FLORESTA 27(1/2):71-82

Ganhos de precisão na alocação ótima em estratificação volumétrica de florestas naturais e plantações florestais

SYLVIO PÉLLICO NETTO*
CARLOS ROBERTO SANQUETTA**

RESUMO

Esta pesquisa apresenta a conceituação teórica para se avaliar os ganhos em precisão, quando uma população florestal é estratificada usando-se a variável volume como referência para a separação das sub-populações, ou estratos. A metodologia usada foi ilustrada com a estratificação das florestas naturais de araucária no Estado do Paraná, onde foi possível demonstrar concretamente o conteúdo teórico, em termos práticos, dos ganhos em precisão, quando a alocação ótima é aplicada, ou quando a amostragem é alocada proporcional à variância e custos dos estratos. Conclusivamente, foi apresentado quantitativamente os ganhos em precisão, quando as médias e variâncias dos estratos mostram-se ser diferentes estatisticamente.
Palavras-chave: Amostragem estratificada, inventário florestal

ABSTRACT

Precision gain in the optimum allocation in volumetric stratification of natural and man-made forests. This paper presents a theory to evaluate gain in precision in forest inventory when a population is divided in strata on the basis of the variable timber volume. The method proposed was illustrated through the stratification of a natural araucaria forest in Parana State, where it was possible to have a real example of gain in precision when the optimum allocation is applied, i.e., when the sample is allocated proportional to the variance and the costs of each stratum. Conclusively, it was shown the gain in precision when the stratified means and variances are statistically different.
Key words: Stratified sampling, forest inventory

*Eng. Flor., M Sc, Dr., Professor Sênior do Curso de Pós-Graduação em Engenharia Florestal da UFPR, Professor Adjunto da PUC-PR, Bolsista do CNPq
**Eng. Flor., M Sc, Dr., Professor de Biometria e Inventário Florestal, Departamento de Ciências Florestais da UFPR, Bolsista do CNPq

INTRODUÇÃO

A estratificação de uma população florestal, utilizando-se a variável volume, normalmente pode ser eficiente se esta é estruturada apropriadamente. Assim, se a amostragem for alocada proporcional ao tamanho dos estratos, a redução da variância da média ocorre apenas pela redução do componente de variação entre as médias dos estratos. Se, entretanto, for possível conhecer as variâncias dentro dos estratos e a amostragem for alocada proporcional a elas, então a redução da variância da média será máxima e conhecida na literatura como alocação ótima.

Em muitas circunstâncias, não se consegue ter nas populações florestais todos os conhecimentos dos estratos, para se conseguir todas as reduções no cálculo da precisão resultante da amostragem estratificada. Ocorre até mesmo situações em que a alocação da amostragem não é feita com proporcionalidade ao tamanho dos estratos, dado estes não serem precisamente conhecidos previamente. Este é caso de não existirem mapas florestais no momento em que se distribui as unidades amostrais para a realização do inventário florestal.

O que se pretende apresentar neste trabalho é a formulação teórica e operacional dos ganhos de precisão dos estimadores, quando se pode aplicar a alocação ótima das unidades que compõem a amostragem nos respectivos estratos.

COMPONENTES E EFEITOS DOS DESVIOS DOS ESTIMADORES DE PRECISÃO EM RELAÇÃO À ALOCAÇÃO ÓTIMA

Num inventário florestal em que se aplicará a amostragem estratificada com alocação ótima, pode-se assumir que a precisão será máxima, ou seja o erro padrão da estimativa da média estratificada será mínimo.

A derivação da expressão que permite calcular a variância da média na estratificação volumétrica encontra-se apresentada em COCHRAN (1977) e resultou em:

$$s_{\bar{x}e}^2 = \sum_{h=1}^{L} \frac{W_h^2 \cdot s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h \cdot s_h^2}{N} \tag{1}$$

onde:

$s_{\bar{x}e}^2 =$ variância da média estratificada;

$W_h =$ peso dos estratos;

$s_h^2 =$ variância dos estratos;

$N =$ número total de unidades amostrais na população;

$n_h =$ número de unidades amostradas nos estratos.

Para se minimizar a variância da média conforme definida em (1), TSCHUPROW (1923) desenvolveu a prova para a alocação da amostragem proporcional à variância dentro dos estratos, ou seja:

$$n_h = n \frac{W_h \cdot s_h}{\sum_{h=1}^{L} W_h \cdot s_h} \tag{2}$$

Esta formulação encontra-se também na literatura citada como desenvolvida por NEYMAN (1934).

Fixando-se *n*, ou seja, a intensidade de amostragem a ser alocada na população total e substituindo n_h na fórmula (1), obtém-se a formulação para se calcular a variância da média mínima na estratificação volumétrica.

$$s^2_{\bar{x}e(min)} = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 - \sum_{h=1}^{L} \frac{W_h \cdot s_h^2}{N} \tag{3}$$

Suponha agora, que a alocação da amostragem foi feita na população subjetivamente, ou até mesmo com alocação apenas proporcional à grandeza dos estratos. Nessas condições, então, a perda de precisão será obtida pela diferença entre as variâncias definidas nas formulações (3) e (1).

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \sum_{h=1}^{L} \frac{W_h^2 \cdot s_h^2}{\hat{n}_h} - \frac{1}{n}\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 \tag{4}$$

onde:

$\hat{n}_h =$ alocação real feita nos estratos.

Se for considerado o termo $W_h^2 \cdot s_h^2$ da fórmula (4) e substituindo o resultado na fórmula (2), onde n_h é a intensidade amostral ótima por estrato, que para fins diferenciais será redefinido como n'_h, tem-se:

$$n'_h = n \frac{(W_h \cdot s_h)}{\sum_{h=1}^{L} W_h \cdot s_h}$$

e

$$n'_h \sum_{h=1}^{L} W_h \cdot s_h = n(W_h \cdot s_h)$$

ou elevando-se ambos os termos ao quadrado tem-se:

$$n'^2_h \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 = n^2 (W_h \cdot s_h)^2$$

Dividindo-se ambos os termos por $\hat{n}_h$ e somando-se para todos os estratos tem-se:

$$\sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 = n^2 \frac{(W_h \cdot s_h)^2}{\hat{n}_h}$$

ou isolando-se o termo

$$\sum_{h=1}^{L} \frac{(W_h \cdot s_h)^2}{\hat{n}_h}$$

$$\sum_{h=1}^{L} \frac{(W_h \cdot s_h)^2}{\hat{n}_h} = \frac{1}{n^2} \sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 \qquad (5)$$

Substituindo-se (5) em (4), tem-se:

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \frac{1}{n^2} \sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 - \frac{1}{n} \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2$$

Racionalizando-se tem-se:

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2 \left[\frac{1}{n^2} \sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} - \frac{1}{n}\right]$$

Tomando-se o mínimo tem-se:

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2}\left(\sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} - n\right)$$

Como

$$n = \sum_{h=1}^{L} n'_h$$

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2}\left(\sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h} - \sum_{h=1}^{L} n'_h\right)$$

ou ainda

$$s^2_{\bar{x}e} - s^2_{\bar{x}e(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2}\left(\sum_{h=1}^{L} \hat{n}_h - 2\sum_{h=1}^{L} n'_h + \sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h}\right)$$

Esta expansão é possível pelo fato de que

$$\sum_{h=1}^{L} \hat{n}_h = \sum_{h=1}^{L} n'_h = n$$

conforme anteriormente definido.

Uniformizando-se o mínimo para todos os termos da expressão tem-se

$$s^2_{\bar{x}e} - s^2_{\bar{x}(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2}\left(\sum_{h=1}^{L} \frac{\hat{n}^2_h}{\hat{n}_h} - 2\sum_{h=1}^{L} \frac{\hat{n}_h \cdot n'_h}{\hat{n}_h} + \sum_{h=1}^{L} \frac{n'^2_h}{\hat{n}_h}\right)$$

Como se pode observar, o segundo termo tornou-se um quadrado da diferença completo e portanto

$$s^2_{\bar{x}e} - s^2_{\bar{x}(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2} \sum_{h=1}^{L} \frac{\left(\hat{n}_h^2 - 2\hat{n}_h \cdot n'_h + n'^2_h\right)}{\hat{n}_h}$$

ou finalmente

$$s^2_{\bar{x}e} - s^2_{\bar{x}(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n^2} \sum_{h=1}^{L} \frac{\left(\hat{n}_h - n'_h\right)^2}{\hat{n}_h}$$

Considere a equação (3) e despreze o fator de correção de populações finitas. Neste caso

$$s^2_{\bar{x}(min)} = \frac{\left(\sum_{h=1}^{L} W_h \cdot s_h\right)^2}{n}$$

Assim o aumento proporcional da variância da média resulta, em relação à alocação ótima, em:

$$\frac{s^2_{\bar{x}e} - s^2_{\bar{x}(min)}}{s^2_{\bar{x}(min)}} = \frac{1}{n} \frac{\sum_{h=1}^{L} \left(\hat{n}_h - n'_h\right)^2}{\hat{n}_h} \tag{6}$$

Adicionalmente pode-se ainda, a partir dos dados da estratificação, obter a informação da variância da média que seria obtida da população sem subdividi-la em estratos, conforme apresentado por COCHRAN (1977)

$$S^2_{\bar{x}\,aleat} = \frac{N-n}{n(N-1)} \left[\sum_{h=1}^{L} W_h \cdot s_h^2 - \sum_{h=1}^{L} \frac{W_h^2 \cdot s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h \cdot s_h^2}{N} + \sum_{h=1}^{L} W_h \cdot \bar{x}_h^2 - \left(\sum_{h=1}^{L} W_h \cdot \bar{x}_h \right)^2 \right] \tag{7}$$

EXEMPLO APLICATIVO:
GANHOS DE PRECISÃO EM UM CASO DE ESTRATIFICAÇÃO DE UMA FLORESTA NATURAL DE ARAUCÁRIA NO SUL DO BRASIL

Em 1965 realizou-se no Estado do Paraná o Inventário das Florestas Naturais de *Araucaria angustifolia*. O trabalho foi publicado por DILLEWIJN (1966).

A estratificação da floresta foi efetuada pela separação de três sub-populações, a saber:

- Tipo I: Floresta Natural Intocada de *A. angustifolia*;
- Tipo II: Floresta Natural Explorada de *A. angustifolia*;
- Tipo III: Floresta Natural com baixa ocorrência de *A. angustifolia*.

A área total do inventário compreendeu 2.918.726,62 ha, distribuídos como segue:

- Tipo I: 120.776,56;
- Tipo II: 444.642,20 ha;
- Tipo III: 2.353.307,86 ha.

A intensidade total de amostragem definida para o inventário foi de 321 unidades, que por motivos estratégicos, foram distribuídas com maior intensidade no Tipo I, por ser este o mais importante reduto da Floresta de Araucária e que, segundo Dillewijn, deveria ser o mais intensivamente amostrado. É claro que a alocação do total das unidades, nessas circunstâncias, foi subjetivamente efetuada, não tendo portanto sido utilizada nem a alocação proporcional nem a ótima.

PÉLLICO NETTO (1979) propôs se avaliar, neste caso, qual foi a perda de precisão em relação às duas formas clássicas de repartição da intensidade amostral estratificada. É claro que tais avaliações somente podem ser efetuadas após o inventário estar concluído.

Os resultados do inventário acima mencionado estão apresentados no Quadro 1.

Quadro 1 - Resultados da amostragem do Inventário do Pinheiro do Paraná - 1966
Table 1 - Sampling results of the Araucaria forest inventory in Paraná - 1966

Estrato	A_h (ha)	W_h	n_h	$\bar{X}_h$ (m^3)	s^2_h	s_h
Tipo I	120.776,56	0,04	138	492,68	61.862,10	248,72
Tipo II	444.642,20	0,15	80	304,97	8.315,06	91,19
Tipo III	2.353.307,86	0,81	103	134,56	3.754,92	61,28
Σ	2.918.726,62	1,00	321	-	-	-

Para a obtenção dos estimadores de precisão, os cálculos foram sumarizados no Quadro 2.

Quadro 2 - Síntese dos resultados da amostragem com estratificação do Inventário do Pinheiro do Paraná - 1966
Table 2 - Results of the stratified sampling of the Araucaria forest inventory in Paraná - 1966

Estrato	$W_h \bar{x}_h$	$W_h \bar{x}_h^2$	$W_h s^2_h$	$W_h s^2_h / n_h$	$W_h^2 s^2_h / n_h$
Tipo I	19,71	9.709,34	2.474,48	17,93	0,72
Tipo II	45,75	13.951,01	1.247,26	15,59	2,34
Tipo III	108,99	14.666,18	3.041,49	29,53	23,92
Σ	174,45	38.326,53	6.763,23	63,05	26,98

Considerando-se a conversão da área total e o número potencial de unidades amostrais de ¼ de ha, tem-se que N = 11.674 e os resultados finalmente foram obtidos.

$$s^2_{\bar{x}e} = 26,98 - \frac{6.763,23}{11.674.904} = 26,98$$

e para uma amostragem inteiramente aleatória

$$s^2_{\bar{x}aleat} = \frac{11.674.904 - 321}{321(11.674.904 - 1)} \left[6.763,23 - 63,05 + 26,98 + 38.326,53 - 30.432,80\right]$$

$$s^2_{\bar{x}aleat} = 0,0031[14.620,89] = 45,32$$

A partir destes resultados, pode-se verificar que a estratificação determinou substancial redução da variância da média e, conseqüentemente, o ganho em precisão pode ser avaliado em:

$$Ps^2_{\bar{x}e}\ \% = \frac{s^2_{\bar{x}\ aleat} - s^2_{\bar{x}e}}{s^2_{\bar{x}e}} = \frac{45,32 - 26,98}{26,98} 100 = 67,98\%$$

Se a redução for calculada em função dos erros padrões tem-se

$$Ps^2_{\bar{x}e}\ \% = \frac{6,73 - 6,19}{5,19} 100 = 29,67\%$$

Esses resultados foram tão expressivos, dado a confirmação de significância tanto entre as médias, como entre as variâncias dos estratos, condições necessárias para ocorrência da redução na variância da média estratificada.

Através de uma análise de variância pode-se confirmar a significância entre as médias, conforme está apresentado no Quadro 3. A significância a nível de 99% entre todas as médias também pode ser visualizada através da realização do teste de Tukey, conforme está apresentado no Quadro 4.

Quadro 3 - Análise de variância para os estratos volumétricos das florestas de *Araucaria* no Estado do Paraná

Table 3 - Analysis of variance of the volumetric stratification of the Araucaria forests in Paraná State

Fonte de Variação	Graus de liberdade	Soma de Quadrados	Média Quadrática	F
Entre os estratos	2	7.636.216,40	3.818.108,20	127,60**
Dentro dos estratos	318	9.514.999,20	29.921,38	-
Total	320	17.151.215,60	-	-

Quadro 4 - Resultados do teste de Tukey para as médias dos estratos das florestas de *Araucaria* no Estado do Paraná

Table 4 - Tukey's test of the strata means of the Araucaria forests in Paraná State

Tipo Florestal / Médias	III 134,56	II 304,97	I 492,68
III 134,56	-	**	**
II 304,97	-	-	**
I 492,68	-	-	-

Adicionalmente, pode-se ainda comprovar a significância entre as variâncias dos estratos, através da aplicação do teste de Bartlett, conforme apresentado por STEEL & TORRIE (1960) e sumarizado no Quadro 5.

Quadro 5 - Teste de homogeneidade de variância para os estratos volumétricos da amostragem do Inventário do Pinheiro do Paraná - 1966

Table 5 - Test of homogeneity of variance of volumetric strata of the Araucaria forest inventory in Paraná - 1966

Estrato	Graus de liberdade	Soma de Quadrados	s^2_h	$log\ s^2_h$	$(n_h-1)\ log\ (s^2_h)$	$1/(n_h-1)$
Tipo I	137	8.475.107,70	61.862,10	4,79142	656,4245	0,00723
Tipo II	79	656.889,74	8.315,06	3,91986	309,6689	0,01266
Tipo III	102	383.001,84	3.754,92	3,57460	364,6092	0,00980
Total	318	9.514.999,20			1.330,7026	0,02969
Ponderado			29.921,38	4,47598	1.423,3616	

$$c^{2} = 2{,}3026\left\{\left[\sum_{h=1}^{L}(n_h - 1)\right] log\ s_h^2 - \sum_{h=1}^{L}(n_h - 1)\ log\ s_h^2\right\}$$

c^2 = 2,3026 (1.423,3616-1.330,7026) = 213,3566

com 2 graus de liberdade.

Usando-se o fator de correção tem-se:

$$FC = 1 + \frac{1}{3(k-1)}\left[\sum_{h=1}^{L}\frac{1}{n_h - 1} - \frac{1}{\sum_{h=1}^{L}(n_h - 1)}\right]$$

e

$$c^{2} corrigido = \frac{213{,}3566}{1{,}004424} = 212{,}4168**$$

A significância entre as variâncias se deu ao nível de 99% de probabilidade.

Finalmente, pela aplicação da formulação apresentada na equação (6), pode-se obter os adicionais ganhos em precisão, relativos às alocações proporcional e ótima aplicadas para a estratificação da população. Esta comparação está também feita entre a alocação proporcional e ótima, para se configurar o ganho de precisão, quando as variâncias são heterogêneas, conforme comprovado pelo teste de Bartlett. Os resultados para a condução desta demonstração estão sumarizados no Quadro 6.

Quadro 6 - Resultados para o cálculo do aumento proporcional da variância da média em estratificação volumétrica na amostragem do Inventário do Pinheiro do Paraná – 1966

Table 6 - Results for the calculation of the proportional gain in precision on the stratified volumetric means of the Araucaria forest inventory in Paraná – 1966

Estrato	Intensidade amostral efetuada $\hat{n}_h$	Intensidade amostral proporcional n_h	Intensidade amostral ótima n'_h	$W_h s_h$	$n(W_h s_h)$	$\frac{(\hat{n}_h - n_h)^2}{n_h}$	$\frac{(\hat{n}_h - n'_h)^2}{n_h}$	$\frac{(n_h - n'_h)^2}{n_h}$
Tipo I	138	13	44	9,95	3.193,95	113,22	64,03	73,92
Tipo II	80	48	60	13,68	4.391,28	12,80	5,00	3,00
Tipo III	103	260	217	49,64	15.934,44	239,31	126,17	7,11
Σ	321	321	321	73,27	-	365,33	195,20	84,03

Os ganhos de precisão foram obtidos como segue:

- Entre a alocação subjetiva e a alocação proporcional

$$R^2 \bar{x}e = \frac{365,33}{321} 100 = 113\%$$

- Entre a alocação subjetiva e a alocação ótima

$$R^2 \bar{x}e = \frac{195,20}{321} 100 = 60,81\%$$

- Entre a alocação proporcional e a alocação ótima

$$R^2 \bar{x}e = \frac{84,03}{321} 100 = 26,18\%$$

Estes resultados mostram a grande vantagem, que o aprofundamento da estratificação com alocação ótima pode resultar.

CONCLUSÕES

A estratificação volumétrica nas florestas de *A. angustifolia* no Estado do Paraná resultou em uma eficiente técnica, como processo de amostragem aplicado em inventário de nível regional;

A população apresentou diferenças expressivas entre médias e heterogeneidade de variâncias em todos os estratos.

Os ganhos de precisão foram expressivas em estratificar versus realizar uma amostragem aleatória, da ordem de 30%.

Entre fazer alocação subjetiva e proporcional à área dos estratos, o ganho é da ordem de 113% em termos de variância da média;

Entre fazer alocação ótima e proporcional à área dos estratos, poderá resultar em um ganho da ordem de 26,18% em termos de variância da média, em condições similares às da população amostrada.

A diferença de precisão entre a alocação subjetiva e alocação ótima foi menor em termos de variância da média, devido esta ter se aproximado mais da alocação ótima que a proporcional e foi da ordem de 60,81%.

Os resultados anteriormente apresentados mostram que o conhecimento prévio de estimadores de estratos, tanto de médias, de variâncias como de suas áreas, permitem reduzir as intensidades amostrais nos estratos, para uma precisão especificada. Associado a isto obter-se-á uma redução nos custos de amostragem, dado o que, teoricamente, a alocação proporciona aos estimadores do inventário.

BIBLIOGRAFIA CITADA

COCHRAN, W.G. 1977. **Sampling Techniques**. 3rd. ed. John Wiley & Sons. Inc. New York and London, 413 p.

DILLEWIJN, F.J. van. 1966. **Inventário do Pinheiro do Paraná**. Cerena. CODEPAR. Curitiba. 104 p.

NEYMAN, J. 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. **Journal of the Royal Statistical Society**., 97: 558-606.

PÉLLICO NETTO, S.P. 1979. Die Forstinventuren in Brasilien. Neue Entwicklungen und ihr Beitrag für eine geregelte Forstwirtschaft. Tese de Doutorado. Mitteilungen aus dem Arbeitskreis für Forstliche Biometrie. Freiburg i Br. 232 p.

STEEL, R.G.D. & TORRIE, J.H. 1960. **Principles and procedures of statistics**. Mc Graw-Hill Book Company Inc., New York. 481 p.

TSCHUPROW, A.A. 1923. On the mathematical expectation of the moments of frequency distributions in the case of correlated observations. **Metron**, 2:461-493, 646-683.