

ANÁLISE DE COVARIÂNCIA COM ENFOQUE EM MODELOS LINEARES MISTOS DE SIMETRIA COMPOSTA

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra. Carla Cláudia da Rocha Rêgo Monteiro e aprovada pela Comissão Julgadora.

Campinas, 15 de março de 1996

Profa. Dra. Cláudia Azurido de Luna Freire
Orientadora

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em ESTATÍSTICA.

UNIDADE	BC
N.º CHAMADA:	UNICAMP
V.	M764a
I.	27745
P.	667196
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	22/05/96
N.º UFD	

CM-000 9000 3-4

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Monteiro. Carla Cláudia da Rocha Rego

M764a Análise de covariância com enfoque em modelos lineares mistos de simetria composta / Carla Cláudia da Rocha Rego Monteiro. -- Campinas, [S.P. :s.n.], 1996.

Orientadora : Clarice Azevedo de Luna Freire

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Ciência da Computação.

1. Análise de variância. 2. Análise de covariância. 3. Estimativa de parametro. 4. Modelos lineares (Estatística). I. Freire, Clarice Azevedo de Luna. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Ciência da Computação. III. Título.

Aos meus pais,

Vilma e José

AGRADECIMENTOS

Meu agradecimento aos que contribuíram para a realização desta dissertação.

Aos meus familiares pelo incentivo e apoio que foram essenciais para realização deste trabalho.

Ao Cristiano Ferraz pelo companheirismo e apoio.

À Profa. Clarice Azevedo de Luna Freire pela orientação, colaboração e amizade.

Aos amigos pelos momentos de alegria.

Ao Cherubino Colares pelas ajudas nas realizações dos programas.

Ao Marcelo do Nascimento pelo apoio logístico e confecção da capa.

SUMÁRIO

INTRODUÇÃO	1
CAPÍTULO 1 - Teoria Básica de Modelos Lineares	3
1.1 Modelo de Gauss - Markov	3
1.2 Modelo de Aitken	19
1.3 Modelo de Gauss - Markov Restrito	22
1.4 Estimador da Variância do Erro	27
1.5 Análise de Variância - ANOVA	33
1.6 Testes de Hipóteses	38
1.6.1 Comparando modelos Completos e reduzidos	39
1.6.2 Testando uma Hipótese	45
1.6.3 Comparando Modelos Restritos e Reduzidos ou Testando uma Hipótese ?	49
1.7 Reparametrização	53
CAPÍTULO 2 - Análise de Covariância	59
2.1 Modelo de Análise de Covariância	60
2.1.1 Modelo de Análise de Covariância generalizado	69
2.1.2 Teste de Hipóteses no Modelo de Análise de Covariância	74
2.2 Modelo com um fator e uma covariável	78
2.3 Ilustrações	87
CAPÍTULO 3 - Modelo Linear Misto de Laird - Ware	115
3.1 Modelo Linear Misto de Laird-Ware de Simetria Composta.....	116

3.1.1	Estimação : Máxima Verossimilhança	119
3.1.2	Teste de Hipóteses	126
3.1.4	Predição dos efeitos aleatórios do modelo de Laird - Ware de Simetria Composta	126
3.2	Modelo de Laird - Ware de Simetria Composta com Uma Covariável	129
3.3	Ilustações	131
APÊNDICE A		137
APÊNDICE B		149
REFERÊNCIAS BIBLIOGRÁFICAS		163

INTRODUÇÃO

A motivação desta dissertação se originou pelo surgimento de um problema para análise estatística, na área de medicina. Tratava-se de um estudo observacional onde precisávamos comparar as médias de duas populações, uma de pessoas saudáveis e outra, de pessoas acometidas de osteoartrose na cabeça do fêmur, doença caracterizada pela degeneração óssea. A variável de interesse era uma medida da irrigação local do sangue. Havia também o registro de uma segunda variável, medida da irrigação do sangue do grande trocanter (outra região do fêmur). Seria ideal uma análise estatística utilizando esta segunda medida como covariável, numa análise de covariância. Os dados já coletados eram de tal forma, que a medida de irrigação local de sangue foi observada várias vezes para indivíduos da população acometida por osteoartrose e apenas uma vez para indivíduos na outra população. Gostaríamos portanto de analisar o ajuste de um modelo que considerasse as várias observações dos indivíduos, bem como levasse em consideração a segunda medida de irrigação. Uma solução apresentada por nós foi o uso de um modelo linear misto de simetria composta com uma covariável.

Neste trabalho, o Capítulo 1 apresenta partes da teoria básica de modelos lineares que justificam os resultados utilizados nos capítulos seguintes. Em sua apresentação, várias demonstrações úteis ao entendimento do Capítulo 2 são explicitadas, uma vez que estas são usualmente omitidas na literatura da metodologia de análise de covariância. As provas de teoremas e resultados de Álgebra Linear e Teoria distribucional foram deslocadas para o Apêndice A.

O Capítulo 2 tem como objetivo a apresentação do modelo de análise de

covariância. Uma reparametrização do mesmo é apresentada, permitindo de forma mais fácil a obtenção dos estimadores dos parâmetros estimáveis. Alguns exemplos foram escolhidos como ilustração.

Por fim, no Capítulo 3 apresentamos um modelo linear que se utiliza de efeitos fixos e aleatórios: Modelo de Laird-Ware de Simetria Composta. O passo seguinte foi acrescentar uma covariável a este modelo de modo a compatibilizar o problema médico acima exposto.

CAPÍTULO 1

Teoria Básica de Modelos Lineares

Este Capítulo apresenta resultados básicos da teoria de modelos lineares. O nosso objetivo foi fazer uma seleção, dentre os resultados na literatura, de aspectos fundamentais e expô-los de maneira clara e formal. Assim, o conteúdo deste Capítulo é tido como suporte para resultados dos Capítulos seguintes.

Iniciamos com a apresentação do modelo de Gauss - Markov. Antes de mencionar a estimação dos parâmetros do modelo, o conceito de estimabilidade é apresentado. O método de estimação abordado é o de mínimos quadrados. Dois outros tipos de modelos, relacionados ao modelo de Gauss - Markov são analisados. Fazemos um destaque especial para o procedimento de teste de hipóteses, ou, alternativamente, escolha de modelo, com a organização em tabelas de análise de variância. Finalmente, abordamos a técnica de reparametrização.

Muitos resultados de Álgebra Linear são constantemente utilizados, como, por exemplo, espaço coluna e espaço linha de uma matriz projeções ortogonais. Para continuidade da apresentação dos resultados da modelagem estatística, resultados exclusivos de Álgebra Linear foram deslocados para o Apêndice A.

1.1 Modelo de Gauss - Markov

Geralmente estamos interessados em estudar uma variável aleatória Y com base em p variáveis $X_1, \dots, X_p$. Na literatura, Y é denominada variável dependente e $X_1, \dots, X_p$ variáveis independentes, porque os valores destas últimas são considerados

fixos, a priori. Na nossa dissertação, iremos estudar a relação linear apresentada a seguir.

O modelo linear de *Gauss - Markov*(GM) é definido como :

$$y = x'\beta + e \quad (1.1)$$

onde

y é a variável aleatória de interesse ;

$\underline{x}' = [x_1, \dots, x_p]$ é um vetor cujos elementos $x_1, \dots, x_p$ são os valores das variáveis independentes;

$\underline{\beta}' = [\beta_1, \dots, \beta_p]$ é um vetor de parâmetros desconhecidos, de dimensão p e

e é a variável aleatória erro, com $E(e) = 0$ e $\text{Var}(e) = \sigma^2$, $\sigma > 0$ desconhecido.

Observamos que o modelo deve ser linear nos parâmetros, e não necessariamente nas variáveis independentes. Logo, se considerarmos um termo quadrático para uma variável independente, por exemplo :

$$y = x_1\beta_1 + x_2\beta_2 + x_2^2\beta_3 + e, \quad E(e) = 0 \text{ e } \text{Var}(e) = \sigma^2,$$

temos um caso particular do modelo Gauss - Markov.

O modelo amostral de Gauss - Markov correspondente é :

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{e} \quad (1.2)$$

onde

$\underset{\sim}{y}' = [y_1, \dots, y_n]$ é um vetor de n observações da variável y ;

$\underset{\sim}{X} = [\underset{\sim}{x}_1, \dots, \underset{\sim}{x}_p]$ é uma matriz conhecida, $n \times p$, onde $\underset{\sim}{x}_i, i = 1, \dots, p$, são

vetores correspondentes às variáveis independentes;

$\underset{\sim}{\beta}' = [\beta_1, \dots, \beta_p]$ é o vetor de p parâmetros desconhecidos e

$\underset{\sim}{e} = [e_1, \dots, e_n]$ é um vetor aleatório, tal que $E(\underset{\sim}{e}) = \underset{\sim}{0}$ e $\text{Var}(\underset{\sim}{e}) = \sigma^2 I$.

Definição 1.1

Seja o modelo de GM e $\underset{\sim}{\lambda}$ um vetor $p \times 1$. A combinação linear $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$ é estimável, se existe um estimador de $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$ não viciado, função linear de $\underset{\sim}{y}$, isto é, se existe um vetor $\underset{\sim}{a}$, $n \times 1$, tal que $E(\underset{\sim}{a}' \underset{\sim}{y}) = \underset{\sim}{\lambda}' \underset{\sim}{\beta}$.

Uma consequência imediata da definição é que $\underset{\sim}{a}' \underset{\sim}{y}$ é um estimador não viciado para $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$, se, e somente se, $\underset{\sim}{\lambda}' = \underset{\sim}{a}' \underset{\sim}{X}$, isto é, se, e somente se, $\underset{\sim}{\lambda}' \in \mathcal{L}(X)$, onde $\mathcal{L}(X)$ é o espaço linha de X.

Existem outras condições necessárias e suficientes (equivalentes) para $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$ ser estimável. Uma destas é apresentada abaixo.

Resultado 1.1

Seja $s = p - r(X)$, onde $r(X)$ denota o posto da matriz X . Sejam $c_1, \dots, c_s$ s vetores, $p \times 1$, linearmente independentes, tais que $Xc_1 = \dots = Xc_s = \mathbf{0}$. $\lambda' \beta$ é estimável se, e somente se, $\lambda' c_1 = \dots = \lambda' c_s = 0$.

Prova : Harville, D. A. (1985).

■

O resultado 1.1 apresenta uma maneira prática de investigação da estimabilidade de $\lambda' \beta$, especialmente quando o valor de s é pequeno. Por exemplo, se $s = 1$, precisamos evidenciar apenas uma relação entre as colunas de X e verificar se a mesma relação ocorre para os correspondentes elementos do vetor λ' .

Algumas propriedades das funções estimáveis são listadas abaixo:

- todas combinações lineares das funções estimáveis são estimáveis;
- combinações lineares de funções não estimáveis podem ser estimáveis;
- se $r(X) = p$, então, todas funções $\lambda' \beta$ são estimáveis, inclusive, $\beta_1, \dots, \beta_p$;
- se $r(X) < p$, então, pelo menos um dos elementos de β não é estimável.

Exemplo 1.1 : Modelo de Classificação

Em alguns estudos é usada uma subdivisão do conjunto das unidades observadas em classes, baseada em algum critério. Estas classes são disjuntas e exaustivas. Ao critério usado é dado o nome de *Fator*, enquanto que, as classes de unidades correspondem aos *Níveis do Fator*. Por exemplo, suponha que o diretor, de um determinado colégio, deseja comparar o rendimento de alunos submetidos a um dentre

três métodos de ensino, A, B e C. O fator neste exemplo é o método de ensino e seus níveis A, B e C.

Se neste caso a variável de interesse, y , é modelada conforme o modelo de GM, a matriz X é denominada *matriz de delineamento* e usualmente sua primeira coluna é um vetor unitário, e as demais colunas são variáveis indicadoras dos níveis do fator.

Assim, para um fator com p níveis e n_i observações na classe do i -ésimo nível, $i = 1, \dots, p$, temos :

$$X = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ \sim & \sim & \sim & & \sim \\ 1 & 0 & 1 & \sim & 1 \\ \sim & \sim & \sim & & \sim \\ 1 & 1 & \sim & \sim & 0 \\ & & & & \sim \\ 1 & 0 & \dots & 0 & 1 \\ \sim & \sim & & \sim & \sim \end{bmatrix} \begin{array}{l} \} n_1 \text{ linhas} \\ \\ \} n_2 \text{ linhas} \\ \\ \\ \} n_p \text{ linhas} \end{array}$$

e

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix},$$

notando - se que o posto da matriz X é p e usando o resultado 1.1, com $s = 1$, $\lambda' \beta$ será estimável, para :

$$\lambda' = \left[\sum_{i=1}^p \lambda_i, \lambda_1, \dots, \lambda_p \right] .$$

Por exemplo, $\beta_0 + \beta_1$, $\beta_1 - \beta_2$, $\beta_1 + \beta_2 - 2\beta_3$ são estimáveis, no entanto β_0 , β_1 , ..., β_p , isoladamente, não são estimáveis.

□

Exemplo 1.2 : Modelo de Regressão

Na nossa natureza sempre estamos querendo explicar algum fenômeno. Há ocasiões, em que esta explicação pode ser obtida através de fatos relacionados a este fenômeno. Na nossa linguagem, nós chamamos estes fatos de variáveis explicativas, ou ainda, variáveis regressoras, as quais irão explicar uma determinada variável dependente. Simplificando tudo, nós queremos um modelo que mostre esta relação. A este modelo daremos o nome de modelo de regressão .

Um modelo de regressão com p variáveis regressoras é o modelo de Gauss - Markov, onde a matriz X é denominada *matriz de regressão* e usualmente sua primeira coluna é um vetor unitário, sendo as demais colunas compostas pelos valores da i -ésima variável regressora, $i = 1, \dots, p$.

Geralmente, a matriz X é de posto coluna completo, logo, $\lambda' \beta$ é estimável para qualquer vetor λ . Assim, β_0 e β_1 são estimáveis isoladamente para o modelo de GM

com uma variável regressora, onde:

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{e} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}.$$

□

Voltando a uma situação geral, sejam $\lambda_1', \dots, \lambda_m'$, vetores $1 \times p$, e seja

$$\Lambda_{m \times p} = \begin{bmatrix} \lambda_1' \\ \vdots \\ \lambda_m' \end{bmatrix} \quad \Rightarrow \quad \Lambda \beta = \begin{bmatrix} \lambda_1' \beta \\ \vdots \\ \lambda_m' \beta \end{bmatrix}.$$

Dizemos que o vetor $\Lambda \beta$ é estimável, se cada elemento $\lambda_i' \beta$ é estimável. Ou ainda, $\Lambda \beta$ é estimável se, e somente se, existe uma matriz A , $m \times n$, tal que $\Lambda = AX$. Então, considerando $A = I$, temos que $X \beta$ é sempre estimável.

Seja $Q: \mathbb{R}^p \rightarrow \mathbb{R}$ uma função definida por $Q(\mathbf{b}) = (y - X\mathbf{b})'(y - X\mathbf{b})$. O nosso objetivo é achar todos vetores de $\mathbb{R}^p$ que minimizem $Q(\mathbf{b})$. Uma maneira de achar o mínimo de $Q(\mathbf{b})$ é calcular as derivadas parciais de $Q(\mathbf{b})$, com respeito a $\mathbf{b}$, $\partial Q(\mathbf{b}) / \partial \mathbf{b}$, e resolver o sistema $\partial Q(\mathbf{b}) / \partial \mathbf{b} = \mathbf{0}$. Seria ainda necessário investigar

se os pontos extremos encontrados são pontos de mínimo. Assim, obtemos

$$\begin{aligned}
 \underline{\mathbf{Q}}(\underline{\mathbf{b}}) / \partial \underline{\mathbf{b}} &= -2\underline{\mathbf{X}}'\underline{\mathbf{y}} + 2\underline{\mathbf{X}}'\underline{\mathbf{X}}\underline{\mathbf{b}} \\
 \Rightarrow -2\underline{\mathbf{X}}'\underline{\mathbf{y}} + 2\underline{\mathbf{X}}'\underline{\mathbf{X}}\underline{\mathbf{b}} &= \underline{\mathbf{0}} \\
 \Rightarrow \underline{\mathbf{X}}'\underline{\mathbf{X}}\underline{\mathbf{b}} &= \underline{\mathbf{X}}'\underline{\mathbf{y}}.
 \end{aligned} \tag{1.3}$$

O sistema acima, em $\underline{\mathbf{b}}$, é denominado sistema de *Equações Normais* (EN).

Proposição 1.1

As equações normais são consistentes, isto é, existe pelo menos uma solução para o sistema (1.3).

Prova :

Pela proposição A2, apêndice A,

$$\underline{\mathbf{X}}' = \underline{\mathbf{X}}'\underline{\mathbf{P}}_{\mathbf{x}} = \underline{\mathbf{X}}'\underline{\mathbf{X}}(\underline{\mathbf{X}}'\underline{\mathbf{X}})^{-}\underline{\mathbf{X}}',$$

onde $(\underline{\mathbf{X}}'\underline{\mathbf{X}})^{-}$ é qualquer inversa generalizada de $\underline{\mathbf{X}}'\underline{\mathbf{X}}$ e $\underline{\mathbf{P}}_{\mathbf{x}}$ é a matriz de projeção sobre o espaço coluna de $\underline{\mathbf{X}}$. Então,

$$(\underline{\mathbf{X}}'\underline{\mathbf{X}})[(\underline{\mathbf{X}}'\underline{\mathbf{X}})^{-}\underline{\mathbf{X}}'\underline{\mathbf{y}}] = \underline{\mathbf{X}}'\underline{\mathbf{P}}_{\mathbf{x}}\underline{\mathbf{y}} = \underline{\mathbf{X}}'\underline{\mathbf{y}},$$

assim, $(\underline{\mathbf{X}}'\underline{\mathbf{X}})^{-}\underline{\mathbf{X}}'\underline{\mathbf{y}}$ é uma solução do sistema de EN.

■

Resultado 1.2

a) A função $Q(\underline{b})$ atinge o valor mínimo em $\underline{\hat{b}}$, se, e somente se $\underline{\hat{b}}$ é uma solução das equações normais.

b) $\underline{\lambda}'\underline{\beta}$ é estimável, se, e somente, se $\underline{\lambda}'\underline{\hat{b}}$ é invariante para qualquer solução, $\underline{\hat{b}}$, das EN.

Prova : *Harville, D. A. (1985).*

■

Os resultados apontados em 1.2 permitem a definição de um estimador para uma combinação linear dos elementos de $\underline{\beta}$, dada a seguir.

Definição 1.2

O estimador de mínimos quadrados ordinário, MQO, da função estimável $\underline{\lambda}'\underline{\beta}$ é definido por

$$\underline{\lambda}'\underline{\hat{b}} = \underline{\lambda}'(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}.$$

O estimador MQO do vetor de funções estimáveis $\underline{\Lambda}\underline{\beta}$ é definido por

$$\underline{\Lambda}\underline{\hat{b}} = \underline{\Lambda}(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}.$$

Proposição 1.2

O estimador de MQO de $X\beta$ é dado pela projeção ortogonal de y em $\mathcal{L}(X)$,

$$X\hat{\beta} = P_X y.$$

Prova :

Seja $\Lambda = X$. Então pela definição 1.2, temos que o estimador MQO de $X\beta$ é

$$\Lambda \hat{\beta} = X\hat{\beta} = X(X'X)^{-1}X'y = P_X y.$$

■

Exemplo 1.1 (Continuação)

Voltando ao modelo de classificação, onde a matriz do modelo e o vetor de observações são respectivamente

$$X = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ \sim & \sim & \sim & & \sim \\ 1 & 0 & 1 & \sim & \vdots \\ \sim & \sim & \sim & & \\ \vdots & \vdots & \sim & \sim & 0 \\ & & & & \sim \\ 1 & 0 & \dots & 0 & 1 \\ \sim & \sim & & \sim & \sim \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ \vdots \\ y_{p1} \\ \vdots \\ y_{pn_p} \end{bmatrix}$$

temos

$$X'X = \begin{bmatrix} \sum_{i=1}^k n_i & n_1 & n_2 & \dots & n_p \\ n_1 & n_1 & 0 & \dots & 0 \\ n_2 & 0 & n_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n_p & 0 & \dots & 0 & n_p \end{bmatrix}$$

$$\Rightarrow (X'X)^{-} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{n_1} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{n_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \frac{1}{n_p} \end{bmatrix},$$

uma inversa generalizada de $X'X$. Agora, temos que

$$X' y = \begin{bmatrix} \sum_{i=1}^p \sum_{j=1}^{n_i} y_{ij} \\ n_1 \sum_{j=1}^{n_1} y_{1j} \\ \vdots \\ n_p \sum_{j=1}^{n_p} y_{pj} \end{bmatrix}$$

logo, uma solução das EN é :

$$(X'X)^{-1}X'y = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{n_1} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{n_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \frac{1}{n_p} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^p \sum_{j=1}^{n_i} y_{ij} \\ n_1 \sum_{j=1}^{n_1} y_{1j} \\ \vdots \\ n_p \sum_{j=1}^{n_p} y_{pj} \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix},$$

e o estimador de MQO de $X\beta$ é

$$X(X'X)^{-1}X'y = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \sim & \sim & \sim & & \sim \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} \\ \sim & \sim & \sim & & \sim \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{1} \\ \sim & \sim & & \sim & \sim \end{bmatrix} \begin{bmatrix} 0 \\ \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix} = \begin{bmatrix} \bar{y}_{1.} \\ \vdots \\ \bar{y}_{1.} \\ \vdots \\ \bar{y}_{p.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix}.$$

Os estimadores de MQO das funções $\beta_0 + \beta_1$, $\beta_1 - \beta_2$ e $\beta_1 + \beta_2 - 2\beta_3$ apresentadas anteriormente são :

$$\bar{y}_{1.},$$

$$\bar{y}_{1.} - \bar{y}_{2.} \quad \text{e}$$

$$\bar{y}_1 + \bar{y}_2 - 2\bar{y}_3,$$

respectivamente. □

Exemplo 1.2 (Continuação)

Agora, voltando ao modelo de regressão, onde a matriz do modelo é

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},$$

temos

$$X'X = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}$$

$$\Rightarrow (X'X)^{-1} = \begin{bmatrix} \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} & \frac{-\sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i - \bar{x})^2} \\ \frac{-\sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i - \bar{x})^2} & \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{bmatrix},$$

a inversa de $X'X$.

Agora temos que

$$X' y = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}.$$

Logo, o estimador de MQO de β é

$$(X'X)^{-1}X' y = \begin{bmatrix} \frac{\sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j}{n \sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\sum_{i=1}^n x_i \sum_{j=1}^n x_j y_j}{n \sum_{i=1}^n (x_i - \bar{x})^2} \\ \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\sum_{i=1}^n x_i \sum_{j=1}^n y_j}{n \sum_{i=1}^n (x_i - \bar{x})^2} \end{bmatrix}$$

Mas,

$$\begin{aligned} & \frac{\sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j}{n \sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\sum_{i=1}^n x_i \sum_{j=1}^n x_j y_j}{n \sum_{i=1}^n (x_i - \bar{x})^2} = \\ & = \frac{\sum_{j=1}^n y_j (\sum_{i=1}^n x_i^2 - x_j \sum_{i=1}^n x_i)}{n \sum_{i=1}^n (x_i - \bar{x})^2} = \end{aligned}$$

$$= \frac{\sum_{j=1}^n y_j (\sum_{i=1}^n (x_i - \bar{x})^2 + n\bar{x}^2 - n x_j \bar{x})}{n \sum_{i=1}^n (x_i - \bar{x})^2}, \text{ porque } \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2,$$

$$= \bar{y} - \bar{x} \frac{\sum_{j=1}^n y_j (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} =$$

$$= \bar{y} - \bar{x} \frac{\sum_{j=1}^n (y_j - \bar{y})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

e

$$\frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i - \bar{x})^2} =$$

$$= \frac{n \sum_{i=1}^n y_i x_i - n \bar{x} \sum_{i=1}^n y_i}{n \sum_{i=1}^n (x_i - \bar{x})^2} =$$

$$\begin{aligned}
 &= \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\
 &= \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} .
 \end{aligned}$$

Então :

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = (X'X)^{-1} X'y = \begin{bmatrix} \bar{y} - \bar{x} \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{bmatrix} .$$

Agora, seja r o coeficiente de correlação amostral entre y e x , dado por

$$\frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{[\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2]^{\frac{1}{2}}} ,$$

então,

$$\hat{\beta}_1 = \frac{\left[\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} \right]^{\frac{1}{2}}}{\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right]^{\frac{1}{2}}} r.$$

□

Nesta Seção, abordamos a estimação de mínimos quadrados para contexto do modelo Gauss - Markov. Em seções seguintes iremos abordar a estimação de mínimos quadrados para contexto dos modelos de Aitken e Gauss - Markov restrito. O modelo de Aitken é uma generalização do modelo de GM, no tocante à matriz de variâncias e covariâncias do vetor de erros. O modelo de GM restrito adiciona uma restrição com respeito ao vetor de parâmetros, no contexto do modelo de GM. A questão da estimabilidade de combinação linear de parâmetros dos modelos de Aitken e Gauss - Markov restrito será investigada, tomando-se por base o conteúdo desta Seção.

1.2 Modelo de Aitken

Quando definimos o modelo de GM, supomos que, para qualquer observação da variável y , o erro correspondente tem a mesma variância. Algumas vezes, porém, não é adequada a suposição de que a matriz de variâncias e covariâncias do erro seja $\sigma^2 \mathbf{I}$. Consideremos todas as suposições do modelo GM, modificando apenas o seguinte: $\text{Var}(\mathbf{e}) = \sigma^2 \mathbf{V}$, para $\mathbf{V}$, matriz positiva definida (p.d.) conhecida. Tal modelo é conhecido como *modelo de Aitken*. A seguir vamos mostrar que o modelo de Aitken pode ser expresso como um modelo de GM. Desta forma, podemos usar os resultados

já apresentados na Seção anterior.

A matriz V é positiva definida. Portanto, sua inversa existe e nós podemos fatorá-la como $V^{-1} = Q'Q$, para alguma matriz Q , $n \times n$, não singular. Então, se premultiplicarmos ambos lados da equação (1.2) por Q , nós teremos

$$\underset{\sim}{Q} \underset{\sim}{y} = \underset{\sim}{Q} \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{Q} \underset{\sim}{e},$$

ou,

$$\underset{\sim}{y}^* = \underset{\sim}{X}^* \underset{\sim}{\beta} + \underset{\sim}{e}^*, \quad (1.4)$$

onde $\underset{\sim}{y}^* = \underset{\sim}{Q} \underset{\sim}{y}$, $\underset{\sim}{X}^* = \underset{\sim}{Q} \underset{\sim}{X}$ e $\underset{\sim}{e}^* = \underset{\sim}{Q} \underset{\sim}{e}$, observando que $E(\underset{\sim}{e}^*) = \underset{\sim}{0}$ e $\text{Var}(\underset{\sim}{e}^*) = \sigma^2 \underset{\sim}{Q} \underset{\sim}{V} \underset{\sim}{Q}' = \sigma^2 \underset{\sim}{I}$. Isto é, nós expressamos o modelo de Aitken, como um modelo de Gauss Markov.

Usando os resultados já apresentados sobre o modelo de GM, observamos que no modelo de Aitken, $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$ é estimável se $\underset{\sim}{\lambda} \in \mathcal{L}(\underset{\sim}{X}^*)$, ou seja, se $\underset{\sim}{\lambda}' \in \mathcal{L}(\underset{\sim}{Q} \underset{\sim}{X})$. Mas, a matriz Q é não singular, portanto, $\mathcal{L}(\underset{\sim}{Q} \underset{\sim}{X}) = \mathcal{L}(\underset{\sim}{X})$. Assim, podemos dizer que $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$ é estimável, sob o modelo de Aitken, se $\underset{\sim}{\lambda}' \in \mathcal{L}(\underset{\sim}{X})$.

O estimador de mínimos quadrados para $\underset{\sim}{\lambda}' \underset{\sim}{\beta}$, estimável, sob o modelo de GM de equação (1.4), é $\underset{\sim}{\lambda}' \underset{\sim}{\hat{b}}$, onde $\underset{\sim}{\hat{b}}$ é qualquer solução das EN,

$$\underset{\sim}{X}^{**} \underset{\sim}{X}^* \underset{\sim}{b} = \underset{\sim}{X}^{**} \underset{\sim}{y}^*,$$

conforme definição 1.2, da Seção anterior.

Definição 1.3

O estimador de mínimos quadrados para $\lambda'\beta$, estimável, sob o modelo de GM da equação (1.4), denominado de estimador de mínimos quadrados generalizado (MQG), com base em V , é

$$\lambda'(X^*X^*)^{-1}X^*y^*,$$

expressando em termos de X , y e V , temos

$$\lambda'(X^*X^*)^{-1}X^*y^* = \lambda'(X'Q'QX)^{-1}X'Q'Qy = \lambda'(X'V^{-1}X)^{-1}X'V^{-1}y.$$

O estimador MQG do vetor de funções estimáveis $\Lambda\beta$ é definido por

$$\Lambda(X^*X^*)^{-1}X^*y^* = \Lambda(X'Q'QX)^{-1}X'Q'Qy = \Lambda(X'V^{-1}X)^{-1}X'V^{-1}y.$$

As equações normais expressas em função de X , y e V , denominadas *equações normais de Aitken* (ENA) são :

$$(QX)'(QX)b = (QX)'(Qy),$$

ou,

$$X'V^{-1}Xb = X'V^{-1}y. \quad (1.5)$$

Proposição 1.3

O estimador de MQG de $\underline{\beta}$ é dado pela projeção não ortogonal, com base em V , de $\underline{y}$ em $\mathcal{L}(X)$, ou seja,

$$\underline{X\hat{\beta}} = X(X'V^{-1}X)^{-1}X'V^{-1}\underline{y} = \underline{A y},$$

onde $A = X(X'V^{-1}X)^{-1}X'V^{-1}$.

Prova :

Seja $\underline{\Lambda} = X$. Então, pela definição 1.3, temos que o estimador de MQG de $\underline{\beta}$ é

$$\underline{\Lambda\hat{\beta}} = \underline{X\hat{\beta}} = X(X'V^{-1}X)^{-1}X'V^{-1}\underline{y} = \underline{A y}.$$

Pelo Apêndice A, Proposição A11, temos que A é a matriz de projeção não ortogonal sobre $\mathcal{L}(X)$.

■

1.3 Modelo de Gauss - Markov Restrito

O modelo GM, com restrição no vetor de parâmetros $\underline{\beta}$, denominado modelo restrito, é definido por todas suposições do modelo de GM, com o acréscimo de

$$\underline{C\beta} = \underline{c},$$

onde o sistema, em $\underline{b}$, $\underline{Cb} = \underline{c}$ é consistente.

Sob o modelo restrito, vamos, agora, obter

$$\min_{\tilde{\mathbf{b}}} \mathbf{Q}(\tilde{\mathbf{b}}) = \min_{\tilde{\mathbf{b}}} (\tilde{\mathbf{y}} - \mathbf{X}\tilde{\mathbf{b}})'(\tilde{\mathbf{y}} - \mathbf{X}\tilde{\mathbf{b}}),$$

para $\tilde{\mathbf{b}} \in \mathbb{R}^p$, satisfazendo o sistema consistente $\tilde{\mathbf{C}}\tilde{\mathbf{b}} = \tilde{\mathbf{c}}$. Isto equivale a achar o mínimo em $\tilde{\mathbf{b}}$ e $\tilde{\mathbf{k}}$, para $\tilde{\mathbf{b}} \in \mathbb{R}^p$ e $\tilde{\mathbf{k}} \in \mathbb{R}^k$, da função

$$\mathbf{Q}^1(\tilde{\mathbf{b}}, \tilde{\mathbf{k}}) = (\tilde{\mathbf{y}} - \mathbf{X}\tilde{\mathbf{b}})'(\tilde{\mathbf{y}} - \mathbf{X}\tilde{\mathbf{b}}) + 2\tilde{\mathbf{k}}'(\tilde{\mathbf{C}}\tilde{\mathbf{b}} - \tilde{\mathbf{c}}),$$

onde $\tilde{\mathbf{k}}$ é um vetor de multiplicadores de Lagrange. As equações normais no modelo restrito são definidas por

$$\partial \mathbf{Q}(\tilde{\mathbf{b}}, \tilde{\mathbf{k}}) / \partial \tilde{\mathbf{b}} = -2\mathbf{X}'\tilde{\mathbf{y}} + 2\mathbf{X}'\mathbf{X}\tilde{\mathbf{b}} + 2\tilde{\mathbf{C}}'\tilde{\mathbf{k}} = \mathbf{0}$$

e

$$\partial \mathbf{Q}(\tilde{\mathbf{b}}, \tilde{\mathbf{k}}) / \partial \tilde{\mathbf{k}} = 2(\tilde{\mathbf{C}}\tilde{\mathbf{b}} - \tilde{\mathbf{c}}) = \mathbf{0}.$$

Assim, nós teremos o seguinte sistema de *equações normais restritas* (ENR):

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \tilde{\mathbf{C}}' \\ \tilde{\mathbf{C}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{b}} \\ \tilde{\mathbf{k}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\tilde{\mathbf{y}} \\ \tilde{\mathbf{c}} \end{bmatrix}$$

¹Por simplicidade, usamos a mesma notação $\mathbf{Q}(\cdot)$ para esta função de $\tilde{\mathbf{b}}$ e $\tilde{\mathbf{k}}$.

ou ainda,

$$\begin{cases} \tilde{X}'\tilde{X}\tilde{b} + \tilde{C}'\tilde{k} = \tilde{X}'\tilde{y} \\ \tilde{C}\tilde{b} = \tilde{c} \end{cases} \quad (1.6)$$

Agora apresentaremos alguns resultados importantes, considerando o modelo restrito, que serão úteis para o presente trabalho.

Resultado 1.3

a) As equações normais restritas são consistentes.

b) $Q(\tilde{b}, \tilde{k})$ atinge o mínimo em $\tilde{b} = \hat{\tilde{b}}_r$ e $\tilde{k} = \hat{\tilde{k}}$ se, e somente se, as equações normais restritas tem uma solução cujas primeiras p componentes são $\hat{\tilde{b}}_r$.

Prova : Harville, D. A. (1985) .

■

Sob o modelo de GM restrito, $\tilde{\lambda}'\tilde{\beta}$ é estimável, se, e somente se, $\tilde{\lambda}' \in \mathcal{Q}\begin{bmatrix} \tilde{X} \\ \tilde{C} \end{bmatrix}$. Portanto, todas funções $\tilde{\lambda}'\tilde{\beta}$, estimáveis sob o modelo GM, irrestrito, são também estimáveis no modelo restrito.

Proposição 1.4

Se a matriz C é de posto linha completo e $C\beta$ é estimável, os vetores

$$\hat{k} = [C(X'X)^{-1}C']^{-1}(C\hat{b} - c)$$

e

$$\hat{b}_r = \hat{b} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C\hat{b} - c)$$

compõe uma solução para as ENR, do modelo de GM sujeito à restrição $C\beta = c$, onde $\hat{b}$ é qualquer solução do sistema de EN do modelo GM (irrestrito).

Prova :

Notemos que C é de posto linha completo $\Rightarrow C(X'X)^{-1}C'$ invertível.

Agora a primeira parte do sistema de ENR é :

$$X'X\hat{b} + C'\hat{k} = X'y.$$

Substituindo os vetores $\hat{b}_r$ e $\hat{k}$ acima, em $\hat{b}$ e $\hat{k}$, temos :

$$\begin{aligned} X'X\hat{b}_r + C'\hat{k} &= X'X(\hat{b} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C\hat{b} - c)) \\ &\quad + C'([C(X'X)^{-1}C']^{-1}(C\hat{b} - c)). \end{aligned} \quad (1.7)$$

Notemos que ,

$$X'X\hat{b} = X'y$$

e

$$\begin{aligned} X'X(X'X)^{-1}C' &= X'X(X'X)^{-1}X'W', \text{ porque } C \in \mathcal{C}(X) \Rightarrow C = WX, \\ &\text{para alguma matriz } W, \\ &= X'W' \\ &= C'. \end{aligned}$$

Logo, (1.7) é dado por :

$$X' \underset{\sim}{y} - C'[C(X'X)^{-1}C']^{-1}(C \underset{\sim}{\hat{b}} - \underset{\sim}{c}) + C'[C(X'X)^{-1}C']^{-1}(C \underset{\sim}{\hat{b}} - \underset{\sim}{c}) = X' \underset{\sim}{y},$$

ficando a primeira parte do sistema de ENR satisfeita .

A segunda parte do sistema de ENR é :

$$C \underset{\sim}{b} = \underset{\sim}{c}.$$

Substituindo o vetor $\underset{\sim}{\hat{b}}$, no lugar de $\underset{\sim}{b}$, temos :

$$\begin{aligned} C[\underset{\sim}{\hat{b}} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C \underset{\sim}{\hat{b}} - \underset{\sim}{c})] &= C \underset{\sim}{\hat{b}} - C(X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C \underset{\sim}{\hat{b}} - \underset{\sim}{c}) \\ &= C \underset{\sim}{\hat{b}} - (C \underset{\sim}{\hat{b}} - \underset{\sim}{c}) \\ &= \underset{\sim}{c}, \end{aligned}$$

ficando satisfeita a segunda parte do sistema de ENR.

Assim, o vetor das primeiras p componentes de uma solução das ENR é igual ao vetor de solução das EN (irrestritas), menos uma certa quantidade.

Definição 1.4

O estimador de mínimos quadrados da função estimável $\lambda' \beta$, sob o modelo de GM restrito, denominado estimador de *mínimos quadrados restrito* (MQR), é definido por $\lambda' \hat{\beta}_r$, onde $\hat{\beta}_r$ corresponde às primeiras p componentes de qualquer solução das equações normais restritas.

Em ocasiões especiais, dada uma restrição $C\beta = 0$, podemos definir um modelo denominado *modelo reduzido*. Este modelo incorpora ao antigo vetor de parâmetros a condição $C\beta = 0$ e utiliza uma nova matriz de modelo, cujo espaço coluna é subconjunto do espaço coluna da matriz do modelo irrestrito.

1.4 Estimador da variância do erro

Considerando o modelo de GM, verificamos que a esperança de y , $X\beta$, é sempre estimável e o estimador de MQO é $P_x y$, onde $P_x = X(X'X)^- X'$, para qualquer inversa generalizada de $X'X$, $(X'X)^-$.

Definimos o vetor de erro², $\hat{e}$, como a diferença entre o vetor de observações y

² Chamaremos de vetor de erro o que seria de fato o vetor de resíduos. Como estamos usando a notação $\hat{e}$, não haverá confusão com o vetor erro, e . Desta forma, estamos seguindo o usual na literatura.

e o estimador de sua esperança :

$$\begin{aligned}\underline{\hat{e}} &= \underline{y} - \underline{\hat{E}}(\underline{y}) \\ &= \underline{y} - \underline{P_x y} \\ &= (\underline{I} - \underline{P_x})\underline{y}\end{aligned}\tag{1.8}$$

Uma suposição relativa ao erro foi $\text{Var}(\underline{e}) = \sigma^2 \underline{I}$. Logo, é intuitivo usarmos o vetor de erro $(\underline{I} - \underline{P_x})\underline{y}$ na estimação de σ^2 . Consideremos

$$(\underline{y} - \underline{X}\hat{\underline{\beta}})'(\underline{y} - \underline{X}\hat{\underline{\beta}}) = (\underline{y} - \underline{P_x y})'(\underline{y} - \underline{P_x y}) = \underline{y}'(\underline{I} - \underline{P_x})'(\underline{I} - \underline{P_x})\underline{y} = \underline{y}'(\underline{I} - \underline{P_x})\underline{y},$$

denominada *soma de quadrados do erro* (SSE). Seja

$$\underline{y}'(\underline{I} - \underline{P_x})\underline{y} / r(\underline{I} - \underline{P_x}),$$

denominado *quadrado médio do erro* (MSE), e $r(\underline{I} - \underline{P_x})$, denominado *graus de liberdade* associados ao erro. A estatística MSE é um estimador não viciado de σ^2 , como apresentaremos no teorema 1.1.

Teorema 1.1

Um estimador não viciado de σ^2 , no contexto do modelo de GM, é

$$\underline{y}'(\underline{I} - \underline{P_x})\underline{y} / (n - r(\underline{X}))$$

Prova :

Pelas suposições do modelo de GM e pelo teorema A2, apêndice A,

$$E[\underline{y}'(I - P_x)\underline{y}] = \text{Tr}(\sigma^2(I - P_x)) + \underline{\beta}'X(I - P_x)X\underline{\beta}.$$

Mas,

$$\text{Tr}(\sigma^2(I - P_x)) = \sigma^2 \text{Tr}(I - P_x) = \sigma^2 r(I - P_x)$$

e

$$(I - P_x)X = 0$$

$$\Rightarrow \underline{\beta}'X(I - P_x)X\underline{\beta} = 0.$$

Portanto

$$E[\underline{y}'(I - P_x)\underline{y}] = \sigma^2 r(I - P_x)$$

e

$$E[\underline{y}'(I - P_x)\underline{y} / r(I - P_x)] = \sigma^2.$$

Mas, no apêndice A, proposição A6, $r(I - P_x) = n - r(X)$. Logo, segue o resultado. ■

Consideremos agora o estimador de σ^2 no contexto do modelo de Aitken. Na Seção 1.2, fizemos uma transformação no modelo de Aitken, de maneira que passamos a usar os resultados do modelo Gauss - Markov. Vamos agora utilizar essa transformação para a estimação de σ^2 .

³ Tr é a notação para o traço de uma matriz.

Proposição 1.5

No contexto do modelo de Aitken, o quadrado médio do erro (MSE_A),

$$\frac{\underline{y}'(I - A)V^{-1}(I - A)\underline{y}}{n - r(X)},$$

é um estimador não viciado de σ^2 , onde $A = X(X'V^{-1}X)^{-1}X'V^{-1}$.

Prova :

Pelo teorema 1.1, temos o seguinte estimador não viciado de σ^2 , após transformação do modelo de Aitken :

$$\underline{y}^*(I - P_{X^*})\underline{y}^* / (n - r(X^*)),$$

onde $\underline{y}^* = Q\underline{y}$, $X^* = QX$, e, $Q'Q = V^{-1}$.

Expressando em termos de $\underline{y}$, X , V e

$$A = X(X'V^{-1}X)^{-1}X'V^{-1},$$

$$P_{X^*} = X^*(X^{*'}X^*)^{-1}X^{*'} = QX(X'V^{-1}X)^{-1}X'Q'$$

temos que

$$\begin{aligned}\underline{y}^*(I - P_{X^*})\underline{y}^* &= \underline{y}'Q'[I - QX(X'V^{-1}X)^{-1}X'Q']Q\underline{y} \\ &= \underline{y}'[Q'Q - Q'QX(X'V^{-1}X)^{-1}X'Q']\underline{y} \\ &= \underline{y}'[V^{-1} - Q'QX(X'V^{-1}X)^{-1}X'V^{-1}]\underline{y} \\ &= \underline{y}'[V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}]\underline{y} \\ &= \underline{y}'[V^{-1} - V^{-1}A]\underline{y}\end{aligned}$$

$$\begin{aligned}
&= \underline{y}' \underline{V}^{-1} [\underline{I} - \underline{A}] \underline{y} \\
&= \underline{y}' \underline{V}^{-1} \underline{y} - \underline{y}' \underline{V}^{-1} \underline{A} \underline{y}.
\end{aligned} \tag{1.9}$$

Alternativamente, notamos que

$$\underline{V}^{-1} \underline{A} = \underline{V}^{-1} \underline{X} (\underline{X}' \underline{V}^{-1} \underline{X})^{-1} \underline{X}' \underline{V}^{-1} = \underline{A}' \underline{V}^{-1}$$

e usando o fato de $\underline{I} - \underline{A}$ ser idempotente,

$$\begin{aligned}
\underline{V}^{-1} [\underline{I} - \underline{A}] &= \underline{V}^{-1} [\underline{I} - \underline{A}] [\underline{I} - \underline{A}] \\
&= [\underline{V}^{-1} - \underline{V}^{-1} \underline{A}] [\underline{I} - \underline{A}] \\
&= [\underline{V}^{-1} - \underline{A}' \underline{V}^{-1}] [\underline{I} - \underline{A}] \\
&= [\underline{I} - \underline{A}]' \underline{V}^{-1} [\underline{I} - \underline{A}] .
\end{aligned} \tag{1.10}$$

Assim, podemos expressar o numerador do estimador de σ^2 como :

$$\underline{y}' (\underline{I} - \underline{A})' \underline{V}^{-1} (\underline{I} - \underline{A}) \underline{y}.$$

Também temos

$$r(\underline{X}^*) = r(\underline{X}),$$

logo, um estimador não viciado para σ^2 é dado por

$$(\underline{y}' \underline{V}^{-1} \underline{y} - \underline{y}' \underline{V}^{-1} \underline{A} \underline{y}) / (n - r(\underline{X})),$$

ou

$$\underline{y}' (I - A)' V^{-1} (I - A) \underline{y} / (n - r(X)),$$

onde $A = X(X'V^{-1}X)^{-1}X'V^{-1}$.

■

Consideremos agora o estimador de σ^2 no contexto do modelo GM restrito. Seja $\hat{\beta}_r$ uma solução da primeira parte do sistema das ENR e $\hat{\beta}$ uma solução das EN, sob o modelo de GM (irrestrito), então,

$$\begin{aligned} SSE_r &= (\underline{y} - X\hat{\beta} + X(\hat{\beta} - \hat{\beta}_r))' (\underline{y} - X\hat{\beta} + X(\hat{\beta} - \hat{\beta}_r)) \\ &= SSE + (\hat{\beta} - \hat{\beta}_r)' X' X (\hat{\beta} - \hat{\beta}_r) + 2(\hat{\beta} - \hat{\beta}_r)' X' (\underline{y} - X\hat{\beta}), \end{aligned}$$

mas $X'(\underline{y} - X\hat{\beta}) = X'\underline{y} - X'X\hat{\beta} = 0$, pois $X'X\hat{\beta} = X'\underline{y}$, EN do modelo de GM (irrestrito), logo,

$$SSE_r = SSE + (\hat{\beta} - \hat{\beta}_r)' X' X (\hat{\beta} - \hat{\beta}_r). \quad (1.11)$$

Notemos que a forma quadrática da segunda parcela não é negativa, e observemos que a imposição de restrições provoca, em geral, um aumento na soma de quadrados do erro, pois dificilmente $C\hat{\beta} = \underline{c}$. Isto é, observamos que a soma de quadrados do erro no modelo de GM restrito é geralmente maior que a soma de quadrados do erro no modelo de GM (irrestrito).

Consideremos agora, $C\beta$ estimável, sendo a matriz C de posto linha completo.

Desta forma $C \in \mathcal{L}(X)$. A soma de quadrados do erro (SSE_r), sob o modelo de GM restrito, expressa em termos de uma solução das EN, sob o modelo de GM (irrestrito), $\hat{\underline{\beta}}$, e usando

$$D = (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C\hat{\underline{\beta}} - \underline{c}),$$

será:

$$\begin{aligned} SSE_r &= SSE + (\hat{\underline{\beta}} - \hat{\underline{\beta}}_r)'X'X(\hat{\underline{\beta}} - \hat{\underline{\beta}}_r), \quad \text{por (1.11),} \\ &= SSE + [\hat{\underline{\beta}} - (\hat{\underline{\beta}} - D)]'X'X[\hat{\underline{\beta}} - (\hat{\underline{\beta}} - D)], \quad \text{pela proposição 1.4,} \\ &= SSE + D'X'XD \\ &= SSE + (C\hat{\underline{\beta}} - \underline{c})'[C(X'X)^{-1}C']^{-1}(C\hat{\underline{\beta}} - \underline{c}). \end{aligned} \quad (1.12)$$

Resultado 1.4

No contexto do modelo de GM, restrito a $C\underline{\beta} = \underline{c}$, C de posto completo, o estimador abaixo é não viciado para σ^2 :

$$MSE_r = \frac{SSE + (C\hat{\underline{\beta}} - \underline{c})'[C(X'X)^{-1}C']^{-1}(C\hat{\underline{\beta}} - \underline{c})}{n - r(X) + r(C)}.$$

Prova : Harville, D. A. (1985).

■

1.5 Análise de Variância - ANOVA

Seja $\underline{y}$, $n \times 1$, um vetor de observações .

Sejam $A_1, \dots, A_k$, k matrizes, $n \times n$, com $r(A_i) = r_i$, $i = 1, \dots, k$.

Uma análise de variância, ANOVA, é definida como uma partição da soma de quadrados total do vetor $\underline{y}$, isto é, $\underline{y}'\underline{y} = \underline{y}'\underline{A}_1\underline{y} + \dots + \underline{y}'\underline{A}_k\underline{y}$, tal que $r_1 + \dots + r_k = n$. Notemos que esta igualdade implica em $\sum_{i=1}^k \underline{A}_i = \underline{I}$.

Esta partição é comumente organizada em uma tabela, a qual chamaremos de *Tabela de ANOVA*.

Tabela ANOVA 2.1

Fonte de Variação(FV)	Graus de Liberdade(GL)	Soma de Quadrados(SS)	Quadrado Médio (MS)
1	r_1	$\underline{y}'\underline{A}_1\underline{y}$	$\underline{y}'\underline{A}_1\underline{y}/r_1$
...	...	...	...
k	r_k	$\underline{y}'\underline{A}_k\underline{y}$	$\underline{y}'\underline{A}_k\underline{y}/r_k$
Total	n	$\underline{y}'\underline{y}$	—

Resultado 1.5

Sejam $\underline{A}_1, \dots, \underline{A}_k$, matrizes $n \times n$, simétricas e $\sum_{i=1}^k \underline{A}_i = \underline{I}$. Então, as seguintes condições são equivalentes.

a) $\underline{A}_i \underline{A}_j = 0$, $\forall i \neq j$ e $i, j = 1, \dots, k$.

b) $\underline{A}_i$ idempotente, $i = 1, \dots, k$.

c) $\sum_{i=1}^k r(\underline{A}_i) = n$.

Prova : *Graybill, F. A.*(1961) .

O resultado 1.5 nos proporciona meios alternativos para verificarmos se uma partição de $\tilde{y}'\tilde{y}$ é de fato uma análise de variância.

Proposição 1.6

Sob o modelo GM ,

$$\tilde{y}'\tilde{y} = \tilde{y}'P_x\tilde{y} + \tilde{y}'(I - P_x)\tilde{y}$$

é uma ANOVA, onde P_x é a matriz de projeção ortogonal sobre $\mathcal{L}(X)$.

Prova :

Temos que

$$P_x \text{ e } I - P_x$$

são simétricas e

$$P_x + I - P_x = I \text{ e } P_x (I - P_x) = 0 ,$$

então, pela condição (a) do resultado 1.5, provamos a proposição acima.

■

Vimos na Seção 1.4 que $\tilde{y}'(I - P_x)\tilde{y}$ é a soma de quadrados do erro. Logo o que resta de $\tilde{y}'\tilde{y}$, $\tilde{y}'P_x\tilde{y}$, é atribuído ao modelo. Por isso, $\tilde{y}'P_x\tilde{y}$ é denominada *soma de quadrados do modelo*.

Proposição 1.7

Sob o modelo GM,

$$\underset{\sim}{y}' \underset{\sim}{y} = \underset{\sim}{y}' \underset{\sim}{P}_{x_0} \underset{\sim}{y} + \underset{\sim}{y}' (\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0}) \underset{\sim}{y} + \underset{\sim}{y}' (\underset{\sim}{I} - \underset{\sim}{P}_x) \underset{\sim}{y}$$

é uma ANOVA, onde $\mathcal{E}(X_0) \subset \mathcal{E}(X)$ e P_{x_0} é a projeção ortogonal em $\mathcal{E}(X_0)$.

Prova :

Temos que P_{x_0} , $P_x - P_{x_0}$ e $I - P_x$ são simétricas e

$$P_{x_0} + (P_x - P_{x_0}) + (I - P_x) = I.$$

Temos $P_{x_0} P_x = (P_{x_0} P_x)' = P_x P_{x_0} = P_{x_0}$

porque as colunas de P_{x_0} pertencem a $\mathcal{E}(X_0) \subset \mathcal{E}(X)$. Logo,

$$P_{x_0} (P_x - P_{x_0}) = 0,$$

$$P_{x_0} (I - P_x) = 0$$

e

$$(P_x - P_{x_0})(I - P_x) = P_x - P_x - P_{x_0} + P_{x_0} P_x = 0,$$

o resultado segue pela condição (a) do resultado 1.5.

■

Proposição 1.8

Sob o modelo GM,

$$\underline{\underline{y}}' \underline{\underline{y}} = \frac{1}{n} \underline{\underline{y}}' (\underline{\underline{1}} \underline{\underline{1}}') \underline{\underline{y}} + \underline{\underline{y}}' \left(\underline{\underline{P}}_x - \frac{1}{n} \underline{\underline{1}} \underline{\underline{1}}' \right) \underline{\underline{y}} + \underline{\underline{y}}' (\underline{\underline{I}} - \underline{\underline{P}}_x) \underline{\underline{y}}$$

é ANOVA se e somente se $\underline{\underline{1}} \in \mathcal{L}(X)$.

Prova :

Seja $X_0 = \underline{\underline{1}}$. Então, a matriz de projeção ortogonal em $\mathcal{L}(X_0)$ é

$$\underline{\underline{P}}_{x_0} = \underline{\underline{1}} (\underline{\underline{1}}' \underline{\underline{1}})^{-1} \underline{\underline{1}}' = \frac{1}{n} \underline{\underline{1}} \underline{\underline{1}}'.$$

Temos que

$$\underline{\underline{X}}_0 = \underline{\underline{1}} \in \mathcal{L}(X) \Leftrightarrow \mathcal{L}(X_0) \subset \mathcal{L}(X),$$

e o resultado segue a proposição 1.7. ■

O seguinte teorema é um clássico na literatura estatística: Teorema de Cochran. Ele pode ser utilizado para definir as distribuições das variáveis aleatórias correspondentes às parcelas da análise de variância, divididas por σ^2 , sob o modelo GM normal.

Teorema 1.2**Teorema de Cochran**

Seja $\underline{y}$ um vetor aleatório, $n \times 1$, com distribuição $N(\underline{\mu}, \sigma^2 \underline{I})$.

Sejam $\underline{A}_1, \dots, \underline{A}_k$ matrizes simétricas, tal que $\sum_{i=1}^k \underline{A}_i = \underline{I}_n$, ou equivalentemente, tal

que $\sum_{i=1}^k \underline{y}' \underline{A}_i \underline{y} = \underline{y}' \underline{y}$ para todos valores de $\underline{y}$.

Seja $r(\underline{A}_i) = r_i$, $i = 1, \dots, k$, e $\sum_{i=1}^k r_i = n$.

a) $\underline{y}' \underline{A}_i \underline{y} \sim \chi^2_{(r_i; \underline{\mu}' \underline{A}_i \underline{\mu})}$, $i = 1, \dots, k$.

b) $\underline{y}' \underline{A}_1 \underline{y}, \dots, \underline{y}' \underline{A}_k \underline{y}$ são distribuídas independentemente.

Prova : *Harville, D. A. (1985).*

■

Vimos nesta seção que, quando decompomos a soma de quadrados total do vetor $\underline{y}$ em uma análise de variância, estamos lidando com um parcelamento bem especial. Com o conhecimento da distribuição das somas de quadrados, torna-se possível a comparação de certos modelos, conforme será visto na próxima Seção.

1.6 Teste de Hipóteses

Esta Seção apresenta resultados sobre testes de hipóteses sobre os parâmetros da equação básica do modelo GM. Na literatura encontramos duas versões: uma versão trata da escolha entre um modelo mais complexo e uma simplificação deste, e, a outra versão define um teste de hipóteses da maneira convencional, baseado em uma variável pivô. As duas versões são aqui expostas e apontamos em seguida que, se para uma

dada hipótese podemos exibir o modelo reduzido correspondente à hipótese , então, as duas versões são equivalentes. A obtenção do valor observado da estatística do teste (ambas versões) pode ser organizada através da tabela de ANOVA.

1.6.1 Comparando Modelos Completo e Reduzido

Em *Christensen* (1987), Capítulo III, é fornecida uma versão intuitiva para o uso de uma estatística no processo de escolha entre modelos, um uma simplificação do outro. Apresentamos um resumo seguido de um teorema com resultados distribucionais de tal estatística.

Consideremos o modelo de GM, normal,

$$\underline{y} = \underline{X}\underline{\beta} + \underline{e} , \quad \underline{e} \sim N(\underline{0}, \sigma^2 \underline{I}),$$

definido em (1.2). Seja agora , um modelo de GM reduzido, uma simplificação do modelo anterior :

$$\underline{y} = \underline{X}_0 \underline{\delta} + \underline{e} , \quad \underline{e} \sim N(\underline{0}, \sigma^2 \underline{I}), \quad (1.15)$$

reduzido, porque $\mathcal{L}(\underline{X}_0) \subset \mathcal{L}(\underline{X})$, sendo $\underline{\delta}$ um vetor $p_0 \times 1$, $p_0 < p$.

Exemplo 1.1 (Continuação)

Considere o modelo de GM exemplo 1.1 onde

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \sim & \sim & \sim & & \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \sim & \vdots \\ \sim & \sim & \sim & & \\ \vdots & \vdots & \sim & \sim & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{1} \\ \sim & \sim & \sim & & \end{bmatrix} \quad \text{e} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Um modelo de GM reduzido é dado por

$$X_0 = \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \sim & \sim \\ \mathbf{1} & \mathbf{1} \\ \sim & \sim \\ \vdots & \vdots \\ \mathbf{1} & \mathbf{1} \\ \sim & \sim \end{bmatrix} \quad \text{e} \quad \delta = \begin{bmatrix} \delta_0 \\ \delta_1 \end{bmatrix},$$

notando que $\mathcal{L}(X_0) \subset \mathcal{L}(X)$.

□

Seja P_{x_0} a matriz de projeção ortogonal sobre $\mathcal{L}(X_0)$ e P_x a matriz de projeção ortogonal sobre $\mathcal{L}(X)$. Sob o modelo (1.15) ambos $P_{x_0}y$ e $P_x y$ são estimadores não viciados da mesma quantidade, porque :

$$E(P_{x_0}y) = P_{x_0}X_0\delta = X_0\delta$$

e

$$E(\underset{\sim}{P_x} \underset{\sim}{X_0} \underset{\sim}{y}) = E(\underset{\sim}{P_x} \underset{\sim}{P_{x_0}} \underset{\sim}{X_0} \underset{\sim}{y}) = E(\underset{\sim}{P_{x_0}} \underset{\sim}{X_0} \underset{\sim}{y}) = \underset{\sim}{X_0} \underset{\sim}{\delta},$$

pelo apêndice A, proposição A2. Assim sob o modelo (1.15),

$$\underset{\sim}{P_x} \underset{\sim}{y} - \underset{\sim}{P_{x_0}} \underset{\sim}{y} = (\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y}$$

deveria ser relativamente pequena. A escolha pelo modelo mais simplificado (1.15), depende da "grandeza" do vetor $(\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y}$, grandeza esta que poderia ser medida por função da soma dos quadrados de seus componentes,

$$[(\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y}]' [(\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y}] = \underset{\sim}{y}' (\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y}.$$

A estatística

$$\underset{\sim}{y}' (\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}}) \underset{\sim}{y} / r(\underset{\sim}{P_x} - \underset{\sim}{P_{x_0}})$$

incorporaria aspectos dos espaços $\mathcal{L}(X)$ e $\mathcal{L}(X_0)$.

Teorema 1.3

Sejam dois modelos GM normais , para o mesmo vetor de observações :

$$\underset{\sim}{y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{e} , \quad \underset{\sim}{e} \sim N(\underset{\sim}{0}, \sigma^2 I), \quad \text{modelo completo}$$

e

$$\underset{\sim}{y} = \underset{\sim}{X}_0 \underset{\sim}{\delta} + \underset{\sim}{e} , \quad \underset{\sim}{e} \sim N(\underset{\sim}{0}, \sigma^2 I), \quad \text{modelo reduzido}$$

onde $\mathcal{L}(\underset{\sim}{X}_0) \subset \mathcal{L}(\underset{\sim}{X})$.

Sejam P_x e P_{x_0} as matrizes de projeção ortogonal sobre $\mathcal{L}(\underset{\sim}{X})$ e $\mathcal{L}(\underset{\sim}{X}_0)$, respectivamente. Então,

a) sob o modelo completo ,

$$\frac{\underset{\sim}{y}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{y} / r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})}{\underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{y} / r(\underset{\sim}{I} - \underset{\sim}{P}_x)} \sim F_{(r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0}), r(\underset{\sim}{I} - \underset{\sim}{P}_x), \underset{\sim}{\beta}'\underset{\sim}{X}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{X}\underset{\sim}{\beta})} ;$$

b) sob o modelo reduzido

$$\frac{\underset{\sim}{y}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{y} / r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})}{\underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{y} / r(\underset{\sim}{I} - \underset{\sim}{P}_x)} \sim F_{(r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0}), r(\underset{\sim}{I} - \underset{\sim}{P}_x), 0)} . \quad (1.16)$$

Prova :

a) Pela proposição 1.7, a partição

$$\underset{\sim}{y}'\underset{\sim}{y} = \underset{\sim}{y}'\underset{\sim}{P}_{x_0}\underset{\sim}{y} + \underset{\sim}{y}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{y} + \underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{y}$$

é uma ANOVA . E pelo teorema de Cochran

$$\frac{\underset{\sim}{y}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{y}}{\underset{\sim}{\sigma}^2} \sim \chi^2_{(r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0}), [E(\frac{1}{\underset{\sim}{\sigma}^2})]'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})[E(\frac{1}{\underset{\sim}{\sigma}^2})]}$$

e

$$\frac{\underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{y}}{\underset{\sim}{\sigma}^2} \sim \chi^2_{(r(\underset{\sim}{I} - \underset{\sim}{P}_x), [E(\frac{1}{\underset{\sim}{\sigma}^2})]'(\underset{\sim}{I} - \underset{\sim}{P}_x)[E(\frac{1}{\underset{\sim}{\sigma}^2})])}$$

além de serem independentes. Agora, sob o modelo completo, $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{\beta}$, logo,

$$\frac{\underset{\sim}{y}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{y}}{\underset{\sim}{\sigma}^2} \sim \chi^2_{(r(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0}), \frac{1}{\underset{\sim}{\sigma}^2}\underset{\sim}{\beta}'\underset{\sim}{X}'(\underset{\sim}{P}_x - \underset{\sim}{P}_{x_0})\underset{\sim}{X}\underset{\sim}{\beta})}$$

e

$$\frac{\underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{y}}{\underset{\sim}{\sigma}^2} \sim \chi^2_{(r(\underset{\sim}{I} - \underset{\sim}{P}_x), \frac{1}{\underset{\sim}{\sigma}^2}\underset{\sim}{\beta}'\underset{\sim}{X}'(\underset{\sim}{I} - \underset{\sim}{P}_x)\underset{\sim}{X}\underset{\sim}{\beta})}$$

Notando que, $\underset{\sim}{\beta}' X' (I - P_x) X \underset{\sim}{\beta} = \underset{\sim}{\beta}' (X' - X' P_x) X \underset{\sim}{\beta} = \underset{\sim}{\beta}' (X' - X') X \underset{\sim}{\beta} = 0$,

$$\frac{\underset{\sim}{y}' (I - P_x) \underset{\sim}{y}}{\sigma^2} \sim \chi^2_{(r(I - P_x), 0)}.$$

O resultado (a) segue porque temos a razão entre duas distribuições Qui - quadradas independentes, divididas pelos respectivos graus de liberdade, sendo a do denominador, Qui - quadrada central.

b) Sob o modelo reduzido,

$$E\left(\frac{1}{\sigma} \underset{\sim}{y}\right) = \frac{1}{\sigma} X_0 \underset{\sim}{\delta}.$$

Logo, os parâmetros de não - centralidade são

$$\begin{aligned} E\left(\frac{1}{\sigma} \underset{\sim}{y}\right)' (P_x - P_{x0}) E\left(\frac{1}{\sigma} \underset{\sim}{y}\right) &= \frac{1}{\sigma^2} \underset{\sim}{\delta}' X_0' (P_x - P_{x0}) X_0 \underset{\sim}{\delta} \\ &= \frac{1}{\sigma^2} \underset{\sim}{\delta}' (X_0' P_x - X_0' P_{x0}) X_0 \underset{\sim}{\delta} \\ &= 0, \end{aligned}$$

e

$$E \left(\frac{1}{\sigma^2} \mathbf{y}' \right) (I - P_x) E \left(\frac{1}{\sigma^2} \mathbf{y} \right) = \frac{1}{\sigma^2} \boldsymbol{\delta}' \mathbf{X}_0' (I - P_x) \mathbf{X}_0 \boldsymbol{\delta} \\ = 0.$$

Logo, segue o resultado (b). ■

Então, a escolha pelo modelo reduzido, ao invés do completo, pode ser feita quando a estatística (1.16) é menor do que algum quantil da distribuição

$$F_{((r(P_x - P_{x_0}), r(I - P_x))}.$$

1.6.2 Testando uma Hipótese

Definição 1.6

Considere o modelo de GM. Seja Λ uma matriz $m \times p$. A hipótese $\Lambda \boldsymbol{\beta} = \mathbf{d}$ é testável se:

- $\Lambda \boldsymbol{\beta}$ é estimável e
- $r(\Lambda) = m$.

Esta definição é clara e intuitiva, pois, se $\Lambda \boldsymbol{\beta}$ não fosse estimável, nós não poderíamos definir um pivô para um teste desta hipótese e achar estimativas da estatística do teste. Se $r(\Lambda) < m$, ou o sistema, em $\boldsymbol{\beta}$, $\Lambda \boldsymbol{\beta} = \mathbf{d}$, é inconsistente, ou é redundante, assim, devemos requerer $r(\Lambda) = m$.

Pela Seção 1.4, a soma de quadrados do erro, sob o modelo de GM restrito, com

a restrição $\Lambda \underline{\beta} = \underline{d}$,

$$SSE_r = SSE + SQ,$$

por (1.12), onde $SQ = (\Lambda \underline{\hat{\beta}} - \underline{d})' (\Lambda(X'X)^- \Lambda')^{-1} (\Lambda \underline{\hat{\beta}} - \underline{d})$ e $\Lambda \underline{\hat{\beta}} = \Lambda(X'X)^- X' y$.

Temos

$$(\Lambda \underline{\hat{\beta}} - \underline{d}) \sim N(\Lambda \underline{\beta} - \underline{d}; \sigma^2 \Lambda(X'X)^- \Lambda').$$

Então, pelo teorema A4, apêndice A,

$$SQ / \sigma^2 = (\Lambda \underline{\hat{\beta}} - \underline{d})' (\Lambda(X'X)^- \Lambda')^{-1} (\Lambda \underline{\hat{\beta}} - \underline{d}) / \sigma^2 \sim \chi^2_{(r(\Lambda))}; (\Lambda \underline{\beta} - \underline{d})' (\Lambda(X'X)^- \Lambda')^{-1} (\Lambda \underline{\beta} - \underline{d}) / \sigma^2.$$

Agora, vamos reescrever SQ e SSE de maneira que eles tenham formas quadráticas com o mesmo vetor.

Proposição 1.9

$$a) SSE = (\underline{y} - X\Lambda'(\Lambda\Lambda')^- \underline{d})' (I - P_x)(\underline{y} - X\Lambda'(\Lambda\Lambda')^- \underline{d})$$

$$b) SQ = (\underline{y} - X\Lambda'(\Lambda\Lambda')^- \underline{d})' [(X(X'X)^- \Lambda'(\Lambda(X'X)^- \Lambda')^{-1} \Lambda(X'X)^- X)] (\underline{y} - X\Lambda'(\Lambda\Lambda')^- \underline{d}).$$

Prova:

(a) O resultado segue

$$X'(I - P_x) = X' - X'P_x = \underline{0}.$$

(b)

$$\begin{aligned} SQ &= (\Lambda \hat{\underline{\beta}} - \underline{d})' (\Lambda(X'X)^- \Lambda')^{-1} (\Lambda \hat{\underline{\beta}} - \underline{d}) \\ &= (\Lambda(X'X)^- X' \underline{y} - \underline{d})' (\Lambda(X'X)^- \Lambda')^{-1} (\Lambda(X'X)^- X' \underline{y} - \underline{d}), \end{aligned}$$

mas ,

$$\begin{aligned} [(\Lambda\Lambda')^{-1}\Lambda X']X(X'X)^- \Lambda' &= (\Lambda\Lambda')^{-1}\Lambda X'X(X'X)^- X' \Lambda', \text{ pois } \Lambda = AX \text{ para alguma matriz } A, \\ &= (\Lambda\Lambda')^{-1}\Lambda X'P_X \Lambda' \\ &= (\Lambda\Lambda')^{-1}\Lambda X' \Lambda' \\ &= (\Lambda\Lambda')^{-1}\Lambda\Lambda' = I_m, \end{aligned}$$

logo,

$$SQ = (\underline{y} - X\Lambda'(\Lambda\Lambda')^{-1}\underline{d})'[(X(X'X)^- \Lambda'(\Lambda(X'X)^- \Lambda')^{-1}\Lambda(X'X)^- X')](\underline{y} - X\Lambda'(\Lambda\Lambda')^{-1}\underline{d}).$$

■

Proposição 1.10

$$F = \frac{SQ / r(\Lambda)}{SSE / (n - r(X))} \sim F_{(r(\Lambda); n - r(X); (\Lambda\hat{\underline{\beta}} - \underline{d})'(\Lambda(X'X)^- \Lambda')^{-1}(\Lambda\hat{\underline{\beta}} - \underline{d})/\sigma^2)}$$

e sob a hipótese $\Lambda \underline{\beta} = \underline{d}$,

$$F = \frac{SQ / r(\Lambda)}{SSE / (n - r(X))} \sim F_{(r(\Lambda); n - r(X); 0)}.$$

Prova:

Como :

$$(I - P_X)[X(X'X)^{-1}\Lambda'(\Lambda(X'X)^{-1}\Lambda)^{-1}\Lambda(X'X)^{-1}X'] = 0,$$

$$SQ / \sigma^2 \sim \chi^2_{(r(\Lambda); (\Lambda\beta - d)(\Lambda(X'X)^{-1}\Lambda)^{-1}(\Lambda\beta - d)/\sigma^2)} ,$$

$$SSE / \sigma^2 \sim \chi^2_{(n - r(X); 0)} ,$$

e, SQ e SSE são independentes, o resultado segue porque temos a razão entre duas distribuições Qui - quadradas independentes, divididas pelos respectivos graus de liberdade, sendo a do denominador, Qui - quadrada central.

■

Logo, para um nível de significância α , a estatística de um teste para:

$$H_0 : \Lambda\beta = d$$

versus

$$H_1 : \Lambda\beta \neq d$$

é dada por :

$$F = \frac{SQ / r(\Lambda)}{SSE / (n - r(X))} \sim F_{(r(\Lambda); n - r(X); 0)} , \text{ sob } H_0 ,$$

onde rejeitamos a hipótese $\Lambda\beta = d$, se $F > F_{((1 - \alpha), (r(\Lambda), n - r(X))}$.

1.6.3 Comparando Modelos Restrito e Reduzido ou Testando uma Hipótese?

Voltando ao exemplo 1.1, temos a matriz de delineamento e o vetor de parâmetros do modelo completo :

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \sim & \sim & \sim & & \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \sim & \vdots \\ \sim & \sim & \sim & & \\ \vdots & \vdots & \sim & \sim & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{1} \\ \sim & \sim & \sim & & \end{bmatrix} \quad \text{e} \quad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} ;$$

e

$$X_0 = \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \sim & \sim \\ \mathbf{1} & \mathbf{1} \\ \sim & \sim \\ \vdots & \vdots \\ \mathbf{1} & \mathbf{1} \\ \sim & \sim \end{bmatrix} \quad \text{e} \quad \underset{\sim}{\delta} = \begin{bmatrix} \delta_0 \\ \delta_1 \end{bmatrix} ,$$

para um modelo reduzido. Observemos que obtemos o modelo reduzido a partir do completo fazendo

$$\beta_1 = \beta_2 = \dots = \beta_p = \delta_1 ,$$

ou seja, impondo

$$\beta_1 - \beta_2 = \beta_1 - \beta_3 = \dots = \beta_1 - \beta_p = 0,$$

$$\Lambda \tilde{\beta} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 & -1 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Esta última expressão é estimável. Logo, escolher o modelo de GM reduzido é o mesmo que escolher o modelo GM restrito com a restrição acima. Pela proposição 1.9 uma estatística F central, sob a hipótese $\Lambda \tilde{\beta} = \tilde{0}$, é dada por

$$F = \frac{\tilde{y}'X(X'X)^{-1}\Lambda'(\Lambda(X'X)^{-1}\Lambda')^{-1}\Lambda(X'X)^{-1}X'y / r(\Lambda)}{SSE / (n - r(X))}. \quad (1.17)$$

Agora, a proposição 1.7, evidencia a seguinte ANOVA :

$$\tilde{y}'\tilde{y} = \tilde{y}'P_{x0}\tilde{y} + \tilde{y}'(P_x - P_{x0})\tilde{y} + \tilde{y}'(I - P_x)\tilde{y} \quad (1.18)$$

$$\Rightarrow \tilde{y}'(P_x - P_{x0})\tilde{y} = \tilde{y}'(I - P_{x0})\tilde{y} - \tilde{y}'(I - P_x)\tilde{y},$$

onde

$\tilde{y}'(I - P_{x0})\tilde{y}$ é a soma de quadrados do erro, sob o modelo de GM reduzido, e

$\underline{y}'(I - P_x)\underline{y}$ é soma de quadrados do erro, sob o modelo de GM, irrestrito.

No entanto, se especificamos uma hipótese (testável) causadora da redução do modelo completo, temos

$$\underline{y}'(I - P_{x0})\underline{y} = SSE_r,$$

a soma de quadrados do erro, sob o modelo de GM restrito.

Assim,

$$\begin{aligned}\underline{y}'(P_x - P_{x0})\underline{y} &= \underline{y}'(I - P_{x0})\underline{y} - \underline{y}'(I - P_x)\underline{y} \\ &= SSE_r - SSE \\ &= SSE + SQ - SSE, \text{ pela equação (1.12),} \\ &= SQ \\ &= (\Lambda \hat{\delta})'(\Lambda(X'X)^{-1}\Lambda')^{-1}(\Lambda \hat{\delta}).\end{aligned}$$

A matriz $P_x - P_{x0}$ é uma matriz idempotente, logo, usando a igualdade de posto e traço para matrizes idempotentes, temos:

$$\begin{aligned}r(P_x - P_{x0}) &= \text{Tr}(P_x - P_{x0}) \\ &= \text{Tr}(P_x) - \text{Tr}(P_{x0}) \\ &= r(X) - r(X_0) \\ &= m = r(\Lambda).\end{aligned}$$

Então podemos dizer que as estatísticas (1.16) e (1.17) são equivalentes, quando temos ambos: modelo reduzido e hipótese testável associada.

Resumindo, sob o modelo de GM, seja $H_0: \Lambda \beta = 0$ uma hipótese testável, onde Λ é uma matriz de m linhas e de posto m . E seja X_0 a matriz do modelo reduzido correspondente. Denominando :

$SSH_0 = y'X(X'X)^{-1}\Lambda'(\Lambda(X'X)^{-1}\Lambda(X'X)^{-1}X'y = y'(P_x - P_{x_0})y$, a soma de quadrados sob a hipótese H_0 ;

$SSM_r = y'P_{x_0}y$, a soma de quadrados devido ao modelo de GM restrito (reduzido);

$SSE = y'(I - P_x)y$ a soma de quadrados do erro, sob o modelo de GM irrestrito e

$SSTot = y'y$, a soma de quadrados total,

por (1.18), temos a seguinte ANOVA

$$SSTot = SSM_r + SSH_0 + SSE,$$

apresentada na seguinte tabela .

Tabela ANOVA 1.7.1

FV	GL	SS	MS
Modelo Restrito	$r(X) - r(\Lambda)$	SSM_r	$SSM_r / (r(X) - r(\Lambda))$
Hipótese $\Lambda\beta = 0$	$r(\Lambda)$	SSH_0	$MSH_0 = SSH_0 / r(\Lambda)$
Erro	$n - r(X)$	SSE	$MSE = SSE / (n - r(X))$
Total	n	$SSTot$	-

Logo, para obtermos a estatística F em (1.16), basta usarmos a razão entre MSH_0 e MSE da tabela de ANOVA.

Frequentemente é utilizada a tabela de ANOVA com a soma total do vetor y ajustada. Neste caso, juntando os resultados das proposições 1.7 e 1.8, temos a ANOVA :

$$y'y = \frac{1}{n} y'(\mathbf{1} \mathbf{1}')y + y'(P_{x0} - \frac{1}{n} \mathbf{1} \mathbf{1}')y + y'(P_x - P_{x0})y + y'(I - P_x)y$$

$$\Rightarrow y'y - \frac{1}{n} y'(\mathbf{1} \mathbf{1}')y = y'(P_{x0} - \frac{1}{n} \mathbf{1} \mathbf{1}')y + y'(P_x - P_{x0})y + y'(I - P_x)y,$$

apresentada na seguinte tabela .

Tabela ANOVA 1.6.2

FV	GL	SS
Modelo Restrito	$r(X) - r(\Lambda) - 1$	$SSM_r - n\bar{y}_{..}^2$
Hipótese $\Lambda\beta = 0$	$r(\Lambda)$	SSH_0
Erro	$n - r(X)$	SSE
Total Ajustado	$n - 1$	$SSTot - n\bar{y}_{..}^2$

Notando que apenas a primeira e a quarta linha se modificaram.

1.7 Reparametrização

Algumas vezes são usadas estruturas diferentes para modelos GM para o mesmo conjunto de observações. Dependendo da modificação entre duas modelagens é possível "transferir " os resultados obtidos sob um dos modelos para o outro. Vamos expor a seguir, quando é que torna-se possível o intercâmbio de resultados entre dois

modelos.

Definição 1.7

Considere dois modelos de Gauss - Markov, para o mesmo vetor de observações

$\underline{y}$,

$$\underline{y} = \underline{X} \underline{\beta} + \underline{e}, \text{ onde } \underline{\beta} \text{ é um vetor } p \times 1 \text{ (modelo I)}$$

e

$$\underline{y} = \underline{W} \underline{\alpha} + \underline{e}, \text{ onde } \underline{\alpha} \text{ é um vetor } q \times 1 \text{ (modelo II)}.$$

Por definição, o modelo II é uma reparametrização do modelo I, se para

$\forall \underline{b} \in \mathbb{R}^p$, existe $\underline{a} \in \mathbb{R}^q$, tal que $\underline{W} \underline{a} = \underline{b}$, e para $\forall \underline{a} \in \mathbb{R}^q$, existe $\underline{b} \in \mathbb{R}^p$, tal que $\underline{b} = \underline{W} \underline{a}$.

Notemos que o conceito de reparametrização é simétrico.

Uma condição necessária e suficiente para o modelo II ser uma reparametrização do modelo I, e vice versa, é $\mathcal{L}(\underline{X}) = \mathcal{L}(\underline{W})$. Consequentemente, existem matrizes $\underline{F}$, $p \times q$, e $\underline{G}$, $q \times p$, tais que $\underline{W} = \underline{X}\underline{F}$ e $\underline{X} = \underline{W}\underline{G}$.

Proposição 1.11

Se os modelos I e II são reparametrizações, e, se a função $\underline{\lambda}' \underline{\beta}$ é estimável sob o modelo I, então :

a) $\underline{\lambda}' \underline{F} \underline{\alpha}$ é estimável sob o modelo II;

b) o estimador de MQO de $\lambda' \beta$, sob o modelo I, é o estimador de MQO de $\lambda' F \alpha$, sob o modelo II.

Prova :

a) $\lambda' \beta$ é estimável, sob o modelo I, logo

$$\Rightarrow \lambda' \in \mathcal{L}(X)$$

$$\Rightarrow \lambda' = a'X, \text{ para algum } a$$

$$\Rightarrow \lambda' F = a'XF = a'W$$

$$\Rightarrow \lambda' F \in \mathcal{L}(W)$$

$$\Rightarrow \lambda' F \alpha \text{ é estimável sob modelo II.}$$

b) Sob o modelo II, o estimador MQO de $\lambda' F \alpha$ é

$$\lambda' F (W'W)^{-1} W' y.$$

Sob o modelo I, o estimador MQO de $\lambda' \beta$ estimável é

$$\begin{aligned} \lambda' (X'X)^{-1} X' y &= a' X (X'X)^{-1} X' y = a' P_x y = a' P_w y, \text{ pelo apêndice A, proposição A7,} \\ &= a' W (W'W)^{-1} W' y \\ &= a' XF (W'W)^{-1} W' y \\ &= \lambda' F (W'W)^{-1} W' y. \end{aligned}$$

■

Exemplo 1.1 (Continuação)

Voltando ao exemplo de classificação, temos :

$$X = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ \sim & \sim & \sim & & \\ 1 & 0 & 1 & \sim & 1 \\ \sim & \sim & \sim & & \\ \vdots & \vdots & \vdots & \sim & \vdots \\ 1 & 0 & \dots & 0 & 1 \\ \sim & \sim & \sim & & \end{bmatrix} \quad \text{e} \quad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \text{ modelo (I)}$$

Usualmente, textos básicos trabalham com a reparametrização

$$W = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \sim & \sim & & \sim \\ 0 & 1 & \sim & 1 \\ \sim & \sim & & \sim \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 \\ \sim & & \sim & \sim \end{bmatrix} \quad \underset{\sim}{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}. \text{ (II)}$$

Assim,

$$\underset{\sim}{\hat{\alpha}} = (W'W)^{-1}W'y = \begin{bmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix}$$

é solução única das EN do modelo II.

Usando

$$F = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

$\lambda' \beta$ é estimado por

$$\underset{\sim}{\lambda}' F \underset{\sim}{\hat{\alpha}} = \begin{bmatrix} 0 \\ \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix},$$

para $\underset{\sim}{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_p)$, onde $\lambda_0 = \lambda_1 + \dots + \lambda_p$.

CAPÍTULO 2

Análise de Covariância

O modelo de análise de covariância (ANCOVA) incorpora características dos modelos de classificação e de regressão. Na maioria das vezes, os objetivos de análises de covariância coincidem com os objetivos de modelos de classificação. No entanto, o uso de covariáveis (aspectos do modelo de regressão) permite, por exemplo, a comparação de grupos, " tirando o efeito " de alguma variável secundária, relevante. É como que fizéssemos uma comparação, nivelando previamente com respeito à variável secundária (covariável).

Além da vantagem acima citada, o uso do modelo de análise de covariância pode acarretar em um melhor ajuste do modelo linear. Tradicionalmente, a análise de covariância tem sido usada como uma ferramenta em análise de experimento. Quando as unidades experimentais são homogêneas entre si, realiza-se um plano completamente aleatorizado. No entanto, o erro experimental pode ser grande caso as unidades experimentais sejam heterogêneas. Neste caso, o uso de blocos poderia ser uma solução, isto é, agrupar as unidades experimentais em blocos que contenham unidades as mais homogêneas possíveis e consequentemente, que os blocos sejam os mais heterogêneos possíveis. Mas, nem sempre uso do bloco é viável. E uma maneira alternativa para reduzir este erro seria usar a análise de covariância. Isto é, introduzir uma covariável no modelo, como será visto neste Capítulo. Neste caso, a covariável teria o mesmo objetivo dos blocos.

Este Capítulo é iniciado com a apresentação do modelo de Análise de covariância. Antes de mencionarmos a estimação das funções estimáveis dos

parâmetros do modelo, uma reparametrização do mesmo é apresentada, permitindo de forma mais fácil a obtenção dos estimadores destes parâmetros. O método de estimação abordado, como no Capítulo 1, é o de mínimos quadrados. Em seguida abordamos a análise de covariância com a matriz de variância e covariâncias do erro igual a $\sigma^2 V$, seguido de Teste de Hipóteses. Finalmente, apresentamos cinco exemplos para um melhor entendimento.

Para obtermos a análise dos exemplos apresentados, foram feitos programas em SAS utilizando procedimentos ANOVA e programação em linguagem IML, os quais são encontrados no Apêndice B.

2.1 Modelo de Análise de Covariância

No capítulo anterior nos referimos ao *modelo de classificação*, quando todos as variáveis independentes eram indicadoras e ao *modelo de regressão*, quando as variáveis independentes eram variáveis gerais. Neste Capítulo iremos tratar de um modelo de GM onde são usadas simultaneamente variáveis classificatórias e regressoras. Este modelo é denominado de *modelo de análise de covariância*, onde às variáveis regressoras daremos o nome de variáveis *concomitantes*, ou ainda, *covariáveis*. Sendo um caso particular do modelo de Gauss Markov, quando usamos o modelo de análise de covariância estaremos interessados em estudar uma variável aleatória dependente Y com base em p variáveis independentes $X_1, \dots, X_p$ e com base em s covariáveis $Z_1, \dots, Z_s$.

O modelo de análise de covariância amostral em forma matricial é dado por

$$\underset{\sim}{y} = [\underset{\sim}{X} \quad \underset{\sim}{Z}] \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\gamma} \end{bmatrix} + \underset{\sim}{e}, \quad (2.1)$$

onde

$\underset{\sim}{y}' = [y_1, \dots, y_n]$ é um vetor de n observações ;

$\underset{\sim}{X} = [\underset{\sim}{x}_1, \dots, \underset{\sim}{x}_p]$ é uma matriz de delineamento, $n \times p$;

$\underset{\sim}{Z} = [\underset{\sim}{z}_1, \dots, \underset{\sim}{z}_s]$ é uma matriz de regressão $n \times s$;

$\underset{\sim}{\beta}' = [\beta_1, \dots, \beta_p]$ é um vetor de parâmetros desconhecidos, de dimensão p ;

$\underset{\sim}{\gamma}' = [\gamma_1, \dots, \gamma_s]$ é um vetor de parâmetros desconhecidos, de dimensão s ;

$\underset{\sim}{e}' = [e_1, \dots, e_n]$ é um vetor aleatório, tal que $E(\underset{\sim}{e}) = \underset{\sim}{0}$ e $\text{Var}(\underset{\sim}{e}) = \sigma^2 I$.

Portanto, o modelo de análise de covariância é apenas um modelo de GM onde particionamos a matriz do modelo e o vetor de parâmetros.

Vamos, agora, reescrever a equação do modelo (2.1), somando e subtraindo a parcela $P_x Z$:

$$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{\gamma} + \underset{\sim}{e} \\ &= \underset{\sim}{X}\underset{\sim}{\beta} + P_x \underset{\sim}{Z}\underset{\sim}{\gamma} + (I - P_x)\underset{\sim}{Z}\underset{\sim}{\gamma} + \underset{\sim}{e}, \end{aligned}$$

onde P_x é a matriz de projeção sobre o espaço coluna de X , $\mathcal{L}(X)$.

Proposição 2.1

Os modelos de análise de covariância com equações dadas por

$$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{P_x}Z\underset{\sim}{\gamma} + (I - \underset{\sim}{P_x})Z\underset{\sim}{\gamma} + \underset{\sim}{e} \\ &= \begin{bmatrix} \underset{\sim}{X} & \underset{\sim}{P_x}Z & (I - \underset{\sim}{P_x})Z \end{bmatrix} \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\gamma} \\ \underset{\sim}{\gamma} \end{bmatrix} + \underset{\sim}{e} \end{aligned} \quad (2.2)$$

e

$$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{X}\underset{\sim}{\delta} + (I - \underset{\sim}{P_x})Z\underset{\sim}{\theta} + \underset{\sim}{e} \\ &= \begin{bmatrix} \underset{\sim}{X} & (I - \underset{\sim}{P_x})Z \end{bmatrix} \begin{bmatrix} \underset{\sim}{\delta} \\ \underset{\sim}{\theta} \end{bmatrix} + \underset{\sim}{e}, \end{aligned} \quad (2.3)$$

onde $\underset{\sim}{P_x}$ é a matriz de projecção sobre $\mathcal{L}(\underset{\sim}{X})$, são reparametrizações.

Prova :

Pela proposição A2, apêndice A,

$$\mathcal{L}(\underset{\sim}{X}) = \mathcal{L}(\underset{\sim}{P_x}).$$

As colunas da matriz $\underset{\sim}{P_x}Z$ são combinações lineares das colunas da matriz $\underset{\sim}{P_x}$, logo,

$$\begin{aligned}
\mathcal{L}(P_x Z) &\subset \mathcal{L}(P_x) = \mathcal{L}(X) \\
\Rightarrow \mathcal{L}(X : P_x Z) &= \mathcal{L}(X) \\
\Rightarrow \mathcal{L}(X : P_x Z : (I - P_x)Z) &= \mathcal{L}(X : (I - P_x)Z).
\end{aligned}$$

Portanto, pela Seção 1.7, Capítulo 1, os dois modelos são reparametrizações.

Note que X e $(I - P_x)Z$ são ortogonais, isto é, $X'(I - P_x)Z = \mathbf{0}$. De maneira análoga ao modelo (1.2), Capítulo 1, as equações normais referentes ao modelo (2.3) são :

$$\begin{bmatrix} X' & (I - P_x)Z' \end{bmatrix} \begin{bmatrix} X & (I - P_x)Z \end{bmatrix} \begin{bmatrix} \delta \\ \theta \end{bmatrix} = \begin{bmatrix} X' & (I - P_x)Z' \end{bmatrix} y$$

ou

$$\begin{bmatrix} X'X & 0 \\ 0 & Z'(I - P_x)Z \end{bmatrix} \begin{bmatrix} \delta \\ \theta \end{bmatrix} = \begin{bmatrix} X'y \\ Z'(I - P_x)y \end{bmatrix},$$

ou dois sistemas com vetores de incógnitas distintos,

$$\begin{cases} X'X\delta = X'y \\ Z'(I - P_x)Z\theta = Z'(I - P_x)y \end{cases}, \quad (2.4)$$

de modo que uma solução da primeira parte do sistema é

$$\hat{\underline{\delta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y},$$

e uma solução da segunda parte do sistema é

$$\hat{\underline{\theta}} = [\underline{Z}'(\underline{I} - \underline{P}_x)\underline{Z}]^{-1} \underline{Z}'(\underline{I} - \underline{P}_x)\underline{y}.$$

A matriz $\underline{Z}$ consiste de colunas de variáveis regressoras, portanto, em muitos casos $\underline{Z}'(\underline{I} - \underline{P}_x)\underline{Z}$ será tipicamente não singular. Logo, para $\underline{Z}$ de posto coluna completo, a solução da segunda parte do sistema de EN (2.4) será

$$\hat{\underline{\theta}} = [\underline{Z}'(\underline{I} - \underline{P}_x)\underline{Z}]^{-1} \underline{Z}'(\underline{I} - \underline{P}_x)\underline{y},$$

sendo esta solução única.

A reparametrização (2.3) é bastante útil. Uma maneira de determinarmos a estimabilidade de uma função

$$\underline{\lambda}'_1 \underline{\delta} + \underline{\lambda}'_2 \underline{\theta},$$

pelo resultado 1.1, Capítulo 1, é observarmos em $[\underline{\lambda}'_1 : \underline{\lambda}'_2]$ as mesmas relações lineares presentes entre as colunas de $[\underline{X} : (\underline{I} - \underline{P}_x)\underline{Z}]$. E, se $(\underline{I} - \underline{P}_x)\underline{Z}$ é de posto coluna completo, as relações lineares residem nas colunas de $\underline{X}$. Desta forma, neste caso, é suficiente verificar se as relações entre as colunas de $\underline{X}$ prevalecem entre os elementos de $\underline{\lambda}'_1$.

Sendo X uma matriz de delineamento, não é necessariamente de posto coluna completo, e temos que δ não é necessariamente estimável. No entanto, $X\delta$ é sempre estimável e seu estimador é

$$\hat{X\delta} = X(X'X)^{-1}X'y = P_x y.$$

Uma vez que o modelo (2.3) é uma reparametrização do modelo (2.2), o estimador da função estimável

$$A \begin{bmatrix} \beta \\ \gamma \\ \gamma \\ - \end{bmatrix}, \text{ sob o modelo (2.2),}$$

é igual ao estimador da função estimável

$$AF \begin{bmatrix} \delta \\ \theta \\ - \end{bmatrix}, \text{ sob o modelo (2.3),}$$

como visto na seção 1.5, Capítulo 1, onde F é uma matriz tal que

$$[X : (I - P_x)Z] = [X : P_x Z : (I - P_x)Z] F.$$

Proposição 2.2

Uma matriz F tal que $[X : (I - P_x)Z] = [X : P_x Z : (I - P_x)Z] F$ é

$$\begin{bmatrix} I_p - (X'X)^{-1}X'Z \\ 0 & I_s \\ - & \\ 0 & I_s \\ - & \end{bmatrix}$$

Prova:

$$\begin{aligned} [X : P_x Z : (I - P_x)Z] F &= [X : P_x Z : (I - P_x)Z] \begin{bmatrix} I_p - (X'X)^{-1}X'Z \\ 0 & I_s \\ - & \\ 0 & I_s \\ - & \end{bmatrix} \\ &= [X : (I - P_x)Z] \end{aligned}$$

■

Sob o modelo (2.2), de matriz de modelo

$$[X : P_x Z : (I - P_x)Z],$$

se Z é matriz de posto coluna completo, a estimabilidade de

$$\lambda_1' \beta + \lambda_2' \gamma + \lambda_3' \gamma$$

se verificará se os componentes de λ_1' atenderem às mesmas relações lineares existentes entre as colunas de X , conforme Capítulo 1, Seção 1.1.

Proposição 2.3

Os estimadores de MQO de $X\beta$ e γ , para o caso da matriz Z de posto coluna completo, são

$$\begin{bmatrix} X\hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X\hat{\delta} - P_x Z\hat{\theta} \\ \hat{\theta} \end{bmatrix} = \begin{bmatrix} P_x y - P_x Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x)y \\ [Z'(I - P_x)Z]^{-1}Z'(I - P_x)y \end{bmatrix}.$$

Prova :

Os componentes do primeiro bloco de colunas em

$$\begin{bmatrix} X & 0 & 0 \\ 0 & I_r & 0 \end{bmatrix}$$

atendem às mesmas relações lineares existentes entre as colunas de X , logo,

$$\begin{bmatrix} X & 0 & 0 \\ 0 & I_r & 0 \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \\ \gamma \end{bmatrix}$$

é estimável sob o modelo (2.2) e seu estimador é

$$\begin{bmatrix} X & 0 & 0 \\ 0 & I_r & 0 \end{bmatrix} \begin{bmatrix} I_p - (X'X)^{-1}X'Z \\ 0 & I_r \\ 0 & I_r \end{bmatrix} \begin{bmatrix} \hat{\delta} \\ \hat{\theta} \\ \hat{\theta} \end{bmatrix} = \begin{bmatrix} X & -P_x Z \\ 0 & I_r \end{bmatrix} \begin{bmatrix} \hat{\delta} \\ \hat{\theta} \end{bmatrix},$$

pela proposição 1.11, Capítulo 1, e pela proposição 2.2. Logo o resultado segue. ■

Proposição 2.4

No contexto do modelo de análise de covariância, um estimador não viciado para σ^2 é

$$MSE_{\sigma} = \frac{\tilde{y}'(I - P)\tilde{y}}{n - r([X : Z])},$$

onde P é a matriz de projeção ortogonal sobre $\mathcal{L}([X : Z])$.

Prova:

O resultado segue o teorema 1.1, Capítulo 1, porque o modelo de análise de covariância é um modelo de GM, onde particionamos a matriz do modelo e o vetor de parâmetros, sendo a matriz do modelo (2.1) igual a $[X : Z]$. ■

Proposição 2.5

A soma de quadrados do erro (SSE_e), sob o modelo de análise de covariância pode ser decomposta em duas parcelas

$$\tilde{y}'(I - P)\tilde{y} = \tilde{y}'(I - P_x)\tilde{y} - \tilde{y}'(I - P_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x)\tilde{y}.$$

Prova :

Pelas proposições A8 e A9, apêndice A,

$$P = P_x + (I - P_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x),$$

é a matriz de projeção ortogonal do espaço coluna da matriz do modelo de análise de covariância, $\mathcal{C}([X : Z])$. Logo,

$$\begin{aligned} \tilde{y}'(I - P)\tilde{y} &= \tilde{y}'(I - P_x - (I - P_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x))\tilde{y} \\ &= \tilde{y}'(I - P_x)\tilde{y} - \tilde{y}'(I - P_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x)\tilde{y}. \end{aligned}$$

■

Pela proposição acima observamos que a forma quadrática da segunda parcela não é negativa, portanto, a soma de quadrados do erro, sob o modelo de análise de covariância, é menor que a soma de quadrados do erro sob o modelo GM correspondente sem covariáveis. Além disso, a proposição 2.5 indica uma maneira de usar matrizes de menores dimensões, ao usar P_x ao invés de P .

2.1.1 Modelo de Análise de Covariância Generalizado

Consideremos todas as suposições do modelo de análise de covariância, modificando apenas o seguinte : $\text{Var}(\tilde{e}) = \sigma^2 V$, para V matriz positiva definida conhecida. Tal modelo é conhecido como modelo de análise de covariância generalizado, sendo um caso particular do modelo de Aitken, visto na seção 1.2,

Capítulo 1. Pré-multiplicando ambos lados da equação (2.2) do modelo por Q , onde $V^{-1} = Q'Q$, nós teremos

$$Q\underset{\sim}{y} = Q \begin{bmatrix} X & Z \end{bmatrix} \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\gamma} \end{bmatrix} + Q\underset{\sim}{e}$$

ou

$$\underset{\sim}{y}^* = \begin{bmatrix} X & Z \end{bmatrix}^* \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\gamma} \end{bmatrix} + \underset{\sim}{e}^* \quad (2.5)$$

onde $\underset{\sim}{y}^* = Q\underset{\sim}{y}$, $\begin{bmatrix} X & Z \end{bmatrix}^* = Q\begin{bmatrix} X & Z \end{bmatrix}$ e $\underset{\sim}{e}^* = Q\underset{\sim}{e}$, observando que $E(\underset{\sim}{e}^*) = \mathbf{0}$ e $\text{Var}(\underset{\sim}{e}^*) = \sigma^2 Q'VQ = \sigma^2 I$. Isto é expressamos o modelo de análise de covariância generalizado como um modelo de análise de covariância.

Usando os resultados já apresentados sobre o modelo de análise de covariância e sobre o modelo de Aitken, observamos que no modelo de análise de covariância generalizado, $\lambda'\underset{\sim}{\beta}$ é estimável se

$$\underset{\sim}{\lambda}' \in \mathcal{Q}(\underset{\sim}{X}).$$

Proposição 2.6

Os estimadores de mínimos quadrados para $\underset{\sim}{X}'\underset{\sim}{\beta}$ e $\underset{\sim}{\gamma}$, para Z de posto coluna completo, sob o modelo de análise de covariância generalizado, denominados estimadores de *mínimos quadrados generalizados*, com base em V , são

$$\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{A}\underset{\sim}{y} - \underset{\sim}{A}\underset{\sim}{Z}[Z'(I - A)'V^{-1}(I - A)Z]^{-1}Z'(I - A)'V^{-1}(I - A)\underset{\sim}{y}$$

e

$$\underset{\sim}{\hat{y}} = [Z'(I - A)'V^{-1}(I - A)Z]^{-1} Z'(I - A)'V^{-1}(I - A)\underset{\sim}{y}$$

onde

$A = X(X'V^{-1}X)^{-1}X'V^{-1}$ é uma matriz de projeção não ortogonal sobre o espaço coluna de X

Prova :

Pela proposição 2.3, o estimador de mínimos quadrados de $X'\beta$ da equação (2.5) é

$$\underset{\sim}{X}'\underset{\sim}{\hat{\beta}} = P_{\underset{\sim}{x}}^*\underset{\sim}{y}^* - P_{\underset{\sim}{x}}^*Z^*[Z^*(I - P_{\underset{\sim}{x}}^*Z^*)]^{-1}Z^*(I - P_{\underset{\sim}{x}}^*)\underset{\sim}{y}^*,$$

expressando em termos de X , $\underset{\sim}{y}$ e V temos

$$\begin{aligned}\underset{\sim}{QX}\underset{\sim}{\hat{\beta}} &= \underset{\sim}{QX}(X'Q'QX)^{-1}X'Q'Q\underset{\sim}{y} \\ &- \underset{\sim}{QX}(X'Q'QX)^{-1}X'Q'Q\underset{\sim}{Z}[Z'Q'(I - \underset{\sim}{QX}(X'Q'QX)^{-1}X'Q)QZ]^{-1} \\ &Z'Q'(I - \underset{\sim}{QX}(X'Q'QX)^{-1}X'Q)Q\underset{\sim}{y},\end{aligned}$$

Pré- multiplicando por Q^{-1} :

$$\begin{aligned}\underset{\sim}{X}\underset{\sim}{\hat{\beta}} &= X(X'V^{-1}X)^{-1}X'V^{-1}\underset{\sim}{y} \\ &- X(X'V^{-1}X)^{-1}X'V^{-1}\underset{\sim}{Z}[Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})Z]^{-1} \\ &Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})\underset{\sim}{y}\end{aligned}$$

$$\begin{aligned}
&= \underset{\sim}{A} \underset{\sim}{y} - AZ[Z'V^{-1}(I - A)Z]^{-1} Z'V^{-1}(I - A) \underset{\sim}{y} \\
&= \underset{\sim}{A} \underset{\sim}{y} - AZ[Z'(I - A)'V^{-1}(I - A)Z]^{-1} Z'(I - A)'V^{-1}(I - A) \underset{\sim}{y}, \text{ por (1.10).}
\end{aligned}$$

Também pela proposição 2.3, o estimador de mínimos quadrados de $\underset{\sim}{\gamma}$ da equação (2.5) é

$$\underset{\sim}{\hat{\gamma}} = [Z^{*'}(I - P_{\underset{\sim}{X}}^{*})Z^{*}]^{-1}Z^{*'}(I - P_{\underset{\sim}{X}}^{*}) \underset{\sim}{y}^{*},$$

expressando em termos de $\underset{\sim}{X}$, $\underset{\sim}{y}$ e $\underset{\sim}{V}$, temos

$$\begin{aligned}
\underset{\sim}{\hat{\gamma}} &= [Z'Q'(I - QX(X'Q'QX)^{-1}X'Q')QZ]^{-1}Z'Q'(I - QX(X'Q'QX)^{-1}X'Q')Q \underset{\sim}{y} \\
&= [Z'(I - A)'V^{-1}(I - A)Z]^{-1}Z'(I - A)'V^{-1}(I - A) \underset{\sim}{y}
\end{aligned}$$

■

Proposição 2.7

No contexto do modelo de análise de covariância generalizado, para Z de posto coluna completo,

$$MSE_{\underset{\sim}{g}} = \frac{\underset{\sim}{y}'(I - A_{\underset{\sim}{g}})'V^{-1}(I - A_{\underset{\sim}{g}})\underset{\sim}{y}}{(n - r([X : Z]))}$$

é um estimador não viciado de σ^2 , onde

$$A_g = A + (I - A)Z[Z'(I - A)'V^{-1}(I - A)Z]^{-1}Z'(I - A)'V^{-1}$$

é a matriz de projeção não ortogonal sobre $\mathcal{L}([X : Z])$.

Prova :

Pela proposição 2.4, temos o seguinte estimador não viciado de σ^2 , após transformação do modelo de covariância generalizado :

$$\tilde{y}^* (I - P^*) \tilde{y}^* / (n - r([X^* : Z^*]))$$

onde $y^* = Qy$, $X^* = QX$, $Z^* = QZ$ e $Q'Q = V^{-1}$.

Expressando em termos de X , y , V^{-1} ,

$$A = X(X'V^{-1}X)^{-1}X'V^{-1},$$

$$P_x^* = QX(X'Q'QX)^{-1}X'Q', \text{ pelo teorema A1, Apêndice A e}$$

$$P^* = [QX(X'Q'QX)^{-1}X'Q' + ((I - QX(X'Q'QX)^{-1}X'Q')QZ)$$

$$[Z'Q'(I - QX(X'Q'QX)^{-1}X'Q')QZ]^{-1}$$

$$((I - QX(X'Q'QX)^{-1}X'Q')QZ)']], \text{ pelas proposições A8 e A9, Apêndice A.}$$

temos que

$$\begin{aligned} SSE_g &= y^* (I - P^*) y^* \\ &= \tilde{y}'Q' \{ I - [\tilde{Q}X(X'Q'QX)^{-1}X'Q' + ((I - QX(X'Q'QX)^{-1}X'Q')QZ) \\ &\quad [Z'Q'(I - QX(X'Q'QX)^{-1}X'Q')QZ]^{-1} \\ &\quad ((I - QX(X'Q'QX)^{-1}X'Q')QZ)']] \} Qy \end{aligned}$$

$$\begin{aligned}
&= \underset{\sim}{y}' \{ \underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1} \underset{\sim}{X} (\underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{Q}' - ((\underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1} \underset{\sim}{X} (\underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{V}^{-1}) \underset{\sim}{Z}) \\
&\quad [\underset{\sim}{Z}' (\underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1} \underset{\sim}{X} (\underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{V}^{-1}) \underset{\sim}{Z}]^{-1} \\
&\quad ((\underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1} \underset{\sim}{X} (\underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{V}^{-1}) \underset{\sim}{Z})' \} \underset{\sim}{y} \\
&= \underset{\sim}{y}' \underset{\sim}{V}^{-1} [\underset{\sim}{I} - \underset{\sim}{A} - (\underset{\sim}{I} - \underset{\sim}{A}) \underset{\sim}{Z} [\underset{\sim}{Z}' \underset{\sim}{V}^{-1} (\underset{\sim}{I} - \underset{\sim}{A}) \underset{\sim}{Z}]^{-1} \underset{\sim}{Z}' (\underset{\sim}{I} - \underset{\sim}{A})' \underset{\sim}{V}^{-1}] \underset{\sim}{y} \\
&= \underset{\sim}{y}' \underset{\sim}{V}^{-1} (\underset{\sim}{I} - \underset{\sim}{A}_g) \underset{\sim}{y} \\
&= \underset{\sim}{y}' (\underset{\sim}{I} - \underset{\sim}{A}_g)' \underset{\sim}{V}^{-1} (\underset{\sim}{I} - \underset{\sim}{A}_g) \underset{\sim}{y}, \text{ por (1.10).} \tag{2.6}
\end{aligned}$$

Agora,

$$\begin{aligned}
\mathcal{L}([QX : QZ]) &= \mathcal{L}([X : Z]), \text{ pois } Q \text{ é uma matriz não singular,} \\
&= \mathcal{L}(X : (I - A)Z), \text{ pela proposição A10, apêndice A.}
\end{aligned}$$

E como $AX = X$, X e $(I - A)Z$ são ortogonais. Logo, pela proposição A9, apêndice A, A_g é a matriz de projeção não ortogonal sobre $\mathcal{L}([X : Z])$.

Temos também, $r([X : Z]^*) = r([X : Z])$. Então, segue o resultado.

■

2.1.2 Teste de Hipóteses no Modelo de Análise de covariância

Denomina-se análise de covariância uma análise de variância sob o modelo de análise de covariância.

Proposição 2.8

A partição da soma de quadrados total, sob o modelo de análise de covariância, para Z de posto coluna completo,

$$\underset{\sim}{y}'\underset{\sim}{y} = \underset{\sim}{y}'\underset{\sim}{P}_x\underset{\sim}{y} + \underset{\sim}{y}'(I - \underset{\sim}{P}_x)Z[Z'(I - \underset{\sim}{P}_x)Z]^{-1}Z'(I - \underset{\sim}{P}_x)\underset{\sim}{y} + \underset{\sim}{y}'(I - P)\underset{\sim}{y},$$

é uma ANCOVA.

Prova :

Temos que

$$\underset{\sim}{P}_x, (I - \underset{\sim}{P}_x)Z[Z'(I - \underset{\sim}{P}_x)Z]^{-1}Z'(I - \underset{\sim}{P}_x) \text{ e } (I - P)$$

são simétricas ,

$$\underset{\sim}{P}_x + (I - \underset{\sim}{P}_x)Z[Z'(I - \underset{\sim}{P}_x)Z]^{-1}Z'(I - \underset{\sim}{P}_x) + (I - P) = I$$

e

$$\underset{\sim}{P}_x (I - \underset{\sim}{P}_x)Z[Z'(I - \underset{\sim}{P}_x)Z]^{-1}Z'(I - \underset{\sim}{P}_x) = 0$$

$$\underset{\sim}{P}_x (I - P) = 0$$

$$\text{e } (I - \underset{\sim}{P}_x)Z[Z'(I - \underset{\sim}{P}_x)Z]^{-1}Z'(I - \underset{\sim}{P}_x) (I - P) = 0 .$$

Então, pela condição (a) do resultado 1.5, Capítulo 1, segue o resultado .



Uma Tabela de ANCOVA que destaca a soma de quadrados devida à covariável é dada por :

Tabela ANCOVA 2.1

FV	GL	SS
Modelo(sem Z)	$r(X)$	$y' P_x y$
Covariáveis	$r(Z)$	$\tilde{y}'(I - \tilde{P}_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x)\tilde{y}$
Erro	$n - r(X) - r(Z)$	$\tilde{y}'(I - P)\tilde{y}$
Total	n	$\tilde{y}'\tilde{y}$

onde

$\tilde{y}'(I - P_x)Z[Z'(I - P_x)Z]^{-1}Z'(I - P_x)\tilde{y}$ é a soma de quadrados devida à covariável Z;

$\tilde{y}' P_x \tilde{y}$ é a soma de quadrados do modelo sem a covariável Z ;

$\tilde{y}'(I - P)\tilde{y}$ é a soma de quadrados do erro.

Uma estatística do teste para verificar uma hipótese testável $H_0 : \Lambda \beta = 0$ é análoga às estatísticas (1.16) e (1.17) do Capítulo 1, alterando o espaço coluna, que antes era o espaço coluna X e agora é o espaço coluna $[X : Z]$. Consequentemente alteramos a matriz de projeção ortogonal .

Quanto ao modelo de análise de covariância generalizado, uma estatística do teste ainda será análoga às estatísticas (1.16) e (1.17), mas agora o espaço coluna é $[QX : QZ]$.

Seja X_0 uma matriz de delineamento reduzida, onde $\mathcal{L}(X_0) \subset \mathcal{L}(X)$. E sejam $P_{X_0}^*$ e A_{X_0} projeções ortogonal e não ortogonal sobre $\mathcal{L}([X_0 : Z])$. Assim,

expressando o numerador em termos de $\tilde{X}$, $\tilde{y}$, V , A_g e A_{g0} temos

$$\begin{aligned}
 \tilde{y}'(P^* - P_{g0}^*)\tilde{y} &= \tilde{y}'Q' \{ [QX(X'Q'QX)^{-1}X'Q' + ((I - QX(X'Q'QX)^{-1}X'Q')QZ) \\
 &\quad [Z'Q'(I - QX(X'Q'QX)^{-1}X'Q')QZ]^{-1}((I - QX(X'Q'QX)^{-1}X'Q')QZ)'] \\
 &\quad - [QX_0(X_0'Q'QX_0)^{-1}X_0'Q' + ((I - QX_0(X_0'Q'QX_0)^{-1}X_0'Q')QZ) \\
 &\quad [Z'Q'(I - QX_0(X_0'Q'QX_0)^{-1}X_0'Q')QZ]^{-1} \\
 &\quad ((I - QX_0(X_0'Q'QX_0)^{-1}X_0'Q')QZ)'] \} Q\tilde{y} \\
 &= \tilde{y}' \{ [V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} + ((V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})Z) \\
 &\quad [Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})Z]^{-1}((V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})Z)'] \\
 &\quad - [V^{-1}X_0(X_0'V^{-1}X_0)^{-1}X_0'V^{-1} + ((V^{-1} - V^{-1}X_0(X_0'V^{-1}X_0)^{-1}X_0'V^{-1})Z) \\
 &\quad [Z'(V^{-1} - V^{-1}X_0(X_0'V^{-1}X_0)^{-1}X_0'V^{-1})Z]^{-1} \\
 &\quad ((V^{-1} - V^{-1}X_0(X_0'V^{-1}X_0)^{-1}X_0'V^{-1})Z)'] \} \tilde{y} \\
 &= \tilde{y}'V^{-1} \{ [A + (I - A)Z[Z'(I - A)V^{-1}Z]^{-1}Z'(I - A)V^{-1}] \\
 &\quad - [A_0 + (I - A_0)Z[Z'(I - A_0)V^{-1}Z]^{-1}Z'(I - A_0)V^{-1}] \} \tilde{y} \\
 &= \tilde{y}'V^{-1} \{ [A + (I - A)Z[Z'(I - A)V^{-1}(I - A)Z]^{-1}Z'(I - A)V^{-1}] \\
 &\quad - [A_0 + (I - A_0)Z[Z'(I - A_0)V^{-1}(I - A_0)Z]^{-1}Z'(I - A_0)V^{-1}] \} \tilde{y} \\
 &= \tilde{y}'V^{-1}(A_g - A_{g0})\tilde{y}, \\
 &= \tilde{y}'(A_g - A_{g0})'V^{-1}(A_g - A_{g0})\tilde{y}, \text{ por (1.10),}
 \end{aligned}$$

Sabemos por (2.6) que

$$\tilde{y}'(I - P^*) \tilde{y}^* = \tilde{y}'(I - A_g)V^{-1}(I - A_g) \tilde{y} ,$$

logo, a estatística do teste , sob o modelo de covariância generalizado, é

$$F = \frac{\tilde{y}'(A_g - A_{g0})'V^{-1}(A_g - A_{g0})\tilde{y} / r(A_g - A_{g0})}{\tilde{y}'(I - A_g)'V^{-1}(I - A_g)\tilde{y} / r(I - A_g)} .$$

2.2 Modelo com um Fator e uma Covariável

Vamos analisar o modelo de classificação em p níveis de um fator, além de uma covariável. Este é o modelo (2.1) onde

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \sim & \sim & \sim & & \sim \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \sim & \vdots \\ \sim & \sim & \sim & & \sim \\ \vdots & \vdots & \sim & \sim & \mathbf{0} \\ & & & & \sim \\ \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{1} \\ \sim & \sim & & \sim & \sim \end{bmatrix}, \quad Z = \begin{bmatrix} z_{11} \\ \vdots \\ z_{1ni} \\ \vdots \\ z_{p1} \\ \vdots \\ z_{pnp} \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \text{ e } \underline{\gamma} = [\gamma] .$$

A equação do modelo para a j-ésima observação da variável y, pertencente ao i-ésimo nível, fixando a covariável no valor z_{ij} é

$$y_{ij} = \beta_0 + \beta_i + \gamma z_{ij} + e_{ij} . \quad (2.7)$$

A interpretação deste modelo é que o valor esperado da variável y é dado por uma reta de inclinação γ , independente do nível do fator, reta esta, função da variável Z , cujo intercepto varia conforme o nível do fator.

Proposição 2.9

Os estimadores de MQO de $X\beta$ e γ são

$$X\hat{\beta} = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_1 \\ \vdots \\ \bar{y}_p \\ \vdots \\ \bar{y}_p \end{bmatrix} - \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_1 \\ \vdots \\ \bar{z}_p \\ \vdots \\ \bar{z}_p \end{bmatrix} \frac{\sum_{i=1}^p \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(z_{ij} - \bar{z}_i)}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2}$$

e

$$\hat{\gamma} = \frac{\sum_{i=1}^p \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(z_{ij} - \bar{z}_i)}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2}.$$

Prova :

Como a matriz Z tem apenas uma coluna, pela proposição 2.3, temos

$$\begin{bmatrix} X\hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} P_x y - P_x z [z'(I - P_x)z]^{-1} z'(I - P_x) y \\ [z'(I - P_x)z]^{-1} z'(I - P_x) y \end{bmatrix}.$$

Vamos agora expressar estes estimadores em termos das variáveis $\underline{y}$ e $\underline{z}$. Pelo exemplo 1.1, Capítulo 1, página 14, temos que

$$P_x \underline{y} = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_1 \\ \vdots \\ \bar{y}_p \\ \vdots \\ \bar{y}_p \end{bmatrix} \Rightarrow (I - P_x) \underline{y} = \begin{bmatrix} y_{11} - \bar{y}_1 \\ \vdots \\ y_{1n_1} - \bar{y}_1 \\ \vdots \\ y_{p1} - \bar{y}_p \\ \vdots \\ y_{pn_p} - \bar{y}_p \end{bmatrix}.$$

De forma análoga a $P_x \underline{y}$ e $(I - P_x) \underline{y}$ temos :

$$P_x \underline{z} = \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_1 \\ \vdots \\ \bar{z}_p \\ \vdots \\ \bar{z}_p \end{bmatrix} \text{ e } (I - P_x) \underline{z} = \begin{bmatrix} z_{11} - \bar{z}_1 \\ \vdots \\ z_{1n_1} - \bar{z}_1 \\ \vdots \\ z_{p1} - \bar{z}_p \\ \vdots \\ z_{pn_p} - \bar{z}_p \end{bmatrix}$$

$$\Rightarrow \underline{z}' (I - P_x) \underline{z} = [\underline{z}' (I - P_x)] [\underline{z}' (I - P_x)]'$$

$$= \sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2$$

$$\Rightarrow \quad [\underline{z}' (I - P_X) \underline{z}]^{-1} = \frac{1}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} .$$

Analogamente à $\underline{z}' (I - P_X) \underline{z}$ temos :

$$\begin{aligned} \underline{z}' (I - P_X) \underline{y} &= [\underline{z}' (I - P_X)] [\underline{y}' (I - P_X)]' \\ &= \sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.}) (y_{ij} - \bar{y}_{i.}) . \end{aligned}$$

Logo o resultado segue. ■

Proposição 2.10

A esperança da variável y , para o i -ésimo nível do fator considerado, fixando a covariável em z_{ij} , é estimada por

$$\bar{y}_{i.} + (z_{ij} - \bar{z}_{i.}) \hat{\gamma} . \quad (2.8)$$

Prova :

$$\begin{aligned} \hat{E}(y_{ij}) &= (X\hat{\beta})_{ij} + \hat{\gamma} z_{ij} \\ &= \bar{y}_{i.} - \bar{z}_{i.} \hat{\gamma} + \hat{\gamma} z_{ij} , \text{ pela proposição 2.9} \\ &= \bar{y}_{i.} + (z_{ij} - \bar{z}_{i.}) \hat{\gamma} . \end{aligned}$$

■

Sejam os modelos de GM, de regressão linear simples, de y em z , separadamente para cada nível do fator, com equações básicas dadas por:

$$y_{ij} = \delta_{0i} + \delta_{1i}z_{ij} + e_{ij}, \quad j=1, \dots, n_i, \text{ separadamente para } i = 1, \dots, p.$$

Então, as retas de mínimos quadrados, pelo exemplo 1.1, Capítulo 1, página 18, são dadas por:

$$\begin{aligned} \hat{E}(y_i) &= \bar{y}_{i.} + (z_{ij} - \bar{z}_{i.}) \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})(z_{ij} - \bar{z}_{i.})}{\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} \\ &= \bar{y}_{i.} + (z_{ij} - \bar{z}_{i.}) \hat{\delta}_{1i} \end{aligned} \quad (2.9)$$

Ou ainda, podemos reescrever $\hat{E}(y_i)$ em termos do coeficiente de correlação amostral

$$r_i = \frac{\sum_{j=1}^n (y_{ij} - \bar{y}_{i.})(z_{ij} - \bar{z}_{i.})}{[\sum_{i=1}^n (z_i - \bar{z}_{i.})^2 \sum_{i=1}^n (y_i - \bar{y}_{i.})^2]^{\frac{1}{2}}},$$

isto é,

$$\hat{E}(y_i) = \bar{y}_{i.} + (z_{ij} - \bar{z}_{i.}) \left[\frac{\sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2}{\sum_{j=1}^n (z_{ij} - \bar{z}_{i.})^2} \right]^{\frac{1}{2}} r_i. \quad (2.10)$$

No Capítulo 1, página 14, vimos que $\hat{E}(y_{ij}) = \bar{y}_{i.}$, $i = 1, \dots, p$, para o modelo de classificação de um fator em p níveis. Observando (2.10), notamos que se a correlação linear entre y e z for próxima de zero, os valores esperados de y_i serão aproximadamente iguais sob os modelos com ou sem a covariável.

Proposição 2.11

O estimador de MQO da inclinação γ no modelo de análise de covariância, $\hat{\gamma}$, é uma combinação linear dos estimadores de MQO dos parâmetros de inclinação das retas separadas por nível

$$\hat{\delta}_{11}, \dots, \hat{\delta}_{ip}. \quad (2.11)$$

Prova:

Pela proposição 2.9, temos

$$\begin{aligned} \hat{\gamma} &= \frac{\sum_{i=1}^p \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})(z_{ij} - \bar{z}_{i.})}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} \\ &= \sum_{i=1}^p \left[\frac{\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} \right] \left[\frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})(z_{ij} - \bar{z}_{i.})}{\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} \right] \end{aligned}$$

$$= \sum_{i=1}^p \left[\frac{\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2}{\sum_{i=1}^p \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2} \right] \hat{\delta}_{1i}$$

■

O resultado da proposição 2.11 é razoável porque se assumimos um modelo que especifica a mesma inclinação para a reta y em z , independente do nível do fator, o estimador da inclinação deveria ser alguma função dos estimadores das inclinações das retas individuais por nível. Observamos que maiores pesos são dados aos estimadores de inclinação de reta individual para níveis onde há maior espalhamento dos valores da variável z .

Quanto aos interceptos da reta do modelo de análise de covariância e do modelo de reta individual, para o i -ésimo nível do fator, os estimadores de MQO tem a estrutura comum :

$$\bar{y}_{i.} - \bar{z}_{i.} \hat{\gamma}, \text{ sob modelo de análise de covariância}$$

e

$$\bar{y}_{i.} - \bar{z}_{i.} \hat{\delta}_{1i}, \text{ sob modelo de reta no } i\text{-ésimo nível,}$$

lembrando que $\hat{\gamma}$ e $\hat{\delta}_{1i}$ estimam a mesma quantidade, mas, $\hat{\gamma}$ é baseado em um número maior de observações.

Observemos ainda que por (2.8) temos que o valor esperado da variável y , dentro do i -ésimo nível, para um valor fixo z da variável regressora, é dado pela reta

$\bar{y}_{i.} + (z - \bar{z}_{i.})\hat{\gamma}$, $i=1, \dots, p$, sob modelo de análise de covariância.

Comparando estes valores esperados, para, por exemplo, $z = \bar{z}_{..}$, a média de todos os valores utilizados da variável regressora, temos

$$\begin{aligned} &\bar{y}_{1.} + (\bar{z}_{..} - \bar{z}_{1.})\hat{\gamma} \\ &\vdots \\ &\bar{y}_{p.} + (\bar{z}_{..} - \bar{z}_{p.})\hat{\gamma} \end{aligned} \tag{2.12}$$

como valores esperados para os vários níveis.

Os estimadores em (2.12) não são importantes por si próprios, mas sim as diferenças destes, que estimam as diferenças entre níveis. Vamos comparar os estimadores em (2.12) com $\bar{y}_{1.}, \dots, \bar{y}_{p.}$, estimadores das médias dos níveis, sob o respectivo modelo sem covariável.

Em primeiro lugar notemos que se os valores $\bar{z}_{1.}, \dots, \bar{z}_{p.}$ forem praticamente iguais, os estimadores em (2.12) serão aproximadamente iguais a $\bar{y}_{1.}, \dots, \bar{y}_{p.}$, respectivamente. No entanto, se a correlação entre a z e y for forte, a soma de quadrados do erro será bem menor no modelo com covariável, e também o quadrado médio do erro, apesar de ter um denominador menor. Desta forma, uma diferença significativa entre os níveis poderia ficar evidente apenas no modelo de análise de covariância. E uma diferença já significativa no modelo sem covariável, seria mais acentuada no modelo com covariável. Por outro lado, se a correlação não é forte, os quadrados médios do erro podem ser aproximadamente iguais nos dois modelos e neste caso os dois modelos teriam o mesmo desempenho.

O que ocorre se os valores $\bar{z}_{1.}, \dots, \bar{z}_{p.}$ forem bem dispersos ?

Se observarmos $\hat{\gamma}$ positivo,

$$\bar{z}_{i.} > \bar{z}_{..} \Rightarrow \bar{y}_{i.} + (\bar{z}_{..} - \bar{z}_{i.})\hat{\gamma} < \bar{y}_{i.}$$

e

$$\bar{z}_{i.} < \bar{z}_{..} \Rightarrow \bar{y}_{i.} + (\bar{z}_{..} - \bar{z}_{i.})\hat{\gamma} > \bar{y}_{i.},$$

sendo a diferença entre $\bar{y}_{i.} + (\bar{z}_{..} - \bar{z}_{i.})\hat{\gamma}$ e $\bar{y}_{i.}$ mais pronunciada para os níveis correspondentes aos maiores e menores valores do conjunto $\bar{z}_{1.}, \dots, \bar{z}_{p.}$. Se a correlação entre y e a covariável é forte (positiva), aos menores e maiores valores do conjunto $\bar{z}_{1.}, \dots, \bar{z}_{p.}$ podem corresponder os menores e maiores valores do conjunto $\bar{y}_{1.}, \dots, \bar{y}_{p.}$, respectivamente. Assim as estimativas de (2.12) seriam menos dispersas em comparação às estimativas sob o modelo sem covariável. Por esta razão, às vezes, diferenças que parecem significantes, sob o modelo sem covariável, ficam diluídas sob o prisma do modelo de análise de covariância, mesmo que neste a soma de quadrados do erro seja bem menor. Os dois modelos concordariam, no entanto, se sob o modelo sem covariável não houver evidência de diferenças entre os níveis. Se, no entanto, a correlação entre y e a covariável não é forte, os quadrados médios do erro podem vir a ser aproximadamente iguais, de modo que sob os dois modelos teríamos resultados similares.

Nesta seção consideramos a inclinação γ independente do nível do fator. Por isso antes de fazermos inferências, ou mesmo aplicar o teste da hipótese de igualdade entre os níveis do fator, deveríamos verificar como se comportam as inclinações em cada nível do fator.

Considere o modelo

$$y_{ij} = \beta_0 + \beta_i + \gamma z_{ij} + \sum_{i=1}^p \theta_i z_{ij} x_{ij} + e_{ij}, i = 1, \dots, p, \quad (2.13)$$

onde

x_{ij} é o j -ésimo valor de variável indicadora do i -ésimo nível .

Este modelo é o mesmo modelo (2.7) acrescentado das interações entre os níveis dos fatores e a covariável z . Isto é, estamos considerando agora, que para cada nível do fator existe uma reta com uma inclinação diferente. Então, para verificarmos se estas retas tem uma mesma inclinação , iremos testar a seguinte hipótese :

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_p \text{ ou } H_0 : \theta_2 - \theta_1 = \dots = \theta_p - \theta_1 = \mathbf{0}, \text{ ou ainda}$$

$$H_0 : \Lambda \theta = \begin{bmatrix} 0 & 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 & -1 \end{bmatrix} \begin{bmatrix} \gamma \\ \theta_1 \\ \vdots \\ \theta_p \end{bmatrix} = \mathbf{0},$$

hipótese testável, pois $\lambda'_j = [0, \sum_{i=1}^p \lambda_{ij}, \lambda_{1j}, \dots, \lambda_{pj}]$, $j = 1, \dots, p - 1$.

2.3 Ilustrações

Nesta seção são apresentados inicialmente quatro exemplos para avaliar o desempenho de covariáveis quando no modelo de classificação por um fator . Em seguida é apresentado o exemplo comentado na introdução desta dissertação para

ilustração do modelo de análise de covariância quando conhecemos a matriz V . No Capítulo 3 este exemplo será mais explorado. Em todos os exemplos deste Capítulo a suposição de normalidade e homogeneidade da variância do erro foi verificada, não sendo apresentadas as análises por não haver nada especial a ser destacado. Também foi verificado, nos exemplos, a suposição de igualdade de inclinação das retas dos diferentes níveis.

Exemplo 1 :

Snedecor, G. W. & Cochran, W. G. (1982) apresentam um exemplo de um experimento realizado sobre o uso de drogas no tratamento de certo tipo de lepra, num hospital de crianças (Eversley Childs Sanitarium, nas Filipinas). Numa etapa preliminar em cada paciente foram selecionadas seis regiões do corpo, infectadas por leprosy bacilli. Foi, então, medida através de testes laboratoriais a quantidade de leprosy bacilli nessas regiões sendo produzido um escore, antes de ser realizado um tratamento com drogas. Meses após um tratamento com drogas, a mesma medida, observada na etapa preliminar ao tratamento, foi observada, obtendo-se um escore para a quantidade de leprosy bacilli.

No experimento foram utilizadas três drogas, dois antibióticos A e B e C, uma droga neutra usada como controle. Dez pacientes foram selecionados para cada droga. O escore da quantidade de leprosy bacilli, medida antes do tratamento, é considerado como uma covariável (z) e o escore da quantidade de leprosy bacilli, medida após o tratamento, como a variável dependente (y). Os dados estão apresentados na tabela 2.1.

Tabela 2.1
Escore para leprosy bacilli antes (z) e após (y) o tratamento

A		B		C	
z	y	z	y	z	y
11	6	6	0	16	13
8	0	6	2	13	10
5	2	7	3	11	18
14	8	8	1	9	5
19	11	18	18	21	23
6	4	8	4	16	12
10	13	19	14	12	5
6	1	8	9	12	16
11	8	5	1	7	1
3	0	15	9	12	20
$\bar{z}_{1.} = 9,3$	$\bar{y}_{1.} = 5,3$	$\bar{z}_{2.} = 10,0$	$\bar{y}_{2.} = 6,0$	$\bar{z}_{3.} = 12,9$	$\bar{y}_{3.} = 10,3$

Esquecendo a covariável e usando o modelo de análise de variância, nós teremos o modelo de GM (1.2), Capítulo 1, onde

y é o vetor dos escores da quantidade de leprosy bacilli, medida após o tratamento, em cada um dos 30 pacientes ;

X é uma matriz de delineamento (30×4), do tipo da página 77 e

$\beta' = [\beta_0, \beta_1, \beta_2, \beta_3]'$, onde β_0 representa o efeito médio das drogas e β_i , $i = 1, 2, 3$, representam os efeitos adicionais individuais das drogas .

Usando os resultados apresentados na página 82, as estimativas para o escore

esperado após tratamento com as drogas A, B e C são, respectivamente,

$$\bar{y}_{1.} = 5,3, \quad \bar{y}_{2.} = 6,0 \quad \text{e} \quad \bar{y}_{3.} = 10,3.$$

Resta investigar se os escores médios são estatisticamente diferentes. Então, para testar se existem diferenças entre as drogas A, B e C consideramos as hipóteses

$$H_0: \beta_1 = \beta_2 = \beta_3 = \delta \text{ ou } \Lambda \underline{\beta} = \underline{0}, \text{ onde } \Lambda = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

versus

$$H_1: \beta_i \neq \beta_j, \text{ para pelo menos um } i \neq j, \text{ ou } \Lambda \underline{\beta} \neq \underline{0}, \quad (2.14)$$

e obtemos o seguinte resultado.

Tabela ANOVA

FV	GL	SS	MS
Modelo(droga)	2	293,6	146,8
Erro	27	995,1	36,856
Total corrigido	29	1288,7	—

Logo, $F = 3,98 > F_{(0,05; 2; 27)} = 3,35$ implica em termos alguma evidência em rejeitar a hipótese H_0 , ou seja, temos evidência de diferenças entre as drogas.

Agora, considerando a covariável, nosso modelo passa a ser o modelo (2.1), deste Capítulo, onde

$\underline{z}$ é o vetor dos escores da quantidade de leprosy bacilli, medida antes do tratamento, em cada um dos 30 pacientes.

Obtemos para estimativas do escore médio, para $\bar{z}_{..} = 10,7$, os seguintes

valores, para as drogas A, B e C :

$$\begin{aligned}
 A & 5,3 + (10,7 - 9,3) 0,99 = 6,686 \\
 B & 6,0 + (10,7 - 10,0) 0,99 = 6,693 \\
 C & 10,3 + (10,7 - 12,9) 0,99 = 8,122 .
 \end{aligned} \tag{2.15}$$

Resta, agora, analisar se são estatisticamente diferentes. Assim, considerando a covariável no modelo, obtivemos a seguinte tabela de análise de covariância para o teste do par de hipóteses em (2.14)

Tabela ANCOVA

FV	GL	SS	MS
Modelo restrito	1	803,000	803,000
Hipótese $\Delta \beta = 0$	2	68,554	34,277
Erro	26	417,203	16,046
Total corrigido	29	1288,757	—

Logo, $F = 2,14 < F_{(0.05; 2; 26)} = 3,37$ implica em não rejeitarmos a hipótese H_0 , ou seja, não temos evidência de que existem diferenças entre as drogas.

Assim, observamos divergências com o uso dos dois modelos. Ao que tudo indica, as diferenças que antes se apresentavam, poderiam não ser devidas às diferentes drogas, mas a uma outra fonte de variação. Para entendermos melhor o que está acontecendo observemos os próximos gráficos .

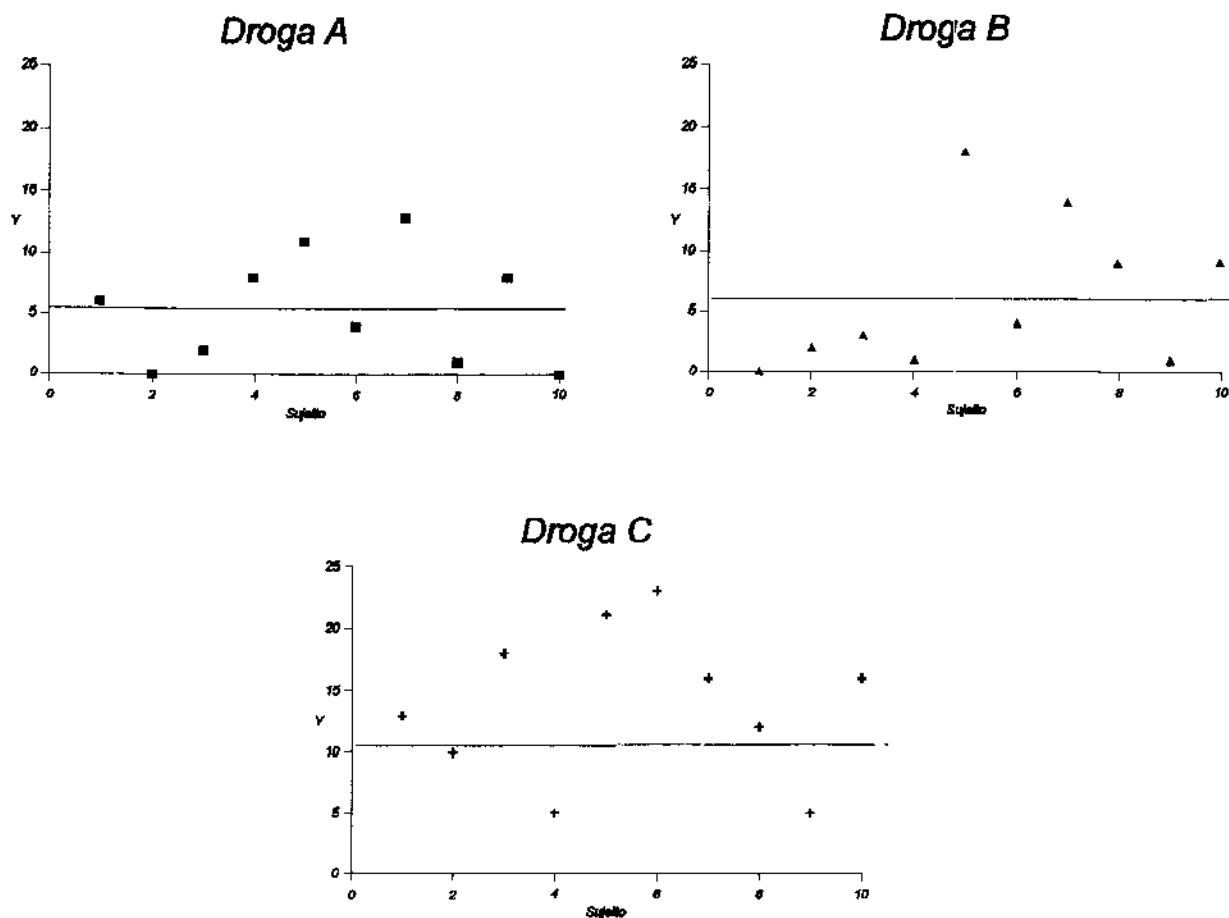


Gráfico 2.1 : Escores de leprosy bacilli, após meses de tratamento (10 crianças por droga). Reta de estimativa do escore esperado.

O gráfico 2.1 ilustra a informação usada na análise de variância (sem a covariável). As retas indicam as estimativas de escores médios e os resíduos que geram o valor do quadrado médio do erro de 36,86 podem ser visualizados. Notemos que a estimativa do escore médio, sob a droga C, foi mais elevada do que as estimativas das outras drogas. Mesmo em meio à variação ilustrada, um teste, a nível 0,05, conclui pela

rejeição da hipótese de igualdade dos efeitos das drogas.

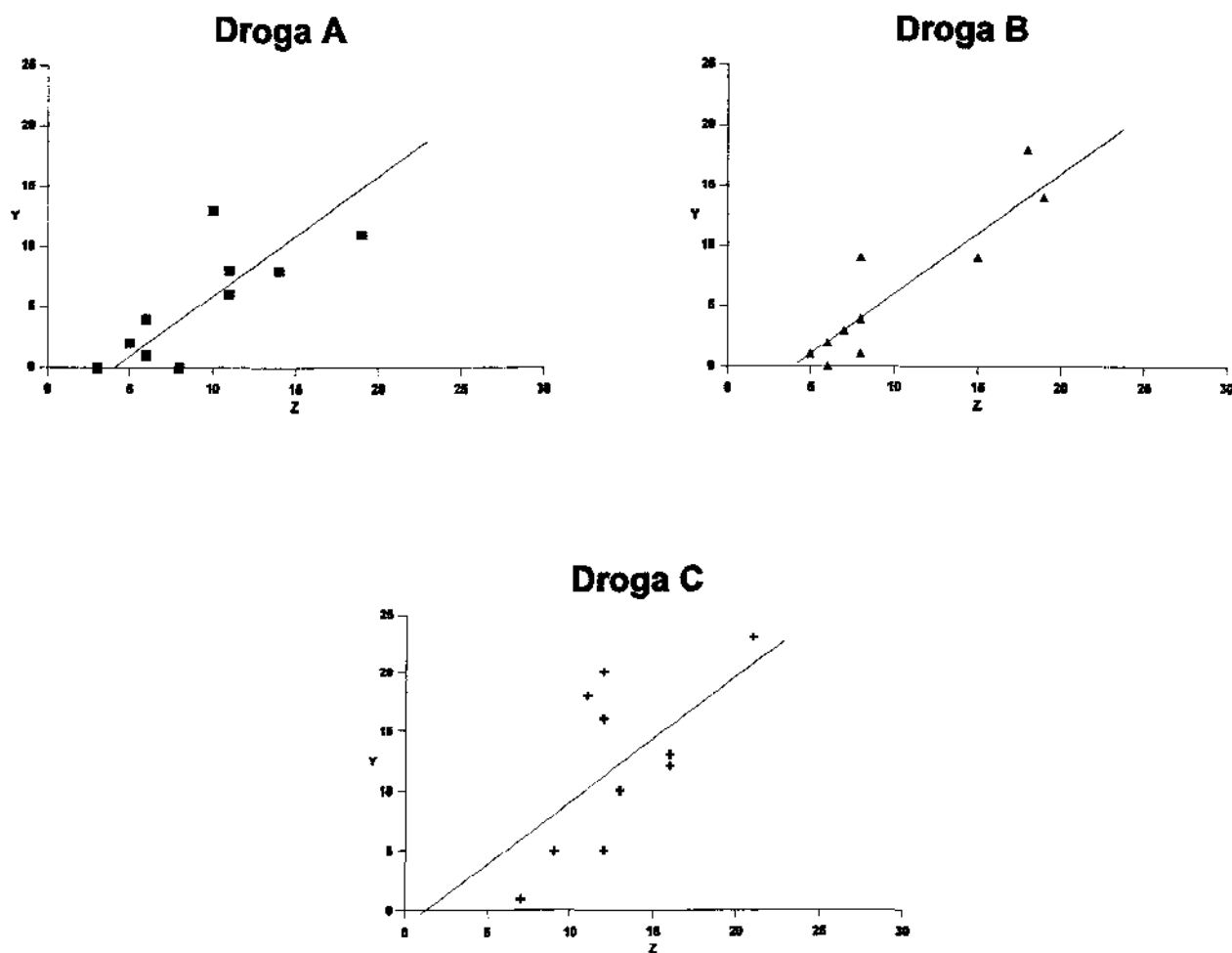


Gráfico 2.2 : Escores de leprosy bacilli antes, z (eixo horizontal), e depois, y (eixo vertical), de tratamento. Retas de mínimos quadrados, sob o modelo de análise de covariância.

O gráfico 2.2 ilustra a informação utilizada na análise de covariância. Podemos notar que a maioria das crianças que receberam a droga C (Controle) já estavam a nível mais

avanzado da doença, contrário ao observado com as crianças que receberam as drogas A e B, no início dos tratamentos. Observando a linha inferior da tabela 2.1, confirmamos isto, verificando que os escores médios da quantidade de leprosy bacilli, antes do tratamento, são bem próximos para A e B, sendo inferiores ao valor de C . Como as medidas anteriores e após o tratamento parecem ser diretamente correlacionadas, os valores em (2.15), usados para estimativas das diferenças entre escores esperados, estão menos dispersos do que os valores obtidos pelo modelo sem covariável. Isto contribuiu para um numerador menor na estatística F do teste de igualdade de efeitos das drogas, e, apesar de observarmos uma menor variação em torno das retas, ou seja, um MSE menor, o valor observado da estatística F não superou o ponto crítico para rejeição.

□

Exemplo 2

Neter, J. & Wasserman, W. (1977) apresentam um exemplo sobre o estudo de uma empresa que deseja analisar o efeito de três esquemas promocionais de vendas para seus biscoitos. Os esquemas são :

Esquema 1. Colocar nas lojas estante com amostras grátis do produto para os consumidores.

Esquema 2. Aumentar o espaço que as lojas reservam regularmente para expor o produto aos clientes.

Esquema 3. Colocar material publicitário sobre o produto (cartazes, panfletos, ...) no final do corredor da loja onde os biscoitos ficam normalmente expostos.

Para o experimento foram selecionadas 15 lojas (unidades experimentais). Cada

loja foi aleatoriamente designada para trabalhar com um dos três esquemas promocionais de vendas. Em todas as lojas manteve-se o mesmo preço do produto e padrão básico de propagandas.

Os dados referentes às vendas do produto no " período promocional ", denotados por y , são mostrados na tabela 2.2, onde também se encontram os dados referentes às vendas do produto num período precedente, denotados por z .

Tabela 2.2
Quantidades de vendas no período precedente (z) e durante (y) as promoções

Esquema 1		Esquema 2		Esquema 3	
z	y	z	y	z	y
21	38	34	43	23	24
26	39	26	38	29	32
22	36	29	38	30	31
28	45	18	27	16	21
19	33	25	34	29	28
$\bar{z}_1 = 23,2$	$\bar{y}_1 = 38,2$	$\bar{z}_2 = 26,4$	$\bar{y}_2 = 36,0$	$\bar{z}_3 = 25,4$	$\bar{y}_3 = 27,2$

Esquecendo a covariável e usando o modelo de análise de variância, nós teremos o modelo de GM (1.2), Capítulo 1, onde

y é o vetor das quantidades de vendas durante a promoção nas quinze lojas.

X é uma matriz de delineamento (15×4), do tipo da página 77 e

$\beta' = [\beta_0, \beta_1, \beta_2, \beta_3]$, onde β_0 representa o efeito médio dos esquemas e β_i ,

$i = 1, 2, 3$, representam os efeitos adicionais individuais dos esquemas.

Usando os resultados apresentados na página 82, as estimativas para a quantidade

de vendas esperadas após os esquemas 1, 2 e 3 são, respectivamente,

$$\bar{y}_{1.} = 38,2 \quad , \quad \bar{y}_{2.} = 36,0 \quad \text{e} \quad \bar{y}_{3.} = 27,2.$$

Resta investigar se esses resultados são estatisticamente diferentes. Então, para testar se existem diferenças entre os esquemas 1, 2 e 3 consideramos as mesmas hipóteses (2.14), do exemplo anterior, e obtemos o seguinte resultado.

Tabela ANOVA

FV	GL	SS	MS
Modelo(droga)	2	338,80	169,40
Erro	12	307,60	25,63
Total corrigido	14	646,40	—

Logo, $F = 6,63 > F_{(0,05; 2; 12)} = 3,88$ implica em termos alguma evidência em rejeitar a hipótese H_0 , ou seja, temos evidência de diferenças entre os esquemas promocionais.

Agora, considerando a covariável, nosso modelo passa a ser o modelo (2.1), deste Capítulo, onde

$\mathbf{z}$ é o vetor das quantidades de vendas no período precedente ao esquema, em cada uma das 15 lojas,

obtemos para estimativas das quantidades de vendas média, para $\bar{z}_{..} = 25$, os seguintes valores, para os esquemas 1, 2 e 3 :

$$\begin{aligned} 1 & \quad 38,2 + (25 - 23,2)0,99 = 39,82 \\ 2 & \quad 36,0 + (25 - 26,4)0,99 = 34,74 \\ 3 & \quad 27,2 + (25 - 25,4)0,99 = 26,84, \end{aligned} \quad (2.16)$$

que são mais dispersos do que as estimativas sob o modelo sem covariável.

Considerando a covariável no modelo, obtivemos a seguinte tabela de análise de covariância para o teste do par de hipóteses em (2.12).

Tabela ANCOVA

FV	GL	SS	MS
Modelo restrito	1	198,678	198,678
Hipótese $\Lambda\beta = 0$	2	417,151	208,575
Erro	11	38,571	3,506
Total corrigido	14	646,400	-

Assim, $F = 59,48 > F_{(0,05; 2; 11)} = 3,98$ implica em rejeitarmos a hipótese H_0 . Neste caso não observamos divergências com o uso dos dois modelos. Porém, com o uso da covariável temos mais evidência em rejeitarmos a hipótese (2.12). Observemos os próximos gráficos.

O gráfico 2.3 ilustra a análise de variância. Notemos que as estimativas das quantidades de vendas média, sob os esquemas 1 e 2 foram mais elevadas do que a estimativa sob o esquema 3. Aliado ao fato de que o valor estimado do quadrado médio do erro 25,63 indicar pouca variação, um teste, a nível de 0,05, conclui pela rejeição da hipótese de igualdade dos efeitos dos tratamentos.

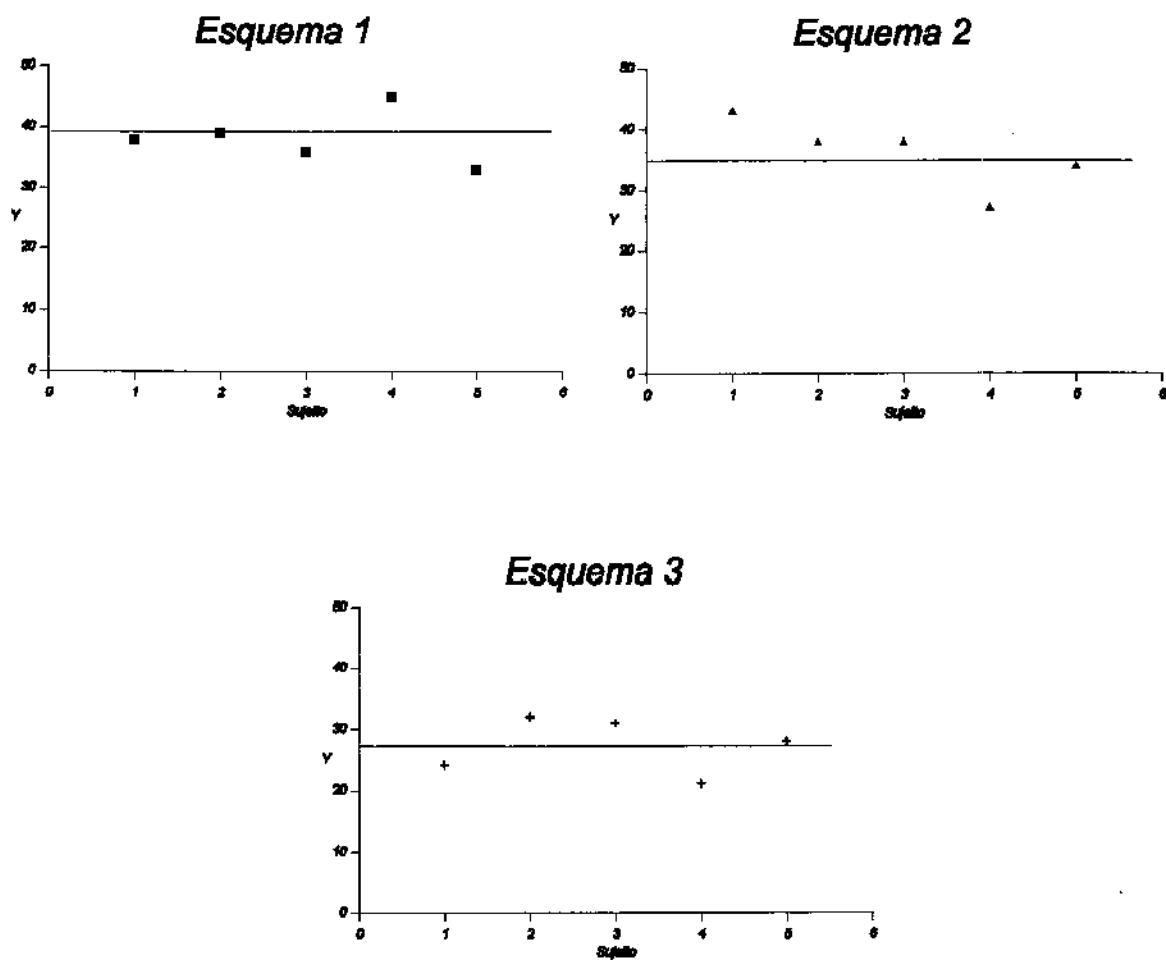


Gráfico 2.3 : Quantidades de vendas após esquema (5 lojas por Esquema). Reta de estimativa da quantidade de vendas esperada.

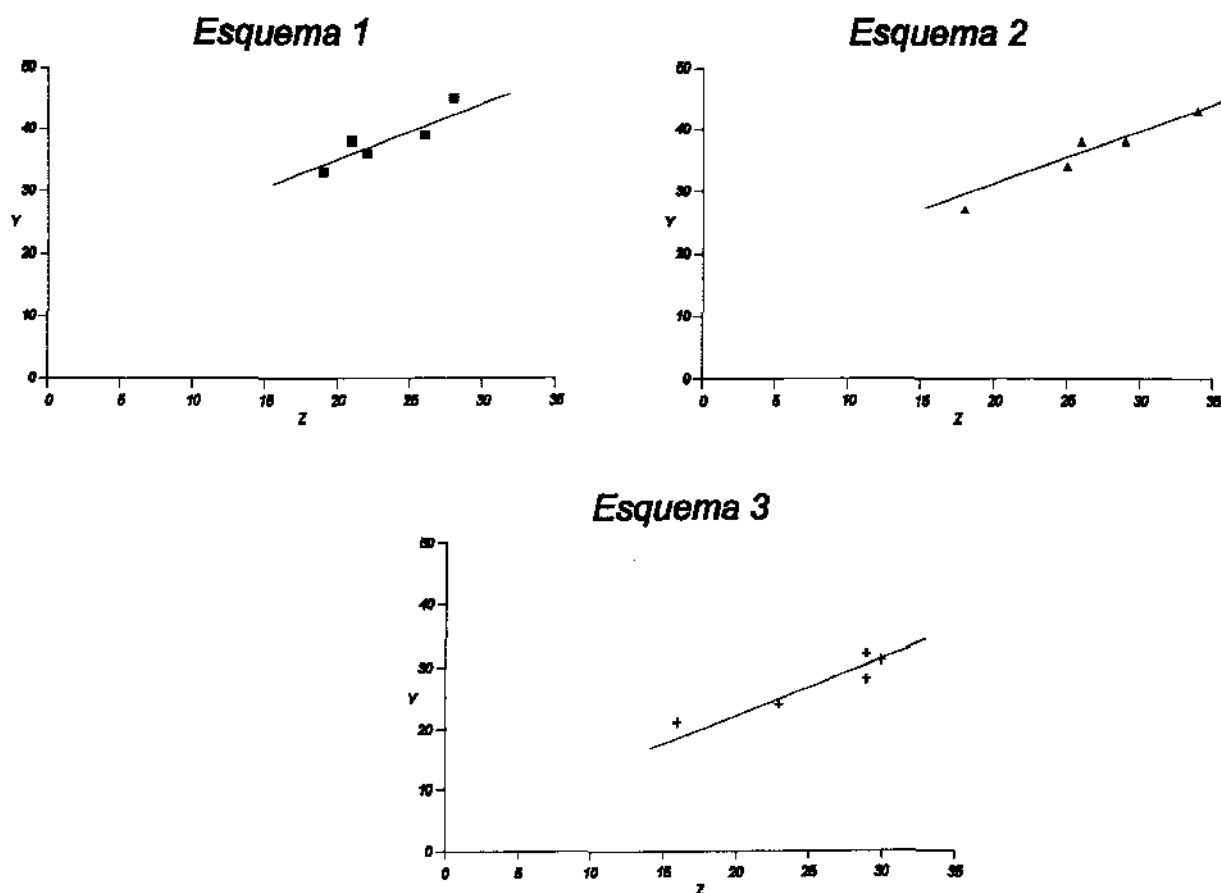


Gráfico 2.4 : Quantidade de vendas antes, z , e após, y , esquema promocional. Retas de mínimos quadrados , sob o modelo de análise de covariância.

Observando a linha inferior da tabela 2.2, notamos que a venda média, anterior ao esquema promocional 1 foi menor. Dada a correlação positiva observada entre vendas pré e pós esquema promocionais, os valores em (2.16), usados para estimativas das diferenças entre as quantidades de vendas esperadas, estão mais dispersos do que os valores obtidos pelo modelo sem covariável, contribuindo para um numerador maior na estatística F do teste de igualdade dos efeitos dos tratamentos . Além do acréscimo

do numerador, no denominador houve uma redução razoavelmente grande, de modo que o valor observado da estatística F supera o ponto crítico para rejeição, com mais evidência.

□

Exemplo 3

Kleinbaum, D. G., Kupper, L. L. & Muller, K. E. (1988) apresentam um exemplo sobre o estudo do efeito do sódio na urina de cachorros. Para o experimento foram selecionados 37 cachorros, onde 30 receberam um tratamento com droga (grupo controle) e 7 não receberam nenhuma droga (grupo experimental). De cada cachorro foi retirada uma amostra de urina e obtida a quantidade de sódio. Os dados referentes às quantidades percentuais de sódio na urina após os tratamentos, denotados por y , são mostrados nas tabelas 2.3a e 2.3b, onde também se encontram os dados referentes às quantidades percentuais de sódio na urina num período precedente aos tratamentos, denotados por z .

Tabela 2.3b

Quantidades de sódio na urina no período precedente (z) e posterior (y) ao tratamento 2

Tratamento 2	
y	z
11,1	9,4
5,1	5,9
6,5	14,8
17,2	15,5
11,8	23,4
6,6	7,3
4,1	8,2
$\bar{y}_2 = 12,1$	$\bar{z}_1 = 8,9$

Tabela 2.3a

Quantidades de sódio na urina no período precedente (z) e posterior (y) ao tratamento 1

Tratamento 1		Tratamento 1	
z	y	z	y
17,5	22,1	9,4	11,5
9,4	12,0	8,2	8,4
10,0	15,2	6,3	12,7
7,4	23,1	9,7	17,1
8,8	9,8	7,1	9,5
18,9	26,9	7,2	11,0
10,8	11,1	5,3	8,2
8,8	13,6	14,3	15,8
8,8	12,8	7,9	9,5
9,2	7,5	14,1	14,7
8,1	8,1	12,8	17,0
10,3	27,5	12,8	20,2
10,1	11,2	10,7	13,9
7,3	11,0	5,9	11,8
11,1	15,3	3,8	9,0
		$\bar{z}_{1.} = 9,7$	$\bar{y}_{1.} = 13,9$

Esquecendo a covariável e usando o modelo de análise de variância, nós teremos o modelo de GM (1.2), Capítulo 1, onde

$\mathbf{y}$ é o vetor das quantidades de sódio após tratamento nas 37 amostras de urina dos cachorros.

$\mathbf{X}$ é uma matriz de delineamento (37×3), do tipo da página 77 e

$\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2]$, onde β_0 representa o efeito médio dos tratamentos e β_1 ,

$i = 1, 2$, representam os efeitos adicionais individuais dos tratamentos.

Usando os resultados apresentados na página 82, as estimativas para a quantidade de sódio esperada após tratamento 1 e 2 são, respectivamente,

$$\bar{y}_{1.} = 13,9 \text{ e } \bar{y}_{2.} = 12,1.$$

Resta investigar se esses resultados são estatisticamente diferentes. Então, para testar se existem diferenças entre os tratamentos 1 e 2 consideramos as hipóteses

$$H_0: \beta_1 - \beta_2 = \Lambda\beta = 0 \text{ versus } H_0: \Lambda\beta \neq 0, \text{ onde } \Lambda = [0 \ 1 \ -1] \quad (2.17)$$

e obtemos o seguinte resultado.

Tabela ANOVA

FV	GL	SS	MS
Modelo(droga)	1	19,465	19,465
Erro	35	1067,408	30,497
Total corrigido	36	1086,873	—

Logo, $F = 0.64 < F_{(0.05; 1; 35)} = 4,12$ implica que não temos evidência para rejeitar a hipótese H_0 , ou seja, não temos evidência de diferenças entre os tratamentos.

Considerando a covariável, nosso modelo passa a ser o modelo (2.1), deste Capítulo, onde

z é o vetor das quantidades de sódio no período anterior ao tratamento nas 37 amostras de urina dos cachorros.

Obtemos para estimativas das quantidades de vendas média, para $\bar{z}_{..} = 9,3$ e os seguintes valores, para os tratamentos 1 e 2 :

$$\begin{aligned} 1 & \quad 13,9 + (9,3 - 9,7) 0,96 = 13,516 \\ 2 & \quad 12,1 + (9,3 - 8,9) 0,96 = 12,484, \end{aligned} \quad (2.18)$$

que são praticamente iguais às estimativas sob o modelo sem covariável. Assim, considerando a covariável no modelo, obtivemos a seguinte tabela de análise de covariância para o teste do par de hipóteses em (2.17).

Tabela ANCOVA

FV	GL	SS	MS
Modelo restrito	1	434,487	434,487
Hipótese $\Delta\beta = 0$	1	6,376	6,376
Erro	34	646,010	19,000
Total corrigido	36	1086,873	—

Portanto, $F = 0,34 < F_{(0,05; 1; 34)} = 4,14$ implica novamente em não rejeitarmos a hipótese H_0 . Assim, não observamos divergências com o uso dos dois modelos. Porém, com o uso da covariável tivemos um modelo mais ajustado, dado o valor menor do MSE. Observemos os próximos gráficos.

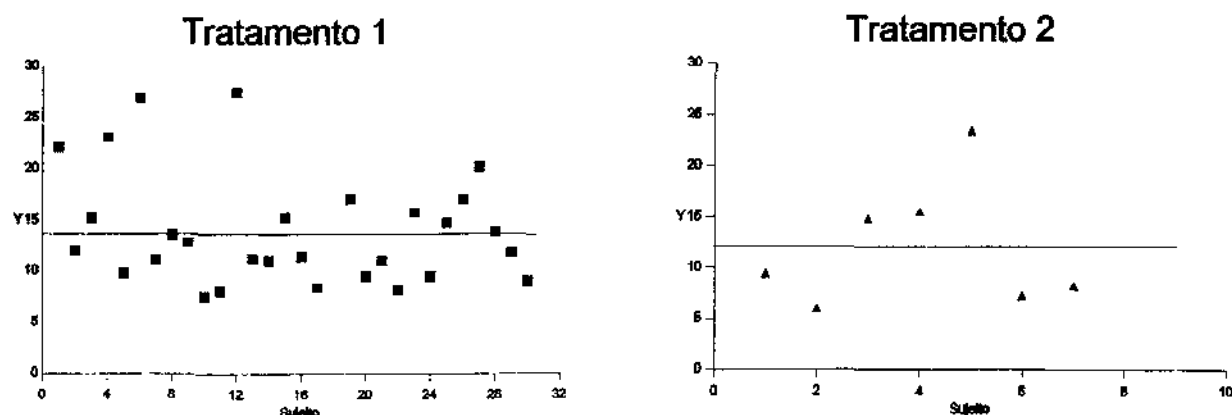


Gráfico 2.5 : Quantidade de sódio na urina , após tratamento . Reta de estimativa da quantidade de sódio esperada.

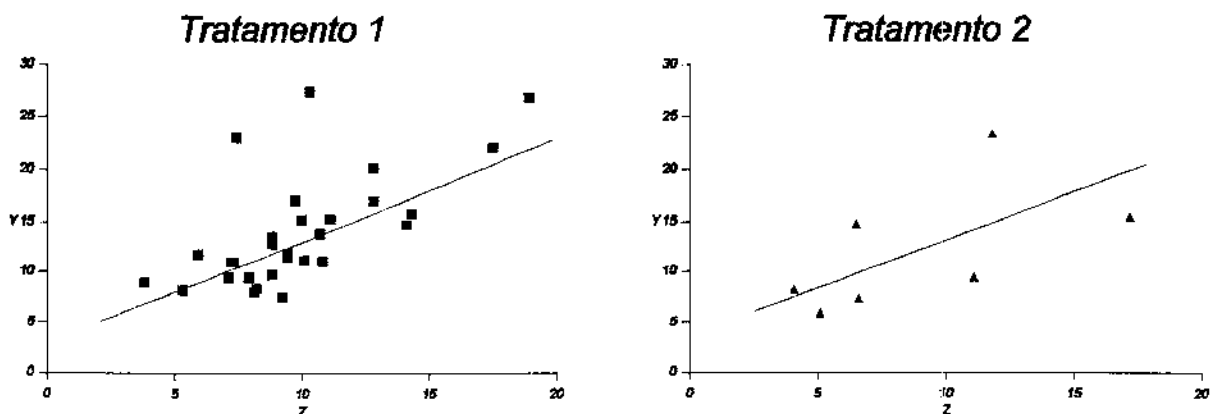


Gráfico 2.6 : Quantidade de sódio na urina antes, z , e após, y , os tratamentos. Retas de mínimos quadrados , sob o modelo de análise de covariância.

No gráfico 2.5 notamos que as estimativas das quantidades de sódio média , sob o grupo sob a droga foi um pouco mais elevada que sob o grupo controle. Porém há grande variação em torno destes valores .

Através do gráfico 2.6, e observando as linhas inferiores das tabelas 2.3a e 2.3b, notamos que a quantidade de sódio média, anterior ao uso da droga foi apenas um pouco menor no grupo controle. Os valores em (2.18) estão tão dispersos quanto os valores obtidos pelo modelo sem covariável. Observarmos uma menor variação em torno das retas, mas apesar disso, o valor observado da estatística F não superou o ponto crítico para rejeição.

□

Exemplo 4 :

Este exemplo é fictício. Através dele onde desejamos ilustrar a verificação da existência da diferença entre três tratamentos, segundo um determinado escore. Os dados estão apresentados na tabela 2.4.

Tabela 2.4
Escore antes (z) e após (y) o tratamento

A		B		C	
z	y	z	y	z	y
3,4	7,2	5,2	7	6,2	6
1,2	5	3,2	5	4,2	4,5
2,3	6,1	4	6,2	5,2	5,4
5,1	9	2,3	4,4	3,2	3,4
4,2	8,3	6	8,1	7	7,1
$\bar{z}_{1.} = 3,24$	$\bar{y}_{1.} = 7,12$	$\bar{z}_{2.} = 4,14$	$\bar{y}_{2.} = 6,14$	$\bar{z}_{3.} = 5,16$	$\bar{y}_{3.} = 5,28$

Esquecendo a covariável e usando o modelo de análise de variância, nós teremos o modelo de GM (1.2), Capítulo 1, onde

y é o vetor dos escores após o tratamento, em cada um dos 15 pacientes ;

X é uma matriz de delineamento (15×4), do tipo da página 77 e

$\beta' = [\beta_0, \beta_1, \beta_2, \beta_3]'$, onde β_0 representa o efeito médio dos tratamentos e

$\beta_i, i = 1, 2, 3$, representam os efeitos adicionais individuais dos tratamentos .

Usando os resultados apresentados na página 82, as estimativas para o escore

esperado após tratamento A, B e C são, respectivamente,

$$\bar{y}_{1.} = 7,12 \quad , \quad \bar{y}_{2.} = 6,14 \quad \text{e} \quad \bar{y}_{3.} = 5,28.$$

Resta investigar se os escores médios são estatisticamente diferentes. Então, para testar se existem diferenças entre os tratamentos A, B e C consideramos as hipóteses (2.14) e obtemos o seguinte resultado.

Tabela ANOVA

FV	GL	SS	MS
Modelo(droga)	2	8,476	4,230
Erro	12	27,368	2,280
Total corrigido	14	35,844	-

Logo, $F = 1,86 < F_{(0,05; 2; 12)} = 3,88$ implica em termos alguma evidência para não rejeitar a hipótese H_0 , ou seja, não temos evidência de diferenças entre os tratamentos.

Agora, considerando a covariável, nosso modelo passa a ser o modelo (2.1), deste Capítulo, onde

z é o vetor dos escores medida antes do tratamento, em cada um dos 15 pacientes,

obtemos para estimativas do escore médio, para $\bar{z}_{..} = 4,18$, os seguintes valores, para os tratamentos A, B e C :

$$\begin{aligned} \text{A} \quad & 7,12 + (4,18 - 3,24) 0,99 = 8,050 \\ \text{B} \quad & 6,14 + (4,18 - 4,14) 0,99 = 6,180 \\ \text{C} \quad & 5,28 + (4,18 - 5,16) 0,99 = 4,310 \end{aligned} \quad (2.15)$$

Resta, agora, analisar se são estatisticamente diferentes. Assim, considerando a covariável no modelo, obtivemos a seguinte tabela de análise de covariância para o teste do par de hipóteses em (2.14)

Tabela ANCOVA

FV	GL	SS	MS
Modelo restrito	1	9,269	9,269
Hipótese $\Lambda \beta = 0$	2	26,222	13,111
Erro	11	0,353	0,032
Total corrigido	14	35,844	-

Logo, $F = 408,2 > F_{(0,05; 2; 11)} = 3,98$ implica em rejeitarmos a hipótese H_0 , ou seja, temos evidência de que existem diferenças entre os tratamentos

Assim, observamos divergências com o uso dos dois modelos. Ao que tudo indica, as diferenças que antes não se apresentavam, poderiam não ser devidas aos diferentes tratamentos, mas, a dispersão em torno do valor esperado. Para entendermos melhor o que está acontecendo observemos os próximos gráficos.

O gráfico 2.7 ilustra a informação usada na análise de variância (sem a covariável). As retas indicam as estimativas de escores médios e os resíduos que geram o valor do quadrado médio do erro de 2,28 podem ser visualizados. Notemos que a estimativa do escore médio, sob o tratamento A, foi mais elevada do que as estimativas dos outros tratamentos. Um teste, a nível 0,05, conclui pela não rejeição da hipótese de igualdade dos efeitos dos tratamentos.

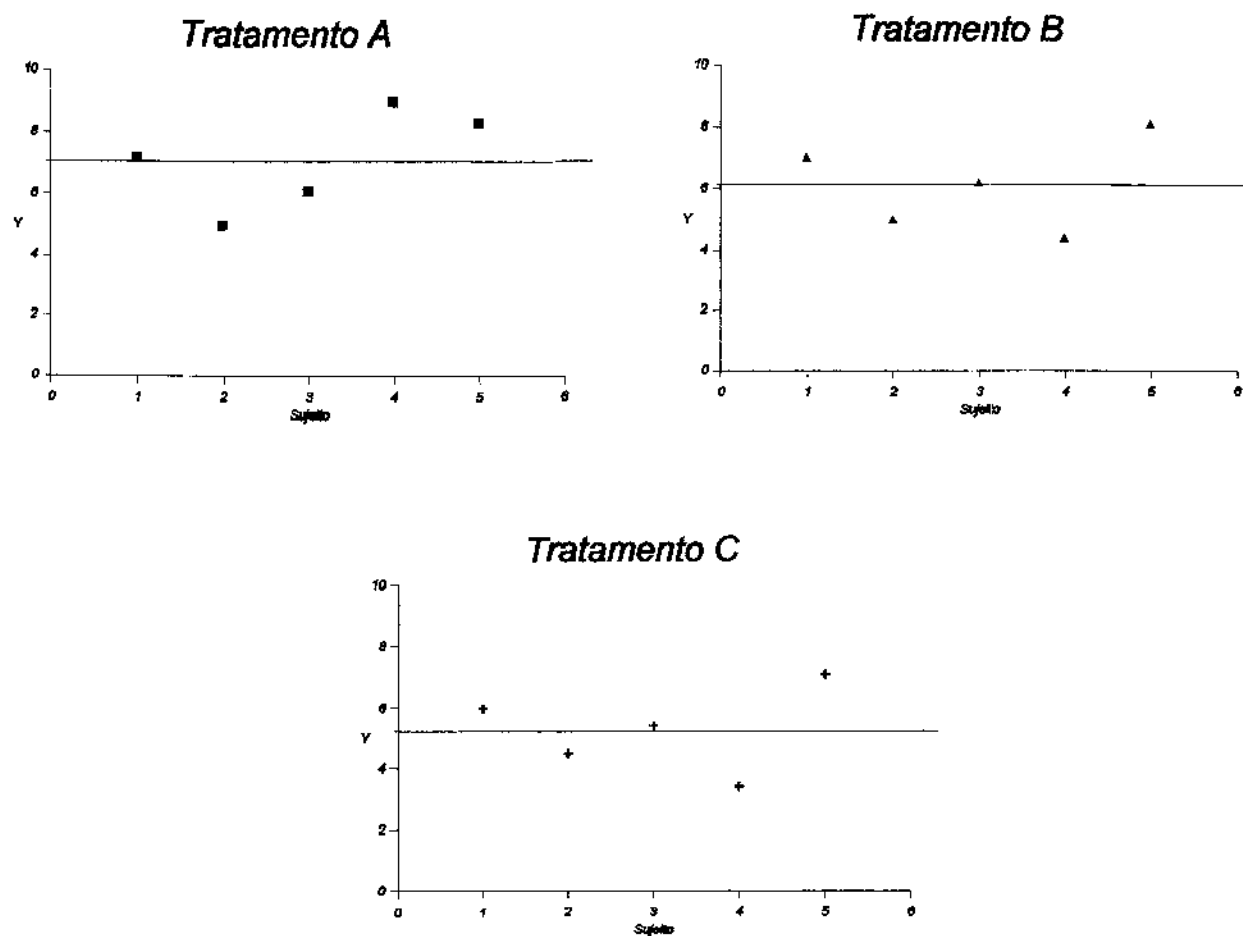


Gráfico 2.7 : Escores após tratamento (5 pacientes por tratamento). Reta de estimativa do escore esperado.

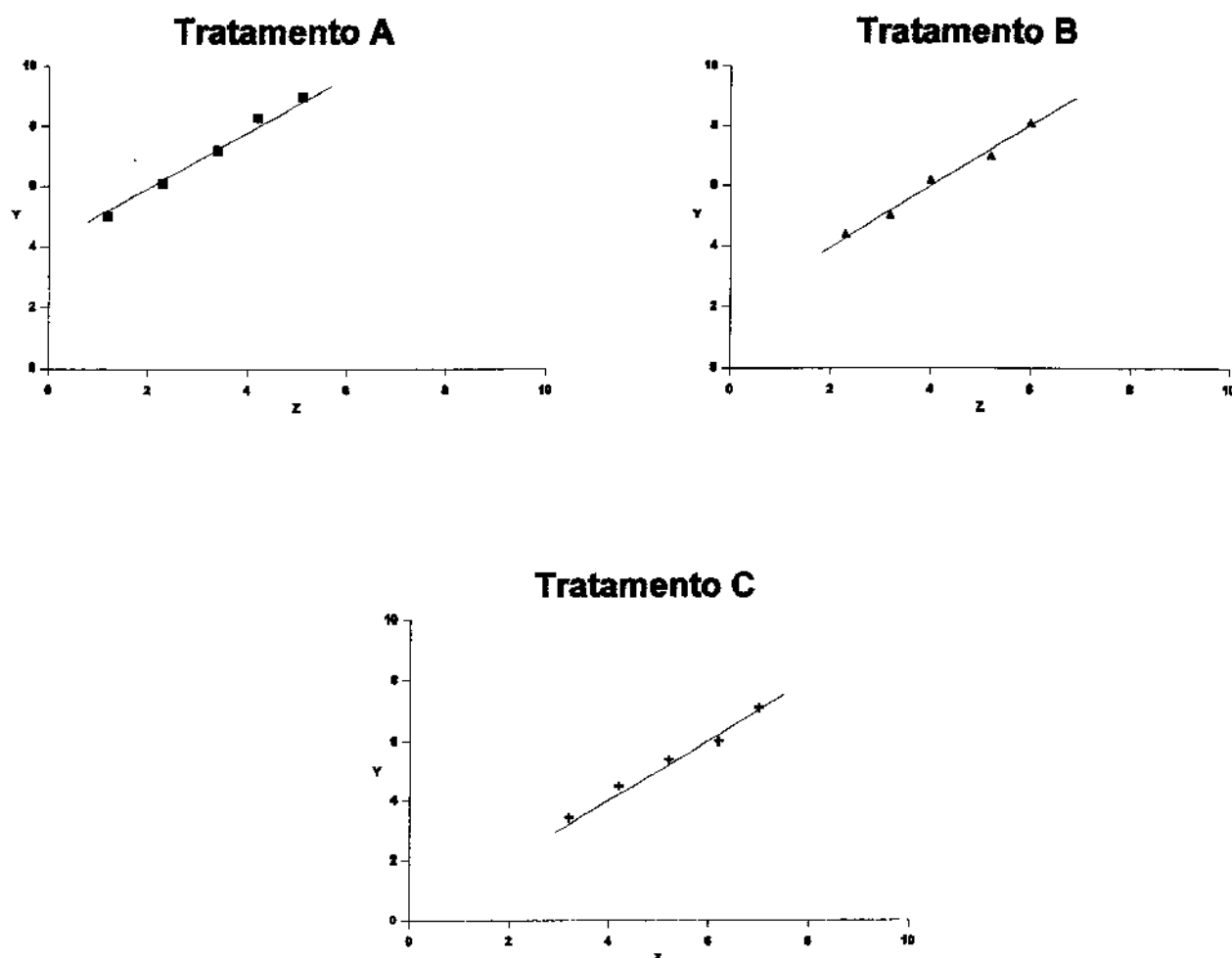


Gráfico 2.8 : Escores antes, z (eixo horizontal), e depois, y (eixo vertical), de tratamento. Retas de mínimos quadrados, sob o modelo de análise de covariância.

O gráfico 2.8 ilustra a informação utilizada na análise de covariância. Observando a linha inferior da tabela 2.4, verificamos que os escores médios antes do tratamento, são maiores para B e C. Como as medidas anteriores e após o tratamento parecem ser diretamente correlacionadas, os valores em (2.15), usados para estimativas das

diferenças entre escores esperados, estão mais dispersos do que os valores obtidos pelo modelo sem covariável. Isto contribuiu para um numerador maior na estatística F do teste de igualdade de efeitos dos tratamentos. Além do acréscimo do numerador, no denominador houve uma grande redução, de modo que o valor observado da estatística F supera o ponto crítico para rejeição.

□

Exemplo 5

Neste exemplo apresentamos um estudo (não publicado) conduzido por Dr. João Camargo, sobre o nível de irrigação do sangue na cabeça do fêmur em pessoas acometidas de osteoartrose (OA). Os ortopedistas questionam se pessoas acometidas de OA, na cabeça do fêmur, podem estar com a circulação do sangue reduzida nesta área afetada. Um grupo de 12 pessoas acometidas de osteoartrose é observado (grupo OA). Essas pessoas estão situadas na faixa etária de 50 a 85 anos de idade. Um segundo grupo de 11 pessoas, da mesma faixa etária, é observado (grupo controle), tendo estas pessoas sofrido acidentes com fraturas do fêmur, devido a quedas em suas próprias residências. Uma medida da circulação do sangue é observada após a injeção de substância radioativa, sendo definida como a razão entre as duas quantidades :

contagem de radiação, por grama de osso, num período de 100 segundos e a quantidade de radiação por grama de células vermelhas administrada.

Duas localidades no fêmur são amostradas: a cabeça do fêmur (objetivo principal do estudo), sendo a medida de irrigação considerada como a variável dependente, y , e o grande trocanter, sendo a medida de irrigação considerado como uma covariável, z , sendo este último área de grande irrigação sanguínea, e , portanto, um tipo de controle,

dentro do mesmo indivíduo. No grupo controle, duas medidas são observadas para cada uma das 11 pessoas que o compõe, uma na cabeça do fêmur e outra no grande trocanter. No grupo de osteoartrose, uma medida é observada no grande trocanter e, na maioria dos casos, mais de uma medida são observadas na cabeça do fêmur, sendo medidas repetidas da mesma unidade. Os dados obtidos são apresentados abaixo.

Tabela 2.5

Medida de irrigação sanguínea na cabeça do fêmur e no grande trocanter nos grupos controle e OA

Grupo Controle		Grupo OA		
GT (z)	Cabeça(y)	GT(z)	Cabeça(y)	Cabeça Média
49	13	-	18, 14, 15, 19, 19	17
73	12	46	12, 17, 31, 26	21,5
58	9	50	15, 9	12,0
33	9	91	24, 9, 8, 10, 26, 13	15,0
84	10	16	13, 19, 6, 3, 12	10,6
15	8	97	31, 35	33,0
43	9	102	19, 20, 25	21,3
38	11	73	19, 19	19,0
90	28	-	14, 19, 18, 24, 18, 11	17,3
43	12	67	10	10
71	28	70	9, 13, 14, 10	11,5
		56	18	18

Uma maneira de trabalharmos com esses dados é considerarmos somente as médias das observações, de modo que teremos apenas uma observação por unidade. Agora, quando seguirmos um modelo de análise de covariância supomos que todos os y_{ij} provêm de uma distribuição normal com mesma média e variância. Porém, se

usarmos as médias, estaremos tratando de $\text{Var}(\bar{y}_{ij}) = \sigma^2 / n_{ij}$, ignorando covariâncias, onde n_{ij} é o número de observações na i -ésima unidade do i -ésimo nível, isto é, estaremos agora tratando do modelo de análise de covariância generalizado, onde

$$V = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & & & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & \frac{1}{4} & 0 & \dots & 0 \\ \vdots & & & \vdots & 0 & \frac{1}{2} & \dots & \vdots \\ \vdots & & & \vdots & \vdots & \vdots & \dots & 0 \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \} \text{ grupo controle} \\ \\ \\ \\ \\ \} \text{ grupo OA} \end{matrix}$$

e a covariável é a medida do grande trocanter. Observemos que no grupo OA estão faltando duas observações no grande trocante, por isso, só iremos considerar os indivíduos restantes.

Assim, verificar se a média de irrigação da cabeça do fêmur no grupo controle difere da média de irrigação da cabeça do fêmur do grupo OA, equivale a testar as hipóteses

$$H_0: \beta_1 - \beta_2 = 0$$

versus

$$H_1: \beta_1 - \beta_2 \neq 0. \quad (2.18)$$

De modo que obtemos a seguinte tabela .

Tabela ANCOVA

FV	GL	SS	MS
Modelo Restrito	1	1734,0957	1734,0957
Hipótese $H_0: \beta_1 - \beta_2 = 0$	1	21,3033	21,3033
Erro	18	1221,1089	67,8394
Total corrigido	20	2976,5079	-

Logo $F = 0,3140 < F_{(0,05; 1; 18)} = 4,4139$ não temos evidência para rejeitar a hipótese H_0 .

□

CAPÍTULO 3

Modelo Linear Misto de Laird - Ware

Neste Capítulo tratamos de modelos direcionados a análises onde as unidades experimentais ou observacionais são observadas várias vezes. Os modelos lineares apresentados nos Capítulos 1 e 2 poderiam ser adaptados, de modo a tratar adequadamente a existência de dependência entre as observações de uma mesma unidade. Além disso, a heterogeneidade entre unidades pode ser modelada através do uso de efeitos individuais aleatórios. Um modelo linear geral que se utiliza de efeitos fixos e aleatórios é o modelo de *Laird, N. M. & Ware, J. H. (1982)*. Abordamos especificamente o modelo de Laird - Ware com estrutura de covariância constante entre as observações de uma mesma unidade, denominado *modelo de Laird - Ware de simetria composta*. Este modelo pode vir a ser adequado para situações onde uma enumeração das observações não tenha nenhum significado.

Apresentamos os resultados de estimação de máxima verossimilhança, que coincidem em parte com a estimação sob o método de mínimos quadrados. Temos apenas um parâmetro, ao qual atribuímos valores arbitrários, no cálculo da função de verossimilhança e não há problemas em localizar um máximo global.

A utilização de efeitos aleatórios no modelo, sob a suposição de normalidade, nos fornece mais uma maneira de verificar a adequação do modelo (*Ryan, L. M. & Lange, N. (1989)*). Assim, para o problema que motivou este estudo, ajustamos modelos com e sem covariável. O gráfico dos efeitos aleatórios padronizados evidência a grande utilidade da covariável, pois, no modelo com covariável a configuração dos efeitos aleatórios parece bem razoável. O uso de covariável nos pareceu indispensável

na análise deste problema.

No Apêndice B apresentamos o programa usado para a obtenção das diversas estimativas. Este programa pode ser modificado para modelagens diversas de modelo de covariância com mais de uma covariável ou outros modelos polinomiais com as covariáveis.

3.1 Modelo Linear Misto de Laird -Ware de Simetria Composta

Os modelos apresentados até agora consideram efeitos dos tratamentos (níveis do fator) fixos. Iremos agora considerar uma abordagem onde utilizamos efeitos fixos e aleatórios.

Laird, N. M. & Ware, J. H. (1982) propuseram um modelo, denominado modelo linear misto de Laird-Ware, a partir do trabalho de Harville, D. A. (1977), englobando uma variedade de modelos com efeitos fixos e aleatórios e com variedade de estruturas de variâncias e covariâncias. O modelo linear misto de Laird- Ware definido para a i -ésima unidade do grupo k é dado por :

$$\underset{\sim}{y}_{ik} = \underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{L}_{ik} \underset{\sim}{\theta}_{ik} + \underset{\sim}{e}_{ik}, \quad (3.1)$$

onde

$\underset{\sim}{y}_{ik} = [y_{ik1}, \dots, y_{ikn_i}]'$ é um vetor coluna, $n_i \times 1$, das n_i observações obtidas da unidade i ;

m_k é o número de unidades do grupo k , com $\sum_{k=1}^p m_k = m$

X_{ik} é uma matriz conhecida $n_i \times p^1$, onde $i = 1, \dots, m_k$;

L_{ik} é uma matriz conhecida $n_i \times q$, onde $i = 1, \dots, m_k$;

$\beta = [\beta_1, \dots, \beta_p]'$ é um vetor de g parâmetros fixos desconhecidos;

$\theta_{ik} = [\theta_{ik1}, \dots, \theta_{ikq}]'$ é um vetor de $q \times 1$ de efeitos aleatórios que representam alguma característica, particular à unidade, onde a dimensão q é fixa para qualquer unidade, e $\theta_{ik} \sim N(\mathbf{0}; \sigma^2 B)$, onde σ^2 um escalar, e B uma matriz $q \times q$ desconhecida; $e_{ik} = [e_{ik1}, \dots, e_{ikn_i}]'$ é um vetor aleatório, $n_i \times 1$, tal que $e_{ik} \sim N(\mathbf{0}; \sigma^2 W_{ik})$, onde W_{ik} pode ser conhecida ou não, e, e_{ik} são independentes entre diferentes unidades e independentes dos vetores de efeitos aleatórios.

Notemos que sob este modelo para cada unidade i podemos ter diferentes números de observações. Notemos ainda que $\theta_{ik} \sim N(\mathbf{0}; \sigma^2 B)$ e $e_{ik} \sim N(\mathbf{0}; \sigma^2 W_{ik})$ implicam em

$$\begin{aligned} \text{Var}(\mathbf{y}_{ik}) &= \sigma^2 (L_{ik} B L_{ik}' + W_{ik}) = \sigma^2 V_{ik} \\ E(\mathbf{y}_{ik}) &= X_{ik} \beta \\ \Rightarrow \mathbf{y}_{ik} &\sim N(X_{ik} \beta; \sigma^2 V_{ik}). \end{aligned} \quad (3.2)$$

Observemos que a matriz de variâncias e covariâncias do vetor de observações da i -ésima unidade no grupo k evidencia duas fontes de variabilidade B e W_i . E, ainda

¹ O número de colunas da matriz X_{ik} , bem como o número de parâmetros fixos, aqui denotados por p , não são necessariamente o número de grupos, também denotado por p neste Capítulo.

se $B=0$ trata-se do modelo de GM normal ou Aitken, definido no Capítulo 1, onde cada unidade tem n_{ik} observações.

O modelo de Laird - Ware de Simetria Composta é o modelo que apresenta uma estrutura de matriz de variâncias e covariâncias do vetor de observações de uma unidade bem particular : covariância constante entre quaisquer observações de uma mesma unidade e efeito aleatório unidimensional. Uma situação para qual este modelo parece ser adequado é quando tomamos várias observações numa mesma unidade sendo que a ordem dessas observações não tem nenhum significado. Nesta dissertação só será considerada esta estrutura. O modelo de Laird-Ware de simetria composta para a i - ésima unidade do grupo k é dado por (3.1), com as seguintes particularidades

$$\underset{\sim}{y}_{ik} = \underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{1} \underset{\sim}{\theta}_{ik} + \underset{\sim}{e}_{ik}, \quad (3.3)$$

onde

$$\underset{\sim}{X}_{i1} = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{0} & \dots & \underset{\sim}{0} \\ \underset{\sim}{\sim} & \underset{\sim}{\sim} & & \underset{\sim}{\sim} \end{bmatrix} \text{ } n_{i1} \text{ linhas, } i = 1, \dots, m_1;$$

$$\underset{\sim}{X}_{i2} = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{1} & \underset{\sim}{0} & \dots & \underset{\sim}{0} \\ \underset{\sim}{\sim} & \underset{\sim}{\sim} & \underset{\sim}{\sim} & & \underset{\sim}{\sim} \end{bmatrix} \text{ } n_{i2} \text{ linhas, } i = 1, \dots, m_2;$$

$$\underset{\sim}{X}_{i3} = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{1} & \underset{\sim}{0} & \dots & \underset{\sim}{0} \\ \underset{\sim}{\sim} & \underset{\sim}{\sim} & \underset{\sim}{\sim} & \underset{\sim}{\sim} & & \underset{\sim}{\sim} \end{bmatrix} \text{ } n_{i3} \text{ linhas, } i = 1, \dots, m_3;$$

$\vdots$

$$\underset{\sim}{X}_{ip} = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{0} & \dots & \underset{\sim}{0} & \underset{\sim}{1} \\ \underset{\sim}{\sim} & \underset{\sim}{\sim} & & \underset{\sim}{\sim} & \underset{\sim}{\sim} \end{bmatrix} \text{ } n_{ip} \text{ linhas, } i = 1, \dots, m_p;$$

$\theta_{11}, \dots, \theta_{m1}, \dots, \theta_{1p}, \dots, \theta_{mp}$ são variáveis aleatórias unidimensionais $N(0; \sigma^2 b)$; $e_{ik}, i = 1, \dots, m_k$ e $k = 1, \dots, p$ são variáveis aleatórias multidimensionais $N(\mathbf{0}; \sigma^2 \tilde{I}_{i,k})$ e

os efeitos aleatórios $\theta_{11}, \dots, \theta_{mp}$ são referidos como "efeitos aleatórios entre as unidades". As variáveis $e_{11}, \dots, e_{mp}$ representam os fatores não controláveis e são referidas como "efeitos aleatórios dentro das unidades".

A definição das matrizes X_{ik} implica em que β_1 , o primeiro elemento do vetor β , seja o valor esperado para o grupo 1 e que $\beta_1 + \beta_k$, seja o valor esperado para o k -ésimo grupo, $k = 2, 3, \dots, p$.

Para esta estrutura de matriz de variâncias e covariâncias temos que

$$V_{ik} = \mathbf{1} \mathbf{b} \mathbf{1}' + I_{i,k} =$$

$$= \begin{bmatrix} b & \dots & \dots & b \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b & \dots & \dots & b \end{bmatrix} + \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} = \begin{bmatrix} b+1 & b & \dots & b \\ b & b+1 & \dots & \vdots \\ \vdots & \vdots & \ddots & b \\ b & \dots & b & b+1 \end{bmatrix}. \quad (3.4)$$

3.1.1 Estimação - Máxima Verossimilhança

No Capítulo 1 apresentamos o método de mínimos quadrados para estimar alguns parâmetros. Neste Capítulo iremos apresentar o método de máxima verossimilhança, que para o caso que estamos tratando equivale ao método de mínimos

quadrados generalizado, quando supomos as matrizes V_{ik} conhecidas, isto é, quando temos b fixo.

Por (3.2) e considerando a estrutura de simetria composta temos que a função de densidade conjunta das observações da i -ésima unidade, y_{ik} , é

$$f(\underline{y}_{ik}) = \frac{1}{(2\pi)^{\frac{n_i}{2}} (\sigma^2 |V_{ik}|)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\underline{y}_{ik} - X_{ik}\underline{\beta})' V_{ik}^{-1} (\underline{y}_{ik} - X_{ik}\underline{\beta}) \right\}. \quad (3.5)$$

Seja $L(\underline{\beta}, \sigma^2, b)$ a função de verossimilhança, a qual é obtida pelo produtório das densidades conjuntas (3.5) para todas unidades, considerando os parâmetros como variáveis e as observações fixas. Então,

$$L(\underline{\beta}, \sigma^2, b) = \prod_{k=1}^p \prod_{i=1}^{m_k} \frac{1}{(2\pi)^{\frac{n_i}{2}} (\sigma^2 |V_{ik}|)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\underline{y}_{ik} - X_{ik}\underline{\beta})' V_{ik}^{-1} (\underline{y}_{ik} - X_{ik}\underline{\beta}) \right\}, \quad (3.6)$$

de modo que os estimadores de máxima verossimilhança são obtidos maximizando $L(\underline{\beta}, \sigma^2, b)$, ou, equivalentemente, minimizando $-2\ln L = \ell$.

Resultado 3.1

O estimador de máxima verossimilhança de $\underline{\beta}$, fixado b , é

$$\hat{\underline{\beta}} = \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X_{ik}' V_{ik}^{-1} X_{ik} \right)^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X_{ik}' V_{ik}^{-1} \underline{y}_{ik} \right), \quad (3.7)$$

assumindo a existência da inversa.

Prova : Jones, R. H.(1993) .

■

Neste Capítulo só iremos usar a matriz $(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik})$ de posto completo. Logo, notemos que $\tilde{\beta}$ é estimável ou qualquer combinação linear dos elementos de $\tilde{\beta}$.

Resultado 3.2

O estimador de máxima verossimilhança da variância do erro, fixado b, é

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^p \sum_{i=1}^{m_k} (\tilde{y}_{ik} - X_{ik} \tilde{\hat{\beta}})' V_{ik}^{-1} (\tilde{y}_{ik} - X_{ik} \tilde{\hat{\beta}}), \quad (3.8)$$

onde n é total do número de observações.

Prova : Jones, R. H.(1993) .

■

Os resultados acima consideram b fixo, mas na realidade não conhecemos o valor b. Porém, pelo método de máxima verossimilhança, escolhemos arbitrariamente um valor positivo para b e em seguida calculamos $\tilde{\hat{\beta}}$, $\hat{\sigma}^2$ e ℓ , onde

$$\ell = -2 \ln L = n \ln (2\pi \hat{\sigma}^2) + \sum_{k=1}^p \sum_{i=1}^{m_k} \ln (|V_{ik}|) + n, \quad (3.9)$$

efetuando-se este procedimento para diversos valores para b, até ℓ ser minimizada.

Uma vez que ℓ seja mínima, temos as estimativas de máxima verossimilhança de $\tilde{\beta}$, σ^2 e b. Observemos que, fixado b, conhecemos as matrizes V_{ik} , logo, o modelo de

Laird-Ware equivale ao modelo de Aitken, normal, com a seguinte equação

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{m_p p} \end{bmatrix} = \begin{bmatrix} X_{11} \\ \vdots \\ X_{m_p p} \end{bmatrix} \beta + \begin{bmatrix} 1 \theta_{11} + e_{11} \\ \vdots \\ 1 \theta_{m_p p} + e_{m_p p} \end{bmatrix}, \quad (y = X\beta + e) \quad (3.10)$$

onde

$$E \begin{bmatrix} 1 \theta_{11} + e_{11} \\ \vdots \\ 1 \theta_{m_p p} + e_{m_p p} \end{bmatrix} = \mathbf{0};$$

$$\text{Var} \begin{bmatrix} 1 \theta_{11} + e_{11} \\ \vdots \\ 1 \theta_{m_p p} + e_{m_p p} \end{bmatrix} = \sigma^2 \begin{bmatrix} V_{11} & 0 & \dots & 0 \\ 0 & V_{21} & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & V_{m_p p} \end{bmatrix} = \sigma^2 V.$$

De modo que os estimadores de MQG de $\hat{\beta}$ e $\hat{\sigma}^2$ são respectivamente, pelo Capítulo 1, seções 1.2 e 1.4,

$$\hat{\beta} = (X'V^{-1}X)^{-1}(X'V^{-1}y) = \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} y_{ik} \right),$$

e

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{(n - r(X))} (\underline{y} - X \hat{\underline{\beta}})' V^{-1} (\underline{y} - X \hat{\underline{\beta}}) = \\ &= \frac{1}{(n - r(X))} \sum_{k=1}^p \sum_{i=1}^{m_k} (\underline{y}_{ik} - X_{ik} \hat{\underline{\beta}})' V_{ik}^{-1} (\underline{y}_{ik} - X_{ik} \hat{\underline{\beta}}),\end{aligned}$$

supondo que a inversa de $(X'V^{-1}X)^{-1}$ exista.

Logo, percebemos que o estimador de máxima verossimilhança para $\underline{\beta}$ é exatamente o estimador de MQG e o estimador de máxima verossimilhança de $\tilde{\sigma}^2$ é equivalente ao estimador de MQG (denominador menor para o MQG).

Vamos expressar esses estimadores em função dos totais, ou médias das unidades e de b, para o caso do modelo de Laird- Ware de simetria composta.

Resultado 3.3

No contexto do modelo de simetria composta, o estimador de máxima verossimilhança do valor esperado do grupo K é dado por :

$$\frac{\sum_{i=1}^{m_k} (y_{ik} / 1 + b n_{ik})}{\sum_{i=1}^{m_k} (n_{ik} / 1 + b n_{ik})} = \begin{cases} \hat{\beta}_1 & \text{se } k=1 \\ \hat{\beta}_1 + \hat{\beta}_k & \text{se } k \neq 1 \end{cases}$$

onde y_{ik} é a soma das observações da unidade i, no grupo k.

Prova : Guimarães, Paulo R. B. (1994) .

Resultado 3.4

No contexto do modelo de simetria composta, o estimador de máxima verossimilhança de σ^2 é dado por :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^p \sum_{i=1}^{m_k} \left(\frac{\tilde{y}_{ik} \tilde{y}_{ik}}{\tilde{y}_{ik.}} - \frac{(\sqrt{\hat{b}} \tilde{y}_{ik.})^2}{(1 + \hat{b} n_{ik})} \right),$$

onde $\tilde{y}_{ik} = y_{ik} - X_{ik} \hat{\beta}$ é o vetor de resíduos da i-ésima unidade do grupo k e

$\tilde{y}_{ik.}$ é a soma dos elementos do vetor $\tilde{y}_{ik}$.

Prova : *Guimarães, Paulo R. B.*(1994) .

■

Exemplo 3.1

Para ilustrarmos o resultado 3.3, consideremos que num grupo k, $k \neq 1$, temos 3 unidades, com $n_{1k} = 5$, $n_{2k} = 20$ e $n_{3k} = 50$ unidades. Vamos analisar o que ocorre se a estimativa de máxima verossimilhança de b for maior do que 1 (a) ou menor do que 1 (b).

a) Suponha $\hat{b} = 1,2$, então,

$$\begin{aligned} \hat{\beta}_1 + \hat{\beta}_k &= \left[\frac{y_{1k.}}{1 + 5(1,2)} + \frac{y_{2k.}}{1 + 20(1,2)} + \frac{y_{3k.}}{1 + 50(1,2)} \right] + \left[\frac{5}{7} + \frac{20}{25} + \frac{50}{61} \right] \\ &= \frac{y_{1k.}}{6,3} + \frac{y_{2k.}}{58,3} + \frac{y_{3k.}}{142,4} \end{aligned}$$

$$= 0,30 \bar{y}_{1k} + 0,34 \bar{y}_{2k} + 0,36 \bar{y}_{3k},$$

onde $\bar{y}_{ik}$ é a média das observações da i-ésima unidade do grupo k.

b) Suponha $\hat{\delta} = 0,6$, então,

$$\begin{aligned}\hat{\beta}_1 + \hat{\beta}_k &= \left[\frac{y_{1k.}}{1 + 5(0,6)} + \frac{y_{2k.}}{1 + 20(0,6)} + \frac{y_{3k.}}{1 + 50(0,6)} \right] + \left[\frac{5}{4} + \frac{20}{13} + \frac{50}{31} \right] \\ &= \frac{y_{1k.}}{17,61} + \frac{y_{2k.}}{57,22} + \frac{y_{3k.}}{136,44} \\ &= 0,28 \bar{y}_{1k} + 0,35 \bar{y}_{2k} + 0,37 \bar{y}_{3k}.\end{aligned}$$

□

Pelo exemplo é ilustrado que o estimador das médias dos grupos são médias ponderadas das médias por unidade, sempre com maior peso para aquelas unidades com maior número de observações. Observemos que se $\hat{\delta}$ é maior do que 1, estamos inferindo que a variância do erro é menor do que a variância do efeito aleatório e o contrário se $\hat{\delta}$ é menor do que 1. Quando a variância do erro é maior do que a variância do efeito aleatório (caso b), ou seja, mais variação dentro de unidades do que entre as médias das unidades, a ponderação dada à média é maior quanto mais observações houver. Assim, quando $\hat{\delta}$ é menor que 1, os pesos das médias dos indivíduos são mais dispersos, em comparação ao caso com $\hat{\delta}$ maior do que 1.

3.1.2 Teste de Hipóteses

Quanto à estatística do teste para uma hipótese $\Lambda \boldsymbol{\beta} = \mathbf{0}$, ela é análoga às estatísticas (1.16) e (1.17) do Capítulo 1, onde o MSE apresentado naquele Capítulo equivale ao $\hat{\sigma}^2$ deste Capítulo, com pequena diferença no denominador. Uma observação a ser feita é que neste Capítulo quando quisermos testar a igualdade das médias dos grupos, devemos testar a hipótese $\beta_2 = \dots = \beta_p = 0$, porque $\beta_2, \dots, \beta_p$ são os incrementos das médias dos grupos em relação à média do grupo 1.

3.1.3 Predição dos Efeitos Aleatórios do Modelo de Laird - Ware de Simetria Composta

Resultado 3.5

Um preditor do efeito aleatório da i -ésima unidade do grupo k , θ_{ik} , no contexto do modelo geral de Laird-Ware é

$$\hat{\theta}_{ik} = BL'_i V_{ik}^{-1} (y_{ik} - X_{ik} \hat{\boldsymbol{\beta}}).$$

Prova : Harville, D. A. (1976) .

■

Proposição 3.1

O preditor do efeito aleatório da i -ésima unidade do grupo k , θ_{ik} , no contexto do modelo de Laird-Ware de simetria composta é dado por

$$\hat{\theta}_{ik} = \frac{1}{1/\hat{\sigma}^2 + n_{ik}} y_{ik}.$$

Prova :

No contexto do modelo de simetria composta $B = b$ e $L_i = \mathbf{1}_i$, logo pelo resultado 3.5, segue que

$$\hat{\theta}_{ik} = \hat{b} \mathbf{1}_i' V_{ik}^{-1} (y_{ik} - X_{ik} \hat{\beta}) . \quad (3.11)$$

Temos em *Guimarães, P. R. B.* (1994) que no modelo geral de Laird-Ware

$$V_{ik}^{-1} = W_i^{-1} - W_i^{-1} L_i [B^{-1} + L_i' W_i^{-1} L_i]^{-1} L_i' W_i^{-1} .$$

Logo, temos no modelo de simetria composta

$$\begin{aligned} V_{ik}^{-1} &= I_i - \mathbf{1}_i [1/\hat{b} + \mathbf{1}_i' \mathbf{1}_i]^{-1} \mathbf{1}_i' \\ &= I_i - \frac{\hat{b} \mathbf{1}_i \mathbf{1}_i'}{1 + \hat{b} n_{ik}} . \end{aligned} \quad (3.12)$$

Sabemos que $(y_{ik} - X_{ik} \hat{\beta}) = y_{ik}$ é o vetor de resíduos da i -ésima unidade. Assim,

substituindo (3.12) e y_{ik} em (3.11), temos

$$\hat{\theta}_{ik} = \hat{b} \mathbf{1}_i' \left[I_i - \frac{\hat{b} \mathbf{1}_i \mathbf{1}_i'}{1 + \hat{b} n_{ik}} \right] y_{ik}$$

$$\begin{aligned}
&= \hat{b} \begin{bmatrix} \hat{b} \mathbf{1}' \mathbf{1} \mathbf{1}' \\ \sim \quad \sim \quad \sim \\ 1 + \hat{b} n_{ik} \end{bmatrix} \begin{matrix} \sim \\ y_{ik} \\ \sim \end{matrix} \\
&= \hat{b} \begin{bmatrix} 1 - \frac{\hat{b} n_{ik}}{1 + \hat{b} n_{ik}} \end{bmatrix} \begin{matrix} \sim \\ \mathbf{1}' y_{ik} \\ \sim \end{matrix} \\
&= \hat{b} \begin{bmatrix} 1 - \frac{\hat{b} n_{ik}}{1 + \hat{b} n_{ik}} \end{bmatrix} y_{ik} \\
&= \hat{b} \begin{bmatrix} \frac{1 + \hat{b} n_{ik} - \hat{b} n_{ik}}{1 + \hat{b} n_{ik}} \end{bmatrix} y_{ik} \\
&= \frac{\hat{b}}{1 + \hat{b} n_{ik}} y_{ik} \\
&= \frac{1}{1/\hat{b} + n_{ik}} y_{ik}
\end{aligned}$$

■

Observemos que $\hat{\theta}_{ik}$ é quase uma média dos resíduos da i -ésima unidade do grupo k .

Ou seja,

se $\hat{b} > 1 \Rightarrow \frac{1}{\hat{b}} < 1$ e $\hat{\theta}_{ik}$ é praticamente a média dos elementos de y_{ik} ;

se $\hat{b} < 1 \Rightarrow \frac{1}{\hat{b}} > 1$ e $\hat{\theta}_{ik}$ é menor que a média dos elementos de y_{ik} .

Uma outra observação sobre os efeitos aleatórios é que usamos a suposição de normalidade, logo, devemos verificá-la. Uma maneira de fazer esta verificação é desenhando o gráfico normal ponderado dos efeitos aleatórios padronizados preditos, *Dempster, A. P. & Ryan, Louise M. (1985)*.

3.2 Modelo de Laird - Ware de Simetria Composta com Uma Covariável

Consideremos o modelo de Laird-Ware de Simetria Composta e acrescentemos uma covariável no modelo :

$$\underset{\sim}{y}_{ik} = \underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{z}_{ik} \underset{\sim}{\gamma} + \underset{\sim}{1} \underset{\sim}{\theta}_{ik} + \underset{\sim}{e}_{ik}, \quad (3.13)$$

onde

$\underset{\sim}{z}_{ik}$ é o vetor da covariável correspondente à i -ésima unidade do grupo k e

$\underset{\sim}{\gamma}$ é o parâmetro associado à covariável.

Notemos que agora

$$\begin{aligned} E(\underset{\sim}{y}_{ik}) &= \underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{z}_{ik} \underset{\sim}{\gamma} \\ \Rightarrow \underset{\sim}{y}_{ik} &\sim N(\underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{z}_{ik} \underset{\sim}{\gamma}; \sigma^2 \underset{\sim}{V}_{ik}), \end{aligned} \quad (3.14)$$

onde a estrutura de variâncias e covariâncias permanece a mesma do modelo de

Laird-Ware de Simetria Composta.

Proposição 3.1

Os estimadores de máxima verossimilhança de β e γ , fixado b , no contexto modelo de Laird-Ware de Simetria Composta com uma covariável são respectivamente

$$\hat{\beta} = \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} y_{ik} \right) - \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} z_{ik} \right)$$

$$\left[\sum_{k=1}^p \sum_{i=1}^{m_k} z'_{ik} V_{ik}^{-1} (I_{ik} - A_{ik})^{-1} z_{ik} \right]^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} z'_{ik} V_{ik}^{-1} (I_{ik} - A_{ik})^{-1} y_{ik} \right)$$

e

$$\hat{\gamma} = \left[\sum_{k=1}^p \sum_{i=1}^{m_k} z'_{ik} V_{ik}^{-1} (I_{ik} - A_{ik})^{-1} z_{ik} \right]^{-1} \left(\sum_{k=1}^p \sum_{i=1}^{m_k} z'_{ik} V_{ik}^{-1} (I_{ik} - A_{ik})^{-1} y_{ik} \right),$$

assumindo a existência da inversa e definindo A_{ik} como a matriz de projeção não ortogonal sobre $\mathcal{L}(X_{ik})$ para a i -ésima unidade.

Prova :

Observemos que o modelo de Laird-Ware de Simetria Composta, fixado b , com uma covariável é equivalente ao modelo de análise de covariância generalizado. De modo que, de forma análoga ao modelo de Aitken, os estimadores de MQG do modelo de análise de covariância generalizado, Capítulo 2, coincidirão com os estimadores de máxima verossimilhança do modelo de Laird-Ware de Simetria Composta com uma covariável. Assim, o resultado segue.

Proposição 3.2

Os estimadores de máxima verossimilhança de σ^2 , fixado b , no contexto modelo de Laird-Ware de Simetria Composta com uma covariável é

$$\sigma^2 = \frac{1}{n} \sum_{k=1}^p \sum_{i=1}^{m_k} (\underset{\sim}{y}_{ik} - \underset{\sim}{X}_{ik} \underset{\sim}{\hat{\beta}} - \underset{\sim}{z}_{ik} \underset{\sim}{\hat{\gamma}})' \underset{\sim}{V}_{ik}^{-1} (\underset{\sim}{y}_{ik} - \underset{\sim}{X}_{ik} \underset{\sim}{\hat{\beta}} - \underset{\sim}{z}_{ik} \underset{\sim}{\hat{\gamma}}).$$

Prova :

De maneira análoga ao modelo sem covariável, no modelo com a covariável o estimador de MQG de σ^2 , Capítulo 2, coincide com o estimador de máxima verossimilhança, a menos de r ($[X | Z]$) unidades no denominador.

■

Observemos que no modelo exposto sem a covariável o resíduo da parte sistemática variava em torno da média dos grupos, isto é, variava em torno de uma constante, enquanto, que no modelo agora apresentado o resíduo da parte sistemática varia em torno de uma reta ajustada.

3.3 Ilustrações

Exemplo 1

Consideremos o exemplo 5 do Capítulo 2 e usando o modelo de Laird - Ware de Simetria Composta,

$$\underset{\sim}{y}_{ik} = \underset{\sim}{X}_{ik} \underset{\sim}{\beta} + \underset{\sim}{1} \underset{\sim}{\theta}_{ik} + \underset{\sim}{e}_{ik},$$

para a i -ésima unidade do grupo k ,

onde

$X_{i1} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \end{bmatrix}$ é a i-ésima matriz de delineamento do grupo controle;

$X_{i2} = \begin{bmatrix} \mathbf{1} & \mathbf{1} \end{bmatrix}$ é a i-ésima matriz de delineamento do grupo OA e

$\beta = \begin{bmatrix} \beta_1, \beta_2 \end{bmatrix}$ é o vetor de parâmetros fixos.

Agora, para verificarmos se há diferença entre os valores esperados dos grupos, ou seja para testar as hipóteses

$$H_0 : \beta_2 = 0$$

versus

$$H_0 : \beta_2 \neq 0, \quad (3.15)$$

nós obtemos a seguinte estatística do teste

$$F = \frac{(\Lambda \hat{\beta})' (\Lambda (\sum_{k=1}^2 \sum_{i=1}^{m_k} X_{ik}' V_i^{-1} X_{ik})^{-1} \Lambda)^{-1} (\Lambda \hat{\beta})}{\hat{\sigma}^2} = 0,08920,$$

onde $\Lambda = [0 \ 1]$,

Assim, ao nível de 5%, $F_{(0,05;1;39)} = 4,09$, logo, não temos evidências de que os dois grupos tenham médias de irrigação na cabeça do fêmur diferente.

Agora, quando exibimos o modelo de Laird-Ware de Simetria Composta, supomos que os efeitos aleatórios vinham de uma distribuição normal. Assim, vamos verificar esta suposição observando como se comportam os valores dos efeitos aleatórios padronizados preditos no gráfico 3.1.

Notemos que no grupo controle os valores preditos dos efeitos aleatórios estão concentrados em valores negativos, não havendo um espalhamento homogêneo como é de se esperar, caso a suposição de normalidade com mesma variância seja atendida. Além do mais, os pontos não estão muito centralizados na faixa demarcada. Logo, há evidência de que o uso deste modelo não é adequado.

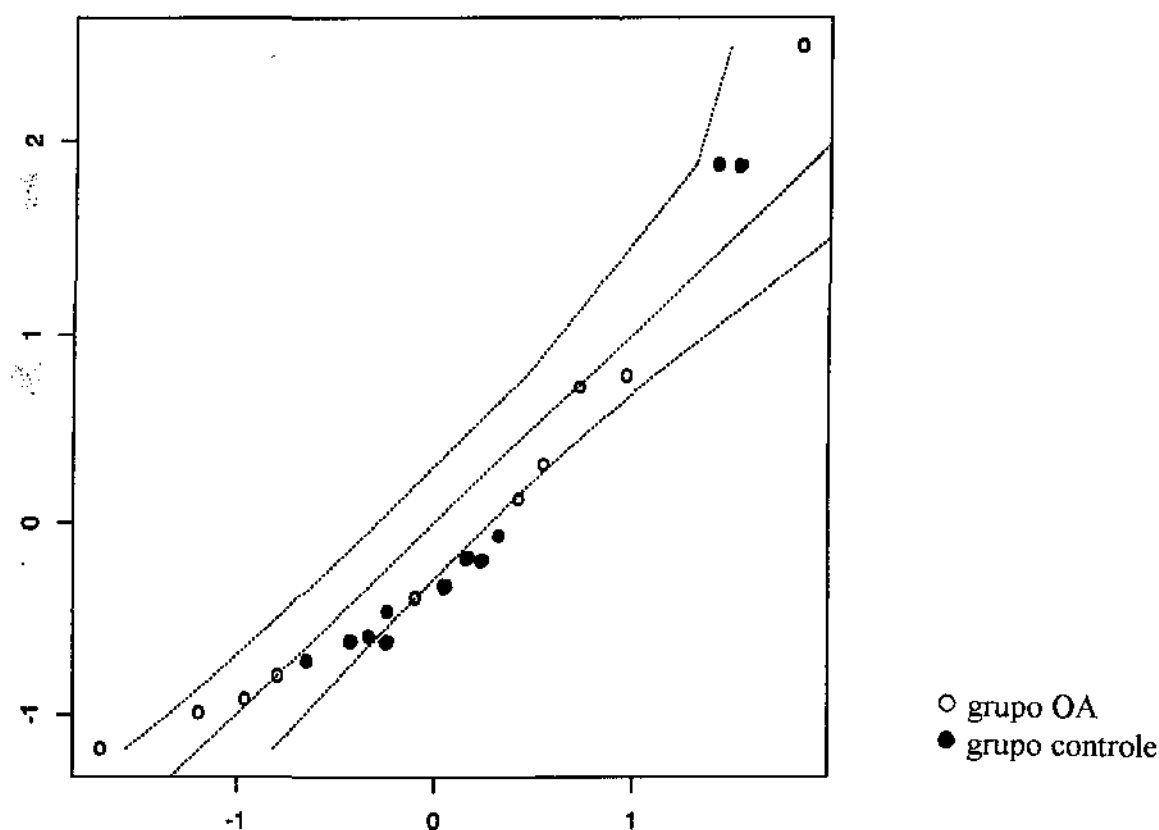


Gráfico 3.1: Gráfico de probabilidade normal ponderado para os efeitos aleatórios padronizados preditos, sob o modelo de Laird - Ware de Simetria Composta sem covariável. □

Exemplo 2

Continuando o exemplo 1, ou seja, considerando agora a covariável GT no modelo de Laird-Ware de Simetria Composta, nos obtemos a seguinte estatística do teste para as hipóteses (3.15),

$$F = \frac{(\Lambda \hat{\beta})' (\Lambda (SQ) \Lambda')^{-1} (\Lambda \hat{\beta})}{\hat{\sigma}^2} = 0,0176,$$

$$\text{onde } SQ = \left(\sum_{k=1}^2 \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1} - \left(\sum_{k=1}^2 \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1} \sum_{k=1}^2 \sum_{i=1}^{m_k} (X'_{ik} V_{ik}^{-1} \tilde{z}_{ik})$$

$$\left[\sum_{k=1}^2 \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} (I - A_{ik}) \tilde{z}_{ik} \right]^{-1} \sum_{k=1}^2 \sum_{i=1}^{m_k} (X'_{ik} V_{ik}^{-1} \tilde{z}_{ik})' \left(\sum_{k=1}^2 \sum_{i=1}^{m_k} X'_{ik} V_{ik}^{-1} X_{ik} \right)^{-1};$$

$\tilde{z}_{ik}$ é o vetor $n_i \times 1$ da covariável do i -ésimo indivíduo do grupo k , $k = 1, 2$.

Neste exemplo o vetor correspondente ao indivíduo i possui todos os seus elementos iguais, $i = 1, \dots, m_k$;

V_{ik} é uma matriz do tipo (3.4) ;

X_{ik} é a i -ésima matriz de delineamento do grupo k e

A_{ik} é a matriz de projeção não ortogonal sobre $\mathcal{L}(X_{ik})$ para o i -ésimo indivíduo do grupo k .

Assim, ao nível de 5%, $F_{(0,05; 1; 38)} = 4,10$, logo, novamente temos evidências de que os dois grupos tenham a mesma média de irrigação na cabeça do fêmur. Porém, com o uso da covariável os efeitos aleatórios dos grupos ficaram mais dispersos, provocando uma maior mistura entre eles, além de quase todos caírem dentro da faixa de normalidade, como é observado no gráfico abaixo. De modo que o uso da covariável nos parece ser indispensável na análise deste problema.

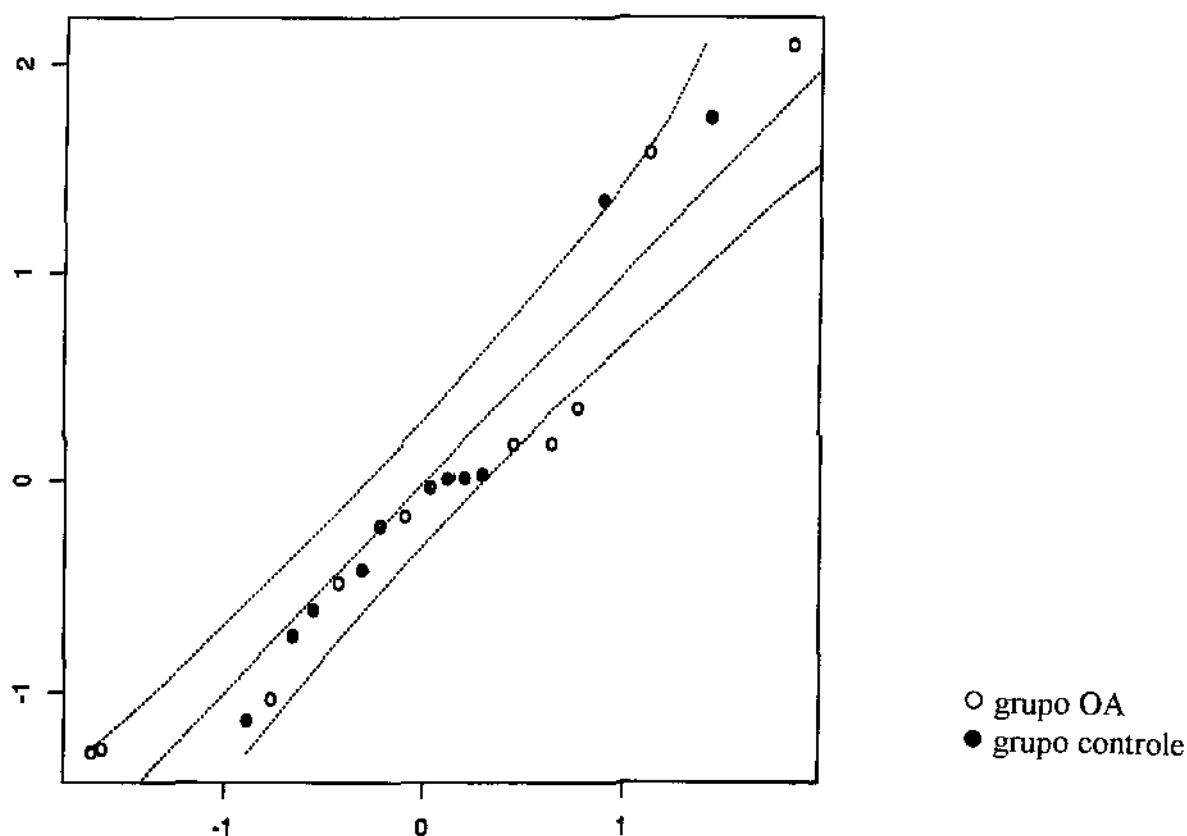


Gráfico 3.2: Gráfico de probabilidade normal ponderado para os efeitos aleatórios padronizados preditos, sob o modelo de Laird - Ware de Simetria Composta com covariável.

□

Nos exemplos anteriores verificamos que não há evidência de diferença entre as médias de irrigação na cabeça do fêmur dos dois grupos. Logo, abaixo iremos apresentar as estimativas dos parâmetros do modelo de Laird - Ware de Simetria composta com covariável considerando as populações de pessoas acometidas por osteoartrose e saudáveis como uma única população, para $\hat{b} = 0,409$.

$$\cdot \hat{\sigma}^2 = 34,37 \quad \cdot \hat{\beta}_1 = 6,335 \quad \cdot \hat{\gamma} = 0,1517$$

Observemos o gráfico 3.3 e notemos que os valores preditos dos efeitos aleatórios padronizados dos grupos continuaram bem distribuídos e quase todos caíram dentro da faixa de normalidade, apresentando indícios de que o uso deste novo modelo é adequado.

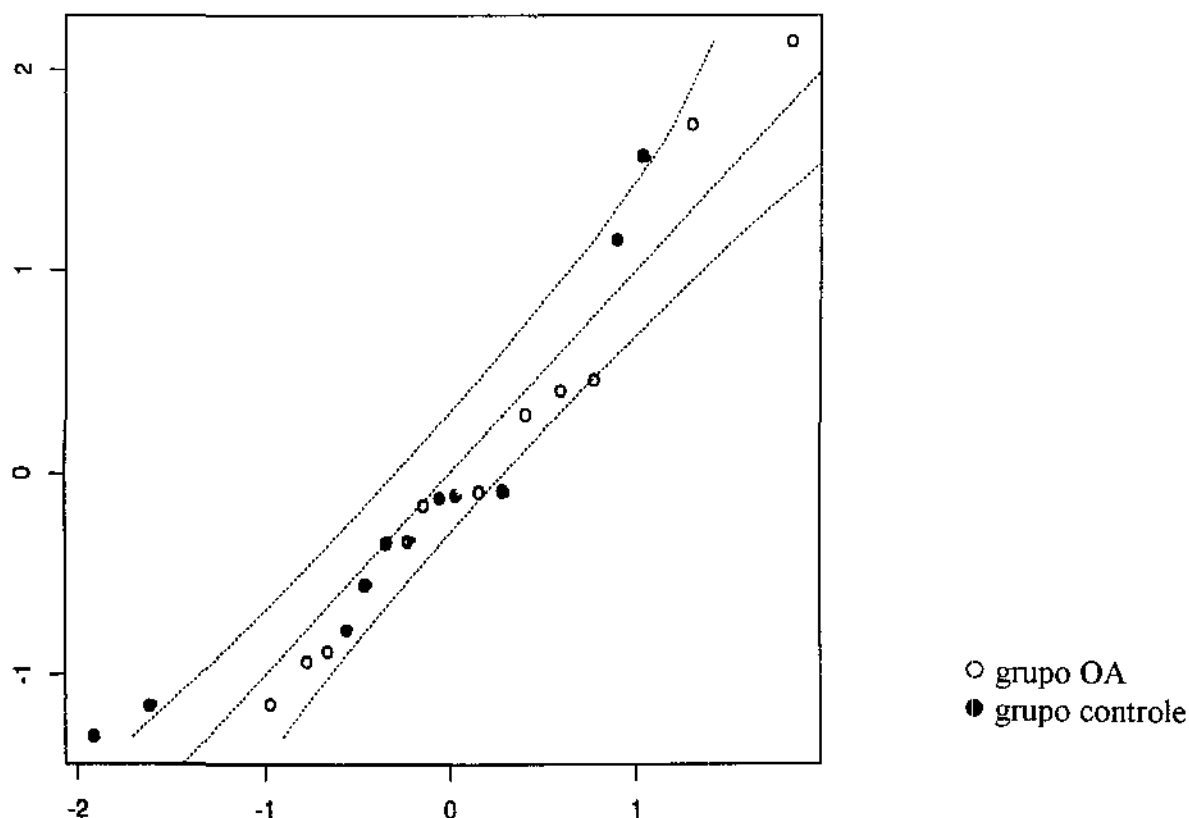


Gráfico 3.3: Gráfico de probabilidade normal dos efeitos aleatórios padronizados preditos sob modelo de Laird - Ware de Simetria Composta com covariável considerando uma única população.

APÊNDICE A

No nosso estudo utilizamos várias vezes conceitos e resultados de álgebra linear e de teoria distribucional. Apesar de supormos o conhecimento destas teorias, pelo leitor, para o entendimento da nossa dissertação, iremos apresentar alguns resultados, pois ainda que apareçam na literatura, estão bastante dispersos.

Definição A1

Seja A uma matriz, $n \times m$. O *espaço coluna* de A , $\mathcal{L}(A)$, é o conjunto de todos vetores coluna, $n \times 1$, que sejam combinações lineares das colunas de A .

Assim, se $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_m]$, um vetor $\mathbf{a}$ pertence ao espaço coluna de A se existem $\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_m \in \mathbb{R}$, tal que, $\mathbf{a} = \tilde{c}_1 \mathbf{a}_1 + \dots + \tilde{c}_m \mathbf{a}_m$, isto é, se existe um vetor $\tilde{\mathbf{c}} \in \mathbb{R}^m$, tal que, $\tilde{\mathbf{a}} = A \tilde{\mathbf{c}}$.

Analogamente, o *espaço linha*, $\mathcal{L}(A)$, de uma matriz A , $n \times m$, é definido como o conjunto de todos vetores linha, $1 \times m$, que sejam combinações lineares das linhas de A .

Proposição A1

Seja A uma matriz $n \times m$ e B uma matriz $n \times p$. O espaço coluna de A está contido no espaço coluna de B , se, e somente se, existe uma matriz F , $p \times m$, tal que,

$$A = BF.$$

Prova :

1) $\Leftarrow$: Suponha que exista uma matriz F , $p \times m$, tal que $A = BF$. Seja $\tilde{\mathbf{a}}$, um vetor $n \times 1$,

tal que $\underline{a} \in \mathcal{L}(A)$, então, pela definição A1, existe um vetor $\underline{c}$ tal que

$$\underline{a} = A \underline{c} = BF \underline{c} = B (F \underline{c}),$$

implicando em

$$\underline{a} \in \mathcal{L}(B).$$

2) $\Rightarrow$: Suponha $\mathcal{L}(A) \subset \mathcal{L}(B)$. Então, cada uma das colunas de A, $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_m$ de A, pertence ao $\mathcal{L}(A)$, e, portanto, pertence ao $\mathcal{L}(B)$. Logo, existem vetores $\underline{f}_i$, tais que $\underline{a}_i = B \underline{f}_i$, $i = 1, \dots, m$. Logo,

$$\underline{A} = [\underline{a}_1 \ \underline{a}_2 \ \dots \ \underline{a}_m] = [B \underline{f}_1, \dots, B \underline{f}_m] = B[\underline{f}_1, \dots, \underline{f}_m] = BF,$$

para $F = [\underline{f}_1, \dots, \underline{f}_m]$.

■

Definição A2

Seja P uma matriz $n \times n$. A matriz P é denominada matriz de projeção ortogonal sobre o espaço coluna de uma matriz X, se e somente se

- a) $\underline{a} \in \mathcal{L}(X)$ implica em $P \underline{a} = \underline{a}$. (Projeção)
- b) $\underline{c} \perp \mathcal{L}(X)$ implica em $P \underline{c} = \underline{0}$. (Ortogonalidade)

Proposição A2

Seja P uma matriz de projeção sobre o espaço coluna de uma matriz X, então,

$$PX = X.$$

Prova:

Pela definição A2, $\tilde{x} \in \mathcal{L}(X)$ implica em $P\tilde{x} = \tilde{x}$. Logo, para

$X = [\tilde{x}_1 \dots \tilde{x}_p]$, onde $\tilde{x}_i \in \mathcal{L}(X)$, para todo $i = 1, \dots, p$,

teremos

$$PX = X.$$

■

Proposição A3

Seja P uma matriz $n \times n$. A matriz P é denominada matriz de projeção ortogonal sobre o espaço coluna de uma matriz X , se e somente se

- a) P é idempotente . (Projeção)
- b) P é simétrica . (Ortogonalidade)

Prova :

Diretamente da definição A2.

■

Definição A3

Uma inversa generalizada de uma matriz A é qualquer matriz G , tal que

$$AGA = A.$$

No presente trabalho usamos a notação A^+ para designar uma inversa generalizada de uma matriz A .

Teorema A1

A matriz $X(X'X)^-X'$ é matriz de projeção ortogonal sobre o espaço coluna da matriz X .

Prova :

Para demonstrar este teorema, basta conferir a validade das condições (a) e (b), da definição A2. Pela definição A1, sabemos que $\underline{a} \in \mathcal{C}(X)$ implica na existência de um vetor $\underline{b}$, tal que $\underline{a} = X\underline{b}$, então,

$$\begin{aligned} X'X(X'X)^-X'\underline{a} &= X'X(X'X)^-X'X\underline{b} = X'X\underline{b} \\ \Rightarrow X(X'X)^-X'\underline{a} &= X\underline{b} = \underline{a}. \end{aligned}$$

Temos que $\underline{c} \perp \mathcal{C}(X)$ implica em $X'\underline{c} = \underline{0}$, logo $X(X'X)^-X'\underline{c} = \underline{0}$.

■

Proposição A4

A matriz $X(X'X)^-X'$ é invariante quanto à escolha da inversa generalizada de $X'X$.

Prova :

Sejam G_1 e G_2 inversas generalizadas de $X'X$, logo,

$$(X'X)G_1(X'X) = X'X$$

e

$$(X'X)G_2(X'X) = X'X.$$

Pelo teorema A1 as matrizes $P_1 = XG_1X'$ e $P_2 = XG_2X'$ são matrizes de projeção sobre

o espaço coluna de X . Então,

$$\begin{aligned} X'XG_1X' &= (X'X)G_2(X'X)G_1X' \\ &= X'XG_2X'(XG_1X') = X'XG_2X'P_1 = X'XG_2X', \end{aligned}$$

$$\Rightarrow XG_1X' = XG_2X'.$$

■

Proposição A5

Seja X uma matriz $n \times p$ e P uma matriz de projeção sobre $\mathcal{C}(X)$. Então,

$$r(P) = r(X),$$

onde $r(A)$ denota o posto da matriz A .

Prova :

O posto de X é o número de colunas linearmente independente de X , logo,

$$r(X) \geq r(X(X'X)^+X') = r(P) \geq r(PX) = r(X)$$

que implica em

$$r(P) = r(X).$$

■

Proposição A6

Seja X uma matriz $n \times p$ e P uma matriz de projeção sobre $\mathcal{C}(X)$. Então,

$$r(I - P) = n - r(X),$$

onde I denota a matriz identidade de ordem n .

Prova :

$$(I - P)(I - P) = I - P - P + PP = I - P,$$

logo, $(I - P)$ é idempotente. Então, usando a igualdade de posto e traço (Tr) para matrizes idempotentes, temos :

$$r(I - P) = \text{Tr}(I - P) = \text{Tr}(I) - \text{Tr}(P) = n - r(P) = n - r(X),$$

pela proposição A5.

■

Proposição A.7

A matriz de projeção ortogonal sobre $\mathcal{L}(X)$, P_x , é igual à matriz de projeção ortogonal sobre $\mathcal{L}(W)$, P_w , se e somente se,

$$\mathcal{L}(X) = \mathcal{L}(W).$$

Prova :

⇒ Seja $P_x = P_w$. Então,

$$X = P_x X = P_w X = W(W'W)^- W'X = WG,$$

onde $G = (W'W)^- W'X$. Analogamente, $W = XF$, onde $F = (X'X)^- X'W$. Logo, pela proposição A1,

$$\mathcal{L}(X) \subset \mathcal{L}(W) \text{ e } \mathcal{L}(W) \subset \mathcal{L}(X),$$

o que implica em

$$\mathcal{L}(X) = \mathcal{L}(W).$$

⇐ Seja $\mathcal{L}(X) = \mathcal{L}(W)$. Então existem F e G , matrizes tais que $X = WG$ e $W = XF$.

Logo,

$$\begin{aligned} P_x P_w &= P_x W (W'W)^{-1} W' \\ &= P_x X F (W'W)^{-1} W' \\ &= X F (W'W)^{-1} W' \\ &= W (W'W)^{-1} W' = P_w . \end{aligned}$$

Analogamente $P_w P_x = P_x$. Então,

$$P_w = P_w' = (P_x P_w)' = P_w' P_x' = P_w P_x = P_x .$$

■

Proposição A8

O espaço coluna da matriz $[X : Z]$ é igual ao espaço coluna da matriz $[X : (I - P_x)Z]$, onde P_x é a matriz de projeção ortogonal sobre $\mathcal{L}(X)$.

Prova :

Para provarmos esta proposição basta mostrarmos que se um vetor pertence ao $\mathcal{L}(X)$, ou $\mathcal{L}(Z)$, ou ao $\mathcal{L}((I - P_x)Z)$, então ele pertence ao

$\mathcal{L}([X : Z])$ e ao $\mathcal{L}([X : (I - P_x)Z])$.

$$a) \underline{x} \in \mathcal{L}(X) \Rightarrow \underline{x} \in \mathcal{L}([X : Z]) \text{ e } \underline{x} \in \mathcal{L}([X : (I - P_x)Z]) .$$

$$b) \underline{x} \in \mathcal{L}(Z) \Rightarrow \underline{x} \in \mathcal{L}([X : Z]) ,$$

mas ,

$$\underline{x} \in \mathcal{L}(Z) \Leftrightarrow \underline{x} = Z \underline{a} ,$$

e fazendo $Z = P_x Z + (I - P_x)Z$, temos

$$\underset{\sim}{x} = P_x Z \underset{\sim}{a} + (I - P_x)Z \underset{\sim}{a},$$

agora, $P_x Z = X(X'X)^{-1}X'Z$ é uma combinação linear das colunas de X , então,

$$\underset{\sim}{x} \in \mathcal{L}([X : (I - P_x)Z]).$$

$$c) \quad \underset{\sim}{x} \in \mathcal{L}((I - P_x)Z) \Rightarrow \underset{\sim}{x} \in \mathcal{L}([X : (I - P_x)Z]),$$

mas

$$\begin{aligned} \underset{\sim}{x} \in \mathcal{L}((I - P_x)Z) &\Rightarrow \underset{\sim}{x} = (I - P_x)Z \underset{\sim}{a} = Z \underset{\sim}{a} - P_x Z \underset{\sim}{a} \\ &\Rightarrow \underset{\sim}{x} \in \mathcal{L}([Z : P_x Z]), \\ &\Rightarrow \underset{\sim}{x} \in \mathcal{L}([X : Z]). \end{aligned}$$

■

Proposição A9

Seja P uma matriz de projeção sobre $\mathcal{L}([X_1 : X_2])$, onde $X_1 \perp X_2$,

então P é a soma das matrizes de projeções em $\mathcal{L}(X_1)$ e $\mathcal{L}(X_2)$,

$$P = P_{X_1} + P_{X_2}.$$

Prova :

Seja $X_1 \perp X_2$ então $X_1'X_2 = 0$, $X_2'X_1 = 0$ e

$$P = [X_1 : X_2]([X_1 : X_2]'[X_1 : X_2])^{-1}[X_1 : X_2]'$$

$$= [X_1 : X_2] \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1} [X_1 : X_2]'$$

$$= [X_1 : X_2] \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix} [X_1 : X_2]'$$

$$= X_1(X_1'X_1)^{-1}X_1' + X_2(X_2'X_2)^{-1}X_2'$$

$$= P_{X_1} + P_{X_2}.$$

■

Proposição A10

O espaço coluna da matriz $[X : Z]$ é igual ao espaço coluna da matriz $[X : (I - A)Z]$, onde $A = X(X'V^{-1}X)^{-1}X'V^{-1}$ para V não singular.

Prova :

Como $AZ = X(X'V^{-1}X)^{-1}X'V^{-1}Z$ é uma combinação linear das colunas de X , temos

$$\mathcal{C}(AZ) \subset \mathcal{C}(X).$$

Logo, a prova segue similarmente à prova da proposição A8.

■

Proposição A11

A matriz de $A = X(X'V^{-1}X)^{-1}X'V^{-1}$ é uma matriz de projeção não ortogonal sobre X , onde V é uma matriz não singular.

Prova :

Seja $V^{-1} = Q'Q$.

Seja $A = X(X'V^{-1}X)^{-1}X'V^{-1}$. Então $\mathcal{L}(A) \subset \mathcal{L}(X)$, pela proposição A1. Pela proposição A2, $(QX)[(QX)'(QX)]^{-1}(QX)'(QX) = (QX)$, mas

$$(QX)[(QX)'(QX)]^{-1}(QX)'(QX) = QX(X'V^{-1}X)^{-1}X'V^{-1}X = QAX$$

$$\Rightarrow QAX = QX,$$

$$\Rightarrow AX = X$$

$$\Rightarrow \mathcal{L}(A) \supset \mathcal{L}(X),$$

de modo que $\mathcal{L}(A) = \mathcal{L}(X)$.

Agora

$$AA = AX(X'V^{-1}X)^{-1}X'V^{-1} = X(X'V^{-1}X)^{-1}X'V^{-1} = A$$

e

$A = X(X'V^{-1}X)^{-1}X'V^{-1} \neq A' = V^{-1}X(X'V^{-1}X)^{-1}X'$. Logo, pela proposição A3, A é uma matriz de projeção não ortogonal sobre $\mathcal{L}(X)$.

Teorema A2

Seja y um vetor aleatório com $E(\underset{\sim}{y}) = \underset{\sim}{\mu}$ e $\text{Var}(\underset{\sim}{y}) = V$, então,

$$E(\underset{\sim}{y}'A\underset{\sim}{y}) = \text{Tr}(AV) + \underset{\sim}{\mu}'A\underset{\sim}{\mu},$$

onde $\text{Tr}(\cdot)$ é o traço de $(\cdot)$.

Prova :

$$(\underline{y} - \underline{\mu})' A (\underline{y} - \underline{\mu}) = \underline{y}' A \underline{y} - \underline{\mu}' A \underline{y} - \underline{y}' A \underline{\mu} + \underline{\mu}' A \underline{\mu}.$$

Então,

$$\begin{aligned} E[(\underline{y} - \underline{\mu})' A (\underline{y} - \underline{\mu})] &= E[\underline{y}' A \underline{y}] - \underline{\mu}' A \underline{\mu} - \underline{\mu}' A \underline{\mu} + \underline{\mu}' A \underline{\mu} \\ &= E[\underline{y}' A \underline{y}] - \underline{\mu}' A \underline{\mu}. \end{aligned}$$

Assim, $E[\underline{y}' A \underline{y}] = E[(\underline{y} - \underline{\mu})' A (\underline{y} - \underline{\mu})] + \underline{\mu}' A \underline{\mu}$. Mas,

$$\begin{aligned} E[(\underline{y} - \underline{\mu})' A (\underline{y} - \underline{\mu})] &= E[\text{Tr}((\underline{y} - \underline{\mu})' A (\underline{y} - \underline{\mu}))] \\ &= E[\text{Tr}(A (\underline{y} - \underline{\mu})(\underline{y} - \underline{\mu})')] \\ &= \text{Tr}(E[A (\underline{y} - \underline{\mu})(\underline{y} - \underline{\mu})']) \\ &= \text{Tr}(A E[(\underline{y} - \underline{\mu})(\underline{y} - \underline{\mu})']) = \text{Tr}(A V). \end{aligned}$$

Portanto ,

$$E(\underline{y}' A \underline{y}) = \text{Tr}(A V) + \underline{\mu}' A \underline{\mu}$$

■

APÊNDICE B

B1 Programa para Análises de Variância e Covariância

Este programa foi utilizado nos exemplos 1, 2, 3 e 4, Capítulo 2. Para facilitar o entendimento do leitor apresentaremos o programa com os dados do exercício 1.

```
DATA DROGA ;
```

```
INPUT DROG $ z y @@;
```

```
CARDS ;
```

```
A 11 6 A 8 0 A 5 2 A 14 8 A 19 11
```

```
A 6 4 A 10 13 A 6 1 A 11 8 A 3 0
```

```
B 6 0 B 6 2 B 7 3 B 8 1 B 18 18
```

```
B 8 4 B 19 14 B 8 9 B 5 1 B 15 9
```

```
C 16 13 C 13 10 C 11 18 C 9 5 C 21 23
```

```
C 16 12 C 12 5 C 12 16 C 7 1 C 12 20
```

```
;
```

```
PROC ANOVA;
```

```
CLASS DRUG;
```

```
MODEL Y = DRUG;
```

```
RUN ;
```

```
PROC GLM;
```

```
CLASS DRUG;
```

```
MODEL Y=DRUG Z / SOLUTION;
```

```
LSMEANS DRUG / STDERR PDIFF ;
```

```
RUN;
```

B2 Programa para Análises de Variâncias e Covariâncias quando a matriz de variâncias e covariâncias do erro é $\sigma^2 V$.

O presente programa foi desenvolvido no procedimento IML do SAS, para resolução do exemplo 5, Capítulo 2.

```
PROC IML ;
```

```
    /* Geração da matriz de delineamento completa X, de dimensão  
(21×3) */
```

```
    X1 = shape({1 1 0}, 11, 3) ;
```

```
    X2 = shape({1 0 0}, 10, 3) ;
```

```
    X = X1//X2 ;
```

```
    /* Geração da matriz de delineamento reduzida X0, de dimensão  
(21×3) */
```

```
    X0 = shape({1 1 }, 21 , 2 ) ;
```

```
    /* Especificação do vetor da variável dependente, Y, de dimensão  
(21×1) */
```

```
    Y1 = { 13 12 9 9 10 8 9 11 28 12 28  
          21.5 12 15 10.6 33 21.3 19  
          10 11.5 18 } ;
```

```
    Y = Y1` ;
```

```
/* Especificação do vetor da covariável Z, de dimensão  
(21×1) */
```

```
      Z = { 49, 73, 58, 33, 84, 15, 43, 38, 90, 43, 71,  
            46, 50, 91, 16, 97, 102, 73, 67, 70, 56 } ;
```

```
/* Especificação dos elementos da diagonal da matriz inversa de  
variâncias e covariâncias */
```

```
      W = { 1 1 1 1 1 1 1 1 1 1 1  
            4 2 6 5 2 3 2 1 4 1 } ;
```

```
/* Cálculo da inversa da matriz de variâncias e covariâncias, V */
```

```
      V = SHAPE( 0, 21 , 21 ) ;  
      V = DIAG(W) ;
```

```
/* Especificação dos elementos da diagonal da raiz da matriz  
inversa de variâncias e covariâncias */
```

```
      Q1 = { 1 1 1 1 1 1 1 1 1 1 1  
            2 1,4142 2,4494 2,2361 1,4142  
            1,7320 1,4142 1 2 1 } ;
```

```
/* Cálculo da inversa da raiz da matriz de variâncias e  
covariâncias, V */
```

```
      Q = SHAPE( 0, 21, 21 ) ;  
      Q = DIAG(Q1) ;
```

```
/* Especificação da matriz identidade I */
```

```
      I = I(21) ;
```

```
/* Cálculos para achar a estatística do teste F */
```

```
/* Cálculos do modelo completo */
```

```
B = SHAPE( 0, 3, 3 ) ;
```

```
B = X`*V*X ;
```

```
C = SHAPE( 0, 3, 3 ) ;
```

```
C = GINV(B) ;
```

```
D = SHAPE( 0, 21, 21 ) ;
```

```
D = X*C*X`*V; /* matriz de projeção sobre  $\mathcal{E}(X)$  */
```

```
E = SHAPE( 0, 21, 21 ) ;
```

```
E = I - D ;
```

```
G = SHAPE( 0, 1, 1 ) ;
```

```
G = Z`*E`*V*E*Z ;
```

```
H = SHAPE( 0, 1, 1 ) ;
```

```
H = GINV(G) ;
```

```
L = SHAPE( 0, 21, 21 ) ;
```

```
L = E*Z*H*Z`*E`*V ;
```

```
A = SHAPE( 0, 21, 21 ) ;
```

```
A = D + L ; /* matriz de projeção sobre  $\mathcal{E}(X, Z)$  */
```

```
/* Cálculos do modelo reduzido */
```

```
Bo = SHAPE( 0, 1, 1 ) ;
```

```
Bo = Xo`*V*Xo ;
```

```
Co = SHAPE( 0, 1, 1 ) ;
```

```

Co = GINV(Bo) ;

Do = SHAPE( 0, 21, 21 ) ;

Do = Xo*Co*Xo`*V; /* matriz de projecao sobre  $\mathcal{E}(Xo)$  */
Eo = SHAPE( 0, 21 ,21 ) ;

Eo = I - Do ;

Go = SHAPE( 0, 1 , 1 ) ;

Go = Z`*Eo`*V*Eo*Z ;

Ho = SHAPE( 0, 1 , 1 ) ;

Ho = GINV(Go) ;

Lo = SHAPE( 0, 21 ,21 ) ;

Lo = Eo*Z*Ho*Z`*Eo`*V ;

Ao = SHAPE( 0, 21 ,21 ) ;

Ao = Do + Lo ; /* matriz de projecao sobre  $\mathcal{E}(Xo , Z)$  */

/* Cálculo do numerador da ANOVA SQHa =Y'(D - Do)V(D - Do)Y/1 e
da ANCOVA    AQHc = Y'(A - Ao)V(A - Ao)Y/1 */

M = SHAPE( 0, 21, 21 ) ;

M = A - Ao ;

SQHc = SHAPE( 0, 1, 1 ) ;

SQHc = Y`*M`*V*M*Y ;

U = SHAPE( 0, 21, 21 ) ;

U = D - Do ;

SQHa = SHAPE( 0, 1, 1 ) ;

SQHa = Y`*U`*V*U*Y ;

```

```
/* Cálculo do MSEa( ANOVA) e MSEC ( ANCOVA) */

O = SHAPE( 0, 21, 21 ) ;

O = I - A ;

SSEc = SHAPE( 0, 1, 1 ) ;

SSEc = Y`*O`*V*O*Y ;

SSEa = SHAPE( 0 ,1 ,1 ) ;

SSEa = Y`*E`*V*E*Y ;

MSEa = SHAPE( 0, 1 ,1 ) ;

MSEa = SSEa / 19 ;

MSEC = SHAPE( 0, 1, 1 ) ;

MSEC = SSEc / 18 ;

Fa = SQHa / MSEa ;

Fc = SQHa / MSEC ;

/* Fa =Fo (1 - α, 1 , 19) */

/* Fc =Fo (1 - α, 1 , 18) */

/* Cálculo da soma de quadrado total ajustada pela média, SQTa .*/

J = SHAPE( {1}, 21,1) ;

SQT = Y`*V*Y ;

AJ = Y`*Q1`*J ;

AJUST = AJ*AJ/21 ;

SQTa = SQT - AJUST ;
```



```
PRINT  Y Z, SSEa  MSEa , SQHa  Fa,
      SSEc  MSEc , SQHc  Fc ,
      SQT AJUST SQTa ;
```

B3 Programa para o Modelo de Laird - Ware de Simetria Composta com ou sem Covariável

Este programa foi desenvolvido para resolução dos exemplos 1 e 2 , Capítulo 3.

```
Proc Iml ;
```

```
Start DADOS ;
```

```
/* Especificação do vetor de observações */
```

```
Y = { 13, 12, 9, 9, 10, 8, 9, 11, 28, 12, 28,
      12, 17, 31, 26,
      15, 9,
      24, 9, 8, 10, 26, 13,
      13, 19, 6, 3, 12,
      31, 35,
      19, 20, 25,
      19, 19,
      10,
      9, 13, 14, 10,
      18
      } ;
```

```
/* Especificação do vetor de covariáveis */
```

```
Z = { 49, 73, 58, 33, 84, 15, 43, 38, 90, 43, 71,  
      46, 46, 46, 46,  
      50, 50,  
      91, 91, 91, 91, 91, 91,  
      16, 16, 16, 16, 16,  
      97, 97,  
      102, 102, 102,  
      73, 73,  
      67,  
      70, 70, 70, 70,  
      56 } ;
```

```
/* Geração da matriz de delineamento */
```

```
X1 = shape( {1 0}, 11, 2) ;
```

```
X2 = shape( {1 1}, 30, 2) ;
```

```
X = X1//X2 ;
```

```
Finish ;
```

```
Start MATRIZES ;
```

y_i */
~

/* Construção das matrizes de variâncias e covariâncias de

V11 = (b+1)*I(11) ;

W12 = shape(0, 4,4) ;

Do i = 1 to 4 ;

Do j = 1 to 4 ;

W12[i,j] = b ; end ; end;

W13 = shape(0, 2,2) ;

Do i = 1 to 2 ;

Do j = 1 to 2 ;

W13[i,j] = b ; end ; end;

W14 = shape(0, 6, 6) ;

Do i = 1 to 6 ;

Do j = 1 to 6 ;

W14[i,j] = b ; end ; end;

W15 = shape(0, 5,5) ;

Do i = 1 to 5 ;

Do j = 1 to 5 ;

W15[i,j] = b ; end ; end;

```
W16 = shape(0, 2,2) ;  
    Do i = 1 to 2 ;  
        Do j = 1 to 2 ;  
W16[i,j] = b ; end ; end;  
  
W17 = shape(0, 3,3) ;  
    Do i = 1 to 3 ;  
        Do j = 1 to 3 ;  
        W17[i,j] = b ; end ; end;  
  
W18 = shape(0, 2,2) ;  
    Do i = 1 to 2 ;  
        Do j = 1 to 2 ;  
        W18[i,j] = b ; end ; end;  
  
V19 = b + 1 ;  
  
W20 = shape(0, 4,4) ;  
    Do i = 1 to 4 ;  
        Do j = 1 to 4 ;  
        W20[i,j] = b ; end ; end;  
  
V21 = b + 1 ;  
V12 = W12 + I(4) ;  
V13 = W13 + I(2) ;  
V14 = W14 + I(6) ;
```

```
V15 = W15 + I(5) ;
```

```
V16 = W16 + I(2) ;
```

```
V17 = W17 + I(3) ;
```

```
V18 = W18 + I(2) ;
```

```
V20 = W20 + I(4) ;
```

```
Finish ;
```

```
Start EXECUT ;
```

```
/*Construção da inversa da matriz de variâncias e  
covariâncias de  $\tilde{\mathbf{y}}$  */
```

```
W = block(V11, V12, V13, V14, V15, V16, V17, V18, V19,  
          V20, V21) ;
```

```
V = inv(W) ;
```

```
D = det(W) ;
```

```
/* Cálculo das estimativas de  $\gamma$ ,  $\beta$ ,  $\beta_0$  ( com a covariável)  $\sigma^2$  e  
 $\sigma^2_0$  ( com a covariável) */
```

```
A = X`*V*X ;
```

```
Ao = inv(A) ;
```

```
C = X`*V*Y ;
```

```
Beta = Ao*C ;
```

```
H = X`*V*Z ;
```

```
P = X*Ao*X`*V ;
```

```
M = Z`*V*(I(41) - P)*Z ;
```

```
M1 = Z`*V*(I(41) - P)*Y ;
```

```

Mo =inv(M) ;
Gama = Mo*M1 ;
No = Ao*H*Mo*M1 ;
Betac = beta - No ;
E = Y - X*Beta ;
Sigma = (E`*V*E)/N ;
zGama = Z*Gama ;
Ec = Y - X*Betac ;
Ecz = Ec - zGama ;
Sigmac = (Ecz`*V*Ecz)/N ;

```

```
/* Cálculo de  $\ell$  , sob modelo com e sem covariável*/
```

```

Lat = N*log(2*3.1415*Sigma) + log(D) + N ;
Latc = N*log(2*3.1415*Sigmac) + log(D) + N ;

T = {0 1} ;
Covc = Ao - Ao*H*Mo*H`*Ao` ;

```

```
/* Cálculo da estatística do teste F com e sem covariável*/
```

```

F = ( (T*Beta)`*INV(T*Ao*T`)*(T*Beta) ) / Sigma ;
Fc = ( (T*Betac)`*INV(T*Covc*T`)*(T*Betac) ) / Sigmac ;

```

```
Print b Beta Sigma Lat F Fob ,
```

```
      b Betac Gama Sigmac Latc Fc Fobc ;
```

```
Finish ;
```

Do ;

W = SHAPE (0, 41, 41) ;

A = SHAPE (0, 2, 2) ;

C = SHAPE (0, 2, 1) ;

Beta = SHAPE (0, 2, 1) ;

Betac = SHAPE (0, 2, 1) ;

H = SHAPE (0, 2, 1) ;

P = SHAPE (0, 41, 41) ;

No = SHAPE (0 , 2, 1) ;

E = SHAPE (0, 41, 1) ;

Ec = SHAPE (0, 41, 1) ;

Ecz = SHAPE (0, 41, 1) ;

Covc = Shape(0, 2, 2) ;

b = 0.1 ;

Run DADOS ;

Run MATRIZES ;

Lant = 10000 ;

Lantc = 10000 ;

Run EXECUT ;

Do I = 1 to 100 WHILE(Lat<Lant) ;

Lant = Lat ;

b = b + 0.05 ;

Run MATRIZES ;

```
        Run EXECUT ;

End;

Do j = 1 to 100 WHILE(Latc<Lantc) ;
    Lantc = Latc ;
    b = b + 0.05 ;
    Run MATRIZES ;
    Run EXECUT ;

End;

End ;
```


REFERÊNCIAS BIBLIOGRÁFICAS

1. *Bhattacharyya, G. K. & Johnson, R. A. (1977). Statistical Concepts and Methods, John Wiley, New York.*
2. *Christensen, R. A. (1987) . Plane Answers to Complex Questions, Springer - Verlag, New York .*
3. *Cox, D. R. & McCullagh, P. (1982) . Some Aspects of Analysis of Covariance, BIOMETRICS, Vol. 38, 541 - 561.*
4. *Dempster, A. P. & Ryan, L. M. (1985) . Weighted Normal Plots. Journal of the American Statistical Association, Vol. 80, 845 - 850.*
5. *Federer, W. T. & Meredith, M. P. (1992). Covariance Analysis for Split - plot and Split- Block Designs, The American Statistician, Vol. 46, No. 2, 155 - 162.*
6. *Guimarães, P. R. B. (1994). Modelo Linear Misto de Laird - Ware : Predição de Efeitos Aleatórios e Estimação de Parâmetros via Filtro de Kalman, Tese de Mestrado, UNICAMP.*
7. *Graybill, F. A. (1961) . An Introduction to Linear Statistical Models, McGraw - Hill Book Company , New York .*

8. *Graybill, F. A.* (1961). *Matrices with Applications in Statistics*, Wadsworth International Group, Belmont .
9. *Harville , D. A.* (1976). Extension of the Gauss - Markov Theorem to Include the Estimation of Random Effects. *The Annals of Statistics*, Vol. 4, 384 - 395.
10. *Harville , D. A.* (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems, *JASA*, Vol. 72, No. 358, 320-340.
11. *Harville , D. A.* (1985). *Notas de Aulas, não publicadas.*
12. *Healy, M. J. R.* (1986). *Matrices for Statistics*, Clarendon Press, Oxford.
13. *Jones, R. H.* (1993). *Longitudinal Data with Serial Correlation : A State Space Approach*, Chapman & Hall , London.
14. *Kemphorne, O.* (1980). *Notas de Aulas, não publicadas.*
15. *Kleinbaum, D. G., Kupper, L. L. & Muller, K. E.* (1988). *Applied Regression Analysis and other Multivariable Methods*, PWS, Kent.
16. *Koch, G. G. , Amara, I. A. , Davis, G. W. & Gillings, D. B.* (1982). A Review of Some Statistical Methods for Covariance Analysis of Categorical Data, *BIOMETRICS*, VOL. 38, 563 - 595.

17. Kravitz , L. , Cizar, G. J. , Christensen, C. L. & Setterlund, S. S. (1993). The Physiological Effects of Step Training with and without Handweights, The Journal of Sports Medicine and Physical Fitness, Vol. 33, No. 4, 348 - 358 .
18. Laird, N. M. & Ware, J. H. (1982). Random-Effects Models for Longitudinal Data, BIOMETRICS, Vol. 38, 963 - 974.
19. Lane, P. W. & Nelder, J. A. (1982). Analysis of Covariance and Standardization as Instances of Prediction, BIOMETRICS, vol. 38, 613 - 621.
20. McArdle, J. J. & Hamagami, F. (1994). Logit and Multilevel Logit Modeling of College Graduation for 1984 - 1985 Freshman Student-Athletes, Journal of the American Statistical Association, Vol. 89, No. 427, 1107- 1123.
21. Miller Jr. , Rupert G. (1986). Beyond Anova, Basics of Applied Statistics , John Wiley, New York.
22. Montgomery, D. C. (1991). Design and Analysis of Experiments, John Wiley, New York.
23. Neter, J. & Wasserman, W. (1974). Applied Linear Statistical Models, Richard D. Yrwin, New York.

24. Ridder, G. , Cramer, J. S. & Hopstaken, P. (1994). Down to Ten : Estimating the Effect of a Red Card in Soccer, Journal of the American Statistical Association, Vol. 89, No. 427, 1124- 1127.

25. Rutherford, Andrew (1992). Alternatives to Traditional Analysis of Covariance, British Journal of Mathematical and Statistical Psychology, Vol. 45, Part 2, 197 - 223.

26. Ryan, L. M. & Lange, N. (1989). Assessing Normality in Random Effects Models. The Annals of Statistics. Vol. 17, 624 - 642.

27. Searle, S. R. (1971). Linear Models, John Wiley, New York.

28. Searle, S. R. (1987). Linear Models for Unbalanced Data, John Wiley, New York.

29. Seber, G. A. F. (1977). Linear Regression Analysis, John Wiley, New York.

30. Singer, J. M. (1986). Análise de Dados Longitudinais, VII SINAPE .

31. Snedecor, G. W. & Cochran, W. G. (1982). Statistical Methods, The Iowa State University Press, Ames.

32. Tufte, E. R. (1974). Data Analysis for Politics and Policy, Prentice Hall, New Jersey.

33. *Urquhart, N. S.* (1982). Adjustment in Covariance when One Factor Affects the Covariate, *BIOMETRICS*, VOL. 38, 651-660.