ESTIMABILIDADE NO MODELO LINEAR EM CLASSIFICAÇÃO
HIERÁRQUICA COM s ESTÁGIOS

CARLOS ROBERTO PADOVANI

Orientador: Prof. Dr. DÉCIO BARBIN

Tese apresentada à Escola Superior de Agricultura "Luiz de Queiroz", da Universidade de São Paulo, para obtenção do título de Doutor em Agronomia. Área de concentração: Estatística e Experimentação Agronômica.

P I R A C I C A B A
Estado de São Paulo - Brasil
Janeiro, 1984

À memória de meus pais:

Carlos e Conceição,

DEDICO

À minha esposa,

Sílvia Helena,

e aos meus filhos:

Carlos Roberto,

Flávia Helena,

Paulo Roberto,

OFEREÇO

AGRADECIMENTOS

Ao Dr. *Décio Barbin*, Professor Titular do Departamento de Matemática e Estatística da ESALQ-USP, pela constante e eficiente orientação.

Aos *Professores* do Departamento de Matemática e Estatística da ESALQ-USP, pelos ensinamentos, apoios e incentivos oferecidos.

À *CAPES/PICD*, pelo auxílio financeiro, concedido para a realização do programa de Doutorado.

À Dra. *Sheila Zambello de Pinho*, Professora Assistente Doutora do Departamento de Bioestatística do IBBMA-UNESP, pelo auxílio na parte computacional e sugestões proporcionadas.

À Dra. *Martha Maria Mischán*, Professora Assistente Doutora do Departamento de Bioestatística do IBBMA-UNESP, pela leitura do manuscrito e sugestões dispensadas.

À Dra. *Maria Aparecida Mourão Brasil*, Professora Assistente Doutora do Departamento de Agricultura e Silvicultura da FCA-UNESP, pela cessão dos dados.

Aos *Colegas* do Departamento de Bioestatística do IBBMA-UNESP, pelo apoio, incentivo, compreensão e, principalmente, pelo ambiente de trabalho.

Às Funcionárias do Departamento de Bioestatística do IBBMA-UNESP, *Laura Aparecida Lopes* e *Célia Maria Nogueira da Silva*, pela colaboração prestada e *Elza Ivone de Almeida Mota*, pelo serviço datilográfico.

Ao Sr. *Décio Ranzani da Silva*, Professor Efetivo de Português do Ensino Oficial do Estado de São Paulo, pela revisão final do texto.

A todos aqueles que, de uma forma ou outra, contribuíram para a realização deste trabalho.

Í N D I C E

	Página
1. INTRODUÇÃO	1
2. REVISÃO DA LITERATURA	3
3. DESENVOLVIMENTO TEÓRICO	14
3.1 - Modelo Matemático	14
3.2 - Forma Matricial	18
3.3 - Equações Normais	22
3.4 - Solução de Mínimos Quadrados	26
3.5 - Análise de Variância	29
3.6 - Funções Estimáveis	34
3.7 - Teste da hipótese $q'b = m$	41
4. PROGRAMA COMPUTACIONAL	45
4.1 - Características do Programa	45
4.2 - Entrada de Dados	50
4.3 - Listagem do Programa	52
5. APLICAÇÃO	58
6. CONCLUSÕES	72
7. REFERÊNCIAS BIBLIOGRÁFICAS	73
8. APÊNDICE	76

ESTIMABILIDADE NO MODELO LINEAR EM CLASSIFICAÇÃO
HIERÁRQUICA COM s ESTÁGIOS

Autor: Carlos Roberto Padovani

Orientador: Prof. Dr. Décio Barbin

RESUMO

Considera-se o modelo linear em classificação hierárquica, com s -estágios, onde as subclasses em cada um dos s -estágios, podem possuir um número diferente de unidades experimentais.

Através do conceito de função linear paramétrica estimável no contexto de modelo linear, estabelece-se um teorema que resulta numa regra simples e prática para identificar a estimabilidade no modelo considerado. Baseado nesse teorema, conclui-se que são estimáveis as diferenças entre duas médias de classes e todas as diferenças entre duas médias de subclasses de um j -ésimo estágio, dentro da subclasse do $(j-1)$ -ésimo estágio.

Faz-se ainda a programação computacional da metodologia apresentada e ilustra-se o desenvolvimento teórico com uma aplicação a dados de comprimentos de fibras de madeira.

ESTIMABILITY IN LINEAR MODEL IN HIERARCHICAL
CLASSIFICATION WITH s STAGES

Author: Carlos Roberto Padovani,

Adviser: Prof. Dr. Dêcio Barbin

SUMMARY

The linear model with hierarchical classification in s stages with possible different sub-unit numbers of observations is considered in this work.

Through estimability of linear function of parameters concept, a theorem is established resulting in a simple and practical rule under which estimability of the parameters of the model can be identified. Based on this theorem, the estimability of the differences among classes averages and all differences between two j^{th} sub-classes within a the $(j-1)^{\text{th}}$ sub-class can be obtained.

Computer programs for the methodology are presented and the theoretical development is exemplified by an application to wood fiber measurements.

1. INTRODUÇÃO

Entre os modelos lineares usados no campo da Estatística Experimental, encontra-se o modelo em classificação hierárquica, ou classificação aninhada. Contudo, quando a experimentação é feita com mais de três estágios, não são encontradas regras práticas sobre as condições de estimabilidade para as funções lineares dos parâmetros envolvidos no modelo matemático desse delineamento e, conseqüentemente, quais dessas funções são testáveis.

Em modelos de até três estágios, existem, na literatura, textos que apresentam quadros, contendo os tipos das funções estimáveis. O julgamento da estimabilidade de uma função linear paramétrica por um pesquisador é feito comparando-se a função dada com as existentes nesses quadros. Tal procedimento pode não ser o mais prático, para se verificar a estimabilidade de uma função. O ideal é estabelecer as condições de estimabilidade e verificar se são satisfeitas para a função

proposta.

Baseando-se nestes fatos, o trabalho é desenvolvido, tendo em vista um número qualquer de estágios e considerando o experimento com dados não-balanceados (o balanceado passa a ser um caso particular), com o objetivo de estabelecer as condições de estimabilidade para as funções lineares paramétricas. Uma consequência imediata do estabelecimento dessas condições é a identificação das funções testáveis.

Usando-se esses resultados, faz-se um programa computacional em FORTRAN IV, que verifica se a função proposta pelo experimentador é estimável e, caso seja, faz-se o respectivo teste de hipótese. Este é um aspecto de grande interesse prático, pois um pesquisador, mesmo não sendo um especialista da área de Experimentação, pode fazer uso do programa, sem o risco de estar testando hipóteses que não são testáveis.

Finalmente, com o objetivo de ilustrar o desenvolvimento teórico e evidenciar o manuseio do programa computacional, faz-se uma aplicação considerando um experimento com medida do comprimento de fibras de eucalipto, composto de quatro estágios.

2. REVISÃO DA LITERATURA

SEARLE (1961) considera o modelo de componentes de variância para a classificação hierárquica com dois critérios (estágios), não-balanceado, e deduz as expressões das variâncias das estimativas dos componentes de variância. Para esta dedução, utiliza resultados, que já havia obtido em outros dois artigos (1956 e 1958), para esse modelo. O autor mostra ainda que as expressões obtidas para as variâncias se tornam as já bem conhecidas, se se considera o experimento com dados balanceados.

GATES e SHIUE (1962) consideram o modelo de componentes de variância para o delineamento de classificação hierárquica com s-estágios, não-balanceado. Os autores, usando os resultados obtidos por Ganguli (1941), são que de uma forma mais compacta, estabelecem uma regra para determinar os coeficientes dos componentes das variâncias das expressões das es-

peranças dos quadrados médios das causas de variações. Após estabelecerem fórmulas dos coeficientes dos componentes das variâncias, são consideradas algumas situações particulares para esse modelo de componentes de variância e mostram-se as simplificações que surgem nestas fórmulas. Os autores enfocam ainda dois aspectos interessantes. O primeiro é que, em 1961, Graybill apresenta essa forma compactada, mas considerando o delineamento de classificação hierárquica não-balanceado apenas com dois estágios. O segundo é que Pulley (1957) desenvolveu um programa computacional para calcular as somas de quadrados, quadrados médios e todos os coeficientes das esperanças dos quadrados médios para a análise de variância do delineamento de classificação hierárquica com até quatro estágios.

GOWER (1962) aborda os coeficientes dos componentes de variância considerando o modelo de componentes de variância aninhado, não-balanceado, com h níveis hierárquicos. O autor desenvolve um método sistemático para calcular os coeficientes dos componentes de variância, nas esperanças dos quadrados médios de uma análise de variância hierárquica. Ilustra este método com a aplicação em uma análise de variância com $h = 4$ níveis e finaliza discutindo brevemente a programação computacional do método desenvolvido.

MAHAMUNULU (1963) considera os resultados obtidos por Searle (1956, 1958 e 1961) e estende esses resultados para o caso do modelo de componentes de variância para a clas

sificação hierárquica a três critérios. A ordem de apresentação, em seu artigo, é exatamente a mesma da apresentada por Searle, para o hierárquico a dois critérios.

EISEN (1966) considerou o modelo misto hierárquico, não-balanceado, $y_{ijkm} = \mu + G_i + S_{j(i)} + D_{k(ij)} + e_{m(ijk)}$, onde o efeito médio geral μ e o efeito de "line" G_i são fixos e os demais, aleatórios. O autor constrói o quadro da análise de variância para esse modelo e, por inspeção das esperanças dos quadrados médios, nota que não existe uma maneira direta de testar a significância do efeito de "line". Baseado nisso, propõe um método que permite a obtenção de um teste de significância, aproximado do efeito fixo principal não-aninhado, baseado em uma análise de variância aninhada não-balanceada. O processo usado para obtenção do teste de significância, segundo o autor, é o indicado por Anderson, em 1960, o qual está baseado nas investigações de Cochran (1951) e de Satterthwaite (1946). Conclui também não parecer que se possa obter uma expressão geral para o denominador da estatística, usada no teste de significância, fato que impediria uma generalização do modelo para um número qualquer de aninhamentos.

ANDERSON e CRUMP (1967) consideram o modelo de componentes de variância para a classificação hierárquica, com dois estágios, e discutem procedimentos de estimação para a média geral μ , o componente de variância entre classes σ_a^2 , o

componente de variância dentro de classes σ_b^2 e $\rho = \frac{\sigma_a^2}{\sigma_b^2}$. Os

autores comparam delineamentos de classificação hierárquica com dois estágios com o objetivo de encontrar o número ótimo de classes e de observações por classe para estimar σ_a^2 e ρ .

KUSSMAUL & ANDERSON (1967) consideram que cada observação Y_{ij} é descrita por $Y_{ij} = m + A_i + B_{ij}$, $i = 1, \dots, a$ e $j = 1, \dots, r_i$, onde m é um efeito fixo, A_i é o efeito correspondente à i -ésima classe, e B_{ij} é o efeito correspondente a uma composição amostral, isto é,

$$B_{ij} = \sum_{k=1}^{n_{ij}} B_{ijk},$$

onde B_{ijk} representa o indivíduo amostral, antes do agrupamento. Os autores supõem A_i e B_{ijk} normais, independentemente distribuídos, com médias nulas e variâncias σ_A^2 e σ_B^2 , respectivamente, e discutem três processos de estimação: método das médias ponderadas, método das médias das classes não ponderadas e método das médias não ponderadas. Os parâmetros estimados são: μ , σ_A^2 , σ_B^2 e $\sigma_A^2 + \sigma_B^2$.

TIAO e BOX (1967) consideram o modelo de componentes de variância para o delineamento de classificação hierárquica, balanceado, com três componentes: duas diferentes espécies de efeitos aleatórios, a_i e b_{ij} e e_{ijk} um erro alea-

tório ou distúrbio. Os autores fazem inferências sobre os pa
râmetros do modelo do ponto de vista Bayesiano, que são: dis
cussão de várias características da distribuição "a posteriori" dos componentes de variância $\sigma_1^2 = \text{var}(e_{ijk})$, $\sigma_2^2 = \text{var}(b_{ij})$
 e $\sigma_3^2 = \text{var}(a_i)$, contribuição relativa dos componentes, magnitude dos componentes individuais e solução para o problema da estimação das variâncias agrupadas. Os autores esclarecem ainda que o delineamento de classificação hierárquica com dois componentes já fora abordado, nesse aspecto, por Hill (1965) e Tiao e Tan (1965). Os autores desenvolvem ainda um método de aproximação para os componentes de variância individuais, o qual pode ser prontamente estendido para o modelo hierárquico com q-componentes.

GAYLOR e HARTWELL (1969) consideram uma população de A classes, B subclasses dentro de cada uma das A clas
ses, e C subsubclasses dentro de cada uma das B subclasses. Nota-se que a população é balanceada e que nenhuma restrição é feita no sentido de que A, B e C não possam ser infinitos. Baseados nessa população, eles consideram o modelo de compo
nentes de variância aninhado com três níveis:

$$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \delta_{k(ij)} + \rho_{l(ijk)}$$

e são selecionados casualmente a amostras das A classes, b_i das B subclasses da i-ésima classe de A, e c_{ij} das C subsubclasses

da j -ésima classe de B, dentro da i -ésima classe amostrada. São também amostradas casualmente n_{ijk} observações entre um número infinito de observações possíveis. Observa-se também que, embora a população tenha sido considerada balanceada, a amostra não necessariamente o será. São, então, calculadas as esperanças dos quadrados médios para as fontes de variações. Embora as esperanças sejam condicionais, por causa da aleatoriedade do número de amostras obtidas, elas não o serão quando se considera uma qualquer particular amostra, pois a , b_i , s , c_{ij} , s e n_{ijk} , s são considerados fixos para essa amostra.

SEARLE (1971) estabelece as funções lineares paramétricas estimáveis, considerando o modelo linear de classificação hierárquica com dois estágios para o caso não-balanceado. O autor discute ainda vários outros aspectos relacionados com este modelo, que são solução de mínimos quadrados, análise da variância, estimadores e variâncias dos estimadores das funções lineares, hipóteses testáveis e modelos com restrições.

LAMOTE (1972) apresenta uma representação conveniente sobre a matriz de covariância de um modelo geral de análise de variância hierárquico, aleatório com k -níveis. O autor propõe expressões, para a inversa e para o determinante da matriz de covariância obtida, baseada nessa representação conveniente.

EDELMAN (1974) faz uma extensão do modelo hierárquico com dois estágios, com amostras compostas, considerado inicialmente por KUSSMAUL & ANDERSON (1967). O modelo é estendido para a composição de classes e também é incluído um termo de medida do erro casual, fato negligenciado inicialmente. O modelo, considerado por Edelman, onde existe composição entre classes ou dentro de subclasses, é:

$$y_{ijk} = \mu + \frac{1}{s_i} \sum_{\ell=1}^{s_i} a_{i\ell} + \frac{1}{s_i t_{ij}} \sum_{\ell=1}^{s_i} \sum_{m=1}^{t_{ij}} b_{j(i)\ell m} + e_{ijk},$$

$i = 1, \dots, r$; $j = 1, \dots, c_i$ e $k = 1, \dots, n_{ij}$, onde

i : número de classes do modelo composto,

c_i : número de classes do modelo não-composto dentro da i -ésima classe do modelo composto,

t_{ij} : número de subclasses dentro da i -ésima classe do modelo não-composto, que serão combinadas, para formar a j -ésima subclasse do modelo composto,

n_{ij} : repetições da j -ésima subclasse dentro da i -ésima classe do modelo composto.

Excetuando-se a média, os demais efeitos são todos aleatórios e há também uma restrição no modelo, qual seja: se duas ou mais classes são compostas para formar uma classe única, então cada uma das classes compostas contém o mesmo número de sub

classes, isto é,

$$s_i > 1 \text{ implica } t_{i_1} = t_{i_2} = \dots = t_{i_{c_i}}.$$

Baseado neste modelo de componentes de variância hierárquico, formado pela composição de amostras, o autor descreve uma metodologia geral, para estimar a média e os componentes de variância e faz, em seguida, uma comparação numérica de quatro processos diferentes de estimação a fim de determinar em que condições, qual é o melhor.

THOMPSON e ANDERSON (1975), consideram o modelo de componentes de variância hierárquico, não-balanceado, com dois estágios, e comparam dois estimadores truncados, baseados na usual análise da variância, ponderado e não-ponderado, e o estimador de máxima verossimilhança para o delineamento (k_1, k_2) . O delineamento (k_1, k_2) é definido como um experimento hierárquico, não-balanceado, com dois estágios, tendo k classes, tais que, $k_1 + k_2 = k$ e com uma observação em cada uma das k_1 classes e duas observações em cada uma das outras k_2 classes. Os autores mostram que os estimadores de máxima verossimilhança resultam um procedimento mais eficiente (menor quadrado médio residual) que os outros, para a estimação, nessa classe de delineamentos. Os autores salientam ainda que, se o delineamento é balanceado, uma modificação no estimador de máxima verossimilhança fornece um estimador substancialmente

te melhor que os estimadores, baseados ou na usual análise da variância ou nos de máxima verossimilhança.

SAHAI (1976), considera o modelo hierárquico de efeitos aleatórios, balanceado, com três-estágios, dado por: $y_{ijk} = \mu + a_i + b_{ij} + e_{ijk}$, sob as suposições padrões de normalidade e independência dos efeitos aleatórios. Vários estimadores dos componentes de variância σ_a^2 , σ_b^2 e σ_e^2 são comparados, usando o quadrado médio do erro, como a medida do desempenho. O autor prova ainda vários resultados, demonstrando as relações de desigualdades do quadrado médio residual dos estimadores não-viciados, máxima verossimilhança e máxima verossimilhança restrita, e introduz também diversas modificações nos estimadores de máxima verossimilhança não-viciados, incluindo também estimadores do tipo Stein. O autor lembra ainda que Klotz e colaboradores (1969), já haviam feito uma comparação similar de estimadores de componentes de variância, considerando o modelo hierárquico de efeitos aleatórios em dois-estágios.

SEELY (1977) considera o modelo linear $\underline{y} = X\underline{\beta} + \underline{e}$ e discute problemas encontrados em hipóteses lineares, quando nossa atenção não se limita a funções paramétricas estimáveis. O propósito principal do artigo é dar uma justificativa alternativa para aplicação restrita a vetores paramétricos estimáveis em hipótese lineares. Justificativa baseada na

disjunção das classes de distribuições sob H_0 e H_1 . O autor estabelece ainda uma proposição secundária, que é apontar um caminho alternativo de examinar funções paramétricas estimáveis, baseado na condição de identificabilidade discutida por Reiersol, em 1963.

ALALOUF e STYAN (1979) consideram o modelo linear geral $E(\underline{y}) = X\underline{\beta}$ e estudam condições de estimabilidade para um vetor $A\underline{\beta}$. Os autores caracterizam a estimabilidade, com base ou na matriz X , ou na matriz $X'X$.

VIDMAR e BRUNDEN (1980) consideram o modelo

$$y_{ijke} = \mu + \alpha_i + \beta_j(i) + \delta_{k(ij)} + \pi_{e(ijk)}$$

balanceado, referido pelos autores como experimento inteiramente casualizado com dois grupos de tratamentos ($i = 1, 2$) e três níveis de subamostragem (β, δ, π), sendo μ e α fixos e os demais aleatórios. Quando se faz o teste de significância do efeito de grupos de tratamentos, a estimativa do quadrado médio residual usada é formada por uma função linear das estimativas das variâncias dos níveis de subamostragem. Segundo os autores, Marcuse, em 1949, propôs um procedimento que minimiza o custo de obtenção dessas amostras, se o grau de precisão desejado para a esperança do quadrado médio residual usado é fixado; tal procedimento não considera o poder do teste para a diferença entre grupos de tratamentos. Baseados nisto, os

autores propõem um método em que o poder, antes que a precisão, é fixado a um nível crítico e o custo total é minimizado, sujeito a essa restrição.

3. DESENVOLVIMENTO TEÓRICO

3.1 - Modelo Matemático

Considera-se o delineamento de classificação hierárquica com s -estágios e com pelo menos uma das subclasses de um dos estágios, tendo número diferente de unidades experimentais, ou seja, o delineamento de classificação hierárquica não-balanceado.

O modelo matemático para esse delineamento é dado por:

$$y_{i_1 i_2 \dots i_s i_{s+1}} = \mu + \alpha_{i_1} + \beta_{i_1 i_2} + \theta_{i_1 i_2 i_3} + \dots + \delta_{i_1 i_2 \dots i_{s-1}} + \gamma_{i_1 i_2 \dots i_s} + \epsilon_{i_1 i_2 \dots i_{s+1}} \quad (3.1),$$

onde:

- μ : efeito médio geral,
 α_{i_1} : efeito da i_1 -ésima classe,
 $\beta_{i_1 i_2}$: efeito da i_2 -ésima subclasse dentro da i_1 -ésima classe,
 $\theta_{i_1 i_2 i_3}$: efeito da i_3 -ésima sub-subclasse dentro da i_2 -ésima subclasse, dentro da i_1 -ésima classe,
 $\vdots$
 $\delta_{i_1 i_2 \dots i_{s-1}}$: efeito da i_{s-1} -ésima sub... subclasse dentro da i_{s-2} -ésima sub... subclasse... dentro da i_1 -ésima classe,
 $\gamma_{i_1 i_2 \dots i_s}$: efeito da i_s -ésima sub... subclasse dentro da i_{s-1} -ésima sub... subclasse... dentro da i_1 -ésima classe,
 $\epsilon_{i_1 i_2 \dots i_{s+1}}$: efeito residual (entre as unidades experimentais da i_s -ésima sub... subclasse),
 $y_{i_1 i_2 \dots i_s i_{s+1}}$: i_{s+1} -ésima observação da i_s -ésima sub... subclasse dentro da i_{s-1} -ésima sub... subclasse... dentro da i_1 -ésima classe.

Supõe-se que os parâmetros μ , α_{i_1} , $\beta_{i_1 i_2}$, ...,

$\delta_{i_1 i_2 \dots i_{s-1}}$, $\gamma_{i_1 i_2 \dots i_s}$ são todos de efeito fixo, isto é:

$$\left. \begin{aligned} E(\eta) &= \eta \\ &\text{para } \eta = \mu, \alpha_{i_1}, \beta_{i_1 i_2}, \dots, \delta_{i_1 i_2 \dots i_{s-1}}, \gamma_{i_1 i_2 \dots i_s} \\ E(\eta^2) &= \eta^2 \end{aligned} \right\} (3.2)$$

e $\epsilon_{i_1 i_2 \dots i_{s+1}}$ é normalmente distribuído com

$$\left. \begin{aligned} E(\epsilon_{i_1 i_2 \dots i_{s+1}}) &= 0 \\ E(\epsilon_{i_1 i_2 \dots i_{s+1}}^2) &= \sigma^2 \\ E(\epsilon_{i_1 i_2 \dots i_{s+1}} \epsilon_{i_1 i_2 \dots i_{s'+1}}) &= 0 \quad s \neq s'. \end{aligned} \right\} (3.3)$$

Seja também:

$i_1 = 1, \dots, a$ (a : número de classes),

$i_2 = 1, \dots, b_{i_1}$ (b_{i_1} : número de subclasses dentro da i_1 -ésima classe),

$i_3 = 1, \dots, c_{i_1 i_2}$ ($c_{i_1 i_2}$: número de subsubclasses dentro da i_2 -ésima subclasse dentro da i_1 -ésima classe),

$\vdots$

$i_{s-1} = 1, \dots, m_{i_1 \dots i_{s-2}}$ ($m_{i_1 \dots i_{s-2}}$: número de sub...subclasses dentro da i_{s-2} -ésima sub...subclasse dentro da i_1 -ésima classe),

$i_s = 1, \dots, t_{i_1 \dots i_{s-1}}$ ($t_{i_1 \dots i_{s-1}}$: número de sub...subclasses dentro da i_{s-1} -ésima sub...subclasse... dentro da i_1 -ésima classe),

$i_{s+1} = 1, \dots, n_{i_1 \dots i_s}$ ($n_{i_1 \dots i_s}$: número de unidades experimentais da i_s -ésima sub...subclasse).

A seguinte notação será usada:

$$n_{i_1 i_2 \dots i_{s-1} \bullet} = \sum_{i_s=1}^{t_{i_1 \dots i_{s-1}}} n_{i_1 i_2 \dots i_{s-1} i_s}$$

$$n_{i_1 i_2 \dots i_{s-2} \bullet \bullet} = \sum_{i_{s-1}=1}^{m_{i_1 \dots i_{s-2}}} \sum_{i_s=1}^{t_{i_1 \dots i_{s-1}}} n_{i_1 i_2 \dots i_{s-1} i_s}$$

$$n_{\bullet \dots \bullet} = \sum_{i_1=1}^a \sum_{i_2=1}^{b_{i_1}} \dots \sum_{i_s=1}^{t_{i_1 \dots i_{s-1}}} n_{i_1 i_2 \dots i_{s-1} i_s}$$

(3.4)

$$b_{\bullet} = \sum_{i_1=1}^a b_{i_1}$$

$$c_{\bullet\bullet} = \sum_{i_1=1}^a \sum_{i_2=1}^{b_{i_1}} c_{i_1 i_2}$$

$$t_{\bullet\cdots\bullet} = \sum_{i_1=1}^a \sum_{i_2=1}^{b_{i_1}} \cdots \sum_{i_{s-1}=1}^{m_{i_1 i_2 \dots i_{s-2}}} t_{i_1 i_2 \dots i_{s-1}}$$

$$p = 1 + a + b_{\bullet} + c_{\bullet\bullet} + \dots + t_{\bullet\cdots\bullet}$$

3.2 - Forma Matricial

Uma maneira mais interessante e mais prática é escrever o modelo, dado em (3.1), na forma matricial:

$$\underline{y} = X\underline{b} + \underline{e}, \quad (3.5)$$

onde $\underline{y}(n_{\bullet\cdots\bullet} \times 1)$ é o vetor das observações; $X(n_{\bullet\cdots\bullet} \times p)$ é a matriz de valores conhecidos-no presente caso, zeros ou uns (X também é chamada matriz de planejamento ou matriz de incidência); $\underline{b}(p \times 1)$ é o vetor dos parâmetros e $\underline{e}(n_{\bullet\cdots\bullet} \times 1)$ é o vetor dos erros experimentais (vetor dos efeitos residuais), ou seja:

$$\tilde{y} = \begin{bmatrix} y_{11\dots 11} \\ \vdots \\ y_{11\dots 1t_1\dots 1} \\ \vdots \\ y_{1b_1\dots t_{1b_1}\dots} \\ \vdots \\ y_{a1\dots 11} \\ \vdots \\ y_{ab_a\dots t_{ab_a}\dots} \end{bmatrix} = \begin{bmatrix} y_1 \\ \sim \\ n_{1\dots x1} \\ \vdots \\ y_a \\ \sim \\ n_{a\dots x1} \end{bmatrix},$$

onde

$$y_{i_1} = \begin{bmatrix} y_{i_1 1\dots t_{i_1 1}\dots} \\ \vdots \\ n_{i_1\dots x1} \\ y_{i_1 b_{i_1}\dots t_{i_1 b_{i_1}}\dots} \end{bmatrix};$$

$$\tilde{b} = \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \\ \vdots \\ \gamma_{11\dots 1} \\ \vdots \\ \gamma_{ab_a\dots t_{ab_a}\dots} \end{bmatrix} = \begin{bmatrix} \tilde{\mu} \\ 1 \times 1 \\ \alpha \\ \alpha \times 1 \\ \beta \\ \beta \times 1 \\ \vdots \\ \gamma \\ t \bullet \dots \bullet x1 \end{bmatrix} ;$$

$$\tilde{e} = \begin{bmatrix} \varepsilon \\ 11\dots 11 \\ \vdots \\ \varepsilon_{1b_1\dots t_{1b_1}\dots} \\ \vdots \\ \varepsilon_{a1\dots 11} \\ \vdots \\ \varepsilon_{ab_a\dots t_{ab_a}\dots} \end{bmatrix} = \begin{bmatrix} e_1 \\ \tilde{1} \\ n_{1 \bullet \dots \bullet x1} \\ \vdots \\ e_a \\ \tilde{a} \\ n_{a \bullet \dots \bullet x1} \end{bmatrix} ;$$

$$X = \begin{bmatrix} X_\mu & X_\alpha & X_\beta & \dots & X_\delta & X_\gamma \end{bmatrix} ,$$

onde

$$\begin{aligned}
 & X_{\mu} = 1 \quad ; \\
 & n_{\dots x1} \\
 & X_{\alpha} = \begin{bmatrix} X_{\alpha 1} & X_{\alpha 2} & \dots & X_{\alpha a} \end{bmatrix}, \text{ com } X_{\alpha i_1} = \begin{bmatrix} 0 \\ 1 \\ n_{i_1 \dots x1} \\ 0 \end{bmatrix} \rightarrow i_1\text{-ésima posição;} \\
 & n_{\dots x a} \\
 & X_{\beta} = \begin{bmatrix} X_{\beta 11} & \dots & X_{\beta 1b_1} & \dots & X_{\beta a1} & \dots & X_{\beta ab_a} \end{bmatrix}, \\
 & n_{\dots x b} \\
 & \text{com } X_{\beta i_1 i_2} = \begin{bmatrix} 0 \\ 1 \\ n_{i_1 i_2 \dots x1} \\ 0 \end{bmatrix} \rightarrow i_2\text{-ésima posição dentro da } i_1\text{-ésima;} \\
 & n_{\dots x b}
 \end{aligned}$$

e assim, sucessivamente, até

$$\begin{aligned}
 & X_{\gamma} = \begin{bmatrix} X_{\gamma 11 \dots 1} & \dots & X_{\gamma ab_a \dots t_{ab_a}} \end{bmatrix}, \\
 & n_{\dots x t} \\
 & \text{com } X_{\gamma i_1 \dots i_s} = \begin{bmatrix} 0 \\ 1 \\ n_{i_1 \dots i_s x1} \\ 0 \end{bmatrix} \rightarrow i_s\text{-ésima posição dentro da } i_{s-1}\text{-ésima posição} \dots \text{ dentro da } i_1\text{-ésima posição.} \\
 & n_{\dots x t}
 \end{aligned}$$

Tem-se, também:

$$\left. \begin{aligned} E(\underline{e}) &= \underline{0} \quad e \quad E(\underline{e}\underline{e}') = \sigma^2 I_{n \dots} \\ e, \text{ conseqüentemente,} \\ E(\underline{y}) &= X\underline{b} \quad e \quad \text{Var}(\underline{y}) = \sigma^2 I_{n \dots} \end{aligned} \right\} \quad (3.6)$$

3.3 - Equações Normais

Conforme descrito em RAO (1965), as equações normais, correspondentes ao modelo (3.5), podem ser deduzidas pelo processo da soma dos quadrados mínimos e, para

$$\text{Var}(\underline{e}) = \sigma^2 I_{n \dots},$$

elas se tornam:

$$X'X\underline{b} = X'\underline{y}. \quad (3.7)$$

Para o delineamento em estudo, tem-se:

$$X'X = \begin{bmatrix} X'X_{\mu\mu} & X'X_{\mu\alpha} & X'X_{\mu\beta} & \dots & X'X_{\mu\gamma} \\ 1x1 & 1xa & 1xb & & 1xt \\ X'X_{\alpha\mu} & X'X_{\alpha\alpha} & X'X_{\alpha\beta} & \dots & X'X_{\alpha\gamma} \\ ax1 & axa & axb & & axt \\ \vdots & \vdots & \vdots & & \vdots \\ X'X_{\gamma\mu} & X'X_{\gamma\alpha} & X'X_{\gamma\beta} & \dots & X'X_{\gamma\gamma} \\ t \dots x1 & t \dots xa & t \dots xb & & t \dots xt \dots \end{bmatrix} \quad (3.8)$$

Efetutando os produtos matriciais $X_i'X_j$, para $i, j = \mu, \alpha, \beta, \dots, \gamma$, obtêm-se a seguinte matriz $X'X(p \times p)$, descrita em função do número de observações em cada subclasse.

Nesta matriz, onde há omissão de elementos, subentendem-se zeros.

[illegible]

Desde que:

$$n_{\bullet \dots \bullet} = \sum_{i_1} n_{i_1 \bullet \dots \bullet},$$

tem-se que a primeira coluna de $X'X$ é combinação linear das a colunas seguintes;

$$n_{i_1 \bullet \dots \bullet} = \sum_{i_2} n_{i_1 i_2 \bullet \dots \bullet}$$

para $i_1 = 1, \dots, a$, tem-se que cada uma das colunas desde a segunda até a $(1+a)$ -ésima é formada pela combinação linear de b_{i_1} colunas, que se encontram a partir da coluna $[(1+a)+1]$ -ésima até a $[(1+a)+b_{\bullet}]$ -ésima. Desta forma, tem-se até a coluna de ordem $1+a+b_{\bullet}$, $1+a$ colunas linearmente dependentes (ou seja: $(1+a+b_{\bullet}) - (1+a) = b_{\bullet}$ linearmente independentes).

Continuando este procedimento até a última sub-classe, tem-se facilmente o posto da matriz $X'X$, que é dado por:

$$\rho(X'X) = p - (1 + a + \dots + m_{\bullet \dots \bullet}) = t_{\bullet \dots \bullet} < p.$$

Sendo $X'X$ de posto incompleto, o sistema de equações normais admite infinitas soluções, onde cada uma dessas soluções dependerá da inversa generalizada $G = (X'X)^{-}$ usada para a resolução do sistema. Desta forma, não se tem o vetor $\hat{b}$, estimador de mínimos quadrados, mas sim o vetor b^0 , uma so

lução de mínimos quadrados de (3.7).

3.4 - Solução de Mínimos Quadrados

Considerando a equação (3.7), com $\hat{\underline{b}}$ trocado por $\underline{b}^0$, tem-se que o segundo membro pode ser escrito como:

$$\begin{matrix} X' \underline{y} = \\ px1 \end{matrix} \begin{bmatrix} X'_{\mu} \\ X'_{\alpha} \\ X'_{\beta} \\ \vdots \\ X'_{\gamma} \end{bmatrix} \quad \underline{y} = \begin{bmatrix} X'_{\mu} \underline{y} \\ X'_{\alpha} \underline{y} \\ X'_{\beta} \underline{y} \\ \vdots \\ X'_{\gamma} \underline{y} \end{bmatrix} = \begin{bmatrix} \underline{y}_{\dots\dots\dots} \\ 1x1 \\ \underline{y}_{\alpha\dots\dots\dots} \\ ax1 \\ \underline{y}_{\alpha\beta\dots\dots\dots} \\ b \ x1 \\ \vdots \\ \vdots \\ \underline{y}_{\alpha\beta\dots\dots\dots} \\ t \dots\dots\dots x1 \end{bmatrix} \quad (3.10)$$

Consideram-se, então, as seguintes partições:

$$\begin{matrix} X \\ n \dots\dots xp \end{matrix} = \left[\begin{matrix} X_1 \\ n \dots\dots x(1+a+\dots+m \dots\dots) \end{matrix} \quad \begin{matrix} X_2 \\ n \dots\dots xt \dots\dots \end{matrix} \right], \quad (3.11)$$

onde $X_1 = \begin{bmatrix} X_{\mu} & X_{\alpha} & X_{\beta} & \dots & X_{\delta} \end{bmatrix}$ e $X_2 = \begin{bmatrix} X_{\gamma} \end{bmatrix}$;

e

$$\underset{px1}{\tilde{b}}^o = \begin{bmatrix} \underset{\sim}{b}_1^o \\ (1+a+\dots+m \dots) \times 1 \\ \underset{\sim}{b}_2^o \\ t \dots \times 1 \end{bmatrix},$$

onde

$$\underset{\sim}{b}_1^{o'} = \begin{bmatrix} \underset{\sim}{\mu}^{o'} & \underset{\sim}{\alpha}^{o'} & \underset{\sim}{\beta}^{o'} & \dots & \underset{\sim}{\delta}^{o'} \end{bmatrix} \text{ e } \underset{\sim}{b}_2^{o'} = \begin{bmatrix} \underset{\sim}{\gamma}^{o'} \\ \underset{\sim}{\tau} \end{bmatrix}$$

Segundo essas partições, tem-se

$$X'X = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix},$$

$$\text{com } \rho(X'X) = \rho(X_2'X_2) = t \dots,$$

e

$$X' \underset{\sim}{y} = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \underset{\sim}{y} = \begin{bmatrix} X_1' \underset{\sim}{y} \\ X_2' \underset{\sim}{y} \end{bmatrix}$$

onde

$$\underset{\sim}{X}_1' \underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{y} \dots \dots \\ \vdots \\ \underset{\sim}{y}_{\alpha\beta \dots \delta \dots} \end{bmatrix} \text{ e } \underset{\sim}{X}_2' \underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{y}_{\alpha\beta \dots \gamma \dots} \end{bmatrix}$$

Uma das possíveis soluções do sistema de equações normais pode ser obtida considerando G uma inversa generalizada de $X'X$, como:

$$G = \begin{bmatrix} 0 & 0 \\ (1+a+\dots+m) \dots x(1+a+\dots+m) \dots (1+a+\dots+m) \dots x t \dots \\ 0 & D \\ t \dots x(1+a+\dots+m) \dots t \dots x t \dots \end{bmatrix} \quad (3.12)$$

onde D é uma matriz diagonal dada por $D = \text{diag}(1/n_{i_1} \dots i_s)$. Essa inversa generalizada foi encontrada segundo o processo descrito em RAO & MITRA (1971a,b). De acordo com este processo, suponha ser possível encontrar as linhas dependentes de $X'X$. Por omissão dessas linhas e das correspondentes colunas, obtém-se uma matriz $t \dots x t$ de posto $t \dots$, que admite, portanto, uma inversa tradicional. Em seguida, aumenta-se a ordem desta matriz inversa, pela introdução de linhas e colunas de zeros nos lugares correspondentes às linhas e colunas removidas. A matriz, obtida dessa forma, é uma inversa generalizada de $X'X$, denotada por $(X'X)^-$ ou G.

Nestas condições, tem-se:

$$\underline{b}^0 = G X' \underline{y} = \begin{bmatrix} 0 \\ DX_2' \underline{y} \end{bmatrix} \quad \text{ou seja: } \underline{b}^0 = \begin{bmatrix} 0 \\ \underline{y}_{\alpha\beta\dots\gamma_0} \end{bmatrix} \quad (3.13)$$

onde

$$\bar{y}_{\alpha\beta\dots\gamma} = \begin{bmatrix} \frac{\sum_{i_{s+1}} y_{1\dots 1 i_{s+1}}}{n_{1\dots 1}} \\ \vdots \\ \frac{\sum_{i_{s+1}} y_{ab_a \dots t_{ab_a}}}{n_{ab_a \dots t_{ab_a}}} \end{bmatrix}$$

3.5 - Análise de Variância

A notação $R(\cdot)$, descrita em SEARLE (1971) e GRAYBILL (1976), significa a redução na soma de quadrados total, devida ao ajustamento dos parâmetros de um modelo particular, ou seja, $R(\cdot)$ é a medida da variação em $y_{i_1 \dots i_{s+1}}$, considerada por aquele modelo.

A notação $R(\cdot)$ representa a soma de quadrados de parâmetros, com a vantagem de mostrar que modelo foi ajustado.

Baseia-se, então, no modelo dado em (3.1), para o ajustamento seqüencial dos parâmetros (bem como a hierarquia). Esses ajustamentos são feitos baseados nos modelos particulares, definidos a partir de (3.1), cujos valores nessas

reduções são calculados a seguir:

$$y_{i_1 i_2 \dots i_s i_{s+1}} = \mu + \varepsilon_{i_1 \dots i_s i_{s+1}} \rightarrow R(\mu) = \frac{y^2_{\bullet \dots \bullet}}{n_{\bullet \dots \bullet}},$$

onde $R(\mu)$ é a redução da SQ Total, devido ao ajuste de μ ;

$$y_{i_1 i_2 \dots i_s i_{s+1}} = \mu + \alpha_{i_1} + \varepsilon_{i_1 \dots i_s i_{s+1}} \rightarrow R(\alpha, \mu) = \sum_{i_1} \frac{y^2_{i_1 \bullet \dots \bullet}}{n_{i_1 \bullet \dots \bullet}},$$

onde $R(\alpha, \mu)$ é a redução da SQ Total, pelo ajustamento de α e μ .

Logo, o ajustamento de α após μ , indicado por $R(\alpha|\mu)$ é dado por:

$$R(\alpha|\mu) = R(\alpha, \mu) - R(\mu) = \sum_{i_1} \frac{y^2_{i_1 \bullet \dots \bullet}}{n_{i_1 \bullet \dots \bullet}} - \frac{y^2_{\bullet \dots \bullet}}{n_{\bullet \dots \bullet}};$$

$$y_{i_1 i_2 \dots i_s i_{s+1}} = \mu + \alpha_{i_1} + \beta_{i_1 i_2} + \varepsilon_{i_1 \dots i_s i_{s+1}} \rightarrow$$

$$R(\beta:\alpha, \alpha, \mu) = \sum_{i_1} \sum_{i_2} \frac{y^2_{i_1 i_2 \bullet \dots \bullet}}{n_{i_1 i_2 \bullet \dots \bullet}},$$

onde $R(\beta:\alpha, \alpha, \mu)$ é a redução da SQ Total em razão do ajustamento de $\beta:\alpha$ (β dentro de α), α e μ .

O ajustamento de $\beta:\alpha$ após α e μ , indicado por $R(\beta:\alpha \mid \alpha, \mu)$ é obtido de maneira similar à anterior, isto é;

$$R(\beta:\alpha \mid \alpha, \mu) = R(\beta:\alpha, \alpha, \mu) - R(\alpha, \mu), \text{ ou}$$

$$R(\beta:\alpha \mid \alpha, \mu) = \sum_{i_1} \sum_{i_2} \frac{y_{i_1 i_2 \dots}^2}{n_{i_1 i_2 \dots}} - \sum_{i_1} \frac{y_{i_1 \dots}^2}{n_{i_1 \dots}} ;$$

e assim, sucessivamente, até atingir o modelo geral

$$y_{i_1 i_2 \dots i_s i_{s+1}} = \mu + \alpha_{i_1} + \dots + \gamma_{i_1 i_2 \dots i_s} + \epsilon_{i_1 \dots i_s i_{s+1}} \text{ com a}$$

redução dada por:

$$R(\gamma:\dots: \beta:\alpha \mid \delta:\dots: \beta:\alpha, \dots, \beta:\alpha, \alpha, \mu) =$$

$$= \sum_{i_1} \dots \sum_{i_s} \frac{y_{i_1 \dots i_s}^2}{n_{i_1 \dots i_s}} - \sum_{i_1} \dots \sum_{i_{s-1}} \frac{y_{i_1 i_2 \dots i_{s-1}}^2}{n_{i_1 \dots i_{s-1}}} .$$

Observam-se, ainda, os seguintes fatos, nessas somas de quadrados:

$$\sum_{i_1} \frac{y_{i_1 \dots}^2}{n_{i_1 \dots}} - \frac{y_{\dots}^2}{n_{\dots}} = \sum_{i_1} \dots \sum_{i_{s+1}} (\bar{y}_{i_1 \dots} - \bar{y}_{\dots})^2 . \text{ Logo,}$$

pelo conceito de graus de liberdade, pode concluir-se que os g.l., associados a $R(\alpha|\mu)$, são $a-1$;

$$\sum_{i_1} \sum_{i_2} \frac{y_{i_1 i_2 \dots}^2}{n_{i_1 i_2 \dots}} - \sum_{i_1} \frac{y_{i_1 \dots}^2}{n_{i_1 \dots}} =$$

$$= \sum_{i_1} \dots \sum_{i_{s+1}} (\bar{y}_{i_1 i_2 \dots} - \bar{y}_{i_1 \dots})^2,$$

da mesma forma que, anteriormente, os g.l., associados a $R(\beta:\alpha|\alpha, \mu)$, são $b-a$. Por um procedimento sucessivo e análogo, tem-se que:

$$\sum_{i_1} \dots \sum_{i_s} \frac{y_{i_1 \dots i_s}^2}{n_{i_1 \dots i_s}} - \sum_{i_1} \dots \sum_{i_{s+1}} \frac{y_{i_1 \dots i_{s-1}}^2}{n_{i_1 \dots i_{s-1}}} =$$

$$= \sum_{i_1} \dots \sum_{i_{s+1}} (\bar{y}_{i_1 \dots i_s} - \bar{y}_{i_1 \dots i_{s-1}})^2, \text{ e}$$

os g.l., associados a $R(\gamma:\dots:\beta:\alpha, \dots, \alpha, \mu)$, são

$$t - m.$$

Pode-se, então, construir o seguinte quadro:

Quadro I: Análise da variância de experimentos hierárquicos.

Causa de Variação	G.L	QM
Média	1	$R(\mu)$
α após μ	$a - 1$	$\frac{R(\alpha \mu)}{a - 1}$
$\beta:\alpha$ após α e μ	$b - a$	$\frac{R(\beta:\alpha \alpha, \mu)}{b - a}$
----- (3.14)		
$\gamma: \delta: \dots: \alpha$ após $\delta: \dots: \beta: \alpha$ e α e μ	$t \dots - m$	$\frac{R(\gamma: \dots: \beta: \alpha \delta: \dots: \alpha, \dots, \alpha, \mu)}{t \dots - m}$
Resíduo	$n \dots - t$	$\frac{\sum_{i_1} \dots \sum_{i_{s+1}} y_{i_1 \dots i_{s+1}}^2 - R(\mu, \alpha, \dots, \gamma: \dots: \alpha)}{n \dots - t}$
Total	$n \dots$	

3.6 - Funções Estimáveis

O conceito de função paramétrica estimável no contexto de modelo linear foi introduzido por BOSE (1944), o qual exige que tal função possua um estimador linear não-viciado.

Seja $\underline{q}$ ($p \times 1$) um vetor de coeficientes, denotado por

$$\begin{array}{c} \underline{q} \\ \sim \\ p \times 1 \end{array} = \begin{bmatrix} \text{coef}(\mu) \\ \sim \\ 1 \times 1 \\ \\ \text{coef}(\alpha) \\ \sim \\ a \times 1 \\ \\ \vdots \\ \\ \text{coef}(\delta) \\ \sim \\ m \dots \times 1 \\ \hline \text{coef}(\gamma) \\ \sim \\ t \dots \times 1 \end{bmatrix} = \begin{bmatrix} q_1 \\ \sim \\ (1+\dots+m \dots) \times 1 \\ \hline q_2 \\ \sim \\ t \dots \times 1 \end{bmatrix}, \quad (3.15)$$

onde

$$\text{coef}'(\mu) = [\text{coef}(\mu)] ,$$

$$\text{coef}'(\alpha) = [\text{coef}(\alpha_1) \quad \dots \quad \text{coef}(\alpha_a)]$$

$$\text{coef}'(\gamma) = [\text{coef}(\gamma_{11\dots 1} \quad \dots \quad \text{coef}(\gamma_{ab_a \dots t_{ab_a} \dots})]$$

e que a mesma partição se faça no vetor dos parâmetros, isto é:

$$\underset{\sim}{b} = \begin{bmatrix} \underset{\sim}{\mu} \\ \underset{\sim}{\alpha} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{\delta} \\ \hline \underset{\sim}{\gamma} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{b}_1 \\ \hline \underset{\sim}{b}_2 \end{bmatrix}.$$

RAO (1965) e SEARLE (1971) afirmam que uma função linear dos parâmetros, $\underset{\sim}{q}'\underset{\sim}{b}$, é estimável se ela é idêntica a alguma função linear dos valores esperados do vetor de observações $\underset{\sim}{y}$; isto é, se existir um vetor $\underset{\sim}{t}$ tal que

$$\underset{\sim}{t}'E(\underset{\sim}{y}) = \underset{\sim}{q}'\underset{\sim}{b}.$$

O primeiro destes autores ainda demonstra um teorema relativo à condição necessária e suficiente para a estimabilidade de funções lineares dos parâmetros, que é:

"A função linear paramétrica $\underset{\sim}{q}'\underset{\sim}{b}$ é estimável se, e somente se: $\underset{\sim}{q}' = \underset{\sim}{q}'H$, onde $H = GX'X^{-1}$ ".

Usa-se, então, este teorema para explicitar as condições de estimabilidade das funções lineares paramétricas, considerando o modelo matemático dado em (3.1).

Nestas condições, tem-se:

$$\begin{aligned}
 H_{p \times p} &= GX'X = \begin{bmatrix} 0 & 0 \\ DX_2'X_1 & DX_2'X_2 \end{bmatrix}, \\
 \underline{q}'H &= \begin{bmatrix} \underline{q}_2' DX_2'X_1 & \underline{q}_2' DX_2'X_2 \end{bmatrix}
 \end{aligned} \tag{3.16}$$

e, portanto,

$$\begin{cases} \underline{q}_2' DX_2'X_1 = \underline{q}_1' & (3.17.a) \\ \underline{q}_2' DX_2'X_2 = \underline{q}_2' & (3.17.b) \end{cases}$$

Observa-se que a equação (3.17.b) sempre é verificada, para qualquer que seja o vetor $\underline{q}$ escolhido; portanto, usa-se somente a equação (3.17.a), para especificar as condições de estimabilidade.

Dessa forma, tem-se:

$$DX_2^1 X_1 = \text{diag} \left(\underbrace{1}_{n_1 \dots i_s} \right)$$

μ	α	β	δ
$n_{11} \dots 1$ $\vdots$ $n_{1b_1} \dots t_{1b_1} \dots$ $\vdots$ $n_{a1} \dots 1$ $\vdots$ $n_{ab_a} \dots t_{ab_a} \dots$	$n_{11} \dots 1$ $\vdots$ $n_{1b_1} \dots t_{1b_1} \dots$ $\vdots$ $n_{a1} \dots 1$ $\vdots$ $n_{ab_a} \dots t_{ab_a} \dots$	$n_{111} \dots 1$ $\vdots$ $n_{11c_{11}} \dots t_{11c_{11}} \dots$ $\vdots$ $n_{1b_1} \dots 1$ $\vdots$ $n_{1b_1 c_{1b_1}} \dots t_{1b_1 c_{1b_1}} \dots$ $\vdots$ $n_{a11} \dots 1$ $\vdots$ $n_{a1c_{a1}} \dots t_{a1c_{a1}} \dots$ $\vdots$ $n_{ab_a} \dots 1$ $\vdots$ $n_{ab_a c_{ab_a}} \dots t_{ab_a c_{ab_a}} \dots$	$n_{111} \dots 1$ $\vdots$ $n_{111} \dots t_{111} \dots$ $\vdots$ $n_{ab_a c_{ab_a}} \dots 1$ $\vdots$ $n_{ab_a c_{ab_a}} \dots t_{ab_a c_{ab_a}} \dots$

(3.18)

ou seja,

$$DX'_2 X_1 = \begin{bmatrix} X_\mu^* & X_\alpha^* & X_\beta^* & \dots & X_\delta^* \\ t_{\dots} x_{1+a+\dots+m} & t_{\dots} x_1 & t_{\dots} x_a & t_{\dots} x_b & t_{\dots} x_m \end{bmatrix};$$

onde

$$X_\mu^* = \begin{bmatrix} 1 \\ \sim \end{bmatrix};$$

$$X_\alpha^* = \begin{bmatrix} X_{\alpha 1}^* & X_{\alpha 2}^* & \dots & X_{\alpha a}^* \end{bmatrix},$$

com $X_{\alpha i_1}^* = \begin{bmatrix} 0 \\ \sim \\ 1 \\ t_{i_1} \dots x_1 \\ 0 \\ \sim \end{bmatrix} \rightarrow i_1\text{-ésima posição};$

$$X_\beta^* = \begin{bmatrix} X_{\beta 11}^* & \dots & X_{\beta 1b_1}^* & \dots & X_{\beta a1}^* & \dots & X_{\beta ab_a}^* \end{bmatrix},$$

com $X_{\beta i_1 i_2}^* = \begin{bmatrix} 0 \\ \sim \\ 1 \\ t_{i_1 i_2} \dots x_1 \\ 0 \\ \sim \end{bmatrix}$

$$\left. \begin{aligned}
 \sum_{i_1} \text{coef}(\alpha_{i_1}) &= \text{coef}(\mu) \\
 \sum_{i_2} \text{coef}(\beta_{i_1 i_2}) &= \text{coef}(\alpha_{i_1}) & i_1 &= 1, \dots, a \\
 \sum_{i_3} \text{coef}(\theta_{i_1 i_2 i_3}) &= \text{coef}(\beta_{i_1 i_2}) & i_1 &= 1, \dots, a \\
 & & i_2 &= 1, \dots, b_{i_1} \\
 \dots & \dots & & \\
 \sum_{i_s} \text{coef}(\gamma_{i_1 \dots i_s}) &= \text{coef}(\delta_{i_1 \dots i_{s-1}}) & i_1 &= 1, \dots, a \\
 & & & \dots \\
 & & i_{s-1} &= 1, \dots, m_{i_1 \dots i_{s-2}}
 \end{aligned} \right\} (3.21)$$

3.7 - Teste da Hipótese $q'b \approx m$

Como pode ser visto em SEARLE (1971), o estima
dor de mínimos quadrados coincide com o "best linear unbiased
estimator", ou simplesmente "b.l.u.e", (MELI: melhor est
imador linear imparcial). Considera-se ainda como sendo o melhor est
imador linear imparcial (não-viciado) o estimador que é uma fun
ção linear das observações e que, dentro da classe dos impar
ciais, possuir variância mínima.

Com base neste autor, pode-se afirmar que toda
função linear $q'b$ que satisfaz às condições estabelecidas pelo
Teorema 1, tem seu MELI dado por:

$$\text{MELI}(\underline{q}'\underline{b}) = \widehat{\underline{q}'\underline{b}} = \underline{q}'\underline{b}^0 ,$$

ou simplesmente:

$$\widehat{\underline{q}'\underline{b}} = \underline{q}'_2 \bar{y}_{\alpha\beta\dots\gamma} = \sum_{i_1} \dots \sum_{i_s} \text{coef}(\gamma_{i_1\dots i_s}) \bar{y}_{i_1 i_2 \dots i_s} \quad (3.22)$$

com $\underline{q}'\underline{b}^0$, sendo uma função linear das observações, não-viciada, para $\underline{q}'\underline{b}$ e de variância mínima.

A variância do $\text{MELI}(\underline{q}'\underline{b})$ é obtida por:

$$\begin{aligned} \text{var}(\widehat{\underline{q}'\underline{b}}) &= \text{var}(\underline{q}'\underline{b}^0) \\ &= \underline{q}' \text{var}(\underline{b}^0) \underline{q} \\ &= \underline{q}' \text{GX'XG}' \underline{q} \sigma^2 , \end{aligned}$$

ou simplesmente:

$$\text{var}(\widehat{\underline{q}'\underline{b}}) = \underline{q}'_2 \text{D}' \underline{q}_2 \sigma^2 = \underline{q}'_2 \text{D} \underline{q}_2 \sigma^2 = \sum_{i_1} \dots \sum_{i_s} \frac{\text{coef}(\gamma_{i_1\dots i_s})^2}{n_{i_1\dots i_s}} \sigma^2 . \quad (3.23)$$

Mostra-se em BOSE (1944), que tanto o $\text{MELI}(\underline{q}'\underline{b})$ quanto sua variância, $\underline{q}'\text{GX'XG}'\underline{q}\sigma^2$, são invariantes, em qualquer escolha feita para a inversa generalizada de X'X . Opta-se, no presente caso, por G da forma dada em (3.12), pela praticidade e simplicidade de sua determinação.

Sabe-se que uma hipótese é testável quando pode ser expressa em termos de funções estimáveis. Isto é, a hipó-

tese $H_0: \underline{q}'\underline{b} = m$ deve ser invariante para qualquer que seja a G escolhida, e isso acontece se $\underline{q}'\underline{b}$ é uma função estimável: função que satisfaz às condições estabelecidas pelo Teorema 1.

Para o teste da hipótese testável $H_0: \underline{q}'\underline{b} = m$, usa-se o procedimento descrito em RAO (1965), SEARLE (1971) e GRAYBILL (1976) que, no presente caso, resulta na estatística:

$$F(H) = \frac{(\underline{q}'_2 \underline{b}_2^0 - m)' (\underline{q}'_2 D \underline{q}_2)^{-1} (\underline{q}'_2 \underline{b}_2^0 - m)}{\hat{\sigma}^2} \quad (3.24)$$

com distribuição $F_{(1, n_{\bullet\bullet\bullet\bullet} - t_{\bullet\bullet\bullet\bullet})}$ sob a veracidade de H_0 .

A estimativa do erro residual, dada por $\hat{\sigma}^2$, é obtida de (3.14) por:

$$\hat{\sigma}^2 = \frac{\sum_{i_1} \dots \sum_{i_{s+1}} y_{i_1 \dots i_{s+1}}^2 - R(\mu, \alpha, \beta: \alpha, \dots, \gamma: \sigma: \dots: \beta: \alpha)}{n_{\bullet\bullet\bullet\bullet} - t_{\bullet\bullet\bullet\bullet}}$$

Uma forma equivalente de se escrever (3.24) é dada por:

$$F(H) = \frac{(\sum_{i_1} \dots \sum_{i_s} \text{coef}(\gamma_{i_1 \dots i_s}) \bar{y}_{i_1 i_2 \dots i_s} - m)^2}{\text{coef}(\gamma_{i_1 \dots i_s})^2 \left(\sum_{i_1} \dots \sum_{i_s} \frac{1}{n_{i_1 \dots i_s}} \right) \hat{\sigma}^2} \quad (3.25)$$

Estabelecido o nível de significância α , a regra de decisão é a habitual, ou seja:

$$\text{se } F(H) > F_{\alpha, (1, n_{\bullet \dots \bullet} - t_{\bullet \dots \bullet})} ,$$

o resultado do teste estatístico é significativo e, portanto, rejeita-se H_0 ; caso contrário, não se rejeita.

4. PROGRAMA COMPUTACIONAL

4.1 - Características do Programa

O programa computacional foi desenvolvido em linguagem FORTRAN IV, e será oportunamente implantado no Centro de Comunicação e Informática da UNESP (CCI - UNESP).

Este programa analisa dados em classificação hierárquica, balanceados ou não, considerando um máximo de dez estágios. As somas de quadrados foram obtidas, segundo SEARLE (1971). A estatística F para o teste do ajuste de cada um dos efeitos aninhados, foi obtida pela razão entre o quadrado médio correspondente a esse efeito, e o quadrado médio residual.

Para simplificar a impressão dos resultados, no quadro da análise da variância, usou-se a seguinte terminologia, na coluna causa de variação: "classes", para indicar o ajuste de α após μ ; "subclasses (1): classes", para indicar o ajuste de $\beta:\alpha$ após α e μ ; "subclasses (2): subclasses (1)", para in

dicar o ajuste de $\theta:\beta:\alpha$ após $\beta:\alpha$, α e μ , e assim sucessivamente, até "subclasses (s): subclasses (s-1)", para indicar o ajuste de $\gamma:\delta:\dots:\beta:\alpha$ após $\delta:\dots:\beta:\alpha$, ..., $\beta:\alpha$, α e μ .

Em seguida, são impressos o desvio-padrão, o coeficiente de variação, a média geral e todas as médias, dentro de cada um dos s -estágios.

Considerando o vetor das médias amostrais,

$$\begin{matrix} \bar{y} \\ \sim \\ (1+a+\dots+t) \times 1 \end{matrix} = \begin{bmatrix} \bar{y}_{\bullet \dots \bullet} \\ 1 \times 1 \\ \\ \bar{y}_{\alpha \bullet \dots \bullet} \\ a \times 1 \\ \vdots \\ \vdots \\ \vdots \\ \bar{y}_{\alpha \beta \dots \gamma \bullet} \\ t \dots \times 1 \end{bmatrix}, \quad (4.1)$$

onde

$$\bar{y}_{\bullet \dots \bullet} = \left[\frac{\sum_{i_1} \dots \sum_{i_{s+1}} y_{i_1 i_2 \dots i_{s+1}}}{n \dots} \right],$$

$$\begin{matrix} \bar{y}_{\alpha \dots \alpha} \\ \sim \\ ax1 \end{matrix} = \left[\begin{array}{c} \frac{\sum_{i_2} \dots \sum_{i_{s+1}} y_{1i_2 \dots i_{s+1}}}{n_{1 \dots \dots}} \\ \vdots \\ \frac{\sum_{i_2} \dots \sum_{i_{s+1}} y_{ai_2 \dots i_{s+1}}}{n_{a \dots \dots}} \end{array} \right],$$

$$\begin{matrix} \bar{y}_{\alpha \beta \dots \gamma} \\ \sim \\ t \dots x 1 \end{matrix} = \left[\begin{array}{c} \frac{\sum_{i_{s+1}} y_{11 \dots 1i_{s+1}}}{n_{1 \dots 1}} \\ \vdots \\ \frac{\sum_{i_{s+1}} y_{ab_a \dots t_{ab_a}}}{n_{ab_a \dots t_{ab_a}}} \end{array} \right],$$

foi elaborado um quadro, contendo todos os contrastes possíveis entre duas médias, dentro de cada um dos s estágios. Neste quadro, os índices que aparecem entre parênteses, na coluna denominada "CONTRASTE", correspondem às posições ocupadas pelas médias amostrais, no vetor dado pela expressão (4.1). Por exemplo, o valor estimado do contraste $M(2) - M(3)$ é

$$\bar{y}_{1 \dots t} - \bar{y}_{2 \dots t} ,$$

que são, respectivamente, o segundo e o terceiro elementos do vetor $\bar{y}$, de (4.1). Este quadro se completa com o valor da variância e o valor obtido pela estatística F, usada para testar a significância do contraste.

Estas comparações podem ser realizadas, pois qualquer contraste entre duas médias, dentro de cada um dos s estágios, resulta uma função linear paramétrica estimável, de acordo com as condições aqui estabelecidas.

O programa possibilita, também, verificar a estimabilidade de qualquer função linear paramétrica, segundo as condições aqui estabelecidas. Não sendo estimável, imprime-se a mensagem "FUNCAO NAO ESTIMAVEL". Caso ela o seja, calcula-se sua estimativa, variância e o valor da estatística F, correspondente ao teste de hipótese $q'b = 0$, relacionado com essa função estimável. Torna-se oportuno observar que os contrastes entre duas médias, dentro de cada estágio, são, na realidade, um caso particular das funções anteriormente referidas.

É impressa, também, a matriz diagonal,

$$D^{-1} = \text{diag} (n_{i_1 \dots i_s})$$

de ordem $t \dots t$, necessária para a determinação de conjuntos de contrastes ortogonais, a critério do pesquisador. Segundo BOSE (1944), dois contrastes $q^{(i)'}_b$ e $q^{(j)'}_b$ são ortogonais

nais se $q^{(i)'} G q^{(j)} = 0$ para $i \neq j$. De acordo com as considerações empregadas neste trabalho, a condição de ortogonalidade se verifica se $q_2^{(i)'} D q_2^{(j)} = 0$ para $i \neq j$.

Isto é, considerando

$$q^{(k)} = \begin{bmatrix} \text{coef}^{(k)}(\mu) \\ \sim \\ 1 \times 1 \\ \vdots \\ \vdots \\ \text{coef}^{(k)}(\delta) \\ m \quad \sim \quad \times 1 \\ \bullet \dots \bullet \\ \hline \text{coef}^{(k)}(\gamma) \\ t \quad \sim \quad \times 1 \\ \bullet \dots \bullet \end{bmatrix} = \begin{bmatrix} q_1^{(k)} \\ \hline q_2^{(k)} \end{bmatrix}, \text{ com } k = i, j; \quad (4.2)$$

tem-se que a condição de ortogonalidade fica:

$$\sum_{i_1} \dots \sum_{i_s} \frac{\text{coef}^{(i)}(\gamma_{i_1 \dots i_s}) \text{coef}^{(j)}(\gamma_{i_1 \dots i_s})}{n_{i_1 \dots i_s}} = 0. \quad (4.3)$$

4.2 - Entrada de Dados

O experimento é composto de s estágios. O primeiro estágio é composto de a classes; o segundo é composto de

$$b_{i_1} = \sum_{i_1=1}^1 b_{i_1}$$

subclasses (1), onde b_{i_1} é o número de subclasses (1), dentro da i_1 -ésima classe; o terceiro de

$$c_{i_1 i_2} = \sum_{i_1} \sum_{i_2} c_{i_1 i_2}$$

subclasses (2), onde $c_{i_1 i_2}$ é o número de subclasses (2), dentro da i_2 -ésima subclasse (1) e i_1 -ésima classe, e assim por diante. O número total de unidades experimentais é $n_{\dots}$.

A ordem de entrada de dados para o programa é como segue:

1 ^a leitura	col:1-80	comentário FORMAT(20A4).
2 ^a leitura	col:1-10	variáveis NS e NCONT FORMAT(16I5),
	col:1-5	NS: número de estágios (NS ≤ 10),
	col:6-10	NCONT: número de funções desejadas.
3 ^a leitura	col:1-50	veter (NN(I), I=1, NS+1) FORMAT(16I5),

	col:1-5	NN(1)=a, número de classes no estágio 1,
	col:6-10	NN(2)=b _• , número de subclasses no estágio 2,
	⋮	⋮
	col:(5*NS+1)-5*(NS+1)	NN(NS+1)=n _{•••••} , número total de unidades experimentais.
4 ^a leitura	col:1-60	vetor dos valores dos índices de cada unidade experimental (II(J), J=1,10) e Y: valor da respectiva unidade experimental FORMAT(10I5,F10.0),
	col:1-5	II(1): valor do índice de classe,
	col:6-10	II(2): valor do índice de sub-classe do estágio 2,
	⋮	⋮
	col:(5*(NS-1)+1)-5*NS	II(NS): valor do índice de sub-classe do último estágio,
	col:51-60	Y: valor da unidade experimental.
5 ^a leitura	col:1-80	(QL(J),J=1,N), vetor dos coeficientes dos parâmetros que definem a função desejada, onde N=número total de parâmetros do modelo (dimensão de $\tilde{b}$) FORMAT(8F10.0),

col:1-10	QL(1): coeficiente do primeiro parâmetro ($\text{coef}(\mu)$),
col:11-20	QL(2): coeficiente do segundo parâmetro ($\text{coef}(\alpha_1)$),
col:21-30	QL(3): coeficiente do terceiro parâmetro ($\text{coef}(\alpha_2)$),
⋮	⋮

Obs: 1) Se $N > 8$ continuar na
linha seguinte, usando
tantas linhas quantas
necessárias,
2) se $NCONT=0$, omitir es-
ta leitura.

4.3 - Listagem do Programa

A seguir, encontra-se a listagem do programa
"CLAH11", em linguagem FORTRAN IV.

```

C MODELO LINEAR DE EFEITOS FIXOS EM CLASSIFICACAO HIERARQUICA
C COM NS ESTAGIOS - (CLAH11)
C ANALISE DE VARIANCIA, FUNCOES ESTIMAVEIS E TESTE DE HIPOTHESES
C NS= NUMERO DE ESTAGIOS (NS MENOR OU IGUAL A 10)
C NN(1)= NUMERO DE CLASSES DO ESTAGIO 1
C NN(I),I=2,NS : NUMERO DE SUBCLASSES EM CADA ESTAGIO
C NN(NS+1)= NUMERO TOTAL DE UNIDADES EXPERIMENTAIS
C II(J),J=1,10 : VETOR DOS VALORES DOS INDICES DE CADA UNIDADE
C                EXPERIMENTAL
C Y= VALOR DE CADA UNIDADE EXPERIMENTAL
C NCONT= NUMERO DE FUNCOES DESEJADAS
C QL(J),J=1,N = VETOR DOS COEFICIENTES DAS FUNCOES, ONDE N= NUMERO DE
C                PARAMETROS DO MODELO
      IMPLICIT REAL*8 (A-H, O-Z)
      DIMENSION COM(20),NN(11),II(11),NV(1000),XLY(1000),SQ(12),XM(1000)
      DIMENSION SQT(11),IGT(12),QMT(11),FT(11),QL(500)
      READ(2,1000)NCAS
1000  FORMAT(I5)
      NCACO=0
1001  READ(2,100)(COM(I),I=1,20)
      100  FORMAT(20A4)
      READ(2,101)NS,NCONT
101  FORMAT(16I5)
      NF=NS+1
      READ(2,101)(NN(I),I=1,NF)
      NT=NN(NF)
      N=1
      DO 1 I=1,NS
1  N=N+NN(I)
      SX2=0.
      DO 2 J=1,N
      NV(J)=0
2  XLY(J)=0.
      DO 3 I=1,NT
      READ(2,102)(II(J),J=1,10),Y
102  FORMAT(10I5,F10.0)
      SX2=SX2+Y**2
      NV(1)=NV(1)+1
      XLY(1)=XLY(1)+Y
      IND=1
      DO 4 J=1,NS
      IND=IND+II(J)
      NV(IND)=NV(IND)+1
      XLY(IND)=XLY(IND)+Y
      IND=IND-II(J)+NN(J)
4  CONTINUE
3  CONTINUE
C CALCULO DAS SOMAS DE QUADRADOS E MEDIAS
      LI=1
      LS=1

```

```

      DO 5 I=1,NF
      SQ(I)=0
      DO 6 J=LI,LS
      ENE=NV(J)
      SQ(I)=SQ(I)+XLY(J)**2/ENE
6     XM(J)=XLY(J)/ENE
      LI=LS+1
      LS=LS+NN(I)
5     CONTINUE
      SQ(NF+1)=SX2
      DO 7 I=1,NF
      SQT(I)=SQ(I+1)-SQ(I)
      SQ(NF+1)=SQ(NF+1)-SQ(1)
C     CALCULO DOS GRAUS DE LIBERDADE
      IGT(NF+1)=NT-1
      IGT(1)=NN(1)-1
      DO 8 I=2,NF
      IGT(I)=NN(I)-NN(I-1)
C     CÁLCULO DOS QUADRADOS MEDIOS
      DO 9 I=1,NF
      D=IGT(I)
      9     QMT(I)=SQT(I)/D
      DO 10 I=1,NS
      10    FT(I)=QMT(I)/QMT(NF)
      S=DSQRT(QMT(NF))
      CV=100.*S/XM(1)
C     SAIDA DA ANALISE DE VARIANCIA
      WRITE(6,200)(COM(I),I=1,20)
200    FORMAT('1',///,20A4,///)
      WRITE(6,201)
201    FORMAT(10X,'ANALISE DA VARIANCIA',/)
      WRITE(6,205)
      WRITE(6,202)
202    FORMAT(10X,'CAUSA DE VARIACAO',19X,'GL',10X,'SQ',18X,'QM',18X,'F')
      WRITE(6,205)
205    FORMAT(10X,100('*'))
      WRITE(6,211)IGT(1),SQT(1),QMT(1),FT(1)
211    FORMAT(10X,'CLASSES',28X,I4,1X,3E20.10)
      WRITE(6,212)IGT(2),SQT(2),QMT(2),FT(2)
212    FORMAT(10X,'SUBCLASSES( 1) : CLASSES',11X,I4,1X,3E20.10)
      NF2=NF-2
      IF(NF2-3)12,13,13
13     CONTINUE
      DO 11 I=3,NS
      K=I-1
      J=I-2
      11    WRITE(6,203)K,J,IGT(I),SQT(I),QMT(I),FT(I)
203    FORMAT(10X,'SUBCLASSES(',I2,') : SUBCLASSES(',I2,')',4X,I4,1X,3E20
1.10)
12     CONTINUE

```

```

WRITE(6,213) IGT(NF),SQT(NF),QMT(NF)
213 FORMAT(10X,'RESIDUO',28X,I4,1X,2E20.10)
WRITE(6,205)
I1=NF+1
WRITE(6,206) IGT(I1),SQ(I1)
206 FORMAT(10X,'TOTAL',30X,I4,1X,E20.10,)
WRITE(6,205)
WRITE(6,207) S,CV,XM(1)
207 FORMAT(//,10X,
1'DESVIO PADRAO =',E20.10,5X,'COEFICIENTE DE VARIACAO ='
2,E20.10,/,10X,'MEDIA GERAL =',E20.10,/)
C SAIDA DAS MEDIAS
LSS=N-NN(NS)
IFA=0
LI=1
LS=1
I=2
NC=Nv(1)
DO 30 J=1,LSS
IF(NC-Nv(1))31,35,35
35 NC=0
IFA=IFA+1
WRITE(6,208) IFA
208 FORMAT(//,10X,'MEDIAS DENTRO DO ESTAGIO',I3,/)
31 CONT=1
NA=0
32 NA=NA+Nv(I)
I=I+1
IF(NA-Nv(J))33,34,34
33 CONT=CONT+1
GO TO 32
34 LI=LS+1
LS=LS+CONT
DO 36 K=LI,LS
36 NC=NC+Nv(K)
WRITE(6,209) (XM(K),K=LI,LS)
209 FORMAT(5E20.10)
WRITE(6,210)
210 FORMAT(//)
30 CONTINUE
WRITE(6,220)
220 FORMAT('1',///,10X,'CONTRASTES ENTRE DUAS MEDIAS DENTRO DE CADA ES
1TAGIO',/)
C CALCULO DOS CONTRASTES ENTRE DUAS MEDIAS DENTRO DE CADA ESTAGIO
LSS=N-NN(NS)
IFA=0
LI=1
LS=1
I=2
NC=Nv(1)

```

```

      DO 80 J=1,LSS
      IF(NC-NV(1))81,85,85
85  NC=0
      IFA=IFA+1
      WRITE(6,221)IFA
221  FORMAT(/,10X,'ESTAGIO',I3,/)
      WRITE(6,222)
222  FORMAT(10X,'CONTRASTE',16X,'VALOR. ESTIMADO',8X,'VARIANCIA',
      112X,'TESTE F')
      WRITE(6,223)
223  FORMAT(10X,80('*'))
81  CONT=1
      NA=0
82  NA=NA+NV(I)
      I=I+1
      IF(NA-NV(J))83,84,84
83  CONT=CONT+1
      GO TO 82
84  LI=LS+1
      LS=LS+CONT
      DO 86 K=LI,LS
86  NC=NC+NV(K)
      IF(LI-LS)88,80,80
88  CONTINUE
C  CALCULO DO CONTRASTE, VARIANCIA E TESTE F
      LIM=LS-1
      DO 87 K=LI,LIM
      LL=K+1
      DO 87 L=LL,LS
      CON=XM(K)-XM(L)
      VAR=QMT(NF)*(1./NV(K)+1./NV(L))
      FF=CON**2/VAR
      WRITE(6,224)K,L,CON,VAR,FF
224  FORMAT(10X,'M(',I2,') - M(',I2,')',7X,3E20.10)
87  CONTINUE
      WRITE(6,223)
80  CONTINUE
C  CALCULO DE OUTRAS FUNCOES ESTIMAVEIS
      IF(NCONT)90,90,91
91  WRITE(6,230)
230  FORMAT('1',///,10X,'OUTROS TIPOS DE FUNCOES',/)
      WRITE(6,231)
231  FORMAT(5X,'COEFICIENTES DA FUNCAO',32X,'VALOR ESTIMADO ',8X,'VARIA
      INCIA',12X,'TESTE F',/)
      WRITE(6,232)
232  FORMAT(5X,110('*'))
      DO 92 L=1,NCONT
      READ(2,103)(QL(J),J=1,N)
103  FORMAT(8F10.0)
      WRITE(6,233)(QL(J),J=1,N)

```

```

233 FORMAT(5X,5F10.5)
C  VERIFICACAO DA ESTIMABILIDADE DA FUNCAO
    LSS=N-NN(NS)
    LI=1
    LS=1
    I=2
    DO 55 J=1,LSS
      CONT=1
      NA=0
52  NA=NA+NV(I)
      I=I+1
      IF(NA-NV(J))56,57,57
56  CONT=CONT+1
      GO TO 52
57  LI=LS+1
      LS=LS+CONT
      SS=0
      DO 58 K=LI,LS
58  SS=SS+QL(K)
      IF(DABS(SS-QL(J))-0.1D-04)55,55,53
53  WRITE(6,234)
234 FORMAT(60X,'FUNCAO NAO ESTIMAVEL')
      GO TO 92
55  CONTINUE
      LI=LSS+1
      LS=N
      I=1
      CON=0
      VAR=0
      DO 59 K=LI,LS
        CON=CON+QL(K)*XM(K)
59  VAR=VAR+QL(K)**2/NV(K)
      VAR=VAR*QMT(NF)
      F=CON**2/VAR
      WRITE(6,235)CON,VAR,F
235 FORMAT(55X,3E20.10)
92  CONTINUE
      WRITE(6,232)
90  CONTINUE
      LI=LSS+1
      WRITE(6,240)
240 FORMAT(////,10X,'MATRIZ DIAGONAL D INVERSA',/)
      WRITE(6,241)(NV(I),I=LI,N)
241 FORMAT(10X,20I5)
      NCACO=NCACO+1
      IF(NCAS.GT.NCACO) GO TO 1001
      RETURN
      END

```


5. APLICAÇÃO

A metodologia desenvolvida neste trabalho é ilustrada, utilizando-se parte dos dados de comprimento de fibras de eucalipto, apresentados em BRASIL (1976) e reproduzidos no Quadro II do Apêndice. Neste experimento, a espécie utilizada foi o Eucalyptus grandis Hill ex Maiden, aos três anos de idade, de povoamentos pertencentes à Duratex S/A Indústria e Comércio, instalado no Horto da Mina no município de Itupeva (LC1: localidade 1) e à Champion Papel e Celulose S/A, instalado no Horto Santa Teresinha, no município de Moji-Guaçu (LC2: localidade 2).

Consideraram-se três árvores da LC1 e duas da LC2, todas com 10,0 cm de diâmetro à altura do peito (DAP). Foram retiradas secções transversais (discos) das árvores, ao nível do DAP, e, nestes discos, consideraram-se os incrementos anuais de crescimento, denominados: posição 1, (PS1), região

próxima à medula; posição 3, (PS3), região próxima à casca; e posição 2, (PS2), região intermediária, correspondendo, respectivamente, ao 1º, 3º e 2º anos de crescimento.

Para cada árvore, foram obtidas três amostras correspondentes às posições 1, 2 e 3, relativas aos anos de crescimento. A partir deste material macerado, foi procedida a montagem de lâminas em geléia de glicerina. Nas lâminas, foram realizadas medições de fibras inteiras, totalmente ao acaso. O comprimento das fibras foi medido em micra, em microscópio com aumento de cem vezes.

O modelo matemático para uma medida do comprimento das fibras dos eucaliptos é descrito como:

$$y_{i_1 i_2 i_3 i_4 i_5} = \mu + \alpha_{i_1} + \beta_{i_1 i_2} + \delta_{i_1 i_2 i_3} + \gamma_{i_1 i_2 i_3 i_4} + \epsilon_{i_1 i_2 i_3 i_4 i_5}$$

onde

μ : efeito médio geral,

α_{i_1} : efeito da i_1 -ésima localidade,

$\beta_{i_1 i_2}$: efeito da i_2 -ésima árvore dentro da i_1 -ésima localidade,

$\delta_{i_1 i_2 i_3}$: efeito da i_3 -ésima posição dentro da i_2 -ésima árvore dentro da i_1 -ésima localidade,

$\gamma_{i_1 i_2 i_3 i_4}$: efeito da i_4 -ésima lâmina dentro da i_3 -ésima posição dentro da i_2 -ésima árvore dentro da i_1 -ésima localidade,

$\varepsilon_{i_1 i_2 i_3 i_4 i_5}$: efeito da i_5 -ésima fibra dentro da i_4 -ésima lâmina dentro da i_3 -ésima posição da i_2 -ésima árvore dentro da i_1 -ésima localidade.

A variação de cada um dos índices do modelo é a seguinte:

$$i_1 = 1, 2 ;$$

$$i_2 = 1, \dots, b_{i_1}, \text{ com:}$$

$$b_1 = 3$$

$$b_2 = 2$$

$$b_{\bullet} = 5 ;$$

$$i_3 = 1, \dots, m_{i_1 i_2}, \text{ com:}$$

$$m_{11} = 3$$

$$m_{12} = 3$$

$$m_{13} = 3$$

$$m_{21} = 3$$

$$m_{22} = 3$$

$$m_{\bullet\bullet} = 15 ;$$

$$i_4 = 1, \dots, t_{i_1 i_2 i_3}, \text{ com:}$$

$$t_{111} = 3$$

$$t_{112} = 1$$

$$t_{113} = 2$$

$$t_{121} = 2$$

$$t_{122} = 4$$

$$t_{123} = 2$$

$$t_{131} = 4$$

$$t_{132} = 2$$

$$t_{133} = 3$$

$$t_{211} = 2$$

$$t_{212} = 1$$

$$t_{213} = 4$$

$$t_{221} = 3$$

$$t_{222} = 2$$

$$t_{223} = 4$$

$$t_{\bullet\bullet\bullet} = 39 ;$$

$$i_5 = 1, \dots, n_{i_1 i_2 i_3 i_4}, \text{ com:}$$

$n_{1111} = 3$	$n_{1112} = 2$	$n_{1113} = 4$	
$n_{1121} = 4$			
$n_{1131} = 2$	$n_{1132} = 3$		
$n_{1211} = 4$	$n_{1212} = 3$		
$n_{1221} = 4$	$n_{1222} = 2$	$n_{1223} = 1$	$n_{1224} = 3$
$n_{1231} = 2$	$n_{1232} = 3$		
$n_{1311} = 2$	$n_{1312} = 3$	$n_{1313} = 3$	$n_{1314} = 4$
$n_{1321} = 2$	$n_{1322} = 3$		
$n_{1331} = 1$	$n_{1332} = 3$	$n_{1333} = 4$	
$n_{2111} = 2$	$n_{2112} = 3$		
$n_{2121} = 4$			
$n_{2131} = 2$	$n_{2132} = 4$	$n_{2133} = 4$	$n_{2134} = 3$
$n_{2211} = 3$	$n_{2212} = 3$	$n_{2213} = 4$	
$n_{2221} = 3$	$n_{2222} = 2$		
$n_{2231} = 4$	$n_{2232} = 4$	$n_{2233} = 4$	$n_{2234} = 2$
$n_{111\bullet} = 9$	$n_{112\bullet} = 4$	$n_{113\bullet} = 5$	
$n_{121\bullet} = 7$	$n_{122\bullet} = 10$	$n_{123\bullet} = 5$	

$$\begin{array}{lll}
n_{131\bullet} = 12 & n_{132\bullet} = 5 & n_{133\bullet} = 8 \\
n_{211\bullet} = 5 & n_{212\bullet} = 4 & n_{213\bullet} = 13 \\
n_{221\bullet} = 10 & n_{222\bullet} = 5 & n_{223\bullet} = 14 \\
n_{11\bullet\bullet} = 18 & n_{12\bullet\bullet} = 22 & n_{13\bullet\bullet} = 25 \\
n_{21\bullet\bullet} = 22 & n_{22\bullet\bullet} = 29 & \\
n_{1\bullet\bullet\bullet} = 65 & n_{2\bullet\bullet\bullet} = 51 & \\
n_{\bullet\bullet\bullet\bullet} = 116 & &
\end{array}$$

Os resultados, dados a seguir, referem-se à análise da variância, teste de hipótese das comparações entre pares de médias e testes da condição de estimabilidade e de hipótese de algumas funções requeridas.

COMPRIMENTO DAS FIBRAS DA MADEIRA DE EUCALYPTUS GRANDIS HILL EX MALDEN

ANALISE DA VARIANCIA

```

*****
CAUSA DE VARIACAO          GL          SQ          QM          I
*****
CLASSES                    1          0.1275557542D+00    0.1275557542D+00    0.7734761634D+01
SUBCLASSES( 1) : CLASSES    3          0.5323566558D+00    0.1774522186D+00    0.1002995614D+02
SUBCLASSES( 2) : SUBCLASSES( 1) 10          0.1460671429D+01    0.1460671429D+00    0.5914516616D+01
SUBCLASSES( 3) : SUBCLASSES( 2) 24          0.2503868272D+00    0.1043278447D-01    0.6367156130D+00
RESIDUO                    77          0.1261669083D+01    0.1635531277D-01
*****
TOTAL                      115          0.3632639750D+01
*****

```

DESVIO PADRAO = 0.1280051279D+00 COEFICIENTE DE VARIACAO = 0.1391738276D+00
 MEDIA GERAL = 0.9197500000D+00

MEDIAS DENTRO DO ESTAGIO 1

0.9491230769D+00 0.8823137255D+00

MEDIAS DENTRO DO ESTAGIO 2

0.8438333333D+00 0.9916181818D+00 0.9873600000D+00

0.9636818182D+00 0.8205862069D+00

MEDIAS DENTRO DO ESTAGIO 3

0.7218888889D+00 0.9795000000D+00 0.9548000000D+00

0.8261428571D+00 0.1006600000D+01 0.1194200000D+01

0.9014166667D+00 0.9974000000D+00 0.1110000000D+01

0.7482000000D+00 0.1125750000D+01 0.9966923077D+00

0.7787000000D+00 0.1016400000D+01 0.7805714286D+00

MEDIAS DENTRO DO ESTAGIO 4

0.6570000000+00	0.7465000000+00	0.7582500000+00	
0.9795000000+00			
0.9235000000+00	0.9756666667+00		
0.8645000000+00	0.7750000000+00		
0.9265000000+00	0.1106000000+01	0.9740000000+00	0.1056000000+01
0.1082500000+01	0.1268666667+01		
0.9825000000+00	0.8696666667+00	0.9323333333+00	0.8615000000+00
0.1059500000+01	0.9560000000+00		
0.1038000000+01	0.1125000000+01	0.1116750000+01	
0.7700000000+00	0.7336666667+00		
0.1125750000+01			
0.9735000000+00	0.9972500000+00	0.1053250000+01	0.9360000000+00
0.8356666667+00	0.7573333333+00	0.7520000000+00	
0.9856666667+00	0.1062500000+01		
0.7522500000+00	0.7465000000+00	0.8482500000+00	0.7700000000+00

CONTRASTES ENTRE DUAS MEDIAS DENTRO DE CADA ESTAGIO

ESTAGIO 1

CONTRASTE	VALOR ESTIMADO	VARIANCIA	TESTE F
M(2) - M(3)	0.668093E143D-01	0.5733623433D-03	0.7784762090D+01

ESTAGIO 2

CONTRASTE	VALOR ESTIMADO	VARIANCIA	TESTE F
M(4) - M(5)	-0.1479848485D+00	0.1655081822D-02	0.1323168141D+02
M(4) - M(6)	-0.1435266667D+00	0.1565707474D-02	0.1315693026D+02
M(5) - M(6)	0.4458181818D-02	0.1400199096D-02	0.1419468501D-01
M(7) - M(8)	0.1430956113D+00	0.1309797322D-02	0.1563322327D+02

ESTAGIO 3

CONTRASTE	VALOR ESTIMADO	VARIANCIA	TESTE F
M(9) - M(10)	-0.2576111111D+00	0.5916918392D-02	0.1121588641D+02
M(9) - M(11)	-0.2329111111D+00	0.5097652558D-02	0.1064167969D+02
M(10) - M(11)	0.2470090000D-01	0.7373390551D-02	0.8274212664D-01
M(12) - M(13)	-0.1804571429D+00	0.3979289519D-02	0.8163566562D+01
M(12) - M(14)	-0.3680571429D+00	0.5617821186D-02	0.2411362981D+02
M(13) - M(14)	-0.1876000000D+00	0.4915593050D-02	0.7159616275D-01
M(15) - M(16)	-0.9598333333D-01	0.4642504764D-02	0.1984460853D+01
M(15) - M(17)	-0.2085833333D+00	0.3413606502D-02	0.1274517345D+02
M(16) - M(17)	-0.1126000000D+00	0.5325226455D-02	0.2380866542D+01
M(18) - M(19)	-0.3775500000D+00	0.7373390551D-02	0.1933221922D+02
M(18) - M(20)	-0.2484923077D+00	0.4537470958D-02	0.1360855566D+02
M(19) - M(20)	0.1290576923D+00	0.5356736792D-02	0.3109334787D+01
M(21) - M(22)	-0.2377000000D+00	0.4915593050D-02	0.1149429772D+02
M(21) - M(23)	-0.1871428571D-02	0.2808909617D-02	0.1246834315D-02
M(22) - M(23)	0.2358285714D+00	0.4447441284D-02	0.1250496894D+02

ESTAGIO 4

CONTRASTE	VALOR ESTIMADO	VARIANCIA	TESTE F
M(24) - M(25)	-0.8950000000D-01	0.1365442698D-01	0.5866412417D+00
M(24) - M(26)	-0.1012500000D+00	0.9558098791D-02	0.1072552474D+01
M(25) - M(26)	-0.1175000000D-01	0.1228893458D-01	0.1123465483D-01
M(28) - M(29)	-0.5216666667D-01	0.1365442698D-01	0.1993024784D+00

CONTRASTE	VALOR ESTIMADO	VARIANCI A	TESTE F
M(30) - M(31)	0.8950000000D-01	0.9558098791D-02	0.8330539253D+00
M(32) - M(33)	-0.1795000000D+00	0.1228893458D-01	0.2621880575D+01
M(32) - M(34)	-0.4750000000D-01	0.2048164096D-01	0.1101596305D+00
M(32) - M(35)	-0.1315000000D+00	0.9558098791D-02	0.1809172554D+01
M(33) - M(34)	0.1320000000D+00	0.2457796916D-01	0.7039275721D+00
M(33) - M(35)	0.4800000000D-01	0.1365442698D-01	0.1687364840D+00
M(34) - M(35)	-0.8400000000D-01	0.2184707849D-01	0.3229722457D+00
M(36) - M(37)	-0.1861666667D+00	0.1365442698D-01	0.2538226454D+01
M(38) - M(39)	0.1128333333D+00	0.1365442698D-01	0.9323980513D+00
M(38) - M(40)	0.5016666667D-01	0.1365442698D-01	0.1843134426D+00
M(38) - M(41)	0.1210000000D+00	0.1228893458D-01	0.1191392170D+01
M(39) - M(40)	-0.6266666667D-01	0.1092354120D-01	0.3595089761D+00
M(39) - M(41)	0.8166666667D-02	0.9558098791D-02	0.6977794005D-02
M(40) - M(41)	0.7083333333D-01	0.9558098791D-02	0.5249329622D+00
M(42) - M(43)	0.1035000000D+00	0.1365442698D-01	0.7845257815D+00
M(44) - M(45)	-0.8700000000D-01	0.2184707849D-01	0.3464536462D+00
M(44) - M(46)	-0.7875000000D-01	0.2048164096D-01	0.3027864081D+00
M(45) - M(46)	0.8250000000D-02	0.9558098791D-02	0.7120924516D-02
M(47) - M(48)	0.3633333333D-01	0.1365442698D-01	0.9668008132D-01
M(50) - M(51)	-0.2375000000D-01	0.1228893458D-01	0.4589984603D-01
M(50) - M(52)	-0.7975000000D-01	0.1228893458D-01	0.5175417431D+00
M(50) - M(53)	0.3750000000D-01	0.1365442698D-01	0.1029885767D+00
M(51) - M(52)	-0.5600000000D-01	0.8192656385D-02	0.3827818295D+00
M(51) - M(53)	0.6125000000D-01	0.9558098791D-02	0.3925009123D+00
M(52) - M(53)	0.1172500000D+00	0.9558098791D-02	0.1438315590D+01
M(54) - M(55)	0.7833333333D-01	0.1092354120D-01	0.5617327752D+00
M(54) - M(56)	0.8366666667D-01	0.9558098791D-02	0.7323748440D+00
M(55) - M(56)	0.5333333333D-02	0.9558098791D-02	0.2975952129D-02
M(57) - M(58)	-0.7683333333D-01	0.1365442698D-01	0.4323404503D+00
M(59) - M(60)	0.5750000000D-02	0.8192656385D-02	0.4035626352D-02
M(59) - M(61)	-0.9600000000D-01	0.8192656385D-02	0.1124909866D+01
M(59) - M(62)	-0.1775000000D-01	0.1228893458D-01	0.2563779766D-01
M(60) - M(61)	-0.1017500000D+00	0.8192656385D-02	0.1263700321D+01
M(60) - M(62)	-0.2350000000D-01	0.1228893458D-01	0.4493861934D-01
M(61) - M(62)	0.7825000000D-01	0.1228893458D-01	0.4982561790D+00

COEFICIENTES DA FUNCAO					VALOR ESTIMADO	VARIANCIA	TESTE F

0.0	1.00000	-1.00000	0.27000	0.33990			
0.38300	-0.43000	-0.57000	0.13960	0.06260			
0.07700	0.10500	0.15400	0.07750	0.13300			
0.07700	0.12300	-0.09300	-0.05500	-0.25600			
-0.19600	-0.09500	-0.27600	0.04600	0.01100			
0.06200	0.06200	0.03100	0.04600	0.06250			
0.04600	0.06200	0.03100	0.01500	0.06250			
0.03100	0.04600	0.03100	0.04500	0.04500			
0.06200	0.03100	0.04600	0.01500	0.04500			
0.06200	-0.03500	-0.05900	-0.07500	-0.03500			
-0.07800	-0.07500	-0.05900	-0.05800	-0.05900			
-0.07800	-0.05900	-0.03900	-0.07900	-0.07900			
-0.07900	-0.03900						
					0.6719266670-01	0.57339972410-03	0.76736316210+01
-0.08900	5.02700	-5.11600	1.33600	1.49900			
2.19300	-1.86200	-3.25400	0.53200	0.43750			
0.37200	0.29500	0.66200	0.53300	0.00600			
0.57400	0.61300	-0.20000	-0.03500	-1.02700			
-1.07500	-1.02300	-1.17600	0.15400	0.21500			
0.12300	0.43200	0.23500	0.13700	0.05600			
0.21200	0.12500	0.12300	0.17200	0.24200			
0.32100	0.21700	0.61500	0.04700	0.03500			
0.30900	0.25900	0.31500	0.15300	0.45000			
0.21000	-0.37000	-0.43000	-0.03500	-0.03500			
-0.06600	-0.67400	-0.25000	-0.23000	-0.47300			
-0.37200	-0.25100	-0.75200	-0.25300	-0.20700			
-0.30100	-0.41500						

FUNCAO NAO ESTIMAVEL							

MATRIZ DIAGONAL D INVERSA							
3	2	4	4	2	3	4	3
1	3	4	2	3	4	2	4

Pelo quadro da análise da variância, nota-se que somente o resultado do teste F da causa de variação "lâmina dentro de posição dentro de árvore dentro de localidade" foi não significativo, ao nível $\alpha = 0,05$.

Analisando-se os contrastes entre duas médias dentro de cada um dos estágios, verifica-se que existem alguns resultados significativos ao nível adotado. Por exemplo, é significativo o contraste representado por $M(4) - M(5)$, que compara o efeito da árvore 1 com o da árvore 2, dentro da localidade 1 (Itupeva).

Considerando-se os outros tipos de funções, observa-se que algumas são estimáveis e outras não. Por exemplo, a primeira delas é não estimável, pois

$$\text{coef}(\beta_{11}) = 3 \quad \text{e} \quad \sum_{i_3=1}^3 \text{coef}(\delta_{11i_3}) = 5 ,$$

e portanto

$$\sum_{i_3=1}^3 \text{coef}(\delta_{11i_3}) \neq \text{coef}(\beta_{11}) ,$$

que não satisfaz (3.21).

A quarta dessas funções é estimável. Ou seja,

$$\sum_{i_1=1}^2 \text{coef}(\alpha_{i_1}) = \text{coef}(\mu) = 0 ,$$

$$\sum_{i_2=1}^3 \text{coef}(\beta_{1i_2}) = \text{coef}(\alpha_1) = 0 ,$$

$$\sum_{i_2=1}^2 \text{coef}(\beta_{2i_2}) = \text{coef}(\alpha_2) = 0 ,$$

$$\sum_{i_3=1}^3 \text{coef}(\delta_{11i_3}) = \text{coef}(\beta_{11}) = 1 ,$$

$$\sum_{i_3=1}^3 \text{coef}(\delta_{12i_3}) = \text{coef}(\beta_{12}) = -1 ,$$

$$\sum_{i_3=1}^3 \text{coef}(\delta_{13i_3}) = \text{coef}(\beta_{13}) = 0 ,$$

$$\sum_{i_3=1}^3 \text{coef}(\delta_{21i_3}) = \text{coef}(\beta_{21}) = 0 ,$$

$$\sum_{i_3=1}^3 \text{coef}(\delta_{22i_3}) = \text{coef}(\beta_{22}) = 0 ,$$

$$\sum_{i_4=1}^3 \text{coef}(\gamma_{111i_4}) = \text{coef}(\delta_{111}) = 0,5 ,$$

$$\sum_{i_4=1}^1 \text{coef}(\gamma_{112i_4}) = \text{coef}(\delta_{112}) = 0,222 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{113i_4}) = \text{coef}(\delta_{113}) = 0,278 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{121i_4}) = \text{coef}(\delta_{121}) = -0,318 ,$$

$$\sum_{i_4=1}^4 \text{coef}(\gamma_{122i_4}) = \text{coef}(\delta_{122}) = -0,455 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{123i_4}) = \text{coef}(\delta_{123}) = -0,227 ,$$

$$\sum_{i_4=1}^4 \text{coef}(\gamma_{131i_4}) = \text{coef}(\delta_{131}) = 0 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{132i_4}) = \text{coef}(\delta_{132}) = 0 ,$$

$$\sum_{i_4=1}^3 \text{coef}(\gamma_{133i_4}) = \text{coef}(\delta_{133}) = 0 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{211i_4}) = \text{coef}(\delta_{211}) = 0 ,$$

$$\sum_{i_4=1}^1 \text{coef}(\gamma_{212i_4}) = \text{coef}(\delta_{212}) = 0 ,$$

$$\sum_{i_4=1}^4 \text{coef}(\gamma_{213i_4}) = \text{coef}(\delta_{213}) = 0 ,$$

$$\sum_{i_4=1}^3 \text{coef}(\gamma_{221i_4}) = \text{coef}(\delta_{221}) = 0 ,$$

$$\sum_{i_4=1}^2 \text{coef}(\gamma_{222i_4}) = \text{coef}(\delta_{222}) = 0 \text{ e}$$

$$\sum_{i_4=1}^4 \text{coef}(\gamma_{223i_4}) = \text{coef}(\delta_{223}) = 0 .$$

Observa-se também, que o resultado do teste F relativo ao teste da hipótese $g'b = 0$, associado a essa função é significativo ao nível α adotado.

6. CONCLUSÕES

Conclui-se que, considerando o modelo linear de classificação hierárquica não-balanceado ou balanceado, uma função linear dos parâmetros $q'b$ é estimável se e somente se os elementos do vetor q (coeficientes do vetor dos parâmetros) satisfazem as condições estabelecidas no Teorema 1.

Deduz-se ainda, por este teorema, que são estimáveis as diferenças entre duas médias de classes e todas as diferenças entre duas médias de subclasses de um j -ésimo estágio, dentro da subclasse do $(j-1)$ -ésimo estágio.

Com relação à estimabilidade e testabilidade de funções lineares paramétricas, no modelo linear de classificação hierárquica, poderão ser agora prontamente estabelecidas e verificadas, através do programa computacional desenvolvido.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALALOUF, I.S. e P.H. STYAN, 1979. Characterizations of estimability in the general linear model. Annals of Statistics. San Francisco, 7(1):194-200.
- ANDERSON, R.L. e P.P. CRUMP, 1967. Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process. Technometrics. Richmond, 9(4): 499 - 516.
- BOSE, R.C., 1944. The fundamental theorem of linear estimation. In: Proceedings of the Thirty-first Indian Science Congress, p. 2 - 3.
- BRASIL, M.A.M., 1976. Densidade Básica e Características das Fibras da Madeira de *Eucalyptus grandis* Hill ex Maiden aos 3 anos de Idade. Piracicaba, ESALQ - USP, 126 pp. (Tese de Doutorado).
- EDELMAN, D.A., 1974. Three - stage nested designs with composited samples. Technometrics. Richmond, 16(3): 409-417.

- EISEN, E.J., 1966. The quasi-F test for an unnested fixed factor in an unbalanced hierarchical design with a mixed model. Biometrics. Raleigh, 22(4): 937-942.
- GATES, C.E. e C.J. SHIUE, 1962. The analysis of variance on the s - stage hierarchical classification. Biometrics. Raleigh, 18: 529-536.
- GAYLOR, D.W. e T.O. HARTWELL, 1969. Expected mean squares for nested classifications. Biometrics. Raleigh, 25(2): 427 - 430.
- GOWER, J.C., 1962. Variance component estimation for unbalanced hierarchical classifications. Biometrics. Raleigh, 18: 537-542.
- GRAYBILL, F.A., 1976. Theory and Application of the Linear Model. Massachusetts, Duxbury Press. 704 pp.
- KUSSMAUL, K. e R.L. ANDERSON, 1967. Estimation of variance components in two - stage nested designs with composite samples. Technometrics. Richmond, 9(3): 373-389.
- LaMOTE, L.R., 1972. Notes on the covariance matrix of a random, nested ANOVA model. Ann. Math. Statist.. Baltimore, 43(2): 659-662.
- MAHAMUNULU, D.M., 1963. Sampling variances of the estimates of variance components in the unbalanced 3 - way nested classification. Ann. Math. Statist.. Baltimore, 34: 521-527.
- RAO, C.R., 1965. Linear Statistical Inference and its Applications. New York, John Wiley & Sons, Inc.. 625 pp.

- RAO, C.R. e S.K. MITRA, 1971a. Generalized inverse of a matrix and applications. In: Sixth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, p. 601-620.
- RAO, C.R. e S.K. MITRA, 1971b. Generalized Inverse of Matrices and Its Applications. New York, John Wiley & Sons, Inc.. 240 pp.
- SAHAI, H., 1976. A comparison of estimators of variance components in the balanced three-stage nested random effects model using mean squared error criterion. J. Am. Statist. Ass. Washington, 71(354): 435-444.
- SEARLE, S.R., 1961. Variance components in the unbalanced 2-way nested classification. Ann. Math. Statistic.. Baltimore, 32: 1161-1166.
- SEARLE, S.R., 1971. Linear Models. New York, John Wiley & Sons, Inc.. 532 pp.
- SEELY, J., 1977. Estimability and linear hypothesis. American Statistician. Washington, 31(3): 121-123.
- THOMPSON, W.O. e R.L. ANDERSON, 1975. A comparison of designs and estimators for the two-stage nested random model. Technometrics. Richmond, 17(1): 37-44.
- TIAO, G.C. e G.E.P. BOX, 1967. Baysean analysis of a three-component hierarchical design model. Biometrika. Londres, 54(1 and 2): 109-125.
- VIDMAR, T.J. e M.N. BRUNDEN, 1980. Optimal allocation with fixed power in a completely randomized design with leves of subsampling. Commun. Statist. - Theor. Meth.. New York, A9(7): 757-763.

8. APÊNDICE

Considerando-se a seguinte notação: LC para indicar localidades, AV para árvores, PS para posição, LM para lâminas e FB para fibras, pode, então, ser elaborado o seguinte quadro de dados dos comprimentos das fibras de eucaliptos:

Quadro II : Unidade experimental e valor do seu respectivo comprimento.

LC	AV	PS	LM	FB	Valor do comprimento
1	1	1	1	1	0,720
				2	0,620
				3	0,631
			2	1	0,738
				2	0,755
			3	1	0,844
				2	0,714
				3	0,673
				4	0,802
		2	1	1	1,097
				2	1,086
				3	0,826
				4	0,909
		3	1	1	0,791
				2	1,056
			2	1	1,121
				2	0,832
				3	0,974
	2	1	1	1	0,850
				2	0,944
				3	0,956
				4	0,708
			2	1	0,726
				2	0,844
				3	0,755
		2	1	1	1,033
				2	0,926

Quadro II - continuação

LC	AV	PS	LM	FB	Valor do comprimento
				3	0,844
				4	0,903
			2	1	1,115
				2	1,097
			3	1	0,974
				1	1,103
			4	2	0,979
				3	1,092
		3	1	1	1,050
				2	1,115
			2	1	1,357
				2	1,269
				3	1,180
	3	1	1	1	0,903
				2	1,062
			2	1	0,915
				2	0,856
				3	0,838
			3	1	0,920
				2	0,856
				3	1,021
			4	1	0,873
				2	0,909
				3	0,779
				4	0,885
		2	1	1	0,974
				2	1,145
			2	1	0,903
				2	0,915

Quadro II - continuação

LC	AV	PS	LM	FB	Valor do comprimento
				3	1,050
		3	1	1	1,038
			2	1	1,121
				2	1,204
				3	1,050
			3	1	0,944
				2	1,292
				3	1,080
				4	1,151
2	1	1	1	1	0,791
				2	0,749
			2	1	0,649
				2	0,791
				3	0,761
		2	1	1	1,729
				2	0,915
				3	0,856
				4	1,003
		3	1	1	1,162
				2	0,785
			2	1	0,909
				2	1,062
				3	0,915
				4	1,103
			3	1	0,979
				2	0,962
				3	1,239
				4	1,033

Quadro II - continuação

LC	AV	PS	LM	FB	Valor do comprimento
	2	1	4	1	0,979
				2	0,991
				3	0,838
			1	1	0,944
				2	0,820
				3	0,743
			2	1	0,702
				2	0,655
				3	0,915
			3	1	0,708
				2	0,861
				3	0,743
				4	0,696
		2	1	1	0,968
				2	0,956
				3	1,033
			2	1	1,210
				2	0,915
		3	1	1	0,761
				2	0,720
				3	0,820
				4	0,708
			2	1	0,726
				2	0,791
				3	0,714
				4	0,755
			3	1	0,773
				2	0,885
				3	0,791

Quadro II - continuação

LC	AV	PS	LM	FB	Valor do comprimento
				4	0,944
			4	1	0,920
				2	0,620