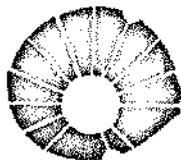


COMPARAÇÃO DE PROCEDIMENTOS DE COMPARAÇÕES
MÚLTIPLAS: PARAMÉTRICOS E NÃO PARAMÉTRICOS.



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO

CAMPINAS - SÃO PAULO
BRASIL

COMPARAÇÃO DE PROCEDIMENTOS DE COMPARAÇÕES MÚLTIPLAS: PARAMÉTRICOS E
NÃO PARAMÉTRICOS

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. Manoel Raimundo de Sena Júnior e aprovada pela Comissão Julgadora.

Campinas, 02 de setembro, de 1987.

Prof. Dr.



Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Estatística.

A meu pai, Manoel Raimundo de Sena
(in memoriam)

A minha mãe e irmãos

AGRADECIMENTOS

. Ao Prof. Dr. Ademir José Petenate, pela orientação deste trabalho, manifestando paciência, compreensão e competência em todas as fases de elaboração.

. À Inára Ignacchitti de Sena pelo apoio nos bons e maus momentos.

. À direção e aos colegas do Departamento de Estatística e Informática da UFPB pela credibilidade em mim depositada.

. Ao CNPq, PICD (CAPES), pelo apoio financeiro.

. À bibliotecária Marisabel Regina Rodrigues do Amaral, pela revisão das referências bibliográficas.

. Ao Antonio Sérgio Andó pelo apoio computacional.

. À Maria Fatima L. Dallantonia, pelo serviço de datilografia.

Aos colegas da pós-graduação do IMECC que participaram de alguma forma para a realização deste trabalho.

Í N D I C E

SUMÁRIO	i
CAPÍTULO I	1
1. PROCEDIMENTO DE COMPARAÇÕES MÚLTIPLAS (PCM)	1
1.1 - Introdução	1
1.2 - Os Modelos	1
1.3 - As Definições	2
1.4 - As Técnicas Estatísticas	4
1.4.1.1 - Estatística do Sinal (Steel)	4
1.4.1.2 - Teste do Sinal (Steel)	7
1.4.2.1 - Estatística dos Postos (Steel)	10
1.4.2.2 - Estatística dos Postos (Steel-Dwass)	13
1.4.3 - Estatística dos Postos Assinalados ...	15
1.4.4 - Estatística de Kruskal - Wallis (Nemenyi)	18
1.4.5 - Estatística dos Postos de Friedman (Nemenyi)	21
1.4.6 - Teste da Alternativa Ordenada (Jonck heere, Terpstra)	23
1.4.7 - Teste t (Dunnett)	26
1.4.8 - Teste Studentizado (Tukey-Kramer)	28
CAPÍTULO II	31
2. COMPARAÇÃO DE (PCM's)	31
2.1 - Introdução	31
2.2 - Razão de Erro	32

2.3 - Poder	33
2.4 - Eficiência Relativa	36
CAPÍTULO III	39
3. METODOLOGIA DO TRABALHO	39
3.1 - Teste dos Postos (Steel)	41
3.1.1 - Comparação - Um Controle Contra Vários Tratamentos	41
3.1.2 - Comparação - Tratamento Contra Tratamen <u>to</u> - (Steel-Dwass)	42
3.2 - Teste de Kruskal-Wallis - (Nemenyi)	44
3.3 - Teste do Sinal - (Steel)	46
3.3.1 - Comparação - Um Controle Contra Vários Tratamentos	46
3.3.2 - Comparação - Tratamento Contra Tratamen <u>to</u> - (Steel)	47
3.4 - Testes dos Postos Assinalados - (Nemenyi)	47
3.5 - Teste de Friedman - (Nemenyi)	48
3.6 - Teste da Alternativa Ordenada - (Jonckheere-Terps <u>tra</u>)	49
3.7 - Teste Studentizado - (Tukey-Kramer)	50
3.8 - Teste "t" - (Dunnett)	50
CAPÍTULO IV	53
4. DISCUSSÃO ASSOCIADA AOS RESULTADOS	53
4.1 - Introdução e Objetivo	53
4.2 - O Problema do Empate com a Distribuição de Pois <u>son</u>	53

4.2.1 - Discussão Com a Presença do Controle ... 54

4.2.2 - Discussão Sem a Presença do Controle ... 54

4.3 - Comentários e Conclusões 63

4.3.1 - Comparação: Um Controle Contra Vários
Tratamentos 63

4.3.2 - Comparação: Tratamento Contra Tratamento 65

REFERÊNCIAS BIBLIOGRÁFICAS 67

ANEXO 73

S U M Á R I O

Procedimentos de comparações múltiplas (PCM's) para os pares de médias (incluindo ou não um tratamento sob controle), paramétricos e não paramétricos, têm sido amplamente discutidos nos últimos anos. Dentre esses, os procedimentos paramétricos mais comumente recomendados são os de Dunnett (1964) e Tukey-Kramer (1956), respectivamente, para a inclusão e não inclusão de um tratamento sob controle. Porém, restrições a ambos os métodos devem ser consideradas, tais como, normalidade da população, independência, balanceamento e homocedasticidade das observações. Entretanto, dentre os não paramétricos algumas destas restrições não são necessárias.

O nosso estudo investiga o vício da razão de erro do tipo I, em situações onde a condição de normalidade não está satisfeita.

Alguns PCM's não paramétricos e os dois (paramétricos) citados acima são apresentados e comparados entre si, considerando os aspectos de conservantismo, adequabilidade e optimalidade.

No Capítulo I apresentamos os PCM's que serão discutidos no nosso estudo; no Capítulo II colocamos os métodos de comparação mais comumente usados; já no Capítulo III demonstramos a metodologia empregada na elaboração desse estudo e finalmente no Capítulo IV investigamos a significância de cada procedimento na presença da variação da probabilidade de empate entre as

observações, juntamente com as nossas conclusões.

C A P Í T U L O I

1. PROCEDIMENTO DE COMPARAÇÕES MÚLTIPLAS (PCM)

1.1 - INTRODUÇÃO

Através do teste F da Análise de Variância podemos testar, a um nível de significância α , a hipótese básica H_0 que todas as médias μ_i de k tratamentos sejam iguais, ou que todas k médias sejam iguais a uma média μ_0 dita como a média de uma situação sob controle. Se a hipótese básica H_0 é rejeitada, estaremos admitindo que, para pelo menos um par, $\mu_i \neq \mu_0$, ou, $\mu_i \neq \mu_{i'}$; $i \neq i'$. É de interesse seguir nessa investigação afim de localizar as diferenças entre as médias dentre os diferentes tratamentos. Essa continuidade pode ser feita através da técnica denominada Comparações Múltiplas.

Um PCM é um procedimento no qual, com base nas observações amostrais, toma-se um número (digamos N_d) de afirmações (decisões) envolvendo os parâmetros de interesse. Os procedimentos são realizados com base na hipótese básica H_0 , na qual, uma particular combinação de afirmações, no conjunto de todas possíveis afirmações, são feitas usando o procedimento, levando à aceitação ou rejeição da hipótese básica: O termo "Procedimento de Comparações Múltiplas" é usado quando N_d for maior que 1.

1.2 - OS MODELOS

Basicamente associaremos alguns modelos, como por

exemplo, para a comparação de um controle contra vários tratamentos, quando as observações chegam em blocos:

$$p_{Y_{ij}}(x) = p(x - \theta_i - \beta_j) \quad \begin{array}{l} i=0,1,\dots,k \\ j=1,2,\dots,n \end{array} \quad (1.2.1)$$

ou quando as observações não chegam obrigatoriamente em blocos,

$$p_{Y_{ij}}(x) = p(x - \theta_i) \quad \begin{array}{l} i=0,1,\dots,k \\ j=1,2,\dots,n \end{array} \quad (1.2.2)$$

onde: Y_{ij} : j -ésima observação do tratamento i

$p_{Y_{ij}}$: função de densidade de Y_{ij}

$p(x)$: uma função de densidade (desconhecida fixada)

θ_i : desvio populacional do i -ésimo tratamento

β_j : desvio populacional do j -ésimo bloco.

Ainda com base nestes modelos estaremos comparando todos os tratamentos entre si, usamos também estes modelos quando as observações chegam ou não obrigatoriamente em blocos, respectivamente.

Aqui, o interesse está no desvio populacional dos tratamentos, tanto em relação a um controle (θ_0) como nos demais tratamentos. Ou seja, estaremos interessados nas quantidades $(\theta_i - \theta_0)$ (no caso da comparação de um controle contra os k tratamentos) e $(\theta_i - \theta_{i'})$ $i \neq i'$ (no caso da comparação de todos tratamentos entre si).

1.3 - AS DEFINIÇÕES

Consideramos neste estudo, a importância da clareza dos comentários e definições dos seguintes temas:

- . **HIPÓTESE NULA (H_0):** Uma hipótese nula (H_0) determina se uma distribuição básica (ou distribuições) é membro de uma determinada classe, comumente estabelecida pela especificação de um parâmetro.
- . **DISTRIBUIÇÃO NULA:** A distribuição nula de uma estatística é a distribuição assumida quando a hipótese nula é verdadeira.
- . **TESTE DE HIPÓTESE:** É uma regra de decisão, a qual, com base nas observações amostrais, aceita-se ou rejeita-se a hipótese nula.
- . **NÍVEL DE SIGNIFICÂNCIA:** O nível de significância de um teste é a probabilidade de rejeitar falsamente a hipótese nula, ou seja, a probabilidade de rejeitar a hipótese nula, quando de fato, ela é verdadeira.
- . **REGIÃO CRÍTICA:** Uma região crítica de um teste de hipótese, é uma região de valores amostrais que leva à rejeição da hipótese nula. Especificamente, uma região crítica é escolhida para ter-se:
 - a) baixa probabilidade quando a hipótese nula é verdadeira e;
 - b) alta probabilidade quando a alternativa de interesse é verdadeira. Especificar um nível de significância α , é equivalente a especificar que a região crítica tem uma probabilidade α sob a hipótese nula.

Partindo dessa linha de definições é importante também deixar claro os conceitos, pelos quais, estaremos conduzindo nosso estudo.

Um PCM é dito **CONSERVATIVO** se apresentar um vício ne

gativo no nível de significância, é dito OTIMISTA se, esse vício no nível de significância for positivo e diremos ADEQUADO se, esse vício for considerado irrelevante.

Apresentaremos a seguir alguns PCM's paramétricos e não paramétricos, cujos os desempenhos serão comparados através da razão de erro do tipo I.

1.4 - AS TÉCNICAS ESTATÍSTICAS

1.4.1.1 - Estatística do Sinal (Steel)

(Controle x Tratamento)

Seja Y_{ij} (i, j ; $j=1, \dots, n$; $i=0, 1, \dots, k$) n amostras independentes de tamanho $k+1$, com uma observação em cada amostra de uma das $k+1$ diferentes populações. As $k+1$ observações em cada amostra são assumidas independentes. A população controle corresponde a $i=0$ e $i=1, \dots, k$ corresponde aos tratamentos. Cada amostra $j=1, \dots, n$ corresponde aos n blocos. O interesse é saber se todos os tratamentos dentro de um bloco têm a mesma densidade.

O modelo usado é:

$$p_{Y_{ij}}(x) = p(x - \theta_i - \beta_j) \quad \begin{array}{l} i=0, \dots, k \\ j=1, \dots, n \end{array} \quad (1.4.1)$$

onde $p_{Y_{ij}}$ é a densidade de Y_{ij} , com $p(x)$ fixada e $\theta_0 = \beta_1 = 0$.

A hipótese nula será:

$$H_0: \theta_i = 0 \quad i=1, \dots, k \quad (1.4.2)$$

isto é, o tratamento não difere do controle.

Definimos:

$$D_{ij} = Y_{ij} - Y_{0j} \quad i=1,2,\dots,k \quad (1.4.3)$$

Então, sob (1.4.1) temos:

$$p_{D_{ij}}(x) = q(x - \theta_i) \quad i=1,2,\dots,k$$

onde $q(x) = \int_{-\infty}^{+\infty} p(y)p(y+x)dy$ e θ_i é a mediana das densidades de

$p_{D_{ij}}(x)$ portanto a hipótese nula pode ser colocada como:

$$H_0: \text{med } D_{ij} = 0 \quad \begin{matrix} i=1,\dots,k \\ j=1,\dots,n \end{matrix} \quad (1.4.4)$$

contra a alternativa bilateral

$$H_1: \theta_i \neq 0 \quad i=1,2,\dots,k \quad (1.4.5)$$

ou unilateral

$$H_1: \theta_i > 0 \quad \text{para algum } i \quad (1.4.6)$$

$$H_1: \theta_i < 0$$

O teste consiste da estatística do sinal para cada uma das k variáveis $\{D_{ij}; j=1,\dots,n\}$, comparando o valor máximo dessas k estatísticas com um ponto crítico apropriado.

Especificamente:

$$S_i^+ = \sum_{j=1}^n D_{ij}^+ \quad \text{onde} \quad D_{ij}^+ = \begin{cases} 1 & \text{se } D_{ij} > 0 \\ 0 & \text{c.c.} \end{cases} \quad (1.4.7)$$

O teste unilateral do desvio positivo rejeita H_0 quando $S^+ \geq S_+^\alpha$ onde $S^+ = \text{Max} \{S_1^+, \dots, S_k^+\}$ e S_+^α depende de n , k e α .

A correspondente contagem negativa será:

$$S_i^- = \sum_{j=1}^n D_{ij}^- \quad \text{onde} \quad D_{ij}^- = \begin{cases} 1 & \text{se } D_{ij} < 0 \\ 0 & \text{c.c.} \end{cases} \quad (1.4.8)$$

- . O teste unilateral negativo rejeita H_0 quando $S^- \geq S^{\alpha} (= S_+^{\alpha})$, onde $S^- = \text{Max} \{S_1^-, \dots, S_k^-\}$.
- . O teste bilateral rejeita H_0 quando $S_{\pm} \geq S_{\pm}^{\alpha}$, onde $S_{\pm} = \text{Max} \{S_1^{\pm}, \dots, S_k^{\pm}\}$ e $S_{\pm}^{\alpha} = \text{Max} \{S_{\pm}^+, S_{\pm}^-\} = \text{Max} \{S_{\pm}^+, n - S_{\pm}^+\}$, e S_{\pm}^{α} é o ponto crítico que depende de n , k e α .

Sob o modelo (1.4.1) as constantes críticas podem ser usadas para a construção de Intervalos de Confiança Simultâneos (I.C.S.). Um Intervalo de Confiança Simultâneo bilateral para θ_i é dado por:

$$[D_{i(n-S_{\pm}^{\alpha}+1)}; D_{i(S_{\pm}^{\alpha})}] \quad (1.4.9.1)$$

Os $D_{i(1)}, \dots, D_{i(k)}$ são estatísticas de ordem que correspondem a D_{i1}, \dots, D_{in} . Quando todos os k intervalos são tomados conjuntamente eles tem "aproximadamente" probabilidade $(1-\alpha)$ de cobrir todos os k desvios $\theta_i, \dots, \theta_k$. O termo aproximadamente é usado, pois, o ponto crítico S_{\pm}^{α} é calculado com base em aproximações para grandes amostras. Os Intervalos de Confiança Simultâneos unilaterais para os θ_i são:

$$[D_{i(n-S_+^{\alpha}); +\infty}] \text{ e } [-\infty; D_{i(S_-^{\alpha})}] \quad (1.4.9.2)$$

O número de observações em cada tratamento deve ser igual ao do controle. Se as observações não ocorrem naturalmente em blocos, a estatística é feita com outra abordagem, agrupando aleatoriamente observações dentro dos blocos.

Um valor aproximado para S_+^{α} é o menor inteiro, maior ou igual a:

$$\frac{n}{2} + \frac{1}{2} + m(k, \alpha, \rho) \cdot \frac{\sqrt{n}}{2} \quad (1.4.10)$$

onde $m_{(k, \alpha, \rho)}$ é o ponto percentil α superior da distribuição do Máximo de k Variáveis Aleatórias normais igualmente correlacionadas ($\rho = 1/3$) e o termo " $1/2$ " é a usual correção de continuidade.

Gupta (1963) tabelou $m_{(k, \alpha, \rho)}$ para $\rho = 1/3$. Tabelas para S_{\pm}^{α} são encontradas em Miller (1981) para $\alpha = .05, 0.1$; $k = 2(1)10$; $n = 6(1)20$; (5)50; 100. As tabelas originais de Steel foram baseadas em Dunnett (1961) para $\rho = 1/2$, pois, não havia ainda avaliado com $\rho = 1/3$. As aproximações são boas desde que $0 \leq \rho \leq 1/2$ (excelente para n pequeno).

O ponto crítico S_{\pm}^{α} pode ser calculado similarmente. Rejeitar H_0 para grandes valores de S_{\pm} é equivalente a rejeitar H_0 para grandes ou pequenos valores de S_{\pm} , pois, $S^{-} + S^{+} = n$ para todo i . Aproximações para grandes amostras sugerem que S_{\pm}^{α} seja aproximado pelo menor inteiro igual ou superior a:

$$\frac{n}{2} + \frac{1}{2} + |m|_{(k, \alpha, \rho)} \frac{\sqrt{n}}{2} \quad (1.4.11)$$

onde $|m|_{(k, \alpha, \rho)}$ é o ponto percentil α superior da distribuição do Máximo Valor absoluto de k Variáveis Aleatórias normais igualmente correlacionadas ($\rho = 1/3$). Tabelas baseadas em $\rho = 1/2$ são encontradas em Miller (1981), pois, ainda não foi avaliado para $\rho = 1/3$. Nemenyi (1963) deu uma curta tabela para S_{\pm}^{α} com $k = 2, 3$, baseado no cálculo de $|m|_{(k, \alpha, \rho)}$ com $\rho = 1/3$ de Thigpen e David (1961).

1.4.1.2 - Teste do Sinal (Steel)

(Tratamento x Tratamento)

A estrutura deste teste é a mesma que no caso uniamostal. A comparação é feita entre todos tratamentos.

Seja $\{Y_{ij}: i=1, \dots, k; j=1, \dots, n\}$ n amostras independentes de k populações independentes (uma de cada k diferentes populações).

$$P_{Y_{ij}}(x) = p(x - \theta_i - \beta_j) \quad (1.4.12)$$

onde $p(x)$ é uma densidade fixada; θ_i é o desvio da população i e β_j é o desvio do bloco j . A quantidade paramétrica de interesse é $\theta_i - \theta_{i'}$, $i \neq i'$.

Seja:

$$D_{ii';j} = Y_{ij} - Y_{i'j} \quad \begin{matrix} i \neq i' & i, i' = 1, \dots, k \\ j = 1, 2, \dots, n \end{matrix} \quad (1.4.13)$$

tendo densidade:

$$P_{D_{ii';j}}(x) = q(x - (\theta_i - \theta_{i'})) \quad \begin{matrix} i, i' = 1, \dots, k & i \neq i' \\ j = 1, \dots, n \end{matrix}$$

onde $q(x) = \int_{-\infty}^{+\infty} p(y)p(x+y)dy$, q é simétrica sobre zero, portanto

$\theta_i - \theta_{i'}$ é a mediana populacional dos $D_{ii';j}$.

A hipótese nula é:

$$H_0: \theta_i = \theta_{i'} \quad \forall i, i' \quad (1.4.14)$$

isto é, as k populações são as mesmas, contra a alternativa usual

$$H_i: \theta_i \neq \theta_{i'} \quad \forall i, i' \quad (1.4.15)$$

o teste consiste em calcular:

$$S_{ii'}^+ = \sum_{j=1}^n D_{ii';j}^+ \quad \text{onde} \quad D_{ii';j} = \begin{cases} 1 & \text{se } D_{ii';j} > 0 \\ 0 & \text{c.c.} \end{cases}$$

$$S_{ii'} = \text{Max}\{S_{ii'}^+, n - S_{ii'}^+\} \quad \text{e} \quad S = \text{Max}_{i, i'}\{S_{ii'}\} \quad (1.4.16)$$

O teste rejeita H_0 se $S \geq S^\alpha$ onde S^α depende de n , k e α . A relação entre S e S^\pm é que S toma o valor máximo de (i, i') com $i \neq i'$ e S^\pm toma o máximo em $(i, 0)$.

Os Intervalos de Confiança Simultâneos dos dois testes são semelhantes, apenas com a troca de S^\pm por S^α , como:

$$[D_{ii', (n-S^\alpha+1)}; D_{ii', (S^\alpha)}] \quad i, i' = 1, 2, \dots, k \quad (1.4.17)$$

Quando tomados os $\binom{k}{2}$ intervalos eles têm uma probabilidade (aproximada) de $1-\alpha$ de cobrirem todas as diferenças $\theta_i - \theta_{i'}$.

As observações devem estar em blocos de mesmo tamanho.

Rejeitar H_0 quando S é grande, é equivalente a rejeitar H_0 quando:

$$\text{Max}\left\{\left|S_{ii'}^+ - \frac{n}{2}\right|\right\} \quad i, i' = 1, 2, \dots, k \quad (1.4.18)$$

é grande, pois, para grandes ou pequenos valores de $S_{ii'}$, deixam $S_{ii'}$ grande, isto é, grandes valores de $n-S_{ii'}$ deixam $S_{ii'}$ grande também.

Miller (1981) sugere uma aproximação para S como:

$$\frac{n}{2} + \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot q(k, \alpha, \infty) \cdot \frac{\sqrt{n}}{2} \quad (1.4.19)$$

onde $q(k, \alpha, \infty)$ é o percentil α superior da distribuição da amplitude de k Variáveis Aleatórias normais padrões e apresenta uma tabela baseada nesta aproximação, e esta aproximação é boa quando conhecermos o valor crítico exato de S^α .

1.4.2.1 - Estatística dos Postos (Steel)

(Controle x Tratamento)

Seja Y_{ij} ($i=0,1,\dots,k; j=1,\dots,n_i$) $k+1$ amostras independentes de tamanhos n_0, n_1, \dots, n_k respectivamente onde a população $i=0$, representa o controle e as outras $i=1, \dots, k$ são os tratamentos.

Consideramos o caso em que $m=n_0$ e $n=n_i; i=1, \dots, k$.

Supondo que Y_{ij} tem densidade $p_{Y_{ij}}(x) = p(x-\theta_i)$

$$(1.4.21)$$

onde:

- . $p(x)$ é uma função densidade fixada (desconhecida) com $\theta_0=0$.
- . θ_i são desvios populacionais em relação ao controle.

O interesse é, saber se todos os tratamentos são idênticos ao controle. A hipótese, a ser testada é:

$$H_0: \theta_i \neq 0 \quad i=1, \dots, k \quad (1.4.22)$$

contra as alternativas:

1. $H_1: \theta_i \neq 0$
2. $H_1: \theta_i > 0$ ou (1.4.23)
3. $H_1: \theta_i < 0$ para algum i .

Para o i -ésimo tratamento tomamos juntamente com o controle uma amostra $\{Y_{i1}, \dots, Y_{in}, Y_{01}, \dots, Y_{0m}\}$ combinada de tamanho $n+m$ e arranjadas em ordem crescente independente de onde elas venham.

Seja R_{i1}, \dots, R_{in} os postos das observações dos tratamentos e $R_i = \sum_{j=1}^n R_{ij}$

$$(1.4.24)$$

Testar H_0 contra H_1 em (1.4.23) é rejeitar H_0 quando:

$$R \geq r^\alpha \quad \text{onde} \quad R = \text{Max}\{R_1, \dots, R_k\} \quad (1.4.25)$$

O ponto crítico r^α depende de n , m , k e α .

Qualquer $R_i \geq r^\alpha$ é considerado como desvio positivo sob o controle.

Para a alternativa bilateral a região crítica consiste de pontos amostrais com:

$$R^* = \text{Max}\{R_1^*, \dots, R_k^*\} \geq r^* \quad (1.4.26)$$

onde

$$R_i^* = \text{Max}\{R_i, n(m+n+1) - R_i\} \quad i=1, 2, \dots, k$$

e r^* depende de n , m , k e α , e $n(m+n+1) - R_i$ é a soma dos postos do i -ésimo tratamento se as observações forem assinaladas em ordem decrescente.

A estrutura desse teste é a mesma que do teste "t" de Dunnett's (controle x tratamento) exceto a suposição de normalidade. Os tratamentos não necessitam chegar em blocos como no teste do sinal (controle x tratamento).

Sob H_0 as Variáveis Aleatórias $Y_{01}, \dots, Y_{0m}, Y_{11}, \dots, Y_{1n}, \dots, Y_{k1}, \dots, Y_{kn}$ são *iid*. Existem $(m+k*n)!$ ordenações possíveis e sob H_0 cada uma é igualmente possível (isto é, cada uma tem $\text{prob} = 1/(m+k*n)!$ de ser escolhida).

R_{ij} e R_i dependem das observações somente pela sua ordenação. Para cada ordenação $R_i = r_i$ pode ser ordenado em vetor $r = (r_1, \dots, r_k)$, cada ordenação produzindo um vetor r , podendo existir mais que uma ordenação dando o vetor r . A estrutura de probabilidade de um vetor $R = (R_1, \dots, R_k)$ é:

$$P(R = r) = \frac{n^{\circ} \text{ de ordenação dando } r}{(m + nk)!} \quad (1.4.27)$$

Sob H_0 R é distribuído assintomaticamente normal, onde a média e variância são detalhadas em Miller (1981).

Uma razoável aproximação para r é dada por:

$$r^{\alpha} = n \frac{(m+n+1)}{2} + \frac{1}{2} + m(k, \alpha, \rho) \sqrt{\frac{nm(n+m+1)}{12}} \quad (1.4.28)$$

onde $m(k, \alpha, \rho)$ é o percentil α superior da distribuição do Máximo de k Variáveis Aleatórias normais padrão, igualmente correlacionadas (a constante $1/2$ que aparece em (1.4.28) é usual correção de continuidade) e ρ é dado por:

$$\rho = \frac{n}{n+m+1} \quad (1.4.29)$$

Se $m=n$ (1.4.28) se reduz a:

$$r^{\alpha} = \frac{n(2n+1)}{2} + \frac{1}{2} + m(k, \alpha, 1/2) \sqrt{\frac{n^2(2n+1)}{12}} \quad (1.4.30)$$

Tabelas são apresentadas em Miller (1981). Esta aproximação de (1.4.30) e para alguns valores exatos de k e n (ambos pequenos), são boas (mesmo no pior caso).

Para o ponto crítico bilateral, uma aproximação apropriada é dada por:

$$r^{\alpha} = n \frac{(n+m+1)}{2} + \frac{1}{2} + |m|(k, \alpha, \rho) \sqrt{\frac{nm(n+m+1)}{12}} \quad (1.4.31)$$

onde $|m|(k, \alpha, \rho)$ é o percentil α superior da distribuição do Máximo Valor absoluto de k Variáveis Aleatórias normais padrões igualmente correlacionadas (1.4.29). Infelizmente tabelas para esta estrutura de correlação não existem. No caso $m=n$ o ρ

tende assintoticamente para $1/2$ (partindo de $1/3$). Steel sugere uma aproximação para $|m|_{(k, \alpha, \rho)}$ pelo ponto crítico da estatística "t" de Dunnett com "v" graus de liberdade. Esta aproximação encontra-se tabelada em Miller (1981).

1.4.2.2 - Estatística dos Postos (Steel-Dwass)

(Tratamento x Tratamento)

O modelo deste teste é semelhante ao caso de controle contra tratamentos (controle x tratamento), quando são tomados todos os pares de tratamentos populacionais.

Seja Y_{ij} ($i=1, \dots, k; j=1, \dots, n$) k amostras independentes com n observações cada. O caso não-balanceado não é considerado devido a impossibilidade do cálculo do ponto crítico.

Seja a densidade de Y_{ij} dada por:

$$p_{Y_{ij}}(x) = p(x - \theta_i) \quad (1.4.32)$$

onde $p(x)$ é uma função fixada e desconhecida com $\theta_i - \theta_{i'}$, sendo o parâmetro de interesse. A hipótese nula será:

$$H_0: \theta_i = \theta_{i'}, \quad \forall i, i' \quad i \neq i' \quad (1.4.33)$$

contra a simples negação

$$H_i: \theta_i \neq \theta_{i'}, \quad \forall i, i'$$

Esta é uma derivação da estatística do posto máximo de Wilcoxon (biamostrual) sob todos os pares populacionais. As observações para os tratamentos i e i' formam uma amostra $\{Y_{i1}, \dots, Y_{in}, Y_{i'1}, \dots, Y_{i'n}\}$ de tamanho $2n$.

Seja $R_{i1, i'}, \dots, R_{in, i'}$ os postos das observações Y_{i1}, \dots, Y_{in} , com respeito a amostra combinada de tamanho $2n$. A

soma dos postos para i -ésima amostra com respeito a amostra i' é:

$$R_{ii'} = \sum_{j=1}^n R_{ij,i'} \quad i, i' = 1, \dots, k \quad (1.4.34)$$

o teste rejeita H_0 quando:

$$R^{**} \geq r_{**}^{\alpha} \quad (1.4.35)$$

onde:

$$R^{**} = \text{Max}_{ii'} \{R_{ii'}^*\} \quad (1.4.36)$$

e

$$R_{ii'}^* = \text{Max}_{ii'} \{R_{ii'}, R_{i'i}\} \quad (1.4.37)$$

Note que para qualquer par de comparação apenas uma soma de postos é necessária, pois:

$$R_{i'i} = n(2n+1) - R_{ii'} \Rightarrow R_{ii'}^* = R_{i'i} \quad (1.4.38)$$

Assim a estatística dos postos (tratamento x tratamento) é o máximo dessas $\binom{k}{2}$ soma de postos simétricos.

A distribuição do módulo máximo de uma normal multivariada com a estrutura de correlação dessas variáveis [ver: Miller (1981) equações (69), (70) e (71)], não foi tabelada, no entanto, uma aproximação é feita pensando em substituir os valores apresentados na estrutura de correlação que são de $\{-1/3, 0, 1/3\}$ por $\{-1/2, 0, 1/2\}$, respectivamente, então a Variável Aleatória teria a mesma distribuição da amplitude de k Variáveis Aleatórias normais padrão, isto sugere uma aproximação para r_{**}^{α} por:

$$r_{**}^{\alpha} = \frac{n(2n+1)}{2} + \frac{1}{2} + q(k, \alpha, \infty) \sqrt{\frac{n^2(2n+1)}{24}} \quad (1.4.39)$$

onde $q(k, \alpha, \infty)$ é o percentil α superior da distribuição da amplitude de k Variáveis Aleatórias normais padrão, tabela para esta aproximação aparece em Miller (1981).

A estrutura deste teste é idêntica ao teste da amplitude studentizada de Tukey, apenas no caso de uma classificação. Um teste que poderia ser um competidor para este é do tipo Kruskal-Wallis.

1.4.3 - Estatística dos Postos Assinalados

A estrutura deste teste não difere entre os casos de controle contra tratamento e tratamento contra tratamento. O procedimento deste teste pode ser definido diretamente, mas ele não tem distribuição nula livre, a estatística teste depende da distribuição geradora.

Considere o problema controle contra tratamento com dois tratamentos e um controle. Seja Y_{ij} ($i=0,1,2; j=1,\dots,n$) $3n$ observações independentes com densidade dada por:

$$p_{Y_{ij}}(x) = p(x - \theta_i - \beta_i) \quad \begin{matrix} i=0,1,2 \\ j=1,\dots,n \end{matrix} \quad (1.4.40)$$

onde $p(x)$ é fixada (desconhecida) com $\theta_0 = \beta_1 = 0$. A hipótese nula é:

$$H_0: \theta_1 = \theta_2 = 0 \quad (1.4.41)$$

A estrutura do teste para cada (controle x tratamento) é definida como usual problema biamostrais. Sejam:

$$\begin{aligned} U_j &= Y_{1j} - Y_{0j} \\ V_j &= Y_{2j} - Y_{0j} \end{aligned} \quad j=1, \dots, n \quad (1.4.42)$$

considerando os U_j 's calcule os $|U_1|, \dots, |U_n|$ e os postos de ordem de tamanho, assinalando 1 para $|U|_{(1)}$, 2 para $|U|_{(2)}$, etc. Para a j -ésima observação U_j seja R_{1j} o posto deste valor. Defina:

$$SR_{1j}^+ = I_{|U_j|, U_j} \cdot R_{1j}; \quad j=1, \dots, n \quad (1.4.43)$$

onde

$$I_{x,z} = \begin{cases} 1 & \text{se } x \leq z \\ 0 & \text{c.c.} \end{cases} \quad (1.4.44)$$

Note que SR_{1j}^+ é zero se U_j é negativo, ou igual ao posto de $|U_j|$ se $U_j \geq 0$. Defina:

$$SR_1^+ = \sum_{j=1}^n SR_{1j}^+ \quad (1.4.45)$$

onde o SR_1 será a soma dos postos dos valores positivos de U_j . Para os V_j 's obtenha a análoga estatística SR_2^+ .

Para a quantidade negativa tome:

$$SR_{ij}^- = -I_{|U_j|, (U_j)^R} i_j \quad \begin{matrix} i=1, 2 \\ j=1, \dots, n \end{matrix} \quad (1.4.46)$$

e

$$SR_i^- = \sum_{j=1}^n SR_{ij}^- \quad i=1, 2 \quad (1.4.47)$$

A relação entre as estatísticas SR_{ij}^+ e SR_{ij}^- é dada pela identidade:

$$\frac{n(n+1)}{2} = SR_i^+ + |SR_i^-| \quad i=1,2 \quad (1.4.48)$$

Para o teste do desvio positivo a estatística teste será:

$$SR^+ = \text{Max}\{SR_1^+, SR_2^+\} \quad (1.4.49)$$

e rejeita-se H_0 quando:

$$SR^+ \geq sr^{\alpha} \quad (1.4.50)$$

onde sr^{α} depende de n e α .

As coordenadas SR_i^+ $i=1,2$ têm individualmente distribuição livre. Cada coordenada é uma estatística dos postos assinalados biamostrais padrão. Todavia, sua distribuição conjunta depende da desigualdade básica $p(x)$.

O par de $SR_i^+ = (SR_1^+, SR_2^+)$, nem sempre tem distribuição assintótica livre. Hollander (1965) provou que, assintoticamente SR tem distribuição normal bivariada. Com a média e variância livre da densidade básica, porém, com correlação variando em $p(x)$, dependendo do tamanho amostral e no limite.

A correlação entre SR_1^+ e SR_2^+ pode ser calculada para n finito e no limite quando n tende ao infinito, esta correlação tem valor limitado por:

$$\rho_{\infty} = 12\lambda(f) - 3 \quad (1.4.51)$$

onde $\lambda(f)$ é discutido em Miller (1981) e Lehmann (1964).

Lehmann mostrou que $\lambda(f)$ varia entre $1/4$ e $7/24$. Isto implica que o limite de (1.4.51) varia entre 0 e $1/2$. E ainda em alguns casos particulares de $\lambda(f)$, ele calculou o limite superior e fixando em quatro casas decimais, isto é, 0.2917 .

A aplicação desta técnica não se limita para o caso

onde $k=2$. Quando $k>2$ nos problemas (controle x tratamento) e (tratamento x tratamento), troca-se a distribuição normal bivariada pela multivariada. Substitui-se o ρ por uma estimativa de ρ_m para o cálculo do ponto crítico, onde o procedimento é idêntico ao teste dos postos (controle x tratamento) e (tratamento x tratamento), respectivamente. Com média e variância dada por:

$$E(SR_i^+) = \frac{n(n+1)}{4} \quad \text{e} \quad V(SR_i^+) = \frac{n(n+1)(2n+1)}{24} \quad (1.4.52)$$

Um competidor para este tipo de teste é o próprio teste dos postos (Steel) e os do tipo Kruskal-Wallis.

1.4.4 - Estatística de Kruskal-Wallis (Nemenyi)

Esta técnica baseia-se no método de (Kruskal-Wallis) a uma classificação de postos. Foi proposta e analisada por Nemenyi (1963). Um procedimento similar foi proposto por Dunn (1964).

Considere inicialmente o problema tratamento contra tratamento.

Seja Y_{ij} $\{i=1, \dots, k; j=1, \dots, n_i\}$ k amostras independentes de tamanho n_i , de observações independentes com

$N = \sum_{i=1}^k n_i$, Y_{ij} tem densidade dada por:

$$p_{Y_{ij}}(x) = p(x - \theta_i) \quad \begin{matrix} i=1, \dots, k \\ j=1, \dots, n_i \end{matrix} \quad (1.4.53)$$

onde $p(x)$ é densidade fixada e desconhecida com p arbitrário, os θ_i 's não são unicamente definidos, porém, suas diferenças $\theta_i - \theta_j$ são.

A hipótese considerada é:

$$H_0: \theta_i = \theta_{i'}, \quad \forall i, i' \quad i \neq i' \quad (1.4.54)$$

contra a alternativa geral:

$$H_1: \theta_i \neq \theta_{i'}, \quad \forall i, i' \quad (1.4.55)$$

Os postos das " N " observações em ordem de tamanho $Y_{(1)}, \dots, Y_{(N)}$, são tomados assinalando 1 para $Y_{(1)}$, 2 para $Y_{(2)}$, etc.

Seja R_{ij} o posto de Y_{ij} com as médias dos postos populacionais e geral, respectivamente como:

$$\bar{R}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij} \quad i=1, \dots, k \quad (1.4.56)$$

$$\bar{R}_{..} = \frac{1}{N} \sum_{i,j} R_{ij} = \frac{N+1}{2}$$

o teste rejeita H_0 quando:

$$H_{k-1} \geq h_{k-1}^\alpha$$

onde:

$$H_{k-1} = \frac{12}{N(N+1)} \sum_{i=1}^k n_i (\bar{R}_{i.} - \bar{R}_{..})^2 = \frac{12}{N(N+1)} \sum_{i=1}^k n_i \bar{R}_{i.}^2 - 3(N+1) \quad (1.4.57)$$

Para N grande h_{k-1}^α segue uma distribuição χ_{k-1}^2 , para N pequeno existem tabelas que aparecem em Hollander-Wolfe (1973).

Note que $\frac{H_{k-1}}{N-1}$ é a usual estatística F a uma classificação. Kruskal-Wallis (1952) tabelaram valores de H_{k-1} para $k=3$ e $n_i = 1(1)5$ com $i=1, 2, 3$, para $n_i = 5$ as aproximações pelo χ^2 foi determinado como sendo razoavelmente boa, assim como para $n_i > 5$ esta aproximação também é satisfatória.

No caso particular $n_i = n$ o ponto crítico se reduz para rejeitar H_0 se:

$$|\bar{R}_{i.} - \bar{R}_{i'.}| \geq q(k, \alpha, \infty) \left| \frac{k(nk+1)}{12} \right|^{1/2} \quad \forall i \neq i' \quad (1.4.58)$$

onde $q(k, \alpha, \infty)$ é percentil α superior da distribuição da amplitude de k Variáveis Aleatórias Normais padrão. Pontos críticos para $q(k, \alpha, \infty)$ aparecem em Miller (1981).

Analogamente para o caso de controle contra tratamento, podemos considerar o tamanho de população $i=0$ como " m " e o tamanho de cada tratamento ($i=1, \dots, k$) como " n ".

Para um teste unilateral do desvio positivo, rejeita-se H_0 quando:

$$\bar{R}_{i.} - \bar{R}_{0.} \geq m(k, \alpha, \rho) \left| \frac{N(N+1)}{12} \right|^{1/2} \left(\frac{1}{n} + \frac{1}{m} \right)^{1/2} \quad (1.4.59)$$

onde $m(k, \alpha, \rho)$ é percentil α superior da distribuição de k Variáveis Aleatórias Normais com ρ comum dado por:

$$\rho = \frac{n}{n+m} \quad (1.4.60)$$

Para o teste bilateral, rejeita-se H_0 quando:

$$|\bar{R}_{i.} - \bar{R}_{0.}| \geq |m|(k, \alpha, \rho) \left| \frac{N(N+1)}{12} \right|^{1/2} \left(\frac{1}{n} + \frac{1}{m} \right)^{1/2} \quad (1.4.61)$$

onde $|m|(k, \alpha, \rho)$ é o percentil α superior da distribuição do máximo valor absoluto de k Variáveis Aleatórias Normais com ρ dado por (1.4.60).

Este teste pode ser aplicado desde que não haja efeito de bloco. É um teste não paramétrico análogo ao teste de projeção F de Scheffé para ANOVA a uma classificação, também ao

teste da amplitude studentizada de Tukey, teste "t" de Dunnett's (controle x tratamento) e ainda ao teste dos postos de Steel.

1.4.5 - Estatística dos Postos de Friedman (Nemenyi)

Este teste é descrito pela análise do posto de Friedman a duas classificações, ele foi proposto e analisado por Nemenyi (1963).

A estrutura do teste é idêntica no caso de tratamento contra tratamento e de controle contra tratamento.

Considere o problema do (tratamento x tratamento).

Seja $\{Y_{ij} : i=1, \dots, k; j=1, \dots, n\}$ k amostras independentes de n observações cada, com densidade dada por:

$$p_{ij}(x) = p(x - \theta_i - \beta_j) \quad \begin{matrix} i=1, \dots, k \\ j=1, \dots, n \end{matrix} \quad (1.4.62)$$

$p(x)$ é fixada e desconhecida, os parâmetros θ_i , β_j não são definidos unicamente, porém, $\theta_i - \theta_{i'}$, $\beta_j - \beta_{j'}$ são. Faça $\theta_1 = \beta_1 = 0$ e considere a hipótese nula como:

$$H_0: \theta_i = \theta_{i'}, \quad \begin{matrix} i \neq i' \\ i, i' = 1, \dots, k \end{matrix} \quad (1.4.63)$$

contra a alternativa geral:

$$H_i: \theta_i \neq \theta_{i'}, \quad \forall i, i' \quad (1.4.64)$$

o modelo é de duas classificações com uma observação por cela. Tome R_{ij} como os postos de Y_{ij} relativo a ordenação $Y_{(1)j} < Y_{(2)j} < \dots < Y_{(k)j}$, no bloco j . A média dos postos para população i e geral são: $R_{i.}$ e $R_{..}$, definido por:

$$\bar{R}_{i.} = \frac{1}{n} \sum_{j=1}^n R_{ij}; \quad \bar{R}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^n R_{ij} = \frac{k+1}{2} \quad (1.4.65)$$

onde $N = kn$. A Estatística dos postos de Friedman é:

$$X_r^2 = \frac{12n}{k(k+1)} \sum_{i=1}^k (\bar{R}_{i.} - \bar{R}_{..})^2 = \frac{12n}{k(k+1)} \sum_{i=1}^k \bar{R}_{i.}^2 - 3n(k+1)$$

o teste rejeita H_0 quando $X_r^2 \geq x_r^{2\alpha}$ onde $x_r^{2\alpha}$ está em função de α, k, n .

Tabelas para $x_r^{2\alpha}$ exata são apresentadas para $k = 3, 4, 5$ com $(n=2, \dots, 13)$ ($n=2, \dots, 8$) e $(n=3, 4, 5)$ respectivamente a cada um dos k . E para n grande, as estatísticas $x_r^{2\alpha}$ tem aproximadamente uma distribuição $\chi_{k-1}^2 \cdot \frac{x_r^2}{k-1}$ é a usual estatística F para as populações com duas classificações. Tabelas para valores de $x_r^{2\alpha}$ são encontradas em Hollander-Wolfe (1973), e outras variações de k e n em Siegel (1956).

O teste de Friedman dá acesso a teste simultâneo des de que:

$$|\bar{R}_{i.} - \bar{R}_{i'.}| \leq (x_r^{2\alpha})^{1/2} \cdot \left[\frac{k(k+1)}{6n} \right]^{1/2} \quad i, i'=1, \dots, k$$

(1.4.66.1)

ou quando n tende a infinito:

$$|R_{i.} - R_{i'.}| \leq q(k, \alpha, \infty) \left[\frac{nk(k+1)}{12} \right]^{1/2} \quad i, i'=1, \dots, k$$

(1.4.66.2)

com probabilidade maior que $(1-\alpha)$. Qualquer diferença exceden do este ponto crítico é tomado como evidência de que $\theta_i - \theta_{i'} \neq 0$.

Para o problema de controle contra tratamento, consi dere $\{Y_{ij}: i=0, 1, \dots, k; j=1, \dots, n\}$ como $k+1$ amostra independen te com densidade dada por (1.4.62) com $i=0, 1, \dots, k$ e $j=1, \dots, n$.

O teste unilateral rejeita H_0 quando:

$$R_{i.} - R_{0.} \geq m(k, \alpha, 1/2) \left| \frac{n(k+1)(k+2)}{6} \right|^{1/2} \quad (1.4.67)$$

onde $m(k, \alpha, \rho)$ foi definido anteriormente no teste de postos de Kruskal-Wallis (K.W.) como também o seguinte $|m|(k, \alpha, \rho)$.

O teste bilateral rejeita H_0 quando:

$$|R_{i.} - R_{0.}| \geq |m|(k, \alpha, 1/2) \left| \frac{n(k+1)(k+2)}{6} \right|^{1/2} \quad (1.4.68)$$

A análise deste teste é restrita a duas classificações com uma observação por cela.

Alguns testes usando a teoria da normalidade, podem ser aplicados como um competidor aos testes apresentados, como por exemplo, o teste da amplitude studentizada (tratamento x tratamento), o teste da projeção F (controle x tratamento).

Quando a distribuição é verdadeiramente normal, os testes usando esta teoria de normalização são mais poderosos, isto também acontece quando, existe quase-normalidade. As pesquisas no entanto, se voltam para o caso onde a normalidade não se apresenta.

Um bom competidor não-paramétrico para este teste, é o Teste do Sinal (Steel). Outros testes poderão ser comparados, mas existe a dificuldade diante da distinção de caráter deste com os demais.

1.4.6 - Teste da Alternativa Ordenada (Jonckheere, Terpstra)

Seja Y_{ij} ($i=1, \dots, n_j; j=1, \dots, k$) k amostras independentes de tamanho n_i com modelo básico.

$$Y_{ij} = \mu + \theta_j + e_{ij} \quad (1.4.69)$$

onde μ é conhecida média geral e θ_j é um desconhecido efeito do tratamento i e $\sum_{j=1}^k \theta_j = 0$.

Os e 's (erros) são mutuamente independentes, e cada um vem da mesma população contínua.

A hipótese nula será:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k \quad (1.4.70)$$

contra a alternativa da forma:

$$H_1: \theta_1 \leq \theta_2 \leq \dots \leq \theta_k \quad (1.4.71)$$

pelo menos uma desigualdade é estrita.

A estrutura do teste consiste em calcular todas as $\binom{k}{2}$ combinações Mann-Whitney U_{uv} ; $u < v$ onde:

$$U_{uv} = \sum_{i=1}^{n_u} \sum_{i'=1}^{n_v} \phi(Y_{i'v}, Y_{iu}) \quad (1.4.72)$$

e $\phi(a, b) = 1$ se $a < b$; $\phi(a, b) = 0$ se $a > b$, no caso de empate toma-se $\phi = 1/2$. Isto é, U_{uv} é o número de observações da amostra u menores que da amostra v .

Considere:

$$J = \sum_{u < v} U_{uv} = \sum_{u=1}^{k-1} \sum_{v=u+1}^k U_{uv} \quad (1.4.73)$$

como sendo a soma dessas $\binom{k}{2}$ combinações de Mann-Whitney.

O teste rejeita H_0 quando:

$$J > j[\alpha, k, (n_1, \dots, n_k)] \quad (1.4.74)$$

onde $j[\alpha, k, (n_1, \dots, n_k)]$ satisfaz a equação:

$$P_0(J > j[\alpha, k, (n_1, \dots, n_k)]) = \alpha$$

este valor encontra-se tabelado em Hollander-Wolfe (1973) para $k=3$; $2 < n_1 < n_2 < n_3 < 8$ e $k=4, 5, 6$; $n_i = \dots = n_k = 2(1)6$.

Quando $\text{Min}(n_1, \dots, n_k)$ tende ao infinito e H_0 é verdadeira J^* tem assintoticamente distribuição normal $(0, 1)$ dada por:

$$J^* = \frac{J - E_0(J)}{\sqrt{\text{Var}(J)}} = \frac{J - \{(N^2 - \sum_{i=1}^k n_i^2) / 4\}}{\sqrt{\{[N^2(2N+3) - \sum_{i=1}^k n_i^2(2n_i+3)] / 72\}}} \quad (1.4.75)$$

onde $N = \sum_{i=1}^k n_i$, o teste desta forma rejeita H_0 quando:

$$J^* > Z_{(\alpha)} \quad (1.4.76)$$

Este teste pode ser preferido comparando com o teste de Kruskal-Wallis quando o experimentador espera o desvio de H_0 numa particular direção (note que, se direção esperada não é naturalmente ordenada em H_a , pode-se simplesmente rotular os tratamentos para que siga a ordem natural que acontece aqui), ainda assim a estatística de Kruskal-Wallis não é usada com este tipo de rotulação tomando como mesmo valor para todas as possíveis ($k!$) rótulos dos tratamentos.

Para $k=2$ o procedimento se reduz ao teste unilateral, de Wilcoxon, Hollander-Wolfe (1973) dá maiores detalhes neste sentido. Na condição de consistência do teste quando n_i/N tende a λ_i ($0 < \lambda_i < 1$) $i=1, \dots, k$, é suficiente para assumir que o teste, é consistente contra a alternativa (1.4.70) ver [Terpstra (1952)] e sobre a eficiência do teste Puri (1965) dá maiores detalhes.

1.4.7 - Teste t (Dunnett)

Esta técnica foi desenvolvida por Dunnett (1955) e aprimorada por ele mesmo em 1964. Considere Y_{ij} satisfazendo o usual modelo para ANOVA a uma classificação com $Y_{ij} = \theta_i + e_{ij}$.

Seja $Y_{ij}, \{i=0, 1, \dots, k; j=1, \dots, n_i\}$ k observações independentes de tamanho n_i . Seja também $\bar{Y}_i = \frac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}$ um estimador de θ_i a média do tratamento i , onde os e_{ij} são *iid* normais com média 0 (zero) e variância comum σ^2 . Suponha que S^2 seja um estimador não viciado de σ^2 independente dos \bar{Y}_i , dado por:

$$S^2 = \frac{k}{n} \sum_i \sum_j^{n_i} (Y_{ij} - \bar{Y}_i)^2 \quad (1.4.77)$$

onde $n = \sum_{i=1}^k n_i$. Nosso objetivo é testar a hipótese:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k = \theta_0 \quad (1.4.78)$$

contra a usual alternativa:

$$H: \theta_i \neq \theta_0 \quad (1.4.79)$$

para pelo menos um par, considere inicialmente $k=1$, ou seja, existe apenas um tratamento a ser comparado com o controle, e considere também Z dado por:

$$Z = \frac{(\bar{Y}_1 - \bar{Y}_0) - (\theta_1 - \theta_0)}{\sqrt{\frac{1}{n_1} + \frac{1}{n_0}}} \quad (1.4.80)$$

onde Z é uma Variável Aleatória $N(0, \sigma^2)$, portanto $t=Z/S$ segue uma distribuição t -student com n graus de liberdade. O teste

rejeita H_0 se:

$$t > d' \quad (1.4.81)$$

onde d' satisfaz a equação:

$$P(|t| < d') = 1-\alpha \quad (1.4.82)$$

a constante d' corresponde ao percentual α de pontos da distribuição t-student.

Considere agora o caso geral para k tratamentos e um controle. Podemos escrever o Z_i como:

$$Z_i = \frac{(\bar{Y}_i - \bar{Y}_0) - (\theta_i - \theta_0)}{\sqrt{1/n_i + 1/n_0}} \quad (1.4.83)$$

e $t_i = Z_i/S$, $j=1, \dots, k$. O teste rejeita H_0 quando:

$$|t_i| > d_i, \quad (1.4.84)$$

onde os d_i 's são escolhidos, de modo a satisfazer a equação:

$$P(|t_i| < d_i, \dots, |t_k| < d_k) = 1-\alpha \quad (1.4.85)$$

para escolher os d_i 's satisfazendo (1.4.85) é necessário ao ter a distribuição conjunta dos t_i 's, onde essa distribuição conjunta segue uma normal multivariada vetor de média 0 (zero), variância σ^2 e coeficiente de correlação dado por:

$$\rho_{ii'} = \frac{1}{\sqrt{\left(\frac{n_0}{n_i} + 1\right)\left(\frac{n_0}{n_{i'}} + 1\right)}} \quad i, i'=1, \dots, k \quad (1.4.86)$$

A estrutura do teste rejeitaria H_0 quando:

$$\text{Max}|t_i| > d_g \quad (1.4.87)$$

onde d_g depende de k , n , ρ e α .

Dunnett (1964) tabelou os valores de d_g quando o número de observações do controle é a mesma que o número de observações dos tratamentos e usou uma aproximação razoável, quando o número de observações do controle é maior que o número de observações dos tratamentos, mantendo ainda o número de observações dos tratamentos constante.

1.4.8 - Teste Studentizado (Tukey-Kramer)

Ordenação múltipla usando a amplitude studentizada, tem sido frequentemente usada para se fazer comparações entre média de tratamentos emparelhados e experimentos k -amostrais [ver Duncan (1951), Keuls (1952), Newman (1939) e Tukey (1953) para diversas versões do teste] considere Y_{ij} satisfazendo o usual modelo da ANOVA a uma classificação, com $Y_{ij} = \theta_i + e_{ij}$.

Seja Y_{ij} $\{i=1, \dots, k; j=1, \dots, n_i\}$ k amostras independentes de tamanho n_i , onde os e_{ij} são $iid N(0, \sigma^2)$ com parâmetro de média 0 e variância σ^2 desconhecida. A proposta desse experimento é provar que as inferências podem ser feitas sobre os valores desses parâmetros. O interesse particular está voltado para diferença $\theta_i - \theta_{i'}$, $(i \neq i')$ $1 < i, i' < k$.

Intervalos de confiança podem ser usados como teste baseados nos resultados do experimento, ou teste de hipótese quando assumido um específico valor. O procedimento de comparações múltiplas é usado quando controlada a razão de erro experimental.

O teste de múltipla ordenação de Tukey (1953), conhecido como método T , é um P.C.M. que propõe o seguinte Intervalo

de Confiança para estimar a quantidade $\theta_{i'} - \theta_i$,

$$\bar{Y}_{i.} - \bar{Y}_{i'}. \pm SR_{(\alpha, k, v)} s / \sqrt{n} \quad i, i' = 1, 2, \dots, k \quad (1.4.88)$$

onde $\bar{Y}_{i.}$ é a média do i -ésimo tratamento, $SR_{(\alpha, k, v)}$ é o percentil α superior da distribuição da amplitude de k Variáveis Aleatórias Normais padrão, s^2 é a usual estimativa não viciada de σ^2 e v é o número de graus de liberdade.

Os tamanhos amostrais $n_{i'}$ s são assumidos iguais a n sob a condição de comum variância.

Para testar a hipótese:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k \quad (1.4.89)$$

contra a usual alternativa:

$$H_1: \theta_i \neq \theta_{i'} \quad \forall i, i' \quad (1.4.90)$$

o método T , propõe rejeitar H_0 se:

$$\max_{i, i'} \frac{|\bar{Y}_{i.} - \bar{Y}_{i'}.|}{s / \sqrt{n}} \geq SR_{(\alpha, k, v)} \quad (1.4.91)$$

onde os valores de $SR_{(\alpha, k, v)}$ e s foram definidos anteriormente.

Uma limitação do método T está no fato de que os tamanhos amostrais devem ser iguais em todos os tratamentos. Sempre em planejamento de mesmo tamanho amostral resultam em acidentes nos tamanhos dos $n_{i'}$ s, portanto, há necessidade de um planejamento detalhando os tamanhos amostrais desiguais.

Um método para adaptação do método T (aliás, não só deste método como também do teste de Newman-Keuls e o teste da múltipla ordenação de Duncan's), foi dado por Kramer (1956) e aparentemente sem o conhecimento de Kramer e outros autores, foi também proposto por Tukey (1953), este método propõe a tro

ca da quantidade seguinte ao \pm em (1.4.88) por $SR/\sqrt{2}$ vezes o erro padrão da diferença entre as duas médias amostrais e (1.4.88) ficaria:

$$\bar{Y}_{i.} - \bar{Y}_{i'.} \pm SR_{(\alpha, k, v)} s[(n_i^{-1} + n_{i'}^{-1})/2]^{1/2} \quad (1.4.92)$$

e o teste rejeitaria H_0 quando:

$$\max_{i, i'} \frac{|\bar{Y}_{i.} - \bar{Y}_{j.}|}{s \sqrt{\frac{1}{2}(\frac{1}{n_i} + \frac{1}{n_{i'}})}} \geq SR_{(\alpha, k, v)} \quad (1.4.93)$$

onde as quantidades foram definidas anteriormente. Note que (1.4.92) se reduz em (1.4.88) quando os tamanhos amostrais são iguais, e também quando $k=2$ (1.4.92) se reduz ao usual teste t de student.

C A P Í T U L O I I

2. COMPARAÇÃO DE (PCM's)

2.1 - INTRODUÇÃO

A escolha do critério de comparação é o primeiro problema na avaliação dos PCM's. Um dos critérios é o poder que um determinado procedimento tem em detectar falhas na hipótese básica, um outro critério é o nível de erro admitido pelo procedimento.

Existe na literatura uma grande quantidade de publicações com relação ao problema de comparações múltiplas. O'Neill e Wetherill (1971) apresentam em suas referências bibliográficas a citação de 200 artigos abordando os diversos segmentos desse assunto. Os procedimentos PARAMÉTRICOS mais usados são: Duncan (1955); Tukey (1949); Dunnett (1955); Newman (1939); Keuls (1952); Scheffé (1953) e Waller & Duncan (1969). Estes procedimentos aplicam basicamente as teorias da normalidade, limitados pelos planejamentos do tipo ANOVA. Dunnett (1980) comparou a razão de erro de seis PCM's entre médias de k tratamentos, com variâncias homogêneas, usando simulações de Monte Carlo, e concluiu que o método de Tukey Kramer (1956) é o que apresenta menor vício da razão de erro comparado ao valor nominal α .

As vantagens dos procedimentos NÃO PARAMÉTRICOS, têm sido apresentadas por Steel (1959); Wilcoxon & Wilcox (1964);

Hollander (1966); Miller (1981); Mc Donald & Thompson (1967) e Sen (1968), considerando estatísticas dos postos ou sinal. Já Boyett e Schuster (1977 e 1979) aplicam os testes de permutações ou de postos estratificados.

Neave & Grauger (1968) comparou alguns testes para diferença de médias entre duas populações. Baseados em simulações de Monte Carlo, eles simularam a distribuição básica sob diversas situações experimentais, envolvendo procedimentos PARA MÉTRICOS e NÃO PARAMÉTRICOS. Fizeram estimativas para o poder de cada teste juntamente com medidas de extensão, as quais davam resultados similares. Ainda no artigo, eles apresentam conclusões e recomendações que devem ser levadas em consideração, sempre que se deseje pesquisar neste sentido.

2.2 - RAZÃO DE ERRO

Uma das melhores ferramentas usadas na comparação de PCM's é através do cálculo da razão de erro. Para o teste de significância a primeira espécie de erro usado é o que cometemos ao rejeitar a hipótese H_0 , quando de fato, ela é verdadeira.

Existem dois caminhos básicos que podemos utilizar no cálculo da razão de erro: por inferência (comparações) ou por experimentos. No primeiro caso, cada inferência é tratada independentemente juntamente com o erro cometido. Entretanto, em grandes números de experimentos, onde muitas inferências são feitas, existe, em média, maiores chances de erros, do que em números pequenos de experimentos. No segundo caso, a razão está relacionada com a validade simultânea do conjunto geral de

inferências que forem realizadas.

O'Neill & Wetherill (1971), usam dois critérios para comparação dos procedimentos:

- i) A probabilidade que uma família contenha pelo menos um erro de inferência;
- ii) O número esperado de erro de inferência, dentro de uma família.

Com três possibilidades de erros:

a) Por Comparação:

$$\frac{\text{Número de falsas inferências}}{\text{Número de inferências}}$$

b) Por Experimento:

- i) $\frac{\text{Nº de experimentos com pelo menos uma inferência falsa}}{\text{Número de experimentos}}$
- ii) $\frac{\text{Número de falsas inferências}}{\text{Número de experimentos}}$

A escolha da definição de razão de erro, pode ser, mais decisivo na opção por um determinado procedimento, que uma prevista significância quando na verdade não é.

2.3 - PODER

O poder de um teste é usualmente definido como a probabilidade que o teste tem em rejeitar a hipótese nula. Basicamente, quando estudamos qualquer tipo de erro, estamos também associando a este estudo o poder que um determinado teste tem, em detectar possíveis diferenças entre tratamentos. Em particular, a razão de erro do tipo II, definida como a probabilidade

de falsamente aceitar a hipótese nula, compara a sensibilidade que cada procedimento tem em detectar grandes diferenças. Alguns autores preferem usar um terceiro tipo de erro, que é definido como a probabilidade que dois parâmetros populacionais sejam declarados significativamente diferentes, quando de fato são, mas numa outra direção.

Ramsey (1978) estudou o poder de alguns testes, baseado em procedimentos simultâneos e passo-a-passo, considerando a definição de poder alternativo como a probabilidade de detectar pelo menos uma verdadeira diferença entre pares de médias. Ele considera que se diferenças entre populações são declaradas significativas, mas nenhuma corresponde a verdadeira diferença entre as populações (supondo a existência da verdadeira diferença), então a convencional definição é a mais poderosa do que em pares. Ele ainda avalia a sensibilidade relativa dos PCM's em detectar pequenas diferenças de médias, resumindo a definição para todos os subconjuntos de pares de médias, onde define poder como probabilidade de detectar todas as diferenças, entre pares de médias. Ele também apresenta resultados para o poder, não apenas para os pares, mas também para trios, quádruplos e quintuplos de médias. Ele considera a probabilidade de rejeição para qualquer dos pares diferentes no experimento, enquanto que Einott & Gabriel (1975) consideram a probabilidade de rejeição para um dado par.

Welsch (1977) compara o poder de alguns testes controlando a razão de erro do tipo I. Ele usa a complementação da razão de erro tipo II, segundo Cramer & Swanson (1973), para as verdadeiras diferenças significativas, isto é, $1 - \text{razão de erro do tipo II}$.

Em termos não paramétricos, Campbell & Killings (1985) discutiu alguns PCM's para comparação de k tratamentos ($k \geq 2$), com particular ênfase ao procedimento passo-a-passo. Comparam um novo procedimento proposto por eles, juntamente com o de Einott & Gabriel, ambos passo-a-passo, e também alguns convencionais.

A preocupação deles está no fato de que pesquisas relacionadas ao procedimento proposto por Nemenyi (1963) falha ao controlar a razão de erro do tipo I a um nível α , particularmente, as situações onde, por exemplo, se tenha uma população diferente e outras $(k-1)$ populações iguais, sempre que a variação estiver em termos de locação [ver Ouder Voshaar (1980)]. No artigo, ainda aparece outro exemplo, onde o mesmo procedimento de Nemenyi (1963) é analisado segundo a sensibilidade do poder em detectar pequenas diferenças com distribuições de conjunto de postos. Neste caso, eles procuravam ilustrar uma possível falha do procedimento, cuja dependência entre postos pode causar baixo poder.

Através de simulações de Monte Carlo, eles também estudaram o poder juntamente com a razão de erro tipo I em PCM's não paramétricos. Consideram o desvio populacional como parâmetro, o qual segue uma distribuição uniforme com tamanhos amostrais iguais. Apresentam as vantagens proporcionadas para cada método comparando os três procedimentos: o convencional, Einott & Gabriel (1975) e o deles mesmo. No convencional eles dividiram em dois grupos, considerando a distribuição conjunta e separada dos postos, proposta por Nemenyi (1963), Steel (1960, 1961) e Daws (1960), respectivamente. Esses dois últimos procedimentos foram discutidos em Miller (1981) e aparecem no Capítulo I

dessa monografia. Nos procedimentos de Einott-Gabriel eles usam a teoria básica dos dois procedimentos acompanhado de um terceiro bem conhecido como de Kruskal-Wallis. Para o procedimento proposto por eles, são usados passos similares ao procedimento de Newman-Keuls, com algumas modificações, já que, esse procedimento controla a razão de erro do tipo I.

2.4 - EFICIÊNCIA RELATIVA

Outra forma bastante usada na comparação dos procedimentos é a eficiência relativa que um determinado teste tem em relação a outro. Esta é uma forma, mais conveniente, e também mais sugestiva, pois, trata o problema em termos relativos.

A definição da eficiência relativa que um determinado teste T tem em relação a um outro T^* , é colocada como razão de tamanhos amostrais necessários para realizar o mesmo desempenho, ou seja, para um fixado erro do tipo I e II comum aos dois testes, a eficiência relativa é definida pela razão n^*/n , onde sob mesmas suposições n é o número de observações requerido para T^* realizar o mesmo desempenho que T . Este valor pode ser maior ou menor que 1, de acordo com o número de observações necessárias para T^* realizar o mesmo desempenho de T . Se esse valor se aproxima de 1, dizemos que a eficiência relativa dos dois testes são iguais.

Este conceito foi inicialmente introduzido por Pitman (1948). Entretanto, esta definição propõe diversos problemas, como por exemplo, o fato de que n^* não é uma variável contínua, pois, particularmente não existe um número inteiro n^* , onde o desempenho de T^* coincida exatamente com o desempenho de T . Neste tipo de comparação, em geral, são necessários estudos deta

lhados da função poder (comumente chamada eficiência do poder), causando grandes complicações no seu desenvolvimento. Pratt & Gibbons (1981) coloca como o maior problema, a especificação de "realizar o mesmo desempenho". Cada interpretação produz seu particular n^* e conseqüentemente sua eficiência relativa.

Puri & Sen (1971) discutiram largamente a eficiência relativa assintótica de Pitman (1948). As descrições consideram as exposições apresentadas, de modo generalizado em Noether (1955) e Hoettding & Roseblatt (1955), juntamente com medidas propostas recentemente. Além de fazerem relevâncias ao problema da fixação do nível de significância, eles compararam esta medida com o grau de associação dos testes.

Pratt & Gibbons (1981) definiram eficiência relativa assintótica de dois procedimentos como o limite das eficiências relativas, com o tamanho amostral tendendo ao infinito. Sem tomar essa definição como definitiva, eles deixam claro que definições alternativas podem ser consideradas em outros casos possíveis.

Através da distribuição *k-generalizada*, descrita sob a densidade:

$$f(y) = V_{(\sigma, k)} \exp\left\{-\left|\frac{y-\theta}{\sigma}\right|^k\right\}, \quad (2.4.1)$$

onde $V_{(\sigma, k)} = [2\Gamma(1+1/k)\sigma]^{-1}$, y um número real, $\sigma > 0$, $k > 0$ e θ um fixado número real, Pratt & Gibbons (1981) ainda discutiram a eficácia e eficiência relativa, assintoticamente para alguns testes paramétricos e não paramétricos da locação do parâmetro θ , generalizando a eficiência de cada teste, de acordo com a variação da medida de curtose, neste caso representada por k .

Após definir a eficiência relativa assintótica por:

$$e_{T, T^*} = \lim_{n \rightarrow \infty} \frac{n^*}{n}, \quad (2.4.2)$$

a eficácia assintótica foi convenientemente definida para cada procedimento usado, dependendo do parâmetro a ser estimado, média, mediana, etc. Os resultados desses estudos aparecem comentados também em Negrillo (1985), de modo generalizado.

A eficiência de cada teste é discutida de acordo com a medida de curtose, onde esta medida determina o escore a ser usado, tornando assim o procedimento mais eficiente.

CAPÍTULO III

3. METODOLOGIA DO TRABALHO

Os procedimentos não paramétricos usados aqui, não apresentam bom desempenho, quando nas observações amostrais, a probabilidade de se obter empates é alta, ou seja, quando os valores de cada observação se repete seguidamente. Um exemplo comum, acontece quando a população básica se distribui segundo uma geométrica ou ainda uma binomial com um número (N) pequeno de provas de Bernoulli. Sempre que a probabilidade de sucesso (ou fracasso), nessas distribuições se afasta de $1/2$.

Através do processo de Monte Carlo, foram realizadas 2.000 simulações independentes, segundo rotinas apresentadas em Goelzer (1970), escritas em FORTRAN-VAX, para as seguintes distribuições: POISSON com ($\lambda = 0.1, 1, 30$); NORMAL ($0, 1$) com percentagens de contaminação ($p = .0, .10, .15$) da NORMAL ($0, 25$); EXPONENCIAL DUPLA com variância igual a 1 e valores absolutos das médias ($0, 1, 3$); BINOMIAL com $n=1$ e $n=20$ com probabilidades de sucessos ($p = .01, .2, .5$); GEOMÉTRICA com as mesmas probabilidades de sucessos da BINOMIAL; UNIFORME ($3, 6$), EXPONENCIAL (1) e GAMA ($1, 20$). Discutimos ainda, alguns Procedimentos de Comparações Múltiplas (PCM) não paramétricos, juntamente com o teste t (Dunnett) e o teste de Tukey-Kramer.

Dividimos nossos resultados em duas partes, onde comparamos inicialmente, um controle contra vários tratamentos e a seguir comparamos todos os tratamentos entre si.

Quando as comparações envolvem um controle, tomamos $k=4$ com a população $i=0$ representando o controle e as demais populações como os quatro tratamentos. Se as comparações não envolvem o controle, desaparece a população $i=0$, e as outras $k=5$ populações representam os tratamentos.

Os tamanhos amostrais $(n_i; i=0,1,\dots,k)$, no primeiro caso, é tomada inicialmente como iguais em todos os tratamentos, inclusive o controle ($m = n_0 = 20$) ($n_i = n = 20$). A seguir fixamos o tamanho amostral dos tratamentos em 10 observações e mantendo o tamanho amostral do controle em 20 observações, para obedecer a proposição $m = 2n$. Finalmente, variamos o tamanho amostral de todos os tratamentos n_i , decrescendo desde 10 até 7, respectivamente para cada tratamento, mantendo também o número de 20 observações no controle.

Quando não consideramos o controle nas comparações, tomamos inicialmente 20 observações em cada tratamento. A seguir, variamos os tamanhos amostrais de modo a obedecer a estrutura dos $n_{i,s}$ como $(20, 10, 9, 8, 7)$, respectivamente para cada tratamento ($i; i=1,\dots,5$).

Cada tipo de comparação foi realizado, segundo um valor esperado α para o nível de significância de 1% e 5%. Nossas conclusões se baseiam em intervalos, construídos a partir dos resultados, que contenham o valor esperado. Aqui, a cada resultado conseguido para estimar o nível de significância, associamos ao centro do intervalo. Isto é, para $\alpha = 0.01$, subtrairemos e adicionaremos a quantidade 0.00436 ao valor estimado, e para $\alpha = 0.05$ a quantidade 0.00955, e assim construiremos um intervalo de 95% de confiança para cada particular procedimento. Sob a hipótese nula, supomos que a cada particular decisão, dentre

as 2.000 simuladas, são provas independentes de Bernoulli, individualmente, com $p = 0.01, 0.05$ respectivamente. A soma, portanto, se distribui segundo uma binomial com $n = 2.000$ e mesmas respectivas probabilidades (independentes) de sucessos. Tabelas com os nossos resultados são apresentadas no Anexo.

3.1 - TESTE DOS POSTOS (STEEL)

3.1.1 - Comparação - Um Controle Contra Vários Tratamentos

Existe uma dificuldade na aproximação amostral do ponto crítico no teste dos postos propostos por Steel (1959), devido ao fato de que a distribuição do máximo valor (absoluto ou não) de k variáveis aleatórias normais igualmente correlacionadas $|m|_{(k, \alpha, \rho)}$ ou $m_{(k, \alpha, \rho)}$, não se encontra tabelada para alguns valores de ρ . Gupta (1961) construiu tabelas de $m_{(k, \alpha, \rho)}$ com $k = 1(1)12$ e $\rho = .125(.125).875$, $\rho = 1/3, 2/3$ e para tabelas $|m|_{(k, \alpha, 1/2)}$ com $k = 1(1)12, 15, 20$. O cálculo do ponto, somente é correntemente possível, quando os tamanhos amostrais dos tratamentos são iguais. Neste caso, o coeficiente de correlação $\rho = \frac{n}{(n+m+1)}$, tende monotonicamente para $1/2$ (partindo de $1/3$).

Steel tabelou valores críticos para r^α usando a aproximação dada em (1.4.32), tabelas para $\alpha = .01, .05$ são apresentadas em Miller (1981) para $k = 2(1)9, n = 4(1)20$. Ele classificou como boas as aproximações para k e n ambos pequenos. Nosso estudo usou esta aproximação no caso balanceado, e aproximou $|m|_{(k, \alpha, 10/31)}$, por $|m|_{(k, \alpha, 1/2)}$ no caso $m = 2n = 20$.

O interesse do nosso estudo, está voltado para o fato de que, na prática, nem sempre é possível assegurar tamanhos amostrais iguais em todos os tratamentos. Usando o último

procedimento ($m = 2n$), podemos aplicar a desigualdade de Bonferroni, em cada comparação isolada de um tratamento contra o controle, através de k comparações independentes. Consideramos " E_i " como sendo o evento que rejeita a i -ésima hipótese nula, quando de fato ela é verdadeira, temos que, para quaisquer eventos E_1, E_2, \dots, E_k :

$$p\left(\bigcup_{i=1}^k E_i\right) = p(E_1 \text{ ou } E_2 \text{ ou } \dots \text{ ou } E_k) \leq \sum_{i=1}^k p(E_i)$$

onde $H_0: \theta_i - \theta_0 = 0$. Aqui escolhemos $\alpha_i = \alpha/k$, onde α_i é a significância esperada para cada comparação individual e α é a significância esperada do teste, de modo a se obter $\alpha = (\alpha_1 + \alpha_2 + \dots + \alpha_k)$. Entretanto, não existe aqui, a influência de um terceiro tratamento nas comparações individuais, desse modo, a distribuição $|m|_{(1, \alpha, \rho)}$ é encarada como uma normal padrão em (1.4.32).

Sempre que consideramos distribuições com baixa probabilidade de empates entre as observações, os resultados, em geral, para o caso balanceado parecem satisfatórios. Nos dois últimos casos, apresentam tendências conservativas, já que os pontos críticos envolvem aproximações para grandes amostras, além da própria definição, no caso da desigualdade de Bonferroni.

3.1.2 - Comparação - Tratamento Contra Tratamento-(Steel-Dwass)

O problema desse procedimento é análogo ao da seção anterior. O ponto crítico somente é correntemente possível, quando os tamanhos amostrais são iguais. Aqui, a única hipótese testada é que as populações não diferem.

O problema da distribuição dos postos entre tratamentos é semelhante ao teste studentizado de Tukey-Kramer, exceto

que, o teste dos postos não envolve a suposição que as observações se distribuam segundo uma normal.

A modificação proposta nesse estudo, se confunde com a empregada na seção anterior, encarando a possibilidade de se perder informações. O procedimento não balanceado sai um pouco das características, para o caso de iguais tamanhos amostrais, pois, não utiliza a distribuição da amplitude studentizada, nas aproximações dos pontos críticos. O processo usado aqui, aplica a desigualdade de Bonferroni, em comparações individuais entre os $k(k-1)/2$ pares de tratamentos. Consideramos o procedimento da seção anterior, onde o tratamento do par que tiver maior tamanho amostral, é tomado como se estivesse sob controle. Com $k(k-1)/2$ comparações separadas de pares de tratamentos, os individuais esperados α_i ($i=1, \dots, k(k-1)/2$), são escolhidos de modo a satisfazer a equação:

$$\alpha^* \leq \alpha_{1*} + \dots + \alpha_{[k(k-1)/2]*} \quad (3.1.1)$$

onde α^* e α_{i*} são as desejadas significâncias para os testes individuais e geral, respectivamente. A aproximação do ponto crítico, envolve, inicialmente, substituir $q(k, \alpha, \infty)$ por $|m|_{(1, \alpha, \rho)}$, já que, o método é descrito para a comparação de um controle e um tratamento. E assim, a estatística $|m|_{(1, \alpha, \rho)}$ é substituída pela normal $(0, 1)$, no caso $k=1$, na equação (1.4.31). Decidindo-se rejeitar a i -ésima hipótese se qualquer R^{**} for maior ou igual a:

$$\frac{n(n+m+1)}{2} + \frac{1}{2} + Z_{(\alpha_i)} \sqrt{\frac{n+m(m+n+1)}{12}} \quad (3.1.2)$$

onde $\alpha_i = \alpha/k(k-1)$ e $Z_{(\alpha)}$ é o percentil α superior da distribuição normal padrão, e 'm' e 'n' são os números de observações do

maior e menor tamanhos amostrais, respectivamente.

Nossos resultados para o caso balanceado parece confirmar as proposições dos autores. Onde o conservativismo aparece ante a possibilidade crescente de acontecer empates entre as observações, além de aparecer, de modo menos claro, nos outros casos.

Quando as nossas modificações, acreditamos que pesquisas nesse sentido, ainda se fazem necessárias, pois, os níveis de significâncias estimados parecem seguir a forte tendência conservativa, mesmo nos casos onde a distribuição básica é normal padrão.

3.2 - TESTE DE KRUSKAL-WALLIS - (Nemenyí)

Os procedimentos apresentados nesta seção, não apresentam grandes diferenças, quanto ao problema da comparação de tratamentos, incluindo ou não um controle. No caso, onde se inclui o controle, ele é tomado como a população amostral $i=0$, e as restantes, como tratamentos a serem testados contra ele. Caso contrário, esta população amostral $i=0$ é eliminada e as comparações são feitas dentre pares de tratamentos.

O procedimento exato é apresentado apenas, se experimento for balanceado com um número de tratamentos (k) reduzidos. Na presença do controle, a estatística se reduz ao procedimento de Wilcoxon para $k=1$. O mesmo acontece se as comparações não incluem o controle, no caso $k=2$. Tabelas para valores críticos exatos, para os dois procedimentos, são apresentadas em Hollander-Wolfe (1973), com alguns valores de n (número de observações nos tratamentos) e $k=3$.

Nosso estudo baseou-se em aproximações para grandes

amostras. Através da equação (1.4.61) sugerida por Miller (1981), com coeficiente de correlação dado por (1.4.60), aproximou os pontos críticos quando as comparações envolvam o controle, a dificuldade na aproximação desses pontos críticos, também se apresenta aqui, devido ao fato que a estatística $|m|_{(k, \alpha, \rho)}$, não ter sido avaliada quando $\rho \neq 1/2$. A aproximação, entretanto, usou $\rho = 1/2$ para os dois casos onde os tamanhos dos tratamentos são iguais, diferindo ou não do tamanho amostral do controle. Ainda com relação a comparação envolvendo o controle, Dunn (1964) propôs, substituir $|m|_{(k, \alpha, \rho)}$ por $Z_{(\alpha/2k)}$, quando os tamanhos amostrais dos tratamentos são diferentes. Esta aproximação também, foi usada no nosso estudo, decidindo-se rejeitar a hipótese nula, se pelo menos uma diferença:

$$|\bar{R}_i - \bar{R}_0| \geq Z_{(\alpha_i)} \left| \frac{N(N+1)}{12} \right|^{1/2} \left(\frac{1}{n_i} + \frac{1}{n_0} \right)^{1/2} \quad (3.2.1)$$

acontecer onde $\alpha_i = \alpha/2k$, $N = n_0 + n_1 + \dots + n_k$ e $Z_{(\alpha)}$ é o percentil α superior da distribuição normal padrão, note que α e α_i são as respectivas significâncias esperadas, de modo $\alpha \leq \alpha_1 + \dots + \alpha_k$.

Quanto ao problema da comparação, onde não se inclui o controle, nosso estudo usou a aproximação baseada em (1.4.58), também sugerida por Miller (1981). E também Dunn (1964) sugeriu substituir $q(k, \alpha, \omega)$ em (1.4.58) por $Z_{(\alpha/k(k-1))}$, onde decide-se rejeitar a hipótese nula, se acontecer qualquer uma das desigualdades, note porém, que α_i neste caso será $(\alpha_i = \alpha/k(k-1))$ e $N = n_1 + \dots + n_k$.

Consideramos razoavelmente satisfatórios, os resultados conseguidos por este método. Principalmente nos casos balanceados, com ou não a inclusão do controle nas comparações,

desde que, consideremos apenas as distribuições com baixas probabilidades de empates, entre as observações. Notamos entretanto, alguma tendência conservativa, nos casos não balanceados, a qual torna-se evidente quando excluimos o controle das comparações.

3.3 - TESTE DO SINAL - (Steel)

3.3.1 - Comparação - Um Controle Contra Vários Tratamentos

Sob a suposição que as observações são independentes e suas densidades dadas por (1.4.1), a hipótese nula ($H_0: \theta_i - \theta_0 = 0$) implica que qualquer arranjo das $k+1$ observações, em cada bloco, é igualmente possível. Os resultados das estatísticas D_{ij}^+ e D_{ij}^- , dentro de um específico bloco j , depende das observações, somente pela sua ordenação.

O cálculo dos pontos críticos usados aqui, foi baseado em aproximações, devido a distribuição exata de S_{\pm}^0 torna-se impraticável, exceto para valores pequenos de k e n . Os pontos críticos usados aqui, são aproximados por (1.4.11) com $|m|(k, \alpha, 1/2)$, segundo tabelas avaliadas por Gupta (1963). O caso não balanceado, não foi considerado, devido a necessidade das observações chegarem em blocos.

Os resultados conseguidos por este teste, apresentam tendências conservativas. Miller (1981) coloca como boas as aproximações usando apenas os casos n e k ambos pequenos, isto é, $k \leq 3$ e $n \leq 6$, pela possibilidade no cálculo exato de S_{\pm} . Pensamos que este teste poderia causar grandes desvios, pela própria proposta da aproximação contínua para um teste com cortes discretos. Para o nosso objetivo, este foi o teste, no

gênero, que apresentou maior tendência conservativa.

3.3.2 - Comparação - Tratamento Contra Tratamento - (Steel)

A estrutura deste teste, é análoga ao da seção anterior exceto que, o interesse é a comparação entre todos os tratamentos. As análises são feitas apenas quando as observações chegam em blocos, com o mesmo número de observações em cada tratamento.

A distribuição exata de S^{α} , aqui também torna-se muito complicada, exceto para um número pequeno de comparações ($k \leq 4$), com número também pequeno de blocos ($n \leq 6$).

Nossos resultados são baseados na aproximação dada em (1.4.). Eles parecem sugerir a mesma tendência conservativa dos resultados da seção anterior. Baseado em programas escritos por Howard Givner, Nemenyi (1963), elaborou tabelas exatas de S com $k=3$; $n=7(1)16$ e $\alpha = .01$ e $.05$. Ele classifica, as aproximações como boas desde que, comparadas a valores críticos de S^{α} conhecidos. Para nossos objetivos, este também foi o teste que apresentou maior tendência conservativa dentre todos os seus competidores.

3.4 - TESTES DOS POSTOS ASSINALADOS - (Nemenyi)

Este procedimento não difere na comparação de tratamentos, incluindo ou não um controle nas comparações. O cálculo do ponto crítico é somente possível no caso balanceado, pois, utiliza a estatística do sinal no seu desenvolvimento, supondo que as observações devam chegar em blocos. Mais detalhes desse teste, são apresentados em Miller (1981).

O nosso estudo usou a aproximação do ponto crítico

através das equações (1.4.31) e (1.4.39), substituindo a esperança e a variância dada em (1.4.52), respectivamente, similar ao procedimento usado no teste dos postos (Steel), onde o primeiro caso aproxima o ponto crítico pela estatística $|m|(k, \alpha, 1/2)$, e $q(k, \alpha, \infty)$ no segundo caso.

Nossos resultados também apresentam analogia nos dois tipos de comparações, sugerindo uma leve tendência conservativa em ambos os casos. Entretanto, devem figurar como bons competidores, embora a tendência conservativa seja um tanto quanto mais presente no segundo caso. Este teste ainda apresenta a maior sensibilidade, nos casos, onde os empates entre as observações são mais prováveis, isto comparado entre os procedimentos não paramétricos usados no nosso estudo. Embora, nesse caso, a tendência conservativa, ainda seja evidente.

3.5 - TESTE DE FRIEDMAN - (Nemenyi)

Este procedimento proposto e avaliado por Nemenyi (1963), também não faz grandes distinções quanto aos dois problemas de comparação. Propõe comparar as diferenças entre as médias dos postos com um ponto crítico adequado. É necessário que as observações cheguem em blocos. Para comparação envolvendo um controle, as observações Y_{ij} $\{i=0, 1, \dots, k; j=1, \dots, n\}$ devem ter mesmo tamanho amostral em cada tratamento. Analogamente no outro caso os tamanhos amostrais dos tratamentos, são supostos iguais, pois, necessitam chegar em blocos e as observações são definidas com Y_{ij} $\{i=1, \dots, k; j=1, \dots, n\}$.

Tabelas para pequenos valores de k e n são encontradas em Hollander-Wolfe (1973), para os dois casos. O nosso estudo usou a aproximação para grandes amostras dadas em (1.4.66.2)

e (1.4.68), respectivamente.

Os resultados conseguidos através deste teste, nos pareceu satisfatório, sempre considerando a baixa probabilidade de empates entre as observações, embora, exista nestes casos, evidente tendência conservativa, um pouco mais visível no caso de comparação entre pares de tratamento.

3.6 - TESTE DA ALTERNATIVA ORDENADA - (Jonckheere-Terpstra)

Este procedimento é aplicado somente na comparação entre k tratamentos, consiste no cálculo dos $k(k-1)/2$ contadores $U_{u,v}$ de Mann-Whitney onde $U_{u,v}$ é dado em (1.4.72). Aqui, existe a restrição de que a alternativa tenha uma natural ordenação. Isto limita a comparação desse teste junto aos demais.

Tabelas exatas deste teste com alguns valores de k e n , são encontradas em Hollander-Wolfe (1981). Nosso estudo baseou os resultados em aproximações de grandes amostras, que visam a normalização de J dadas em (1.4.73) por (1.4.75).

Este procedimento é preferido aos demais competidores não paramétricos, quando o investigador estiver interessado numa particular direção da hipótese nula. Se a hipótese alternativa não estiver naturalmente ordenada, os autores sugerem simplesmente rotular os tratamentos, colocando-os na ordem desejada como propõe este teste. Procedimentos do tipo Kruskal-Wallis não utilizam essa informação, pois supõem iguais probabilidades na alternativa, com os tratamentos rotulados.

Para as proposições dos autores, os resultados são considerados satisfatórios, embora os competidores não tenham sido testados sob essa conotação. Entretanto, em termos de razão de erro experimental, os resultados apresentam bom desempe

nho em quase todas as distribuições contínuas e, quando a probabilidade de empates entre as observações, é baixa.

3.7 - TESTE STUDENTIZADO - (Tukey-Kramer)

Este procedimento é aplicado somente as comparações entre tratamentos. Ele supõe que as observações se distribuam segundo uma normal, necessitando assim, uma padronização de todas as outras distribuições usadas neste estudo.

A grande utilização deste teste, está voltada ao fato de que ele é bastante sensível a tamanhos amostrais diferentes entre os tratamentos. As pesquisas neste sentido têm apresentado este fato quando as distribuições básicas são normais ou quase normais. Sendo inclusive, fortemente recomendável junto a seus competidores que usam a teoria normal.

Nosso objetivo aqui, é estudar o seu comportamento diante das situações adversas as suas suposições, através de distribuições básicas, claramente, não normais. Como também nos interessava o comportamento dos seus competidores não paramétricos em situações onde se encontra a presença da normalidade.

Os resultados conseguidos, parecem mostrar que ele continua com bom desempenho, nas situações onde a distribuição básica não seja normal. Nos casos balanceados ou não, os resultados sugerem uma boa aproximação, mesmo nas situações onde a probabilidade de empate, entre as observações, seja alta, como por exemplo, se a distribuição básica segue uma Bernoulli.

3.8 - TESTE "t" - (Dunnett)

Este procedimento para tamanhos de amostras iguais,

proposto para comparação de um controle contra vários tratamentos, supõe que as observações seguem uma distribuição normal, segundo Dunnett (1955). Como extensão deste procedimento, ele em 1964, sugeriu a avaliação numérica da expressão da integral dupla proposta por Gupta-Sobel (1957), e construiu tabelas para testar hipóteses bilaterais, quando os tamanhos amostrais dos tratamentos são menores, ainda que iguais entre si, que o tamanho amostral do controle.

Através do atual ponto crítico (quando todos os tratamentos, inclusive o controle, têm os mesmos tamanhos amostrais), ele propõe multiplicar a quantidade $(1-2\rho)/(1-\rho)$ ou $(1-n_t/n_c)$ por um fator de correção, resultando na percentagem que deverá ser acrescida ao atual ponto crítico, se o coeficiente de correlação ρ estiver entre .725 e .5.

O argumento usado no teste dos postos (Steel), para tamanhos amostrais diferentes, foi também usada aqui. Aplicamos aqui a desigualdade de Bonferroni, de modo a assegurar um nível de significância estimado ($\hat{\alpha}$), menor ou igual ao esperado (α). Formulando k hipóteses com $\hat{\alpha}_i = \alpha/k$, de que $\theta_i = \theta_0$ ($i=1, \dots, k$), cada uma individualmente, usando a aproximação pela distribuição t-student com $n_0 + n_i - 2$ graus de liberdade.

Quando a distribuição básica é normal ou quase normal o desempenho deste teste, em geral, nos pareceu satisfatório. Esse desempenho, parece manter-se em quase todas as distribuições contínuas, com algumas exceções, como por exemplo, a distribuição uniforme e exponencial, particularmente quando aplicamos a desigualdade de Bonferroni, onde se apresenta alguma tendência otimista e conservativa, respectivamente.

Quando a distribuição básica não é contínua, mas se

distribui segundo uma forma equilibrada (probabilidade de sucesso $1/2$) o teste continua apresentando bom desempenho em todos os três casos. Apenas quando esta probabilidade se afasta rapidamente de $1/2$, nota-se algumas tendências conservativas, caso por exemplo, quando a distribuição básica é binomial com $n=1$ ou geométrica, principalmente no último caso, onde o tamanho amostral varia entre os tratamentos.

C A P Í T U L O I V

4. DISCUSSÃO ASSOCIADA AOS RESULTADOS

4.1 - INTRODUÇÃO E OBJETIVO

Esta parte do estudo procura colocar, de modo simples, o objetivo do nosso trabalho. Através da distribuição Poisson, desejamos deixar claro que a alta probabilidade de empates entre as observações torna, geralmente, alguns dos testes não paramétricos conservativos. Nos dois casos apresentados (controle x tratamento e tratamento x tratamento), as conclusões dos nossos estudos são discutidas com base na razão de erro do tipo I. O nosso objetivo é questionar os procedimentos que apresentam grandes desvios no nível de significância.

4.2 - O PROBLEMA DO EMPATE COM A DISTRIBUIÇÃO DE POISSON

De acordo com o algoritmo apresentado em Goelzer (1970), simulamos a distribuição de Poisson para vários valores do parâmetro de média λ . Nossa preocupação foi determinar uma lei de formação, para um particular problema, onde a medida que a probabilidade de empates diminuisse, teríamos uma maior evidência conservativa dos testes. Aqui, simulamos a distribuição com um acréscimo do parâmetro λ a cada situação, com $\lambda = .1, .5(.5) 15(5) 30$. Os testes usados, seguem os desenvolvimentos originalmente propostos pelos autores, com mesmo tamanho amostral em todos os tratamentos (incluindo ou não o controle).

4.2.1 - Discussão Com a Presença do Controle

Os Gráficos 4.1.1 e 4.1.2, sugerem valores mais adequados para o procedimento paramétrico. Embora, algumas estimativas parece ultrapassar os nossos limites de confiança, especialmente para o valor esperado de 0.05 , esse procedimento apresenta uma maior estabilidade diante da variação do parâmetro e o problema do empate parece não afetar, em grandes proporções, o seu desempenho.

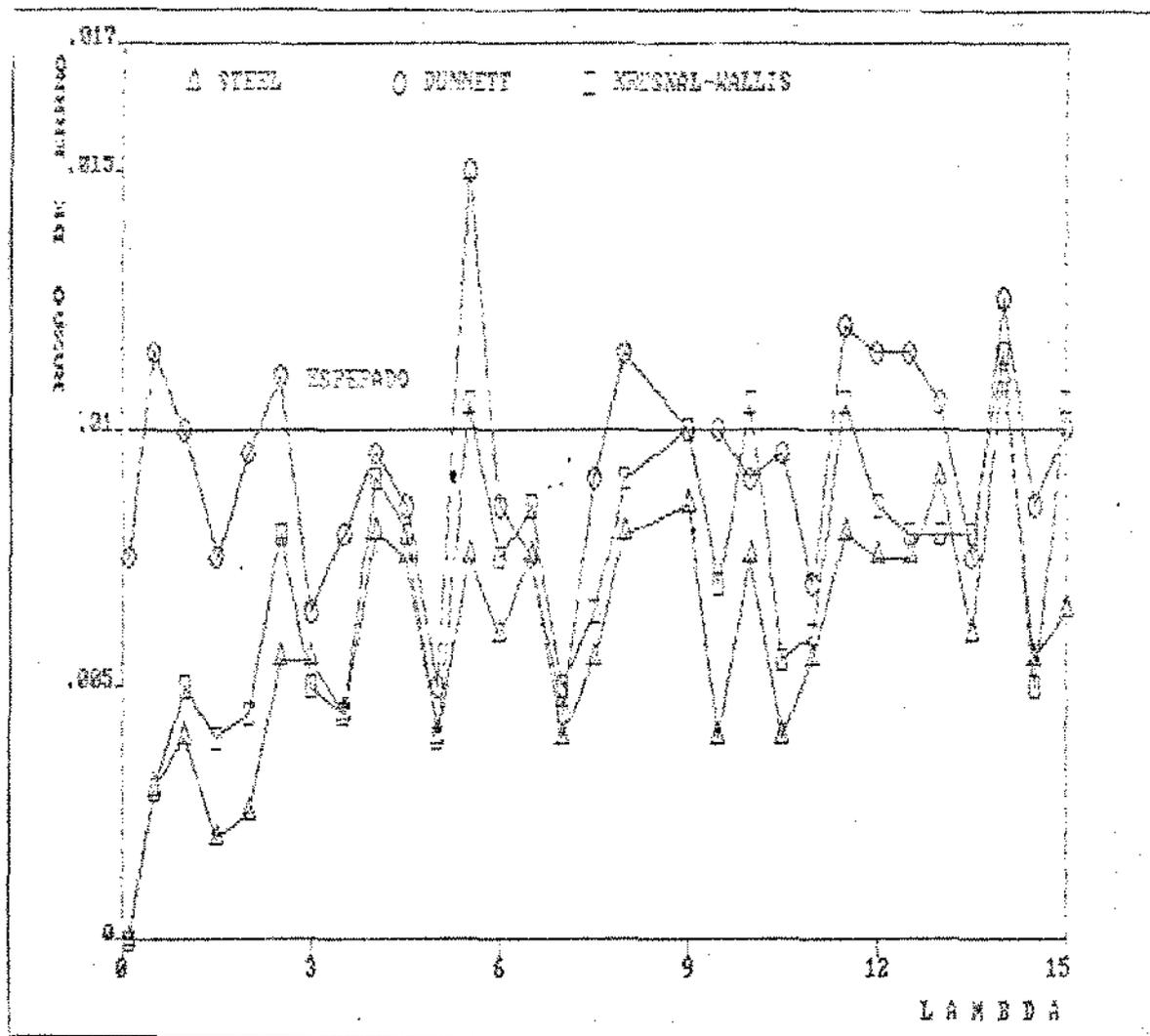
Para os procedimentos não paramétricos, notamos que os resultados só apresentam bons desempenhos a partir de $\lambda=3.0$. Aqui, dividimos os resultados em dois grupos: os que supõem que as observações cheguem em blocos e os que não necessitam dessa suposição. No primeiro grupo, o teste de Friedman parece se destacar dos outros, particularmente, quando esperamos $\alpha = 0.05$. Os dois testes, o do sinal e dos postos assinalados, parecem manter a tendência conservativa, ainda mais clara no teste do sinal. Já no segundo grupo, os resultados chegam até a se confundirem entre os dois testes. Notamos, porém, que o teste de Kruskal-Wallis é ligeiramente superior ao teste dos postos, devido a inclusão de quase todos os pontos, a partir de $\lambda = 2.5$, do valor esperado.

4.2.2 - Discussão Sem a Presença do Controle

Os Gráficos 4.2.1 e 4.2.2 parecem surgir as mesmas idéias já mencionadas na seção anterior. O procedimento paramétrico, porém, se destaca mais claramente dos demais, chegando inclusive, a apresentar uma maior tendência otimista no caso de 5% , mas mantendo o bom desempenho para o caso de 1% .

GRÁFICOS (*)

GRÁFICO 4.1.1.a - (C x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 1%.



(*) A idéia básica de cada gráfico é comparar os procedimentos quanto a razão de erro esperada. A distinção entre a) e b) tem apenas o sentido de melhorar a visualização do desempenho de cada procedimento. Portanto, o primeiro e segundo número após o algarismo 4 representa se a comparação inclui ou não um controle e se a razão de erro esperada é de 1% ou 5%, respectivamente.

GRÁFICO 4.1.1.b - (C x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 1%.

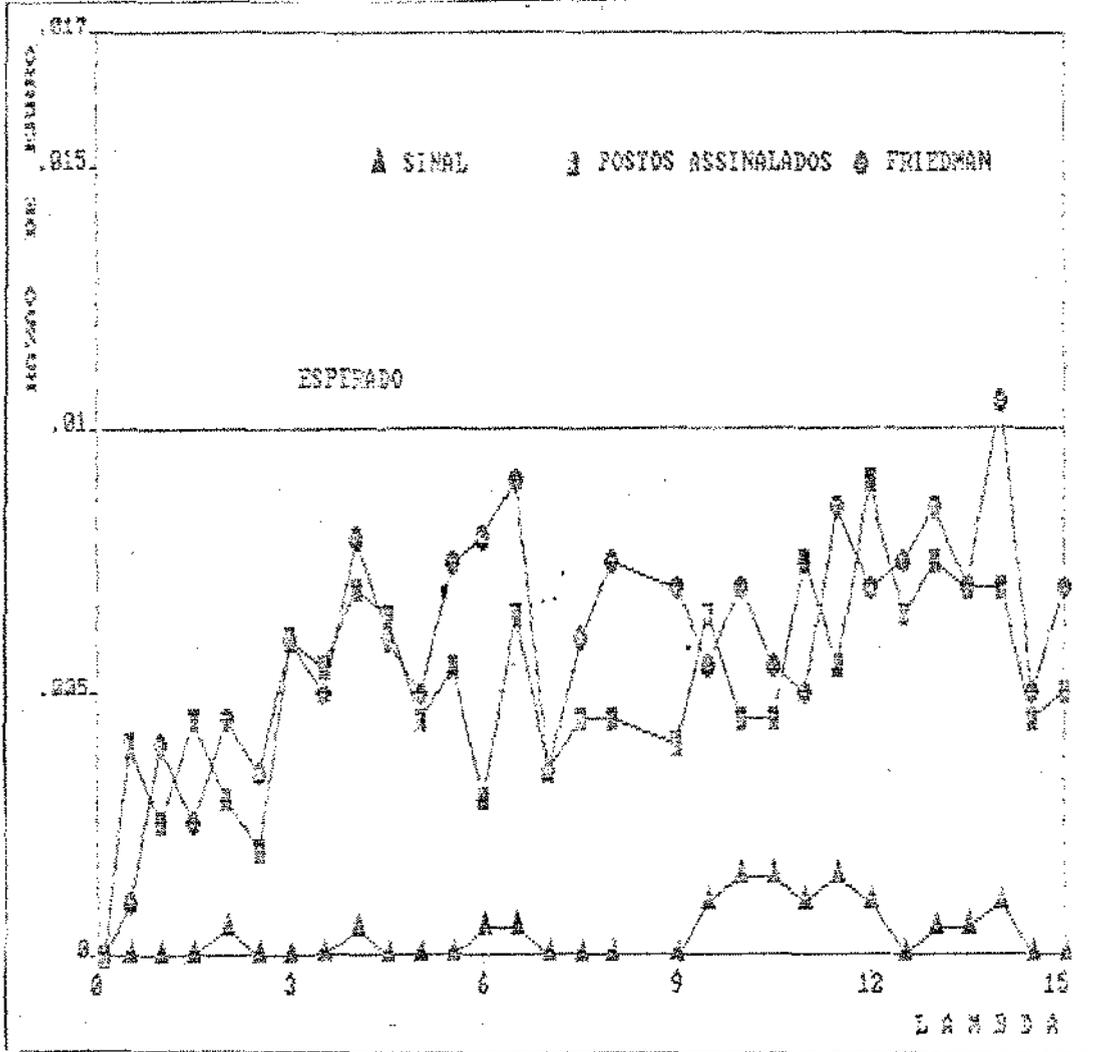


GRÁFICO 4.1.2.a - (C x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 5%.

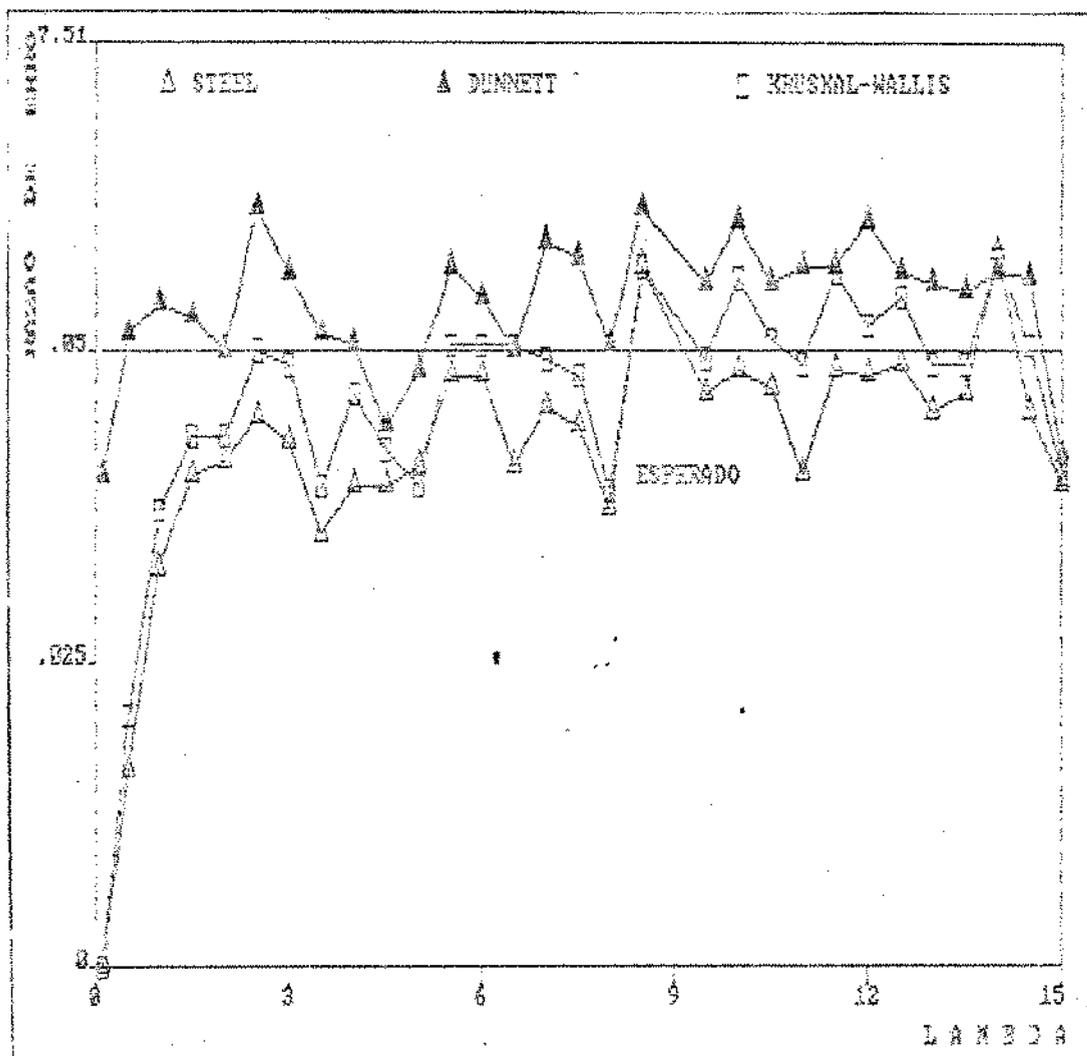


GRÁFICO 4.1.2.b - (T x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 5%.

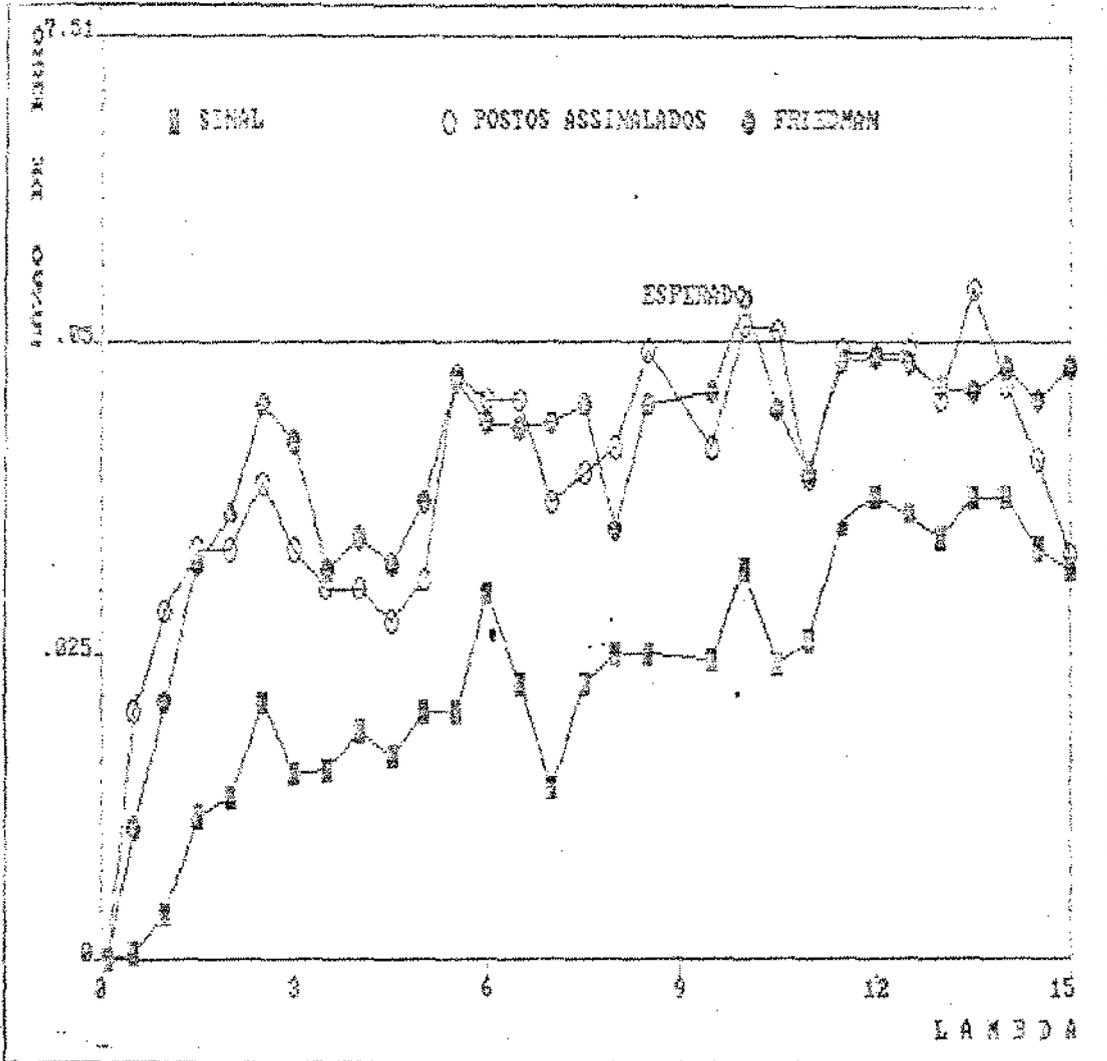


GRÁFICO 4.2.1.a - (T x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 1%.

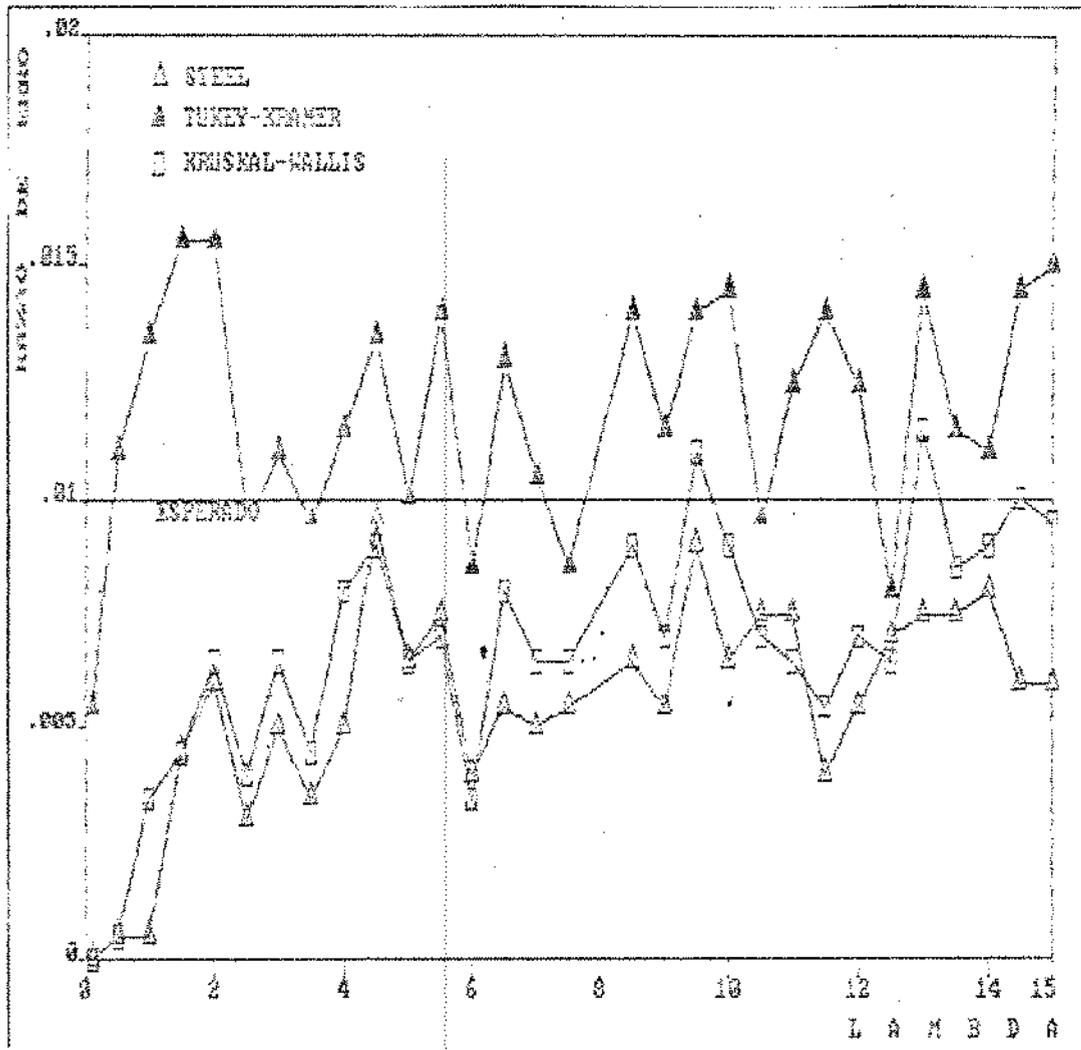


GRÁFICO 4.2.1.b - (T x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 1%.

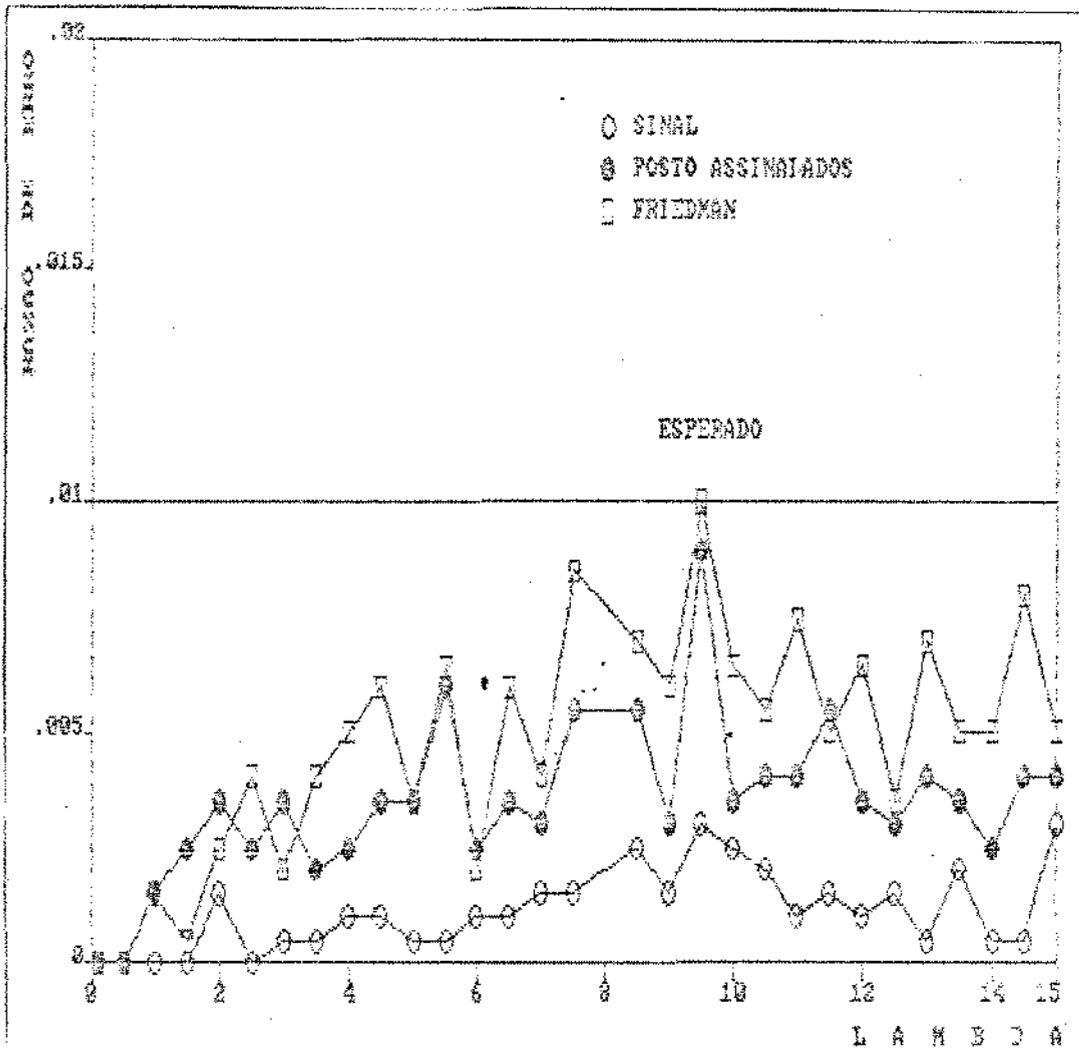


GRÁFICO 4.2.2.a - (T x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 5%.

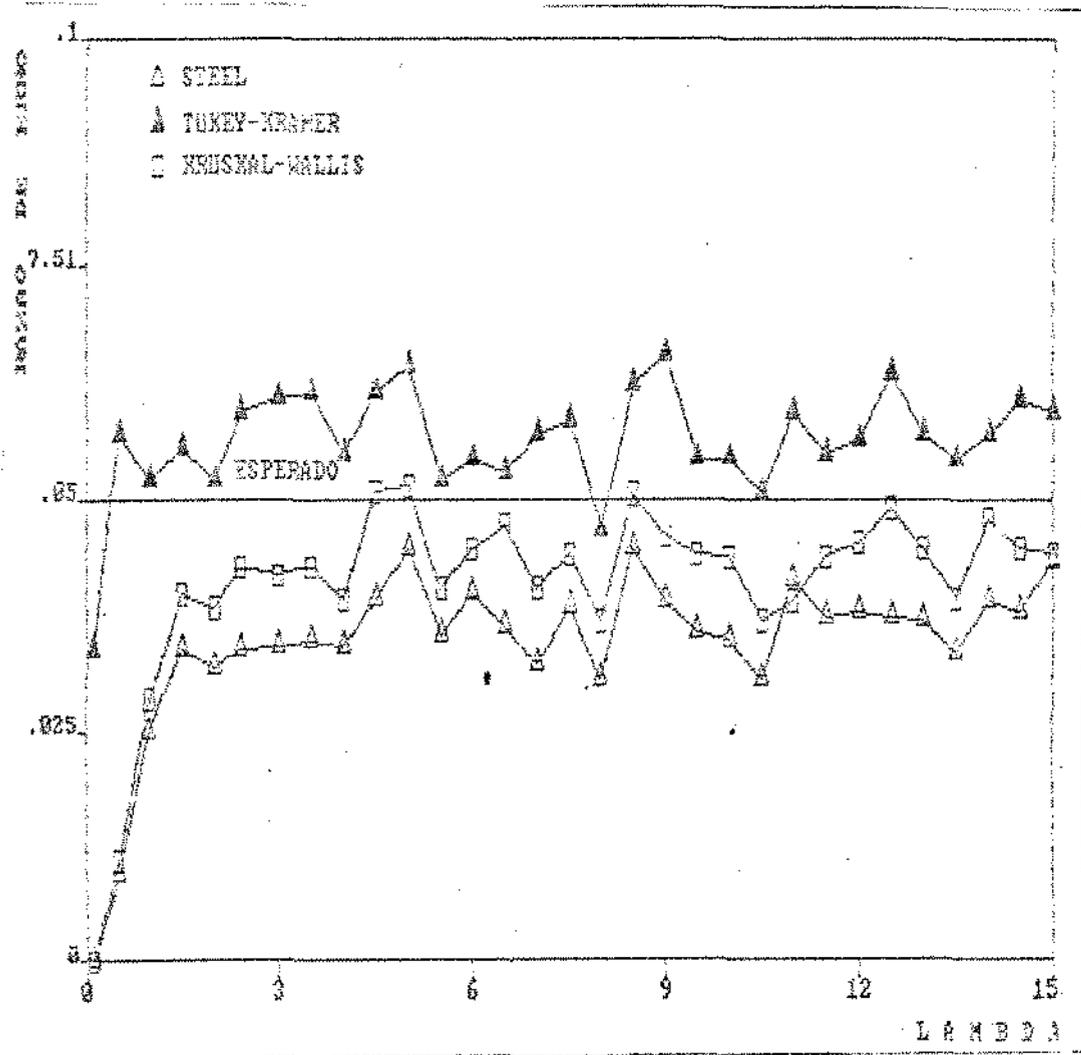
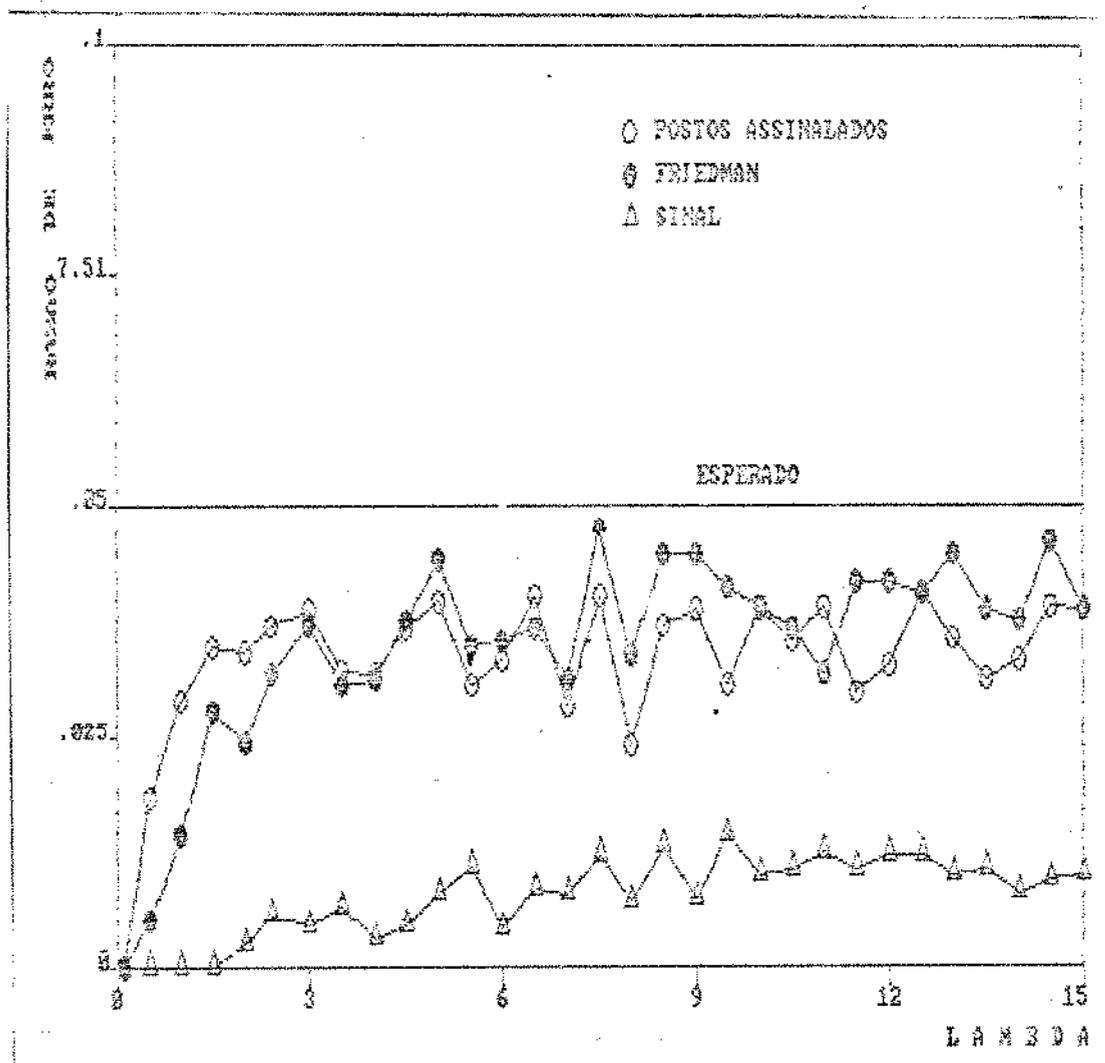


GRÁFICO 4.2.2.b - (T x T) SIGNIFICÂNCIA ESPERADA x LAMBDA (POISSON) 5%.



Quanto aos procedimentos não paramétricos, os gráficos também sugerem os mesmos desempenhos para os dois casos. Para o primeiro grupo, parece não ficar muito claro o melhor desempenho do teste de Friedman em relação ao teste dos postos assinalados, embora se perceba algumas diferenças. Apenas para valores de alto λ (> 4), ele apresenta estimativas que cobrem os valores esperados, o que parece não ocorrer com o teste dos postos assinalados. No segundo grupo, o teste de Kruskal Wallis, mantém a vantagem ao teste dos postos, mostrando inclusive bons resultados para pequenos valores de λ ($= 2$).

4.3 - COMENTÁRIOS E CONCLUSÕES

De modo geral, nosso estudo mostra que os procedimentos não paramétricos apresentam maiores desvios em relação aos níveis de significância propostos. Entretanto, na utilização desses procedimentos se faz necessário um estudo mais detalhado da distribuição das observações. Nossas conclusões, consideram apenas a razão de erro do tipo I. Outras formas de comparação têm sido apresentadas, seja por meio do poder, eficiência relativa, robustez, etc.

Dividimos nossas conclusões e discussões em duas partes: inicialmente a comparação envolvendo um tratamento sob controle, a seguir a comparação de todos os tratamentos entre si.

4.3.1 - Comparação: Um Controle Contra Vários Tratamentos

Particularmente, se a probabilidade de empates entre as observações é baixa, os testes não paramétricos apresentam, em geral, bons desempenhos. Com exceção do teste do sinal, os intervalos de 95% de confiança, construídos a partir das alter

nativas, conseguem cobrir os valores esperados para α (nível de significância proposto de 1% e 5%). Todavia, se a probabilidade de dos empates aumenta, a exemplo das distribuições discretas, quando a probabilidade de sucesso p se afasta de $1/2$, esses desempenhos são considerados insatisfatórios.

O teste t (Dunnett), parece não sentir essa variação, exceto nos extremos (distribuição Geométrica ou Bernoulli com $p = 0.01$). E, nessas situações, mesmo sendo considerado conservativo, ainda apresenta melhor desempenho.

A aplicação de desigualdade de Bonferroni, em dois testes, teve como objetivo uma generalização para tamanhos amostrais diferentes. Acreditamos porém, que a utilização dessa ferramenta não contribuiu, de modo evidente, em ganhos reais. No caso do teste t , os resultados parecem não manter a estabilidade das estimativas, no caso do teste ser realizado sob balanceamento (ainda que o controle tenha o dobro de observações). Mesmo assim, pode ser considerado como um bom competidor paramétrico ao teste de Kruskal-Wallis. Já no caso do teste dos postos, as estimativas parecem ficar sempre abaixo dos valores esperados para α , inclusive dos níveis estimados pelo teste de Kruskal-Wallis.

Quanto aos testes que supõem que as observações devam chegar em blocos, não conseguimos notar diferenças entre os rendimentos dos testes de Friedman e o dos postos assinalados. Apesar desses testes não apresentarem resultados satisfatórios, observamos neles menores desvios que no teste do sinal. Particularmente a Seção 4.2 indica um melhor desempenho do teste de Friedman em relação ao teste dos postos assinalados, quando tratamos apenas da distribuição de Poisson. Isto nem sempre pare

ce ocorrer, principalmente, quando a distribuição for discreta, tornando difícil nós determinarmos qual deles apresenta melhor desempenho.

4.3.2 - Comparação: Tratamento Contra Tratamento

De maneira geral, a linha de conclusão quando comparamos um controle contra vários tratamentos, não sofre modificação ao compararmos os tratamentos entre si. Alguns procedimentos, caso haja baixa probabilidade de empates entre as observações, se tornam conservativos a medida que essa probabilidade aumenta. A presença ou não de empates (a menor de casos extremos), também não afeta os resultados para o procedimento paramétrico de Tukey-Kramer.

Entretanto, outras conclusões podem ser tomadas em relação, principalmente, ao teste paramétrico, pois, os resultados indicam estimativas mais adequadas, mesmo no caso de tamanhos amostrais diferentes.

Outro resultado importante aparece nos testes que usam a suposição que as observações devam chegar em blocos. Aqui, o teste de Friedman apresenta melhor desempenho que o teste dos postos assinalados, embora no caso em que a distribuição básica seja discreta (maior probabilidade de empates), o teste dos postos assinalados leva alguma vantagem. Ambos ainda mantêm melhor rendimento que o teste do sinal.

Quanto à modificação, proposta apenas ao teste dos postos, já que o procedimento paramétrico tem definição para tamanhos amostrais diferentes, acreditamos que pesquisas nesse sentido devam ser feitas. Pois, apesar do teste de Kruskal-Wallis apresentar melhor desempenho, ainda assim, consideramos

ambos como conservativos.

Finalmente, o teste de Jonckheere-Terpstra mostra, apesar de não ter as mesmas alternativas dos demais, melhor de sempenho comparado ao teste de Kruskal-Wallis. Devendo ser preferido, caso o pesquisador esteja interessado numa particular direção parâmetro.

REFERÊNCIAS BIBLIOGRÁFICAS

- BOYETT, J.M. and SHUSTER, J.J. - Nonparametric one-side tests in multivariate analysis with medical applications. *J. Am. Statist. Assoc.* 72: 665-668, 1977.
- BOYETT, J.M. and SHUSTER, J.J. - Nonparametric multiple comparison procedures. *J. Am. Statist. Assoc.* 74: 379-382, 1979.
- CAMPBELL, G. and SKILLINGS, J.H. - Nonparametric stepwise multiple comparison procedures. *J. Am. Statist. Assoc.* 80: 998-1003, 1985.
- CARMER, S.G. and SWANSON, M.R. - An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. *J. Am. Statist. Assoc.* 68: 66-74, 1973.
- DUNCAN, D.B. - Multiple-range and multiple-F tests. *Biometrics* 11: 1-42, 1955.
- DUNCAN, D.B. - An bayesian approach to multiple comparisons. *Technometrics* 7: 171-222, 1965.
- DUNN, O.J. - Multiple comparisons usine rank sum. *Technometrics* 6: 241-252, 1964.
- DUNNETT, C.W. - A multiple comparison procedure for comparing several treatments with a control. *J. Am. Statist. Assoc.* 50: 1096-1121, 1955.

DUNNETT, C.W. - Query: Multiple comparisons between treatments and control. *Biometrics* 17: 324-326, 1961.

DUNNETT, C.W. - New table for multiple comparisons with control. *Biometrics* 20: 482-491, 1964.

DUNNETT, C.W. - Pairwise multiple comparisons in the homogeneous variance, unequal sample size case. *J. Am. Statist. Assoc.* 75: 789-795, 1980.

DWASS, M. - Some k - sample rank-order tests. In: OLKIN, I. *et al*, contributions to probability and statistics, (eds.) Stanford University Press, Stanford, 1960.

EINOT, I. and GABRIEL, K.R. - A study of the powers of several methods of multiple comparison. *J. Am. Statist. Assoc.* 70: 547-583, 1975.

GOELZER, L. - Técnicas de Simulação, EAESP da FGV, São Paulo, 1970.

GUPTA, S.S. - Probability integrals of multivariate normal and multivariate t. *Ann. Math. Statist.* 34: 792-828, 1963.

GUPTA, S.S. and SOBEL, M. - On a statistic which arises in selection and ranking problems. *Ann. Math. Statist.* 28: 957-967, 1957.

HOEFFDING, N. and ROSENBLATT, J.R. - The efficiency of test. *Ann. Math. Statist.* 26: 52-63, 1955.

HOLLANDER, M. - Rank tests for randomized blocks when the, *al*

ternatives have an a prior ordering, 1965, apud MILLER Jr., R.G., Simultaneous statistical inference, (eds.) McGraw-Hill, New York, (2 rd. ed.), 1981, p. 129-188.

HOLLANDER, M. - An assymptotically distribution free multiple comparison procedure - treatment versus control. *Ann. Math. Statist.* 37: 735-738, 1966.

HOLLANDER, M. and WOLFE, D.A. - Nonparametric statistical methods, (eds.), John Wiley, New York, 1973, 503 p.

JONCKHEERE, A.R. - A distribution-free k-sample test against ordered alternatives. *Biometrika* 41: 133-145, 1954.

KEUS, M. - The use of the studentized range in connection with an analysis of variance. *Euphytica* 1: 112-122, 1952.

KRAMER, A. - Extensions of multiple range tests to group means with unequal numbers of replication. *Biometrics* 12: 307-310, 1956.

KRUSKAL, W.H. - A nonparametric test of the several sample problem. *Ann. Math. Statist.* 23: 525-540, 1952.

KRUSKAL, W.H. and WALLIS, W.A. - Use of ranks in one-criterion variance analysis. *J. Am. Statist. Assoc.* 47, 583-621, 1952.

LEHMANN, E.L. - Assimptotica nonparametric inference in some linear model with one observation per cell. *Ann. Math. Statist.* 35: 726-734, 1964.

- MANN, H.B. and WHITNEY, D.R. - On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* 18: 50-60, 1947.
- McDONALD, B.J. and THOMPSON, W.A. - Rank sum multiples comparisons in one-and two-way classifications. *Biometrika* 54: 487-497, 1967.
- MILLER Jr., R.G. - Simultaneous statistical inference, (eds.) McGraw-Hill, New York, (2rd.ed.), 1981, p. 129-188.
- NEAVE, H.R. and GRANGER, C.W.J. - A Monte Carlo study comparing various two sample tests for differences in mean. *Technometric* 10: 509-522, 1968.
- NEGRILLO, B.G. - Métodos não paramétricos uni e multivariados, Simpósio de Estatística Aplicada à Experimentação Agronômica, 1ª e Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria, 30ª de 22 a 26 de julho de 1985, Piracicaba, 82 p.
- NEMENYI, P. - Distribution-free multiple comparisons, 1963, apud MILLER Jr., R.G., Simultaneous statistical inference, (eds.) McGraw-Hill, New York, (2rd.ed.), 1981, p. 129-188.
- NEWMAN, D. - The distribution of the range in samples from a normal population expressed in terms of an independent estimate of the standard deviation. *Biometrika* 31:20-30, 1939.
- NOETHER, G.E. - On a theorem of Pitman. *Ann. Math. Statist.* 26: 64-68, 1955.
- O'NEILL, R. and WETHERILL, G.B. - The present state of multiple

comparison methods. *Royal Statistical Society, ser. B* 2: 218-250, 1971.

OUDE VOSHAAR, J.H. - (k-1) mean significance levels of nonparametric multiple comparisons procedures. *The Annals of Statistics* 8: 75-86, 1980.

PRATT, J.W. and GIBBONS, J.D. - Concepts of nonparametric theory, (eds.) Springer-Verlag, Harvard University, 1981, p. 345-444.

PITMAN, E.J.G. - Notes on nonparametric statistical inference, 1948, apud, PURI, M.L. and SEN, P.K., Nonparametric methods in multivariate analysis, (eds.) John Wiley, New York, 1971, p. 107-124.

PURI, M.L. - Some distribution-free k sample rank tests of homogeneity against ordered alternatives. *Communications on Pure and Applied Mathematics* 18: 51-63, 1963.

PURI, M.L. and SEN, P.K. - Nonparametric methods in multivariate analysis, (eds.) John Wiley, New York, 1971, p. 107-124.

RAMSEY, P.H. - Power differences between pairwise multiple comparisons. *J. Am. Statist. Assoc.* 73: 479-485, 1978.

SHEFFÉ, H. - Statistical inference nonparametric case. *Ann. Math. Statist.* 14: 305-332, 1943.

TUKEY, J.W. - The problem of multiple comparisons, 1953, apud, KRAMER, A., Extensions of multiple range tests to group means with unequal numbers of replication. *Biometrics* 12: 307-310, 1956.

WALLER, R.A. and DUNCAN, D.B. - A bayes rule for the symmetric multiple comparisons problem. *J. Am. Statist. Assoc.* 64: 1484-1503.

WELSCH, R.E. - Stepwise multiple comparison procedures. *J. Am. Statist. Assoc.* 72: 566-575, 1977.

WILCOXON, F. and WILCOX, R.A. - Some rapid approximate statistical procedures, Pearl River, New York, 1964, apud BOYETT, J. M. e SHUSTER, J.J. - Nonparametric multiple comparison procedures. *J. Am. Statist. Assoc.* 74: 379-382, 1979.

' A N E X O

NIVEIS DE SIGNIFICANCIAS ESTIMADOS DE ACORDO COM A $\alpha = 0.05$
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

(*) OBS: As observacoes feitas abaixo de cada bloco de simulacoes, sao analogas para todos os casos. Elas devem ser identificadas pelo numero de asteriscos entre parenteses, respectivamente.

(CONTROLE x TRATAMENTO)

TESTES	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0430	0.0495	0.0475	0.0230	0.0410	0.0510	POISSON (30)
	0.0445	0.0555	0.0495	0.0220	0.0440	0.0420	BINOMIAL (20, 0.5)
	0.0220	0.0445	0.0250	0.0030	0.0390	0.0175	GEOMETRICA (0.5)
	0.0080	0.0580	0.0210	0.0005	0.0210	0.0080	BERNOULLI (0.5)

(**) OBS: Os procedimentos usados aqui estao de acordo com as propostas dos autores.

TESTES	STEEL	DUNNETT	K.W.	
M=2N	0.0330	0.0485	0.0455	POISSON (30)
	0.0320	0.0505	0.0360	BINOMIAL (20, 0.5)
	0.0195	0.0450	0.0245	GEOMETRICA (0.5)
	0.0075	0.0490	0.0075	BERNOULLI (0.5)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M+N	0.0380	0.0545	0.0445	POISSON (30)
	0.0340	0.0555	0.0405	BINOMIAL (20, 0.5)
	0.0140	0.0290	0.0180	GEOMETRICA (0.5)
	0.0060	0.0655	0.0080	BERNOULLI (0.5)

(***)OBS: Apenas o teste de K.W. esta' de acordo com a proposta dos seus autores, os demais estao aproximado pela desigualdade de Bonferroni.

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. $\alpha = 0.05$
 (TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0365	0.0535	0.0475	0.0550	0.0110	0.0365	0.0420	POIS(30)
	0.0315	0.0555	0.0390	0.0475	0.0055	0.0335	0.0270	B(20,0.5)
	0.0220	0.0565	0.0260	0.0360	0.0025	0.0280	0.0170	GEOM(0.5)
	0.0100	0.0475	0.0100	0.0255	0.0005	0.0180	0.0100	BERN(0.5)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni≠N	0.0125	0.0460	0.0285	0.0465	POISSON (30)
	0.0130	0.0450	0.0255	0.0550	BIN. (20,0.5)
	0.0075	0.0395	0.0140	0.0340	GEOM. (0.5)
	0.0015	0.0455	0.0030	0.0320	BERN. (0.5)

(****) OBS: Apenas o teste dos postos de Steel esta' aproximado pela
 desigualdade de Bonferroni, os demais seguem as propostas
 dos autores.

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.
 ALTA PROBABILIDADE DE EMPATES.

alpha=0.01

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0080	0.0110	0.0105	0.0040	0.0085	0.0090	POISSON(30)
	0.0095	0.0120	0.0110	0.0025	0.0060	0.0090	BINOMIAL (20,0.5)
	0.0030	0.0075	0.0045	0.0005	0.0020	0.0010	GEOMETRICA(0.5)
	0.0005	0.0080	0.0005	0.0000	0.0015	0.0005	BERNOULLI(0.5)

OBS:(**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0050	0.0065	0.0065	POISSON (30)
	0.0030	0.0080	0.0050	BINOMIAL (20,0.5)
	0.0010	0.0085	0.0025	GEOMETRICA (0.5)
	0.0015	0.0085	0.0015	BERNOULLI (0.5)

OBS:(**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0030	0.0100	0.0040	POISSON (30)
	0.0015	0.0100	0.0030	BINOMIAL (20,0.5)
	0.0005	0.0055	0.0020	GEOMETRICA (0.5)
	0.0005	0.0185	0.0005	BERNOULLI (0.5)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.
 ALTA PROBABILIDADE DE EMPATES.

alpha=0.01

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0080	0.0125	0.0090	0.0100	0.0005	0.0020	0.0040	POIS(30)
	0.0015	0.0100	0.0050	0.0070	0.0000	0.0035	0.0045	B(20,0.5)
	0.0025	0.0140	0.0030	0.0080	0.0005	0.0035	0.0015	GEOM(0.5)
	0.0005	0.0190	0.0020	0.0050	0.0000	0.0005	0.0005	BERN(0.5)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni+N	0.0015	0.0025	0.0025	0.0065	POISSON (30)
	0.0010	0.0055	0.0025	0.0065	BIN. (20,0.5)
	0.0005	0.0080	0.0010	0.0040	GEOM. (0.5)
	0.0000	0.0075	0.0000	0.0035	BERN. (0.5)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. alpha=0.05
 ALTA PROBABILIDADE DE EMPATES.

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N		0.0325	0.0540	0.0355	0.0035	0.0380	0.0210 POIS(1)
		0.0455	0.0525	0.0485	0.0250	0.0440	0.0400 BIN(20,0.2)
		0.0010	0.0420	0.0015	0.0000	0.0035	0.0005 GEOM(0.2)
		0.0005	0.0480	0.0035	0.0000	0.0050	0.0005 BERN(0.2)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N		0.0290	0.0575	0.0370 POISSON (1)
		0.0270	0.0470	0.0375 BIN(20,0.2)
		0.0010	0.0425	0.0015 GEOM(0.2)
		0.0000	0.0450	0.0000 BERN(0.2)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N		0.0185	0.0475	0.0250 POISSON (1)
		0.0240	0.0490	0.0320 BIN(20,0.2)
		0.0020	0.0175	0.0030 GEOM(0.2)
		0.0020	0.0310	0.0020 BERN(0.2)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOATRAL. alpha=0.05
 ALTA PROBABILIDADE DE EMPATES.

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0250	0.0520	0.0285	0.0485	0.0005	0.0290	0.0140	POIS(1)
	0.0310	0.0560	0.0395	0.0440	0.0065	0.0350	0.0300	B(20,0.2)
	0.0000	0.0400	0.0005	0.0080	0.0000	0.0045	0.0000	GEOM(0.2)
	0.0010	0.0520	0.0010	0.0090	0.0000	0.0040	0.0010	BER(0.2)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni≠N	0.0090	0.0380	0.0205	0.0390	POISSON (1)
	0.0165	0.0530	0.0305	0.0505	BIN(20,0.2)
	0.0010	0.0345	0.0010	0.0065	GEOM(0.2)
	0.0000	0.0240	0.0000	0.0100	BER(0.2)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUCAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. alpha=0.01
 ALTA PROBABILIDADE DE EMPATES.

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0040	0.0100	0.0050	0.0000	0.0025	0.0040	POISSON (1)
	0.0090	0.0150	0.0125	0.0000	0.0035	0.0050	BIN(20,0.2)
	0.0000	0.0085	0.0000	0.0000	0.0000	0.0000	GEOM(0.2)
	0.0000	0.0070	0.0000	0.0000	0.0000	0.0000	BERN(0.2)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0020	0.0085	0.0035	POISSON (1)
	0.0015	0.0075	0.0030	BIN(20,0.2)
	0.0000	0.0075	0.0000	GEOM(0.2)
	0.0000	0.0065	0.0000	BERN(0.2)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0010	0.0095	0.0010	POISSON (1)
	0.0030	0.0100	0.0055	BIN(20,0.2)
	0.0000	0.0010	0.0000	GEOM(0.2)
	0.0000	0.0055	0.0000	BERN(0.2)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. $\alpha=0.01$
 ALTA PROBABILIDADE DE EMPATES.

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FIRED.	
Ni=N	0.0005	0.0135	0.0035	0.0055	0.0000	0.0015	0.0015	POISSON (1)
	0.0045	0.0125	0.0060	0.0060	0.0010	0.0040	0.0025	BIN(20,0.2)
	0.0000	0.0060	0.0000	0.0005	0.0000	0.0000	0.0000	GEOM(0.2)
	0.0000	0.0105	0.0005	0.0005	0.0000	0.0000	0.0000	BER(0.2)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni≠N	0.0005	0.0095	0.0015	0.0035	POISSON (1)
	0.0015	0.0085	0.0045	0.0085	BIN(20,0.2)
	0.0000	0.0040	0.0000	0.0000	GEOM(0.2)
	0.0000	0.0060	0.0000	0.0000	BER(0.2)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. $\alpha=0.05$
 ALTA PROBABILIDADE DE EMPATES.

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0000	0.0400	0.0000	0.0000	0.0000	0.0000	POISSON (0.1)
	0.0005	0.0480	0.0005	0.0000	0.0035	0.0005	BIN(20,0.01)
	0.0000	0.0025	0.0000	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0030	0.0000	0.0000	0.0000	0.0000	BERN(0.01)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0000	0.0495	0.0000	POISSON (0.1)
	0.0010	0.0470	0.0015	BIN(20,0.01)
	0.0000	0.0075	0.0000	GEOM(0.01)
	0.0000	0.0110	0.0000	BERN(0.01)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0005	0.0040	0.0005	POISSON (0.1)
	0.0020	0.0170	0.0020	BIN(20,0.01)
	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0000	0.0000	BERN(0.01)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. alpha=0.05
 ALTA PROBABILIDADE DE EMPATES.

(TRATAMENTO x TRATAMENTO)

ESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
	0.0000	0.0340	0.0000	0.0005	0.0000	0.0000	0.0000	POIS(0.1)
	0.0000	0.0505	0.0000	0.0075	0.0000	0.0010	0.0000	B(20,0.01)
i=N	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	BER(0.01)

OBS: (***)

ESTE	STEEL	T.K.	K.W.	J.T.	
	0.0000	0.0255	0.0000	0.0000	POIS. (0.1)
	0.0000	0.0205	0.0005	0.0085	BIM. (20,0.01)
Ni+N	0.0000	0.0040	0.0000	0.0000	GEOM. (0.01)
	0.0000	0.0000	0.0000	0.0000	BER. (0.01)

OBS: (****)

DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. alpha=0.01

ALTA PROBABILIDADE DE EMPATES.

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0000	0.0075	0.0000	0.0000	0.0000	0.0000	POISSON(0.1)
	0.0000	0.0060	0.0000	0.0000	0.0000	0.0000	BIN(20,0.01)
	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	BERN(0.01)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0000	0.0080	0.0000	POISSON (0.1)
	0.0000	0.0095	0.0000	BIN(20,0.01)
	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0005	0.0000	BERN(0.01)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0000	0.0020	0.0005	POISSON (0.1)
	0.0005	0.0035	0.0005	BIN(20,0.01)
	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0000	0.0000	BERN(0.01)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL. alpha=0.01

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0000	0.0055	0.0000	0.0000	0.0000	0.0000	0.0000	POIS(0.1)
	0.0000	0.0100	0.0000	0.0000	0.0000	0.0000	0.0000	B(20,0.01)
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	GEOM(0.01)
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	BER(0.01)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni≠N	0.0000	0.0060	0.0000	0.0000	POISSON (0.1)
	0.0000	0.0020	0.0000	0.0005	B(20,0.01)
	0.0000	0.0000	0.0000	0.0000	GEOM. (0.01)
	0.0000	0.0000	0.0000	0.0000	BER. (0.01)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.05

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0500	0.0525	0.0490	0.0385	0.0435	0.0400	NORMAL (0,1)
	0.0505	0.0545	0.0545	0.0405	0.0440	0.0525	EXP.DUP.(0,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0460	0.0590	0.0555	NORMAL (0,1)
	0.0335	0.0500	0.0380	EXP.DUP.(0,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0360	0.0555	0.0390	NORMAL (0,1)
	0.0355	0.0390	0.0400	EXP.DUP.(0,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATE TENDENDO A ZERO.

alpha= 0.05

(TRATAMENTO x TRATAMENTO)

ESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
i=N	0.0410	0.0580	0.0485	0.0545	0.0245	0.0360	0.0460	NORMAL (0,1)
i	0.0470	0.0505	0.0495	0.0375	0.0260	0.0335	0.0435	EXP.D. (0,1)

OBS: (**)

ESTE	STEEL	T.K.	K.W.	J.T.	
i=N	0.0180	0.0405	0.0305	0.0390	NORMAL (0,1)
i	0.0260	0.0480	0.0380	0.0530	EXP.D. (0,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.01

(CONTROLE x TRATAMENTO)

ESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED.	
	0.0055	0.0075	0.0090	0.0080	0.0035	0.0090	NORMAL (0,1)
	0.0115	0.0120	0.0125	0.0095	0.0050	0.0125	EXP.DUP.(0,1)
OBS: (**)							

ESTE	STEEL	DUNNETT	K.W.	
	0.0055	0.0105	0.0110	NORMAL (0,1)
	0.0075	0.0120	0.0060	EXP.DUP.(0,1)
OBS: (**)				

ESTE	STEEL	DUNNETT	K.W.	
	0.0065	0.0105	0.0040	NORMAL (0,1)
	0.0045	0.0050	0.0065	EXP.DUP.(0,1)
OBS: (***)				

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.01

(TRATAMENTO x TRATAMENTO)

ESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
i=N	0.0070	0.0105	0.0065	0.0100	0.0050	0.0035	0.0080	NORMAL (0,1)
	0.0070	0.0115	0.0090	0.0080	0.0020	0.0030	0.0075	EXP.D. (0,1)

OBS: (**)

ESTE	STEEL	T.K.	K.W.	J.T.	
i=N	0.0005	0.0090	0.0035	0.0065	NORMAL (0,1)
	0.0020	0.0090	0.0065	0.0110	EXP.D. (0,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.05

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0475	0.0625	0.0525	0.0385	0.0455	0.0440	N(0,1)C(P=0.10)
	0.0505	0.0500	0.0545	0.0405	0.0435	0.0525	EXP.DUP.(1,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0440	0.0640	0.0535	N(0,1)C(P=0.10)
	0.0335	0.0440	0.0380	EXP.DUP.(1,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0360	0.0260	0.0400	N(0,1)C(P=0.10)
	0.0355	0.0600	0.0400	EXP.DUP.(1,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
 DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.05

(TRATAMENTO x TRATAMENTO)

STE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
	0.0420	0.0665	0.0485	0.0445	0.0275	0.0390	0.0480	N(0,1)C(P=0.10)
1#N	0.0470	0.0540	0.0495	0.0375	0.0260	0.0345	0.0435	EXP.D.(1,1)

OBS: (**)

STE	STEEL	T.K.	K.W.	J.T.	
	0.0295	0.0605	0.0405	0.0635	N(0,1)C(P=0.10)
1#N	0.0260	0.0465	0.0380	0.0530	EXP.D.(1,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.01

(CONTROLE x TRATAMENTO)

ESTE	STEEL DUNNETT	K.W.	SINAL	R.S.	FRIED.		
=N	0.0075	0.0110	0.0085	0.0010	0.0040	0.0115	N(0,1)C(P=0.10)
	0.0115	0.0135	0.0125	0.0015	0.0055	0.0125	EXP.DUP.(1,1)

OBS: (**)

ESTE	STEEL DUNNETT	K.W.	SINAL	R.S.	FRIED.		
M=2N	0.0080	0.0155	0.0115			0.0115	N(0,1)C(P=0.10)
	0.0075	0.0105	0.0060				EXP.DUP.(1,1)

OBS: (**)

ESTE	STEEL DUNNETT	K.W.	SINAL	R.S.	FRIED.		
M+N	0.0035	0.0040	0.0070			0.0070	N(0,1)C(P=0.10)
	0.0045	0.0155	0.0065				EXP.DUP.(1,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.01

(TRATAMENTO x TRATAMENTO)

STE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FIRED.	
1	0.0065	0.0135	0.0115	0.0075	0.0035	0.0070	0.0105	N(0,1)C(P=0.10)
1	0.0070	0.0135	0.0090	0.0080	0.0020	0.0045	0.0075	EXP.D.(1,1)

OBS: (**)

STE	STEEL	T.K.	K.W.	J.T.	
1	0.0035	0.0090	0.0040	0.0100	N(0,1)C(P=0.10)
1	0.0020	0.0085	0.0065	0.0110	EXP.D.(1,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBICOES DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.05

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED		
M=N		0.0515	0.0555	0.0545	0.0410	0.0385	0.0485	N(0,1)C(P=0.15)
		0.0505	0.0520	0.0545	0.0405	0.0455	0.0525	EXP.D. (3,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.		
M=2N		0.0480	0.0685	0.0580	N(0,1)C(P=0.15)
		0.0335	0.0475	0.0380	EXP.D. (3,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.		
M≠N		0.0350	0.0215	0.0400	N(0,1)C(P=0.15)
		0.0355	0.0720	0.0400	EXP.D. (3,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A

DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.05

(TRATAMENTO x TRATAMENTO)

STE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
=N	0.0440	0.0690	0.0475	0.0390	0.0225	0.0430	0.0520	N(0,1)C(P=.15)
	0.0470	0.0590	0.0495	0.0375	0.0260	0.0350	0.0435	EXP.D.(3,1)

OBS:(***)

STE	STEEL	T.K.	K.W.	J.T.				
#N	0.0305	0.0605	0.0400	0.0625	NORMAL	(0,1)	CONT.	(P=0.15)
	0.0260	0.0480	0.0380	0.0530	EXP.D.	(3,1)		

OBS:(****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO. alpha= 0.01

(CONTROLE x TRATAMENTO)

STE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED.	
	0.0075	0.0090	0.0095	0.0015	0.0055	0.0100	N(0,1)C(P=0.15)
N							
	0.0115	0.0145	0.0125	0.0015	0.0060	0.0125	EXP.D.(3,1)
OBS: (**)							

STE	STEEL	DUNNETT	K.W.	
	0.0090	0.0190	0.0120	N(0,1)C(P=0.15)
2N				
	0.0075	0.0115	0.0060	EXP.D.(3,1)
OBS: (**)				

ESTE	STEEL	DUNNETT	K.W.	
	0.0035	0.0020	0.0065	N(0,1)C(P=0.15)
N				
	0.0045	0.0200	0.0065	EXP.D.(3,1)
OBS: (***)				

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

PROBABILIDADE DE EMPATES TENDENDO A ZERO.

alpha= 0.01

(TRATAMENTO x TRATAMENTO)

STE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
1	0.0080	0.0150	0.0105	0.0080	0.0045	0.0065	0.0095	N(0,1)C(P=0.15)
1	0.0070	0.0115	0.0090	0.0080	0.0020	0.0045	0.0075	EXP.D.(3,1)

OBS: (**)

ESTE	STEEL	T.K.	K.W.	J.T.	
1	0.0035	0.0135	0.0035	0.0085	N(0,1)C(P=0.15)
1	0.0020	0.0095	0.0065	0.0110	EXP.D.(3,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

BAIXA PROBABILIDADEDE EMPATES.

alpha=0.05

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0430	0.0495	0.0475	0.0230	0.0410	0.0510	POISSON (30)
	0.0500	0.0525	0.0490	0.0385	0.0435	0.0400	NORMAL (0,1)
	0.0505	0.0545	0.0545	0.0405	0.0440	0.0525	EXP.DUP.(0,1)
	0.0425	0.0475	0.0480	0.0370	0.0385	0.0400	UNIFOR (3,6)
	0.0410	0.0455	0.0465	0.0420	0.0420	0.0460	EXPONEN (1)
	0.0500	0.0550	0.0560	0.0430	0.0425	0.0425	GAMA (20,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0330	0.0485	0.0455	POISSON (30)
	0.0460	0.0590	0.0555	NORMAL (0,1)
	0.0335	0.0500	0.0380	EXP.DUP.(0,1)
	0.0365	0.0455	0.0470	UNIFOR (3,6)
	0.0365	0.0490	0.0470	EXPONEN (1)
	0.0365	0.0450	0.0430	GAMA (20,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0380	0.0545	0.0445	POISSON (30)
	0.0360	0.0555	0.0390	NORMAL (0,1)
	0.0355	0.0390	0.0400	EXP.DUP.(0,1)
	0.0370	0.0645	0.0425	UNIFOR (3,6)
	0.0370	0.0320	0.0425	EXPONEN (1)
	0.0355	0.0520	0.0390	GAMA (20,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

BAIXA PROBABILIDADE DE EMPATES.

alpha=0.05.

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0125	0.0460	0.0285	0.0465	0.0000	0.0000	0.0000	POIS(30)
	0.0410	0.0580	0.0485	0.0545	0.0245	0.0360	0.0460	NORMAL(0,1)
	0.0470	0.0505	0.0495	0.0375	0.0260	0.0335	0.0435	EXP.D.(0,1)
	0.0425	0.0625	0.0515	0.0445	0.0210	0.0385	0.0445	UNIFOR(3,6)
	0.0425	0.0515	0.0515	0.0515	0.0240	0.0390	0.0445	EXPONEN(1)
	0.0380	0.0520	0.0430	0.0440	0.0240	0.0325	0.0400	GAMA(20,1)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni+N	0.0365	0.0535	0.0475	0.0550	POISSON(30)
	0.0180	0.0405	0.0305	0.0390	NORMAL(0,1)
	0.0260	0.0480	0.0380	0.0530	EXP.D.(0,1)
	0.0200	0.0410	0.0285	0.0500	UNIFOR(3,6)
	0.0200	0.0385	0.0285	0.0485	EXPONEN(1)
	0.0245	0.0475	0.0340	0.0510	GAMA(20,1)

OBS: (****)

NIVEIS DE SIGNIFICANCIAS ESTIMANDOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

BAIXA PROBABILIDADE DE EMPATES.

alpha= 0.01

(CONTROLE x TRATAMENTO)

TESTE	STEEL	DUNNETT	K.W.	SINAL	R.S.	FRIED	
M=N	0.0080	0.0110	0.0105	0.0015	0.0085	0.0090	POISSON (30)
	0.0055	0.0075	0.0090	0.0020	0.0035	0.0090	NORMAL (0,1)
	0.0115	0.0120	0.0125	0.0015	0.0050	0.0125	EXP.DUP.(0,1)
	0.0080	0.0095	0.0095	0.0020	0.0045	0.0125	UNIFOR (3,6)
	0.0080	0.0080	0.0095	0.0020	0.0040	0.0125	EXPONEN (1)
	0.0110	0.0135	0.0110	0.0010	0.0055	0.0075	GAMA (20,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M=2N	0.0050	0.0065	0.0065	POISSON (30)
	0.0055	0.0105	0.0110	NORMAL (0,1)
	0.0075	0.0120	0.0060	EXP.DUP.(0,1)
	0.0075	0.0105	0.0090	UNIFOR (3,6)
	0.0075	0.0110	0.0090	EXPONEN (1)
	0.0040	0.0075	0.0055	GAMA (20,1)

OBS: (**)

TESTE	STEEL	DUNNETT	K.W.	
M≠N	0.0030	0.0100	0.0040	POISSON (30)
	0.0065	0.0105	0.0040	NORMAL (0,1)
	0.0045	0.0050	0.0065	EXP.DUP.(0,1)
	0.0045	0.0165	0.0065	UNIFOR (3,6)
	0.0045	0.0040	0.0065	EXPONEN (1)
	0.0030	0.0085	0.0045	GAMA (20,1)

OBS: (***)

NIVEIS DE SIGNIFICANCIAS ESTIMADOS, DE ACORDO COM A
DISTRIBUICAO DE PROBABILIDADE E O TAMANHO AMOSTRAL.

BAIXA PROBABILIDADE DE EMPATES.

alpha= 0.01

(TRATAMENTO x TRATAMENTO)

TESTE	STEEL	T.K.	K.W.	J.T.	SINAL	R.S.	FRIED.	
Ni=N	0.0080	0.0125	0.0090	0.0100	0.0005	0.0020	0.0040	POIS(30)
	0.0070	0.0105	0.0065	0.0100	0.0050	0.0035	0.0080	NORMAL(0,1)
	0.0070	0.0115	0.0090	0.0080	0.0020	0.0030	0.0075	EXP.D.(0,1)
	0.0070	0.0130	0.0110	0.0080	0.0035	0.0050	0.0090	UNIFOR(3,6)
	0.0070	0.0095	0.0110	0.0095	0.0035	0.0035	0.0090	EXPONEN(1)
	0.0065	0.0090	0.0080	0.0060	0.0030	0.0030	0.0075	GAMA(20,1)

OBS: (**)

TESTE	STEEL	T.K.	K.W.	J.T.	
Ni≠N	0.0015	0.0025	0.0025	0.0065	POISSON (30)
	0.0005	0.0090	0.0035	0.0065	NORMAL (0,1)
	0.0020	0.0090	0.0065	0.0110	EXP.D.(0,1)
	0.0035	0.0080	0.0060	0.0115	UNIFOR (3,6)
	0.0035	0.0060	0.0060	0.0095	EXPONEN (1)
	0.0015	0.0085	0.0040	0.0085	GAMA (20,1)

OBS: (****)