

4- Modelos com Matriz de Planejamento de Fatores Incompleto (Cap 13 - Graybill)

127

4.1- Introdução

Efeito de dois compostos químicos e quatro diferentes métodos de aplicações na produção de milho

Y_{ij} - produção associada ao i -ésimo composto no j -ésimo método
 $i=1,2 \quad j=1,2,3,4$

$$Y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij} \rightarrow \text{análise de variância com dois fatores (modelo aditivo)}$$

μ - média geral

α_i - efeito do i -ésimo composto

τ_j - efeito do j -ésimo método de aplicação

↳ Parametrizações dos desvios médios

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}$$

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{24} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \end{bmatrix}$$

$\underbrace{\quad}_{X_0} \quad \underbrace{\quad}_{X_1} \quad \underbrace{\quad}_{X_2}$

Posto (X) = 7-2 = 5, não completo

128

Outra Parametrização (da tabela de referência)

$$\tilde{Y} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + \tilde{\epsilon} = \begin{bmatrix} \mu \\ \mu + \gamma_1 \\ \mu + \gamma_2 \\ \mu + \gamma_3 \\ \mu + \alpha \\ \mu + \alpha + \gamma_1 \\ \mu + \alpha + \gamma_2 \\ \mu + \alpha + \gamma_3 \end{bmatrix} + \tilde{\epsilon}$$

X é de posto completo = 5 5 parâmetros

$$\gamma_j = \begin{cases} 0 & 1^\circ \text{ composto} \\ 1 & 2^\circ \text{ composto} \end{cases}$$

$$(\alpha_2, \alpha_3, \alpha_4) = \begin{cases} (0, 0, 0) & 1^\circ \text{ método} \\ (1, 0, 0) & 2^\circ \text{ método} \\ (0, 1, 0) & 3^\circ \text{ método} \\ (0, 0, 1) & 4^\circ \text{ método} \end{cases}$$

↳ 2 categorias, uma var. dummy

↳ 4 categorias, 3 var. dummy

Parametrização dos desvios médios com restrições $\Rightarrow$ matriz de planejamento de posto completo

$$Y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij} \quad i=1,2 \quad j=1,2,3,4$$

$$\alpha_1 + \alpha_2 = 0 \quad \Rightarrow \alpha_2 = -\alpha_1$$

$$\tau_1 + \tau_2 + \tau_3 + \tau_4 = 0 \quad \Rightarrow \tau_4 = -\tau_1 - \tau_2 - \tau_3$$

$$\beta = \begin{bmatrix} \mu \\ \alpha \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}$$

$$\alpha_1 = \alpha$$

$$\alpha_2 = -\alpha$$

$$Y = X\beta = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} = \begin{bmatrix} \mu + \alpha + \tau_1 \\ \mu + \alpha + \tau_2 \\ \mu + \alpha + \tau_3 \\ \mu + \alpha - \underbrace{\tau_1 - \tau_2 - \tau_3}_{\tau_4} \\ \mu - \alpha + \tau_1 \\ \mu - \alpha + \tau_2 \\ \mu - \alpha + \tau_3 \\ \mu - \alpha - \tau_1 - \tau_2 - \tau_3 \end{bmatrix}$$

X - posto completo = 5 = n.º de parâmetros a serem estimados
(Todos os elementos de β podem ser estimados, tópico 3 do curso)

Definição

Seja o modelo linear geral $Y = X\beta + \epsilon$, Y vetor aleatório $n \times 1$, X matriz $n \times p$ de posto k ($n > p \geq k$), β vetor $p \times 1$ de parâmetros desconhecidos, ϵ vetor aleatório $n \times 1$.

Este modelo é um modelo de planejamento de experimento se e somente se os elementos de X são zeros ou uns e X puder ser particionada como

$[X_0 \ X_1 \ \dots \ X_q]$, onde X_j é $n \times q_j$ satisfazendo as seguintes condições;

(1) Para cada submatriz X_j , $j=0,1,\dots,q$ cada linha contém exatamente um elemento igual a 1 e os restantes são zero

(2) Para cada X_j , $j=0,1,\dots,q$, cada coluna contém no mínimo um elemento não nulo.

Ex: $X_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Obs:

1) De acordo com esta definição, o vetor β fica particionado como $\beta' = [\beta_0 \ \beta_1 \ \dots \ \beta_q]$ e cada equação do modelo

$Y = X\beta + \epsilon$ contém exatamente um elemento de cada vetor $\beta_0, \beta_1, \dots, \beta_q$.

2) Se $Y = X\beta + \epsilon$ é um modelo de planejamento de experimentos, o correspondente modelo reparametrizado pode não ser um modelo de planejamento de experimentos.

(ver exemplo)

Nas parametrizações da casela de referência e dos desvios médios com restrições ~~na~~ a matriz X não satisfaz estas propriedades.

3) Se $K < p$, o modelo é de posto não completo.

4.2 - Estimativas Pontuais

Modelo

$$Y = X\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I)$$

X - matriz de planejamento $n \times p$, $r(X) = k < p < n$

Estimadores de Máxima Verossimilhança

$$L(\beta, \sigma^2 | y, X) = (2\pi\sigma^2)^{-n/2} \exp \left[-\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta) \right]$$

$$\frac{\partial L}{\partial \beta} = 0 \quad \frac{\partial L}{\partial \sigma^2} = 0 \Rightarrow$$

$$\begin{cases} X'X \hat{\beta}_{mv} = X'Y \\ \hat{\sigma}_{mv}^2 = \frac{1}{n} (Y'Y - \hat{\beta}_{mv}' X'Y) \end{cases}$$

$r(X) = k < p \Rightarrow r(X'X) = k < p$, $X'X$ é não inversível

$$\hat{\beta} = (X'X G)^{-1} X'Y + (X'X G X'X - X'X) z$$

onde $X'X G X'X = X'X$ e z satisfaz ao sistema

Estimadores de Mínimos Quadrados

Equações Normais: $X'X \hat{\beta} = X'Y$ (por $X'X$ não ser invertível, não podemos simplesmente dividir ambos os lados da equação por $X'X$)

$\hat{\beta}^* = G X'Y$ onde G é uma inversa generalizada de $X'X$

Solução Geral (Teorema 4 de Inversas Generalizadas)

$$\hat{\beta}^* = G X'Y + (G X'X - I) z \quad z \in \mathbb{R}^p$$

Observações

1) Existem infinitas soluções $\hat{\beta}^*$. Verifica-se que todas maximizam L e minimizam $S = \sum e_i^2$.

2) Qualquer valor $\hat{\beta}^*$ leva a um único valor de $\hat{\sigma}_{mv}^2$

$$\hat{\sigma}_{mv}^2 = \frac{1}{n} \left(\underline{Y}' \underline{Y} - \underline{\hat{\beta}}_{mv}' \underline{X}' \underline{Y} \right)$$

G' também é I.G. de $X'X$

pois $\underline{\hat{\beta}}_{mv}' \underline{X}' \underline{Y} = \underline{Y}' \underline{X} \underline{G}' \underline{X}' \underline{Y}$ e $\underline{X} \underline{G}' \underline{X}'$ é invariante
 (é constante qualquer que seja $\underline{G}$ inversa generalizada de $X'X$)

3) Não vai existir um estimador (m.v. ou m.q.) único para todos os elementos de β . No entanto, existirá estimador único para algumas funções dos parâmetros presentes em

-||-||-

Um modelo estatístico pode ser escrito como

$$\mathcal{P} = \{ P_{\underline{\theta}} : \underline{\theta} \in \textcircled{H} \}$$

onde $P_{\underline{\theta}}$ distribuição de probabilidade dos dados observados

$\textcircled{H}$ espaço paramétrico

$\underline{\theta}$ parâmetro (uni ou multidimensional)

Muitas vezes, uma parametrização pode ser tal que

$$\underline{\theta}_1 \neq \underline{\theta}_2 \Rightarrow P_{\underline{\theta}_1} = P_{\underline{\theta}_2}$$

Esta parametrização é dita não identificável. X

Definição

Seja $\underline{Y}$ um vetor aleatório com função densidade de probabilidade $f_Y(\underline{y}, \underline{\theta})$, $\underline{\theta} \in \Theta$. A parametrização utilizada e o correspondente parâmetro $\underline{\theta}$ é dito identificável se e somente se

$$\underline{\theta}_1 \neq \underline{\theta}_2 \Rightarrow P_{\underline{\theta}_1} \neq P_{\underline{\theta}_2}$$

Exemplo 1

$$\underline{Y} = (Y_1, Y_2)'$$

Y_1 e Y_2 v.a. independentes

$$Y_1 \sim N(\alpha_1 + \beta, 1)$$

$$Y_2 \sim N(\alpha_2 + \beta, 1)$$

$$\alpha_1, \alpha_2, \beta \in \mathbb{R}$$

Se $\underline{\theta} = (\alpha_1, \alpha_2, \beta)$ é o parâmetro da distribuição de probabilidades de $\underline{Y}$, $\underline{\theta}$ é não identificável pois

para $\underline{\theta}_1 = (0, 0, 2)$ e $\underline{\theta}_2 = (1, 1, 1)$, $Y_1 \sim N(2, 1)$

$$Y_2 \sim N(2, 1) \quad [P_{\underline{\theta}_1} = P_{\underline{\theta}_2}] \text{ mas } \underline{\theta}_1 \neq \underline{\theta}_2$$

Exemplo 2

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

$$r(X) = k < p \quad X - n \times p$$

$$\underline{\theta} = [\underline{\beta}', \sigma^2]$$

Se $\underline{\mu}_0$ é um vetor $n \times 1$ tal que o sistema $X\underline{\beta} = \underline{\mu}_0$ tenha uma solução ^{tem $\underline{\beta}$} <sub>n equações
p incógnitas</sub>. X

$k < p \Rightarrow$ o sistema tem infinitas soluções $\Rightarrow$
[$\exists$ infinitos $\underline{\beta}$ | $X\underline{\beta} = \underline{\mu}_0$]

$\underline{Y} \sim N_n(\underline{\mu}_0, \sigma^2 I)$ para um número infinito de diferentes valores de $\underline{\beta}$ (e: $\underline{\theta}$). Portanto $\underline{\theta}$ não é identificável. Determinaremos quais funções de $\underline{\theta}$ são identificáveis.

Definições: Funções Estimáveis

No modelo $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim N(0, \sigma^2 I)$, $r(X) = k < p < n$ ($X - n \times p$) uma função linear do parâmetro $\underline{\beta}$, $\underline{L}'\underline{\beta}$ é uma função estimável se existir um estimador não viesado de $\underline{L}'\underline{\beta}$ que é linear em $\underline{Y}$.

Isto equivale a dizer que existe $\underline{t}'$ tal que

$$E(\underline{t}'\underline{Y}) = \underline{L}'\underline{\beta}.$$

Teorema 1

No modelo $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$, X de posto incompleto, se $L'\beta$ é uma função linear de β , então

a) $L'\beta$ é uma função estimável se e somente se existir um vetor $\underline{t}$, $n \times 1$, tal que $L' = \underline{t}'X$. desnecessário
↑

b) $L'\beta$ é identificável se e somente se $L'\beta$ é estimável.

$$\Rightarrow L'\beta \text{ estim.} \Leftrightarrow E(\underline{t}'Y) = L'\beta \Leftrightarrow \underline{t}'X\beta = L'\beta \Leftrightarrow \underline{t}'X = L' \quad \forall \beta$$

Exemplo

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

$$i = 1, 2, 3$$

$$j = 1, 2 \dots n_i \quad n_i - \text{número de obs. do } i\text{-ésimo nível do fator}$$

$$n_1 = 3 \quad n_2 = 2 \quad n_3 = 1$$

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}_{6 \times 4} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \varepsilon$$

$$r(X) = 3 = k < p = 4$$

a) O valor esperado de qualquer observação é estimável

é da forma $\underline{l}'\beta$ para $\underline{l}' = [1 \ 1 \ 0 \ 0]$

② $E(Y_{11})$ é estimável pois tomando $\underline{t}' = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$

$$E(\underline{t}'\underline{Y}) = E(Y_{11})$$

Verificamos aqui pela definição
 Se um est. linear $\underline{t}'\underline{Y}$ | $E(\underline{t}'\underline{Y}) = ao$
 que desejest

Se um estimador $\underline{t}'\underline{Y}$ linear não viado de $E(Y_{11})$
 Pelo Teorema

$$\textcircled{1} E(Y_{11}) = \underline{l}'\beta = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \mu + \alpha_1$$

$\downarrow$
 $\underline{t}'X$

$= \mu + \alpha_1$

$\therefore E(Y_{11})$ é estimável

③ $\underline{l}' = \underline{t}'X$ para $\underline{t}' = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$

substituir por *

b) $E(Y_{11}) - E(Y_{21}) = \alpha_1 - \alpha_2$ é estimável

Tomar $\underline{t}' = [1 \ 0 \ 0 \ -1 \ 0 \ 0]$, $\underline{t}'\underline{Y} = Y_{11} - Y_{21}$

$$E(\underline{t}'\underline{Y}) = E(Y_{11} - Y_{21}) = \alpha_1 - \alpha_2 \quad \therefore \alpha_1 - \alpha_2 \text{ é estimável pela definição}$$

c) μ não é estimável

$$\underline{l}'\beta \text{ é estimável} \Leftrightarrow \underline{l}' = \underline{t}'X$$

$$\mu = \underline{l}'\beta \text{ para } \underline{l}' = [1 \ 0 \ 0 \ 0]$$

$$[1 \ 0 \ 0 \ 0] = [t_1 \ t_2 \ t_3 \ t_4 \ t_5 \ t_6] \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{cases} t_1 + t_2 + t_3 + t_4 + t_5 + t_6 = 1 \\ t_1 + t_2 + t_3 = 0 \\ t_4 + t_5 = 0 \\ t_6 = 0 \end{cases}$$

impossível

$$\nexists \underline{t}' \mid \underline{l}' = \underline{t}'X$$

Não consigo sem est. lin. não vic. p. 111

$E(Y_{11}) = \mu + \alpha_1$ e é da forma $\underline{\ell}'\beta$ para

*
136'

$$\underline{\ell}' = [1 \ 1 \ 0 \ 0]$$

$E(Y_{11})$ é estimável pois tomando $\underline{x}' = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$

$$E(\underline{x}'\underline{Y}) = E\left\{ [1 \ 0 \ 0 \ 0 \ 0 \ 0] \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{31} \end{bmatrix} \right\} = E(Y_{11})$$

Verificamos aqui, pela definição, que $E(Y_{11})$ é estimável

($\exists \underline{x}'\underline{Y} \mid E(\underline{x}'\underline{Y}) = \text{parâmetro que desejamos estimar}$)

Realmente $\underline{\ell}' = \underline{x}'X$ para

$$\underline{x}' = [1 \ 0 \ 0 \ 0 \ 0 \ 0] \text{ (verificar)}$$

Definições

$\underline{L}'_1 \beta \quad \underline{L}'_2 \beta \quad \dots \quad \underline{L}'_m \beta$ é um conjunto de m funções estimáveis linearmente independentes se cada uma delas é estimável e os vetores $p \times 1$, $\underline{L}_1, \underline{L}_2, \dots, \underline{L}_m$ são linearmente independentes (o posto de $L = [\underline{L}_1 \quad \underline{L}_2 \quad \dots \quad \underline{L}_m]$ é m).

Teorema 2

O número de funções estimáveis linearmente independentes $\underline{L}' \beta$ é $r(X) = K$.

Obs:

No modelo de posto completo, qualquer função $\underline{L}' \beta$ é estimável. (Daí que este conceito não tem importância no modelo de posto completo).

Teorema 3

No modelo $\underline{Y} = X\beta + \underline{\epsilon}$, $\underline{\epsilon} \sim N(0, \sigma^2 I)$, $r(X) = K$

i) Se $\underline{L}' \beta$ é estimável, então $\underline{L}' \hat{\beta}^*$ é invariante para qualquer solução $\hat{\beta}^*$ das equações normais.

$$\underline{L}' \beta \text{ est} \Rightarrow \underline{L}' = \underline{L}' X \Rightarrow \underline{L}' \hat{\beta}^* = \underline{L}' X G X' \underline{Y}$$

$\underbrace{\hspace{10em}}_{\text{inv em } G}$

ii) Se $\underline{L}'\beta$ é estimável então $\underline{L}'\hat{\beta}^*$ é NVUVM para $\underline{L}'\beta$, onde $\hat{\beta}^*$ é qualquer solução das equações normais.

iii) $\hat{\sigma}^2 = \frac{\underline{Y}'\underline{Y} - \hat{\beta}^{*'}\underline{X}'\underline{Y}}{n-k}$ é o estimador NVUVM de σ^2

NVU $\rightarrow$

(é invariante)

iv) Cada linha de $\underline{X}\beta$ é uma função estimável $= E(y_{ij})$.

—||—||—

IG de $\underline{X}'\underline{X}$

$$\underline{X}'\underline{X} \underline{G} \underline{X}'\underline{X} = \underline{X}'\underline{X}$$

$$\begin{aligned} \underline{X}\underline{G}\underline{X}'\underline{X} &= \underline{X} \\ \underline{X}\underline{G}\underline{X}' &\text{é uma} \\ \underline{G}' &\leq \underline{I}\underline{G} \end{aligned}$$

Lembrando

$\underline{L}'\beta$ é estimável se existir $\underline{t}'$ tal que $E(\underline{t}'\underline{Y}) = \underline{L}'\beta$

Obs:

1) A estimabilidade de $\underline{L}'\beta$ garante a invariância de $\underline{L}'\hat{\beta}^*$, $\forall \hat{\beta}^*$ solução das equações normais.

2) No modelo de posto completo, qualquer função $\underline{L}'\beta$ é estimável. (Por isto, este conceito não tem importância no modelo de posto completo).

3) $\hat{\beta}^* = \underline{G}\underline{X}'\underline{Y} \sim N_p(\underline{G}\underline{X}'\underline{X}\beta, \underline{G}\underline{X}'\underline{X}\underline{G}'\sigma^2)$

$\hookrightarrow$ pode ser singular

invariante

4) $\underline{X}\beta$ (n dimensional) é estimável e

$\therefore \underline{X}\hat{\beta}^*$ é invariante, $\underline{X}\hat{\beta}^* \sim N_n(\underbrace{\underline{X}\underline{G}\underline{X}'\underline{X}}_{\underline{X}}\beta, \underbrace{\underline{X}\underline{G}\underline{X}'\underline{X}\underline{G}'\underline{X}'}_{\underline{X}}\sigma^2)$

$$\underline{X}\underline{G}'\underline{X}' = \underline{X}\underline{G}\underline{X}'$$

$$\sim N_n(\underline{X}\beta, \underline{X}\underline{G}\underline{X}'\sigma^2)$$

invariante em $\underline{G} = \underline{X}$

$$i) \underset{\sim}{L}' \underset{\sim}{\beta} \text{ estimável} \Rightarrow \underset{\sim}{L}' = \underset{\sim}{L}' X \Rightarrow$$

$$\underset{\sim}{L}' \underset{\sim}{\hat{\beta}}^* = \underset{\sim}{L}' X \underset{\sim}{G} X' \underset{\sim}{Y}$$

invariante em G

Os:



iv) Cada linha de $X \underset{\sim}{\beta}$ é uma média dos Y
 Todas as médias de Y são estimáveis

G $\perp G$ de $X'X$

$$1) XGX'X = X$$

$$2) XGX' \text{ é invariante em } G$$

$$3) G' \text{ é } \perp G \text{ de } X'X$$

Exemplo

Consideremos o peso de 6 plantas de três diferentes tipos

Normal	atípica	aberrante
101	84	32
105	88	
94		

Y_{ij} - peso da j -ésima planta do i -ésimo tipo

$$\begin{bmatrix} 101 \\ 105 \\ 94 \\ 84 \\ 88 \\ 32 \end{bmatrix} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{31} \end{bmatrix}$$

$$X'X = \begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} Y_{11} + Y_{12} + Y_{13} + Y_{21} + Y_{22} + Y_{31} \\ Y_{11} + Y_{12} + Y_{13} \\ Y_{21} + Y_{22} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} Y_{1..} \\ Y_{1j.} \\ Y_{2.} \\ Y_{3.} \end{bmatrix}$$

$$= \begin{bmatrix} 504 \\ 300 \\ 172 \\ 32 \end{bmatrix}$$

Equações Normais

$$\begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \alpha_1^0 \\ \alpha_2^0 \\ \alpha_3^0 \end{bmatrix} = \begin{bmatrix} 504 \\ 300 \\ 172 \\ 32 \end{bmatrix}$$

$$AGA = A$$

Verifica-se que $G = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ é uma inversa generalizada de A e é a única transposta.

transponha o resultado
 $XGX'X = X$

linhas de $X'X$

$$XGX' \text{ é inv}$$

$$G' \text{ é } IG$$

$$\hat{\beta}^* = GX'Y = [0 \quad 100 \quad 86 \quad 32]'$$

vide

$$\hat{\sigma}^2 = \frac{(Y'Y - \hat{\beta}^{*'}X'Y)}{n-k}$$

$$k = r(X) = 3$$

$$Y'Y = 45886 \quad \hat{\beta}^{*'}X'Y = 45.816$$

$$\hat{\sigma}^2 = \frac{70}{3}$$

$\alpha_1 - \alpha_2$ é estimável

$$\alpha_1 - \alpha_2 = [0 \quad 1 \quad -1 \quad 0] \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

$$\hat{\alpha}_1 - \hat{\alpha}_2 = [0 \quad 1 \quad -1 \quad 0] \hat{\beta}^*$$

é invariante com $\hat{\beta}^*$

$X'X$ é de posto 3

- 1) Determine uma submatriz de posto 3 não singular M
- 2) Calcule $(M^{-1})'$
- 3) Em $X'X$ substitua cada elemento de M pelo correspondente de $(M^{-1})'$
- 4) Substitua os demais por zero
- 5) Transponha o resultado. A matriz obtida é uma IG de $X'X$.

Portanto

141

$$\hat{\alpha}_1 - \alpha_2 = \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 100 \\ 86 \\ 32 \end{bmatrix} = 14$$

1) Verifica-se que α_1 e α_2 não são estimáveis.

2) $E(Y_{11}) = E(Y_{12}) = E(Y_{13}) = \mu + \alpha_1$ é estimável
(1ª, 2ª e 3ª linhas de X(2))

$$\therefore \hat{\mu + \alpha_1} = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \hat{\beta}^* = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 100 \\ 86 \\ 32 \end{bmatrix} = 100$$

$$\left(\mu + \alpha_1 = \underline{L}' \beta \text{ p/ } \underline{L}' = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \right)$$

Observar que $\hat{\mu + \alpha_1} = \bar{Y}_1 = \frac{Y_{1.}}{3} = \frac{300}{3}$ (média do grupo) ^{amostral} _{normal}

3) $\text{Var}(\hat{\alpha}_1 - \alpha_2)$

$$\text{Var}(\hat{\beta}^*) = \text{Var}(GX'Y) = GX'XG'\sigma^2$$

$$\text{Var}(q'\hat{\beta}^*) = q'\text{Var}(\hat{\beta}^*)q = q'GX'XG'q\sigma^2$$

Para $q'\beta$ estimável, $q' = t'X \Rightarrow$

$$\text{Var}(q'\hat{\beta}^*) = q'GX'XG'X't\sigma^2 = q'GX't\sigma^2 =$$

$$= q'GXq\sigma^2 = \text{Var}(q'\hat{\beta}^*) \text{ é constante } \forall q \text{ pois } XGX' \text{ é invar.}$$

usar p/ cálculo

$\text{Var}(\mathbf{q}' \hat{\beta}^*)$ é estimável e . . .

$$\hat{\text{Var}}(\mathbf{q}' \hat{\beta}^*) = \mathbf{q}' \mathbf{G} \mathbf{q} \hat{\sigma}^2$$

invariante

$\text{Var}(\mathbf{q}' \hat{\beta}^*) = \mathbf{q}' \mathbf{G} \mathbf{q} \sigma^2$ cte $\forall \mathbf{G}$
pois vale $\underbrace{\mathbf{Z}' \mathbf{X} \mathbf{G} \mathbf{X}' \mathbf{Z}}_{\text{inv.}}$ constante
Média normalizada com corte 5 sobre 26

