Ministério da Agricultura e do Abastecimento

Embrapa

Pesquisa em andamento

Número 33 | 4p. | 100 exemplares | dez./1999 | ISSN 1517-4921

UM ALGORITMO PARA CONSTRUÇÃO DE DELINEAMENTOS EFICIENTES

Lúcio José Vivaldi

Nos delineamentos em blocos incompletos, o número de parcelas por bloco é menor do que o de tratamentos e, como conseqüência, há uma perda de informação porque alguns contrastes entre tratamentos estão confundidos (parcial ou totalmente) com efeitos de blocos. Espera-se que essa perda seja compensada por um ganho em precisão pelo uso de blocos menores. Os delineamentos em blocos incompletos, por sua vez, são comparados com os em blocos casualizados, desde que estes não apresentam confundimento e são de fácil análise; a comparação é feita mediante definições de eficiência de um delineamento. No que vem a seguir, b é o número de blocos, v o número de tratamentos, k o tamanho de cada bloco (número de parcelas por bloco), r o número de repetições e p_{ih} é o número de blocos em que os tratamentos i e h aparecem juntos. Também, neste trabalho serão considerados:

1) Os delineamentos em que um tratamento ocorre uma vez no bloco ou está ausente, chamados de delineamentos binários;

2) Os delineamentos resolvíveis, em que cada bloco tem k parcelas e s blocos formam uma repetição, tal que $v = ks$;

2) Por último, somente serão vistos os delineamentos conectados que permitem a estimação de qualquer contraste entre tratamento.

O modelo matemático para o delineamento em blocos incompletos é

$$\mathbf{Y} = \mathbf{1}\mu + \mathbf{F}\tau + \mathbf{L}\beta + \mathbf{e},\quad E(\mathbf{Y}) = \mathbf{1}\mu + \mathbf{F}\tau + \mathbf{L}\beta \tag{1}$$

$$\text{Var}(\mathbf{Y}) = \mathbf{I}\sigma^2 \tag{2}$$

onde E é o operador de esperança matemática; Y (n x 1) é o vetor das observações, μ é a média geral, 1 (n x 1) é um vetor de **1 s**, **F** (n x v) é a matriz de incidência dos tratamentos nas parcelas; τ (v x 1) é o vetor de efeitos de tratamento; L (n x b) é a matriz de incidência dos blocos; β (b x 1) é o vetor de efeitos de blocos; **e** (n x 1) é o vetor dos erros; Var significa variância; **I** é uma matriz identidade; σ^2 é o erro experimental. Seja

$$\mathbf{N} = \mathbf{F'L} \tag{3}$$

a matriz de incidência dos tratamentos nos blocos, seguindo a notação de John (1971) e aplicando o método dos quadrados mínimos ao modelo (1) tem-se o sistema de equações :

$$\mathbf{C}\hat{\tau} = \mathbf{Q} \tag{4}$$

$$\mathbf{C} = r\mathbf{I} - \frac{1}{k}\mathbf{NN'} \tag{5}$$

[1] Pesquisador da Embrapa Cerrados.

$$Q = (F' - \frac{1}{k}F'LL')Y \qquad (6)$$

A matriz C (v x v) é a matriz de informação de tratamentos, é da forma **X'X**, tem posto igual a (v-1) e suas linhas são contrastes. Uma solução para a equação (4) é dada por

$$\hat{\tau} = C^{-}Q \qquad (7)$$

onde $\mathbf{C}^{-}$ é uma inversa generalizada de **C**. A variância de $\mathbf{p}'\hat{\tau}$ a estimativa de um contraste entre tratamentos, é dada por

$$\mathrm{Var}(\mathbf{p}'\hat{\tau}) = \mathbf{p}'\mathbf{C}^{-}\mathbf{p}\sigma^2 \qquad (8)$$

Uma das formas de se estudar eficiência é tomar como base a comparação entre dois tratamentos e existem $\frac{v(v-1)}{2}$ diferentes comparações desse tipo. A média das variâncias dessas comparações é dada por :

$$VM = \frac{v\sum_{1}^{v-1}\frac{1}{\lambda_i}}{\frac{v(v-1)}{2}}\sigma^2 = 2\frac{\sum_{1}^{v-1}\frac{1}{\lambda_i}}{v-1}\sigma^2 \qquad (9)$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_{v-1}$ são os autovalores positivos da matriz **C**. A variância da diferença entre dois tratamentos num delineamento em blocos ao acaso com *r* blocos e *v* tratamentos é dada por:

$$\mathrm{Var}(\hat{t}_i - \hat{t}_h) = \frac{2\sigma^2}{2}; \qquad (10)$$

assim , dividindo a expressão (10) por (9) e assumindo que σ^2 é o mesmo nas duas situações, tem-se uma definição de eficiência, EF, de um bloco incompleto em relação aos blocos casualizados, onde :

$$EF = \frac{v-1}{r\sum_{1}^{v-1}\frac{1}{\lambda_i}} \qquad (11)$$

conhecido com fator de eficiência da média harmônica (harmonic mean efficient factor – HMEF; ver , por exemplo John 1971). No caso de blocos incompletos balanceados, $p_{ih} = p$ (p_{ih} é o número de vezes em que os tratamentos (i h) aparecem juntos em um bloco), para todo par (*i h*) de tratamentos e referido por Yates (1936).

$$HMEF = \frac{pv}{rk} \qquad (12)$$

Os delineamentos em blocos incompletos balanceados, entretanto, não atendem às exigências dos pesquisadores, porque para um grande número de combinações *r,v,k* não existe o delineamento; em conseqüência passou-se a estudar com mais detalhes os delineamentos em blocos incompletos parcialmente balanceados. A dificuldade de se conseguir um delineamento eficiente dentro dessa classe pode ser ilustrada quando r=4, v=20 e k=10; o total de delineamentos é de aproximadamente 10^{21} e entre eles está o mais eficiente (ou os mais eficientes). Calcular HMEF para cada um e escolher o melhor é inviável. Uma solução para esse problema é estabelecer um processo de busca de um delineamento que tenha EF próximo do máximo; nessa direção surgem duas perguntas:

1) Qual será o valor máximo de EF para a coleção de delineamentos *r*, *k* e *v*?

2) Como procurar pelo melhor?

Em relação à primeira, a estratégia utilizada com maior freqüência é a de obter um limite superior da coleção de delineamentos para EF e ao longo do tempo foram surgindo limites cada vez mais refinados ; no que tange à segunda pergunta, respostas têm sido dadas mediante os algoritmos de construção e de busca. Entre os limites de EF, será usado o de o Williams & Patterson (1977) , onde para um delineamento resolvível (*r*, *v*, *s*, *k*), tem-se que:

$$EF \leq \left(\frac{1}{U_Q} + \frac{\beta^2 r^2 d(s-d)}{v-1} + \frac{\beta^3 rz}{(r-1)(v-1)} \right)^{-1} \qquad (13)$$

Os parâmetros de (13), não definidos aqui, estão no trabalho de Williams & Patterson (1977).

Entre os métodos de construção de blocos incompletos resolvíveis , será utilizado o método proposto por Patterson & Williams (1976) (outros métodos ainda serão estudados) e que os autores chamaram de "latices generalizados"; são delineamentos semelhantes aos latices de Yates (1936), porém, mais flexíveis. O primeiro algoritmo tem as seguintes fases:

1) Construir uma matriz , chamada ALFA, (k x r), onde os elementos de ALFA pertencem ao conjunto $\{0,1,2,..., s-1\}$;

2) Cada coluna de ALFA é usada para gerar ($s-1$) outras colunas, através de substituição cíclica; a matriz resultante tem dimensão (k x rs) , chamada BETA;

3) Adicionar s a todos os elementos da linha 2 de BETA, $2s$ a todos os elementos da linha 3 de BETA e assim por diante.

4) As colunas são os blocos do delineamento e os tratamentos são os números de 0 a ($v-1$).

Os delineamentos gerados podem ser os $\alpha(0,1)$, onde $p_{ih} = 0$ $p_{ih} = 1$ ou, e os $\alpha(0,1,2)$. A chave do algoritmo está na escolha da matriz ALFA; algumas delas estão no trabalho dos autores.

Este algoritmo foi implementado em linguagem SAS (1997), e utilizado para gerar delineamentos eficientes para o subprojeto 04.0.94.321-25, ano 1999/2000. Um delineamento quase ótimo (EF=0.8965) gerado consta na Tabela 1; nas colunas da tabela estão os tratamentos.

TABELA 1. Delineamento gerado pelo algoritmo.

Repetição 1 bloco			Repetição 2 bloco			Repetição 3 bloco		
1	2	3	1	2	3	1	2	3
1	2	3	1	2	3	1	2	3
4	5	6	5	6	4	6	4	5
7	8	9	9	7	8	8	9	7
10	11	12	10	11	12	12	10	11
13	14	15	14	15	13	13	14	15
16	17	18	17	18	16	18	16	17
19	20	21	21	19	20	20	21	19
22	23	24	22	23	24	22	23	24
25	26	27	26	27	25	26	27	25

REFERÊNCIAS BIBLIOGRÁFICAS

WILLIAMS, E.R. ; PATTERSON, H.D. Upper bounds for efficiency factors in block designs. **Australian Journal of Statistics**, Canberra, v.19, p.194-201, 1977.

PATTERSON, H.D.; WILLIAMS, E.R. A new class of resolvable incomplete block design. **Biometrika**, London, v.63, n.1, p.82-92, 1976.

YATES, F. A new method of arranging variety trials involving a large number of varities. **Journal of Agricultural Science**, Cambridge, v.26, p.424-455, 1936.

JOHN, P.W.M. **Statistical design and analysis of experiments**. London, Macmillam,1971.

SAS INSTITUTE INC (Cary, NC). **Release 6.12**. Cary, 1997.

Embrapa

Empresa Brasileira de Pesquisa Agropecuária
Embrapa Cerrados
Ministério da Agricultura e do Abastecimento
BR 020, km 18, Rodovia Brasília/Fortaleza, Caixa Postal 08223
CEP 73301-970, Planaltina, DF
Telefone: (61) 388-9898 FAX: (61) 388-9879