166

25/05

Pasta ~~15/05~~ 16

Nº cópias 23

4.5 - Análise de Variância com dois fatores, Uma observação por casela sem interação

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

$i = 1,2 \dots a > 1$ $\qquad j = 1,2 \dots b > 1$ $\qquad n = ab$

$$\underset{\sim}{Y} = X \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

$$\underset{\sim}{\beta} = \begin{bmatrix} \mu \\ \underset{\sim}{\alpha} \\ \underset{\sim}{\tau} \end{bmatrix} \qquad \underset{\sim}{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \\ \alpha_a \end{bmatrix} \qquad \underset{\sim}{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \\ \tau_b \end{bmatrix}$$

A\B	B_1	B_2	…	B_b	
A_1	Y_{11}	Y_{12}		Y_{1b}	$Y_{1.}$
A_2	Y_{21}	Y_{22}		Y_{2b}	$Y_{2.}$
A_a	Y_{a1}	Y_{a2}		Y_{ab}	$Y_{a.}$
	$Y_{.1}$	$Y_{.2}$		$Y_{.b}$	$Y_{..}$

$$X = \left[\begin{array}{c|c} 1 & \overbrace{\begin{matrix} 1 & 0 & \dots & 0 \end{matrix}}^{a} \quad \overbrace{\begin{matrix} 1 & 0 & \dots & 0 \end{matrix}}^{b} \\ 1 & \begin{matrix} 1 & 0 & \dots & 0 \end{matrix} \quad \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \\ & \vdots \\ 1 & \begin{matrix} 1 & 0 & \dots & 0 \end{matrix} \quad \begin{matrix} 0 & 0 & \dots & 1 \end{matrix} \\ \\ 1 & \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \quad \begin{matrix} 1 & 0 & \dots & 0 \end{matrix} \\ 1 & \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \quad \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \\ \\ 1 & \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \quad \begin{matrix} 0 & 0 & \dots & 1 \end{matrix} \\ \\ \vdots \\ \\ 1 & \begin{matrix} 0 & 0 & & 1 \end{matrix} \quad \begin{matrix} 1 & 0 & \dots & 0 \end{matrix} \\ 1 & \begin{matrix} 0 & 0 & & 1 \end{matrix} \quad \begin{matrix} 0 & 1 & \dots & 0 \end{matrix} \end{array}\right. \begin{array}{l} \left.\begin{array}{l} \\ 1 \\ b \text{ linhas} \\ \\ \end{array}\right\} \\ \\ \left.\begin{array}{l} \\ 2 \\ b \text{ linhas} \\ \\ \end{array}\right\} \\ \\ \\ \\ \\ a \end{array}$$

$$\underset{\sim}{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1b} \\ \hline Y_{21} \\ Y_{22} \\ \\ Y_{2b} \\ \hline \\ Y_{a1} \\ Y_{a2} \\ \\ Y_{ab} \end{bmatrix}_{ab \times 1}$$

$$X'X = \begin{bmatrix} ab & b & b & \dots & b & a & a & \dots & a \\ & b & 0 & & 0 & 1 & 1 & \dots & 1 \\ & & b & \dots & 0 & 1 & 1 & \dots & 1 \\ & & & & b & 1 & 1 & \dots & 1 \\ & & & & & a & 0 & \dots & 0 \\ & & & & & & a & \dots & 0 \\ & & & & & & & & 0 \end{bmatrix}_{(a+b+1)\times(a+b+1)}$$

$$r(X) = r(X'X) = a+b-1$$

$$x_1 + x_2 + \dots + x_a = y_1 + y_2 + \dots + y_b$$

$$\Rightarrow y_b = x_1 + \dots + x_a - y_1 \dots - y_{b-1}$$

Equações Normais

$$\begin{cases} ab\hat{\mu} + b\sum \hat{\alpha}_i + a\sum \hat{\tau}_j = Y_{..} \\ b\hat{\mu} + b\hat{\alpha}_i + \sum \hat{\tau}_j = Y_{i.} & i = 1,2\dots a \\ a\hat{\mu} + \sum \hat{\alpha}_i + a\hat{\tau}_j = Y_{.j} & j = 1,2\dots b \end{cases}$$

a+b+1 parâmetros, a+b+1 equações

Uma solução das equações normais é

$$\hat{\mu}^* = \overline{Y}_{..}$$

$$\hat{\alpha}^*_i = \overline{Y}_{i.} - \overline{Y}_{..} \qquad i = 1,2\dots a$$

$$\hat{\tau}^*_j = \overline{Y}_{.j} - \overline{Y}_{..} \qquad j = 1,2,,,b$$

$$\overline{Y}_{..} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} Y_{ij}}{ab} = \frac{Y_{..}}{ab}$$

$$\overline{Y}_{i.} = \frac{\sum_{j=1}^{b} Y_{ij}}{b} = \frac{Y_{i.}}{b}$$

$$\overline{Y}_{.j} = \frac{\sum_{i=1}^{a} Y_{ij}}{a} = \frac{Y_{.j}}{a}$$

$r(X) = r(X'X) \leq a+b-1$

Podem existir mais dependências lineares.

Para verificar isso Graybill pag 535 sugere analisar o sistema de equações normais da seguinte maneira;

1- Dividir a 1ª linha por ab

2- Subtrair a linha α_i da linha α_1 e dividir por b, $i = 2, 3 \dots a$

3- Subtrair a linha τ_j da linha τ_1 e dividir por a, $j = 2, 3 \dots b$

Ignorando as linhas α_1 e τ_1 obtem-se

$$\hat{\mu} + \frac{\sum \hat{\alpha}_i}{a} + \frac{\sum \hat{\tau}_j}{b} = \bar{Y}_{..}$$

$$\hat{\alpha}_1 - \hat{\alpha}_i = \bar{Y}_{1.} - \bar{Y}_{i.} \qquad i = 2, 3 \dots a$$

$$\hat{\tau}_1 - \hat{\tau}_j = \bar{Y}_{.1} - \bar{Y}_{.j} \qquad j = 2, 3 \dots b$$

Temos $1 + a - 1 + b - 1 = a + b - 1$ eq. LI porque somente a 1ª equação contem μ e as restantes contem parâmetros distintos.

Ex: $a = 2$ 2 tipos de feijão
$b = 4$ 4 tipos de fertilizante

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j \qquad i = 1,2 \quad j = 1,2,3,4$$

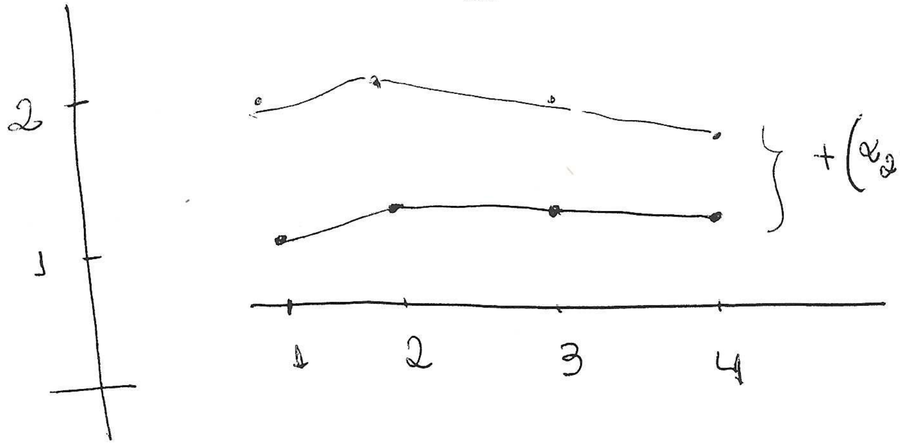
$$E(\underset{\sim}{Y}) = \begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) = \\ E(Y_{24}) = \end{bmatrix} = \begin{bmatrix} \mu + \alpha_1 + \tau_1 \\ \mu + \alpha_1 + \tau_2 \\ \mu + \alpha_1 + \tau_3 \\ \mu + \alpha_1 + \tau_4 \\ \mu + \alpha_2 + \tau_1 \\ \mu + \alpha_2 + \tau_2 \\ \mu + \alpha_2 + \tau_3 \\ \mu + \alpha_2 + \tau_4 \end{bmatrix}$$

$+\alpha_2 - \alpha_1$

$+\alpha_2 - \alpha_1$

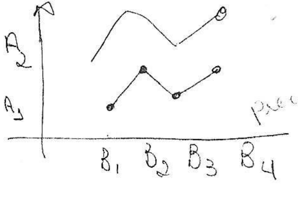
$+\alpha_2 - \alpha_1$

$\alpha_2 - \alpha_1$



$\} + (\alpha_2 - \alpha_1)$

$\forall$ nível do fator B



? médias (? por) Como é sem int; precisa 4+1 porém $b + a - 1$

O modelo é superp. porque usa $\underbrace{a+b}_{7}+1$ por uso

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

$$= E(Y_{22}) - E(Y_{12}) = E(Y_{23}) - E(Y_{13}) = E(Y_{24}) - E(Y_{14})$$

acréscimo quando passo do nível 1 p/ 2 do fator A fixado o nível de B é o mesmo para os 4 níveis de B.

$$E(Y_{ij}) = \mu + \alpha_i + \tau_j + \gamma_{ij}$$

[illegible] $\tau + \gamma - (\mu + \alpha_1 + \tau_1 + \gamma_{11}) = \alpha_2 - \alpha_1 + (\gamma_{22} - \gamma_{11})$

Verifica-se que

$$\bar{\alpha} = \frac{\sum_{i=1}^{a} \alpha_i}{a} \qquad \bar{\tau} = \frac{\sum_{j=1}^{b} \tau_j}{b}$$

$$\mu + \bar{\alpha} + \bar{\tau}$$

$$\alpha_1 - \alpha_i \qquad i = 2, 3 \ldots a$$

$$\tau_1 - \tau_j \qquad j = 2, 3 \ldots b$$

graybill pag 536

formam um conjunto base de funções LI estimáveis no modelo.

Os correspondentes estimadores ótimos destas quantidades são

$$\hat{\mu}^* + \frac{\sum_{i=1}^{a} \hat{\alpha}_i^*}{a} + \frac{\sum_{j=1}^{b} \hat{\tau}_j^*}{b} = \bar{Y}_{..}$$

$$\hat{\alpha}_1^* - \hat{\alpha}_i^* = \bar{Y}_{1.} - \bar{Y}_{i.} \qquad i = 2, 3 \ldots a$$

$$\hat{\tau}_1^* - \hat{\tau}_j^* = \bar{Y}_{.1} - \bar{Y}_{.j} \qquad j = 2, 3 \ldots b$$

$\hat{\mu}^* = \bar{Y}_{..}$ $\quad \hat{\alpha}_i^* = \bar{Y}_{i.} - \bar{Y}_{..}$ $\quad \hat{\tau}_j^* = \bar{Y}_{.j} - \bar{Y}_{..}$ solução das equações normais

$$\bar{Y}_{..} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} Y_{ij}}{ab} = \frac{Y_{..}}{ab} \qquad \bar{Y}_{i.} = \frac{\sum_{j=1}^{b} Y_{ij}}{b} = \frac{Y_{i.}}{b}$$

$$\bar{Y}_{.j} = \frac{\sum_{i=1}^{a} Y_{ij}}{a} = \frac{Y_{.j}}{a}$$

Teorema

No modelo de análise de variância com dois fatores, sem interação, uma observação por casela, $\sum c_i \alpha_i$ é estimável se e somente se $\sum c_i \alpha_i$ é um contraste nos α_i's. O estimador NVUVM de $\sum c_i \alpha_i$ é $\sum_{i=1}^{a} c_i \hat{\alpha}_i^* = \sum_{i=1}^{a} c_i \bar{Y}_{i.}$ Analogamente, com relação aos T_j, o estimador NVUVM do contraste $\sum_{j=1}^{b} d_j T_j$ é $\sum d_j \hat{T}_j^* = \sum_{j=1}^{b} d_j \bar{Y}_{.j}$.

Prova

$\sum c_i \alpha_i$ estimável $\Rightarrow \sum_{i=1}^{a} c_i \alpha_i = d_0(\mu + \bar{\alpha} + \bar{T}) + \sum_{i=1}^{a} d_i(\alpha_1 - \alpha_i) + \sum_{j=1}^{b} h_j (T_1 - T_j)$

$d_0 = 0 \quad h_j = 0 \quad j = 1, 2 \ldots b$

$= \sum_{i=1}^{a} d_i \alpha_1 - \sum_{i=1}^{a} d_i \alpha_i =$

$= \alpha_1 \left(\sum_{i=1}^{a} d_i - d_1 \right) - \sum_{i=2}^{a} d_i \alpha_i = \alpha_1 \sum_{i=2}^{a} d_i - \sum_{i=2}^{a} d_i \alpha_i$

Estimável p/

$c_1 = \sum_{i=2}^{a} d_i \qquad c_2 = -d_2,\ c_3 = -d_3 \quad ,,, \quad c_a = -d_a$

$, \quad \sum_{i=1}^{a} c_i = \sum_{i=2}^{a} d_i - d_2 - d_3 \ldots - d_a = 0$

Estimador NVUVM de $\sum c_i \alpha_i$ $\quad \sum c_i \hat{\alpha}_i^*$, $\quad \hat{\alpha}_i^* = \bar{Y}_{i.} - \bar{Y}_{..}$ soluções das eq. normais

Cálculo de estimadores NVUVM de σ^2

$$\hat{\sigma}^2 = \frac{1}{n-k}\left(\underset{\sim}{Y}'\underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^{*\prime} X' \underset{\sim}{Y}\right)$$

$$n = ab \qquad k = a+b-1$$

$$n-k = ab-a-b+1 = a(b-1)-(b-1) = (a-1)(b-1)$$

$$\underset{\sim}{Y}'\underset{\sim}{Y} = \sum_{i=1}^{a}\sum_{j=1}^{b} Y_{ij}^2$$

$$X'\underset{\sim}{Y} = \begin{bmatrix} Y_{\cdot\cdot} \\ Y_{1\cdot} \\ \vdots \\ Y_{a\cdot} \\ Y_{\cdot 1} \\ \vdots \\ Y_{\cdot b} \end{bmatrix} \qquad \hat{\underset{\sim}{\beta}}^{*} = \begin{bmatrix} \overline{Y}_{\cdot\cdot} \\ \overline{Y}_{1\cdot} - \overline{Y}_{\cdot\cdot} \\ \vdots \\ \overline{Y}_{a\cdot} - \overline{Y}_{\cdot\cdot} \\ \overline{Y}_{\cdot 1} - \overline{Y}_{\cdot\cdot} \\ \vdots \\ \overline{Y}_{\cdot b} - \overline{Y}_{\cdot\cdot} \end{bmatrix}$$

$$\hat{\underset{\sim}{\beta}}^{*\prime} X' \underset{\sim}{Y} = \hat{\mu}^* Y_{\cdot\cdot} + \sum_{i=1}^{a} \hat{\alpha}_i^* Y_{i\cdot} + \sum_{j=1}^{b} \hat{\tau}_j^* Y_{\cdot j}$$

$$= \frac{Y_{\cdot\cdot}^2}{ab} + \sum_{i=1}^{a} Y_{i\cdot}\left(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot}\right) + \sum_{j=1}^{b} Y_{\cdot j}\left(\overline{Y}_{\cdot j} - \overline{Y}_{\cdot\cdot}\right)$$

$$= \frac{Y_{\cdot\cdot}^2}{ab} + \left(\sum_{i=1}^{a}\frac{Y_{i\cdot}^2}{b} - \frac{Y_{\cdot\cdot}}{ab}\underbrace{\sum_{i=1}^{a} Y_{i\cdot}}_{Y_{\cdot\cdot}}\right) + \left(\sum_{j=1}^{b}\frac{Y_{\cdot j}^2}{a} - \frac{Y_{\cdot\cdot}}{ab}\sum_{j=1}^{b} Y_{\cdot j}\right)$$

$$\underset{\sim}{Y}'\underset{\sim}{Y} - \hat{\beta}^{*'} X' \underset{\sim}{Y} \overset{\nearrow \text{cálculos}}{=}$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{b} \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}\right)^2$$

$$\therefore \; \hat{\sigma}^2 = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}\right)^2}{(a-1)(b-1)}$$

Teorema

No modelo $Y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij}$, $i = 1, 2 \dots a$ $\quad j = 1, 2 \dots b$

$\epsilon_{ij} \sim N(0, \sigma^2)$ $\quad Y_{ij}$ v.a. independentes

a) $Z_1 = \widehat{\sum c_i \alpha_i} = \sum c_i \bar{Y}_{i.} \sim N\left(\sum c_i \alpha_i, \sigma^2 \dfrac{\sum c_i^2}{b}\right)$ se $\sum c_i = 0$

b) $Z_2 = \widehat{\sum d_j \tau_j} = \sum d_j \bar{Y}_{.j} \sim N\left(\sum d_j \tau_j, \sigma^2 \dfrac{\sum d_j^2}{a}\right)$ se $\sum d_j = 0$

c) $U = \dfrac{(a-1)(b-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(a-1)(b-1)}$

d) U, Z_1 e Z_2 são v.a. independentes

e) $\mathrm{Cov}\left[\widehat{\sum c_i \alpha_i}, \widehat{\sum c_i^* \alpha_i}\right] = \sigma^2 \dfrac{\sum c_i c_i^*}{b}$ se $\sum c_i = \sum c_i^* = 0$

f) $\mathrm{Cov}\left[\widehat{\sum d_j \tau_j}, \widehat{\sum d_j^* \tau_j}\right] = \sigma^2 \sum d_j d_j^*$ $\quad \sum d_j = \sum d_j^* = 0$

Consequência

Os intervalos de confiança com coeficiente $1-\alpha$ para

$\sum c_i \alpha_i$ com $\sum c_i = 0$ e $\sum d_j \tau_j$ com $\sum \tau_j = 0$

são respectivamente

$$\sum c_i \bar{Y}_{i.} \pm t_{\alpha/2, (a-1)(b-1)} \, \hat{\sigma} \sqrt{\frac{\sum c_i^2}{b}}$$

$$\sum d_j \bar{Y}_{.j} \pm t_{\alpha/2, (a-1)(b-1)} \, \hat{\sigma} \sqrt{\frac{\sum d_j^2}{a}}$$

Teste de Hipóteses

$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a$

H_a : Pelo menos uma das igualdades acima não se verifica

A estatística do teste da razão de Verossimilhança generalizada é

$$W_A = \frac{\dfrac{\boxed{\sum \sum^{b} (\bar{Y}_{i.} - \bar{Y}_{..})^2}}{a-1}}{\dfrac{\sum \sum (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2}{(a-1)(b-1)}}$$

Rejeitamos Ho se $W_A \geqslant w$, $P(F_{a-1, (a-1)(b-1)} \geqslant w) = \alpha$

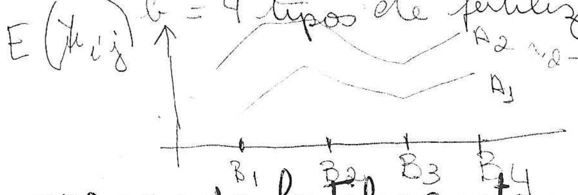
No exemplo, se $H_0: \alpha_1 = \alpha_2$ é aceita temos

$$Y_{ij} = \mu + \alpha + T_j + \varepsilon_{ij}$$

$$E(Y_{ij}) = \mu + \alpha + T_j$$

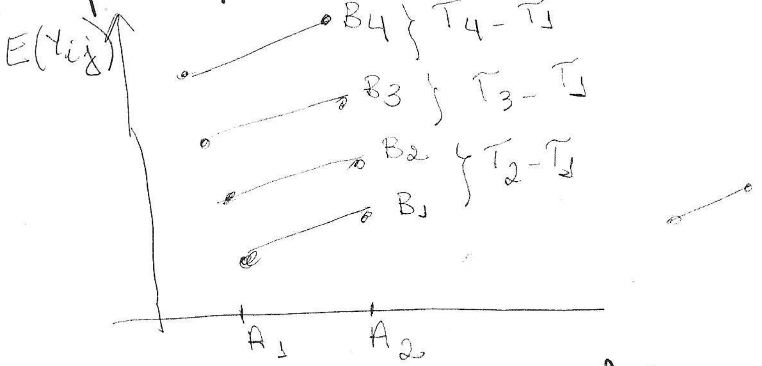
a = 2 tipos de feijão

b = 4 tipos de fertilizante



(a produção média depende apenas do fertilizante.

Fixado o fertilizante, a produção média e idêntica para os dois tipos de feijão.)



$H_0: T_1 = T_2 = \dots = T_b$

H_a: Pelo menos uma das igualdades não se verifica

$$W_B = \frac{\dfrac{\sum\sum(\bar{Y}_{.j} - \bar{Y}_{..})^2}{b-1}}{\dfrac{\sum\sum(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2}{(a-1)(b-1)}}$$

Rejeitamos H_0 se $W_B \geqslant w$, $P(F_{b-1,(a-1)(b-1)} \geqslant w) = \alpha$.

Sob H_0, a produção média depende somente do tipo de feijão

Fixado o tipo de feijão, a produção média é idêntica para os quatro fertilizantes

4.6 - Modelo de Análise de Variância com dois fatores, sem interação, m observações por casela

	B_1	B_2	$\cdots$	B_b
A_1	$Y_{111} \cdots Y_{11m}$	$Y_{121} \cdots Y_{12m}$	$\cdots$	$Y_{1b1} \cdots Y_{1bm}$
A_2	$Y_{211} \cdots Y_{21m}$	$Y_{221} \cdots Y_{22m}$	$\cdots$	$Y_{2b1} \cdots Y_{2bm}$
$\vdots$				
A_a	$Y_{a11} \cdots Y_{a1m}$	$Y_{a21} \cdots Y_{a2m}$	$\cdots$	$Y_{ab1} \quad Y_{abm}$

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \epsilon_{ijk}$$

$i = 1, 2 \ldots a \quad j = 1, 2 \ldots b \quad k = 1, 2 \ldots m \qquad a > 1 \quad b > 1 \quad m > 1$

$\epsilon_{ijk} \sim N(0, \sigma^2)$

Sistema de Equações Normais

$$abm\,\hat{\mu} + bm\sum\hat{\alpha}_i + am\sum\hat{\tau}_j = Y_{\cdots}$$

$$bm\,\hat{\mu} + bm\,\hat{\alpha}_i + m\sum\hat{\tau}_j = Y_{i\cdot\cdot} \qquad i = 1, 2 \ldots a$$

$$am \quad + m\sum\hat{\alpha}_i + am\,\hat{\tau}_j = Y_{\cdot j \cdot} \qquad j = 1, 2 \ldots b$$

$$Y_{\cdots} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m} Y_{ijk} \qquad Y_{i\cdot\cdot} = \sum_{j=1}^{b}\sum_{k=1}^{m} Y_{ijk} \qquad Y_{\cdot j \cdot} = \sum_{i=1}^{a}\sum_{k=1}^{m} Y_{ijk}$$

Teorema

1) O modelo admite $K = a+b-1$ funções estimáveis L.I.

2) $\sum c_i \alpha_i$ é estimável $\iff \sum c_i = 0$ e o estimador NVUVM de

$\sum c_i \alpha_i$ é $\sum c_i \bar{Y}_{i\cdot\cdot}$ $\quad \bar{Y}_{i\cdot\cdot} = \dfrac{\sum_{j=1}^{b}\sum_{k=1}^{m} Y_{ijk}}{bm}$

$\alpha_2 - \alpha_1$ é um contraste: ≠ entre as produções médias dos 2 tipos de feijões fixado o tipo do fertilizante

3) $\sum d_j \tau_j$ é estimável $\iff \sum d_j = 0$ e o estimador NVUVM de

$\sum d_j \tau_j$ é $\sum d_j \bar{Y}_{\cdot j \cdot}$ $\quad \bar{Y}_{\cdot j\cdot} = \dfrac{\sum_{i=1}^{a}\sum_{k=1}^{m} Y_{ijk}}{am}$

4)
$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} m\left(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}\left(Y_{ijk} - \bar{Y}_{ij\cdot}\right)^2}{abm-a-b+1}$$

é NVUVM de σ^2 $\quad \bar{Y}_{ij\cdot} = \dfrac{\sum_{k=1}^{m} Y_{ijk}}{m}$

5) $Z_1 = \sum c_i \bar{Y}_{i\cdot\cdot} \sim N\left(\sum c_i \alpha_i, \ \sigma^2 \dfrac{\sum c_i^2}{mb}\right)$ se $\sum c_i = 0$

(6) $Z_2 = \sum d_j \bar{Y}_{\cdot j\cdot} \sim N\left(\sum d_j \tau_j, \ \dfrac{\sigma^2 \sum d_j^2}{ma}\right)$ se $\sum d_j = 0$

(7) $U = \dfrac{(abm-a-b+1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{abm-a-b+1}$

Teorema

(1) Um intervalo de confiança com coeficiente $1-\alpha$ para $\sum_{i=1}^{a} c_i \alpha_i$ com $\sum c_i = 0$ é

$$\sum_{i=1}^{a} c_i \overline{Y}_{i..} \pm t_{\alpha/2, N} \, \hat{\sigma} \sqrt{\frac{\sum c_i^2}{mb}} \qquad N = abm - a - b + 1$$

(2) analogamente, um intervalo de confiança para $\sum_{j=1}^{b} d_i \tau_j$ com $\sum_{j=1}^{b} d_j = 0$ é

$$\sum_{j=1}^{b} d_j \overline{Y}_{.j.} \pm t_{\alpha/2, N} \, \hat{\sigma} \sqrt{\frac{\sum d_j^2}{ma}} \qquad N = abm - a - b + 1$$

(3) O teste da razão de verossimilhança generalizada para $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_a$ contra H_a: pelo menos uma das igualdades em H_0 não se verifica rejeita H_0 para

$$W = bm \sum_{i=1}^{a} \frac{(\overline{Y}_{i..} - \overline{Y}_{...})^2}{(a-1)\hat{\sigma}^2} \geqslant F_{c, a-1, N}$$

$$N = abm - a - b + 1$$

(4) Para $H_0: \tau_1 = \tau_2 = \dots = \tau_b$, rejeita-se H_0 se

$$W = am \sum_{j=1}^{b} \frac{(\overline{Y}_{.j.} - \overline{Y}_{...})^2}{(b-1)\hat{\sigma}^2} \geqslant F_{c, b-1, N}$$

$$N = abm - a - b + 1$$

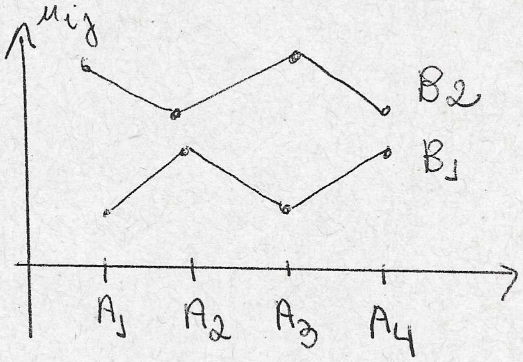
4.7- Modelo de Análise de Variância com dois fatores, interação e $m>1$ observações por casela

$Y_{ijk}=\mu+\alpha_i+\tau_j+\gamma_{ij}+\epsilon_{ijk}$ $\epsilon_{ijk}\sim N(0,\sigma^2)$, independentes,

para $i=1,2\ldots a$ $j=1,2\ldots b$ $k=1,2\ldots m$

$E(Y_{ijk})=\mu_{ij}=\mu+\alpha_i+\tau_j+\gamma_{ij}$

$E(Y_{ij'k})=\mu_{ij'}=\mu+\alpha_i+\tau_{j'}+\gamma_{ij'}$



$\mu_{ij}-\mu_{ij'}=\tau_j-\tau_{j'}+\gamma_{ij}-\gamma_{ij'}$ (depende de i, não é constante para todos os níveis de A).

Dizemos que não há interação se $\gamma_{ij}=0$, $i=1,2\ldots a$, $j=1,2\ldots b$

Para $a=2$ $b=2$

$$X=\begin{array}{c}\\ m\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ m\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ m\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ m\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ \\ \end{array}\begin{array}{ccccccccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ \vdots \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ \vdots \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \vdots \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ \vdots \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ & x_1 & x_2 & y_1 & y_2 & x_1y_1 & x_1y_2 & x_2y_1 & x_2y_2 \end{array}$$

$X_{abm\times(ab+a+b+1)}$

$$\underset{\sim}{\beta}=\begin{bmatrix}\mu\\ \alpha_1\\ \vdots\\ \alpha_a\\ \beta_1\\ \vdots\\ \beta_b\\ \gamma_{11}\\ \vdots\\ \gamma_{1a}\\ \gamma_{21}\\ \vdots\\ \gamma_{ab}\end{bmatrix}$$

Sistema de equações Normais

$$m a b\hat{\mu} + m b\sum_{i=1}^{a}\hat{\alpha}_i + m a\sum_{j=1}^{b}\hat{\tau}_j + m\sum_{i=1}^{a}\sum_{j=1}^{b}\hat{\gamma}_{ij} = Y_{\ldots}$$

I $$m b\hat{\mu} + m b\hat{\alpha}_i + m\sum_{j=1}^{b}\hat{\tau}_j + m\sum_{j=1}^{b}\hat{\gamma}_{ij} = Y_{i\ldots}$$

$i = 1, 2, \ldots, a$

II $$m a\hat{\mu} + m\sum_{i=1}^{a}\hat{\alpha}_i + m a\hat{\tau}_j + m\sum_{i=1}^{a}\hat{\gamma}_{ij} = Y_{.j.}$$

$j = 1, 2, \ldots, b$

III $$m\hat{\mu} + m\hat{\alpha}_i + m\hat{\tau}_j + m\hat{\gamma}_{ij} = Y_{ij.}$$

$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b$

Posto de $X = r(X) = r(X'X) = k$

- $p = ab + b + a + 1 = (a+1)(b+1)$ equações
- Cada a i-ésima equação de I é a soma em j de b equações de III com o mesmo i,
- A j-ésima equação de II é a soma em i de a equações de III com o mesmo j.
- A primeira equação é a soma em i e j das a.b equações do grupo III, Portanto $r(X) \leq ab$.

Cada equação do grupo III possui um diferente elemento γ_{ij}, portanto $r(X'X) \geq ab \Longrightarrow k = r(X) = a.b$

Verifica-se que $\mu + \alpha_i + \tau_j + \gamma_{ij}$ $\quad i=1,2\dots a$, $j=1,2\dots b$

é um conjunto base de funçõs estimáveis

$\hat{\mu}=0 \quad \hat{\alpha}_1=0 \quad \hat{\alpha}_2=0 \dots \quad \hat{\alpha}_a=0 \quad \tau_1=0 \dots \hat{\tau}_b=0$

$\hat{\gamma}_{ij} = \bar{Y}_{ij.} \qquad \bar{Y}_{ij.} = \dfrac{\sum_{k=1}^{m} Y_{ijk}}{m}$ é uma solução

$i=1,2\dots a$

$j=1,2\dots b$

do sistema de equaçõs normais.

Consequência desses fatos e do teorema 3:

a) O estimador NVUVM de $E(Y_{ijk}) = \mu + \alpha_i + \tau_j + \gamma_{ij}$

é $\bar{Y}_{ij.}$ (média da casela).

b) O estimador NVUVM de σ^2 é

$$\hat{\sigma}^2 = \frac{1}{\underbrace{ab(m-1)}_{n - r(X) = abm - ab}} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{m} \left(Y_{ijk} - \bar{Y}_{ij.}\right)^2$$

Os parâmetros $\mu, \alpha_1 \dots \alpha_a, \tau_1, \dots \tau_b, \gamma_{ij}$ não são estimáveis. Por esse motivo, é comum ~~utilizar~~ reescrever o modelo como

$$Y_{ijk} = \mu^* + \alpha_i^* + \tau_j^* + \gamma_{ij}^* + \epsilon_{ijk}$$

com $\mu^* = \bar{\mu}_{..}$ $\alpha_i^* = \bar{\mu}_{i.} - \bar{\mu}_{..}$

$\tau_j^* = \bar{\mu}_{.j} - \bar{\mu}_{..}$ $\gamma_{ij}^* = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}$

e $$\bar{\mu}_{..} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} \mu_{ij}}{a\,b}$$

$$\bar{\mu}_{i.} = \frac{1}{b} \sum_{j=1}^{b} \mu_{ij}$$

$$\bar{\mu}_{.j} = \frac{1}{a} \sum_{i=1}^{a} \mu_{ij}$$

No caso i) $\mu^* + \alpha_i^* + \tau_j^* + \gamma_{ij}^* =$

$= \bar{\mu}_{..} + \bar{\mu}_{i.} - \bar{\mu}_{..} + \bar{\mu}_{.j} - \bar{\mu}_{..} + \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}$

$= \mu_{ij}$

e $\sum_{i=1}^{a} \alpha_i^* = 0$ $\sum_{j=1}^{b} \tau_j^* = 0$ $\sum_{i=1}^{a} \gamma_{ij}^* = 0$ $j = 1, 2, .. b$

$\sum_{j=1}^{b} \gamma_{ij}^* = 0$ $i = 1, 2 ... a$

ii) μ^*, α_i^*, τ_j^* e γ_{ij}^* são estimáveis

iii) O modelo será sem interação $\Longleftrightarrow \gamma_{ij} = 0 \quad \forall i, j$

$$\Longleftrightarrow \sum_{i=1}^{a} \sum_{j=1}^{b} \gamma_{ij}^{*\,2} = 0$$

iv) São testáveis as hipóteses

$$H_0: \sum_{i=1}^{a} \sum_{j=1}^{b} \gamma_{ij}^{*\,2} = 0 \text{ (inexistência de interação)}$$

$$H_a: \sum_{i=1}^{a} \sum_{j=1}^{b} \gamma_{ij}^{*2} \neq 0$$

$$W = \frac{\dfrac{\sum\sum m\left(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...}\right)^2}{(a-1)(b-1)}}{\hat{\sigma}^2}$$

Rejeitamos H_0 se $W > F_{(a-1)(b-1),\, ab(m-1)}$

Se H_0 é aceita, posseguir a análise no modelo sem interação

Se H_0 é rejeitada, constrói-se análises de Variância (com um fator) dentro dos níveis de B e de A.

v) Total de parâmetros μ_{ij}: ab

Se $m=1$, $n=ab$, não é possível estimar σ^2 ($\hat{\sigma}^2 = 0$).

Se não existir interação, temos $a+b-1$ parâmetros e é possível estimar σ^2 mesmo que $m=1$.

4.7- Análise de Variância com dois fatores, modelo hierárquico

Searle pag 249

Os dados a seguir correspondem às notas dadas por alunos relativas à introdução de uma facilidade computacional em diferentes seções de cursos de Inglês e Geologia.

Curso	Seção	Observações	Total	média
Inglês	1	5	5	5
	2	8, 10, 9	27	9
Geologia	1	8, 10	18	9
	2	6, 2, 1, 3	12	3
	3	3, 7	10	5

Observa-se que as seções não são as mesmas nos curso de Inglês e Geologia.

Experimento: 2 fatores, curso e seção, seção é hierárquico dentro do curso.

Modelo

$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \epsilon_{ijk}$ $i = 1, 2 \dots a$

Y_{ijk} - k-ésima observação na j-ésima seção do i-ésimo curso

$i = 1, 2$ $j = 1, 2 \dots b_i$, $b_1 = 2$ e $b_2 = 3$

$k = 1, 2, \dots n_{ij}$ ($n_{11} = 1$ $n_{12} = 3$ $n_{21} = 2$
$n_{22} = 4$ $n_{23} = 2$)

μ - média geral

α_i - efeito do i-ésimo curso

β_{ij} - efeito da j-ésima seção no i-ésimo curso

ou

$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk}$

$$\underset{\sim}{Y} = \begin{bmatrix} Y_{111} \\ Y_{121} \\ Y_{122} \\ Y_{123} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{223} \\ Y_{224} \\ Y_{231} \\ Y_{232} \end{bmatrix}$$

$$\underset{\sim}{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix}_{8 \times 1}$$

$$X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{12 \times 8}$$

Sistema de equações Normais

$$\left[\begin{array}{cccccccc}
n_{..} & n_{1.} & n_{2.} & n_{11} & n_{12} & n_{21} & n_{22} & n_{23} \\
\hline
n_{1.} & n_{1.} & 0 & n_{11} & n_{12} & 0 & 0 & 0 \\
n_{2.} & 0 & n_{2.} & 0 & 0 & n_{21} & n_{22} & n_{23} \\
\hline
n_{11} & n_{11} & 0 & n_{11} & 0 & 0 & 0 & 0 \\
n_{12} & n_{12} & 0 & 0 & n_{12} & 0 & 0 & 0 \\
\hline
n_{21} & 0 & n_{21} & 0 & 0 & n_{21} & 0 & 0 \\
n_{22} & 0 & n_{22} & 0 & 0 & 0 & n_{22} & 0 \\
n_{23} & 0 & n_{23} & 0 & 0 & 0 & 0 & n_{23}
\end{array}\right]_{8\times 8}
\left[\begin{array}{c}
\mu^0 \\
\hline
\alpha_1^0 \\
\alpha_2^0 \\
\hline
\beta_{11}^0 \\
\beta_{12}^0 \\
\hline
\beta_{21}^0 \\
\beta_{22}^0 \\
\beta_{23}^0
\end{array}\right]
=
\left[\begin{array}{c}
Y_{...} \\
\hline
Y_{1..} \\
Y_{2..} \\
\hline
Y_{11.} \\
Y_{12.} \\
\hline
Y_{21.} \\
Y_{22.} \\
Y_{23.}
\end{array}\right]$$

$L2 + L3 = L1$

$L4 + L5 = L2$

$L6 + L7 + L8 = L3$

$r(X'X) = 8 - 3 = 5$

$$X_{n \times (1+a+b_.)}$$

$a = n^{\underline{o}}$ de níveis do fator principa

$$\downarrow$$

$$b_. = \sum_{i=1}^{a} b_i \nearrow \quad p = 1 + a + b_.$$

De modo geral, $r(X'X) = 1 + a + b_. - (1+a) = b_.$

Possível solução do sistema:

$p - r(X'X) = 1+a$ soluções nulas e resolve-se o sistema nos ~~Me~~ restantes

Tomar $\mu^0 = \alpha_1^0 = \dots = \alpha_a^0 = 0$

Obtem-se

$$\beta_{ij}^0 = \bar{Y}_{ij.}$$

$\mu+\alpha_i+\beta_{ij}$ é um conjunto de funções estimáveis LI portanto é uma base de funções estimáveis nesse modelo.

Portanto, o estimador NVUVM de $E(Y_{ijk}) = \mu+\alpha_i+\beta_{ij}$ é $\widehat{\mu+\alpha_i}+\beta_{ij} = \bar{Y}_{ij.}$

$$\hat{\sigma}^2 = \left(\sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} Y_{ijk}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b_i} \frac{Y_{ij.}^2}{n_{ij}} \right) \frac{1}{\underbrace{n-r(X'X)}_{n-\underset{\sum_{i=1}^{a} b_i}{\underset{\|}{b_.}}}}$$

$$E(Y_{ijk}) - E(Y_{ij'k}) = \mu + \alpha_i + \beta_{ij} - \mu - \alpha_i - \beta_{ij'}$$

$= \beta_{ij} - \beta_{ij'}$ $\quad \forall i,\ j \neq j'$ são estimáveis

Portanto

$H_0: \beta_{i1} = \beta_{i2} = \dots = \beta_{ib_i}$ $\quad i = 1, 2 \dots a$ é testável

Estatística de Teste

$$W = \frac{\displaystyle\sum_{i=1}^{a}\sum_{j=1}^{b_i} \frac{Y_{ij.}^2}{n_{ij}} - \sum_{i=1}^{a} \frac{Y_{i..}^2}{n_{i.}}}{(b_. - a)\,\hat{\sigma}^2}$$

$\hookrightarrow \sum_{i=1}^{a} (b_i - 1) =$ nº de hipóteses a serem testadas

Rejeita-se Ho se $W \geq F_{c,\, b_. - a,\, n - b_.}$

Verifica-se que $\alpha_i - \alpha_{i'}$ não são estimáveis. Portanto, testes de hipóteses só são possíveis sob restrições.