

## COMPARAÇÕES MÚLTIPLAS NÃO-PARAMÉTRICAS PARA O DELINEAMENTO COM UM FATOR DE CLASSIFICAÇÃO SIMPLES

Antonio Carlos Fonseca PONTES<sup>1</sup>

José Eduardo CORRENTE<sup>2</sup>

- **RESUMO:** Uma das principais dificuldades na utilização da Estatística Experimental Não-Paramétrica é a obtenção de resultados confiáveis. As tabelas disponíveis para o Teste de Kruskal-Wallis e para as comparações múltiplas são pouco abrangentes, em especial nos casos de número diferente de repetições entre tratamentos, fazendo com que o pesquisador seja obrigado a recorrer a aproximações. Estas aproximações diferem dependendo do autor a ser consultado, podendo levar a resultados contraditórios. Além disso, tais tabelas não consideram empates, mesmo no caso de pequenas amostras. Assim, o objetivo deste trabalho é apresentar um programa, em linguagem C, que realiza os testes de Kruskal-Wallis e de comparações múltiplas entre todos os tratamentos (bilateral) e entre os tratamentos e o controle (uni e bilateral) considerando todas as configurações sistemáticas dos postos ou com um número fixo de configurações aleatórias, dependendo do número total de permutações possíveis. Os resultados apresentados através da utilização do programa mostram ainda que os testes utilizando as permutações aleatórias podem ser bons substitutos nos casos em que o número de permutações sistemáticas é muito grande.
- **PALAVRAS-CHAVE:** Testes não-paramétricos e de distribuição livre, Teste de Kruskal-Wallis, comparações múltiplas, testes de permutação.

---

1 Departamento de Matemática e Estatística – Universidade Federal do Acre – CEP 69915-900, Rio Branco, Acre, Brasil. E-mail: acfpontes@bol.com.br.

2 Departamento de Ciências Exatas – ESALQ/USP – CEP 13418-900, Piracicaba, São Paulo, Brasil. E-mail: jecorren@carpa.ciagri.usp.br.

## 1 Introdução

Testes não-paramétricos ou de distribuição livre têm sido amplamente utilizados em substituição aos testes paramétricos usuais, em especial quando as pressuposições do modelo não se verificam, ou seja, quando os dados provenientes de um experimento não possuem normalidade ou homogeneidade de variâncias ou em situações em que a aplicação de testes paramétricos não é possível ou torna-se muito complicada, seja pela falta de informações a respeito da forma da distribuição da população, seja pela dificuldade de obtenção de estimativas confiáveis dos parâmetros populacionais. Kendall & Stuart (1952) utilizam os termos distribuição livre (*distribution-free*) e não-paramétrico de formas distintas: não-paramétrico é a descrição do problema e distribuição livre é o método usado para resolver o problema. Entretanto, atualmente os termos não-paramétrico ou distribuição livre para designar um teste são usados indistintamente. Os métodos não-paramétricos são relacionados ao desenvolvimento de procedimentos de inferência estatística que não fazem qualquer suposição explícita sobre a forma da distribuição dos dados, tendo, portanto, menores exigências. Além disso, os procedimentos da estatística não-paramétrica em geral são dados sob a ótica do desenvolvimento inicial da distribuição exata, o que leva a um maior conhecimento das vantagens e desvantagens do teste que está sendo utilizado. A base dos testes não-paramétricos está na ordenação (*ranks*) dos dados e não em seu valor intrínseco, e na aleatorização, onde se consideram todas as possíveis permutações (rearranjos) dos dados. Se, por um lado, perde-se em precisão na troca dos valores da variável por seus respectivos postos, ganha-se em eficiência e facilidade no entendimento dos resultados. Como os métodos não-paramétricos têm como base as estatísticas de ordem, isso significa que, dada uma amostra aleatória, as estatísticas de ordem dessa amostra aleatória são os valores ordenados e, portanto, dependentes, o que torna as deduções muito mais difíceis.

Segundo Conover & Iman (1981), existem várias formas pelas quais os postos podem ser assinalados às observações, destacando os seguintes tipos: atribuição conjunta de postos ao conjunto completo de observações, com a menor delas tendo posto 1, a segunda menor posto 2 e assim por diante (*rank transformation 1* ou *RT-1*); atribuições de postos às variáveis particionadas em subconjuntos, de forma independente em cada um deles (*rank transformation 2* ou *RT-2*); aplicação da *RT-1* após uma reexpressão apropriada dos dados (*rank transformation*

3 ou RT-3) e, finalmente, a aplicação da RT-2 após uma transformação de dados (rank transformation 4 ou RT-4). Outros tipos de atribuição de postos podem ser utilizados para casos específicos.

Dentre os testes não-paramétricos utilizados para delineamentos inteiramente casualizados, o mais comum é o teste proposto por Kruskal & Wallis (1952), em que a atribuição de postos se faz conjuntamente sem levar em conta os tratamentos (ou seja, utiliza-se o RT-1, de acordo com Conover & Iman, 1981). Assim, seja  $X_{ij}$ ,  $i=1, \dots, k$  e  $j=1, \dots, n_i$ , um conjunto de variáveis aleatórias independentes e  $F_i(x)$  a distribuição contínua de  $X_{ij}$  e considere o modelo  $X_{ij} = \mu + \tau_i + \varepsilon_{ij}$ , onde  $\mu$  é a média geral e  $\tau_i$  ( $i=1, \dots, k$ ) é o efeito do  $i$ -ésimo tratamento com a restrição  $\sum_{i=1}^k n_i \tau_i = 0$ . Os parâmetros  $\mu + \tau_i$  são os parâmetros de

locação e os erros  $\varepsilon_{ij}$  são considerados independentes e identicamente distribuídos. Suponha ainda que  $F_i(x) = G(x - \tau_i)$ . A hipótese a ser testada é a de que não há nenhuma diferença entre os efeitos de tratamentos, ou seja,  $H_0: \tau_1 = \dots = \tau_k$  contra a alternativa  $H_a: \tau_i \neq \tau_j$  para ao menos um par  $i, j \in \{1, \dots, k\}$ , ( $i \neq j$ ).

Considere  $N$  elementos submetidos a  $k$  tratamentos, cada um deles com  $n_i$  repetições, ou seja,  $N = \sum_{i=1}^k n_i$  e denote o posto de  $X_{ij}$  na amostra combinada por  $r_{ij}$ . Define-se a estatística  $H$  de Kruskal-Wallis como

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k n_i \left( \bar{R}_{i.} - \frac{N+1}{2} \right)^2 = \frac{12}{N(N+1)} \left( \sum_{i=1}^k \frac{R_{i.}^2}{n_i} \right) - 3(N+1)$$

em que  $R_{i.} = \sum_{j=1}^{n_i} R_{ij}$ ;  $R_{..} = \sum_{i=1}^k R_{i.}$ ;  $\bar{R}_{i.} = \frac{R_{i.}}{n_i}$  e  $\bar{R}_{..} = \frac{R_{..}}{N} = \frac{N+1}{2}$ .

Formas equivalentes de  $H$  também podem ser encontradas em diversos autores. Se a hipótese nula não é verdadeira, então  $H$  deverá tomar valores grandes e, neste caso, rejeita-se  $H_0$  em favor de  $H_a$ , isto é, com um nível de significância  $\alpha$  rejeitamos  $H_0$  se  $H \geq c_N$ , onde  $P_{H_0} [H \geq c_N] = \alpha$ . Sob  $H_0$ , pode-se considerar a amostra combinada  $X_{ij}$ ,  $i=1, \dots, k$  e  $j=1, \dots, n_i$  como sendo proveniente de uma mesma população. Portanto, se  $c(n_1, \dots, n_k; h)$  denota o número de partições de  $N$  observações dentro de todos os  $k$  subgrupos possíveis de tamanhos  $n_1, \dots, n_k$  tais que  $H = h$ , então  $P(H = h) = c(n_1, \dots, n_k; h) \frac{n_1! \dots n_k!}{N!}$ .

A aproximação normal através do  $\chi^2$  com  $(k-1)$  graus de liberdade utilizada para o teste de Kruskal-Wallis quando os grupos são de tamanhos iguais é adequada, segundo Lehmann & D'abrera (1974), exceto para níveis de significância baixos. Para grupos de tamanhos desiguais ou ainda nos casos onde ocorrem empates, tal aproximação também é utilizada, mas o grau de precisão diminui à medida que se aumenta o número de empates ou ainda as diferenças entre os tamanhos dos grupos. Ainda, no caso de ocorrência de empates, uma correção adequada à estatística de Kruskal-Wallis é recomendada.

Ao se rejeitar  $H_0$ , existe ainda o interesse em saber quais dos tratamentos diferem através de comparações múltiplas e o emprego de tais comparações não-paramétricas pode ser encarado como uma complementação ao teste de Kruskal-Wallis. Dentre os testes de comparações múltiplas não-paramétricos destacam-se os propostos por Steel (1960) e Dunn (1964). Steel (1960) apresenta um teste de comparações múltiplas para a comparação simultânea de todos os pares de tratamentos num delineamento inteiramente casualizado balanceado, baseado na atribuição de postos de forma conjunta aos tratamentos pareados  $i$  e  $j$  ( $i=1, \dots, k-1$ ;  $j=i+1, \dots, k$ ), assinalando posto 1 à menor observação e posto  $2n$  à maior. Dunn (1964) apresenta testes de comparações múltiplas utilizando atribuição de postos conjunta a todos os tratamentos. Discussões referentes à escolha do tipo de teste a ser efetuado foram feitas por Skillings (1983), Fairley & Pearl (1984), Fligner (1985) e Critchlow & Fligner (1991), não havendo uma clara vantagem de um sobre outro em todos os casos. Alguns tipos de aproximações para grandes amostras, tanto para os testes de comparações múltiplas entre todos os tratamentos como comparações entre tratamentos e a testemunha, são apresentados em Campos (1983), Conover (1999) e Hollander & Wolfe (1999), utilizando diferentes tipos de aproximações (através da distribuições normal, normal multivariada,  $t$  de Student, amplitude de variáveis independentes com distribuição normal padrão, dentre outras). Valores críticos para o comparações múltiplas baseadas em Dunn (1964), para o caso especial de tamanhos de amostras iguais, foram obtidos em McDonald & Thompson (1967) e Damico & Wolfe (1987) estendem para o caso de delineamentos desbalanceados. Damico & Wolfe (1989) apresentam procedimentos para a obtenção de tabelas para comparações múltiplas entre os tratamentos e o controle em delineamentos inteiramente casualizados. Uma extensão também é feita para o caso de tamanhos de amostras diferentes.

Um dos grandes problemas ao se trabalhar com métodos não-paramétricos é a obtenção dos níveis ou probabilidades de significância exatos para os testes utilizados. A construção de tabelas exatas abrangentes é trabalhosa, exigindo atenção para cada caso em particular. Esta lacuna nas tabelas é sentida especialmente quando o número de repetições dos tratamentos não são iguais. Quando pretende-se efetuar comparações múltiplas a dificuldade é maior ainda já que as tabelas são ainda menos abrangentes. Tanto nas tabelas para o Teste de Kruskal-Wallis como para as comparações múltiplas, os empates não são contemplados e isto leva à utilização de aproximações em delineamentos com um número razoavelmente pequeno de tratamentos e poucas observações, podendo resultar em erros nas conclusões.

Assim, obter os níveis de significância, tanto para o teste de Kruskal-Wallis como para a realização das comparações múltiplas, com uma precisão razoável torna-se bastante difícil. Vários *softwares* podem ser utilizados para a finalidade de obter resultados para testes não-paramétricos em delineamentos experimentais. Cita-se, dentre eles, o *SAS*, o *S-Plus*, o *Statistica*, etc., que contemplam apenas aproximações normais para a maioria dos testes não-paramétricos e não fazem qualquer menção às comparações múltiplas. Outros *softwares*, específicos para métodos não-paramétricos, como o *NPSTAT*, o *StatXact*, o *SENP* e o *XploRe*, são pouco conhecidos e utilizados. Alguns deles trabalham com os testes na sua forma exata ou ainda com testes aleatórios mas, no que se refere à Análise de Variância, a maioria deles se limita aos testes gerais, não trabalhando com métodos de comparações múltiplas na sua forma exata.

Assim, o objetivo deste trabalho é apresentar um programa, construído em linguagem *C*, que forneça níveis de significância para o Teste de Kruskal-Wallis e para as comparações múltiplas. Na Seção 2, é apresentada a metodologia usada para a construção do programa através de testes de permutação sistemática e aleatória e na Seção 3 são discutidos alguns exemplos e mostrada a utilização do programa.

## 2 Material e Métodos

Inicialmente, deve-se fazer uma distinção entre testes de permutação e testes não-paramétricos. Apesar dos testes não-paramétricos utilizarem, na sua forma exata, a permutação para a obtenção dos

níveis de significância para a realização dos testes de hipóteses, os testes de permutação são mais abrangentes, compreendendo uma gama ampla de testes, paramétricos ou não. No teste de permutação a significância é calculada com base na distribuição das estatísticas de teste geradas pela permutação dos dados. Quando a base para a permutação dos dados é a atribuição aleatória então este teste é denominado teste de permutações aleatórias e nos casos em que são realizadas todas as permutações possíveis, este teste é denominado teste de permutações sistemáticas. O teste é realizado da seguinte maneira: uma estatística de teste é calculada para os dados experimentais, então os dados são permutados (rearranjados) repetidamente e a estatística de teste é calculada para cada uma das permutações resultantes. Os dados permutados constituem o conjunto referência para determinar a significância. A proporção de dados permutados no conjunto referência que têm a estatística de teste com valores maiores ou iguais (ou, em certos casos, menores ou iguais) ao valor dos resultados obtidos experimentalmente fornece o  $p$ -valor (significância ou valor de probabilidade). Cada conjunto de dados permutados no conjunto de referência representa os resultados que teriam sido obtidos para uma particular atribuição no caso da hipótese nula ser verdadeira. Observa-se que o número de permutações necessárias para a realização deste teste cresce rapidamente com o aumento do tamanho da amostra impossibilitando a obtenção de todas elas. Com o advento dos computadores, entretanto, a utilização de testes de permutações aleatórias tornou-se possível para um número razoavelmente grande de configurações aleatórias. Por causa da versatilidade destes testes, eles podem ser utilizados tanto para assegurar a validade de testes estatísticos existentes como para desenvolver novos testes para propósitos específicos, dando portanto grande flexibilidade no planejamento e análise de dados experimentais.

Para delineamentos inteiramente casualizados, supondo  $k$  tratamentos com  $n_1, \dots, n_k$  repetições respectivamente, e ainda  $N = \sum_{i=1}^k n_i$ , procede-se da seguinte maneira: inicialmente, calcula-se o valor  $H_C$ , a estatística de Kruskal-Wallis com os dados do experimento transformados em postos  $R_{ij}$ , obtidos utilizando-se a transformação  $RT-I$ ; em seguida, seleciona-se aleatoriamente  $n_1$  elementos para o primeiro tratamento,  $n_2$  dos indivíduos restantes para o segundo tratamento e assim por diante, e qualquer indivíduo pode ser atribuído a qualquer dos tratamentos, de forma aleatória. O número total de permutações é  $NPerm = N! / (n_1! \dots n_k!)$  e se este número for adequado à capacidade

de processamento do programa e do computador que se está trabalhando, pode-se indexar os resultados obtidos de  $1$  a  $N$  e permutar os números indexados, ao invés de usar os resultados, de forma sistemática, garantindo a não redundância na listagem. Para cada permutação obtida, utilizando-se os valores indexados, pode-se calcular o valor da  $h_r$  ( $r=1, \dots, NPerm$ ) estatística de Kruskal-Wallis obtendo-se assim a distribuição de  $H$ . Comparando-se o valor de  $H_C$  com os valores de  $h_r$ , é possível calcular o  $p$ -valor exato para os dados do experimento. Pode-se diminuir o tempo de processamento, utilizando-se como fator de comparação os valores da estatística  $\sum_{i=1}^k R_i^2$  ou  $\sum_{i=1}^k (R_i^2/n_i)$ , para amostras iguais e desiguais, respectivamente, já que os outros valores envolvidos são constantes. Como o teste de Kruskal-Wallis é bilateral, se  $n_i = n_j, \forall i, j$  pode-se utilizar um subconjunto do conjunto completo de permutações, ou seja, as

$$NPerm_{(n_i=n_j)} = NPerm / k! = N! / (n_1! \dots n_k! k!).$$

Quando o número de permutações possíveis cresce, pode-se utilizar o método de permutações aleatórias, no qual um número fixo de permutações é selecionado aleatoriamente, e o  $p$ -valor é baseado apenas nestas permutações. A inclusão ou não da configuração resultante dos dados originais é irrelevante quando o número de permutações geradas é grande.

Ambos os procedimentos, permutações sistemáticas e aleatórias, podem ser adaptados para a utilização nos testes de comparações múltiplas, ou seja, todos os casos possíveis são obtidos através da permutação dos postos (no caso de pequenas amostras) ou obtém-se um conjunto relativamente grande de permutações aleatórias dos postos. Para cada uma das configurações obtidas calcula-se a média das somas dos postos de cada tratamento e utiliza-se o método adequado a cada caso, comparações entre todos os tratamentos ou comparações entre tratamentos e o controle. Quando tem-se o número de repetições iguais para todos os tratamentos e ainda testes de comparações múltiplas bilaterais, pode-se utilizar a mesma redução do número de casos vista para o Teste de Kruskal-Wallis. Para amostras de tamanhos desiguais ou amostras de tamanhos iguais e comparações entre tratamentos e testemunha (ou controle) ou ainda para hipóteses unilaterais, não é possível utilizar esta redução.

Os empates serão tratados em ambos os casos de forma semelhante, ou seja, calcula-se os postos médios para cada grupo de obser-

vações empatadas e a partir deste conjunto de postos, obtém-se todas as configurações (para o caso de pequenas amostras) ou um número relativamente grande de configurações aleatórias.

O programa, que pode ser solicitado através de e-mail aos autores ou ainda encontrado na página <http://ce.esalq.usp.br/~jec>, foi projetado com o intuito de fornecer as probabilidades de significância, tanto para o teste de Kruskal-Wallis, como para as comparações múltiplas de todos os tratamentos e dos tratamentos com o controle. A idéia central é simples: obtém-se todas as configurações referentes à distribuição dos postos às unidades experimentais e a partir destas configurações calcula-se o nível de significância para o teste desejado. Casos com empates podem ser tratados normalmente neste programa.

Na inicialização do programa, é preciso fornecer: número de tratamentos; número de repetições para cada tratamento, observando que o número total de unidades experimentais deve ser menor ou igual a 100; entrada dos dados do experimento, que deve ser feita com os postos atribuídos às unidades experimentais; verificação se existe interesse em proceder às comparações múltiplas e o tipo de comparação a ser efetuada, se de todos os tratamentos ou de tratamentos com o controle; neste último caso, pede-se informar qual tratamento será utilizado como controle. Com essas informações, o programa realiza todas as configurações desde que esse número não ultrapasse 10.000.000. Caso contrário, utiliza-se do teste de permutações aleatórias, ou seja, são construídas 1.000.000 de configurações aleatórias e, a partir delas, obtém-se a probabilidade de significância desejada. O tempo para a realização destas configurações aleatórias depende do tamanho do experimento devido ao processamento da aleatorização. O programa elaborado pode ser facilmente adaptado para obtenção de tabelas completas, tanto para o teste de Kruskal-Wallis como para comparações múltiplas. Os resultados apresentados são:

a) Valor da Estatística de Kruskal-Wallis ( $H_C$ ) para os dados de entrada;

b) Probabilidade de Significância de  $H_C$ , exata, através de permutações sistemáticas, ou aproximada pelas permutações aleatórias, dependendo do número de configurações;

c) Níveis de significância para as comparações múltiplas, de todos os tratamentos (bilateral) ou dos tratamentos com o controle (unilateral ou bilateral), de acordo com a escolha feita pelo usuário; em ambos os casos, dois níveis de significância são apresentados: o primeiro, denotado  $DW$ , onde a diferença entre tratamentos (em módulo) é



comparada com a diferença máxima (em módulo) de cada configuração, de acordo com Damico & Wolfe (1987) ou Damico & Wolfe (1989), para comparações entre todos os tratamentos ou comparações entre os tratamentos e o controle, respectivamente; o segundo, denominado *Geral*, em que cada módulo da diferença entre tratamentos é comparado com todas as diferenças entre tratamentos ou entre os tratamentos e o controle (em módulo) de todas as configurações.

### 3 Resultados e Discussão

Para mostrar a estruturação das tabelas de níveis de significância, tanto para o teste de Kruskal-Wallis como para comparações múltiplas, considere, por exemplo,  $k=3$  tratamentos e  $N=6$ , com  $n_1=1$ ,  $n_2=2$  e  $n_3=3$  que geram  $N!/(n_1!n_2!n_3!)=60$  permutações ou configurações possíveis, obtidas de forma sistemática, que são mostradas na Tabela 1.

Para cada uma das configurações possíveis foi calculada as somas dos postos e o valor  $h_S$  da estatística de Kruskal-Wallis. Obteve-se então a distribuição nula exata para a estatística de Kruskal-Wallis ( $H$ ), com os valores possíveis  $h_S$  de  $H$ , o número de casos em que  $H = h_S$ , as probabilidades  $P_0(H = h_S)$  e as probabilidades acumuladas  $P_0(H \geq h_S)$  que representam os níveis de significância desejados (Tabela 2).

Para obter a Tabela 3, referente às comparações múltiplas utilizou-se as configurações construídas e as médias dos tratamentos  $\bar{R}_i$  ( $i=1,2,3$ ). Considerou-se ainda as  $NC = C_k^2 = 3$  diferenças entre médias de tratamentos  $d_{ij} = |\bar{R}_i - \bar{R}_j|$  e as diferenças padronizadas  $R^*$ , tomando valores  $r_t^*$  ( $t=1,2,3$ ) com  $r_1^* = N^*d_{12}$ ,  $r_2^* = N^*d_{13}$  e  $r_3^* = N^*d_{23}$ , onde  $N^*$  é o mínimo múltiplo comum entre  $n_1$ ,  $n_2$  e  $n_3$ .

Tabela 1 – Configurações do delineamento inteiramente casualizado com três tratamentos ( $n_1=1, n_2=2, n_3=3$ ) e os respectivos valores da estatística de Kruskal-Wallis ( $h_s$ ).

Trat	1	2	3	$h_s$	1	2	3	$h_s$	1	2	3	$h_s$	1	2	3	$h_s$	1	2	3	$h_s$
	1	2	4		1	2	3		1	2	3		1	2	3		1	3	2	
		3	5			4	5			5	4			6	4			4	5	
			6				6				6				5				6	
$R_i$	1	5	15	4,286	1	6	14	3,095	1	7	13	2,381	1	8	12	2,143	1	7	13	2,381
	1	3	2		1	3	2		1	4	2		1	4	2		1	5	2	
		5	4			6	4			5	3			6	3			6	3	
			6				5				6			5					4	
$R_i$	1	8	12	2,143	1	9	11	2,381	1	9	11	2,381	1	10	10	3,095	1	11	9	4,286
	2	1	4		2	1	3		2	1	3		2	1	3		2	3	1	
		3	5			4	5			5	4			6	4			4	5	
			6				6				6				5				6	
$R_i$	2	4	15	3,857	2	5	14	2,381	2	6	13	1,381	2	7	12	0,857	2	7	12	0,857
	2	3	1		2	3	1		2	4	1		2	4	1		2	5	1	
		5	4			6	4			5	3			6	3			6	3	
			6				5				6			5					4	
$R_i$	2	8	11	0,810	2	9	10	1,238	2	9	10	1,238	2	10	9	2,143	2	11	8	3,524
	3	1	4		3	1	2		3	1	2		3	1	2		3	2	1	
		2	5			4	5			5	4			6	4			4	5	
			6				6				6				5				6	
$R_i$	3	3	15	4,286	3	5	13	1,238	3	6	12	0,429	3	7	11	0,095	3	6	12	0,429
	3	2	1		3	2	1		3	4	1		3	4	1		3	5	1	
		5	4			6	4			5	2			6	2			6	2	
			6				5				6			5					4	
$R_i$	3	7	11	0,095	3	8	10	0,238	3	9	9	0,857	3	10	8	1,952	3	11	7	3,524
	4	1	3		4	1	2		4	1	2		4	1	2		4	2	1	
		2	5			3	5			5	3			6	3			3	5	
			6				6				6				5				6	
$R_i$	4	3	14	3,524	4	4	13	1,952	4	6	11	0,238	4	7	10	0,095	4	5	12	0,857
	4	2	1		4	2	1		4	3	1		4	3	1		4	5	1	
		5	3			6	3			5	2			6	2			6	2	
			6				5				6			5					3	
$R_i$	4	7	10	0,095	4	8	9	0,429	4	8	9	0,429	4	9	8	1,238	4	11	6	4,286

Continuação

	5 1 3		5 1 2		5 1 2		5 1 2		5 2 1	
	2 4		3 4		4 3		6 3		3 4	
	6		6		6		4		6	
R <sub>i</sub>	5 3 13	3,524	5 4 12	2,143	5 5 11	1,238	5 7 9	0,857	5 5 11	1,238
	5 2 1		5 2 1		5 3 1		5 3 1		5 4 1	
	4 3		6 3		4 2		6 2		6 2	
	6		4		6		4		3	
R <sub>i</sub>	5 6 10	0,810	5 8 8	1,381	5 7 9	0,857	5 9 7	2,381	5 10 6	3,857
	6 1 3		6 1 2		6 1 2		6 1 2		6 2 1	
	2 4		3 4		4 3		5 3		3 4	
	5		5		5		4		5	
R <sub>i</sub>	6 3 12	4,286	6 4 11	3,095	6 5 10	2,381	6 6 9	2,143	6 5 10	2,381
	6 2 1		6 2 1		6 3 1		6 3 1		6 4 1	
	4 3		5 3		4 2		5 2		5 2	
	5		4		5		4		3	
R <sub>i</sub>	6 6 9	2,143	6 7 8	2,381	6 7 8	2,381	6 8 7	3,095	6 9 6	4,286

Tabela 2 – Distribuição nula da estatística de Kruskal-Wallis para o delineamento inteiramente casualizado com três tratamentos e  $n_1=1, n_2=2$  e  $n_3=3$

$h_s$	# (H = $h_s$ )	P (H = $h_s$ )	P (H ≥ $h_s$ )
4,286	6	3/30	0,1000
3,857	2	1/30	0,1333
3,524	4	2/30	0,2000
3,095	4	2/30	0,2667
2,381	10	5/30	0,4333
2,143	6	3/30	0,5333
1,952	2	1/30	0,5667
1,381	2	1/30	0,6000
1,238	6	3/30	0,7000
0,857	6	3/30	0,8000
0,810	2	1/30	0,8333
0,429	4	2/30	0,9000
0,238	2	1/30	0,9333
0,095	4	2/30	1,0000

Tabela 3 – Médias de Tratamentos ( $\bar{R}_i$ ), diferenças entre os tratamentos  $i$  e  $j$  ( $d_{ij}$ ), diferenças padronizadas ( $r_i^*$ ) e diferença padronizada máxima ( $r^*$ ).

$\bar{R}_1$	$\bar{R}_2$	$\bar{R}_3$	$d_{12}$	$d_{13}$	$d_{23}$	$r_1^*$	$r_2^*$	$r_3^*$	$r^*$	$\bar{R}_1$	$\bar{R}_2$	$\bar{R}_3$	$d_{12}$	$d_{13}$	$d_{23}$	$r_1^*$	$r_2^*$	$r_3^*$	$r^*$
1,00	2,50	5,00	1,50	4,00	2,50	9	24	15	24	1,00	3,00	4,67	2,00	3,67	1,67	12	22	10	22
1,00	3,50	4,33	2,50	3,33	0,83	15	20	5	20	1,00	4,00	3,00	3,00	2,00	1,00	18	12	6	18
1,00	3,50	4,33	2,50	3,33	0,83	15	20	5	20	1,00	4,00	3,00	3,00	2,00	1,00	18	12	6	18
1,00	4,50	3,67	3,50	2,67	0,83	21	16	5	21	1,00	4,50	3,67	3,50	2,67	0,83	21	16	5	21
1,00	5,00	3,33	4,00	2,33	1,67	24	14	10	24	1,00	5,50	3,00	4,50	2,00	2,50	27	12	15	27
2,00	2,00	5,00	0,00	3,00	3,00	0	18	18	18	2,00	2,50	4,67	0,50	2,67	2,17	3	16	13	16
2,00	3,00	4,33	1,00	2,33	1,33	6	14	8	14	2,00	3,50	4,00	1,50	2,00	0,50	9	12	3	12
2,00	3,50	4,00	1,50	2,00	0,50	9	12	3	12	2,00	4,00	3,67	2,00	1,67	0,33	12	10	2	12
2,00	4,50	3,33	2,50	1,33	1,17	15	8	7	15	2,00	4,50	3,33	2,50	1,33	1,17	15	8	7	15
2,00	5,00	3,00	3,00	1,00	2,00	18	6	12	18	2,00	5,50	2,67	3,50	0,67	2,83	21	4	17	21
3,00	1,50	5,00	1,50	2,00	3,50	9	12	21	21	3,00	2,50	4,33	0,50	1,33	1,83	3	8	11	11
3,00	3,00	4,00	0,00	1,00	1,00	0	6	6	6	3,00	3,50	3,67	0,50	0,67	0,17	3	4	1	4
3,00	3,00	4,00	0,00	1,00	1,00	0	6	6	6	3,00	3,50	3,67	0,50	0,67	0,17	3	4	1	4
3,00	4,00	3,33	1,00	0,33	0,67	6	2	4	6	3,00	4,50	3,00	1,50	0,00	1,50	9	0	9	9
3,00	5,00	2,67	2,00	0,33	2,33	12	2	14	14	3,00	5,50	2,33	2,50	0,67	3,17	15	4	19	19
4,00	1,50	4,67	2,50	0,67	3,17	15	4	19	19	4,00	2,00	4,33	2,00	0,33	2,33	12	2	14	14
4,00	3,00	3,67	1,00	0,33	0,67	6	2	4	6	4,00	3,50	3,33	0,50	0,67	0,17	3	4	1	4
4,00	2,50	4,00	1,50	0,00	1,50	9	0	9	9	4,00	3,50	3,33	0,50	0,67	0,17	3	4	1	4
4,00	4,00	3,00	0,00	1,00	1,00	0	6	6	6	4,00	4,00	3,00	0,00	1,00	1,00	0	6	6	6
4,00	4,50	2,67	0,50	1,33	1,83	3	8	11	11	4,00	5,50	2,00	1,50	2,00	3,50	9	12	21	21
5,00	1,50	4,33	3,50	0,67	2,83	21	4	17	21	5,00	2,00	4,00	3,00	1,00	2,00	18	6	12	18
5,00	2,50	3,67	2,50	1,33	1,17	15	8	7	15	5,00	3,50	3,00	1,50	2,00	0,50	9	12	3	12
5,00	2,50	3,67	2,50	1,33	1,17	15	8	7	15	5,00	3,00	3,33	2,00	1,67	0,33	12	10	2	12
5,00	4,00	2,67	1,00	2,33	1,33	6	14	8	14	5,00	3,50	3,00	1,50	2,00	0,50	9	12	3	12
5,00	4,50	2,33	0,50	2,67	2,17	3	16	13	16	5,00	5,00	2,00	0,00	3,00	3,00	0	18	18	18
6,00	1,50	4,00	4,50	2,00	2,50	27	12	15	27	6,00	2,00	3,67	4,00	2,33	1,67	24	14	10	24
6,00	2,50	3,33	3,50	2,67	0,83	21	16	5	21	6,00	3,00	3,00	3,00	3,00	0,00	18	18	0	18
6,00	2,50	3,33	3,50	2,67	0,83	21	16	5	21	6,00	3,00	3,00	3,00	3,00	0,00	18	18	0	18
6,00	3,50	2,67	2,50	3,33	0,83	15	20	5	20	6,00	3,50	2,67	2,50	3,33	0,83	15	20	5	20
6,00	4,00	2,33	2,00	3,67	1,67	12	22	10	22	6,00	4,50	2,00	1,50	4,00	2,50	9	24	15	24

Calculou-se ainda a diferença padronizada máxima  $r^* = N^* \max(d_{ij}) = \max(r_i^*)$  para cada configuração e foram obtidas as probabilidades ou níveis de significância

$\alpha_{DW} = P[R^* \geq N^* \max(d_{ij})] = P[R^* \geq \max(r_i^*)]$  e  $\alpha'_G = P(R^* \geq N^* d_{ij}) = P(R^* \geq r_i^*)$ , para  $t=1,2,3$ .

A Tabela 4 fornece, portanto, as tabelas da distribuição nula de  $R^*_{DW}$  e  $R^*_G$  para as comparações múltiplas entre todos os tratamentos, com  $k=3$ ,  $n_1=1$ ,  $n_2=2$  e  $n_3=3$ , sem a ocorrência de empates, onde  $f_{Max}=\#[R^*=\max(r_i^*)]$ ,  $F_{Max}=\#[R^*\geq\max(r_i^*)]$ ,  $f_{Geral}=\#[R^*=r_i^*]$  e  $F_{Geral}=\#[R^*\geq r_i^*]$ . O símbolo # denota o número de casos em que a condição ocorre.

Tabela 4 – Frequências absolutas e acumuladas, probabilidades de ocorrência para os valores das diferenças padronizadas, comparadas com a diferença máxima (Max) e todas as diferenças (Geral) para as configurações da Tabela 1.

$r_i^*$	$f_{Max}$	$F_{Max}$	$\alpha_{DW}$	$f_{Geral}$	$F_{Geral}$	$\alpha'_G$	$r_i^*$	$f_{Max}$	$F_{Max}$	$\alpha_{DW}$	$f_{Geral}$	$F_{Geral}$	$\alpha'_{Geral}$
27	2	2	0,0333	2	2	0,0111	11	2	48	0,8000	2	84	0,4667
24	4	6	0,1000	4	6	0,0333	10	0	48	0,8000	6	90	0,5000
22	2	8	0,1333	2	8	0,0444	9	2	50	0,8333	12	102	0,5667
21	8	16	0,2667	8	16	0,0889	8	0	50	0,8333	8	110	0,6111
20	4	20	0,3333	4	20	0,1111	7	0	50	0,8333	4	114	0,6333
19	2	22	0,3667	2	22	0,1222	6	6	56	0,9333	16	130	0,7222
18	8	30	0,5000	12	34	0,1889	5	0	56	0,9333	8	138	0,7667
17	0	30	0,5000	2	36	0,2000	4	4	60	1,0000	10	148	0,8222
16	2	32	0,5333	6	42	0,2333	3	0	60	1,0000	12	160	0,8889
15	4	36	0,6000	14	56	0,3111	2	0	60	1,0000	6	166	0,9222
14	4	40	0,6667	6	62	0,3444	1	0	60	1,0000	4	170	0,9444
13	0	40	0,6667	2	64	0,3556	0	0	60	1,0000	10	180	1,0000
12	6	46	0,7667	18	82	0,4556							

Verifica-se, através da Tabela 4, que  $\alpha_{DW} \leq NC \times \alpha_G$ , ou seja, neste caso, em que são realizadas três comparações entre tratamentos, nem sempre a igualdade  $\alpha_{DW} = 3 \times \alpha_G$  ocorre devido à maneira como estes níveis são calculados. Enquanto que para a obtenção do nível *GERAL* faz-se todas as comparações, para o nível *DW* são realizadas comparações apenas com o máximo de cada configuração, gerando assim perda de informações.

O programa apresentado fornece como respostas, para este caso em que há poucas repetições e tratamentos, o nível de significância

exato tanto para o teste de Kruskal-Wallis como para as comparações múltiplas.

Para mostrar a utilização do programa num caso em que o obtenção dos níveis exatos de significância não é viável, utilizou-se um exemplo apresentado por Noether (1991) onde são testadas cinco marcas de pneus (A, B, C, D e E) com o intuito de verificar se existem diferenças entre as marcas no que se refere à distância percorrida pelo carro até a parada final após freá-lo numa determinada velocidade fixa. Com as marcas A, B e C foram realizadas cinco repetições do experimento ( $n_1=n_2=n_3=5$ ) e para as marcas D e E, seis repetições ( $n_4=n_5=6$ ), num total de vinte e sete repetições. Os dados apresentados na Tabela 5 referem-se às distâncias percorridas ( $X_{ij}$ ), em pés, e os respectivos postos ( $R_{ij}$ ) para cada um dos resultados obtidos do experimento, as somas dos postos para cada tratamento ( $R_{ij}$ ) e suas médias ( $\bar{R}_i$ ).

Tabela 5 – Distância percorrida (em pés) ( $X_{ij}$ ), seus respectivos postos ( $R_{ij}$ ), somas de postos ( $R_i$ ) e médias dos postos ( $\bar{R}_i$ ) para as marcas de pneus A, B, C, D e E.

	A		B		C		D		E	
	$X_{1j}$	$R_{1j}$	$X_{2j}$	$R_{2j}$	$X_{3j}$	$R_{3j}$	$X_{4j}$	$R_{4j}$	$X_{5j}$	$R_{5j}$
	151	15,0	157	18,0	135	2,0	147	11,5	146	9,5
	143	7,0	158	19,0	146	9,5	174	26	171	25,0
	159	20,0	150	14,0	142	5,5	179	27	167	24,0
	152	16,0	142	5,5	129	1,0	163	21	145	8,0
	156	17,0	140	4,0	139	3,0	148	13	147	11,5
							165	22	166	23,0
$R_i$		75,0		60,5		21,0		120,5		101,0
$\bar{R}_i$		15,0		12,1		4,2		20,08		16,83

O autor obtém, inicialmente, a estatística de Kruskal-Wallis  $H=12,3$ . Baseado neste valor e devido a inexistência de tabelas, Noether (1991) utilizou a aproximação através da distribuição de  $\chi^2$  com quatro graus de liberdade, obtendo um nível de significância de  $0,01533l$ . Para efetuar as comparações múltiplas entre todos os tratamentos, o autor utilizou a aproximação normal, em que, para

cada par  $i, j$  ( $i < j$ ) calcula-se  $z_{ij} = (\bar{R}_i - \bar{R}_j) / \sigma_{ij}$ , onde

$$\sigma_{ij} = \sqrt{\frac{N(N+1)}{12} \left( \frac{1}{n_i} + \frac{1}{n_j} \right)},$$

considerando ainda  $\alpha=0,25$  e que o número de comparações a serem efetuadas é dado por  $nc=k(k+1)/2=10$ , ou seja,  $\alpha'=\alpha/nc=0,025$ . Com este nível de significância, concluiu que a marca C difere das marcas D e E, requerendo menores distâncias para frear. Com relação às outras comparações, não detectou diferenças significativas.

Com o objetivo de comparar resultados, foram realizadas duas execuções do programa, utilizando o procedimento de permutações aleatórias, a primeira com 1.000.000 e a segunda com 2.000.000 de configurações, tendo em vista a impossibilidade de realizar o teste com as permutações sistemáticas (o número de configurações possíveis para este caso é  $1,22 \times 10^{16}$ ). Desse modo, obteve-se a estatística  $H = 12,27712$  com níveis de significância iguais a 0,006651 e 0,006672, respectivamente para a primeira e segunda execução. A utilização da correção para empates modifica o valor da estatística de Kruskal-Wallis ( $H_{Emp} = 12,28837$ ), mas não o nível de significância obtido através do programa. Estes níveis de significância mostram a fragilidade da aproximação através da distribuição de  $\chi^2$  neste caso, onde tem-se número de repetições diferentes para os tratamentos e a ocorrência de empates, conforme salientam Lehmann & D'Abbrera (1974). Para um nível de significância igual a 5%, em ambos os casos tem-se a rejeição da hipótese de igualdade entre os efeitos de tratamentos. Entretanto, se o nível de significância adotado para o teste de Kruskal-Wallis fosse de 1%, não haveria rejeição da hipótese pela aproximação através da distribuição de  $\chi^2$ , enquanto que a aproximação pelas permutações aleatórias levaria à rejeição, ou seja, as conclusões obtidas seriam totalmente díspares.

Foram realizadas também as comparações múltiplas entre todos os tratamentos nas duas execuções do programa, obtendo-se as probabilidades de significância  $DW$  ( $\alpha_{DW}$ ) e  $Geral$  ( $\alpha'_G$ ), para cada execução aleatória. Foram acrescentados ainda os valores de  $\alpha_G$  para o nível de significância  $Geral$  e os resultados obtidos por Noether (1991) ( $\alpha'_N$ ), para comparações. Neste exemplo, como temos dez comparações sendo efetuadas, teste bilateral, então  $\alpha_G = 10\alpha'_G$  e  $\alpha_N = 10\alpha'_N$ . Novamente aqui a desigualdade  $\alpha_{DW} \leq \alpha_G$  ocorre, pelos motivos relacionados à forma de cálculo dos mesmos.

Tabela 6 – Níveis de significância para as comparações múltiplas entre todos os tratamentos obtidos nas duas execuções do programa (Aleat1 e Aleat2) e resultados dos níveis de significância obtidos por aproximação normal.

Marcas	Aprox. Normal		$\alpha_{DW1}$ (Aleat1)	Geral (Aleat1)		$\alpha_{DW2}$ (Aleat2)	Geral (Aleat2)	
	$\alpha'_N$	$\alpha_N$		$\alpha'_{G1}$	$\alpha_{G1}$		$\alpha'_{G2}$	$\alpha_{G2}$
A-B	0,5635	n.s.	0,97963	0,56168	n.s.	0,97951	0,56206	n.s.
A-C	0,0314	0,314	0,16530	0,02317	0,23166	0,16505	0,02313	0,23129
A-D	0,2902	n.s.	0,85053	0,30305	n.s.	0,85034	0,30302	n.s.
A-E	0,7029	n.s.	0,99632	0,71054	n.s.	0,99625	0,71068	n.s.
B-C	0,1156	n.s.	0,49773	0,10385	n.s.	0,49806	0,10408	n.s.
B-D	0,0967	0,967	0,48884	0,10101	n.s.	0,48914	0,10102	n.s.
B-E	0,3247	n.s.	0,88010	0,33694	n.s.	0,88025	0,33693	n.s.
C-D	0,0010	0,010	0,00321	0,00033	0,00325	0,00319	0,00032	0,00323
C-E	0,0086	0,086	0,05645	0,00662	0,06624	0,05675	0,00666	0,06664
D-E	0,4782	n.s.	0,96836	0,51296	n.s.	0,96815	0,51309	n.s.

Os resultados apresentados mostram que a aproximação normal falha nos casos em que os tratamentos a serem testados têm número de repetições diferentes e os níveis de significância são muito baixos, podendo levar a erros nas conclusões. Considerando-se, por exemplo,  $\alpha = 0,25$  ( $\alpha' = 0,025$ ), através da aproximação normal a marca C difere apenas das marcas D e E. Entretanto, com a utilização do programa é detectada uma diferença significativa também entre as marcas A e C. Além de não detectar esta diferença, nos casos em que os níveis de significância são baixos, os valores obtidos através da aproximação normal são muito distantes dos valores obtidos pelo programa. Em relação às permutações aleatórias, observa-se que os resultados são bastante estáveis, com as diferenças entre as duas execuções realizadas não ultrapassando 0,7%. Em relação aos níveis apresentados pelo programa, *DW* e *GERAL*, recomenda-se a utilização de  $\alpha_{DW}$  quando o interesse do pesquisador for referente ao nível de significância conjunto do experimento, pois, além de ter uma distribuição assintótica conhecida para os casos de tratamentos com o mesmo número de repetições, é um pouco menos conservativo que o nível *GERAL*. Entretanto, quando busca-se os níveis individualizados,  $\alpha'_G$  é mais adequado pois, para sua obtenção, são feitas todas as comparações dentro de cada configuração, tornando-se, por este motivo, mais confiável.



## 4 Conclusões

A utilização de testes não-paramétricos na análise de variância, nos casos em que as pressuposições do modelo não são satisfeitas é uma prática que não é muito comum entre os pesquisadores, algumas vezes por desconhecimento dos métodos, outras pela dificuldade de obtenção de resultados confiáveis. Apesar dos livros-textos apresentarem um grande número de páginas dedicadas exclusivamente às tabelas, elas são limitadas e não apresentam resultados referentes aos casos em que ocorrem empates, levando ao uso de aproximações, mesmo no caso de pequenas amostras. O uso da aproximação normal através do  $\chi^2$  para o teste de Kruskal-Wallis não é adequado nos casos em que temos grupos de tamanhos diferentes ou ainda quando os níveis de significância são baixos, podendo induzir o pesquisador a erros consideráveis em tais casos. Aproximações para as comparações múltiplas são confusas e dependem, não só do tipo de comparação a ser efetuada e do fato de ter-se ou não número de repetições iguais nos tratamentos, mas também do método, que difere de autor para autor. O programa apresentado neste trabalho substitui, com vantagens, as tabelas existentes, fornecendo os níveis de significância exatos para o teste de Kruskal-Wallis e para os testes de comparações múltiplas para amostras de tamanhos iguais ou não, com ou sem empates, no caso de pequenas amostras. Já, para o caso de grandes amostras, é apresentado o uso do teste de permutações aleatórias para a obtenção de níveis de significância aproximados. Estas aproximações nos testes efetuados se mostraram confiáveis, apresentando erros pequenos em relação aos níveis exatos.

## Agradecimentos

Os autores agradecem às contribuições e sugestões feitas pelos revisores.

PONTES, A. C. F, CORRENTE, J. E. Nonparametric multiple comparison for a one-way layout. *Rev. Mat. Estat.* (São Paulo), v.19, p.179-97, 2001.

- **ABSTRACT:** *The main difficulty in using nonparametric methods on experimental design is the lack of good results. The tables providing the quantiles for Kruskal-Wallis test and for multiple comparisons are not wide, mainly when the number of replications are different. In these cases, an approximation is used. These approximated values differ depending on where the table is consulted. Moreover, is difficult to find tables that take account tied observations even though in small samples. Then the aim of this work is to present a routine, built in C language, that runs Kruskal-Wallis and multiple comparison among treatments (bi-tailed test) and between treatment and control (uni and bi-tailed test), considering all systematic configurations of ranks or 1.000.000 of random configurations depending on the total number of permutations. Results are presented using the routine and show that the random permutation tests can be successfully used when number of systematic permutation is very large.*
- **KEYWORDS:** *Nonparametric methods, distribution-free tests, Kruskal-Wallis test, multiple comparisons, permutation tests.*

### **Referências bibliográficas**

CAMPOS, H. Estatística experimental não-paramétrica, 4. ed. Piracicaba: FEALQ, 1983. 349p.

CONOVER, W. J. Practical nonparametric statistics, 3. ed. New York: John Wiley & Sons, 1999. 584p.

CONOVER, W. J., IMAN, R. L. Rank transformations as a bridge between parametric and nonparametric statistics. *Am. Stat.*, v.35, n.3, p.124-9, 1981.

CRITCHLOW, D. E., FLIGNER, M. A. On distribution-free multiple comparisons in the one-way analysis of variance. *Commun. Stat. – Theory Meth.*, v.20, n.1, 127-39, 1991.

DAMICO, J. A., WOLFE, D. A. Extended tables of the exact distribution of a rank statistic for all treatments multiple comparisons in one-way layout designs. *Commun. Stat. – Theory Meth.*, v.16, n.8, 2343-60, 1987.

- DAMICO, J. A., WOLFE, D. A. Extended tables of the exact distribution of a rank statistic for treatments versus control multiple comparisons in one-way layout designs. *Commun. Stat. – Theory Meth.*, v.18, n.9, 3327-53, 1989.
- DUNN, O. J. Multiple comparisons using rank sums. *Technometrics*, v.6, n.3, p.241-52, 1964.
- EDGINGTON, E. S. Randomization tests. 3. ed. New York: Marcel Dekker, 1995, 409p.
- FAIRLEY, D., PEARL, D. K. The Bahadur efficiency of paired versus joint ranking procedures for pairwise multiple comparisons. *Commun. Stat. – Theory Meth.*, v.13, n.12, 1471-81, 1984.
- FLIGNER, M. A. Pairwise versus joint ranking: another look at the Kruskal-Wallis statistic. *Biometrika*, v.72, n.3, p.705-9, 1985.
- HOLLANDER, M., WOLFE, D. A. Nonparametric statistical methods. 2. ed. New York: John Wiley & Sons, 1999. 787p.
- KENDALL, M. G., STUART, A. The advanced theory of Statistics. London: Charles Griffin, 1952. v.2, 690p.
- KRUSKAL, W. H., WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Am. Stat. Assoc.*, v.47, n.260, p.583-621, 1952.
- LEHMANN, E. L., D'ABRERA, H. J. M. Nonparametrics: statistical based on ranks. San Francisco: Holden-Day, 1974. 457p.
- MANOUKIAN, E. B. Mathematical nonparametric statistics. Glasgow: Gordon and Breach Science, 1986. 326p.
- McDONALD, B. J., THOMPSON Jr., W. A. Rank sum multiple comparisons in one- and two-way classifications. *Biometrika*, v.54, n.3, 487-97, 1967.
- NOETHER, G. E. Introduction to statistics: the nonparametric way. New York: Springer-Verlag, 1991. 414p.
- SKILLINGS, J. H. Nonparametric approaches to testing and multiple comparisons in a one-way ANOVA. *Commun. Stat. – Simul. Comput.*, v.12, n.4, 373-87, 1983.
- STEEL, R. G. D. A rank sum test for comparing all pairs of treatments. *Technometrics*, v.3, n.2, p.197-207, 1960.

Recebido em 19.5.2000