

ESTIMAÇÃO DE COMPONENTES DE VARIÂNCIA

- OS MÉTODOS DE HENDERSON -

LUIZ CARLOS BASSO

DISSERTAÇÃO APRESENTADA

AO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DA

UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE MESTRE

EM

ESTATÍSTICA

ORIENTADOR:

PROF. DR. ADOLPHO WALTER PIMAZONI CANTON

“ SÃO PAULO, OUTUBRO DE 1981 -

À  
meus Pais  
e  
À  
Sílvia

## AGRADECIMENTOS

Queremos aqui registrar nossos sinceros agradecimentos a todos os que direta ou indiretamente colaboraram para a concretização desse trabalho, e em particular

- ao Prof. Dr. Adolpho Walter P. Canton, nosso professor durante a pós-graduação, orientador e amigo, pelo trabalho de orientação e pela ajuda constante face às numerosas dificuldades com que deparamos durante a execução deste trabalho;
- ao Prof. Dr. Clóvis de Araujo Peres, pela discussão e valiosas sugestões apresentadas;
- à minha esposa Sílvia Edna Sousa, pelas sugestões e principalmente pela compreensão e colaboração de todas as horas;
- aos colegas e professores do IME-USP, pelo convívio e pelo aprendizado (formal ou informal);
- às bibliotecárias do IME-USP, pela gentileza e atenção, e em especial a Edmari G. Teixeira que muito colaborou no trabalho bibliográfico;
- à Dna. Selma Lemos de Melo secretária da C.P.G. do IME-USP pela atenção e solicitude desde o primeiro momento em que recorremos à secretaria da C.P.G.;
- à Regina Helena da Silva pelo paciente e meticoloso trabalho de datilografia.

## SUMÁRIO

	Página
CAPITULO 1 - INTRODUÇÃO	
CAPÍTULO 2 - A ANÁLISE DE VARIÂNCIA	
2.1 - O MODELO E AS SUPOSIÇÕES . . . . .	4
2.2 - AS EQUAÇÕES NORMAIS E AS SOMAS DE QUADRADOS . . . . .	6
2.3 - OS MODELOS I, II E III NA ANÁLISE DE VARIÂNCIA . . . . .	8
2.3.1 - Modelo de Fatores Fixos ou Modelo I . . . . .	8
2.3.2 - Modelo de Fatores Aleatórios ou Modelo II. . . . .	8
2.3.3 - Modelo III - Modelo Misto . . . . .	10
2.4 - REDUÇÕES NAS SOMAS DE QUADRADOS - A NOTAÇÃO $R(\ )$ . . . . .	10
2.5 - DADOS BALANCEADOS E DADOS NÃO BALANCEADOS. . . . .	11
2.6 - A ESPERANÇA DE FORMAS QUADRÁTICAS - ALGUNS TÓPICOS . . . . .	13
CAPÍTULO 3 - A ESTIMAÇÃO PONTUAL DOS COMPONENTES DE VARIÂNCIA	
3.1 - DADOS BALANCEADOS . . . . .	16
3.1.1 - O Método da Análise de Variância . . . . .	16
3.1.2 - Vantagens e Desvantagens do Método de Análise de Variância . . . . .	18
3.1.3 - O Método de Máxima Verossimilhança (Maximum Likelihood). . . . .	19

	Página
3.2 - DADOS NÃO-BALANCEADOS . . . . .	22
3.2.1 - Método 1 de Henderson ou Método da Análise de Variância. . . . .	23
3.2.2 - Método 2 de Henderson . . . . .	24
3.2.3 - Método 3 de Henderson - Método de Ajustar Constantes. . . . .	24
3.2.4 - Variações sobre o Método da Análise de Variância . . . . .	25
3.2.5 - Método de Máxima Verossimilhança. . . . .	26
3.3 - AS PROPRIEDADES DESEJÁVEIS DOS ESTIMADORES. . . . .	27
3.3.1 - Estimador Quadrático Não-Viesado de Variância Mínima. . . . .	27
3.3.2 - Viés. . . . .	27
 CAPÍTULO 4 - OS MÉTODOS 1, 2 E 3 DE HENDERSON	
4.1 - O MÉTODO 1 DE HENDERSON. . . . .	28
4.1.1 - Somas de Quadrados . . . . .	29
4.1.2 - As Esperanças das S.Q. . . . .	30
4.1.3 - Um Exemplo Numérico . . . . .	35
4.1.4 - Método de Síntese de Hartley. . . . .	39
4.2 - O MÉTODO 2. . . . .	41
4.2.1 - Método 2 Generalizado. . . . .	42
4.2.2 - Método 2 Simplificado. . . . .	43
4.2.2.1 - A Invariância do Método . . . . .	46
4.2.2.2 - O Método 2 de Henderson . . . . .	46
4.2.3 - Exemplo Numérico . . . . .	49

	Página
4.3 - O MÉTODO 3 DE HENDERSON - MÉTODO DE AJUSTAR CONSTANTES. . . . .	52
4.3.1 - O Método . . . . .	52
4.3.2 - A Aplicação do Método 3 . . . . .	55
4.3.2.1 - Exemplo Numérico. . . . .	56
4.3.2.2 - Um Tratamento Matricial . . . . .	60
4.3.3 - Vantagens e desvantagens do Método 3 . . . . .	61
CONCLUSÃO . . . . .	64
BIBLIOGRAFIA. . . . .	66

## CAPÍTULO 1

### INTRODUÇÃO

No modelo linear geral

$$\underline{y} = \underline{X}\underline{b} + \underline{e} \quad (1.1)$$

os efeitos dos fatores, isto é, os elementos de  $\underline{b}$  podem ser tratados *quantitativa* ou *qualitativamente*. No primeiro caso, quando os elementos de  $\underline{X}$  podem assumir qualquer valor real, dizemos que temos *modelo de análise de regressão linear*. No segundo caso, quando cada elemento de  $\underline{X}$  assume o valor zero ou o valor um, o modelo é chamado *modelo de análise de variância*. Quando alguns efeitos dos fatores são tratados quantitativamente e outros qualitativamente, o modelo é chamado *modelo de análise de covariância* (Scheffé [1976]).

Neste trabalho usaremos modelos de análise de variância com os efeitos dos fatores explicitados (ex.:  $y_{ij} = \mu + \alpha_i + e_{ij}$ ), resultando em modelos de posto incompleto. Alguns autores preferem trabalhar com modelos reparametrizados (ex.:  $y_{ij} = \mu_i + e_{ij}$ ), obtendo, assim, posto completo.

Desde que Fisher desenvolveu a técnica de Análise de

Variância, na década de 20, o processo ganhou grande popularidade na comunidade científica e os procedimentos baseados na tabela da "anova" foram muito utilizados. Entretanto, quase sempre o interesse era inferir sobre o valor dos parâmetros, e relativamente pouca atenção era dedicada a uma outra classe de problemas: a inferência sobre componentes de variância.

Um perfil dessas duas classes de problemas é visto no capítulo 2 desse trabalho, juntamente com as suposições necessárias em cada caso. A importância dessas suposições e as condições em que podem ser dispensadas foram analisadas por Eisenhart [1947], que introduziu a linguagem: modelo I, modelo II, modelo III.

Apesar da grande quantidade de trabalhos hoje existentes em análise de variância, e do grande refinamento matemático alcançado, reina certa discordância quanto à análise estatística adequada ao experimento, quando os dados são não balanceados, ou seja, quando o número de observações não é o mesmo em todas as caselas.

No caso de balanceamento há um consenso geral sobre a "anova" adequada, sobre quais hipóteses estão sendo testadas, sobre as propriedades de cada método (momentos, mínimos quadrados, etc.) e assim por diante. No caso de não-balanceamento há dificuldades iniciais relacionadas à distribuição de certas formas quadráticas, e também dificuldades provenientes do próprio fato do número desigual de observações por casela, que ocasiona uma complexidade algébrica, por vezes proibitiva. Além disso, há na literatura muita confusão com respeito aos métodos adequados a cada caso. Um exemplo dessa situação, para o particular caso de dois fatores com interação e modelo fixo, está em Speed, Hocking e Hackney [1978].

Em particular, no caso de modelos de componentes de variância - modelos aleatórios - o não balanceamento restringe, e mesmo impede, o uso de certos métodos de estimação (máxima ve

rossimilhança, por exemplo).

É nesse contexto que se situa o tema deste trabalho : os métodos de Henderson para a estimação dos componentes de variância. Estes métodos têm sido muito utilizados desde sua publicação em 1953, e apesar de certas deficiências, talvez sejam ainda a melhor opção em termos de estimação pontual de componentes de variância com dados não balanceados.

O plano deste trabalho é o seguinte:

- no capítulo 2 são apresentados alguns conceitos fundamentais em análise de variância, definidos modelos fixos e aleatórios, estabelecidas as suposições dos modelos, definida a notação  $R(\quad)$ , e discutida a situação de balanceamento e não balanceamento.

- no capítulo 3 é analisada a estimação dos componentes de variância, sendo dada uma descrição dos métodos existentes tanto no caso balanceado como no não balanceado e são discutidas as propriedades desejáveis dos estimadores.

O capítulo 4 analisa em detalhe os métodos 1, 2 e 3 de Henderson para estimação dos componentes de variância no caso não-balanceado. Aplicações e exemplos acompanham a exposição.

## CAPÍTULO 2

### A ANÁLISE DE VARIÂNCIA

O objetivo desse capítulo é definir o modelo que usaremos, suas suposições e principalmente a diferença entre modelo fixo (I) e modelo com componentes de variância (II e III). A questão de balanceamento e não balanceamento dos dados é também introduzida, e um importante teorema sobre a esperança de formas quadráticas é aplicado a cada um dos três modelos (I, II e III).

#### 2.1 - O MODELO E AS SUPOSIÇÕES

O modelo com o qual trabalharemos é

$$\underline{y} = \underline{X}\underline{b} + \underline{e} \quad (2.1)$$

onde  $\underline{y}$  é o vetor  $N \times 1$  de observações,  $\underline{b}$  é o vetor  $p \times 1$  de parâmetros,  $\underline{X}$  é uma matriz  $N \times p$  de posto incompleto, formada por zeros e uns, chamada matriz de planejamento, e  $\underline{e}$  um vetor  $N \times 1$  de erros (aleatórios).

São suposições do modelo:

a)  $E(\underline{e}) = \underline{0}$

b)  $\text{var}(\underline{e}) = E(\underline{e}\underline{e}') = \sigma_e^2 \underline{I}_N$ , ou seja, cada elemento de  $\underline{e}$  tem variância  $\sigma_e^2$  e covariância nula com qualquer outro elemento.

Assim,  $\underline{e} \sim (0, \sigma_e^2 \underline{I}_N)$

Essas suposições permitem estimarmos algumas funções dos parâmetros, que são os elementos de  $\underline{b}$ . Uma função linear  $\underline{q}'\underline{b}$  é definida como estimável se, e somente se,  $\underline{q}'\underline{b} = \underline{t}'E(\underline{y})$ , para algum vetor  $\underline{t}$ . Basicamente, a implicação dessa definição é que a função estimável é uma função linear dos parâmetros para a qual pode ser encontrado um estimador não viesado.

Além disso, outra suposição que em geral é colocada no modelo, é:

c) Os erros tem distribuição normal, isto é:

$$\underline{e} \sim N(\underline{0}, \sigma_e^2 \underline{I}_N)$$

Com essas 3 suposições todos os procedimentos da análise de variância para estimação, testes e intervalos de confiança relacionados com os parâmetros podem ser aplicados. Essas suposições foram explicitamente formalizadas em Eisenhart [1947] que analisa detalhadamente possíveis situações em que alguma ou algumas das suposições não são válidas, e também as suposições no caso em que o interesse não está nos elementos de  $\underline{b}$ , mas sim nos componentes de variância.

Antes de estabelecer as equações normais, vejamos a explicitação do modelo da equação (2.1), através de um exemplo, de modelo de um fator ( $\alpha$ ) com 3 níveis e 2 ou 3 observações em cada nível, onde

$$\begin{aligned}
 Y_{11} &= \mu + \alpha_1 && + e_{11} \\
 Y_{12} &= \mu + \alpha_1 && + e_{12} \\
 Y_{21} &= \mu &+ \alpha_2 & + e_{21} \\
 Y_{22} &= \mu &+ \alpha_2 & + e_{22} \\
 Y_{31} &= \mu && + \alpha_3 + e_{31} \\
 Y_{32} &= \mu && + \alpha_3 + e_{32} \\
 Y_{33} &= \mu && + \alpha_3 + e_{33}
 \end{aligned}$$

são as equações obtidas aplicando-se o modelo aos dados:

níveis de $\alpha$	$\alpha_1$	$\alpha_2$	$\alpha_3$
	$Y_{11}$	$Y_{21}$	$Y_{31}$
observações	$Y_{12}$	$Y_{22}$	$Y_{32}$
			$Y_{33}$

Essas equações escritas na forma matricial produzem:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{31} \\ Y_{32} \\ Y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \\ e_{33} \end{bmatrix}$$

## 2.2 - AS EQUAÇÕES NORMAIS E AS SOMAS DE QUADRADOS

As equações normais do modelo obtidas pelo método de mínimos quadrados, Searle [1971], são:

$$\underline{\underline{X}}' \underline{\underline{X}} \hat{\underline{\underline{b}}} = \underline{\underline{X}}' \underline{\underline{y}} \quad (2.2)$$

e no caso de modelo de posto incompleto a solução  $\hat{\underline{\underline{b}}}$  não é única, sendo então representada por  $\underline{\underline{b}}^0$ . Dessa forma, escrevemos

$$\underline{\underline{X}}' \underline{\underline{X}} \underline{\underline{b}}^0 = \underline{\underline{X}}' \underline{\underline{y}} \quad (2.3)$$

e a solução é

$$\underline{\underline{b}}^0 = \underline{\underline{G}} \underline{\underline{X}}' \underline{\underline{y}} \quad , \quad (2.4)$$

onde  $\underline{\underline{G}}$  é uma inversa generalizada de  $\underline{\underline{X}}' \underline{\underline{X}}$ .

O valor estimado de  $\underline{\underline{y}}$ , correspondente ao observado é

$$\hat{\underline{\underline{y}}} = \underline{\underline{X}} \underline{\underline{G}} \underline{\underline{X}}' \underline{\underline{y}} \quad (2.5)$$

A soma de Quadrados Residual é

$$\begin{aligned} \text{SQE} &= \underline{\underline{y}}' \underline{\underline{y}} - \underline{\underline{b}}^0' \underline{\underline{X}}' \underline{\underline{y}} \\ &= \underline{\underline{y}}' \underline{\underline{y}} - \underline{\underline{y}}' \underline{\underline{X}} \underline{\underline{G}} \underline{\underline{X}}' \underline{\underline{y}} \end{aligned} \quad (2.6)$$

A Soma de Quadrados Total é

$$\text{SQT} = \underline{\underline{y}}' \underline{\underline{y}} = \sum y^2 \quad (2.7)$$

A Soma de Quadrados devido a ajustar o modelo é

$$\text{SQR} = \underline{\underline{b}}^0 \underline{\underline{X}}' \underline{\underline{y}} = \underline{\underline{y}}' \underline{\underline{X}} \underline{\underline{G}} \underline{\underline{X}}' \underline{\underline{y}} \quad (2.8)$$

A Soma de Quadrados devido a ajustar a média é

$$\text{SQM} = N \bar{y}^{-2} \quad (2.9)$$

As Somas de Quadrados corrigidas pela média são

$$SQR_M = SQR - SQM \quad (2.10)$$

$$SQT_M = SQT - SQM \quad (2.11)$$

## 2.3 OS MODELOS I II e III NA ANÁLISE DE VARIÂNCIA

### 2.3.1 - Modelo de Fatores Fixos ou Modelo I

Dizemos que um fator é fixo se os níveis, desse fator, que estão sendo considerados no experimento, são exatamente aqueles de interesse. Por exemplo, se o fator é fertilizante e estamos interessados especificamente em comparar nitrogênio, potássio e fósforo quanto à produtividade de certo cereal, então esse fator é fixo. O modelo aqui pode ser escrito explicitamente como:

$$y_{ij} = \mu + A_i + e_{ij} ,$$

onde  $y_{ij}$  é a  $j$ -ésima observação do  $i$ -ésimo nível (tratamento) do fator A,  $\mu$  a média,  $A_i$  o efeito do  $i$ -ésimo tratamento e  $e_{ij}$  o erro associado a  $y_{ij}$ . As suposições desse modelo são as enumeradas em 2.1.

Nosso interesse no caso de modelo fixo está na comparação dos efeitos dos fatores e no estudo das relações entre eles.

### 2.3.2 - Modelo de Fatores Aleatórios ou Modelo II

Um fator é chamado aleatório quando os níveis desse fator que estão sendo considerados representam uma amostra aleatória de todos os possíveis níveis. O interesse não é nos efei-

tos dos fatores, mas sim na sua variabilidade. Por exemplo, num experimento sobre a produtividade de certa máquina, escolhemos 5 operários e observamos a produção diária de cada um. Cada operário representa um nível do fator e não estamos interessados em comparar ou verificar relações entre os efeitos desses particulares níveis. Iremos apenas verificar como o fator operário contribui para a variância total da produção.

O modelo pode ser escrito como:

$$y_{ij} = \mu + \alpha_i + e_{ij}$$

onde  $y_{ij}$  é a produção do  $i$ -ésimo operário no  $j$ -ésimo dia,  $\mu$  é a média,  $\alpha_i$  o efeito do  $i$ -ésimo operário sobre a produção ( $i$ -ésimo nível do fator  $\alpha$ ) e  $e_{ij}$  o erro associado a  $y_{ij}$ .

As suposições do modelo são:

$$e_{ij} \sim (0, \sigma_e^2)$$

$$\alpha_i \sim (0, \sigma_\alpha^2)$$

$$E(\alpha_i \alpha_j) = 0, \text{ para } i \neq j$$

e

$$E(\alpha_i e_{i',j'}) = 0 \quad \forall i, i', j'$$

Com isso

$$\text{cov}(y_{ij}, y_{i',j'}) = \begin{cases} \sigma_\alpha^2 + \sigma_e^2, & i = i', j = j' \\ \sigma_\alpha^2, & i = i', j \neq j' \\ 0, & i \neq i' \end{cases}$$

Assim,  $\sigma_\alpha^2$  representa o componente de variância devido ao fator operário.

### 2.3.3 - Modelo Misto ou Modelo III

O modelo é chamado misto quando alguns dos fatores são fixos e outros aleatórios. No caso de dados não balanceados a estimação dos componentes de variância pelo método tradicional da "ANOVA" (3.1.1) não é satisfatória pois produz estimadores viesados, como veremos adiante.

### 2.4 - REDUÇÕES NAS SOMAS DE QUADRADOS - A NOTAÇÃO $R(\ )$

As somas de quadrados devido a ajustar um ou mais fatores são reduções na soma total de quadrados (SQT), e fornecem uma medida da adequação do modelo que está sendo usado. Encarando cada soma de quadrados como uma redução na SQT, podemos trabalhar com as diversas reduções, cuja notação é  $R(\ )$ , indicando entre os parênteses os parâmetros do modelo que está sendo ajustado. Assim, por exemplo no modelo  $y_{ij} = \mu + \alpha_i + e_{ij}$  a redução correspondente a ajustar o modelo é  $R(\mu, \alpha)$  e no modelo  $Y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$  é  $R(\mu, \alpha, \beta)$ . A redução devido à média, SQM é  $R(\mu)$ , que é correspondente a ajustar o modelo  $y_i = \mu + e_i$ .

A equação (2.10) pode ser escrita como:

$$SQR_M = SQR - SQM = R(\mu, \alpha) - R(\mu).$$

Sob o ponto de vista de reduções,  $SQR_M$  pode ser considerada como a redução devido a ajustar um modelo contendo  $\mu$  e  $\alpha$  após ter sido ajustado um modelo contendo  $\alpha$ , e assim uma notação adequada é:  $R(\alpha/\mu)$ , ou seja:

$$R(\alpha/\mu) = R(\mu, \alpha) - R(\mu)$$

$$\text{Analogamente, } R(\beta/\mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha).$$

No caso de fator hierárquico a notação é  $R(\mu, \alpha, \beta:\alpha)$

indicando que  $\beta$  é hierárquico dentro de  $\alpha$ .

Ainda, no modelo  $\underline{y} = \underline{X}_1 \underline{b}_1 + \underline{X}_2 \underline{b}_2 + \underline{e}$  a redução devido a ajustá-lo é  $R(\underline{b}_1, \underline{b}_2)$ ; a redução devido a ajustar o sub-modelo  $\underline{y} = \underline{X}_1 \underline{b}_1 + \underline{e}$  é  $R(\underline{b}_1)$ , e  $R(\underline{b}_2/\underline{b}_1)$  é a redução devido a ajustar o modelo contendo  $\underline{b}_1$  e  $\underline{b}_2$  após ter ajustado o modelo contendo  $\underline{b}_1$ , ou seja

$$R(\underline{b}_2/\underline{b}_1) = R(\underline{b}_1, \underline{b}_2) - R(\underline{b}_1).$$

Essa notação tem a vantagem de ser sucinta, indicando claramente qual modelo está sendo ajustado, e é particularmente útil em análise dos componentes de variância no caso de dados não balanceados e modelo misto, onde precisamos ajustar modelos e sub-modelos para obter estimadores (capítulo 4). Cumpre salientar que há 2 diferentes maneiras de calcular os termos  $R(\ )$ , uma no caso nosso de modelos de posto incompleto e outra no caso de modelo reparametrizado para modelo de posto completo, como em Harvey [1968], Overall e Klett [1972] e Carlson e Timm [1974].

A diferença entre os dois enfoques é apresentada em Speed e Hocking [1976], onde há também uma colocação das limitações do uso da notação  $R(\ )$ , ou seja, que ela não deixa claro qual hipótese está sendo testada, e que não pode ser usada para testar uma hipótese linear geral. Entretanto, no caso de estimação de componentes de variância não há problemas em usar esta notação, e ela, será bastante útil no método 3 de Henderson.

## 2.5 DADOS BALANCEADOS E DADOS NÃO BALANCEADOS

Quando o experimento com que estamos lidando é planejado e realizado com dados balanceados, a análise do experimento é grandemente simplificada, sob vários aspectos.

No caso de estimação dos componentes de variância (mo

delos II e III), temos uma única "anova", e os estimadores obtidos pelo método de análise de variância (que consistem em soluções de um conjunto de equações obtidas igualando-se cada quadrado médio à sua esperança) são bem definidos e possuem as propriedades desejáveis dos estimadores (capítulo 3). Isso não acontece no caso de dados não balanceados, quando então, o conjunto de equações não é único (por exemplo, no modelo 2 fatores temos 2 "anovas" uma correspondendo a ajustar  $\alpha$  antes de  $\beta$  e outra,  $\beta$  antes de  $\alpha$ ). Além disso, o método da análise de variância no modelo III produz estimadores viesados pelos efeitos fixos.

No caso de testes de hipóteses para os componentes de variância, a situação não é simples, pois mesmo no caso balanceado, nem sempre há testes exatos. Assim, nos modelos de um fator, ou de dois fatores cruzados há testes exatos, e nos de 3 ou mais fatores não há testes para alguns dos efeitos (principais, no modelo 3 fatores), mas existe uma aproximação, (Scheffé [1959]) baseada em combinação linear de quadrados médios, desenvolvida primeiramente por Satterthwaite [1946], e que se tornou comum na literatura (Anderson [1960], Eisen [1966]; uma discussão dos critérios para seu uso está em Gaylor e Hoper [1969]). No caso não balanceado, simplesmente não há testes ótimos, nem mesmo no modelo de 1 fator: "A SQ (ponderada) entre os níveis do fator não tem distribuição proporcional a  $\chi^2$ , quaisquer que sejam os pesos usados". (Scheffé [1959]). O problema é que a teoria da distribuição se torna muito complicada.

Nos modelos hierárquicos não há problemas. Nos não balanceados com 2 fatores, idem, mas para 3 ou mais fatores não há testes exatos para todos os fatores. Para os modelos 3 fatores Tietjen [1974], faz uma análise, concluindo que o teste convencional F é uma aproximação razoável, embora os quadrados médios da análise de variância não sejam variáveis com distribuição  $\chi^2$  e independentes.

Estimativas das perturbações no nível de significância

devidas a essa aproximação, são apresentadas por Cummings e Gaylor [1974] , mostrando que o nível de significância é pouco afetado para uma grande variedade de razões entre componentes de variância em situações de acentuado não balanceamento. Notemos que para construção de intervalos de confiança a dificuldade é a mesma, já que testes de hipóteses e I.C. estão estreitamente relacionados.

No caso do modelo I, observemos que a situação também se complica quando os dados não são balanceados.

Existe na literatura, um grande número de métodos de análise, sugerindo diversas hipóteses a serem testadas, com as correspondentes somas de quadrados. Uma visão desses métodos com correspondentes programas de computação para modelos de dois fatores está em Speed, Hocking e Hackney [1978] , e uma abordagem diferente, também para modelo 2 fatores, classificando os métodos em dois grupos, método de soma de quadrados aditivas e método de mínimos quadrados, está em Gosslee e Lucas [1965] .

## 2.6 A ESPERANÇA DE FORMAS QUADRÁTICAS - ALGUNS TÓPICOS

Apresentamos aqui o resultado de um teorema, que nos permite obter a esperança de formas quadráticas, acompanhado da correspondente aplicação a cada um dos três modelos.

Teorema - Quando  $\underline{y} \sim (\underline{\mu}, \underline{V})$  ,

$$E(\underline{y}' \underline{Q} \underline{y}) = \text{tr}(\underline{Q} \underline{V}) + \underline{\mu}' \underline{Q} \underline{\mu} . \quad (2.12)$$

A aplicação do teorema a cada um dos 3 modelos, fixo, aleatório ou misto, é:

## a) Modelo Fixo

Aqui  $\underline{b}$  é um vetor de efeitos fixos com  $E(\underline{y}) = \underline{X}\underline{b}$  e

$$\underline{v} = \sigma_e^2 \underline{I}_N$$

$$E(\underline{y}' \underline{Q} \underline{y}) = \underline{b}' \underline{X}' \underline{Q} \underline{X} \underline{b} + \sigma_e^2 \text{tr}(\underline{Q}) . \quad (2.13)$$

Em particular, quando  $\underline{Q} = \underline{I}_N$ , temos

$$E(\underline{y}' \underline{y}) = \underline{b}' \underline{X}' \underline{X} \underline{b} + N \sigma_e^2 \quad (2.14)$$

e quando  $\underline{Q} = \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'$ , então

$$\begin{aligned} E(\underline{y}' \underline{Q} \underline{y}) &= E[R(\underline{b})] = \underline{b}' \underline{X}' \underline{X} \underline{b} + \sigma_e^2 \text{tr}[\underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'] \\ &= \underline{b}' \underline{X}' \underline{X} \underline{b} + \sigma_e^2 r(\underline{X}) \end{aligned} \quad (2.15)$$

## b) Modelo Misto

Aqui  $\underline{b}$ , que contém os efeitos dos fatores fixos e dos fatores aleatórios pode ser particionado como  $\underline{b}' = [\underline{b}'_1 \ \underline{b}'_A \ \underline{b}'_B \ \dots \ \underline{b}'_K]$  onde  $\underline{b}'_1$  representa os fatores fixos, incluindo  $\mu$ ,  $\underline{b}'_i$  é o vetor dos efeitos do fator aleatório  $i$ , ( $i = A, \dots, K$ ), e desse modo o modelo (2.1) pode ser escrito como

$$\underline{y} = \underline{X}_1 \underline{b}_1 + \sum_{i=A}^k \underline{X}_i \underline{b}_i + \underline{e} \quad (2.16)$$

$$\text{onde } E(\underline{y}) = \underline{X}_1 \underline{b}_1 \quad (2.17)$$

$$\text{e } \text{var}(\underline{y}) = \underline{V} = \sum_{i=A}^k \underline{X}_i \text{var}(\underline{b}_i) \underline{X}_i' + \sigma_e^2 \underline{I}_N \quad (2.18)$$

De acordo com as suposições do modelo, os efeitos são não correlacionados, e

$$\text{var}(\underline{b}_i) = \sigma_i^2 \underline{I}_{N_i} .$$

$$\text{Assim, } \underline{\underline{V}} = \sum_{i=A}^k \underline{\underline{X}}_i \underline{\underline{X}}_i' \sigma_i^2 + \sigma_e^2 \underline{\underline{I}}_N . \quad (2.19)$$

Portanto

$$E(\underline{\underline{y}}' \underline{\underline{Q}} \underline{\underline{y}}) = (\underline{\underline{X}}_j \underline{\underline{b}}_j)' \underline{\underline{Q}} \underline{\underline{X}}_l \underline{\underline{b}}_l + \sum_{i=A}^k \sigma_i^2 \text{tr}(\underline{\underline{Q}} \underline{\underline{X}}_i \underline{\underline{X}}_i') + \sigma_e^2 \text{tr}(\underline{\underline{Q}}) \quad (2.20)$$

### c) Modelo Aleatório

No modelo aleatório  $\underline{\underline{b}}_1$  corresponde ao escalar  $\mu$  e  $\underline{\underline{X}}_1$  ao vetor  $\underline{\underline{1}}$ , logo,

$$E(\underline{\underline{y}}' \underline{\underline{Q}} \underline{\underline{y}}) = \mu^2 \underline{\underline{1}}' \underline{\underline{Q}} \underline{\underline{1}} + \sum_{i=A}^k \sigma_i^2 \text{tr}(\underline{\underline{Q}} \underline{\underline{X}}_i \underline{\underline{X}}_i') + \sigma_e^2 \text{tr}(\underline{\underline{Q}}) \quad (2.21)$$

Nos capítulos subsequentes, referir-nos-emos com frequência a esses resultados, pois a estimação dos componentes de variância será baseada nas esperanças dos quadrados médios, que são formas quadráticas.

## CAPÍTULO 3

### A ESTIMAÇÃO PONTUAL DOS COMPONENTES DE VARIÂNCIA

Este capítulo é basicamente dividido em duas partes : caso balanceado e caso não balanceado. Em cada um dos casos são apresentados os métodos existentes, com suas vantagens, desvantagens e os problemas na aplicabilidade.

Além disso, são discutidas as propriedades desejáveis dos estimadores, e as limitações existentes na estimação.

#### 3.1 - DADOS BALANCEADOS

##### 3.1.1 - O Método da Análise de Variância

Quando há balanceamento, a estimação dos componentes de variância é feita usualmente, pelo método da "Anova" que consiste em obter um sistema de equações lineares, igualando-se cada quadrado médio a seu valor esperado, e as soluções desse sistema são os estimadores dos componentes de variância.

Para a obtenção dos quadrados médios (QM), que são so

mas de quadrados (SQ) divididos pelos graus de liberdade, usa-se a mesma "anova" do modelo I (portanto, esses quadrados médios são baseados nos métodos de mínimos quadrados) e para a obtenção das E(QM) pode-se usar o teorema de 2.6 (pois os QM são formas quadráticas nas observações), ou calcular diretamente a esperança substituindo-se nas fórmulas dos QM as várias médias das observações ( $\bar{y} \dots, \bar{y}_i \dots, \bar{y}_{i.}$ , etc) escritas em função do modelo considerado (Searle [1971]).

Além disso as E(QM) podem ser também obtidas pelas "regras" apropriadas que são fornecidas nos livros de planejamento de experimentos, como por exemplo, Searle [1971], Scheffé [1959] ou John [1971].

Como ilustração do método vejamos a estimação no modelo aleatório de 1 fator com  $a$  níveis e  $n$  observações por casela, onde

$$y_{ij} = \mu + \alpha_i + e_{ij} \quad (3.1)$$

e

$$\begin{aligned} E(QMA) &= n \sigma_{\alpha}^2 + \sigma_e^2 \\ E(QME) &= \sigma_e^2 \end{aligned} \quad (3.2)$$

e o conjunto de equações é

$$QME = \hat{\sigma}_e^2 \quad e \quad QMA = n \hat{\sigma}_{\alpha}^2 + \hat{\sigma}_e^2 \quad (3.3)$$

dando as soluções

$$\begin{aligned} \hat{\sigma}_e^2 &= QME \\ \hat{\sigma}_{\alpha}^2 &= (QMA - QME)/n \end{aligned} \quad (3.4)$$

Nos modelos de 2 ou mais fatores, seja o modelo misto ou aleatório, o procedimento é análogo e o método produz estimadores não-viesados e de variância mínima.

### 3.1.2 - Vantagens e Desvantagens do Método da Análise de Variância

O método é, em termos computacionais, relativamente simples e produz estimadores não-viesados:

Seja  $\underline{m}$  o vetor dos QM tais que  $E(\underline{m})$  não envolva efeitos fixos e seja  $\underline{\sigma}^2$  o vetor de componentes de variância a ser estimado, com  $E(\underline{m}) = \underline{P} \underline{\sigma}^2$ , sendo  $\underline{P}$  não-singular. Então as equações a serem resolvidas são:  $\underline{m} = \underline{P} \underline{\sigma}^2$ , com solução:

$$\hat{\underline{\sigma}}^2 = \underline{P}^{-1} \underline{m} \quad (3.5)$$

$$\text{logo } E(\hat{\underline{\sigma}}^2) = \underline{P}^{-1} E(\underline{m}) = \underline{P}^{-1} \underline{P} \underline{\sigma}^2 = \underline{\sigma}^2 . \quad (3.6)$$

Além disso os estimadores de (3.5) são os que tem menor variância dentre todos os estimadores não viesados que são formas quadráticas nas observações. (Graybill e Hultquist [1961]) Ou seja, os estimadores da "Anova" são estimadores "BQUE" (best quadratic unbiased est. - melhor estimador quadrático não-viesado), que é o análogo ao "b.l.u.e." (do modelo I) para os componentes de variância. E, ainda, de acordo com Graybill [1954] e Graybill e Wortham [1956], sob suposições de normalidade, os estimadores de (3.5) tem menor variância dentre todos os estimadores não-viesados, sejam ou não formas quadráticas nas observações.

A deficiência do método é que os estimadores de (3.5) podem ser negativos, o que é evidentemente um problema sério. As sugestões para contornar o problema quando obtemos estimativas negativas são:

a) Aceitar o valor negativo. Entretanto, é claro que o problema persiste, especialmente se precisamos obter uma soma estimada de componentes envolvendo a estimativa negativa.

b) Supor que o verdadeiro valor é zero. Embora essa

seja uma solução mais razoável (logicamente falando), destrói a propriedade de não-viés.

c) Ignorar a componente de variância estimada como negativa tirando-a do modelo e recalculando os outros componentes. Thompson [1961], [1962], Bezerra [1976], Thompson e Moore [1963].

d) Procurar um outro modelo que explique os dados.

e) Coletar mais dados e analisá-los separada ou juntamente com os anteriores.

Além dessas opções diante de estimativas negativas, há ainda a opção de usarmos outros métodos que não o da "Anova": ou métodos Bayesianos ou estimação por máxima verossimilhança.

### 3.1.3 - O Método de Máxima Verossimilhança (Maximum likelihood)

No modelo I o estimador de mínimos quadrados (m.q.), e de máxima verossimilhança (m.v.) e o "b.l.u.e." (best linear unbiased estimator) para  $t'b$  coincidem (sob suposição de normalidade). Aqui no modelo II, entretanto, tal não ocorre pois o espaço paramétrico, onde a função L, de m.v., será maximizada, é não-negativo quando os parâmetros a serem estimados são componentes de variância. Os estimadores de m.v., portanto, serão sempre não negativos, e sua obtenção não é tão direta como no caso do modelo I. Vejamos a aplicação do método no modelo de 1 fator, lembrando a necessidade de suposições sobre a distribuição.

O modelo é dado em (3.1) e as suposições são:

$$\left\{ \begin{array}{l} \alpha_i \sim N(0, \sigma_\alpha^2) \\ e_{ij} \sim N(0, \sigma_e^2) \end{array} \right. \quad \left\{ \begin{array}{ll} \sigma_\alpha^2 + \sigma_e^2 & i = i' \quad j = j' \\ \sigma_\alpha^2 & i = i' \quad j \neq j' \\ 0 & i \neq i' \end{array} \right. \quad (3.7)$$

onde  $i = 1, 2, \dots, a$  e  $j = 1, 2, \dots, n$ .

Logo, sendo  $\tilde{I}_n$  a matriz identidade e  $\tilde{J}_n$  a matriz cujos elementos são todos 1, a matriz de variância de  $\tilde{y}$  é:

$$\tilde{V} = \begin{bmatrix} \sigma_e^2 \tilde{I} + \sigma_\alpha^2 \tilde{J} & 0 & 0 & \dots & 0 \\ 0 & \sigma_e^2 \tilde{I} + \sigma_\alpha^2 \tilde{J} & 0 & \dots & 0 \\ 0 & 0 & \sigma_e^2 \tilde{I} + \sigma_\alpha^2 \tilde{J} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_e^2 \tilde{I} + \sigma_\alpha^2 \tilde{J} \end{bmatrix} \quad (3.8)$$

que pode ser denotada como

$$\tilde{V} = \tilde{I}_a \otimes (\sigma_e^2 \tilde{I}_n + \sigma_\alpha^2 \tilde{J}_n), \quad (3.8a)$$

onde  $\otimes$  representa o produto de Kronecker.

A função de máxima verossimilhança é:

$$L = (2\pi)^{\frac{1}{2} an} |\tilde{V}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{y} - \underline{\mu})' \tilde{V}^{-1} (\underline{y} - \underline{\mu}) \right\} \quad (3.9)$$

onde

$$|\tilde{V}| = \prod_{i=1}^a \left| \begin{pmatrix} \sigma_e^2 & \mathbf{I}_n \\ \sigma_\alpha^2 & \mathbf{J}_n \end{pmatrix} \right| = \left[ \sigma_e^{2(n-1)} (\sigma_e^2 + n \sigma_\alpha^2) \right]^a \quad (3.10)$$

e

$$\tilde{V}^{-1} = \frac{1}{\sigma_e^2} \otimes \left[ \frac{1}{\sigma_e^2} \mathbf{I}_n - \frac{\sigma_\alpha^2}{\sigma_e^2 (\sigma_e^2 + n \sigma_\alpha^2)} \mathbf{J}_n \right] \quad (3.11)$$

Assim,

$$L = \frac{\exp - \frac{1}{2} \left[ \frac{SQE}{\sigma_e^2} + \frac{SQA}{\sigma_e^2 + n \sigma_\alpha^2} + \frac{an(\bar{y}.. - \mu)^2}{\sigma_e^2 + n \sigma_\alpha^2} \right]}{(2\pi)^{\frac{1}{2} an} (\sigma_e^2)^{\frac{1}{2} a(n-1)} (\sigma_e^2 + n \sigma_\alpha^2)^{\frac{1}{2} a}} \quad (3.12)$$

Derivando  $\log L$  em relação a  $\mu$ ,  $\sigma_\alpha^2$  e  $\sigma_e^2$  e igualando a zero, obtemos:

$$\hat{\mu} = \bar{y}.. \quad (3.13)$$

e

$$a(\hat{\sigma}_e^2 + n \hat{\sigma}_\alpha^2) = SQA \quad \text{e} \quad a(n-1) \hat{\sigma}_e^2 = SQE \quad (3.14)$$

Como não houve restrição da maximização a valores positivos, as soluções

$$\hat{\sigma}_e^2 = \frac{SQE}{a(n-1)} = QME$$

e

(3.15)

$$\hat{\sigma}_{\alpha}^2 = \frac{SQA/a - \hat{\sigma}_e^2}{n} = \frac{(1 - \frac{1}{a}) QMA - QME}{n}$$

não são os estimadores de máxima verossimilhança (ou seja, estes não são obtidos diretamente com a aplicação da "fórmula"), se  $(1 - \frac{1}{a}) QMA < QME$ . Nesse caso, conforme Herbach [1959], o estimador de m.v. para  $\sigma_{\alpha}^2$  é 0 e para  $\sigma_e^2$  é  $\frac{SQT}{an}$ . Dessa forma, a comparação entre os estimadores de "Anova" e de m.v. no modelo de 1 fator pode ser resumida da seguinte forma:

Método	Condições	$\hat{\sigma}_{\alpha}^2$	$\hat{\sigma}_e^2$
"Anova"	Nenhuma	$\frac{QMA - QME}{n}$	QME
m.v.	$\frac{a-1}{a} QMA \geq QME$	$\frac{\frac{a-1}{a} QMA - QME}{n}$	QME
	$\frac{a-1}{a} QMA < QME$	0	$\frac{SQT}{an}$

Há ainda um estimador de m.v. modificado chamado "REML" (restricted maximum likelihood) desenvolvido por Thompson [1962] e um estudo comparativo do e.m.v. com o e.m.v.r. (REML) e com uma modificação pseudo-Bayesiana do e.m.v.r., para o modelo misto, está em Harville [1978].

### 3.2 DADOS NÃO-BALANCEADOS

Os métodos para estimação dos componentes de variância costumemente usados nesse caso são os métodos 1, 2 e 3 de

de Henderson, que são análogos ao método da "anova". Um problema não resolvido com esses métodos é que no caso não-balanceado não há uma "anova" única, e também não há um conjunto único de equações, como no caso balanceado, quando igualamos cada forma quadrática com seu valor esperado.

### 3.2.1 - Método 1 de Henderson ou Método da Análise de Variância

É basicamente o mesmo método da "anova" usado para dados balanceados, somente com alterações nas fórmulas das S.Q., onde estão envolvidos os totais das caselas. O método, que está desenvolvido e exemplificado no capítulo seguinte, só pode ser usado para modelos aleatórios, pois no caso misto os estimadores são viesados pelos fatores fixos. Por exemplo, no modelo de 2 fatores com  $\alpha$  fixo a  $E(SQA)$  e a  $E(SQB)$  conterão os seguintes termos, que envolvem fatores fixos:

$$E(SQA) \text{ contém } \sum_{i=1}^a n_i \alpha_i^2 - \frac{(\sum_{i=1}^a n_i \alpha_i)^2}{N}$$

e

$$E(SQB) \text{ contém } \sum_{j=1}^b \frac{(\sum_{i=1}^a n_{ij} \alpha_i)^2}{n_{.j}} - \frac{(\sum_{i=1}^a n_i \alpha_i)^2}{N},$$

ou seja, os termos envolvendo fatores fixos são diferentes em  $E(SQA)$  e  $E(SQB)$ , e assim não se cancelam em  $E(SQA - SQB)$ .

Essa dificuldade pode ser contornada de 2 maneiras: ou tirando os efeitos fixos do modelo, ou encarando-os como aleatórios. Evidentemente ambas são insatisfatórias pois os estimadores serão viesados. Entretanto, caso essas aproximações possam ser consideradas razoáveis, o uso do procedimento pode ser cogitado, tendo-se em conta que o método 1 é, computacionalmente, o mais fácil.

### 3.2.2 - Método 2 de Henderson

Consiste em estimar os fatores fixos, corrigir os dados de acordo com essa estimativa e a seguir usar o método 1.

O método apresenta dois inconvenientes sérios: primeiro, não pode ser usado quando há interação entre fatores fixos e aleatórios, e segundo, o método nem sempre é invariante em relação ao estimador  $\tilde{b}_f$  dos fatores fixos, dependendo de certa relação de postos (capítulo 4).

O método inicialmente proposto por Henderson [1953] foi generalizado por Searle [1968], [1971], de forma tal que perde a primeira das limitações, mas os estimadores obtidos não são, em geral, unicamente definidos.

### 3.2.3 - Método 3 de Henderson - Método de Ajustar Constantes

Este método não usa as somas de quadrados da análise de variância, mas sim reduções nas SQ devido a ajustar diferentes modelos e sub-modelos. A idéia básica é particionar o vetor  $\underline{b}$  em dois vetores  $\underline{b}_1$  e  $\underline{b}_2$  e ajustar 2 modelos: o modelo completo e um sub-modelo  $\underline{y} = \underline{x}_1 \underline{b}_1 + \underline{e}$ , com as notações  $R(\underline{b}_1, \underline{b}_2)$  e  $R(\underline{b}_1)$  para as respectivas reduções correspondentes a esses ajustes. Como mostraremos no capítulo 4, sendo  $R(\underline{b}_2/\underline{b}_1) = R(\underline{b}_1, \underline{b}_2) - R(\underline{b}_1)$ ,  $E[R(\underline{b}_2/\underline{b}_1)]$  é uma função simplesmente de  $E(\underline{b}_2 \underline{b}_2')$  e  $\sigma_e^2$ , não envolvendo  $E(\underline{b}_1 \underline{b}_1')$  nem  $E(\underline{b}_1 \underline{b}_2')$ ; ou seja se  $\underline{b}_1$  contém todos os fatores fixos, então  $E[R(\underline{b}_2/\underline{b}_1)]$  não contém termos envolvendo os fatores fixos. Assim, ao aplicarmos o processo de igualar cada Redução a seu valor esperado e resolver o sistema de equações assim obtido, os estimadores serão não viesados por fatores fixos. A obtenção do conjunto de equações necessário é feita usando-se reduções que vão envolvendo sucessivamente mais fatores, de forma a introduzir um componente de variância (na esperança das reduções) de cada vez.

Além disso, outra característica importante do método é que quando aplicado ao modelo aleatório, ou seja,  $b_1$  contendo fatores aleatórios,  $E[R(b_2/b_1)]$  não contém termos envolvendo  $\text{var}(b_1)$  e nem covariância entre elementos de  $b_1$  e  $b_2$ . Assim o método pode ser usado, mesmo que haja correlação não nula entre elementos de  $b_1$  e de  $b_2$ , sem que os estimadores fiquem viesados por essa correlação.

O método apresenta, basicamente, duas desvantagens: primeiro, envolve cálculo de matrizes inversas generalizadas que terão ordem grande em modelos com grande número de fatores, tanto ao calcularmos as reduções, como os coeficientes dos componentes de variância nas esperanças; entretanto essa dificuldade é em parte contornável pelo método de síntese de Hartley; segundo, o número de equações é maior que o necessário, de forma que haverá mais de um estimador para alguns componentes de variância, e o método não fornece indicações sobre qual conjunto de equações usar.

#### 3.2.4 - Variações sobre o Método da Análise de Variância

No caso de não balanceamento, além dos métodos 1, 2 e 3 de Henderson, há mais dois métodos baseados em equações onde a esperança é uma combinação linear dos componentes de variância.

Um deles, análise de médias, consiste em calcular as médias de cada casela (não pode haver caselas vazias), obter as S.Q. da Análise de Variância com esses valores, calcular as  $E(QM)$  e montar o sistema de equações. Há duas alternativas: usar médias simples, ou ponderadas (com pesos proporcionais ao número de observações de cada casela). Os cálculos são baseados nas "a novas" sugeridas em Yates [1934] para modelos fixos.

Outro, desenvolvido por Koch [1967], [1968], baseia-se em produtos cruzados de observações, e não em somas de quadrados, e no fato de que as esperanças desses produtos cruzados tam

bem são combinações lineares dos componentes de variância.

Desses métodos, evidentemente, o de análises de médias não ponderadas é o de mais fácil computação.

### 3.2.5 - Método de Máxima Verossimilhança

No caso do modelo III, o método de m.v., pode ser usado para a estimação dos fatores fixos, caso a matriz de variância  $\underline{V}$  seja conhecida, ou ainda, em combinação com o método de ajustar constantes, estimando-se os componentes de variância e a seguir estimando os fatores fixos por m.v. usando a matriz  $\underline{V}$  estimada.

Entretanto, para a estimação dos componentes de variância, no modelo II ou no III é impraticável, pois a dedução algébrica torna-se extremamente complicada. Por exemplo, no caso de modelo aleatório 1 fator, as equações fornecidas pelo método são:

$$\hat{\mu} = \frac{\sum_{i=1}^a \frac{n_i \bar{y}_{i.}}{\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2}}{\sum_{i=1}^a \frac{n_i}{\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2}}$$

$$\frac{N-a}{\hat{\sigma}_e^2} + \sum_{i=1}^a \frac{1}{\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2} - \frac{\sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{\hat{\sigma}_e^4} - \sum_{i=1}^a \frac{n_i (\bar{y}_{i.} - \hat{\mu})^2}{(\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2)} = 0$$

e

$$\sum_{i=1}^a \frac{n_i}{\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2} - \sum_{i=1}^a \frac{n_i^2 (\bar{y}_{i.} - \hat{\mu})^2}{(\hat{\sigma}_e^2 + n_i \hat{\sigma}_\alpha^2)^2} = 0$$

Obviamente, essas equações não permitem estimadores para  $\mu$ ,  $\sigma_e^2$  e  $\sigma_\alpha^2$ , e mesmo que houvesse solução, a restrição de espaço paramétrico não-negativo teria de ser considerada, e a ob-

tenção dos estimadores não seria direta.

### 3.3 - AS PROPRIEDADES DESEJÁVEIS DOS ESTIMADORES

#### 3.3.1 - Estimador Quadrático Não-Viesado de Variância Mínima

Como vimos em 3.1.2, no caso balanceado os estimadores da "anova" são os estimadores quadráticos não-viesados de mínima variância. No caso não-balanceado entretanto, a obtenção de estimadores assim é muito difícil e às vezes não existe um estimador que tenha sempre variância mínima para todos os valores dos componentes de variância. Em alguns casos particulares, a literatura apresenta aproximações ou estimadores "locais" (com essas propriedades).

#### 3.3.2 - Viés

A propriedade de não-viés dos estimadores do modelo fixo é desejável com base no conceito de repetição dos dados provenientes do mesmo conjunto de efeitos. No caso de dados não-balanceados, entretanto, há repetição de dados, mas não necessariamente com o mesmo padrão de não balanceamento, e os dados não resultam do mesmo conjunto de efeitos. Assim, o conceito de não-viés, nesse caso, é questionável. Dentro desse contexto, Blischke [1968] lembra que esses estimadores considerados (com exceção dos de m.v.) são estimadores do método de momentos, que a maioria deles é, provavelmente, consistente, e sendo essa propriedade assintótica, ela é mantida, mesmo que a estimativa negativa seja substituída por zero. Outra sugestão, em Searle [1968], é usar-se um estimador modal.

O fato é que no caso de não-balanceamento, não há uma teoria completa e geral, e muitos dos resultados obtidos são particulares. A discussão em Searle [1968] mostra bem esse aspecto da questão.

## CAPÍTULO 4

### OS MÉTODOS 1 2 E 3 DE HENDERSON

Neste capítulo apresentamos detalhadamente os 3 métodos de Henderson para estimação pontual de componentes de variância. É apresentado também o método de síntese, de Hartley, que é um procedimento para facilitar a obtenção dos coeficientes nas equações.

O método 2 de Henderson é mostrado como um caso particular de uma simplificação do método 2 Generalizado de Searle. Discute-se o problema de unicidade da solução.

Todos os métodos são acompanhados de aplicação, inclusive numérica (e resolução).

#### 4.1 - O MÉTODO 1 DE HENDERSON

O método 1 é basicamente o método da "anova" (usado para dados balanceados), e consiste em obter SQ não corrigidas, análogas àquelas, e em igualar cada SQ à sua esperança, obtendo assim um sistema de equações cujas incógnitas são os estimadores

dos componentes de variância. Para ilustrar o método, usaremos o modelo:

$$Y_{ijk\ell} = \mu + \alpha_i + \beta_j + \gamma_K + (\alpha\beta)_{ij} + e_{ijk\ell} \quad (4.1)$$

onde:

$$i = 1, \dots, a ; j = 1, \dots, b ; K = 1, \dots, r$$

$$\ell = 1, \dots, n_{ijk} \quad \sum_i \sum_j \sum_K n_{ijk} = n \dots = N$$

ou em forma matricial:

$$\underline{y} = \underline{1} + \underline{X}_1 \underline{b}_1 + \underline{X}_2 \underline{b}_2 + \underline{X}_3 \underline{b}_3 + \underline{X}_4 \underline{b}_4 + e \quad (4.2)$$

onde  $\underline{b}_1$  é o vetor dos efeitos  $\alpha$ ,  $\underline{b}_2$  dos efeitos  $\beta$ , etc..

#### 4.1.1 - Somas de Quadrados

Para a obtenção das S.Q. é conveniente o cálculo das S.Q. não corrigidas pela média, (que chamaremos, termos T), sendo cada termo T, uma S.Q. correspondente ao fator indexado. A seguir cada termo T é corrigido por  $T_\mu$ , fornecendo as S.Q. devidas a cada fator (ou interação).

As S.Q. são análogas às SQ para dados balanceados; por exemplo, no caso balanceado  $SQA = \sum_{i=1}^a \frac{Y_{i\dots}^2}{brn} - \frac{Y_{\dots}^2}{N}$  (onde  $N = abrn$ ) e aqui temos:  $SQA = \sum_{i=1}^a \frac{Y_{i\dots}^2}{n_i} - \frac{Y_{\dots}^2}{N}$ .

Os termos T são:

$$T_A = \sum_{i=1}^a \frac{Y_{i\dots}^2}{n_{i\dots}} \quad T_B = \sum_{j=1}^b \frac{Y_{\dots j\dots}^2}{n_{\dots j\dots}} \quad T_C = \sum_{K=1}^r \frac{Y_{\dots \dots K}^2}{n_{\dots \dots K}}$$

$$T_{AB} = \sum_{i=1}^a \sum_{j=1}^b \frac{Y_{ij\dots}^2}{n_{ij\dots}} \quad (4.3)$$

A S.Q. total e o fator de correção pela média são respectivamente

$$T_0 = \frac{a}{\sum_{i=1}^r} \frac{b}{\sum_{j=1}^r} \frac{r}{\sum_{K=1}^r} \frac{n_{ijK}}{\sum_{\ell=1}^2} y_{ijK\ell}^2 \quad \text{e} \quad T = \frac{Y^2 \dots}{n \dots} \quad (4.4)$$

As S.Q. devidas a cada fator, são então:

$$\begin{aligned} SQA &= T_A - T_\mu \\ SQB &= T_B - T_\mu \\ SQC &= T_C - T_\mu \\ SQAB &= T_{AB} - T_A - T_B + T_\mu \\ SQE &= T_0 - T_C - T_{AB} + T_\mu \end{aligned} \quad (4.5)$$

Usando o modelo (4.2) escrevemos as SQ como reduções nas S.Q. devido a ajustar submodelos. Assim por exemplo  $T_C$  é a redução na SQT devido a ajustar o modelo  $\underline{y} = \underline{X}_3 \underline{b}_3 + \underline{e}$  e, portanto,

$$\sum_{K=1}^r \frac{y_{..K}^2}{n_{..K}} = R(\underline{b}_3) = \underline{y}' \underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' \underline{y} \quad (4.6)$$

#### 4.1.2 - As Esperanças das S.Q.

A obtenção das  $E(SQ)$  é feita calculando-se diretamente a esperança de cada termo, após substituir o termo de  $\underline{y}$  de cada fator pelo modelo (4.1) devidamente modificado. Por exemplo,

$$\begin{aligned}
E(T_A) &= E \sum_{i=1}^a \frac{Y_{i\dots}^2}{n_{i\dots}} = \sum_{i=1}^a E \frac{Y_{i\dots}^2}{n_{i\dots}} = \sum_{i=1}^a E(n_{i\dots} \mu + n_{i\dots} \alpha_1 + \\
&+ \sum_j n_{ij} \beta_j + \sum_k n_{i.K} \gamma_K + \sum_j n_{ij} (\alpha\beta)_{ij} + e_{i\dots})^2 / n_{i\dots} = \\
&= \sum_{i=1}^a E(n_{i\dots}^2 \mu^2 + n_{i\dots}^2 \alpha_1^2 + \sum_j n_{ij}^2 \beta_j^2 + \sum_K n_{i.K}^2 \gamma_K^2 + \sum_j n_{ij}^2 (\alpha\beta)_{ij}^2 + \\
&+ e_{i\dots}^2 + \Sigma (\text{duplos produtos, todos tendo esperança nula}) / n_{i\dots} = \\
&= \sum_{i=1}^a (n_{i\dots}^2 \mu^2 + n_{i\dots}^2 \sigma_\alpha^2 + \frac{\sum_j n_{ij}^2}{n_{i\dots}} \sigma_\beta^2 + \frac{\sum_K n_{i.K}^2}{n_{i\dots}} \sigma_\gamma^2 + \\
&+ \frac{\sum_j n_{ij}^2}{n_{i\dots}} \sigma_{\alpha\beta}^2 + \frac{\sum_j \sum_K \sum_l \sigma_e^2}{n_{i\dots}}) \\
&= N\mu^2 + N \sigma_\alpha^2 + \sum_i \frac{\sum_j n_{ij}^2}{n_{i\dots}} \sigma_\beta^2 + \sum_i \frac{\sum_K n_{i.K}^2}{n_{i\dots}} \sigma_\gamma^2 + \\
&+ \sum_i \frac{\sum_j n_{ij}^2}{n_{i\dots}} \sigma_{\alpha\beta}^2 + a \sigma_e^2 \tag{4.7}
\end{aligned}$$

O processo é um tanto trabalhoso, e há duas alternativas: usar um método desenvolvido por Hartley, e exposto mais adiante, chamado método de síntese, que fornece os valores numéricos dos coeficientes, ou usar uma expressão geral para os coeficientes dos termos T, como faremos a seguir.

$$\text{Sendo o modelo } \underline{y} = \mu \underline{1} + \sum_{\theta=A}^K \underline{x}_{\theta} \underline{b}_{\theta} + \underline{e} \quad (4.8)$$

definimos

$y.(A_i)$  = total das observações no  $i$ -ésimo nível do fator A

$n(A_i)$  = número de observações no  $i$ -ésimo nível do fator A

$$e \quad T_A = \sum_{i=1}^{N_A} \frac{[y.(A_i)]^2}{n(A_i)} \quad , \quad \text{onde } N_A = \text{n}^{\circ} \text{ de níveis do fator A.} \quad (4.9)$$

Definimos ainda:

$n(A_i, \theta_j)$  = número de observações no  $i$ -ésimo nível do fator A e  $j$ -ésimo nível do fator  $\theta$

$b_{\theta_j}$  =  $j$ -ésimo elemento de  $\underline{b}_{\theta}$

$e.(A_i)$  = total do erro correspondente a  $y.(A_i)$ .

Então, substituindo (4.8) em (4.9) obtemos:

$$T_A = \sum_{i=1}^{N_A} \frac{[n(A_i) \mu + \sum_{\theta=A}^K \sum_{j=1}^{N_{\theta}} n(A_i, \theta_j) b_{\theta_j} + e.(A_i)]^2}{n(A_i)} \quad (4.10)$$

A esperança de  $T_A$  contém:

a) um termo em  $\mu^2$

$$\sum_{i=1}^{N_A} \frac{[n(A_i)]^2 \mu^2}{n(A_i)} = \mu^2 \sum_{i=1}^{N_A} n(A_i) = N \mu^2 \quad (4.11)$$

b) um termo em  $\sigma_\theta^2$ , para  $\theta = A, B, \dots, K$  :

$$K(\sigma_\theta^2, E(T_A)) \sigma_\theta^2 = \sum_{i=1}^{N_A} \frac{\sum_{j=1}^{N_\theta} [n(A_i, \theta_j)]^2}{n(A_i)} \sigma_\theta^2 \quad (4.12)$$

onde  $K(\sigma_\theta^2, E(T_A))$  é a notação para o coeficiente de  $\sigma_\theta^2$  na  $E(T_A)$ .

Notar, como regra para facilitar a expressão das fórmulas, que o índice da 1ª somatória (no caso,  $i$ ) refere-se ao fator  $A$  e que o índice da 2ª somatória (no caso,  $j$ ) refere-se ao fator  $\theta$ .

c) um termo em  $\sigma_e^2$ :

$$\sum_{i=1}^{N_A} \frac{n(A_i) \sigma_e^2}{n(A_i)} = \sigma_e^2 \sum_{i=1}^{N_A} 1 = N_A \sigma_e^2 \quad (4.13)$$

Trabalhando com reduções na SQT, ao invés de usar os termos  $T$ , o mesmo resultado é obtido aplicando-se o teorema de 2.6 à redução considerada. Por exemplo, para a  $E[R(\underline{b}_3)]$ , supondo  $\underline{b}_1$  como vetor de efeitos fixos, temos:

$$\begin{aligned} E[R(\underline{b}_3)] &= E[\underline{y}' \underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' \underline{y}] = \\ &= (\mu \underline{1} + \underline{X}_1 \underline{b}_1)' \underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' (\mu \underline{1} + \underline{X}_1 \underline{b}_1) + \\ &+ \sum_{i=2}^4 \sigma_i^2 \text{tr}[\underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' \underline{X}_i \underline{X}_i'] + \sigma_e^2 \text{tr}[\underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3'] \quad (4.14) \end{aligned}$$

No caso do modelo aleatório, a expressão se reduz a

$$E[R(\underline{b}_3)] = \mu^2 \underline{1}' \underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' \underline{1} + \sum_{i=1}^4 \sigma_i^2 \text{tr}[\underline{X}_3 (\underline{X}_3' \underline{X}_3)^{-1} \underline{X}_3' \underline{X}_i \underline{X}_i'] + \sigma_e^2 r(\underline{X}_3) \quad (4.15)$$

O desenvolvimento dos produtos de matrizes tem como resultado os termos dados por (4.11), (4.12) e (4.13). No caso do modelo (4.1), que estamos usando, os coeficientes de  $\mu^2$  e dos componentes de variância nas esperanças dos termos T são dados na tabela 4.1:

	$\mu^2$	$\sigma_\alpha$	$\sigma_\beta$	$\sigma_\gamma$	$\sigma_{\alpha\beta}$	$\sigma_e^2$
$T_0$	N	N	N	N	N	N
$T_A$	N	N	$K_1$	$K_2$	$K_1$	a
$T_B$	N	$K_3$	N	$K_4$	$K_3$	b
$T_{AB}$	N	N	N	$K_5$	N	s
$T_C$	N	$K_6$	$K_7$	N	$K_8$	r
$T_\mu$	N	$K_9$	$K_{10}$	$K_{11}$	$K_{12}$	1

s = número total de caselas não vazias no cruzamento de  $\alpha$  e  $\beta$ .

TABELA 4.1

onde

$$K_1 = \sum_i \frac{\sum_j n_{ij}^2}{n_{i..}}$$

$$K_2 = \sum_i \frac{\sum_K n_{i.K}^2}{n_{i..}}$$

$$K_3 = \sum_j \frac{\sum_i n_{ij}^2}{n_{.j}}$$

$$K_4 = \sum_j \frac{\sum_K n_{i.K}^2}{n_{.j}}$$

$$K_5 = \sum_i \sum_j \frac{\sum_K n_{ijk}^2}{n_{ij.}}$$

$$K_6 = \sum_K \frac{\sum_i n_{i.K}^2}{n_{..K}}$$

$$\begin{aligned}
 K_7 &= \sum_K \frac{\sum_j n_{.jK}^2}{n_{.jK}} & K_8 &= \sum_K \frac{\sum_i \sum_j n_{ijK}^2}{n_{..K}} & K_9 &= \frac{\sum_i n_{i..}^2}{N} \\
 K_{10} &= \frac{\sum_j n_{.j.}^2}{N} & K_{11} &= \frac{\sum_K n_{..K}^2}{N} & K_{12} &= \frac{\sum_i \sum_j n_{ij.}^2}{N}
 \end{aligned}
 \tag{4.16}$$

#### 4.1.3 - Um Exemplo Numérico

Um exemplo numérico ao qual o modelo dado pela equação (4.1) se aplica é dado em Henderson [1953], e é baseado na seguinte situação:

No estado de N.Y. uma organização cooperativa (New York Artificial Breeders Cooperative, Inc.), que fornece semen bovino, tem aproximadamente 60 touros em atividade. O touro que fornece semen a uma particular vaca é escolhido aleatoriamente, o que, ao lado do grande número de descendentes (fêmeas) de cada touro, torna as observações da produção leiteira dessas fêmeas particularmente adequadas ao estudo das diferenças genéticas entre touros e também para estimar a magnitude de outras fontes de variação na produção de leite. É visível a utilidade de "boas" estimativas dessas variâncias para o planejamento de programas de seleção eficientes.

Henderson salienta que ao estimar os componentes de variância nessa situação, as dificuldades eram provenientes dos seguintes aspectos:

a) Os dados provinham de observações feitas em diversos anos, e o efeito do tempo (como tendência) é um fator importante a ser considerado.

b) Os dois principais critérios de classificação dos dados eram: reprodutor e reprodutora ("pai" e "mãe"), sendo o nú-

mero de reprodutores maior que 100 e o de reprodutoras maior que 2.000.

c) Não-balanceamento: o número de observações nas caselas reprodutor x reprodutora variava, a maioria sendo zero.

As estimativas obtidas pelos métodos 1 e 2 foram essencialmente as mesmas. A partir dessa massa de dados, Henderson selecionou uma pequena amostra que é apresentada em seu artigo e que apresentamos a seguir.

Os dados, que estão dispostos na tabela 2, são o número de observações, da quantidade de gordura ("butterfat") na primeira lactação, em cada casela (do cruzamento pai x mãe x ano) e também a soma das observações em cada casela.

"MÃE"	"PAI"	A N O				TOTAL
		1	2	3	4	
1	1	3-1414	2- 981			5-2395
1	2		4-1766	2- 862		6-2628
1	3				5-1609	5-1609
2	1	1- 404	3-1270			4-1674
2	2			5-2109		5-2109
2	3			4-1563	2- 740	6-2303
3	1		3-1705			3-1705
3	2		4-2310	2-1134		6-3444
4	1	3-1113	5-1951			8-3064
4	3			3-1291	6-2457	9-3748
TOTAL		7-2931	21-9983	16-6959	13-4806	57-24679

TABELA 4.2

O modelo adequado, evidentemente, é o da equação (4.1)

$$y_{ijK\ell} = \mu + \alpha_i + \beta_j + \gamma_K + (\alpha\beta)_{ij} + e_{ijK\ell}$$

onde

$$i = 1, 2, 3, 4 \quad j = 1, 2, 3 \quad K = 1, 2, 3, 4$$

$$\ell = 1, \dots, n_{ijK} \quad \sum_i \sum_j \sum_K n_{ijK} = n \dots = N = 57 .$$

$y_{ijK\ell}$  é a quantidade de gordura do leite na primeira lactação da  $i$ -ésima descendente do  $j$ -ésimo reprodutor com a  $i$ -ésima reprodutora, medida no  $K$ -ésimo ano.

$\mu$  é a média geral,  $\alpha_i$  é o efeito da  $i$ -ésima reprodutora sobre a produção de gordura,  $\beta_j$  é o efeito do  $j$ -ésimo reprodutor,  $\gamma_K$  é o efeito do  $K$ -ésimo ano na medida,  $(\alpha\beta)_{ij}$  é o efeito de interação entre o  $j$ -ésimo reprodutor e a  $i$ -ésima reprodutora e  $e_{ijK}$  é o erro associado à  $i$ -ésima medida da casela  $ijK$ , que é um elemento aleatório.

As suposições são as usuais: todos os fatores do modelo são v.a. não correlacionadas, com média zero e respectivas variâncias:  $\sigma_\alpha^2$ ,  $\sigma_\beta^2$ ,  $\sigma_{\alpha\beta}^2$ ,  $\sigma_\gamma^2$ ,  $\sigma_e^2$ . A matriz de variância de  $\underline{y}$ ,  $\underline{V}$  é semelhante à matriz (3.8) (onde os componentes de variância eram  $\sigma_\alpha^2$  e  $\sigma_e^2$ ).

Os coeficientes de  $\mu^2$  e dos componentes de variância nas esperanças dos termos T, indicados na tabela 4.1, e calculados conforme (4.16), estão na tabela 4.3, que apresenta, também, os valores dos termos T, na última coluna.

TABELA 4.3

	$\mu^2$	$\sigma_\alpha^2$	$\sigma_\beta^2$	$\sigma_C^2$	$\sigma_{AB}^2$	$\sigma_e^2$	Valor de T
$T_0$	57	57	57	57	57	57	11.124.007
$T_A$	57	57	24,04	21,49	24,04	4	10.893.666
$T_B$	57	18,51	57	30,33	18,51	3	10.776.278
$T_{AB}$	57	57	57	37,35	57	10	10.970.369
$T_C$	57	19,51	39,22	57	15,10	4	10.776.451
$T_\mu$	57	14,93	19,11	16,05	6,19	1	10.685.141

Portanto, as equações a serem resolvidas são as apresentadas na forma da tabela 4.4.

TABELA 4.4

	$\sigma_A^2$	$\sigma_B^2$	$\sigma_C^2$	$\sigma_{AB}^2$	$\sigma_e^2$	Valor das diferenças
$T_A - T_\mu$	42,07	4,93	5,44	17,85	3	208.525
$T_B - T_\mu$	3,58	37,89	14,28	12,32	2	91,137
$T_C - T_\mu$	4,58	20,11	40,95	8,91	3	91,310
$T_{AB} - T_A - T_B + T_\mu$	-3,58	-4,93	1,58	20,64	4	-14,434
$T_0 - T_C - T_{AB} + T_\mu$	-4,58	-20,11	-21,30	-8,91	44	62,328

Isto é, a primeira equação é:

$$42,07 \hat{\sigma}_A^2 + 4,93 \hat{\sigma}_B^2 + 5,44 \hat{\sigma}_C^2 + 17,85 \hat{\sigma}_{AB}^2 + 3 \hat{\sigma}_e^2 = 208.525, \text{ e as}$$

outras são escritas analogamente. Resolvendo o sistema obtemos:

$$\begin{aligned} \hat{\sigma}_A^2 &= 4531 & \hat{\sigma}_B^2 &= 1587 & \hat{\sigma}_C^2 &= 763 \\ \hat{\sigma}_{AB}^2 &= -164 & \hat{\sigma}_e^2 &= 2950 \end{aligned} \quad (4.17)$$

como  $\hat{\sigma}_{AB}^2$  é negativo vamos considerar  $\sigma_{AB}^2 = 0$  e com isso a solução do sistema fica:

$$\hat{\sigma}_A^2 = 4468 \quad \hat{\sigma}_B^2 = 1542 \quad \hat{\sigma}_C^2 = 756 \quad \text{e} \quad \hat{\sigma}_e^2 = 2952. \quad (4.18)$$

O componente de variância devido ao fator A (reprodutora) é o que tem maior peso na variância total, indicando que na seleção para obtenção de leite mais "gordo" esse fator deve ser mais controlado.

#### 4.1.4 - Método de Síntese de Hartley

Hartley [1967] desenvolveu um método por ele denominado de "synthesis", que fornece os valores numéricos dos coeficientes dos componentes da variância, em termos como  $E(T_A)$ , sem a necessidade de se obter a expressão algébrica.

Consiste em calcular o coeficiente do componente  $\sigma_\theta^2$  em  $E(T_A)$  de maneira análoga ao cálculo da SQ  $T_A$ , mas usando a matriz de planejamento  $\tilde{X}$  ao invés da matriz  $y$ . Isto é, como vimos em (4.9) o termo  $T_A$  é:

$$T_A = \sum_{i=1}^{N_A} \frac{[y \cdot (A_i)]^2}{n(A_i)}$$

que é a soma dos quadrados das observações de cada nível de A, dividida por  $n(A_i)$ , e o coeficiente de  $\sigma_\theta^2$  em  $E(T_A)$  é a soma dos quadrados, para todos os níveis de  $\theta$ , de  $T_A[\tilde{X}(\theta, j)]$  que é o análogo de  $T_A$  usando a matriz de planejamento.

Nessa notação  $\tilde{X}(\theta, j)$  é a  $j$ -ésima coluna de  $\tilde{X}_\theta$ , e

$$T_A[\tilde{X}(\theta, j)] = \sum_{i=1}^{N_A} \frac{[n(A_i, \theta_j)]^2}{n(A_i)} \quad (4.19)$$

ou seja,  $T_A[\tilde{X}(\theta, j)]$  é a soma, em  $i$ , dos quadrados dos números de "uns" em  $\tilde{X}(\theta, j)$  que provêm de observações que estão no  $i$ -ésimo nível de  $A$ .

O coeficiente de  $\sigma_\theta^2$  em  $E(T_A)$  é:

$$K[\sigma_\theta^2, E(T_A)] = \sum_{j=1}^{N_\theta} T_A[\tilde{X}(\theta, j)] = \sum_{j=1}^{N_\theta} \sum_{i=1}^{N_A} \frac{[n(A_i, \theta_j)]^2}{n(A_i)} \quad (4.20)$$

A proposição de Hartley é de efetuarmos essas somas com dados obtidos em  $\tilde{X}$ . Entretanto, o método pode ser usado consultando diretamente uma tabela que forneça o número de observações nos diversos cruzamentos, como a tabela 4.2.

Assim, por exemplo, o cálculo do coeficiente  $K_1$  de  $\sigma_\beta^2$  em  $E(T_A)$ , que foi apresentado na tabela 4.3 pode ser obtido da seguinte maneira:

a) Em cada um dos quatro níveis de  $\alpha$ , verificamos qual o número de observações nos cruzamentos com cada um dos níveis de  $\beta$ , e dividimos a soma dos quadrados desses números pela soma simples (que é o número total de observações nesse nível de  $\alpha$ ). Dessa forma, a tabela 4.2 nos fornece:

TABELA 4.5

nível de $\alpha$	1	2	3	4
$\sum_{j=1}^3 \frac{[n(A_1, \beta_j)]^2}{n(A_1)}$	$\frac{5^2+6^2+5^2}{16}$	$\frac{4^2+5^2+6^2}{15}$	$\frac{3^2+6^2}{9}$	$\frac{8^2+9^2}{17}$

b) Somamos esses 4 resultados, conforme 4.20:

$$K_1 = \frac{5^2+6^2+5^2}{16} + \frac{4^2+5^2+6^2}{15} + \frac{3^2+6^2}{9} + \frac{8^2+9^2}{17} = 24,04 \quad (4.21)$$

Notemos que se na expressão 4.20, que fornece os coeficientes por síntese, trocarmos as somatórias obtemos a expressão (4.9) que fornece diretamente os coeficientes:

$$K[\sigma_\theta^2, E(T_A)] = \sum_{i=1}^{N_A} \sum_{j=1}^{N_\theta} \frac{[n(A_i, \theta_j)]^2}{n(A_i)} \quad (4.22)$$

A dedução, formalização e exposição teórica do método encontra-se em Hartley [1967], e uma extensão, para a obtenção dos coeficientes nas fórmulas das esperanças, variâncias e covariâncias dos quadrados médios da "anova", para qualquer matriz de planejamento e para o caso de modelo misto, em Rao [1968] .

#### 4.2 - O MÉTODO 2

Como vimos em 3.2.1., quando o modelo é misto, os estimadores do método 1 são viesados pelos fatores fixos. A idéia, então, é fazer uma correção nas observações, de forma que as observações corrigidas sejam explicadas por um modelo composto só por fatores aleatórios, ou seja, se no modelo dado em (4.2)  $b_1$  é o vetor dos fatores fixos, o modelo ajustado aos dados corrigidos  $\tilde{z} = \tilde{y} - \tilde{X}_1 b_1$  é  $\tilde{z} = \mu \tilde{1} + \tilde{X}_2 b_2 + \tilde{X}_3 b_3 + \tilde{X}_4 b_4 + e$ . Como  $b_1$  não é conhecido, vamos estimá-lo e corrigir  $\tilde{y}$  usando  $\tilde{b}_1$ . Após essa correção aplicamos o método 1.

Conforme a maneira de estimarmos os fatores fixos teremos dois procedimentos diferentes, um que chamaremos método 2 generalizado e outro, método 2 simplificado, sendo o método 2 de Henderson apenas uma das maneiras de executar esse último.

## 4.2.1 - Método 2 Generalizado

Vamos representar o vetor dos fatores fixos (excluindo  $\mu$ ) por  $\underline{b}_f$ , o dos aleatórios por  $\underline{b}_r$  e as correspondentes matrizes da partição de  $\underline{X}$  por  $\underline{X}_f$  e  $\underline{X}_r$ . Dessa forma podemos escrever o modelo como:

$$\underline{y} = \mu \underline{1} + \underline{X}_f \underline{b}_f + \underline{X}_r \underline{b}_r + \underline{e} . \quad (4.23)$$

O estimador de  $\underline{b}_f$  será  $\hat{\underline{b}}_f$  e o vetor de dados corrigido será:

$$\underline{z} = \underline{y} - \underline{X}_f \hat{\underline{b}}_f \quad (4.24)$$

e terá a forma:

$$\underline{z} = \mu \underline{t} + \underline{X}_r^* \underline{b}_r + \underline{W} \underline{e} \quad (4.25)$$

Suponhamos que o estimador de  $\underline{b}_f$  seja

$$\hat{\underline{b}}_f = \underline{L} \underline{y} \quad (4.26)$$

Então

$$\begin{aligned} \underline{z} &= \underline{y} - \underline{X}_f \hat{\underline{b}}_f \\ &= (\underline{I} - \underline{X}_f \underline{L}) \underline{y} \\ &= \mu (\underline{I} - \underline{X}_f \underline{L}) \underline{1} + (\underline{X}_f - \underline{X}_f \underline{L} \underline{X}_f) \underline{b}_f + \\ &\quad + (\underline{X}_r - \underline{X}_f \underline{L} \underline{X}_r) \underline{b}_r + (\underline{I} - \underline{X}_f \underline{L}) \underline{e} \end{aligned} \quad (4.27)$$

É claro que  $\underline{b}_f$  não aparece em (4.27) se

$$\underline{X}_f = \underline{X}_f \underline{L} \underline{X}_f , \quad (4.28)$$

ou seja, se  $\underline{L}$  é uma inversa generalizada de  $\underline{X}_f$ .

Portanto, as matrizes  $\underline{t}$ ,  $\underline{X}_r^*$  e  $\underline{W}$  da eq. (4.25) são:

$$\underline{t} = (\underline{I} - \underline{X}_f \underline{L}) \underline{l} \quad (4.29)$$

$$\underline{X}_r^* = (\underline{X}_r - \underline{X}_f \underline{L} \underline{X}_r) \quad (4.30)$$

e

$$\underline{W} = (\underline{I} - \underline{X}_f \underline{L}) \quad (4.31)$$

O método 2 generalizado consiste em estimarmos  $\underline{b}_f$  como  $\underline{L} \underline{y}$  (eq. (4.26)), sendo  $\underline{L}$  uma inversa generalizada de  $\underline{X}_f$ , corrigirmos  $\underline{y}$  como na eq. (4.24), obtendo  $\underline{z}$  dado em (4.25) com  $\underline{t}$ ,  $\underline{X}_r^*$  e  $\underline{W}$  de (4.29), (4.30) e (4.31), e aplicarmos a  $\underline{z}$  o método 1. Notar que para cada  $\underline{L}$  diferente que usarmos obteremos diferentes valores para  $\underline{t}$ ,  $\underline{X}_r^*$  e  $\underline{W}$ , e diferentes estimadores para os componentes de variância. Nesse sentido dizemos que o método não é unicamente definido. Entretanto, no método 2 simplificado temos certa classe de matrizes  $\underline{X}$  tais que os estimadores dos componentes de variância são invariantes a qualquer solução  $\hat{\underline{b}}^0$  de (4.26).

#### 4.2.2 - Método 2 Simplificado

Uma simplificação que podemos fazer em (4.25) é forçar  $\mu \underline{t}$  a ter uma forma  $\mu^* \underline{l}$  (para algum escalar  $\mu^*$  não necessariamente igual a  $\mu$ ) e que  $\underline{X}_r^* = \underline{X}_r$ . Com isso  $\underline{z}$  da eq. (4.25) se reduz a

$$\underline{z} = \mu^* \underline{l} + \underline{X}_r \underline{b}_r + \underline{W} \underline{e} \quad (4.32)$$

Para obtermos tais condições, é preciso que, ao invés de exigirmos a validade de (4.28) ou seja, que  $\underline{L}$  seja inversa

generalizada de  $\tilde{X}_f$ , e também de (4.29) (que é consequência de (4.28), tenhamos:

$$\mu^* \underline{1} = \mu (\underline{I} - \tilde{X}_f \tilde{L}) \underline{1} + (\tilde{X}_f - \tilde{X}_f \tilde{L} \tilde{X}_f) \underline{b}_f \quad (4.33)$$

e ao invés de (4.30), tenhamos:

$$\tilde{X}_f \tilde{L} \tilde{X}_r = \underline{0} . \quad (4.34)$$

Agora, para que (4.33) seja satisfeita é preciso que

$$\tilde{X}_f \tilde{L} \tilde{X}_f \text{ tenha a mesma soma em cada linha} \quad (4.35)$$

$$\text{e } \tilde{X}_f - \tilde{X}_f \tilde{L} \tilde{X}_f \text{ tenha todas as linhas iguais} \quad (4.36)$$

pois, (4.35) implica em que

$$\mu (\underline{I} - \tilde{X}_f \tilde{L}) \underline{1} = \mu (\underline{1} - \tilde{X}_f \tilde{L} \underline{1}) \text{ seja } = \mu \lambda \underline{1} , \quad \lambda \in \mathbb{R} \text{ e} \quad (4.36), \text{ em que}$$

$$(\tilde{X}_f - \tilde{X}_f \tilde{L} \tilde{X}_f) \underline{b}_f = K \underline{1} , \quad K \in \mathbb{R} , \text{ de}$$

forma que o membro direito de (4.33) se reduz a

$$\mu \lambda \underline{1} + K \underline{1} = (\mu \lambda + K) \underline{1} = \mu^* \underline{1}$$

Assim, no método 2 simplificado, não exigimos que  $\tilde{L}$  seja inversa generalizada de  $\tilde{X}_f$ , condição (4.28), e nem que os consequentes valores de  $t$  e de  $\tilde{X}_r^*$  de (4.29) e (4.30) valham; o conjunto de condições que devem vigorar é o dado pelas equações (4.34), (4.35) e (4.36) e também (4.31), ou seja:

$$\tilde{X}_f \tilde{L} \tilde{X}_r = \underline{0}$$

$\tilde{X}_f \tilde{L}$  tem a mesma soma em cada linha

$\tilde{X}_f - \tilde{X}_f \tilde{L} \tilde{X}_f$  tem todas as linhas iguais

e conseqüentemente

$$\mu^* = \mu\lambda + K, \quad \lambda \text{ e } K \text{ escalares} \quad (4.37)$$

e

$$\tilde{W} = (\tilde{I} - \tilde{X}_f \tilde{L})$$

no modelo (4.32).

Usar (4.32) é mais simples (computacionalmente) que usar (4.25), mas envolve uma limitação: não pode haver interação entre fatores fixos e aleatórios, conforme demonstramos a seguir:

Suponhamos que haja interação desse tipo e consideremo-la como fator aleatório. Nesse caso algumas colunas de  $\tilde{X}_f$  serão somas de colunas de  $\tilde{X}_r$ , ou seja  $\tilde{X}_f$  pode ser particionada em  $[\tilde{X}_{f_1}, \tilde{X}_{f_2}]$  onde  $\tilde{X}_{f_2} = \tilde{X}_r \tilde{M}$  para alguma matriz  $\tilde{M}$ .

Então, (4.34),

$$\tilde{X}_f \tilde{L} \tilde{X}_r = \tilde{0} \Rightarrow \tilde{X}_f \tilde{L} \tilde{X}_r \tilde{M} = \tilde{0} \Rightarrow \tilde{X}_f \tilde{L} \tilde{X}_{f_2} = \tilde{0} \quad (4.38)$$

A matriz de (4.36) pode ser escrita como

$$\begin{aligned} \tilde{X}_f - \tilde{X}_f \tilde{L} \tilde{X}_f &= \begin{bmatrix} \tilde{X}_{f_1} & \tilde{X}_{f_2} \end{bmatrix} - \begin{bmatrix} \tilde{X}_f \tilde{L} \tilde{X}_{f_1} & \tilde{X}_f \tilde{L} \tilde{X}_{f_2} \end{bmatrix} = \\ &= \begin{bmatrix} \tilde{X}_{f_1} - \tilde{X}_f \tilde{L} \tilde{X}_{f_1} & \tilde{X}_{f_2} - \tilde{X}_f \tilde{L} \tilde{X}_{f_2} \end{bmatrix} \\ &= \begin{bmatrix} \tilde{X}_{f_1} - \tilde{X}_f \tilde{L} \tilde{X}_{f_1} & \tilde{X}_{f_2} \end{bmatrix} \end{aligned} \quad (4.39)$$

E assim a validade de (4.36) exige que  $X_{f_2}$  tenha todas as linhas iguais, o que é absurdo.

Se considerarmos a interação como fator fixo, então  $X_r$  pode ser particionada em  $(X_{r_1}, X_{r_2})$  com  $X_{r_2} = X_f M$  para alguma matriz  $m$  e por (4.34)  $X_{f_2} L X_f M = 0$  e multiplicando (4.36) por  $M$ , à direita, obtemos que  $X_f M$  tem todas as linhas iguais, o que também é absurdo pois  $X_f M = X_{r_2}$  é submatriz de  $X_r$ .

#### 4.2.2.1 - A Invariância do Método

O método 2 simplificado é invariante qualquer que seja o estimador  $b_f$  se a matriz  $X$  for tal que

$$r[X_{f_r}] = r[X_f] + r[X_r] - 1 \quad (4.40)$$

conforme Henderson et al., [1974]; continua valendo entretanto a restrição quanto à interação entre fatores fixos e fatores aleatórios e também para modelos em que há confundimento entre fatores fixos e aleatórios.

#### 4.2.2.2 - O Método 2 de Henderson

O método 2 descrito por Henderson [1953] é simplesmente uma forma do método 2 simplificado, que descrevemos a seguir.

O método inicialmente obtém as equações normais

$$\begin{bmatrix} 1'1 & 1' X_f & 1' X_r \\ X_f'1 & X_f' X_f & X_f' X_r \\ X_r'1 & X_r' X_f & X_r' X_r \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{b}_f \\ \hat{b}_r \end{bmatrix} = \begin{bmatrix} 1' y \\ X_f' y \\ X_r' y \end{bmatrix} \quad (4.41)$$

O problema fundamental é como resolvê-las de forma que a solução satisfaça (4.34), (4.35) e (4.36). Considerando apenas as soluções para as quais  $\hat{\mu} = 0$ , (4.41) se reduz a

$$[\underset{\sim}{X}'\underset{\sim}{X}] \begin{bmatrix} \hat{\underset{\sim}{b}}_f \\ \hat{\underset{\sim}{b}}_r \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'_f \underset{\sim}{X}_f & \underset{\sim}{X}'_f \underset{\sim}{X}_r \\ \underset{\sim}{X}'_r \underset{\sim}{X}_f & \underset{\sim}{X}'_r \underset{\sim}{X}_r \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{b}}_f \\ \hat{\underset{\sim}{b}}_r \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'_f \underset{\sim}{y} \\ \underset{\sim}{X}'_r \underset{\sim}{y} \end{bmatrix} \quad (4.42)$$

Sendo

$$\underset{\sim}{P} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1} = \begin{bmatrix} P_{\sim 11} & P_{\sim 12} \\ P_{\sim 21} & P_{\sim 22} \end{bmatrix}, \quad (4.43)$$

a solução de (4.42) é

$$\begin{bmatrix} \hat{\underset{\sim}{b}}_f \\ \hat{\underset{\sim}{b}}_r \end{bmatrix} = \underset{\sim}{P} \begin{bmatrix} \underset{\sim}{X}'_f \underset{\sim}{y} \\ \underset{\sim}{X}'_r \underset{\sim}{y} \end{bmatrix} \quad (4.44)$$

dando, portanto, para  $\hat{\underset{\sim}{b}}_f = \underset{\sim}{L} \underset{\sim}{y}$ ,

$$\underset{\sim}{L} = P_{\sim 11} \underset{\sim}{X}'_f + P_{\sim 12} \underset{\sim}{X}'_r = (P_{\sim 11} \ P_{\sim 12}) \underset{\sim}{X}' \quad (4.45)$$

Nem todas as soluções  $\underset{\sim}{L}$  dadas por (4.45) satisfazem (4.34), (4.35) e (4.36), e para que isso aconteça o método 2 de Henderson obtém  $\underset{\sim}{P}$  da seguinte maneira: em  $\underset{\sim}{X}'\underset{\sim}{X}$  é escolhida uma submatriz principal,  $\underset{\sim}{S}$ , de posto e ordem iguais ao posto de  $\underset{\sim}{X}$ . Sua inversa  $\underset{\sim}{S}^{-1}$  substitui  $\underset{\sim}{S}$  em  $\underset{\sim}{X}'\underset{\sim}{X}$ , elemento por elemento, e os elementos restantes de  $\underset{\sim}{X}'\underset{\sim}{X}$  são substituídos por zero. Até aqui o procedimento é o usual para obtenção de  $(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$ ; a particularidade no método 2 de Henderson é que as linhas e colunas de  $\underset{\sim}{X}'\underset{\sim}{X}$  descartadas para deixar  $\underset{\sim}{S}$  devem tanto quanto possível passar por

$X'_f X_f$ . Assim,  $\tilde{P}$  pode ser escrita como

$$\tilde{P} = \begin{bmatrix} P_{\tilde{f}} \\ P_{\tilde{r}} \end{bmatrix}$$

onde  $P_{\tilde{f}}$  e  $P_{\tilde{r}}$  são as linhas de  $\tilde{P}$  correspondentes respectivamente as linhas que passam por  $X'_f X_f$  e  $X'_r X_r$  de  $X'_X$ .

Com isso  $\tilde{L}$  de (4.45) pode ser escrita como

$$\tilde{L} = (P_{\tilde{f}} \ P_{\tilde{r}}) X' = P_{\tilde{f}} X' \quad (4.46)$$

A demonstração de que  $\tilde{L}$  assim obtida satisfaz (4.34), (4.35) e (4.36) é dada em Searle, [1968]- Apêndice.  $\tilde{L}$  pode, portanto, ser usada no método 2 simplificado, i.é., aplicamos o método 1 a  $\tilde{z} = \tilde{y} - X_f \tilde{L} \tilde{y} = \mu^* \tilde{1} + X_r b_r + (I - X_f \tilde{L}) \tilde{e}$ . No caso do modelo dado na eq. (4.2), chamando  $X_{\tilde{1}} = X_f$  e  $X_{\tilde{r}} = [X_2 \ X_3 \ X_4]$ , por exemplo,

$$R(\tilde{b}_3) = \tilde{z}' X_{\tilde{3}} (X'_{\tilde{3}} X_{\tilde{3}})^{-1} X'_{\tilde{3}} \tilde{z} \quad (4.47)$$

e

$$\begin{aligned} E[R(\tilde{b}_3)] &= \mu^{*2} \tilde{1}' X_{\tilde{3}} (X'_{\tilde{3}} X_{\tilde{3}})^{-1} X'_{\tilde{3}} \tilde{1} + \\ &+ \sum_{i=2}^4 \sigma_i^2 \text{tr} [X_{\tilde{3}} (X'_{\tilde{3}} X_{\tilde{3}})^{-1} X'_{\tilde{3}} X_i X_i'] \\ &+ \sigma_e^2 \text{tr} [(I - X_f \tilde{L})' X_{\tilde{3}} (X'_{\tilde{3}} X_{\tilde{3}})^{-1} X'_{\tilde{3}} (I - X_f \tilde{L})] \end{aligned} \quad (4.48)$$

Como  $\tilde{L} X_{\tilde{r}} = \tilde{0}$  então  $\tilde{L} X_{\tilde{3}} = \tilde{0}$  e o termo em  $\sigma_e^2$  simplifica, dando

$$\begin{aligned}
E[R(\underline{b}_3)] &= \mu^* 2 \underline{1}' \underline{X}_3 (\underline{X}'_3 \underline{X}_3)^{-1} \underline{X}'_3 \underline{1} + \sum_{i=2}^4 \sigma_i^2 \text{tr} [\underline{X}_3 (\underline{X}'_3 \underline{X}_3)^{-1} \underline{X}'_3 \underline{X}_3 \underline{X}'_3] + \\
&+ \sigma_e^2 r(\underline{X}_3) + \sigma_e^2 \text{tr} [\underline{L} \underline{L}' \underline{X}'_f \underline{X}_3 (\underline{X}'_3 \underline{X}_3)^{-1} \underline{X}'_3 \underline{X}_f] \quad (4.49)
\end{aligned}$$

Notar que os coeficientes dos  $\sigma_i^2$  são os mesmos obtidos pelo método 1, eq. (4.12), com exceção do de  $\sigma_1^2$  (fator fixo) que, naturalmente, é zero.

#### 4.2.3 - Exemplo Numérico

Um tratamento matricial dos dados do exemplo de Henderson [1953], visto em 4.1.3, envolve matrizes de ordem maior que 21. Por isso, optamos, para ilustração, por um exemplo fictício, porém esclarecedor dado em Searle [1968].

Temos 6 observações do modelo

$$Y_{ijK} = \mu + a_i + r_j + c_k + e_{ijK} \quad (4.50)$$

com  $i = 1, 2$   
 $j = 1, 2, 3$   
 $K = 1, 2$

(4.50) pode ser escrito como

$$\underline{y} = \mu \underline{1} + \underline{X}_1 \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \underline{X}_2 \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} + \underline{X}_3 \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \underline{e} \quad (4.51)$$

Numericamente:

$$\tilde{y} = \begin{bmatrix} 100 \\ 200 \\ 300 \\ 400 \\ 800 \\ 600 \end{bmatrix} = \begin{bmatrix} 1 & \underbrace{X_{\tilde{1}}} & \underbrace{X_{\tilde{2}}} & \underbrace{X_{\tilde{3}}} \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ a_1 \\ a_2 \\ r_1 \\ r_2 \\ r_3 \\ c_1 \\ c_2 \end{bmatrix} + \tilde{e} \quad (4.52)$$

O método 2 generalizado requer simplesmente que  $\tilde{b}_f = \tilde{L} \tilde{y}$ ,  $\tilde{L}$  seja inversa generalizada de  $\tilde{X}_f = \tilde{X}_{\tilde{1}}$ . Uma solução é:

$$\tilde{L}_{\tilde{1}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\tilde{z}_1$  é obtido como  $(\tilde{I} - \tilde{X}_{\tilde{1}} \tilde{L}_{\tilde{1}}) \tilde{y}$ , e

$$\tilde{z}_1' = [ 0 \quad 0 \quad 200 \quad 200 \quad 600 \quad 500 ]$$

O modelo para  $\tilde{z}$  é o de (4.25) obtido com (4.29), (4.30) e (4.31), e a ele é aplicado o método 1.

A questão de o método 2 generalizado não ser unicamente definido é exemplificada tomando outra matriz,  $\tilde{L}_{\tilde{2}}$

$$\tilde{L}_{\tilde{2}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

com a qual

$$\tilde{z}_2' = [ -500 \quad -600 \quad -300 \quad -400 \quad 0 \quad 0 ],$$

mostrando a dependência das soluções em relação a  $\underline{L}$ .

O método 2 simplificado não admite nem  $\underline{L}_1$ , nem  $\underline{L}_2$  como soluções, pois não satisfazem (4.34):  $\underline{X}'_f \underline{L} \underline{X}_r = 0$ . Uma matriz  $\underline{L}$  que satisfaz as três condições, é dada pelo método 2 de Henderson:

$$\underline{L} = \underline{P}_f \underline{X}' = \frac{1}{4} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 4 & -1 & -1 & 1 & 0 \end{bmatrix}, \quad (4.53)$$

onde  $\underline{P}_f$  é a partição de  $\underline{P}$  constituída pelas primeiras duas linhas,  $\underline{P} = \underline{S}^{-1}$  e  $\underline{S}$  é a matriz obtida suprimindo a primeira linha e coluna de  $\underline{X}'\underline{X}$ , e a seguir a primeira e última linhas e colunas.

Com esse resultado,

$$\underline{z} = \underline{y} - \underline{X}'_f \underline{L} \underline{y} = \begin{bmatrix} 100 \\ 200 \\ 300 \\ 400 \\ 800 \\ 600 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 0 \\ 600 \end{bmatrix} = \begin{bmatrix} 100 \\ 50 \\ 300 \\ 250 \\ 650 \\ 600 \end{bmatrix} \quad (4.54)$$

$$\text{Assim, } R(\underline{b}_3) = \underline{z}' \underline{X}_3 (\underline{X}'_3 \underline{X}_3)^{-1} \underline{X}'_3 \underline{z} = 701 \ 250 \quad (4.55)$$

e  $E[R(\underline{b}_3)]$  é obtida pela expressão (4.49), onde

$$\mu^* \underline{1} = (\mu + a_1) \underline{1}$$

e

$$\mu^{*2} \underline{1}' \underline{X}_3 (\underline{X}'_3 \underline{X}_3)^{-1} \underline{X}'_3 \underline{1} = 6(\mu + a_1)^2$$

Portanto,

$$E[R(\underline{b}_3)] = 6(\mu + a_1)^2 + 2,5 \sigma_2^2 + 6 \sigma_3^2 + (2 + \frac{21}{8}) \sigma_e^2 \quad (4.56)$$

De maneira análoga são obtidas as esperanças e os valores das demais reduções, formando o sistema de equações cuja solução são os componentes de variância.

#### 4.3 O MÉTODO 3 DE HENDERSON : MÉTODO DE AJUSTAR CONSTANTES

##### 4.3.1 - O Método

O método 3 consiste em usar as reduções na SQ, devido a ajustar o modelo e um sub-modelo que é parametrizado pelos fatores fixos, aplicando a elas o método 1, ou seja, iguala-se cada redução na SQ a seu valor esperado obtendo-se um sistema de equações cujas soluções são os estimadores dos componentes de variância.

O método, inicialmente descrito em Henderson [1953] em forma não matricial, é aqui tratado matricialmente.

O modelo geral

$$\underline{y} = \underline{X}\underline{b} + \underline{e} \quad (4.57)$$

será escrito como

$$\underline{y} = \underline{X}_1\underline{b}_1 + \underline{X}_2\underline{b}_2 + \underline{e} \quad (4.58)$$

e a redução na SQ devido a ajustar esse modelo será denotada por  $R(\underline{b}_1, \underline{b}_2)$ .

Ajustaremos também o sub-modelo

$$\underline{y} = \underline{X}_1\underline{b}_1 + \underline{e} \quad (4.59)$$

sendo a correspondente redução na SQ denotada  $R(\underline{b}_1)$ .

As esperanças destas duas reduções serão tomadas sob o modelo completo (4.58).

Denotaremos  $R(\underline{b}_1, \underline{b}_2) - R(\underline{b}_1)$  por  $R(\underline{b}_2 | \underline{b}_1)$  e mostraremos que  $E[R(\underline{b}_2 | \underline{b}_1)]$  envolve apenas  $\sigma_e^2$  e  $E(\underline{b}_2 \underline{b}_2')$ , e não envolve  $\underline{b}_1$ . (Nota-se que  $E(\underline{b}_2 \underline{b}_2') = \text{Var}(\underline{b}_2) + E(\underline{b}_2) E(\underline{b}_2')$ ).

A implicação é que se  $\underline{b}_1$  é o vetor de todos os fatores fixos e  $\underline{b}_2$  de todos os aleatórios, então  $E[R(\underline{b}_2 | \underline{b}_1)]$  envolverá somente fatores aleatórios e portanto os estimadores dos componentes de variância não serão afetados pelos fatores fixos.

$$\begin{aligned}
 \text{Temos que } E(\underline{y}' \underline{Q} \underline{y}) &= \text{tr}(\underline{Q} \underline{V}) + E(\underline{y}') \underline{Q} E(\underline{y}) \\
 &= \text{tr}[\underline{Q}\{\underline{X} \text{var}(\underline{b}) \underline{X}' + \sigma_e^2 \underline{I}\}] + \\
 &\quad + E(\underline{b}') \underline{X}' \underline{Q} \underline{X} E(\underline{b}) \\
 &= \text{tr}[\underline{Q}\{\underline{X} E(\underline{b} \underline{b}') \underline{X}'\} + \sigma_e^2 \underline{Q} - \underline{Q}\{\underline{X} E(\underline{b}) E(\underline{b}') \underline{X}'\}] + E(\underline{b}) \underline{X}' \underline{Q} \underline{X} E(\underline{b}) \\
 &= \text{tr}[\underline{X}' \underline{Q} \underline{X} E(\underline{b} \underline{b}')] + \sigma_e^2 \text{tr}(\underline{Q}) \tag{4.60}
 \end{aligned}$$

Ajustando o modelo completo a redução na SQ é

$$R(\underline{b}_1, \underline{b}_2) = \underline{y}' \underline{X} (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{y} \tag{4.61}$$

e por (4.60) a esperança de (4.61) é:

$$\begin{aligned}
 E[R(\underline{b}_1, \underline{b}_2)] &= \text{tr}\{(\underline{X}' \underline{X})^{-1} E(\underline{b} \underline{b}')\} + \sigma_e^2 r(\underline{X}) \\
 &= \text{tr} \left\{ \begin{bmatrix} \underline{X}' \underline{X}_{11} & \underline{X}' \underline{X}_{12} \\ \underline{X}' \underline{X}_{21} & \underline{X}' \underline{X}_{22} \end{bmatrix} E(\underline{b} \underline{b}') \right\} + \sigma_e^2 r(\underline{X}) \tag{4.62}
 \end{aligned}$$

Ajustando o sub-modelo (4.59) a redução na SQ é

$$R(\underline{b}_1) = \underline{y}' \left[ \begin{array}{c} \underline{X}_1' (\underline{X}_1' \underline{X}_1)^{-1} \underline{X}_1' \\ \underline{X}_2' \end{array} \right] \underline{y} \quad (4.63)$$

e a esperança de (4.63) é

$$\begin{aligned} E[R(\underline{b}_1)] &= \text{tr} \left\{ \underline{X}' \left[ \begin{array}{c} \underline{X}_1' (\underline{X}_1' \underline{X}_1)^{-1} \underline{X}_1' \\ \underline{X}_2' \end{array} \right] \underline{X} E(\underline{b} \underline{b}') \right\} + \sigma_e^2 r(\underline{X}_1) \\ &= \text{tr} \left\{ \begin{bmatrix} \underline{X}_1' \underline{X}_1 \\ \underline{X}_2' \underline{X}_1 \end{bmatrix} (\underline{X}_1' \underline{X}_1)^{-1} \begin{bmatrix} \underline{X}_1' \underline{X}_1 & \underline{X}_1' \underline{X}_2 \end{bmatrix} E(\underline{b} \underline{b}') \right\} + \sigma_e^2 r(\underline{X}_1) \\ &= \text{tr} \left\{ \begin{bmatrix} \underline{I} \\ \underline{X}_2' \underline{X}_1 (\underline{X}_1' \underline{X}_1)^{-1} \end{bmatrix} \begin{bmatrix} \underline{X}_1' \underline{X}_1 & \underline{X}_1' \underline{X}_2 \end{bmatrix} E(\underline{b} \underline{b}') \right\} + \sigma_e^2 r(\underline{X}_1) \\ &= \text{tr} \left\{ \begin{bmatrix} \underline{X}_1' \underline{X}_1 & \underline{X}_1' \underline{X}_2 \\ \underline{X}_2' \underline{X}_1 & \underline{X}_2' \underline{X}_1 (\underline{X}_1' \underline{X}_1)^{-1} (\underline{X}_1' \underline{X}_2) \end{bmatrix} E(\underline{b} \underline{b}') \right\} + \sigma_e^2 r(\underline{X}_1) \end{aligned} \quad (4.64)$$

Portanto a esperança de  $R(\underline{b}_2/\underline{b}_1)$  é

$$\begin{aligned} E[R(\underline{b}_2|\underline{b}_1)] &= E[R(\underline{b}_1, \underline{b}_2) - R(\underline{b}_1)] \\ &= \text{tr} \left\{ \underline{X}_2' \left[ \underline{I} - \underline{X}_1 (\underline{X}_1' \underline{X}_1)^{-1} \underline{X}_1' \right] \underline{X}_2 E(\underline{b}_2 \underline{b}_2') \right\} + \sigma_e^2 [r(\underline{X}) - r(\underline{X}_1)] \end{aligned} \quad (4.65)$$

e como afirmáramos não contem termos em  $\underline{b}_1$ .

## 4.3.2 A Aplicação do Método 3

Consideremos o modelo de 2 fatores com interação:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (4.66)$$

Supondo o modelo aleatório, temos quatro componentes de variância a serem estimados:  $\sigma_\alpha^2$ ,  $\sigma_\beta^2$ ,  $\sigma_{\alpha\beta}^2$  e  $\sigma_e^2$ , e para isso podemos usar as seguintes reduções na soma de quadrados ( tabela 4.6. )

Reduções na SQT	Fórmula Computacional	Valores esperados
$R(\alpha, \beta, \alpha\beta/\mu)$	$T_{AB} - T_\mu$	$h_1 \sigma_\alpha^2 + h_2 \sigma_\beta^2 + h_3 \sigma_{\alpha\beta}^2 + (s-1) \sigma_e^2$
$R(\beta, \alpha\beta/\mu, \alpha)$	$T_{AB} - T_A$	$h_4 \sigma_\beta^2 + h_5 \sigma_{\alpha\beta}^2 + (s-a) \sigma_e^2$
$R(\alpha\beta/\mu, \alpha, \beta)$	$T_{AB} - R(\mu, \alpha, \beta)$	$h_6 \sigma_{\alpha\beta}^2 + s^* \sigma_e^2$
SQE	$T_0 - T_{AB}$	$(N-s) \sigma_e^2$

Tabela 4.6

Os termos T são os usuais (equação 4.9),  $R(\mu, \alpha, \beta)$  é dado em Searle e Henderson [1961] e os coeficientes são obtidos como em 4.1.2.

Os estimadores são:

$$\begin{aligned}\hat{\sigma}_e^2 &= \text{SQE}/N-s \\ \hat{\sigma}_{\alpha\beta}^2 &= [R(\alpha\beta/\mu, \alpha, \beta) - (s-a-b+1)\hat{\sigma}_e^2]/h_6 \\ \hat{\sigma}_\beta^2 &= [R(\beta, \alpha\beta/\mu, \alpha) - h_5 \hat{\sigma}_{\alpha\beta}^2 - (s-a) \hat{\sigma}_e^2]/h_4 \\ \hat{\sigma}_\alpha^2 &= [R(\alpha, \beta, \alpha\beta/\mu) - h_2 \hat{\sigma}_\beta^2 - h_3 \hat{\sigma}_{\alpha\beta}^2 - (s-1) \hat{\sigma}_e^2]/h_1\end{aligned}\tag{4.67}$$

Um conjunto de equações alternativo que pode ser obtido substituindo-se  $R(\beta, \alpha\beta/\mu, \alpha)$  por  $R(\alpha, \alpha\beta/\mu, \beta)$ , está na tabela 4.7.

Redução na SQ	Fórmula Computacional	Valores esperados
$R(\alpha, \beta, \alpha\beta/\mu)$	$T_{AB} - T_\mu$	$h_1 \sigma_\alpha^2 + h_2 \sigma_\beta^2 + h_3 \sigma_{\alpha\beta}^2 + (s-1) \sigma_e^2$
$R(\alpha, \alpha\beta/\mu, \beta)$	$T_{AB} - T_B$	$h_7 \sigma_\alpha^2 + h_8 \sigma_{\alpha\beta}^2 + (s-b) \sigma_e^2$
$R(\alpha\beta/\mu, \alpha, \beta)$	$T_{AB} - R(\mu, \alpha, \beta)$	$h_6 \sigma_{\alpha\beta}^2 + s^* \sigma_e^2$
SQE	$T_0 - T_{AB}$	$(N-s) \sigma_e^2$

Tabela 4.7

No caso de o modelo ser misto, o fator fixo é a constante a ser ajustada (o vetor  $\underline{b}_1$  do caso geral). Nesse exemplo, se  $\alpha$  é o fator fixo, usamos as três últimas equações da tabela 4.6, que são as que tem  $\alpha$  como fator ajustado no sub-modelo, e se o fator fixo é  $\beta$ , usamos as três últimas equações da tabela 4.7, que são as que tem  $\beta$  ajustado no sub-modelo.

#### 4.3.2.1 - Exemplo numérico

Vejamos a aplicação do método ao exemplo de Henderson,

com os dados da tabela 4.2, supondo  $\gamma$  como fator fixo. De acordo com Searle [1968] as reduções poderiam ser:

---


$$\begin{aligned}
 R(\alpha, \alpha\beta / \beta, \gamma) &= R(\alpha, \beta, \gamma, \alpha\beta) - R(\beta, \gamma) \\
 R(\beta, \alpha\beta / \alpha, \gamma) &= R(\alpha, \beta, \gamma, \alpha\beta) - R(\alpha, \gamma) \\
 R(\alpha\beta / \alpha, \beta, \gamma) &= R(\alpha, \beta, \gamma, \alpha\beta) - R(\alpha, \beta, \gamma) \\
 \text{SQE} &= \sum_{ijkl} Y_{ijkl}^2 - R(\alpha, \beta, \gamma, \alpha\beta)
 \end{aligned}$$


---

Tabela 4.8

cuja computação é feita usando

$$\begin{aligned}
 R(\alpha, \beta, \gamma, \alpha\beta) &= \underline{y}' [\underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'] \underline{y} \\
 R(\alpha, \beta, \gamma) &= \underline{y}' [\underline{X}_{\alpha, \beta, \gamma}(\underline{X}'_{\alpha, \beta, \gamma} \underline{X}_{\alpha, \beta, \gamma})^{-1} \underline{X}'_{\alpha, \beta, \gamma}] \underline{y} \\
 R(\alpha, \gamma) &= \underline{y}' [\underline{X}_{\alpha, \gamma}(\underline{X}'_{\alpha, \gamma} \underline{X}_{\alpha, \gamma})^{-1} \underline{X}'_{\alpha, \gamma}] \underline{y} \\
 R(\beta, \gamma) &= \underline{y}' [\underline{X}_{\beta, \gamma}(\underline{X}'_{\beta, \gamma} \underline{X}_{\beta, \gamma})^{-1} \underline{X}'_{\beta, \gamma}] \underline{y}
 \end{aligned} \tag{4.68}$$

onde  $\underline{X}$  é a matriz de planejamento correspondente aos fatores  $\alpha, \beta, \gamma, \alpha\beta$  e cada uma das outras matrizes é a partição de  $\underline{X}$  com as colunas correspondentes aos fatores indicados nos índices.

Observemos que esse conjunto de reduções é uma das possibilidades. Outra, por exemplo, seria o mesmo conjunto de reduções, mas com  $R(\alpha, \beta, \alpha\beta / \mu, \gamma)$  ao invés de  $R(\alpha, \alpha\beta / \mu, \beta, \gamma)$ . Comentaremos mais adiante sobre essa multiplicidade de conjuntos de reduções.

Notar que os cálculos indicados em (4.68) na forma matricial podem ser dados em termos de fórmulas computacionais co

mo nas tabelas 4.6 e 4.7.

Conforme vemos na tabela 4.8 as reduções são sempre diferenças entre dois termos  $R(\ )$ , um dos quais corresponde ao modelo completo e o outro a um sub-modelo, o que é coerente com o resultado expresso pela equação (4.65), e frisado por Searle [1971] que faz algumas transformações algébricas, onde necessá-rio, para que os 2 termos  $R(\ )$  sejam dessa maneira. Henderson [1953], entretanto, faz as reduções incluindo um só fator por vez. Por exemplo,  $R(\beta, \alpha\beta / \mu, \alpha, \gamma)$  não é usado e sim  $R(\beta/\mu, \alpha, \gamma)$ , e dessa forma o cálculo da redução é feito por diferença entre 2 sub-modelos. Aqui poderia haver dois problemas: primeiro, os coeeficientes serem inadequados, o que não ocorre, pois a esperança é calculada para a redução que está sendo usada (isto é, os va- lores obtidos para as esperanças das reduções são outros, mas os coeficientes dos componentes de variância também são); e se- gundo, a esperança das diferenças entre duas reduções de sub-mo- delos conter fator fixo, o que também não ocorre, pois a expres- são que contém os fatores fixos, que é  $N \mu^2 + 2 \sum_K n_{..K} \mu \gamma_K + \sum_K n_{..K} \gamma_K^2$  está presente em todos os pares de reduções, e, quando é feita a diferença, fica cancelada. Observemos, aqui, que  $\mu$  faz parte da expressão que desaparece, sendo porisso indiferente se, na ta- bela 4.8,  $\mu$  entra ou não nas reduções.

Tendo isso em vista, vamos desenvolver o exemplo calcu- lando as seguintes reduções:

---


$$R(\alpha/\mu, \beta, \gamma) = R(\mu, \alpha, \beta, \gamma) - R(\mu, \beta, \gamma) \quad (2 \text{ sub-modelos})$$

$$R(\beta/\mu, \alpha, \gamma) = R(\mu, \alpha, \beta, \gamma) - R(\mu, \alpha, \gamma) \quad (2 \text{ sub-modelos})$$

$$R(\alpha\beta / \mu, \alpha, \beta, \gamma) = R(\mu, \alpha, \beta, \gamma, \alpha\beta) - R(\mu, \alpha, \beta, \gamma) \quad (\text{modelo e sub- modelo})$$

$$SQE = \sum Y_{ijK\ell}^2 = R(\mu, \alpha, \beta, \gamma, \alpha\beta)$$


---

Tabela 4.9

O cálculo desses valores pode ser feito usando as correspondentes formas quadráticas, como indicado em (4.68) ou, diretamente, as fórmulas computacionais como as das tabelas 4.6 e 4.7.

A escolha desse conjunto de reduções é uma escolha adequada, pois

$$E[R(\alpha/\mu, \beta, \gamma)] = E[R(\mu, \alpha, \beta, \gamma) - R(\mu, \beta, \gamma)] \text{ envolve } \sigma_{\alpha}^2, \sigma_{\alpha\beta}^2 \text{ e } \sigma_e^2$$

$$E[R(\beta/\mu, \alpha, \gamma)] = E[R(\mu, \alpha, \beta, \gamma) - R(\mu, \alpha, \gamma)] \text{ envolve } \sigma_{\beta}^2, \sigma_{\alpha\beta}^2 \text{ e } \sigma_e^2$$

$$E[R(\alpha\beta/\mu, \alpha, \beta, \gamma)] = E[R(\mu, \alpha, \beta, \gamma, \alpha\beta) - R(\mu, \alpha, \beta, \gamma)] \text{ envolve } \sigma_{\alpha\beta}^2 \text{ e } \sigma_e^2$$

e  $E[SQE]$  só envolve  $\sigma_e^2$ .

Quanto aos coeficientes dos componentes de variância nas  $E[R(\cdot)]$  podem ser obtidos pela fórmula (4.12) ou diretamente com o uso das expressões matriciais dadas em (4.62) e (4.64).

Os valores das reduções e os coeficientes dos componentes de variância nas esperanças dessas reduções, assim obtidos, estão na tabela 4.10.

Reduções	Coeficientes de				Valor das Reduções
	$\sigma_{\alpha}^2$	$\alpha_{\beta}^2$	$\sigma_{\alpha\beta}^2$	$\sigma_e^2$	
$R(\alpha/\mu, \beta, \gamma)$	34,43	0	15,72	3	120.428
$R(\beta/\mu, \alpha, \gamma)$	0	14,71	6,58	2	1.409
$R(\alpha\beta/\mu, \alpha, \beta, \gamma)$	0	0	18,71	4	52.411
SSE	0	0	0	44	150.489

Tabela 4.10

Igualando-se cada esperança com a redução na S.Q. correspondente, obtemos um conjunto de 4 equações, cuja solução é:

$$\begin{aligned}\hat{\sigma}_e^2 &= 3.420 \\ \hat{\sigma}_{\alpha\beta}^2 &= 2.070 \\ \hat{\sigma}_\beta^2 &= -1.295 \\ \hat{\sigma}_\alpha^2 &= 2.255\end{aligned}\tag{4.69}$$

Essas são as estimativas dos componentes de variância, e obtivemos um valor negativo para  $\sigma_\beta^2$ , o que nos coloca novamente na discussão feita em 3.1.2.

#### 4.3.2.2 - Um tratamento Matricial

Um tratamento matricial do exemplo de Henderson envolve matrizes de ordem superior a 21; entretanto, ao exemplo explorado em 4.2.3, de Searle [1968], aplicamos agora o método 3. Um conjunto adequado de reduções é

$$\begin{aligned}R(r/a, c) &= R(a, r, c) = R(a, c) \\ R(c/a, r) &= R(a, r, c) - R(a, r)\end{aligned}\tag{4.70}$$

$$\text{e } SQE = \underline{y}'\underline{y} - R(a, r, c)$$

pois

$$\begin{aligned}E[R(r/a, c)] &\text{ envolve } \sigma_e^2 \text{ e } \sigma_r^2 \\ E[R(c/a, r)] &\text{ envolve } \sigma_e^2 \text{ e } \sigma_c^2\end{aligned}\tag{4.71}$$

$$\text{e } E[SQE] \text{ envolve } \sigma_e^2 .$$

Os calculos, feitos como em (4.68) fornecem:

$$R(a,r,c) = \underline{y}' \underline{X}_{\sim arc} (X'_{\sim arc} X_{\sim arc})^{-1} X'_{\sim arc} \underline{y} = 1.290.000 \quad (4.72)$$

$$R(a,c) = \underline{y}' \underline{X}_{\sim ac} (X'_{\sim ac} X_{\sim ac})^{-1} X'_{\sim ac} \underline{y} = \frac{12.650.000}{12} \quad (4.73)$$

$$e \quad R(r/a,c) = \frac{2.830.000}{12} \quad (4.74)$$

O valor esperado de  $R(r/a,c)$  é, por (4.65),

$$\begin{aligned} E[R(r/a,c)] &= \text{tr}\{X'_{\sim 2} X_{\sim 2} - X'_{\sim 2} [X_{\sim 1}, X_{\sim 3}] \left[ \begin{array}{c} X'_{\sim 1} \\ X'_{\sim 3} \end{array} [X_{\sim 1}, X_{\sim 3}] \right]^{-1} \left[ \begin{array}{c} X'_{\sim 1} \\ X'_{\sim 3} \end{array} \right] X_{\sim 2}\} \sigma_r^2 + \\ &+ \sigma_e^2 \{r[X_{\sim 1}, X_{\sim 2}, X_{\sim 3}] - r[X_{\sim 1}, X_{\sim 3}]\} = \\ &= 4,5 \sigma_r^2 + 2 \sigma_e^2 . \end{aligned} \quad (4.75)$$

Logo, com  $\hat{\sigma}_e^2 = (\underline{y}' \underline{y} - 1.290.000)/(6-5) = 10.000$ ,

$$\hat{\sigma}_r^2 = 47.962,963$$

A determinação de  $\sigma_c^2$  é analoga, usando-se  $R(c/a,r)$ .

#### 4.3.3 - Vantagens e Desvantagens do Método 3

Como vimos pela equação (4.65), a  $E[R(b_2/b_1)]$  não contém termos em  $b_1$ , e dessa forma os estimadores dos componentes de variância obtidos são não viesados pelos fatores fixos tornando o método 3 o ideal para o caso de modelo misto (lembramos

que o procedimento de estimar os fatores fixos e então estimar os componentes de variância usando os estimadores obtidos para os fatores fixos, que é o método 2, embora produza estimadores de componentes de variância não viesados nem sempre pode ser usado e nem sempre produz uma estimativa única). Além disso, se o modelo é misto,  $E(b_{\tilde{2}}/b_{\tilde{1}})$  não contém termos que envolvam  $\text{var}(b_{\tilde{1}})$  e, mais importante, não contém termos envolvendo covariância entre elementos de  $b_{\tilde{1}}$  com  $b_{\tilde{2}}$ . Essa, portanto, é uma grande vantagem do método, pois, mesmo que no modelo haja correlação entre elementos de  $b_{\tilde{1}}$  e  $b_{\tilde{2}}$ , a  $E(b_{\tilde{2}}/b_{\tilde{1}})$  não envolve essa correlação.

A grande desvantagem do método 3 é o relativamente grande volume de cálculo, mesmo no caso de usarmos métodos abreviados como o de síntese de Hartley, descrito em 4.1.4. É verdade, contudo, que o advento dos computadores simplificou um pouco a questão e que há perspectiva de que esse aspecto torne-se cada vez menos significativo (pacotes, etc). Notemos, a título de ilustração, que no exemplo numérico de 4.3.2.1 surgem operações com matrizes de ordem maior que 24.

A grande questão em aberto, contudo, a respeito do método 3 é que o número de equações de que podemos dispor é maior que o número de componentes de variância a estimar, e não há indicação de qual conjunto de equações usarmos. Nas tabelas 4.6 e 4.7 temos 2 conjuntos de equações para estimar os mesmos componentes de variância e os estimadores de  $\sigma_e^2$  e  $\sigma_{\alpha\beta}^2$  são os mesmos usando qualquer das 2 tabelas, mas  $\hat{\sigma}_\alpha^2$  e  $\hat{\sigma}_\beta^2$ , não. Além disso quanto maior o número de fatores do modelo, maior ainda o número de conjuntos de equações (No caso de 3 fatores há seis conjuntos). No caso de 2 fatores Harville [1967] e Low [1964] propõe usarmos as últimas 2 linhas da tabela 4.6 e 4.7 que são iguais, e mais a 2ª linha de cada tabela.

Um procedimento que poderíamos usar é o de empregar reduções cuja soma seja  $SQT_M = \sum y^2 - T_{11}$ . Com isso teríamos estimadores provenientes de reduções que medem a variabilidade total dos valores  $y$ .

A verdade é que, como essas soluções para casos particulares indicam, não há um procedimento geral para escolher o conjunto de equações adequado.

### CONCLUSÃO

Este trabalho foi baseado principalmente no artigo de Henderson [1953] e no de Searle [1968], tendo este último consistido em um aprimoramento do primeiro, introduzindo um tratamento matricial e generalizando alguns tópicos, como o método 2.

Embora nosso objetivo último tenha sido a estimação de componentes de variância, e em particular os três métodos de Henderson, procuramos situar o assunto dentro do contexto mais amplo da análise de variância. E dentro deste contexto há muitos tópicos, que poderiam "de per si" constituir-se em tema de outro trabalho.

Assim, em relação a modelos fixos com dados não balanceados, o trabalho de Speed, Hocking e Hackney [1978] poderia constituir-se num ponto de partida, devido às numerosas indicações que apresenta dos vários métodos de análise, e com as hipóteses que podem ser testadas em cada um. Nesse mesmo aspecto, o trabalho já citado de Gosslee e Lucas [1965], também é substancial.

Os vários métodos para estimação de  $\underline{b}$  que têm sido desenvolvidos recentemente apresentam, também, campo para novos trabalhos. Nesse particular, podemos citar os métodos de máxima

verossimilhança, bem como modificações dos mesmos, e também os métodos Bayesianos e pseudo-Bayesianos (embora esses últimos não contem com a aprovação geral - in Searle [1968] - discussão).

Um tópico que nos pareceu insuficientemente explorado foi a questão de posto completo, incompleto e reparametrização.

Em relação a modelos com componentes de variância e dados não balanceados, trabalhos envolvendo testes de hipóteses, parece-nos, seriam uma complementação deste nosso trabalho de estimação.

## BIBLIOGRAFIA

- ANDERSON, R.L. 1960. Uses of variance component analysis in the interpretation of biological experiments. *Bulletin de l'Institut International de Statistique*, 37(3):71-90.
- BEZERRA, R.C.F. 1976. *Componentes de variancia o problema de estimativas negativas*. Sao Paulo. 68p. Dissertação (Mestrado) - IME-USP.
- BLISCHKE, W.R. 1966. Variances of estimates of variance components in a three way classification. *Biometrics*, 22(3):553-565.
- CARLSON, J.E. & TIMM, N.H. 1974. Analysis of nonorthogonal fixed effects designs. *Psychological Bulletin*, 81:563-570.
- COCHRAN, W.G. 1951. Testing a linear relation among variances. *Biometrics*, 7(1):17-32.
- CUMMINGS, W.B. & GAYLOR, D.W. 1974. Variance component testing in unbalanced nested design. *Journal of the American Statistical Association*, 69(347):765-771.
- EISEN, E.J. 1966. The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model. *Biometrics*, 22(4):937-942.
- EISENHART, C. 1947. The assumptions underlying the analysis of variance. *Biometrics*, 3(1):1-21.
- FEDERER, W.T. 1955. *Experimental design: theory and application*. New York, Macmillan. 590p.

- FEDERER, W.T. 1968. Non-negative estimators for components of variance. *Applied Statistics*, 17(2):171-174.
- GAYLOR, D.W. & HOPPER, F.N. 1969. Estimating the degrees of freedom from linear combinations of mean squares by Satterthwaite's formula. *Technometrics*, 11(4):691-706.
- GOSSLEE, D.G. & LUCAS, H.L. 1965. Analysis of variance of disproportionate data when interaction is present. *Biometrics*, 21(1):115-133.
- GOWER, J.C. 1962. Variance component estimation for unbalanced hierarchical classifications. *Biometrics*, 18(4):537-542.
- GRAYBILL, F.A. 1954. On quadratic estimates of variance components. *Annals of Mathematical Statistics*, 25(2):367-372.
- GRAYBILL, F.A. & HULTQUIST, R.A. 1961. Theorems concerning Eisenhart's model II. *Annals of Mathematical Statistics*, 32(1):261-269.
- GRAYBILL, F.A. & WORTHAM, A.W. 1956. A note on uniformly best unbiased estimators for variance components. *Journal of the American Statistical Association*, 51(274):266-268.
- HARTLEY, H.O. 1967. Expectations, variances and covariances of ANOVA mean squares by 'synthesis'. *Biometrics*, 23(1):105-114. Corrections, *Biometrics*, 23(4):823, 1967.
- HARTLEY, H.O. & HOCKING, R.R. 1971. The analysis of incomplete data. *Biometrics*, 27(4):783-823.
- HARVEY, W.R. 1968. Least squares analysis of data with unequal subclass numbers. Agricultural Research Service, ARS, 20-8, July.

- HARVILLE, D.A. 1967. Estimability of variance components for the two-way classification with interaction. *Annals of Mathematical Statistics*, 38(5):1508-1519.
- HARVILLE, D.A. 1978. Alternative formulations and procedures for the two-way mixed model. *Biometrics*, 34:441-453.
- HEMMERLE, W.J. 1974. Nonorthogonal analysis of variance using iterative improvement and balanced residuals. *Journal of the American Statistical Association*, 69(347):772-778.
- HEMMERLE, W.J. 1979. Balanced hypothesis and unbalanced data. *Journal of the American Statistical Association*, 74(368):794-798.
- HEMMERLE, W.J. 1980. Recognizing balance with unbalanced data. *Communications in Statistics Theory and Methods*, A9(2):201-211.
- HENDERSON, C.R. 1953. Estimation of variance and covariance components. *Biometrics*, 9(2):226-252.
- HENDERSON, C.R.; SEARLE, S.R.; SCHAEFFER, L.R. 1974. The invariance and calculation of method 2 for estimating variance components. *Biometrics*, 30(4):583-588.
- HERBACH, L.H. 1959. Properties of model II-type analysis of variance tests, A: optimum nature of the F-test for model II in the balanced case. *Annals of Mathematical Statistics*, 30(4):939-959.
- JEYARATNAM, S. & GRAYBILL, F.A. 1980. Confidence intervals on variance components in three-factor cross classification models. *Technometrics*, 22(3):375-380.
- JOHN, F.W.M. 1971. *Statistical design and analysis of experiments*. New York, Macmillan. 356p.

- KOCH, G.G. 1967. A general approach to the estimation of variance components. *Technometrics*, 9(1):93-118.
- KOCH, G.G. 1967. A procedure to estimate the population mean in random effects models. *Technometrics*, 9(4):577-585.
- KOCH, G.G. 1968. Some further remarks concerning "a general approach to the estimation of variance components". *Technometrics*, 10(3):551-558.
- LAMOTTE, L.R. 1973. On non-negative quadratic unbiased estimation of variance components. *Journal of the American Statistical Association*, 68(343):728-730.
- LOW, L.Y. 1964. Sampling variances of estimates of components of variance from a non-orthogonal two-way classification. *Biometrika*, 51(3/4):491-494.
- MENDENHALL, W. 1968. *Introduction to linear models and the design and analysis of experiments*. Belmont, Wadsworth. 465p.
- OVERALL, J.E. & KLETT, J.C. 1972. *Applied multivariate analysis*. New York, McGraw-Hill. 500p.
- RAO, J.N.K. 1968. On expectations, variances and covariances of ANOVA mean squares by 'synthesis'. *Biometrics*, 24(4):963-978.
- SAHAI, H. 1979. A bibliography on variance components. *International Statistical Review*, 47(2):177-222.
- SATTERTHWAITE, F.E. 1946. An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2:110-114.
- SCHEFFE, H. 1959. *The analysis of variance*. New York, John Wiley. 477p.

- SEARLE, S.R. 1966. *Matrix algebra for the biological sciences: including applications in statistics*. New York, John Wiley. 296p.
- SEARLE, S.R. 1968. Another look at Henderson's methods of estimating variance components. *Biometrics*, 24(4):749-787.
- SEARLE, S.R. 1971. *Linear models*. New York, John Wiley. 532p. (Wiley Series in Probability and Mathematical Statistics)
- SEARLE, S.R. 1971. Topics in variance component estimation. *Biometrics*, 27(1):1-76.
- SEARLE, S.R. & HENDERSON, C.R. 1961. Computing procedures for estimating components of variance in the two-way classification, mixed model. *Biometrics*, 17(4):607-616. Corrections, *Biometrics*, 23(4):852, 1967.
- SEBER, G.A.F. 1977. *Linear regression analysis*. New York, John Wiley. 465p.
- SEEGER, P. 1970. A method of estimating variance components in unbalanced designs. *Technometrics*, 12(2):207-218.
- SNEDECOR, G.W. & COCHRAN, W.G. 1967. *Statistical methods*. 6.ed. Ames, Iowa State University Press. 503p.
- SPEED, F.M. & HOCKING, R.R. 1976. The use of the R()-notation with unbalanced data. *The American Statistician*, 30(1):30-33.
- SPEED, F.M.; HOCKING, R.R.; HACNEY, O.P. 1978. Methods of analysis of linear models with unbalanced data. *Journal of the American Statistical Association*, 73(361):105-112.

- THOMPSON JR., W.A. 1961. Negative estimates of variance components: an introduction. *Bulletin de l'Institut International de Statistique*, 39(3):181-184.
- THOMPSON, JR., W.A. 1962. The problem of negative estimates of variance components. *Annals of Mathematical Statistics*, 33(1):273-289.
- THOMPSON, JR., W.A. & MOORE, J.R. 1963. Non-negative estimates of variance components. *Technometrics*, 5(4):441-449.
- TIETJEN, G.L. 1974. Exact and approximate tests for unbalanced random effects designs. *Biometrics*, 30(4):573-581.
- TIETJEN, G.L. & MOORE, R.H. 1968. On testing significance of components of variance in the unbalanced nested analysis of variance. *Biometrics*, 24(2):423-429.
- YATES, F. 1934. The analysis of multiple classifications with unequal numbers in the different classes. *Journal of the American Statistical Association*, 29(185):51-66.