用点抽样和乘积估计值法测定林分蓄积量

宋新民

关键词 点抽样、乘积估计值、次级样本

用角规点抽样，林分每公顷蓄积量为：$M=FZ\overline{R}$（式中，F 为角规断面积因子，Z 为样点计数木株数，$\overline{R}$ 为林木平均形高）。为了获得 $\overline{R}$ 值，必须采用角规控制检尺和一元立木材积表。目前用点抽样估计 $\overline{R}$ 的方法有：(1)对每个样点上计数木全部检尺；(2)从全部样点中随机抽取一部分样点，对该次级样点上计数木全部控制检尺；(3)从全部样点的计数木中随机（系统）地抽取一部分样木，组成次级样本，用来估计均值 $\overline{R}$。

根据抽样理论，方法(1)用简单随机抽样方法可估计其抽样误差，方法(2)与(3)一般采用双重比估计抽样方法估计其方差。上述估计误差的方法，不仅计算麻烦，有时并不能完全有效地估计抽样误差，所以本文提出了一种新的点抽样误差估计方法——乘积估计值法。其理论依据就是在 $M=G\overline{R}$[1]式中，把 G 与 $\overline{R}$ 视为两个相互独立变量，M 的方差用两个独立变量之积的误差传播公式近似估计。实验结果表明，用同样的样本资料，乘积估计法不仅计算简便而且比双重比估计法的效果更好。

1 乘积估计值法的原理

林分每公顷的实际蓄积量可以用下式表示：

$$M=\frac{M}{G}\cdot G=\frac{\sum_{j}^{N}V_j}{\sum_{j}^{N}g_j}\cdot G,\quad 令\quad R=\sum_{j}^{N}V_j/\sum_{j}^{N}g_j \tag{1}$$

则

$$M=G\cdot R \tag{2}$$

∴

$$R=M/G \tag{3}$$

式中，V_j 为第 j 株木材积；g_j 为第 j 株木断面积；N 为每公顷林木总株数。

当在林分中随机地抽取 n 个角规点，平均每公顷蓄积估计值($\overline{m}$)为

$$\overline{m}=\overline{G}\,\overline{R} \tag{4}$$

当测定样点上全部计数木形高(R_j)时，第 i 个样点林分蓄积量估计值为

$$m_i=F\sum_{j}^{Z}R_j\text{[1]} \tag{5}$$

$$E(m_i)=FE(\sum_{j}^{Z}R_j)=F\sum_{j}^{N}E(R_jZ_j)$$

根据 $g_j=FZ_j$， $g_j/F=Z_j$

∴

$$E(m_i)=F\sum_{j}^{N}(\frac{V_j}{g_j})(\frac{g_j}{F})=\sum_{j}^{N}V_j=M$$

1993—11—23收稿。
宋新民副教授（北京林业大学森林资源与环境学院 北京 100083）。

当第 j 株树在第 i 个样点被抽中时 $Z_j=1$,未抽中时 $Z_j=0$。

$$\overline{m} = \frac{1}{n}\sum_i^n m_i \tag{6}$$

$$E(\overline{m}) = \frac{1}{n}\sum_i^n E(m_i) = M$$

同理,第 i 个样点每公顷断面积估计值为

$$G_i = FZ_i = F\sum_j^Z Z_j$$

$$E(\overline{G}) = E(\frac{1}{n}\sum_i^n G_i) = F\,\frac{1}{n}\sum_i^n[\sum_j^N E(Z_j)]$$

$$= F\sum_j^N(\frac{g_j}{F}) = \sum_j^N g_j = G \tag{7}$$

在(4)中 $\overline{R}$ 是 R 的无偏估计值,即

$$E(\overline{R}) = R \tag{8}$$

这个结论 Lahiri[2]在论述按单元值大小成比例的不等概抽样方法中已有证明。角规点抽样正是一种典型的不等概(PPS)抽样,$\overline{R}$ 是无偏的。只要在各样点采用重复抽样(同一株树在几个样点重复计数)则 $\overline{R}$ 值与样点本身无关。

如果从初级样本($Z=\sum_i^n Z_i$)的样木中,随机(系统)地抽取一个次级样本,设次级样本(木)单元数为 L,第 K 株样木的形高为 $R_K(K=1,2,\cdots,L)$,可以证明,次级样木平均形高($\overline{R}_2$)的条件数学期望 $E_2(\overline{R}_2)$是全部被抽中(Z)株林木的平均值 $\overline{R}$,也等于总体 R。

因为 $\overline{R}_2=\frac{1}{L}\sum_K^L R_K$

$$E(\overline{R}_2) = E_1[E_2(\overline{R}_2)] = E_1[\frac{1}{Z}\sum_j^Z E_2(\overline{R}_2)]$$

$$= E_1[\frac{1}{Z}\sum_j^z \overline{R}_j] = R$$

由上可以得出:点抽样林分形高估计值,不论是测定全部计数木或抽取一部分次级样点测定计数木形高,还是从全部计数木(Z)中随机抽取一部分次级样木测其形高,都可以得到林分(总体)形高 R 的无偏估计值。

当用无偏估计值 $\overline{G}$ 乘上 R 的有条件无偏估计值 $\overline{R}_2$便可得到蓄积 M 的无偏估计值。

$$\overline{m}_2 = \overline{G}\cdot\overline{R}_2 \tag{9}$$

$$E(\overline{m}_2) = E_1E_2(\overline{m}_2) = E_1(\overline{m}) = M \tag{10}$$

根据两个变量乘积的方差公式,则(4)或(9)式的方差为

$$S_{\overline{m}}^2 = \overline{G}^2S_{\overline{R}}^2 + \overline{R}^2S_{\overline{G}}^2 + 2\rho S_G S_R{}^{[3]} \tag{11}$$

上式中,如果能证明最后一项非常小可略而不计或样点计数株数(或 G_i)与样点平均形高($\overline{R}_i$)之间不存在相关关系,则(11)式可写为

$$S_{\overline{m}}^2 = \overline{G}^2S_{\overline{R}}^2 + \overline{R}^2S_{\overline{G}}^2 \tag{12}$$

式中, $S_{\overline{G}}^2=F^2\frac{1}{n(n-1)}\sum_i^n(Z_i-\overline{Z})^2$

$S_{\overline{R}}^2=\frac{1}{L(L-1)}\sum_K^L(R_K-\overline{R}_2)^2$

若用变动系数$C_{\bar{m}}$来表示，(12)式化为

$$C_{\bar{m}}^{2} = C_{\bar{G}}^{2} + C_{\bar{R}}^{2} \tag{13}$$

则相对误差(抽样误差)$E\%$为

$$E_{\bar{m}} = \sqrt{\frac{C_{\bar{G}}^{2}}{n_1} + \frac{C_{\bar{R}}^{2}}{n_2}} \tag{14}$$

这里n_1与n_2分别表示估计断面积的样点个数和估计形高$\bar{R}$的计数木株数。

现在来讨论$\bar{G}$与$\bar{R}$统计的独立性问题，从理论上两个变量是相互不独立的。因为次级样点(木)是由初级样点(木)中抽取的。两者是否可按独立变量对待，只能通过实验，检验样点计数木株数随样点$\bar{R}_i$的分布。根据Palley和Horwitz[4]的研究结果，证明样点计数木株数(或G)与样点$\bar{R}_i$之间不存在显著相关关系。这是因为PPS抽样中，即不论形高(R_j)值大小如何，其抽中计数值均为1。另一方面，(11)式最后误差项一般都很小，尤其是同一林分相同树种的形高变动系数多在10%～20%[5]。基于上述原因，乘积估计值的方差可用(12)式作近似估计。

2　实验及结果

2.1　实验方法

1987年在华盛顿大学林学院实验林场，在两林分中曾分别设置120和137个角规点，树种是花旗松(*Pseudotsuga menziesii* (Mirb.) Franco)和异叶铁杉(*Tsuga heterophylla* (Rafi.) Sarg.)，$F=2.3\ m^2/hm^2$。用立木材积方程获得样木形高。全部样点计数木都控制检尺。内业分析中为比较乘积估计值法和双重比估计法的效率，分别从样点中各抽取1/10样点并采用这些样点上全部样木形高，进行双重比估计。另外，从120和137个样点的计数木中系统地抽取1/20样木，用以进行乘积估计分析。

2.2　实验结果(见表1，2)

表1　两个林分点抽样估计结果

项　目	样点(样木)	平均值	标准差	标准误	E%
		A　林　分			
断面积(m^2)	120个点	36.16	11.94	1.09	3.01
形　高(m^3/m^2)	945株	12.05	2.23	0.073	0.61
点估计(m^3)	120个点	435.72	159.11	14.53	3.33
乘积估计(m^3)	120(945)	435.72	—	13.40	3.07
		B　林　分			
断面积(m^2)	137个点	29.56	14.29	1.22	4.13
形　高(m^3/m^2)	882株	11.10	1.92	0.065	0.58
点估计(m^3)	137个点	328.12	175.31	14.98	4.56
乘积估计(m^3)	137(882)	328.12	—	13.70	4.17

表2　双重比估计抽样和乘积估计法估计结果

林　分	1/10样点双重比估计法			1/20样木乘积估计值法		
	蓄积量	标准误	E%	蓄积量	标准误	E%
A	435.88	22.03	5.05	435.84	17.65	4.05
B	351.31	18.90	5.77	328.21	16.05	4.89

3 结论与讨论

(1)用点抽样估计林分蓄积量,用测全部计数木形高或测次级样点计数木或抽取系统样本(次级样木)测其计数木形高,都可以得到形高的无偏估计值。

(2)用乘积估计值法(4)式,林分平均蓄积是无偏的。这种方法并不要求 $\bar{G}$ 与 $\bar{R}$ 是独立的,但对于提出一个有效方差计算方法,这种独立性是必要的。研究表明,样点计数木株数与 $\bar{R}_i$ 之间存在着非常弱相关,且形高变动系数很小,在实际工作中可以用两独立变量乘积误差(14)式近似计算抽样误差。

(3)双重(点)比估计法,均值,方差也无偏,且不受相关性影响[6]。表2表明两种估计方法结果近似。但当样点内形高变动小于样点间时,双重比估计结果可能劣于乘积估计值法。

(4)系统抽取1/20样木组成的次级样本(木)用乘积估计值法,其结果精度高于双重比估计(见表2),其原因可能是该实验林分,样点间的形高变动大于样点内变动所致。

(5)点抽样误差主要来自样点计数木株数的变动,可以通过增加样点数(费用较低)减少抽样误差,相反,形高变动小(10%~20%),可以用较少样木获得。由表2两个对比可以看,在同样精度下,用乘积估计值法所需外业调查费用远比用双重点比估计抽样要少,此例中1/10样点检尺株数比乘积估计值法约多2倍。

参 考 文 献

1 北京林业大学主编. 测树学. 北京:中国林业出版社,1987.

2 科克伦 W G(张尧庭,吴辉译). 抽样技术. 北京:中国统计出版社,1984.

3 斯图尔特 L 迈耶(周鎏元,顾兆良译). 科技工作中的数据分析. 北京:原子能出版社,1983.

4 Palley M N, Horwitz L G. Properties of some random and systematic point sampling estimators. Forest Science, 1961, 16(3):240~246.

5 宋新民. 点抽样误差来源的研究. 北京林业大学学报,1990,12(2):21~29.

6 洛茨 F,哈勒 K E,佐勒(林昌庚,沙琢等译校). 森林资源清查. 北京:中国林业出版社,1985.

Timber Cruising with Point Sampling and Product Estimator

Song Xinmin

Abstract A sampling procedure for volume estimation is presented, in which volume is computed as the product of average basal area per hectare and average volume-basal-area-ratio(R). The basal area estimate is obtained by point sampling, and R is estimated by subsampling point/sampled trees. Expression for the variance of the product estimator is also given. Simulated tests show that for the same degree of precision, using this sampling method is over that of the double sampling for ratio estimator.

Key words point sampling, product estimator, subsampling

Song Xinmin, Associate Professor (Forest Resources College, Beijing Forestry University Beijing 100083).