EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA

DEPARTAMENTO DE MÉTODOS QUANTITATIVOS

ANÁLISES DA VARIÂNCIA E DA COVARIÂNCIA LINEAR DE DADOS DE UMA CLASSIFICAÇÃO DUPLA NÃO BALANCEADA

Amauri de Almeida Machado

DMQ/A/56

Julho, 1982

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA
DEPARTAMENTO DE MÉTODOS QUANTITATIVOS

ANÁLISES DA VARIÂNCIA E DA COVARIÂNCIA LINEAR DE DADOS DE UMA CLASSIFICAÇÃO DUPLA NÃO BALANCEADA

Amauri de Almeida Machado

DMQ/A/56

Julho, 1982

DEPARTAMENTO DE MÉTODOS QUANTITATIVOS/EMBRAPA
Super Center Venâncio 2000 - Sala 722
70333 Brasília, DF
Brasil

Machado, Amauri de Almeida
Análise da variância e da covariância linear de dados de uma classificação dupla não balanceada. Brasília, EMBRAPA-DMQ, 1982.
93p (EMBRAPA-DMQ/A/56)

1. Estatística matemática - Análise de variância 2. Estatística matemática - Análise de covariância I. Empresa Brasileira de Pesquisa Agropecuária. Departamento de Métodos Quantitativos. Brasília, D.F. II. Título III. Série.

CDD: 519.5352

A meu pai *(in memoriam)*
A minha mãe,
A Dida e a Carolina,
D E D I C O.

ANÁLISES DA VARIÂNCIA E DA COVARIÂNCIA LINEAR DE DADOS DE UMA CLASSIFICAÇÃO DUPLA NÃO BALANCEADA

AMAURI DE ALMEIDA MACHADO
Engenheiro Agrônomo

Orientador: Dr. Humberto de Campos

Dissertação apresentada à Escola Superior de Agricultura "Luiz de Queiroz", da Universidade de São Paulo, para obtenção do título de Mestre em Estatística e Experimentação Agronômica.

PIRACICABA
Estado de São Paulo - Brasil
Março, 1981

AGRADECIMENTOS

Ao Prof. Dr. Humberto de Campos, pela amizade e orientação.

Ao Prof. Élio Paulo Zonta, a quem muito deve este trabalho.

Aos Professores Paulo Silveira Júnior e João Baptista da Silva, pela amizade e constante incentivo.

Aos Professores do Departamento de Matemática e Estatística da ESALQ, pelos ensinamentos.

À Universidade Federal de Pelotas, pela oportunidade concedida.

Ao Prof. Antonio C.S.A. Barros, pela versão do resumo para o inglês.

Aos colegas de curso, pelo excelente convívio.

À Maria Izalina Ferreira Alves e família, pela amizade.

Aos funcionários do Departamento de Matemática e Estatística da ESALQ, pela constante colaboração.

Í N D I C E

Pág.

ANÁLISES DA VARIÂNCIA E DA COVARIÂNCIA LINEAR DE DADOS DE UMA CLASSIFICAÇÃO DUPLA NÃO BALANCEADA

Autor: AMAURI DE ALMEIDA MACHADO
Orientador: Dr. Humberto de Campos

R E S U M O

O objetivo do presente trabalho é estabelecer uma teoria para as análises da variância e da covariância em classificações duplas não balanceadas. Para tanto utilizou-se o Método do Ajustamento de Constante, introduzido por YATES (1934), como gerador de estimativas e somas de quadrados.

O modelo linear básico é:

$$y_{ijk} = \mu + a_i + b_j + \gamma_{ij} + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk}$$

onde γ_{ij} representa a interação entre os níveis dos fatores A e B.

Entretanto, para maior facilidade nas deduções teóricas, utilizou-se o modelo sem interação, ou seja:

$$y_{ijk} = \mu + a_i + b_j + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk}$$

ou, na forma matricial,

$$\underset{\sim}{y} = \underset{\sim}{j}\mu + X_1\underset{\sim}{a} + X_2\underset{\sim}{b} + X_3\underset{\sim}{c} + \underset{\sim}{e}$$

Mesmo assim, além de um estudo completo acerca dos testes de significância para os efeitos principais e para a regressão, é apresentado, ainda, um procedimento no sentido de verificar a significância da interação.

As principais conclusões deste trabalho são:

a) nos casos onde a interação não está presente no modelo o teste de significância para os efeitos principais é um teste exato;

b) a presença da interação no modelo dificulta sobremaneira a interpretação das hipóteses, além de tornar aproximados os testes para os efeitos principais;

c) as hipóteses devem ser formuladas, preferentemente, em termos de funções lineares estimáveis. Caso contrário, deverão ser associadas à essas funções as restrições não estimáveis que possibilitaram expressá-la como tal.

VARIANCE AND LINEAR COVARIANCE ANALYSES FOR AN UNBALANCED TWO-WAY CLASSIFICATION

Author: AMAURI DE ALMEIDA MACHADO
Adviser: Dr. Humberto de Campos

SUMMARY

The objective of this dissertation is to establish a theory for the Analyses of Variances and Covariances in unbalanced crossed classifications. For this the Fitting Constants Method, due to YATES (1934) was used as a generator for estimates sums of squares.

The basic linear model was

$$y_{ijk} = \mu + a_i + b_j + \gamma_{ij} + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk}$$

where γ_{ij} represents the interactions between factors A and B.

However, for simplicity in theoretical developments, the model without interection parameters was first considered, that is

$$y_{ijk} = \mu + a_i + b_j + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk}$$

or, in matrix form,

$$\underset{\sim}{y} = \underset{\sim}{j}\mu + X_1\underset{\sim}{a} + X_2\underset{\sim}{b} + X_3\underset{\sim}{c} + \underset{\sim}{e} \quad .$$

In addition to a study on significance tests for main effects and regression, a proceding to test interaction is also presented.

The main conclusions of this work are:

a) when the interaction is absent in the model, the test of significance for the main effects is sharp;

b) the presence of the interaction in the model turn dif ficult the understanding of the hypothesis, turning the tests close for the main effects;

c) the hypothesis must be formulated, mainly, in term of estimables linear functions. Otherwise, must be associated to those functions, the nonestimables conditions, that permited the expression such as.

1. INTRODUÇÃO

As classificações não balanceadas podem ser subdivididas em três categorias principais. A primeira é aquela onde cada observação pode ser representada, matematicamente, através de um modelo de classificação simples. Neste caso, o número diferente de observações para cada tratamento, em nada modifica a análise tradicionalmente feita através dos totais marginais.

A segunda é aquela onde se tem um modelo de classificação dupla, cujo número de observações por tratamento é diferente mas proporcional. A proporcionalidade entre o número de observações mantém a ortogonalidade das estimativas, obtidas através dos totais marginais, mantendo, consequentemente, a aditividade das somas de quadrados.

Finalmente, uma terceira categoria onde o número de observações, além de desigual, é desproporcional. Os dados aqui enquadrados, requerem um tipo especial de análise para que as estimativas obtidas não resultem confundidas.

Uma classificação é dita balanceada quando o número de observações é constante para todas as combinações dos níveis dos fatores ou tratamentos, caso contrário, ela é dita não balanceada.

Nas classificações balanceadas, as estimativas e as somas de quadrados necessárias aos diversos testes de significância podem ser facilmente obtidas através dos totais marginais e, além do mais, existe um concenso geral sobre as hipóteses que estão sendo testadas.

Os métodos de análise de classificações não balanceadas são, geralmente, extensões da metodologia de análise das classificações balanceadas. Entretanto, essas extensões podem ser feitas de muitos modos, que, diga-se de passagem, não convergem a resultados únicos, fato esse que pode conduzir a alguma confusão na interpretação dos resultados.

Em geral, tais métodos são mais complexos, se comparados àquele apropriado para a análise de classificações balanceadas, tanto na fase de cálculos como na interpretação das hipóteses testadas. Assim, o estudo de alguns desses métodos não pode estar, de forma nenhuma, desvinculado das hipóteses a esses métodos associadas.

O presente trabalho estabelece a teoria geral das análises da variância e da covariância de classificações duplas não balanceadas, pelo Método do Ajustamento de Constantes, considerando um modelo do tipo I, ou de efeitos fixos, e a presença de p variáveis auxiliares.

2. REVISÃO DE LITERATURA

YATES (1934) apresenta diversos métodos apropriados para a análise de classificações não balanceadas, conforme se verificam a seguir:

a) Método das Médias Não Ponderadas

Este método tem como pressuposição básica a homogeneidade entre as variâncias das médias das subclasses, a qual só é conseguida se o número de observações em cada subclasse ou célula não diferem muito entre si. A proposição deste método deve-se à sua facilidade de cálculo e tem como principal desvantagem o fato de que as suas somas de quadrados divididas por σ^2 (SQ/σ^2) não tem distribuição χ^2. O quadro da análise da variância pode ser resumido como na tabela 1.

Na notação utilizada na tabela 1,

$$\bar{n}_h = ab/(\sum_{ij} 1/n_{ij})$$

$$\bar{x}_{ij} = \bar{y}_{ij.} = \sum_k y_{ijk}/n_{ij}$$

sendo a o número de níveis de A, b o número de níveis de B e SQR, a soma de quadrados do resíduo, constante para todos os métodos.

Tabela 1. Esquema de análise da variância associado ao Método das Médias Não Ponderadas.

Causas da Variação	Somas de Quadrados
A	$\bar{n}_h \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2$
B	$\bar{n}_h \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2$
AB	$\bar{n}_h \sum_{ij} (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$
Resíduo	SQR

b) Método da Soma de Quadrados de Médias Ponderadas

Este método, segundo o autor, é apropriado para a análise da variância de classificações duplas não balanceadas onde deva ser considerada a interação. Todas as somas de quadrados a ele associadas, divididas por σ^2, têm distribuição χ^2 permitindo, dessa forma, testes exatos tanto para a interação como para os efeitos principais. O método é resumido na tabela 2, onde:

$$\bar{x}_1 = \sum_i w_i \bar{x}_{i.} / \sum_i w_i \; , \qquad \bar{x}_2 = \sum_j v_j \bar{x}_{.j} / \sum_j v_j \; ,$$

$$w_i = b^2(\sum_j 1/n_{ij})^{-1} \quad \text{e} \quad v_j = a^2(\sum_i 1/n_{ij})^{-1}$$

Tabela 2. Esquema de análise da variância associado ao Método da Soma de Quadrados de Médias Ponderadas

Causas da Variação	Somas de Quadrados
A	$\sum_i w_i(\bar{x}_{i.} - \bar{x}_1)^2$
B	$\sum_j v_j(\bar{x}_{.j} - \bar{x}_2)^2$
AB	$R(\gamma/\mu,\alpha,\beta)$

c) Método do Ajustamento de Constantes

O Método do Ajustamento de Constantes, também denominado por alguns autores de Análise de Mínimos Quadrados, tem como vantagens o fato de ser um método rigoroso e proporcionar um teste exato para a interação e, caso não esteja presente, um teste exato para os efeitos principais. Na Tabela 3 vê-se as causas da variação e as somas de quadrados associadas a esse método, onde $R(\mu,\alpha)$, por exemplo, é a redução na soma de quadrados devida ao ajustamento do modelo

$$y_{ijk} = \mu + \alpha_i + e_{ijk}.$$

SEARLE (1971) considera que a grande vantagem do Método das Somas de Quadrados de Médias Ponderadas é que somas de quadrados, por ele geradas, divididas por σ^2, tem distribuição χ^2 permitindo assim, a utilização da estatística F nos testes de hipóteses. Segundo o autor

Tabela 3. Esquema de análise da variância associado ao Método do Ajustamento de Constantes

Causas da Variação	Somas de Quadrados
A (não ajustado)	$R(\alpha/\mu)$
A (ajustado p/B)	$R(\alpha/\mu,\beta)$
B (não ajustado)	$R(\beta/\mu)$
B (ajustado p/A)	$R(\beta/\mu,\alpha)$
AB	$R(\gamma/\mu,\alpha,\beta)$

$$E(QMA) = \frac{1}{a-1} \sum_{i=1}^{a} w_i \left[\alpha_i + \bar{\gamma}_{i.} - \frac{\sum_i w_i(\alpha_i + \bar{\gamma}_{i.})}{\sum_i w_i} \right]^2 + \sigma^2$$

$$E(QMB) = \frac{1}{b-1} \sum_{j=1}^{b} v_j \left[\beta_j + \bar{\gamma}_{.j} - \frac{\sum_j v_j(\beta_j + \bar{\gamma}_{.j})}{\sum_j v_j} \right]^2 + \sigma^2$$

De forma que a razão,

$$F = \frac{QMA}{QMR}$$

tem distribuição F e possibilita testar a hipótese:

$$H_0: \ \alpha_i + \bar{\gamma}_{i.} = \alpha_{i'} + \bar{\gamma}_{i'.} \ , \text{ para todo } i \neq i' .$$

Caso o modelo inclua a restrição $\gamma_{i.} = \gamma_{.j} = 0$, a hipótese testada é:

$$H_0: \ \alpha_i = \alpha_{i'} \ , \text{ para todo } i \neq i' .$$

Do mesmo autor é a notação R( ), muito útil na representação de somas de quadrados, especialmente quando se trata de casos não balanceados.

O termo R( ) é definido, por SEARLE (1971), como sendo a redução na soma de quadrados devida ao ajustamento de um determinado modelo. Assim, considere-se o modelo

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

Sejam $\underset{\sim}{\beta}' = [\underset{\sim}{\beta}_1', \underset{\sim}{\beta}_2']$ e $X = [X_1, X_2]$ partições convenientes de $\underset{\sim}{\beta}$ e X, respectivamente.

Pela definição, para o modelo

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

a redução na soma de quadrados será:

$$R(\underset{\sim}{\beta}) = \hat{\underset{\sim}{\beta}}'X'\underset{\sim}{y}$$

e para o modelo,

$$\underset{\sim}{y} = X_1\underset{\sim}{\beta}_1 + \underset{\sim}{e}$$

a redução será:

$$R(\underset{\sim}{\beta}_1) = \hat{\underset{\sim}{\beta}}_1'X_1'\underset{\sim}{y}$$

onde $\hat{\underset{\sim}{\beta}}$ e $\hat{\underset{\sim}{\beta}}_1$ são, respectivamente, uma das soluções dos sistemas

$$X'X\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{y}$$

e

$$X_1'X_1\hat{\underset{\sim}{\beta}}_1 = X_1'\underset{\sim}{y} .$$

Assegura o autor que, se $\underset{\sim}{y} \cap N(\underset{\sim}{\mu}, \sigma^2 I)$ então $R(\ )/\sigma^2$ tem distribuição χ^2 não central independentemente de SQR.

Finalmente, o autor define que:

$$R(\underset{\sim}{\beta}_1/\underset{\sim}{\beta}_2) = R(\underset{\sim}{\beta}) - R(\underset{\sim}{\beta}_2)$$

ou
$$R(\underset{\sim}{\beta}) = R(\underset{\sim}{\beta}_1/\underset{\sim}{\beta}_2) + R(\underset{\sim}{\beta}_2)$$

donde se conclui que $R(\underset{\sim}{\beta}_1/\underset{\sim}{\beta}_2)/\sigma^2$ e $R(\underset{\sim}{\beta}_2)/\sigma^2$ têm, independentemente, distribuição χ^2 não central.

Segundo HOCKING e SPEED (1975), a análise de modelos lineares pelos processos usuais, embora esteja amparada por uma rigorosa teoria matemática, apresentam sérias lacunas, especialmente na área dos testes de hipóteses formuladas para a análise de classificações não balanceadas. Para tanto, os autores propuseram a análise através do modelo

$$\underset{\sim}{y} = W\underset{\sim}{\mu} + \underset{\sim}{e}$$

sujeito a restrição

$$G\underset{\sim}{\mu} = \underset{\sim}{0}$$

onde $G\underset{\sim}{\mu}$ expressa relações conhecidas entre as médias que compoem o vetor $\underset{\sim}{\mu}$.

Como exemplo, pode-se citar o modelo básico apropriado para análise de classificações duplas:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

onde γ representa a interação entre os efeitos principais A e B.

Considerando que $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$, o modelo anterior pode ser reescrito como:

$$y_{ijk} = \mu_{ij} + e_{ijk}$$

Caso haja evidências sobre a inexistência da interação, então esse modelo estará sujeito à restrição

$$\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0$$

que seria equivalente ao modelo

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$$

SPEED *et alii* (1978), revendo os métodos para a análise de classificações não balanceadas, distinguem tais métodos pelas hipóteses a eles associadas e asseguram que a escolha do método deve basear-se somente nas hipóteses de interesse e não em facilidades de cálculo ou exigências de ortogonalidade de formas quadráticas. Assim, com relação ao Método das Médias Não Ponderadas, a simples facilidade de cálculo não justifica, de forma nenhuma, sua aplicação.

Num modelo de classificação dupla, as hipóteses comumente utilizadas na análise da variância, segundo esses autores, estão na tabela 4, colocadas sob a forma do modelo μ_{ij}.

Tabela 4. Hipóteses comumente utilizadas na análise da variância, tendo como base o modelo de classificação dupla

Fator A
$H_1: \bar{\mu}_{i.} = \bar{\mu}_{i'.}$
$H_2: \sum_j n_{ij}\mu_{ij}/n_{i.} = \sum_j n_{i'j}\mu_{i'j}/n_{i'.}$
$H_3: \sum_j n_{ij}\mu_{ij} = \sum_{i'} \sum_j n_{ij}n_{i'j}\mu_{i'j}/n_{.j}$
$H_4: \mu_{i1} = \mu_{i'1}$

Tabela 4. continuação

Fator B
$H_5: \bar{\mu}_{.j} = \bar{\mu}_{.j'}$
$H_6: \sum_i n_{ij}\mu_{ij}/n_{.j} = \sum_i n_{ij'}\mu_{ij'}/n_{.j'}$
$H_7: \sum_i n_{ij}\mu_{ij} = \sum_i \sum_{j'} n_{ij}n_{ij'}\mu_{ij'}/n_{i.}$
$H_8: \mu_{1j} = \mu_{1j'}$

Interação AB
$H_9: \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0$

As hipóteses associadas aos efeitos principais A, B e à interação AB são, respectivamente, H_1, H_5 e H_9 no Método das Somas de Quadrados de Médias Ponderadas. Pode-se notar perfeitamente que este método tem como exigência fundamental a existência de, pelo menos, uma observação em cada célula, ou seja, exige que $n_{ij} > 0$, para qualquer ij.

Porém a mesma exigência não é necessária para a aplicação do Método do Ajustamento de Constantes. As hipóteses associadas aos efeitos principais e à interação, na exata ordem em que se encontram na tabela 4, são: H_2, H_3, H_6, H_7 e H_9. Assim, $R(\alpha/\mu)$, calculada através do ajustamento do modelo de classificação simples, é a soma de quadrados apropriada para o teste de H_2. Alguns autores, entre eles SEARLE (1971), se referem a $R(\alpha/\mu)$ como "soma de quadrados para o fator A, ignorando o fator B e a interação AB".

Com relação à inclusão de variáveis auxiliares no modelo, FEDERER (1957) agrupa as análises da variância e da covariância em três casos: o primeiro é quando a interação não é considerada no modelo, o segundo é quando a interação está presente e os efeitos são considerados fixos e, finalmente, um terceiro caso onde a interação está presente e pelo menos um dos efeitos principais que a compõe é considerado aleatório.

Para a análise dos dois primeiros casos o autor emprega, respectivamente, o Método do Ajustamento de Constantes, o Método da Soma de Quadrados de Médias Ponderadas e, para o terceiro caso, apresenta um procedimento baseado na estimação dos componentes de variância pelos métodos propostos por HENDERSON (1953). Todos os casos são ilustrados, pelo autor, considerando apenas o caso de uma variável auxiliar no modelo. Além disso, indica processos a serem aplicados à classificações mais complexas com $\underset{\sim}{p}$ variáveis.

KATTI (1965) parte do modelo

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{b} + \underset{\sim}{e}$$

reescrevendo-o sob a forma:

$$\underset{\sim}{y} = X(\underset{\sim}{\beta} + \Delta\underset{\sim}{b}) + (Z - X\Delta)\underset{\sim}{b} + \underset{\sim}{e} =$$

$$= X\underset{\sim}{\beta}^* + (Z - X\Delta)\underset{\sim}{b} + \underset{\sim}{e}$$

onde Δ deve satisfazer o sistema $X'(Z - X\Delta) = \Phi$.

Pondera o autor que, se $X'X$ é não singular, então $\hat{\Delta} = (X'X)^{-1}X'Z$. No caso geral, entretanto, $\hat{\Delta} = (X'X)^{-}X'Z$, é uma solução do sistema $X'(Z - X\Delta) = \Phi$.

As estimativas obtidas através deste procedimento são:

$$\hat{\underset{\sim}{\beta}} = \hat{\underset{\sim}{\beta}}^* - \hat{\Delta}\hat{\underset{\sim}{b}} = (X'X)^{-}X'\underset{\sim}{y} - (X'X)^{-}X'Z\hat{\underset{\sim}{b}}$$

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

enquanto que a redução na soma de quadrados devida ao ajusta -

mento do modelo

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{b} + \underset{\sim}{e}$$

é: $$R(\underset{\sim}{\beta},\underset{\sim}{b}) = R(\underset{\sim}{\beta}^*,\underset{\sim}{b}) = R(\underset{\sim}{\beta}^*) + R(\underset{\sim}{b})$$

onde $R(\underset{\sim}{\beta}^*)$ é devida ao ajustamento de $\underset{\sim}{y} = X\underset{\sim}{\beta}^* + \underset{\sim}{e}$ e, $R(\underset{\sim}{b})$ devida ao ajustamento de $\underset{\sim}{y} = (Z - X\Delta)\underset{\sim}{b} + \underset{\sim}{e}$.

SEARLE (1971) argumenta que a análise de covariância pode ser feita utilizando-se o modelo $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{e}$, onde X é uma matriz composta de valores 0 e 1 e dos valores observados das covariáveis.

No entanto, para uma melhor distinção entre os dois tipos de parâmetros, o mesmo autor subdividiu a matriz X e o vetor $\underset{\sim}{\beta}$ em duas partes. O modelo, dessa forma, fica:

$$\underset{\sim}{y} = X\underset{\sim}{a} + Z\underset{\sim}{b} + \underset{\sim}{e}$$

onde X é uma matriz composta de valores da variável binária e Z dos valores da variável auxiliar, suposta de posto coluna completo, cujas colunas são independentes das colunas de X. Finalmente, o vetor $\underset{\sim}{a}$ é o vetor que contem a média geral $\underset{\sim}{\mu}$ mais os níveis dos fatores e suas interações; $\underset{\sim}{b}$ é o vetor dos coeficientes de regressão e $\underset{\sim}{e}$, o vetor dos erros, supostos independentes, com distribuição normal e variância comum σ^2.

Os estimadores obtidos pelo Método dos Mínimos Quadrados são:

$$\hat{\underset{\sim}{a}} = \underset{\sim}{a}^* - (X'X)^- X'Z\hat{\underset{\sim}{b}}$$

e $$\hat{\underset{\sim}{b}} = (Z'PZ)^{-1}Z'P\underset{\sim}{y}$$

sendo $\underset{\sim}{a}^* = (X'X)^- X'\underset{\sim}{y}$ e $P = I - X(X'X)^- X'$.

O mesmo estimador $\hat{\underset{\sim}{b}}$ pode, segundo o mesmo autor, ser obtido através do modelo:

$$y = R_z\underset{\sim}{b} + \underset{\sim}{e}$$

ou seja, $$\hat{\underset{\sim}{b}} = (R_z'R_z)^{-1}R_z'\underset{\sim}{y}$$

onde $R_z'R_z$ é a matriz das somas de quadrados e produtos do resíduo, relativas às variáveis auxiliares.

Dessa forma, o vetor solução do sistema de equações normais, na análise de covariância, é:

$$\begin{bmatrix} \hat{\underset{\sim}{a}} \\ \\ \hat{\underset{\sim}{b}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{a}^* - (X'X)^-X'Z\hat{\underset{\sim}{b}} \\ \\ (R_z'R_z)^{-1}R_z'\underset{\sim}{y} \end{bmatrix}$$

As variâncias e a covariância entre os estimadores são dadas por:

$$V(\hat{\underset{\sim}{a}}) = \left[(X'X)^- + (X'X)^-X'Z(R_z'R_z)^{-1}Z'X(X'X)^-\right]\sigma^2$$

$$V(\hat{\underset{\sim}{b}}) = (R_z'R_z)^{-1}\sigma^2$$

$$Cov(\hat{\underset{\sim}{a}},\hat{\underset{\sim}{b}}) = -(X'X)^-X'Z(R_z'R_z)^{-1}\sigma^2$$

A soma de quadrados do resíduo, nesse caso, é:

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - R(\underset{\sim}{a},\underset{\sim}{b})$$

ou $$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{y}'X(X'X)^-X'\underset{\sim}{y} - \underset{\sim}{y}'R_z(R_z'R_z)^{-1}R_z'\underset{\sim}{y}$$

ou $$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - R(\underset{\sim}{a}) - S.Q.Regr.$$

com $N - \rho(X) - \rho(Z)$ graus de liberdade, sendo N o número de observações, $\rho(X)$ e $\rho(Z)$, respectivamente, os postos de X e de Z.

O autor assegura, ainda, que as hipóteses H_1: $\underset{\sim}{b} = \underset{\sim}{0}$, H_2: $K'\underset{\sim}{a} = \underset{\sim}{0}$ e H_3: $K'\left[\underset{\sim}{a} + (X'X)^-X'Z\underset{\sim}{b}\right] = \underset{\sim}{0}$, são testadas, respectivamente, por:

$$F(H_1) = \frac{R(\underline{b}/\underline{a})/\rho(Z)}{SQR^*/[N - \rho(X) - \rho(Z)]} = \frac{S.Q.Regr./\rho(Z)}{QMR^*}$$

$$F(H_2) = \frac{Q/\rho(X)}{QMR^*}$$

$$F(H_3) = \frac{Q^*/\rho(X)}{QMR^*}$$

sendo,

$$Q = (K'\hat{\underline{a}})'[K'(X'X)^-K + K'(X'X)^-X'Z(Z'PZ)^{-1}Z'X(X'X)^-K]^{-1}K'\hat{\underline{a}}$$

e, $Q^* = (K'\underline{a}^*)'[K'(X'X)^-K]^{-1}K'\underline{a}^*$

DIAS (1981), estudando a análise de covariância intrablocos, com $\underline{p}$ variáveis auxiliares, no delineamento em blocos incompletos, caso particular dos ensaios equilibrados, considera o modelo matemático

$$y_{ij} = m + t_i + b_j + \sum_{w=1}^{p} a_w x_{ij}^{(w)} + e_{ij}$$

O principal objetivo do referido autor é justificar os fundamentos teóricos de tal análise. Assim sendo, demonstra que a soma de quadrados do resíduo, dividida por σ^2, ajustada para a regressão, tem distribuição χ^2 central e as demais, χ^2 não central, já que o modelo é suposto de efeitos fixos. Ademais, afirma que é correta a utilização do teste F para o teste das hipóteses de que os efeitos de tratamentos e os coeficientes de regressão são todos nulos.

Finalmente, determina a matriz de dispersão para as estimativas dos efeitos de tratamentos ajustados para blocos e regressão. Determina, ainda, uma fórmula para o cálculo da variância média do contraste entre duas médias de tratamentos ajustadas para blocos e regressão.

3. DESENVOLVIMENTO TEÓRICO

3.1. Modelo matemático

Supondo uma relação linear entre as variáveis independentes e a variável dependente, o modelo matemático considerado neste trabalho é:

$$y_{ijk} = \mu + a_i + b_j + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk} \qquad (1)$$

onde,

y_{ijk}: valor da k-ésima observação relativa à combinação ij, $(k = 0, 1, 2, \ldots, n_{ij})$;

μ: média geral teórica;

a_i: efeito do nível i do fator A ou primeira classificação, $(i = 1, 2, \ldots, v)$;

b_j: efeito do nível j do fator B ou segunda classificação, $(j = 1, 2, \ldots, r)$;

$c^{(g)}$: coeficiente de regressão linear múltipla , onde g representa o grupo de variáveis auxiliares, $(g = 1, 2, \ldots, p)$;

$x_{ijk}^{(g)}$: desvio de cada observação $X_{ijk}^{(g)}$ em relação à média geral do respectivo grupo, $\overline{X}^{(g)}$;

e_{ijk}: erro experimental associado à respectiva observação y_{ijk}, supostos independentes, normalmente distribuídos com média zero e variância constante σ^2.

Esse modelo de análise de covariância linear múltipla é o mesmo utilizado por FEDERER (1957) e semelhante àquele sugerido por ZELEN (1957), também utilizado por DIAS (1981). Estes últimos, entretanto, consideraram o caso dos blocos incompletos onde $n_{ij} = 0$ ou $n_{ij} = 1$, ou, de maneira mais completa, o caso dos blocos incompletos balanceados.

Optando-se pelo esquema matricial, o modelo (1) pode ser expresso como:

$$\underline{y} = \underline{j}\mu + X_1\underline{a} + X_2\underline{b} + X_3\underline{c} + \underline{e} \qquad (2)$$

sendo,

$\underline{y}$: vetor $(n_{..} \times 1)$ das observações da variável dependente;

$\underline{j}$: vetor $(n_{..} \times 1)$ composto de 1's referentes aos coeficientes da média geral μ;

X_1: matriz $(n_{..} \times v)$ composta de 0's e de 1's, coeficientes de a_i;

$\underline{a}$: vetor $(v \times 1)$ dos efeitos a_i;

X_2: matriz $(n_{..} \times r)$, semelhante a X_1, relativa aos coeficientes de b_j;

$\underline{b}$: vetor $(r \times 1)$ dos efeitos b_j;

X_3: matriz $(n_{..} \times p)$ cujos componentes são os valores da variável independente ou covariável suposta de colunas independentes entre si e das colunas de X_1 e X_2;

$\underline{c}$: vetor $(p \times 1)$ dos coeficientes de regressão;

$\underline{e}$: vetor $(n_{..} \times 1)$ dos erros experimentais.

Um procedimento que facilita as deduções teóricas é considerar o efeito da média somado ao efeito de uma das classificações. Embora a notação continue a mesma, supor-se-á, sem perda de generalidade, que o efeito da média está somado ao efeito da segunda classificação ou ao fator B.

Assim, pode-se reescrever (2) da seguinte forma:

$$\underset{\sim}{y} = X_1\underset{\sim}{a} + X_2\underset{\sim}{b} + X_3\underset{\sim}{c} + \underset{\sim}{e} \qquad (3)$$

Nos casos onde deva ser considerada a interação entre os fatores ou classificações, o modelo será:

$$\underset{\sim}{y} = X_1\underset{\sim}{a} + X_2\underset{\sim}{b} + X_3\underset{\sim}{c} + X_4\underset{\sim}{\gamma} + \underset{\sim}{e} \qquad (4)$$

onde X_4 é uma matriz ($n_{..}$ x s), de definição semelhante a X_1 e X_2 e γ é o vetor (s x 1) das interações entre os níveis dos fatores, sendo s o número de células da tabela de dupla entrada que contém pelo menos uma observação.

Como a soma de quadrados para a interação pode ser obtida por uma diferença entre somas de quadrados, como pode ser visto no item 3.10 deste trabalho, todas as demais deduções serão feitas considerando o modelo (3).

3.2. Sistema de equações normais

Os modelos (2) e (3) podem ser postos na forma

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{e} \qquad (5)$$

onde $X = [X_1, X_2, X_3]$ e $\underset{\sim}{\beta}' = [\underset{\sim}{a}', \underset{\sim}{b}', \underset{\sim}{c}']$.

Considerando o modelo (2), o sistema de equações normais $X'X\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{y}$, obtido pelo Método dos Quadrados Mínimos, pode ser representado por:

$$\begin{bmatrix} \underset{\sim}{j}'\underset{\sim}{j} & \underset{\sim}{j}'X_1 & \underset{\sim}{j}'X_2 & \underset{\sim}{j}'X_3 \\ X_1'\underset{\sim}{j} & X_1'X_1 & X_1'X_2 & X_1'X_3 \\ X_2'\underset{\sim}{j} & X_2'X_1 & X_2'X_2 & X_2'X_3 \\ X_3'\underset{\sim}{j} & X_3'X_1 & X_3'X_2 & X_3'X_3 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\underset{\sim}{a}} \\ \hat{\underset{\sim}{b}} \\ \hat{\underset{\sim}{c}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{j}'\underset{\sim}{y} \\ X_1'\underset{\sim}{y} \\ X_2'\underset{\sim}{y} \\ X_3'\underset{\sim}{y} \end{bmatrix}$$

A notação melhora sensivelmente se for considerado o esquema:

$$\begin{bmatrix} n_{..} & \underset{\sim}{r}' & \underset{\sim}{k}' & \underset{\sim}{0}' \\ \underset{\sim}{r} & R & N & A \\ \underset{\sim}{k} & N' & K & D \\ \underset{\sim}{0} & A' & D' & E \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\underset{\sim}{a}} \\ \hat{\underset{\sim}{b}} \\ \hat{\underset{\sim}{c}} \end{bmatrix} = \begin{bmatrix} y_{...} \\ \underset{\sim}{T} \\ \underset{\sim}{B} \\ \underset{\sim}{P} \end{bmatrix} \qquad (6)$$

As submatrizes de X'X têm a composição que segue:

$$\underset{\sim}{r}' = [n_{1.}, n_{2.}, \ldots, n_{v.}]$$

$$\underset{\sim}{k}' = [n_{.1}, n_{.2}, \ldots, n_{.r}]$$

$$R = \mathrm{diag}\{n_{1.}, n_{2.}, \ldots, n_{v.}\}$$

$$K = \mathrm{diag}\{n_{.1}, n_{.2}, \ldots, n_{.r}\}$$

$$N = {}_{v}[n_{ij}]_{r}$$

sendo n_{ij} o número de observações que cada combinação $\underline{ij}$ fornece.

$$A = \begin{bmatrix} x_{1..}^{(1)} & x_{1..}^{(2)} & \ldots & x_{1..}^{(p)} \\ x_{2..}^{(1)} & x_{2..}^{(2)} & \ldots & x_{2..}^{(p)} \\ \ldots & \ldots & \ldots & \ldots \\ x_{v..}^{(1)} & x_{v..}^{(2)} & \ldots & x_{v..}^{(p)} \end{bmatrix}$$

$$D = \begin{bmatrix} x^{(1)}_{.1.} & x^{(2)}_{.1.} & \cdots & x^{(p)}_{.1.} \\ x^{(1)}_{.2.} & x^{(2)}_{.2.} & \cdots & x^{(p)}_{.2.} \\ \cdots & \cdots & \cdots & \cdots \\ x^{(1)}_{.r.} & x^{(2)}_{.r.} & & x^{(p)}_{.r.} \end{bmatrix}$$

$$E = \begin{bmatrix} \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(1)}_{ijk}x^{(1)}_{ijk} & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(1)}_{ijk}x^{(2)}_{ijk} & \cdots & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(1)}_{ijk}x^{(p)}_{ijk} \\ \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(2)}_{ijk}x^{(1)}_{ijk} & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(2)}_{ijk}x^{(2)}_{ijk} & \cdots & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(2)}_{ijk}x^{(p)}_{ijk} \\ \cdots & \cdots & \cdots & \cdots \\ \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(p)}_{ijk}x^{(1)}_{ijk} & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(p)}_{ijk}x^{(2)}_{ijk} & \cdots & \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(p)}_{ijk}x^{(p)}_{ijk} \end{bmatrix}$$

enquanto que

$$\hat{\underset{\sim}{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \cdots \\ \hat{a}_v \\ \hat{b}_1 \\ \hat{b}_2 \\ \cdots \\ \hat{b}_r \\ \hat{c}_1 \\ \hat{c}_2 \\ \cdots \\ \hat{c}_p \end{bmatrix} \quad \text{e} \quad X'\underset{\sim}{y} = \begin{bmatrix} y_{...} \\ y_{1..} \\ y_{2..} \\ \cdots \\ y_{v..} \\ y_{.1.} \\ y_{.2.} \\ \cdots \\ y_{.r.} \\ \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(1)}_{ijk}y_{ijk} \\ \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(2)}_{ijk}y_{ijk} \\ \cdots \\ \underset{ijk}{\Sigma\Sigma\Sigma}\, x^{(p)}_{ijk}y_{ijk} \end{bmatrix}$$

sendo $n_{i.} = \sum_j n_{ij}$; $n_{.j} = \sum_i n_{ij}$ e $n_{..} = \sum_i \sum_j n_{ij}$.

Da mesma forma,

$$y_{i..} = \sum_j \sum_k y_{ijk} \; ; \; y_{.j.} = \sum_i \sum_k y_{ijk} \text{ e } y_{...} = \sum_i \sum_j \sum_k y_{ijk} .$$

Através de (3) obtém-se o seguinte sistema de equações normais:

$$\begin{cases} R\hat{\underset{\sim}{a}} + N\hat{\underset{\sim}{b}} + A\hat{\underset{\sim}{c}} = \underset{\sim}{T} & (7) \\ N'\hat{\underset{\sim}{a}} + K\hat{\underset{\sim}{b}} + D\hat{\underset{\sim}{c}} = \underset{\sim}{B} & (8) \\ A'\hat{\underset{\sim}{a}} + D'\hat{\underset{\sim}{b}} + E\hat{\underset{\sim}{c}} = \underset{\sim}{P} & (9) \end{cases}$$

A estimação simultânea de todos os parâmetros frequentemente torna-se por demais trabalhosa devido à presença de matrizes de ordem muito elevada. Pode-se lançar mão, entretanto, de um processo descrito em alguns textos como, por exemplo, SEARLE (1971), HARVEY (1975), dentre outros, denominado de "absorção de equações". Essa "absorção" é feita de forma que sejam "absorvidas" as equações referentes ao fator que tenha maior número de níveis. Supor-se-á, inicialmente, que a segunda classificação ou fator B possui maior número de níveis.

Isolando $\hat{\underset{\sim}{b}}$ em (8) e substituindo em (7) e (9), tem-se:

$$\begin{cases} R\hat{\underset{\sim}{a}} + NK^{-1}\underset{\sim}{B} - NK^{-1}D\hat{\underset{\sim}{c}} - NK^{-1}N'\hat{\underset{\sim}{a}} + A\hat{\underset{\sim}{c}} = \underset{\sim}{T} \\ A'\hat{\underset{\sim}{a}} + D'K^{-1}\underset{\sim}{B} - D'K^{-1}D\hat{\underset{\sim}{c}} - D'K^{-1}N'\hat{\underset{\sim}{a}} + E\hat{\underset{\sim}{c}} = \underset{\sim}{P} \end{cases}$$

Reunindo os termos semelhantes, obtém-se:

$$\begin{cases} (R - NK^{-1}N')\underset{\sim}{\hat{a}} + (A - NK^{-1}D)\underset{\sim}{\hat{c}} = \underset{\sim}{T} - NK^{-1}\underset{\sim}{B} \\ (A' - D'K^{-1}N')\underset{\sim}{\hat{a}} + (E - D'K^{-1}D)\underset{\sim}{\hat{c}} = \underset{\sim}{P} - D'K^{-1}\underset{\sim}{B} \end{cases}$$

Fazendo,

$$C = R - NK^{-1}N'$$
$$Z = A - NK^{-1}D$$
$$W = E - D'K^{-1}D$$
$$\underset{\sim}{Q} = \underset{\sim}{T} - NK^{-1}\underset{\sim}{B}$$
$$\underset{\sim}{S} = \underset{\sim}{P} - D'K^{-1}\underset{\sim}{B}$$

o sistema fica:

$$\begin{cases} C\underset{\sim}{\hat{a}} + Z\underset{\sim}{\hat{c}} = \underset{\sim}{Q} & (10) \\ Z'\underset{\sim}{\hat{a}} + W\underset{\sim}{\hat{c}} = \underset{\sim}{S} & (11) \end{cases}$$

equivalente ao sistema (6), absorvida a equação,

$$\underset{\sim}{j}\hat{\mu} + \underset{\sim}{\hat{b}} = K^{-1}[\underset{\sim}{B} - N'\underset{\sim}{\hat{a}} - D\underset{\sim}{\hat{c}}] \qquad (12)$$

Supondo W não singular, obtém-se, através de (11), que:

$$\underset{\sim}{\hat{c}} = W^{-1}\underset{\sim}{S} - W^{-1}Z'\underset{\sim}{\hat{a}}$$

Substituindo em (10) fica:

$$C\underset{\sim}{\hat{a}} + Z(W^{-1}\underset{\sim}{S} - W^{-1}Z'\underset{\sim}{\hat{a}}) = \underset{\sim}{Q}$$

Reunindo os termos semelhantes, tem-se:

$$(C - ZW^{-1}Z')\underset{\sim}{\hat{a}} = \underset{\sim}{Q} - ZW^{-1}\underset{\sim}{S}$$

Fazendo,

$$C - ZW^{-1}Z' = C^*$$

e,
$$\underset{\sim}{Q} - ZW^{-1}\underset{\sim}{S} = \underset{\sim}{Q}^*$$

o sistema reduzido torna-se,

$$C^*\underset{\sim}{\hat{a}} = \underset{\sim}{Q}^* \tag{13}$$

As matrizes envolvidas nesse sistema têm a composição dada a seguir:

$$C = {}_v[c_{ii'}]_v$$

onde, $c_{ii} = n_{i.} - \sum_j n_{ij}^2/n_{.j}$, para $i = i'$

e $c_{ii'} = - \sum_j n_{ij}n_{i'j}/n_{.j}$, para $i \neq i'$.

$$Z = \begin{bmatrix} x_{1..}^{(1)} - \sum_j \frac{n_{1j}x_{.j.}^{(1)}}{n_{.j}} & x_{1..}^{(2)} - \sum_j \frac{n_{1j}x_{.j.}^{(2)}}{n_{.j}} & \cdots & x_{1..}^{(p)} - \sum_j \frac{n_{1j}x_{.j.}^{(p)}}{n_{.j}} \\ x_{2..}^{(1)} - \sum_j \frac{n_{2j}x_{.j.}^{(1)}}{n_{.j}} & x_{2..}^{(2)} - \sum_j \frac{n_{2j}x_{.j.}^{(2)}}{n_{.j}} & \cdots & x_{2..}^{(p)} - \sum_j \frac{n_{2j}x_{.j.}^{(p)}}{n_{.j}} \\ \cdots & \cdots & \cdots & \cdots \\ x_{v..}^{(1)} - \sum_j \frac{n_{vj}x_{.j.}^{(1)}}{n_{.j}} & x_{v..}^{(2)} - \sum_j \frac{n_{vj}x_{.j.}^{(2)}}{n_{.j}} & \cdots & x_{v..}^{(p)} - \sum_j \frac{n_{vj}x_{.j.}^{(p)}}{n_{.j}} \end{bmatrix}$$

$$W = \begin{bmatrix} \sum\limits_{ijk}\sum\sum x^{(1)}_{ijk}x^{(1)}_{ijk} - \sum\limits_{j} \frac{x^{(1)}_{.j.}x^{(1)}_{.j.}}{n_{.j}} & \sum\limits_{ijk}\sum\sum x^{(1)}_{ijk}x^{(2)}_{ijk} - \sum\limits_{j} \frac{x^{(1)}_{.j.}x^{(2)}_{.j.}}{n_{.j}} & \cdots & \sum\limits_{ijk}\sum\sum x^{(1)}_{ijk}x^{(p)}_{ijk} - \sum\limits_{j} \frac{x^{(1)}_{.j.}x^{(p)}_{.j.}}{n_{.j}} \\ \sum\limits_{ijk}\sum\sum x^{(2)}_{ijk}x^{(1)}_{ijk} - \sum\limits_{j} \frac{x^{(2)}_{.j.}x^{(1)}_{.j.}}{n_{.j}} & \sum\limits_{ijk}\sum\sum x^{(2)}_{ijk}x^{(2)}_{ijk} - \sum\limits_{j} \frac{x^{(2)}_{.j.}x^{(2)}_{.j.}}{n_{.j}} & \cdots & \sum\limits_{ijk}\sum\sum x^{(2)}_{ijk}x^{(p)}_{ijk} - \sum\limits_{j} \frac{x^{(2)}_{.j.}x^{(p)}_{.j.}}{n_{.j}} \\ \cdots & \cdots & \cdots & \cdots \\ \sum\limits_{ijk}\sum\sum x^{(p)}_{ijk}x^{(1)}_{ijk} - \sum\limits_{j} \frac{x^{(p)}_{.j.}x^{(1)}_{.j.}}{n_{.j}} & \sum\limits_{ijk}\sum\sum x^{(p)}_{ijk}x^{(2)}_{ijk} - \sum\limits_{j} \frac{x^{(p)}_{.j.}x^{(2)}_{.j.}}{n_{.j}} & \cdots & \sum\limits_{ijk}\sum\sum x^{(p)}_{ijk}x^{(p)}_{ijk} - \sum\limits_{j} \frac{x^{(p)}_{.j.}x^{(p)}_{.j.}}{n_{.j}} \end{bmatrix}$$

$$\underset{\sim}{Q} = \begin{bmatrix} y_{1..} - \sum_j \frac{n_{1j} y_{.j.}}{n_{.j}} \\ y_{2..} - \sum_j \frac{n_{2j} y_{.j.}}{n_{.j}} \\ \cdots \\ y_{v..} - \sum_j \frac{n_{vj} y_{.j.}}{n_{.j}} \end{bmatrix}$$

$$\underset{\sim}{S} = \begin{bmatrix} \sum\sum\sum_{ijk} x^{(1)}_{ijk} y_{ijk} - \sum_j \frac{x^{(1)}_{.j.} y_{.j.}}{n_{.j}} \\ \sum\sum\sum_{ijk} x^{(2)}_{ijk} y_{ijk} - \sum_j \frac{x^{(2)}_{.j.} y_{.j.}}{n_{.j}} \\ \cdots \\ \sum\sum\sum_{ijk} x^{(p)}_{ijk} y_{ijk} - \sum_j \frac{x^{(p)}_{.j.} y_{.j.}}{n_{.j}} \end{bmatrix}$$

Pela composição dessas matrizes pode-se notar perfeitamente que os vetores da matriz Z têm a mesma forma do vetor $\underset{\sim}{Q}$, com a diferença de que o último refere-se à variável dependente e os primeiros às variáveis independentes. Dessa forma, vê-se facilmente que:

$$\underset{\sim}{j}'C = \underset{\sim}{0}'$$

$$\underset{\sim}{j}'Z = \underset{\sim}{0}'$$

$$\underset{\sim}{j}'\underset{\sim}{Q} = 0$$

De modo que (10) é um sistema consistente e indeterminado, já que $j'C = 0'$ e $j'(\underset{\sim}{Q} - Z\hat{\underset{\sim}{c}}) = 0$.

Adotando, dentre os muitos existentes, o procedimento proposto por PIMENTEL GOMES (1968), seja $\bar{A}$ tal que:

a) $\rho(\bar{A}) = 1$, onde $\rho(\bar{A})$ é o posto ou característica de $\bar{A}$.

b) $\bar{A}\hat{\underset{\sim}{a}} = \underset{\sim}{0}$.

c) $\bar{A}\underset{\sim}{a}$ não pode ser combinação linear estimável de parâmetros.

Assim sendo,

$$C\hat{\underset{\sim}{a}} = \underset{\sim}{Q} - Z\hat{\underset{\sim}{c}}$$

$$\bar{A}\hat{\underset{\sim}{a}} = \underset{\sim}{0}.$$

Efetuando-se a subtração, tem-se:

$$(C - \bar{A})\hat{\underset{\sim}{a}} = \underset{\sim}{Q} - Z\hat{\underset{\sim}{c}}$$

Fazendo $C - \bar{A} = M$, onde M é suposta não singular, uma das infinitas soluções que satisfazem (10) é:

$$\hat{\underset{\sim}{a}} = M^{-1}\underset{\sim}{Q} - M^{-1}Z\hat{\underset{\sim}{c}}$$

Substituindo em (11), fica:

$$(W - Z'M^{-1}Z)\hat{\underset{\sim}{c}} = \underset{\sim}{S} - Z'M^{-1}\underset{\sim}{Q}$$

Fazendo,

$$R_1 = W - Z'M^{-1}Z$$

e $$\underset{\sim}{R}_2 = \underset{\sim}{S} - Z'M^{-1}\underset{\sim}{Q}$$

o sistema fica:

$$R_1 \hat{\underset{\sim}{c}} = \underset{\sim}{R}_2 .$$

Ora, facilmente se verifica que $Z'M^{-1}Z$ é a matriz das somas de quadrados e produtos relativas à primeira classificação ou fator A, ajustadas para a segunda, consideradas as variáveis auxiliares. De forma que $R_1 = E - D'K^{-1}D - Z'M^{-1}Z$ é a matriz dos resíduos das somas de quadrados e produtos, suposta não singular, relativas às variáveis auxiliares, enquanto que, $\underset{\sim}{R}_2 = \underset{\sim}{P} - D'K^{-1}\underset{\sim}{B} - Z'M^{-1}\underset{\sim}{Q}$ é o vetor dos resíduos das somas de produtos, consideradas as variáveis independentes e a variável dependente.

Adaptando o procedimento utilizado por PIMENTEL GOMES (1968), os sistemas (13) e (14) podem ser mais facilmente obtidos definindo:

$$V = \begin{bmatrix} V_1 \\ -- \\ V_2 \end{bmatrix} = \left[\begin{array}{ccc} I_v & (ZW^{-1}D'K^{-1} - NK^{-1}) & -ZW^{-1} \\ \hline -Z'M^{-1} & (Z'M^{-1}NK^{-1} - D'K^{-1}) & I_p \end{array}\right]$$

Pré-multiplicando o sistema de equações normais $X'X\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{y}$ por V, obtém-se:

$$VX'X\hat{\underset{\sim}{\beta}} = VX'\underset{\sim}{y}$$

donde se comprova que:

$$\begin{bmatrix} V_1X'X \\ \\ V_2X'X \end{bmatrix} = \begin{bmatrix} C^* & \Phi & \Phi \\ \\ \Phi & \Phi & R_1 \end{bmatrix} \qquad (15)$$

sendo Φ uma matriz nula e C^* e R_1 as matrizes definidas em (13) e (14), respectivamente.

Da mesma forma:

$$\begin{bmatrix} V_1 X'\underset{\sim}{y} \\ V_2 X'\underset{\sim}{y} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{Q}^* \\ \underset{\sim}{R}_2 \end{bmatrix} \tag{16}$$

Assim, através de:

$$V_1 X'X\hat{\underset{\sim}{\beta}} = V_1 X'\underset{\sim}{y}$$

obtém-se o sistema $C^*\hat{\underset{\sim}{a}} = \underset{\sim}{Q}^*$, e de,

$$V_2 X'X\hat{\underset{\sim}{\beta}} = V_2 X'\underset{\sim}{y}$$

o sistema $R_1\hat{\underset{\sim}{c}} = \underset{\sim}{R}_2$.

Como R_1 é suposta não singular, o melhor estimador linear imparcial de $\underset{\sim}{c}$ é:

$$\hat{\underset{\sim}{c}} = R_1^{-1}\underset{\sim}{R}_2 \tag{17}$$

Já foi visto que $\underset{\sim}{j}'C^* = \underset{\sim}{0}'$, logo $\rho(C^*) < v$.Assim sendo, $\underset{\sim}{a}$ só poderá ser estimado se for imposta uma restrição às estimativas $\hat{a}_i$. Logicamente, existem infinitas restrições mas, dentre elas, será escolhida a restrição $\underset{\sim}{j}'\hat{\underset{\sim}{a}} = \sum_i \hat{a}_i = 0$.

Assim, seja A^*, uma matriz de propriedades semelhantes à $\bar{A}$, construída de maneira que a restrição $\underset{\sim}{j}'\hat{\underset{\sim}{a}} = 0$ seja satisfeita, ou seja,

$$A^*\hat{\underset{\sim}{a}} = 0 .$$

Isto dito, uma das muitas soluções que satisfazem (13) é:

$$\underset{\sim}{\hat{a}} = M^{*^{-1}} \underset{\sim}{Q}^* \qquad (18)$$

sendo $M^* = C^* - A^*$.

3.3. Matriz de dispersão dos estimadores

3.3.1. Matriz de dispersão de $\underset{\sim}{\hat{a}}$

Por definição:

$$D(\underset{\sim}{\hat{a}}) = E\{[\underset{\sim}{\hat{a}} - E(\underset{\sim}{\hat{a}})][\underset{\sim}{\hat{a}} - E(\underset{\sim}{\hat{a}})]'\}$$

Viu-se, em (18), que:

$$\underset{\sim}{\hat{a}} = M^{*^{-1}} \underset{\sim}{Q}^*$$

Como $\underset{\sim}{Q}^* = V_1 X'\underset{\sim}{y}$, fica:

$$\underset{\sim}{\hat{a}} = M^{*^{-1}} V_1 X'\underset{\sim}{y}$$

ou $$\underset{\sim}{\hat{a}} = M^{*^{-1}} V_1 X'(X\underset{\sim}{\beta} + \underset{\sim}{e})$$

donde $$\underset{\sim}{\hat{a}} = M^{*^{-1}} V_1 X'X\underset{\sim}{\beta} + M^{*^{-1}} V_1 X'\underset{\sim}{e}$$

Logo,

$$E(\underset{\sim}{\hat{a}}) = M^{*^{-1}} V_1 X'X\underset{\sim}{\beta}$$

ou ainda, por (15)

$$E(\underset{\sim}{\hat{a}}) = M^{*^{-1}} C^*\underset{\sim}{a}$$

De forma que,

$$\hat{\underset{\sim}{a}} - E(\hat{\underset{\sim}{a}}) = M^{*^{-1}} V_1 X' \underset{\sim}{e}$$

e $$[\hat{\underset{\sim}{a}} - E(\hat{\underset{\sim}{a}})][\hat{\underset{\sim}{a}} - E(\hat{\underset{\sim}{a}})]' = M^{*^{-1}} V_1 X' \underset{\sim}{e}\underset{\sim}{e}' X V_1 M^{*^{-1}}{}'$$

Pela definição:

$$D(\hat{\underset{\sim}{a}}) = E[M^{*^{-1}} V_1 X' \underset{\sim}{e}\underset{\sim}{e}' X V_1' M^{*^{-1}}{}']$$

resultando

$$D(\hat{a}) = M^{*^{-1}} C^* M^{*^{-1}}{}' \sigma^2 \tag{19}$$

uma vez obedecidas as pressuposições sobre os erros e sabendo que $V_1 X' X V_1' = C^*$.

3.3.2. Matriz de dispersão de $\hat{\underset{\sim}{c}}$

Por (17),

$$\hat{\underset{\sim}{c}} = R_1^{-1} \underset{\sim}{R}_2$$

ou $$\hat{\underset{\sim}{c}} = R_1^{-1} V_2 X' \underset{\sim}{y}$$

ou, ainda, $$\hat{\underset{\sim}{c}} = R_1^{-1} R_1 \underset{\sim}{c} + R_1^{-1} V_2 X' \underset{\sim}{e}$$

resultando que,

$$E(\hat{\underset{\sim}{c}}) = \underset{\sim}{c}$$

À semelhança de $\hat{\underset{\sim}{a}}$,

$$D(\hat{\underset{\sim}{c}}) = E\{[\hat{\underset{\sim}{c}} - E(\hat{\underset{\sim}{c}})][\hat{\underset{\sim}{c}} - E(\hat{\underset{\sim}{c}})]'\}$$

$$D(\underset{\sim}{\hat{c}}) = E[R_1^{-1}V_2X'\underset{\sim}{e}\underset{\sim}{e}'XV_2'R_1^{-1}]$$

pois $\underset{\sim}{\hat{c}} - E(\underset{\sim}{\hat{c}}) = R_1^{-1}V_2X'\underset{\sim}{e}$.

Como $V_2X'XV_2' = \underset{\sim}{R}_1$ e $E(\underset{\sim}{e}\underset{\sim}{e}') = \sigma^2 I$, conclui-se que:

$$D(\underset{\sim}{\hat{c}}) = R_1^{-1}\sigma^2$$

3.4. Distribuição dos estimadores

Considerando que, por hipótese, $\underset{\sim}{y}\ \Omega\ N(X\underset{\sim}{\beta}, \sigma^2 I)$ e que:

$$\underset{\sim}{\hat{a}} = M^{*-1}\underset{\sim}{Q}^* = M^{*-1}V_1X'\underset{\sim}{y}$$

vê-se que $\hat{a}$ é uma função linear dos y, donde resulta que:

$$\underset{\sim}{\hat{a}}\ \Omega\ N(M^{*-1}C^*\underset{\sim}{a},\ M^{*-1}C^*M^{*-1}{}'\sigma^2)\ .$$

Do mesmo modo:

$$\underset{\sim}{\hat{c}} = R_1^{-1}\underset{\sim}{R}_2 = R_1^{-1}V_2X'\underset{\sim}{y}\ \Omega\ N(\underset{\sim}{c},\ R_1^{-1}\sigma^2)\ .$$

3.5. Médias ajustadas dos níveis dos fatores

Admitiu-se, inicialmente, sem perda de generalidade, que a segunda classificação ou fator B, tinha um número maior de níveis e, isto considerado, as equações absorvidas foram:

$$\underset{\sim}{j}\hat{\mu} + \underset{\sim}{\hat{b}} = K^{-1}[\underset{\sim}{B} - N'\underset{\sim}{\hat{a}} - D\underset{\sim}{\hat{c}}] \tag{21}$$

que nada mais são que as equações das médias dos níveis de B ajustadas para os demais efeitos.

Pré-multiplicando ambos os membros por $\underset{\sim}{j}'$, tem-se:

$$\underset{\sim}{j}'\underset{\sim}{j}\hat{\mu} + \underset{\sim}{j}'\hat{\underset{\sim}{b}} = \underset{\sim}{j}'K^{-1}[\underset{\sim}{B} - N'\hat{\underset{\sim}{a}} - D\hat{\underset{\sim}{c}}]$$

Considerada a restrição $\underset{\sim}{j}'\hat{\underset{\sim}{b}} = \sum_j \hat{b}_j = 0$, um estimador para μ é obtido facilmente, ou seja:

$$\hat{\mu} = \frac{1}{r}\underset{\sim}{j}'K^{-1}[\underset{\sim}{B} - N'\hat{\underset{\sim}{a}} - D\hat{\underset{\sim}{c}}] \tag{22}$$

Obtidos $\hat{\mu}$, $\hat{a}$ e $\hat{c}$, $\hat{b}$ é imediatamente obtido através de:

$$\hat{\underset{\sim}{b}} = K^{-1}[\underset{\sim}{B} - N'\hat{\underset{\sim}{a}} - D\hat{\underset{\sim}{c}}] - \underset{\sim}{j}\hat{\mu} \tag{23}$$

De forma idêntica, são determinadas as médias dos níveis de A ajustadas para o fator B e para a regressão, ou seja:

$$\underset{\sim}{j}\hat{\mu} + \hat{\underset{\sim}{a}} = R^{-1}[\underset{\sim}{T} - N\hat{\underset{\sim}{b}} - A\hat{\underset{\sim}{c}}] \tag{24}$$

3.6. Somas de quadrados

3.6.1. Soma de quadrados do resíduo

A soma de quadrados do resíduo ajustada para a regressão, considerando o modelo (5), é dada por:

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{\beta}}'X'\underset{\sim}{y}$$

ou

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{a}}'\underset{\sim}{T} - \hat{\underset{\sim}{b}}'\underset{\sim}{B} - \hat{\underset{\sim}{c}}'\underset{\sim}{P}$$

Considerando que, por (8),

$$\hat{\underset{\sim}{b}}' = \underset{\sim}{B}'K^{-1} - \hat{\underset{\sim}{c}}'D'K^{-1} - \hat{\underset{\sim}{a}}'NK^{-1}$$

resulta que:

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{a}}'(\underset{\sim}{T} - NK^{-1}\underset{\sim}{B}) - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - \hat{\underset{\sim}{c}}'(\underset{\sim}{P} - D'K^{-1}\underset{\sim}{B})$$

ou, ainda,

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{a}}'\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - \hat{\underset{\sim}{c}}'\underset{\sim}{S}$$

Mas, considerando que:

$$\hat{\underset{\sim}{a}}' = \underset{\sim}{Q}'M^{-1'} - \hat{\underset{\sim}{c}}'Z'M^{-1'}$$

e, pós-multiplicando por $\underset{\sim}{Q}$, fica:

$$\hat{\underset{\sim}{a}}'\underset{\sim}{Q} = \underset{\sim}{Q}'M^{-1'}\underset{\sim}{Q} - \hat{\underset{\sim}{c}}'Z'M^{-1'}\underset{\sim}{Q}$$

ou $$\hat{\underset{\sim}{a}}'\underset{\sim}{Q} = \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{Q}'M^{-1}Z\hat{\underset{\sim}{c}}$$

Substituindo em SQR*,

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - (\underset{\sim}{S}' - \underset{\sim}{Q}'M^{-1}Z)\hat{\underset{\sim}{c}}$$

sendo o mesmo que:

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2 \qquad (25)$$

Caso se queira, pode-se introduzir a correção para a média, fazendo:

$$SQR^* = (\underset{\sim}{y}'\underset{\sim}{y} - C) - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - (\underset{\sim}{B}'K^{-1}\underset{\sim}{B} - C) - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2 \qquad (26)$$

onde: $C = (\underset{\sim}{j}'\underset{\sim}{j})^{-1}(\underset{\sim}{j}'\underset{\sim}{y})^2 = \frac{1}{n_{..}}(\sum\sum\sum_{ijk} y_{ijk})^2$;

$\underset{\sim}{y}'\underset{\sim}{y} - C$ é a soma de quadrados total;

$\underset{\sim}{Q}'M^{-1}\underset{\sim}{Q}$ é a soma de quadrados para A, ajustada para B, ignorando as covariáveis;

$\underset{\sim}{B}'K^{-1}\underset{\sim}{B} - C$ é a soma de quadrados para B, ignorando A e a regressão; e

$\hat{\underset{\sim}{c}}'\underset{\sim}{R}_2$ é a soma de quadrados da regressão.

3.6.2. Soma de quadrados para o teste da hipótese H_1: $a_i - a_{i'} = 0$, para todo $i \neq i'$

A soma de quadrados referente ao numerador do teste F apropriado para o teste da hipótese H_1: $a_i - a_{i'}=0$, para todo $i \neq i'$, ou, de forma equivalente, H_1: $K'\underset{\sim}{a} = \underset{\sim}{0}$, é:

$$SQH_1 = (K'\hat{\underset{\sim}{a}})'(K'M^{*-1}K)^{-1}K'\hat{\underset{\sim}{a}} \qquad (27)$$

onde K' é uma matriz de posto linha completo. Considerando a hipótese de igualdade entre os níveis de A, a matriz K' pode ter a forma:

$$K' = [I_{v-1} \quad -\underset{\sim}{j}]$$

Assim, $K'\underset{\sim}{a} = \underset{\sim}{0}$, expressa um conjunto de afirmações acerca dos níveis de A cuja validade se deseja testar. No caso presente, essas afirmações são:

$$H_1: a_i - a_v = 0, \text{ para } i = 1, 2, \ldots, v-1 \qquad (28)$$

Uma forma equivalente a (28), na notação matricial, é:

$$H_1: \underset{\sim}{a} - \underset{\sim}{j}a = \underset{\sim}{0}$$

Substituindo em (27) a expressão de $\hat{\underset{\sim}{a}}$, dada em (18), obtém-se:

$$SQH_1 = \underset{\sim}{Q}^{*\prime}M^{*^{-1}\prime}K(K'M^{*^{-1}}K)^{-1}K'M^{*^{-1}}\underset{\sim}{Q}^{*}$$

Partindo da igualdade,

$$M^{*^{-1}\prime}K(K'M^{*^{-1}}K)^{-1}K'M^{*^{-1}} = M^{*^{-1}\prime}$$

e, pós-multiplicando ambos os membros por K, resulta:

$$M^{*^{-1}\prime}K(K'M^{*^{-1}}K)^{-1}K'M^{*^{-1}}K = M^{*^{-1}\prime}K$$

donde: $$M^{*^{-1}\prime}K = M^{*^{-1}\prime}K$$

ou seja, a igualdade que representou o ponto de partida é verdadeira.

Portanto, (27) pode ser reescrita como:

$$SQH_1 = \underset{\sim}{Q}^{*\prime}M^{*^{-1}\prime}\underset{\sim}{Q}^{*}$$

ou $$SQH_1 = \underset{\sim}{Q}^{*\prime}M^{*^{-1}}\underset{\sim}{Q}^{*}$$

ou, ainda, $$SQH_1 = \hat{\underset{\sim}{a}}'\underset{\sim}{Q}^{*} \qquad (29)$$

3.6.3. Soma de quadrados para o teste da hipótese H_2: $\underset{\sim}{c} = \underset{\sim}{0}$

Neste caso, como R_1 é não singular, o vetor $\underset{\sim}{c}$ é estimável, de maneira que H_2 pode ser expressa como $H_2: K'\underset{\sim}{c} = \underset{\sim}{0}$, onde $K' = I_p$.

De maneira análoga a (27), facilmente se verifica que:

$$SQH_2 = \hat{\underset{\sim}{c}}'R_1\hat{\underset{\sim}{c}}$$

ou $$SQH_2 = \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2 \qquad (30)$$

Pode-se notar aqui, uma perfeita identidade entre a fórmula (30) e aquela obtida na determinação de SQR*.

De maneira que a soma de quadrados obtida em (30) será, algumas vezes, denominada de soma de quadrados para a regressão e a obtida em (29) de soma de quadrados para o fator A ajustado para B e para regressão.

3.6.4. Somas de quadrados para o teste das sub-hipóteses H_i: $\underline{k}_i'\underline{a} = 0$

Viu-se em 3.6.2. que a soma de quadrados para o teste da hipótese H_1: $K'\underline{a} = 0$ é $SQH_1 = \hat{\underline{a}}'\underline{Q}^*$. No entanto, pode ser desejável decompor H_1 em sub-hipóteses independentes do tipo H_i: $\underline{k}_i'\underline{a} = 0$, $i = 1, 2, \ldots, \rho(K')$.

Sejam q_i a i-ésima e q_j a j-ésima parte de $\hat{\underline{a}}'\underline{Q}^*$. Por (27), tem-se que:

$$SQH_1 = \hat{\underline{a}}'K(K'M^{*-1}K)^{-1}K'\hat{\underline{a}}$$

Analogamente,

$$SQH_i = \hat{\underline{a}}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'\hat{\underline{a}} \qquad (31)$$

e
$$SQH_j = \hat{\underline{a}}'\underline{k}_j(\underline{k}_j'M^{*-1}\underline{k}_j)^{-1}\underline{k}_j'\hat{\underline{a}} \qquad (32)$$

Considerando (16) e (18) as duas somas de quadrados ficam:

$$SQH_i = \underline{y}'XV_1'M^{*-1}{}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'M^{*-1}V_1X'\underline{y}$$

e
$$SQH_j = \underline{y}'XV_1'M^{*-1}{}'\underline{k}_j(\underline{k}_j'M^{*-1}\underline{k}_j)^{-1}\underline{k}_j'M^{*-1}V_1X'\underline{y}$$

Assim, (31) e (32) são somas de quadrados apropriadas para o teste das sub-hipóteses H_i: $\underline{k}_i'\underline{a} = 0$ e H_j: $\underline{k}_j'\underline{a} = 0$.

Essas formas quadráticas são independentes, suposto que $\underline{y}\ \Omega\ N(X\underline{\beta},\sigma^2 I)$, se:

$$XV_1'M^{*^{-1}}{}'\underline{k}_i(\underline{k}_i'M^{*^{-1}}\underline{k}_i)^{-1}\underline{k}_i'M^{*^{-1}}V_1X'XV_1'M^{*^{-1}}{}'\underline{k}_j(\underline{k}_j'M^{*^{-1}}\underline{k}_j)^{-1}\underline{k}_j'M^{*^{-1}}V_1X'=\Phi$$

Para tanto, é suficiente que:

$$\underline{k}_i'M^{*^{-1}}V_1X'XV_1'M^{*^{-1}}{}'\underline{k}_j = 0$$

Mas, por (15), $V_1X'XV_1' = C^*$, o que resulta,

$$\underline{k}_i'M^{*^{-1}}C^*M^{*^{-1}}{}'\underline{k}_j = 0$$

Ver-se-á mais adiante, que se $\underline{k}_j'\underline{a}$ é estimável, então $\underline{k}_j' = \underline{\lambda}'C^*$, de maneira que:

$$\underline{k}_i'M^{*^{-1}}C^*M^{*^{-1}}{}'C^*\underline{\lambda} = 0.$$

Ora, $M^{*^{-1}}$ é uma inversa generalizada de C^*, PIMENTEL GOMES (1968), logo $C^*M^{*^{-1}}C^* = C^*$, donde:

$$\underline{k}_i'M^{*^{-1}}C^*\underline{\lambda} = 0\ .$$

Finalmente a condição se reduz a,

$$\underline{k}_i'M^{*^{-1}}\underline{k}_j = 0.$$

Nos casos onde $M^* = C^* + A^*$ é uma matriz diagonal, (33) se reduz a:

$$\frac{1}{m}\underline{k}_i'\underline{k}_j = 0$$

onde $\underline{m}$ é uma constante. Ou seja, para que (33) ocorra é necessário e suficiente que $\underline{k}_i'\underline{k}_j = 0$.

Por esse motivo, $\underset{\sim}{k}_i'\underset{\sim}{a}$ e $\underset{\sim}{k}_j'\underset{\sim}{a}$ são, frequentemente, denominados contrastes ortogonais. A condição exigida, no entanto, é (33).

Se (33) se verifica, então,

$$(K'M^{*^{-1}}K)^{-1} = \text{diag}\{(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\},\ i = 1, 2, \ldots, v-1 .$$

Substituindo em (27), fica:

$$\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = (K'\hat{\underset{\sim}{a}})'\text{diag}\{(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\}K'\hat{\underset{\sim}{a}}$$

ou $$\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = \sum_{i=1}^{v-1} \frac{(\underset{\sim}{k}_i'\hat{a})^2}{\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i}$$

ou, ainda, $$\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = \sum_{i=1}^{v-1} q_i$$

3.7. Variância de uma combinação linear de estimativas

A variância de uma combinação linear de estimativas $\underset{\sim}{k}'\hat{\underset{\sim}{a}}$ é dada por:

$$V(\underset{\sim}{k}'\hat{\underset{\sim}{a}}) = \underset{\sim}{k}'V(\hat{\underset{\sim}{a}})\underset{\sim}{k}$$

ou, ainda, $$V(\underset{\sim}{k}'\hat{\underset{\sim}{a}}) = \underset{\sim}{k}'D(\hat{\underset{\sim}{a}})\underset{\sim}{k}$$

onde $D(\hat{\underset{\sim}{a}})$ é obtida através de (19).

Se $\underset{\sim}{k}'$ for tal que $\underset{\sim}{k}'\hat{\underset{\sim}{a}} = \hat{a}_i - \hat{a}_{i'}$, então:

$$V(\underset{\sim}{k}'\hat{\underset{\sim}{a}}) = (d_{ii} + d_{i'i'} - 2d_{i'i})\sigma^2 \qquad (34)$$

Entretanto, como $D(\hat{\underset{\sim}{a}}) = M^{*^{-1}}C^*M^{*^{-1}}{}'\sigma^2$, verifica-se que:

$$V(\underset{\sim}{k}'\hat{\underset{\sim}{a}}) = \underset{\sim}{k}'M^{*^{-1}}\underset{\sim}{k}\sigma^2$$

3.8. Esperança matemática das somas de quadrados

As esperanças das somas de quadrados serão tomadas considerando que β, em (5), é um vetor de efeitos fixos e que e é um vetor de erros aleatórios, independentes, de média zero e homocedásticos.

3.8.1. Esperança matemática da soma de quadrados do resíduo ajustada para a regressão

Viu-se, anteriormente, que:

$$SQR^* = y'y - \hat{\beta}'X'y .$$

Seja $\hat{\beta} = (X'X)^- X'y$, onde $(X'X)^-$ é uma inversa generalizada de $X'X$. Assim,

$$SQR^* = y'y - y'X(X'X)^- X'y$$

ou

$$SQR^* = y'[I - X(X'X)^- X']y \tag{35}$$

Segundo SEARLE (1971), a esperança matemática de uma forma quadrática $y'Qy$ é:

$$E(y'Qy) = E(y')QE(y) + tr(QV) \tag{36}$$

sendo Q o núcleo da forma quadrática, $V = V(y)$ e tr indica o operador traço de uma matriz. Se for considerado, em (5), que $E(y) = X\beta$ e $V = \sigma^2 I$, a expressão dada por (36) fica:

$$E(y'Qy) = \beta'X'QX\beta + \sigma^2 tr(Q) \tag{37}$$

Dois casos particulares importantes se verificam quando $Q = I_{n..}$ e $Q = X(X'X)^- X'$. Nestes, tem-se:

$$E(\underset{\sim}{y}'\underset{\sim}{y}) = \underset{\sim}{\beta}'X'X\underset{\sim}{\beta} + \sigma^2 tr(I_{n_{..}}) \tag{38}$$

e $$E[\underset{\sim}{y}'X(X'X)^{-}X'\underset{\sim}{y}] = \underset{\sim}{\beta}'X'X(X'X)^{-}X'X\underset{\sim}{\beta} + \sigma^2 tr[X(X'X)^{-}X'] =$$

$$= \underset{\sim}{\beta}'X'X\underset{\sim}{\beta} + \sigma^2\rho(X) \tag{39}$$

já que $$tr[X(X'X)^{-}X'] = \rho[X(X'X)^{-}X'] = \rho(X)$$

por ser $X(X'X)^{-}X'$ uma matriz idempotente.

No caso particular do modelo (3), obtém-se, então:

$$E(SQR^*) = [n_{..} - \rho(X)]\sigma^2.$$

Viu-se no item 3.2. deste trabalho que, partindo do sistema de equações normais $X'X\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{y}$, pode-se determinar sistemas reduzidos perfeitamente equivalentes. Isso significa que:

$$\begin{bmatrix} n_{..} & \underset{\sim}{r}' & \underset{\sim}{k}' & \underset{\sim}{0}' \\ \underset{\sim}{r} & R & N & A \\ \underset{\sim}{k} & N' & K & D \\ \underset{\sim}{0} & A' & D' & E \end{bmatrix} \sim \begin{bmatrix} 0 & \underset{\sim}{0}' & \underset{\sim}{0}' & \underset{\sim}{0}' \\ \underset{\sim}{0} & C^* & \Phi & \Phi \\ \underset{\sim}{k} & N' & K & D \\ \underset{\sim}{0} & \Phi & \Phi & R_1 \end{bmatrix}$$

Assim, vê-se perfeitamente que:

$$\rho(X) = \rho(X'X) = p + r + \rho(C^*) .$$

Adaptando para este caso as conclusões de GRAYBILL (1961) e, sabendo que todas as funções lineares dos c's são estimáveis, se todos os contrastes entre os níveis de A forem estimáveis, então $\rho(C^*) = v - 1$, donde se conclui que:

$$\rho(X) = p + r + (v - 1) \tag{40}$$

Logo,

$$E(SQR^*) = (n_{..} - p - r - v + 1)\sigma^2 \tag{41}$$

De maneira que, um estimador imparcial de σ^2, é:

$$QMR^* = \frac{1}{n_{..} - \rho(X)} SQR^* .$$

3.8.2. Esperança matemática da soma de quadrados para A eliminando o fator B e a regressão.

O passo inicial é reescrever $\underset{\sim}{\hat{a}}'\underset{\sim}{Q}^*$ como uma forma quadrática em $\underset{\sim}{y}$. Por (19), tem-se que:

$$\underset{\sim}{\hat{a}} = M^{*^{-1}}\underset{\sim}{Q}^*$$

mas, por (16), $\underset{\sim}{Q}^* = V_1X'\underset{\sim}{y}$, donde:

$$\underset{\sim}{\hat{a}}'\underset{\sim}{Q}^* = \underset{\sim}{Q}^{*\prime}M^{*^{-1}}\underset{\sim}{Q}^* = \underset{\sim}{y}'XV_1'M^{*^{-1}}V_1X'\underset{\sim}{y}$$

De acordo com (37),

$$E(\underset{\sim}{\hat{a}}'\underset{\sim}{Q}^*) = \underset{\sim}{\beta}'X'XV_1'M^{*^{-1}}V_1X'X\underset{\sim}{\beta} + \sigma^2 tr(XV_1'M^{*^{-1}}V_1X') .$$

Considerando que,

$$V_1X'X\underset{\sim}{\beta} = C^*\underset{\sim}{a}$$

e que,

$$\begin{aligned} tr(XV_1'M^{*^{-1}}V_1X') &= tr(M^{*^{-1}}V_1X'XV_1') = \\ &= tr(M^{*^{-1}}C^*) = \\ &= \rho(M^{*^{-1}}C^*) = \rho(C^*) \end{aligned}$$

já que $M^{*^{-1}}C^*$ é idempotente e $\rho(C^*) < \rho(M^{*^{-1}})$.

Resulta que:

$$E(\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*) = \underset{\sim}{a}'C^*\underset{\sim}{a} + \sigma^2\rho(C^*) .$$

Caso (40) se verifique,

$$E(\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*) = \underset{\sim}{a}'C^*\underset{\sim}{a} + (v - 1)\sigma^2 \qquad (42)$$

3.8.3. Esperança matemática de SQH_i

Por (31), tem-se que:

$$SQH_i = \hat{\underset{\sim}{a}}'\underset{\sim}{k}_i(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'\hat{\underset{\sim}{a}}$$

ou, ainda,

$$SQH_i = \underset{\sim}{y}'XV_1'M^{*^{-1}}{}'\underset{\sim}{k}_i(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}V_1X'\underset{\sim}{y} = \underset{\sim}{y}'A_i\underset{\sim}{y}$$

e, por (37),

$$E(SQH_i) = \underset{\sim}{\beta}'X'XV_1'M^{*^{-1}}{}'\underset{\sim}{k}_i(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}V_1X'X\underset{\sim}{\beta} +$$

$$+ \sigma^2 tr[XV_1'M^{*^{-1}}\underset{\sim}{k}_i(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}V_1X']$$

mas, $V_1X'X\underset{\sim}{\beta} = C^*\underset{\sim}{a}$ e,

$$tr(A_i) = tr[(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}V_1X'XV_1'M^{*^{-1}}{}'\underset{\sim}{k}_i] =$$

$$= tr[(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}C^*M^{*^{-1}}{}'\underset{\sim}{k}_i]$$

donde,

$$E(SQH_i) = \frac{1}{\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i}\underset{\sim}{a}'C^*M^{*^{-1}}{}'\underset{\sim}{k}_i\underset{\sim}{k}_i'M^{*^{-1}}C^*\underset{\sim}{a} +$$

$$+ \sigma^2 tr[(\underset{\sim}{k}_i'M^{*^{-1}}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*^{-1}}C^*M^{*^{-1}}{}'\underset{\sim}{k}_i]$$

Mas,

$$\underset{\sim}{k}_i' M^{*^{-1}} C^* \underset{\sim}{a} = \underset{\sim}{k}_i' E(\hat{\underset{\sim}{a}}) = E(\underset{\sim}{k}_i' \hat{\underset{\sim}{a}}) = \underset{\sim}{k}_i' \underset{\sim}{a}$$

uma vez que $\underset{\sim}{k}_i' \underset{\sim}{a}$ é estimável e,

$$\underset{\sim}{k}_i' M^{*^{-1}} C^* M^{*^{-1}}{}' \underset{\sim}{k}_i = \underset{\sim}{k}_i' M^{*^{-1}} \underset{\sim}{k}_i$$

resulta,

$$E(SQH_i) = \frac{1}{\underset{\sim}{k}_i' M^{*^{-1}} \underset{\sim}{k}_i} (\underset{\sim}{k}_i' \underset{\sim}{a})' (\underset{\sim}{k}_i' \underset{\sim}{a}) + \sigma^2 tr(I_1)$$

ou, ainda,

$$E(SQH_i) = \frac{1}{\underset{\sim}{k}_i' M^{*^{-1}} \underset{\sim}{k}_i} (\underset{\sim}{k}_i' \underset{\sim}{a})' (\underset{\sim}{k}_i' \underset{\sim}{a}) + \sigma^2 \tag{43}$$

para i = 1, 2, ..., $\rho(K')$.

3.8.4. Esperança matemática da soma de quadrados da regressão.

Por (17), tem-se,

$$\hat{\underset{\sim}{c}} = R_1^{-1} \underset{\sim}{R}_2$$

ou $$\hat{\underset{\sim}{c}} = R_1^{-1} V_2 X' \underset{\sim}{y}$$

donde $$\hat{\underset{\sim}{c}}' \underset{\sim}{R}_2 = \underset{\sim}{y}' X V_2' R_1^{-1} V_2 X' \underset{\sim}{y}.$$

Considerando (37), resulta:

$$E(\hat{\underset{\sim}{c}}' \underset{\sim}{R}_2) = \underset{\sim}{c}' R_1 \underset{\sim}{c} + \sigma^2 tr(I_p) = \underset{\sim}{c}' R_1 \underset{\sim}{c} + p\sigma^2 \tag{44}$$

3.9. Distribuição de formas quadráticas

Os teoremas sobre a independência e distribuição de formas quadráticas podem ser reunidos da forma que segue.

Sejam $\underset{\sim}{y}'A_i\underset{\sim}{y}$ e $\underset{\sim}{y}'A_j\underset{\sim}{y}$ duas formas quadráticas com $\underset{\sim}{y} \;\Omega\; N(\underset{\sim}{\mu}, \sigma^2 I)$, então:

a) $\underset{\sim}{y}'A_i\underset{\sim}{y}/\sigma^2$ e $\underset{\sim}{y}'A_j\underset{\sim}{y}/\sigma^2$ são χ^2's não centrais, independentes, com parâmetros de não centralidade λ_i e λ_j, respectivamente se e somente se:

$$A_i^2 = A_i; \qquad A_j^2 = A_j \quad \text{e} \quad A_iA_j = A_jA_i = \Phi \tag{45}$$

b) Se (45) ocorre e $\lambda_j = 0$, isto é, $\underset{\sim}{y}'A_j\underset{\sim}{y}/\sigma^2$ tem distribuição χ^2 central, então a razão,

$$\frac{\underset{\sim}{y}'A_i\underset{\sim}{y}/\rho(A_i)}{\underset{\sim}{y}'A_j\underset{\sim}{y}/\rho(A_j)}$$

tem distribuição F não central, com parâmetro de não centralidade de λ_i.

3.9.1. Distribuição de $\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*/\sigma^2$

Viu-se anteriormente que:

$$\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = \underset{\sim}{y}'XV_1'M^{*^{-1}}V_1X'\underset{\sim}{y}$$

Fazendo $A_1 = XV_1'M^{*^{-1}}V_1X'$, então:

$$A_1^2 = XV_1'M^{*^{-1}}V_1X'XV_1'M^{*^{-1}}V_1X' .$$

Lembrando que $V_1X'XV_1' = C^*$ e que $C^* = M^* + A^*$, fica:

$$A_1^2 = XV_1'M^{*^{-1}}(M^* + A^*)M^{*^{-1}}V_1X'$$

$$A_1^2 = XV_1'M^{*^{-1}}V_1X' + XV_1'M^{*^{-1}}A^*M^{*^{-1}}V_1X'.$$

Entretanto, sabe-se que:

$$A^*\hat{\underline{a}} = \underline{0}$$

ou
$$A^*M^{*^{-1}}V_1X'\underline{y} = \underline{0}$$

Para que a igualdade seja válida para todo $\underline{y}$, tem-se, obrigatoriamente, que:

$$A^*M^{*^{-1}}V_1X' = \Phi$$

donde se conclui, finalmente, que:

$$A_1^2 = XV_1'M^{*^{-1}}V_1X' = A_1.$$

O número de graus de liberdade é dado por:

$$\rho(A_1) = tr(A_1)$$

por ser A_1 idempotente. Assim:

$$tr(A_1) = tr(XV_1'M^{*^{-1}}V_1X') =$$

$$= tr(M^{*^{-1}}V_1X'XV_1') =$$

$$= tr(M^{*^{-1}}C^*).$$

Sendo $M^{*^{-1}}C^*$ idempotente, fica:

$$tr(A_1) = \rho(M^{*^{-1}}C^*) = \rho(C^*) = v - 1$$

e o parâmetro de não centralidade, segundo GRAYBILL (1960), por:

$$\lambda_1 = \frac{1}{2\sigma^2} \underline{\mu}'A_1\underline{\mu}$$

como $\underline{\mu} = E(\underline{y}) = X\underline{\beta}$, fica:

$$\lambda_1 = \frac{1}{2\sigma^2} \underline{\beta}'X'XV_1'M^{*-1} V_1 X'X\underline{\beta}$$

e, sendo $V_1X'X\underline{\beta} = C^*\underline{a}$, resulta, finalmente, que:

$$\lambda_1 = \frac{1}{2\sigma^2} \underline{a}'C^*\underline{a} \ .$$

De forma mais completa, pode-se dizer que:

$$SQH_1/\sigma^2 = \hat{\underline{a}}'\underline{Q}^*/\sigma^2 \ \cap \ \chi^2[\rho(A_1), \lambda_1] \qquad (46)$$

sendo $\rho(A_1) = v - 1$ e $\lambda_1 = \frac{1}{2\sigma^2} \underline{a}'C^*\underline{a}$.

3.9.2. Distribuição de SQH_i/σ^2

Escrevendo SQH_i como uma forma quadrática em $\underline{y}$, fica:

$$SQH_i = \underline{y}'XV_1'M^{*-1}{}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'M^{*-1}V_1X'\underline{y} = \underline{y}'A_i\underline{y}$$

Como:

$$A_i^2 = XV_1'M^{*-1}{}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'M^{*-1}V_1X'XV_1'M^{*-1}{}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'M^{*-1}V_1X' =$$

$$= XV_1'M^{*-1}{}'\underline{k}_i(\underline{k}_i'M^{*-1}\underline{k}_i)^{-1}\underline{k}_i'M^{*-1}V_1X' = A_i$$

pois

$$\underset{\sim}{k}_i' M^{*-1} V_1 X'XV_1' M^{*-1\,'} \underset{\sim}{k}_i = \underset{\sim}{k}_i' M^{*-1} \underset{\sim}{k}_i$$

se $\underset{\sim}{k}_i' \underset{\sim}{a}$ é estimável e,

$$\rho(A_i) = tr(A_i) = tr[XV_1' M^{*-1\,'} \underset{\sim}{k}_i (\underset{\sim}{k}_i' M^{*-1} \underset{\sim}{k}_i)^{-1} \underset{\sim}{k}_i' M^{*-1} V_1 X'] = tr(I_1) = 1$$

e, ainda,

$$\lambda_i = \frac{1}{2\sigma^2} \underset{\sim}{\mu}' A_i \underset{\sim}{\mu} = \frac{1}{2\sigma^2} \underset{\sim}{\beta}' X'XV_1' M^{*-1\,'} \underset{\sim}{k}_i (\underset{\sim}{k}_i' M^{*-1} \underset{\sim}{k}_i)^{-1} \underset{\sim}{k}_i' M^{*-1} V_1 X'X \underset{\sim}{\beta}$$

ou
$$\lambda_i = \frac{1}{2\sigma^2 \underset{\sim}{k}_i' M^{*-1} \underset{\sim}{k}_i} (\underset{\sim}{k}_i' \underset{\sim}{a})' (\underset{\sim}{k}_i' \underset{\sim}{a})$$

como foi obtido no item 3.8.3., pode-se concluir que:

$$SQH_i/\sigma^2 \;\Omega\; \chi^2[1, \lambda_i] \tag{47}$$

3.9.3. Distribuição de $\hat{\underset{\sim}{c}}' \underset{\sim}{R}_2/\sigma^2$

De forma semelhante ao item 3.9.1.,

$$\hat{\underset{\sim}{c}}' \underset{\sim}{R}_2 = \underset{\sim}{y}' XV_2' R_1^{-1} V_2 X' \underset{\sim}{y} = \underset{\sim}{y}' A_2 \underset{\sim}{y}$$

Como,

$$A_2^2 = XV_2' R_1^{-1} V_2 X'XV_2' R_1^{-1} V_2 X' = XV_2' R_1^{-1} V_2 X' = A_2$$

e,
$$\rho(A_2) = \rho(XV_2' R_1^{-1} V_2 X')$$

ou,
$$\rho(A_2) = tr(XV_2' R_1^{-1} V_2 X')$$

que, pelas propriedades do operador traço, fica:

$$\rho(A_2) = tr(R_1^{-1}V_2X'XV_2') = tr(I_p)$$

ou, $$\rho(A_2) = p$$

e, ainda,

$$\lambda_2 = \frac{1}{2\sigma^2}\underline{\mu}'A_2\underline{\mu}$$

de onde, substituidos $\underline{\mu}$ e A_2, obtém-se:

$$\lambda_2 = \frac{1}{2\sigma^2}\underline{c}'R_1\underline{c}$$

Pode-se dizer, portanto, que:

$$SQH_2/\sigma^2 = \underline{\hat{c}}'\underline{R}_2/\sigma^2 \;\Omega\; \chi^2[\rho(A_2),\ \lambda_2] \tag{48}$$

3.9.4. Distribuição de SQR^*/σ^2

Viu-se que:

$$SQR^* = \underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}$$

ou, ainda, $$SQR^* = \underline{y}'[I - X(X'X)^-X']\underline{y}.$$

Considerando que:

$$A_3^2 = [I - X(X'X)^-X'][I - X(X'X)^-X'] =$$

$$= I - X(X'X)^-X' - X(X'X)^-X' + X(X'X)^-X'X(X'X)^-X' =$$

$$= I - X(X'X)^-X'$$

pois $X(X'X)^{-}X'X = X$, e que o número de graus de liberdade é:

$$\rho(A_3) = \rho[I_{n_{..}} - X(X'X)^{-}X'] = tr(I_{n_{..}}) - tr[X(X'X)^{-}X']$$

que, por (40), fica:

$$\rho(A_3) = n_{..} - \rho(X) = n_{..} - p - r - v + 1$$

e, considerando, ainda, o parâmetro de não centralidade,

$$\lambda_3 = \frac{1}{2\sigma^2}\, \underset{\sim}{\beta}'X'[I - X(X'X)^{-}X']X\underset{\sim}{\beta}$$

ou, $\lambda_3 = 0$

já que $X'[I - X(X'X)^{-}X'] = \Phi$, conclui-se que SQR^*/σ^2 tem distribuição χ^2 central com $n_{..} - p - r - v + 1$ graus de liberdade ou, simbolicamente,

$$SQR^*/\sigma^2 \; \Omega \; \chi^2[\rho(A_3),\ 0] \qquad (49)$$

3.9.5. Distribuição da razão $\dfrac{\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*/\rho(A_1)}{SQR^*/\rho(A_3)}$

Viu-se que:

$$SQH_1 = \hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = \underset{\sim}{y}'XV_1'M^{*^{-1}}V_1X'\underset{\sim}{y}$$

e que $SQR^* = \underset{\sim}{y}'[I - X(X'X)^{-}X']\underset{\sim}{y}$

A condição necessária e suficiente para que SQH_1 e SQR^* sejam independentes é:

$$A_1A_3 = A_3A_1 = \Phi.$$

Dessa forma, a condição exige que:

$$XV_1'M^{*-1}V_1X'[I - X(X'X)^{-}X'] = \Phi$$

Ora, foi visto que:

$$X'[I - X(X'X)^{-}X'] = \Phi$$

Logo,

$$A_1A_3 = \Phi$$

Pode-se, então, concluir que:

$$\frac{SQH_1/\rho(A_1)}{SQR^*/\rho(A_3)} \; \Omega \; F[\rho(A_1),\ \rho(A_3),\ \lambda_1] \tag{50}$$

ou, ainda, o quociente tem distribuição F não central com $\rho(A_1) = v - 1$ e $\rho(A_3) = n_{..} - p - r - v + 1$ graus de liberdade e parâmetro de não centralidade $\lambda_1 = \frac{1}{2\sigma^2}\underset{\sim}{a}'C^*\underset{\sim}{a}$.

3.9.6. Distribuição de $\dfrac{SQH_i}{SQR^*/\rho(A_3)}$

A independência entre SQH_i e SQR^* é facilmente verificada pois:

$$A_iA_3 = XV_1'M^{*-1}{}'\underset{\sim}{k}_i(\underset{\sim}{k}_i'M^{*-1}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'M^{*-1}V_1X'[I - X(X'X)^{-}X'] = \Phi$$

já que $X'[I - X(X'X)^{-}X'] = \Phi$.

Assim, pode-se dizer que:

$$\frac{SQH_i}{SQR^*/\rho(A_3)} \; \Omega \; F[1,\ (n_{..} - p - r - v + 1),\ \lambda_i] \tag{51}$$

ou seja, a razão tem distribuição F com 1 e $n_{..} - p - r - v + 1$ graus de liberdade e parâmetro de não centralidade λ_i.

3.9.7. Distribuição da razão $\dfrac{\hat{\underset{\sim}{c}}'\underset{\sim}{R}_2/\rho(A_2)}{SQR^*/\rho(A_3)}$

De maneira semelhante ao item anterior pode-se verificar, facilmente, que:

$$A_2A_3 = A_3A_2 = \Phi .$$

Assim,

$$\frac{SQH_2/\rho(A_2)}{SQR^*/\rho(A_3)} \;\Omega\; F[\rho(A_2),\ \rho(A_3),\ \lambda_2] \tag{52}$$

isto é, a razão tem distribuição F não central com $\rho(A_2) = p$ e $\rho(A_3) = n_{..} - p - r - v + 1$ graus de liberdade e parâmetro de não centralidade $\lambda_2 = \dfrac{1}{2\sigma^2}\, \underset{\sim}{c}'R_1\underset{\sim}{c}$.

3.10. Soma de quadrados e teste de significância para a interação

Caso haja necessidade de um teste de significância para a interação, a soma de quadrados apropriada para o numerador do teste F é:

$$S.Q.Int^* = R(\underset{\sim}{a},\ \underset{\sim}{b},\ \underset{\sim}{\gamma},\ \underset{\sim}{c}) - R(\underset{\sim}{a},\ \underset{\sim}{b},\ \underset{\sim}{c})$$

Lembrando que:

$$R(\underset{\sim}{a},\ \underset{\sim}{b},\ \underset{\sim}{c}) = \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} + \underset{\sim}{B}'K^{-1}\underset{\sim}{B} + \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2$$

tem-se:

$$S.Q.Int^* = R(\underset{\sim}{a}, \underset{\sim}{b}, \underset{\sim}{\gamma}, \underset{\sim}{c}) - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2$$

ou, ainda,

$$S.Q.Int^* = \sum_i \sum_j \frac{y_{ij.}^2}{n_{ij}} + \hat{\underset{\sim}{c}}^{*\prime}\underset{\sim}{R}_2^* - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B} - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2 \quad (53)$$

Uma outra forma de escrever (53) é:

$$S.Q.Int^* = \left(\sum_i \sum_j \frac{y_{ij.}^2}{n_{ij}} - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - \underset{\sim}{B}'K^{-1}\underset{\sim}{B}\right) + \hat{\underset{\sim}{c}}^{*\prime}\underset{\sim}{R}_2^* - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2$$

ou, ainda,

$$S.Q.Int^* = S.Q.Int. + \hat{\underset{\sim}{c}}^{*\prime}\underset{\sim}{R}_2^* - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2 \quad (54)$$

sendo que $\hat{\underset{\sim}{c}}^*$ e $\underset{\sim}{R}_2^*$ são obtidos de forma análoga à $\hat{\underset{\sim}{c}}$ e $\underset{\sim}{R}_2$, devendo, agora, ser considerada a interação para o cálculo das somas de quadrados e produtos referentes ao resíduo, consideradas as co variáveis e o termo $\sum_i \sum_j y_{ij.}^2/n_{ij}$ é a redução na soma de quadrados devido ao ajustamento das constantes a, b e γ.

O sistema $R_1^*\hat{\underset{\sim}{c}}^* = \underset{\sim}{R}_2^*$ pode ser obtido de maneira análoga ao caso sem interação, ou seja, tomando R_1^* e $\underset{\sim}{R}_2^*$ conforme dadas a seguir.

Admitindo que exista R_1^{*-1}, então:

$$\hat{\underset{\sim}{c}}^* = R_1^{*-1}\underset{\sim}{R}_2^*$$

e, como essas reduções, divididas por σ^2, têm, independentemente, distribuição χ^2, pode-se concluir que a razão,

$$\frac{S.Q.Int^*/(s - r - v + 1)}{SQR^*/(n_{..} - s - p)} \quad (55)$$

tem distribuição F não central, sendo s o número de células on-

$$
R_i^* = \begin{bmatrix}
\sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(1)} x_{ijk}^{(1)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(1)} x_{ij.}^{(1)}}{n_{ij}} & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(1)} x_{ijk}^{(2)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(1)} x_{ij.}^{(2)}}{n_{ij}} & \cdots & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(1)} x_{ijk}^{(p)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(1)} x_{ij.}^{(p)}}{n_{ij}} \\
\sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(2)} x_{ijk}^{(1)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(2)} x_{ij.}^{(1)}}{n_{ij}} & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(2)} x_{ijk}^{(2)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(2)} x_{ij.}^{(2)}}{n_{ij}} & \cdots & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(2)} x_{ijk}^{(p)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(2)} x_{ij.}^{(p)}}{n_{ij}} \\
\cdots & \cdots & \cdots & \cdots \\
\sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(p)} x_{ijk}^{(1)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(p)} x_{ij.}^{(1)}}{n_{ij}} & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(p)} x_{ijk}^{(2)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(p)} x_{ij.}^{(2)}}{n_{ij}} & \cdots & \sum\limits_{i}\sum\limits_{j}\sum\limits_{k} x_{ijk}^{(p)} x_{ijk}^{(p)} - \sum\limits_{i}\sum\limits_{j} \dfrac{x_{ij.}^{(p)} x_{ij.}^{(p)}}{n_{ij}}
\end{bmatrix}
$$

$$\underset{\sim}{R}_2^* = \begin{bmatrix} \sum\limits_{ijk}\sum\sum x_{ijk}^{(1)} y_{ijk} - \sum\limits_{ij}\sum \dfrac{x_{ij.}^{(1)} y_{ij.}}{n_{ij}} \\ \\ \sum\limits_{ijk}\sum\sum x_{ijk}^{(2)} y_{ijk} - \sum\limits_{ij}\sum \dfrac{x_{ij.}^{(2)} y_{ij.}}{n_{ij}} \\ \\ \cdots \\ \\ \sum\limits_{ijk}\sum\sum x_{ijk}^{(p)} y_{ijk} - \sum\limits_{ij}\sum \dfrac{x_{ij.}^{(p)} y_{ij.}}{n_{ij}} \end{bmatrix}$$

de $n_{ij} > 0$ e,

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - R(\underset{\sim}{a}, \underset{\sim}{b}, \underset{\sim}{\gamma}, \underset{\sim}{c})$$

ou

$$SQR^* = \underset{\sim}{y}'\underset{\sim}{y} - \sum_{ij}\sum \frac{y_{ij.}^2}{n_{ij}} - \hat{\underset{\sim}{c}}^{*\prime}\underset{\sim}{R}_2^* \qquad (56)$$

4. RESULTADOS E DISCUSSÃO

4.1. Estimadores dos parâmetros

4.1.1. Estimador de $\underset{\sim}{c}$

Por (17) vê-se que:

$$\hat{\underset{\sim}{c}} = R_1^{-1} \underset{\sim}{R}_2$$

onde R_1 é a matriz dos resíduos das somas de quadrados e produtos referentes às variáveis independentes e $\underset{\sim}{R}_2$ é o vetor dos resíduos referentes às somas de produtos das variáveis independentes e à variável dependente. Resultado idêntico foi obtido por SEARLE (1971) e DIAS (1981), dentre outros.

4.1.2. Estimador de $\underset{\sim}{a}$

O estimador de mínimos quadrados para $\underset{\sim}{a}$, obtido pelo Método do Ajustamento de Constantes, como se pode observar em (18) é:

$$\hat{\underset{\sim}{a}} = M^{*^{-1}}\underset{\sim}{Q}^{*}$$

sendo $M^* = C^* + A^*$ e A^* uma matriz construida segundo uma restrição previamente estabelecida.

4.2. Matriz de Dispersão

4.2.1. Matriz de dispersão de $\hat{\underset{\sim}{c}}$

Como se pode notar por (20), a matriz de dispersão de $\hat{\underset{\sim}{c}}$ é:

$$D(\hat{\underset{\sim}{c}}) = R_1^{-1}\sigma^2.$$

Igual resultado foi obtido por ZELLEN (1957), SEARLE (1971) e DIAS (1981), dentre outros.

4.2.2. Matriz de dispersão de $\hat{\underset{\sim}{a}}$

A matriz de dispersão de $\hat{\underset{\sim}{a}}$, conforme (19), é:

$$D(\hat{a}) = M^{*^{-1}}C^{*}M^{*^{-1}\prime}\sigma^2.$$

4.3. Médias ajustadas dos níveis dos fatores

As médias dos níveis de A, ajustadas para o efeito do fator B e da regressão são obtidas de:

$$j\hat{\mu} + \hat{a} = R^{-1}[T - N\hat{b} - A\hat{c}]$$

enquanto que as dos níveis de B, ajustadas para o fator A e para a regressão, resultam de:

$$j\hat{\mu} + \hat{b} = K^{-1}[B - N'\hat{a} - D\hat{c}]$$

de acordo com (24) e (21), respectivamente.

O estimador $\hat{\mu}$ pode ser obtido por:

$$\hat{\mu} = \frac{1}{v} R^{-1}[T - N\hat{b} - A\hat{c}]$$

ou por:

$$\hat{\mu} = \frac{1}{r} K^{-1}[B - N'\hat{a} - D\hat{c}]$$

conforme (22), desde que impostas as restrições $\sum_i \hat{a}_i = 0$ e $\sum_j \hat{b}_j = 0$, respectivamente.

Igual procedimento foi utilizado por FEDERER (1957)

4.4. Variância do contraste entre duas médias ajustadas

A variância do contraste entre as médias ajustadas dos níveis do fator A, por exemplo, é:

$$V(\hat{a}_i - \hat{a}_{i'}) = (d_{ii} + d_{i'i'} - 2d_{ii'})\sigma^2$$

onde d_{ii}, $d_{i'i'}$ e $d_{ii'}$ são elementos da matriz de dispersão $D(\hat{a})$, dada por (19).

4.5. Testes de hipóteses

Discutir-se-á, inicialmente, o caso onde $x_{ijk}=0$, para todo $\underline{i}$, $\underline{j}$ e $\underline{k}$ em (4), ou seja, serão discutidas as hipóteses comumente testadas na análise da variância de classificações duplas.

É importante tornar claro o conceito de hipótese testável. Segundo SEARLE (1971), uma hipótese só é testável quando for expressa através de funções lineares estimáveis.

No modelo de classificação dupla, incluida a interação, a função linear estimável básica é $\mu_{ij} = E(y_{ijk})$, ou ainda, $\mu_{ij} = \mu + a_i + b_j + \gamma_{ij}$. Dessa forma, todas as hipóteses serão formuladas em termos dos μ_{ij} pois, além de facilitar a compreensão sobre o que está sendo testado, toda combinação linear dos μ_{ij} é estimável, logo é testável.

Para o caso da análise da variância através do Método do Ajustamento de Constantes, as hipóteses comumente formuladas são:

$$H_1: \sum_j n_{ij}\mu_{ij}/n_{i.} = \sum_j n_{i'j}\mu_{i'j}/n_{i'.}$$

$$H_2: \sum_j n_{ij}\mu_{ij} = \sum_{i'} \sum_j n_{ij}n_{i'j}\mu_{i'j}/n_{.j}$$

$$H_3: \sum_i n_{ij}\mu_{ij}/n_{.j} = \sum_i n_{ij'}\mu_{ij'}/n_{.j'}$$

$$H_4: \sum_i n_{ij}\mu_{ij} = \sum_{j'} \sum_i n_{ij}n_{ij'}\mu_{ij'}/n_{i.}$$

$$H_5: \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0$$

Substituindo μ_{ij} em H_1, pela sua expressão, fica:

$$H_1: a_i - a_{i'} + \frac{1}{n_{i.}} \sum_j n_{ij}(b_j + \gamma_{ij}) - \frac{1}{n_{i'.}} \sum_j n_{i'j}(b_j + \gamma_{i'j}) = 0$$

para todo $i \neq i'$.

Similarmente, em H_3, obtém-se:

$$H_3: b_j - b_{j'} + \frac{1}{n_{.j}} \sum_i n_{ij}(a_i + \gamma_{ij}) - \frac{1}{n_{.j'}} \sum_i n_{ij'}(a_i + \gamma_{ij'}) = 0$$

De modo que, o teste F para H_1, $F(H_1)$, testa a hipótese de que os níveis do fator A, na presença de uma média ponderada dos b's e γ's, cujos pesos de ponderação são os valores n_{ij}, são iguais para todo i. Este teste recebe, de alguns autores, o nome de "teste para A ignorando B e a interação", e a soma de quadrados apropriada para o numerador de $F(H_1)$ é $R(a/\mu)$.

Segundo SEARLE (1971), o melhor estimador linear imparcial (MELI) de μ_{ij} é:

$$\hat{\mu}_{ij} = \hat{\mu} + \hat{a}_i + \hat{b}_j + \hat{\gamma}_{ij} = \bar{y}_{ij.}$$

onde $\bar{y}_{ij.} = \sum_k y_{ijk}/n_{ij}$.

De maneira que o MELI da expressão associada à H_1 é:

$$\sum_j \frac{n_{ij}}{n_{i.}} \hat{\mu}_{ij} - \sum_j \frac{n_{i'j}}{n_{i'.}} \hat{\mu}_{i'j} = \bar{y}_{i..} - \bar{y}_{i'..}$$

Em H_2, a substituição de μ_{ij} pela sua expressão faz com que:

$$H_2: \sum_j n_{ij}(\mu + a_i + b_j + \gamma_{ij}) - \sum_{ij}\sum \frac{n_{ij}n_{i'j}}{n_{.j}}(\mu + a_{i'} + b_j + \gamma_{i'j}) = 0$$

Desenvolvendo a expressão, obtém-se:

$$H_2:\ (n_{i.} - \sum_j \frac{n_{ij}^2}{n_{.j}})a_i - \sum_{i \neq i'} \sum_j \frac{n_{ij}n_{i'j}}{n_{.j}} a_{i'} +$$

$$+ \sum_j (n_{ij} - \frac{n_{ij}^2}{n_{.j}})\gamma_{ij} - \sum_{i \neq i'} \sum_j \frac{n_{ij}n_{i'j}}{n_{.j}} \gamma_{i'j} = 0$$

Analogamente, a substituição em H_4, resulta:

$$H_4:\ (n_{.j} - \sum_i \frac{n_{ij}^2}{n_{i.}})b_j - \sum_{j \neq j'} \sum_i \frac{n_{ij}n_{ij'}}{n_{i.}} b_{j'} +$$

$$+ \sum_i (n_{ij} - \frac{n_{ij}^2}{n_{i.}})\gamma_{ij} - \sum_{j \neq j'} \sum_i \frac{n_{ij}n_{ij'}}{n_{i.}} \gamma_{ij'} = 0$$

As somas de quadrados apropriadas para os numeradores de $F(H_2)$ e $F(H_4)$ são $R(a/\mu,b)$ e $R(b/\mu,a)$, respectivamente, obtidas pelo ajustamento do modelo,

$$y_{ijk} = \mu + a_i + b_j + e_{ijk} .$$

Finalmente, com relação a H_5, tem-se:

$$H_5:\ \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0$$

ou, equivalentemente,

$$H_5:\ \gamma_{ij} - \gamma_{i'j} - \gamma_{ij'} + \gamma_{i'j'} = 0.$$

Através do exposto, questões interessantes podem ser levantadas. Caso se considere a interação no modelo, os contrastes do tipo $a_i - a_{i'}$ ou $b_j - b_{j'}$, não são estimáveis já que não existe nenhum contraste entre os μ_{ij} que os represente, ou seja, nenhum contraste entre os a_i, ou entre os b_j, é possível sem a presença, pelo menos, da interação.

Em nenhuma das hipóteses associadas ao Método do Ajustamento de Constantes, a não ser H_5, exigem que $n_{ij} > 0$. A exigência de que $n_{ij} > 0$, em H_5, é fundamental já que μ_{ij} só é estimável nessas condições. Entretanto, mesmo existindo funções lineares não estimáveis, entre os μ_{ij} que compoem H_5, poderá existir uma combinação linear estimável dessas funções. Assim, por exemplo, seja:

$$\theta_{11,22} = \mu_{11} - \mu_{12} - \mu_{21} + \mu_{22}$$

$$\theta_{13,33} = \mu_{11} - \mu_{13} - \mu_{31} + \mu_{33}$$

com $n_{11} = 0$. Ora, se $n_{11} = 0$, nem $\theta_{11,22}$, nem $\theta_{13,33}$ são estimáveis, já que μ_{11} não o é. Mas, tomando-se

$$\theta_{11,22} - \theta_{13,33} = \mu_{22} - \mu_{12} - \mu_{21} + \mu_{13} + \mu_{31} - \mu_{33}$$

verifica-se perfeitamente que $\theta_{11,22} - \theta_{13,33}$ é estimável pois μ_{11} não está presente.

Um outro aspecto a ser considerado é o estabelecimento de restrições não estimáveis aos parâmetros do modelo. Sabe-se que estimativas de funções estimáveis e testes de hipóteses testáveis são invariantes à escolha de condições não estimáveis. No entanto, o estabelecimento de condições não estimáveis altera completamente a formulação das hipóteses. HARVEY (1975), cita o caso, também comentado por HOCKING e SPEED(1975), onde o estabelecimento de algumas condições não estimáveis mudam completamente a hipótese testada. Uma prática recomendável é associar essas restrições, caso existam, às hipóteses formula -

das, ou, melhor ainda, formulá-las em termos de funções lineares estimáveis.

Pode-se notar, pelas hipóteses formuladas, que a interpretação dos resultados é grandemente facilitada caso não seja considerada a interação.

Em (42), vê-se que se $x_{ijk} = 0$, para todo i,j e k, então:

$$E(\hat{a}'Q) = a'Ca + (v - 1)\sigma^2.$$

A hipótese de igualdade dos efeitos dos níveis do fator A é dada por:

$$H_1: a_i - a_{i'} = 0, \text{ para todo } i \neq i',$$

ou, de outra forma:

$$H_1: a - ja = 0.$$

Caso H_1 se verifique, tem-se que:

$$E(\hat{a}'Q) = a^2 j'Cj + (v - 1)\sigma^2 .$$

Como $j'C = 0'$, pode-se verificar que $\hat{a}'Q/(v-1)$, sob H_1, é um estimador imparcial de σ^2. Ou seja, $F(H_1)$ testa a hipótese de que o efeito dos níveis de A é constante. Raciocínio análogo vale para o fator B.

Já, a soma de quadrados $R(a/\mu)$ está associada a hipótese:

$$H: a_i + \frac{1}{n_{i.}} \sum_j n_{ij} b_j \text{ são iguais para todo } i,$$

ou seja, $R(a/\mu)$ é a soma de quadrados apropriada para o teste da hipótese em que os efeitos dos níveis de A, na presença de uma média ponderada dos efeitos dos níveis de B, são iguais para todo i.

Na análise de covariância, as hipóteses a serem testadas podem ser colocadas de maneira análoga às da análise da variância. Nesta discussão, o modelo básico a ser considerado será:

$$y_{ijk} = \mu + a_i + b_j + \sum_{g=1}^{p} c^{(g)} x_{ijk}^{(g)} + e_{ijk}$$

ou seja, o modelo sem interação. A significância desta poderá, se necessário, ser testada pelo procedimento indicado através de (55).

Conforme visto anteriormente, hipóteses testáveis são aquelas expressas em termos de funções lineares estimáveis. Torna-se interessante, portanto, verificar a estimabilidade de funções lineares de parâmetros na análise de covariância.

Sabe-se que, através da teoria das funções lineares estimáveis e considerando-se o sistema de equações normais $C^*\underline{\hat{a}} = \underline{Q}^*$, um conjunto de funções lineares $K'\underline{a}$ será estimável se houver uma matriz Λ tal que o sistema:

$$K' = \Lambda' C^* \qquad (57)$$

tenha solução.

Ora, pelas características de C^*, vê-se facilmente que $C^*\underline{j} = \underline{0}$, de forma que (57) será compatível se K' satisfaz a condição:

$$K'\underline{j} = \underline{0}.$$

Assim, conclui-se que $K'\underline{a}$ é estimável sempre que for um contraste entre os níveis de A. Idêntico resultado pode ser estendido aos níveis de B.

Com relação ao parâmetro $\underline{c}$, toda e qualquer função linear dos coeficientes de regressão é estimável já que $\underline{c}$

também o $\underset{\sim}{e}$. Uma hipótese comumente formulada acerca desses coeficientes é que $K'\underset{\sim}{c} = \underset{\sim}{0}$, onde $K' = I_p$, o que corresponde a tornar

$$H_2: \underset{\sim}{c} = \underset{\sim}{0}.$$

Por (42), vê-se que:

$$E(\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*) = \underset{\sim}{a}'C^*\underset{\sim}{a} + (v - 1)\sigma^2$$

e, por (46),

$$\lambda_1 = \frac{1}{2\sigma^2}\underset{\sim}{a}'C^*\underset{\sim}{a}$$

Caso $H_1: \underset{\sim}{a} - \underset{\sim}{j}a = \underset{\sim}{0}$ se verifique, pode-se notar que $\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*/(v-1)$ é um estimador imparcial de σ^2 e que a razão:

$$F(H_1) = \frac{\hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*/(v - 1)}{QMR^*}$$

tem distribuição F central com $(v - 1)$ e $(n_{..} - r - p - v + 1)$ graus de liberdade.

Observando os resultados obtidos por DIAS(1981) pode-se notar, ressalvando-se a diferença de notação, que as somas de quadrados obtidas pelo Processo do Resíduo Condicional e pelo Método do Ajustamento de Constantes, testam a mesma hipótese. Nota-se ainda, que a hipótese de nulidade formulada por esse autor, é:

$$H_1: \underset{\sim}{a} = \underset{\sim}{0}$$

se for utilizada a notação deste trabalho. Essa mesma hipótese poderia ser escrita como:

$$H_1: K'\underset{\sim}{a} = \underset{\sim}{0}$$

onde $K' = I_v$.

Ora, se observada a condição (57) de estimabilidade, observa-se que não existe Λ' tal que $\Lambda'C^* = I_v$, já que C^* é uma matriz singular. De modo que a referida hipótese deve ser formulada sob a forma de funções lineares estimáveis, uma das quais é dada por (28).

Com relação à regressão, através de (44), tem-se:

$$E(\hat{\underline{c}}'\underline{R}_2) = \underline{c}'R_2\underline{c} + p\sigma^2$$

e, por (48),

$$\lambda_2 = \frac{1}{2\sigma^2}\,\underline{c}'R_2\underline{c}$$

Da mesma forma, sob H_2: $\underline{c} = \underline{0}$, $\hat{\underline{c}}'\underline{R}_2/p$ é um estimador imparcial do resíduo e a razão:

$$F(H_2) = \frac{\hat{\underline{c}}'\underline{R}_2/p}{QMR^*}$$

tem distribuição F central com $\underline{p}$ e $(n_{..} - r - p - v + 1)$ graus de liberdade.

A soma de quadrados $\hat{\underline{a}}'\underline{Q}^*$, conforme se demonstrou no item 3.6.4., pode ser decomposta num conjunto de $(v - 1)$ partes independentes, desde que a condição $\underline{k}_i'M^{*-1}\underline{k}_j = 0$ seja satisfeita. Como $\hat{\underline{a}}'\underline{Q}^*/\sigma^2$ tem distribuição χ^2, cada q_i/σ^2 também se distribui independentemente como um χ^2, o que permite dizer que a razão:

$$\frac{S.Q.H_i}{QMR^*}$$

sob a hipótese $H_i : \underline{k}_i'\underline{a} = 0$ tem distribuição F central com $\underline{1}$ e $(n_{..} - r - p - v + 1)$ graus de liberdade.

Se for observada a formulação das hipóteses na análise da covariância, nota-se que são as mesmas da análise da variância. Assim, por analogia, pode-se dizer que a hipótese testada em (55) é:

$$H_5: \gamma_{ij} - \gamma_{i'j} - \gamma_{ij'} + \gamma_{i'j'} = 0.$$

Da mesma forma, pode-se dizer aqui, que se H_5 não se verifica, os testes de significância para os efeitos principais e para a interação podem ser obtidos pela análise da covariância através do Método das Somas de Quadrados de Médias Ponderadas.

4.6. Exemplo numérico

O exemplo que aqui será discutido refere-se a um experimento onde a variável dependente y é o número de rosas vendáveis, sendo que, cada repetição está localizada em uma estufa. A variável $X^{(1)}$ foi incluida com a finalidade de estabelecer um tipo de controle sobre a variabilidade dentro de cada repetição, de acordo com FEDERER (1957), de onde foram retirados os dados aqui apresentados. A variável $X^{(2)}$ é fictícia e foi incluida a fim de melhor exemplificar as fórmulas propostas por este trabalho.

Na tabela 5 encontram-se os valores das variáveis y, $X^{(1)}$ e $X^{(2)}$, sendo que, na repetição I, foram tomadas duas medidas para o tratamento 1 e uma medida para os demais, enquanto que, na repetição II, no primeiro tratamento, fez-se apenas uma observação e duas para os restantes.

Tabela 5. Valores observados das variáveis y, $X^{(1)}$ e $X^{(2)}$.

Tratamentos	Repetição I		Repetição II	
y_1	27	18	19	
$X^{(1)}$	5	5	4	
$X^{(2)}$	4	4	3	
y_2	31		52	57
$X^{(1)}$	3		5	5
$X^{(2)}$	2		4	4
y_3	34		52	33
$X^{(1)}$	2		1	1
$X^{(2)}$	1		5	3
y_4	38		60	45
$X^{(1)}$	4		2	2
$X^{(2)}$	3		2	1
y_5	31		50	40
$X^{(1)}$	1		3	3
$X^{(2)}$	5		1	2

4.6.1. Sistema de equações normais

O sistema de equações normais $X'X\hat{\underline{\beta}} = X'\underline{y}$, desconsiderando a média, considerados os dados da tabela 5, é:

$$\left[\begin{array}{ccccc:cc:cc}
3 & 0 & 0 & 0 & 0 & 2 & 1 & 4,8 & 2,2 \\
0 & 3 & 0 & 0 & 0 & 1 & 2 & 3,8 & 1,2 \\
0 & 0 & 3 & 0 & 0 & 1 & 2 & -5,2 & 0,2 \\
0 & 0 & 0 & 3 & 0 & 1 & 2 & -1,2 & -2,8 \\
0 & 0 & 0 & 0 & 3 & 1 & 2 & -2,2 & -0,8 \\
\hdashline
2 & 1 & 1 & 1 & 1 & 6 & 0 & 1,6 & 1,4 \\
1 & 2 & 2 & 2 & 2 & 0 & 9 & -1,6 & -1,4 \\
\hdashline
4,8 & 3,8 & -5,2 & -1,2 & -2,2 & 1,6 & -1,6 & 32,93 & 5,06 \\
2,2 & 1,2 & 0,2 & -2,8 & -0,8 & 1,4 & -1,4 & 5,06 & 26,93
\end{array}\right]
\left[\begin{array}{c}
\hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hdashline \hat{b}_1 \\ \hat{b}_2 \\ \hdashline \hat{c}_1 \\ \hat{c}_2
\end{array}\right]
=
\left[\begin{array}{c}
64 \\ 140 \\ 119 \\ 143 \\ 121 \\ \hdashline 179 \\ 408 \\ \hdashline -45,13 \\ -29,86
\end{array}\right]$$

onde a linha pontilhada permite uma melhor visualização das matrizes e vetores representados anteriormente.

A partir do sistema geral $X'X\hat{\underline{\beta}} = X'\underline{y}$ poderão ser obtidos outros sistemas menores, perfeitamente equivalentes. Estes sistemas são:

a) Equações normais para estimação dos níveis de A ajustados para B, ignorando as variáveis auxiliares.

$$\frac{1}{18}\begin{bmatrix}
40 & -10 & -10 & -10 & -10 \\
-10 & 43 & -11 & -11 & -11 \\
-10 & -11 & 43 & -11 & -11 \\
-10 & -11 & -11 & 43 & -11 \\
-10 & -11 & -11 & -11 & 43
\end{bmatrix}
\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{bmatrix}
=
\begin{bmatrix} -41,0 \\ 19,5 \\ -1,5 \\ 22,5 \\ 0,5 \end{bmatrix} \qquad (58)$$

Esse sistema é obtido por (10) fazendo $x_{ijk} = 0$ para todo $\underline{i}$, $\underline{j}$ e k.

b) Equações normais para estimação dos coeficientes de regressão.

Considerando (14), os coeficientes de regressão para o caso dos dados da tabela 5, são obtidos do sistema de equações:

$$\begin{bmatrix} 28 & -4 \\ -4 & 65,6 \end{bmatrix} \begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} = \begin{bmatrix} 27,333 \\ 32,600 \end{bmatrix} \quad (59)$$

c) Equações normais para estimação dos efeitos dos níveis de A, ajustados para B e para a regressão.

Através de (13), esse sistema é:

$$\begin{bmatrix} 1,5460 & -1,1287 & 0,1008 & -0,2751 & -0,2430 \\ -1,1287 & 1,8982 & -0,0225 & -0,4031 & -0,3439 \\ 0,1008 & -0,0225 & 1,5407 & -0,6894 & -0,9296 \\ -0,2751 & -0,4031 & -0,6894 & 2,0945 & -0,7269 \\ -0,2430 & -0,3439 & -0,9296 & -0,7269 & 2,2434 \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{bmatrix} = \begin{bmatrix} -37,9300 \\ 22,1082 \\ -4,5208 \\ 21,2644 \\ -0,9219 \end{bmatrix} \quad (60)$$

Facilmente se verifica que os sistemas (58) e (60) não têm solução única, enquanto que (59) tem apenas uma solução. Estimativas únicas para a_i podem ser obtidas aplicando-se a restrição $\underline{j}'\hat{\underline{a}} = 0$. Assim, os vetores das estimativas dos efeitos dos níveis de A, ajustados para B; dos níveis de A, ajustados para B e para a regressão e dos coeficientes de regressão são:

$$\hat{\underline{a}} = \begin{bmatrix} -14,7600 \\ 6,7733 \\ -0,2266 \\ 7,7733 \\ 0,4400 \end{bmatrix} \quad ; \quad \hat{\underline{a}} = \begin{bmatrix} -20,9758 \\ 2,0630 \\ 5,1318 \\ 10,5904 \\ 3,1907 \end{bmatrix} \quad \text{e} \quad \hat{\underline{c}} = \begin{bmatrix} 3,1692 \\ 1,6841 \end{bmatrix}$$

4.6.2. Análise da variância

De acordo com as fórmulas apresentadas anteriormente, tem-se que:

$$SQR^* = (\underset{\sim}{y}'\underset{\sim}{y} - C) - \underset{\sim}{Q}'M^{-1}\underset{\sim}{Q} - (\underset{\sim}{B}'K^{-1}\underset{\sim}{B} - C) - \hat{\underset{\sim}{c}}'\underset{\sim}{R}_2, \quad \text{por (26)}$$

donde,

$$SQR^* = 2435{,}7333 - 912{,}70 - 864{,}8999 - 141{,}5251 = 516{,}6082$$

$$S.Q.A(\text{ajust.para B e regr.}) = \hat{\underset{\sim}{a}}'\underset{\sim}{Q}^*, \qquad \text{por (29)}$$

ou, $$S.Q.A(\text{ajust.para B e regr.}) = 1040{,}2786$$

Alternativamente, poderia ser feito:

$$S.Q.H_1 = \hat{\underset{\sim}{a}}'\underset{\sim}{Q}^* = (K'\hat{\underset{\sim}{a}})'(K'M^{*-1}K)^{-1}K'\hat{\underset{\sim}{a}}, \qquad \text{por (27)}$$

cujos valores numéricos do exemplo são:

$$K'M^{*-1}K = \begin{bmatrix} 1{,}357352078 & 0{,}888986525 & 0{,}094463972 & 0{,}380492090 \\ 0{,}888986525 & 1{,}137082600 & 0{,}127416521 & 0{,}377562975 \\ 0{,}094463972 & 0{,}127416521 & 0{,}775483303 & 0{,}292179261 \\ 0{,}384920900 & 0{,}377562975 & 0{,}292179261 & 0{,}696250730 \end{bmatrix}$$

$$(K'M^{*-1}K)^{-1} = \begin{bmatrix} 1{,}546042365 & -1{,}128651060 & 0{,}100780379 & -0{,}275139353 \\ -1{,}128651060 & 1{,}898216278 & -0{,}225195118 & -0{,}403121516 \\ 0{,}100780379 & -0{,}022519512 & 1{,}540691194 & -0{,}689409141 \\ -0{,}275139353 & -0{,}403121516 & -0{,}689409141 & 2{,}094537355 \end{bmatrix}$$

$$K'\hat{\underline{a}} = \begin{bmatrix} -24,166520230 \\ -1,127709453 \\ 1,941124787 \\ 7,399677792 \end{bmatrix}$$

sendo que:

$$K' = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

e $M^* = C^* - A^*$, onde A^* é a matriz de restrição, tomada como sendo:

$$A^* = (-1,12865106)J \tag{61}$$

onde J é uma matriz 5 x 5, cujos elementos constituintes são todos iguais a unidade.

Efetuando o produto, obtém-se:

$$\hat{\underline{a}}'\underline{Q}^* = (K'\hat{\underline{a}})'(K'M^{*-1}K)^{-1}(K'\hat{\underline{a}}) = 1040,2786$$

A tabela 6 sumariza os resultados obtidos através da aplicação das fórmulas apresentadas anteriormente aos dados da tabela 5.

Se B fosse considerado um fator de interesse, a análise se faria de modo idêntico àquele feito para o fator A. A tabela 7 resume esses resultados.

Tabela 6. Análise da covariância linear dos dados apresentados na tabela 5, onde o fator A refere-se à tratamentos e B à repetições.

Causas da Variação	GL	S.Q.	Q.M.	F
B(não ajustado)	1	864,8999		
A(ajust.para B)	4	912,7000		
Regressão	2	141,5251	70,7625	0,958 ns
Resíduo(ajust.p/regr.)	7	516,6082	73,8011	
A(aj.para B e p/regr.)	4	1040,2786	260,0696	3,523 ns
Total	14	2435,7333		

Tabela 7. Análise da covariância linear dos dados apresentados na tabela 5, sendo B o fator ajustado.

Causas da Variação	GL	S.Q.	Q.M.	F
A(não ajustado)	4	1344,3999		
B(ajust.para A)	1	433,2000		
Regressão	2	141,5251	70,7625	0,958 ns
Resíduo(ajust.p/regr.)	7	516,6082	73,8011	
B(aj.para A e p/regr.)	1	456,3692	456,3692	6,183 *
Total	14	2435,7333		

4.6.3. Médias ajustadas dos níveis dos fatores A e B

As médias ajustadas dos níveis de B são, por (21):

$$\underset{\sim}{j}\hat{\mu} + \hat{\underset{\sim}{b}} = K^{-1}\left[\underset{\sim}{B} - N'\hat{\underset{\sim}{a}} - D\hat{\underset{\sim}{c}}\right]$$

$$\underset{\sim}{B} - N'\hat{\underset{\sim}{a}} - D\hat{\underset{\sim}{c}} = \begin{bmatrix} 192,5474517 \\ \\ 394,4525483 \end{bmatrix}$$

donde,

$$j\hat{\mu} + \hat{\underset{\sim}{b}} = \begin{bmatrix} 32,09124195 \\ \\ 43,82806092 \end{bmatrix} \tag{62}$$

Consequentemente,

$$\hat{\mu} = 37,9597$$

e,

$$\hat{\underset{\sim}{b}} = \begin{bmatrix} -5,8684 \\ \\ 5,8684 \end{bmatrix}$$

Pelo mesmo raciocícinio, obtém-se, por (24):

$$j\hat{\mu} + \hat{\underset{\sim}{a}} = \begin{bmatrix} 16,9838 \\ 40,0226 \\ 43,0914 \\ 48,5500 \\ 41,1503 \end{bmatrix} = R^{-1}\left[\underset{\sim}{T} - N\hat{\underset{\sim}{b}} - A\hat{\underset{\sim}{c}}\right]$$

Naturalmente,

$$\hat{\mu} = 16,9836 + \dots + 41,1504 = 37,9596$$

e os $\hat{a}_i$ são os mesmos apresentados anteriormente.

4.6.4. Matriz de dispersão para $\hat{\underline{a}}$

De acordo com (19), a matriz de dispersão para $\hat{\underline{a}}$ é:

$$D(\hat{\underline{a}}) = M^{*^{-1}} C^* M^{*^{-1}\prime} \sigma^2 .$$

Para $M^* = C^* - A^*$, sendo A^* dada por (61) e C^* por (60), tem-se:

$$D = \begin{bmatrix} 0,60036906 & 0,17005272 & -0,37616871 & -0,18152899 & -0,21272407 \\ 0,17005272 & 0,45619800 & -0,30516695 & -0,14640890 & -0,17467486 \\ -0,37616871 & -0,30516695 & 0,59120093 & 0,01650849 & 0,07362624 \\ -0,18152899 & -0,14640890 & 0,01650849 & 0,32919156 & -0,01776215 \\ -0,21272407 & -0,17467486 & 0,07362624 & -0,01776215 & 0,33153485 \end{bmatrix} \sigma^2$$

Assim, por exemplo, a variância do contraste

$$\hat{Y} = \hat{m}_4 - \hat{m}_1 = 48,5503 - 16,9836 = 31,5667$$

ou, ainda,

$$\hat{Y} = \hat{a}_4 - \hat{a}_1 = 31,5667, \text{ é:}$$

$$V(\hat{Y}) = (d_{11} + d_{44} - 2d_{14})\sigma^2$$

ou, $$\hat{V}(\hat{Y}) = (1,292618628)(73,8011) = 95,3967.$$

Vê-se, facilmente, que o contraste do tipo $a_i - a_j$ é estimável, pois o sistema

$$C^*\underline{\lambda} = \underline{k}$$

tem solução, já que 1,5460 + ... + (-0,2430) = 0 e 1 + 0 + ... - 1 + 0 = 0, ou seja, $\underset{\sim}{j}'\underset{-}{c}^* = 0$ e $\underset{\sim}{j}'\underset{-}{k} = 0$.

4.6.5. Matriz de dispersão para $\hat{\underline{c}}$

A matriz de dispersão para $\hat{\underline{c}}$, considerando (20), é:

$$D(\hat{\underline{c}}) = R_1^{-1}\sigma^2$$

Considerando os dados do exemplo:

$$D(\hat{\underline{c}}) = \begin{bmatrix} 0{,}108084359 & 0{,}006590510 \\ & \\ 0{,}006590510 & 0{,}046133568 \end{bmatrix}\sigma^2$$

5. CONCLUSÕES

a) Caso seja (2) o modelo adequado, a interpretação das hipóteses testadas é feita sem grandes dificuldades tanto na análise da variância como da covariância, mesmo que exista algum $n_{ij} = 0$. Assim,

$$F(H_1) = \frac{\hat{\underline{a}}'\underline{Q}^*/(v - 1)}{QMR^*}$$

testa a hipótese H_1: $a_i - a_{i'} = 0$, para todo $i \neq i'$,

$$F(H_2) = \frac{\hat{\underline{c}}'\underline{R}_2/p}{QMR^*}$$

testa a hipótese H_2: $\underline{c} = \underline{0}$

e, respeitadas as condições de estimabilidade e que:

$\underline{k}_i' M^{*-1} \underline{k}_j = 0$, para todo $i \neq j$, a razão

$$F(H_i) = \frac{S.Q.H_i}{QMR^*}$$

possibilita testar as sub-hipóteses H_i: $\underline{k}_i' \underline{a} = 0$, $i = 1, 2, \ldots, v-1$.

b) A presença da interação no modelo altera completamente as hipóteses, tanto na sua formulação, como na sua utilidade prática. Assim, a única hipótese de interesse para ser aquela que expressa a não significância da interação, já que as demais, além de difícil compreensão, dificilmente se aplicam na prática. Além disso, H_5 depende de n_{ij} para que a estimabilidade se verifique, o que torna a hipótese extremamente dependente das condições particulares de cada ensaio.

c) A formulação das hipóteses deve ser feita, de preferência, em termos de funções lineares estimáveis. Caso contrário, devem ser associadas à elas, as restrições não estimáveis que possibilitaram expressá-la como tal.

d) O Método do Ajustamento de Constantes tem como única desvantagem, além da maior dificuldade de cálculo, o fato de produzir somas de quadrados viciadas para os efeitos principais. Entretanto, mesmo nos casos de classificações balanceadas, onde tal fato não ocorre, o estudo da interação é que, realmente, tem maior interesse.

e) As hipóteses a serem testadas nos casos de classificações balanceadas são facilmente obtidas, bastando que se faça $n_{ij} = n$.

6. LITERATURA CITADA

DIAS, J.F.S., 1981. Análise de Covariância Intrablocos, com p Variáveis Auxiliares, para Delineamentos em Blocos Incompletos Equilibrados. Piracicaba, ESALQ/USP, 96 p. (Tese de Doutoramento).

FEDERER, R.A., 1957. Variance and covariance analyses for unbalanced classifications. Biometrics, Raleigh, 13: 333--362.

GRAYBILL, F.A., 1961. An Introduction to Linear Statistical Models. Nova York, McGraw-Hill Book Company, Inc. Vol. 1. 463 p.

HARVEY, W.R., 1975. Least Squares Analysis of Data with Unequal Subclass Numbers. ARS-20-8, U.S.D.A. Maryland.

HENDERSON, C.R., 1953. Estimation of variance and covariance components. Biometrics, Raleigh, 9:226-252.

HOCKING, R.R. e F.M. SPEED, 1975. A full rank analysis of some linear model problems. J. Amer. Stat. Assoc., Boston, 351:706-712.

KATTI, S.K., 1965. Multiple covariate analysis. Biometrics, Raleigh, 21:957-974.

PIMENTEL GOMES, F., 1968. The solution of normal equations of experiments in incomplete blocks. Ciência e Cultura, São Paulo, 20:733-746.

SEARLE, S.R., 1971. Linear Models. Nova York, John Wiley & Sons, Inc., 532 p.

SPEED, F.M., R.R. HOCKING e O.P. HACKNEY, 1978. Methods of analysis of linear models with unbalanced data. J. Amer. Stat. Assoc., Boston, 73:105-112.

YATES, F., 1934. The analysis of multiple classifications with unequal numbers in the different classes. J. Amer. Stat. Assoc., Boston, 29:52-66.

ZELEN, M., 1957. The analysis of covariance for incomplete block designs. Biometrics, Raleigh, 13:309-332.

A P Ê N D I C E

1. LISTAGEM DO PROGRAMA PARA AS ANÁLISES DA VARIÂNCIA E DA COVARIÂNCIA DE DADOS DE UMA CLASSIFICAÇÃO DUPLA.

```
// JOB     191A 191B 191C 191D 191E

LOG DRIVE    CART SPEC    CART AVAIL   PHY DRIVE
  0000         191A          191A         0001
  0001         191B          191B         0002
  0002         191C          191C         0003
  0003         191D          191D         0004
  0004         191E          191E         0005

V2 M11    ACTUAL 16K   CONFIG 16K

*EQUAT(PRNTZ,PRNZ)

// FOR
*LIST SOURCE PROGRAM
*ONE WORD INTEGERS
*EXTENDED PRECISION
*IOCS(CARD,1132PRINTER,DISK)
C     *********************************************************
C     *                                                       *
C     * ANALISES DA VARIANCIA E DA COVARIANCIA LINEAR DE      *
C     * DADOS DE UMA CLASSIFICACAO DUPLA NAO BALANCEADA       *
C     *                                                       *
C     *********************************************************
C     XNOME=TITULO DA ANALISE
C     NTA=NUMERO DE TRATAMENTOS A
C     NTB=NUMERO DE TRATAMENTOS B(OU NUMERO DE REPETICOES)
C     NX=NUMERO DE VARIAVEIS X
C     IA=CODIGO DE TRANSFORMACAO DA VARIAVEL Y DE ACORDO COM A
C        SUBROTINA TRANS
C     ALFA=CONSTANTE DE TRANSFORMACAO
C     N(I,J)=NUMERO DE OBSERVACOES DOS TRATAMENTOS I E J
C     Y(I,J)=VALORES DAS VARIAVEIS Y ESGOTANDO TODAS AS OBSERVACOES DENTRO
C           DOS TRATAMENTOS B
C     NAB=0 - QUANDO O FATOR B REPRESENTA BLOCOS AO ACASO
C     NAB=1 - FATORIAL A X B EM DELINEAMENTO INTEIRAMENTE AO ACASO
C     NTB=1, NAB=0 - DELINEAMENTO INTEIRAMENTE AO ACASO
C     INT=0 - NAO CALCULA A INTERACAO
C     INT=1 - CALCULA A INTERACAO
```

```
      REAL M(10,10),MI(10,10)
      DIMENSION XNOME(20),N(10,12),NJ(12),P(5),B(12),D(12,5),AT(3),X(5,5
     1),NI(10),YY(50),XX(5,50),A(10,5),T(10),Z(10,5),W(5,5),S(5),LX(10),
     2MX(10),C(10,10),ZW(10,5),YMED(10),Q(10),TAL(10),WI(5,5),DA(10,10),
     3BETA(5),QQ(10),TALAS(10),CAV(19),R2A(5),R1A(5,5),CAST(5),XM(5)
      DATA CAV/'BLOCOS','NAD ','AJUST.','TRATS.','P/','REGR.','BL.','E '
     1,'RESIDU','O ','AJ.','A ','B ',' ',' ',' ','REGRL','SSA','
     2O '/
      DATA AT/'NS','* ','**'/
      DEFINE FILE 8(35,69,U,MV)
      DEFINE FILE 9(35,69,U,MV)
      DEFINE FILE 10(35,63,U,KL)
    1 READ(2,2)(XNOME(J),J=1,20)
    2 FORMAT(20A4)
      READ(2,3)NTA,NTB,NX,IA,ALFA,NAR,INT
    3 FORMAT(4I5,F5.0,2I5)
      DO 4 I=1,NTA
    4 READ(2,5)(N(I,J),J=1,NTB)
    5 FORMAT(16I5)
      WRITE(3,6)(XNOME(J),J=1,20)
    6 FORMAT('1',T21,20A4)
      WRITE(3,7)
    7 FORMAT(T17,86('-'))
      WRITE(3,8)NTA,NTB,NX,IA,ALFA,NAR,INT
    8 FORMAT(T21,4I5,F5.1,2I5)
      DO 9 I=1,NTA
    9 WRITE(3,10)(N(I,J),J=1,NTB)
   10 FORMAT(T21,16I5)
      IF(NX)78,78,79
   78 NY=0
      NX=1
      GO TO 121
   79 NY=NX
  121 DO 11 I=1,NTB
      NJ(I)=0
      B(I)=0.
      DO 11 J=1,NX
      R2A(J)=0.
      P(J)=0.
      CAST(J)=0.
      D(I,J)=0.
      XM(J)=0.
      DO 11 K=1,NX
      R1A(J,K)=0.
   11 X(J,K)=0.
      DO 19 I=1,NTA
      DO 19 J=1,NX
   19 A(I,J)=0.
      SQINT=0.
      G=0.
      KV=1
```

```
      SOTOT=0
      NC=0
      DO 21 I=1,NTA
      KK=0
      DO 12 J=1,NTB
      KK=KK+N(I,J)
   12 NJ(J)=NJ(J)+N(I,J)
      NI(I)=KK
      READ(2,13)(YY(K),K=1,KK)
   13 FORMAT(8F10.0)
      IF(NY)74,74,75
   75 DO 14 J=1,NX
   14 READ(2,13)(XX(J,K),K=1,KK)
      GO TO 180
   74 DO 80 JJ=1,KK
   80 XX(1,JJ)=0.
  180 WRITE(3,15)(YY(K),K=1,KK)
   15 FORMAT(T17,'Y',8F10.3)
      IF(IA)182,182,181
  181 CALL TRANS(YY,KK,IA,ALFA)
  182 IF(NY)76,76,77
   77 DO 16 J=1,NX
      KT=1
      KV=8
      IF(KK-KV)183,184,184
  183 KV=KK
  184 WRITE(3,17)J,(XX(J,K),K=KT,KV)
   17 FORMAT(T17,'X',I1,8F10.3)
      IF(KK-KV)16,16,185
  185 KT=KT+8
      KV=KV+8
      IF(KK-KV)183,184,184
   16 CONTINUE
   76 TT=0.
      DO 18 J=1,KK
      TT=TT+YY(J)
   18 SOTOT=SOTOT+YY(J)**2
      T(I)=TT
      G=G+TT
      KT=0
      DO 21 J=1,NTB
      DO 275 JJ=1,NX
  275 QO(JJ)=0.
      KK=N(I,J)
      IF(KK)21,21,190
  190 TT=0.
      NC=NC+1
      DO 81 K=1,KK
      KT=KT+1
      B(J)=B(J)+YY(KT)
      TT=TT+YY(KT)
```

```
      DO 20 L=1,NX
      QQ(L)=QQ(L)+XX(L,KT)
      XM(L)=XM(L)+XX(L,KT)
      P(L)=P(L)+YY(KT)*XX(L,KT)
      A(I,L)=A(I,L)+XX(L,KT)
      D(J,L)=D(J,L)+XX(L,KT)
      DO 20 MM=1,NX
   20 X(L,MM)=X(L,MM)+XX(L,KT)*XX(MM,KT)
   81 CONTINUE
      IF(INT)21,21,257
  257 DO 251 ML=1,NX
      R2A(ML)=R2A(ML)+QQ(ML)*TT/FLOAT(KK)
      DO 251 MM=1,NX
  251 R1A(ML,MM)=R1A(ML,MM)+QQ(ML)*QQ(MM)/FLOAT(KK)
      SQINT=SQINT+TT**2/KK
   21 CONTINUE
      SQIN=SQINT
      IF(INT)259,259,258
  258 DO 252 I=1,NX
      R2A(I)=P(I)-R2A(I)
      DO 252 J=1,NX
  252 R1A(I,J)=X(I,J)-R1A(I,J)
      KK=0
      DO 253 I=1,NX
      DO 253 J=1,NX
      KK=KK+1
  253 YY(KK)=R1A(I,J)
      CALL MINV(YY,NX,DELT,LX,MX)
      KK=0
      DO 254 I=1,NX
      DO 254 J=1,NX
      KK=KK+1
  254 R1A(I,J)=YY(KK)
      DO 255 I=1,NX
      DO 255 J=1,NX
  255 CAST(I)=CAST(I)+R1A(I,J)*R2A(J)
      SQREI=0.
      DO 256 I=1,NX
  256 SQREI=SQREI+CAST(I)*R2A(I)
      QMREI=SQREI/FLOAT(NX)
  259 NRO=0
      NTOT=0
      DO 22 I=1,NTA
   22 NTOT=NTOT+NI(I)
      CC=G**2/NTOT
      SQTOT=SQTOT-CC
      WRITE(3,7)
      DO 26 I=1,NX
      P(I)=P(I)-XM(I)*G/NTOT
      DO 23 J=1,NX
   23 X(I,J)=X(I,J)-XM(I)*XM(J)/NTOT
```

```
      DO 24 K=1,NTA
   24 A(K,I)=A(K,I)-NI(K)*XM(I)/NTOT
      DO 25 L=1,NTB
   25 D(L,I)=D(L,I)-NJ(L)*XM(I)/NTOT
   26 CONTINUE
  143 DO 28 I=1,NTA
      DO 28 J=1,NX
      TT=0.
      DO 27 K=1,NTB
   27 TT=TT+N(I,K)*D(K,J)/NJ(K)
   28 Z(I,J)=A(I,J)-TT
      DO 30 I=1,NX
      DO 30 J=1,NX
      TT=0.
      DO 29 K=1,NTB
   29 TT=TT+D(K,I)*D(K,J)/NJ(K)
   30 W(I,J)=X(I,J)-TT
      DO 32 I=1,NX
      TT=0.
      DO 31 J=1,NTB
   31 TT=TT+D(J,I)*B(J)/NJ(J)
   32 S(I)=P(I)-TT
      KK=0.
      DO 33 I=1,NX
      DO 33 J=1,NX
      KK=KK+1
   33 YY(KK)=W(I,J)
      CALL MINV(YY,NX,DELT,LX,MX)
      KK=0.
      DO 34 I=1,NX
      DO 34 J=1,NX
      KK=KK+1
   34 WI(I,J)=YY(KK)
      DO 38 I=1,NTA
      DO 38 J=1,NTA
      TT=0.
      DO 35 K=1,NTB
   35 TT=TT+FLOAT(N(I,K)*N(J,K))/FLOAT(NJ(K))
      IF(I-J)36,37,36
   36 C(I,J)=-TT
      GO TO 38
   37 C(I,J)=NI(I)-TT
   38 CONTINUF
      DO 40 I=1,NTA
      DO 40 J=1,NX
      TT=0.
      DO 39 K=1,NX
   39 TT=TT+Z(I,K)*WI(K,J)
   40 ZW(I,J)=TT
      DO 42 I=1,NTA
      TT=0.
```

```
      DO 58 I=1,NTA
      DO 58 J=1,NTA
      KK=KK+1
   58 M(I,J)=YY(KK)
      DO 60 I=1,NTA
      TT=0.
      DO 59 J=1,NTA
   59 TT=TT+M(I,J)*QQ(J)
   60 TALAS(I)=TT
      SQTBR=0.
      DO 61 I=1,NTA
   61 SQTBR=SQTBR+TALAS(I)*QQ(I)
      DO 63 I=1,NX
      DO 63 J=1,NTA
      TT=0.
      DO 62 K=1,NTA
   62 TT=TT+Z(K,I)*MI(K,J)
   63 ZW(J,I)=TT
      DO 65 I=1,NX
      TT=0.
      DO 64 J=1,NTA
   64 TT=TT+ZW(J,I)*Q(J)
   65 S(I)=S(I)-TT
      DO 67 I=1,NX
      DO 67 J=1,NX
      TT=0.
      DO 66 K=1,NTA
   66 TT=TT+ZW(K,I)*Z(K,J)
   67 W(I,J)=W(I,J)-TT
      KK=0
      DO 68 I=1,NX
      DO 68 J=1,NX
      KK=KK+1
   68 YY(KK)=W(I,J)
      CALL MINV(YY,NX,DELT,LX,MX)
      KK=0
      DO 69 I=1,NX
      DO 69 J=1,NX
      KK=KK+1
   69 W(I,J)=YY(KK)
      DO 71 I=1,NX
      TT=0.
      DO 70 J=1,NX
   70 TT=TT+W(I,J)*S(J)
   71 BETA(I)=TT
      NGLTA=NTA-1
      NGLTO=NTOT-1
      SOBLO=0.
      IF(NTB-1)196,196,72
   72 DO 73 I=1,NTB
   73 SQBLO=SQBLO+B(I)**2/NJ(I)
      SQBLO=SQBLO-CC
  196 NGLTB=NTB-1
```

```
      NGLIN=0
      SQREG=0.
      NGLRE=0
      IF(NY)197,197,82
   82 NGLRE=NX
      DO 83 I=1,NX
   83 SQREG=SQREG+BETA(I)*S(I)
      QMREG=SQREG/NX
  197 IF(INT)286,286,84
   84 NGLIN=NC-NTA-NTB+1
      IF(NRO)240,240,85
  240 SQINT=SQINT-SQTAB-SQBLO-CC+SQREI-SQREG
      QMINT=SQINT/NGLIN
      GO TO 85
  286 SQINT=0.
   85 IF(INT)260,260,261
  261 NGRAR=NTOT-NC-NY
      IF(NY)270,270,271
  271 SQRAR=SQTOT-SQIN-SQREI+CC
      GO TO 264
  260 NGRAR=NGLTO-NGLTA-NGLTB-NGLIN-NGLRE
  270 SQRAR=SQTOT-SQBLO-SQTAB-SQINT-SQREG
  264 QMRAR=SQRAR/NGRAR
      QMTBR=SQTBR/NGLTA
      QMT=SQTAB/NGLTA
      DO 155 I=1,NTA
      DO 155 J=1,NTA
      TT=0.
      DO 154 K=1,NTA
  154 TT=TT+C(I,K)*M(J,K)
  155 MI(I,J)=TT
      DO 157 I=1,NTA
      DO 157 J=1,NTA
      TT=0.
      DO 156 K=1,NTA
  156 TT=TT+M(I,K)*MI(K,J)
  157 DA(I,J)=TT
      CALL CABEC(TT)
      K1=12+NRO
      K2=13-NRO
      IF(NTB-1)86,86,87
   87 IF(NAB)88,88,90
   88 WRITE(3,89)NGLTB,SQBLO
   89 FORMAT(/,T19,'BLOCOS NAO AJUST.',T44,I2,F19.4)
      GO TO 86
   90 IF(NRO)91,91,93
   91 WRITE(3,92)CAV(4),CAV(13),CAV(2),CAV(3),NGLTB,SQBLO
   92 FORMAT(/,T19,A6,A2,A4,A6,T44,I2,F19.4)
      GO TO 86
   93 WRITE(3,92)CAV(4),CAV(12),CAV(2),CAV(3),NGLTB,SQBLO
   86 IF(NY)95,95,97
   95 FC=QMT/QMRAR
```

```
      KK=TEFE(NGLTA,NGRAR,FC)
      IF(NTB-1)96,96,102
   96 WRITE(3,94)CAV(4),CAV(3),CAV(14),CAV(15),NGLTA,SQTAB,QMT,FC,AT(KK)
   94 FORMAT(/,T19,A6,A6,A2,A6,T44,I2,F19.4,F17.4,T87,F11.3,A2)
      GO TO 106
  102 IF(NAB)98,98,99
   98 WRITE(3,94)CAV(4),CAV(3),CAV(5),CAV(1),NGLTA,SQTAB,QMT,FC,AT(KK)
      GO TO 106
   99 WRITE(3,100)CAV(4),CAV(K1),CAV(3),CAV(5),CAV(K2),NGLTA,SQTAB,QMT,F
     1C,AT(KK)
  100 FORMAT(/,T19,A6,A2,A6,A2,A2,T44,I2,F19.4,F17.4,T87,F11.3,A2)
      GO TO 106
   97 IF(NTB-1)101,101,103
  101 WRITE(3,94)CAV(4),CAV(3),CAV(14),CAV(15),NGLTA,SQTAB
      GO TO 106
  103 IF(NAB)104,104,105
  104 WRITE(3,94)CAV(4),CAV(3),CAV(5),CAV(1),NGLTA,SQTAB
      GO TO 106
  105 WRITE(3,100)CAV(4),CAV(K1),CAV(3),CAV(5),CAV(K2),NGLTA,SQTAB
  106 IF(INT)109,109,107
  107 FC=QMINT/QMRAR
      KK=TEFE(NGLIN,NGRAR,FC)
      WRITE(3,108)NGLIN,SQINT,QMINT,FC,AT(KK)
  108 FORMAT(/,T19,'INTERACAO A X B',T44,I2,F19.4,F17.4,T87,F11.3,A2,)
  109 IF(NY)116,116,115
  115 IF(INT)262,262,263
  263 SQREG=SQREI
      QMREG=QMREI
  262 FC=QMREG/QMRAR
      KK=TEFE(NGLRE,NGRAR,FC)
      WRITE(3,111)CAV(17),CAV(18),CAV(19),CAV(14),CAV(14),CAV(14),CAV(14
     1),NGLRE,SQREG,QMREG,FC,AT(KK)
  111 FORMAT(/,T19,A5,A3,A2,A6,A2,A3,A2,T44,I2,F19.4,F17.4,T87,F11.3,A2)
      GO TO 228
  116 WRITE(3,227)CAV(9),CAV(10),CAV(16),CAV(14),CAV(15),NGRAR,SQRAR,QMR
     1AR
  227 FORMAT(/,T19,A6,A2,A3,A2,A6,T44,I2,F19.4,F17.4,/)
      GO TO 229
  228 WRITE(3,227)CAV(9),CAV(10),CAV(11),CAV(5),CAV(6),NGRAR,SQRAR,QMRAR
  229 WRITE(3,7)
      IF(NY)236,236,230
  230 FC=QMTBR/QMRAR
      KK=TEFE(NGLTA,NGRAR,FC)
      IF(NTB-1)238,238,239
  238 WRITE(3,94)CAV(4),CAV(3),CAV(5),CAV(6),NGLTA,SQTBR,QMTBR,FC,AT(KK)
      GO TO 235
  239 IF(NAB)231,231,233
  231 WRITE(3,232)CAV(4),CAV(11),CAV(5),CAV(7),CAV(8),CAV(6),NGLTA,SQTBR
     1,QMTBR,FC,AT(KK)
  232 FORMAT(/,T19,A6,A3,A2,A3,A2,A6,T44,I2,F19.4,F17.4,T87,F11.3,A2)
      GO TO 235
  233 WRITE(3,234)CAV(4),CAV(K1),CAV(11),CAV(5),CAV(K2),CAV(8),CAV(6),NG
```

```
     1LTA,SQTBR,QMTBR,FC,AT(KK)
234 FORMAT(/,T19,A6,A2,A3,A2,A2,A2,A6,T44,I2,F19.4,F17.4,I87,F11.3,A2)
235 WRITE(3,7)
236 WRITE(3,237)NGLTO,SQTOT
237 FORMAT(T19,'T O T A L',T43,I3,F19.4)
    WRITE(3,7)
    DO 160 I=1,NTB
    TT=0.
    DO 161 J=1,NX
161 TT=TT+D(I,J)*BETA(J)
160 YMED(I)=TT
    DO 119 I=1,NTB
    TT=0.
    DO 118 J=1,NTA
118 TT=TT+N(J,I)*TALAS(J)
119 YMED(I)=(B(I)-TT-YMED(I))/NJ(I)
    XMI=0.
    DO 120 I=1,NTB
120 XMI=XMI+YMED(I)
    XMI=XMI/NTB
    DO 250 I=1,NTA
250 YMED(I)=XMI+TALAS(I)
    WRITE(3,127)XMI
127 FORMAT(//,T21,'MEDIA GERAL = ',F12.6)
    IF(NAB)212,212,213
212 KK=14
    GO TO 215
213 KK=12+NRO
215 WRITE(3,214)CAV(KK)
214 FORMAT(//,T34,'TESTE DE DUNCAN PARA MEDIAS AJUSTADAS DE TRATAMENTO
   1S ',A2)
    DO 249 I=1,NTA
249 LX(I)=NI(I)
    CALL DDED6(YMED,LX,NTA,IA,ALFA,NGRAR,QMRAR,DA)
    WRITE(3,128)
128 FORMAT(//,T41,'M A T R I Z    D E    D I S P E R S A O')
    WRITE(3,7)
    DO 166 I=1,NTA
    KT=1
    KV=4
125 IF(NTA-KV)162,162,163
162 KV=NTA
163 WRITE(3,164)I,(DA(I,J),J=KT,KV)
164 FORMAT(T17,'LINHA',I3,4F18.9)
    IF(NTA-KV)166,166,165
165 KT=KT+4
    KV=KV+4
    GO TO 125
166 CONTINUE
    WRITE(3,7)
    IF(NY)216,216,217
217 WRITE(3,218)
```

```
218 FORMAT(//,T27,'M A T R I Z   D E   D I S P E R S Ã O   D E   R
   1 E G R E S S A O')
    WRITE(3,7)
    DO 219 I=1,NX
    KT=1
    KV=4
223 IF(NX-KV)220,220,221
220 KV=NX
221 WRITE(3,164)I,(W(I,J),J=KT,KV)
    IF(NX-KV)219,219,222
222 KT=KT+4
    KV=KV+4
    GO TO 223
219 CONTINUE
    WRITE(3,7)
216 IF(NRO)131,131,142
131 IF(NAB)142,142,132
132 NRO=1
    DO 133 I=1,NTA
    LX(I)=NI(I)
133 YY(I)=T(I)
    DO 134 I=1,NTB
    NI(I)=NJ(I)
134 T(I)=B(I)
    DO 135 I=1,NTA
    NJ(I)=LX(I)
135 B(I)=YY(I)
    DO 139 I=1,NX
    DO 136 J=1,NTA
136 YY(J)=A(J,I)
    DO 137 J=1,NTB
137 A(J,I)=D(J,I)
    DO 138 J=1,NTA
138 D(J,I)=YY(J)
139 CONTINUE
    IF(NTA-NTB)246,246,247
246 KV=NTB
    GO TO 248
247 KV=NTA
248 DO 140 I=1,KV
    DO 140 J=1,KV
    IF(I-J)245,140,140
245 KT=N(I,J)
    N(I,J)=N(J,I)
    N(J,I)=KT
140 CONTINUE
    DO 141 I=1,NTA
141 LX(I)=0.
    KK=NTA
    NTA=NTB
    NTB=KK
    KV=1
```

```
      GO TO 143
  142 GO TO 1
      END

.FEATURES SUPPORTED
 ONE WORD INTEGERS
 EXTENDED PRECISION
 IOCS

CORE REQUIREMENTS FOR
 COMMON        0  VARIABLES     3758  PROGRAM     5564

END OF COMPILATION

// XEQ           3

*FILES(10,DUNC)

*FILES(8,TABF1)

*FILES(9,TABF5)
```

2. SUBROTINAS E FUNÇÕES UTILIZADAS PELO PROGRAMA

- Subroutine MINV - subrotina para a inversão de matrizes.
- Subroutine CABEC - subrotina que escreve o cabeçário da análise da variância.
- Subroutine ODED6 - subrotina que efetua o teste de Duncan.
- Function TEFE - função que executa o teste F.

3. EXEMPLO, ATRAVÉS DO PROGRAMA, DA ANÁLISE DE COVARIÂNCIA DOS DADOS DA TABELA 5.

```
      EXEMPLO DE APLICACAO DA ANALISE DE COVARIANCIA   MODELO SEM INTERACAO
-------------------------------------------------------------------------------
          5      2      2      0   0.0      1      0
          2      1
          1      2
          1      2
          1      2
          1      2
Y      27.000     18.000     19.000
X1      5.000      5.000      4.000
X2      4.000      4.000      3.000
Y      31.000     52.000     57.000
X1      3.000      5.000      5.000
X2      2.000      4.000      4.000
Y      34.000     52.000     33.000
X1      2.000      1.000      1.000
X2      1.000      5.000      3.000
Y      38.000     60.000     45.000
X1      4.000      2.000      2.000
X2      3.000      2.000      1.000
Y      31.000     50.000     40.000
X1      1.000      3.000      1.000
X2      5.000      1.000      2.000
-------------------------------------------------------------------------------
```

QUADRO DA ANALISE DA VARIANCIA

CAUSAS DA VARIACAO	GL	S.Q.	Q.M.	F
TRATS.B NAO AJUST.	1	864.8999		
TRATS.A AJUST.P/B	4	912.7000		
REGRESSAO	2	141.5251	70.7625	0.958NS
RESIDUO AJ.P/REGR.	7	516.6082	73.8011	
TRATS.A AJ.P/B E REGR.	4	1045.2786	26[illegible].[illegible]696	3.523NS
T O T A L	14	2435.7333		

MEDIA GERAL = 37.959651

TESTE DE DUNCAN PARA MEDIAS AJUSTADAS DE TRATAMENTOS A
NIVEL DE 5 POR CENTO DE PROBABILIDADE

NUM.DE ORDEM	NUM.DOS TRATS.	NUM.DE REPET.	MEDIAS	MEDIAS ORIG.
1	4	3	48.550014	48.550014
2	3	3	43.091461	43.091461
3	5	3	41.150335	41.150336
4	2	3	40.022627	40.022627
5	1	3	16.983816	16.983816

OS TRATAMENTOS DE NUMERO 4 A 1 NAO DIFEREM ENTRE SI

OS TRATAMENTOS DE NUMERO 5 A 2 NAO DIFEREM ENTRE SI

M A T R I Z   D E   D I S P E R S A O

LINHA 1	0.600369967	0.170052722	-0.376168715	-0.181528998
LINHA 1	-0.212724076			
LINHA 2	0.170052723	0.456198007	-0.305166959	-0.146408904
LINHA 2	-0.174674867			
LINHA 3	-0.376168715	-0.305166958	0.591200936	0.016508494
LINHA 3	0.073626244			
LINHA 4	-0.181528998	-0.146408904	0.016508494	0.329191554
LINHA 4	-0.017762155			
LINHA 5	-0.212724077	-0.174674867	0.073626244	-0.017762155
LINHA 5	0.331534855			

M A T R I Z   D E   D I S P E R S A O   D E   R E G R E S S A O

LINHA 1	0.108084358	0.006590509
LINHA 2	0.006590509	0.046133567

QUADRO DA ANALISE DA VARIANCIA

CAUSAS DA VARIACAO	GL	S.Q.	Q.M.	F
TRATS.A NAO AJUST.	4	1344.3999		
TRATS.B AJUST.P/A	1	433.2000		
REGRESSAO	2	141.5251	70.7625	0.958NS
RESIDUO AJ.P/REGR.	7	516.6082	73.8011	
TRATS.B AJ.P/A E REGR.	1	456.3692	456.3692	6.183*
T O T A L	14	2435.7333		

MEDIA GERAL = 37.959651

TESTE DE DUNCAN PARA MEDIAS AJUSTADAS DE TRATAMENTOS B
NIVEL DE 5 POR CENTO DE PROBABILIDADE

NUM.DE ORDEM	NUM.DOS TRATS.	NUM.DE REPET.	MEDIAS	MEDIAS ORIG.
1	2	9	43.828060	43.828060
2	1	6	32.091241	32.091241

TODOS OS TRATAMENTOS DIFEREM ENTRE SI

M A T R I Z   D E   D I S P E R S A O

LINHA 1	0.075461335	-0.075461335
LINHA 2	-0.075461335	0.075461335

M A T R I Z   D E   D I S P E R S A O   D E   R E G R E S S A O

LINHA 1	0.108084358	0.006590509
LINHA 2	0.006590509	0.046133567