UTILIZACAO DE "SOFTWARES" ESTATISTICOS NA INTERPRETACAO
DE HIPOTHESES COM DADOS DESBALANCEADOS

EUFRÁZIO DE SOUZA SANTOS
Licenciatura em Matemática

Orientador: Prof.Dr. Antonio Francisco Iemma

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

P I R A C I C A B A
Estado de São Paulo - Brasil
Março - 1994

Ficha catalográfica preparada pela Seção de Livros da
Divisão de Biblioteca e Documentação - PCLQ/USP

Santos, Eufrázio de Souza
S237u Utilização de "softwares" estatísticos na interpretação de hipóteses com dados desbalanceados. Piracicaba, 1994.
175p. ilustr.

Tese - ESALQ
Bibliografia.

1. Dado estatístico 2. Estatística - Informática
3. Informática aplicada 4. Modelo matemático I. Escola Superior de Agricultura Luiz de Queiroz, Piracicaba

CDD 519.50285

UTILIZACAO DE "SOFTWARES" ESTATISTICOS NA INTERPRETACAO
DE HIPOTHESES COM DADOS DESBALANCEADOS

EUFRÁZIO DE SOUZA SANTOS

APROVADA EM: 13/04/94

COMISSÃO JULGADORA :

PROF. DR. ANTONIO FRANCISCO IEMMA	USP\PIRACICABA
PROF. ^a DR. ^a MARIA CRISTINA STOLF NOGUEIRA	USP\PIRACICABA
PROF. DR. JOSÉ EDUARDO CORRENTE	USP\PIRACICABA
PROF. DR. LÚCIO BENEDITO KROLL	UNESP\BOTUCATU
PROF. DR. FERNANDO FERRARI	UNESP\S.J.R.P.

Prof. Dr. Antonio Francisco Iemma
Orientador

A meus pais, José e Helena, *in memoriam*

OFEREÇO

À minha esposa, Elza e ao
nosso filho, Érico

DEDICO

AGRADECIMENTOS

- A Deus, pela vida.
- Ao Professor Dr. Antonio Francisco Iemma, pela orientação eficiente, séria e honesta; pela confiança e amizade, o que contribuiu de forma decisiva para realização e conclusão deste trabalho.
- Ao Professor Dr. Décio Barbin pela compreensão e apoio sempre presentes.
- À Universidade Federal de Pernambuco (UFPE) e à Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ/USP), pela oportunidade.
- À CAPES/PICD, pela bolsa concedida.
- Aos Professores do Departamento de Estatística da UFPE, pelo apoio e amizade.
- Aos professores e funcionários do Departamento de Matemática e Estatística da ESALQ pelos ensinamentos e pela amizade.

- Aos colegas de curso, Ana, Amauri, Augusto, Carlos, Cecília, César, Creusa, Gener, Gilmar, Guidoni, Iza, João Gil, Joel, Maurício, Paulo Cecon, Pilar, Rosa, Rosana Rosimére (*in memoriam*), Rosely, Rui Verneck, Sônia, Tadeu e Tiemi pelo apoio e incentivo.
- À minha esposa Elza e ao nosso filho Érico pela compreensão, apoio e incentivo.
- Finalmente, a todos que, direta ou indiretamente, contribuíram para a realização e conclusão deste trabalho.

SUMÁRIO

	Página
RESUMO.....	ix
SUMMARY.....	xi
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Introdução	3
2.2. O modelo linear	8
2.2.1. Estimação	9
2.2.2. Formas do Modelo Linear	11
2.2.2.1. O Modelo Superparametrizado..	11
2.2.2.2. O Modelo de Médias de Caselas	13
2.2.2.3. O Modelo com restrições para-	
métricas do tipo Σ	14
3. METODOLOGIA	16
3.1. Introdução	16
3.2. Hipóteses e Somas de Quadrados Associadas ...	18
3.2.1. Todas as caselas ocupadas.....	20
3.2.1.1. O modelo linear com interação	20
3.2.1.2. O modelo linear sem interação	23
3.2.2. Algumas caselas vazias	27
3.2.2.1. Conectividade.....	28
3.2.2.2. Hipóteses sobre médias não	
ponderadas de linhas, colunas	
e interações.....	29
3.2.3. Resumo das hipóteses	32
3.3. Métodos computacionais.....	34
3.3.1. Introdução	34
3.3.2. O método das médias não ponderadas....	34
3.3.3. O método dos quadrados de médias pon-	
deradas	36
3.3.4. O método do ajustamento de constantes.	37
3.3.5. Os métodos de Henderson	38

3.3.5.1.	Método 1 de Henderson (H1)...	39
3.3.5.2.	Método 3 de Henderson (H3)...	39
3.3.6.	Os métodos de Overall e Spiegel (1969)	40
3.3.6.1.	Método dos mínimos quadrados completo (OS1).....	40
3.3.6.2.	Método do planejamento experimental (OS2).....	40
3.3.6.3.	Método da ordenação a priori, (OS3).....	41
3.3.7.	Método Iterativo de Hermmerle.....	41
3.4.	Performance de "softwares" (pacotes) com relação a dados desbalanceados.....	41
3.4.1.	GLM DO SAS. (1986).....	42
3.4.2.	SOC (SISTEMA OPERACIONAL C) (1982)....	48
3.4.3.	SYSTAT (THE SYSTEM FOR STATISTICS).... (1986).....	50
3.4.4.	STATGRAPHICS (STATISTICAL GRAPHICS... SYSTEM) (1985).....	50
3.4.5.	SAEG (SISTEMA PARA ANÁLISES ESTATÍSTICAS) 91986).....	51
3.4.6.	MSTAT (A MICROCOMPUTER FOR THE DESIGN, MANAGEMENT AND ANALYSIS OF AGRONOMIC.. RESEARCH EXPERIMENTS) (1984).....	51
3.4.7.	MINITAB (MINITAB STATISTICAL..... SOFTWARES) (1986).....	52
3.4.8.	BMDP (BIOMEDICAL STATISTIC PACKAGE)... (1983).....	52
3.4.9.	HARVEY (MIXED MODEL LEAST-SQUARES AND. MAXIMUM LIKELIHOOD COMPUTER PR) (1987)	55
3.4.10.	SPSS (STATISTICAL PACKAGES) (1986)....	56
3.4.11.	GLIM (GENERALISED LINEAR INTERACTIVE.. MODELING SYSTEM) (1984).....	57
3.4.12.	SANEST (SISTEMA DE ANÁLISE ESTATÍSTICA) (1986).....	57
4.	EXEMPLO NUMÉRICO	59

4.1. O Modelo Superparametrizado (S).....	61
4.1.1. Com interação na ordem A - B - A*B.....	61
4.1.1.1. Equações normais e suas solu- ções.....	62
4.1.1.2. Somas de quadrados (S.Q).....	64
4.1.1.3. Análise de Variância.....	65
4.1.2. Sem interação na ordem A - B.....	67
4.1.2.1. Equações normais e suas solu- ções.....	67
4.1.2.2. Somas de quadrados.....	69
4.1.2.3. Análise de Variância.....	70
4.2. O Modelo de Médias de Caselas.....	71
4.2.1. Com interação na ordem A - B - A*B....	71
4.2.1.1. Equações normais sua solução.	72
4.2.1.2. Somas de quadrados.....	74
4.2.2. Sem interação na ordem A - B.....	75
4.2.2.1. Equações normais e suas solu- ções.....	76
4.2.2.2. Somas de quadrados.....	77
4.3. O Modelo - Σ	77
4.3.1. Com interação na ordem A - B - A*B....	77
4.3.1.1. Equações normais e suas so- luções.....	78
4.3.1.2. Somas de quadrados.....	79
4.4. Hipóteses.....	82
5. RESULTADOS E DISCUSSÃO.....	107
6. CONCLUSÕES.....	124
6.1. Para modelos com interação.....	124
6.1.1. Dados balanceados	124
6.1.2. Dados não balanceados	124
6.2. Para modelos sem interação	126
6.2.1. Dados balanceados	126
6.2.2. Dados não balanceados	126
7. REFERENCIAS BIBLIOGRÁFICAS.....	129

8.	APENDICES		138
8.1	APÊNCICDE.1	GLIM.	139
8.2	APÊNCICDE.2	HARVEY.	141
8.3	APÊNCICDE.3	MINITAB.	143
8.4	APÊNCICDE.4	MSTAT.	146
8.5	APÊNCICDE.5	SAEG.	150
8.6	APÊNCICDE.6	SANEST.	156
8.7	APÊNCICDE.7	SAS.	160
8.8	APÊNCICDE.8	SOC.	165
8.9	APÊNCICDE.9	SPSS.	170
8.10	APÊNCICDE.10	STATGRAPHICS.	174
8.11	APÊNCICDE.11	SYSTAT.	175

UTILIZACAO DE "SOFTWARES" ESTATISTICOS NA INTERPRETACAO DE HIPOTHESES COM DADOS DESBALANCEADOS

Autor: Eufrázio de Souza Santos

Orientador: Prof. Dr. Antonio Francisco Iemma

RESUMO

Este trabalho teve como objetivo a estruturação de hipóteses e somas de quadrados a elas associadas, com dados desbalanceados, visando orientar a utilização de alguns pacotes estatísticos por usuários não iniciados em estatística.

Para tanto, foram revistos métodos de análise de dados desbalanceados, e adotado o modelo estatístico: $y = X\theta + e$, com e sem interação, caracterizado respectivamente por:

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

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

para $i = 1, \dots, a$ linhas; $j = 1, \dots, b$ colunas; e n_{ij} elementos em cada casela, onde

- y_{ijk} : valor observado na k-ésima repetição do i-ésimo nível do fator A e j-ésimo nível do fator B.
 μ : constante inerente a todas as observações
 α_i : efeito do i-ésimo nível do fator A;
 β_j : efeito do j-ésimo nível do fator B;
 γ_{ijk} : efeito da interação entre o i-ésimo nível do fator A e o j-ésimo nível do fator B.
 e_{ijk} : erro aleatório atribuído à observação y_{ijk} .

No desenvolvimento da metodologia, foram estabelecidas as expressões analíticas correspondentes às hipóteses e somas de quadrados associadas, sobre linhas, colunas e interações, bem como, sobre os métodos computacionais envolvidos nas mesmas. Ainda nesse tópico, foram descritos doze pacotes estatísticos, e suas principais características relativamente a tais hipóteses e somas de quadrados foram apresentadas..

Através de exemplo numérico, mostrou-se o desempenho desses pacotes no trato com dados desbalanceados, com algumas caselas vazias, expondo-se suas hipóteses e somas de quadrados associadas; e apresentou-se uma comparação entre os resultados por eles fornecidos.

USING STATISTICAL SOFTWARES TO UNDERSTAND HYPOTHESIS-TESTING WITH UNBALANCED DATA

Author: Eufrázio de Souza Santos

Adviser: Prof. Dr. Antonio Francisco Iemma

SUMMARY

The purpose of this work was to study the structure of the hypothesis-testing and associated sum-of-squares for linear models and for unbalanced data, as well as to set up the relationship among these hypothesis and the output of some statistical softwares.

The methods of analysis for unbalanced data have been revised. For this, the following statistical model was considered:

$y = X\theta + e$, characterized by:

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

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

for $i = 1, \dots, a$ rows; $j = 1, \dots, b$ columns; $e_{n_{ij}}$ elements for each cell, where:

- y_{ijk} : denote the observation in k-th repetition of the i-th level of factor A and the j-th level of factor B;
- μ : denote the effect of the constant (grand mean);
- α_i : denote the i-th level of the factor A;
- β_j : denote the j-th level of the factor B;
- γ_{ijk} : denote the interaction between the i-th level of factor A and the j-th level of factor. B;
- e_{ijk} : denote the random error inherent to y_{ijk} observation.

It has been established the analitic forms and computational methods relatively to the hypothesis and associated sum-of-squares of rows, columns and interaction as well in the methodology development. Also, statistical softwares and its details were exposed.

A numerical simulated example, with some empty cells was presented in order to illustrate the methodology and the objective proposed, including a comparison among the outputs of the softwares.

1. INTRODUÇÃO

Tanto no setor Agropecuário como em qualquer outro setor, quando se realiza uma pesquisa, para se atingir o objetivo colimado, é de fundamental importância formular de maneira precisa as hipóteses, que através de uma análise estatística irão dar cunho científico à essa pesquisa.

A interpretação das verdadeiras hipóteses que são testadas através das somas de quadrados obtidas por qualquer dos métodos disponíveis, pode trazer problemas quando se tem dados desbalanceados. Assim, a estrutura de desbalanceamento pode induzir a hipóteses equivocadas, isto é, pode-se imaginar estar testando uma hipótese $H_o^{(1)}$, quando na realidade se está testando uma hipótese $H_o^{(2)}$, que não corresponde ao objetivo da pesquisa. Um outro agravante é o uso inadequado dos pacotes estatísticos, porquanto, via de regra, o usuário de pacotes não conhece exatamente as hipóteses envolvidas, acarretando na maioria das vezes, interpretações distorcidas. Um problema para usuários das rotinas da análise de variância é que as expressões das somas de quadrados, raramente se apresentam de uma forma clara e inequívoca nas saídas dos pacotes. Por exemplo, na análise de

dados através de um modelo de classificação dupla com interação, precisa-se conhecer a expressão exata da soma de quadrados que é simplesmente rotulada como "A". Apenas como ilustração, ela pode ser do tipo $R(\alpha|\mu)$, $R(\alpha|\mu,\beta)$, SSA_w ou $R(\alpha|\mu,\beta,\phi)$, entre outros.

Pretende-se com este trabalho, rever os métodos de análise para dados desbalanceados, mais citados na literatura e estabelecer uma relação clara entre somas de quadrados e as hipóteses que estão sendo testadas.

Saídas de pacotes como: SAS (1982), BMDP (1983), SYSTAT (1985), SOC (1982), STATGRAPHICS (1985), SANEST (1986), SAEG (1986), GLIM (1984), HARVEY (1987), MINITAB (1986), MSTAT (1984), SPSS (1986), e outros de interesse para o setor Agro-Pecuário, são aqui estudadas, visando estabelecer as hipóteses envolvidas e suas respectivas somas de quadrados. Explicitamente, o presente trabalho se propõe a:

1. Estruturação de hipóteses, utilizando modelos lineares para dados desbalanceados.
2. Estudo de programas estatísticos para dados desbalanceados, com o fito de determinar as hipóteses envolvidas.
3. Comparação das saídas geradas pelos pacotes estatísticos mais utilizados, visando a uma orientação na análise e formulação das hipóteses.

2. REVISÃO DE LITERATURA

2.1. Introdução

A história da análise de dados desbalanceados, pode ser descrita, segundo Herr (1986), desde as publicações originais de Yates (1933, 1934), Hendricks (1934), Snedecor & Cox (1935), Comstock (1943), Hazel (1946), Patterson (1946), Vajda (1974), Yates (1947), Stevens (1948), Henderson (1953), Kramer (1955), até o início dos anos sessenta, com a revolução computacional. Nesta década, conforme Herr (1986), entre outros, surgiram artigos de Searle (1965), Grosslee & Lucas (1965) e Federer & Zelen (1966).

ARVESEN & LAYARD (1975) obtiveram um teste assintoticamente robusto para a hipótese $\Delta \geq \Delta_0$ contra $\Delta > \Delta_0$, onde Δ é quociente de variâncias em um modelo com um fator, desbalanceado, segundo ANOVA (Analysis of Variance) Modelo II. O teste é baseado em uma extensão de Arvesen (1969) usando "jackknife". A robustez do teste é mostrada através de resultados teóricos e simulados utilizando o método de Monte Carlo.

SPEED & HOCKING (1976) estabeleceram diferenças entre dois procedimentos que denominaram "Procedimento 1" e "Procedimento 2", concernentes ao uso da notação $R(\cdot)$, mostrando que um não é um caso especial do outro. Enfatizaram em seu trabalho que o Procedimento 1 é aplicado a modelos de posto incompleto, enquanto o Procedimento 2 é aplicado a modelos reparametrizados de posto completo. Ressaltaram ainda que algumas das hipóteses a serem testadas quando do uso da notação $R(\cdot)$ nos dois procedimentos não são claras e outras não podem ser testadas.

SPEED, HOCKING & HACKNEY (1978) consideraram a existência de métodos para analisar modelos de delineamento experimental com dados desbalanceados e os relacionaram à existência de programas computacionais. Os métodos são caracterizados pelas hipóteses associadas com as somas de quadrados que são geradas. A escolha de um método seria baseada na adequacidade da hipótese ao invés da conveniência computacional ou a ortogonalidade das formas quadráticas. As somas de quadrados são descritas através da notação $R(\cdot)$ aplicada ao modelo linear superparametrizado. As hipóteses são estabelecidas em termos do modelo de médias de caselas de posto completo. A situação de casela de frequência zero é abordada sucintamente.

URQUHART & WEEKS (1978) apresentaram alternativas para análise de valores perdidos, através de modelos

lineares, sugerindo que pesquisadores em colaboração com biometricistas poderiam definir funções paramétricas relevantes em termos de médias de caselas. Ressaltaram ainda que o modelo de média de caselas dá ao pesquisador a flexibilidade de especificar precisamente as funções de interesse.

ALALOUF & STYAN (1979) baseados no fato de que no modelo linear geral $E(y) = X\theta$, o vetor $A\theta$ é estimável se existe uma matriz B tal que $E(By) = A\theta$, apresentaram diversas caracterizações de estimabilidade, com provas simples e rápidas. As caracterizações envolvem igualdades de postos, inversas generalizadas, complementos de Schur e matrizes partidas.

HOCKING, SPEED & COLEMAN (1980) propiciaram um critério para escolher as hipóteses a serem testadas no caso de desbalanceamento, e apresentaram um algoritmo para obter as hipóteses desejadas. A metodologia consiste em iniciar com aquelas hipóteses que teriam sido testadas se o experimento fosse balanceado. Em seguida, baseando-se naquelas hipóteses seria determinada uma hipótese "efetiva".

BRYCE, SCOTT & CARTER (1980) desenvolveram uma abordagem para a análise de modelos de efeito fixo balanceados através de uma reparametrização do modelo de médias das caselas.

FRANE (1980) fez uma discussão dos diversos algoritmos usados no BMDP (Biomedical Computer Programs)

focalizando as hipóteses decorrentes do modelo linear geral.

GOODNIGHT (1980) usou o conceito de estimabilidade para testar hipóteses em modelos lineares com múltiplos fatores de efeitos fixos sem violar as pressuposições usuais. Definiu três tipos de funções estimáveis e tratou cada tipo com número desigual de observações, caselas vazias e qualquer grau de confundimento para qualquer modelo linear de efeitos fixos.

JORN & KLOTZ (1980) consideraram o problema de calcular soma de quadrados em modelos de análise de variância com múltiplos fatores, com número desigual, mas não zero por casela. Apresentaram um algoritmo baseado no método de contrastes de Scheffé.

SEARLE (1980) mostrou como os quatro tipos de funções estimáveis da saída "output" do SAS ("Statistical Analysis System") são obtidas das expressões gerais para testar hipóteses através da tradicional estatística F para dados desbalanceados e como elas são usadas.

MONLEZUN & SPEED (1980) discorrem que a abordagem clássica para modelos lineares é de natureza algébrica, com atenção voltada sobre o espaço de parâmetros. O vetor de parâmetros é estimado minimizando-se uma forma quadrática. O numerador da soma de quadrados para testar uma hipótese nula no modelo restrito é dado pela diferença entre duas formas quadráticas na estimativa do vetor de parâmetros

para o modelo irrestrito. O instrumento para a abordagem geométrica é o espaço das observações. O vetor de médias é estimado diretamente através de uma projeção ortogonal do vetor de observações sobre um subespaço apropriado, se o modelo é ou não restrito. A estimativa do vetor de parâmetros é obtida dessa estimativa por uma transformação linear apropriada. A geometria do espaço das observações revela que o numerador da soma de quadrados para testar a hipótese nula no modelo restrito pode ser expressa como uma só forma quadrática na estimativa do vetor de parâmetros para o modelo restrito. Esta forma quadrática simples é computacionalmente mais eficiente do que a diferença entre as duas formas quadráticas citadas inicialmente.

BURDICK & HERR (1980) mostraram através de exemplos, fracassos decorrentes da análise de dados desbalanceados, quando se usam as abordagens da interpretação de parâmetros e da comparação de modelos. Na primeira, o fracasso consiste na não ortogonalidade, enquanto na segunda, o principal fracasso é decorrente da dependência das hipóteses testadas sobre o tamanho das caselas.

IEMMA (1983) tomou três casos clássicos de blocos incompletos balanceados, BIB tipo I, BIB tipo II e BIB tipo III, para estudar testes de hipóteses nos experimentos com parcelas subdivididas.

MORTAZA & JENNRICH (1988) consideraram uma nova classe de algoritmos baseada em uma métrica de subclasses proporcionais definida pelos problemas de ANOVA. Obtiveram novos resultados teóricos sobre convergência fazendo algumas comparações empíricas, e, dando também um tratamento de análise de covariância.

SOUZA & IEMMA (1990) estudaram dados longitudinais incompletos, usando modelos de curvas de crescimento. Supondo que os resíduos seguem um processo autoregressivo de 1ª ordem, estimaram-se os parâmetros do modelo proposto pelo método de máxima verossimilhança. Utilizando-se a estatística de Wald realizaram testes de hipóteses da forma $C\xi U = \phi$.

SUBRAMANI (1992) fez uma tentativa para reduzir as complexidades computacionais envolvidas na estimação de vários valores perdidos. Como um resultado ele mostrou que se pode estimar m valores perdidos desenvolvendo apenas k ($\leq m$) equações lineares, onde m e k são respectivamente o número de valores perdidos e caselas vazias. O método é também ilustrado com o auxílio de um exemplo numérico.

2.2. O Modelo Linear

Aborda-se de forma sucinta alguns conceitos da teoria dos modelos lineares, segundo IEMMA (1991), SEARLE

(1987), HOCKING (1985), ALALOUF & STYAN (1979), CARVALHO FILHO & IEMMA (1990), GRAYBILL (1961), MARSAGLIA & STYAN (1974a e 1974b),

Considere o modelo linear

$$E(y) = X\theta \quad (2.1)$$

onde

X é uma matriz conhecida de dimensão $(n \times p)$,
 θ é um vetor de parâmetros desconhecidos de dimensão $(p \times 1)$ e

y é um vetor de variáveis aleatórias observáveis.

Como y e $E(y)$ não são em geral iguais, defini-se a diferença

$$e = y - E(y) \quad (2.2)$$

e considera-se essa diferença como um vetor de erros. De (2.2) tem-se que $y = E(y) + e$, que combinado com (2.1) produz

$$y = X\theta + e. \quad (2.3)$$

2.2.1. Estimação

Deseja-se estimar o vetor de dimensão $(s \times 1)$

$$A\theta \quad (2.4)$$

onde A é uma matriz dada $(s \times p)$. Quando existe uma matriz B $(s \times n)$, tal que

$$E(B y) = A\theta, \quad (2.5)$$

então $A\theta$ é dito *estimável*.

Usando (2.1), (2.5) é verdade se, e somente se

$$A = BX \quad (2.6)$$

ou equivalentemente,

$$\rho \begin{bmatrix} X \\ A \end{bmatrix} = \rho(X) \quad (2.7)$$

onde $\rho(\cdot)$ denota posto.

ROY & ROY (1959) mostraram que a hipótese nula $A\theta = 0$ é completamente testável se

$$\rho(XT) = \rho(X) - \rho(A) \quad (2.8)$$

onde a matriz T gera o espaço nulo de A .

MILLIKEN (1971) mostrou que (2.6) e (2.8) são equivalentes.

Outra caracterização de (2.6).

Se X^- é qualquer inversa generalizada de X satisfazendo a

$$X X^- X = X, \quad (2.9)$$

então (2.6) se verifica se e somente se

$$A X^- X = A. \quad (2.10)$$

SEARLE (1982) provou (2.10) fazendo $X^l = (X' X)^- X$, que é uma inversa generalizada de mínimos quadrados satisfazendo a

$$X^l = (X' X)^- X \quad (2.11)$$

e

$$(X X^l)' = X X^l \quad (2.12)$$

Seguindo RAO & MITRA (1971), diz-se que X^- é uma inversa generalizada reflexiva se ambas, (2.9) e (2.11) são satisfeitas, e então $X^- = X_r^-$. Diz-se que X^- é uma inversa generalizada de quadrados mínimos se (2.9) e (2.12) são satisfeitas, e então $X^- = X_l^-$. Mostra-se que uma inversa generalizada é reflexiva se e somente se $\rho(X^-) = \rho(X)$. Além disso, $\theta = Ky$ é uma das equações normais $X'X\theta = X'y$ se e somente se K é uma inversa generalizada de quadrados mínimos de X , conforme RAO & MITRA (1971), IEMMA (1989), IEMMA & PALM (1991/1992).

Assumindo apenas que $\rho(X) < \min(n, p)$, as caracterizações de estimabilidade caem facilmente nos casos que envolvem X e $X'X$.

2.2.2. Formas do Modelo Linear

Adota-se a estrutura de Gauss-Markov para o modelo (2.3), isto é, $e \sim N(0, I_n)$, com vista a futuras aplicações. Esse modelo será abordado em três formas: Modelo superparametrizado (Modelo - S), Modelo de médias de caselas (Modelo - M), Modelo com restrição paramétrica do tipo sigma (Modelo - Σ).

2.2.2.1. O Modelo Superparametrizado

Diz-se que um modelo é superparametrizado, se

para cada efeito dos fatores envolvidos ele exibe seu respectivo parâmetro. Aqui ele será expresso pelo modelo (2.3), isto é,

$$y = X\theta + e.$$

Consoante interesse do pesquisador, o Modelo-S pode assumir diferentes caracterizações. Assim, parametrizando sucessivamente desde o modelo mais simples até o modelo com dois fatores e interação, tem-se:

$$y = X_1\theta_1 + e \Rightarrow y_i = \mu + e_i \quad (S.1)$$

$$y = X_2\theta_2 + e \Rightarrow y_{ij} = \mu + \alpha_i + e_{ij} \quad (S.2)$$

$$y = X_3\theta_3 + e \Rightarrow y_{ij} = \mu + \beta_j + e_{ij} \quad (S.3)$$

$$y = X_4\theta_4 + e \Rightarrow y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \quad (S.4)$$

$$y = X_5\theta_5 + e \Rightarrow y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \quad (S.5)$$

para $i = 1, \dots, a$ linhas; $j = 1, \dots, b$ colunas; e n_{ij} elementos em cada casela.

Nota-se que, exceto (S.1) todas as outras caracterizações de (2.3) são de posto coluna incompleto.

Relativamente aos modelos acima, (S.1) corresponde a um modelo aditivo, isto é, todos os dados podem ser descritos como a adição de uma constante e um erro aleatório. Os modelos (S.2) e (S.3) decorrem de se adicionar um efeito fixo (α_i ou β_j) ao modelo anterior, e são chamados de modelos de classificação simples. Os modelos (S.4) e (S.5) são conhecidos como modelos de classificação dupla cruzada, sem e com interação respectivamente.

2.2.2.2. O Modelo de Médias de Caselas

Os modelos superparametrizados, descritos antes, segundo IEMMA (1991), são parte integrante da história dos modelos lineares e tem sido de grande valia para os pesquisadores das ciências aplicadas; porquanto, eles podem mostrar explicitamente os parâmetros μ , α , β e γ , sobre os quais estão em geral, os reais anseios de tais pesquisadores. Segundo HOCKING (1985) e SEARLE (1987), entre outros, eles apresentam alguns inconvenientes como, por exemplo, o de possuir mais parâmetros que médias de caselas para estimá-los. Esse problema pode ser sensivelmente agravado no caso de caselas vazias. Detalhes podem ser obtidos em HOCKING (1985), SEARLE (1987) e IEMMA (1991,1993). Ressalte-se ainda, em face aos objetivos deste trabalho, os modelos de médias apresentam ainda a vantagem de descrever hipóteses com grande simplicidade, em relação aos Modelos - S.

Segundo SPEED, HOCKING & HACKNEY (1978), define-se o Modelo - M, como o Modelo Linear

$$y = W\mu + e \quad (2.13)$$

onde

W é uma matriz ($n \times k$) conhecida,

μ é um vetor de médias ($k \times 1$) das populações das quais são retiradas as amostras aleatórias que compõem as caselas e,

y_i e e_i são como no Modelo - S.

Em correspondência com o Modelo - S tem-se:

$$y_i = \mu + e_i \quad \rightarrow \quad y_i = \mu + e_i \quad (M.1)$$

$$y_{ij} = \mu + \alpha_i + e_{ij} \quad \rightarrow \quad y_{ij} = \mu_i + e_{ij} \quad (M.2)$$

$$y_{ij} = \mu + \beta_j + e_{ij} \quad \rightarrow \quad y_{ij} = \mu_j + e_{ij} \quad (M.3)$$

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \quad \rightarrow \quad y_{ijk} = \mu_i + \beta_j + e_{ijk} \quad (M.4)$$

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \quad \rightarrow \quad y_{ijk} = \mu_{ij} + e_{ijk} \quad (M.5)$$

2.2.2.3. O Modelo com restrições paramétricas do tipo Σ

O Modelo- Σ é um modelo linear superparametrizado com restrições paramétricas do tipo "soma igual a zero". IEMMA (1991) denomina-o por:

$$y = Z \theta + e \quad (2.14)$$

e sua caracterização correspondente a (S.5) é dada por:

$$y_{ijk} = \mu + \dot{\alpha}_i + \beta_j + \dot{\gamma}_{ij} + e_{ijk} \quad (2.15)$$

onde,

$$\dot{\alpha}_i = \alpha_i - \frac{\sum_i \alpha_i}{a} \quad \Rightarrow \quad \sum_i \dot{\alpha}_i = 0$$

$$\beta_j = \beta_j - \frac{\sum_j \beta_j}{b} \Rightarrow \sum_j \beta_j = 0$$

$$\dot{\gamma}_{ij} = \gamma_{ij} - \frac{\sum_i \gamma_{ij}}{a} - \frac{\sum_j \gamma_{ij}}{b} + \frac{\sum_{ij} \gamma_{ij}}{ab},$$

consequentemente,

$$\sum_i \dot{\gamma}_{ij} = 0 \text{ para todo } j, \text{ e } \sum_j \dot{\gamma}_{ij} = 0 \text{ para todo } i.$$

3. METODOLOGIA

3.1. Introdução.

Nesta seção, abordam-se as hipóteses e as somas de quadrados associadas, a partir do modelo com dois fatores cruzados, com e sem interação. Interpretam-se os termos das respectivas somas de quadrados, os testes de linhas, colunas e interação. A metodologia usada será a seguida por IEMMA (1991), SEARLE (1987) e HOCKING (1985).

O modelo linear geral é representado como em (2.3) por $y = X\theta + e$, onde:

y é um vetor de observações de dimensão $(n \times 1)$;

X é uma matriz conhecida de dimensão $(n \times p)$;

θ é um vetor de parâmetros desconhecidos de dimensão $(p \times 1)$

e é um vetor $(n \times 1)$ de termos aleatórios, assumindo a matriz $\sigma^2 I_{(n)}$, sendo $I_{(n)}$ a matriz identidade de ordem n .

Uma solução das equações normais $X'X\theta = X'y$ é dada por $\theta^o = (X'X)^G X'y$, onde $(X'X)^G$ é qualquer matriz satisfazendo $X'X(X'X)^G X'X = X'X$, e $r = \rho(X)$ é o posto de X . Um estimador não viesado de σ^2 é

$$\begin{aligned}\hat{\sigma}^2 &= \frac{y'(I - X(X'X)^{-1}X')y}{(n - r)} \\ &= y'(I - P)y/(n - r)\end{aligned}\quad (3.1)$$

Para B' , uma matriz de posto coluna completo com $s = \rho(B)$ e, cada uma das s funções de θ em $B'\theta$ é estimável, a hipótese

$$H_0: B'\theta = m \quad (3.2)$$

é dita testável, e sob normalidade

$$F = Q/\hat{\sigma}^2, \text{ com } Q = (B'\theta^0 - m)'(B'\theta - m) \quad (3.3)$$

tem distribuição de F com s e $(n - r)$ graus de liberdade e pode ser usada para testar H_0 (SEARLE, 1971).

Um procedimento alternativo consiste na obtenção das somas de quadrados das hipóteses de interesse. Assim, para a hipótese

$$H_0: B'\theta = \phi \quad (3.4)$$

onde $B'\theta$ é estimável e B' tem posto linha completo, tem-se

$$S.Q.H_0 = (B'\theta^0)'B'(X'X)^{-1}B(B'\theta^0) \quad (3.5)$$

Verifica-se também que

$$\frac{S.Q.H_0}{\hat{\sigma}^2} \sim \chi^2_{(\rho(B); \delta)} \quad (3.6)$$

onde $\delta = \frac{1}{2\hat{\sigma}^2} \theta^0 X' P^* X \theta^0$ é o parâmetro de não centralidade e

P^* é um projetor ortogonal tal que $y'P^*y = S.Q.H_0$. Ressalte-se ainda que, se $P^*(I - P^*) = \phi$ então as formas quadráticas associadas são independentes e o quociente dos respectivos quadrados médios tem distribuição F não central

$$\frac{Q.M.H_0}{Q.M.Res} \sim F_{(p(B); p(I - B); \delta)} \quad (3.7)$$

$$E[S.Q.H_0] = E[y'P^*y] = \text{Tr}[P^*] \sigma^2 + e'X'P^*Xe \quad (3.8)$$

Assim, com base no parâmetro de não centralidade δ de (3.6) ou na esperança da S.Q.H₀ de (3.8) pode-se verificar qual a verdadeira hipótese que está sendo testada. (CIEMMA 1991 e 1992).

O uso adequado deste procedimento para teste de hipóteses impõe inicialmente a formulação da hipótese de acordo com (3.2) ou (3.4), e então, testá-la usando (3.3) ou (3.7).

3.2. Hipóteses e Somas de Quadrados Associadas

Nesta seção serão abordadas as hipóteses mais usadas para testar efeitos de linhas, colunas e interações. Na seção seguinte serão vistos os métodos mais utilizados para testá-las e na subsequente será discutido o comportamento de diversos pacotes estatísticos em relação a tais métodos e hipóteses.

Geralmente, para testar os efeitos de linhas (Fator A) existem quatro hipóteses mais comuns:

- i) a hipótese sobre médias ponderadas de linhas $\bar{L}_i$;
- ii) a hipótese sobre médias ponderadas de linhas ajustadas para colunas, $(\bar{L}/C)_i$;

iii) a hipótese sobre médias não ponderadas de linhas,

$$\bar{l}_i.$$

iv) a hipótese do tipo iv ,contrastes sobre linhas,de interesse do pesquisador, baseados nas caselas ocupadas.

(i = 1 , 2, ... ,a).

Analogamente, para testar os efeitos das colunas (Fator B):

i) a hipótese sobre médias ponderadas de colunas $\bar{c}_j$;

ii) a hipótese sobre médias ponderadas de colunas ajustadas para linhas, $(\bar{c}/L)_j$;

iii) a hipótese sobre médias não ponderadas de colunas,

$$c_j.$$

iv) a hipótese do tipo iv ,contrastes sobre colunas,de interesse do pesquisador, baseados nas caselas ocupadas.

(j = 1 , 2, ... ,b).

Também será visto como testar as hipóteses sobre a interação, A*B. Estas hipóteses serão estabelecidas considerando-se os modelos de classificação dupla cruzada com e sem interação, em casos nos quais todas as caselas estão ocupadas, bem como nos casos nos quais algumas caselas são vazias.

3.2.1 Todas as caselas ocupadas

Considera-se inicialmente o caso no qual apesar do desbalanceamento, todas as caselas estão ocupadas.

3.2.1.1. O modelo linear com interação.

a) *Hipóteses mais comuns sobre linhas.*

a.1) Hipóteses sobre as médias ponderadas de linhas.

$H_0^1: \bar{L}_i = \bar{L}_v \quad \forall \quad (i, v)$ que pode ser posto na forma

$$H_0^1: B_1' \mu = \phi$$

com

$$\begin{aligned} \text{S.Q.} H_0^1 &= (B_1' \hat{\mu})' B_1' (W'W)^{-1} B_1 \hat{\mu} \\ &= R(\alpha \mid \mu) \\ &= \sum_i (A_i^2 / n_{i.}) - C \end{aligned}$$

a.2) Hipóteses sobre as médias ponderadas de linhas ajustadas para colunas

$$\begin{aligned} H_0^{(2)}: \bar{L}_i &= (\bar{L} / C)_i, \quad \forall i \text{ ou} \\ &: \sum_j n_{ij} \mu_{ij} = \sum_j \sum_i (n_{ij} n_{.j} \mu_{ij}) / n_{.j} \end{aligned}$$

com

$$\text{S.Q.} H_0^{(2)} = R(\alpha \mid \mu, \beta)$$

a.3) Hipóteses sobre médias não ponderadas de linhas.

$$H_O^{(3)} : \bar{1}_i = \bar{1}_v, \forall (i, v)$$

com

$$S.Q.H_O^{(3)} = R(\alpha \mid \dot{\mu}, \dot{\beta}, \dot{\gamma})$$

a.4) Hipóteses do tipo iv (hipóteses sobre contrastes entre médias populacionais das caselas que estão na mesma linha, iniciando-se a comparação sempre pela última coluna).

$H_O^{(4)}$ é formulada como na expressão (3.4), e

$S.Q.H_O^{(4)}$ como em (3.5).

b) *Hipóteses mais comuns sobre colunas.*

b.1) Hipóteses sobre médias ponderadas de colunas

$$H_O^{(5)} : \bar{c}_j = \bar{c}_{j'}, \forall (j, j')$$

com

$$S.Q.H_O^{(5)} = R(\beta \mid \mu) = \sum_j (\beta_j^2 / n_{.j}) - c$$

Observa-se que, para a obtenção de $S.Q.H_O^{(5)}$ através de (3.5) é necessário notar a existência de b colunas e conseqüentemente $(b-1)$ graus de liberdade. Ressalte-se ainda, que existe várias maneiras de escolher B' tal que se tenha

$$H_O^{(5)} : B'\mu = \phi$$

resultando em hipóteses equivalentes.

b.2) Hipóteses sobre médias ponderadas de colunas ajustadas para linhas.

Analogamente a (a.2), tem-se

$$H_O^{(6)} : \bar{c}_j = (\bar{C} / L)_j, \forall j$$

com

$$S.Q.H_O^{(6)} = R(\beta \mid \mu, \alpha)$$

b.3) Hipóteses sobre médias não ponderadas de colunas

$$H_O^{(7)} : \bar{c}_j = \bar{c}_{j'} \quad \forall (j, j')$$

com

$$S.Q.H_O^{(7)} = R(\beta \mid \mu, \alpha, \gamma)$$

b.4) Hipóteses do tipo iv (hipóteses sobre contrastes entre médias populacionais das caselas que estão na mesma coluna, iniciando-se a comparação sempre pela última linha).

$H_O^{(8)}$ é formulada como na expressão (3.4), e

$S.Q.H_O^{(8)}$ como em (3.5).

c) *Hipóteses mais comuns sobre a interação.*

$H_0^{(9)}$: Qualquer vetor de $(ab-a-b+1)$ funções estimáveis de

$$\Theta_{ij,ij} = \gamma_{ij} - \gamma_{i.} - \gamma_{.j} + \gamma_{..}$$

ou de combinações lineares estimáveis de e 's não estimáveis, é nulo.

Com

$$S.Q.H_0^{(9)} = R(\gamma \mid \mu, \alpha, \beta) = R(\dot{\gamma} \mid \dot{\mu}, \dot{\alpha}, \dot{\beta})$$

OBS. Os parâmetros envolvidos na expressão anterior já foram definidos no modelo (2.14).

3.2.1.2. O modelo linear sem interação.

Quando se tem dados dispostos em caselas, de modo que essas caselas são todas ocupadas, uma regra na escolha entre um modelo com e sem interação baseia-se em testar $H_0^{(9)}$. Caso $H_0^{(9)}$ seja verdadeira pode-se adotar o modelo sem interação, embora essa regra não seja adequada quando existem caselas vazias, como se verá na seção seguinte.

Como se pode observar, as hipóteses até aqui apresentadas não testam diretamente a igualdade entre efeitos de linhas ou entre efeitos de colunas, porquanto, os contrastes simples entre efeitos de linhas e entre efeitos de colunas não são estimáveis e conseqüentemente não são

testáveis no modelo com interação. No modelo objeto desta seção pode-se testar hipóteses do tipo:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_a \quad \text{ou} \quad H_0: \beta_1 = \beta_2 = \dots = \beta_b$$

Tomando-se como base o modelo de médias de caselas (Modelo - M)

$$y_{ijk} = \mu_i + \beta_j + e_{ijk}$$

onde

y_{ijk} e e_{ijk} já foram definidos anteriormente;

$\mu_i = \mu + \alpha_i$ representa os efeitos de linhas e

β_j descreve os efeitos de colunas.

a) Hipóteses mais comuns sobre linhas

a.1) Hipóteses sobre médias ponderadas de linhas

$$H_0^{(10)}: \mu_i + \sum_j n_{ij} \beta_j / n_{i.} = \mu_{i'} + \sum_j n_{i'j} \beta_j / n_{i'.$$

$$\forall (i, i')$$

com soma de quadrados dada por

$$S.Q.H_0^{(10)} = R(\alpha \mid \mu) = \sum_i (A_i^2 / n_{i.}) - C$$

ressalte-se que

$$S.Q.H_0^{(1)} = S.Q.H_0^{(10)} = R(\alpha \mid \mu)$$

embora, $S.Q.H_0^{(1)}$ e $S.Q.H_0^{(10)}$ sejam fundamentalmente diferentes. Para isto, basta expressar as referidas hipóteses em termos do modelo-S.

$$H_O^{(1)}: \alpha_i + \sum_j n_{ij}(\beta_j + \gamma_{ij}) / n_{i.} \text{ iguais } \forall i$$

$$H_O^{(10)}: \alpha_i + \sum_j n_{ij} \beta_j / n_{i.} \text{ iguais } \forall i.$$

a.2) Hipóteses sobre médias ponderadas de linhas ajustadas para colunas.

$$H_O^{(11)}: \mu_i = \mu_{i'} \quad \forall (i, i')$$

ou equivalente, no modelo - S

$$H_O^{(11)}: \alpha_i = \alpha_{i'} \quad \forall (i, i')$$

com soma de quadrados dada por

$$S.Q.H_O^{(11)} = R(\alpha \mid \mu, \beta)$$

que analogamente ao caso anterior,

$$S.Q.H_O^{(11)} = S.Q.H_O^{(2)},$$

embora, $H_O^{(11)}$ e $H_O^{(2)}$ sejam intrinsecamente diferentes.

a.3) Hipóteses sobre médias não ponderadas de linhas.

$$H_O^{(12)}: \mu_i = \mu_{i'} \quad \forall (i, i')$$

ou equivalente, no modelo - S

$$H_O^{(12)}: \alpha_i = \alpha_{i'} \quad \forall (i, i')$$

com soma de quadrados dada por

$$S.Q.H_O^{(12)} = R(\alpha \mid \bar{\mu}, \bar{\beta}) = R(\alpha \mid \mu, \beta)$$

Como pode ser observado, no modelo sem interação as hipóteses sobre as médias de linhas ajustadas para colunas e as hipóteses sobre as médias não ponderadas de linhas, reduzem-se à mesma hipótese. Ressalte-se ainda, que

a soma de quadrados devida às médias não ponderadas de linhas, dada conforme acima, envolve o modelo - Σ , ou seja,

$$\sum_i \dot{\alpha}_i = 0, \quad \dot{\alpha}_i = \alpha_i - \bar{\alpha}.$$

a.4) Hipóteses do tipo iv (hipóteses sobre contrastes entre médias populacionais das caselas que estão na mesma linha, iniciando-se a comparação sempre pela última coluna).

$H_O^{(13)}$: é formulada como na expressão (3.4), e
 $S.Q.H_O^{(13)}$ como em (3.5).

b) *Hipóteses mais comuns sobre colunas.*

b.1) Hipóteses sobre médias ponderadas de colunas.

$$H_O^{(14)}: \beta_j + \sum_i n_{ij} \mu_i / n_{.j} = \beta_{j'} + \sum_i n_{ij'} \mu_i / n_{.j'} \\ \forall (j, j')$$

com soma de quadrados dada por

$$S.Q.H_O^{(14)} = R(\beta \mid \mu) = \sum_j (\beta_j^2 / n_{.j}) - C$$

b.2) Hipóteses sobre médias ponderadas de colunas ajustadas para linhas.

$$H_O^{(15)}: \beta_j = \beta_{j'} \quad \forall (j, j')$$

com soma de quadrados dada por

$$S.Q.H_O^{(15)} = R(\beta \mid \mu, \alpha)$$

b.3) Hipóteses sobre médias não ponderadas de colunas

$$H_O^{(16)}: \beta_j = \beta_{j'} \quad \forall \quad (j, j')$$

com soma de quadrados dada por

$$S.Q.H_O^{(16)} = R(\dot{\beta} \mid \dot{\mu}, \dot{\alpha}) = R(\dot{\beta} \mid \mu, \alpha)$$

b.4) Hipóteses do tipo iv (hipóteses sobre contrastes entre médias populacionais das caselas que estão na mesma coluna, iniciando-se a comparação sempre pela última linha).

$H_O^{(17)}$ é formulada como na expressão (3.4), e

$S.Q.H_O^{(17)}$ como em (3.5).

Relativamente ao modelo sem interação, observa-se que, mesmo existindo desbalanceamento, se todas as caselas são ocupadas, as hipóteses sobre médias ponderadas ajustadas e as hipóteses sobre médias não ponderadas são as mesmas, ou seja:

$$H_O^{(11)} = H_O^{(12)} \quad \text{e} \quad H_O^{(15)} = H_O^{(16)}.$$

3.2.2. Algumas caselas vazias

Viu-se na seção anterior que, quando todas as caselas são ocupadas, utilizando-se o modelo de médias, a matriz X é de posto coluna completo e possibilita testar hipóteses muito convenientemente acerca das médias populacionais. Contudo, quando existem caselas vazias, algumas médias

populacionais correspondentes as caselas vazias não são estimáveis. Nesse caso, X pode não ser mais de posto coluna completo e surgem vários problemas motivados pela presença das caselas vazias.

Dentre esses problemas, segundo IEMMA (1991) pode-se relacionar os seguintes:

- i) Podem ocorrer desconexões (entre as caselas);
- ii) As hipóteses sobre médias não ponderadas de linhas e colunas, em geral, são diferentes daquelas vistas com caselas ocupadas;
- iii) A hipótese sobre a interação não testa, como em $H_0^{(g)}$ que todas as interações são nulas;
- iv) A interpretação das hipóteses está condicionada à posição e ao número das caselas vazias;
- v) Os graus de liberdade da interação devem ser escolhidos com cautela.

3.2.2.1. Conectividade.

O conceito de conectividade foi introduzido por BOSE & SRIVASTAVA (1964) na forma de uma cadeia entre blocos e tratamentos. Um bloco e um tratamento são conectados se pode formar uma cadeia de tratamentos e blocos da forma: tratamento - bloco - tratamento - bloco - ... - trata-

mento, começando com o fator A (tratamento) passando através da cadeia (bloco) para associar-se. Dessa forma, vários blocos na cadeia são associados mediante tratamentos adjacentes entre si. WEEKS & WILLIAMS (1964) definem conectividade da seguinte forma: "Um modelo é conectado se todo contraste simples de diferenças de dois níveis de um fator é estimável. RAGHAVARAO (1971) define conectividade como: "Um delineamento é dito conectado se todos os contrastes elementares são estimáveis. Se isto não acontece, o delineamento é dito desconexo." Discussão detalhada sobre o assunto poderá ser encontrada em SEARLE (1987), VILCA & IEMMA (1990), IEMMA (1991), e IEMMA (1993a).

3.2.2.2. Hipóteses sobre médias não ponderadas de linhas, colunas e interações.

As expressões para as hipóteses sobre linhas, colunas e interações quando se tem caselas vazias, apesar de serem análogas aquelas vistas anteriormente, podem ser frontalmente diferentes. Para dados com todas as caselas ocupadas pode-se definir

$$\rho_i = \frac{1}{b} \sum_j \mu_{ij}$$

que é a média não ponderada das médias da linha i . Considerando os dados do conjunto A a seguir, tem-se:

Conjunto A

	j = 1	j = 2	j = 3
i = 1	μ_{11}	μ_{12}	μ_{13}
i = 2	μ_{21}	μ_{22}	-

$$\rho_1 = \frac{1}{3} (\mu_{11} + \mu_{12} + \mu_{13})$$

$$\rho_2 = \frac{1}{3} (\mu_{21} + \mu_{22} + \mu_{23})$$

porém, como não existe μ_{23} , então ρ_2 não é estimável. Surge então o problema: como comparar as linhas 1 e 2, usando-se médias não ponderadas de linhas. SEARLE (1987) diz que é estimável qualquer μ_{ij} para toda casela ocupada e qualquer combinação linear de μ 's de caselas ocupadas. Contudo, outra questão se apresenta: qual combinação linear de μ_{ij} 's de caselas ocupadas pode ser de interesse para comparação de linhas. Usando-se as caselas ocupadas do Conjunto A, pode-se definir

$$\rho'_1 = \frac{1}{2} (\mu_{11} + \mu_{12}) \quad \text{e} \quad \rho'_2 = \frac{1}{2} (\mu_{21} + \mu_{22})$$

que são funções lineares estimáveis, pois são constituídas por μ_{ij} 's estimáveis. Então, para comparar as linhas 1 e 2 pode-se tomar

$$\widehat{\rho'_1 - \rho'_2} = \frac{1}{2} (\mu_{11} + \mu_{12}) - \frac{1}{2} (\mu_{21} + \mu_{22})$$

onde o BLUE de $(\rho'_1 - \rho'_2)$ provê informações sobre a diferença entre as linhas 1 e 2, embora usando apenas as caselas das

colunas 1 e 2, ou seja

$$(\rho'_1 - \rho'_2) = \frac{1}{2} (\bar{y}_{11.} + \bar{y}_{12.}) - \frac{1}{2} (\bar{y}_{21.} + \bar{y}_{22.})$$

Convém salientar que este procedimento vale para outras comparações de interesse, utilizando apenas parte dos dados. Ressalte-se ainda, que quando todas as caselas estão ocupadas existem $ab(a-1)(b-1)$ interações θ 's, das quais no máximo, $(a-1)(b-1)$ são linearmente independentes (L. I.). Também isto é válido para dados com algumas caselas vazias. No primeiro caso, qualquer conjunto das $(a-1)(b-1)$ interações L.I. de $\theta_{ij,ij'}$ pode ser estimável, enquanto no segundo caso, apenas algumas das $(a-1)(b-1)$ interações L.I. são estimáveis. Finalmente, concluindo esta discussão, pode-se utilizar as mesmas expressões das hipóteses H_0^1 a $H_0^{1?}$ quando se tem caselas vazias, contudo, deve-se ter em mente que essas expressões correspondem a hipóteses diferentes. IEMMA (1991 e 1993b), MEAD; BANCROFT & HAN (1975), FRANCIS (1973), FREUND (1980) e DODGE (1985), entre outros, abordam o problema dessas hipóteses relacionadas com o desbalanceamento.

3.2.3. Resumo das hipóteses

Tabela 3.1. Hipóteses nos modelos com dois fatores com interação.

EFEITOS		
<u>Linhas</u>		
$H_o^{(1)} : \bar{L}_{i.} = \bar{L}_{i' .} \quad \forall (i, i')$		ponderadas
$H_o^{(2)} : \bar{L}_{i.} = (\bar{L} / C)_{i.} \quad \forall i$		ponderadas (□)
$H_o^{(3)} : \bar{l}_{i.} = \bar{l}_{i' .} \quad \forall (i, i')$		não ponderadas
$H_o^{(4)} : B'\theta = \phi$		tipo iv
<u>Colunas</u>		
$H_o^{(5)} : \bar{C}_{.j} = \bar{C}_{.j'} \quad \forall (j, j')$		ponderadas
$H_o^{(6)} : \bar{C}_{.j} = (\bar{C} / L)_{.j} \quad \forall j$		ponderadas (■)
$H_o^{(7)} : \bar{c}_{.j} = \bar{c}_{.j'} \quad \forall (j, j')$		não ponderadas
$H_o^{(8)} : B'\theta = \phi$		tipo iv
<u>Interações</u>		
$H_o^{(9)} : \theta_{ij, i'j'} = \gamma_{ij} - \gamma_{i'j} - \gamma_{ij'} + \gamma_{i'j'}$		

(□) - ajustadas para colunas ; (■) - ajustadas para linhas

Tabela 3.2 Hipóteses nos Modelos com dois fatores sem interação.

EFEITOS	
<u>Linhas</u>	
$H_O^{(10)}: \mu_i + \sum_j n_{ij} \beta_j / n_{i.} = \mu_{i'} + \sum_j n_{i'j} \beta_j / n_{i'}. \quad \forall (i, i')$	ponderadas
$H_O^{(11)}: \mu_i = \mu_{i'}, \quad \forall (i, i')$	ponderadas (□)
$H_O^{(12)}: \tilde{\mu}_i = \tilde{\mu}_{i'}, \quad \forall (i, i')$	não ponderadas
$H_O^{(13)}: B'e = \phi$	tipo iv
<u>Colunas</u>	
$H_O^{(14)}: \beta_j + \sum_i n_{ij} \mu_i / n_{.j} = \beta_{j'} + \sum_i n_{i'j'} \mu_i / n_{.j'}, \quad \forall (j, j')$	ponderadas
$H_O^{(15)}: \beta_j = \beta_{j'}, \quad \forall (j, j')$	ponderadas (■)
$H_O^{(16)}: \tilde{\beta}_j = \tilde{\beta}_{j'}, \quad \forall (j, j')$	não ponderadas
$H_O^{(17)}: B'e = \phi$	tipo iv

(□) - ajustadas para colunas

(■) - ajustadas para linhas

3.3. Métodos computacionais

3.3.1. Introdução

Descrevem-se aqui, alguns métodos mais utilizados, considerando-se para cada método sua principal característica.

3.3.2. O método das médias não ponderadas.

O Método das médias não ponderadas foi descrito por YATES (1934) como um método aproximado, porém tido como computacionalmente simples para aquela época. Referências mais recentes podem ser encontradas em ANDERSON & BANCROFT (1952), SNEDECOR & COCHARAN (1967), OVERALL & SPIEGEL (1969), SEARLE (1971, 1987), LEVY, NARULA & ABRAMI (1973), SPEED & MONLEZUN (1976), SPEED, HOCKING & HACKNEY (1978), HOCKING (1985) e IEMMA (1991), entre outros.

Este método é aproximado, porquanto, as somas de quadrados por ele produzidas para testar linhas (A), colunas (B) e interação (A*B), não tem distribuição exata de qui-quadrado, exceto no caso particular do fatorial 2^k . Detalhes podem ser vistos em SPEED & MONLEZUN (1976). Atualmente, este método não tem sido usado, mesmo porque, com a disponibilidade de recursos computacionais de alta tecnolo-

gia, a simplicidade de cálculos deixou de ser uma grande vantagem. Ademais, segundo SPEED, HOCKING & HACKNEY (1978), os argumentos de que essa aproximação é razoável, como visto por exemplo, em LEVY, NARULA & ABRAMI (1973), não são muito convincentes.

A Tabela 3.3 mostra a análise de variância fornecida pelo método em questão. Infelizmente, nenhuma das hipóteses apresentadas no item 3.2 é testada por qualquer de suas somas de quadrados.

Tabela 3.3 Análise de variância de experimentos com dois fatores e amostras desbalanceadas, segundo o método das médias não ponderadas de YATES (1934).

VARIAÇÕES CONSIDERADAS	SOMAS DE QUADRADOS	HIPÓTESES DESCRITAS EM (3.2)
Linhas (A)	$\bar{n}_h (b) \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2$	nenhuma
Colunas (B)	$\bar{n}_h (a) \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2$	nenhuma
Interação (AB)	$\bar{n}_h \sum_{ij} (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$	nenhuma
Resíduo	$\sum_{ijk} (\bar{x}_{ijk} - \bar{x}_{ij.})^2$	

onde ,

$$\bar{n}_h = \frac{ab}{\sum_{ij} \frac{1}{n_{ij}}} \quad \text{e} \quad \bar{x}_{ij.} = \bar{y}_{ij.}$$

3.3.3. O método dos quadrados de médias ponderadas

Este método também foi descrito por YATES (1934), e mais recentemente por ANDERSON & BANCROFT (1952), STELL & TORRIE (1960), ELSTON & BUSH (1964), FEDERER & ZELEN (1966), SCHEFFÉ (1959), SEARLE (1971), NETER & WASSERMAN (1974) e IEMMA (1991) entre outros.

Este método utiliza as hipóteses $H_0^{(3)}$, $H_0^{(6)}$, $H_0^{(7)}$, para linhas, colunas e interações respectivamente, além das hipóteses análogas, (*), quando se tem caselas vazias, conforme mostra a Tabela 3.4., a seguir.

Tabela 3.4 Análise de variância de experimentos com dois fatores e amostras desbalanceadas, segundo o método de quadrados de médias ponderadas de YATES (1934).

VARIAÇÕES CONSIDERADAS	SOMAS DE QUADRADOS	HIPÓTESES TESTADAS
Linhas (A)	$\sum_i w_i (\bar{x}_{i.} - \bar{x}_{[1]})^2 = R(\alpha \mid \mu, \beta, \gamma)$	$H_0^{(3)}$
Colunas (B)	$\sum_j v_j (\bar{x}_{.j} - \bar{x}_{[2]})^2 = R(\beta \mid \mu, \alpha, \gamma)$	$H_0^{(6)}$
Interação (AB)	$R(\gamma \mid \mu, \alpha, \beta) = R(\gamma \mid \mu, \alpha, \beta)$	$H_0^{(7)}$
Resíduo	$\sum_{ijk} (\bar{x}_{ijk} - \bar{x}_{ij.})^2$	

onde,

$$w_i = \left[\frac{1}{b_i^2} \quad \sum_j \frac{1}{n_{ij}} \right]^{-1} ; \quad v_j = \left[\frac{1}{a_j^2} \quad \sum_i \frac{1}{n_{ij}} \right]^{-1}$$

$$\bar{x}_{[1]} = \sum_i \frac{\bar{w}_i x_{i.}}{w_i} ; \quad \bar{x}_{[2]} = \sum_j \frac{\bar{v}_j x_{.j}}{v_j}$$

3.3.4. O método do ajustamento de constantes

Este método foi descrito por YATES (1934) e posteriormente discutido por ANDERSON & BANGROFT (1952), STEEL & TORRIE (1960), RAO (1965), SNEDECOR & COCHRAN (1967), CHAKRAVARTI, LAHA & ROY (1967), SEARLE (1971), WINER (1971), DAGNELIE (1973), HOCKING (1985) e IEMMA (1991) entre outros.

Consiste este método de uma parametrização sucessiva do modelo (2.3) e uso da notação $R(\cdot)$. Utiliza as hipóteses $H_o^{(1)}$, $H_o^{(2)}$, $H_o^{(8)}$, $H_o^{(9)}$ e suas análogas (*) para testar os efeitos de linhas, $H_o^{(4)}$, $H_o^{(5)}$, $H_o^{(11)}$, $H_o^{(12)}$ e suas análogas para testar os efeitos de colunas, e $H_o^{(7)}$ e similares para testar as interações, como mostra a Tabela 3.5.

Tabela 3.5 Análise de variância de experimentos com dois fatores e amostras desbalanceadas, segundo o método do ajustamento de constantes de YATES (1934).

VARIAÇÕES CONSIDERADAS	SOMAS DE QUADRADOS	HIPÓTESES TESTADAS
Linhas (não ajustadas)	$R(\alpha \mid \mu)$	$H_0^{(1)}$ e $H_0^{(8)}$
Linhas (ajustadas)	$R(\alpha \mid \mu, \beta)$	$H_0^{(2)}$ e $H_0^{(9)}$
Colunas (não ajustadas)	$R(\beta \mid \mu)$	$H_0^{(4)}$ e $H_0^{(11)}$
Colunas (ajustadas)	$R(\beta \mid \mu, \alpha)$	$H_0^{(5)}$ e $H_0^{(12)}$
Interação	$R(\gamma \mid \mu, \alpha, \beta)$	$H_0^{(7)}$
Resíduo	$\sum_{ijk} (x_{ijk} - \bar{x}_{ij.})^2$	

3.3.5. Os métodos de Henderson.

HENDERSON (1953) propôs três métodos para gerar somas de quadrados para serem usadas na estimação de componentes de variância em modelos de efeitos aleatórios. Desde que estas somas de quadrados podem também ser usadas para testar hipóteses no modelo de efeito fixo, SPEED, HOCKING & HACKNEY (1978) incorporaram os métodos 1 e 3 aos métodos utilizados na análise de variância de dados desbalanceados.

3.3.5.1. Método 1 de Henderson (H1).

Analogamente ao método do ajustamento de constantes de Yates, as somas de quadrados dos efeitos principais são $R(\alpha|\mu)$ e $R(\beta|\mu)$ e portanto, suas hipóteses associadas para efeito de linhas e colunas são aquelas descritas na tabela 3.5. A soma de quadrados devida à interação pode ser obtida por subtração.

3.3.5.2. Método 3 de Henderson (H3).

Embora esse método possa ser usado com qualquer tabela de análise de variância, geralmente ele está associado com as somas de quadrados sugeridas pelo método de ordenação a priori, ou seja, as hipóteses utilizadas são aquelas constantes da Tabela 3.6.

Tabela 3.6 Análise de variância de experimentos com dois fatores e amostras desbalanceadas, segundo o método 3 de Henderson (H3) (1953).

VARIAÇÕES CONSIDERADAS	HIPÓTESES TESTADAS	$R(\gamma)$	SOMAS DE QUADRADOS DO TIPO I
A (não ajustado)	$H_0^{(1)}$	$R(\alpha \mu)$	
B (ajustado)	$H_0^{(5)}$	$R(\beta \mu, \alpha)$	
AB	$H_0^{(7)}$	$R(\gamma \mu, \alpha, \beta)$	
Resíduo		$\sum_{ijk} (\bar{y}_{ijk} - \bar{y}_{ij.})^2$	

3.3.6. Os métodos de Overall e Spiegel (1969).

OVERALL & SPIEGEL (1969) propuseram três métodos para análise de dados desbalanceados que podem ser apropriados sob diferentes circunstâncias. Eles generalizaram e formalizaram os principais métodos até então, destacando-se o fato que esses três métodos foram incorporados em 1979 pelo pacote "SAS", muito embora eles tenham também sido propostos anteriormente por outros autores como Yates e Henderson, dentre outros.

3.3.6.1. Método dos mínimos quadrados completo (OS1).

Esse método corresponde ao método dos quadrados de médias ponderadas de Yates, e as somas de quadrados e hipóteses associadas são análogas aquelas da Tabela 3.4.

3.3.6.2. Método do planejamento experimental (OS2).

Corresponde ao método do ajustamento de constantes de Yates, e suas somas de quadrados e hipóteses associadas são aquelas expressas na Tabela 3.5. Por outro lado, os autores sugerem que o método do planejamento experimental é a generalização adequada da análise para dados balanceados.

3.3.6.3. Método da ordenação a priori (OS3).

Esse é também um caso especial do método do ajustamento de constantes no qual se usa $R(\alpha|\mu)$ e $R(\beta|\mu, \alpha)$ como somas de quadrados dos efeitos principais para linhas e colunas respectivamente, podendo-se usar alternativamente $R(\beta|\mu)$ e $R(\alpha|\mu, \beta)$. Ressalte-se ainda que, quando se tem dados desbalanceados aconselha-se o uso desse método, com a ordenação de maior interesse para o pesquisador.

3.3.7. O método Iterativo de Hemmerle.

Esse método foi proposto por HEMMERLE (1980) para calcular $R(\beta_1|\beta_2)$. Os cálculos são executados iterativamente; conseqüentemente, as quantias obtidas são aproximações de somas de quadrados. As hipóteses testadas dependem da partição do vetor de parâmetros β em β_1 e β_2 e é baseada na escolha de condições de não estimabilidade. As hipóteses testadas pelo programa de Hemmerle são análogas as constantes da Tabela 3.4.

3.4. Performance de "softwares" com relação a dados desbalanceados.

Sob este tópico, estudam-se alguns dos princi-

país pacotes estatísticos usados no setor agropecuário e nas ciências biológicas e sociais. Pretende-se estabelecer aqui, de forma clara, as hipóteses associadas às somas de quadrados envolvidas nesses pacotes e apresentar uma orientação precisa de como comparar análises estatísticas baseadas em saídas de diferentes pacotes. Utilizam-se como fonte os manuais disponíveis dos pacotes citados na introdução e adota-se a metodologia de IEMMA (1991), SEARLE (1987), HOCKING (1985) e SPEED, HOCKING e HACKNEY (1978).

3.4.1. GLM DO SAS (1986)

Para uma perfeita interpretação da rotina do GLM (General Linear Models) do SAS (Statistical Analysis System), faz-se necessário um domínio perfeito sobre os conceitos de estimabilidade e testabilidade. Aqui, face aos objetivos deste trabalho, serão abordados os quatro tipos de funções estimáveis e as hipóteses a eles associadas. Maiores detalhes podem ser encontrados no manual

a) Somas de Quadrados do Tipo I.

As somas de quadrados do tipo I são obtidas através do método da ordenação a priori (H_g ou OS3). Assim sendo, elas são dependentes da ordem dos parâmetros no mode-

lo. Cada efeito é ajustado somente para os efeitos precedentes no modelo. Consequentemente elas são usadas para testar hipóteses sobre médias ponderadas ajustadas ou não. (Tabelas 3.6 e 3.7).

Tabela 3.6 Análise de variância de experimentos, através de um modelo com dois fatores e interação. Somas de quadrados e hipóteses do tipo I para a ordenação A, B, A*B.

VARIAÇÕES CONSIDERADAS	HIPÓTESES TESTADAS	$R(\gamma)$ SOMAS DE QUADRADOS DO TIPO I
A (não ajustado)	$H_0^{(1)}$	$R(\alpha \mid \mu)$
B (ajustado)	$H_0^{(5)}$	$R(\beta \mid \mu, \alpha)$
A*B	$H_0^{(7)}$	$R(\gamma \mid \mu, \alpha, \beta)$

Tabela 3.7 Análise de variância de experimentos, através de um modelo com dois fatores e interação. Somas de quadrados e hipóteses do tipo I para a ordenação B, A, A*B.

VARIAÇÕES CONSIDERADAS	HIPÓTESES TESTADAS	$R(\gamma)$ SOMAS DE QUADRADOS DO TIPO I
B (não ajustado)	$H_0^{(4)}$	$R(\beta \mid \mu)$
A (ajustado)	$H_0^{(2)}$	$R(\alpha \mid \mu, \beta)$
A*B	$H_0^{(7)}$	$R(\gamma \mid \mu, \alpha, \beta)$

b) Somas de Quadrados do Tipo II.

As somas de quadrados tipo II correspondem a notação $R(\)$ em que cada efeito, é ajustado para todos os outros possíveis efeitos, exceto para interações. Desse modo, as S.Q. do tipo II são da forma $R(\alpha \mid \mu, \beta)$ e $R(\beta \mid \mu, \alpha)$. Ressalte-se ainda que, enquanto para as S.Q. do tipo I tem-se:

$$\begin{aligned} R(\mu) + R(\alpha \mid \mu) + R(\beta \mid \mu, \alpha) + R(\gamma \mid \mu, \alpha, \beta) &= \\ R(\mu) + R(\beta \mid \mu) + R(\alpha \mid \mu, \beta) + R(\gamma \mid \mu, \alpha, \beta) &= \\ &= R(\mu, \alpha, \beta, \gamma) \end{aligned}$$

para as S.Q. do tipo II, isto em geral não ocorre, ou seja, $R(\mu) + R(\alpha \mid \mu) + R(\beta \mid \mu, \alpha) + R(\gamma \mid \mu, \alpha, \beta) \neq R(\mu, \alpha, \beta, \gamma)$, devido ao fato que as somas de quadrados do tipo II nem sempre são provenientes de uma partição ortogonal da S.Q. dos parâmetros. (Tabela 3.8).

Tabela 3.8 Análise de variância de experimentos, através de um modelo com dois fatores e interação. Somas de quadrados e hipóteses do tipo II para a ordenação A - B - A*B.

VARIAÇÕES CONSIDERADAS	HIPÓTESES TESTADAS	$R(\)$ SOMAS DE QUADRADOS DO TIPO II
A (ajustado)	$H_0^{(2)}$	$R(\alpha \mid \mu, \beta)$
B (ajustado)	$H_0^{(5)}$	$R(\beta \mid \mu, \alpha)$
A*B	$H_0^{(7)}$	$R(\gamma \mid \mu, \alpha, \beta)$

c) Somas de Quadrados do Tipo III.

As somas de quadrados do tipo III podem ser obtidas, dentre outros, através dos métodos dos quadrados de médias ponderadas, Y_2 , de YATES (1934) ou dos mínimos quadrados completos, OS1 de OVERALL & SPIEGEL (1969), ou da inversa da parte da inversa de SEARLE (1971). Nos modelos fatoriais, quando as caselas são vazias, a S.Q. do tipo III coincide com aquelas produzidas pelos programas de modelos lineares de efeitos fixos de HARVEY (veja o método de HARVEY (1987) no SUIGI ("Supplemental Library Uses Guide", versions 5 edition). As hipóteses associadas a essas somas de quadrados são aquelas que constam da Tabela 3.9 e dizem respeito às médias não ponderadas de linhas e /ou de colunas.

Tabela 3.9 Análise de variância de experimentos, através de um modelo com dois fatores e interação. Somas de quadrados e hipóteses do tipo III para a ordenação A - B - A*B.

VARIAÇÕES CONSIDERADAS	HIPÓTESES TESTADAS	$R(\cdot)$ SOMAS DE QUADRADOS DO TIPO III
A	$H_O^{(3)}$	$R(\dot{\alpha} \mid \dot{\mu}, \dot{\beta}, \dot{\gamma})$
B	$H_O^{(6)}$	$R(\dot{\beta} \mid \dot{\mu}, \dot{\alpha}, \dot{\gamma})$
A*B	$H_O^{(7)}$	$R(\dot{\gamma} \mid \dot{\mu}, \dot{\alpha}, \dot{\beta}) = R(\dot{\gamma} \mid \dot{\mu}, \dot{\alpha}, \dot{\beta})$

NOTA. Para a ordenação $B - A - A*B$, os resultados são os mesmos, da Tabela 3.9, mudando apenas a ordem de apresentação, isto é, $R(\beta \mid \mu, \alpha)$, $R(\alpha \mid \mu, \beta)$ e $R(\gamma \mid \mu, \alpha, \beta)$.

d) Somas de Quadrados do Tipo IV.

As somas de quadrados do tipo I, II, III e IV são derivadas ao se ajustar um modelo e diferentes sub-modelos. Por outro lado, as S.Q. do tipo IV testam hipóteses que são geradas automaticamente pelo GLM do SAS, que podem usualmente ser interpretadas. Contudo, uma interpretação apropriada não pode ser feita sem primeiro examinar as funções estimáveis do tipo IV para se ver que hipóteses são geradas e testadas. Isto é, não existe uma única interpretação apropriada para todo o conjunto de dados. Se não existem caselas vazias, as somas de quadrados do tipo IV são similares as do tipo III; se no entanto existe pelo menos uma casela vazia, então as somas de quadrados dos tipos III e IV são em geral diferentes e referem-se a diferentes hipóteses.

Observa-se ainda que, se existem caselas vazias, as somas de quadrados do tipo IV podem não ser únicas, vez que elas dependem da posição e do número de caselas vazias. Geralmente, IEMMA (1991), pode-se interpretar as hipóteses do tipo IV como hipóteses sobre contrastes entre médias populacionais das caselas que estão na mesma coluna

(linha) iniciando-se a comparação sempre pela última linha (coluna). Como exemplo, vejam-se os conjuntos de dados I, II

Conjunto I

	j = 1	j = 2	j = 3	j = 4
i = 1	-	μ_{12}	μ_{13}	μ_{14}
i = 2	μ_{21}	μ_{22}	-	-
i = 3	μ_{31}	-	μ_{33}	μ_{34}

Conjunto II

	j = 1	j = 2	j = 3	j = 4
i = 1	-	μ_{12}	μ_{13}	μ_{14}
i = 2	μ_{31}	-	μ_{33}	μ_{34}
i = 3	μ_{21}	μ_{22}	-	-

Verifica-se de imediato, que se trata de um único conjunto de dados, vez que o conjunto II foi obtido permutando-se as linhas 2 e 3 do conjunto I. É de se esperar portanto, por exemplo, que as hipóteses sobre linhas sejam iguais para os dois conjuntos. Contudo, para hipóteses do tipo IV tem-se:

Para o conjunto I

$$H_0: \begin{cases} \mu_{33} + \mu_{34} = \mu_{13} + \mu_{14} \\ \mu_{31} = \mu_{21} \end{cases}$$

Para o conjunto II

$$H_0: \begin{cases} \mu_{22} = \mu_{12} \\ \mu_{21} = \mu_{31} \end{cases}$$

que são hipóteses diferentes para testar os efeitos de linha do mesmo conjunto de dados. Consequentemente, a essas hipóteses estão associadas diferentes somas de quadrados.

3.4.2. SOC (SISTEMA OPERACIONAL C) (1982)

Através do procedimento MODLIN (Modelo Linear) o SOC analisa dados por meio de ajuste de modelos lineares univariados utilizando o método dos mínimos quadrados. Calcula dois tipos de somas de quadrados: sequencial e parcial.

a) Somas de Quadrados Sequenciais.

As somas de quadrados sequenciais são obtidas através de modelos construídos com a introdução sequencial dos efeitos e correspondem às do tipo I. (Tabela 3.6).

Exemplo: $y_{ijk} = \mu + \alpha_i + e_{ijk}$

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

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

b) Somas de Quadrados Parciais.

As somas de quadrados parciais são obtidas a partir do ajuste do modelo com todos os efeitos. Obtida a estimativa do vetor de parâmetros, essas somas de quadrados são calculadas através da expressão geral

$$(Cb)'[C'(X'X)^G C]^{-1} (Cb)$$

onde C é uma matriz apropriada, de dimensão $(s \times p)$, com posto $s \leq p$ que é utilizada para descrever hipóteses de interesse na forma geral

$$H_0 : CB = \phi.$$

Como consequência desta forma de obtenção de soma de quadrados, a soma de quadrado parcial para cada efeito do modelo fica ajustada pelos demais efeitos. Para certas estruturas de dados, como por exemplo, dados completos e balanceados, os valores desses dois tipos de somas de quadrados coincidem, ao passo que quando se trata de dados desbalanceados, elas são geralmente diferentes.

Levando-se em consideração um modelo com dois fatores e interação, nas ordens A , B e AB , o SOG fornece somas de quadrados equivalentes à soma de quadrados tipo I do SAS, bem como somas de quadrados tipo II, tipo III e tipo IV iguais a do SAS.

3.4.3. SYSTAT (THE SYSTEM FOR STATISTICS) (1986)

Para dados desbalanceados, esse software fornece através da sua saída "Análise de Variância" somas de quadrados equivalentes às somas de quadrados dos tipos I, II, III e IV do SAS, para modelos com dois fatores sem interação. Entretanto, em modelos com caselas vazias e interação, o processo é interrompido e nenhuma informação adicional é fornecida.

3.4.4. STATGRAPHICS (STATISTICAL GRAPHICS SYSTEM) (1985)

O procedimento ANOVA (Analysis of variance) para experimento desbalanceados inclui em sua saída os resultados padrões para os efeitos principais e interações. A soma total de quadrados representa a soma de contribuições de todos os efeitos principais no modelo, excluindo-se os termos das interações. A contribuição de cada efeito principal é também estimada com a pressuposição de que todos os outros efeitos principais tenham, de fato, sido incluídos no modelo. Usando este resultado, pode-se determinar se um fator individual pode ser removido do modelo. Ressalte-se porém, que na presença de caselas vazias, no modelo sem interação, os resultados gerados por esse "software" são equivalentes às somas de quadrados dos tipos II, III e IV do SAS,

nas ordens A, B e B, A. Contudo, na presença da interação A*B, o mesmo não executa.

3.4.5. SAEF (SISTEMA PARA ANÁLISES ESTATÍSTICAS) (1986)

O procedimento SAEF para dados desbalanceados e para modelos com dois fatores e interação não produz saída. Entretanto, para o modelo com dois fatores sem interação produz somas de quadrados equivalentes as somas de quadrados tipo II = tipo III = tipo IV do SAS.

3.4.6. MSTAT (A MICROCOMPUTER FOR THE DESIGN, MANAGEMENT AND ANALYSIS OF AGRONOMIC RESEARCH EXPERIMENTS) (1984)

Através do procedimento NONORTH (NONORTHONAL ANALYSIS OF VARIANCE), dados desbalanceados são analisados por este "software". Sua saída fornece cálculo da frequência, soma e ou médias para uma ou mais variáveis em uma tabela dois por dois. Para tabelas com dados não ortogonais, médias ajustadas em ambas as direções podem ser calculadas junto com uma análise de variância, fornecendo quadrados médios ajustados para ambos os grupos (fatores), interações e dentro de caselas.

No modelo com dois fatores e interação produz

somas de quadrados equivalentes às somas de quadrados do tipo I e II do SAS.

3.4.7. MINITAB (MINITAB STATISTICAL SOFTWARE) (1986)

O software MINITAB através do procedimento GLM ajusta o modelo linear geral. Assim, ele pode ser usado para fazer análise de variância de experimentos balanceados e não balanceados. Faz testes para modelos de efeitos fixos automaticamente. O subcomando TEST permite calcular o teste-F apropriado para modelos mistos. Sua saída contém somas de quadrados sequenciais e somas de quadrados ajustadas. Automaticamente, os testes são feitos usando somas de quadrados ajustadas, assumindo todos os outros fatores fixados. Essas somas de quadrados são rotuladas de MT1 e MT2 respectivamente, e são equivalentes as somas de quadrados do tipo I e tipo II do SAS.

3.4.8. BMDP (BIOMEDICAL STATISTIC PACKAGE) (1983)

O BMDP10 é o primeiro programa BMDP que trata da análise de dados não balanceados de uma forma razoavelmente conveniente. A primeira versão do BMDP10 foi desenvolvida em início dos anos 60 com o nome de BMDX64.

O modelo utilizado no BMDP10 é:

$$y = X\beta + e,$$

onde X é uma matriz ($n \times p$), incluindo as covariáveis, as variáveis binárias definindo os índices da análise de variação. Suponha que se tenha dois fatores, A e B, onde A e B são cruzados. Os possíveis componentes da análise de variação são:

Constante

A

B

A*B

O termo (A*B) representando a interação é opcional. A variável binária para a média é um vetor de 1's. Seja A com dois níveis. Então seu efeito principal é definido pela variável binária:

$$D_{A1} = \begin{cases} 1 & \text{se } A=1 \\ -1 & \text{se } A=2 \end{cases}$$

Seja B com três níveis. Seus efeitos principais são definidos pelas duas variáveis :

$$D_{B1} = \begin{cases} 1 & \text{se } B=1 \\ 0 & \text{se } B=2 \\ -1 & \text{se } B=3 \end{cases} \quad D_{B2} = \begin{cases} 0 & \text{se } B=1 \\ 1 & \text{se } B=2 \\ -1 & \text{se } B=3 \end{cases}$$

Todas as variáveis e as covariáveis definem as colunas de X .

Existem muitas filosofias acerca dos testes de

hipóteses na ANOVA de classificação dupla ou superiores. A filosofia utilizada no BMD e BMDP é mais facilmente descrita em termos das médias populacionais das caselas. Num esquema fatorial 2×3 , representa-se as médias das caselas por:

$$\begin{array}{ccc} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \end{array}$$

As hipóteses do BMDP são independentes das frequências das caselas. Para dados experimentais, as hipóteses usualmente de interesse são:

$$\text{Efeito de linhas: } \mu_{11} + \mu_{12} + \mu_{13} = \mu_{21} + \mu_{22} + \mu_{23}$$

$$\text{Efeito colunas : } \mu_{11} + \mu_{21} = \mu_{12} + \mu_{22} = \mu_{13} + \mu_{23}$$

$$\text{Interação : } \mu_{11} + \mu_{22} = \mu_{12} + \mu_{21}$$

$$\mu_{11} + \mu_{23} = \mu_{13} + \mu_{21}$$

As hipóteses testadas no BMDP para número desigual de observações são as mesmas testadas no caso de número constante apresentadas em muitos textos e programas de computador. O modelo do BMDP, nesse caso, pode ser calculado numa tabela de ANOVA como:

$$E(y_{ijk}) = \mu_{ij} \quad (a)$$

ou

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad (b)$$

com as restrições "usuais":

$$\sum \alpha_i = \sum \beta_j = \sum \gamma_{ij} = 0.$$

ou

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad (c)$$

com as restrições "usuais":

$$\sum \alpha_i = \sum \beta_j = \sum \gamma_{ij} = 0.$$

Observe que as expressões (a), (b) e (c) são três formas diferentes de expressar o mesmo modelo. Quando (b) é utilizado, as hipóteses podem ser escritas como:

Efeito de linha : $\alpha_i = 0$

Efeito de colunas: $\beta_j = 0$

Interação : $\gamma_{ij} = 0$

Observe ainda que as hipóteses acima são as mesmas colocadas em termos das médias populacionais μ_{ij} .

3.4.9. HARVEY (MIXED MODEL LEAST-SQUARES AND MAXIMUM LIKELIHOOD COMPUTER PR.) (1987)

O "software" HARVEY, através do procedimento LSMLMW leva em conta o tipo de modelo, se é fixo, aleatório ou misto. Foi elaborado por geneticistas e seu objetivo principal é praticamente específico para soluções de problemas na área de genética. Seus modelos podem envolver covariáveis, além do que, estimam os componentes de variância dos efeitos aleatórios envolvidos no modelo. Associam somas de quadrados a algumas hipóteses de interesse.

A saída rotulada "CONSTANT ESTIMATES" fornece uma solução das equações normais que satisfaz as restrições- Σ . Testa as hipóteses sobre linhas e colunas. Para testar interações no modelo com dois fatores cruzados, é necessário que pelo menos uma linha ou uma coluna seja completa, do contrário, sua saída não produz a soma de quadrados relativa à interação.

3.4.10. SPSS (STATISTICAL PACKAGE) (1986)

Nenhuma solução das equações normais no seu output (saída), relativo à análise de variância, mas uma seção originalmente rotulada "Multiple Classification Analysis" sob o título "Adjusted for Independents" e subtitulada DEV'N, contém um vetor de solução das equações normais depois de retirar as interações do modelo. Tão bom como sendo para esta situação de não interação, os elementos da solução satisfazem também as restrições de ponderação, tais como, $\sum n_{i.} e_i^0 = 0$. Nota-se que, mesmo quando a parte da análise de variância do "output" é para um modelo com interação, elas são omitidas dos cálculos do "output" intitulado DEV'N como foi descrito.

O SPSS usa o Modelo- Σ e exige que os dados tenham pelo menos uma linha e uma coluna com todas as células ocupadas, e esta, deve ser a primeira linha e a primeira

coluna. Para todas as caselas ocupadas, as somas de quadrados tipo III e IV são as mesmas.

3.4.11. GLIM (GENERALISED LINEAR INTERACTIVE MODELING SYSTEM) (1984)

O GLIM através da sua saída fornece somente um tipo de soma de quadrados, na forma sequencial, ajustando cada fator para os precedentes no modelo. Corresponde a soma de quadrados do tipo I do SAS. Sua saída constitui-se de um único quadro para os modelos com e sem interação, e a diferença entre eles, constitui-se na inclusão ou não da hipótese $H_0^{(?)}$ (interação).

3.4.12. SANEST (SISTEMA DE ANÁLISE ESTATÍSTICA) (1986)

Considerando um modelo com dois fatores, através da saída do SANEST, pode-se notar que a ordem com que se entra com os fatores A e B do modelo no programa é irrelevante. As somas de quadrados para os fatores A e B apresentam-se segundo as opções:

Na ordem A , B (modelo sem interação)

B não ajustado

B ajustado

Na ordem B , A (modelo sem interação)

B não ajustado

B ajustado

Na ordem A , B (modelo com interação)

B não ajustado

B ajustado

Na ordem B , A (modelo com interação)

B não ajustado

B ajustado,

e correspondem as somas de quadrados do tipo I e II do SAS.

OBS.

- 1) O fator A é ajustado automaticamente, ao passo que, para o fator B, faz-se a opção pelo ajuste ou não.
- 2) O SANEST trabalha no máximo com dois fatores.

4. EXEMPLO NUMÉRICO

Objetivando ilustrar a metodologia aplicada, utilizou-se dos dados provenientes do exemplo simulado a seguir.

Num ensaio com adubação fosfatada em cana-de-açúcar (*saccharum officinarum* L.) consorciada com feijão (*phaseolus*) vigna, foram consideradas quatro variedades (V1 = CB45-3; V2 = CB45-15; V3 = C0997 e V4 = C0331) e três níveis de fósforo em Kg/ha (P0 = 0; P1 = 60 e P2 = 120).

Com o objetivo precípua de estudar o efeito das variedades de cana-de-açúcar e dos níveis de fósforo (P) sobre a produção de feijão vigna, foram observadas (em tonelada/ha) os dados constantes da Tabela 4.1.

Tabela 4.1 Produção em toneladas/ha de feijão vigna num experimento com adubação fosfatada em cana-de-açúcar consorciada com feijão vigna.

	j = 1 VARIED 1	j = 2 VARIED 2	j = 3 VARIED 3	j = 4 VARIED 4	
i = 1	10	-	14	9	
P ₀ = 0	15			13	
	11				
	36(3)12	-	14(1)14	22(2)11	72(6)12
i = 2	8	14	-	12	
P ₁ = 60	14	16		14	
				13	
	22(2)11	30(2)15	-	39(3)13	91(7)13
i = 3	-	11	16	-	
P ₂ = 120			18		
	-	11(1)11	34(2)17	-	45(3)15
	58(5)11.6	41(3)13.7	48(3)16	61(5)12.2	208(16)13

Conforme é bem sabido, em geral as hipóteses usuais de interesse são aquelas sobre:

- i) efeitos de linhas (fator A = níveis de fósforo)
- ii) efeitos de colunas (fator B = variedades de feijão vigna)
- iii) efeito da interação A * B.

No entanto, essas hipóteses ainda não foram formuladas, em face dos dados disponíveis, e podem surgir algumas questões, tais como:

- a) Elas são independentes das frequências das caselas?
- b) Perdas acidentais de unidades experimentais alteram as hipóteses a serem testadas?
- c) Qual o modelo apropriado para análise dos dados?
- d) Que "softwares" (pacotes) realizam essa tarefa?

Objetivando responder as questões acima, adotar-se-á para análise dos dados da Tabela 4.1 o modelo (2.3) definido no item 2, nas formas S , M e Σ , através de suas caracterizações.

4.1. O Modelo Superparametrizado (S)

4.1.1. Com interação na ordem A - B - A*B

Seja modelo (2.3), $y = Xe + e$, caracterizado por:

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

onde,

$$i = 1, 2, 3 ;$$

$$j = 1, 2, 3, 4 ;$$

$$k = 1, \dots, n_{ij}$$

e $n_{ij} \geq 0$ (repetições dentro das caselas)

Expresso na forma matricial de (2.3) fica:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{131} \\ y_{141} \\ y_{142} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{241} \\ y_{242} \\ y_{243} \\ y_{321} \\ y_{331} \\ y_{332} \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 11 \\ 14 \\ 9 \\ 13 \\ 8 \\ 14 \\ 14 \\ 16 \\ 12 \\ 14 \\ 13 \\ 11 \\ 16 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \theta + e$$

(4.1)

4.1.1.1. Equações normais e suas soluções

Para o exemplo em questão, o sistema de equações normais tem a forma:

$$X'X\theta = X'y$$

onde, $\theta' = (\theta_1, \theta_2, \theta_3, \theta_4)$ e

$\theta_1 = \mu$; $\theta_2' = (\mu, \alpha_1, \alpha_2, \alpha_3)$; $\theta_3' = (\mu, \beta_1, \beta_2, \beta_3, \beta_4)$;

$\theta_4' = (\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \beta_4)$, e a matriz X também é

particionada seguindo os e's em X_1 (coluna 1 de X), X_2 (colunas 1 a 4 de X), X_3 (colunas 1, 5, 6, 7 e 8 de X) e X_4 (colunas 1 a 8 de X), que em termos numéricos fica conforme segue:

16	6 7 3	5 3 3 5	3 1 2 2 2 3 1 2	μ	208
6	6 0 0	3 0 1 2	3 1 2 0 0 0 0 0	α_1	72
7	0 7 0	2 2 0 3	0 0 0 2 2 3 0 0	α_2	91
3	0 0 3	0 1 2 0	0 0 0 0 0 0 1 2	α_3	45
5	3 2 0	5 0 0 0	3 0 0 2 0 0 0 0	β_1	58
3	0 2 1	0 3 0 0	0 0 0 0 2 0 1 0	β_2	41
3	1 0 2	0 0 3 0	0 1 0 0 0 0 0 2	β_3	48
5	2 3 0	0 0 0 5	0 0 2 0 0 3 0 0	β_4	61
3	3 0 0	3 0 0 0	3 0 0 0 0 0 0 0	γ_{11}	36
1	1 0 0	0 0 1 0	0 1 0 0 0 0 0 0	γ_{13}	14
2	2 0 0	0 0 0 2	0 0 2 0 0 0 0 0	γ_{14}	22
2	0 2 0	2 0 0 0	0 0 0 2 0 0 0 0	γ_{21}	22
2	0 2 0	0 2 0 0	0 0 0 0 2 0 0 0	γ_{22}	30
3	0 3 0	0 0 0 3	0 0 0 0 0 3 0 0	γ_{24}	39
1	0 0 1	0 1 0 0	0 0 0 0 0 0 1 0	γ_{32}	11
2	0 0 2	0 0 2 0	0 0 0 0 0 0 0 2	γ_{33}	34

(4.2)

Como se está tratando de um problema, onde X não é de posto coluna completo, a matriz $X'X$ não é não singular e o sistema de equações normais, sempre consistente, é indeterminado. Suas soluções exatas entre outras podem ser

$$^1R(\mu) = \Theta_1^0 X_1' y = 2704 \quad (4.7)$$

$$R(\mu, \alpha) = \Theta_2^0 X_2' y = 2722 \quad (4.8)$$

Assim, com base em (4.7) e (4.8) pode-se obter a soma de quadrados de linhas, ajustadas para μ , ignorando-se β e γ

$$R(\alpha \mid \mu) = R(\mu, \alpha) - R(\mu) = 2722 - 2704 = 18 \quad (4.9)$$

De modo análogo, obtém-se

$$R(\mu, \beta) = \Theta_3^0 X_3' y = 2745,33 \quad (4.10)$$

$$R(\beta \mid \mu) = R(\mu, \beta) - R(\mu) = 41,33 \quad (4.11)$$

$$R(\mu, \alpha, \beta) = \Theta_4^0 X_4' y = 2750,23 \quad (4.12)$$

$$R(\alpha \mid \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta) = 4,9 \quad (4.13)$$

$$R(\beta \mid \mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha) = 28,23 \quad (4.14)$$

$$R(\mu, \alpha, \beta, \gamma) = \Theta^0 X' y = 2768 \text{ como em (4.4)}$$

$$R(\mu, \alpha, \beta, \gamma \mid \mu) = R(\mu, \alpha, \beta, \gamma) - R(\mu) = 64 \quad (4.15)$$

$$R(\gamma \mid \mu, \alpha, \beta) = R(\mu, \alpha, \beta, \gamma) - R(\mu, \alpha, \beta) = 17,77 \quad (4.16)$$

4.1.1.3 Análise de Variância.

Do exposto na seção 3.1, depreende-se que existem várias formas de se fazer uma análise de variância, e isto fica mais claro, após a obtenção das diversas somas

¹ Observe que (4.4) é a soma de quadrados devida a todos os parâmetros descritos no modelo adotado. Através de parametrizações sucessivas e de subtrações pode-se obter as somas de quadrados de interesse.

de quadrados, descritas no item anterior. Também, a ordem em que se toma o modelo diz respeito a forma da análise. As tabelas a seguir foram elaboradas guardando este aspecto, bem como aos diversos "output" dos pacotes constantes do apêndice.

Tabela 4.2 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores e interação. (Comum às ordens A-B-A*B e B-A-A*B).

C.V	g.l.	S.Q.		Q.M.	F
Média	1	$R(\mu)$	= 2704	2704,00	470,26
Modelo	$s-1=7$	$R(\alpha, \beta, \gamma \mu)$	= 64	9,14	1,59
Resíduo	$n-s=8$	S.Q.Res.	= 46	5,75	
Total	$n=16$	.Q.T.	= 2814		

a.1) Ordem A - B - A*B

$n = 16, s = \rho[X] = 8$

Tabela 4.3 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores e interação na ordem A - B - A*B.

C.V	g.l.	S.Q.		Q.M.	F
Média	1	$R(\mu)$	= 2704,00	2704,00	470,26
$(\alpha \mu)$	$a-1=2$	$R(\alpha \mu)$	= 18,00	9,00	1,56
$(\beta \mu, \alpha)$	$b-1=3$	$R(\beta \mu, \alpha)$	= 28,33	9,11	1,58
$(\gamma \mu, \alpha, \beta)$	$s-a-b+1=2$	$R(\gamma \mu, \alpha, \beta)$	= 17,77	8,85	1,54
Resíduo	$n-s=8$	S.Q.Res.	= 46,00	5,75	
Total	$n=16$	S.Q.T.	= 2814,00		

a.2) Ordem B - A - A*B

Tabela 4.4 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores e interação, na ordem B - A - A*B

C.V	g.l.	S.Q.		Q.M.	F
Média	1	R (μ)	= 2704,00	2704,00	470,3
($\beta \mu$)	b-1=3	R ($\beta \mu$)	= 41,33	13,78	2,4
($\alpha \mu,\beta$)	a-1=2	R ($\alpha \mu,\beta$)	= 4,90	2,49	0,4
($\gamma \mu,\alpha,\beta$)	s-a-b+1=2	R ($\gamma \mu,\alpha,\beta$)	= 17,77	8,85	1,5
Resíduo	n-s=8	S.Q. Res.	= 46,00	5,75	
Total	n=16	S.Q.T.	= 2814,00		

4.1.2. Sem interação na ordem A - B.

Neste modelo, caracterizado por:

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

matricialmente, tem-se:

$$y = X_4 \theta_4 + e,$$

onde, X_4 é formada pelas colunas 1 a 8 da matriz X no modelo anterior e,

$$\theta_4' = [\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \beta_4]$$

como anteriormente.

4.1.2.1. Equações normais e suas soluções

As equações normais $X_4' X_4 \theta = X_4' y$, que numericamente em termos do exemplo em questão fica:

$$\begin{bmatrix} 16 & 6 & 7 & 3 & 5 & 3 & 3 & 5 \\ 6 & 6 & 0 & 0 & 3 & 0 & 1 & 2 \\ 7 & 0 & 7 & 0 & 2 & 2 & 0 & 3 \\ 3 & 0 & 0 & 3 & 0 & 1 & 2 & 0 \\ 5 & 3 & 2 & 0 & 5 & 0 & 0 & 0 \\ 3 & 0 & 2 & 1 & 0 & 3 & 0 & 0 \\ 3 & 1 & 0 & 2 & 0 & 0 & 3 & 0 \\ 5 & 2 & 3 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 208 \\ 72 \\ 91 \\ 45 \\ 58 \\ 41 \\ 48 \\ 61 \end{bmatrix}$$

Analogamente à caracterização anterior do modelo (2.3) X_4 não é de posto coluna $C(X)$ completo, e aqui cabe o mesmo tipo de comentário, contudo, no caso anterior, por exemplo a função paramétrica $\lambda'\theta = \alpha_1 - \alpha_2$ não é estimável e, conseqüentemente não testável, ao passo que, no presente caso ela o é. Assim.

$$\theta_4^o = (X_4' X_4)^G X_4' y$$

$$\begin{bmatrix} 0.43 & -0.28 & -0.22 & 0.00 & -0.17 & -0.28 & -1.00 & -0.18 \\ -0.28 & 0.84 & 0.66 & 0.00 & -0.49 & -0.16 & 0.00 & -0.45 \\ -0.22 & 0.66 & 0.84 & 0.00 & -0.51 & -0.34 & 0.00 & -0.55 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -0.17 & -0.49 & -0.51 & 0.00 & 0.87 & 0.51 & 1.00 & 0.67 \\ -0.28 & -0.16 & -0.34 & 0.00 & 0.51 & 0.84 & 1.00 & 0.55 \\ 1.00 & -0.00 & -0.00 & 0.00 & -1.00 & -1.00 & 0.00 & -1.00 \\ -0.18 & -0.45 & -0.55 & 0.00 & 0.67 & 0.55 & 1.00 & 0.89 \end{bmatrix} \begin{bmatrix} 208 \\ 72 \\ 91 \\ 45 \\ 58 \\ 41 \\ 48 \\ 61 \end{bmatrix}$$

ou,

$$\mathbf{e}_4^{\circ} = (\mathbf{X}_4' \mathbf{X}_4)^G \mathbf{X}_4' \mathbf{y} = \begin{bmatrix} -31.95 \\ -0.15 \\ 1.15 \\ 0.00 \\ 43.18 \\ 44.85 \\ 48.00 \\ 43.52 \end{bmatrix}$$

4.1.2.2. Somas de quadrados:

$$R(\mathbf{e}_4) = \mathbf{e}_4^{\circ} \mathbf{X}_4' \mathbf{y} = \text{S.Q. Parâmetros} = 2750.23 \quad (4.17)$$

$$\text{S.Q.Total} = \mathbf{y}'\mathbf{y} = \sum_{ijk} y_{ijk}^2 = 2814 \quad \text{como em (4.5)}$$

$$\text{S.Q.Res.} = \mathbf{y}'\mathbf{y} - \mathbf{e}_4^{\circ} \mathbf{X}_4' \mathbf{y} = 2814 - 2750.23 = 63.77 \quad (4.18)$$

$$R(\mu) = \mathbf{e}_1^{\circ} \mathbf{X}_1' \mathbf{y} = 2704 \quad \text{como em (4.7)}$$

$$R(\mu, \alpha) = \mathbf{e}_2^{\circ} \mathbf{X}_2' \mathbf{y} = 2722 \quad \text{como em (4.8)}$$

$$R(\mu, \beta) = \mathbf{e}_3^{\circ} \mathbf{X}_3' \mathbf{y} = 2745.33 \quad \text{como em (4.10)}$$

$$R(\beta \mid \mu) = R(\mu, \beta) - R(\mu) = 41.33 \quad \text{como em (4.11)}$$

$$R(\mu, \alpha, \beta) = \mathbf{e}_4^{\circ} \mathbf{X}_4' \mathbf{y} = 2750.23 \quad \text{como em (4.12)}$$

$$R(\alpha \mid \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta) = 4,9 \text{ como em (4.13)}$$

$$R(\beta \mid \mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha) = 28,23 \text{ como em (4.14)}$$

$$R(\alpha, \beta \mid \mu) = R(\mu, \alpha, \beta) - R(\alpha \mid \mu) = 46,23 \quad (4.19)$$

4.1.2.3. Análise de variância:

Tabela 4.5 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores sem interação. (Comum às ordens A-B e B-A).

C.V	g.l.	S.Q.		Q.M.	F
Média	1	$R(\mu)$	= 2704	2704,00	470,26
Modelo	$s-1=7$	$R(\alpha, \beta, \mid \mu)$	= 46,23	9,25	1,59
Resíduo	$n-s=8$	S.Q.Res.	= 63,77	6,38	
Total	$n=16$	.Q.T.	= 2814		

b.1) Ordem A - B.

Tabela 4.6 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores sem interação na ordem A - B.

C.V	g.l.	S.Q.		Q.M.	F
Média	1	$R(\mu)$	= 2704	2704	424,49
$(\alpha \mid \mu)$	$a-1=2$	$R(\alpha \mid \mu)$	= 18	9	1,41
$(\beta \mid \mu, \alpha)$	$b-1=3$	$R(\beta \mid \mu, \alpha)$	= 28,23	9,41	1,48
Resíduo	10	S.Q.Res.	= 63,77	6,37	
Total	$n=16$	S.Q.T.	= 2814		

4.2) Ordem B - A.

Tabela 4.7 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores sem interação na ordem B - A.

G.V	g.l.	S.Q.	Q.M.	F
Média	1	$R(\mu) = 2704$	2704	424,49
$(\beta \mu)$	$b-1=3$	$R(\beta \mu) = 41,33$	13,77	2,16
$(\alpha \mu\beta)$	$a-1=2$	$R(\alpha \mu\beta) = 4,90$	2,45	0,38
Resíduo	10	$S.Q. Res. = 63,77$	6,37	
Total	$n=16$	$S.Q.T. = 2814$		

4.2. O Modelo de Médias de Caselas

4.2.1 Com interação na ordem A - B - A*B

Seja o modelo (2.13), $y = W\mu + e$,

caracterizado por:

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

e expresso na forma matricial de (2.13) conforme segue:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{131} \\ y_{141} \\ y_{142} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{241} \\ y_{242} \\ y_{243} \\ y_{321} \\ y_{331} \\ y_{332} \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 11 \\ 14 \\ 9 \\ 13 \\ 8 \\ 14 \\ 14 \\ 16 \\ 12 \\ 14 \\ 13 \\ 11 \\ 16 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mu + e$$

A Matriz W nesse caso, corresponde às colunas 9 a 16 no modelo-S.

4.2.1.1. Equações normais e sua solução

As equações normais expressas por:

$$W'W \mu = W'y,$$

em termos do problema em questão, numericamente, fica:

$$\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{13} \\ \mu_{14} \\ \mu_{21} \\ \mu_{22} \\ \mu_{24} \\ \mu_{32} \\ \mu_{33} \end{bmatrix} = \begin{bmatrix} 36 \\ 14 \\ 22 \\ 22 \\ 30 \\ 39 \\ 11 \\ 34 \end{bmatrix}$$

O sistema de equações normais $W'W\mu = W'y$ é, em geral, muito mais simples que $X'X\theta = X'y$ do modelo - S. Para o sistema acima, tem-se por exemplo um sistema de solução única pois W tem posto coluna completo e então $W'W$ é positiva definida. Consequentemente, sua solução exata $\hat{\mu} = (W'W)^{-1}W'y$, que é solução aproximada para o sistema inconsistente $y = W\mu$, coincide com o estimador de mínimos quadrados para μ , ou seja, de acordo com o teorema de Gauss-Markov.

Teorema de Gauss-Markov: "Se o modelo linear geral de posto completo, $y = W\mu + e$ é tal que $E(e) = 0$ e $E(ee') = \sigma^2 I$, então, o melhor estimador linear de mínimos quadrados, não viado para μ é dado por $\hat{\mu} = (W'W)^{-1}W'y$."

Assim, o BLUE de μ_{ij} é dado por $\hat{\mu}_{ij} = \bar{y}_{ij}$, isto é, o BLUE de μ_{ij} é a média da casela (i, j) . Naturalmente, $\widehat{\lambda'\mu} = \lambda'\hat{\mu}$ como no Modelo - S e analogamente para a estimação por intervalo e por região. Logo,

$$\hat{\mu} = \begin{bmatrix} 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 36 \\ 14 \\ 22 \\ 22 \\ 30 \\ 39 \\ 11 \\ 34 \end{bmatrix} = \begin{bmatrix} \hat{\mu}_{11}^0 \\ \hat{\mu}_{13}^0 \\ \hat{\mu}_{14}^0 \\ \hat{\mu}_{21}^0 \\ \hat{\mu}_{22}^0 \\ \hat{\mu}_{24}^0 \\ \hat{\mu}_{32}^0 \\ \hat{\mu}_{33}^0 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\mu}_{11}^0 \\ \hat{\mu}_{13}^0 \\ \hat{\mu}_{14}^0 \\ \hat{\mu}_{21}^0 \\ \hat{\mu}_{22}^0 \\ \hat{\mu}_{24}^0 \\ \hat{\mu}_{32}^0 \\ \hat{\mu}_{33}^0 \end{bmatrix} = \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

4.2.1.2. Somas de quadrados.

$$R(\mu) = \phi_1^0 X_1^0 y = n(\bar{y}_{..})^2 = 2704, \text{ como em (4.7)}$$

$$R(\mu_{ij}) = R(\mu, \alpha, \beta, \gamma) = 2768$$

$$R(\mu_i) = R(\mu, \alpha) = 2722$$

$$R(\mu_j) = R(\mu, \beta) = 2745,33$$

Assim, com base nas expressões acima, pode-se obter as somas de quadrados:

$$R(\mu_i | \mu) = R(\mu_i) - R(\mu) = R(\alpha | \mu) = 2722 - 2704 = 18,$$

$$R(\mu_j | \mu) = R(\mu_j) - R(\mu) = R(\beta | \mu) = 2745,3 - 2704 = 41,3$$

e as demais somas de quadrados necessárias à formação dos quadros de análise de variância, cujos valores são análogos aos encontrados no modelo-S.

4.2.2. Sem interação na ordem A - B

Seja o modelo (2.13), $y = W \mu + e$, caracterizado por:

$$y_{ijk} = \mu_i + \beta_j + e_{ijk}$$

e expresso na forma matricial de (2.13) conforme abaixo:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{131} \\ y_{141} \\ y_{142} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{241} \\ y_{242} \\ y_{243} \\ y_{321} \\ y_{331} \\ y_{332} \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 11 \\ 14 \\ 9 \\ 13 \\ 8 \\ 14 \\ 14 \\ 16 \\ 12 \\ 14 \\ 13 \\ 11 \\ 16 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + e$$

4.2.2.1. Equações normais e suas soluções

As equações normais $W'W \mu = W'y$,

$$\begin{bmatrix} 6 & 0 & 0 & 3 & 0 & 1 & 2 \\ 0 & 7 & 0 & 2 & 2 & 0 & 3 \\ 0 & 0 & 3 & 0 & 1 & 2 & 0 \\ 3 & 2 & 0 & 5 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 3 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 3 & 0 \\ 2 & 3 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 72 \\ 91 \\ 45 \\ 58 \\ 41 \\ 48 \\ 61 \end{bmatrix}$$

e, uma solução é

$$\begin{bmatrix} 0.71 & 0.59 & 0.15 & -0.66 & -0.44 & -1.00 & -0.63 \\ 0.59 & 0.83 & 0.21 & -0.68 & -0.62 & -1.00 & -0.73 \\ 0.15 & 0.21 & 0.43 & -0.17 & -0.28 & -1.00 & -0.18 \\ -0.66 & -0.68 & -0.17 & 0.87 & 0.51 & 1.00 & 0.67 \\ -0.44 & -0.62 & -0.28 & 0.51 & 0.84 & 1.00 & 0.55 \\ 1.00 & 1.00 & 1.00 & -1.00 & -1.00 & 0.00 & -1.00 \\ -0.63 & -0.73 & -0.18 & 0.67 & 0.55 & 1.00 & 0.90 \end{bmatrix} \begin{bmatrix} 72 \\ 91 \\ 45 \\ 58 \\ 41 \\ 48 \\ 61 \end{bmatrix} =$$

$$\begin{bmatrix} \mu_1^o \\ \mu_2^o \\ \mu_3^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \\ \beta_4^o \end{bmatrix} = \begin{bmatrix} -32.10 \\ -30.80 \\ -31.95 \\ 43.18 \\ 44.85 \\ 48.00 \\ 43.52 \end{bmatrix}$$

4.2.2.2 Somas de quadrados

De modo análogo ao anterior, obtém-se:
 $R(\mu_i | \beta_j) = R(\alpha | \mu) = 2722 - 2704 = 18$, e demais somas de quadrados necessárias para compor os quadros de análise de variância na ordem A, B e B, A, cujos valores são iguais aos das tabelas 4.6 e 4.7.

4.3. O Modelo - Σ

4.3.1. Com interação na ordem A - B - A*B

Seja modelo (2.14), $y = Z\hat{\theta} + e$, caracterizado por:

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

Expresso na forma matricial de (2.14) fica:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{131} \\ y_{141} \\ y_{142} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{241} \\ y_{242} \\ y_{243} \\ y_{321} \\ y_{331} \\ y_{332} \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 11 \\ 14 \\ 9 \\ 13 \\ 8 \\ 14 \\ 14 \\ 16 \\ 12 \\ 14 \\ 13 \\ 11 \\ 16 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 0 & 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \hat{\theta} + e$$

4.3.1.1. Equações normais e suas soluções

$$\begin{bmatrix} 16 & 3 & 4 & 0 & -2 & -2 & 2 & -1 \\ 3 & 9 & 3 & 1 & -3 & -3 & 1 & 0 \\ 4 & 3 & 10 & -1 & -2 & -5 & 1 & 2 \\ 0 & 1 & -1 & 10 & 5 & 5 & 0 & -1 \\ -2 & -3 & -2 & 5 & 8 & 5 & -1 & -2 \\ -2 & -3 & -5 & 5 & 5 & 8 & -1 & -2 \\ 2 & 1 & 1 & 0 & -1 & -1 & 10 & 5 \\ -1 & 0 & 2 & -1 & -2 & -2 & 5 & 11 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\gamma}_{11} \\ \hat{\gamma}_{19} \end{bmatrix} = \begin{bmatrix} 208 \\ 27 \\ 46 \\ -3 \\ -20 \\ -13 \\ 31 \\ -22 \end{bmatrix}$$

e uma solução das equações normais é

$$\begin{bmatrix} 0.08 & -0.01 & -0.03 & -0.01 & 0.02 & -0.01 & -0.02 & 0.03 \\ -0.01 & 0.18 & -0.01 & -0.09 & 0.08 & 0.07 & -0.01 & 0.03 \\ -0.03 & -0.01 & 0.18 & -0.03 & -0.04 & 0.14 & 0.01 & -0.03 \\ -0.01 & -0.09 & -0.03 & 0.22 & -0.10 & -0.14 & -0.00 & -0.02 \\ 0.02 & 0.08 & -0.04 & -0.10 & 0.28 & -0.09 & -0.01 & 0.04 \\ -0.01 & 0.07 & 0.14 & -0.14 & -0.09 & 0.38 & 0.01 & 0.01 \\ -0.02 & -0.01 & 0.01 & -0.00 & -0.01 & 0.01 & 0.14 & -0.07 \\ 0.03 & 0.03 & -0.03 & -0.02 & 0.04 & 0.01 & -0.07 & 0.14 \end{bmatrix} \begin{bmatrix} 208 \\ 27 \\ 46 \\ -3 \\ -20 \\ -13 \\ 31 \\ -22 \end{bmatrix} = \hat{\theta}$$

Portanto,

$$\hat{\hat{\theta}} = \begin{bmatrix} 12.958 \\ -0.733 \\ 1.067 \\ -1.625 \\ -0.325 \\ 3.075 \\ 1.400 \\ -1.300 \end{bmatrix}$$

4.3.1.2. Somas de quadrados

De modo análogo a (4.7 à 4.16), obtem-se:

$$R(\hat{\mu}) = \hat{\hat{\theta}}_1' Z_1' y = 2704$$

$$R(\hat{\mu}, \hat{\alpha}) = \hat{\hat{\theta}}_2' Z_2' y = 2722$$

$$R(\hat{\alpha} \mid \hat{\mu}) = R(\hat{\mu}, \hat{\alpha}) - R(\hat{\mu}) = 2722 - 2704 = 18$$

$$R(\hat{\mu}, \hat{\beta}) = \hat{\hat{\theta}}_3' Z_3' y = 2745,33$$

$$R(\hat{\beta} \mid \hat{\mu}) = R(\hat{\mu}, \hat{\beta}) - R(\hat{\mu}) = 41,33$$

$$R(\hat{\mu}, \hat{\alpha}, \hat{\beta}) = \hat{\hat{\theta}}_4' Z_4' y = 2750,23$$

$$R(\hat{\alpha} \mid \hat{\mu}, \hat{\beta}) = R(\hat{\mu}, \hat{\alpha}, \hat{\beta}) - R(\hat{\mu}, \hat{\beta}) = 4,9$$

$$R(\hat{\beta} \mid \hat{\mu}, \hat{\alpha}) = R(\hat{\mu}, \hat{\alpha}, \hat{\beta}) - R(\hat{\mu}, \hat{\alpha}) = 27,33$$

$$R(\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \hat{\hat{\theta}}' Z' y = 2768$$

$$R(\hat{\gamma} \mid \hat{\mu}, \hat{\alpha}, \hat{\beta}) = R(\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) - R(\hat{\mu}, \hat{\alpha}, \hat{\beta}) = 17,77$$

As somas de quadrados acima produzem tabelas de análise de variância semelhantes às tabelas 4.3, 4.4, na presença de interação, e 4.5, 4.6, sem interação. Contudo,

para os objetivos deste trabalho será mais importante no modelo- Σ , conforme IEMMA (1991) as somas de quadrados resultantes do ajuste de um parâmetro em relação a todos os outros do modelo, por exemplo:

$$R(\hat{\alpha} \mid \hat{\mu}, \beta, \hat{\gamma}), R(\hat{\beta} \mid \hat{\mu}, \hat{\alpha}, \hat{\gamma}) \text{ e } R(\hat{\gamma} \mid \hat{\mu}, \hat{\alpha}, \hat{\beta}) \quad (4.19)$$

em lugar do ajuste sequencial.

Para tanto, usando o processo da inversa de partes da inversa de SEARLE (1971), descrita também por IEMMA (1991), obter-se-á as somas de quadrados ajustadas. Assim, colocando-se a inversa $(Z'Z)^{-1}$ na forma abaixo:

$$(Z'Z)^{-1} = \begin{bmatrix} A_1 & M_1 & M_2 & M_3 \\ M'_1 & A_2 & M_4 & M_5 \\ M'_2 & M'_4 & A_3 & M_6 \\ M'_3 & M'_5 & M'_6 & A_4 \end{bmatrix}$$

onde, M_k , $k = 1 \dots 6$, são matrizes sem interesse e

$$A_1 = 0,08$$

$$A_2 = \begin{bmatrix} 0.18 & -0.01 \\ -0.01 & 0.18 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0.22 & -0.10 & -0.14 \\ -0.10 & 0.28 & -0.09 \\ -0.14 & -0.09 & 0.38 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 0.14 & -0.07 \\ -0.07 & 0.14 \end{bmatrix}$$

ter-se-á as somas de quadrados relativas à expressão (4.19), as quais explicitadas segundo os dados do exemplo em questão ficam:

$$\begin{aligned} R(\hat{\alpha} \mid \hat{\mu}, \beta, \hat{\gamma}) &= \hat{\hat{\epsilon}}_2' \mathbf{A}_2^{-1} \hat{\hat{\epsilon}}_2 = 8.71 \\ R(\hat{\beta} \mid \hat{\mu}, \hat{\alpha}, \hat{\gamma}) &= \hat{\hat{\epsilon}}_3' \mathbf{A}_3^{-1} \hat{\hat{\epsilon}}_3 = 26.36 \\ R(\hat{\gamma} \mid \hat{\mu}, \hat{\alpha}, \hat{\beta}) &= \hat{\hat{\epsilon}}_4' \mathbf{A}_4^{-1} \hat{\hat{\epsilon}}_4 = 17.77, \end{aligned}$$

onde, os $\hat{\hat{\epsilon}}$'s são aqueles definidos na pg. 79 e as matrizes envolvidas são as inversas das matrizes definidas anteriormente e tomadas da matriz $(\mathbf{Z}'\mathbf{Z})^{-1}$ com três casas decimais.

Essas somas de quadrados assim definidas, produzem a tabela de análise de variância a seguir:

Tabela 4.8 Análise de variância dos dados da Tabela 4.1, através de um modelo com dois fatores com interação na ordem A - B - A*B.

C.V	g.l.	S.Q.	Q.M.	F
$(\hat{\alpha} \mid \hat{\mu}, \beta, \hat{\gamma})$	2	$R(\hat{\alpha} \mid \hat{\mu}, \beta, \hat{\gamma}) = 8,71$	4,36	0,76
$(\hat{\beta} \mid \hat{\mu}, \hat{\alpha}, \hat{\gamma})$	3	$R(\hat{\beta} \mid \hat{\mu}, \hat{\alpha}, \hat{\gamma}) = 26,36$	8,79	1,53
$(\hat{\gamma} \mid \hat{\mu}, \hat{\alpha}, \hat{\beta})$	2	$R(\hat{\gamma} \mid \hat{\mu}, \hat{\alpha}, \hat{\beta}) = 17.76$	8,88	1,54
Resíduo	8	S.Q.Res. = 46,00	5,75	

4.4. Hipóteses

A expressão (3.2) pode ser usada diretamente para hipóteses sobre os termos $\mu + \alpha_i + \beta_j + \gamma_{ij}$. Quando se está trabalhando com o modelo de médias de caselas, usa-se $\mu + \alpha_i + \beta_j + \gamma_{ij}$ para μ_{ij} nas caselas ocupadas na formulação da hipótese $B'e = m$. Assim, qualquer hipótese colocada em termos dos μ_{ij} pode ser redefinido em termos de $\mu + \alpha_i + \beta_j + \gamma_{ij}$.

Segundo SEARLE (1987), um aspecto importante do teste de hipóteses decorrente do conceito de funções estimáveis é que, hipóteses sobre diferenças de efeitos de linhas (colunas) não são testáveis. Isto significa que hipóteses sobre combinações lineares de diferenças tais como $(\alpha_i - \alpha_h)$ e $(\beta_j - \beta_k)$ não são testáveis, ou seja, hipóteses concernentes a contrastes de linhas (colunas) não são testáveis. A razão para isto é que as únicas hipóteses que podem ser testadas são aquelas explicitadas por $B'e = m$, com os elementos de $B'e$ estimáveis.

Modelo - M com interação na ordem A - B - A*B.

$y = W\mu + e$, caracterizado por $y_{ijk} = \mu_{ij} + e_{ijk}$.

a) Hipóteses do Tipo I:

$H_0^{(1)}$ - Médias ponderadas de linhas, com

$$S.Q.H_0^{(1)} = R(\alpha | \mu) = R(\mu, \alpha) - R(\mu)$$

$H_o^{(5)}$ - Médias ponderadas de colunas, com

$$S.Q.H_o^{(5)} = R(\beta \mid \mu) = R(\mu, \beta) - R(\mu)$$

Dados da Tabela 4.1 em termos das médias.

	j = 1	j = 2	j = 3	j = 4
i = 1	$\mu_{11}^{(3)}$	-	$\mu_{13}^{(1)}$	$\mu_{14}^{(2)}$
i = 2	$\mu_{21}^{(2)}$	$\mu_{22}^{(2)}$	-	$\mu_{24}^{(3)}$
i = 3	-	$\mu_{32}^{(1)}$	$\mu_{33}^{(2)}$	

i = 3 $\Rightarrow$ 2 G.L (graus de liberdade) para linhas

j = 4 $\Rightarrow$ 3 G.L para colunas

$$H_o^{(1)}: \begin{cases} \frac{3\mu_{11} + \mu_{13} + 2\mu_{14}}{6} = \frac{2\mu_{21} + 2\mu_{22} + 3\mu_{24}}{7} \\ \frac{3\mu_{11} + \mu_{13} + 2\mu_{14}}{6} = \frac{\mu_{32} + 2\mu_{33}}{3} \end{cases}$$

$$B'\mu_o = \begin{bmatrix} 3/6 & 1/6 & 2/6 & -2/7 & -2/7 & -3/7 & 0 & 0 \\ 3/6 & 1/6 & 2/6 & 0 & 0 & 0 & -1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 3.9376 & -1.3125 \\ -1.3125 & 2.4375 \end{bmatrix}$$

$$S.Q.H_o^{(1)} = (B'\mu^o)'[B'(W'W)^{-1}B]^{-1}(B'\mu^o) = 18 \text{ (Apêndices: 1, 3, 4, 6, 7 e 8).}$$

$$H_o^{(5)} = \begin{cases} \frac{3\mu_{11} + 2\mu_{21}}{5} = \frac{2\mu_{22} + \mu_{32}}{3} \\ \frac{3\mu_{11} + 2\mu_{21}}{5} = \frac{\mu_{13} + 2\mu_{33}}{3} \\ \frac{3\mu_{11} + 2\mu_{21}}{5} = \frac{2\mu_{14} + 3\mu_{24}}{5} \end{cases}$$

$$B'\mu^o = \begin{bmatrix} 3/5 & 0 & 0 & 2/5 & -2/3 & 0 & -1/3 & 0 \\ 3/5 & -1/3 & 0 & 2/5 & 0 & 0 & 0 & -2/3 \\ 3/5 & 0 & -2/5 & 2/5 & 0 & -3/5 & 0 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$= \begin{bmatrix} -2.0668 \\ -4.4001 \\ -0.6000 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 2.4375 & -1.5625 & -0.9375 \\ -0.5625 & 2.4375 & -0.9375 \\ -0.9375 & -0.9375 & 3.4375 \end{bmatrix}$$

$$S.Q.H_o^{(5)} = (B'\mu^o)'[B'(W'W)^{-1}B]^{-1}(B'\mu^o) = 41.3357 \text{ (Apêndices: 1, 3, 4, 6, 7 e 8).}$$

1

¹Observa-se que as hipóteses $H_o^{(1)}$ e $H_o^{(5)}$ para testarem os efeitos de linhas e colunas respectivamente, apesar de não apresentar explicitamente em suas expressões, elas envolvem efeitos de fatores secundários, como se pode ver através do Modelo-S.

Modelo - S com interação:

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

$$H_o^{(1)} \left\{ \begin{array}{l} \alpha_1 - \alpha_2 + \frac{9\beta_1 - 12\beta_2 + 7\beta_3 - 4\beta_4}{42} + \\ \frac{21\gamma_{11} + 7\gamma_{13} + 14\gamma_{14} - 12\gamma_{21} - 12\gamma_{22} - 18\gamma_{24}}{42} \\ = 0. \\ \alpha_1 - \alpha_3 + \frac{3\beta_1 - 2\beta_2 - 3\beta_3 + 2\beta_4}{6} + \\ \frac{3\gamma_{11} + \gamma_{13} + 2\gamma_{14} - 2\gamma_{32} - 4\gamma_{33}}{6} = 0 \end{array} \right.$$

$$H_o^{(5)} \left\{ \begin{array}{l} \beta_1 - \beta_2 + \frac{9\alpha_1 - 4\alpha_2 - 5\alpha_3}{15} + \\ \frac{9\gamma_{11} + 6\gamma_{21} - 10\gamma_{22} - 5\gamma_{32}}{15} = 0 \\ \beta_1 - \beta_3 + \frac{4\alpha_1 + 6\alpha_2 - 10\alpha_3}{15} + \\ \frac{9\gamma_{11} - 5\gamma_{13} + 6\gamma_{21} - 10\gamma_{33}}{15} = 0 \\ \beta_1 - \beta_4 + \frac{\alpha_1 - \alpha_2 + 3\gamma_{11} - 2\gamma_{14} + 2\gamma_{21} - 3\gamma_{24}}{5} = 0 \end{array} \right.$$

Através das expressões acima, observa-se que não se consegue testar os efeitos dos níveis dos fatores principais isoladamente, vez que, a hipótese absorve também os efeitos dos fatores secundários e da interação. Caso a

interação não seja significativa, alguns especialistas aconselham adotar o modelo sem interação. Assim fazendo, ter-se-á a hipótese do tipo I.

Modelo - M sem interação: $y_{ijk} = \mu_i + \beta_j + e_{ijk}$,

Dados da Tabela 4.1 em termos das médias.

	j = 1	j = 2	j = 3	j = 4
i = 1	$\mu_1 + \beta_1^{(3)}$	-	$\mu_1 + \beta_3^{(1)}$	$\mu_1 + \beta_4^{(2)}$
i = 2	$\mu_2 + \beta_1^{(2)}$	$\mu_2 + \beta_2^{(2)}$	-	$\mu_2 + \beta_4^{(3)}$
i = 3	-	$\mu_3 + \beta_2^{(1)}$	$\mu_3 + \beta_3^{(2)}$	-

i = 3 $\Rightarrow$ 2 G.L (graus de liberdade) para linhas

j = 4 $\Rightarrow$ 3 G.L para colunas

$$H_o^{(10)} \begin{cases} \mu_1 - \mu_2 + \frac{1}{42} [9\beta_1 - 12\beta_2 + 7\beta_3 - 4\beta_4] = 0 \\ \mu_1 - \mu_3 + \frac{1}{6} [3\beta_1 - 2\beta_2 - 3\beta_3 + 2\beta_4] = 0 \end{cases}$$

$$H_{\bullet}^{(14)} \begin{cases} \beta_1 - \beta_2 + \frac{9\mu_1 - 4\mu_2 - 5\mu_3}{15} = 0 \\ \beta_1 - \beta_3 + \frac{4\mu_1 + 6\mu_2 - 10\mu_3}{15} = 0 \\ \beta_1 - \beta_4 + \frac{\mu_1 - \mu_2}{5} = 0 \end{cases}$$

Modelo - S sem interação: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$,

$$H_o^{(10)} \left[\begin{array}{l} \alpha_1 - \alpha_2 + \frac{9\beta_1 - 12\beta_2 + 7\beta_3 - 4\beta_4}{42} = 0 \\ \alpha_1 - \alpha_3 + \frac{3\beta_1 - 2\beta_2 - 3\beta_3 + 2\beta_4}{6} = 0 \end{array} \right.$$

$$H_o^{(14)} \left[\begin{array}{l} \beta_1 - \beta_2 + \frac{9\alpha_1 - 4\alpha_2 - 5\alpha_3}{15} = \\ \beta_1 - \beta_3 + \frac{4\alpha_1 + 6\alpha_2 - 10\alpha_3}{15} \\ \beta_1 - \beta_4 + \frac{\alpha_1 - \alpha_2}{5} = 0 \end{array} \right.$$

Percebe-se mais uma vez que os efeitos dos níveis dos fatores principais estão contaminados com a presença de parâmetros "indesejáveis", o que dificulta o teste sobre diferenças entre os níveis desses fatores.

Com isto, conclui-se para este exemplo e casos similares, que a hipótese do tipo I testa apenas a interação, e, caso esta não seja significativa, aconselha-se repensar o problema e tentar outras alternativas, e isto se constitui uma dificuldade a mais para os não iniciados em Estatística.

b) Hipóteses do Tipo II:

$H_o^{(2)}$ - Médias ponderadas de linhas, ajustadas para colunas; conseqüentemente o fator A ajustado.

$$S.Q.H_o^{(2)} = R(\alpha \mid \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta)$$

$H_o^{(6)}$ - Médias ponderadas de colunas, ajustadas para linhas; conseqüentemente, o fator B ajustado.

$$S.Q.H_o^{(6)} = R(\beta \mid \mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha)$$

Modelo-M com interação: $y_{ijk} = \mu_{ij} + e_{ijk}$

$$H_o^{(2)} : \begin{cases} \frac{3\mu_{11} + \mu_{13} + 2\mu_{14}}{6} = \frac{3\bar{c}_1 + \bar{c}_3 + 2\bar{c}_4}{6} \\ \frac{2\mu_{21} + 2\mu_{22} + 3\mu_{24}}{7} = \frac{2\bar{c}_1 + 2\bar{c}_2 + 3\bar{c}_4}{7} \end{cases}$$

Onde,

$$\bar{c}_1 = \frac{3\mu_{11} + 2\mu_{21}}{5}$$

$$\bar{c}_2 = \frac{2\mu_{22} + \mu_{32}}{3}$$

$$\bar{c}_3 = \frac{\mu_{13} + 2\mu_{33}}{3}$$

$$\bar{c}_4 = \frac{2\mu_{14} + 3\mu_{24}}{5}$$

então, tem-se as seguintes expressões para $H_{\circ}^{(2)}$

$$H_{\circ}^{(2)} : \begin{cases} \frac{18\mu_{11} + 10\mu_{13} + 18\mu_{14}}{15} = \frac{18\mu_{21} + 18\mu_{24} + 10\mu_{33}}{15} \\ \frac{18\mu_{21} + 10\mu_{22} + 18\mu_{24}}{15} = \frac{18\mu_{11} + 18\mu_{14} + 10\mu_{32}}{15} \end{cases}$$

$$B'\mu^{\circ} = \begin{bmatrix} 6/5 & 2/3 & 6/5 & -6/5 & 0 & -6/5 & 0 & -2/3 \\ -6/5 & 0 & -6/5 & 6/5 & 2/3 & 6/5 & -2/3 & 0 \end{bmatrix}$$

12
14
11
11
15
13
11
17

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 0.8415 & 0.6585 \\ 0.6585 & 0.8415 \end{bmatrix}$$

$$S.Q.H_{\circ}^{(2)} = (B'\mu^{\circ})'[B'(W'W)^{-1}B]^{-1}(B'\mu^{\circ}) = 4.90088 \text{ (Apêndices: 1, 3, 4, 6, 7, 8, 9).}$$

Para $H_{\circ}^{(6)}$ tem-se:

$$H_o^{(6)} \left\{ \begin{array}{l} \frac{3\mu_{11} + 2\mu_{21}}{5} = \frac{3\bar{L}_1 + 2\bar{L}_2}{5} \\ \frac{2\mu_{22} + \mu_{32}}{3} = \frac{2\bar{L}_2 + \bar{L}_3}{3} \\ \frac{\mu_{13} + 2\mu_{33}}{3} = \frac{\bar{L}_1 + 2\bar{L}_3}{3} \end{array} \right.$$

onde,

$$\bar{L}_1 = \frac{3\mu_{11} + \mu_{13} + 2\mu_{14}}{6}$$

$$\bar{L}_2 = \frac{2\mu_{21} + 2\mu_{22} + 3\mu_{24}}{7}$$

$$\bar{L}_3 = \frac{\mu_{32} + 2\mu_{33}}{3}$$

Substituindo-se estes valores nas expressões de $H_o^{(6)}$, acima, tem-se:

$$H_o^{(6)} \left\{ \begin{array}{l} \frac{63\mu_{11} + 60\mu_{21}}{42} = \frac{21\mu_{13} + 42\mu_{14} + 24\mu_{22} + 36\mu_{24}}{42} \\ \frac{30\mu_{22} + 14\mu_{32}}{21} = \frac{12\mu_{21} + 18\mu_{24} + 14\mu_{33}}{21} \\ \frac{5\mu_{13} + 4\mu_{33}}{6} = \frac{3\mu_{11} + 2\mu_{14} + 4\mu_{32}}{6} \end{array} \right.$$

$$B'\mu^{\circ} = \begin{bmatrix} 1.50 & -0.50 & -1.00 & 1.43 & -0.57 & -0.86 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & -0.57 & 1.43 & -0.86 & 0.67 & -0.67 \\ -0.50 & 0.83 & -0.332 & 0.00 & 0.00 & 0.00 & -0.67 & 0.67 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$= \begin{bmatrix} -4.00 \\ -0.20 \\ 6.01 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 0.4139 & 0.1821 & 0.2197 \\ 0.1821 & 0.6349 & 0.3457 \\ 0.2197 & 0.3457 & 0.8951 \end{bmatrix}$$

$$S.Q.H_{\circ}^{(\phi)} = (B'\mu^{\circ})'[B'(W'W)^{-1}B]^{-1}(B'\mu^{\circ}) = 28.335 \quad (\text{Apêndices: 1, 3, 4, 6, 7, 8 e 9}).$$

Modelo - S com interação:

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

Para a hipótese $H_{\circ}^{(2)}$, tem-se a seguinte expressão:

$$H_{\circ}^{(2)} \left\{ \begin{array}{l} \frac{1}{6} \left[3(\mu + \alpha_1 + \beta_1 + \gamma_{11}) + (\mu + \alpha_1 + \beta_3 + \gamma_{13}) + \right. \\ \quad \left. 2(\mu + \alpha_1 + \beta_4 + \gamma_{14}) \right] = \frac{1}{6} \left[3\bar{c}_1 + \bar{c}_3 + 2\bar{c}_4 \right] \\ \\ \frac{1}{7} \left[2(\mu + \alpha_2 + \beta_1 + \gamma_{21}) + 2(\mu + \alpha_2 + \beta_2 + \gamma_{22}) + \right. \\ \quad \left. 3(\mu + \alpha_2 + \beta_4 + \gamma_{24}) \right] = \frac{1}{7} \left[2\bar{c}_1 + 2\bar{c}_2 + 3\bar{c}_4 \right] \end{array} \right.$$

onde,

$$\bar{c}_1 = \frac{1}{5} \left[3(\mu + \alpha_1 + \beta_1 + \gamma_{11}) + 2(\mu + \alpha_2 + \beta_1 + \gamma_{21}) \right]$$

$$\bar{c}_2 = \frac{1}{3} \left[2(\mu + \alpha_2 + \beta_2 + \gamma_{22}) + (\mu + \alpha_3 + \beta_2 + \gamma_{32}) \right]$$

$$\bar{c}_3 = \frac{1}{3} \left[(\mu + \alpha_1 + \beta_3 + \gamma_{13}) + 2(\mu + \alpha_3 + \beta_3 + \gamma_{33}) \right]$$

$$\bar{c}_4 = \frac{1}{5} \left[2(\mu + \alpha_1 + \beta_4 + \gamma_{14}) + 3(\mu + \alpha_2 + \beta_4 + \gamma_{24}) \right]$$

substituindo-se vem

$$H_o^{(2)} \cdot \begin{cases} \frac{1}{15} \left[46\alpha_1 - 36\alpha_2 - 10\alpha_3 + 18\gamma_{11} + 10\gamma_{13} + 18\gamma_{14} - 18\gamma_{21} \right. \\ \quad \left. - 18\gamma_{24} - 10\gamma_{33} \right] = 0 \\ \frac{1}{15} \left[-36\alpha_1 + 46\alpha_2 - 10\alpha_3 - 18\gamma_{11} - 18\gamma_{14} + 18\gamma_{21} + 18\gamma_{22} \right. \\ \quad \left. + 18\gamma_{24} - 10\gamma_{32} \right] = 0 \end{cases}$$

Para a Hipótese $H_o^{(6)}$ tem-se:

$$\begin{cases} \frac{3(\mu + \alpha_1 + \beta_1 + \gamma_{11}) + 2(\mu + \alpha_2 + \beta_1 + \gamma_{21})}{5} = \frac{3\bar{L}_1 + 2\bar{L}_2}{5} \\ \frac{2(\mu + \alpha_2 + \beta_2 + \gamma_{22}) + (\mu + \alpha_3 + \beta_2 + \gamma_{32})}{3} = \frac{2\bar{L}_2 + 2\bar{L}_3}{3} \\ \frac{(\mu + \alpha_1 + \beta_3 + \gamma_{13}) + 2(\mu + \alpha_3 + \beta_3 + \gamma_{33})}{3} = \frac{\bar{L}_1 + 2\bar{L}_3}{3} \end{cases}$$

onde,

$$\begin{aligned}\bar{L}_1 &= \frac{3(\mu + \alpha_1 + \beta_1 + \gamma_{11}) + (\mu + \alpha_1 + \beta_3 + \gamma_{13}) + 2(\mu + \alpha_1 + \beta_4 + \gamma_{14})}{6} \\ \bar{L}_2 &= \frac{2(\mu + \alpha_2 + \beta_1 + \gamma_{21}) + 2(\mu + \alpha_2 + \beta_2 + \gamma_{22}) + 3(\mu + \alpha_2 + \beta_4 + \gamma_{24})}{7} \\ \bar{L}_3 &= \frac{(\mu + \alpha_3 + \beta_2 + \gamma_{32}) + 2(\mu + \alpha_3 + \beta_3 + \gamma_{33})}{3}\end{aligned}$$

Tem-se então,

$$H^{(6)} \left\{ \begin{aligned} &\frac{1}{14} \left[41\beta_1 - 8\beta_2 - 7\beta_3 - 26\beta_4 + 21\gamma_{11} - 7\gamma_{13} - 14\gamma_{14} \right. \\ &\quad \left. + 20\gamma_{21} - 8\gamma_{22} - 12\gamma_{24} \right] = 0 \\ &\frac{1}{21} \left[-12\beta_1 + 44\beta_2 - 14\beta_3 - 18\beta_4 - 12\gamma_{21} + 30\gamma_{22} - 18\gamma_{24} \right. \\ &\quad \left. + 14\gamma_{32} - 14\gamma_{33} \right] = 0 \\ &\frac{1}{6} \left[-3\beta_1 - 4\beta_2 + 9\beta_3 - 2\beta_4 - 3\gamma_{11} + 5\gamma_{13} - 2\gamma_{14} \right. \\ &\quad \left. - 4\gamma_{32} + 4\gamma_{33} \right] = 0 \end{aligned} \right.$$

Observa-se através das expressões para as hipóteses $H^{(2)}_0$ e $H^{(6)}_0$, que constituem a hipótese do tipo II, que não é possível isolar os efeitos principais, porquanto, em ambos os casos $H^{(2)}_0$ e $H^{(6)}_0$, esses efeitos são acompanhados de uma função da interação, $f(\gamma)$. Consequentemente, pode-se testar somente a interação. Caso esta, seja não significativa, estrutura-se as hipóteses do tipo II para o modelo sem interação, conforme segue:

Modelo-M , sem interação: $y_{ijk} = \mu_i + \beta_j + e_{ijk}$

$$H_o^{(11)} \left\{ \begin{array}{l} \frac{3(\mu_1 + \beta_1) + (\mu_1 + \beta_3) + 2(\mu_1 + \beta_4)}{6} = \frac{3\bar{c}_1 + \bar{c}_3 + 2\bar{c}_4}{6} \\ \frac{2(\mu_2 + \beta_1) + 2(\mu_2 + \beta_2) + 3(\mu_2 + \beta_4)}{7} = \frac{2\bar{c}_1 + 2\bar{c}_2 + 3\bar{c}_4}{7} \end{array} \right.$$

onde,

$$\bar{c}_1 = \frac{1}{3} \left[3(\mu_1 + \beta_1) + 2(\mu_2 + \beta_1) \right]$$

$$\bar{c}_2 = \frac{1}{3} \left[2(\mu_2 + \beta_2) + (\mu_3 + \beta_2) \right]$$

$$\bar{c}_3 = \frac{1}{3} \left[(\mu_1 + \beta_3) + 2(\mu_3 + \beta_3) \right]$$

$$\bar{c}_4 = \frac{1}{5} \left[2(\mu_1 + \beta_4) + 3(\mu_2 + \beta_4) \right]$$

substituindo-se esses valores na expressão de $H_o^{(11)}$ acima,

tem-se

$$H_o^{(11)} \left\{ \begin{array}{l} 46\mu_1 - 36\mu_2 - 10\mu_3 = 0 \\ -36\mu_1 + 46\mu_2 - 10\mu_3 = 0 \end{array} \right.$$

Para a Hipótese $H_o^{(15)}$ vem

$$H_{\circ}^{(15)} \left[\begin{array}{lcl} \frac{1}{5} \left(3(\mu_1 + \beta_1) + 2(\mu_2 + \beta_1) \right) & = & \frac{1}{5} \left(3\bar{L}_1 + 2\bar{L}_2 \right) \\ \frac{1}{3} \left(2(\mu_2 + \beta_2) + (\mu_3 + \beta_2) \right) & = & \frac{1}{3} \left(2\bar{L}_2 + \bar{L}_3 \right) \\ \frac{1}{3} \left((\mu_1 + \beta_3) + 2(\mu_3 + \beta_3) \right) & = & \frac{1}{3} \left(\bar{L}_1 + 2\bar{L}_3 \right) \end{array} \right]$$

onde,

$$\bar{L}_1 = \frac{3(\mu_1 + \beta_1) + (\mu_1 + \beta_3) + 2(\mu_1 + \beta_4)}{6}$$

$$\bar{L}_2 = \frac{2(\mu_2 + \beta_1) + 2(\mu_2 + \beta_2) + 3(\mu_2 + \beta_4)}{7}$$

$$\bar{L}_3 = \frac{(\mu_3 + \beta_2) + 2(\mu_3 + \beta_3)}{3}$$

substituindo-se vem

$$H_{\circ}^{(15)} \left\{ \begin{array}{lcl} 41\beta_1 - 8\beta_2 - 7\beta_3 - 26\beta_4 & = & 0 \\ -12\beta_1 + 44\beta_2 - 14\beta_3 - 18\beta_4 & = & 0 \\ -3\beta_1 - 4\beta_2 + 9\beta_3 - 2\beta_4 & = & 0 \end{array} \right.$$

Como se pode observar, essas expressões permitem testar isoladamente os efeitos principais dos fatores.

Modelo-S , sem interação

$$H_o^{(11)} \left\{ \begin{array}{l} \frac{1}{6} \left[3(\mu + \alpha_1 + \beta_1) + (\mu + \alpha_1 + \beta_3) + 2(\mu + \alpha_1 + \beta_1) \right] \\ \qquad \qquad \qquad = \frac{1}{6} \left[3\bar{c}_1 + \bar{c}_3 + 2\bar{c}_4 \right] \\ \\ \frac{1}{7} \left[2(\mu + \alpha_2 + \beta_1) + 2(\mu + \alpha_2 + \beta_2) + 3(\mu + \alpha_2 + \beta_4) \right] \\ \qquad \qquad \qquad = \frac{1}{7} \left[2\bar{c}_1 + 2\bar{c}_2 + 3\bar{c}_4 \right] \end{array} \right.$$

onde,

$$\bar{c}_1 = \frac{1}{5} \left[3(\mu + \alpha_1 + \beta_1) + 2(\mu + \alpha_2 + \beta_1) \right]$$

$$\bar{c}_2 = \frac{1}{3} \left[2(\mu + \alpha_2 + \beta_2) + (\mu + \alpha_3 + \beta_2) \right]$$

$$\bar{c}_3 = \frac{1}{3} \left[(\mu + \alpha_1 + \beta_3) + 2(\mu + \alpha_3 + \beta_3) \right]$$

$$\bar{c}_4 = \frac{1}{5} \left[2(\mu + \alpha_1 + \beta_4) + 3(\mu + \alpha_2 + \beta_4) \right]$$

substituindo-se vem

$$H_o^{(11)} \left\{ \begin{array}{l} \frac{46\alpha_1 - 36\alpha_2 - 10\alpha_3}{15} = 0 \\ \frac{-36\alpha_1 + 46\alpha_2 - 10\alpha_3}{15} = 0 \end{array} \right.$$

Para a hipótese $H_0^{(15)}$, tem-se

$$H_0^{(15)} \left\{ \begin{array}{l} \frac{3(\mu + \alpha_1 + \beta_1) + 2(\mu + \alpha_2 + \beta_1)}{5} = \frac{3\bar{L}_1 + 2\bar{L}_2}{5} \\ \frac{2(\mu + \alpha_2 + \beta_2) + (\mu + \alpha_3 + \beta_2)}{3} = \frac{2\bar{L}_2 + \bar{L}_3}{3} \\ \frac{(\mu + \alpha_1 + \beta_3) + 2(\mu + \alpha_3 + \beta_3)}{3} = \frac{\bar{L}_1 + 2\bar{L}_3}{3} \end{array} \right.$$

onde,

$$\bar{L}_1 = \frac{3(\mu + \alpha_1 + \beta_1) + (\mu + \alpha_1 + \beta_3) + 2(\mu + \alpha_1 + \beta_4)}{6}$$

$$\bar{L}_2 = \frac{2(\mu + \alpha_2 + \beta_1) + 2(\mu + \alpha_2 + \beta_2) + 3(\mu + \alpha_2 + \beta_4)}{7}$$

$$\bar{L}_3 = \frac{(\mu + \alpha_3 + \beta_2) + 2(\mu + \alpha_3 + \beta_3)}{3}$$

substituindo-se vem

$$H_0^{(15)} \left\{ \begin{array}{l} \frac{(41\beta_1 - 8\beta_2 - 7\beta_3 - 26\beta_4)}{14} = 0 \\ \frac{(-12\beta_1 + 44\beta_2 - 14\beta_3 - 18\beta_4)}{21} = 0 \\ \frac{(-3\beta_1 - 4\beta_2 + 9\beta_3 - 2\beta_4)}{6} = 0 \end{array} \right.$$

Pelo exposto, observa-se que no modelo sem interação, para a hipótese do tipo II, obter-se-á diferenças entre os níveis dos fatores principais sem a presença de parâmetros indesejáveis, utilizando-se funções estimáveis.

c) Hipóteses do Tipo III:

$H_0^{(3)}$ - Médias não ponderadas de linhas.

$$S.Q.H_0^{(3)} = R(\alpha \mid \mu, \beta, \gamma)$$

$H_0^{(7)}$ - Médias não ponderadas de colunas.

$$S.Q.H_0^{(7)} = R(\beta \mid \mu, \alpha, \gamma)$$

Ressalta-se que essas hipóteses são médias não ponderadas de linhas e colunas, e suas respectivas somas de quadrados associadas são aquelas constantes de (4.19). As expressões analíticas dessas hipóteses podem ser obtidas:

1) sobre médias não ponderadas de linhas (colunas) nas colunas (linhas) completas;

2) utilizando o seguinte algoritmo (SEARLE 1987 p.367):

"Seja s o número de caselas ocupadas pelos dados. Seja B o vetor de parâmetros no modelo-S e seja B_S (de ordem s) o vetor de parâmetros no modelo- Σ , de posto completo. Se X e X_S são relacionadamente definidas por $E(y) = XB$ e $E(y) = X_S B_S$, então define-se T e T_S como matrizes de s linhas distintamente diferentes de X e X_S , respectivamente. Então B e B_S são relacionados igualando-se as médias das caselas: $T_S B_S = TB$. Quando existem todas as possíveis

interações, T_S^{-1} existe, e assim

$$B_S = T_S^{-1} TB. \quad (4.20)$$

3) utilizando-se as funções estimáveis do tipo III do SAS, obtidas através da opção E3. Estas funções são implicitamente baseadas no uso de restrições usuais como $\sum_i^a \alpha_i = 0$, $\sum_j^b \beta_j = 0$, $\sum_i^a \gamma_{ij} = 0$ e $\sum_j^b \gamma_{ij} = 0$, quando da solução das equações normais.

No Modelo -M, com interação para $H_o^{(3)}$ tem-se:

$$H_o^{(3)} : \begin{cases} \frac{\mu_{11} + 3\mu_{13} + \mu_{14} + 2\mu_{22}}{5} = \frac{\mu_{21} + \mu_{24} + 2\mu_{32} + 3\mu_{33}}{5} \\ \frac{2\mu_{13} + \mu_{21} + 3\mu_{22} + \mu_{24}}{5} = \frac{\mu_{11} + \mu_{14} + 3\mu_{32} + 2\mu_{33}}{5} \end{cases}$$

$$B'\mu^o = \begin{bmatrix} 1/5 & 3/5 & 1/5 & -1/5 & 2/5 & -1/5 & -2/5 & -3/5 \\ -1/5 & 2/5 & -1/5 & 1/5 & 3/5 & 1/5 & -3/5 & -2/5 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$= \begin{bmatrix} -0.4 \\ 1.4 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 2.9195 & -2.2529 \\ -2.2529 & 2.9195 \end{bmatrix}$$

$$S.Q.H_o^{(3)} = (B'\mu^o)'[B'(W'W)^{-1}B]^{-1}(B'\mu^o) = 8.7126$$

(Apêndice: 7)

Analogamente, no Modelo -S, com interação tem-se:

$$H_o^{(3)}: \begin{cases} \alpha_1 - \alpha_3 = \frac{1}{5} \left[-\gamma_{11} - 3\gamma_{13} - \gamma_{14} + \gamma_{21} - \gamma_{22} + \gamma_{24} + 2\gamma_{32} + 3\gamma_{33} \right] \\ \alpha_2 - \alpha_3 = \frac{1}{5} \left[\gamma_{11} - 2\gamma_{13} + \gamma_{14} - \gamma_{21} - \gamma_{22} - \gamma_{24} + 3\gamma_{32} + 2\gamma_{33} \right] \end{cases}$$

$$B' = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & -\frac{1}{5} & -\frac{3}{5} & -\frac{1}{5} & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} & -\frac{1}{5} & -\frac{3}{5} & -\frac{1}{5} & \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

$$B'e^o = \begin{bmatrix} -0.4 \\ 1.4 \end{bmatrix}$$

$$S.Q.H_o^{(3)} = (B'e^o)' [B'(X'X)^{-1}B]^{-1} (B'e^o) = 8.7126$$

(Apêndice: 7)

OBS. Para obtenção da expressão analítica de $H_o^{(3)}$, bem como para a de $H_o^{(7)}$ a seguir, optou-se pelas funções estimáveis do tipo III do SAS.

No modelo - M, para $H_o^{(7)}$ tem-se:

$$H_o^{(7)}: \begin{cases} \frac{\mu_{11} + \mu_{21}}{2} = \frac{\mu_{14} + \mu_{24}}{2} \\ \frac{\mu_{11} + 3\mu_{13} + 7\mu_{22} + 3\mu_{32}}{10} = \frac{4\mu_{14} + \mu_{21} + 6\mu_{24} + 6\mu_{33}}{10} \\ \frac{7\mu_{13} + \mu_{21} + 3\mu_{22} + 3\mu_{33}}{10} = \frac{\mu_{11} + 6\mu_{14} + 4\mu_{24} + 3\mu_{32}}{10} \end{cases}$$

$$B'\mu = \begin{bmatrix} 0.5 & 0 & -0.5 & 0.5 & 0 & -0.5 & 0 & 0 \\ 0.1 & 0.3 & -0.4 & -0.1 & 0.7 & -0.6 & 0.3 & -0.3 \\ -0.1 & 0.7 & -0.6 & 0.1 & 0.3 & -0.4 & -0.3 & 0.3 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 2.9220 & -0.5540 & -0.4950 \\ -0.5540 & 2.0030 & -0.6800 \\ -0.4950 & -0.6800 & 1.4960 \end{bmatrix}$$

$$S.Q.H_o^{(7)} = (B'\mu^o)' [B'(W'W)^{-1}B]^{-1} (B'\mu^o) = 26.3621$$

(Apêndice: 7)

Analogamente, no Modelo -S tem-se:

$$H_0^{(7)}: \begin{cases} \beta_1 - \beta_4 = \frac{1}{2} \left[-\gamma_{11} - \gamma_{21} + \gamma_{14} + \gamma_{24} \right] \\ \beta_2 - \beta_4 = \frac{1}{10} \left[-\gamma_{11} - 3\gamma_{13} + \gamma_{14} + \gamma_{21} - 7\gamma_{22} + 6\gamma_{24} - 3\gamma_{32} + 3\gamma_{33} \right] \\ \beta_3 - \beta_4 = \frac{1}{10} \left[\gamma_{11} - 7\gamma_{13} + 6\gamma_{14} - \gamma_{21} - \gamma_{22} + 4\gamma_{24} + 3\gamma_{32} - 3\gamma_{33} \right] \end{cases}$$

$$B'e^0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & \frac{1}{2}(1 & 0 & -1 & 1 & 0 & -1 & 0 & 0) \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & \frac{1}{10}(1 & 3 & -1 & -1 & 7 & -6 & 3 & -3) \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & \frac{1}{10}(1 & 7 & -6 & 1 & 1 & -4 & -3 & 3) \end{bmatrix} e^0$$

$$B'e^0 = \begin{bmatrix} -0.4 \\ 1.4 \end{bmatrix}$$

$$S.Q.H_0^{(7)} = (B'e^0)' [B'(X'X)^{-1}B]^{-1} (B'e^0) = 26.3621$$

(Apêndice: 7)

d) Hipóteses do Tipo IV:

$H_0^{(4)}$ - Médias de caselas que estão na mesma coluna, começando na última linha.

$$S.Q.H_0^{(4)} = R(\alpha \mid \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta)$$

$H_0^{(8)}$ - Médias de caselas que estão na mesma linha, começando pela última coluna.

$$S.Q.H_0^{(8)} = R(\alpha \mid \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta)$$

No Modelo - M:

Para a hipótese $H_0^{(4)}$ tem-se a seguinte expressão:

$$H_0^{(4)}: \begin{cases} \mu_{13} - \mu_{33} = 0 \\ \mu_{22} - \mu_{32} = 0 \end{cases}$$

$$B'\mu_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$= \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 0.67 & 0 \\ 0 & 0.67 \end{bmatrix}$$

$$S.Q.H_0^{(4)} = (B'\mu_0)'[B'(W'W)^{-1}B]^{-1}(B'\mu_0) = 16.67$$

(Apêndice: 7)

Para a hipótese $H_0^{(8)}$ tem-se:

$$H_0^{(8)}: \begin{cases} (\mu_{11} - \mu_{14}) + (\mu_{21} - \mu_{24}) = 0 \\ \mu_{22} - \mu_{24} = 0 \\ \mu_{13} - \mu_{14} = 0 \end{cases}$$

$$B'\mu_o = \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 0.7320 & -0.2930 & -0.2440 \\ -0.2930 & 1.3170 & 0.0980 \\ -0.2440 & 0.0980 & 0.7480 \end{bmatrix}$$

$$S.Q.H_c^{(8)} = (B'\mu_o)'[B'(W'W)^{-1}B]^{-1}(B'\mu_o) = 16.54$$

(Apêndice: 7)

No Modelo - S:

Para a hipótese $H_o^{(4)}$ tem-se:

$$H_o^{(4)}: \begin{cases} (\alpha_1 - \alpha_3 = \gamma_{33} - \gamma_{13}) \\ (\alpha_2 - \alpha_3 = \gamma_{32} - \gamma_{22}) \end{cases}$$

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

$$S.Q.H_c^{(4)} = (B'e_o)'[B'(X'X)^G B]^{-1}(B'e_o) = 16.67$$

Para a hipótese $H_{\bullet}^{(8)}$ tem-se:

$$H_0^{(8)}: \begin{cases} \beta_1 - \beta_4 = \frac{1}{2}(-\gamma_{11} + \gamma_{14} - \gamma_{21} + \gamma_{24}) \\ \beta_2 - \beta_4 = -\gamma_{22} + \gamma_{24} \\ \beta_3 - \beta_4 = -\gamma_{13} + \gamma_{14} \end{cases}$$

Observando as hipóteses no modelo - M ter-se-á a impressão de ser possível testar diferenças entre linhas e entre colunas. Todavia, comparando as hipóteses equivalentes no modelo - S, fica claro que não é possível comparar as diferenças entre os fatores principais. Existem parâmetros indesejáveis. Contudo, pode-se testar o efeito da interação

$$B' = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0.5 & 0 & -0.5 & 0.5 & 0 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$S.Q.H_0^{(8)} = (B'e^0)' [B'(X'X)^{-1}B]^{-1} (B'e^0) = 16.67$$

(Apêndice: 7)

e) Hipóteses sobre a interação:

$$H_0^{(9)}: \theta_{ij, i'j'} = \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'}$$

$$S.Q.H_0^{(9)} = R(\gamma \mid \mu, \alpha, \beta)$$

Modelo - M:

$$H_0^{(9)} = \begin{cases} \mu_{11} - \mu_{14} - \mu_{21} + \mu_{24} = 0 \\ \mu_{13} - \mu_{14} - \mu_{22} + \mu_{24} + \mu_{32} - \mu_{33} = 0 \end{cases}$$

$$B' \mu^{\circ} = \begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 11 \\ 11 \\ 15 \\ 13 \\ 11 \\ 17 \end{bmatrix}$$

$$[B'(W'W)^{-1}B]^{-1} = \begin{bmatrix} 0.673 & -0.146 \\ -0.146 & 0.293 \end{bmatrix}$$

$$S.Q.H_o^{(D)} = (B'\mu^{\circ})'[B'(W'W)^{-1}B]^{-1}(B'\mu^{\circ}) = 17.77$$

(Apêndice: 7)

Modelo - S:

$$H_o^{(D)} = \begin{cases} \gamma_{11} - \gamma_{14} - \gamma_{21} + \gamma_{24} = 0 \\ \gamma_{13} - \gamma_{14} - \gamma_{22} + \gamma_{24} + \gamma_{32} - \gamma_{33} = 0 \end{cases}$$

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

$$[B'(X'X)^G B]^{-1} = \begin{bmatrix} 0.673 & -0.146 \\ -0.146 & 0.293 \end{bmatrix}$$

$$S.Q.H_o^{(D)} = (B'e^{\circ})'[B'(X'X)^G B]^{-1}(B'e^{\circ}) = 17.77$$

(Apêndice: 7)

5. RESULTADOS E DISCUSSÃO

No desenvolvimento da metodologia e posterior ilustração com exemplo numérico, foi possível constatar a problemática de testar e interpretar hipóteses na análise de variância de dados desbalanceados, que pode não ser de trivial solução, mormente para os usuários de pacotes estatísticos que não são iniciados nas idéias básicas sobre a análise estatística de dados.

Através de tabelas, mostra-se o envolvimento dos pacotes estudados nos quatro tipos de somas de quadrados e hipóteses associadas, conforme abaixo:

a) Se as amostras são equilibradas tem-se:

$$S.Q.TIPO I = S.Q.TIPO II = S.Q.TIPO III = S.Q.TIPO IV.$$

b) Se todas as caselas são ocupadas tem-se:

$$S.Q.TIPO III = S.Q.TIPO IV.$$

c) Se o modelo não contém interação tem-se:

$$S.Q.TIPO II = S.Q.TIPO III = S.Q.TIPO IV.$$

d) Se o modelo contém interação, e se existe caselas vazias, então $S.Q.TIPO III \neq S.Q.TIPO IV$.

- Nas tabelas 5.1 e 5.2 estão os resultados das análises sob o modelo (3) com interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B. As somas de quadrados apresentadas nessas tabelas correspondem às somas de quadrados do tipo I do SAS, e as

hipóteses a elas associadas são: $H_0^{(1)}$, $H_0^{(6)}$, $H_0^{(5)}$, $H_0^{(2)}$, relativas a linhas e colunas e $H_0^{(9)}$ a interseção. Essas somas de quadrados são baseadas nos métodos H3 e OS3 que adotam ordenação à priori.

Outros pacotes além do SAS que apresentam os mesmos resultados são: SOC, MSTAT, MINITAB, GLIM e SANEST, como pode ser visto nos apêndices 8, 4, 3, 1 e 6 respectivamente.

Tabela 5.1 Resumo da Análise de Variância, expondo somas de quadrados tipo I e respectivas hipóteses associadas. Modelo (3) com interação na ordem A - B - A*B.

C.V	GL	S.Q.TIPO I	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu) = 18.00$	9.00	$H_0^{(1)}$	1.57	0.27
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_0^{(6)}$	1.64	0.26
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_0^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS $\equiv$ SOC $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ GLIM $\equiv$ SANEST

Tabela 5.2 Resumo da Análise de Variância, expondo somas de quadrados tipo I e respectivas hipóteses associadas. Modelo (3) com interação na ordem B - A - A*B.

C.V	GL	S.Q.TIPO I	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu) = 41.33$	13.77	$H_{\circ}^{(5)}$	2.40	0.14
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_{\circ}^{(2)}$	0.43	0.67
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_{\circ}^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS $\equiv$ SOC $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ GLIM $\equiv$ SANEST

Nas tabelas 5.3 e 5.4 estão os resultados das análises sob o modelo (3) sem interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$, nas ordens A - B - e B - A. As somas de quadrados apresentadas nessas tabelas correspondem às somas de quadrados do tipo I do SAS, e as hipóteses a elas associadas são: $H_{\circ}^{(10)}$, $H_{\circ}^{(15)}$, $H_{\circ}^{(14)}$, $H_{\circ}^{(11)}$, relativas à linhas e colunas.

Outros pacotes além do SAS que apresentam os mesmos resultados são: SOC, MSTAT, MINITAB, GLIM e SANEST, como pode ser observado nos apêndices 8, 4, 3, 1 e 6. respectivamente.

Tabela 5.3 Resumo da Análise de Variância, expondo somas de quadrados tipo I e respectivas hipóteses associadas. Modelo (3) sem interação na ordem A - B.

C.V	GL	S.Q.TIPO I	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu)$ = 18.00	9.00	$H_o^{(10)}$	1.41	0.29
Varied.	3	$R(\beta \mu, \alpha)$ = 28.23	9.41	$H_o^{(15)}$	1.48	0.28
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ GLIM $\equiv$ SANEST

Tabela 5.4 Resumo da Análise de Variância, expondo somas de quadrados tipo I e respectivas hipóteses associadas. Modelo (3) sem interação na ordem B - A.

C.V	GL	S.Q.TIPO I	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu)$ = 41.33	13.78	$H_o^{(14)}$	2.16	0.16
Fosf.	2	$R(\alpha \mu, \beta)$ = 4.90	2.45	$H_o^{(11)}$	0.38	0.69
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ GLIM $\equiv$ SANEST

- Nas tabelas 5.5 e 5.6 estão os resultados das análises sob o modelo (3) com interação, caracterizado por:

$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B. As somas de quadrados, todas ajustadas, apresentadas nessas tabelas correspondem às somas de quadrados do tipo II do SAS, e as hipóteses a elas associadas são: $H_o^{(11)}$ e $H_o^{(15)}$ relativas a linhas e colunas e $H_o^{(9)}$ a interação.

Outros pacotes além do SAS que apresentam os mesmos resultados são: SPSS, MSTAT e SANEST, apêndices 9, 4 e 6.

Tabela 5.5 Resumo da Análise de Variância, expondo somas de quadrados tipo II e respectivas hipóteses associadas. Modelo (3) com interação na ordem A - B - A*B.

G.V	GL	S.Q.TIPO II	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu \beta) = 4.90$	2.45	$H_o^{(11)}$	0.43	0.67
Varied.	3	$R(\beta \mu \alpha) = 28.23$	9.41	$H_o^{(15)}$	1.64	0.26
Fosf.* Var.	2	$R(\gamma \mu \alpha \beta) = 17.77$	8.89	$H_o^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS $\equiv$ SPSS $\equiv$ MSTAT $\equiv$ SANEST

Tabela 5.6 Resumo da Análise de Variância, expondo somas de quadrados tipo II e respectivas hipóteses associadas. Modelo (3) com interação na ordem B - A - A*B.

C.V	GL	S.Q.TIPO II	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(15)}$	1.64	0.26
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(11)}$	0.43	0.67
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_o^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS $\equiv$ SPSS $\equiv$ MSTAT $\equiv$ SANEST

- Nas tabelas 5.7 e 5.8 estão os resultados das análises sob o modelo (3) sem interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$, nas ordens A - B - e B - A. As somas de quadrados apresentadas nessas tabelas correspondem às somas de quadrados do tipo II do SAS, e as hipóteses a elas associadas são: $H_o^{(11)}$ e, $H_o^{(15)}$, relativas a linhas e colunas.

Essas tabelas correspondem também aos resultados obtidos pelos pacotes: SOC, SAEG, SPSS, SANEST, MSTAT, MINITAB, SYSTAT, STATGRAPHICS e HARVEY, apêndices 8, 5, 9, 6, 4, 3, 11, 10 e 2 respectivamente.

Tabela 5.7 Resumo da Análise de Variância, expondo somas de quadrados tipo II e respectivas hipóteses associadas. Modelo (3) sem interação na ordem A - B.

C.V	GL	S.Q.TIPO II	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(11)}$	0.38	0.69
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(15)}$	1.48	0.28
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

Tabela 5.8 Resumo da Análise de Variância, expondo somas de quadrados tipo II e respectivas hipóteses associadas. Modelo (3) sem interação na ordem B - A.

C.V	GL	S.Q.TIPO II	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(15)}$	1.48	0.28
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(11)}$	0.38	0.69
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

- Nas tabelas 5.9 e 5.10 tem-se os resultados das análises sob o modelo (3) com interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B. As somas de quadrados e hipóteses associadas $H_o^{(3)}$, $H_o^{(7)}$, $H_o^{(9)}$, correspondem às somas de quadrados do tipo III do SAS, e são baseadas nos métodos OS1 e Y2, o que implica na utilização do Modelo- Σ .

Tabela 5.9 Resumo da Análise de variância, expondo somas de quadrados tipo III e respectivas hipóteses associadas. Modelo (3) com interação na ordem A - B - A*B.

G.V	GL	S.Q.TIPO III	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu, \beta, \gamma) = 8.71$	4.36	$H_o^{(3)}$	0.76	0.50
Varied.	3	$R(\beta \mu, \alpha, \gamma) = 26.36$	8.79	$H_o^{(7)}$	1.53	0.28
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_o^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

Tabela 5.10 Resumo da Análise de Variância, expondo somas de quadrados tipo III e respectivas hipóteses associadas. Modelo (3) com interação na ordem B - A - A*B.

G.V	GL	S.Q.TIPO III	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha, \gamma) = 26.36$	8.79	$H_o^{(7)}$	1.53	0.28
Fosf.	2	$R(\alpha \mu, \beta, \gamma) = 8.71$	4.36	$H_o^{(3)}$	0.76	0.50
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_o^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

- Nas tabelas 5.11 e 5.12 estão os resultados das análises sob o modelo (3) sem interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$, nas ordens A - B e B - A. As somas de quadrados apresentadas nessas tabelas correspondem às somas de quadrados do tipo III do SAS, e as hipóteses a elas associadas são: $H_o^{(12)}$ e $H_o^{(16)}$, relativas a linhas e colunas.

Essas tabelas correspondem também aos resultados obtidos pelos pacotes: SOG, SAEG, SPSS, SANEST, MSTAT, MINITAB, SYSTAT, STATGRAPHICS e HARVEY.

Tabela 5.11 Resumo da Análise de Variância, expondo somas de quadrados tipo III e respectivas hipóteses associadas. Modelo (3) sem interação na ordem A - B.

C.V	GL	S.Q.TIPO III	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(12)}$	0.38	0.69
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(16)}$	1.48	0.28
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

Tabela 5.12 Resumo da Análise de Variância, expondo somas de quadrados tipo III e respectivas hipóteses associadas. Modelo (3) sem interação na ordem B - A.

C.V	GL	S.Q.TIPO III	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(16)}$	1.48	0.28
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(12)}$	0.38	0.69
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

- Nas tabelas 5.13 e 5.14 tem-se os resultados das análises sob o modelo (3) com interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B. As somas de quadrados e hipóteses associadas $H_0^{(13)}$, $H_0^{(17)}$, $H_0^{(8)}$, $H_0^{(4)}$ correspondem às somas de quadrados do tipo IV do SAS, e são baseadas nos métodos OS1 e Y2, o que implica na utilização do Modelo- Σ .

Tabela 5.13 Resumo da Análise de Variância, expondo somas de quadrados tipo IV e respectivas hipóteses associadas. Modelo (3) com interação na ordem A - B - A*B.

C.V	GL	S.Q.TIPO IV	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu, \beta, \gamma) = 16.67$	8.33	$H_0^{(13)}$	1.45	0.29
Varied.	3	$R(\beta \mu, \alpha, \gamma) = 16.54$	5.51	$H_0^{(17)}$	0.96	0.46
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_0^{(8)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS

Tabela 5.14 Resumo da Análise de Variância, expondo somas de quadrados tipo IV e respectivas hipóteses associadas. Modelo (3) com interação na ordem B - A - A*B.

G.V	GL	S.Q.TIPO IV	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha, \gamma) = 16.54$	5.51	$H_o^{(8)}$	0.96	0.46
Fosf.	2	$R(\alpha \mu, \beta, \gamma) = 16.67$	8.33	$H_o^{(4)}$	1.45	0.29
Fosf.* Var.	2	$R(\gamma \mu, \alpha, \beta) = 17.77$	8.89	$H_o^{(9)}$	1.54	0.27
Erro	8	S.Q.Res. = 46.00	5.75			

SAS

- Nas tabelas 5.15 e 5.16 estão os resultados das análises sob o modelo (3) sem interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$, nas ordens A - B e B - A. As somas de quadrados apresentadas nessas tabelas correspondem às somas de quadrados do tipo IV do SAS, e as hipóteses a elas associadas são: $H_o^{(13)}$ e $H_o^{(17)}$, relativas a linhas e colunas.

Essas tabelas correspondem também aos resultados obtidos pelos pacotes : SOC, SAEG, SPSS, SANEST, MSTAT, MINITAB, SYSTAT, STATGRAPHICS e HARVEY.

Tabela 5.15 Resumo da Análise de Variância, expondo somas de quadrados tipo IV e respectivas hipóteses associadas. Modelo (3) sem interação na ordem A - B.

G.V	GL	S.Q.TIPO IV	Q.M	HIP.	F	Pr>F
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(13)}$	0.38	0.69
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(17)}$	1.48	0.28
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

Tabela 5.16 Resumo da Análise de Variância, expondo somas de quadrados tipo IV e respectivas hipóteses associadas. Modelo (3) sem interação na ordem B - A.

G.V	GL	S.Q.TIPO IV	Q.M	HIP.	F	Pr>F
Varied.	3	$R(\beta \mu, \alpha) = 28.23$	9.41	$H_o^{(17)}$	1.48	0.28
Fosf.	2	$R(\alpha \mu, \beta) = 4.90$	2.45	$H_o^{(13)}$	0.38	0.69
Erro	10	S.Q.Res. = 63.77	6.38			

SAS $\equiv$ SOC $\equiv$ SAEG $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ SYSTAT
 $\equiv$ STATGRAPHICS $\equiv$ HARVEY

- Nas tabelas 5.17 e 5.18 tem-se uma versão global dos pacotes utilizados, relativamente aos quatro tipos de somas de quadrados do SAS, resultante das análises sob o modelo (3) com interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B.

Tabela 5.17 Comparação entre somas de quadrados calculadas por diferentes "softwares" (pacotes) considerando o modelo - S e modelo - M com dois fatores com interação na ordem A - B.

SOFTWARES (Pacotes)	SOMAS DE QUADRADOS			
	Tipo I	Tipo II	Tipo III	Tipo IV
SAS	*	*	*	*
SOC	*			
SYSTAT				
SAEG				
STATGRAPHICS				
SPSS		*		
MSTAT	*	*		
MINITAB	*			
BMDP				
GLIM	*			
HARVEY				
SANEST	*	*		

Tabela 5.18 Comparação entre somas de quadrados calculadas por diferentes "softwares" (pacotes) considerando o modelo-S e modelo-M com dois fatores com interação na ordem B - A.

SOFTWARES (Pacotes)	SOMAS DE QUADRADOS			
	Tipo I	Tipo II	Tipo III	Tipo IV
SAS	*	*	*	*
SOC	*			
SYSTAT				
SAEG				
STATGRAPHICS				
SPSS		*		
MSTAT	*	*		
MINITAB	*			
BMDP				
GLIM	*			
HARVEY				
SANEST	*	*		

- Nas tabelas 5.19 e 5.20 tem-se uma versão global dos pacotes utilizados, relativamente aos quatro tipos de somas de quadrados do SAS, resultante das análises sob o modelo (3) sem interação, caracterizado por: $y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$, nas ordens A - B - A*B e B - A - A*B.

Tabela 5.19 Comparação entre somas de quadrados calculadas por diferentes "softwares" (pacotes) considerando o modelo-S e modelo-M com dois fatores sem interação na ordem A - B.

SOFTWARES (Pacotes)	SOMAS DE QUADRADOS			
	Tipo I	Tipo II	Tipo III	Tipo IV
SAS	*	*	*	*
SOC	*	*	*	*
SYSTAT		*	*	*
SAEG		*	*	*
STATGRAPHICS		*	*	*
SPSS		*	*	*
MSTAT	*	*	*	*
MINITAB	*	*	*	*
BMDP				
GLIM	*			
HARVEY		*	*	*
SANEST	*	*	*	*

Tabela 5.20 Comparação entre somas de quadrados calculadas por diferentes "softwares" (pacotes) considerando o modelo S(M) com dois fatores sem interação na ordem B - A.

SOFTWARES (Pacotes)	SOMAS DE QUADRADOS			
	Tipo I	Tipo II	Tipo III	Tipo IV
SAS	*	*	*	*
SOC	*	*	*	*
SYSTAT		*	*	*
SAEG		*	*	*
STATGRAPHICS		*	*	*
SPSS		*	*	*
MSTAT	*	*	*	*
MINITAB	*	*	*	*
BMDP				
GLIM	*			
HARVEY		*	*	*
SANEST	*	*	*	*

6. CONCLUSÕES

Face aos resultados obtidos, pode-se concluir que:

6.1. Para modelos com interação

6.1.1 - Dados balanceados

Em geral, para dados balanceados os pacotes fornecem as mesmas somas de quadrados e hipóteses associadas.

6.1.2 - Dados não balanceados

Deve-se considerar duas situações:

i) todas as caselas são ocupadas.

Neste caso, as hipóteses de interesse são:

- Tipo III e IV objetivando testar efeitos puros de cada fator envolvido.
- Hipótese de interação para testar efeitos de interações.

Relativamente a situação anterior, os pacotes utilizados nesse trabalho, produzem os mesmos resultados.

ii) Com algumas caselas vazias.

Neste caso, a hipótese de interesse é a hipótese sobre interação, desde que as demais contém em suas expressões analítica a presença de outros fatores e torna-se difícil sua interpretação por não testar um efeito livre dos outros. A soma de quadrados associada a esta hipótese é do tipo $R(\gamma \mid \mu, \alpha, \beta) = R(\gamma \mid \mu, \alpha, \beta)$.

Neste caso os pacotes estudados apresentam saídas diversificadas, conforme exposição a seguir:

- Para hipóteses do tipo I (tabelas 5.1 e 5.6)

SAS = SOC = MINITAB = GLIM = SANEST

- Para hipóteses do tipo II (tabelas 5.5 e 5.6)

SAS = SPSS = MSTAT = SANEST

- Para hipóteses do tipo III (tabelas 5.9 e 5.10)

Apenas o SAS apresenta saída.

Para hipóteses do tipo IV (tabelas 5.13 e 5.14)

Apenas o SAS apresenta saída.

Ressalte-se, que neste modelo, os pacotes seguintes não apresentam qualquer tipo de saída:

BMDP, HARVEY, SAEG, STATGRAPHICS e SYSTAT.

Como foi observado, no modelo-S ou modelo-M, através das quatro somas de quadrados utilizadas pelo SAS, esse modelo não se presta para testar isoladamente efeitos principais dos fatores, isto é, esse modelo com interação presta-se a testar interações. Uma estratégia prática de análise seria verificar se a interação é significativa; caso seja, deve-se estudar o comportamento de um fator dentro de cada nível fixado do outro. Caso a interação não seja significativa deve-se excluí-la do modelo, e executa-se uma nova análise de variância sem a interação.

6.2. Para modelos sem interação

As hipóteses do tipo II, III e IV são iguais e são as de interesse prático, pois testam o efeito de um fator livre dos efeitos dos outros.

6.2.1 - Dados balanceados

Em geral, para dados balanceados os pacotes fornecem as mesmas somas de quadrados e hipóteses associadas.

6.2.2 - Dados não balanceados.

Deve-se considerar duas situações:

i) todas as caselas são ocupadas.

Neste caso, as hipóteses de interesse são:

Tipo II = III = IV objetivando testar efeitos puros de cada fator envolvido.

Relativamente a situação anterior, os pacotes utilizados nesse trabalho, produzem os mesmos resultados.

ii) Com algumas caselas vazias, os pacotes estudados apresentam os seguintes resultados:

Para hipóteses do tipo I (tabelas 5.3 e 5.4):

SAS $\equiv$ SOC $\equiv$ MSTAT $\equiv$ MINITAB $\equiv$ GLIM $\equiv$ SANEST

Para hipóteses do tipo II (tabelas 5.7 e 5.8):

SAS $\equiv$ SOC $\equiv$ SAE6 $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$
MINITAB $\equiv$ SYSTAT $\equiv$ STATGRAPH $\equiv$ HARVEY

Para hipóteses do tipo III (tabelas 5.11 e 5.12):

SAS $\equiv$ SOC $\equiv$ SAE6 $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$
MINITAB $\equiv$ SYSTAT $\equiv$ STATGRAPH $\equiv$ HARVEY

Para hipóteses do tipo IV (tabelas 5.15 e 5.16):

SAS $\equiv$ SOC $\equiv$ SAE6 $\equiv$ SPSS $\equiv$ SANEST $\equiv$ MSTAT $\equiv$
MINITAB $\equiv$ SYSTAT $\equiv$ STATGRAPH $\equiv$ HARVEY

Um fato importante a ser ressaltado é que o

conhecimento das hipóteses descritas no item (4.4) pode, sem dúvida, propiciar ao pesquisador a correta interpretação das verdadeiras hipóteses testadas quando há desbalanceamento. Pode ainda, permitir a compreensão e a comparação de pesquisas publicadas em revistas internacionais e analisadas através de diferentes pacotes estatísticos e direcionar a escolha do pacote mais indicado na análise de seus dados desbalanceados, mormente no caso de caselas vazias.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALALOUF, I. S. & STYAN, G. P. H. Characterizations of estimability in the general linear model. *The Annals of Statistics*, vol. 7(1) : 19 -200, 1979.
- ANDERSON, R. L. & BANGROFT, T. A. Statistical Theory in Research, New York, McGraw-Hill Book Co., 1952.
- ARVESEN, J. N. & LAYARD, M. W. L. Asymptotically robust tests in unbalanced variance component models. *The Annals of Statistics*, vol. 3(5) : 1122-1134, 1975.
- BMDP, STATISTICAL SOFTWARES, 3ed. Biomedical Computer. Programs. Univ. of California Press, Los Angels - California, 1983.
- BOSE, R. C. & SRIVASTAVA, J. N. Mathematical Theory of factorial Designs, I. Analasys, II Construction, Bull. Int. Stat. Inst.: 40, 780 - 794, 1964.
- BRYCE, G. R.; SCOTT, D. T.; CARTER, M. W. Estimation an Hypothessis Testing in Linear Models - A Reparametrization approach to the cell means model. *Commun. Statist. -Theor. Meth.* A9(2) : 131-150, 1980.
- BURDICK, D. S. & HERR, D. G. Counterexamples in unbalanced two-way analysis of variance. *Commun. Statist. Theor. Meth.*, A9(2): 231-241, 1980.

- CARVALHO FILHO, A. A. & IEMMA, A. F. A aplicação da decomposição pelo valor singular no estudo da estimabilidade de funções lineares em modelos de Gauss-Markov. *Rev. Mat. Estat.*, São Paulo, 8: 37-44, 1990.
- CHAKRAVARTI, I. H.; LAHA, R. G.; ROY, J. Handbook of Methods of Applied Statistics. New York, John Wiley & Sons, 1967.
- COMSTOCK, R.E. "Overestimation of Mean Squares by the Method of Expected Numbers." *Journal of the American Statistical Association*, 38 : 335-340, 1943.
- DAGNELIE, P. Estatística-Teoria e Métodos I., Gembloux, Fac. Agron. Gembloux, 1973, 439p.
- DODGE, Y. Analysis of experiments with missing data. New York, Wiley, 1985, 499p.
- ELSTON, R. C. & BUSH, J. "The hypotheses that can be tested when there are intersections in analysis of variance Model", *Biometrics*, 681-699, 1964.
- FEDERER, W.T., & ZELEN, M. "Analysis of Multifactor Classifications with Unequal Numbers of Observations," *Biometrics*, 22: 525-552, 1966.
- FRANCIS, I. "Comparison of Several Analysis of Variance Programs," *Journal of the American Statistical Association*, 68: 860-865, 1973.
- FRANE, J. W. Some Computing Methods for Unbalanced Analysis of Variance and Covariance. *Commun. Statist. -Theor. Meth.* A9 (2): 151-166, 1980.

FREUND, R.J. The case of missing cell. *The Amer. Statist.*, 34(2): 94-98, 1980.

GLIM, USER'S GUIDE, 2ed. The GLIM System. Royal Statistical Society, London, 1984.

GOODNIGHT, J.H. Test of hypothesis in fixed effect linear models. *Commun. Statist. Theor. Meth.*, A9(2): 167-180, 1980.

GOSSLEE, D.G. & LUCAS, H.L. "Analysis of Variance of Disproportionate Data When Interaction Is Present." *Biometrics*, 21: 115-133, 1965.

GRAYBILL, F. A. An Introduction to linear Statistical Models. New York, McGraw-Hill, 1961. 463p.

HARVEY, ed. User's Guide for LSMLMW. Walter, R. Harvey, Colombus, Ohio. 1987.

HAZEL, L. N. "The Covariance Analysis of Multiple Classification Tables With Unequal Subclass Numbers." *Biometrics*, 2: 21 - 25, 1946.

HENDERSON, C.R. "Estimation of Variance and Covariance Components." *Biometrics*, 9: 226-252, 1953.

HENDRICKS, WALTER A. "Analysis of Variance Considered as an Application of Simple Error Theory. " *Annals of Mathematical Statistics*, 6: 117-126, 1934.

HERR, D.G. On the history of anova in unbalanced, Factorial Designs: The first 30 years. *The Amer. Statist.*, 40(4): 265-270, 1986.

- HOCKING, R.R. The Analysis of Linear Models. **Monterey, Brooks/Cole**, 1985. 385p.
- HOCKING, R. R.; SPEED, F. M.; COLEMAN, A. T Hypotheses tested with unbalanced data. **Commun. Statist. Theor. Meth.**, A9 (2): 116-129, 1980.
- IEMMA, A. F. Testes de hipóteses nos experimentos com parcelas subdivididas em blocos incompletos balanceados. **Dessarrolo Rural en las Am ricas**, 15(2): 143-151, 1983.
- IEMMA, A.F. Uma Introdução à álgebra de matrizes. **Piracicaba. ESALQ/USP**, 1989. 339p.
- IEMMA, A.F. Testes de hipóteses em modelos lineares com amostras desequilibradas. **Belgique. Gembloux**, 1991. 101p.
- IEMMA, A.F. Hypotheses about unbalanced data. **Biometric Bulletin**. 1(10), 1992.
- IEMMA, A.F. Proyectores ortogonales especiales. Diseno de Experimentos. **Universidad Nac de Colombia-Depto Matematicas y Estadistica**. Santafe de Bogota, 1993a.
- IEMMA, A.F. Análisis de Varianza com Datos Desbalanceados. **Universidad Nac de Colombia**, Bogota, 1993b, 102p..
- IEMMA, A.F. & PALM, R. Les Matrices Inverses Généralisées et Leur Utilisation dans Le Modèle Lineaire, **Belgique, Gembloux**, 1991/92. 25p.

- JORN, H. and KLOTZ, J. The Contrast Algorithm for unbalanced multifactor analysis of variance. *Amer. Statist. Assoc.*, **84**(405): 318-324, 1980.
- KRAMER, C. Y. "On the Analysis of Variance of a Two-Way Classification With Unequal Subclass Numbers. " *Biometrics*, **11**: 441-452, 1955.
- LEVY, K. J.; NARULA, S.; ABRAMI, P. "An empirical comparison of the methods of least squares and unweighted means for the analysis of disproportionate cell data". *International Statistical review*, **3**: 335-338, 1973.
- MARSAGLIA, G. & STYAN, G. P. H. Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra* **2**: 269-292, 1974a.
- MARSAGLIA, G. & STYAN, G. P. H. Rank conditions for generalized inverses of partitioned matrices. *Sankhya Ser. A* **36**: 437-442, 1974b.
- MEAD, R.; BANCROFT, T.A.; HAN, C.-PAI Power of Analysis of Variance test procedures for incompletely specified fixed models. *The Annals of Statistics*, **3**(4): 797-808, 1975.
- MILLIKEN, G.A. New criteria for estimability for linear models. *Ann.Math. Statist.*, **42**: 1588 - 1594, 1971.
- MINITAB, REFERENCE MANUAL, ed., Minitab Statistical Computing System, Duxbury Press, Boston 1986.

- MONLEZUN, J. C. & SPEED, F. M. The Geometry of Estimation and Hypothesis Testing in the Constrained Linear Model - The full rank case. **Commun. Statist. -Theor. Meth.** A9 (2): 213-241, 1980.
- MORTAZA J. & JENNRICH, R. I. Nonorthogonal Analysis of Variance using gradient methods. **Jour. Amer. Statist. Assoc.**, 83: 483-489, 1988.
- MSTAT, ed., The User's guide. A Software Program for the Design, Management, and Analysis of Agronomic Research Experiments, **Michigan State University**, Michigan, 1984.
- NETER, J. & WASSERMAN, W. Applied Linear Statistical Models. **Irwin, Homewood**, 1974, 842p.
- OVERALL, J. E. & SPIEGEL, D. K. "Concerning Least Squares Analysis of Experimental Data". **Psychological Bulletin**, 72: 311-322, 1969.
- PATTERSON, R. E. "The Use of Adjusting Factors in Analysis of Data With Disproportionate Subclass Numbers." **Journal of the American Statistical Association**, 41: 334-346, 1946.
- RAGHAVARAO, D. Construction and Combinatorial Problems in designs of Experiments, New York, John Wiley & Sons, 1971
- RAO, C. R. Linear Statistical Inference and Its Applications. New York, John Wiley & Sons, 1965.
- RAO, C. R. & MITRA, S. K. Generalized Inverse of Matrices and its Applications. New York, Wiley, 1971.

- ROY, S. N. & ROY, J. A note on a class of problems in 'normal' multivariate analysis variance. *Ann. Math. Statist.*, 30: 577-581, 1959.
- SAEG, MANUAL, ed., Sistema para Análises Estatísticas. ESAL. Lavras - MG. 1986.
- SANEST, MANUAL DO. ed. Sistema de Análise Estatística para Microcomputadores. CIAGRI/USP, Piracicaba - SP., 1986.
- SAS INSTITUTE, INC., 6ed. SAS user's guide Statistics. SAS Institute, Inc. Cary, N.C., 1982.
- SCHEFFÉ, H. The Analysis of Variance. New York, John Wiley & Sons, 1959.
- SEARLE, S.R. Additional results concerning estimable functions and generalized inverse matrices. *J. Roy. Statist. Soc., Ser. B* 27: 486 - 490, 1965.
- SEARLE, S. R. Arbitrary hypotheses in Linear Models with unbalanced data. *Commun. Statist. -Theor.Meth.*, A9 (2): 181-200, 1980.
- SEARLE, S.R. Linear Models. New York, WILEY, 1971. 532p.
- SEARLE, S.R. Matrix Algebra Useful for Statistics. New York, John Wiley & Sons, 1982.
- SEARLE, S.R. Linear Models for Unbalanced Data. New York, WILEY, 1987. 536p.
- SNEDECOR, G. W. & COCHRAN, W. G. Statistical Methods. 6.ed. Ames, Iowa: The Iowa State University Press, 1967.

- SOC. MANUAL DO, ed. Sistema Operacional C. NTIA - EMBRAPA. Campinas - SP. 1982.
- SOUZA, L. & IEMMA, A. F. Curvas de crescimento: Análise de dados longitudinais incompletos com resíduos autocorrelacionados. *Rev. Mat. Estat., S o Paulo*, 8: 45-55, 1990.
- SPEED, F. M. & MONLEZUN, C. J. "Exact F Tests for the Method of Unweighted Means in a 2^k Experiment". Submitted for publication to *The American Statistical*, 1976.
- SPEED, F.M. & HOCKING, R.R. The use of the $R(\)$ - Notation with unbalanced. *The Amer. Statist.*, 30(1): 30-33, 1976.
- SPEED, F.M.; HOCKING, R.R.; HACKNEY, O.P. Methods of Linear Models with Unbalanced data. *Journal of the American Statistical Association*, 73 (361): 105-112, 1978.
- SPSS, STATISTICAL PACKAGE FOR SOCIAL SCIENCE. SPSS UK Ltda, Hersham, 1986.
- STATGRAPHICS, USER'S GUIDE. ed. Statistical Graphics System. Statistical Graphics Corporation. Rockville, 1985.
- STEEL, R. G. D. & TORRIE, J. M. Principles and Procedures of Statistics. New York, MckGraw-Hill Book Co. 1960.
- SUBRAMANI, J. A Note on Estimation of Several Missing Values in Two Way Classified Data with Many Observations Per Cell. *Biom. J.*, 34(1): 97-102, 1992.
- SYSTAT, INC. ed. The System for Statistics. SYSTAT, Inc. Evanston, IL, 1985.

URQUARDT, N.S. & WEEKS, D.L. Linear Models in Messy Data: Some Problems and Alternatives. **Biometrics**, 34: 896-905, 1978.

VAJDA, S. "Technique of the Analysis of Variance. **Nature**, 160: 27, 1974.

VILCA, M.M. & IEMMA, A.F. Estimability of contrasts in unconnected designs. **Biometrics Bull.**, 7(2): 5, 1990. (Abstracts).

WEEKS, D. L. & WILLIAMS, D. R. A note on the determination of connectedness in an N-way cross classification. **Technometrics**, 6: 319-324, 1964.

WINER, B. J. Statistical Principles in Experimental Design. 2.ed. New York, McGraw-Hill Book Co. 1971.

YATES, F. "Technique of Analysis of Variance." **Nature**, 160: 472-473, 1947.

YATES, F. The analysis of multiple classifications with unequal numbers in the different classes. **Jour. Amer. Statist. Assoc.**, 29: 51-66, 1934.

8. APÊNDICES**A P Ê N D I C E S**

APÊNDICE 1. GLIM

G L I M

[o] GLIM 3.77 (copyright)1984 Royal Statistical Society,
London

```
[i] ? $in 12$
[i] File name? a:eufra.gli
[i] $UNITS 16 $DATA A B Y $READ!
[i] 1 1 10!
[i] 1 1 15!
[i] 1 1 11!
[i] 1 3 14!
[i] 1 4 9!
[i] 1 4 13!
[i] 2 1 8!
[i] 2 1 14!
[i] 2 2 14!
[i] 2 2 16!
[i] 2 4 12!
[i] 2 4 14!
[i] 2 4 13!
[i] 3 2 11!
[i] 3 3 16!
[i] 3 3 18!
[i] $FACTOR A 3 B 4 $YVAR Y $FIT :+A : +B : +A.B $!
[o] deviance = 110.00
[o]      d.f. = 15
[o] deviance = 92.000 (change = -18.00)
[o]      d.f. = 13      (change = -2 )
[o] deviance = 63.766 (change = -28.23)
[o]      d.f. = 10      (change = -3 )
[o] deviance = 46.000 (change = -17.77)
[o]      d.f. = 8       (change = -2 )
```

```
[i] $FIT : +B : +A : +A.B $!  
[o] deviance = 110.00  
[o]      d.f. = 15  
[o]  
[o] deviance = 68.667 (change = -41.33)  
[o]      d.f. = 12      (change = -3   )  
[o]  
[o] deviance = 63.766 (change = -4.901)  
[o]      d.f. = 10      (change = -2   )  
[o]  
[o] deviance = 46.000 (change = -17.77)  
[o]      d.f. = 8       (change = -2   )  
[o]  
[i] $FINISH!  
[i] ? $stop
```

APÊNDICE.2 HARVEY

H A R V E Y

MIXED MODEL LEAST-SQUARES AND MAXIMUM LIKELIHOOD COMPUTER PR
 COPYRIGHT 1987 WALTER R. HARVEY
 PARAMETER CARD GENERATION PROGRAM OUTPUT
 LISTING OF INPUT CONTROL CARDS

TITLE 'ANALISE DADOS ORDEM A/B - SEM INTERACAO'/CENTER;

ANALISE DADOS ORDEM A/B - SEM IN

OVERALL MEANS AND STANDARD DEVIATIONS OF RHM

Y MEAN= 13.00000 S.D.= 2.70801

THE DETERMINANT OF THE CORRELATION MATRIX IS .128125000000

ANALISE DADOS ORDEM A/B - SEM IN
 LEAST-SQUARES ANALYSIS OF VARIANCE

Y

SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	F
TOTAL	16	110.000000		
T.REDUCTION	6	46.234146	7.705691	1.208
MU-YM	1	1.013871	1.013871	.159
A	2	4.900813	2.450407	.384
B	3	28.234146	9.411382	1.476
REMAINDER	10	63.765854	6.376585	

MEAN = 13.00000 ERROR ST. DEVIATION = 2.52519 CV = 19.42 R

ANALISE DADOS ORDEM B/A - SEM IN
LEAST-SQUARES ANALYSIS OF VARIANCE

Y

SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	F
TOTAL	16	110.000000		
T.REDUCTION	6	46.234146	7.705691	1.208
MU-YM	1	1.013871	1.013871	.159
B	3	28.234146	9.411382	1.476
A	2	4.900813	2.450407	.384
REMAINDER	10	63.765854	6.376585	

MEAN = 13.00000 ERROR SD. DEVIATION = 2.52519 CV = 19.42 R

MIXED MODEL LEAST-SQUARES AND MAXIMUM LIKELIHOOD COMPUTER PR

COPYRIGHT 1987 WALTER R. HARVEY

PARAMETER CARD GENERATION PROGRAM OUTPUT

LISTING OF INPUT CONTROL CARDS

TITLE 'ANALISE DADOS ORDEM A/B - COM INTERACAO'/CENTER;

THE INTERACTION OF A AND B

HAS A MISSING SUBCLASS IN THE LAST ROW AND/OR LAST

RECODE THE DATA AND/OR SPECIFY THE NOSORT OPTION

END OF PARMCARD OUTPUT

LISTING OF INPUT CONTROL CARDS

TITLE 'ANALISE DADOS ORDEM B/A - COM INTERACAO'/CENTER;

THE INTERACTION OF B AND A

HAS A MISSING SUBCLASS IN THE LAST ROW AND/OR LAST

RECODE THE DATA AND/OR SPECIFY THE NOSORT OPTION

END OF PARMCARD OUTPUT

APENDICE.3 MINITAB

M I N I T A B

```
è MTB > GLM PROD = FOSFO VARIED
```

Factor	Levels	Values
FOSFO	3	1 2 3
VARIED	4	1 2 3 4

Analysis of Variance for prod

Source	DF	Seq SS	Adj SS	Adj MS	F	P
FOSFO	2	18.000	4.901	2.450	0.38	0.691
VARIED	3	28.234	28.234	9.411	1.48	0.280
Error	10	63.766	63.766	6.377		
Total	15	110.000				

Unusual Observations for prod

Obs.	prod	Fit	Stdev.Fit	Residual	St.Resid
7	8.0000	12.3756	1.4544	-4.3756	-2.12R

R denotes an obs. with a large st. resid.

```
MTB > GLM PROD = VARIED FOSFO
```

Factor	Levels	Values
VARIED	4	1 2 3 4
FOSFO	3	1 2 3

Analysis of Variance for prod

Source	DF	Seq SS	Adj SS	Adj MS	F	P
VARIED	3	41.333	28.234	9.411	1.48	0.280
FOSFO	2	4.901	4.901	2.450	0.38	0.691
Error	10	63.766	63.766	6.377		
Total	15	110.000				

Unusual Observations for prod

Obs.	prod	Fit	Stdev.Fit	Residual	St.Resid
7	8.0000	12.3756	1.4544	-4.3756	-2.12R

R denotes an obs. with a large st. resid.

MTB > stop

*** Minitab Release 7.1 *** Minitab, Inc. ***

Storage available 16179

MTB > GLM PROD = FOSFO VARIED FOSFO*VARIED

Factor	Levels	Values			
FOSFO	3	1	2	3	
VARIED	4	1	2	3	4

Analysis of Variance for prod

Source	Model DF	Reduced DF	Seq SS
FOSFO	2	2	18.0000
VARIED	3	3	28.2341
FOSFO*VARIED	6	2+	17.7659
Error	4	8	46.0000
Total	15	15	110.0000

+ Rank deficiency due to empty cells, unbalanced nesting or collinearity.

No storage of results or further analysis will be done.

VARIED	4	1	2	3	4
FOSFO	3	1	2	3	

Analysis of Variance for prod

Source	Model DF	Reduced DF	Seq SS
VARIED	3	3	41.3333
FOSFO	2	2	4.9008
VARIED*FOSFO	6	2+	17.7659
Error	4	8	46.0000
Total	15	15	110.0000

+ Rank deficiency due to empty cells, unbalanced nesting or collinearity.

No storage of results or further analysis will be done.

MTB > STOP

*** Minitab Release 7.1 *** Minitab, Inc. ***

Storage available 16179

APENDICE.4 MSTAT

M S T A T

Data file EUFRA.DAT

Title: ANALISE

Function: NONORTHO

Data case no. 1 to 16

Without selection

Two-way tables sgrouped over the variable 1 and 2

1 VARIABEL I
with values from 1 to 32 VARIABEL J
with values from 1 to 4

Variable 4

VARIABEL

TABLE OF FREQUENCY

Var 1	Var 2				
	1	2	3	4	Total
1	3	0	1	2	6
2	2	2	0	3	7
3	0	1	2	0	3
Total	5	3	3	5	16

TABLE OF SUMS

Var 1	Var 2				
	1	2	3	4	Sum
1	36	0	14	22	72
2	22	30	0	39	91
3	0	11	34	0	45
Total	58	41	48	61	208

TABLE OF MEANS

Var 1	Var 2				Mean
	1	2	3	4	
1	12		14	11	12
2	11	15		13	13
3		11	17		15
Mean	12	14	16	12	13

ANOVA TABLE

SOURCE	DF	SS	MS	F(W)	P %	F(I)	P%
TOTAL	15	110					
BETWEEN	7	64	9.14	1.59	.264		
WITHIN	8	46	5.75				
Var 1 uncor.	2	18	9.00				
Var 2 uncor.	3	41	13.78				
Var 1 cor.	2	5	2.45	0.43		0.28	
Var 2 cor.	3	28	9.41	1.64	.256	1.06	.519
Interaction	2	18	8.88	1.54	.270		

MEANS FOR VAR 4 OVER EACH LEVEL OF VAR 1

variable	1	number	uncorrected mean	corrected mean
	1	6	12.000	12.403
	2	7	13.000	13.694
	3	3	15.000	12.575

MEANS FOR VAR 4 OVER EACH LEVEL OF VAR 2

variable	2	number	uncorrected mean	corrected mean
	1	5	11.600	11.681
	2	3	13.667	13.346
	3	3	16.000	16.482
	4	5	12.200	12.023

Data file EUFRA.DAT

Title: ANALISE

Function: NONORTHO

Data case no. 1 to 16

Without selection

Two-way tables sgrouped over the variable 2 and 1

2 VARIABEL J
with values from 1 to 4

1 VARIABEL I
with values from 1 to 3

Variable 4

VARIABEL

TABLE OF FREQUENCY

Var 2	Var 1			Total
	1	2	3	
1	3	2	0	5
2	0	2	1	3
3	1	0	2	3
4	2	3	0	5
Total	6	7	3	16

TABLE OF SUMS

Var 2	Var 1			Sum
	1	2	3	
1	36	22	0	58
2	0	30	11	41
3	14	0	34	48
4	22	39	0	61
Total	72	91	45	208

TABLE OF MEANS

Var 2	Var 1			Mean
	1	2	3	
1	12	11		12
2		15	11	14
3	14		17	16
4	11	13		12
Mean	12	13	15	13

ANOVA TABLE

SOURCE	DF	SS	MS	F(W)	P %	F(I)	P%
TOTAL	15	110					
BETWEEN	7	64	9.14	1.59	.264		
WITHIN	8	46	5.75				
Var 2 uncor.	3	41	13.78				
Var 1 uncor.	2	18	9.00				
Var 2 cor.	3	28	9.41	1.64	.256	1.06	.519
Var 1 cor.	2	5	2.45	0.43		0.28	
Interaction	2	18	8.88	1.54	.270		

MEANS FOR VAR 4 OVER EACH LEVEL OF VAR 2

variable	2	number	uncorrected mean	corrected mean
	1	5	11.600	11.681
	2	3	13.667	13.346
	3	3	16.000	16.482
	4	5	12.200	12.022

MEANS FOR VAR 4 OVER EACH LEVEL OF VAR 1

variable	1	number	uncorrected mean	corrected mean
	1	6	12.000	12.405
	2	7	13.000	13.697
	3	3	15.000	12.563

APENDICE.5 SAEG

```

***** ESAL-03 15/07/86 ***** S I S T E M A   P A R A
A N A L I S E S   E S T I S T I C A S   *****
*                                                           *
*                                                           * A N A L I S E
*                                                           *

```

```

***** VS 3.0 18:11:56 *****
***** SAEG *****

```

```

V A R I A V E I S   L I D A S

```

```

VARI    VARJ    REP    VARIABEL

```

```

ANOVAG    MODELO=VARIABEL FUNCAO VARJ VARI VARJ*VARI

```

```

CLASSES    (VARJ=1-4) (VARI=1-3)

```


D I S T R I B U I C A O D O S D A D O S

EFEITO	IDENTIFICACAO	DADOS
VARJ	1	5
VARJ	2	3
VARJ	3	3
VARJ	4	5
VARI	1	6
VARI	2	7
VARI	3	3

INTERACOES			IDENTIFICACAO		DADOS	
VARJ	X	VARI	1	1	3	
VARJ	X	VARI	1	2	2	
VARJ	X	VARI	1	3	0	PERDIDO
VARJ	X	VARI	2	1	0	PERDIDO
VARJ	X	VARI	2	2	2	
VARJ	X	VARI	2	3	1	
VARJ	X	VARI	3	1	1	
VARJ	X	VARI	3	2	0	PERDIDO
VARJ	X	VARI	3	3	2	
VARJ	X	VARI	4	1	2	
VARJ	X	VARI	4	2	3	
VARJ	X	VARI	4	3	0	PERDIDO

NOME	MEDIA	DESVIO
VARIAVEL	13.000000	2.708013

C O R R E L A C O E S

NOME	X NOME	PRODUTO CRUZADO	CORRELACAO
VARIAVEL	VARIAVEL	254.0000	1.0000

```

* * * * *
* EXISTEM CLASSES DE VARIAVEIS INDEPENDENTES *
* OU INTERACOES QUE NAO POSSUEM DADOS. *
* * * * *

```

```

***** ESAL-03 15/07/86 ***** S I S T E M A   P A R A
A N A L I S E S   E S T I S T I C A S   *****
*
*
*
*
***** VS 3.0 18:11:56 *****
***** SAEG *****

```

```

ANOVAG      MODELO=VARIABEL FUNCAO VARJ VARI
CLASSES     (VARJ=1-4) (VARI=1-3)

```

PARGEN

```

E S T A T I S T I C A S       S I M P L E S

OBSERVACOES PERDIDAS          =          0
OBSERVACOES DESCARTADAS       =          0
OBSERVACOES CONSIDERADAS      =         16

```

D I S T R I B U I C A O D O S D A D O S

EFEITO	IDENTIFICACAO	DADOS
VARJ	1	5
VARJ	2	3
VARJ	3	3
VARJ	4	5
VARI	1	6
VARI	2	7
VARI	3	3

NOME	MEDIA	DESVIO
VARIAVEL	13.000000	2.708013

C O R R E L A C O E S

NOME	X NOME	PRODUTO CRUZADO	CORRELACAO
VARIAVEL	VARIAVEL	254.0000	1.0000
DETERMINANTE	=	.1281250	

P A R A M E T R O S C A L C U L A D O S

VARIAVEL DEPENDENTE	EFEITO	CLASSES	OBSER.	MEDIA ESTIMADA	CONST
VARIAVEL	MEDIA		16	13.27114	13.27
VARIAVEL	VARJ	1	5	11.56260	-1.71
VARIAVEL	VARJ	2	3	13.23577	-.04
VARIAVEL	VARJ	3	3	16.38211	3.11
VARIAVEL	VARJ	4	5	11.90407	-1.37
VARIAVEL	VARI	1	6	12.79146	-.47
VARIAVEL	VARI	2	7	14.08415	-.81
VARIAVEL	VARI	3	3	12.93781	-.33

A N A L I S E D E V A R I A N C I A

VARIAVEL

F. DE VARIACAO	G.L	S. QUADRADOS	Q. MEDIO	F	SIGNIF.
TOTAL	16	254.0000			
TOTAL DE RED.	6	190.2342	31.70569	4.972	0.0132
MEDIA	1	147.5703	147.57030	23.143	0.0007
VARJ	3	28.2341	9.41138	1.476	0.2797
VARI	2	4.9008	2.45041	.384	*****
RESIDUO	10	63.7658	6.37658		

COEFICIENTE DE VARIACAO = 19.425

```

***** ESAL-03 15/07/86 ***** S I S T E M A   P A R A
A N A L I S E S   E S T I S T I C A S   *****
*                                                           *
*                                                           * A N A L I S E
*                                                           *
***** VS 3.0   18:11:56 *****
***** SAEG *****

```

```

ANOVAG      MODELO=VARIABEL FUNCAO VARI VARJ
CLASSES    (VARI=1-3) (VARJ=1-4)

```

PARGEN

E S T A T I S T I C A S S I M P L E S

```

OBSERVACOES PERDIDAS      =      0
OBSERVACOES DESCARTADAS   =      0
OBSERVACOES CONSIDERADAS  =     16

```

D I S T R I B U I C A O D O S D A D O S

EFEITO	IDENTIFICACAO	DADOS
VARI	1	6
VARI	2	7
VARI	3	3
VARJ	1	5
VARJ	2	3
VARJ	3	3
VARJ	4	5

NOME	MEDIA	DESVIO
VARIAVEL	13.000000	2.708013

C O R R E L A C O E S

NOME	X NOME	PRODUTO CRUZADO	CORRELACAO
VARIAVEL	VARIAVEL	254.0000	1.0000
DETERMINANTE	=	.1281250	

P A R A M E T R O S C A L C U L A D O S

VARIAVEL DEPENDENTE	EFEITO	CLASSES	OBSER.	MEDIA ESTIMADA	CONST
VARIAVEL	MEDIA		16	13.27114	13.27
VARIAVEL	VARI	1	6	12.79146	-0.48
VARIAVEL	VARI	2	7	14.08415	0.81
VARIAVEL	VARI	3	3	12.93781	-0.33
VARIAVEL	VARJ	1	5	11.56260	-1.71
VARIAVEL	VARJ	2	3	13.23577	-0.04
VARIAVEL	VARJ	3	3	16.38211	3.11
VARIAVEL	VARJ	4	5	11.90407	-1.37

A N A L I S E D E V A R I A N C I A

VARIAVEL

F. DE VARIACAO	G.L	S. QUADRADOS	Q. MEDIO	F	SIGNIF.
TOTAL	16	254.0000			
TOTAL DE RED.	6	190.2342	31.70569	4.972	0.0132
MEDIA	1	147.5703	147.57030	23.143	0.0007
VARI	2	4.9008	2.45041	.384	*****
VARJ	3	28.2342	9.41138	1.476	0.2797
RESIDUO	10	63.7658	6.37658		

COEFICIENTE DE VARIACAO = 19.425

APÊNDICE. 6 SANEST

```

*****
*
*      SANEST - SISTEMA DE ANALISE ESTATISTICA
*
*Autores: Elio Paula Zonta - Amaury Almeida Machado
*INSTITUTO AGRONOMICO DE CAMPINAS - I A C - CAMPINAS
*      ANALISE DA VARIABEL PROD - ARQUIVO: EUFRA1
*
*****

```

CODIGO DO PROJETO: ANALISE
 RESPONSÁVEL: EUFRAZIO
 DELINEAMENTO EXPERIMENTAL: 3x4
 OBSERVAÇÕES NÃO TRANSFORMADAS

NOME DOS FATORES

```

FATOR      NOME
  A        FOSFATO
  B        VARIED

```

QUADRO DA ANALISE DA VARIANCIA

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSFATO AJUST.	2	4.90081	2.45041	0.4262	0.6709
VARIED.NAO AJUST.	3	41.33333	13.77778	2.3961	0.1435
FOSF*VARI AJUST.	2	17.76585	8.88293	1.5449	0.2706
RESIDUO	8	46.00000	5.75000		
TOTAL	15	110.00000			

MEDIA GERAL AJUSTADA = 13.271138
 COEFICIENTE DE VARIACAO = 18.069 %
 S.Q. FOSFATO NÃO AJUSTADA = 18.000000

ANALISE DA VARIANCIA FIXANDO NIVEIS

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSF/VAR 1	1	1.20000	1.20000	0.2087	0.81679
FOSF/VAR 2	1	10.66667	10.66667	1.8551	0.21731
FOSF/VAR 3	1	6.00000	6.00000	1.0435	0.39740
FOSF/VAR 4	1	4.80000	4.80000	0.8348	0.52849
VAR/FOSF 1	2	6.00000	3.00000	0.5217	0.61628
VAR/FOSF 2	2	16.00000	8.00000	1.3913	0.30317
VAR/FOSF 3	1	24.00000	24.00000	4.1739	0.05682
RESIDUO	8	46.00000	5.75000		

```

*****
*
*      SANEST - SISTEMA DE ANALISE ESTATISTICA
*
*Autores: Elio Paula Zonta - Amaury Almeida Machado
*INSTITUTO AGRONOMICO DE CAMPINAS - I A C - CAMPINAS
*
*      ANALISE DA VARIABEL PROD - ARQUIVO: EUFRA2
*
*****

```

CODIGO DO PROJETO: ANALISE
 RESPONSAVEL: EUFRAZIO
 DELINEAMENTO EXPERIMENTAL: 3x4
 OBSERVACOES NAO TRANSFORMADAS

NOME DOS FATORES
FATOR NOME
 A FOSFATO
 --B-----VARIED--

QUADRO DA ANALISE DA VARIANCIA

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSFATO AJUST.	2	4.90081	2.45041	0.4262	0.6709
VARIED. AJUST.	3	28.23415	9.41138	1.6368	0.2561
FOSF*VARI AJUST.	2	17.76585	8.88293	1.5449	0.2706
RESIDUO	8	46.00000	5.75000		
TOTAL	15	110.00000			

MEDIA GERAL AJUSTADA = 13.271138
 COEFICIENTE DE VARIACAO = 18.069 %
 S.Q. FOSFATO NAO AJUSTADA = 18.000000
 S.Q. VARIED NAO AJUSTADA = 41.333333

ANALISE DA VARIANCIA FIXANDO NIVEIS

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSF/VAR 1	1	1.20000	1.20000	0.2087	0.81679
FOSF/VAR 2	1	10.66667	10.66667	1.8551	0.21731
FOSF/VAR 3	1	6.00000	6.00000	1.0435	0.39740
FOSF/VAR 4	1	4.80000	4.80000	0.8348	0.52849
VAR/FOSF 1	2	6.00000	3.00000	0.5217	0.61628
VAR/FOSF 2	2	16.00000	8.00000	1.3913	0.30317
VAR/FOSF 3	1	24.00000	24.00000	4.1739	0.05682
RESIDUO	8	46.00000	5.75000		

```

*****
*
*      SANEST - SISTEMA DE ANALISE ESTATISTICA
*Autores: Elio Paula Zonta - Amaury Almeida Machado
*INSTITUTO AGRONOMICO DE CAMPINAS - I A C - CAMPINAS
*      ANALISE DA VARIABEL PROD - ARQUIVO: EUFRA
*
*****

```

CODIGO DO PROJETO: ANALISE
 RESPONSAVEL: EUFRAZIO
 DELINEAMENTO EXPERIMENTAL: 3x4
 OBSERVACOES NAO TRANSFORMADAS

NOME_DOS_FATORES

FATOR-----NOME---

A FOSFATO

--B-----VARIED--

QUADRO DA ANALISE DA VARIANCIA

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSFATO AJUST.	2	4.90081	2.45041	0.4262	0.6709
VARIED.NAO AJUST.	3	41.33333	13.77778	2.3961	0.1435
RESIDUO	10	63.76585	6.37658		
TOTAL	15	110.00000			

MEDIA GERAL AJUSTADA = 13.271138
 COEFICIENTE DE VARIACAO = 19.028 %
 S.Q. FOSFATO NAO AJUSTADA = 18.000000


```

*****
*                                     *
*      SANEST - SISTEMA DE ANALISE ESTATISTICA      *
*  Autores: Elio Paula Zonta - Amaury Almeida Machado  *
* INSTITUTO AGRONOMICO DE CAMPINAS - I A C - CAMPINAS  *
*      ANALISE DA VARIABEL PROD - ARQUIVO: EUFRA      *
*                                     *
*****

```

CODIGO DO PROJETO: ANALISE
 RESPONSAVEL: EUFRAZIO
 DELINEAMENTO EXPERIMENTAL: 3x4
 OBSERVACOES NAO TRANSFORMADAS

NOME DOS FATORES

FATOR NOME

A FOSFATO

B VARIED

QUADRO DA ANALISE DA VARIANCIA

C.DA VARIACAO	G.L.	S.Q.	Q.M.	VALOR F	PROB.>F
FOSFATO AJUST.	2	4.90081	2.45041	0.4262	0.6709
VARIED. AJUST.	3	28.23415	9.41138	1.4759	0.2794
RESIDUO	10	63.76585	6.37658		
TOTAL	15	110.00000			

MEDIA GERAL AJUSTADA = 13.271138
 COEFICIENTE DE VARIAÇÃO = 19.028 %
 S.Q. FOSFATO NAO AJUSTADA = 18.000000
 S.Q. VARIED NAO AJUSTADA = 41.333333

APENDICE.7 SAS

S A S

ordem fosf varied com interacao
General Linear Models Procedure

Dependent Variable: PROD

Source	DF	Sum of Squares	Mean Square	F Value	Pr >F
Model	7	64.00000000	9.14285714	1.59	0.264
Error	8	46.00000000	5.75000000		
C.Total	15	110.00000000			

R-Square	C.V.	Root MSE	PESO Mean
0.581818	18.445506	2.39791576	13.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr >F
FOSF	2	18.00000000	9.00000000	1.57	0.266
VARIED	3	28.23414634	9.41138211	1.64	0.256
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type II SS	Mean Square	F Value	Pr >F
FOSF	2	4.90081301	2.45040650	0.43	0.667
VARIED	3	28.23414634	9.41138211	1.64	0.256
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type III SS	Mean Square	F Value	Pr >F
FOSF	2	8.71264368	4.35632184	0.76	0.499
VARIED	3	26.36213818	8.78737939	1.53	0.280
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type IV SS	Mean Square	F Value	Pr >F
FOSF	2*	16.66666667	8.33333333	1.45	0.290
VARIED	3*	16.53658537	5.51219512	0.96	0.457
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

* NOTE: Other Type IV Testable Hypotheses exist which may yield different SS.

ordem varied fosf com interacao
General Linear Models Procedure

Dependent Variable: PROD

Source	DF	Sum of Squares	Mean Square	F Value	Pr >F
Model	7	64.00000000	9.14285714	1.59	0.264
Error	8	46.00000000	5.75000000		
C.Total	15	110.00000000			

R-Square	C.V.	Root. MSE	PESO Mean
0.581818	18.445506	2.39791576	13.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr >F
VARIED	3	41.33333333	13.77777778	2.40	0.143
FOSF	2	4.90081301	2.45040650	0.43	0.667
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type II SS	Mean Square	F Value	Pr >F
VARIED	3	28.23414634	9.41138211	1.64	0.256
FOSF	2	4.90081301	2.45040650	0.43	0.667
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type III SS	Mean Square	F Value	Pr >F
VARIED	3	26.36213818	8.78737939	1.53	0.280
FOSF	2	8.71264368	4.35632184	0.76	0.499
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

Source	DF	Type IV SS	Mean Square	F Value	Pr >F
VARIED	3*	16.53658537	5.51219512	0.96	0.457
FOSF	2*	16.66666667	8.33333333	1.45	0.290
FOSF*VAR	2	17.76585366	8.88292683	1.54	0.270

* NOTE: Other Type IV Testable Hypotheses exist which may yield different SS.

ordem fosf varied sem interacao
General Linear Models Procedure

Dependent Variable: PROD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	46.23414634	9.24682927	1.45	0.2880
Error	10	63.76585366	6.37658537		
C.Total	15	110.00000000			

R-Square	C.V.	Root MSE	PESO Mean
0.420310	19.424540	2.52519016	13.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
FOSF	2	18.00000000	9.00000000	1.41	0.2885
VARIED	3	28.23414634	9.41138211	1.48	0.2797

Source	DF	Type II SS	Mean Square	F Value	Pr > F
--------	----	------------	-------------	---------	--------

FOSF	2	4.90081301	2.45040650	0.38	0.6906
------	---	------------	------------	------	--------

VARIED	3	28.23414634	9.41138211	1.48	0.2797
--------	---	-------------	------------	------	--------

Source	DF	Type III SS	Mean Square	F Value	Pr > F
--------	----	-------------	-------------	---------	--------

FOSF	2	4.90081301	2.45040650	0.38	0.6906
------	---	------------	------------	------	--------

VARIED	3	28.23414634	9.41138211	1.48	0.2797
--------	---	-------------	------------	------	--------

Source	DF	Type IV SS	Mean Square	F Value	Pr > F
--------	----	------------	-------------	---------	--------

FOSF	2	4.90081301	2.45040650	0.38	0.6906
------	---	------------	------------	------	--------

VARIED	3	28.23414634	9.41138211	1.48	0.2797
--------	---	-------------	------------	------	--------

ordem varied fosf sem interacao

General Linear Models Procedure

Dependent Variable: PROD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
--------	----	----------------	-------------	---------	--------

Model	5	46.23414634	9.24682927	1.45	0.2880
-------	---	-------------	------------	------	--------

Error	10	63.76585366	6.37658537		
-------	----	-------------	------------	--	--

G.Total	15	110.00000000			
---------	----	--------------	--	--	--

R-Square

C.V.

Root MSE

PESO Mean

0.420310

19.424540

2.52519016

13.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
VARIED	3	41.33333333	13.77777778	2.16	0.1559
FOSF	2	4.90081301	2.45040650	0.38	0.6906

Source	DF	Type II SS	Mean Square	F Value	Pr > F
VARIED	3	28.23414634	9.41138211	1.48	0.279
FOSF	2	4.90081301	2.45040650	0.38	0.690

Source	DF	Type III SS	Mean Square	F Value	Pr > F
VARIED	3	28.23414634	9.41138211	1.48	0.279
FOSF	2	4.90081301	2.45040650	0.38	0.690

Source	DF	Type IV SS	Mean Square	F Value	Pr > F
VARIED	3	28.23414634	9.41138211	1.48	0.279
FOSF	2	4.90081301	2.45040650	0.38	0.690

APÊNDICE.8 SOC

S O C

```

-----
obs  fosfo  varied  producao
-----

```

```

  1      1      1      10
  2      1      1      15
  3      1      1      11
  4      1      3      14
  5      1      4      9
  6      1      4      13
  7      2      1      8
  8      2      1      14
  9      2      2      14
 10      2      2      16
 11      2      4      12
 12      2      4      13
 13      2      4      14
 14      3      2      11
 15      3      3      16
 16      3      3      18

```

ordem fosfo varied com interacao

Descricao das variaveis e fatores do modelo

Fatores	# niveis	Niveis			
fosfo	3	1	2	3	
varied	4	1	2	3	4

Variavel dependente: producao

Numero de observacoes lidas : 16

Numero de observacoes utilizadas : 16

Quadro de analise de variancia

Variavel Dependente : producao

F.de variacao	gl	S. quadrados	Q. medio	Valor F	PR >F
Sequencial					
fosfo	2	18.00000000	9.00000000	1.5652	0.267
varied	3	28.23414634	9.41138211	1.6368	0.256
fosfo*varied	2	17.76585366	8.88292683	1.5449	0.271
Residuo	8	46.00000000	5.75000000		
Total	15	110.00000000			

Media : 13.00000000

Raiz Quad. QMres. : 2.39791576

Coef. Variacao : 18.44550586

Atencao: A matriz $X'X$ e' singular. Coluna associada ao efeito

fosfo*varied e' uma combinacao linear de colunas associadas a efeitos que o precedem no modelo. Nesta situacao nao existe solucao unica para as equacoes normais, mesmo com a imposicao das restricoes nos parametros, e as somas de quadrados parciais nao sao fornecidas

ordem varied fosfo com interacao

Descricao das variaveis e fatores do modelo

Fatores	# niveis	Niveis			
varied	4	1	2	3	4
fosfo	3	1	2	3	

Variavel dependente: producao

Numero de observacoes lidas : 16

Numero de observacoes utilizadas : 16

Quadro de analise de variancia

Variavel Dependente : producao

F. variacao	gl	S. quadrados	Q. medio	Valor F	PR > F
Sequencial					
varied	3	41.33333333	13.77777778	2.3961	0.144
fosfo	2	4.90081301	2.45040650	0.4262	0.667
varied*fosfo	2	17.76585366	8.88292683	1.5449	0.271
Residuo	8	46.00000000	5.75000000		
Total	5	110.00000000			

Media : 13.00000000

Raiz Quad. QMres. : 2.39791576

Coef. Variacao : 18.44550586

Atencao: A matriz $X'X$ e' singular. Coluna associada ao efeito

varied*fosfo e' uma combinacao linear de colunas associadas a efeitos que o precedem no modelo. Nesta situacao nao existe solucao unica para as equacoes normais, mesmo com a imposicao das restricoes nos parametros, e as somas de quadrados parciais nao sao fornecidas

ordem fosfo varied sem interacao

Descricao das variaveis e fatores do modelo

Fatores	# niveis	Niveis			
fosfo	3	1	2	3	
varied	4	1	2	3	4

Variavel dependente: producao

Numero de observacoes lidas : 16

Numero de observacoes utilizadas : 16

Quadro de analise de variancia

Variavel Dependente : producao

F.variacao	gl	S. quadrados	Q. medio	Valor F	PR > F
Sequencial					
fosfo	2	18.00000000	9.00000000	1.4114	0.288
varied	3	28.23414634	9.41138211	1.4759	0.280
Residuo	0	63.76585366	6.37658537		
Total	5	110.00000000			

F.variacao	gl	S. quadrados	Q. medio	Valor F	PR > F
Parcial					
fosfo	2	4.90081301	2.45040650	0.3843	0.691
varied	3	28.23414634	9.41138211	1.4759	0.280

Media	:	13.00000000
Raiz Quad. QMres.	:	2.52519016
Coef. Variacao	:	19.42453973

ordem varied fosfo sem interacao

Descricao das variaveis e fatores do modelo

Fatores	# niveis	Niveis			
varied	4	1	2	3	4
fosfo	3	1	2	3	

Variavel dependente: producao

Numero de observacoes lidas : 16

Numero de observacoes utilizadas : 16

Quadro de analise de variancia

Variavel Dependente : producao

F.variacao	gl	S. quadrados	Q. medio	Valor F	PR > F
Sequential					
varied	3	41.33333333	13.77777778	2.1607	0.156
fosfo	2	4.90081301	2.45040650	0.3843	0.691
Residuo	0	63.76585366	6.37658537		
Total	5	110.00000000			

F. variacao	gl	S. quadrados	Q. medio	Valor F	PR > F
Parcial					
varied	3	28.23414634	9.41138211	1.4759	0.280
fosfo	2	4.90081301	2.45040650	0.3843	0.691

Media	13.00000000
Raiz Quad. QMres.	2.52519016
Coef. Variacao	19.42453973

APENDICE.9 SPSS

S P S S

SPSS/PC+ The Statistical Package for IBM PC

2/12/93

* * * A N A L Y S I S O F V A R I A N C E * * *

RESP
BY A
B

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of
Main Effect	46.234	5	9.247	1.608	.262
A	4.901	2	2.450	.426	.667
B	28.234	3	9.411	1.637	.256
2-way Interactions	17.766	2	8.883	1.545	.271
A B	17.766	2	8.883	1.545	.271
Explained	64.000	7	9.143	1.590	.264
Residual	46.000	8	5.750		
Total	110.000	15	7.333		

16 Cases were processed.

0 Cases (.0 PCT) were missing.

This procedure was completed at 21:39:03

anova/variables resp by b(1,4) a(1,3).

'ANOVA' PROBLEM REQUIRES

1780 BYTES OF MEMORY.

* * * A N A L Y S I S O F V A R I A N C E * * *

RESP
BY B
A

Source of Variatio	Sum of Squares	DF	Mean Square	F	Signif of F
Main Effects	46.234	5	9.247	1.608	.262
	4.901	2	2.450	.426	.667
2-way Interactions	17.766	2	8.883	1.545	.271
B A	17.766	2	8.883	1.545	.271
Explained	64.000	7	9.143	1.590	.264
Residual	46.000	8	5.750		
Total	110.000	15	7.333		

16 Cases were processed.

0 Cases (.0 PCT) were missing.

This procedure was completed at 21:39:33

anova/variables resp by b(1,4) a(1,3) /option 3.

'ANOVA' PROBLEM REQUIRES

680 BYTES OF MEMORY.

*** ANALYSIS OF VARIANCE ***

RESP					
BY					
B					
A					
Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Main Effects	46.234	5	9.247	1.450	.288
B	28.234	3	9.411	1.476	.280
A	4.901	2	2.450	.384	.691
Explained	46.234	5	9.247	1.450	.288
Residual	63.766	10	6.377		
Total	110.000	15	7.333		

16 Cases were processed.

0 Cases (.0 PCT) were missing.

This procedure was completed at 21:40:02

anova/variables resp by a(1,3) b(1,4) /option 3.

*** ANALYSIS OF VARIANCE ***

RESP

BY

A

B

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Main Effects	46.234	5	9.247	1.450	.288
A	4.901	2	2.450	.384	.691
B	28.234	3	9.411	1.476	.280
Explained	46.234	5	9.247	1.450	.288
Residual	63.766	10	6.377		
Total	110.000	15	7.333		

16 Cases were processed.

0 Cases (.0 PCT) were missing.

This procedure was completed at 21:40:39
stop.

APENDICE.10 STATGRAPHICS

S T A T G R A P H I C

Analysis of Variance for C:EUFRA.y

Source of variation	Sum of Squares	d.f. level	Mean square	F-ratio	Sig.
MAIN EFFECTS	46.234146	5	9.2468293	1.450	.2880
C:EUFRA.a	4.900813	2	2.4504065	.384	.6906
C:EUFRA.b	28.234146	3	9.4113821	1.476	.2797
RESIDUAL	63.765854	10	6.3765854.		
TOTAL	110.00000	15			
(CORR.)					
0 missing values have been excluded.					

Analysis of Variance for C:EUFRA.y

Source of variation	Sum of Squares	d.f.	Mean square	F-ratio	Sig. level
MAIN EFFECTS	46.234146	5	9.2468293	1.450	.288
C:EUFRA.b	28.234146	3	9.4113821	1.476	.279
C:EUFRA.a	4.900813	2	2.4504065	.384	.690
RESIDUAL	63.765854	10		6.376585	
TOTAL			110.00000		1
(CORR.) 0 missing values have been excluded.					

APÊNDICE.11 SYSTAT

S Y S T A T

ANALISE ORDEM A B - SEM INTERACAO

DEP VAR: Y N: 16 MULTIPLE R: .648 SQUARED MULTIPLE R: .420

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
A	4.901	2	2.450	0.384	0.691
B	28.234	3	9.411	1.476	0.280
ERROR	63.766	10	6.377		

ANALISE ORDEM B A - SEM INTERACAO

DEP VAR: Y N: 16 MULTIPLE R: .648 SQUARED MULTIPLE R: .420

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
B	28.234	3	9.411	1.476	0.280
A	4.901	2	2.450	0.384	0.691
ERROR	63.766	10	6.377		

ANALISE ORDEM A B - COM INTERACAO

ANALISE ORDEM B A - COM INTERACAO