A Technical Note on PPS Sampling with an Application to Fruit and Nuts

by

Frederic A. Vogel

Research and Development Branch
Standards and Research Division
Statistical Reporting Service

March 1970

1. Basic Theory

One of the first problems encountered when designing a sample is to determine what criterion to use as a basis for selecting the sample units. For example, should the sample units be selected with equal probabilities, or with unequal probabilities based on some measure of size? A typical problem is encountered when deciding how to obtain a sample of limbs in a fruit or nut tree for the purpose of estimating the total number of fruit or nuts in the tree. The production per tree is one of the most important components in a forecasting model designed to estimate total production.

A number of research studies have been conducted to determine the most efficient criteria on which to base the sample selection. The limb cross-sectional area (CSA) has been found to show a good correlation with the number of fruit on the limb. The CSA is currently used in several operational surveys as a measure of size for limb selection.

The "random path" PPS method of limb selection as described by Jessen (1955) is commonly used. This method is primarily used when obtaining a sample of only one limb per tree. However, when the sample design makes it necessary to obtain more than one sampling unit per tree, it is not clear how the sample should be drawn, and what form the estimator should take.

In recent years, a large number of selection procedures and estimators have been proposed for sampling with unequal probabilities without replacement when the sample size is greater than one. These were reviewed with the intention of finding one that was both efficient and practical for estimating the number of fruit or nuts in a tree. The following paragraphs outline the basic theory - then an estimator is suggested.

It should be stressed at this time that the success of PPS sampling depends heavily on the reliability of the measures of size. If these are poor, i.e., the limb CSA is not correlated with the number of fruit, it may be no better than sampling with equal probabilities. In fact, the magnitude of the correlation coefficient between y (number of fruit) and x (measure of size) may be unity, and yet PPS sampling may be worse than sampling with equal probabilities. This can occur when there is a negative correlation between y and x. Des Raj (1968) discusses these problems and suggests some methods for determining whether PPS sampling would be appropriate.

Several terms will be repeated so their definitions follow.

1. π_i = The probability that the i^{th} unit is included in a sample selected without replacement (wtr).
2. $\pi_{i'}$ = The probability that unit (i') is included in the sample (wtr).
3. $\pi_{ii'}$ = The probability that both units (i) and (i') are included in the sample (wtr).
4. P_{ij} = The probability that the j^{th} terminal unit from the i^{th} primary unit was drawn into the sample.
5. $P_i(jj')$ = The probability that both terminals j and j' were selected in the i^{th} primary unit.

Primes attached to the subscripts will be used to identify the different sample units within a selection stage.

The general theory was first presented by Horvitz and Thompson (1952). Yates and Grundy (1953) added to this theory by suggesting a variance estimator. The theory applies for any sample size, but only the special case where n = 2 will be discussed. The application of the principle to a single stage case will be discussed first and will be extended to a two-stage sampling scheme.

Suppose that two units are chosen with probabilities of inclusion π_i and $\pi_{i'}$ respectively, and with a joint probability of inclusion $\pi_{ii'}$. Then an unbiased estimator of the population total is $\hat{Y} = \frac{y_i}{\pi_i} + \frac{y_{i'}}{\pi_{i'}}$ and

$$\text{EST. Var}(\hat{Y}) = \frac{\pi_i \pi_{i'} - \pi_{ii'}}{\pi_{ii'}} \left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}} \right)^2$$

This expression for the estimated variance may be negative unless the sampling scheme requires that $\pi_i \pi_{i'} > \pi_{ii'}$. An additional problem is the requirement of the knowledge of π_i, $\pi_{i'}$, and $\pi_{ii'}$. The calculations for obtaining these can become quite cumbersome, especially if the sample size is greater than two. Therefore, a procedure often used in practical situations is to sample in several stages with a sample of size two in each stage.

For example, suppose that two first stage units (primary limbs in the case of fruit or nut trees) are selected with probabilities of inclusion π_i, $\pi_{i'}$, and joint probability $\pi_{ii'}$. Then within each selected first stage unit, a sample of two second stage units (terminal limbs) are selected with probabilities P_{ij}, $P_{ij'}$ and $P_{i(jj')}$, respectively. An unbiased estimator of the population total is then

$$Y = \frac{1}{\pi_i}\left(\frac{y_{ij}}{P_{ij}} + \frac{y_{ij'}}{P_{ij'}} \right) + \frac{1}{\pi_{i'}} \left(\frac{y_{i'j}}{P_{i'j}} + \frac{y_{i'j'}}{P_{i'j'}} \right)$$

where y_{ij} is the value attached to the j^{th} unit in the i^{th} primary. To simplify the notation, let $\left(\frac{y_{ij}}{P_{ij}} + \frac{y_{ij'}}{P_{ij'}} \right) = \hat{Y}_i$ and $\left(\frac{y_{i'j}}{P_{i'j}} + \frac{y_{i'j'}}{P_{i'j'}} \right) = \hat{Y}_{i'}$.

Also consider the total number of primary units to be N, and the total number of secondary units within each primary to be M_i.

Then $\hat{Y} = \frac{\hat{Y}_i}{\pi_i} + \frac{\hat{Y}_{i'}}{\pi_{i'}}$ and

$$\text{Var}(\hat{Y}) = \text{Var}\left[\sum_i^{i'} \frac{1}{\pi_i} \; E(\hat{Y}_i \mid i)\right] + E\left[\text{Var} \sum_i^{i'} \left(\frac{Y_i}{\pi_i} \mid i\right)\right]$$

Since $\hat{Y}_i$, is an unbiased estimate of the primary total Y_i we get

$$\text{Var}(\hat{Y}) = \text{Var}\left[\sum_i^{i'} \frac{Y_i}{\pi_i}\right] + E\left[\sum_i^{i'} \frac{1}{\pi_i}^2 (P_{ij} \; P_{ij'} - P_{ijj'})\left(\frac{y_{ij}}{P_{ij}} - \frac{y_{ij'}}{P_{ij'}}\right)^2\right]$$

$$= \sum_{i<i'}^{N} (\pi_i \pi_{i'} - \pi_{ii'}) \left(\frac{Y_i}{\pi_i} - \frac{Y_{i'}}{\pi_{i'}}\right)^2 \quad + \qquad (1)$$

$$\sum_i^N \frac{1}{\pi_i}\left(\sum_{j<j'}^{M_i} (P_{ij} \; P_{ij'} - P_{ijj'}) \left(\frac{y_{ij}}{P_{ij}} - \frac{y_{ij'}}{P_{ij'}}\right)^2\right) \qquad (2)$$

Notice that Var (Y) is two components:

(1) represents the sampling variance due to selecting primaries.

(2) represents the sampling variance due to subsampling within each primary.

Durbin (1953) suggested a rule for estimating the variance from a sample in the two stage case. His procedure is based on the general theorem on conditional variances which is as follows:

"The total variance in two stage sampling is the sum of two parts. The first part is equal to the estimate of variance calculated on the assumption that the first stage units have been measured without error. The second part is calculated as if the first stage units selected were fixed strata; the contribution to the variance from each first stage unit being multiplied by the probability of that unit's inclusion in the sample."

Using this rule we get

$$\text{EST. Var } (\hat{Y}) = \left(\frac{\pi_i \, \pi_{i'} - \pi_{ii'}}{ii'}\right)\left(\frac{\hat{Y}_i}{\pi_i} - \frac{\hat{Y}_{i'}}{\pi_{i'}}\right)^2 +$$

$$\sum_{i}^{i'} \frac{1}{\pi_i} \frac{(P_{ij} \, P_{ij'} - P_{i(jj')})}{P_{ijj'}} \left(\frac{y_{ij}}{P_{ij}} - \frac{y_{ij'}}{P_{ij'}}\right)^2$$

This method appears to be fairly straight forward. However, as mentioned earlier, the calculations of the π_{i}'s, $\pi_{ii'}$, P_{ij}'s and $P_{i(jj')}$ can become complex and lengthy.

Considerable research has been conducted in an attempt to find ways of calculating the necessary probabilities to use in the above estimators.

A number of different sampling schemes and estimators have been developed. An empirical study by Rao and Bayless (1969) compared several of the available procedures. Their criterion for including a sampling scheme in their study follows:

(1) The variance estimator should be less than the estimate provided by sampling with replacement.

(2) A non-negative, unbiased variance estimator should be available.

(3) Computations should not be complex and lengthy.

They concluded that an estimator developed by Murthy (1957) is preferable over the other methods when a consistent estimator as well as a stable variance estimator are required. The efficiency of his estimator compares favorably with other methods.

The sample selection procedures are fairly straight forward.

(1) Select a PPS sample of size unity from a random array of the population, and remove the selected unit from the population.

(2) Take a PPS sample of size unity from the units remaining in the population. Remove the selected element from the population.

(3) Continue this procedure until n selections are made. This will give a sample selected without replacement and with unequal probabilities. The i^{th} selection is made with probabilities proportionate to the size of the remaining elements.

For $n = 2$, an unbiased estimator of the population total based on the order the sample was drawn is $y_1 = \frac{y_2}{P_2}(1 - P_1)$ (Des Raj, 1968). P_1 and P_2 are the original probabilities of selection as defined below.

Murthy (1957) proves that corresponding to any biased or unbiased ordered estimator, there exists an unordered estimator that is more efficient than the former. First he considered the probabilities associated with the units drawn.

For $n = 2$

P_1 = Probability that unit i is drawn first is $\frac{x_i}{\sum x_i}$.

P_2 = Probability that unit i' is drawn first is $\frac{x_{i'}}{\sum x_i}$.

$P_{s1} = \frac{P_1 P_2}{(1 - P_1)}$ = Probability that unit i is drawn first and unit i' second.

$P_{s2} = \frac{P_2 P_1}{(1 - P_2)}$ = Probability that unit i' is drawn first and unit i second.

$P_s = \frac{P_1 P_2}{(1 - P_1)} + \frac{P_2 P_1}{(1 - P_2)}$ = The probability of selecting i and i' disregarding the order drawn.

$\hat{Y}_o$ = unbiased estimate based on order in which sample was drawn.

$\hat{Y}_u$ = unordered unbiased estimator

If unit i was drawn first and unit i' second $\hat{Y}_{oi} = y_1 + \frac{y_2}{P_2}(1 - P_1)$.

If unit i' is drawn first and i second, $\hat{Y}_{oi'} = y_2 + \frac{y_1}{P_1}(1 - P_2)$.

Then $\hat{Y}u = \left(y_1 + \frac{y_2(1 - P_1)}{P_2}\right)\frac{P_{s1}}{P_s} + \left(y_2 + \frac{y_1(1 - P_2)}{P_1}\frac{P_{s2}}{P_s}\right)$ which is the weighted average of the ordered pairs. Then $\hat{Y}_u$ reduces to:

$$\frac{1}{(2 - P_1 - P_2)}\left(\frac{y_1}{P_1}(1 - P_2) + \frac{y_2}{P_2}(1 - P_1)\right)$$

Notice that $P_1 = \frac{X_i}{\sum X_i}$ and $P_2 = \frac{X_{i'}}{\sum X_i}$ are the original probabilities of selection. He then goes on to show that

$$\text{Var}(\hat{Y}_u) = \sum_{i<i'}^{N} \frac{P_i P_{i'}(1 - P_i - P_{i'})}{(2 - P_i - P_{i'})}\left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2$$ and EST.

$$\text{Var}(\hat{Y}) = \frac{(1 - P_i)(1 - P_{i'})(1 - P_i - P_{i'})}{(2 - P_i - P_{i'})2}\left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2$$

This method has the desirable features that:

(1) P_i and $P_{i'}$ are the probabilities of selection for one unit.

(2) Calculations for computing the estimators are not very cumbersome.

(3) It is a workable procedure for practical use, i.e., sample selection methods are straight-forward.

The method will now be extended to the two stage case. To illustrate, suppose we wish to estimate the number of fruit on a tree by selecting two primaries, and within each selecting two terminals.

N = The number of primary scaffolds in the tree. i = 1... N

$Z_i = \frac{X_i}{\sum X_i}$ = the size of the i^{th} primary divided by the sum of the primary sizes.

M_i = The number of terminal sampling units in the i^{th} primary unit.

$T_{ij} = \frac{X_{ij}}{\sum_j^{M_i}}$ = the size of the j^{th} terminal divided by the sum of the terminal sizes in the i^{th} primary.

$$\hat{Y}_i = \frac{1}{(2 - T_{i1} - T_{i2})}\left(\frac{y_{i1}}{T_{i1}}(1 - T_{i2}) + \frac{y_{i2}}{T_{i2}}(1 - T_{i1})\right)$$ is the estimated total number of fruit in the i^{th} primary, and y_{ij} is the number of fruit in the j^{th} terminal of the i^{th} primary.

Then $$\hat{Y} = \frac{1}{(2 - Z_1 - Z_2)}\left(\frac{\hat{Y}_1}{Z_1}(1 - Z_2) + \frac{\hat{Y}_2}{Z_2}(1 - Z_1)\right)$$

$$\text{EST. Var } Y = \frac{(1 - Z_1)(1 - Z_2)(1 - Z_1 - Z_2)}{(2 - Z_1 - Z_2)^2}\left(\frac{\hat{Y}_1}{Z_1} - \frac{\hat{Y}_2}{Z_2}\right)^2 +$$

$$\frac{1}{(2 - Z_1 - Z_2)} \left(\frac{(1 - Z_2}{Z_1} \quad \frac{(1 - T_{11})(1 - T_{12})(1 - T_{11} - T_{12})}{(2 - T_{11} - T_{12})^2} \left(\frac{y_{11}}{T_{11}} - \frac{y_{12}}{T_{12}} \right)^2 \right) +$$

$$\frac{1}{(2 - Z_1 - Z_2)} \left(\frac{(1 - Z_1)}{Z_2} \quad (1 - \frac{T_{21})(1 - T_{22})(1 - T_{21} - T_{22})}{(2 - T_{21} - T_{22})^2} \left(\frac{y_{21}}{T_{21}} - \frac{y_{22}}{T_{22}} \right)^2 \right)$$

This procedure can be extended to include more than two units at each stage of sampling. To derive on estimator when more than two units are selected, list all possible ordered estimators and the probabilities of getting the respective ordered samples. Then form the weighted average of the ordered estimators.

The scheme can then be extended to include several stages of sampling. For example, a three stage sampling scheme could be trees within blocks, primaries within trees, and terminals within primaries. The first step would be to derive an estimator for each primary as illustrated in the two stage case. Then the tree estimator is based on those for primaries, again as illustrated in the two stage case. The final step is then to combine the tree estimators to obtain the block estimator. In general, one would start with the final stage of selection and work back to the first stage.

2. Simplication of Estimators for Equal Probability Sampling

An interesting comparison between this estimator and the usual estimator for equal probability sampling can be made. Using the single stage estimator for n = 2 to illustrate, we have

$$\hat{Y} = \frac{1}{(2 - P_1 - P_2)} \left(\frac{y_1}{P_1} (1 - P_2) + \frac{y_2}{P_2} (1 - P_1) \right).$$

If the sample was selected with equal probabilities, $P_1 = \frac{1}{N}$ and $P_2 = \frac{1}{N}$.

Then $$\hat{Y} = \frac{1}{(2 - \frac{1}{N} - \frac{1}{N})} \left(\frac{y_1}{\frac{1}{N}} (1 - \frac{1}{N}) + \frac{y_2}{\frac{1}{N}} (1 - \frac{1}{N}) \right)$$

$$= \frac{N}{2(N - 1)} \left((N - 1)(y_1 + y_2) \right) = \frac{N}{2} (y_1 + y_2)$$ which is the unbiased estimate of the population total based on equal probability sampling.

Similarly for $$\text{Var} (\hat{Y}) = \sum_{i\,j}^{\binom{N}{2}} \frac{P_i P_j (1 - P_i - P_j)}{(2 - P_i - P_j)} \left(\frac{y_i}{P_i} - \frac{y_j}{P_j} \right)^2$$

we can assume equal probability sampling and get $$\sum_{i\,j}^{\binom{N}{2}} \frac{\frac{1}{N}^2 \left(\frac{N - 2}{N} \right)}{\frac{(2N - 2)}{N}} (N y_i - N y_j)^2 =$$

$$\frac{N - 2}{2(N - 1)} \sum_{i \neq j}^{\binom{N}{2}} (y_i - y_j)^2 = \frac{N - 2}{2(N - 1)} \left((N - 1) \sum^{N} y_i^2 - \sum_{i \neq j}^{\binom{N}{2}} y_i y_j \right).$$

The identity $\sum_{i \neq j}^{\binom{N}{2}} y_i \, y_j = \sum^{N} y_i^{\,2} + (\sum^{N} y_i)^2$ reduces this to

$$\frac{N - 2}{2\,(N - 1)} \left(N \sum^{N} y_i^2 - (\sum y_i)^2 \right) = N^2 \; \frac{(N - 2)}{2N} \left(\frac{\sum^{N} y_i^2 - (\sum y_i)^2/N}{N - 1} \right)$$

which is the usual variance term for simple random sampling when n = 2.

3. A Numerical Example

An example has been worked out to illustrate how to compute the necessary probabilities and the variance estimates. The basic data was obtained for a cherry tree in Michigan. A total count by terminal limbs was made of all of the fruit in the tree along with CSA measurements for every sampling unit (terminal limb). The data and size measurements are presented in Table 1 which includes the different probabilities of selection.

Two sets of probabilities of limb selection were computed. Suppose that all of the terminal limbs in the tree were randomly arrayed disregarding which primary they belonged to. Then select a sample of two limbs. This represents a single stage sampling scheme. The probability that any terminal is selected on the first draw is the terminal CSA divided by the sum of all terminal CSA's. This is designated as z_{ij} in the Table. Note that

$$\sum_{i}^{N} \sum_{j}^{M_i} z_{ij} = 1.$$

In order to select a sample from such a single stage frame it is necessary to have a size measurement of every terminal limb in the tree. This can be avoided by selecting the sample in two stages. First obtain a sample of primaries - will usually only be one or two. Then within each selected primary draw a sample of terminal limbs. The selection process in each primary is done independently of the other primaries. This has a distinct advantage over the single stage sampling scheme because only the terminal limbs within the <u>selected</u> primaries need be measured. The probability of any primary (P_i) being sampled is the primary CSA divided by the sum of all primary CSA's. The probability that any terminal is drawn from within a selected primary is the terminal CSA divided by the sum of the terminal CSA's in the same

Table 1.--Limb counts of cherries obtained from a tree with CSA's and probabilities of selection, Michigan, June 1968.

	CSA		Probabilities of selection			Terminal
	Primary	Terminal	P_i	T_{ij}	Z_{ij}	fruit Y_{ij}
Primary	9.4		.268			
Terminal 1		1.8		.237	.063	204
		1.1		.145	.039	343
		1.7		.223	.060	615
		1.8		.237	.063	890
		1.2		.158	.042	912
Primary 2	2.1		.060			
		.5		.333	.017	108
		.5		.334	.017	236
		.5		.333	.017	264
Primary 3	7.8		.222			
		.3		.040	.011	11
		.9		.122	.032	123
		.6		.081	.021	177
		.8		.108	.028	204
		.9		.122	.032	260
		1.4		.189	.049	506
		.9		.122	.032	319
		1.6		.216	.056	342
Primary 4	5.8		.165			
		.9		.164	.032	175
		.6		.109	.021	147
		1.2		.218	.042	321
		.5		.091	.017	153
		1.6		.291	.056	595
		.7		.127	.025	214
Primary 5	10.0		.285			
		1.7		.262	.060	302
		.9		.138	.032	120
		1.0		.154	.035	214
		.8		.123	.028	153
		.5		.077	.017	154
		1.6		.246	.056	302
Tree	35.1	28.5	1.000	5.000	1.000	8364

primary. These probabilities are designated by T_{ij} in Table 1. Note that within each primary $\sum_{j}^{M_i} T_{ij} = 1$.

The variance for the single stage case when n = 2 has been computed. To simplify the notation let Z_{ij} and y_{ij} = Z_k and y_k respectively. Then

$$\text{Var}\ (\hat{Y}) = \sum_{k\ k'}^{\binom{28}{2}} Z_k Z_{k'} \frac{(1 - Z_k - Z_{k'})}{(2 - Z_k - Z_{k'})} \left(\frac{y_k}{Z_k} - \frac{y_{k'}}{Z_{k'}} \right)^2 =$$

$$Z_1 Z_2 \frac{(1 - Z_1 - Z_2)}{(2 - Z_1 - Z_2)} \left(\frac{y_1}{Z_1} - \frac{y_2}{Z_2} \right)^2 +$$

$$Z_1 Z_3 \frac{(1 - Z_1 - Z_2)}{(2 - Z_1 - Z_2)} \left(\frac{y_1}{Z_1} - \frac{y_2}{Z_2} \right)^2 +$$

$$\vdots \qquad\qquad \vdots$$

$$Z_{27} Z_{28} \frac{(1 - Z_{27} - Z_{28})}{(2 - Z_{27} - Z_{28})} \left(\frac{y_{27}}{Z_{27}} - \frac{y_{28}}{Z_{28}} \right)^2 = 8{,}336{,}510$$

The variance for the two stage estimator when n = 2 has also been computed. For this case

$$\text{Var}\ (\hat{Y}) = \sum_{i}^{5} \frac{Y_i^2}{P_i} - Y^2_{\cdots} + \sum_{i}^{5} \frac{1}{P_i} \sum_{j<j'}^{M_i} \frac{T_{ij}\, T_{ij'}\, (1 - T_{ij} - T_{ij'})}{(2 - T_{ij} - T_{ij'})} \left(\frac{y_{ij}}{T_{ij}} - \frac{y_{ij'}}{T_{ij}} \right)$$

= 11,821,341. This is greater than the single stage variance. However, when comparing the cost of obtaining counts and measurements of sampling units in the entire tree with that for one primary, the use of the two stage estimator becomes practical.

The variances for the different estimators discussed are summarized in the following Table. Basic data was available from four cherry trees in Michigan.

Table 1.--Variances of equal and unequal probability estimators

Estimator	Tree 1	Tree 2	Tree 3	Tree 4
Single stage				
EP n = 2	17,504,043	22,482,837	20,970,672	2,082,903
PPS n = 2	8,336,510	5,322,750	14,077,800	2,695,190
Two stage				
One prim. PPS	11,821,341	6,200,724	16,502,077	2,505,862
Two prim. PPS				
Number of fruit in trees	8,364	13,250	11,311	3,352

All trees except tree four show a smaller variance for the PPS estimators. Tree four showed a poor correlation between terminal limb size and number of fruit.

In a typical situation, the only values available for calculating the variance are those obtained from the sample. Suppose that terminals y_{13} and y_{45} were selected.

Then $$\hat{Y} = \frac{1}{(2 - z_{13} - z_{45})} \left(\frac{y_{13}}{T_{13}} (1 - z_{45}) + \frac{y_{45}}{T_{45}} (1 - z_{13}) \right)$$

$$= \frac{1}{(2 - .060 - .056)} \left(\frac{615}{.060} \; (1 - .056) + \frac{595}{.056} \; (1 - .060) \right) = 10{,}437$$

and EST. $\text{Var}(\hat{Y}) = \frac{(1 - z_{13})(1 - z_{45})(1 - z_{13} - z_{45})}{(2 - z_{13} - z_{45})^2} \left(\frac{y_{13}}{z_{13}} - \frac{y_{45}}{z_{45}} \right)^2 \doteq 31{,}065$

Now to extend the problem to a two stage case suppose that one primary (say P_3) was selected and then two terminals within this primary are sampled (say y_{32} and y_{38}). Then

$$\hat{Y} = \frac{1}{P_3} \; \frac{1}{(2 - T_{32} - T_{38})} \left(\frac{y_{32}}{T_{32}} \; (1 - T_{38}) + \frac{y_{38}}{T_{38}} \; (1 - T_{32}) \right) = 5{,}886$$

and EST. $\text{Var}(\hat{Y}) = \frac{1}{P_3^2} \; \frac{(1 - T_{32})(1 - T_{38})(1 - T_{32} - T_{38})}{(2 - T_{32} - T_{38})^2} \left(\frac{y_{38}}{T_{38}} - \frac{y_{32}}{T_{32}} \right)^2 =$ 1,110,900

Even though sampling was in two stages, the estimated variance only contains one component because just one first stage unit was selected. If two primary units and two terminals were chosen from each selected primary, then the formulas for the estimated total and estimated variance will be as shown on page 8 above.

References

Durbin, J. (1953). Some results in sampling theory when the units are selected with unequal probabilities. J. Roy. Stat. Soc., B15.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. J. Am. Stat. Assoc., 47.

Jessen, Raymond J. (1955). Determining the fruit count on a tree by randomized branch sampling. Biometrics.

Murthy, M. N. (1957). Ordered and unordered estimators in sampling without replacement. Sankhya - The Indian Journal of Statistics. Volume 18.

Raj, D. Sampling Theory, McGraw Hill, New York, 1968, Chapter 3.

Rao, J. N. K., and D. L. Bayless (1969). An empirical study of the stabilities of estimators and variance estimators in unequal probability sampling of two units per stratum. J. Am. Stat. Assoc., 64.

Yates, F. and P. M. Grundy (1953). Selection without replacement from within strata with probability proportionate to size. J. Roy. Stat. Soc., B15.