



ISSN 0100-3453

CIRCULAR TÉCNICA Nº 178

MAIO 1991

O ÍNDICE DE VARIAÇÃO, UM SUBSTITUTO VANTAJOSO DO COEFICIENTE DE VARIAÇÃO

Frederico Pimentel Gomes*

INTRODUÇÃO

Chama-se coeficiente de variação (CV) de um experimento a estatística:

$$CV = \frac{s}{\hat{m}} 100,$$

onde \hat{m} é a estimativa da média e s é a estimativa do desvio padrão. É idéia antiga e muito difundida a de que o coeficiente de variação “dá uma idéia da precisão do experimento” (PIMENTEL-GOMES, 1987, p.7). É claro que, em igualdade de condições, é mais preciso o experimento com menor coeficiente de variação. Mas é necessário especificar melhor em que consistem essa igualdade de condições. Uma delas seria a igualdade do número de repetições. Com efeito, se num experimento tivermos $CV_1 = 8\%$, com $r_1 = 4$ repetições, e noutro similar $CV_2 = 12\%$, com $r_2 = 16$ repetições, o segundo na verdade é mais preciso do que o primeiro. Realmente, para o primeiro o erro padrão da média de um tratamento, em porcentagem da média, será

$$s(m_1) = \frac{8}{\sqrt{4}} = 4\%,$$

e para o segundo

* Consultor do IPEF

$$s(m_2) = \frac{12}{16} = 3\%$$

Estes valores, que influem diretamente no intervalo de confiança para cada média, no teste F e nos testes de comparação de médias, mostram claramente que é o segundo experimento, de coeficiente de variação mais alto, o que tem maior precisão.

Uma vez que o número de repetições influi na correta avaliação da precisão dos ensaios, uma estatística que o levasse em conta poderia ser mais conveniente do que o coeficiente de variação. Daí surgiu a idéia do índice de variação de um experimento, o qual é definido e estudado neste artigo.

O índice de variação (IV)

Consideremos um ensaio padrão s e r repetições para cada tratamento. Sendo \hat{m}_i a estimativa da média de um tratamento qualquer, o erro padrão desta estimativa é:

$$s(\hat{m}_i) = \frac{s}{\sqrt{r}}$$

Denominaremos índice de variação (IV) do experimento esse erro padrão expresso em porcentagem da estimativa \hat{m} da média geral do experimento, isto é:

$$IV = \frac{s(\hat{m}_i)}{\hat{m}} \times 100 = \frac{s}{\hat{m}\sqrt{r}} \times 100 = \frac{CV}{\sqrt{r}}$$

Nos experimentos em blocos completos, e nos experimentos em blocos incompletos equilibrados (ou balanceados) e também nos reticulados (ou látices) quadrados, todos os tratamentos e caracterizará, pois, por si só, o experimento.

Já nos ensaios inteiramente casualizados, pode haver números distintos de repetições para os diversos tratamentos. Neste caso, o índice de variação poderá variar de um tratamento para outro. Mas isto é pouco comum. Por outro lado, nos experimentos em blocos casualizados com alguns tratamentos comuns (PIMENTEL-GOMES, 1987), há dois tipos de tratamentos: os regulares, cada um com r repetições, e os comuns (ou testemunhas), com gr repetições. Há, pois, neste caso, dois índices de variação:

$$IV = \frac{s}{\hat{m}\sqrt{r}} = \frac{CV}{\sqrt{r}}$$

para os tratamentos regulares, e

$$IV = \frac{s}{\hat{m}\sqrt{gr}} = \frac{CV}{\sqrt{gr}}$$

para os tratamentos comuns.

Interpretação do índice de variação

Consideramos por exemplo um experimento com 8 tratamentos (1, 2,..., 8), em 4 blocos completos casualizados, com a seguinte análise de variância:

C. Variação	G.L.	Q.M.	F
Blocos	3	16,84	1,97
Tratamentos	7	131,31	15,37**
Resíduo	21	8,55	-
Total	31	$s = \sqrt{8,55}$	2,92

Se a média geral for $\hat{m}_i = 19,70$, o coeficiente de variação será:

$$CV = \frac{2,92 \times 100}{19,70} = 14,8\%$$

O índice de variação será:

$$IV = \frac{CV}{\sqrt{r}} = \frac{14,8}{\sqrt{4}} = 7,41\%$$

Para maior facilidade de compreensão, consideremos uma média de tratamento \hat{m}_i , de valor igual ao da média geral m . O intervalo de confiança para essa média de tratamento (\hat{m}_i) é dado pela expressão.

$$\hat{m}_i \pm t \frac{s}{\sqrt{r}}$$

ou, em porcentagem da média $\hat{m}_i = \hat{m}$:

$$\frac{\hat{m}_i}{\hat{m}_i} \times 100 \pm t \frac{s}{\hat{m}_i \sqrt{r}} \times 100 = 100\% \pm t IV$$

No caso presente, esta expressão se torna:

$$100\% \pm 2,08 \times 7,4\% = 100\% \pm 15,4\%,$$

onde o valor de $t = 2,08$ foi tirado da tabela, apropriado ao nível de 5% de probabilidade e com 21 graus de liberdade. O intervalo de confiança tem, pois, os extremos

$$(100 - 15,4; 100 + 15,4) = (84,6\%; 115,5\%)$$

Nestas condições, o intervalo (84,9%; 115,4%) deverá conter, em 95% dos casos, a verdadeira média \hat{m}_i do tratamento.

O comprimento C desse intervalo é:

$$\begin{aligned} C &= 2 \times t \times IV = 2 \times 2,08 \times 7,4 \\ &= 2 \times 15,4\% \\ &= 30,8\% \end{aligned}$$

Este valor, relativamente alto, mostra que o experimento é de precisão razoavelmente baixa.

Imaginemos, porém, que tivéssemos 16 repetições, em vez de 4. O índice de variação seria:

$$IV = \frac{14,8}{\sqrt{16}} = 3,7\%$$

e o comprimento C do intervalo de confiança baixando para a metade, isto é, para 15,4%.

Por outro lado, a diferença mínima significava, pelo teste de Tukey,

$$\Delta = q \frac{s}{\sqrt{r}},$$

daria, em porcentagem da média geral \hat{m} :

$$\Delta = q \frac{s \times 100}{\hat{m} \sqrt{r}} = q \frac{CV}{\sqrt{r}} = q IV$$

Para o teste de Duncan (PIMENTEL-GOMES, 1987) a situação é semelhante, pois temos:

$$D\% = Z \frac{CV}{\sqrt{r}} = Z IV$$

Nota-se que tal como acontece com o coeficiente de variação (CV), cresce a precisão do ensaio quando diminui o índice de variação.

O Problema do Tamanho da Parcela

Nos experimentos florestais as parcelas são evidentemente muito maiores do que nos ensaios com plantas pequenas, como o milho ou o feijão. O mesmo acontece com as parcelas de árvores frutíferas, como a laranjeira ou a mangueira. Como consequência, o tamanho da parcela, que é de pequeno interesse na experimentação com plantas pequenas, torna-se de grande importância quando se passa a trabalhar com árvores.

A experimentação e a teoria demonstram que, na quase totalidade dos casos, o coeficiente de variação decresce quando aumenta o tamanho das parcelas. (PIMENTEL-GOMES, 1984). Isto levou muitos experimentadores a preferir parcelas grandes, para trabalhar com coeficiente de variação menor, sem perceber que as parcelas excessivamente grandes, que acarretam necessariamente menor número de repetições, podem reduzir a precisão do experimento.

Quer pela teoria de SMITH (1938) quer pela de PIMENTEL-GOMES (1984-1988), demonstra-se que, para parcelas sem bordadura, e em experimentos de mesma área, a variância se dá para parcelas de uma só árvore. Ora, se diminui essa variância, também diminui o erro padrão respectivos $s(\hat{m})$ e, pois, se reduz o índice de variação. Isto mostra, pois, que são vantajosas as parcelas pequenas, apesar de levarem a coeficiente de variação maior.

No caso de haver bordadura, a situação é mais complexa e só raramente se recomenda o uso de uma só árvore útil por parcela. O problema é discutido, com detalhes, por PIMENTEL-GOMES(1984, 1988, 1989). Em qualquer caso, porém, o fator decisivo é o valor do índice de variação.

REFERÊNCIAS BIBLIOGRÁFICAS

PIMENTEL-GOMES, F. **Curso de estatística experimental**. 12. ed. São Paulo, Nobel, 1987. 466p.

PIMENTEL-GOMES, F. Novos aspectos do tamanho ótimo das parcelas em experimentos com plantas arbóreas. **Pesquisa agropecuária brasileira**, Brasília, **23**(1): 59-62, 1988.

PIMENTEL-GOMES, F. O problema do tamanho das parcelas em experimentos com plantas arbóreas. **Pesquisa agropecuária brasileira**, Brasília, **19**(12): 1507-12, 1984.

PIMENTEL-GOMES, F.; ROSSETTI, A.G. & VIÉGAS, R.M.F. Tamanho ótimo de parcelas pra experimentação com seringueira. **Pesquisa agropecuária brasileira**, Brasília, **24**(8): 1021-6, 1989.

SMITH, H.F. Na empirical law describing heterogeneity in the yields of agricultural crops. **Journal of agricultural science**, 28: 1-23, 1938.