
SOFTWARE PARA SEPARAÇÃO DE MÉDIAS PELO MÉTODO DE SCOTT-KNOTT

Autores

Romulo Augusto Althaus

E-mail: ralthaus@uepg.br

Vínculo: Universidade Estadual de Ponta Grossa (UEPG)

Endereço: Rua Mário Nogueira, 66CEP: 84031-410, Ponta Grossa, PR

Telefone: 2262695

Marcelo Giovanetti Canteri

E-mail: mgcanter@convoy.com.br

Vínculo: Universidade Estadual de Ponta Grossa (UEPG), Dep. de Informática

Endereço: Av. Gal. Carlos Cavalcante, 4748. CEP 8403000, Ponta Grossa, PR

Telefone: (42) 220-3097

Éder Antônio Gigliotti

E-mail: eder@dbv.cca.ufscar.br

Vínculo: Universidade Federal de São Carlos (UFSCar)

Endereço: Via Anhanguera, Km 174 C.P. 153. CEP: 13600970, Araras, SP

Telefone: (19)5423888

Resumo

Métodos estatísticos de separação de médias são utilizados para classificar dados obtidos em experimentos e para separá-los em grupos. O presente trabalho teve como objetivo desenvolver um algoritmo para separar médias pelo método de Scott-Knott, implementar e validar um sistema capaz de realizar todas as etapas necessárias ao teste, desde a análise de variância até a separação das médias. O sistema foi desenvolvido utilizando-se a linguagem de programação Borland Delphi e foi direcionado a computadores compatíveis com a plataforma PC que utilizam o sistema operacional Microsoft Windows 95 ou posterior. Para validar o sistema foram feitas comparações entre resultados de experimentos calculados de forma manual e pelo sistema, analisados pelo método de Scott-Knott. Também foram comparados os resultados gerados pelo teste de Scott-Knott com os testes de Duncan e Tukey. Quando comparado aos outros testes verificou-se que o método de Scott-Knott facilita a interpretação dos resultados, pois os dados são classificados em grupos distintos e não há sobreposição entre os grupos encontrados. Essa sobreposição é característica de outros métodos bastante utilizados como o de Duncan e o de Tukey. O sistema desenvolvido será útil para viabilizar o uso do teste de Scott-Knott na análise de experimentos. Antes disto, o grande número e a complexidade dos cálculos inviabilizavam sua utilização em experimentos com uma grande quantidade de dados.

Abstract

Statistical methods of means separation are used to classify data obtained in experiments and to separate them in groups. The present work had as objective develop an algorithm to separate means for the Scott-Knott method, to implement

and to validate a system capable to accomplish all the necessary stages of the test, from the variance analysis to the separation of the means. The system was developed being used the programming language Borland Delphi and it was directed to compatible computers with the platform PC that use the operating system Microsoft Windows 95 or above. To validate the system were made comparisons among results of calculated experiments in a manual way and for the system, analyzed by the method of Scott-Knott. Also were compared results generated by the Scott-Knott test with the Duncan's and Tukey's tests. When compared to the other methods of means separation was verified that the Scott-Knott method facilitates the interpretation of the results, because the data are classified in different groups and there is not overlapping among the found groups. That overlapping is characteristic of other methods quite used as Duncan's and Tukey's methods. The developed system can be useful to make possible the use the Scott-Knott test in the analysis of experiments. Before this, the great number and the complexity of the calculations in a manual way, made unfeasible its use in experiments with a great amount of data.

Palavras Chaves

teste de separação de médias, Scott-Knott, testes estatísticos.

1. INTRODUÇÃO

Freqüentemente pesquisadores utilizam-se de métodos estatísticos para analisar os dados gerados por seus experimentos. A análise estatística destes dados exige uma numerosa quantidade de cálculos (Banzatto & Kronka, 1995; Kalil, 1977; Gomes, 1987). Os computadores proporcionam resultados rápidos, permitindo a obtenção de resultados de uma maneira mais precisa, além de diminuir o tempo gasto com cálculos e tarefas repetitivas. A facilidade de acesso aos computadores está permitindo a expansão de seu uso nas mais diversas áreas. Os cálculos estatísticos têm aplicações que vão desde a área de agricultura até a área de manipulação genética além da grande aplicação na área de agronegócio e ciências ambientais (Silva, 1999).

Os objetivos do presente trabalho foram elaborar um software para separação de médias pelo método de Scott-Knott, implementar a análise de variância para delineamento inteiramente casualizado e delineamento em blocos ao acaso, comparar os resultados deste método com outros métodos existentes, principalmente Tukey e Duncan, e executar testes de validação do software.

2. MATERIAIS E MÉTODOS

Para o desenvolvimento do software foi utilizada a linguagem de programação Borland Delphi, versão 4. O software desenvolvido requer como configuração mínima um microcomputador que seja compatível com o sistema operacional Microsoft Windows 95.

O algoritmo do teste de Scott-Knott foi elaborado baseando-se no artigo *A Cluster Analysis Method for Grouping Means in The Analysis of Variance* (Scott & Knott, 1974). Os resultados obtidos no software foram comparados com resultados publicados no artigo *Illustration of a Cluster Analysis Method for Mean Separation*

(Gates & Bilbro, 1978) que demonstrava resultados práticos do teste de Scott-Knott. Os resultados também foram comparados com os resultados obtidos através de um software para ambiente MS-DOS que realiza a separação de médias pelo método de Scott-Knott (Ferreira, 1995).

As aproximações das tabelas estatísticas (χ^2 , teste de F) foram baseadas em algoritmos de Best & Roberts (1975) para a distribuição de χ^2 e de Cooke *et al.* (1981) para a distribuição de F.

3. RESULTADOS E DISCUSSÃO

Durante os estudos das rotinas de cálculo do teste de Scott-Knott foram observados que alguns dados de entrada para a realização dos procedimentos de cálculo só seriam obtidos após a análise de variância do experimento. Passou-se, então a implementação dessa análise, para delineamento em blocos ao acaso. Foi acrescentado, posteriormente, o delineamento inteiramente casualizado. Os resultados da análise de variância ficam disponíveis em uma tabela e os dados dessa análise que serão utilizados no teste de separação de médias são transportados diretamente para os campos correspondentes. Cada etapa (análise de variância, resultado da análise de variância e teste de Scott-Knott) tem sua própria interface com campos para entrada de dados e pode ser selecionada por meio de abas localizadas na parte superior da tela (Figura 1). A etapa do teste de Scott-Knott é independente da análise de variância e uma vez que o usuário já possua os dados necessários ao teste, ele pode simplesmente selecionar a aba correspondente, entrar com os dados e obter os resultados do teste.

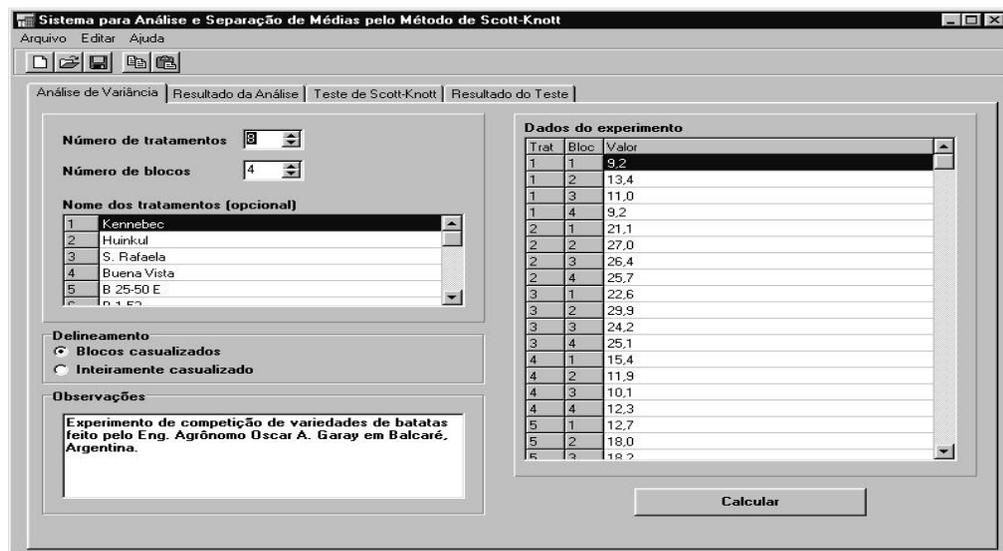


Figura 1 - Interface do programa para separação de médias pelo método de Scott-Knott.

3.1 - ANÁLISE DE VARIÂNCIA

Através da análise da variância, são obtidos alguns dados que serão fundamentais no teste de separação de médias de Scott-Knott, como o quadrado médio do resíduo e os graus de liberdade do resíduo.

A influência dos fatores não controlados (variação ao acaso) pode ser avaliada através do resíduo ou erro. A análise de variância permite identificar os fatores não controlados dos fatores atribuídos aos tratamentos. Num experimento, por exemplo, onde os resultados indicassem que uma variedade de cana-de-açúcar é superior a uma outra variedade, poderíamos identificar se esse resultado se deve realmente a superioridade de uma variedade em relação a outra ou se esse resultado é devido a algum fator não controlado.

A denominação graus de liberdade refere-se ao número de observações no cálculo da variância. A variância poderia ser definida como a medida da dispersão da população. Quanto maior a dispersão nos dados da população, maior a variação ao acaso.

3.2 - MÉTODO DE SEPARAÇÃO DE MÉDIAS DE SCOTT-KNOTT

Para a aplicação do método de separação de médias de Scott-Knott, o software ordena os tratamentos através de suas médias em ordem crescente e realiza o teste estatístico para encontrar conjuntos ou grupos de médias de tratamentos.

Este é um cálculo extremamente trabalhoso, principalmente em experimentos com um grande número de médias pois todas as ramificações dos grupos de médias têm que ser analisadas.

Por exemplo, se houver 8 médias em ordem crescente (1,2,3,4,5,6,7 e 8), onde o teste acusou uma diferença do grupo com as cinco primeiras médias 1,2,3,4,5 e do grupo das três últimas 6,7,8, deve-se separar as médias em dois grupos. Lembrando que as médias num mesmo grupo são consideradas homogêneas, executase os cálculos com as médias 1,2,3,4,5. Se o resultado acusar uma diferença entre o grupo com as médias 1,2,3 e o grupo com as médias 4,5, faz-se o teste para o grupo com as médias 1,2,3. Obtêm-se uma diferença entre os grupos com as médias 1,2 e o grupo com a média 3. Testa-se as médias 1 e 2. Se essas médias forem consideradas homogêneas, então não precisam mais ser particionadas. Todas as outras ramificações restantes serão analisadas até que não existam diferenças entre as médias ou que reste apenas uma média num grupo. O passo seguinte realizado foi o teste do grupo que contém a média 3. Verificado que esta média é única no grupo, passase para o grupo que contém as médias 4 e 5 e assim sucessivamente para os outros grupos. Na Figura 2 estão representados todos os caminhos que foram percorridos para que o algoritmo fizesse a classificação das médias em 5 grupos nesse exemplo hipotético.

3.3 - COMPARAÇÃO ENTRE MÉTODOS DE SEPARAÇÃO DE MÉDIAS

A Tabela 1 apresenta o resultado de três testes de separação de médias diferentes, num experimento de competição de variedades de batatas (Gomes, 1987). A maneira mais usual de classificação de médias atribui letras para diferenciar uma média de outra. Médias que possuem a mesma letra são consideradas equivalentes (não possuem diferença significativa). Quando isso acontece, considerase que a diferença entre os valores das médias equivalentes é devido à variação ao acaso.

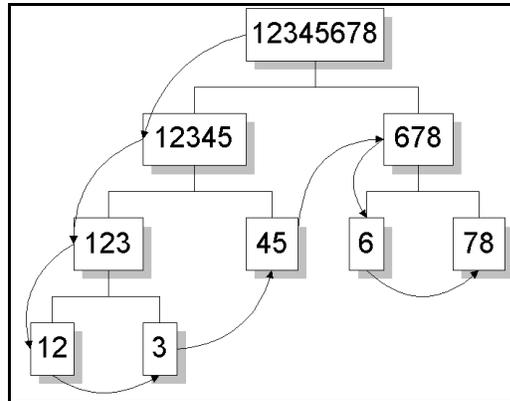


Figura 2 - Caminho percorrido para classificação pelo algoritmo de separação de médias pelo método de Scott-Knott em um exemplo com 8 médias.

Tabela 1 - Resultados de 3 testes de separação de médias num experimento para comparação de produtividade entre variedades de batatas

	Variedades	Médias (t / ha)	Separação por		
			Duncan	Tukey	Scott-Knott
1	S. Rafaela	25,5	a	a	a
2	Huinkul	25,1	a	a	a
3	B 72-53 A	22,8	a	ab	a
4	B 116-51	22,5	a	ab	a
5	B 1-52	22,3	a	ab	a
6	B 25-50 E	16,5	b	bc	b
7	Buena Vista	12,4	bc	c	c
8	Kennebec	10,7	c	c	c

Na análise pelo método de Duncan observou-se que existiam médias com mais de uma letra. A 6ª e a 7ª variedade são consideradas equivalentes. Do mesmo modo, a 7ª e a 8ª variedade também são equivalentes. Se a 6ª e a 7ª variedade e a 7ª e a 8ª variedade são equivalentes, pode-se concluir, erroneamente, que as 3 variedades são equivalentes. Analisando melhor, observa-se que a 6ª e a 8ª variedade não possuem a mesma letra, logo, não são equivalentes. O mesmo raciocínio é aplicado na análise dos resultados do método de Tukey. Essa análise pode tornar-se bastante complicada, em alguns casos, quando tem-se um grande número de médias a serem separadas, como no caso de experimentos com teste de variedades.

Os resultados do método de Scott-Knott apresentam-se de forma direta e não possuem a sobreposição que é característica dos outros métodos de separação de médias, permitindo a separação dos materiais em grupos homogêneos.

O processo de validação indicou que o software desenvolvido mostrou-se eficiente no uso do método de Scott-Knott para separação de médias, principalmente nos casos em que os resultados gerados por outros testes não permitem a separação de grupos de médias. Sugere-se a utilização do software na área agrícola, principalmente em testes de seleção de materiais desenvolvidos pelo melhoramento genético.

4. REFERÊNCIAS

- BANZATTO, David A.; KRONKA, Sérgio do N. Experimentação Agrícola. 3. ed. Jaboticabal : FUNEP, 1995.
- BEST, D.J.; ROBERTS, D.E. Algorithm AS91. The Percentage Points of the Chi-squared Distribution. Applied Statistics, 24, 385-388, 1975.
- COOKE, D.; CRAVEN, A. H.; CLARKE, G.M. Basic Statistical Computing. Londres: Edward Arnold Publ. Co, 1981.
- FERREIRA, Daniel F. Programa para cálculo do agrupamento de Scott & Knott São Paulo, 1995. Disquete 3 1/2, 2 arquivos, 60.382 bytes.
- GATES, C. E.; BILBRO, J. D. Illustration of a cluster analysis method for mean separation. Agronomy Journal, 70: 462-5, 1978.
- GOMES, F. P.. Curso de Estatística Experimental. 12. ed. Piracicaba : Nobel, 1987.
- KALIL, E. B. Princípios de técnicas experimentais com animais. Piracicaba: Curso de Pós-Graduação em Nutrição Animal e Pastagens- ESA - Luis de Queiroz - USP, 1977.
- SCOTT, A.; KNOTT, M. A cluster analysis method for grouping means in the analysis of variance. Biometrics, 30: 507-12, 1974.
- SILVA, Israel P. da. Métodos estatísticos aplicados à pesquisa científica: uma abordagem para profissionais da pesquisa agropecuária. Recife : UFRPE, 1999.