{F}}_{\boldsymbol{Z}^\circ}\varepsilon + q),$$

quando

$$\|\boldsymbol{\mu}\| \geq Max\{u_i(\phi_{1-q}, \frac{\varepsilon}{\sqrt{k}}), \quad i = 1, \dots, n\}.$$

Sejam $\mathbf{G}(\boldsymbol{\mu})$ e $\mathbf{B}$ as matrizes com vectores linha $\boldsymbol{g}_1(\boldsymbol{\mu}), \dots, \boldsymbol{g}_n(\boldsymbol{\mu})$ e $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$. Caso

$$\boldsymbol{g}_i(\boldsymbol{\mu}) \xrightarrow[\|\boldsymbol{\mu}\| \to \infty]{} \boldsymbol{b}_i, \quad i = 1, \dots, n,$$

$F_{\boldsymbol{Y}^\circ}$ convergirá para $F_{\mathring{\boldsymbol{Z}}}$ com

$$\mathring{\boldsymbol{Z}} = B\boldsymbol{e}.$$

Proposição 3.4 Se as funções $g_i(\cdot), \; i = 1, \dots, n$ forem assimptoticamente lineares e se

$$\frac{1}{\|\boldsymbol{g}_i(\boldsymbol{\mu})\|} \boldsymbol{g}_i(\boldsymbol{\mu}) \xrightarrow[\|\boldsymbol{\mu}\| \to \infty]{} \boldsymbol{b}_i, \quad i = 1, \dots, n,$$

qualquer que seja a função continua $\ell(\cdot)$, $F_{\ell(\boldsymbol{Y}^\circ)}$ converge para $F_{\ell(\mathring{Z})}$.

Demonstração

As funções caracteristicas de $\ell(\boldsymbol{Y}^{\circ})$ e $\ell(\boldsymbol{Z}^{\circ})$ são

$$\begin{cases} \varphi_{\ell(\boldsymbol{Y}^{\circ})}(u) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} e^{iu\ell(y)} dF_{\boldsymbol{Y}^{\circ}} \\ \qquad\qquad = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \cos(u\ell(\boldsymbol{y}))dF_{\boldsymbol{Y}^{\circ}} + i \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \sin(u\ell(y))dF_{\mathring{\boldsymbol{Z}}} \\ \varphi_{\ell(\boldsymbol{Z}^{\circ})}(u) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} e^{iu\ell(\boldsymbol{y})} dF_{\mathring{\boldsymbol{Z}}} \\ \qquad\qquad = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \cos(u\ell(\boldsymbol{y}))dF_{\mathring{\boldsymbol{Z}}+i \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \sin(u\ell(z))dF_{\mathring{\boldsymbol{Z}}}}. \end{cases}$$

Ora, atendendo ao lema de Helley Brey tem-se

$$\begin{cases} \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \cos(u\ell(\boldsymbol{y}))dF_{\boldsymbol{y}^{\circ}} \underset{\|\boldsymbol{\mu}\|\rightarrow\infty}{\longrightarrow} \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \cos(u\ell(y))dF_{\mathring{\boldsymbol{Z}}} \\ \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \sin(u\ell(\boldsymbol{y}))dF_{\boldsymbol{y}^{\circ}} \underset{\|\boldsymbol{\mu}\|\rightarrow\infty}{\longrightarrow} \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \sin(u\ell(\boldsymbol{y}))dF_{\mathring{\boldsymbol{Z}}} \end{cases},$$

o que estabelece a tese visto que a convergência das funções características implica a convergência das distribuições, ver por exemplo Lukacs e Laha(1964).

Capítulo 4

Delineamento base ortogonais

4.1 Considerações prévias

Os delineamentos ortogonais com n tratamentos estão associados a partições ortogonais

$$R^n = \boxplus_{j=1}^{m} \bar{w}_j,$$

Na secção seguinte estudamos a estrutura algébrica deste delineamento realçando o papel de matrizes $\mathbf{A}_1, \ldots, \mathbf{A}_m$ cujos vectores linha constituem bases ortonormados para os subespaços $\bar{w}_1, \ldots, \bar{w}_m$. Consideramos, em particular o cruzamento de factores os quais conduzem a um delineamento em que os tratamentos são todas as combinações possíveis de níveis. Mostraremos como construir as matrizes $\mathbf{A}_1, \ldots, \mathbf{A}_m$ para esses delineamento.

Um delineamento ortogonal é equilibrado se tem o mesmo número "r"de observações por tratamentos. Tomamos

$$r = 1$$

já que vamos utilizar a Análise de Variância sem restrições em que se trabalha com um par$(\boldsymbol{Y}, V)$ em que se tem

$$\boldsymbol{Y} \sim \aleph(\boldsymbol{\mu}, \sigma^2 C)$$

independente de

$$V \sim \sigma^2 \chi_g^2,$$

admitindo-se que $\mathbf{C}$ é semi - definida positiva, Mexia (1989 a). Apresentaremos este tipo de ANOVA na secção 4.3. Observe-se que a mesma permite testar, em vez das hipóteses usuais

$$H_{0,j} : \boldsymbol{\mu} \in w_j, \quad j = 1, \ldots, m,$$

as respectivas generalizações

$$H_{0,j}(\boldsymbol{b}) : \boldsymbol{\mu} - \boldsymbol{b} \in w_j \quad j = 1, \ldots, m,$$

vindo os $w_1, \ldots, w_m$ os complementos ortogonais $\bar{w}_1, \ldots, \bar{w}_m$, pondo-se $w_j = \bar{w}_j^{\perp}, \quad j = 1, \ldots, m.$

Esta generalização faz surgir, como veremos, propriedades de dualidade para os testes F. Assim, com

$$\begin{cases} \tilde{\boldsymbol{\eta}}_j & = \mathbf{A}_j \boldsymbol{Y}, \quad j = 1, \ldots, m \\ \boldsymbol{\eta}_j & = \mathbf{A}_j \mu, \quad j = 1, \ldots, m \\ \boldsymbol{\eta}_{j,0} & = \mathbf{A}_j b, \quad j = 1, \ldots, m \end{cases}$$

será possível construir hiperesferas de confiança para os $\boldsymbol{\eta}_j, j = 1, \ldots, m$ e o teste de nível q rejeita $H_{0,j}(\boldsymbol{b})$ se e só se a hiperesfera de confiança de nível $1-q$ não contiver $\boldsymbol{\eta}_{j,0}, j = 1, \ldots, m$. Assim, temos não só um alargamento da família das hipóteses testadas como um enriquecimento dos testes F.

A secção 4.4 representa um desenvolvimento da anterior com vista a aplicação futura a famílias estruturadas do modelo. Em vez do par $(\boldsymbol{Y}, V)$ temos nos modelo múltiplos, $(\boldsymbol{Y}(1), \ldots, \boldsymbol{Y}(\ell), V)$ com

$$\begin{cases} \boldsymbol{Y}(i) \sim & \aleph(\boldsymbol{\mu}(i), \sigma^2 \mathbf{I}_n), \quad i = 1, \ldots, u \\ \\ V \sim & \sigma^2 \chi_g^2, \end{cases}$$

sendo os $\boldsymbol{Y}(1), \ldots, \boldsymbol{Y}(\ell)$ e V independentes. Para compreender a ligação entre estes modelos múltiplos e as famílias estruturadas convém recordar que a informação contida nos modelos dessas famílias é condensada em pares$(\tilde{\boldsymbol{\beta}}_i, V_i), \quad i = 1, \ldots, n$. Podemos agora pensar em tomar

$$\begin{cases} Y_i(d) & = \boldsymbol{c}_d{}^t \tilde{\boldsymbol{\beta}}_i, \quad i = 1, \ldots, n, \quad d = 1, \ldots, \ell \\ \\ V = & \displaystyle\sum_{i=1}^{n} V_i. \end{cases}$$

Segue-se a secção 4.5 em que se aplicam os resultados sobre ANOVA sem restrições a famílias estruturadas de matrizes quase escalares. A informação contida nessas matrizes é condensada em pares $(\tilde{\boldsymbol{\beta}}_i, V_i), \quad i = 1, \ldots, n$ o que permite proceder como foi indicado atrás para obter o vector $\boldsymbol{Y}$ os vectores $\boldsymbol{Y}(1), \ldots, \boldsymbol{Y}(n)$ e uma soma de quadrados de resíduos. Neste trabalho, consideramos um único vector $\boldsymbol{Y}$ deixando os modelos múltiplos para desenvolvimentos futuros. Consideramos para além dessas famílias divididas em subfamílias que correspondem aos tratamentos. Na aplicação que demos às eleições autárquicas de Portugal Continental os tratamentos correspondem a distritos havendo um modelo para cada concelho desses distritos. Assim os modelos agrupam-se em subfamílias, uma por distritos. Como veremos os nossos modelos podem ser aplicados às duas primeiras etapas da metodologia Statis. Temos assim uma subfamílias por distrito e etapa. Veremos como medir a maior ou menor heterogeneidade dessas famílias mostrando-se que a heterogeneidade das subfamílias é transportada de primeira para a segunda etapa.

4.2 Estrutura algébrica

Como referimos, os modelos ortogonais estão associados a partições ortogonais,

$$R^n = \boxplus_{j=1}^{m} \bar{\mathbf{w}}_j$$

sendo n o número de tratamentos. Ponhamos

$$\mathbf{A}_j \longleftrightarrow \bar{\mathbf{w}}_j, \quad j = 1, \ldots, m$$

para indicar que vectores linha de $\mathbf{A}_j$ constituem uma base ortonormada para $\bar{\mathbf{w}}_j$, $j = 1, \ldots, m$.

Teremos então

$$\mathbf{A}_j^t \mathbf{A}_j = I_{g_j}, \quad j = 1, \ldots, m.$$

com

$$g_j = dim(\bar{\mathbf{w}}_j) = car(\mathbf{A}_j), \ j = 1, \ldots, m,$$

sendo a matriz de projecção ortogonal $\mathbf{Q}_j$ sobre $\bar{\mathbf{w}}_j$,$j = 1, \ldots, m$, dada por

$$\mathbf{Q}_j = \mathbf{A}_j^t \mathbf{A}_j, \quad j = 1, \ldots, m.$$

Vê-se ainda que o espaço nulidade $\aleph(\mathbf{A}_j)$ da matriz $\mathbf{A}_j, j = 1, \ldots, m$, é dado por,

$$\aleph(\mathbf{A}_j) = \mathbf{w}_j = \bar{\mathbf{w}}_j^{\perp}, j = 1, \ldots, m.$$

Logo tem-se

$$\boldsymbol{v} - \boldsymbol{b} \in \bar{\mathbf{w}}_j, j = 1, \ldots, m$$

se e só se

$$\mathbf{A}_j\boldsymbol{b} = \mathbf{A}_j\boldsymbol{v}, j = 1, \ldots, m,$$

Em particular

$$\boldsymbol{v} \in \mathbf{w}_j, j = 1, \ldots, m$$

quando e apenas quando

$$\mathbf{A}_j\boldsymbol{v} = \boldsymbol{o},\ \ j = 1, \ldots, m.$$

Como veremos estes resultados algébricos permitem sistematizar a ANOVA sem restrições para estes modelos.

4.2.1 Cruzamento equilibrado de factores

Admitamos que se têm u factores com $a_1, \ldots, a_u$, níveis e que se considerem todas as combinações possíveis dos mesmos. Estas combinações são os tratamentos, havendo pois

$$n = \pi_{i=1}^{u} a_j$$

tratamentos. Além do valor médio geral, temos de considerar os efeitos dos níveis dos vários factores e as interacções para as várias combinações de níveis dos conjuntos de mais de um factores. Identificando os factores com os seus índices os conjuntos de factores serão os sub-conjuntos de

$$\bar{\bar{u}} = \{1, \ldots, u\}.$$

Dado $\varphi \subseteq \bar{\bar{u}}$, se $\#(\varphi) = 0$, temos o conjunto vazio correspondendo-lhe o valor médio geral; se $\#(\varphi) = 1$ corresponder-lhe-ão os efeitos dos niveis do unico factor com indice em φ. Finalmente, se $\#(\varphi) > 1$, a φ corresponderão as interações entre conjunto de níveis dos factores em φ. Temos agora a partição ortogonal

$$R^n = \boxplus_{\varphi \subseteq \bar{\bar{u}}} \bar{\mathbf{w}}(\varphi),$$

onde

$$\bar{\boldsymbol{W}}(\varphi) = \Re(A(\varphi))^t, \quad \varphi \subseteq \bar{\bar{u}},$$

ver Fonseca et al. (2003). Para construir as matrizes $A(\varphi)$ pode recorrer-se no produto de Kronecker $\otimes$ de matrizes. Sendo $\mathbf{C} = [c_{ij}]$ uma matriz $r \times s$ tem-se qualquer que seja a matriz $\mathbf{B}$,

$$\mathbf{C} \otimes B = \begin{bmatrix} c_{11}B \ldots c_{1s}B \\ \vdots \qquad\quad \vdots \\ c_{r1}B \ldots c_{rs}B. \end{bmatrix} \tag{4.2.0}$$

Em particular temos

$$\mathbf{A}(\varphi) = \otimes_{l=1}^{u} \mathbf{A}_l(\varphi)$$

onde

$$\begin{cases} \mathbf{A}_\ell(\varphi) = \frac{1}{\sqrt{a_\ell}} \mathbf{1}_{a_\ell}^t, \quad \ell \notin \varphi \\ \\ \mathbf{A}_\ell(\varphi) = T_{a_\ell}, \quad \ell \in \varphi. \end{cases}$$

sendo T_{a_l} obtida retirando a primeira linha igual a $\dfrac{1}{\sqrt{a_\ell}} \mathbf{1}_{a_\ell}^t$ a uma matriz ortogonal $a_\ell \times a_\ell$. Por exemplo, pode tomar-se

$$\begin{cases} \mathbf{T}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ \\ \mathbf{T}_3 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}}. \end{bmatrix} \end{cases}$$

4.3 ANOVA sem restrições

Nestes modelos temos um vector

$$\boldsymbol{Y} \sim \aleph(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$$

independente de

$$V \sim \sigma^2 \chi^2_g$$

e uma partição ortogonal

$$\Re^n = \boxplus_{j=1}^m \boldsymbol{w}_j.$$

Usualmente interessará testar as hipóteses

$$H_{0j} : \boldsymbol{\mu} \epsilon \boldsymbol{w}_j = \bar{\boldsymbol{w}}_j^{\perp},\ j = 1, \ldots, m$$

no entanto, podemos considerar uma generalização. Passamos então a testar-se as hipóteses

$$H_{0j}(\boldsymbol{b}) : \boldsymbol{\mu} - \boldsymbol{b} \quad \epsilon w_j \quad , j = 1, \ldots, m.$$

Sendo

$$\mathbf{Q}_j = \mathbf{A}_j^t \mathbf{A}_j, \quad j = 1, \ldots, m$$

as matrizes de projecção ortogonal sobre os $\bar{\boldsymbol{w}}_j, \quad j = 1, \ldots, m$, tem-se

$$\mathbf{A}_j \mathbf{A}_j^t = \mathbf{I}_{g_j}, \quad j = 1, \ldots, m$$

com

$$g_j = dim(\bar{\boldsymbol{w}}_j), \quad j = 1, \ldots, m$$

Recorde-se que os vectores linha de $\mathbf{A}_j$, que são vectores coluna de $\mathbf{A}_j^t$, constituiem uma base ortonormada para $\bar{\bar{w}}_j, \quad j = 1, \ldots, m.$

Então com

$$\begin{cases} \tilde{\boldsymbol{\eta}}_j & = \mathbf{A}_j \boldsymbol{Y}, \quad j = 1, \ldots, m \\ \boldsymbol{\eta}_j & = \mathbf{A}_j \mu, \quad j = 1, \ldots, m \\ \boldsymbol{\eta}_{j,0} & = \mathbf{A}_j b, \quad j = 1, \ldots, m \end{cases}$$

temos

$$\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0} \sim \aleph(\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}, \sigma^2 \mathbf{I}_{g_j}), \quad j = 1, \dots, m,$$

bem como

$$\|\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}\|^2 \sim \sigma^2 \chi^2_{g_j, \delta_j}(\boldsymbol{b}), \quad j = 1, \dots, m,$$

com

$$\delta_j(\boldsymbol{b}) = \frac{1}{\sigma^2} \|\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}\|^2, \quad j = 1, \dots, m$$

independente de V pelo que as estatisticas

$$\mathcal{F}_j(\boldsymbol{b}) = \frac{g}{g_j} \frac{\|\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}\|^2}{V}, \quad j = 1, \dots, m,$$

terão distribuições F com g_j e g graus de liberdade e parâmetros de não centralidade $\delta_j(\boldsymbol{b}), j = 1, \dots, m$, tendo-se pois

$$\mathcal{F}_j(\boldsymbol{b}) \sim F(|g_j, g, \delta(\boldsymbol{b})), \quad j = 1, \dots, m.$$

Como a hipótese $H_{0j}(\boldsymbol{b})$ pode ser reescrita como

$$H_{0j}(\boldsymbol{b}) : \delta_j(\boldsymbol{b}) = 0, \quad j = 1, \dots, m,$$

estes testes F serão estritamente não distorcidos, Mexia(1995). Por outro lado, temos as variáveis pivot

$$\mathcal{F}'_j = \frac{g}{g_j} \frac{\|\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_j\|^2}{V}, \quad j = 1, \dots, m$$

com distribuições F centrais com g_j e g graus de liberdade, $j = 1, \dots, m$,

$$F'_j \sim \mathcal{F}(/g_j, g), \quad j = 1, \dots, m.$$

Sendo $f_{p,g_j,g}$ o p-éssimo quantil de $F(|g_j, g), \quad j = 1, \ldots, m$, teremos as hiperesferas de confiança de nivel $1 - q$ para os $\boldsymbol{\eta}_j, \quad j = 1, \ldots, m$, dadas por

$$\|\boldsymbol{\eta}_j - \tilde{\boldsymbol{\eta}}_j\|^2 \leq g_j F_{1-q,g_j,g} \quad \frac{V}{g}, \quad j = 1, \ldots, m$$

já que estas desigualdades são equivalentes às

$$\mathcal{F}\prime \leq f_{1-q,g,g} \quad , j = 1, \ldots, m.$$

Observe-se que estes testes F gozam da dualidade já que o teste de nivel q rejeita $H_{0j}(\underline{b})$ quando e só quando $\boldsymbol{\eta}_{j0}(\boldsymbol{b})$ não está contida na correspondente hiperesfera de confiança de nivel $1 - q$, $j = 1, \ldots, m$.

A completar o estudo da dualidade vamos obter uma versão conveniente do teorema de Scheffé, ver Scheffé (1959) e Mexia(1989 b). Assim, um ponto está no interior dum elipsóide se e só se estiver entre todos os pares de planos paralelos tangentes ao elipsóide, havendo, para toda a direcção, um par desses planos perpendiculares a essa direcção e tangente ao elipsóide. Dado o elipsóide definido por

$$(\boldsymbol{x} - \boldsymbol{a})\mathbf{M}(\boldsymbol{x} - \boldsymbol{a}) \leq d,$$

onde $\mathbf{M}$ é uma matriz $k \times k$, semi-definida positiva e $d > 0$, o ponto $\boldsymbol{b}$ está entre o par dos planos tangentes ao elipsóide e perpendiculares a $\boldsymbol{c}$, se e só se

$$| \boldsymbol{c}^t\boldsymbol{b} - \boldsymbol{c}^t\boldsymbol{a} | \leq \sqrt{d\boldsymbol{c}^t\mathbf{M}\boldsymbol{c}},$$

ver Scheffé (1959). Assim, $\boldsymbol{b}$ estará contido no elipsóide se e só se

$$\bigcap_{\boldsymbol{c}} \left(| \boldsymbol{c}^t\boldsymbol{b} - \boldsymbol{c}^t\boldsymbol{a} | \leq \sqrt{d\boldsymbol{c}^t\mathbf{M}\boldsymbol{c}} \right),$$

com $\bigcap_{\boldsymbol{c}}$ a indicar que se consideram todos os vectores de R^k. Observa-se que, para $\boldsymbol{c} = \mathbf{0}$, se tem a desigualdade trivial, $0 \leq 0$.

Aplicando esta discussão às hiperesferas de confiança obtém-se a versão do teorema de Scheffé dada por

$$pr[\bigcap_{\boldsymbol{c}} \mid \boldsymbol{c}^t\boldsymbol{\eta}_j - \boldsymbol{c}^t\tilde{\boldsymbol{\eta}}_j \mid \leq \sqrt{g_j, F_{1-q,g_j,g}\|\boldsymbol{c}\|^2\frac{V}{g})}] = 1-q, \quad ; j = 2,\ldots,m.$$

Convém observar que esta propriedade de dualidade surgiu ao alargar-se a família das hipóteses testadas. Interessa assim aprofundar a natureza desta extensão para o que seguimos a abordagem apresentada em Mexia(1989 a). Assim definimos em Ω uma relação de congruência $\varrho_j \quad j = 1,\ldots,m$ escrevendo

$$\boldsymbol{v}\varrho_j\boldsymbol{u}, \quad j = 1,\ldots,m$$

quando e só quando

$$\boldsymbol{v} - \boldsymbol{u}\epsilon w_j, \quad j = 1,\ldots,m.$$

Logo $H_{\circ j}(\boldsymbol{b})$ verifica-se se e só se

$$\boldsymbol{\mu}\varrho_j\boldsymbol{b}, \quad j = 1,\ldots,m,$$

vendo-se que $\boldsymbol{b}$ determina a classe de congruência $\varrho_j, j = 1,\ldots,m$, associada a $H_{\circ j}(\boldsymbol{b})$. Sendo $[\boldsymbol{u}]_{\varrho_j}$ a classe de congruência $\varrho_j, \quad j = 1,\ldots,n$, a que $\boldsymbol{u}$ pertence as hipóteses podem ser reescritas como

$$H_{\circ j}(\boldsymbol{b}) : \boldsymbol{\mu}\epsilon[\boldsymbol{b}]_{\varrho_j}, \quad j = 1,\ldots,m,$$

tendo-se em particular

$$H_{\circ j}(\boldsymbol{0}) : \boldsymbol{\mu}\epsilon w_j, \quad j = 1,\ldots,m,$$

já que $w = [0]_{\ell_j} \quad j = 1,\ldots,m.$

Para aplicar os resultados precedentes a um qualquer caso basta construir as matrizes $\mathbf{A}$. Por exemplo com três factores a 3, 2 e 3 níveis que se cruzam, temos

$$\mathbf{A}(\phi) = \frac{1}{\sqrt{3}}\mathbf{1}_3^t \otimes \frac{1}{\sqrt{2}}\mathbf{1}_2^t \otimes \frac{1}{\sqrt{3}}\mathbf{1}_3^t$$

$$\mathbf{A}(\{1\}) = T_3 \otimes \frac{1}{\sqrt{2}}\mathbf{1}_2^t \otimes \frac{1}{\sqrt{3}}\mathbf{1}_3^t$$

$$\mathbf{A}(\{2\}) = \frac{1}{\sqrt{3}}\mathbf{1}_2^t \otimes T_2 \otimes \frac{1}{\sqrt{3}}\mathbf{1}_3^t$$

$$\mathbf{A}(\{1,2\}) = T_3 \otimes T_2 \otimes V\frac{1}{\sqrt{3}}\mathbf{1}_2^t$$

$$\mathbf{A}(\{3\}) = V\frac{1}{\sqrt{3}}\mathbf{1}_2^t \otimes V\frac{1}{\sqrt{2}}\mathbf{1}_2^t \otimes T_3$$

$$\mathbf{A}(1,3) = T_3 \otimes \frac{1}{\sqrt{2}}\mathbf{1}_2^t \otimes T_3$$

$$\mathbf{A}(2,3) = \frac{1}{\sqrt{3}}\mathbf{1}_2^t \otimes T_2 \otimes T3$$

$$\mathbf{A}(1,2,3) = T_3 \otimes T_2 \otimes T_3,$$

o que permite aplicar directamente os resultados anteriores.

4.4 Modelos múltiplos sem restrições

Em vez dum par $(\mathbf{Y}, V)$ temos agora $(\mathbf{Y}(1), \dots, \boldsymbol{Y}(u), V)$ com

$$\boldsymbol{Y}(\ell) \sim \aleph(\boldsymbol{\mu}_{(\ell)}, \sigma^2\mathbf{I}_n), \quad l = 1, \dots, u$$

e $V \sim \sigma^2\chi_g^2$, sendo os $\mathbf{Y}(1), \dots, \mathbf{Y}(u)$ e V independentes. Consideramos a mesma partição

$$R^n = \boxplus_{j=1}^m \bar{\mathbf{w}}_j,$$

para todos os $\boldsymbol{Y}(\ell), \ell = 1, \dots, u$, tendo-se para os mesmos as hipóteses de teste

$$H_{0j}(\boldsymbol{b}_l) : \boldsymbol{\mu}_\ell - \boldsymbol{b}_l \epsilon w_j, \quad j = 1, \dots, m, \quad \ell = 1, \dots, u.$$

Dado $D \subseteq \bar{\bar{u}} = \{1, \ldots, u\}$, ponhamos

$$\underline{\boldsymbol{b}}(D) = \{\boldsymbol{b}_j; \quad j \epsilon D\}$$

o que nos permite definir as hipóteses

$$H_{0j}(\underline{\boldsymbol{b}}(D)) : (\boldsymbol{\mu}_\ell - \boldsymbol{b}_\ell \epsilon w_j; \quad l \epsilon D),$$

isto é $H_{0j}(\underline{\boldsymbol{b}}(D))$ verifica-se se e só se verificar em as $H_{0j}(\boldsymbol{b}_l)$ com $\ell \epsilon D, \ j = 1, \ldots, m$. Teremos agora os vectores

$$\begin{cases} \tilde{\boldsymbol{\eta}}_j(\ell) & = \mathbf{A}_j \boldsymbol{Y}(\ell), \quad \ell = 1, \ldots, u, j = 1, \ldots, m \\ \tilde{\boldsymbol{\eta}}_j(\ell) & = \mathbf{A}_j \boldsymbol{\mu}(\ell), \quad \ell = 1, \ldots, u, j = 1, \ldots, m \\ \boldsymbol{\eta}_{j0}(\boldsymbol{b}_\ell) & = \mathbf{A}_j \boldsymbol{b}_\ell, \quad \ell = 1, \ldots, u, j = 1, \ldots, m, \end{cases}$$

sendo os $\tilde{\boldsymbol{\eta}}_j(1), \ldots, \tilde{\boldsymbol{\eta}}_{j(}u)$ independentes entre si e de V. Assim, atendendo á reprodutividade dos qui-quadrados tem-se

$$\sum_{l \epsilon D} \|\tilde{\boldsymbol{\eta}}_j(\ell) - \boldsymbol{\eta}_{j0}(\boldsymbol{b}_\ell)\|^2 \sim \sigma^2 \chi_{g_j \cdot \#(D), \delta_j(D)}, \quad j = 1, \ldots, m,$$

com

$$\delta_j(D) = \sum_{l \epsilon D} \delta_j(\boldsymbol{b}_l), \quad j = 1, \ldots, m,$$

onde

$$\delta_j(\boldsymbol{b}_\ell) = \frac{1}{\sigma^2} \|\boldsymbol{\eta}_j(l) - \boldsymbol{\eta}_{j0}(\boldsymbol{b}_l) \ \|^2, l \epsilon D, j = 1, \ldots, m.$$

Assim, para testarmos as $H_{0j}(D)$, $j = 1, \ldots, m$ podemos utilizar as estatisticas

$$\mathcal{F}_j(D) = \frac{g}{g_j \cdot \#(D)} \frac{\sum_{\ell \epsilon D} \|\tilde{\boldsymbol{\eta}}_j(\ell) - \boldsymbol{\eta}_{j0}(\boldsymbol{b}_\ell)\|^2}{V}, \quad j = 1, \ldots, m,$$

que têm distribuições F com $g_j \cdot \#(D)$ e g graus de liberdade e parâmetro de não centralidade $\delta_j(D), j = 1 \ldots, m$. Por outro lado, os

$$\sum_{\ell \epsilon D} \|\tilde{\boldsymbol{\eta}}_j(\ell) - \boldsymbol{\eta}_j(\ell)\|^2, \quad j = 1, \ldots, m,$$

serão qui-quadrados centrais com $g_j \cdot \sharp(D)$ graus de liberdade independentes de V, pelo que as variáveis pivot

$$\mathcal{F}'(D) = \frac{g}{g_j \cdot \#(D)} \frac{\sum_{\ell \epsilon D} \|\tilde{\boldsymbol{\eta}}_j(\ell) - \boldsymbol{\eta}_j(\ell)\|^2}{V}, \quad j = 1, \ldots, m.$$

terão distribuições F centrais com $g_j \cdot \#(D)$ e g graus de liberdade, $j = 1, \ldots, m$. Com

$$\begin{cases} \tilde{\boldsymbol{\eta}}_j(D) & = \left[\tilde{\boldsymbol{\eta}}_j(l)^t; \quad l \epsilon D\right]^t; j = 1, \ldots, m; D \subseteq \bar{\bar{u}} \\ \boldsymbol{\eta}_j(D) & = \left[\boldsymbol{\eta}_j(l)^t; \quad l \epsilon D\right]^t; j = 1, \ldots, m; D \subseteq \bar{\bar{u}} \\ \boldsymbol{\eta}_{j0}(D) & = \left[\boldsymbol{\eta}_{j0}(l)^t; \quad l \epsilon D\right]^t; j = 1, \ldots, m; \quad D \subseteq \bar{\bar{u}}, \end{cases}$$

onde $\bar{\bar{u}} = \{1, \ldots, u\}$ teremos

$$\begin{cases} \mathcal{F}_j(D) & = \frac{g}{g_j \#(D)} \frac{\|\tilde{\eta}_j(D) - \eta_{j0}(D)\|^2}{V}; \quad j = 1, \ldots, m; D \subseteq \bar{\bar{u}} \\ \mathcal{F}'_j(D) & = \frac{g}{g_j \#(D)} \frac{\|\tilde{\boldsymbol{\eta}}_j(D) - \boldsymbol{\eta}_j(D)\|^2}{V}; \quad j = 1, \ldots, m; D \subseteq \bar{\bar{u}}. \end{cases}$$

Pode utilizar-se $\mathcal{F}_j(D), j = 1, \ldots, m; D \subseteq \bar{\bar{u}}$, para testar as hipóteses

$$H_{0j}(D) : \boldsymbol{\eta}_j(D) = \eta_{j0}(D); \quad j = 1, \ldots, m; D \subseteq \bar{\bar{u}},$$

vendo-se que

$$H_{0j}(D) = \bigcap_{l \epsilon D} H_{0j}(l),$$

isto é, $H_{0j}(D)$ verifica-se e só se as $H_{0j}(\ell), l \epsilon D$, se verificarem $j = 1, \ldots, m$. Aliás, temos ainda a possibilidade de reescrever estas hipóteses como

$$\begin{cases} H_{0j}(D) & : \delta_j(D) = 0; \quad j = 1, \ldots, m; D \subseteq \bar{\bar{u}} \\ H_{0j}(l) & : \delta_j(\ell) = 0; \quad j = 1, \ldots, m; \ell = 1, \ldots, u \end{cases}$$

e como os $\mathcal{F}_j(D) \quad [\mathcal{F}_j(\ell); \ell = 1, \ldots, u]$ têm distribuições F com $g_j \# D$ e g $[g_j \quad e \quad g]$ graus de liberdade e parâmetro de não centralidade $\delta_j(D) \quad [\delta_j(l),\ l = 1, \ldots, u],\ j =$

$1, \ldots, m, D \subseteq \bar{\bar{u}}$, os testes F com estas estatísticas serão estritamente não distorcidos, ver Mexia (1995). Observe-se que,

$$\mathcal{F}_j(D) = \frac{1}{\#(D)} \sum_{\ell \epsilon D} \mathcal{F}_j(\ell); \quad j = 1, \ldots, m; D \leq \bar{\bar{u}},$$

o que é importante, pois como se verifica facilmente pode ter-se

$$\mathcal{F}_j(D) \leq F_{1-q, g_j \# D, g},$$

embora para um ou mais dos $\ell \epsilon D$ se tenha

$$\mathcal{F}_j(\ell) > F_{1-q, g_j, g},$$

não se rejeitando $H_{0j}(D)$ ao nivel q embora se rejeitem uma ou mais das $H_{0j}(\ell)$ a esse nivel. Assim podemos considerar que os testes para as $H_{0j}(\ell), l \epsilon F, j = 1, \ldots, m$ representam um refinamento dos testes para as $H_{0j}(D),\ j = 1, \ldots, m$.

As variáveis pivot $\mathcal{F}'_j(D)$ $[\mathcal{F}'_j(\ell), \ell = 1, \ldots, u]$ têm distribuições F central com $g_j \#(D)$ e g $[g_j \quad e \quad g]$ graus de liberdade, $j = 1, \ldots, m$, tendo-se pois as hiperesferas de confiança,

$$\begin{cases} \|\boldsymbol{\eta}_j(D) - \tilde{\boldsymbol{\eta}}_j(D)\|^2 & \leq g_j \#(D) \dfrac{V}{g}, \quad j = 1, \ldots, m, D \subseteq \bar{\bar{u}} \\ \\ \|\boldsymbol{\eta}_j(\ell) - \tilde{\boldsymbol{\eta}}_j(\ell)\|^2 & \leq g_j \dfrac{V}{g}, \quad j = 1, \ldots, m, \ell = 1, \ldots, u, \end{cases}$$

paras os $\eta_j(D), D \subseteq \bar{\bar{u}}$, e os $\eta_j(\ell), j = 1, \ldots, m, \ell = l, \ldots, u$. Raciocinando como atrás vê-se ainda que os testes F para as $H_j(D), D \subseteq \bar{\bar{u}}$, e as

$$H_j(l), l = 1, \ldots, n, j = 1, \ldots, m,$$

gozam de dualidade. Tem-se ainda, um nivel de confiança conjunto $1 - q$, os intervalos de confiança simultâneos

$$\begin{cases} \bigcap_{\boldsymbol{d}_j} \left(\mid \boldsymbol{d}_j^t \boldsymbol{\eta}_j(D) - \boldsymbol{d}_j^t \tilde{\boldsymbol{\eta}}_j(D) \mid \leq \quad \sqrt{g_j \#(D) \|d_j\|^2 \frac{V}{g}} \right), \quad j = 1, \ldots, m, D \leq \bar{\bar{u}} \\ \bigcap_{\boldsymbol{d}_{j\ell}} \left(\mid \boldsymbol{d}_{j\ell}^t \boldsymbol{\eta}_{j\ell}(\ell) - \underline{d}_{j\ell}^t \tilde{\boldsymbol{\eta}}_j(\ell) \mid \leq \quad \sqrt{g_j \|d_{j\ell}\| \frac{V}{g}} \right) \quad , j = 1, \ldots, m, \ell = 1, \ldots, u \end{cases}$$

4.5 Aplicação às famílias de matrizes estocásticas quase escalares

Como vimos a informação contida numa matriz simétrica a que se ajusta um modelo da forma

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

com

$$\bar{\mathbf{E}} = \frac{1}{2}(\mathbf{E} + \mathbf{E}^t)$$

e

$$vec(\mathbf{E}) \sim \aleph(\mathbf{0}_{k^2}, \sigma^2 \mathbf{I}_{k^2}),$$

condensa-se num vector de estrutura ajustado

$$\tilde{\boldsymbol{\beta}} = \theta \boldsymbol{\gamma}$$

e numa soma de quadrados dos residuos

$$V = \|\mathbf{M}\|^2 - \theta^2.$$

Vimos ainda que

$$\not{\Sigma}(\tilde{\boldsymbol{\beta}}) \approx \sigma^2 (\mathbf{I}_k + \boldsymbol{\gamma} \boldsymbol{\gamma}^t),$$

no entanto, esta heterocedasticidade é pouca acentuada. Aliás nas famílias estruturadas de matrizes simétricas está-se no caso equilibrado já que as somas de quadrados para os residuos das diferentes matrizes têm

$$g^{\cdot} = \frac{k(k-1)}{2},$$

graus de liberdade. Ora, ver Schefféé (1959) e Ito (1980), a ANOVA e técnicas relacionadas são robustas para a heterocedasticidade no caso equilibrado. Assim, admitimos ter pares $(\tilde{\boldsymbol{\beta}}_\ell, V_\ell)$ independentes com

$$\begin{cases} \tilde{\boldsymbol{\beta}_e} ll \sim & \aleph(\boldsymbol{\beta}_i, \sigma^2 \mathbf{I}_k), \quad i = 1, \dots, n \\ \\ V_i \sim & \sigma^2 \chi^2_{g^{\cdot}}, \quad i = 1, \dots, n, \end{cases}$$

sendo ainda os $\tilde{\boldsymbol{\beta}_e} ll$ independentes das V_i, $i = 1, \dots, u$. A aplicação dos resultados das alíneas anteriores é agora directa. Seguimos assim o caminho já desenvolvido por Oliveira e Mexia (1998,1999a,199b, 2002, 2004, 2007a, 2007 b), Ramos (2006) e Areia (2009).

No caso dos modelos simples há duas variantes interessantes:

- Analise Transversal: em que se trabalha com os vectores constituídos pelas componentes homológas (primeira, segunda,$\dots$,) dos vectores de estrutura. Esta análise tem uma extensão directa para os modelos múltiplos.

- Análise Longitudinal: em que dado um vector $\boldsymbol{c}$ se trabalha com $\boldsymbol{c}^t\tilde{\boldsymbol{\beta}}_1, \dots, \boldsymbol{c}^t\tilde{\boldsymbol{\beta}}_m$. Esta análise diz-se longitudinal pois se as componentes correspondem a instantes, sucessivos as $\boldsymbol{c}^t\tilde{\boldsymbol{\beta}}_i, \quad i = 1, \dots, n$ podem ser utilizadas para medir a evolução. Se quiséssemos podia-se utilizar os resultados da alínea 4.4 para considerar simultaneamente vários contrastes.

Recorde-se que um dos vectores de contrastes tem uma das componentes nula.

4.6 Subfamílias

Ao considerarmos as famílias estruturadas admitimos que, para cada tratamento do modelo base, existia uma matriz da família.

Vamos agora passar a admitir que as famílias se articulam em subfamílias e que são estas que correspondem aos tratamentos. Uma primeira questão a esclarecer será o grau de heterogeneidade das subfamílias.

Admitimos ter h subfamílias com $m_1, \ldots, m_h$ matrizes $k \times k$ cada. Dado estarmos a trabalhar com matrizes estocásticas quase escalares que cada matriz origina um par $\left(\tilde{\boldsymbol{\beta}}_{i,j}, V_{i,j}\right), \quad j = 1, \ldots, m_i; \quad i = 1, \ldots, h$ constituída por um vector de estrutura ajustado e por uma soma de quadrados de resíduos. Mantendo os pressupostos que temos utilizado na inferência temos

$$\tilde{\boldsymbol{\beta}}_{i,j} \sim \aleph\left(\boldsymbol{\beta}_{i,j}, \sigma^2 I_k\right), \quad j = 1, \ldots, m_i; \quad i = 1, \ldots, h$$

independente de

$$V_{i,j} \sim \sigma^2 \chi^2_{\ell} \quad j = 1, \ldots, m_i; \quad i = 1, \ldots, h,$$

sendo ainda independente os pares $\left(\tilde{\boldsymbol{\beta}}_{i,j}, V_{i,j}\right), \quad j = 1, \ldots, m_i; \quad i = 1, \ldots, h$ com

$$\begin{cases} \boldsymbol{\beta}_{i,\circ} & = \dfrac{1}{m_i} \displaystyle\sum_{j=1}^{m_i} \boldsymbol{\beta}_{i,j} \quad i, \ldots, h \\ \tilde{\boldsymbol{\beta}}_{i,\circ} & = \dfrac{1}{m_i} \displaystyle\sum_{j=1}^{m_i} \tilde{\boldsymbol{\beta}}_{i,j}, \quad i, \ldots, h, \end{cases}$$

temos

$$\sum_{i=1}^{m_i} \|\tilde{\boldsymbol{\beta}}_{i,j} - \tilde{\boldsymbol{\beta}}_{i\circ}\|^2 \sim \sigma^2 \chi^2_{k(m_i-1),\delta_i} \quad i = 1, \ldots, h,$$

com

$$\delta_i = \frac{1}{\sigma^2}\sum_{j=1}^{m_i} \|\boldsymbol{\beta}_{i,j} - \boldsymbol{\beta}_{i,\circ}\|^2, i = 1, \ldots, h,$$

independente de

$$V_i = \sum_{i=1}^{m_i} V_{i,j} \sim \sigma^2 \chi_{g_i}^{2}, \quad i = 1, \ldots, h,$$

com, atendendo á produtividade dos qui-quadrados,

$$g_i = m_i \ell \quad i = 1, \ldots, h.$$

Assim a estatística

$$\mathcal{F}_i = \frac{g_i}{k(m_i - 1)} \frac{D_i}{V_i}, \quad i = 1, \ldots, h$$

terá distribuição $F(|k(m_i - 1), g_i, \delta_i), \quad i = 1, \ldots, h$, podendo ser utilizado para testar a hipótese

$$H_{0i} : \tilde{\boldsymbol{\beta}}_{i1} = \ldots, \tilde{\boldsymbol{\beta}}_{i,m_i}, \quad i = 1, \ldots, h.$$

Como δ_i se anula se e só se as $H_{0i}, \quad i = 1, \ldots, h$, se verificare os m testes serão estritamente não destorcidos.

A partir das estatisticas $\mathcal{F}_i, i = 1, \ldots, h$, obtém-se os estimadores centrados

$$\tilde{\boldsymbol{\delta}}_i = \frac{k(m_i - 1) - 2}{k(m_i - 1}\mathcal{F}_i - 1, \quad i = 1, \ldots, k$$

dos δ_i, i=1,...,h, que podemos utilizar para medir a heterogeneidade das subfamílias. Na aplicação que apresentamos temos duas famílias de matrizes havendo uma correspondência 1-1 entre as subfamílias de ambas. Compararemos as heterogeneidade das subfamílias correspondentes.

Capítulo 5

Metodologia Statis

5.1 Introdução

A metodologia Statis ("Structuration de Tableaux à Trois Indices de la Statistique") foi inicialmente introduzida por Escoufier(1973) e L'Hermier des Plantes (1976) e ainda por Escoufier L'Hermier des Plantes(1973),no Laboratório de Probabilidades e Estatísticas da Universidade de Montpellier II, por volta de 1976, e posteriormente desenvolvida por Lavit(1988). A aplicação deste método permite a exploração simultânea de vários quadros de dados quantitativos recolhidos da seguinte forma:

- Segundo o método Statis, os T quadros de dados são recolhidos em diferentes instantes sobre os mesmos objectos; as variáveis podem diferir-se ao longo dos quadros; o método é caracterizado por T estudos.

  $$(\mathbf{X}_i, \mathbf{D}_n, \mathbf{D}_{p_i}),\ i = 1, \ldots, T,$$

  em que $\mathbf{X}_i$ representa as matrizes de dados do tipo $n \times p_i$, $\mathbf{D}_n$ representa a matriz diagonal dos pesos dos objectos e $\mathbf{D}_{pi}$ representa a matriz diagonal dos pesos das variáveis. Com a aplicação desse método consegue-se determinar o nível de proximidade entre os indivíduos.

- Segundo o método Statis Dual, os T quadros de dados são recolhidos em diferentes instantes sobre as mesmas variáveis, em que os indivíduos podem variar ao longo dos quadros. Neste caso o método é caracterizado por T estudos $(\mathbf{X}_i, \mathbf{D}_n, \mathbf{D}_{pi})$ com $i = 1, \ldots, T$, permitindo assim determinar a relação entre as variáveis.

Os métodos **Statis**, desenvolvem-se segundo três etapas fundamentais:

- **Inter- estrutura:** permite fazer uma comparação global dos quadros de dados.

- **Intra- estrutura**: descrever a estrutura comum dos vários quadros de dados através da determinação do compromisso e da respectiva imagem euclidiana.

- **Trajectórias**: No método "Statis"destaca os indivíduos e no "Statis Dual"as variáveis responsáveis, pelas semelhanças ou diferenças entre os quadros de dados, a partir da imagem compromisso, traçam-se as trajectórias que descrevem o comportamento evolutivo de cada individuo ou variável.

5.2 Método Statis

Os T quadros de dados são constituídos por n indivíduos e p_i variáveis quantitativas, com $i = 1, \ldots, T$. No instante i a tabela $\mathbf{X}_i$ representa a matriz de dimensão $n \times p_i$

$$\mathbf{X}_i = \begin{bmatrix} (x_1^1)^i & (x_1^2)^i & ... & (x_1^{p_i})^i \\ (x_2^1)^i & (x_2^2)^i & ... & (x_2^{p_i})^i \\ ... & ... & ... & ... \\ (x_n^1)^i & (x_n^2)^i & ... & (x_n^{p_i})^i \end{bmatrix} \tag{5.2.0}$$

5.2.1 Inter-estrutura

Esta etapa consiste em comparar os T quadros de dados, para tal é necessário definir um operador representativo para cada estudo, uma métrica entre os operadores representativos dos estudos e uma representação geométrica dos estudos. O estudo é representado por uma matriz

$$\mathbf{W}_i = \mathbf{X}_i \mathbf{D}_{p_i} \mathbf{X}_i^t, \ i = 1, \ldots, T.$$

A cada estudo será associado um operador do tipo

$$\mathbf{A}_i = \mathbf{W}_i \mathbf{D}_n, \quad i = 1, \ldots, T,$$

onde $\mathbf{W}_k$, representa os produtoss escalares entre os indivíduos. A comparação entre os estudos poderá ser feita através dos operadores utilizando a norma euclidiana da diferença entre os operadores $\mathbf{A}_i$ e $\mathbf{A}_j$, associados aos estudos. Vejamos como representar geometricamente os estudos em R^T. Para isso consideremos a matriz

$$\mathbf{S} = tr(\mathbf{W}_i \mathbf{D}_n \mathbf{W}_j \mathbf{D}_n) = [s_{ij}], i, j = 1, ..., T$$

simétrica, semi-definida positiva de ordem T(Escoufier,1976), com termo geral

$$\mathbf{S}_{ij} = tr(\mathbf{A}_i \mathbf{A}_j^t), i, j = 1, ..., T,$$

dado pelo produto de Hilbert Schmidt de $\mathbf{A}_i$ e $\mathbf{A}_j$. A matriz $\mathbf{S}$ terá decomposição espectral:

$$\mathbf{S} = \sum_{i=1}^{T} \theta_i \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^t,$$

com os valores próprios $\theta_1 > \theta_2 >, ..., \theta_T$, associados aos vectores próprios $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_T$. As coordenadas do ponto representativo do j-éssimo estudo são as j-éssimas componentes dos vectores $\theta_1\boldsymbol{\gamma}_1, \theta_2\boldsymbol{\gamma}_2, \ldots, \theta_T\boldsymbol{\gamma}_T$. Verifica-se que se $\mathbf{A}_i \approx \mathbf{A}_j$ os pontos representativos dos dois estudos estarão próximos. A matriz $\mathbf{S}$ pode ser considerada como uma matriz de covariâncias entre os T estudos, admite um valor próprio positivo associado a um vector próprio de componentes todas positivas, Lavit (1988). Outras propriedades desta matriz, encontram-se em Escoufier e L'Hermier (1978).

A metodologia Statis modela a inter-estrutura utilizando para isso distâncias euclidianas entre configurações observadas em T situações diferentes. Estas configurações são descritas por variáveis quantitativas ou qualitativas transformadas Oliveira(1996).

5.2.2 Compromisso

Na etapa Compromisso obtem-se uma tabela que resume o conjunto dos objectos da mesma natureza e que representa a média ponderada dos objectos $\mathbf{W}_i$:

$$\mathbf{W} = \sum_{i=1}^{T} \alpha_i \mathbf{W}_i.$$

Para determinar a imagem euclidiana dos T estudos é afectado, a cada um destes, um peso designado de π_i, em que a matriz dos pesos é dada por:

$$\triangle = \begin{bmatrix} \pi_1 & 0 & ... & 0 \\ 0 & \pi_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \pi_T \end{bmatrix} \tag{5.2.0}$$

Se os estudos tiverem a mesma importância, podemos então escrever $\triangle = \frac{1}{T} I_T$. A análise espectral é feita sobre a matriz $\mathbf{S}\triangle$, onde se calculam os valores próprios $\theta_1, \theta_2, \ldots, \theta_T$ associados aos vectores próprios $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_T$, respectivamente.

Se os objectos forem normados o compromisso é definido por:

$$\mathbf{W} = \sum_{i=1}^{T} \alpha_i \frac{\mathbf{W}_i}{\|\mathbf{W}_i\|_{HS}}.$$

A determinação dos coeficientes α_i depende de dois critérios:

- O compromisso $\mathbf{W}$ é o objecto mais correlacionado com os objectos $\mathbf{W}_i$

- $\mathbf{W}$ deve ser um objecto da mesma natureza que os objectos W_i, isto é,

$$\|\mathbf{W}\|_{HS} = \sum_{i=1}^{T} \pi_i \|\mathbf{W}_i\|_{HS}.$$

Se os objectos forem normados então $\|\mathbf{W}\|_{HS} = 1$.

Seja γ_1 o vector próprio de $\mathbf{S}\triangle$, associado ao maior valor próprio θ_1 designado por:

$$\boldsymbol{\gamma}_1 = \begin{bmatrix} \gamma_1^1 \\ \gamma_1^2 \\ \vdots \\ \gamma_1^T. \end{bmatrix} \tag{5.2.0}$$

As coordenadas são todas do mesmo sinal e positivas. Segundo Lavit(1988) e Lavit et al(1999) os coeficientes α_i são determinados da seguinte forma:

- no caso dos objectos $\mathbf{W}_i$;

$$\alpha_i = \frac{1}{\sqrt{\theta_1}} \left(\sum_{l=1}^{T} \pi_l \|\mathbf{W}_l\|_{HS} \right) \pi_i \boldsymbol{\gamma}_1^i$$

- no caso dos objectos normados $W_i/\|\mathbf{W}_i\|_{HS}$

$$\alpha_i = \frac{1}{\sqrt{\theta_1}} \pi_i \boldsymbol{\gamma}_1^i;$$

- no caso dos objectos $\mathbf{W}_i$, o Compromisso, tem a seguinte expressão

$$\mathbf{W} = \sum_{i=1}^{T} \left[\frac{1}{\sqrt{\theta_1}} \left(\sum_{l=1}^{T} \pi_l \|\mathbf{W}_l\|_{HS} \right) \pi_i \gamma_1^i W_i \right]$$

- no caso dos objectos serem normados, $\mathbf{W}_i/\|\mathbf{W}_i\|_{HS}$, tendo-se

$$\mathbf{W} = \sum_{i=1}^{T} \left[\frac{1}{\sqrt{\theta_1}} \pi_i \boldsymbol{\gamma}_1^i \frac{\mathbf{W}_i}{\|\mathbf{W}_i\|_{HS}} \right].$$

A coordenada do compromisso $\mathbf{W}$ sobre o i-éssimo eixo é obtida por combinação linear das coordenadas $\sqrt{\theta_i}\boldsymbol{\gamma}_i^t$ dos pontos A_i sobre o i-éssimo eixo:

$$\sum_{i=1}^{T} \alpha_i \sqrt{\theta_i} \boldsymbol{\gamma}_i^t.$$

Como os vectores próprios da matriz $\mathbf{S}\triangle$ são ortonormados (uma vez que $\mathbf{S}\triangle$ é simétrica) e por definição de α_i, todas as coordenadas do compromisso serão nulas com excepção da primeira.

5.3 Método Statis dual

Este método aplica-se a dados quantitativos, constituídos por um conjunto de tabelas de p variáveis medidas sobre um conjunto de indivíduos, que podem não ser os mesmos de tabela para tabela. Este método, tal como o método Statis, é constituído por três etapas análogas (Inter-estrutura, Compromisso e Trajectórias). Neste caso, Escoufier e L'Hermier (1978), associaram aos estudos as matrizes:

$$\mathbf{B}_i = \mathbf{X}_i^t \mathbf{D}_n X_i \mathbf{D}_p, i = 1, \ldots, T,$$

assim a matriz $\mathbf{S}$ será da forma

$$\mathbf{S}_{ij} = tr(\mathbf{B}_i \mathbf{B}_j^t), \quad i, j = 1, \ldots, T.$$

5.4 Modelação de séries de estudos

Na sua formulação inicial a metodologia Statis, só considerava o estudo de uma única série de estudos consistindo em projecções ortogonais sobre planos paralelos aos dois primeiros vectores próprios da matriz de Hilbert-Schmidt.

Oliveira e Mexia(1998),(1999 a),(1999 b) (2002),(2004), (2007 a),(2007 b), Ramos(2006) e Areia(2009), desenvolveram modelos tendo em vista a inferência para séries isoladas e séries emparelhadas de estudos. Este conceito, "emparelhadas", concretiza-se quando entre duas ou mais séries de estudos existe uma bijecção entre os respectivos conjuntos de estudos.

Nos modelos considerados em Oliveira (2002), a matriz **S** desempenhava um papel central. A abordagem seguia a observação feita por Lavit (1988), de que quando uma série de estudos tem uma estrutura comum os pontos representativos dos estudos se encontram ao longo do primeiro eixo. Assim, os valores próprios com excepção do primeiro são muito pequenos , isto é,

$$\theta_j^2 = \|\theta_j \boldsymbol{\gamma}_j\|^2 \approx 0,\ j = 2, \ldots, T$$

Neste caso a matriz **S** será

$$\mathbf{S} \approx \theta_1 \boldsymbol{\gamma}_1 \boldsymbol{\gamma}_1^t$$

tendo $\theta \boldsymbol{\gamma}_1 \boldsymbol{\gamma}_1^t$ caracteristica um. O primeiro modelo, Oliveira e Mexia(1988), considerado foi da forma

$$\mathbf{S} = \theta \boldsymbol{\gamma} \boldsymbol{\gamma}^t + \mathbf{E}$$

com $\mathbf{E} = (e_{ij})$. Neste modelo admitiu-se que os elementos da matriz **E** são independentes e identicamente distribuídos (i.i.d.), normais, com valor médio nulo e variância σ^2. Este modelo embora extremamente manejável tinha o inconveniente de não respeitar a simetria da matriz $\mathbf{S}$. Admitiu-se posteriormente um outro modelo simétrico

$$\mathbf{S} = \theta\boldsymbol{\gamma}\boldsymbol{\gamma}^t + \overline{\mathbf{E}},$$

com $\overline{E} = \frac{1}{2}(E+E^t)$, com o qual foi possivél obter a generalidade dos resultados obtidos no primeiro modelo. Os modelos adoptados permitiram condensar a informação de uma série de estudos num estimador $\widetilde{\beta}$, vector estrutura, da forma, $\widetilde{\beta} = \theta\boldsymbol{\gamma}$. Estes modelos apenas consideravam séries com estrutura de grau um.

Ramos (2006), admitiu a normalidade na componente aleatória das observações iniciais, considerando a condensação da informação contida numa série de estudos num vector de estrutura $\widetilde{\boldsymbol{\beta}} = \theta_1\boldsymbol{\gamma}_1$ estimador de $\boldsymbol{\beta} = \lambda\alpha$ e na soma dos quadrados dos resíduos

$$\|\mathbf{S} - \sum_{j=1}^{h}\theta_j\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^t\|^2.$$

Na formulação dos modelos para séries isoladas de estudos com estrutura comum de grau um, foi utilizada uma abordagem baseada no conceito de "quase-normalidade". Este conceito permitiu, com recurso a estudos de simulação, mostrar que certas funções não lineares de variáveis normais com desvios padrões muito menor que os módulos dos valores médios pequeno podem ser tratadas como variáveis normais também elas com desvio relativo pequeno.

Areia et al. (2008), admitiu que os elementos das matrizes $A_i, i = 1, \ldots, T$, são normais desenvolvendo, a partir daí, o modelo aplicável às séries de estudos. No desenvolvimento destes modelos considerou-se de uma forma mais abrangente o conceito "quase-linearidade"já utilizado por Ramos (2006). Os modelos em estudo foram generalizados para estruturas de grau h.

No âmbito desta dissertação desenvolveremos modelos aplicáveis a matrizes simétricas. Consideraremos os casos em que essas matrizes são:

- as matrizes **S** de produtos de Hilbert Schimidt da inter-estrutura;
- matrizes compromisso.

Realizaremos assim um estudo integrado das duas primeiras etapas da metodologia Statis (inter-estrutura e compromisso). Em particular para a primeira etapa podemos realizar as análises, transversal e longitudinal e para a segunda a análise transversal. Com efeito, enquanto na primeira etapa se pode considerar uma sucessão de estudos, o que justifica a realização de análise longitudinal, o mesmo não se verifica para a segunda etapa.

Um problema de homogeneidade

Suponhamos que as populações de séries se agrupam em subfamílias. Podemos então aplicar os resultados da alínea 4.6 para medir a heterogeneidade das mesmas. Observa-se que se terá uma medida de heterogeneidade para a primeira e outra para a segunda etapa. Na primeira [segunda] trabalhou-se com os vectores de estrutura ajustados para as matrizes de produtos de Hilbert Schimidt [compromisso].
Interessará ver em que medida é que uma maior ou menor heterogeneidade se transporta para a segunda etapa.
Possivelmente, havendo m pares de sub-famílias, poderá utilizar-se o coeficiente de correlação ordinal de Sperman para testar a preservação da ordem de heterogeneidade.

Capítulo 6

Heterogeneidade em eleições autárquicas

6.1 Dados e objectivos

De maneira a cobrir todo o território Continental Português escolheram-se os distritos de Beja, Braga, Guarda, Leiria, Setúbal e Vila e Real. Deste modo, construímos uma tabela de dados com os resultados apurados ao nível do distrito/freguesias em relação às 6 regiões de Portugal Continental correspondentes aos distritos de Braga (13 Concelhos), Leiria (16 concelhos), Guarda (14 concelhos), Setúbal (12 concelhos), Beja (12 concelhos) e Vila Real (14 concelhos), trabalhando-se todos os concelhos dos mesmos.

Tabela 6.1: Distridos de Portugal em estudo

	Litoral	Interior
Norte	Braga	Vila Real
Centro	Leiria	Guarda
Sul	Setúbal	Beja

Para cada concelho destes distritos tínhamos uma série de estudos logo uma sub-população por distrito no sentido atrás referido.

Essas séries contêm os resultados das sete eleições autárquicas referentes aos distritos.

Para cada concelho os objectos são as freguesias e as variáveis são as opções politicas.

Em algumas eleições participaram PS, PSD e CDS coligados AD, os votos respectivos foram distribuídos pelo PSD e CDS proporcionalmente ao das duas eleições mais próximas em que concorreram separados considerando-se eleições em que concorreram separados.

Resolveu-se estudar a heterogeneidade dessas populações na primeira e segundas etapas da metodologia Statis.

Utilizando-se essa metodologia obtiveram-se para cada um dos concelhos

- a matriz dos produtos de Hilbert Schmidt;
- a matriz compromisso.

6.2 Resultados

A partir dos resultados obtidos obtiveram-se as preponderâncias dos primeiros valores próprios. Estes resultados constam nas Tabelas 6.2 a 6.7.

Tabela 6.2: Valores da preponderância do distrito de Beja

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Aljustrel	124.49	305.89
Almodôvar	344.81	260.41
Beja	152.87	168.62
C. Verde	292.33	241.01
Cuba	388.32	934.90
F. do Alentejo	163.28	279.62
Mértola	254.78	444.80
Moura	434.00	188.58
Odemira	120.53	182.77
Ourique	231.48	81.48
Serpa	107.01	62.57
Vidigueira	186.27	353.97

Tabela 6.3: Valores da preponderância do distrito de Braga

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Amares	11.64	17.65
Barcelos	222.61	98.26
Braga	317.24	172.03
Cab. de Basto	219.39	123.64
Cel. de Basto	7.59	15.08
Esposende	11.24	17.92
Fafe	287.71	104.81
Guimarães	19.23	31.5
Póvoa de Lanhoso	44	37.59
Terras de Bouro	11.96	20.17
Vieira do Minho	96.81	76.35
V. N. de Famalicão	157.12	129
Vila verde	5.36	10.23

Tabela 6.4: Valores da preponderância do distrito da Guarda

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Aguiar da Beira	14.89	20.09
Almeida	2.51	4.37
C. da beira	16.96	26.72
F. C. Rodrigo	23.6	37.47
F. de Algodres	46.01	78.46
Gouveia	303.62	108.12
Guarda	370.83	113.89
Manteiga	18.23	36.15
Meda	15.36	22.93
Pinhel	15.82	25.32
Sabugal	16.8	15.52
Seia	172.96	73.62
Trancoso	97.3	60.95
V. N. Coa	175.08	91.17

Tabela 6.5: Valores da preponderância do distrito de Leiria

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Alcobaça	317.9	133.25
Alvaiázere	30.2	43.53
Ansião	213.31	343.96
Batalha	19.01	44.76
Bombarral	52.27	130.64
Caldas da Rainha	389.72	240.5
Castanheira de Pera	292.97	53.22
Figueiró dos Vinhos	12.23	29.76
Leiria	49.6	68.34
Marinha Grande	343.66	965.7
Nazaré	614.44	346.53
Óbidos	566.33	653.19
Pedrogão Grande	107.23	227.03
Peniche	186.99	279.88
Pombal	351.98	284.27
Porto de Mós	104.84	92.58

Tabela 6.6: Valores da preponderância do distrito de Setúbal

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Alcácer do Sal	366.21	137.62
Alcochete	1297.39	752.98
Almada	135.11	415.56
Barreiro	112.97	286.38
Grândola	109.41	582.27
Moita	49.25	167.73
Montijo	105.15	192.76
Palmela	262.92	491.58
Santiago do Cacém	307.5	344.59
Seixal	133.97	247.71
Sesimbra	112.34	348.44
Setúbal	108.77	272.77

Tabela 6.7: Valores da preponderância do distrito de Vila Real

	Preponderância	
Beja	**Matriz de Hilbert-Schmidt**	**Matriz Compromisso**
Alijo	196.93	132.54
Bóticas	597.58	180.34
Chaves	75.11	93.91
Mesão Frio	1626.39	188.99
Mondim de Basto	10.24	21.29
Montalegre	275.7	87.91
Murça	78.48	117.9
P. da Régua	294.85	139.66
Ribeira de Pena	7.11	17.02
Sabrosa	101.19	70.84
S. Penaguião	356.72	65.22
Valpaços	76.04	75.73
V. de Aguiar	199.42	146.32
Vila Real	237.95	145.99

Daqui constata-se a forte predominância do primeiro valor próprio em relação aos restantes tanto para a matrizes de produtos de Hilbert-Schmidt, como para a matriz compromisso, o que revela que as matrizes consideradas são quase escalares para todos os distritos/concelhos. Os quadros seguintes apresentam os vectores de estrutura ajustado para as matrizes e a soma dos quadrados dos resíduos.

Como vimos quando se quer medir a heterogeneidade dessa subfamília em que m níveis e k estudos a que correspondem os pares $\left(\tilde{\boldsymbol{\beta}}, V_i\right)$ com $i = 1, \ldots, m$ de vectores de estrutura ajustado e de soma de quadrados dos resíduos, utiliza-se a estatística

$$\mathcal{F} = \frac{ml}{k(m-1)} \frac{S_1 - \frac{S_2}{m}}{V}$$

com

$$\begin{cases} S_1 & = \sum_{j=1}^{m} \sum_{i=1}^{k} \beta_{ij}^2 \\ S_2 & = \sum_{i=1}^{k} (\sum_{j=1}^{m} \beta_{ij})^2 \\ V & = \sum_{i=1}^{m} V_i \end{cases}$$

Os vectores de estrutura ajustados encontram-se nos quadros seguintes

Tabela 6.8: Vector estrutura e soma de quadrados dos resíduos do distrito de Beja

	inter-estrutura				
	Aljustrel	Almodôvar	Beja	C. Verde	Cuba
	25830722	23860559	22858554	25279517	27962451
	28416304	17612130	23411813	31030472	33295591
	25907786	17916950	21814336	29139541	30256293
	28334101	20799675	22678400	29126804	29536029
	22908614	19116053	20457135	26922427	27310296
	25006458	19139944	22034528	24124055	24173030
	25179470	18968012	22359796	26721734	25963787
V_i	38016908323367	7902153551154	22664239807922	18203315495401	14633855682961
	compromisso				
	Aljustrel	Almodôvar	Beja	C. Verde	Cuba
	232	701	351	278	65
	1131	1688	1094	1034	1356
	0	0	41	0	0
	190	191	157	171	225
	2380	779	2064	2565	2543
	1650	1836	1656	1600	1516
V_i	31880	28250	49500	42790	11401

Tabela 6.9: Vector estrutura e soma de quadrados dos resíduos do distrito de Beja - Continuação 1

	inter-Estrutura			
	F. do Alentejo	Mértola	Moura	Odemira
	23701440	22622481	24344670	21386827
	22778201	24253653	22514034	20983575
	21993893	24442680	23228085	20628395
	22364961	27138266	28826615	23650796
	21631583	23885376	26892283	21053914
	21370373	22335234	25503076	19627396
	22347083	23890106	25311689	20044764
V_i	21365952297205	15990433509845	10332632308752	25826243266343
	compromisso			
	F. do Alentejo	Mértola	Moura	Odemira
	243	155	188	511
	1719	1149	1276	963
	0	45	0	13
	159	225	173	149
	1791	2075	1675	1802
	1468	1833	2227	1855
V_i	30037	20380	50153	43233

Tabela 6.10: Vector estrutura e soma de quadrados dos resíduos do distrito de Beja - Continuação 2

	Inter - Extrutura			
	Ourique	Serpa	Vidigueira	Total
	18367133	23498989	23021408	
	19067183	22674107	24782185	
	18684366	24293361	26852025	
	22721955	27302856	27875246	
	16164247	26095664	24485778	
	20040508	31637402	22242242	
	18544241	25924346	21723512	
V_i	11115632610877	44442762906732	22593037244906	253087167005464
	compromisso			
	Ourique	Serpa	Vidigueira	Total
	1687	1687	287	
	506	506	1055	
	0	0	0	
	130	130	198	
	1421	1421	2226	
	1400	1400	1728	
V_i	87094	153260	25916	573896

Tabela 6.11: Vector estrutura e soma de quadrados dos resíduos do distrito de Braga

	Inter-estrutura			
	Amares	Barcelos	Braga	Cab. de Basto
	36708945	49611721	50484391	59675041
	9190216	3258694	2586445	3825867
	5224754	564724	1093853	1231739
	1963396	332870	356062	274299
	390629	89039	142657	46662
	102090	11294	25479	8320
	43423	2732	3604	3700
V_i	115777936417038	11056866765081	8034003900171	16231938133801
	compromisso			
	Amares	Barcelos	Braga	Cab. de Basto
	1926	1246	1755	1520
	331	594	861	1072
	850	1169	539	510
	152	115	149	165
	326	244	120	85
	1449	1393	1639	1728
V_i	271818	71231	40271	63065

Tabela 6.12: Vector estrutura e soma de quadrados dos resíduos do distrito de Braga - Continuação 1

	inter-estrutura			
	Cel. de Basto	Esposende	Fafe	Guimarães
	38349530	44784926	55158000	38484159
	13489615	13329854	2639520	8277729
	3332914	698018	1711432	2374542
	769988	550376	815527	1690756
	359934	137402	115209	97410
	148928	17938	10094	61007
	59208	7192	2139	10787
V_i	193826156923374	178494398904746	10574530241567	77031219712687
	compromisso			
	Cel. de Basto	Esposende	Fafe	Guimarães
	2401	1218	1193	1174
	624	1713	2176	1621
	581	187	141	223
	156	195	181	131
	202	200	203	464
	1366	1675	1423	1343
V_i	334801	334195	72723	165764

Tabela 6.13: Vector estrutura e soma de quadrados dos resíduos do distrito de Braga - Continuação 2

	inter-estrutura		
	P. de Lanhoso	Terras de Bouro	Vieira do Minho
	48875963	71011726	48866938
	6055976	20222386	3895009
	4157541	3452188	3069365
	501630	884365	265111
	267286	216761	52698
	53750	38941	36787
	28718	14575	22917
V_i	54286774430597	421693302180851	24667036879110
	compromisso		
	P. de Lanhoso	Terras de Bouro	V. do Minho
	1857	1464	1032
	338	803	781
	681	527	1076
	161	144	188
	462	323	135
	1433	1615	1626
	173971	432208	88260

Tabela 6.14: Valores do teste F da inter-estrutura e compromisso a nível dos distritos

	Teste F	
	Inter-Estrutura	**Compromisso**
Beja	11.91	91.23
Braga	5.15	22.8
Guarda	2.03	15.52
Leiria	10.87	73.01
Setubal	9.99	47.42
Vila real	6.72	42.01

Os valores das estatísticas $\mathcal{F}$ e as suas ordenações encontram-se tabela seguinte Para testar a concordância das ordenação utilizou-se o coeficiente ordinal de Sper-

Tabela 6.15: Estatísticas $\mathcal{F}$ e ordenações

Distritos	**Inter-extrutura**	**Compromisso**	**Ordem de S**	**Ordem do compromisso**	d_i	d_i^2
Beja	11.91	91.23	6	6	0	0
Braga	5.15	22.8	2	2	0	0
Guarda	2.03	15.52	1	1	0	0
Leiria	10.87	73.01	5	5	0	0
Setubal	9.99	47.42	4	4	0	0
Vila Real	6.72	42.01	3	3	0	0

mam tendo-se obtido $R_s = 1$. O coeficiente de correlação ordinal de Spermam é altamente significativo, concluindo-se que a heterogeneidade que se obtém na primeira etapa, inter-estrutura é, transportada para a segunda etapa do compromisso.

Capítulo 7

Considerações finais

Ao longo deste trabalho:

- apresentamos modelos, para matrizes estocásticas quase escalares que se aplicam, sempre que o respectivo primeiro valor próprio, tenha preponderância suficientemente alta;
- construímos um teste para verificar a condição de preponderância suficientemente alta utilizando a linearidade assimptótica;
- mostramos , como analisar também estruturas de matrizes quase escalares associadas a delineamento base ortogonais. Como referimos essas famílias cobrem a generalidade dos casos suficientemente regulares do delineamento base;
- exploramos, a possibilidade desses tratamentos integrados das duas primeiras etapas da metodologia Statis, tendo aplicado essa possibilidade aos resultados das eleições autárquicas de Portugal continental.

Do ponto de vista de interesse estatístico dos resultados apresentados não podemos deixar de referir o ter-se obtido uma distribuição limite F central e construído intervalos de confiança para parâmetros de não centralidade das distribuições $F(|1, g, \delta)$, quando δ é grande. Observa-se que na construção da distribuição limite foi possível eliminar a variância σ^2 que actuava como parâmetro perturbador.

Pretendemos aprofundar o estudo deste caso considerando situações em que tem

$$\begin{cases} \mathbf{X} & = \boldsymbol{\mu} + \boldsymbol{e} \\ (i) & \\ \mathbf{S} & \end{cases}$$

onde $\boldsymbol{e}$, tem distribuição esférica com parâmetro de dispersão θ e $\mathbf{S}$ tem como único parâmetros θ^2, que também actua como parâmetro de dispersão. Pretendemos ainda alargar esse trabalho ao caso de matrizes estocásticas simétricas cuja matriz média tenha característica $(r \geq 1)$.

Bibliografia

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

Areia A.,(2009).*Séries Emparelhadas de Estudos:Estrutura Comum Inferência.* Universidade de de Évora.Portugal

Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. e Nesbet, C. J. (1986),1986.*Actuarial Mathematics*. The Society of Actuaries.

Cantarinha A.,(2012).*Resultados Assintóticos para Famílias Estruturadas de Modelos Colectivos.*Aplicação aos Fogos Florestais em Portugal Continental. Universidade de Évora.Portugal

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

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

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, $23(2) : 175 - 201$.

Ito K.(1980).*Robustness of Anova and Macanova Test Procedures.* P. R Krishnaiah(ed), Handbook of Statistitics, Vol.I North Holland.

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

Lukacs E., Laha R.G.,(1964).*Applications of Characteristic Functions.* Charles Gritten.

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

L'Hermier des Palntes H.,(1976).*Struturation des Tableaus à Trois Indices de la Statistique. Thése de 3^{eme}cycle.UniveritédeMontepllierII&*

Mood A.,Graybill R., Boes D.,(1997).Introduction to the Theory of Statistics.Third Edition. McGRAW-HILL.

Mexia J. T.,(1989 a).Controlled Heterocedasticity, in: Quotient Vector Spaces and F tests for Hypotheses on Mean Vectors,Trabalhos de Investigação, N.1, Departamento de Matemática de Ciência e Tecnologia, Universidade Nova de Lisboa.

Mexia J. T.,(1989 b).Simultaneous Confidence Intervals: Generalization of the Scheffé Theorem.Trabalhos de Investigação, N.2, 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 Scheffés's multiple comparison method in the presence of controlled Heterocedasticity.Journal Comp. Stat. Data Analysis, $10(3) : 271 - 281$.

Mexia, J. T.,(1992).Assiynptotic Chi-Squared Test, Designe and Log-Linear Models. Trabalhos de Investigação,1992 N. 1

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.,(2011).Asymptotic linearity and limit distributions, approximations. Journal of Statistical Planning and Inference Vol. 140, (2), 353-357.

Oliveira M. M., Mexia, J. T.,(1998).Test for the rank of Hilbert-Schmidt product matrices.Advances in Data Science and Classications.(Rizzi, Vichi and Bock, Ed.), *619-625. Springer*

Oliveira M. M., Mexia, J. T.,(1999a).F tests for Hypothesis on the Structure Vectors of Series. Discussiones Mathematicae. Biometrical Letters. Vol. 19, N. 2, 345-353.

Oliveira M. M., Mexia, J. T.,(1999b).Multiple Comparations for Rank one Common Structures. Biometrical Letters. Vol. 36, N. 2, 159-167.

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

Oliveira M. M., Mexia, J. T.,(2004).AIDS in Portugal: endemic versus epidemic forecasting scenarios for mortality.International Journal of Forecasting, N. 20, 131-137.

Oliveira M. M., Mexia, J. T.,(2007).ANOVA like analysis of matched series of studies with a common structure. Journal of Statistical Planning and Inference, N. 137, 1862-1870.

Oliveira M. M., Mexia, J. T.,(2007).Modelling series of studies with a common structure. Computation Statistics and Data Analysis, N. 51, 5876-5885.

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

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

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

Vaquinhas R., Mexia J.,(1992).Convergence of matrices and Sbspaces with Statistical Applications. Distancie 92- Rennes.