

Randomized branch sampling

Randomized branch sampling (RBS) is a special application of multistage probability sampling (*see Sampling, environmental*), which was developed originally by Jessen [3] to estimate fruit counts on individual orchard trees. In general, the method can be used to obtain estimates of many different attributes of trees or other branched plants. The usual objective of RBS is the estimation of a total quantity for a tree or a branch, for example, woody dry matter or volume [5], leaf dry matter [2], or counts of tree-inhabiting insects [1]. RBS also has potential application in the estimation of attributes of river networks (*see Rivers, canals and estuaries*) and other branched structures. Here, our interest is restricted to the above-ground portions of trees.

Terminology

A 'branch' is defined as the entire stem system that develops from a single bud (lateral or terminal) and a 'branch segment' or simply 'segment' is a part of a branch between two consecutive nodes. No distinction is made between the segments of the main stem and side branches. The butt of the main stem of a tree is considered a node and the tree is considered a branch for the purpose of RBS. Terminal shoots are considered to be both branches and branch segments. Thus, any tree or branch can be defined as a population of branch segments.

An 'object branch' is the object of the sampling, i.e. a branch (or tree) for which we need an estimate of the quantity of some attribute or characteristic. The segments that comprise the object branch constitute the population of elements from which we select a sample. The sample that results from RBS has been called a 'path'; it consists of a sequence of connected segments and, possibly, one subterminal branch. A path extends acropetally from the basal segment of the object branch to either a terminal shoot or a subterminal branch. The subterminal branch may contain several segments, including multiple terminal shoots (*see Figure 1*).

Randomization test *see* Computer-intensive methods; Exact randomization technique

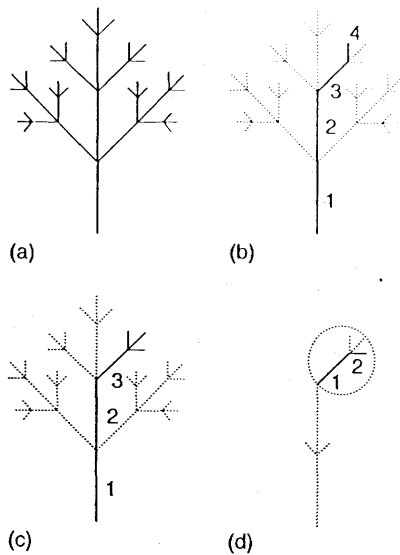


Figure 1 The tree diagrammed in (a) contains 27 possible paths from the butt of the main stem to a terminal shoot (adapted from Figure 1 of Gregoire et al. [2] with permission from *Ecology*). The four branch segments of one possible path are shown in (b). RBS can be stopped at any node, in which case the entire branch [3 in example (c)] is treated as the terminal segment of the path. The sampling can be started from the butt of any branch on the tree [e.g. example (d)], in which case the resultant estimates pertain only to the entire starting branch (encircled), not the entire tree

Path Selection

The butt of the basal segment of the object branch is the starting point or first node of the path and the basal segment is the first segment of the path. If the object branch is purposely selected at the outset, then the probability of selection of the first segment of the path (denoted by q_1) equals 1. The first segment is followed acropetally to the second node where each branch emanating from this node is assigned a probability of selection, and one branch is selected at random by conventional probability sampling procedures. This selection fixes the second segment of the path. The probability of selection of the selected branch is denoted by q_2 . The procedure is repeated until either a small subterminal branch or a terminal shoot is selected at the final (R th) node with probability q_R . If the sampling is terminated upon selection of a subterminal branch, then the entire subterminal branch is

treated as the terminal segment of the path (see Figure 1c).

A selected path is a probability sample of the segments comprising the object branch. The probabilities of selection assigned to branches at nodes are, technically, conditional probabilities because the selection of a branch at a node is conditional upon the path reaching the node from which the branch emanates. The unconditional probability of selection of the r th branch segment of the path (Q_r) is

$$Q_r = q_1 \times q_2 \times \cdots \times q_r = \prod_{k=1}^r q_k, \quad r = 1, 2, \dots, R \quad (1)$$

More than one path may be selected, in which case the unconditional probability of selection of the r th branch segment of the i th of $m > 1$ paths is denoted by $Q_{r(i)}$.

The probabilities of selection assigned to the branches emanating from each node along the path must sum to one. The probability of selection assigned to the b th of B branches at a node ideally should equal $X_b/(X_1 + \cdots + X_B)$, where X_b ($b = 1, \dots, B$) is the quantity of interest borne by the branch. Ordinarily, however, we cannot discern the quantity of interest; otherwise, we would not be sampling. Commonly, the quantity of interest correlates with branch diameter (d), branch length (l), powers of these quantities (d^u or l^v), or a product ($d^u \times l^v$). For example, if the total dry matter, surface area, or count of leaves is of interest, then we may choose to calculate the probabilities from $X = d^2$ because quantities of leaves borne by branches are known to be highly correlated with the diameter squares of the respective branches. Valentine and Hilton [4] used ocular estimates of the foliage borne by the respective branches to calculate probabilities of selection for the objective of estimating leaf count. Ocular estimation is, of course, an option for any RBS objective.

Estimators

Within the population of branch segments that comprise the object branch are subpopulations of branch segments of order r , $r = 1, 2, \dots, R'$, where R' is the longest possible path in terms of number of segments. The r th subpopulation contains every branch segment that could be the r th segment of any possible

path from a given starting node. The reciprocal of the unconditional probability of selection of the r th segment of the i th path, $Q_{r(i)}^{-1}$, is an unbiased estimator of the number of segments in the r th subpopulation. Let $y_{r(i)}$ denote the quantity of interest measured on the r th segment of the i th path, and let Y_r denote the total quantity of interest for all the segments in the r th subpopulation. An unbiased estimator of Y_r is

$$\hat{Y}_{r(i)} = \frac{y_{r(i)}}{Q_{r(i)}} \quad (2)$$

A combined estimator averages across $m > 1$ paths via

$$\hat{Y}_r = \frac{1}{m} \sum_{i=1}^m \hat{Y}_{r(i)} \quad (3)$$

In order to estimate precision it is necessary to select $m > 1$ paths. A sample-based estimator of the variance of $\hat{Y}_r$ is

$$\text{var}(\hat{Y}_r) = \frac{1}{m(m-1)} \sum_{i=1}^m (\hat{Y}_{r(i)} - \hat{Y}_r)^2, \quad m > 1 \quad (4)$$

Subpopulations may correspond to natural populations of different ages. For example, many species of trees have more than one annual cohort of foliage.

Let Y denote the total quantity of interest for all the segments that comprise the object branch. An unbiased estimator of Y based on the i th path is

$$\hat{Y}_i = \sum_{r=1}^R \hat{Y}_{r(i)} = \sum_{r=1}^R \frac{y_{r(i)}}{Q_{r(i)}} \quad (5)$$

and the combined estimator of Y is

$$\hat{Y} = \frac{1}{m} \sum_{i=1}^m \hat{Y}_i \quad (6)$$

If an object branch contains N branch segments, the variance of $\hat{Y}$ is

$$\sigma^2(\hat{Y}) = \frac{1}{m} \left(\sum_{j=1}^N \frac{y_j^2}{Q_j} - Y^2 \right) \quad (7)$$

The sample-based estimator of the variance of $\hat{Y}$ is

$$\text{var}(\hat{Y}) = \frac{1}{m(m-1)} \sum_{i=1}^m (\hat{Y}_i - \hat{Y})^2, \quad m > 1 \quad (8)$$

These various quantities may be used for point or interval estimation of Y_r and/or Y , as appropriate.

References

- [1] Furness, G.O. (1976). The dispersal, age-structure and natural enemies of the long-tailed mealy bug, *Pseudococcus longispinus* (Targoni-Tozzetti) in relation to sampling and control, *Australian Journal of Zoology* 24, 237-247.
- [2] Gregoire, T.G., Valentine, H.T. & Furnival, G.M. (1995). Sampling methods to estimate foliage and other characteristics of individual trees, *Ecology* 56, 1181-1194.
- [3] Jessen, R.J. (1955). Determining the fruit count on a tree by randomized branch sampling, *Biometrics* 11, 99-109.
- [4] Valentine, H.T. & Hilton, S.J. (1977). Sampling oak foliage by the randomized branch method, *Canadian Journal of Forest Research* 7, 295-298.
- [5] Valentine, H.T., Tritton, L.M. & Furnival, G.M. (1984). Subsampling trees for biomass, volume, or mineral content, *Forest Science* 30, 673-681.

(See also Forestry; Multistage design)

HARRY T. VALENTINE

ENCYCLOPEDIA OF ENVIRONMETRICS

Volume 3 M-R

Editors-in-Chief

ABDEL H. EL-SHAARAWI

National Water Research Institute, Canada

WALTER W. PIEGORSCH

University of South Carolina, USA



John Wiley & Sons, Ltd

2002