# NOVOS ASPECTOS DO PROBLEMA DO TAMANHO ÓTIMO DAS PARCELAS EM EXPERIMENTOS COM PLANTAS ARBÓREAS[1]

FREDERICO PIMENTEL GOMES[2]

RESUMO - Em trabalho anterior, introduziu-se novo método de estimação do tamanho ótimo de parcelas experimentais para plantas arbóreas. Esse método, que leva em conta as bordaduras e utiliza o coeficiente de correlação intraclasse ($\rho$) entre árvores úteis dentro das parcelas, define como tamanho ótimo o número k de árvores úteis que minimize a variância da média de cada tratamento, para um número total de árvores (N), considerado fixo. No artigo referido, esse tamanho ótimo foi determinado, para parcelas com meia-bordadura ou com bordadura completa, com uma ou com duas linhas de árvores úteis. No presente artigo, o problema é estudado de modo mais geral, buscando o tamanho ótimo quando varia tanto o número de linhas úteis (n) como o de árvores úteis (k) por parcela. No caso de meia-bordadura, o número ótimo de linhas úteis é $n = \sqrt[3]{(1-\rho)/\rho}$, e no caso de bordadura completa, $n = \sqrt[3]{2(1-\rho)/\rho}$, sendo sempre $k = n^2$, $\rho > 0$. Por outro lado, demonstra-se que, no caso de bordadura completa quando se passa de um ensaio com k árvores úteis em n linhas, com N árvores totais por tratamento numa área A, para outro semelhante com essas características indicadas por k', n', N' e A', sem mudar a variância da média de cada tratamento, temos:

$$\frac{A'}{A} = \frac{N'}{N} = \frac{(1 + 2/n')\,(1 + 2n'/k')\,[1 + (k'-1)\rho]}{(1 + 2/n)\,(1 + 2n/k)\,[1 + (k-1)\rho]}.$$

Termos para indexação: bordaduras, coeficiente de correlação intraclasse.

## A NEW CONTRIBUTION ON THE PROBLEM OF PLOT SIZE IN EXPERIMENTS WITH TREES

ABSTRACT - In a previous paper, a new method was proposed for the estimation of the optimum plot size of experimental plots for trees. The method, which takes in consideration the guard rows and uses the intraclass coefficient of correlation ($\rho$) among test trees within plots, defines as optimum size the number k of test trees which minimizes the variance of the estimate of a treatment mean for a fixed total number (N) of trees per treatment. In that previous paper the optimum size was determined, for plots with one half or complete guard rows, with one or two test rows. In the present paper, the problem is generalized, the optimum size being searched when both the number of test trees (k) and the number of test rows (n) are variable. For the case of one half guard row the optimum number of test rows in $n = \sqrt[3]{(1-\rho)/\rho}$, and for the case of a complete guard row is $n = \sqrt[3]{2(1-\rho)/\rho}$, always with $k = n^2$, $\rho > 0$. On the other hand, it is shown that, with complete guard rows, when changing a trial with k test trees per plot in n rows, with total number of trees per treatment N and area A, into another one with these parameters indicated by k', n', N' and A', keeping constant the variance of the estimate of a treatment mean, we have:

$$\frac{A'}{A} = \frac{N'}{N} = \frac{(1 + 2/n')\,(1 + 2n'/k')\,[1 + (k'-1)\rho]}{(1 + 2/n)\,(1 + 2n/k)\,[1 + (k-1)\rho]}.$$

Index terms: guard rows, intraclass coefficient of correlation.

## INTRODUÇÃO

Em trabalho anterior, Pimentel-Gomes (1984) introduziu novo método para a determinação do tamanho ótimo de parcelas experimentais para plantas arbóreas. O método utiliza o coeficiente de correlação intraclasse ($\rho$) relativo às árvores úteis dentro de cada parcela, e define como tamanho ótimo o número (k) de árvores úteis que minimize a variância da média de um tratamento para um número total de árvores (N), considerado fixo. Isso equivale a minimizar a variância da média de cada tratamento para uma área fixa do ensaio, ou, ao contrário, a tornar mínima a área do ensaio, para obter uma variância dada para a média de cada tratamento.

No trabalho em questão, foram considerados os casos de uma ou de duas linhas de árvores úteis.

[1] Aceito para publicação em 22 de janeiro de 1987.

[2] Eng. - Agr., Dr. em Agronomia, Prof. Catedrático, ESALQ-USP (aposentado), Consultor do IICA/EMBRAPA, CEP 70333, Brasília, DF.

Quando o coeficiente de correlação intraclasse ($\rho$), suposto positivo, não fica próximo de zero ($\rho > 0{,}15$, digamos) as soluções obtidas são excelentes. Mas no caso de valores baixos (positivos) de $\rho$, o número de árvores úteis k da parcela de tamanho ótimo pode crescer e então convém estudar também o número de linhas úteis (n) a ser usado. O presente artigo discute este novo ponto de vista, que aprimora o método em apreço.

Recentemente tomou-se conhecimento de um trabalho de Lin & Binns (1984), que usa também o coeficiente de correlação intraclasse para a determinação do tamanho ótimo de parcelas experimentais. Mas esse trabalho, que não cuida especificamente de experimentos com árvores e baseia a discussão no coeficiente de variação, no parâmetro b de Smith (1938) e no método de Koch & Rigney (1951) para estimação de b, não leva em conta as bordaduras, nem a variação dentro das parcelas.

## TEORIA

Em trabalho anterior (Pimentel-Gomes 1984) foi demonstrado que a variância da média de um tratamento é dada pela expressão:

$$V(\hat{m}) = (\sigma^2/kr)\,[1 + (k-1)\rho],$$

onde r é o número de repetições, k é o número de árvores úteis por parcela, $\rho$ é o coeficiente de correlação intraclasse entre árvores úteis dentro das parcelas, e $\sigma^2$ é a variância residual (de parcelas), quando se toma $k = 1$.

A extensão da teoria anteriormente exposta só tem razão de ser quando as parcelas têm bordadura. Consideraremos dois casos: o de meia bordadura e, principalmente, o de bordadura completa. Por outro lado, salvo menção clara em contrário, só discutiremos o caso de $0 < \rho < 1$.

### Parcelas com meia bordadura

Neste caso o número total (K) de árvores por parcela de k árvores úteis em n linhas é:

$$K = (n + 1)(1 + k/n) = (1 + 1/n)(n + k).$$

Sendo N o número total de árvores por tratamento e r o número de repetições, temos:

$$N = Kr = (1 + 1/n)(n + k)\,r,$$

e a variância da média do tratamento se torna:

$$V(\hat{m}) = (\sigma^2/N)(1 + 1/n)(1 + n/k)\,[1 + (k-1)\rho].$$

Para procurar o mínimo dessa função, igualamos a zero suas derivadas parciais $\partial V/\partial k$ e $\partial V/\partial n$. Temos, assim, as equações:

$$\begin{cases} -n + n\rho + k^2\rho = 0, \\ -k + n^2 = 0, \end{cases}$$

cuja solução é:

$$n = \sqrt[3]{(1-\rho)/\rho}, \qquad k = n^2 \quad (0 < \rho < 1)$$

A matriz hessiana no ponto crítico neste caso é:

$$H = A\begin{bmatrix} 2(n+1)(1-\rho) & -n^2(1-\rho+n^2\rho) \\ -n^2(1-\rho+n^2\rho) & 2n^3(1-\rho+n^2\rho) \end{bmatrix},$$

onde A é uma constante positiva e $n = \sqrt[3]{(1-\rho)/\rho}$. Uma vez que os termos da diagonal principal são positivos para $0 < \rho < 1$, admitida esta hipótese, para provar que a matriz é definida positiva é suficiente demonstrar que seu determinante é positivo. Isto ocorre, pois, como $1 - \rho = n^3\rho$, esse determinante (D) vale:

$$\begin{aligned} D &= (1-\rho+n^2\rho)\,[4(n+1)(1-\rho)n^3 - n^4(1-\rho+n^2\rho) \\ &= (n^3\rho+n^2\rho)\,n^3\,[4(n+1)\,n^3\rho - n(n^3\rho+n^2\rho)] \\ &= 3\,n^8\,(n-1)^2\rho^2 > 0. \end{aligned}$$

Conclui-se, pois, que com $0 < \rho < 1$, temos um ponto de mínimo para os valores de k e de n indicados.

### Parcelas com bordadura completa

Neste caso, temos:

$$K = (n + 2)(2 + k/n) = (1 + 2/n)(2n + k),$$

logo,

$$V(\hat{m}) = (\sigma^2/N)(1 + 2/n)(1 + 2n/k)\,[1 + (k-1)\rho].$$

Igualando-se a zero as derivadas parciais $\partial V/\partial k$ e $\partial V/\partial n$, temos as equações:

$$\begin{cases} k^2\rho - 2n(1-\rho) = 0, \\ -k + n^2 = 0, \end{cases}$$

cuja solução é, com $0 < \rho < 1$:

$$n = \sqrt[3]{2(1-\rho)/\rho}, \qquad k = n^2.$$

## VARIÂNCIA MÍNIMA PARA n FIXO

Quando se fixa *a priori* o número n de linhas úteis, o mínimo de $V(\hat{m})$ se dá para

$$k = \sqrt{n(1-\rho)/\rho},$$

no caso de meia bordadura, e para

$$k = \sqrt{2n(1-\rho)/\rho},$$

no caso de bordadura completa. Essas igualdades continuam válidas quando variam simultaneamente k e n, como se vê pelas equações correspondentes a $\partial V/\partial k = 0$, expostas acima.

## EXEMPLO DE APLICAÇÃO

Tomemos como exemplo um caso de $\rho = 0{,}070$, com bordadura completa. Obtemos, como coordenadas do ponto de mínimo de $V(\hat{m})$:

$$n = \sqrt[3]{2(1 - 0{,}070)/0{,}070} = 2{,}98 \approx 3, k = n^2 = 9.$$

Se fixássemos antecipadamente $n = 1, 2, 3, 4$, teríamos valores mínimos de $V(\hat{m})$ correspondentes aos valores de k indicados a seguir.

$n = 1 \rightarrow k = 5{,}15$, arredondado para 5,
$n = 2 \rightarrow k = 7{,}29$, arredondado para 6 ou 8,
$n = 3 \rightarrow k = 8{,}92$, arredondado para 9,
$n = 4 \rightarrow k = 10{,}30$, arredondado para 8 ou 12,
$n = 6 \rightarrow k = 12{,}61$, arredondado para 12.

Os valores de $V(\hat{m}) = f(n, k)$ serão os seguintes.

| | |
|---|---|
| $f(1, 5) = (\sigma^2/N)\, 5{,}38$ | $f(2, 8) = (\sigma^2/N)\, 4{,}47$ |
| $f(2, 6) = (\sigma^2/N)\, 4{,}50$ | $f(3, 12) = (\sigma^2/N)\, 4{,}43$ |
| $f(3, 9) = (\sigma^2/N)\, 4{,}34$ | $f(2, 12) = (\sigma^2/N)\, 4{,}72$ |
| $f(4, 16) = (\sigma^2/N)\, 4{,}61$ | |

Verifica-se, pois, que realmente o mínimo de $V(\hat{m})$ se dá para k = nove plantas úteis em três linhas. Mas, por outro lado, no caso presente a função varia pouco nas vizinhanças do ponto de mínimo, sendo quase indiferente tomar parcelas de seis plantas úteis em duas linhas, ou de nove plantas úteis em três linhas, ou ainda de 16 plantas úteis em quatro linhas. Mas é claro que as parcelas menores, que exigem maior número de repetições, levam a maior número de graus de liberdade para o resíduo.

## ECONOMIA DE ÁREA

Ao passar de um experimento com k árvores úteis em n linhas para outro com k' árvores úteis em n' linhas, sem mudar a variância da média de um tratamento, devemos ter, no caso de bordadura completa:

$$V(\hat{m}) = (\sigma^2/N)\,(1 + 2/n)\,(1 + 2n/k)\,[1 + (k - 1)\rho]$$
$$= (\sigma^2/N')\,(1 + 2/n')\,(1 + 2n'/k')\,[1 + (k' - 1)\rho],$$

onde $N = rK$, $N' = r'k'$, sendo K e K' os números totais de árvores por parcela, e N e N' os números totais nos ensaios, por tratamento. Temos, portanto:

$$\frac{N'}{N} = \frac{(1 + 2/n')\,(1 + 2n'/k')\,[1 + (k' - 1)\rho]}{(1 + 2/n)\,(1 + 2n/k)\,[1 + (k - 1)\rho]}$$

Como as áreas ocupadas são proporcionais aos números totais de árvores, temos também:

$$\frac{A'}{A} = \frac{(1 + 2/n')\,(1 + 2n'/k')\,[1 + (k' - 1)\rho]}{(1 + 2/n)\,(1 + 2n/k)\,[1 + (k - 1)\rho]}.$$

## O NÚMERO DE REPETIÇÕES

Ao passar de uma parcela com k árvores úteis em n linhas para outra com k' árvores úteis em n' linhas, sem mudar a variância da média de um tratamento, é necessário mudar o número de repetições de r para r'. Temos, porém, $N = rK$, $N' = r'K'$, logo

$$\frac{A}{A'} = \frac{N}{N'} = \frac{rK}{r'K'},$$

isto é

$$\frac{r'}{r} = \frac{K}{K'}\;\frac{A'}{A}.$$

No caso de $\rho = 0{,}30$, por exemplo, se passarmos de uma parcela de $k = 21$ árvores úteis em três linhas para outra de quatro árvores úteis em duas linhas, teremos:

$$\frac{A'}{A} = \frac{N'}{N} = \frac{(1 + 2/2)\,(1 + 2.2/4)\,[1 + 3(0{,}30)]}{(1 + 2/3)\,(1 + 2.3/21)\,[1 + 20(0{,}30)]}$$
$$= 0{,}507 = 50{,}7\%.$$

Haveria, pois, uma redução de 49,3% na área (de 100% para 50,7%).

No primeiro caso, a parcela, com 21 árvores úteis em três linhas, tem um total de $K = 5.9 = 45$ árvores. No segundo, esse número baixa para $K' = 4.4 = 16$ árvores. Com a parcela menor, há economia de área, como vimos, mas aumenta um pouco o número de repetições, pois temos:

$$\frac{r'}{r} = \frac{K}{K'}\;\frac{A'}{A} = (45/16)\,0{,}507 = 1{,}42,$$

Isto é, $r' = 1{,}42\,r$. Assim, se o ensaio de parcela grande tivesse cinco blocos, o de parcela pequena deverá ter 7,1 ou 7 blocos. Mas este aumento de 42% no número de repetições é mais do que compensado pela diminuição da área de cada parcela experimental, de tal sorte que a área total por tratamento cai, em percentagem, de 100 (correspondente a 225 árvores) para 50 (correspondente a 112 árvores).

## O CASO DE $\rho < 0$

No trabalho anterior (Pimentel Gomes 1984), vimos que para $\rho < 0$ a variância $V(\hat{m})$ é função decrescente de k, número de árvores úteis por parcela. Nestas condições, o ideal é tomar k o maior possível, mas compatível com um número razoável de graus de liberdade para o resíduo. No entanto, uma vez fixado o valor de k, o número n de linhas úteis deve ser tão próximo de $\sqrt{k}$ quanto possível, pois o mínimo de $V(\hat{m})$ se dá para $n = \sqrt{k}$. Este resultado, já demonstrado acima, é, por outro lado, bastante intuitivo. Com efeito, em parcelas de k plantas úteis em n linhas, o número total de plantas por parcela é, como vimos:

$K = (1 + 2/n)(2n + k)$,

no caso de haver bordadura completa. A relação g(n) entre o número total de árvores e o número de árvores úteis é:

$g(n) = K/k = (1 + 2/n)(1 + 2n/k)$.

Esta função passa por um mínimo para $k = n^2$, isto é, para $n = \sqrt{k}$. Nestas condições, a fração k/K, das árvores úteis sobre o total de árvores por parcela, é máxima, para K fixo, quando $n = \sqrt{k}$. Ora, isso deve, evidentemente, favorecer a redução de $V(\hat{m})$.

## A VARIÂNCIA DE $\hat{\rho}$

Sendo, numa análise de variância, $V_1$ (com $n_1$ graus de liberdade) o quadrado médio relativo a parcelas, de k árvores úteis, e $V_2$ (com $n_2$ graus de liberdade) o quadrado médio referente a árvores dentro de parcelas, temos (Pimentel-Gomes 1984):

$$\hat{\rho} = (V_1 - V_2)/[V_1 + (k-1)V_2],$$

logo, por diferenciação:

$$d\hat{\rho} = k(V_2\, dV_1 - V_1\, dV_2)/[V_1 + (k-1)V_2]^2.$$

Por outro lado, temos (Pimentel-Gomes 1984):

$$E(V_1) = \sigma^2 [1 + (k-1)\rho], \quad E(V_2) = \sigma^2 (1 - \rho),$$

e portanto:

$$E[V_1 + (k-1)V_2] = k\sigma^2.$$

Tendo em vista o método indicado por Kempthorne (1957), obtemos, pois, a fórmula aproximada:

$$V(\hat{\rho}) = E(d\hat{\rho}^2) = \frac{2(1-\hat{\rho})^2 [1 + (k-1)\hat{\rho}]^2}{k^2} \left[\frac{1}{n_1 + 2} + \frac{1}{n_2 + 2}\right].$$

Com esta estimativa de $V(\hat{\rho})$ podemos calcular intervalos de confiança aproximados para $\rho$.

## REFERÊNCIAS


KEMPTHORNE, O. **An introduction to genetic statistics.** New York, J. Wiley, 1957.

KOCH, E.J. & RIGNEY, J.A. A method of estimating optimum plot size from experimental data. **Agron. J.**, 43:17-21, 1951.

LIN, C.S. & BINNS, M.R. Working rules for determining the plot size and numbers of plots per block in field experiments. **J. Agric. Sci.**, 103:11-5, 1984.

PIMENTEL-GOMES, F. O problema do tamanho das parcelas em experimentos com plantas arbóreas. **Pesq. agropec. bras.**, 19(12):1507-12, 1984.

SMITH, H.F. An empirical law describing heterogeneity in the yield of agricultural crops. **J. Agric. Sci.**, 28: 1-23, 1938.