

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA

**Alternativas para ANOVA e para Comparações Múltiplas
em Modelos Normalmente Distribuídos com Fatores Fixos
que Apresentam Problema de Heterogeneidade de
Variâncias**

Autor: Manoel Mendonça Silveira

Orientador: Suzi Alves Camey

Monografia apresentada para obtenção
do grau de Bacharel em Estatística

PORTO ALEGRE, DEZEMBRO DE 1999.

Dedico à
À Marisol,
À Omayra,
Aos meus pais.

AGRADECIMENTOS

À professora Suzi Alves Camey, pela orientação e incentivo.

Às professoras Jandyra M. G. Fachel e Elsa C. Mundstock que me despertaram para a estatística.

Aos professores Aron Taitelbaum e Maria Medianeira S. Gonzalez que deram condições e tempo para realização do Curso.

Aos professores do Departamento de Estatística por todo os conhecimentos recebidos.

Aos colegas de Curso e colegas funcionários do Instituto de Matemática.

À minha esposa e filha, por estarem sempre presentes nos momentos mais decisivos, pela compreensão e apoio.

SUMÁRIO

1. CONSIDERAÇÕES GERAIS	6
1.1 Introdução	6
1.2 Justificativa	9
1.3 Nota Histórica	10
2. METODOLOGIA	11
3. ANÁLISE DE VARIÂNCIA (ANOVA)	12
3.1 Modelo Matemático	12
3.2 Suposições do Modelo	13
3.3 Análise de Variância com um Fator: Modelo I	14
3.4 Teste de Efeito dos Tratamentos	15
3.5 Teste de Homogeneidade de Variâncias	16
3.5.1 Teste de Levene	16
3.5.2 Teste de Bartlett	17
3.6 Teste F proposto por Welch	18
3.7 Teste F proposto por Brown-Forsythe	19
4. COMPARAÇÕES MÚLTIPLAS EM MODELOS	
ANOVA COM UM FATOR COM VARIÂNCIAS DESIGUAIS	21
4.1 Método de Banerjee	22

4.1.1	Comparações Múltiplas com Controle	23
4.1.2	Comparações de Médias duas a duas e Contrastes Ortogonais	24
4.2	Comparações Múltiplas baseadas no Método de Welch	25
4.2.1	Comparações Múltiplas com Controle	25
4.2.2	Comparação de Médias duas a duas e Contrastes Ortogonais	26
4.3	Procedimento baseado no Método de Hochberg	27
4.3.1	Comparação de Médias duas a duas e Contrastes Não-Ortogonais	27
4.4	Procedimento baseado no Método de Brown-Forsythe	29
4.4.1	Comparação de Médias duas a duas e Contrastes Ortogonais	29
5.	ANOVA COM DOIS FATORES FIXO	30
5.1	Modelo Matemático	30
5.2	Teste de Efeito de Tratamentos	31
6.	APLICAÇÃO	34
7.	ROTINA NO EXCEL	44
8.	CONCLUSÃO	47
9.	REFERÊNCIAS BIBLIOGRÁFICAS	50
10.	APÊNDICE	52

1. CONSIDERAÇÕES GERAIS

1.1. Introdução

Quando temos t grupos de observações independentes, originadas de populações normalmente distribuídas e buscamos compará-las através de suas médias, utilizamos a análise de variância (ANOVA) para determinar se existe pelo menos uma média de um grupo que se diferencie das demais. Quando, pela ANOVA conclui-se que as médias são significativamente diferentes, pode-se comparar as médias dos grupos duas a duas por um simples teste t o que nem sempre é a melhor solução, pois quando são realizados testes com este procedimento, a cada comparação a significância é acrescida de um erro.

Para solucionar este problema, utiliza-se o método das comparações múltiplas, o qual tem sido desenvolvido para vários problemas de comparação, como os descritos abaixo:

P1) Intervalo de confiança para todas as diferenças $\mu_i - \mu_k$ ($1 \leq i \leq t, i \neq k$) onde consideramos μ_k como a média de uma população controle e t o número de populações a comparar.

P2) Intervalo de confiança para todas as diferenças duas a duas $\mu_i - \mu_k$

($1 \leq i \leq t, 1 \leq k \leq t, i \neq k$) e todos os contrastes lineares $\sum_{i=1}^t c_i \mu_i$ onde

c_1, c_2, \dots, c_k é um conjunto arbitrário de constantes reais satisfazendo a condição

$$\sum_{i=1}^k c_i = 0.$$

P3) Intervalo de confiança para todas combinações lineares $\sum_{i=1}^t a_i \mu_i$ onde

a_1, a_2, \dots, a_k é um conjunto qualquer de constantes reais.

Tanto a ANOVA como as comparações múltiplas têm como hipótese a homogeneidade das variância. Neste caso, encontramos vasta literatura descrevendo as suas metodologias.

Quando as variâncias dos grupos são possivelmente heterogêneas, técnicas alternativas foram estudadas e serão apresentadas neste trabalho. Além destas técnicas, outros procedimentos também são eficazes para contornar este tipo de problema, como por exemplo a transformação dos dados observados para uma outra escala que reduza a variabilidade e em onde os dados não satisfazem as exigências para uma análise paramétrica dos dado um procedimento natural é recorrer a análises não-paramétricas.

Mais especificamente daremos alternativas para o problema de heterogeneidade de variâncias na ANOVA quando temos um ou dois fatores fixos. Estas alternativas surgem como uma extensão da solução do problema de Behrens-Fisher. O problema de Behrens-Fisher se caracteriza quando desejamos comparar as

médias de duas amostras independentes, normalmente distribuídas e que não podemos assumir igualdade entre as variâncias dos grupos. Este problema tem uma demonstração difícil e todas as possíveis soluções são geralmente criticadas (Tamhane [6]).

A proposta de fazer inferência sobre dados que apresentassem variâncias heterogêneas recebeu atenção de muitos autores. No caso do teste F convencional (para variâncias homogêneas) citaremos as alternativa de Welch [3], que é uma extensão modificada para o teste t considerando k grupos e a alternativa de Brown-Forsythe [4] que propõe uma discussão para o uso do procedimento de Scheffé aplicando uma aproximação para as distribuições de algumas estatísticas como as usadas no problema sugerido por Behrens-Fisher.

Para as comparações múltiplas serão mostradas os métodos de Welch [5] e o método de Banerjee [6], que são testes aproximados. Também será discutido o teste exato sugerido por Hochberg [7]. Todos os testes de comparações múltiplas mencionados utilizam intervalos de confiança na sua concepção.

No trabalho de Brown-Forsythe [4], existe uma alternativa para o teste F quando temos problema de heterogeneidade de variâncias para a ANOVA com dois fatores fixos, que será mostrada no Capítulo 5.

No Apêndice acrescentamos a alternativa de Welch para o problema de Behrens-Fisher e alguns lemas usados por Banerjee pra obter os intervalos de confiança para comparações múltiplas.

Neste trabalho serão simuladas três situações para comparar os resultados dos teste tradicionais com os alternativos:

S1) grupos com médias diferentes e variâncias levemente heterogêneas, que propiciam uma aceitação da hipótese de igualdade das variâncias mas com uma probabilidade de erro tipo I não muito superior a 0,05;

S2) grupos com médias diferentes e variâncias acentuadamente diferentes; e

S3) grupos com médias iguais e variâncias acentuadamente diferentes.

Para obter os resultados dos testes propostos que serão comparados, foi implementado, no programa Excel, uma rotina de cálculos que a partir dos dados observados dará como resposta primeiramente uma tabela com as estatísticas F calculadas assumindo a homogeneidade dos grupos a serem comparados (F convencional) e mais duas, a F_W e a F_{BF} , sugeridas pelos autores Welch [3] e Brown-Forsythe [4].

Na seqüência serão mostrados os resultados dos procedimentos de comparações múltiplas através de intervalos de confiança construídos pelos métodos sugeridos pelos autores Welch [6], Banerjee [6] e Hochberg [7] e usaremos o teste de Tukey para comparar com as técnicas alternativas.

1.2. Justificativa

Pesquisadores das mais diversas áreas, quando têm por objetivo o teste de efeitos fixos através do modelo de Análise de Variância e posterior comparação múltipla de médias, freqüentemente se deparam com grupos que apresentam variâncias heterogêneas.

Com o objetivo de estudar e trazer à discussão as alternativas para o teste F e comparações múltiplas quando temos um fator fixo e não é possível assumir a

igualdade das variâncias dos grupos, este trabalho realiza um levantamento das principais referências bibliográficas sobre o assunto.

1.3. Nota Histórica

Neste breve histórico, começamos com o procedimento alternativo sugerido por Welch (1938) através de intervalos de confiança que garantem um nível de confiança aproximado e envolve distribuições t student com grau de liberdade aleatório. Banerjee (1961) propõe um conjunto de intervalos que garantem um nível de confiança específico para diferenças entre médias. Em 1962, Banerjee e Slepian implementaram um método para obtenção de intervalos com uma população controle e para contrastes ortogonais.

Uma importante referência é Miller (1966) que faz um apanhado de todos esses métodos de comparações múltiplas, descritos até 1966.

Quando o teste de comparações múltiplas envolvem intervalos e comparações não ortogonais, os procedimentos dados por Banerjee são conservadores, mas como alternativa ainda podemos citar Spjøtvoll (1972), baseado no teste de Scheffé e Hochberg (1976) baseado no teste de Tukey. Procedimentos buscando a solução do problema de Behrens-Fisher são sugeridos através de cálculos em dois estágios, mas essas técnicas não serão abordadas neste trabalho, mas cabe mencionar Stein [15], Chapman [16] e Gosh [17].

3. ANÁLISE DE VARIÂNCIA (ANOVA)

Sabemos que duas médias amostrais podem ser comparadas pelo teste t para duas amostras. A generalização deste problema para k grupos, $k > 2$, nos leva para um conjunto de técnicas conhecido como Análise de Variância. Temos um modelo com um fator quando existe somente uma classificação considerada. Basicamente, existem duas situações distintas, que são modelos fixos (modelo I) e modelos com fatores aleatórios (modelo II). Vamos restringir nosso contexto ao modelo I, quando os grupos da amostra representam grupos específicos nos quais estamos interessados, onde os resultados podem ser considerados apenas para o conjunto especificado de grupos.

3.1. Modelo Matemático

A forma mais simples de um experimento ocorre quando comparamos os níveis de um único fator e a mesma característica é medida em cada indivíduo (Johnson e Leone [8] e Gomes [9]). Suponhamos que temos t grupos de observações $Y_{11}, \dots, Y_{1n_1}, Y_{21}, \dots, Y_{2n_2}, \dots, Y_{t1}, \dots, Y_{tn_t}$ independentes, de k populações normalmente distribuídas com médias μ_1, \dots, μ_k , e com mesma variância σ^2 .

O modelo de análise de variância considerando um fator fixo é dado por

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \begin{cases} i = 1, \dots, t \\ j = 1, \dots, n_i \end{cases} \quad (3.1)$$

onde:

Y_{ij} : observação da variável resposta da j -ésima repetição do tratamento i ;

μ : constante inerente a todos os dados (média geral);

τ_i : efeito do tratamento i , modelo fixo;

ε_{ij} : erro experimental (não observável);

t : número de níveis de um único fator;

n_i : número de repetições para o tratamento i .

3.2 Suposições do modelo

As suposições do modelo para ANOVA fixo com um fator, de acordo com Johnson e Leone [8] e Gomes [9], são as seguintes:

- i) Y_{ij} são funções lineares de ε_{ij} ;
- ii) os erros ε_{ij} são independentes e têm distribuição $N(0, \sigma^2)$;

3.3. Análise de variância com um fator: modelo I

Para proceder o teste F convencional da análise de variância, as condições de normalidade, aditividade e homogeneidade de variâncias devem estar satisfeitas. Esses resultados são convenientemente e convencionalmente resumidos numa tabela de análise de variância.

Quadro 1: Quadro da análise de variância

Causas da Variação	G.L.	Soma de Quadrado	Quadrado Médio	F
Tratamento (entre grupos)	$t - 1$	SQT	QMT	$\frac{QMT}{QME}$
Erro	$\sum_{i=1}^t n_i - t$	SQE	QME	
Total	$\sum_{i=1}^t n_i - 1$	SQtotal		

Onde

$$QME = \frac{\sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{\left(\sum_{i=1}^t n_i\right) - t},$$

$$\bar{y}_i = \sum_{j=1}^{n_i} \frac{y_{ij}}{n_i}, \text{ é a média do tratamento } i, \text{ e}$$

$$\bar{y}_{..} = \frac{\sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^t n_i} \text{ é a média geral.}$$

$$QMT = \frac{\sum_{i=1}^t n_i \left(\bar{y}_{i.} - \bar{y}_{..} \right)^2}{t-1}.$$

Sabemos também que:

$$SQT = \sum_{i=1}^t \left(\bar{y}_{i.} - \bar{y}_{..} \right)^2 = \sum_{i=1}^t \frac{\left(\sum_{j=1}^{n_i} y_{ij} \right)^2}{n_i} - \frac{\left(\sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij} \right)^2}{\sum_{i=1}^t n_i},$$

$$SQE = \sum_{i=1}^t \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_{i.} \right)^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^t \frac{\left(\sum_{j=1}^{n_i} y_{ij} \right)^2}{n_i},$$

$$SQ_{total} = \sum_{i=1}^t \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_{..} \right)^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \frac{\left(\sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij} \right)^2}{\sum_{i=1}^t n_i}$$

o que às vezes pode ser útil para o cálculo das somas de quadrado.

3.4 Teste de Efeito dos Tratamentos

A hipótese de nulidade é

$$H_0: \tau_i = 0 \quad \forall i = 1, 2 \dots, t,$$

que é equivalente a testar $H_0: \mu_i = \mu \quad \forall i = 1, 2 \dots, t$, ou seja, desejamos testar se o efeito dos tratamentos, em média, é nulo. Para testar esta hipótese, usaremos a estatística F , que consta no Quadro 1.

Rejeitaremos a hipótese se o valor da probabilidade de erro tipo I for menor do que o nível de significância adotado.

Cabe salientar que a hipótese alternativa da ANOVA é que $\tau_i \neq 0$ para algum i , isto implica que quando rejeitamos H_0 não sabemos qual tratamento tem efeito não nulo, por isso a necessidade de realizarmos testes de comparações múltiplas que serão descritos mais adiante.

3.5. Teste de homogeneidade de variâncias

Vamos citar as técnicas de Bartlett e Levene para verificar a suposição de homogeneidade de variâncias. Entretanto, percebemos que a estatística F (na ANOVA) é um teste robusto para diferenças de médias desde que:

- a) o número de observações por grupo é superior a dez;
- b) as médias através dos grupos não são correlacionadas com os desvios padrão através dos grupos, o que pode ser facilmente verificado através de um gráfico entre as médias e os desvios padrão de cada tratamento.

3.5.1. Teste de Levene

O teste de Levene é um teste robusto para testar a condição de homogeneidade de variâncias. Neste método, é realizada uma análise de variância para a variável dependente X_{ij} dada por

$$X_{ij} = \left| Y_{ij} - \bar{Y}_{i.} \right|$$

onde $\bar{Y}_{i.}$ é média de cada tratamento.

Se o teste de Levene é estatisticamente significativo, então a hipótese da homogeneidade de variância pode ser rejeitada.

3.5.2. Teste de Bartlett

Outro teste é o de Bartlett que é baseado na distribuição qui-quadrado. A estatística do teste é dada por, Riboldi [22]

$$\chi^2 = \frac{M}{C},$$

onde:

$$M = 2,3026 \left[\sum_{i=1}^t (r_i - 1) \log \bar{s}^2 - \sum_{i=1}^t (r_i - 1) \log s_i^2 \right],$$

o valor 2,3026 converte logaritmo na base decimal para logaritmo na base e,

$$s^2 = \frac{\sum_{i=1}^t (r_i - 1) s_i^2}{\sum_{i=1}^t (r_i - 1)},$$

$$C = 1 + \frac{1}{3(t-1)} \left[\sum_{i=1}^t \frac{1}{(r_i - 1)} - \frac{1}{\sum_{i=1}^t (r_i - 1)} \right],$$

Se $r_1 = r_2 = \dots = r_t = r$, então

$$M = 2,3026(r-1) \left[t \log \bar{s}^2 - \sum_{i=1}^t \log \bar{s}_i^2 \right],$$

$$\bar{s}^2 = \frac{\sum_{i=1}^t s_i^2}{t},$$

$$C = 1 + \frac{t+1}{3t(r-1)}.$$

Sob H_0 $\frac{M}{C} \cap \chi^2_{(t-1)}$, logo rejeita-se H_0 ao nível α de significância se

$$\chi^2_{\text{calculado}} > \chi^2_{\alpha(t-1)}.$$

3.6. Teste F proposto por Welch

Podemos entender a ANOVA com um fator como uma generalização do teste t-student para duas amostras. Considerando t amostras para o modelo (3.1), a estatística F proposta por Welch é dada por

$$F_W = \frac{\left[\sum_{i=1}^t w_i (\bar{y}_i - \bar{y}_{..})^2 \right] / (t-1)}{\left[1 + \frac{2(t-2)}{t^2-1} \sum_{i=1}^t \frac{1}{f_i} \left(1 - \frac{w_i}{\sum_{i=1}^t w_i} \right)^2 \right]} \quad (3.2)$$

onde:

$$w_i = n_i / s_i^2,$$

$$f_i = n_i - 1,$$

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{(n_i - 1)},$$

$$\bar{y}_{..} = \frac{\sum_{i=1}^t w_i \bar{y}_i}{\sum_{i=1}^t w_i} .$$

Sob a hipótese nula temos que

$$F_w \sim F_{(f_1, f_2)}$$

onde:

$$f_1 = t - 1 \text{ e}$$

$$f_2 = \left[\frac{3}{t^2 - 1} \sum_{i=1}^t \frac{1}{f_i} \left(1 - \frac{w_i}{\sum_{i=1}^t w_i} \right)^2 \right]^{-1} .$$

Quando temos $t = 2$, este procedimento se reduz a um teste t para duas amostras com variâncias desiguais.

3.7. Teste F proposto por Brown e Forsythe

Quando não é satisfeita a condição de homogeneidade de variância entre os grupos, podemos proceder o teste proposto por Brown-Forsythe.

A soma de quadrados para o numerador usual da ANOVA com um fator é dado por $\sum_{i=1}^t n_i (\bar{y}_i - \bar{y}_{..})^2$ e seu valor esperado sob a hipótese nula

$$\text{é } \sum_{i=1}^t (1 - n_i/n) \sigma_i^2 \text{ onde } n = \sum_{i=1}^t n_i \text{ e } \bar{y} = \frac{1}{n} \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij} .$$

Então, a hipótese nula pode ser testada pela estatística

$$F_{BF} = \frac{\sum_{i=1}^t n_i (\bar{y}_{i.} - \bar{y}_{..})^2}{\sum_{i=1}^t \left(1 - \frac{n_i}{n}\right) s_i^2} \quad (3.3)$$

Esta estatística pode ser aproximada por uma estatística F com $t - 1$ e f graus de liberdade, onde

$$\frac{1}{f} = \sum_{i=1}^t \frac{c_i^2}{n_i - 1}, \text{ e}$$

$$c_i = \frac{\left(1 - \frac{n_i}{n}\right) s_i^2}{\sum_{i=1}^t \left(1 - \frac{n_i}{n}\right) s_i^2}, \text{ conforme Satterthwaite [23].}$$

4. COMPARAÇÕES MÚLTIPLAS EM MODELOS ANOVA COM UM FATOR COM VARIÂNCIAS DESIGUAIS

Quando, pela ANOVA conclui-se que as médias são significativamente diferentes, não podemos definir quais médias diferem entre si. Procede-se, então um teste de comparações múltiplas, para então, determinar quais médias estão mais próximas e quais realmente são responsáveis pela diferença observada.

Quando as variâncias são heterogêneas e desconhecidas, o problema das comparações múltiplas torna-se relativamente mais complicado. Vários métodos têm sido sugeridos para solucionar este problema. Em especial, destacaremos as sugestões de Banerjee, Welch e Hochberg.

Segundo Tamhane[6] "procedimentos conservadores baseados no método de Banerjee para a solução do problema de Behrens-Fisher foram propostos quando se deseja comparar as médias dos grupos em relação à média de uma população controle ou quando se deseja comparações duas a duas e contrastes lineares para as médias dos grupos. Desde que esse procedimento evidenciou ser muito conservador na prática, um procedimento aproximado baseado nos métodos de Welch aparece como alternativa. Ambas técnicas necessitam somente da tabela de probabilidades da distribuição t-student para sua aplicação e são muito simples na prática".

Os dois procedimentos serão descritos na seqüência deste trabalho, assim como o procedimento proposto por Hochberg.

Considerando o modelo (3.1), trataremos dos problemas de comparações múltiplas através do cálculo de intervalos de confiança para os seguintes casos descritos na introdução:

P1) Intervalo de confiança para todas as diferenças $\mu_i - \mu_k$ ($1 \leq i \leq t, 1 \leq k \leq t, i \neq k$)

onde consideramos μ_k como a média de uma população controle.

P2) Intervalo de confiança para todas as diferenças duas a duas $\mu_i - \mu_k$

($1 \leq i \leq t, 1 \leq k \leq t, i \neq k$) e todos os contrastes lineares $\sum_{i=1}^t c_i \mu_i$ onde c_1, c_2, \dots, c_k é um

conjunto arbitrário de constantes satisfazendo a condição $\sum_{i=1}^k c_i = 0$.

P3) Intervalo de confiança para todas combinações lineares $\sum_{i=1}^t a_i \mu_i$ onde

a_1, a_2, \dots, a_k é um conjunto de constantes reais.

4.1. Método de Banerjee

Os intervalos de confiança proposto por Banerjee são desenvolvidos com base em alguns lemas que serão apresentados no Apêndice.

4.1.1. Comparações Múltiplas com Controle (P1)

Considerando o modelo (3.1) e supondo que μ_k é a média da população controle, temos que os coeficientes de confiança de cada uma das famílias de intervalos de confiança é no mínimo $1 - \alpha$ se $\beta \leq 1 - (1 - \alpha)^{1/(t-1)}$.

(i) intervalo de confiança unilateral superior: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$

$$\mu_i - \mu_k \leq \bar{y}_i - \bar{y}_k + \left(\frac{t_{v_i, \beta}^2 S_i^2}{n_i} + \frac{t_{v_k, \beta}^2 S_k^2}{n_k} \right)^{1/2} \quad (4.1)$$

(ii) intervalo de confiança unilateral inferior: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$

$$\mu_i - \mu_k \geq \bar{y}_i - \bar{y}_k + \left(\frac{t_{v_i, \beta}^2 S_i^2}{n_i} + \frac{t_{v_k, \beta}^2 S_k^2}{n_k} \right)^{1/2} \quad (4.2)$$

(iii) intervalo de confiança bilateral: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$

$$\mu_i - \mu_k \in \left[\bar{y}_i - \bar{y}_k \pm \left(\frac{t_{v_i, \beta/2}^2 S_i^2}{n_i} + \frac{t_{v_k, \beta/2}^2 S_k^2}{n_k} \right)^{1/2} \right] \quad (4.3)$$

Nas equações de [4.1] a [4.3] temos que:

n_i é o número de observações no tratamento i ,

$$\bar{y}_i = \sum_{j=1}^{n_i} Y_{ij} / n_i,$$

$$S_i^2 = \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 / (n_i - 1), \text{ e}$$

$$v_i = n_i - 1.$$

Esta mesma notação será utilizada nos demais métodos.

4.1.2. Comparação de médias duas a duas e contrastes ortogonais

(P2)

i) Intervalo de confiança para comparação de médias duas a duas

Para comparações de médias duas a duas, considerando todos os grupos, Banerjee propõe o seguinte intervalo de confiança para as diferenças

$$\mu_i - \mu_k \in \left[\bar{y}_i - \bar{y}_k \pm \left(\frac{t^2_{v_i, \beta/2} S_i^2}{n_i} + \frac{t^2_{v_k, \beta/2} S_k^2}{n_k} \right)^{1/2} \right] \quad (1 \leq i \leq t, 1 \leq k \leq t, i \neq k) \quad (4.4)$$

e neste caso também temos um nível de confiança de pelo menos $1 - \alpha$ se

$$\beta \leq 1 - (1 - \alpha)^{2/t(t-1)}.$$

ii) Intervalo de confiança para contrastes ortogonais

Os intervalos de confiança para contrastes ortogonais são dados por:

$$\sum_{i=1}^t c_i \mu_i \in \left[\sum_{i=1}^t c_i \bar{y}_i \pm \left\{ \frac{2 \sum_{i \in \mathcal{D}_c} \sum_{j \in \eta_c} c_i (-c_j) \left(\frac{t^2_{v_i, \beta/2} S_i^2}{n_i} + \frac{t^2_{v_k, \beta/2} S_k^2}{n_k} \right)^{1/2}}{\sum_{i=1}^t |c_i|} \right\} \right] \quad (4.5)$$

para todos os contrastes $(c_1 \dots c_t)$; $\sum_{i=1}^t c_i = 0$ e tem coeficiente de confiança de no

mínimo $1 - \alpha$ se $\beta \leq 1 - (1 - \alpha)^{2/t(t-1)}$.

4.2 Comparações múltiplas baseadas no método de Welch

Outro procedimento para comparação de médias quando não é satisfeita a condição de homogeneidade de variância é o proposto por Welch. Esta técnica é análoga a de Banerjee, com algumas modificações

4.2.1. Comparações Múltiplas com Controle (P1)

(i) Intervalo de confiança unilateral superior: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$,

$$\mu_i - \mu_k \leq \bar{y}_i - \bar{y}_k + t_{\nu_{ik}, \beta} \left[\left(\frac{S_i^2}{n_i} + \frac{S_k^2}{n_k} \right) \right]^{1/2} \quad (4.6)$$

(ii) Intervalo de confiança unilateral inferior: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$

$$\mu_i - \mu_k \geq \bar{y}_i - \bar{y}_k + t_{\nu_{ik}, \beta} \left[\left(\frac{S_i^2}{n_i} + \frac{S_k^2}{n_k} \right) \right]^{1/2} \quad (4.7)$$

(iii) Intervalo de confiança bilateral: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$

$$\mu_i - \mu_k \in \left[\bar{y}_i - \bar{y}_k \pm t_{\hat{v}_{ik}, \beta/2} \left(\frac{S_i^2}{n_i} + \frac{S_k^2}{n_k} \right)^{1/2} \right] \quad (4.8)$$

4.2.2. Comparação de médias duas a duas e contrastes ortogonais

(P2)

Vamos buscar um intervalo de confiança para a diferença das médias.

O Intervalo de confiança bilateral: para $(1 \leq i \leq t, 1 \leq k \leq t, i \neq k)$ pela proposta de Welch é dada através de (4.9) e o coeficiente de confiança é aproximadamente $1 - \alpha$.

$$\mu_i - \mu_k \in \left[\bar{y}_i - \bar{y}_k \pm t_{\hat{v}_{ik}, \alpha/2} \left(\frac{S_i^2}{n_i} + \frac{S_k^2}{n_k} \right)^{1/2} \right] \quad (4.9)$$

Nas equações de (4.6) a (4.9) \hat{v}_{ik} é dado por:

$$\hat{v}_{ik} = \frac{\left(\frac{S_i^2}{n_i} + \frac{S_k^2}{n_k} \right)^2}{\left[\frac{S_i^4}{n_i^2(n_i - 1)} + \frac{S_k^4}{n_k^2(n_k - 1)} \right]} \quad (4.10)$$

onde \hat{V}_{ik} é aleatório (não necessariamente um inteiro), e o valor $t_{\hat{V}_{ik}, \alpha/2}$ pode ser interpolado na tabela t. Então, estes intervalos de confiança tem a mesma concepção que o de Banerjee, exceto que o proposto por Welch é utilizado para a aproximação de probabilidades individuais.

4.3. Procedimento baseado no método de Hochberg

4.3.1. Comparação de médias duas a duas e contrastes não ortogonais (P3)

Este procedimento é uma extensão do método de Tukey para condições onde a homogeneidade da variância não é verificada. Aqui, apresentaremos somente o caso para tamanhos de amostras iguais.

Os intervalos de confiança para os contrastes não-ortogonais são obtidos através do Teorema 1:

Teorema 1: O coeficiente de confiança para a família de intervalos de confiança

$$\left\{ \sum_{i=1}^t c_i \mu_i \in \left[\sum_{i=1}^t c_i \bar{y}_i \pm g_{tv}(\alpha) M(b_1 \dots b_t) \right] \right\}, \text{ para todo } l \in R^k \quad (4.3.1)$$

é exatamente $1 - \alpha$.

Onde:

t é número de grupos a serem comparados,

$g_{tv}(\alpha)$ obtido através da Tabela 1,

$$v = n - 1,$$

$$b_i = \frac{c_i S_i}{n^{1/2}}, \text{ e}$$

$$M(c_1 \dots c_t) = \max \left(\sum_{i \in \rho(l)} c_i - \sum_{i \in \pi(l)} c_i \right) \quad (4.3.1)$$

sendo,

$\rho(c)$ o conjunto de índices i para os quais $c_i > 0$,

$\pi(c)$ o conjunto de índices i para os quais $c_i < 0$.

Tabela 1 - Valores de $g_{t, n_i - 1}(\alpha)$

α	n_i	T								
		2	3	4	5	6	7	8	9	10
0,10	5	2,91	3,73	4,26	4,66	4,99	5,27	5,50	5,72	5,91
	8	2,66	3,37	3,81	4,13	4,39	4,61	4,79	4,95	5,09
	10	2,58	3,26	3,68	3,98	4,22	4,42	4,59	4,74	4,87
	15	2,49	3,13	3,52	3,80	4,02	4,19	4,35	4,48	4,59
	20	2,45	3,07	3,44	3,71	3,92	4,09	4,23	4,36	4,47
	30	2,40	3,01	3,37	3,63	3,83	3,99	4,13	4,24	4,35
	40	2,38	2,98	3,34	3,59	3,79	3,94	4,08	4,19	4,29
0,50	60	2,36	2,95	3,30	3,55	3,74	3,90	4,03	4,14	4,24
	5	3,63	4,49	5,05	5,47	5,82	6,12	6,37	6,60	6,80
	8	3,24	3,95	4,40	4,73	4,99	5,21	5,40	5,56	5,71
	10	3,13	3,80	4,22	4,52	4,76	4,96	5,13	5,28	5,41
	15	3,00	3,62	4,00	4,27	4,49	4,66	4,81	4,94	5,05
	20	2,94	3,54	3,90	4,16	4,36	4,53	4,67	4,79	4,89
	30	2,88	3,46	3,81	4,05	4,25	4,40	4,53	4,65	4,74
40	2,85	3,42	3,76	4,00	4,19	4,34	4,47	4,58	4,67	
60	2,82	3,39	3,72	3,95	4,13	4,28	4,41	4,51	4,60	

Referência: Hochberg [7]

4.4. Procedimento baseado no método de Brown-Forsythe

4.4.1. Comparação de médias duas a duas e contrastes ortogonais

(P2)

Este procedimento é uma extensão para o método de Scheffé convencionalmente utilizado quando a condição de homogeneidade de variância não é satisfeita. O nível de confiança obtido para todos os contrastes possíveis é de $1 - \alpha$.

Intervalos de confiança para contrastes ortogonais são dados por:

$$\sum_{i=1}^t c_i \mu_i \in \left[\sum_{i=1}^t c_i \bar{y}_i \pm \sqrt{(t-1)F_{(1-\alpha);t-1;f} \sum_{i=1}^t \frac{d_i s_i^2}{n_i}} \right] \quad (4.4.1)$$

onde

$$\frac{1}{f} = \sum_{i=1}^t \frac{d_i^2}{n_i - 1} \text{ e}$$

$$d_i = \frac{\left(1 - \frac{n_i}{n}\right) s_i^2}{\sum_{i=1}^t \left(1 - \frac{n_i}{n}\right) s_i^2}.$$

5. ANOVA COM DOIS FATORES FIXOS

Quando o experimento em estudo envolve dois fatores, é de interesse determinar o efeito da ação conjunta desses fatores sobre as observações.

Neste capítulo, daremos um opção para a ANOVA quando o modelo envolve dois fatores e não é satisfeita a condição de homogeneidade de variância.

5.1. Modelo Matemático

Admitindo-se um modelo em delineamento completamente casualizado, Riboldi [22]

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk} \quad (5.1)$$

$i = 1, 2, \dots, a$; $j = 1, 2, \dots, b$; $k = 1, 2, \dots, r$, onde:

μ : média geral

α_i : efeito do i -ésimo nível do fator A, definido como $\alpha_i = \mu_{i.} - \mu$

β_j : efeito do j -ésimo nível do fator B, definido como $\beta_j = \mu_{.j} - \mu$

$\alpha\beta_{ij}$: efeito da interação entre o i -ésimo nível de A e o j -ésimo nível de B e é

definido como $\alpha\beta_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu$

ε_{ij} : variável aleatória normalmente distribuída com média zero e variância σ^2 .

Temos que as restrições a seguir, são devidas as definições dos parâmetros do modelo:

$$\sum_{i=1}^a \alpha_i = 0; \quad \sum_{j=1}^b \beta_j = 0$$

$$\left. \begin{array}{l} \sum_{i=1}^a \alpha\beta_{ij} = 0, \forall_j \\ \sum_{j=1}^b \alpha\beta_{ij} = 0, \forall_i \end{array} \right\} \Rightarrow \sum_{i=1}^a \sum_{j=1}^b \alpha\beta_{ij} = 0.$$

E o como suposição do modelo temos $Y_{ijk} \sim N(\mu + \alpha_i + \beta_j + \alpha\beta_{ij}; \sigma^2)$ e independentes e $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

5.2 Teste de efeito de tratamentos

A hipótese nula é dada por:

- i) $H_0 : \alpha_i = 0, \forall_i$, onde testa-se a hipótese de não existir efeito do fator A,
- ii) $H_0 : \beta_j = 0, \forall_j$, onde testa-se a hipótese de não existir efeito do fator B,
- iii) $H_0 : \alpha\beta_{ij} = 0, \forall_i, \forall_j$, onde testa-se a hipótese de não existir efeito da interação.

A seguir, teremos uma tabela de análise de variância para o modelo (5.1)

Causas da Variação	G.L.	Soma de Quadrados	Quadrado Médio	F
Tratamento A	a-1	SQA	QMA	QMA / QME
Tratamento B	b-1	SQB	QMB	QMB / QME
Interação AxB	(a-1)(b-1)	SQAxB	QMAxB	QMAxB / QME
Erro	ab(r-1)	SQE	QME	
Total	abr-1	SQtotal		

Onde:

$$SQA = \sum_{i=1}^a \frac{y_{i..}^2}{rb} - \frac{y_{...}^2}{abr},$$

$$SQB = \sum_{j=1}^b \frac{y_{.j.}^2}{ar} - \frac{y_{...}^2}{abr},$$

$$SQAB = \sum_{i=1}^a \sum_{j=1}^b \frac{y_{ij.}^2}{r} - \frac{y_{...}^2}{abr} - \left[\sum_{i=1}^a \frac{y_{i..}^2}{rb} - \frac{y_{...}^2}{abr} \right] - \left[\sum_{j=1}^b \frac{y_{.j.}^2}{ar} - \frac{y_{...}^2}{abr} \right],$$

$$SQE = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r \left(y_{ijk} - \bar{y}_{ij.} \right)^2,$$

Os quadrados médios são dados por:

$$QMA = SQA / (a - 1),$$

$$QMB = SQB / (b - 1),$$

$$QMA \times B = SQA \times B / (a - 1)(b - 1) \text{ e}$$

$$QME = SQE / ab(r - 1).$$

Para proceder o teste F, quando não podemos assumir a condição de homogeneidade de variância, Brown-Forsythe [4] sugere um teste F com a mesma concepção de um teste F convencional (assumindo variâncias homogêneas) fazendo um correção nos graus de liberdade do erro. Na ANOVA convencional, para testar o efeito do fator A, teríamos uma F com (a - 1) e ab(r-1) graus de liberdade, a sugestão de Brown-Forsythe é fazer uma correção nos graus de liberdade do denominador. Então, teríamos uma estatística com distribuição F aproximada com (a-1) e f graus de liberdade, onde f é dado por:

$$\frac{1}{f} = \sum_{i=1}^a \sum_{j=1}^b \frac{c_{ij}^2}{(r-1)},$$

onde:

$$c_{ij} = \frac{S_{ij}^2}{\sum_{i=1}^a \sum_{j=1}^b S_{ij}^2}$$

Analogamente, para testar o efeito do tratamento B (ou da interação) teríamos uma estatística com distribuição F aproximada com (b - 1) e f (ou (a-1)(b-1) e f) graus de liberdade.

6. APLICAÇÃO

Serão apresentadas três simulações para aplicação numérica para as técnicas sugeridas. Os exemplos 1, 2 e 3 são simulações de valores amostrais teoricamente extraídos de populações normais cujos parâmetros foram intencionalmente estipulados. Em cada situação temos então cinco grupos amostrais com seis repetições cada, aos quais arbitraremos médias e variâncias para posterior discussão. O objetivo é verificar diferenças entre as técnicas de comparações múltiplas e obter algumas conclusões de suas aplicações.

Para gerar estes grupos amostrais, foi utilizado o aplicativo Excel do Windows [20]. Através deste programa foi criada uma rotina que nos fornece os valores e probabilidades da estatística F.

Os intervalos de confiança para cada conjunto de dados não podem ser obtidos diretamente, pois precisam de valores específicos da tabela t. Neste caso, utilizamos o programa Statistica [21] para obtenção dos valores t tabelados.

Para testar a hipótese de homogeneidade de variância entre os tratamentos, usaremos o teste de Levene, o qual será obtido através do pacote estatístico SPSS/8.0.

Em cada uma das simulações somente foram construídos intervalos de confiança para comparar médias duas as duas. Os procedimentos utilizados são os sugeridos por Banerjee, Welch, Hochberg. O teste de Tukey será utilizado para efeito de comparação. Para esses resultados, as diferenças entre médias dos

tratamentos cujos os intervalos construídos não contiverem o valor zero, indicam que não há evidências amostrais que nos levem a rejeitar a hipótese de que essas médias sejam iguais.

Simulação 1.

Neste caso, são simulados cinco grupos de observações com seis repetições cada. Foram geradas amostras oriundas de populações normalmente distribuídas com $\underline{\mu} = (20 ; 23 ; 25 ; 24 ; 27)$ e $\underline{\sigma} = (2,00 ; 2,65 ; 5,48 ; 3,16 ; 1,23)$. Foi simulada uma situação em que os grupos apresentam médias diferentes e variâncias levemente heterogêneas, isto é, aceitamos a igualdade entre variância dos grupos, mas com valor próximo a 0,05 de nível de significância.

Para verificar a condição de homogeneidade dos grupos, foi realizado o teste de Levene que apontou um $p = 0,072$. Podemos aceitar a condição de homogeneidade, mas devemos perceber o valor de p é muito próximo ao 0,05 da significância.

Realizamos o teste ANOVA para verificar se pelo menos uma média se diferencia das demais médias dos grupos. Foi, então, realizado um teste F (anova convencional), F_W (sugerida por Welch) e F_{BF} (sugerida por Brown-Forsythe). Neste caso, somente o teste F modificado sugerido por Welch evidenciou que existe diferença entre as médias dos grupos, conforme Figura 1.

Após verificarmos que existe diferenças entre as médias, precisamos determinar quais médias dos grupos se diferenciam entre si.

Para isso, utilizamos o procedimento de criar intervalos de confiança para verificar se as médias apresentam diferença significativa. Foram obtidos intervalos sugeridos por Banerjee, Welch, Hochberg e Tukey (Figura 2). Podemos observar que os intervalos sugeridos por Banerjee, Welch e Hochberg captaram diferença significativa entre as médias dos grupos 1 e 5, que são a menor e maior média respectivamente. O intervalo de Welch ainda captou diferença significativa entre as médias dos grupo 1 e 4, onde o grupo 4 é a segunda maior média. Foi realizado, também, um intervalo de confiança pelo método de Tukey (embora esse não seja indicado nesta situação) e este não captou diferença entre a média de nenhum dos grupos.

The screenshot shows a Microsoft Excel spreadsheet titled 'SIMULAÇÃO.xls'. The main data is organized into two tables. The first table, 'Valores observados', shows individual data points for five groups across 10 rows. The second table, 'Teste F para ANOVA segundo a sugestão dos autores', summarizes the statistical test results.

Valores observados						
	Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	
1						
2						
3						
4	18,28	24,74	19,82	27,37	25,05	
5	21,06	22,50	12,74	22,39	26,29	
6	21,15	18,39	23,21	23,04	26,66	
7	17,56	22,04	34,78	26,62	27,52	
8	23,53	30,25	28,87	23,52	27,45	
9	21,40	23,46	24,00	26,55	29,78	
10						
11	média amostral	20,50	23,56	23,90	24,92	27,13
12	variância amostral	4,86	15,26	56,95	4,69	2,51
13						
Teste F para ANOVA segundo a sugestão dos autores						
	Fc	Fw	Fbf			
15	f calculado	2,05	7,83	2,31		
16	gl numerador	4	4	4		
17	gl denominador	25	12	10		
18						
19						
20	probabilidade(p)	0,12	0,00	0,13		
21						

Figura 1

Microsoft Excel - SIMULAÇÃO.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 N I S

U17 =

Intervalos de confiança para diferenças entre as médias com confiança de 95%

diferenças	Banerjee		Welch		Hochberg		Tukey	
	LI	LS	LI	LS	LI	LS	LI	LS
M1-M2	-13,29	7,15	-8,10	1,97	-11,12	4,99	-10,01	3,87
M1-M3	-21,32	14,50	-12,94	6,13	-18,97	12,15	-10,35	3,53
M1-M4	-11,46	2,62	-7,74	-1,10 *	-8,96	0,13	-11,36	2,52
M1-M5	-12,81	-0,45 *	-9,61	-3,65 *	-11,17	-2,08 *	-13,57	0,31
M2-M3	-19,70	19,02	-9,88	9,20	-15,90	15,22	-7,28	6,60
M2-M4	-11,53	8,82	-6,37	3,66	-9,41	6,70	-8,29	5,59
M2-M5	-13,16	6,04	-8,45	1,33	-11,62	4,49	-10,50	3,38
M3-M4	-18,90	16,87	-10,53	8,51	-16,57	14,55	-7,95	5,93
M3-M5	-20,79	14,34	-13,17	6,73	-18,78	12,34	-10,16	3,72
M4-M5	-8,32	3,90	-5,16	0,74	-6,67	2,25	-9,15	4,73

* diferença significativa

Comparações Múltiplas / ANOVA / testes / Apresentação

Pronto NUM

Iniciar Microsoft Excel - SIM... Microsoft Word - ensaio8.d... 20:07

Figura 2

Simulação 2.

Foram geradas amostras oriundas de populações normalmente distribuídas com $\underline{\mu} = (20; 15; 25; 24; 27)$ e $\underline{\sigma} = (6,33; 0,78; 5,47; 3,16; 9,49)$. Foi simulada uma situação em que os grupos apresentam médias diferentes e variâncias acentuadamente heterogêneas.

Para verificar a condição de homogeneidade dos grupos, foi realizado o teste de Levene que apontou um $p < 0,001$, portanto, rejeitamos a condição de homogeneidade entre as variâncias.

Realizado testes de análise de variância para este conjunto de observações, percebemos que os testes F, F_W e F_{BF} evidenciaram diferenças entre médias para os grupos (Figura 3).

Para os testes de comparações múltiplas, pelos procedimentos de Banerjee Welch e Hochberg verificou-se diferença entre as médias dos grupos 1 e 2 e os grupos 2 e 4. O método de Welch ainda captou diferença entre os grupos 1 e 4. Os intervalos gerados pelo método de Tukey evidenciam diferença entre a média do grupo 2 em relação ao grupo 5, que são os que apresentam a maior diferença entre médias e possuem a menor e maior variâncias respectivamente (Figura 4).

A seguir mostraremos um resumo dos resultados obtidos na Simulação 2.

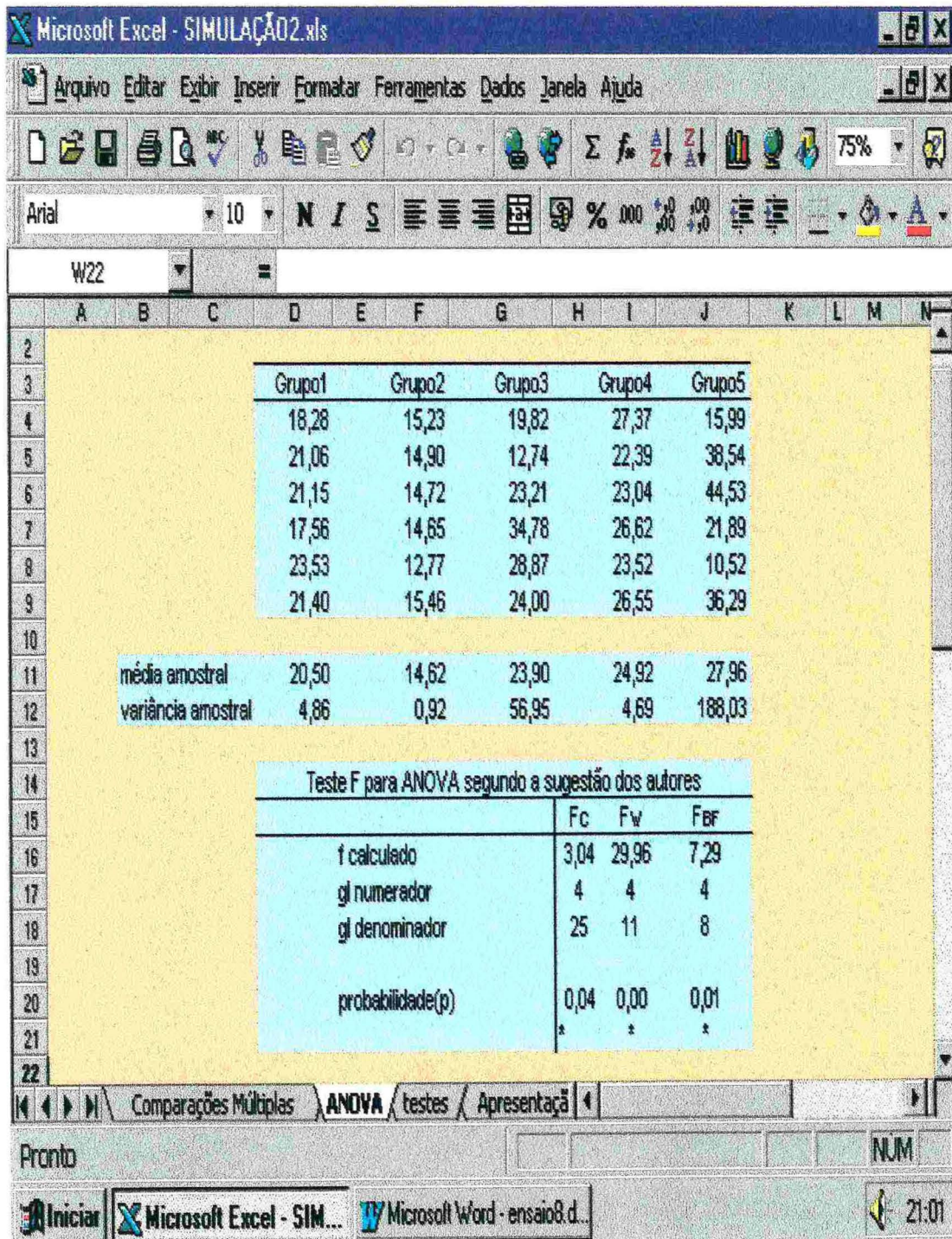


Figura 3

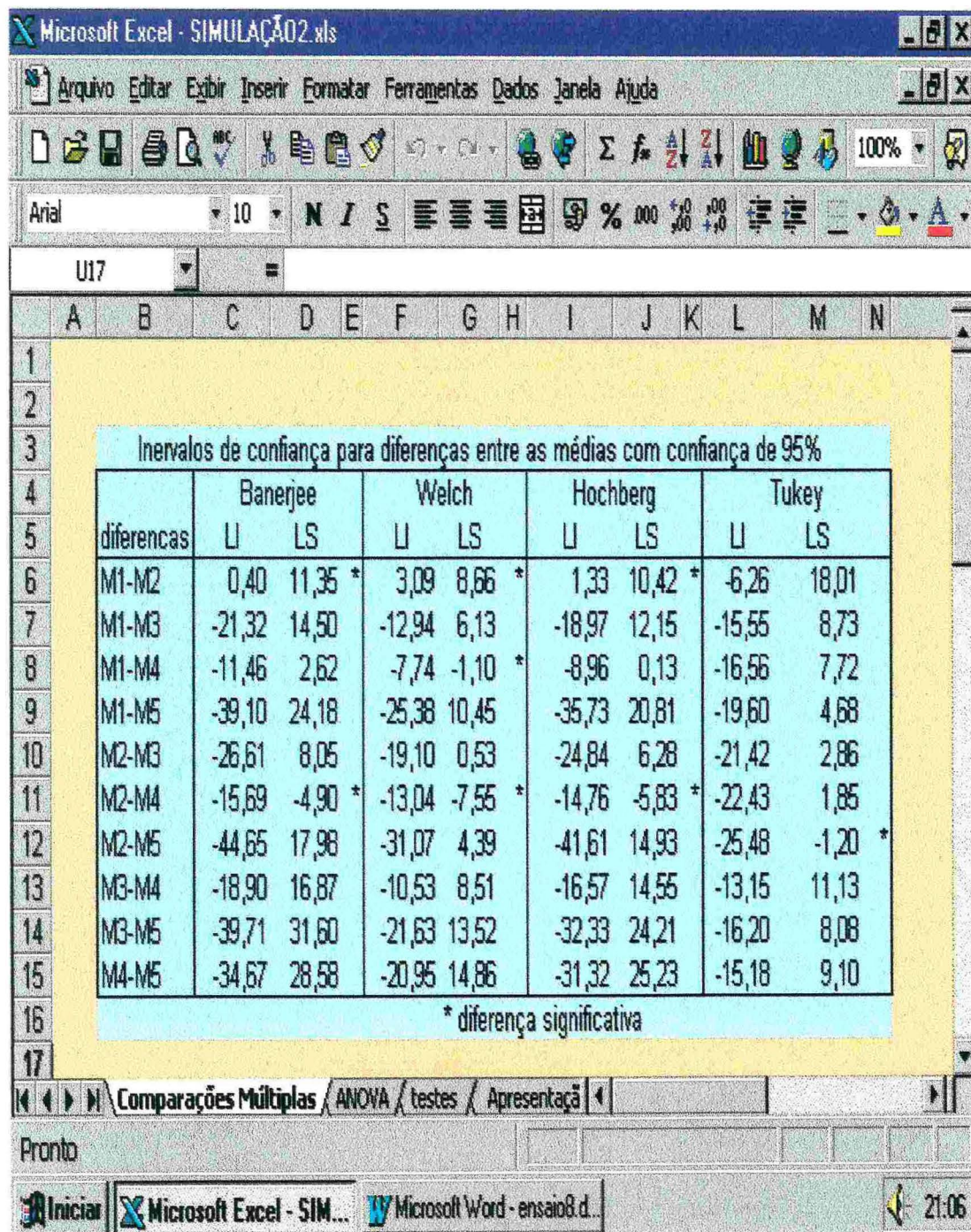


Figura 4

Simulação 3.

Neste caso, foram geradas amostras oriundas de populações normalmente distribuídas com $\underline{\mu} = (20 ; 20 ; 20 ; 20 ; 20)$ e $\underline{\sigma} = (6,33 ; 0,78 ; 5,48 ; 3,16 ; 9,49)$. Foi simulada uma situação em que os grupos apresentam médias iguais e variâncias acentuadamente heterogêneas.

Para verificar a condição de homogeneidade dos grupos, foi realizado o teste de Levene que apontou um $p < 0,001$, logo rejeitamos a condição de homogeneidade entre as variâncias (Figura 5).

Realizado testes de análise de variância para este conjunto de observações, percebemos que nenhum dos testes F, F_W e F_{BF} evidenciaram diferenças entre médias para os grupos. Apesar da alta heterogeneidade entre as variâncias dos grupos, quando as médias são iguais, os três testes são robustos para não rejeitar a hipótese de igualdade. Conseqüentemente, teste de comparações múltiplas, neste caso, não se faz necessário.

Mesmo neste caso, as comparações múltiplas foram feitas e não foi encontrada nenhuma diferença significativa, sendo todos os intervalos robustos para não evidenciar diferença entre as médias.

A seguir mostraremos um resumo dos resultados obtidos na simulação 3.

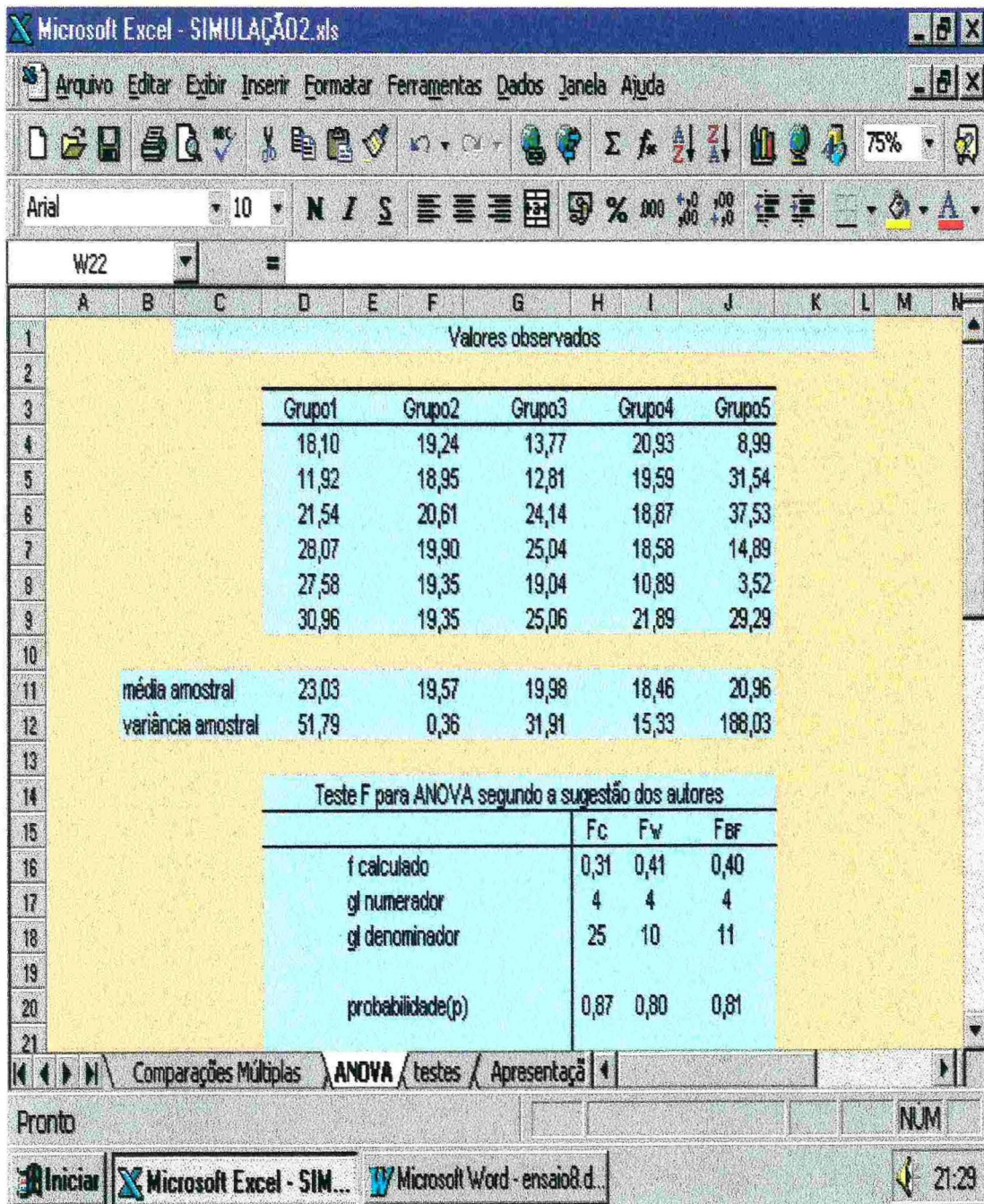


Figura 5

C'omplementando, temos os intervalos

Microsoft Excel - SIMULAÇÃO2.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 N I S % 000 000 000 100%

U17 =

Intervalos de confiança para diferenças entre as médias com confiança de 95%									
diferencas	Banerjee		Welch		Hochberg		Tukey		
	LI	LS	LI	LS	LI	LS	LI	LS	
M1-M2	-12,99	19,91	-5,85	12,78	-11,37	18,30	-9,41	16,34	
M1-M3	-17,79	23,89	-7,00	13,10	-11,78	17,89	-9,82	15,93	
M1-M4	-14,09	23,23	-4,63	13,77	-10,27	19,41	-8,31	17,45	
M1-M5	-33,21	37,35	-15,32	19,45	-26,20	30,34	-10,81	14,94	
M2-M3	-13,35	12,53	-7,74	6,92	-12,06	11,24	-13,29	12,47	
M2-M4	-7,91	10,13	-4,00	6,22	-6,96	9,18	-11,77	13,98	
M2-M5	-32,66	29,87	-19,10	16,31	-29,66	26,88	-14,27	11,48	
M3-M4	-14,14	17,18	-6,03	9,07	-10,13	13,17	-11,36	14,39	
M3-M5	-34,77	32,80	-18,18	16,21	-29,25	27,29	-13,86	11,90	
M4-M5	-34,99	29,98	-19,79	14,74	-30,77	25,77	-15,38	10,38	

* diferença significativa

Comparações Múltiplas / ANOVA / testes / Apresentação

Pronto NUM

Iniciar Microsoft Excel - SIM... Microsoft Word - ensaio8.d... 21:32

Figura 6

7. ROTINA NO EXCEL

Para o teste F e comparações múltiplas, quando temos um fator com cinco tratamentos e seis repetições, foi implementada uma rotina de cálculos no Excel.

A partir da planilha Apresentação temos:

Microsoft Excel - SIMULAÇÃO3.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 12

V27

Valores observados em experimento com um fator, cinco tratamentos com seis repetições

Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
18,28	24,74	19,82	27,37	25,05
21,06	22,5	12,74	22,39	26,29
21,15	18,39	23,21	23,04	26,66
17,56	22,04	34,78	26,62	27,52
23,53	30,25	28,87	23,52	27,45
21,4	23,46	24	26,55	29,78

Os valores observados no experimento devem ser digitados na seqüência.

Intervalos de Welch: entre com os valores de t tabelado

	graus liberdade	alfa/2	leitura valor t
M1-M2	5	0,025	2,75
M1-M3	9	0,025	2,97
M1-M4	8	0,025	2,63
M1-M5	8	0,025	2,69
M2-M3	5	0,025	2,75
M2-M4	5	0,025	2,75
M2-M5	5	0,025	2,84
M3-M4	9	0,025	2,97
M3-M5	7	0,025	3,16
M4-M5	6	0,025	2,69

Os valores de t tabelado devem se digitados de acordo com $\alpha/2$ e os graus de liberdade indicados

Onde M_i é a média do tratamento i

Pronto

Microsoft Excel - SIM... Microsoft Word - ensaio8.d... 21:42

Figura 7

Os cálculos serão desenvolvidos, e os resultados são mostrados através das planilhas ANOVA e Comparações Múltiplas, onde temos:

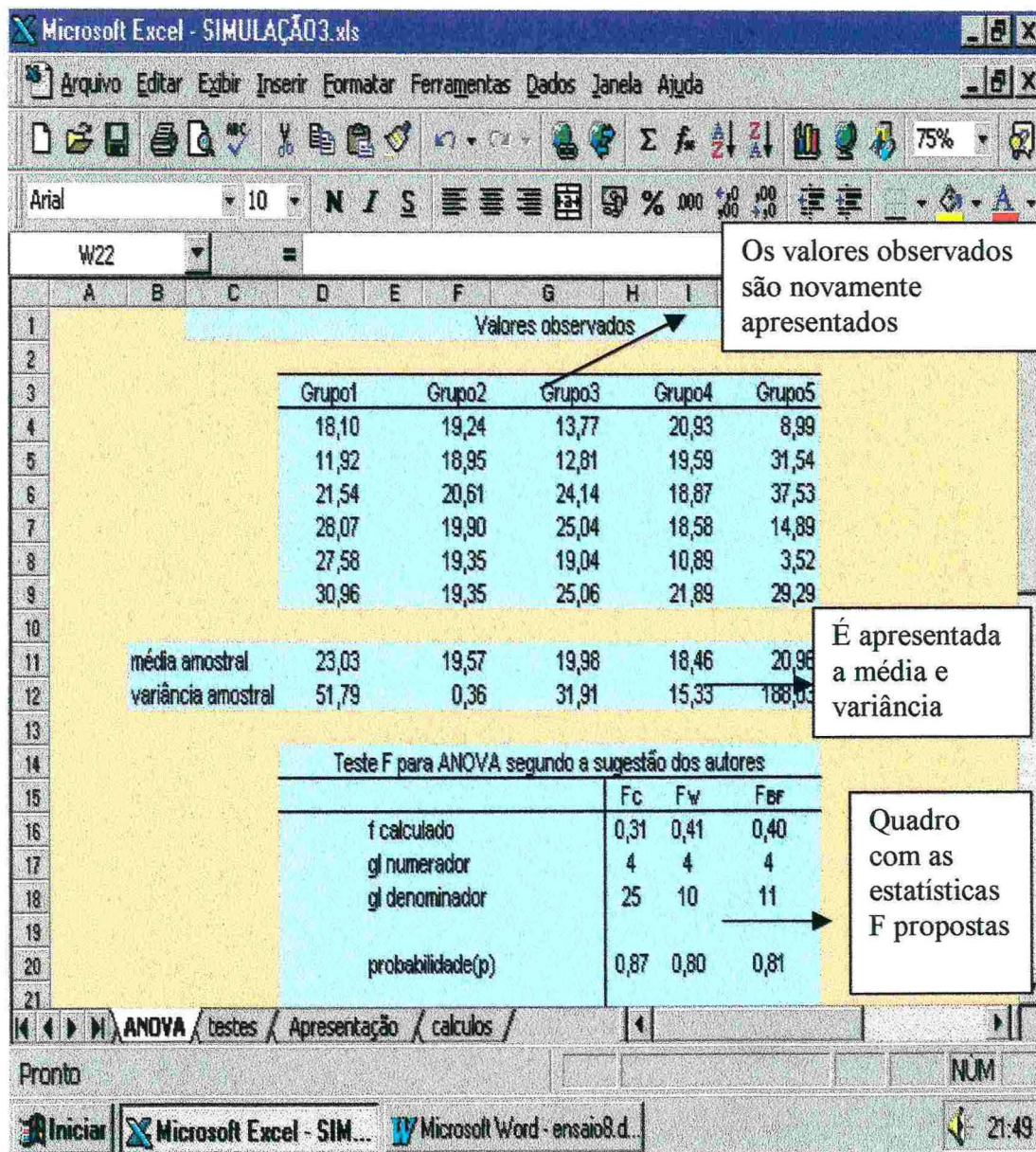


Figura 8

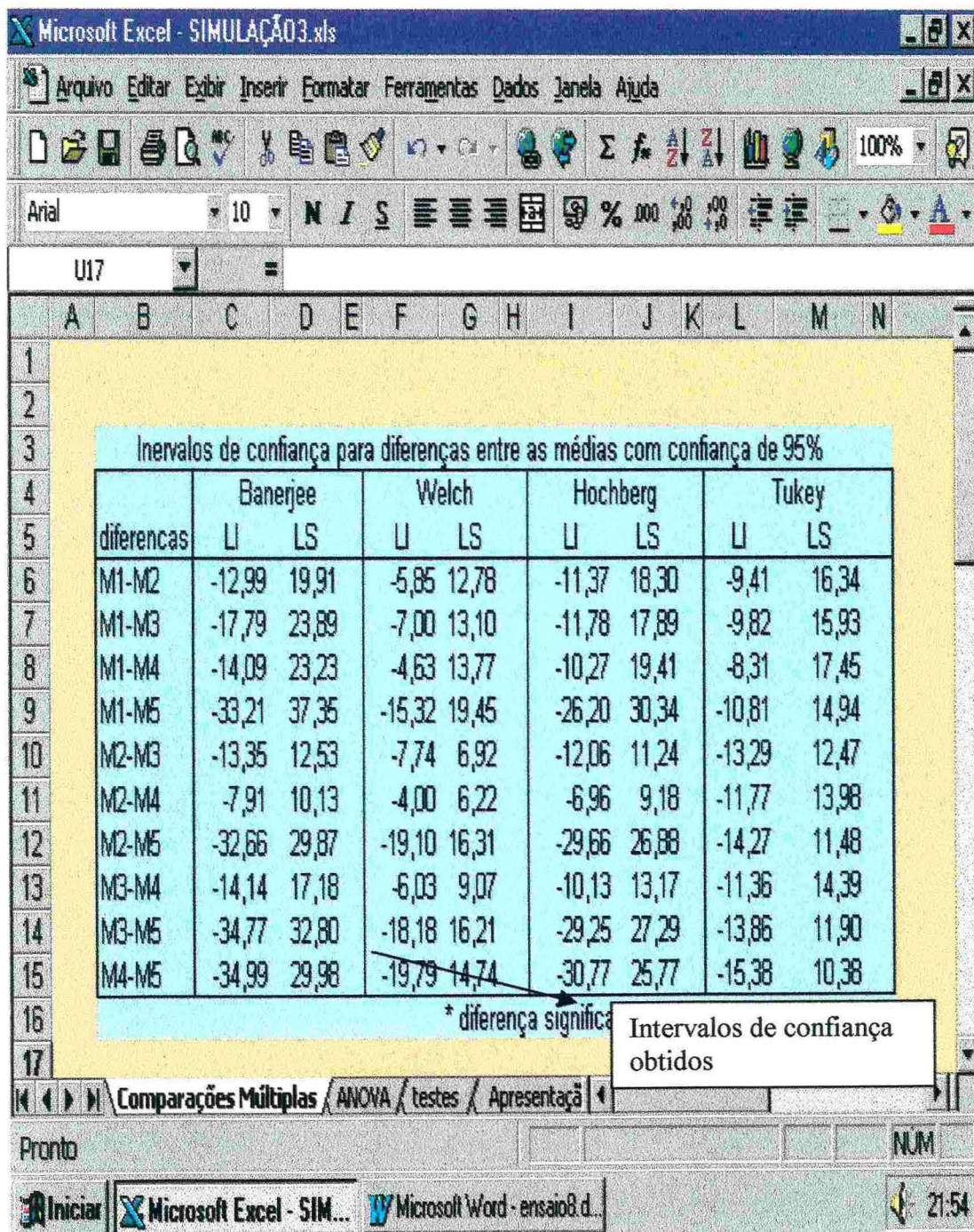


Figura 9

8. CONCLUSÃO

Neste trabalho procuramos abordar um problema muito comum que é a heterogeneidade entre as variâncias dos grupos envolvidos para comparação de médias. Procuramos enfatizar que, especificamente num modelo de um único fator fixo, quando ocorre a não homogeneidade entre as variâncias caímos num caso clássico conhecido como problema de Behrens-Fischer, que é a situação onde temos duas populações com variâncias diferentes. Proposto este problema, foram sugeridas algumas alternativas para contorná-lo.

Existem técnicas eficazes para contornar a heterogeneidade das variâncias, como por exemplo a transformação dos dados observados para uma outra escala que reduza a variabilidade. Outra alternativa seria recorrer a análises não paramétricas, como por exemplo testes de Kruskal-Wallis, Friedman, entre outros.

Com o objetivo de chegar a resultados robustos podemos verificar a normalidade e a heterogeneidade das variâncias. Caso forem satisfeitas estas condições, estamos em condições de continuar a ANOVA. A propriedade da normalidade parece ser o problema de menor relevância, já que ANOVA assume normalidade das médias para grande amostras, e o Teorema Central do Limite dá suporte para o fato de que mesmo populações não normalmente distribuídas têm médias assintoticamente normais. Caso a ANOVA não seja significativa, não temos maiores preocupações se as variâncias são desiguais, já que isso aumentaria o erro

tipo I (rejeitar a hipótese nula quando ela é verdadeira, evidenciando diferenças onde não existe) mas no caso de encontrar diferenças significativas e existir grupos com variâncias heterogêneas, devemos levar em consideração alguns aspectos:

a) ANOVA é robusta mesmo com a violação da homocedasticidade se tivermos dados balanceados e um tamanho grande da amostra.

b) heterogeneidade de variância torna-se um problema quando temos amostras pequenas e particularmente um grupo apresenta variância muito maior em relação aos outros grupos. O problema não é tão relevante se existir heterogeneidade e um dos grupos apresentar variância muito pequena. [10].

Neste trabalho foram abordados testes alternativos para análise de variâncias e posteriores testes de comparações múltiplas, através dos quais podemos chegar a conclusão de que o teste F tradicionalmente utilizado chega aos mesmos resultados de aceitação ou não, do que os testes sugeridos por Welch e por Brown-Fosythe, quando não é assumida a condição de homogeneidade de variância. Através da literatura pesquisada e das simulações efetuadas, percebemos que os referidos testes tem robustez para não evidenciar diferenças entre médias quando estas são muito próximas, mesmo quando existem variâncias heterogêneas.

Foram abordados os testes de comparações múltiplas sugeridos por Welch, Banerjee e Hochberg. O método de Banerjee é notoriamente o mais conservador. Os intervalos propostos por Hochberg tem maior amplitude que os propostos pelos outros métodos, e são exatos.

Recentemente, programas computacionais tem implementado em suas rotinas testes aplicáveis quando a variância dos tratamentos é heterogênea.

O programa SPSS/8.0 traz, no procedimento GLM (general linear model), testes paramétricos de comparações múltiplas. O teste Tamhane's T2 é um teste conservador, baseado no teste t para comparações de médias duas a duas. Dunnett's T3 é baseado no módulo máximo studentizado, o teste de Games-Howell não é muito rigoroso, e o teste Dunnett's C é baseado numa escala studentizada. Todos os testes mencionados, são úteis para comparação de médias duas a duas, sem nenhuma restrição.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] MILLER, R. G. (1966). **Simultaneous Statistical Inference in the Analysis of Variance**. Ann. Math. Statist. 22, 600-2.
- [2] O'NEILL R. and G.B. WETHERILL (1971) **The Present State of Multiple Comparasion Methods**. Journal of the Royal Statistical Society, B, vol. 33 pp 218-241.
- [3] BROWNLEE, K.A. **Statistical Theory and Methodology in Science and Engineering**, John Wiley and Sons, Inc, 1960.
- [4] BROWN, M.B. and FORWYTHE, Alan B., **The ANOVA and Multiple Comparasions for Data with Heterogeneous Variances**, Biometrics, 30 (december 1974), 19-24)
- [5] WELCH, B.L. (1983) **The Significance of the Difference between Two Means when the Population Variances are Unequal**, Biometrika 25, 350-62
- [6] TAMHANE, A. C. (1977). **Multiple Comparasions in model I One-Way. ANOVA with Unequal Variances**. Communications in Statistics, A6, 15-32.
- [7] HOCHBERG, Y. (1976). **A Modification of the T-Method of Multiple Comparasions for a One-Way Layout with Unequal Variances**. Journal of the American Statistical Association, 71, 200-203.
- [8] JOHNSON N.L. and LEONE F.C. (1977) **Statistical and Experimental Design in Engineering and the Physical Sciences**. 2nd edition, John Wilwy & Sons, Inc.
- [9] GOMES, F.P. **Curso de Estatística Experimental**. 10a. edição, Livraria Nobel S.A., Piracicaba, São Paulo, 1982.
- [10] <http://www.earthsystems.org/list/ecolog-1/0594.html>
- [11] MILLER, R.G. Jr (1977). **Developments in Multiple Comparasions**.(1966-976. Journal of the American Statistical Association, vol. 72 pp 779-778.
- [12] NELSON, P.R. (1989) **Multiple Comparasion of Means using Simultaneous Confidence Intervals**. Journal Quality, vol 21, nº 4 pp 232-241.

- [13] FEDERER, W.T. **Experimental Design, Theory and Application**. 1st edition, Copyright, 1955, by the Macmillan Company
- [14] RAY, W.S. **An Introduction to Experimental Design**. Department of Psychology Pennsylvania State University, The Macmillan Company, New York, 1960.
- [15] STEIN, C. (1945). **A Two-sample Teste for a Linear Hypothesis whose Power is Independent of the Variance**. Ann. Math. Statist. 16. 243-58.
- [16] Chapman, D. G. (1950). Some Two-sample Tests. Ann. Math. Statist. 21. 601-6.
- [17] GHOSH, b. K. (1975). **A Two-stage procedure for the Behrens-Fisher Problem**. J.Amer. Statist. Assoc. 70. 457-62.
- [18] ESARY, J.D., PROSCHAN, F. e WaLKUP, D. W. (1967). **Association of Randon of Variables with Applications**. Ann. Math. Statist. 38, 1466-74.
- [19] KIMBALL, A. W. (1951). **On Dependent Tests of Significance in the Analysis of Variance**. Ann. Math. Satist. 22, 600-2.
- [20] EXCEL, Copyright©1985-1997 - Microsoft Corporation
- [21] Statistica for Windows, V. 4.3 - Copyright©StatSoft, Inc. 1993.
- [22] RIBOLDI, J. (1995). Experimentos Fatoriais, UFRGS, Instituto de Matemática, **Cadernos de Matemática e Estatística**, Série B, nº 28.
- [23] SATTERHWAITE, F.E. (1941). **Synthesis of Variance**, Psychometrika, 6.

APÊNDICE

1. Alternativa de Welch para o problema de Behrens-Fisher

Apresentaremos aqui uma opção para o método de Welch para contornar o problema de Behrens-Fisher.

Consideremos duas populações normalmente distribuídas com médias μ_1 e μ_2 e variâncias σ_1^2 e σ_2^2 , respectivamente. Quando desejamos testar a igualdade entre as médias de suas respectivas amostras, devemos considerar dois casos:

i) se σ_1^2 e σ_2^2 podem ser iguais, podemos usar

$$U = \frac{\left(\bar{y}_1 - \bar{y}_2 \right)}{\sqrt{\frac{\sum_{j=1}^{n_1} y_{1j}^2 + \sum_{j=1}^{n_2} y_{2j}^2}{(n_1 + n_2 - 2)} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (8.1)$$

para testar distribuição t com $f = (n_1 + n_2 - 2)$ graus de liberdade,

ii) se a razão de σ_1^2 e σ_2^2 é conhecida, um teste parecido a (8.1) pode ser usado, mas se a razão é desconhecida, não haveria um teste exato para ser utilizado. A solução para esse teste de hipótese foi proposta por R. A. Fischer utilizando o conceito de uma distribuição aproximada. Fischer percebeu a equivalência deste teste com o já proposto por W.V. Behrens.

A validade deste teste foi questionada por M. S. Bartlett. Então, uma alternativa foi sugerida,

$$V = \frac{\left(\bar{y}_1 - \bar{y}_2 \right)}{\sqrt{\frac{\sum_{j=1}^{n_1} y_{1j}^2}{n_1 + (n_1 - 1)} + \frac{\sum_{j=1}^{n_2} y_{2j}^2}{n_2 + (n_2 - 1)}}} \quad (8.2).$$

Esta estatística tem como referência uma tabela da distribuição normal se as amostras forem suficientemente grande. Para amostras pequenas, não é tão claro o teste e a distribuição associada. Quando é assumida a igualdade das variâncias, Fisher verificou que na prática U tem sido satisfatoriamente utilizada, com base numa distribuição t, no sentido de que este controla de maneira segura a chance de rejeitar a hipótese quando esta é realmente verdadeira. U é provavelmente o melhor critério quando desejamos detectar diferenças entre médias considerando variâncias heterogêneas. Entretanto, pode ocorrer que utilizando U pode ocorrer situações em que o teste pode ser realizado com diferenças entre σ_1^2 e σ_2^2 e que isso mascare diferenças entre as médias, apontando não significância quando na realidade existe.

Nem sempre U tem uma distribuição t-student, pois o quadrado do erro padrão de $(\bar{y}_1 - \bar{y}_2)$ é $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$ que é uma quantidade sob a raiz em U um estimador não viciado de $\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{(n_1 + n_2 - 2)} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$. Isso é igual a $\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)$ somente se $\sigma_1^2 = \sigma_2^2$ ou $n_1 = n_2$. Temos que, utilizando V não sofremos essa imposição, mas esta distribuição ainda depende da grandeza do valor obtido em σ_1^2/σ_2^2 . Então, um problema é determinar uma distribuição para U e V. A alternativa é determinar uma distribuição aproximada.

A distribuição aproximada de U e V quando $\mu_1 = \mu_2$ pode ser escrita por

$$\left(\bar{Y}_1 - \bar{Y}_2 \right) = \chi' \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}; \sum_{j=1}^r y_{1j}^2 = \chi_1^2, \sum_{j=1}^r y_{2j}^2 = \chi_2^2 \text{ onde } \chi'^2, \chi_1^2 \text{ e } \chi_2^2, \text{ são}$$

variáveis independentes com distribuição qui-quadrado com $(n_1 - 1)$ e $(n_2 - 1)$ graus de

liberdade respectivamente. É possível, então, escrever U e V sob a forma de

$$Y = \frac{\chi'}{\sqrt{a\chi_1^2 + b\chi_2^2}} = \frac{\chi'}{\sqrt{w}} \quad (8.3)$$

Onde:

$$a = \frac{\sigma_1^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}{(n_1 + n_2 - 2) \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)} \quad \text{e} \quad b = \frac{\sigma_2^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}{(n_1 + n_2 - 2) \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)} .$$

2. Lemas para comparações múltiplas de Banerjee

Para proceder testes de comparações múltiplas a partir das técnicas sugeridas por Banerjee, deve-se utilizar os seguintes lemas (Tamhane [6]).

Lema 1: Consideremos U e V_i variáveis aleatórias com distribuição qui-quadrado com um 1 e ν_i graus de liberdade, respectivamente, sendo U independente de V_i e $1 \leq i \leq t$. Consideremos, também $\lambda_i \geq 0$ ($1 \leq i \leq k$) um conjunto de constantes, sendo

que $\sum_{i=1}^t \lambda_i = 1$. Então,

$$P \left\{ U \leq \sum_{i=1}^t t_{\nu_i, \alpha/2}^2 \left(\frac{\lambda_i V_i}{V_i} \right) \right\} \geq 1 - \alpha .$$

Lema 2: Deixemos $Y_1, \dots, Y_k (Z_1, \dots, Z_k)$ serem variáveis aleatórias normalmente distribuídas com matriz de correlação

$$\{\rho_{ij}\} (\{\eta_{ij}\}).$$

Se $\rho_{ij} \geq \eta_{ij}$ ($1 \leq i \leq k-1$), então para qualquer conjunto de constantes reais

a_1, a_2, \dots, a_k , temos

$$P\{Y_1 \leq a_1, \dots, Y_k \leq a_k\} \geq P\{Z_1 \leq a_1, \dots, Z_k \leq a_k\}.$$

Lema 3: Consideremos Y_1, \dots, Y_k , variável aleatória normalmente distribuída com uma matriz de correlação $\{\rho_{ij}\}$. Então, para qualquer conjunto não negativo de constantes a_1, a_2, \dots, a_k , temos

$$P\{|Y_1| \leq a_1, \dots, |Y_k| \leq a_k\} \geq \prod_{i=1}^k P\{|Y_i| \leq a_i\}.$$

Para desenvolver o próximo lema, é necessário utilizar a seguinte definição devida a Esary, Proschan e Walkup [18].

Definição: Variáveis aleatórias Y_1, \dots, Y_k são ditas associadas (correlacionadas) se para todas funções reais não decrescentes ϕ e ψ de k argumentos, nós temos $\text{cov}(\phi(Y_1, \dots, Y_k), \psi(Y_1, \dots, Y_k)) \geq 0$.

No seguinte lema teremos a generalização do resultado obtido por Kimball [19].

Lema 4: Sejam X_1, X_2, \dots, X_k variáveis aleatórias a valores reais independentes e $\psi_j(X_1, \dots, X_k) (1 \leq j \leq p)$ função não negativa a valores reais as quais são funções não decrescentes para cada um de seus argumentos $X_i (1 \leq i \leq k)$. Então, fazendo $Y_j = \psi_j(X_1, \dots, X_k)$, temos

$$E\left\{\prod_{j=1}^p Y_j\right\} \geq \prod_{j=1}^p E\{Y_j\}$$