UM ALGORITMO PRÁTICO PARA DETERMINAÇÃO DA VARIÂNCIA DO ESTIMADOR DE UM CONTRASTE DE PARÂMETROS[1]

ÉLIO PAULO ZONTA[2] e JOÃO GILBERTO CORRÊA DA SILVA[3]

RESUMO - Testes de significância de contrastes de efeitos de níveis de um fator ou de combinações de níveis de dois ou mais fatores requerem o conhecimento das estimativas das variâncias dos estimadores desses contrastes. As expressões dessas estimativas para experimentos com delineamentos complexos, em geral, não são disponíveis em textos. Sua derivação algébrica a partir da equação do modelo estatístico é trabalhosa. Expõe-se um algoritmo prático para a obtenção da variância do estimador de qualquer contraste de parâmetros para delineamentos completos balanceados. O algoritmo baseia-se nos valores esperados dos quadrados médios.

Termos para indexação: comparações de tratamentos, delineamentos completos balanceados, inferência de experimento.

A PRACTICAL ALGORITHM TO DETERMINE THE VARIANCE OF THE ESTIMATOR OF A CONTRAST OF PARAMETERS

ABSTRACT - Significance tests of contrasts of effects of levels of a factor or of combinations of levels of two or more factors require the knowledge of the estimates of the variances of the estimators of these contrasts. In general, the expressions of these estimates for experiments with complex designs are not available in textbooks. Their algebraic derivation from the equation of the statistical model is tedious. A practical algorithm for obtaining the variance of the estimator of any contrast of parameters for balanced complete designs is presented. The algorithm is based on the expected values of mean squares.

Index terms: treatment comparisons, balanced complete designs, inference from experiment.

INTRODUÇÃO

Freqüentemente, o pesquisador se defronta com o problema de efetuar testes de significância de contrastes de efeitos de níveis de um fator ou de combinações de níveis de dois ou mais fatores para os quais não são disponíveis as expressões das variâncias dos correspondentes estimadores. Isso ocorre, em geral, na análise de experimentos com modelos lineares mistos, com mais de um componente aleatório, como, por exemplo, experimentos

[1] Aceito para publicação em 10 de dezembro de 1998.

[2] Eng. Agr., M.Sc., Prof. Adjunto, Dep. de Matemática, Estatística e Computação, Inst. de Física e Matemática, Univ. Fed. de Pelotas (UFPEL), Caixa Postal 354, CEP 96010-900 Pelotas, RS. E-mail: epzonta@ufpel.tche.br

[3] Eng. Agr., Ph.D., Prof. Titular, UFPEL. Bolsista do CNPq. E-mail: jgcs@ufpel.tche.br

fatoriais com delineamentos da extensa família de parcelas divididas, experimentos de medidas repetidas e experimentos de ampla abrangência no espaço e no tempo.

Recentemente, o Statistical Analysis System (SAS Institute, 1992) implementou procedimento geral de análise para modelo linear misto geral, o PROC MIXED, que permite a execução de testes de significâncias de contrastes para situações mais genéricas, que incluem as que são consideradas aqui. Entretanto, o conhecimento de expressões algébricas de variâncias de estimadores de contrastes pode ser desejável.

Milliken & Johnson (1992) apresentam um método algébrico geral para a determinação da variância de uma diferença de duas médias para delineamento com parcelas divididas balanceado, a partir do modelo estatístico e do conhecimento da composição dos valores esperados dos quadrados médios. Eles afirmam que o método pode ser estendido para situações mais complexas, e o aplicam para delineamentos com observações repetidas no tempo.

Entretanto, a determinação de expressões de variâncias de estimadores de contrastes a partir do modelo estatístico é trabalhosa. Não se tem conhecimento de disponibilidade na literatura de qualquer método prático alternativo. O propósito deste artigo é expor um método expedito, de uso genérico, para a derivação da expressão da variância do estimador de qualquer contraste de parâmetros para delineamentos completos balanceados, ou seja, delineamentos com número de observações nas células constante para qualquer classificação dos dados segundo os fatores presentes no experimento. O algoritmo baseia-se nos valores esperados dos quadrados médios. É uma generalização do método apresentado por Zonta (1989) para a determinação da variância do contraste de duas médias de tratamentos.

DERIVAÇÃO ALGÉBRICA DA VARIÂNCIA DO ESTIMADOR DE UM CONTRASTE DE PARÂMETROS

Sejam $t_1, t_2, ..., t_T$ T parâmetros correspondentes aos T níveis de um fator fixo ou combinações dos níveis de dois ou mais fatores fixos. Então, a combinação linear $C = c_1 t_1 + c_2 t_2 + ...+ c_T t_T$ é uma comparação ou contraste dos parâmetros $t_1, t_2, ..., t_T$ se a soma de seus coeficientes é zero, ou seja, se $c_1 + c_2 + ...+ c_T = 0$.

Um contraste de parâmetros $C = c_1 t_1 + c_2 t_2 + ...+ c_T t_T$ é estimado pelo contraste dos correspondente estimadores dos parâmetros: $\hat{C} = c_1\hat{t}_1 + c_2\hat{t}_2 + ... + c_T\hat{t}_T$, onde $\hat{t}_t$ é o estimador de t_t, t=1, 2,..., T.

A determinação algébrica da variância do estimador de um contraste de parâmetros, a partir do modelo estatístico, é ilustrada, a seguir, para a situação de um experimento com dois fatores fixos *A* e *B* com delineamento blocos casualizados com parcelas divididas, com os fatores *A* e *B* em parcelas e subparcelas, respectivamente, e fator de bloqueamento *R*. Nesse processo, são utilizadas as propriedades referentes a valor esperado, as pressuposições e condições usuais do modelo estatístico e a notação adotada por Silva (1999).

A forma usual da equação do modelo estatístico para este delineamento (Federer, 1955; Hicks, 1973; Steel & Torrie, 1980; Montgomery, 1984; Milliken & Johnson, 1992) é:

$y_{abr} = m + r_r + a_a + ar_{ar} + b_b + ab_{ab} + e_{abr}$, $a = 1, 2, ..., A$, $b = 1, 2, ..., B$, $r = 1, 2, .., R$.

A consideração dos fatores de unidade na derivação das variâncias de estimadores de contrastes de parâmetros é importante para a obtenção das expressões corretas para muitos delineamentos (Silva, 1999).

A equação completa do modelo estatístico para o delineamento em consideração, com a inclusão dos efeitos de unidade, é:

$$y_{abr} = m + [r_r + u'_r] + a_a + [ar_{ar} + u_{a(r)}] + b_b + ab_{ab} + e_{b(ar)},$$

$a = 1, 2, ..., A$, $b = 1, 2, ..., B$, $r = 1, 2, ..., R$, onde $e_{b(ar)}$, $u_{a(r)}$ e u'_r são os efeitos de unidade presentes no delineamento, correspondentes aos fatores de unidade "subparcela", "parcela" e "bloco", respectivamente.

Considerem-se as pressuposições usuais postuladas com o modelo estatístico (Silva, 1999). O estimador do parâmetro a_a, $a = 1, 2, ..., A$, é o desvio da média das respostas para o nível a do fator *A* em relação à média geral das respostas, ou seja:

$$\hat{a}_a = \bar{y}_{a..} - \bar{y}_{...}$$

Logo, o estimador de um contraste genérico dos parâmetros a_a, seja, $C = \sum_a c_a a_a$, é:

$\hat{C} = \sum_a c_a \hat{a}_a = \sum_a c_a \bar{y}_{a..}$, ou seja, o correspondente contraste das médias observadas para os níveis do fator *A*. Mas:

$$\bar{y}_{a..} = \frac{1}{BR}\sum_b\sum_r\left[m + a_a + r_r + u'_r + ar_{ar} + u_{a(r)} + b_b + ab_{ab} + e_{b(ar)}\right]$$

$$= m + a_a + \frac{1}{R}\sum_r r_r + \frac{1}{R}\sum_r u'_r + \frac{1}{R}\sum_r ar_{ar} + \frac{1}{R}\sum_r u_{a(r)} + \frac{1}{BR}\sum_b\sum_r e_{b(ar)},$$

em decorrência das condições impostas para a situação do fator *B* fixo:

$$\sum_b b_b = 0 \text{ e } \sum_b ab_{ab} = 0.$$

Logo, a expressão do estimador de um contraste dos parâmetros é:

$$\hat{C} = \sum_a c_a\left[m + a_a + \frac{1}{R}\sum_r r_r + \frac{1}{R}\sum_r u'_r + \frac{1}{R}\sum_r ar_{ar} + \frac{1}{R}\sum_r u_{a(r)} + \frac{1}{BR}\sum_b\sum_r e_{b(ar)}\right]$$

$$= \sum_a c_a a_a + \frac{1}{R}\sum_a c_a \sum_r ar_{ar} + \frac{1}{R}\sum_a c_a \sum_r u_{a(r)} + \frac{1}{BR}\sum_a c_a \sum_b \sum_r e_{b(ar)} .$$

A variância desse contraste é derivada como segue:

$$Var(\hat{C}) = Var\left[\sum_a c_a a_a + \frac{1}{R}\sum_a c_a \sum_r ar_{ar} + \frac{1}{R}\sum_a c_a \sum_r u_{a(r)} + \frac{1}{BR}\sum_a c_a \sum_b \sum_r e_{b(ar)}\right]$$

$$= \frac{1}{R^2} Var\left[\sum_a c_a \sum_r ar_{ar}\right] + \frac{1}{R^2} Var\left[\sum_a c_a \sum_r u_{a(r)}\right] + \frac{1}{B^2R^2} Var\left[\sum_a c_a \sum_b \sum_r e_{b(ar)}\right]$$

$$= \frac{1}{R^2} R\sum_a c_a^2 \sigma_{AR}^2 + \frac{1}{R^2} R\sum_a c_a^2 \sigma_U^2 + \frac{1}{B^2R^2} BR\sum_a c_a^2\, \sigma_E^2 ,$$

desde que $Var(ar_{ar}) = \sigma_{AR}^2$, $Var(u_{a(r)}) = \sigma_U^2$, $Var(\mathbf{e}_{b(ar)}) = \sigma_E^2$ e as covariâncias são todas nulas. Logo,

$$Var(\hat{C}) = \frac{1}{BR}\sum_a c_a^2 \left[B\sigma_{AR}^2 + B\sigma_U^2 + \sigma_E^2\right].$$

MÉTODO PRÁTICO

A simplicidade da expressão da variância do estimador do contraste de parâmetros obtida na ilustração anterior é resultante das seguintes pressuposições do modelo estatístico:

a) a variância de um efeito fixo é nula;

b) as covariâncias dos efeitos de um termo aleatório correspondentes a diferentes observações são nulas;

c) as covariâncias dos efeitos correspondentes a dois termos aleatórios distintos são nulas.

Essas propriedades permitem o estabelecimento de um procedimento geral para obtenção da variância do estimador de um contraste de parâmetros de um efeito fixo, para qualquer delineamento completo balanceado. Para ilustração do procedimento considere-se o mesmo contraste genérico dos parâmetros a_a do delineamento blocos casualizados com parcelas divididas. O procedimento compreende os seguintes passos:

1) Identificam-se os termos aleatórios do modelo que incluem pelo menos os índices variáveis no contraste de parâmetros. Constrói-se uma tabela com esses termos e as correspondentes variâncias, como a Tabela 1.

2) Para cada termo da Tabela 1, determina-se o produto das amplitudes de seus índices, excluídos os índices dos parâmetros cujo contraste está sendo considerado. Preenche-se esses produtos na última linha da Tabela 1.

3) Efetua-se a soma dos quocientes das variâncias pelos respectivos produtos de amplitudes de índices:

$$\frac{\sigma_E^2}{BR} + \frac{\sigma_U^2}{R} + \frac{\sigma_{AR}^2}{R}$$

4) A variância do contraste é o produto dessa soma pela soma dos quadrados dos coeficientes do contraste:

$$\mathrm{Var}(\hat{C}) = \frac{\sum c_a^2}{BR}\left[\sigma_E^2 + B\sigma_U^2 + B\sigma_{AR}^2\right].$$

De modo geral, a variância do estimador de um contraste de parâmetros pode ser expressa como uma combinação linear de valores esperados de quadrados médios (*E*QM), a partir da inspeção dos *E*QM. Para o delineamento que está sendo considerado para ilustração, os *E*QM completos – incluindo os componentes de variância correspondentes aos efeitos de unidade, determinados pelo método de Silva (1999) –, são dados na Tabela 2.

TABELA 1. **Derivação da variância do estimador de um contraste dos parâmetros a_a do fator *A* em parcelas de um delineamento blocos casualizados com parcelas divididas.**

Termo	$e_{b(ar)}$	$u_{a(r)}$	ar_{ar}
Variância	σ_E^2	σ_U^2	σ_{AR}^2
Divisor	BR	R	R

TABELA 2. ***E*QM para delineamento blocos casualizados com parcelas divididas, com o fator *A* em parcelas, o fator *B* em subparcelas e fator de bloqueamento *R*, considerando os fatores de unidade.**

Fonte de variação /efeito		a F	b F	c A	*E*QM[1]
A:	a_a	0	B	R	$\sigma_E^2 + B[\sigma_{AR}^2 + \sigma_U^2] + BR\sigma_A^2$
R:	r_r	A	B	1	$\sigma_E^2 + B\sigma_U^2 + AB[\sigma_R^2 + \sigma_{U'}^2]$
AR:	ar_{ar}	0	B	1	$\sigma_E^2 + B[\sigma_{AR}^2 + \sigma_U^2]$
B:	b_b	A	0	R	$\sigma_E^2 + AR\sigma_B^2$
AB:	ab_{ab}	0	0	R	$\sigma_E^2 + R\sigma_{AB}^2$
E:	$e_{b(ar)}$	1	1	1	σ_E^2

[1] $\sigma_A^2 = \frac{1}{A-1}\sum_{a=1}^{A} a_a^2$; $\sigma_B^2 = \frac{1}{B-1}\sum_{b=1}^{B} b_b^2$; $\sigma_{AB}^2 = \frac{1}{(A-1)(B-1)}\sum_{a=1}^{A}\sum_{b=1}^{B} ab_{ab}^2$.

Pela inspeção da Tabela 2, observa-se que os componentes de variância que contribuem para a variância do estimador do contraste genérico dos parâmetros a_a aparecem apenas no *E*QM da fonte de variação *AR*, e coincidem exatamente com esse *E*QM. (A fonte de variação *A* não é considerada, pois ela inclui σ_A^2 , que não aparece na variância do contraste.) Nesse caso, conclui-se que a combinação linear de *E*QM que compõe a Var($\hat{C}$) é, simplesmente:

$$\mathrm{Var}(\hat{C}) = \frac{\sum c_a^2}{BR}[E\mathrm{QMAR}].$$

A expressão do estimador da variância do contraste $\hat{C}$ é, então, obtida pela substituição dos *E*QM pelos respectivos QM.

A ilustração do procedimento é estendida, a seguir, para todos os contrastes de parâmetros de interesse do delineamento blocos casualizados com parcelas divididas, com os fatores *A* e *B* fixos, em parcelas e subparcelas, respectivamente, e fator de bloqueamento *R*.

Em experimentos com esse delineamento, freqüentemente, são de interesse os seguintes contrastes de parâmetros dos efeitos fixos:

$a_a - a_{a'}$: diferença de parâmetros para dois níveis do fator *A*;

$b_b - b_{b'}$: diferença de parâmetros para dois níveis do fator *B*;

$ab_{ab} - ab_{ab'}$: diferença de parâmetros para dois níveis do fator *B* em um mesmo nível do fator *A*;

$ab_{ab} - ab_{a'b}$: diferença de parâmetros para dois níveis do fator *A* em um mesmo nível de *B*.

As expressões das variâncias dos estimadores desses quatro contrastes genéricos podem ser derivadas simultaneamente, seguindo os passos do procedimento geral, com o auxílio das Tabelas 2 e 3.

TABELA 3. Derivação das variâncias dos estimadores dos contrastes genéricos dos parâmetros do delineamento blocos casualizados com parcelas divididas com dois fatores fixos *A* e *B*, em parcelas e subparcelas, respectivamente.

Termo	$e_{b(ar)}$	$u_{a(r)}$	ar_{ar}
Variância	σ_E^2	σ_U^2	σ_{AR}^2
Divisor:			
$a_a - a_{a'}$	BR	R	R
$b_b - b_{b'}$	AR	-	-
$ab_{ab} - ab_{ab'}$	R	-	-
$ab_{ab} - ab_{a'b}$	R	R	R

Os índices variáveis nos quatro contrastes genéricos são: a e b. Os termos aleatórios do modelo são: $e_{b(ar)}$, $u_{a(r)}$ e ar_{ar}. As variâncias desses termos e os correspondentes divisores de que tratam os Passos 1 e 2 constam da Tabela 3.

As variâncias dos estimadores dos contrastes são derivadas, a seguir, de acordo com os Passos 3 e 4. Obtém-se, para o primeiro contraste:

$$\mathrm{Var}(\hat{a}_a - \hat{a}_{a'}) = 2\left[\frac{\sigma_E^2}{BR} + \frac{\sigma_U^2}{R} + \frac{\sigma_{AR}^2}{R}\right]$$

$$= \frac{2}{BR}\left[\sigma_E^2 + B\sigma_U^2 + B\sigma_{AR}^2\right].$$

Pela inspeção da Tabela 2, pode-se verificar que:

$$\mathrm{Var}(\hat{a}_a - \hat{a}_{a'}) = \frac{2}{BR}\,E\mathrm{QMAR}.$$

Procedendo de modo semelhante para cada um dos demais contrastes, obtém-se:

$$\mathrm{Var}(\hat{b}_b - \hat{b}_{b'}) = \frac{2}{AR}\;E\mathrm{QME};$$

$$\mathrm{Var}(\hat{ab}_{ab} - \hat{ab}_{ab'}) = \frac{2}{R}\;E\mathrm{QME};$$

$$\mathrm{Var}(\hat{ab}_{ab} - \hat{ab}_{a'b}) = 2\left[\frac{\sigma_E^2}{R} + \frac{\sigma_U^2}{R} + \frac{\sigma_{AR}^2}{R}\right]$$

$$= \frac{2}{BR}\left[B\sigma_E^2 + B\sigma_U^2 + B\sigma_{AR}^2\right]$$

$$= \frac{2}{BR}\left[(B-1)\sigma_E^2 + \sigma_E^2 + B\sigma_U^2 + B\sigma_{AR}^2\right]$$

$$= \frac{2}{BR}\left[(B-1)E\mathrm{QME} + E\mathrm{QMAR}\right].$$

Essas são as expressões das variâncias dos contrastes de estimadores de parâmetros de interesse mais comum em experimentos com delineamento blocos casualizados com parcelas divididas, que se encontram em alguns textos (Federer, 1955; Steel & Torrie, 1980; Milliken & Johnson, 1992).

Observa-se que para o delineamento blocos casualizados, com parcelas divididas, com ambos fatores em parcelas e em subparcelas fixos e fator de bloqueamento aleatório, a consideração dos fatores de unidade não altera as expressões das variâncias dos contrastes de usual interesse.

REFERÊNCIAS

FEDERER, W.T. **Experimental design:** Theory and application. New York: Macmillan, 1955. 538p.

HICKS, C.R. **Fundamental concepts in the design of experiments**. 2.ed. New York: Holt, Rinehart and Winston, 1973. 349p.

MILLIKEN, G.A.; JOHNSON, D.E. **Analysis of messy data**. New York: Chapman & Hall, 1992. 473p.

MONTGOMERY, D.C. **Design and analysis of experiments**. 2.ed. New York: John Wiley, 1984. 538p.

SAS INSTITUTE. The MIXED procedure. In: SAS INSTITUTE. **SAS/STAT software:** changes and enhancements, release 6.07. Cary, NC, 1992. ch.16, p.287-366. (SAS Technical Report P-229).

SILVA, J.G.C. da. A consideração da estrutura das unidades em inferências derivadas do experimento. **Pesquisa Agropecuária Brasileira**, Brasília, v.34, n.6, p.911-925, jun. 1999.

STEEL, R.G.D.; TORRIE, J.H. **Principles and procedures of statistics:** A biometrical approach. 2.ed. New York: McGraw-Hill, 1980. 633p.

ZONTA, E.P. Método prático para a determinação da variância do contraste entre duas médias de tratamentos em experimentos balanceados. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 3. 1989, Lavras. **Anais**... Lavras: ESAL, 1989. p.253-266.