九州大学学術情報リポジトリ
Kyushu University Institutional Repository

NEW DEVELOPMENT OF SAMPLING DESIGNS IN FOREST INVENTORIES

Kinashi, Kenkichi

Nishizawa, Masahisa

Kitagawa, Toshio

https://doi.org/10.15017/14988

出版情報：九州大学農学部演習林報告. 35, pp.1-84, 1962-04-05. 九州大学農学部附属演習林
バージョン：
権利関係：

NEW DEVELOPMENT OF SAMPLING DESIGNS IN FOREST INVENTORIES

Kenkichi KINASHI *
Masahisa NISHIZAWA**
Tosio KITAGAWA ***

Contents

Chapter 1. Introduction

1. Foreward

Sampling methods in forestry have been developed quite rapidly since the publication of Schumacher and Chapman's " Sampling Methods in Forestry and Range Management, 1948. "

In virtue of the development of mathematical statistics, various computation machines, Bitterlich's original method, and wonderful expansion of use of aerial photography, sampling techniques are employed through the forests all over the world.

2. Why does forestry need sampling techniques ?

The answer for this question is very clear in itself. Forest is a population of trees. The science and practice depend on this population. We always need the information of this population. Since forest mensuration started about two centuries ago, almost

* Assistant Professor of Forestry, Kyushu University, Fukuoka.
** Forest Statistician, The Government Forest Experiment Station, Tokyo.
*** Professor of Mathematical Statistics, Kyushu University, Fukuoka.

This article was presented by K. KINASHI at Symposium of Forest Inventories, the Tenth Pacific Science Congress, Honolulu, Hawaii, Aug. 24, 1961.

Special permission for publication by Prof. S. ŌNO, Dean of Kyushu University Forests, is greatly appreciated.

all efforts have been paid for how we gain the information of this population. Recently economic situation is pushing more strongly this tendency. Before the World War II, Japan never used random sampling methods in forestry. Students at that time had been told that since Japanese mountain is so steep, it might be impossible to use them. Selective sampling or purpositive sample area method was necessarily used and nobody believed that it must not be the best way for forest inventory in Japan. Even in the present time, some persons who have much experiences in forestry always say that the selective sample area method is the fast, easy and the best way for Japanese forest inventory. However, these persons must read the following quotation from Bruce and Schumacher's Forest Mensuration, 1950, page 107: "In selective sampling an observer familiarizes himself with the forest in question and then picks out a number of plots which appear to him to be representative of the forest. These plots are then carefully measured and the results combined into an average which is assumed to represent the entire forest. While occasionally very good results have been obtained by this methods, it is completely dependent on the judgement of the man who choose the plots. Further, it is futile to compute the dispersion of the results of the several plots thus selected, for the process of selection in itself is an attempt to minimize this dispersion. There is not slightly relation between the dispersion of the plots and the dispersion of the universe from which they have been chosen. Attempts to calculate the standard error of estimate from such plots lead to no useful conclusions."

So, it is quite clearly understandable that the selective sampling which had been employed in the pioneer days of every country should be disappeared. In spite of this idea, some foresters in Japan still like the selective sampling area method much more. On the other hand, sampling theory applied to forestry is so deeply introduced that more difficult and very complicated discussions appeared in our Journal of Forestry.

Mathematical theory of sampling is always giving a useful background of practical inventory. It should, however, be noticed that these are a quite different point between the mathematical population and the actual field population. The investigation of the latter needs very much physical energy. Sampling design must cover these points. Even though the theory is excellent, it is no useful unless the field work of the investigation is easy.

Chapter 2. Practical Approach of Sampling Inventory in Private and the National Forests in Japan

1. Private Forest

(1) Forest inventory in 1953

Japanese Forest Agency has started to develop the private forest inventory system by issuing the field manual of forest sampling in 1953. In this inventory system, all over Japanese forest land, are divided into 2086 district areas each having 18,000 ha on average. Moreover, a district area is divided into many small patch of stand which will be registered on individual ownership in the official keeping book. Such small stands are usually called as "Hitsu" which means merely a single line of Chinese pencil from old style official book. From a stand of forest policy, large district units are divided into several compartments or blocks. Each individual "Hitsu" is never surveyed in the boundary. At the first time the boundary of compartments are drawn and the area are measured by use of aerial photograph. And then the area of Hitsu may be adjusted. In this case, the multiple regressions were used, in which the independent variables were the adjusted area and the ocular estimate of volume, and the dependent variable was the volume of Hitsu. This procedure was carried out through all the private forest lands in Japan. The number of selected Hitsu was decided as in the following manner. For a district forest area, the total volume estimation was permitted to have 20% sampling error in probability 95%. If we denote as follows:

σ: the standard error,

v: the estimated volume,

$$2\frac{\sigma}{v} \leq 0.2 \qquad \therefore \frac{\sigma}{v} \leq 0.1$$

Now if ρ is the coefficient of multiple correlation between actual volume, adjusted area and eye estimated volume, and the number of the sampling units is n, the following equation may be the relation between the standard deviation σ_v and the standard error of the volume.

$$\sigma = \sqrt{\frac{1-\rho^2}{n}}\sigma_v$$

$$\frac{\sigma_v}{v}\sqrt{\frac{1-\rho^2}{n}} \leq 0.1$$

$$\text{if } \frac{\sigma_v}{v} = 1.6 \text{ and } \rho = 0.9,$$

then we have $n=49$ i.e. n is nearly equal to 50.
Consequently if $n=50$, we can obtain the desired precision of the above mentioned.

For each district forest area, 50 sampling units or Hitsu may be chosen with equal interval in cards of the list which are classified by species and volume. Every tree measurement in sample Hitsu was done. Usually unit Hitsu are about 0.3 ha (slightly larger than one acre) on average. But if these units are too larger, the following number of blocks were sampled within a Hitsu, where the area of block was about 0.1 ha.

Area of Hitsu		Number of Block
less than	3 ha	4
more than less than	3 ha 7 ha	5
more than less than	7 ha 10 ha	6
more than	10 ha	8

Now, if we put eye estimate, adjusted area and actual volume x, y and v respectively for each stand (Hitsu), the equation

$$v = ax + by + c$$

may be set up by the least square solution.

Next, for all stand (Hitsu) in the list of forest keeping book, total number of stands N, the average eye estimation $\bar{X}$, and the average adjusted area Y were given numerically, and then the total volume of the forest district area V_T was decided by the following equation.

$$V_T = (a\bar{X} + b\bar{Y} + c)\,N$$

The confidence limit with the probability 95% was shown as follows:

$$l = V_T \pm 2\sqrt{\frac{1-\rho^2}{n}}\ \sigma_v\ N$$

This inventory system gave the following results:

		Range	Average	Mode
(i)	Coefficient of variation;	0.9–5.0	1.88	1.6

(ii) Correlation coefficient between eye estimate of volume and the actual volume;
the majority of the cases are more than 0.95 and about 10% cases are less than 0.8.

(iii) Total stand number (number of Hitsu);
district area involving maximum number of stands covers 30,000–40,000 Hitsu and average number (mode) 4,000 Hitsu

(iv) Precision aimed at the number of district area in percentage;

less than	20%	88%
more than	20%	12%

(2) Forest inventory in 1954, 1955

Japan has had her experience of sampling survey at the first occasion and she has continued some new development in 1954, 1955 for 44 district forest areas in 10 special provinces, as the test of sampling survey.

District forest area was stratified into soft wood forest, mixed forest, hard wood forest and several volume size classes. Fifty Hitsu (Stand), as same as the previous case, were chosen and measured in volume. But it is a different point from the first survey in which multiple regression was used instead of the linear regression between eye estimate x and actual volume v. This linear regression was based on Hitsu stand, and the formula was

$$v = ax + b$$

If the average volume and the total number of Hitsu were given, the total volume formula was,

$$V_T = (a\bar{x}+b)\ N$$

And the confidence limit was

$$l = V_T \pm 2\sqrt{\frac{1-\rho^2}{n}}\ \sigma_v\ N$$

The result shows:

(i) The mode of actual volume per survey unit in 1953 was about 100 Koku, while that was 100–250 Koku, 1954.

(ii) Coefficient of variation ($\frac{\sigma_v}{v}$) was 0.9–2.5 in the range, 1.57 in average and 1.3 in mode.

(3) Large national survey

Parallel to this test survey, Japan has carried out the large National Forest survey which aimed to obtain the large data of Japanese forest resources covering all over Japanese island, taking one year.

Total forest stock survey over Japan, with less than 5% sampling error in 95% probability, started. The calculation method of the needed plot number was as follows:

$$0.05 \geq \frac{2}{\sqrt{n}} \cdot \frac{\sigma}{\bar{X}}$$

where n is the number of plot, $\bar{X}$ is the average volume, σ is the standard deviation of volume. And the using plot size is 50 m $\times$ 30 m = 0.15 ha. Considering that the distribution of stocking per unit plot are steep L shape type, we assume $\sigma/\bar{X} = 1.3$.

Then we have $0.05 \geq \frac{2 \times 1.3}{\sqrt{n}}$ from the above formula,

$n \geq 2704$

n is nearly equal to 3000.

It is one point per 123 km^2 or (11.1 km)^2. Sampling point was allocated proportionally to the area of province. As the sampling frame, the net of lattice with 1 minute in both latitude and longitude was used. This point is the intersection of both sides of rectangular, about 1800 m $\times$ 1600 m. This net covered the area of each province. Random sampling points were drawn randomly from the population of these intersectional points within the area of province proportionally to that area. A rectangular sample plot was made by taking distance 50 m toward North and distance 30 m toward East respectively, from the sampling point. Within a plot standing trees, field, cultivated area, town, residential area were marked on the field map. For standing trees the diameter at breast height and the total height of tree are measured. For each province, random sampling calculation was made separately and also the sampling error was estimated. At the same time, the percentages and ratio of land class utilization, property of ownership, forest type and forest class were estimated. All results of this sampling inventory were completely finished. However, the obtained results were slightly higher than the numerical data of management plan of the National and private forests. Depending upon this point, the results were not published finally, because they were afraid that the results might introduce some confusion of Japanese forest policy, and the results might give a great influence to 5 year periodical forest inventory.

(4) Since 1957

New forest inventory system has started from 1953, and after it, the 2nd 5 year periodic forest inventory began from 1957. In that time a part of forest inventory system was changed. The major different point of the estimation of the total volume is as follows:

There is a large management plan basic area and under it several district forest area are made. The sample selection is restricted within the district forest area keeping basis as a whole unit. The stratification is made as the following classification.

(i) Cutting area, bare area, bamboo forest, and less than II age class (10 years) forest. (10%)
(ii) Soft wood forest, older than the regular cutting period. (40%)
(iii) Soft wood forest, younger than the regular cutting period. (20%)
(iv) Hardwood forest. (30%)

The above percentages within brackets show sample size in each stratum. Degree of precision is that the reliance percentages are fixed at 4% within the basic area and at 8% within the district area, both with the probability 95%.

For the management plan basic area, about 2,000 plots are necessary and the total sum is 600,000 plots during 5 years.

From the forest keeping book, the working circle is stratified. Section lines from east to west with interval 1 cm and from south to north with interval 2 cm respectively, cover the province or the management plan basic area. From the selected original point, one side 30 m toward North and another side 33.3 m toward west construct the rectangular plot in which every tree measurement is made. For younger stand, plot size is 20 m×25 m (0.05 ha). The calculations are made by the method of stratified random sampling.

2. National Forest Inventory

Since 1958, National Forest began the forest sampling inventory for the management plan to obtain the information of the stocking of National Forest timbers. For a working circle, all stands and forests should be surveyed by stratified random sampling and for no working circle, both the 1st kind forest and the 3rd kind forest should be surveyed by the stratified subsampling.

Especially, in Hokkaido, all forest should be surveyed by the stratified subsampling, because there are all most natural hard wood type mixing with soft woods. By using the aerial photographs and brief mapping methods, the identification of compartments or subcompartment and the boundary of stands may be made fast on the aerial photographs. Mapping of working circle and listing of forest stocking may be adjustable partially. Stratification should be based on working circle, forest type or natural or artificial, main species group and age class.

The size and shape of plot is usually rectangular type 0.1 ha (25 m×40 m) or (20 m×50 m). For younger stand 0.05 ha (20 m×25 m) or 0.04 ha circular plot may be taken. But in Hokkaido, 0.1 ha or 0.15 ha rectangular plot may be preferable. The standard of precision aimed at is the following:

Precision aimed at	Forest objected
less than 10 %	working circle
less than 15 %	3rd kind forest, non working circle
less than 20 %	1st kind forest, non working circle

Sample size may be calculated by the following formula.

$$\text{Sampling intensity}: \varphi = \left(\frac{2}{E}\right)^2 \cdot \frac{\sum N_i s_i^2}{\sum N_i \bar{x}_i}$$

$$\text{Plot size} \quad : n_i = \varphi N_i,$$

where E is a precision aimed at, N_i is the size of i stratum, s_i^2 is the variance of a stratum, and $\bar{x}_i$ is average volume per plot of a stratum. Allocation is proportional. After the first sampling inventory, Neymann's allocation may be taken according to the results of the first year. Working map may be covered with section paper of 5 mm ×5 mm or 10 mm×10 mm interval. Random selection for intersection may be made within each stratum. From the map and photograph, actual point may be found. Growth prediction based on growth increment borer and also the permanent plot may be set, 1 plot per 500 ha in main land and 1 plot per 1000 ha in Hokkaido.

Finally, the Japanese Forest Agency, up-to-date, is planning to push the program of obtaining the information of the total land forest resources. Moreover, based on abounded experiences, the successive sampling will be developed to gain the moving forest stuation and predictable information of cutting and growth or mortality.

There should be two occasion sampling system applied in future of Japanese forest inventory.

Chapter 3. Classification of Forest Measurement Method Mainly from a Point of View of Sampling Techniques

1. Forest Mensuration

H. Arthur Meyer's book (1953) says that " Forest Mensuration includes the measurement of forest products and the determination of timber volume and forest growth. In recent years aerial photographs have been used increasingly by foresters for mapping and other purpose ". Forest mensuration is concerning both time and space. Time conception may be necessary to growth problems. Space conception may be connected to the standing point from where the measurement is projected. Originally tree is standing on the ground. Three angles must be appeared. The first point is the most popular, one from ground, and the second and third points are recently developed from sky or over and under ground.

Above ground measurement or aerial photograph may not be available for growth, while under ground or soil survey may be available for the background of growth.

In this point, we may try some broad classification of forest mensuration and we clarify the connection of sampling methods.

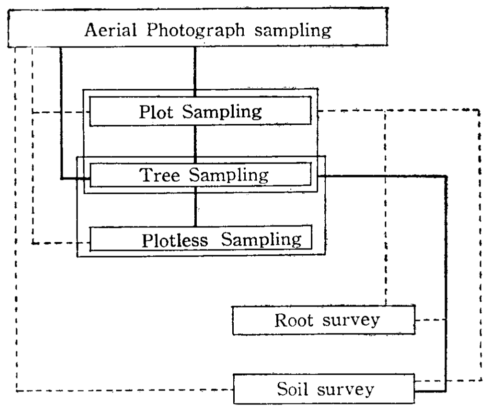
2. Classification of Forest Mensuration

Space	Forest Measurement Method By	Available Measure
Above ground	Aerial photography	Area and volume
Ground	Tree sampling Plot sampling Plotless sampling	Area volume and growth
Under ground	Root survey Soil survey	Growth

It is always necessary to adjust the measurement of both above and under ground to ground. Measurement on ground is of course essential but it is not always sufficient. Stratification and boundary identification on aerial photograph is far superior and growth prediction needs absolutely under ground measurement. Our sampling techniques should use whole information area sufficiently. Combination of them must be excellent.

3. Combination of various kind sampling

The next diagram shows some relationship or connection with various kind of sampling in forest mensuration work. Bold lines represent more strong connection than dotted lines. Three sampling techniques on the ground which situate in the center of the diagram are connected with each other. Of the three sampling, tree sampling is more directly connected to tree measurement and always it is ultimate sampling unit of the other two. Plot sampling is a cluster sampling of trees and of course it is ultimate unit or random sampling unit itself. So tree sampling and plot sampling are the base of forest sampling. Plotless sampling or Bitterlich's sampling is not necessary



basic sampling. In the same meaning, aerial photograph sampling and soil sampling are not direct sampling. If we try two classes of sampling of forest mensuration, we may say that tree and plot sampling belongs to direct sampling, while plotless, aerial photograph and soil survey may belong to indirect sampling. However, our field is so wide and so much variable that one method is not always sufficient. Therefore, we should try various kind of sampling and their combination works. The combination or connection is very important and useful for forestry. Forestry work needs always the sampling of this kind. There are a reasonable way, i.e. regression methods which are used from old time. Now, in the present time, we call them, double sampling and triple sampling, cluster sampling and multiple regression, sometimes yield table, volume table, site table, aerial volume table and so on. They are much concerned to cost and variance. Recently, inventory, research, mensuration work and survey in all field of forestry are depended upon the combination work of this kind, basically regression methods. And in the long run, they are due to the cost and variance. Sampling theory is basically important and useful background of various mensuration works in forestry. Cost represents the economical aspect and variance represents biological or forestry aspect. We should always keep the balance of both.

Chapter 4. Statistical Methodology for Use of Forest Measurement

1. Introduction

Sukhatme defined sampling method in his statistical book " Sampling theory of surveys with applications " as a method of selecting a fraction of the population in a way that the selected sample represents the population, and he said everyone of us has had occasion to use it. Cochran pointed out that the principal advantages of sampling method as compared with complete enumeration are next four items, i.e. reduced cost, greater speed, greater scope and greater accuracy. He stated too that the principal steps in a sample survey are (i) statement of the objectives of the survey, (ii) definition of the population to be sampled, (iii) determination of the data to be collected, (iv) method of measurement, (v) choice of sampling unit, (vi) selection of the sample, (vii) organization of the field work, (viii) summary and analysis of the data, (ix) information gained for future surveys.

Sampling methods are often used in forest measurement, for instance, they are sample tree and sample plot methods. As, in conventional sample tree and plot methods, samples those which it seems average are selected purposely and we estimate the population value by this result, we cannot believe firmly whether we obtained reliable estimate and we cannot evaluate the sampling error which is the difference between the sample estimate and the population value. Whatever be the method of selection, a sample estimate will differ from the one that would be obtained from enumerating the complete population with equal care. We must adopt ourselves the scientific sampling methods in order to obtain the reliable value and able to evaluate the sampling error in forest measurement.

In chapter 4, we will represent various sampling methods in which we will illustrate formally the calculation methods of estimated mean, variance, variance of mean, confidence limits and sample size with desired degrees of precision and some notes. Now, referring to " a dictionary of statistical terms " (Kendall and Buckland), we shall explain simply some of statistical terms necessary to these formulae.

- Statistics (Sample estimate) : Sample value to estimate the population value.
- Unbiased estimate: An estimate will vary from sample to sample, depending upon the units included in the sample. Such sample values as these average over the totality of samples are equal to the corresponding population values are called unbiased estimate of the population values.
- Degrees of freedom: The number of independent coordinate values which are necessary to determine a dynamical system.
- Variance: A mean of squares of variations from the arithmetic mean.
- Covariance: The first product-moment of two variates about their mean values.
- Standard deviation: The measure of dispersion of a frequency distribution and it is equal to the positive square root of the variance.
- Standard error: Positive square root of the variance of sampling distribution of a statistics.
- Standard error of estimate: Standard deviation of the observed values about a regression line.
- Sampling unit: One of the unit into which an aggregate is divided (or regarded as divided) for the purpose of sampling, each unit being regarded as individual and indivisible when the selection is made. The examples of sampling unit are individual tree, a group of tree, subcompartments, compartments etc. on natural basis, or are as defined by grid coordinates on a map (plot) on some arbitrary basis.
- Sampling error: The difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a sample of values is observed.
- Non-sampling error: An error in sample estimate which can not be attributed to sampling fluctuations.
- Sampling fraction: The proportion of the total number of sampling units in the population. It is called sampling ratio too.
- Sample size: The number of sampling units which are to be included in the sample.
- Accuracy: The closeness of computations or estimate to the exact or true values.
- Precision: In general, it is a quality associated with a class of measurements and refers to the way in which repeated observations conform to themselves. In somewhat narrower sense refers to the dispersion of the observations, or some measure of it, whether or not the mean value around which the dispersion is measured approximates to the true value.
- Coefficient of variation: The standard deviation of a distribution divided by the arithmetic mean. It is sometimes multiplied by 100, and it is used for the purpose of comparing the variabilities of frequency distributions, but is sensible to errors in the means and is of limited use.
- Confidence interval: If it is possible to define two statistics t_1 and t_2 (functions of

sample values only) such that, θ being a parameter under estimate

$$P_r(t_1 \leq \theta \leq t_2) = \alpha,$$

where α is some fixed probability, the interval between t_1 and t_2 is called a confidence interval. The assertion that θ lies this interval will be true, on the average, in a proportion α of the cases when the assertion is made.

- Confidence coefficient: The measure of probability α associated with a confidence interval expressing to the probability of the truth of a statement that the interval will include the parameter value.
- Confidence limits: The value t_1 and t_2 which from the upper and lower limits to the confidence interval.
- Cost function: A function giving the cost of obtaining the sample as a function of the relevant factors affecting cost. It may relate to only a part of the entire cost, e.g. by providing for the cost of collecting the sample but not for the cost of tabulation.
- Efficiency: An objective measure of the relative merits of several possible estimators. An estimator is regarded as more efficient than another if it has smaller variance.
- t-distribution: The distribution is usually written the form

$$dF = \frac{\Gamma\, 1/2\,(\nu+1)}{(\nu\pi)\Gamma(1/2\,\nu)}\left(1+\frac{t^2}{\nu}\right)^{-1/2(\nu+1)} dt\,, \qquad -\infty \leq t \leq \infty$$

where ν is the number of degrees of freedom. The distribution is, among other things, that of the ratio of a sample mean (measured from the parent mean) to a sample variance, multiplied by a constant, in sample from a normal population. It is thus independent of the parent scale parameter and can be used to set confidence intervals to the mean independently of the parent variance.

- t-table: A numerical table of t-distribution showing t-values according to confidence coefficients and degrees of freedom.

2. Symbols

Now, we shall explain the principal symbols which are used in 3.

Items	Symbols of	
	Population	Sample
area	A	a
numbers of sampling unit	N	n
numbers of sampling unit in stratum	N_h	n_h
proportion of the numbers of sampling unit of stratum in the population	$W_h = N_n/N$	$w_h = n_h/n$
proportion	P	p
volume (diameter, height, increment etc.)	Y	y
mean volume	$\overline{Y}$ or μ	$\bar{y}$
mean volume in stratum	Y_h	$\bar{y}_h$
variance	$\sigma_Y{}^2$ or $V(Y)$	$v(y)$
variance of mean	$\sigma_{\bar{Y}}{}^2$ or $V(\overline{Y})$	$v(\bar{y})$
standard error	$\sigma_{\bar{Y}}$ or $\sqrt{V(\overline{Y})}$	$\sqrt{v(\bar{y})}$
coefficient of variation	C_Y	c_v
correlation coefficient	ρ	r
regression coefficient	B	b

cost function	C
coefficient of cost functions	c_i
precision aimed at	E
numbers of strata	L

3. Statistical methodology for use of forest mensuration

(1) Random sampling

1. mean $$\bar{y}=\overset{n}{S}(y)/n=\frac{y_1+y_2+\cdots+y_n}{n} \tag{1.1}$$

2. variance $$v(y)=\frac{\overset{n}{S}[(y-\bar{y})^2]}{n-1}$$

$$=\frac{1}{n-1}\left\{\overset{n}{S}(y^2)-\frac{[\overset{n}{S}(y)]^2}{n}\right\}$$

$$=\frac{1}{n-1}\left\{\overset{n}{S}(y^2)-\bar{y}\overset{n}{S}(y)\right\} \tag{1.2}$$

3. variance of mean

$$v(\bar{y})=\frac{N-n}{N}\ \frac{v(y)}{n} \tag{1.3}$$

where $\frac{N-n}{N}$ is called the finite population correction (fpc).

If the fpc is ignored,

$$v(\bar{y})=\frac{v(y)}{n} \tag{1.4}$$

4. confidence limits

$$N\left\{\bar{y}\pm t\sqrt{v(\bar{y})}\right\} \tag{1.5}$$

where t is a value of t table at 95% confidence coefficients based on $(n-1)$ degrees of freedom.

5. n with the desired degree of precision

$$n_0=\left(\frac{tC_v}{E}\right)^2 \tag{1.6}$$

This is adequate if n_0/N is neglected. In otherwise events compute n as

$$n=\frac{n_0}{1+n_0/N} \tag{1.7}$$

It is possible to apply these formulae to sampling for proportions and percentages.

Notes:

1. When we have no informations about the forest population, it is random sampling method that the sampling units are drawn from the population at random and we estimate the desired items of population (i.e. mean d.b.h., mean height, basal area per ha, total volume, total growth, percentage of mortality and percentage of tree species

etc.) by the results of these units.
2. It is often used in the forest inventory of small area and it is important as the basic theory of the various sampling methods. For instance, stratified random sampling method is constructed by random sampling method in each stratum.
3. It is difficult to practical use for the large scale forest inventory and when we have some informations about the population, we had better use stratified random, ratio and regression methods.
4. We can adopt this theory when sample sizes are large (e.g. more than 50), because the sample mean are normally distributed independently of the distribution of universe. When sample sizes are small, we must pay attention to the distribution of universe. When the distribution is Poission or Binomial, the estimation methods of mean and variance are different from those of the random sampling method.

(2) Stratified random sampling

1. mean

$$\bar{y}_{st}=\frac{1}{N}\overset{L}{S}(N_h\bar{y}_h)$$

$$=\overset{L}{S}(W_h\bar{y}_h)$$

$$=\frac{1}{N}\overset{L}{S}\left\{\left(\frac{N_h}{n_h}\right)\overset{n_h}{S}(y_h)\right\} \tag{2.1}$$

2. variance of stratum mean

$$v(\bar{y}_h)=\frac{N_h-n_h}{N_h}\ \frac{v(y_h)}{n_h} \tag{2.2}$$

where

$$v(y_h)=\frac{1}{n_h-1}\overset{n_h}{S}\left[(y_h-\bar{y}_h)^2\right]$$

$$=\frac{1}{n_h-1}\left[\overset{n_h}{S}(y_h{}^2)-\frac{\{\overset{n_h}{S}(y_h)\}^2}{n_h}\right]$$

$$=\frac{1}{n_h-1}\left\{\overset{n_h}{S}(y_h{}^2)-\bar{y}_h\overset{n_h}{S}(y_h)\right\} \tag{2.3}$$

If the fpc is ignored

$$v(\bar{y}_h)=\frac{v(y_h)}{n_h} \tag{2.4}$$

3. variance of mean

$$v(\bar{y}_{st})=\frac{1}{N^2}\overset{L}{S}\left[N_h{}^2v(\bar{y}_h)\right]$$

$$=\frac{1}{N^2}\overset{L}{S}\left\{N_h(N_h-n_h)\frac{v(y_h)}{n_h}\right\}$$

$$=\overset{L}{S}\left\{\frac{W_h{}^2v(y_h)}{n_h}\right\}-\overset{L}{S}\left\{\frac{W_hv(y_h)}{N}\right\} \tag{2.5}$$

If the fpc is ignored.

$$v(\bar{y}_{st}) = \overset{L}{S}\left\{\frac{W_h^2 v(y_h)}{n_h}\right\} \tag{2.6}$$

4. confidence limits

$$N\{\bar{y}_{st} \pm t\sqrt{v(\bar{y}_{st})}\} \tag{2.7}$$

where t is a value of t table at 95% confidence coefficient based on $(n-L)$ degrees of freedom.

5. allocation of the sample sizes n in the respective strata

(i) Optimum allocation with varying costs.

The variance of the population mean in the stratified random sampling is as follows:

$$V(\bar{y}_{st}) = \overset{L}{S}\left\{\frac{W_h^2 V(y_h)}{n_h}\right\} - \overset{L}{S}\left\{\frac{W_h V(y_h)}{N}\right\} \tag{2.8}$$

Now, we consider the next cost function, which expresses the cost of taking the sample in terms of the sample sizes n_h

$$\text{Cost} = C = c_0 + \overset{L}{S}(c_h n_h) \tag{2.9}$$

Within any stratum, the cost mounts directly with the size of sample, but the cost per unit, c_h, may vary from stratum to stratum. The symbol c_0 represents an overhead cost. We may need n which the variance of the estimated mean (2.8) is a minimum with a cost function of the form (2.9). We minimize

$$\begin{aligned}\varphi &= V(\bar{y}_{st}) + \lambda C \\ &= \overset{L}{S}\left\{\frac{W_h^2 V(y_h)}{n_h}\right\} - \overset{L}{S}\left\{\frac{W_h V(y_h)}{N}\right\} + \lambda\{c_0 + \overset{L}{S}(c_h n_h)\}\end{aligned}$$

where λ is the Lagrange multiplier.

Differentiation with respect to n_h gives the equation.

$$-\frac{W_h^2 V(y_h)}{n_h^2} + \lambda c_h = 0$$

$$n_h\sqrt{\lambda} = \frac{W_h\sqrt{V(y_h)}}{\sqrt{c_h}} \tag{2.10}$$

Summing over all strata, we obtain

$$n\sqrt{\lambda} = \overset{L}{S}\left\{\frac{W_h\sqrt{V(y_h)}}{\sqrt{c_h}}\right\} \tag{2.11}$$

The ratio of (2.10) to (2.11) gives

$$\frac{n_h}{n} = \frac{W_h\sqrt{V(y_h)}/\sqrt{c_h}}{\overset{L}{S}\{W_h\sqrt{V(y_h)}/\sqrt{c_h}\}} = \frac{N_h\sqrt{V(y_h)/c_h}}{\overset{L}{S}\{N_h\sqrt{V(y_h)/c_h}\}} \tag{2.12}$$

a) When cost is fixed, we substitute the optimum value of n_h in the cost function (2.9) and solve for n. This gives

$$n = \frac{(C-c_0)\overset{L}{S}\{N_h\sqrt{V(y_h)/c_h}\}}{\overset{L}{S}\{N_h\sqrt{V(y_h)c_h}\}} \tag{2.13}$$

b) When variance is fixed, we substitute the optimum value of n_h in the formula (2.8). We obtain

$$n = \frac{[\overset{L}{S}\{W_h\sqrt{V(y_h)}\sqrt{c_h}\}]\,[\overset{L}{S}\{W_h\sqrt{V(y_h)}/\sqrt{c_h}\}]}{V(\bar{y}_{st}) + \frac{1}{N}\overset{L}{S}\{W_h V(y_h)\}} \tag{2.14}$$

(ii) Optimum allocation with a fixed total size of sample. The problem is to minimize (2.8) subject to the restriction

$$n_1 + n_2 + \cdots + n_L = n$$

Then we minimize

$$\varphi = V(\bar{y}_{st}) + \lambda\,(n_1 + n_2 + \cdots + n_L - n)$$

Similar to that the case of (i), we obtain

$$\frac{n_h}{n} = \frac{W_h\sqrt{V(y_h)}}{\overset{L}{S}\{W_h\sqrt{V(y_h)}\}} = \frac{N_h\sqrt{V(y_h)}}{\overset{L}{S}\{N_h\sqrt{V(y_h)}\}} \tag{2.15}$$

In this case, we have the next minimum variance

$$V(\bar{y}_{st})_{\text{opt}} = \frac{[\overset{L}{S}\{W_h\sqrt{V(y_h)}\}]^2}{n} - \frac{\overset{L}{S}\{W_h V(y_h)\}}{N} \tag{2.16}$$

If the fpc is ignored, we have

$$V(\bar{y}_{st}) = \frac{[\overset{L}{S}\{W_h\sqrt{V(y_h)}\}]^2}{n} \tag{2.17}$$

(iii) Proportional allocation

If we assume that $V(y_h)$ is identical with all strata in (2.15), we have

$$n_h = W_h n \tag{2.18}$$

In this case, we can compute mean and variance of mean simply. The results are as follows:

mean

$$\bar{y}_{st} = \frac{\overset{L}{S}(n_h\bar{y}_h)}{n} = \frac{\overset{L}{S}\overset{n_h}{S}(y_h)}{n} = \frac{\overset{n}{S}(y)}{n} \tag{2.19}$$

variance of mean

$$v(\bar{y}_{st}) = \frac{N-n}{N}\;\frac{\overset{L}{S}[W_h v(y_h)]}{n}$$

$$= \frac{N-n}{N}\;\frac{v(y_w)}{n} \tag{2.20}$$

If the fpc is ignored

$$v(\bar{y}_{st}) = \frac{\overset{L}{S}[W_h v(y_h)]}{n} = \frac{v(y_w)}{n} \tag{2.21}$$

where

$$v(y_w) = \overset{L}{S}[W_h v(y_h)] \tag{2.22}$$

6. n with the desired degree of precision

(i) Optimum allocation

We use the next form of the pooled standard deviation of strata.

$$\bar{\sigma}_w = \overset{L}{S}\left[W_h \sqrt{V(y_h)}\right] \tag{2.23}$$

and we compute the coefficient of variation

$$(C_v)_{st} = \frac{\bar{\sigma}_w}{\bar{Y}_{st}} \tag{2.24}$$

Then we obtain n with the desired degree of precision from the next equation

$$n_0 = \left\{ \frac{t(C_v)_{st}}{E} \right\}^2 \tag{2.25}$$

This is adequate if n_0/N is neglected. In otherwise events we compute n as

$$n = \frac{n_0}{1+n_0/N} \tag{2.26}$$

(ii) Proportional allocation

We use the next form of the pooled variance of strata

$$\sigma_w^2 = \overset{L}{S}\{W_h V(y_h)\} \tag{2.27}$$

and we compute the coefficient of variation

$$(C_v)_{\text{prop}} = \frac{\sigma_w}{\bar{Y}_{st}} \tag{2.28}$$

Then we obtain

$$n_0 = \left\{ \frac{t(C_v)_{\text{prop}}}{E} \right\}^2 \tag{2.29}$$

This is adequate if n_0/N is neglected. In otherwise events we compute n as

$$n = \frac{n_0}{1+n_0/N} \tag{2.30}$$

7. Relative precision of stratified random and simple random sampling. If the fpc is ignored,

variance of random sampling:

$$V_{\text{ran}} = \frac{\sigma^2}{n} \qquad \text{from (1.4)}$$

variance of stratified random sampling by proportional allocation:

$$V_{\text{prop}} = \frac{\overset{L}{S}(W_h \sigma_h^2)}{n} \qquad \text{from (2.21)}$$

variance of stratified random sampling by optimum allocation:

$$V_{\text{opt}} = \frac{\{\overset{L}{S}(W_h \sigma_h)\}^2}{n} \qquad \text{from (2.17)}$$

From the standard algebraic identity for the analysis of variance of the stratified population, we have

$$(N-1)\sigma^2 = \overset{L}{S}\{(N_h-1)\sigma_h^2\} + \overset{L}{S}\{N_h(\mu_h-\mu)^2\}$$

Since terms in $1/N_h$ are negligible, this may be written

$$N\sigma^2 = \overset{L}{S}(N_h\sigma_h^2) + \overset{L}{S}\{N_h(\mu_h - \mu)^2\}$$

Hence

$$V_{\text{ran}} = \frac{\sigma^2}{n} = \frac{\overset{L}{S}(N_h\sigma_h^2)}{nN} + \frac{\overset{L}{S}\{N_h(\mu_h-\mu)^2\}}{nN}$$

$$= \frac{\overset{L}{S}(W_h\sigma_h^2)}{n} + \frac{\overset{L}{S}\{W_h(\mu_h-\mu)^2\}}{n}$$

$$= V_{\text{prop}} + \frac{\overset{L}{S}\{W_h(\mu_h-\mu)^2\}}{n}$$

The difference of V_{opt} and V_{prop} is

$$V_{\text{prop}} - V_{\text{opt}} = \frac{1}{n}\left[\overset{L}{S}(W_h\sigma_h^2) - \frac{\{\overset{L}{S}(W_h\sigma_h)\}^2}{N}\right]$$

$$= \frac{1}{n}\overset{L}{S}\{W_h(\sigma_h - \bar{\sigma})^2\} \geq 0$$

Hence

$$V_{\text{ran}} = V_{\text{opt}} + \frac{\overset{L}{S}\{W_h(\sigma_h-\bar{\sigma})^2\}}{n} + \frac{\overset{L}{S}\{W_h(\mu_h-\mu)^2\}}{n}$$

Then we have

$$V_{\text{opt}} \leq V_{\text{prop}} \leq V_{\text{ran}}$$

where the optimum allocation is for fixed n, i.e. $n_h \propto N_h\sigma_h$.

Notes :

1. We can use the stratified random sampling when all units in the population can be allocated and listed on every stratum. In the forest inventory, forest note and aerial photograph are used as the tool of stratification.
2. The object of the stratified random sampling.

(i) It is often necessary to obtain the information of the every stratum. In the forest inventory, we need the information of tree species, age classes, management classes and working unit, so we must divide our population into these strata. We often call these strata subpopulation.

(ii) Stratification due to the practice and administration of survey. For instance, it is convenient to stratify the administration area of forest office and forest Bureau.

(iii) Stratification for the purpose of increasing the precision of survey. If we stratify our objective forest into the area having resemble variance by aerial photograph, we may expect to increase the precision of survey. When at first we stratify by mean of (ii) and then (iii), we call this stratified method the double stratification. This methods are commonly used in our survey.

(iv) It is possible to apply these above formulae to stratified random sampling for proportions and percentages.

(3) Cluster sampling

1. mean

$$\bar{y}_{cl} = \frac{1}{m} \overset{mN_i}{SS}(y_{ij}) = \frac{1}{m} \overset{m}{S}(y_i) = \frac{\overset{m}{S}(N_i \bar{y}_i)}{m} \tag{3.1}$$

where

$$\bar{y}_i = \frac{1}{N_i} \overset{N_i}{S}(y_{ij}) = \frac{y_i}{N_i}$$

2. variance

$$v(y_{cl}) = \frac{1}{m-1} \overset{m}{S}[(y_i - \bar{y}_{cl})^2]$$

$$= \frac{1}{m-1}\left[\overset{m}{S}(y_i^2) - \frac{\{\overset{m}{S}(y_i)\}^2}{m}\right] \tag{3.2}$$

3. variance of mean

$$v(\bar{y}_{cl}) = \frac{M-m}{M} \frac{v(y_{cl})}{m} \tag{3.3}$$

If the fpc is ignored,

$$v(\bar{y}_{cl}) = \frac{v(y)}{m} = \frac{\overset{m}{S}[(y_i - \bar{y}_{cl})^2]}{m(m-1)} \tag{3.4}$$

4. confidence limits

$$M\{\bar{y}_{cl} \pm t\sqrt{v(\bar{y}_{cl})}\} \tag{3.5}$$

where t is a value of t table at 95% confidence coefficient based on $(m-1)$ degrees of freedom.

5. m with the desired degree of precision

$$m_0 = \left(\frac{tC_v}{E}\right)^2 \tag{3.6}$$

where

$$C_v = \frac{\sqrt{V(y)}}{\bar{Y}_{cl}}$$

This is adequate if m_0/M is neglected. In otherwise events we compute m as

$$m = \frac{m_0}{1 + m_0/M} \tag{3.7}$$

Notes:

1. In the forest inventory, the primary sampling unit is a plot and the secondary sampling unit is a tree. All trees in a plot are measured and plots are selected at random from the population.

2. We must select the plot size which is heterogeneous within plots and homogeneous between plots. An example is the strip survey in the forest. It takes many times to measure a strip, but we are sufficient only to take a few strips.

3. When all secondary sampling units are not surveyed but sampled at random, the theory of subsampling (5) may be applied.

4. It must be noted that in stratified random sampling within strata is homogeneous

and between strata is heterogeneous, on the other hand in cluster sampling the case is reverse.

(4) Subsampling.

1. mean
$$\bar{y}_{sub}=\frac{\overset{mn}{SS}(y_{ij})}{mn} \tag{4.1}$$

where m is numbers of primary sampling unit.
n is numbers of secondary sampling unit.

2. variance

between variance:
$$v(y)_b=\frac{n\overset{m}{S}\{(\bar{y}_i-\bar{y}_{sub})^2\}}{m-1} \tag{4.2}$$

within variance:
$$v(y)_w=\frac{\overset{mn}{SS}\{(y_{ij}-\bar{y}_i)^2\}}{m(n-1)} \tag{4.3}$$

3. variance of mean
$$v(\bar{y}_{sub})=\frac{1}{mn}\left\{\frac{M-m}{M}v(y)_b+\frac{N-n}{N}\frac{m}{M}v(y)_w\right\} \tag{4.4}$$

If the fpc and m/M are ignored,
$$v(\bar{y}_{sub})=\frac{v(y)_b}{mn}=\frac{\overset{m}{S}\{(\bar{y}_i-\bar{y}_{sub})^2\}}{m(m-1)} \tag{4.5}$$

This formula concides with (3.4) when $\bar{y}_{sub}$ equals $\bar{y}_{cl}$.

4. confidence limits
$$MN\{\bar{y}_{sub}\pm t\sqrt{v(\bar{y}_{sub})}\} \tag{4.6}$$

where t is a value of t table at 95% confidence coefficient based on $(m-1)$ degrees of freedom.

5. optimum sampling and subsampling fractions

The type of cost function that has proved is
$$C=c_bm+c_unm \tag{4.7}$$

The first component of cost, c_bm, is proportional to the number of units in the sample (primary sampling unit); the second, c_wnm, to the total number of elements (secondary sampling unit). The variance of the complete population is as follows.
$$V(\bar{y}_{sub})=\frac{M-m}{M}\frac{\sigma_b^2}{m}+\frac{MN-mn}{MN}\frac{\sigma_w^2}{mn}$$

where $\sigma_w^2+N\sigma_b^2$ is the variance between units and σ_w^2 is the one within units between elements.

We minimize
$$\varphi=V(\bar{y}_{sub})+\lambda(C-c_bm-c_wnm)$$
$$=\left(\frac{1}{m}-\frac{1}{M}\right)\sigma_b^2+\left(\frac{1}{k}-\frac{1}{MN}\right)\sigma_w^2+\lambda(C-c_bm-c_unm)$$

where $k=nm$.

Differentiation with respect to m and k gives, respectively,

$$\frac{\sigma_b^2}{m^2} = -\lambda c_b$$

$$\frac{\sigma_w^2}{k^2} = -\lambda c_w$$

Hence

$$\frac{k}{m} = n_{\text{opt}} = \frac{\sigma_w}{\sigma_b}\sqrt{\frac{c_b}{c_w}} \tag{4.8}$$

For practical use, the sample estimate of n is

$$n_{\text{opt}} = \frac{\sqrt{v(y)_w}\ \sqrt{n}}{\sqrt{v(y)_b - v(y)_w}}\sqrt{\frac{c_b}{c_w}} = \frac{\sqrt{n}}{\sqrt{\frac{v(y)_b}{v(y)_w} - 1}}\sqrt{\frac{c_b}{c_w}} \tag{4.9}$$

The value found will be non-integral. It is sufficient to take the nearest integer. If n_{opt} lies between the two integers n, $(n+1)$, we should choose $(n+1)$ if $n^2_{\text{opt}} > n(n+1)$; otherwise we choose n.

Notes:

1. In the forest inventory, each unit (P.S.U.) in the population can be divided into the same number of smaller units, or elements (S.S.U.). A sample of m P.S.U. has been selected. If S.S.U. within a selected P.S.U. give similar results, it seems uneconomical to measure all. In this case it is sufficient only to sampling and measuring a few S.S.U. in any chosen block. We call this method subsampling or two-stage sampling.
2. Subsampling method is commonly used and important in the forest inventory. We choose usually equal number (2–7) of plot (S.S.U.). We have an advantage that we can divide block into equal size in area and can use the formula (4.5) for variance of mean.
3. Sampling more than two-stage is called multi-stage sampling, on the other hand simple random sampling, systematic sampling and stratified random sampling are called one-stage sampling. Two-stage sampling is more flexible than one-stage sampling.
4. If two-stage sampling is to be used when the primary units vary in size, one method is to stratify by size of unit, so that the units within a stratum become equal, or nearly so, in size. There are another methods too that the units may be choosen either with equal probabilities or with probabilities proportional to size or to some estimate of size.

(5) Systematic sampling

We need next some descriptions in systematic sampling.

1. Suppose that the N units in the population are numbered from 1 to N in some order. To select a sample of n units, we take a unit at random from the first units and every kth subsequent unit. Assume that N is divided by n, we have quotient k and surplus r. Namely, it is

$$N = kn + r.$$

If one unit choosen at random from the first k units is i, it is called random start. So series of sampling units are $i, i+k, \cdots, i+jk, \cdots, i+(n-1)k$. This type of sample is called an every kth systematic sample.

2. The advantages of this method over simple random sampling are as follows.

(i) It is easier to draw a sample and often easier to execute without mistakes. Field work of forest inventory and drawing work of sample in the office are easier.

(ii) Intuitively, systematic sampling seems likely to be more precise than simple random sampling. The reasons why it seems so are as follows. See table 1.

Table 1. Composition of the kth systematic samples

	1	2	·	i	·	k	means
1	y_1 (y_{11})	y_2 (y_{21})	·	y_i (y_{i1})	·	y_k (y_{k1})	$\bar{y}_{\cdot 1}$
2	y_{k+1} (y_{12})	y_{k+2} (y_{22})	·	y_{k+i} (y_{i2})	·	y_{2k} (y_{k2})	$\bar{y}_{\cdot 2}$
·	·	·	·	·	·	·	·
j	$y_{(j-1)k+1}$ (y_{1j})	$y_{(j-1)k+2}$ (y_{2j})	·	$y_{(j-1)k+i}$ (y_{ij})	·	y_{jk} (y_{kj})	$\bar{y}_{\cdot j}$
·	·	·	·	·	·	·	·
n	$y_{(n-1)k+1}$ (y_{1n})	$y_{(n-1)k+2}$ (y_{2n})	·	$y_{(n-1)k+i}$ (y_{in})	·	y_{nk} (y_{kn})	$\bar{y}_{\cdot n}$
means	$\bar{y}_{1\cdot}$	$\bar{y}_{2\cdot}$	·	$\bar{y}_{i\cdot}$	·	$\bar{y}_{k\cdot}$	general mean $=\bar{Y}$

It stratifies the population into n strata, which consists of the first k units, the second k units, and so on. We might therefore expect the systematic sample to be about as precise as the corresponding stratified random sample with one unit per stratum. With systematic sample the units all occur at the same relative position in the stratum.

On the other hand we can consider that systematic sampling is a kind of cluster sampling. A single cluster is sampled at random from k cluster each having n units. It seems homogeneous between clusters, so it will be expected more precise than simple random sampling.

1. mean If $N=kn$ or $n>50$

$$\bar{y}_{sy}=\frac{1}{n}\overset{n}{S}(y) \tag{5.1}$$

If $N \neq kn$, this results does not hold, although the bias is unlikely to be important. The bias can be avoided by alloting a higher probability of selection to certain samples.

2. variance of mean

Variance of mean of systematic sampling is shown by next three types. The symbol y_{ij} denotes the jth member of the ith systematic sample, so that $j=1, 2, \cdots n$, $i=1, 2, \cdots k$. The mean of the ith sample is denoted by $\bar{y}_{i\cdot}$ (Refer to table 1).

(i)

$$V(\bar{y}_{sy}) = \frac{N-1}{N}\sigma^2 - \frac{k(n-1)}{N}\sigma^2{}_{wsy} \tag{5.2}$$

where

$$\sigma^2{}_{wsy}=\frac{1}{k(n-1)}\overset{kn}{SS}\{(y_{ij}-\bar{y}_{i\cdot})^2\}$$

is the variance among units which lie within the same systematic sample.

$$\sigma^2=\frac{\overset{k\,n}{SS}[(y_{ij}-\bar{Y})]^2}{N-1}$$

According to (5.2), the mean of systematic sample is more precise than the mean of simple random sample if and only

$$\sigma^2{}_{wsy} > \sigma^2$$

Systematic sampling is precise when units within the same sample are heterogeneous, and is imprecise when they are homogeneous.

(ii) The mean of the jth sample is denoted by $\bar{y}._j$.

$$V(\bar{y}_{sy})=\frac{\sigma^2}{n}\left\{\frac{N-n}{N}+(n-1)\rho_w\right\} \tag{5.3}$$

where $$\rho_w=\frac{2}{kn(n-1)}\overset{k}{S}\underset{j<u}{\overset{n}{S}}\{(y_{ij}-\bar{Y})(y_{iu}-\bar{Y})\}/\sigma^2$$

is the correlation coefficient between pairs of units that are in the same systematic sample. According to (5.3), it shows that positive correlation between units in the same sample inflates the variance of the sample mean. Even small positive correlation may have a large effect, because of the multiplier $(n-1)$.

(iii)

$$V(\bar{y}_{sy})=\frac{\sigma^2{}_{wst}}{n}\left\{\frac{N-n}{N}+(n-1)\rho_{wst}\right\} \tag{5.4}$$

where

$$\sigma^2{}_{wst}=\frac{1}{n(k-1)}\overset{nk}{SS}\{(y_{ij}-\bar{y}._j)^2\}$$

is the variance among units that lies in the same stratum.

$$\rho_{wst}=\frac{1}{kn(n-1)}\overset{k}{S}\underset{j\neq u}{\overset{n}{S}}\{(y_{ij}-\bar{y}._j)(y_{iu}-\bar{y}._u)\}/\sigma^2{}_{wst}$$

is the correlation between the deviations from the stratum means of pairs of items that are in the same systematic sample.

According to (5.4), if $\rho_{wst}=0$, we have

$$V(\bar{y}_{sy})=\left(\frac{N-n}{N}\right)\frac{\sigma^2{}_{wst}}{n}$$

In other words, a systematic sample has the same precision as the corresponding stratified random sample, with one unit per stratum. Following to these three types of variance of mean, a disadvantage of systematic sampling is that no trustworthy method for $V(\bar{y}_{sy})$ from the sample data is known.

3. confidence limits

The conditions that we can set up confidence limits are

(a) sample means are approximately normally distributed,

(b) there are many number of samples.

But we have only k of possible number of samples in systematic sampling. In spite of this fact, when we use systematic sampling to select the secondary sampling unit in subsampling, we can set up confidence limits, because it is proved that sample means are approximately normally distributed. According to (5.2), (5.3) and (5.4), there are no practical methods to calculate the variance of mean from a sample. But we can consider next few methods.

(i) We calculate its variance of mean similar to the one of simple random sampling.

$$v(\bar{y}_{sy})=\frac{N-n}{N}\frac{v(y)}{n} \tag{5.5}$$

If the fpc is ignored,

$$v(\bar{y}_{sy}) = \frac{v(y)}{n}$$

where

$$v(y) = \frac{\overset{n}{S}[(y-\bar{y}_{sy})^2]}{n-1}$$

Then we can set up the confidence limits

$$N\{\bar{y}_{sy} \pm t\sqrt{v(\bar{y}_{sy})}\}$$

where t is a value of t table at 95% confidence coefficient based on $(n-1)$ degrees of freedom. This case applies when we have confidence that the order is essentially random with respect to the items being measured.

(ii) Estimation methods of the variance from a single sample in systematic sampling are proposed by Cochran in the case of the type of the population which is known. As in this case we must know the population type previously, they are difficult to practical use. It is proposed that instead of a single sample we select a few samples of systematic sample and calculate the variance of mean $v(\bar{y}_{sy})$. Even if a single sample, we divide it into two or more samples and calculate $v(\bar{y}_{sy})$. For instance, assume that start is i and intervals of sampling is k, we divide a systematic sample into next two samples.

$$\begin{array}{lll} i, & i+2k, & i+4k \cdots\cdots\cdots \\ i+k, & i+3k, & i+5k \cdots\cdots\cdots \end{array}$$

Notes:

1. Systematic sampling is important to practical use in the forest inventory. The practical results of forest population is necessary to the development of theory of systematic sampling too.

2. Examples of the natural population of forest used in studies of systematic sampling were as follows. Finney used systematic sampling to estimate the volume of salable timber per strip, 3 chains wide and of varying length (Mt. Stuartforest, $N=160$) in 1948, the volume of virgin timber per strip, 2.5 chains wide 80 chains length (Black Mountain forest, $N=288$) in 1948 and the volume of timber per strip, 2 chain wide and of varying width (Dehra Dun forest, $N=292$) in 1950. Johnson used systematic sampling to estimate the number of seedlings per 1-ft-bed-width in 4 beds of hardwood seedbed stock ($N=400$), 1-ft-bed-width in 3 beds of coniferous seedbed stock ($N=400$) and 1-ft-bed-width in 6 beds of coniferous transplant stock ($N=400$) in 1943. Madow studied the efficiency of systematic and random sampling in one bed of hardwood seedling stock in 1946. The bed was 420 ft long, and the units was 1 ft of the bed width.

The studies of systematic sampling becomes pressing more and more because it has easier and simpler practical use in the forest inventory.

(6) Regression and ratio estimate.

(A1) The linear regression estimate

1. mean $\bar{y}_{lr} = \bar{y} + b(\bar{X} - \bar{x})$ (6.1)

where $\bar{y} = \dfrac{\overset{n}{S}(y)}{n}$, $\bar{x} = \dfrac{\overset{n}{S}(x)}{n}$

$$b=\frac{\overset{n}{S}\{(x-\bar{x})(y-\bar{y})\}}{\overset{n}{S}\{(x-\bar{x})^2\}}$$

$$\bar{X}=\frac{\overset{N}{S}(x)}{N}$$

2. variance

$$v(y_{lr})=\frac{S[(y-\bar{y})^2]-b\overset{n}{S}[(x-\bar{x})(y-\bar{y})]}{n-2} \tag{6.2}$$

3. variance of mean

$$v(\bar{y}_{lr})=\frac{N-n}{N}\,v(y_{lr})\left[\frac{1}{n}+\frac{(\bar{X}-\bar{x})^2}{\overset{n}{S}\{(x-\bar{x})^2\}}\right] \tag{6.3}$$

if the fpc is ignored

$$v(\bar{y}_{lr})=v(y_{lr})\left[\frac{1}{n}+\frac{(\bar{X}-\bar{x})^2}{\overset{n}{S}\{(x-\bar{x})^2\}}\right] \tag{6.4}$$

4. confidence limits

$$N\{\bar{y}_{lr}\pm t\sqrt{v(\bar{y}_{lr})}\} \tag{6.5}$$

where t is value of t table at 95% confidence coefficient based on $(n-2)$ degrees of freedom.

5. n with the desired degree of precision

$$n_0=\left(\frac{tC_v}{E}\right)^2(1-\rho^2) \tag{6.6}$$

where ρ is the correlation coefficient of x and y.
C_v is the coefficient of variation of y.

If n_0/N is neglected, this formula is adequate. In otherwise event we compute n as

$$n=\frac{n_0}{1+n_0/N} \tag{6.7}$$

(A2) The regression estimate in stratified sampling

(a) The separate regression estimate

1. mean

$$\bar{y}_{lrs}=\overset{L}{S}\left(\frac{N_h y_{lrh}}{N}\right)=\overset{L}{S}(W_h\bar{y}_{lrh}) \tag{6.8}$$

where $\bar{y}_{lrh}=\bar{y}_h+b_h(\bar{X}_h-\bar{x}_h)$

$$\bar{y}_h=\frac{\overset{n_h}{S}(y)}{n_h},\quad \bar{x}_h=\frac{\overset{n_h}{S}(x)}{n_h},\quad \bar{X}_h=\frac{\overset{N_h}{S}(x)}{N_h}$$

$$b_h=\frac{\overset{n_h}{S}[(x-\bar{x}_h)(y-\bar{y}_h)]}{\overset{n_h}{S}[(x-\bar{x}_h)^2]} \tag{6.9}$$

2. variance

$$v(y_{hs})=\frac{\overset{n_h}{S}[(y-\bar{y}_h)^2]-b_h\overset{n_h}{S}[(x-\bar{x}_h)(y-\bar{y}_h)]}{n_h-2} \tag{6.10}$$

3. variance of mean

$$v(\bar{y}_{lrs}) = \overset{L}{S}\left[\frac{N_h(N_h-n_h)}{N^2 n_h} v(y_{hs})\right]$$

$$= \overset{L}{S}\left[\frac{N_h-n_h}{N_h} \cdot W_h^2 \frac{v(y_{hs})}{n_h}\right] \tag{6.11}$$

If the fpc is ignored,

$$v(\bar{y}_{lrs}) = \overset{L}{S}\left[W_h^2 \frac{v(y_{hs})}{n_h}\right] \tag{6.12}$$

4. confidence limits

$$N\{\bar{y}_{lrs} \pm t\sqrt{v(\bar{y}_{lrs})}\} \tag{6.13}$$

where t is a value of t table at 95% confidence coefficient based on $(n-2L)$ degrees of freedom.

(b) The combined regression estimate

1. mean

$$\bar{y}_{lrc} = \bar{y}_{st} + b(\bar{X} - \bar{x}_{st}) \tag{6.14}$$

where

$$\bar{y}_{st} = \frac{\overset{L}{S}(N_h \bar{y}_h)}{N} = S(W_h \bar{y}_h),$$

$$\bar{x}_{st} = \frac{\overset{L}{S}(N_h \bar{x}_h)}{N} = \overset{L}{S}(W_h \bar{x}_h)$$

$$\bar{y}_h = \frac{\overset{n_h}{S}(y)}{n_h}, \quad \bar{x}_h = \frac{\overset{n_h}{S}(x)}{n_h},$$

$$\bar{X} = \frac{\overset{L}{S}(N_h \bar{X}_h)}{N} = \frac{\overset{L}{S}\overset{N_h}{S}(x)}{N}$$

$$b = \frac{\overset{L}{S}\overset{n_h}{S}[(x-\bar{x}_h)(y-\bar{y}_h)]}{\overset{L}{S}\overset{n_h}{S}[(x-\bar{x}_h)^2]} \tag{6.15}$$

2. variance

$$v(y_{hc}) = \frac{1}{n_h - 1} \overset{n_h}{S}\{[(y-\bar{y}_h) - b(x-\bar{x}_h)]^2\}$$

$$= \frac{1}{n_h-1}\left[\overset{n_h}{S}\{(y-\bar{y}_h)^2\} + b^2 \overset{n_h}{S}\{(x-\bar{x}_h)^2\} - 2b\overset{n_h}{S}\{(x-\bar{x}_h)(y-\bar{y}_h)\}\right] \tag{6.16}$$

3. variance of mean

$$v(\bar{y}_{lrc}) = \overset{L}{S}\left\{\frac{N_h(N_h-n_h)}{N^2 n_h} v(y_{hc})\right\}$$

$$= \overset{L}{S}\left\{\frac{N_h-n_h}{N_h} W_h^2 \frac{v(y_{hc})}{n_h}\right\} \tag{6.17}$$

If the fpc is ignored,

$$v(\bar{y}_{lrc}) = \overset{L}{S}\left\{W_h^2 \frac{v(y_{hc})}{n_h}\right\} \tag{6.18}$$

4. confidence limits

$$N\{\bar{y}_{lrc} \pm t\sqrt{v(\bar{y}_{lrc})}\} \tag{6.19}$$

value t is value of t table at 95% confidence coefficient based on $(n-L)$ degrees of freedom, but it might seem better $(n-L-1)$.

5. optimum allocation of the sample size n_h in the respective stratified regression estimate. By the analogical methods with stratified random sampling (2.5), we have next results.

(i) optimum allocation with varying costs

$$n_h = n \frac{W_h C_{vh}\sqrt{1-\rho_h^2}\,/\sqrt{C_h}}{\overset{L}{S}\{W_h C_{vh}\sqrt{1-\rho_h^2}/\sqrt{C_h}\}} \tag{6.20}$$

where ρ_h is the correlation coefficient of x and y in each stratum. C_{vh} is the coefficient of variation of y in each stratum.

(ii) optimum allocation with a fixed total size of sample

$$n_h = n \frac{W_h C_{vh}\sqrt{1-\rho_h^2}}{\overset{L}{S}\{W_h C_{vh}\sqrt{1-\rho_h^2}\}} = n \frac{N_h C_{vh}\sqrt{1-\rho_h^2}}{\overset{L}{S}\{N_h C_{vh}\sqrt{1-\rho_h^2}\}} \tag{6.21}$$

(iii) proportional allocation

$$n_h = W_h n \tag{6.22}$$

(B1) The ratio estimate

1. mean

If in (6.1) we assume $\bar{y}=b\bar{x}$, we obtain next simple formula.

$$\bar{y}_R = R\bar{X} \tag{6.23}$$

where

$$R = \frac{\overset{n}{S}(y)}{\overset{n}{S}(x)} \tag{6.24}$$

$$\bar{X} = \frac{\overset{N}{S}(x)}{N}$$

This is a special case of the regression estimate having weight $1/x$.

2. variance

$$v(y_R) = \frac{1}{n-1}\left\{\overset{n_h}{S}(y^2) + R^2 \overset{n_h}{S}(x^2) - 2R\overset{n_h}{S}(yx)\right\} \tag{6.25}$$

3. variance of mean

$$v(\bar{y}_R) = \frac{N-n}{N}\,\frac{v(y_R)}{n} \tag{6.26}$$

If the fpc is ignored,

$$v(\bar{y}_R) = \frac{v(y_R)}{n} \tag{6.27}$$

4. confidence limits

$$N\{\bar{y}_R \pm t\sqrt{v(\bar{y}_R)}\} \tag{6.28}$$

where t is a value of t table at 95% confidence coefficient based in $(n-1)$ degrees of freedom.

5. n with the desired degree of precision

$$n_0 = \frac{C_y^2 + C_x^2 - 2\rho C_x C_y}{C_R^2} \tag{6.29}$$

where C_y is the coefficient of variation of y
C_x is 〃 x
C_R is 〃 R
ρ is the correlation coefficient of x and y.

If n_0/N is ignored, this formula is adequate. In otherwise event, we compute n as

$$n = \frac{n_0}{1 + n_0/N}$$

(B2) The ratio estimate in stratified sampling

(a) The separate ratio estimate

1. mean

$$\bar{y}_{Rs} = \overset{L}{S}\left\{\frac{N_h}{N} \frac{\overset{n_h}{S}(y)}{\overset{n_h}{S}(x)} \bar{X}_h\right\} = \overset{L}{S}(W_h R_h \bar{X}_h) \tag{6.31}$$

where $$R_h = \frac{\overset{n_h}{S}(y)}{\overset{n_h}{S}(x)}, \quad \bar{X}_h = \frac{\overset{N_h}{S}(x)}{N_h}$$

2. variance. $$v(y_{R_s}) = \frac{1}{n_h - 1}\{\overset{n_h}{S}(y^2) + R_h^2 \overset{n_h}{S}(x^2) - 2R_h \overset{n_h}{S}(yx)\} \tag{6.33}$$

3. variance of mean.

$$v(\bar{y}_{Rs}) = \overset{L}{S}\left\{\frac{N_h(N_h - n_h)}{N^2 n_h} v(y_{hs})\right\}$$

$$= \overset{L}{S}\left\{\frac{N_h - n_h}{N_h} W_h^2 \frac{v(y_{hs})}{n_h}\right\} \tag{6.34}$$

If the fpc is ignored,

$$v(\bar{y}_{Rs}) = \overset{L}{S}\left\{W_h^2 \frac{v(y_{hs})}{n_h}\right\} \tag{6.35}$$

4. confidence limits

$$N\{\bar{y}_{Rs} \pm t\sqrt{v(\bar{y}_{Rs})}\} \tag{6.36}$$

where t is a value of t table at 95% confidence coefficient based on $(n-L)$ degrees of freedom.

(b) The combined ratio estimate

1. mean $$\bar{y}_{Rc} = \frac{\bar{y}_{st}}{\bar{x}_{st}} \bar{X} = R\bar{X} \tag{6.37}$$

where $$\bar{y}_{st} = \overset{L}{S}\left(\frac{N_h}{N}\bar{y}_h\right) = \overset{L}{S}(W_h\bar{y}_h), \quad \bar{y}_h = \frac{\overset{n_h}{S}(y)}{n_h}$$

$$\bar{x}_{st} = \overset{L}{S}\left(\frac{N_h}{N}\bar{x}_h\right) = \overset{L}{S}(W_h\bar{x}_h), \quad \bar{x}_h = \frac{\overset{n_h}{S}(x)}{n_h}$$

$$\bar{x} = \overset{L}{S}\left(\frac{N_h}{N}\bar{X}_h\right) = \overset{L}{S}(W_h\bar{X}_h), \quad \bar{X}_h = \frac{\overset{N_h}{S}(x)}{N_h}$$

$$R = \frac{\bar{y}_{st}}{\bar{x}_{st}} = \frac{\overset{L}{S}(W_h\bar{y}_h)}{\overset{L}{S}(W_h\bar{x}_h)} = \frac{\overset{L}{S}(N_h\bar{y}_h)}{\overset{L}{S}(N_h\bar{x}_h)}$$

2. variance $$v(y_{Rc}) = \frac{1}{n_h-1}\left[\overset{n_h}{S}(y^2) + R^2\overset{n_h}{S}(x) - 2R\overset{n_h}{S}(yx)\right] \tag{6.38}$$

3. variance of mean

$$v(\bar{y}_{Rc}) = \overset{L}{S}\left\{\frac{N_h(N_h-n_h)}{N^2 n_h}v(y_{hc})\right\}$$

$$= \overset{L}{S}\left\{\frac{N_h-n_h}{N_h}W_h^2\frac{v(y_{hc})}{n_h}\right\} \tag{6.39}$$

If the fpc is ignored,

$$v(\bar{y}_{Rc}) = \overset{L}{S}\left\{W_h^2\frac{v(y_{hc})}{n_h}\right\} \tag{6.40}$$

4. confidence limits

$$N\{\bar{y}_{Rc} \pm t\sqrt{v(\bar{y}_{Rc})}\} \tag{6.41}$$

where t is a value of t table at 95% confidence coefficient based on $(n-1)$ degrees of freedom.

5. optimum allocation of sample sizes n_h of the respective strata in stratified ratio estimate. From (6.34) and (6.39) the population variance of $\bar{y}_R$ is

$$V(\bar{y}_R) = \overset{L}{S}\left\{\frac{N_h(N_h-n_h)}{N^2 n_h}\sigma_{dh}^2\right\}$$

where σ_{dh}^2 is
$\sigma_{yh}^2 + R_h^2\sigma_{xh}^2 - 2R_h\rho_h\sigma_y\sigma_x$ in separate ratio estimate
$\sigma_{yh}^2 + R^2\sigma_{xh}^2 - 2R\rho_h\sigma_y\sigma_x$ in combined ratio estimate

By the analogical method with stratified random sampling (2—5), we have next results.

(i) optimum allocation with varying costs

$$n_h = n\frac{W_h\sigma_{dh}/\sqrt{c_h}}{\overset{L}{S}\{(W_h\sigma_{dh})/\sqrt{c_h}\}}$$

$$= n\frac{N_h\sigma_{dh}/\sqrt{c_h}}{\overset{L}{S}\{(N_h\sigma_{dh})/\sqrt{c_h}\}} \tag{6.42}$$

This formula may appear a little perplexing in practical use, because it seems difficult to speculate about the likely values of σ_{dh}. It is proposed next three formulae in practical use instead of (6.42).

$$n_h = n \frac{N_h \sqrt{\bar{X}_h} \sqrt{c_h}}{\overset{L}{S}\{(N_h \sqrt{\bar{X}_h} / \sqrt{c_h})\}} \tag{6.43}$$

$$n_h = n \frac{N_h \bar{X}_h / \sqrt{c_h}}{\overset{L}{S}\{(N_h X_h) / \sqrt{c_h}\}} \tag{6.44}$$

$$n_h = n \frac{(N_h \sigma_{yh} \sqrt{1-\rho_h}) / \sqrt{c_h}}{\overset{L}{S}\{(N_h \sigma_{dh} \sqrt{1-\rho_h}) / \sqrt{c_h}\}} \tag{6.45}$$

$$= n \frac{(N_h C_{vh} \sqrt{1-\rho_h}) / \sqrt{c_h}}{\overset{L}{S}\{(N_h C_{vh} \sqrt{1-\rho_h}) / \sqrt{c_h}\}}$$

where C_{vh} is the coefficient of variation of y and

ρ is the correlation coefficient of x and y in each stratum.

(ii) optimum allocation with a fixed total size

The formula proposed are as follows.

$$n_h = n \frac{N_h \sigma_{dh}}{\overset{L}{S}(N_h \sigma_{dh})} \tag{6.46}$$

$$n_h = n \frac{N_h \sqrt{\bar{X}_h}}{\overset{L}{S}(N_h \sqrt{\bar{X}_h})} \tag{6.47}$$

$$n_h = n \frac{N_h X_h}{\overset{L}{S}(N_h \bar{X}_h)} \tag{6.48}$$

$$n_h = n \frac{N_h C_{vh} \sqrt{1-\rho_h}}{\overset{L}{S}(N_h C_{vh} \sqrt{1-\rho_h})} \tag{6.48}$$

(iii) proportional allocation

$$n_h = W_h n \tag{6.50}$$

6. comparison of the ratio estimate with the mean per unit

The variance of mean in random sampling: $V(\bar{Y}) = \frac{N-n}{N} \frac{\sigma_y^2}{n}$

" in ratio estimate : $V(\bar{Y}_R) = \frac{N-n}{N} \frac{1}{n} (\sigma_y^2 + R^2\sigma_x^2 - 2R\rho\sigma_x\sigma_y)$

Hence the ratio estimate has the smaller variance if

$$\sigma_y^2 + R^2\sigma_x^2 - 2R\rho\sigma_x\sigma_y < \sigma_y^2$$

i.e. if

$$\rho > \frac{R\sigma_x}{2\sigma_y} = \frac{1}{2}\left(\frac{\sigma_x}{\bar{X}}\right) \Big/ \left(\frac{\sigma_y}{\bar{Y}}\right) = \frac{1}{2}\left(\frac{C_{vx}}{C_{vy}}\right)$$

This results show that

(i) if C_{vx} is more than twice of C_{vy}, the ratio estimate is always less precise, since ρ cannot exceed 1.

(ii) when x is the value of y at some previous time, the two C_v's may be about equal. In this event the ratio estimate is superior if ρ exceeds 0.5.

Notes :

1. We treated one variate in random, stratified, cluster, sub and systematic samplings, but two variates in ratio and regression estimates. When we have the approximate values by a previous survey and other items having high correlation to one desired items, it is efficient to use ratio and regression estimates.
2. The conditions for which the ratio estimate is the best among a wide class of estimates are as follows.

(i) The relation between y and x is a straight line through the origin.

(ii) The variance of y about this line is proportional to x and then bias become negligible in large samples.

3. In the forest inventory, other items having high correlation to our desired volume are volume values of forest notes investigated previously, ocular estimates and basal area computed by breast height diameter which we can measure easily etc. Ratio and regression methods are to estimate the population volume by the population values of these items ($\bar{X}$) and the relation of actual volume (y) and these items (x) in sample.
4. Conventional sample tree methods are a kind of ratio estimates. Illustrations are as follows.

(i) Sample tree method with a single class is ratio estimate in random sampling.

(ii) Hossfeld's method is combined ratio estimate in which the same number of trees are allocated in each stratum (d.b.h. class).

(iii) Draudt's method is combined ratio estimate in which the number of sample trees are allocated proportional to stratum size.

(iv) Urich II's method is combined ratio estimate in which the same number of trees are approximately involved in each stratum.

(v) Hartig's method is separate ratio method in which basal areas and sample sizes are same in each stratum.

(vi) Block's method is combined ratio estimate in which tree number per hectare are same in each stratum.

(vii) Schwappach-Frick's method is combined ratio estimate in which tree numbers in the first 4 strata are each 100, next 3 strata are each 200 and the remainder are each 400, and sample trees obtained from these strata are measured volume by the form-class method.

(viii) Tischendorf's method is a kind of stratified random sampling in which the number of trees are allocated proportionally to the strata volumes obtained by a temporary volume table and sample trees are measured volumes by the form-class method.

(ix) Neubauer's method is combined ratio estimate in which the allocation of sample trees is as same as Tischendorf's method and we use the relation of actual volume and temporary volume of sample trees.

As in these conventional sample tree methods the selection of sample trees are purposive in order to be sampled assumed average sample trees, it is a defect that we cannot appraise the sampling error. If we want to have appropriate estimate and appraise the sampling error, we must use the sampling methods of ratio and regression estimate related above.

(7) Triple and double sampling

(A) Double sampling for stratification

1. mean

The first sample is a simple random sample of size n'.

Let $W_h = N_h/N$ = proportion of population falling into stratum h.

$w_h = n_h'/n'$ = proportion of first sample falling into stratum h.

The w_h is an estimate of W_h. The second sample is a stratified random sample of size n in which y is measured: n_h unit are drawn from stratum h.

The second sample is often a subsample from the first sample, but it may be drawn independently if this is more convenient. Then we can estimate the population mean from next formula.

$$\bar{y}_{st} = \overset{L}{S}(w_h \bar{y}_h) \tag{7.1}$$

where
$$\bar{y}_h = \frac{\overset{n_h}{S}(y)}{n_h}$$

The w_h and $\bar{y}_h$ are both random variables, subject to error.

2. variance

$$v(y_h) = \frac{1}{n_h - 1} \overset{n_h}{S}[(y - \bar{y}_h)^2] \tag{7.2}$$

3. variance of mean

$$v(\bar{y}_{st}) = \frac{n'}{n'-1} \overset{L}{S}\left\{ \left(w_h^2 - \frac{w_h}{n'}\right) \frac{v(y_h)}{n_h} + \frac{w_h(\bar{y}_h - \bar{y}_{st})^2}{n'} \right\} \tag{7.3}$$

If n' is large relative to the n_h, $v(\bar{y}_{st})$ reduces to

$$v(\bar{y}_{st}) = \overset{L}{S}\left\{ w_h^2 \frac{v(y_h)}{n_h} \right\} \tag{7.4}$$

This expression is equivalent to assuming that errors in the strata weights w_h can be ignored.

4. confidence limits

$$N\{\bar{y}_{st} \pm t\sqrt{v(\bar{y}_{st})}\} \tag{7.5}$$

where t is a value of t table at 95% confidence coefficient based on $(n-L)$ degrees of freedom, where $n = \overset{L}{S}(n_h)$.

5. optimum allocation

Neyman suggests taking n_h proportional to $W_h \sigma_h$. Thus,

$$n_h = n \frac{W_h \sigma_h}{\overset{L}{S}(W_h \sigma_h)} \tag{7.6}$$

The population variance of (7.3) is

$$V(\bar{y}_{st}) = \frac{[\overset{L}{S}(W_h \sigma_h)]^2}{n} + \frac{\overset{L}{S}\{W_h(\mu_h - \mu)^2\}}{n'}$$

$$= \frac{X}{n} + \frac{Y}{n'} \tag{7.7}$$

where the term in $W_h(1 - W_h)$ is ignored.

The cost of the two samples is assumed to be of the form

$$C = nc_n + n'c_{n'} \tag{7.8}$$

where c_n is usually large relative to $c_{n'}$.
From both (7.7) and (7.8) we have easily next results.

$$n = n' \frac{\sqrt{Xc_{n'}}}{\sqrt{Yc_n}} \tag{7.9}$$

This equation and (7.8) determine n and n', then we have n_h from (7.6).

(B) Triple and double sampling with regression estimate

(a) Triple sampling

1. mean

The first sample is a simple random sample of size n.

Let $W_h = \dfrac{N_h}{N}$ = proportion of population falling into stratum h.

$w_h = \dfrac{n_h}{n}$ = proportion of first sample falling into stratum h.

The w_h is an estimate of W_h. The second sample is a stratified random sample of size n_{Lh} drawn from stratum h in which x having high correlation with y is measured. The third sample is a random sample of size n_{sh} drawn from the second sample n_{Lh} of stratum h in which both x and y are measured. Then we can estimate the population mean from next formula.

$$\bar{y}_{ts} = \overset{L}{S}(w_h \bar{y}_{lrh}) \tag{7.10}$$

where

$$\bar{y}_{lrh} = \bar{y}_{sh} + b_h(\bar{x}_{Lh} - \bar{x}_{sh}) \tag{7.11}$$

$$\bar{y}_{sh} = \frac{\overset{n_{sh}}{S}(y)}{n_{sh}}, \quad \bar{x}_s = \frac{\overset{n_{sh}}{S}(x)}{n_{sh}}$$

$$b_h = \frac{\overset{n_{sh}}{S}[(x-\bar{x}_{sh})(y-\bar{y}_{sh})]}{\overset{n_{sh}}{S}\{(x-\bar{x}_{sh})^2\}}$$

$$\bar{x}_{Lh} = \frac{\overset{n_{Lh}}{S}(x)}{n_{Lh}}$$

2. variance

residual variance:

$$v(y_{lrh}) = \frac{1}{n_{sh}-1}\left[\overset{n_{sh}}{S}\{(y-\bar{y}_{sh})^2\} - b_h \overset{n_{sh}}{S}\{(x-\bar{x}_{sh})(y-\bar{y}_{sh})\}\right] \tag{7.12}$$

variance of y:

$$v(y_h) = \frac{1}{n_{sh}-1} \overset{n_{sh}}{S}\{(y-\bar{y}_{sh})^2\} \tag{7.13}$$

variance of w_h:

$$v(w_h) = \frac{w_h(1-w_h)}{n-1} \tag{7.14}$$

3. variance of mean

$$v(\bar{y}_{ts})=\overset{L}{S}\left[w_h^2\left(\frac{N_h-n_{sh}}{N_h}v(y_{lrh})\left\{\frac{1}{n_{sh}}+\frac{(\bar{x}_{Lh}-\bar{x}_{sh})^2}{\overset{n_{sh}}{S}(x-\bar{x}_{sh})^2}\right\}\right.\right.$$

$$\left.\left.+\frac{N_h-n_{Lh}}{N_h}\frac{v(y_h)-v(y_{lrh})}{n_{Lh}}\right)+\bar{y}_{lrh}^2\frac{N_h-n_h}{N_h}v(w_h)\right] \tag{7.15}$$

If the fpc is ignored,

$$v(\bar{y}_{ts})=\overset{L}{S}\left[w_h^2\left(v(y_{lrh})\left\{\frac{1}{n_{sh}}+\frac{(\bar{x}_{Lh}-\bar{x}_{sh})^2}{\overset{n_{sh}}{S}(x-\bar{x}_{sh})^2}\right\}\right.\right.$$

$$\left.\left.+\frac{v(y_h)-v(y_{lrh})}{n_{Lh}}\right)+\bar{y}_{lrh}^2v(w_h)\right] \tag{7.16}$$

4. confidence limits

$$N\{\bar{y}_{ts}\pm t\sqrt{v(\bar{y}_{ts})}\} \tag{7.17}$$

where t is a value of t table at 95% confidence coefficient based on (n_s-2L) degrees of freedom.
where

$$n_s=\overset{L}{S}(n_{sh})$$

If n_h is not so large, bias of the regression estimate arises in each stratum, so we had better use the combined regression estimate in this case.

5. optimum allocation of sample size n_s, n_L and n

In large sample, variance of mean $\bar{y}_{ts}$ is

$$V(\bar{y}_{ts})=\overset{L}{S}\left[W_h^2\left\{\frac{\sigma_{yh}^2(1-\rho_h^2)}{n_{sh}}+\frac{\rho_h^2\,\sigma_{yh}^2}{n_{Lh}}\right\}+\bar{y}_{lrh}^2\cdot\frac{W_h(1-W_h)}{n}\right]$$

$$=\frac{X}{n_s}+\frac{Y}{n_L}+\frac{Z}{n} \tag{7.18}$$

where

$$X=\overset{L}{S}\,[W_h^2\sigma_{yh}^2\,(1-\rho_h^2)]$$

$$Y=\overset{L}{S}\,[W_h^2\,\rho_h^2\,\sigma_{yh}^2]$$

$$Z=\overset{L}{S}[\bar{y}_{lrh}^2\,W_h\,(1-W_h)]$$

ρ_h is the correlation coefficient of y and x in h stratum.
Cost function is expressed by

$$C=c_sn_s+c_Ln_L+c_wn \tag{7.19}$$

We minimize

$$\varphi=V(\bar{y}_{ts})+C$$

$$=\frac{X}{n_s}+\frac{Y}{n_L}+\frac{Z}{n}+(c_sn_s+c_Ln_L+c_wn)$$

Following to the same method as stratified random sampling, we have

$$n_L=n_s\sqrt{\frac{Y}{X}}\sqrt{\frac{c_s}{c_L}} \tag{7.20}$$

$$n = n_s\sqrt{\frac{Z}{X}}\sqrt{\frac{c_s}{c_w}} \tag{7.21}$$

(i) when variance is fixed, we substitute the optimum values of n_L and n in the variance formula (7.18) and solve for n_s. This gives

$$n_s = \frac{\sqrt{X}}{v(\bar{y}_{ts})\sqrt{c_s}}(\sqrt{Xc_s}+\sqrt{Yc_L}+\sqrt{Zc_w}) \tag{7.22}$$

(ii) when cost is fixed, we substitute the optimum values of n_L and n in the cost function (7.19). We obtain

$$n_s = \frac{C}{c_s}\sqrt{\frac{c_s}{X}}\frac{1}{\left(\sqrt{\frac{c_s}{X}}+\sqrt{\frac{c_L}{Y}}+\sqrt{\frac{c_w}{Z}}\right)} \tag{7.23}$$

(b) Double sampling

The formula of double sampling are easily induced by means of putting $w_h=1$, $n=0$, $c_w=0$ in the formulae of triple sampling.

1. mean $\quad \bar{y}_{ds} = \bar{y}_s + b(\bar{x}_L - \bar{x}_s)$ (7.24)

where

$$\bar{y}_s = \frac{\overset{n_s}{S}(y)}{n_s}, \quad \bar{x}_s = \frac{\overset{n_s}{S}(x)}{n_s}$$

$$b = \frac{\overset{n_s}{S}[(x-\bar{x}_s)(y-\bar{y}_s)]}{\overset{n_s}{S}\{(x-\bar{x}_s)^2\}}, \quad \bar{x}_L = \frac{\overset{n_L}{S}(x)}{n_L}$$

2. variance

residual variance in small sample:

$$v(y_{lr}) = \frac{1}{n_s-2}\left\{\overset{n_s}{S}[(y-\bar{y}_s)^2] - b\overset{n_s}{S}[(x-\bar{x}_s)(y-\bar{y}_s)]\right\}$$

variance of y in small sample:

$$v(y) = \frac{1}{n_s-1}\overset{n_s}{S}[(y-\bar{y}_s)^2]$$

3. variance of mean

$$v(\bar{y}_{ds}) = \frac{N-n_s}{N}v(y_{lr})\left\{\frac{1}{n_s}+\frac{(\bar{x}_L-\bar{x}_s)^2}{\overset{n_s}{S}[(x-\bar{x}_s)^2]}\right\} + \frac{N-n_L}{N}\frac{v(y)-v(y_{lr})}{n_L} \tag{7.25}$$

If the fpc is ignored,

$$v(\bar{y}_{ds}) = v(y_{lr})\left\{\frac{1}{n_s}+\frac{(\bar{x}_L-\bar{x}_s)^2}{\overset{n_s}{S}[(x-\bar{x}_s)^2]}\right\} + \frac{v(y)-v(y_{lr})}{n_L} \tag{7.26}$$

4. confidence limits

$$N\{\bar{y}_{ds} \pm t\sqrt{v(\bar{y}_{ds})}\} \tag{7.27}$$

where t is a value of t table at 95% confidence coefficient based on (n_s-2) degrees of freedom.

5. optimum sample size of double sampling

$$n_L = n_s\sqrt{\frac{X}{Y}}\sqrt{\frac{c_s}{c_L}}, \quad n_s = \sqrt{\frac{\rho^2}{1-\rho^2}}\sqrt{\frac{c_s}{c_L}} \tag{7.28}$$

(i) when variance is fixed, we have

$$n_s = \frac{\sqrt{X}}{v\,(\bar{y}_{ds})\sqrt{c}}\{\sqrt{Xc_s} + \sqrt{Yc_L}\} \quad \text{or}$$

$$n_s = \left(\frac{tc_v}{E}\right)^2\left[\,(1-\rho^2)+\rho\sqrt{1-\rho^2}\sqrt{\frac{c_L}{c_s}}\,\right] \tag{7.29}$$

at 95% confidence coefficient.

(ii) when cost is fixed, we have

$$n_s = \frac{C}{c_s}\sqrt{\frac{c_s}{X}}\,\frac{1}{\sqrt{\frac{c_s}{X}}+\sqrt{\frac{c_L}{Y}}} \quad \text{or}$$

$$n_s = \frac{C}{c_s}\,\frac{\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}+\rho\sqrt{\frac{c_L}{c_s}}} \tag{7.30}$$

6. double sampling with regression versus single sampling

When the allocation of samples are optimum, the optimum variance in triple sampling is shown by next formula.

$$V(\bar{y}_{ts})_{\text{opt}} = \frac{\sigma_y^2\sqrt{1-\rho^2}\sqrt{c_s} + \rho\sqrt{c_L} + \frac{1}{C_v}\sqrt{\frac{1-W_h}{W_h}}\sqrt{c_w}}{C}$$

when ρ is taken as positive.

If all resources are devoted to a single sample with no adjustment for regression, this sample is of size $n_s = C/c_s$, we put $c_L = c_w = 0$ in formula (7.19) and the variance of its mean is

$$V(\bar{y}) = \frac{\sigma_y^2}{n_s} = \frac{c_s\sigma_y^2}{C}\;. \tag{7.32}$$

Hence, triple sampling gives a smaller variance if

$$c_s > (\sqrt{1-\rho^2}\sqrt{c_s} + \rho\sqrt{c_L} + \frac{1}{c_v}\sqrt{\frac{1-w_h}{w_h}}\sqrt{c_w}\,)^2$$

i.e.

$$\frac{c_s}{c_L} > \frac{\left(\rho + \frac{1}{c_v}\sqrt{\frac{q}{p}}\sqrt{\frac{c_w}{c_L}}\right)^2}{(1-\sqrt{1-\rho^2}\,)^2} \tag{7.33}$$

In double sampling, put $c_w = 0$, so we have

$$\frac{c_s}{c_L} > \frac{\rho^2}{(1-\sqrt{1-\rho^2}\,)^2} = \frac{(1+\sqrt{1-\rho^2})^2}{\rho^2} \tag{7.34}$$

or

$$\rho^2 > \frac{4c_sc_L}{(c_s + c_L)^2} \tag{7.35}$$

Notes:

1. Ratio and regression estimates require a knowledge of the population mean $\bar{X}$. If it is desired to stratify the population according to the values of the x, their frequency distribution must be known. When such information is lacking, it is sometimes relatively cheap to take a large preliminary sample in which x alone is measured. The purpose of this sample is to furnish a good estimate of $\bar{X}$ or the frequency distribution of x. By the estimated knowledge of the frequency distribution of x, we can stratify the population and estimate efficiently the population mean of Y. And by a good estimate of $\bar{X}$ and the relation of x and y in a small sample, we can estimate efficiently the population mean. These techniques are known as double sampling or two-phase sampling.

2. In the forest inventory, an example of utilization of the frequency distribution of x is as follows. When at first large number of sample trees are selected in order to classify the tree species and the second small number of sample trees drawn from these large sample trees or independently are measured tree volumes, we can estimate efficiently the stand volume by species. Examples of utilization of a good estimate of $\bar{X}$ are as follows. When at first large number of trees are measured the diameter breast height in order to obtain the basal area (x) and then small number of trees drawn from these sample trees are measured the tree volumes (y). We can obtain a good estimate of $\bar{X}$ from the large sample and estimate efficiently the mean volume of stand by regression or ratio methods in the small sample. In the stand increment survey, when we obtain the increment of basal area of large sample trees using the increment borer and the increment of volume of small sample trees drawn from large sample trees using the stem analysis, we can estimate the stand increment by the good estimate in large sample trees and the relation of x and y in small sample trees. It is also proposed that we use plotless sampling or air photointerpretating plot for large sample and plot sampling or ground plot for small sample respectively to estimate the population volume.

3. An example of triple sampling with regression estimate in the forest inventory is as follows. The first samples are a large number of photo points in order to classify the forest type. Secondly the samples of forest point are interpreted on photo-plot for the items having high correlation of volume as tree height, crown closure, tree number etc. The third sample drawn from the photo plots are ground plots and they are measured volumes.

(8) Two occasion sampling

A. The estimation of the population mean on the second occasion.

1. mean

(i) when a sample of the same size is taken on each occasion with partial replacement.

$$\bar{y}_w = \frac{\lambda}{1-\mu^2 r^2}\bar{y}_1 + \frac{(1-\mu r^2)}{1-\mu^2 r^2}\bar{y}_2 \tag{8.1}$$

$$= c_1\{\bar{y}' + b(\bar{x} - \bar{x}')\} + (1-c_1)\bar{y}''$$

where

λ is a fraction of units being retained

μ is 〃 being replaced

then $\lambda + \mu = 1$

r is the correlation coefficient between the unit values on the first and second occasion,

x and y show the sample values on the first and the second occasion respectively. x' and y' are the values belonging to units included in both samples. x'' and y'' show the sample values on the first occasion and second occasion only respectively.

If total sample size is n on each occasion,

$$\bar{y}'=\frac{\overset{\lambda n}{S}(y)}{\lambda n}\ ,\quad \bar{x}'=\frac{\overset{\lambda n}{S}(x)}{\lambda n}\ ,\quad \bar{x}=\frac{\overset{n}{S}(x)}{n}=\lambda\bar{x}'+\mu\bar{x}''$$

$$\bar{y}''=\frac{\overset{\mu n}{S}(y)}{\mu n}\ ,\quad \bar{x}''=\frac{\overset{\mu n}{S}(x)}{\mu n}\ ,\quad \bar{y}=\frac{\overset{n}{S}(y)}{n}=\lambda\bar{y}'+\mu\bar{y}''$$

$$b=\frac{\overset{\lambda n}{S}\{(x'-\bar{x}')(y'-\bar{y}')\}}{\overset{\lambda n}{S}\{(x'-\bar{x}')^2\}}\ ,\quad c_1=\frac{\lambda}{1-\mu^2r^2}$$

(ii) when the numbers in the sample on the two occasions are not the same,

$$\bar{y}_w=\frac{n'\{\bar{y}'+b(\bar{x}-\bar{x}')\}+n''(1-r^2)\bar{y}''}{n'+n''(1-r^2)} \tag{8.2}$$

where n' is the number of units re-sampled on the second occasion,

n'' is the number of new units, and μ is the proportion of units sampled on the first occasion which are not re-sampled on the second occasion. Thus,

$$n'=\lambda n\ ,\qquad n''=\mu n$$

2. variance

$$v(y)=\frac{\overset{n}{S}\{(y-\bar{y})^2\}}{n-1}$$

3. variance of mean

(i) in the case of equal number on the two occasions,

$$v(\bar{y}_w)=\frac{(1-\mu r^2)v(y)}{n\,(1-\mu^2r^2)} \tag{8.3}$$

(ii) in the case of unequal number on the two occasions,

$$v(\bar{y}_w)=\frac{(1-\mu r^2)v(y)}{n'+n''(1-\mu r^2)} \tag{8.4}$$

4. confidence limits

$$N\{\bar{y}_w\pm t\sqrt{v(\bar{y}_w)}\} \tag{8.5}$$

where t is a value of t table at 95% confidence coefficient based on $(n'-2)$ degrees of freedom.

5. optimum value of λ

This is found by minimizing (8.3) with respect to variation in λ. This gives

$$\lambda=\frac{\sqrt{1-r^2}}{1+\sqrt{1-r^2}}\quad\text{or}\quad \mu=\frac{1}{1+\sqrt{1-r^2}} \tag{8.6}$$

when the optimum λ is substituted in (8.3), the minimum variance works out as

$$v(\bar{y}_w)_{opt}=\frac{v(y)}{2n}(1+\sqrt{1-r^2}) \tag{8.7}$$

In random sampling, the variance of mean is

$$v(\bar{y})=\frac{v(y)}{n}$$

Then, relative efficiency shows that

$$\text{R.E.}=\frac{v(\bar{y}_w)_{opt}}{v(\bar{y})}=\frac{1+\sqrt{1-r^2}}{2}$$

If $r=0$, R.E.$=1$, i.e. $v(\bar{y}_w)$ accords with $v(\bar{y})$ and in otherwise event $v(\bar{y}_w)$ is always smaller than $v(\bar{y})$.

B. The estimation of the change

1. mean

There are next three types in the estimation of the change.

(i) $$\bar{i}\,(1)=\frac{\lambda}{1-\mu r}(\bar{x}'-\bar{y}')+\frac{\mu(1-r)}{1-\mu r}(\bar{x}''-\bar{y}'')$$
$$=c_2\,(\bar{x}'-\bar{y}')+(1-c_2)(\bar{x}''-\bar{y}'') \tag{8.9}$$

(ii) $$\bar{i}\,(2)=\bar{x}'-\bar{y}' \tag{8.10}$$

(iii) $$\bar{i}\,(3)=\bar{x}-\bar{y} \tag{8.11}$$

2. variance

$$v(x)=\frac{\overset{n}{S}\{(x-\bar{x})^2\}}{n-1}$$

$$v(y)=\frac{\overset{n}{S}\{(y-\bar{y})^2\}}{n-1}$$

3. variance of mean

(i) $$v\{\bar{i}\,(1)\}=\frac{(1-r)\{v(x)+v(y)\}}{n(1-\mu r)} \tag{8.12}$$

(ii) $$v\{\bar{i}\,(2)\}=\frac{(1-r)\{v(x)+v(y)\}}{\lambda n} \tag{8.13}$$

(iii) $$v\{\bar{i}\,(3)\}=\frac{(1-\lambda r)\{v(x)+v(y)\}}{n} \tag{8.14}$$

4. confidence limits

$$N\{\bar{i}\pm t\sqrt{v(\bar{i}\,)}\,\} \tag{8.15}$$

where i is a value of t table at 95% confidence coefficients based on $2(n-1)$ degrees of freedom.

5. relative efficiencies between these 3 types

The relative efficiency between $\bar{i}\,(1)$ and $\bar{i}\,(2)$ is

$$\text{R.E.}=\frac{v\{\bar{i}\,(1)\}}{v\{\bar{i}\,(2)\}}=\frac{\lambda}{1-\mu r}$$

Fig. 1 shows that the relation of r, λ and R.E.

The estimation of $\bar{i}\,(1)$ is more efficient than $\bar{i}\,(2)$ independently of the value r. When the value of r is high and λ is large, estimates of $\bar{i}\,(1)$ and $\bar{i}\,(2)$ are not so different, so we had better use $\bar{i}\,(2)$ in these cases.

The relative efficiency between $\bar{i}\,(1)$ and $\bar{i}\,(3)$ is

$$\text{R.E.}=\frac{v\{\bar{i}(1)\}}{v\{\bar{i}(3)\}}=\frac{1-r}{(1-\mu r)(1-\lambda r)}$$

Fig. 2 shows the relation of r, λ and R.E.

The estimation of $\bar{i}(1)$ is more efficient than $\bar{i}(3)$ except for $r=0$. When $r=0.99$, $v\{\bar{i}(1)\}$ is only 4% of $v\{\bar{i}(3)\}$ approximately. When λ is very small or large, $v\{\bar{i}(1)\}$ accords with $v\{\bar{i}(3)\}$ approximately, and $\lambda=0$ or 1.0 accords completely.

Fig. 1

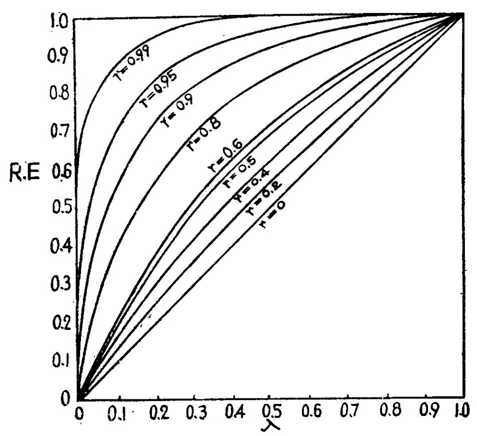
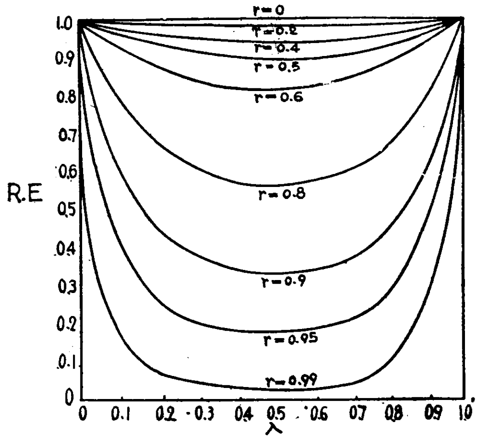


Fig. 2



Notes:

1. We found that the estimation of the population mean on the second occasion by the formula (8.1) or (8.2) is superior to the random sampling and the estimation of change by the formula (8.3) or (8.4) is more efficient. When the correlation of x and y is high, the estimation by the simple formula (8.10) may be adequate.
2. In the forest inventory, we can use these methods for the estimation of the forest resources and its change after the lapse of a fixed period. Assume that at one time plots are sampled to estimate the forest resources, and after a fixed period we must estimate the forest resources and its change. It will be efficient that we measure plots of proportion as the permanent sample plots on both occasions and estimate the forest resources of the second occasion by the formula (8.1) and its change by (8.3) or (8.4).

Chapter 5. Numerical Examples of Plot Sampling

1. Introduction

As we have already mentioned in the previous chapters, the plot sampling is an essential procedure in forestry. We made systematic arrangement of sampling formulae from simple random to two occasion sampling fixed to various inventories. Based on these formulae, we must show calculations which may introduce some practices of application and may introduce simultaneously our forestry numerical variations. Of course, some discussions should be made about the sample plot size. Many reports and discussions concerning the sample plot size have been done, so here details may be abbreviated.

It seems to be preferable that 0.04 ha, or nearly to it may be suitable in Japanese artificial plantation, while 0.06 ha, or about nearly of it, slightly larger relative to the former, may be practical for the natural forest with broad leaved stands. But here, we calculated all computational example based on $10\,m \times 10\,m$ basic unit. Sometimes it might be an ultimate unit and random sampling unit (RSU)——so called, plot may be consisted of four or six basic units.

Now, a sixteen $80\,m \times 80\,m$ block pattern is taken in Shiragadake Cryptomeria stand data of 1950, which is a part of experimental forest published by Kenkichi KINASHI, titled as, " Forest Inventory by sampling Methods " in bulletin of the Kyushu University Forests, No. 23, 1954.

The pattern will be shown as Fig. A.

2. Random sampling

The population is consisted of 1024 plots $10\,m \times 10\,m$.

$$\overset{1024}{S}(y) = 34315$$

Actually 34315 means $34315 \times (1/10) = 3431.5\ m^3$, and we may omit the measurement unit, for simplicity.

Now, we shall select 32 plots at random from the population in order to estimate this population volume.

From the formulae of random sampling mentioned in Chapter 4, we have

1. mean $$\bar{y} = \frac{53+54+\cdots+22}{32} = \frac{1162}{32} = 36.31$$

2. variance $$\begin{aligned} v(y) &= (52992 - 36.3125 \times 1162)/(32-1) \\ &= 10796.8750/31 \\ &= 348.2863 \end{aligned}$$

3. variance of mean

As the fpc is $\frac{N-n}{N} = \frac{1024-32}{1024} = 0.969$, we may use the simple formula (1.4)

$$v(\bar{y}) = \frac{348.2863}{32} = 10.8839$$

Standard error is

$$\sqrt{v(\bar{y})} = 3.299$$

4. confidence limits

$$\begin{aligned} N(\bar{y} \pm t\sqrt{v(\bar{y})}) &= 1024(36.31 \pm 2.040 \times 3.299) \\ &= 37181 \pm 6891 \end{aligned}$$

where t at 95% confidence coefficient based on degrees of freedom 31 is 2.040. Then error percentage of $\bar{y}$ is 18.5%.

5. n with the desired degree of precision

From Fig. A, we have

population mean: $\bar{Y} = \frac{34315}{1024} = 33.51$

Fig. A Shiragadake Cryptomeria Stand (1950)

Column Number

Row Number	3 1	2	3	4	5	6	7	8	4 9	10	11	12	13	14	15	16
3 1	41	20	33	20	36	11	34	40	64	41	39	43	30	17	41	35
2	27	29	33	35	38	25	46	19	53	49	32	43	4	12	36	26
3	55	28	21	44	13	72	44	36	38	30	47	27	7	28	48	34
4	42	20	39	60	41	34	33	31	39	37	41	15	4	48	36	38
5	17	35	30	39	22	37	41	32	38	30	9	9	17	53	34	23
6	24	24	44	46	29	53	34	28	46	18	5	18	20	61	25	30
7	30	12	36	30	31	36	22	27	28	6	13	44	47	39	24	18
8	27	18	41	45	31	31	24	8	6	26	22	45	26	35	23	23
4 9	34	47	39	30	11	12	17	28	54	69	54	42	22	36	23	30
10	29	25	25	17	21	42	49	62	41	42	48	20	26	48	33	21
11	43	47	37	30	42	46	48	48	56	43	52	22	21	34	15	6
12	46	53	57	36	54	40	40	25	39	34	27	16	12	58	10	1
13	52	54	24	34	43	33	40	18	38	20	20	23	22	34	13	4
14	39	29	40	44	43	25	60	32	34	32	33	17	7	29	14	9
15	65	35	45	33	48	14	40	60	30	16	43	30	40	54	27	15
16	34	36	45	47	34	32	39	45	19	36	20	31	36	35	41	15
5 17	34	35	23	25	22	26	27	22	34	23	44	38	44	41	36	17
18	26	25	7	5	30	20	43	24	28	50	31	38	27	41	34	8
19	16	14	10	22	21	23	18	27	54	29	43	43	59	39	37	9
20	16	32	27	35	30	17	33	27	25	33	43	43	55	44	14	16
21	34	48	32	32	42	38	29	26	29	59	35	0	22	13	5	4
22	40	56	28	57	28	35	35	35	41	23	8	24	14	9	10	24
23	39	41	29	35	44	33	42	32	52	58	25	19	8	12	22	58
24	39	36	38	26	34	65	33	22	49	48	20	12	13	22	17	16
6 25	72	52	36	51	41	26	28	62	33	60	3	5	11	21	34	23
26	55	46	63	46	54	27	59	31	29	1	4	21	37	45	41	37
27	50	12	20	24	30	30	16	3	1	12	18	41	41	45	23	34
28	46	24	11	23	48	18	38	41	6	24	31	48	40	41	29	22
29	15	4	8	50	44	21	26	56	19	43	32	44	40	36	24	29
30	6	12	10	49	48	22	39	61	25	49	30	38	49	26	29	14
31	35	34	33	38	30	40	28	37	19	55	25	47	38	27	19	72
32	31	29	33	36	18	38	29	38	37	40	46	43	6	20	46	50

Row Number	5 17	18	19	20	21	22	23	24	6 25	26	27	28	29	30	31	32
3 1	6	13	34	18	32	23	47	39	57	34	28	39	34	39	47	29
2	25	3	5	26	26	25	27	29	25	37	47	41	30	37	40	27
3	11	13	7	21	36	35	32	34	35	29	21	6	28	43	53	40
4	28	14	27	25	41	23	37	37	21	11	12	25	6	14	11	40
5	26	29	34	24	46	25	28	19	5	23	26	25	9	40	41	33
6	25	30	35	43	26	25	29	20	39	34	35	20	14	31	32	20
7	36	25	36	9	21	41	32	34	30	25	28	20	24	27	40	35
8	46	31	3	12	38	34	46	27	43	38	59	29	31	31	40	35
4 9	25	7	2	51	35	32	29	32	29	39	15	45	33	32	33	30
10	2	29	36	27	42	33	23	31	27	40	42	38	42	43	48	35
11	8	66	49	41	36	39	23	34	31	38	29	56	57	30	57	37
12	24	79	43	38	48	47	36	30	29	78	56	56	39	64	36	39
13	22	72	46	32	52	30	22	39	30	45	62	58	70	40	78	54
14	10	40	32	32	51	42	37	28	57	85	35	73	49	69	34	36
15	13	27	47	33	44	37	27	43	65	61	59	90	54	73	66	57
16	8	29	39	51	38	36	38	50	36	73	37	67	58	68	35	32
5 17	6	18	62	48	32	40	36	40	46	46	70	44	49	33	23	27
18	19	33	37	31	36	44	37	46	56	34	61	55	28	31	21	25
19	15	42	22	18	20	30	51	39	56	37	30	21	21	42	43	30
20	20	16	20	21	39	51	35	30	41	47	31	17	19	30	41	28
21	7	36	25	10	38	54	36	43	51	24	42	35	44	30	25	45
22	7	3	3	13	16	26	29	29	17	30	22	41	41	25	36	86
23	30	45	35	26	32	41	26	26	42	27	22	12	45	21	21	7
24	23	24	14	31	33	29	35	11	48	26	35	45	11	12	7	20
6 25	38	35	47	45	28	43	35	34	54	57	22	12	9	6	14	40
26	32	34	31	39	35	50	55	36	85	0	19	10	21	25	21	27
27	21	19	35	51	35	22	48	59	37	54	26	44	33	24	22	30
28	18	10	25	35	36	41	31	11	82	51	40	35	51	57	41	30
29	17	12	17	8	21	14	19	18	66	84	81	30	29	12	14	29
30	35	26	46	38	30	21	34	34	32	58	70	55	71	57	47	70
31	52	22	55	36	35	17	23	25	58	65	34	56	76	60	24	22
32	50	34	52	21	47	35	42	49	44	60	61	91	86	60	48	29

Volume Unit: $0.1m^3$ based on $10\,m \times 10\,m$

population standard deviation: $\sqrt{240.2343}=15.50$

population coefficient of variation: $C_v=\frac{\sigma}{\bar{Y}}\times 100$

$$=\frac{15.50}{33.51}\times 100$$

$$=46.3\%$$

If we may aim at sampling error E as small as 10% in probability 95%, we will obtain the required sample size:

$$n_0=\left(\frac{2\times 46.3}{10}\right)^2=85.7$$

$n_0/N=85.7/1024=0.084>0.05$, then we have next desired n:

$$n=\frac{85.7}{1+85.7/1024}=79.1\doteqdot 80$$

3. Stratified random sampling

This population is quite homogeneous, having coefficient of variation 46%. Usually our artificial plantations are very homogeneous. The variation may be changed, for example, as shown in the following:

Cryptomeria plantation of Shiragadake: C_v = 18% - 44%
Natural pine stand of Kirishima: C_v = 50% - 130%
Natural mixed forests of Takanabe: C_v = 94% - 144%
(Volume based on 10 m$\times$10 m)

There are some variation range from 0 to 90(1/10 m³ unit) in this pattern. From the picture of stand volume pattern, it is recognized that plots having low volume under 10 run toward north-west linearly along very narrow dales and high volume areas over 51 are scattered like spot or cluster. The remains are from 11 to 50, about 80% of total area. Then the assuming stratification will be made.

Stratified population will be shown as the following Fig. B and table 2–1.

Table 2–1

Stratum	Volume range	N_h	$\bar{Y}_h$	σ_h	C_V %	W_h
I	0- 10	69	6.1449	2.72	44.26	0.0673
II	11- 50	825	31.4255	9.76	31.05	0.8057
III	51-100	130	61.2692	9.47	15.45	0.1270
Total		1024				1.0000

$W_h\bar{Y}_h$	c_h	$\frac{W_h\sigma_h}{\sqrt{c_h}}$	$N_h\sigma_h c_h$	$N_h\sigma_h/\sqrt{c_h}$	$W_h\sigma_h$	$W_h\sigma_h^2$
	80	0.0205	1677.86	20.99	0.1833	0.4986
	100	0.7864	80520.00	805.20	7.8636	76.7517
	120	0.1098	13480.54	112.42	1.2027	11.3926
33.5143		0.9167	95678.40	938.61	9.2496	88.6429

c_h=based on 10 m$\times$10 m, Yen

Optimum allocation:

$$\frac{n_{\mathrm{I}}}{n}=\frac{0.0205}{0.9167}=0.022,\qquad \frac{n_{\mathrm{II}}}{n}=\frac{0.7864}{0.9167}=0.858,\qquad \frac{n_{\mathrm{III}}}{n}=\frac{0.1098}{0.9167}=0.120$$

Fig. B Stratified Population

Column Number

Row Number	3 1	2	3	4	5	6	7	8	4 9	10	11	12	13	14	15	16	5 17	18	19	20	21	22	23	24	6 25	26	27	28	29	30	31	32
3 1	8	4	7	4	9	9	5	6	8	9	6	0	10	41	20	33	41	39	43	30	17	41	35	49	29	11	13	21	36	35	32	34
2	6	6	3	5	7	9	3	6	20	36	11	34	40	27	29	33	32	43	12	36	26	38	30	47	28	14	27	25	41	23	37	37
3	6	5	9	6	10	1	4	7	35	38	25	46	19	28	21	44	27	28	48	34	39	37	41	15	26	29	34	24	46	25	28	19
4	9	7	2	2	8	10	8	7	13	44	36	42	20	39	41	34	48	36	38	38	30	17	34	23	25	30	35	43	26	25	29	20
5	5	10	8	9	0	5	4	8	33	31	17	35	30	39	22	37	46	18	18	20	25	30	28	13	36	25	36	21	41	32	34	46
6	9	10	8	6	7	10	7	3	41	32	24	24	44	46	29	34	44	47	39	24	18	26	22	45	31	12	38	34	46	27	34	28
7	3	7	7	3	4	8	6	10	28	30	12	36	30	31	36	22	26	35	23	23	13	34	18	32	39	34	39	47	29	25	37	47
8	3	5	1	4	1	6	6	10	27	27	18	41	45	31	31	24	23	47	39	25	26	26	25	27	41	30	37	40	27	35	29	21
4 9	28	43	40	21	11	12	25	14	40	49	25	24	34	43	33	40	29	14	30	16	43	30	40	27	38	36	38	50	29	39	15	45
10	11	40	23	26	25	40	41	33	18	39	29	40	44	43	25	32	15	19	36	20	31	36	35	41	33	32	33	30	27	40	42	38
11	39	34	35	20	14	31	32	20	35	45	33	48	14	40	34	36	15	25	35	32	29	32	29	36	42	43	48	35	31	38	29	30
12	30	25	28	20	24	27	40	35	45	47	34	32	39	45	42	22	27	42	33	23	31	49	41	36	37	29	39	36	39	30	45	40
13	43	38	29	31	31	40	35	34	36	23	30	41	42	48	20	26	39	23	34	24	43	38	48	47	35	49	34	36	36	37	35	32
14	47	39	30	11	12	17	28	29	48	33	21	43	22	21	34	15	36	30	22	46	32	30	22	39	34	35	23	25	22	26	27	22
15	25	25	17	21	42	49	43	47	39	34	27	16	12	38	20	20	40	32	32	42	37	28	13	27	26	25	30	20	43	24	16	14
16	37	30	42	46	48	48	46	36	23	22	34	13	34	32	33	17	47	33	44	37	27	43	29	39	22	21	23	18	27	16	32	27
5 17	35	30	17	33	27	34	48	32	13	41	23	24	14	24	25	19	23	24	14	31	33	29	35	11	41	26	28	46	46	27	31	50
18	32	42	38	29	26	40	28	28	12	22	49	48	20	12	13	22	46	46	44	49	33	23	27	34	12	20	24	30	30	16	46	24
19	35	35	35	39	41	29	35	44	17	16	18	48	32	40	36	40	28	31	21	25	37	30	21	21	11	23	48	18	38	41	15	50
20	33	42	32	39	36	38	26	34	19	33	37	31	36	44	37	46	42	43	30	41	47	31	17	19	44	21	26	12	49	48	22	39
21	33	22	34	23	44	38	44	41	15	42	22	18	20	30	39	20	30	41	28	24	42	35	44	30	35	34	33	38	40	28	37	31
22	36	17	28	50	31	38	27	41	16	20	21	39	35	30	36	25	25	45	17	30	22	41	41	25	29	33	36	18	38	29	38	33
23	34	29	43	43	39	37	25	33	38	36	43	13	16	26	29	29	36	42	27	22	12	45	21	21	11	21	34	23	29	21	37	45
24	43	43	44	14	16	29	35	22	30	45	35	26	32	41	26	26	48	26	35	45	11	12	20	36	41	37	12	18	41	41	45	23
6 25	34	24	31	48	40	41	29	22	11	17	12	17	21	14	19	18	60	53	64	53	53	61	57	53	59	55	59	52	58	58	62	51
26	19	43	32	44	40	36	24	29	35	26	46	38	30	21	34	34	59	62	53	57	54	52	54	60	51	54	70	56	61	55	56	51
27	25	49	30	38	49	26	29	14	22	36	35	17	23	25	50	34	65	60	54	69	54	56	52	58	86	72	52	51	62	55	63	54
28	19	25	47	38	27	19	37	40	21	47	35	42	49	22	12	14	54	51	66	79	72	52	51	51	59	56	61	60	60	55	72	55
29	46	43	20	46	50	38	35	47	40	19	21	25	21	27	37	26	56	57	57	78	56	56	64	62	51	59	52	55	52	54	57	85
30	45	28	43	35	34	32	34	31	44	33	24	22	30	40	35	41	58	70	78	54	57	85	73	69	54	82	51	51	57	66	84	81
31	39	35	50	36	21	19	35	35	30	30	29	12	14	29	32	47	65	61	59	90	54	73	66	57	58	70	55	71	57	70	58	65
32	22	48	18	25	35	36	41	31	34	24	22	44	48	29	65	72	73	67	58	68	56	57	65	54	56	76	60	60	61	91	86	60

Volume Unit: 0.1 m^3 based on 10 m×10 m

When cost is fixed and if we assume that overhead cost $c_0=0$, we have next sample size from (2.13),

$$n=\frac{938.61}{95678.40}\times C=0.00981\ C$$

where C is total cost. Assuming that C varies from 1,000 to 5,000 Yen we have next n.

Table 2-2

C (Yen)	n (approximately)
1,000	10
2,000	20
3,000	30
3,240	32
5,000	50

In the case of sample size 32, samples proportional to stratum size were selected.

Table 2-3

St.	N_h	n_h	$\bar{y}_h$	$N_h\bar{y}_h$	W_h	$W_h\bar{y}_h$
I	69	2	8.00	552.00	0.06738281	0.5391
II	825	26	28.62	23611.50	0.80576406	23.0581
III	130	4	67.25	8742.50	0.12705313	8.5376
	1024	32		32906.00	1.00000000	32.1348

1. means

$$\bar{y}_{st}=\frac{\overset{L}{S}(N_h\bar{y}_h)}{N}=\frac{32906.00}{1024}=32.135 \text{ or } \bar{y}_{st}=\overset{L}{S}(w_h\bar{y}_h)=32.135$$

2. variance of mean

Table 2-4

$v(y_h)$	$v(\bar{y}_h)$	$W_hv(y_h)$	$N_h^2v(\bar{y}_h)$	$\frac{W_h^2v(y_h)}{n_h}$	$\frac{W_hv(y_h)}{N}$
2.0000	0.9710	0.1348	4622.9310	0.0045	0.0001
148.8062	5.5429	119.8878	3772636.3125	3.7150	0.1171
275.5833	66.7760	34.9962	1128514.4000	1.1104	0.0342
		155.0088	4905773.6435	4.8299	0.1514

$$v(\bar{y}_{st})=\frac{\overset{L}{S}[N_h^2v(\bar{y}_h)]}{N^2}=\frac{4905773.6435}{(1024)^2}=4.6785$$

or

$$v(\bar{y}_{st})=\overset{L}{S}\left[\frac{w_h^2v(y_h)}{n_h}\right]-\overset{L}{S}\left[\frac{w_hv(y_h)}{N}\right]=4.8299-0.1514=4.6785$$

3. confidence limits

$$1024\ (32.134\pm2.045\sqrt{4.6796})$$

$$=32905\pm4529 \qquad df=29$$

Error percentage: 13.76%

4. optimum allocaton

The next pooled standard deviation of strata is obtained from Table 2.1

$$\bar{\sigma}_w=0.1833+7.8636+1.2027=9.2496$$

In this case, coefficient of variation is

$$(C_v)_{st}=\frac{9.2496}{33.5107}=27.6\%$$

The pooled variance of strata is as follows from Table 2.1.

$$\sigma_w{}^2=0.4986+76.7517+11.3926=88.6429$$

In this case, coefficient of variation is

$$(C_v)_{st}=\frac{9.41}{33.5107}=28.0\%$$

From the above results, assume that the coefficient of variation is

$(C_v)_{st}=30\%$ in this population,

Then n with 10% precision is

$$n=\left(\frac{2\times 30}{10}\right)^2=36 \text{ approximately.}$$

The relative precision of this example is as follows:

Variance of random sampling; $V_{\text{ran}}=\frac{240.2343}{32}=7.5073$

Variance of stratified random

sampling by optimum allocation; $V_{\text{opt}}=\frac{(9.2496)^2}{32}=2.6736$

While variance of stratified random sampling by proportional allocation in alternative stratified way, for reference;

$$V_{\text{prop}}=\frac{88.6429}{32}=2.7701$$

From Table 2.1, we have next numerical values.

W_h	σ_h	$\sigma_h-\bar{\sigma}$	$(\sigma_h-\sigma)^2$
0.0674	2.72	−6.53	42.6409
0.8056	9.76	0.51	0.2601
0.1270	9.47	0.22	0.0484

where $\bar{\sigma}=S(W_h\sigma_h)=9.248674=9.25$

From above table, we have

$$\overset{L}{S}\{W_h(\sigma_h-\bar{\sigma})^2\}=3.08968$$

then,

$$\overset{L}{S}\frac{\{W_h(\sigma_h-\bar{\sigma})^2\}}{n}=0.09655$$

$$V_{\text{prop}}-V_{\text{opt}}=2.7701-2.6736=0.0965$$

Difference between V_{prop} and V_{opt} coincides with weighted variance mean. And we have next table from Tabel 2.1.

μ_h	$\mu_h-\mu$	$(\mu_h-\mu)^2$
6.1499	−27.3608	748.6134
31.4255	− 2.0852	4.3481
61.2692	27.7585	770.5343

From the above table, we have

$$\overset{L}{S}[W_h(\mu_h-\mu)^2]=151.8172$$

$$\overset{L}{S}\frac{[W_h(\mu_h-\mu)^2]}{n}=4.7443$$

where

$$\mu=\overset{l}{S}(W_h\mu_h)=33.5107$$

Then,

$$V_{\text{prop}}+\frac{\overset{L}{S}[W_h(\mu_h-\mu)^2]}{n}=2.7701+4.7443=7.5073=V_{\text{ran}}$$

4. Cluster sampling

We shall select 16 units at random each cluster sampling unit having 2 ultimate units, (which is shown by Fig. C) and estimate the population volume. Selected values of units and ultimate units are as follows:

i	j = 1	j = 2	y_i	$\bar{y}_i$
1	28	35	63	31.5
2	43	48	91	45.6
3	35	65	100	50.0
4	48	48	96	48.0
5	9	44	53	26.5
6	17	31	48	24.0
7	44	59	103	51.5
8	40	38	78	39.0
9	26	25	51	25.5
10	32	51	83	41.5
11	54	41	95	47.5
12	36	45	81	40.5
13	37	11	48	24.0
14	35	39	74	37.0
15	28	19	47	23.5
16	66	58	124	62.0
	578	657	1235	

1. mean $\bar{y}_{cl}=\frac{1235}{16}=77.19$

2. variance

$$v(y)=\frac{1}{15}\left(103713-\frac{(1235)^2}{16}\right) \qquad \text{from (3.2)}$$

$$=559.0958$$

3. variance of mean

Fig. C Cluster Sampling Pattern Unit

Column Number

Row Number		3 1	2	3	4	5	6	7	8	4 9	10	11	12	13	14	15	16	5 17	18	19	20	21	22	23	24	6 25	26	27	28	29	30	31	32
3	1	68	49	66	55	74	36	80	59	117	90	71	86	34	29	77	61	31	16	39	44	58	48	74	68	82	71	75	80	64	76	87	56
	2	97	48	60	104	54	106	77	67	77	67	88	42	11	76	84	72	39	27	34	46	77	58	69	71	56	40	33	31	34	57	64	80
	3	41	59	74	85	51	90	75	60	84	48	14	27	37	114	59	53	51	59	69	67	72	50	57	39	44	57	61	45	23	71	73	53
	4	57	30	77	75	62	67	46	35	34	32	35	89	73	74	47	41	82	56	39	21	59	75	78	61	73	63	87	49	55	58	80	70
4	5	63	72	64	47	32	54	66	90	95	111	102	62	48	84	56	51	27	36	38	78	77	65	52	63	56	79	57	83	75	75	81	65
	6	89	100	94	66	96	86	97	73	95	77	79	38	33	92	25	7	32	145	92	79	84	86	59	64	60	116	85	112	96	94	93	76
	7	91	83	64	78	86	58	100	50	72	52	53	40	29	63	27	13	32	112	78	64	103	72	59	67	87	130	97	131	119	109	112	90
	8	99	71	90	80	82	46	79	105	49	52	63	61	76	89	68	30	21	56	86	84	82	73	65	93	101	134	96	157	112	141	101	89
5	9	60	60	30	30	52	46	70	46	62	73	75	76	71	82	70	25	25	51	99	79	68	84	73	86	102	80	131	99	77	64	44	52
	10	32	46	37	57	51	40	51	54	79	62	86	86	114	83	51	25	35	58	42	39	59	81	86	69	97	84	61	38	40	72	84	58
	11	74	104	60	89	70	73	64	61	70	82	43	24	36	22	15	28	14	39	28	23	54	80	65	72	68	54	64	76	85	55	61	131
	12	78	77	67	61	78	98	75	54	101	106	45	31	21	34	39	74	53	69	49	57	65	70	61	37	90	53	57	57	56	33	28	27
6	13	127	98	99	97	95	53	87	93	62	61	7	26	48	66	75	60	70	69	78	84	63	93	90	70	139	57	41	22	30	31	35	67
	14	96	36	31	47	78	48	54	44	7	36	49	89	81	86	52	56	39	29	60	86	71	63	79	70	119	105	66	79	84	81	63	60
	15	21	16	18	99	92	43	65	117	44	92	62	82	89	62	53	43	52	38	63	46	51	35	53	52	98	142	151	85	100	69	61	99
	16	66	63	66	74	78	78	57	75	56	95	71	90	44	47	65	122	102	56	107	57	82	52	65	74	102	125	95	147	162	120	72	51

Volume Unit: 0.1 m^3 based on 10 m×20 m $\overset{512}{S}(y)=34315$, $\overset{512}{S}(y^2)=2655065$

$$v(\bar{y}_{cl})=\frac{512-16}{512}\cdot\frac{559.0958}{16}$$

$$=0.9688\times 34.9435$$

$$=33.8533$$

4. confidence limits

$$512\,(77.19\pm 2.131\times\sqrt{33.8533}\,)$$

$$=39520\pm 6348$$

where $t=2.131$, $df=15$, $\alpha=0.05$

Error percentage is 16.1%.

5. m with the desired degree of precision

In this case population mean $\bar{Y}_{cl}=67.02$. Population standard deviation $=26.366$
Then the coefficient of variation is

$$C_v=\frac{26.366}{67.02}=0.393$$

m_0 with 10% degree of precision at 95% confidence coefficient is calculated as follows:

$$m_0=\left(\frac{2\times 0.393}{0.10}\right)^2=61.9$$

As $m_0/M=\frac{61.9}{512}=0.1209$ is not ignored, we have next desired m

$$m=\frac{61.9}{1+0.1209}=55$$

In this case the coefficient of variation 39% may be noticed, comparing to simple random sampling based on 10 m×10 m, having 46% in coefficient of variation.

5. Sub-sampling

At first we select 16 major plots (20 m×20 m) at random, and then each 2 minor plots (10 m×10 m) are drawn from the selected major plots. The selected values are as follows:

$m=16$ $n=2$

i	j		y_i	$\bar{y}_i$
	1	2		
1	32	28	60	30.0
2	35	34	69	34.5
3	28	33	61	30.5
4	59	16	75	37.5
5	41	35	76	38.0
6	20	23	43	21.5
7	28	23	51	25.5
8	40	48	88	44.0
9	38	9	47	23.5
10	42	22	64	32.0
11	3	3	6	3.0
12	59	11	70	35.0
13	31	31	62	31.0
14	85	45	130	65.0
15	43	28	71	35.5
16	10	33	43	21.5
			1016	

1. mean $\bar{y}_{sub}=\frac{1016}{2\times 16}=31.75$ from (4.1)

2. variance

Between variance; $v(y)_b=\frac{37276-32258}{15}$

$=334.5333$

Within variance; $v(y)_w=\frac{3962}{16(2-1)}$

$=247.6250$

The analysis of variance may be shown as follows:

Source	ss	df	ms
Between	5018	15	334.5333
Within	3962	16	247.6250
Total	8980	31	

3. variance of mean

As the fpc and m/M is not ignored,

$$v(\bar{y}_{sub})=\frac{1}{32}\left(\frac{240}{256}\times 334.5333+\frac{4-2}{4}\cdot\frac{16}{256}\times 247.6250\right)$$

$$=\frac{1}{32}(313.6250+7.7507)$$

$$=\frac{321.3757}{32}$$

$$=10.0430$$

4. confidence limits

$1024\ (31.75\pm 2.131\sqrt{10.0430})$

$=32512\pm 6915$

Error percentage 21.3%

5. optimum sampling

Analysis of variance of subsampling (complete population) is as follows:

Source of variation	ss	df	ms
among primary unit	129181.6320	255	506.5946
among secondary unit	116818.2500	788	152.1071
Total	245999.8820	1024	

$$\sigma_b^2=\frac{506.5946-152.1071}{4}=\frac{354.4875}{4}=88.6219$$

$$V(\bar{y}_{sub})=\frac{256-16}{256}\cdot\frac{88.6210}{16}+\frac{1024-32}{1024}\cdot\frac{152.1071}{32}$$

$$=5.1927+4.6050$$

$$=9.7977$$

Assume that $c_b=5c_w$

so we have next n_{opt}

$$\frac{k}{m}=n_{opt}=\frac{\sqrt{152.1071}}{\sqrt{88.6219}}\sqrt{\frac{50}{10}}=\frac{12.3331}{9.4139}\sqrt{\frac{50}{10}}$$

$$=2.9298$$

$$\doteqdot 3$$

Assume that $C=100c_w$, so we have m from (4.7)

$$100=5\text{m}+3\text{m} \qquad \text{m}=17$$

If, from the previous sample results, we like to know n_{opt} by using formula (4.9)

$$n_{opt}=\frac{\sqrt{247.6250}\ \ \sqrt{2}}{\sqrt{334.5333-247.6250}}\sqrt{5}$$

$$=\frac{\sqrt{2476.250}}{\sqrt{86.9083}}=\sqrt{28.4926756}=5.2$$

$$n_{opt}{}^2=27.04$$

$$n(n+1)=5\times 6=30$$

As $n_{opt}{}^2<n(n+1)$, we choose $n_{opt}=5$

6. Systematic sampling

In the forest inventory, quite often, we are taking systematic sampling. Now we take 32 samples with 32 size, which will be shown in Fig. D.

Sample	mean	standard deviation
1	34.2	18.1
2	37.9	19.3
3	33.6	18.5
4	31.5	15.4
5	32.2	13.9
6	29.6	12.2
7	30.6	13.3
8	31.3	13.0
9	32.4	17.0
10	32.1	17.9
11	31.2	17.4
12	34.4	15.4
13	32.3	14.6
14	34.5	14.0
15	35.5	11.8
16	32.5	16.8
17	36.0	17.2
18	33.3	16.9
19	32.3	13.0
20	33.4	16.6

Sample	mean	standard deviation
21	35.8	15.7
22	36.1	14.4
23	33.2	13.4
24	33.4	16.4
25	33.2	15.7
26	35.8	17.9
27	34.2	14.7
28	36.3	17.2
29	34.5	17.7
30	37.8	14.9
31	32.9	11.1
32	28.6	11.1

Range of means, 28.6–37.9 and the range of standard deviation, 11.1–19.3 and its analysis of variance will be shown as follows:

Source of variation	*df*	*ss*	*ms*
Between strata	31	49511.225	1597.1363
Within strata	992	196488.657	198.0732
Total	1023	245999.882	240.4691

Fig. D Systematic samples from Fig. A population

Strata Number

Sample Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	41	20	33	20	36	11	34	40	27	29	33	35	38	25	46	19	55	28	21	44	13	72	44	36	42	20	39	60	41	34	33	31
2	17	35	30	39	22	37	41	32	24	24	44	46	29	53	34	28	30	12	36	30	31	36	22	27	27	18	41	45	31	31	24	8
3	34	47	39	30	11	12	17	28	29	25	26	17	21	42	49	62	43	47	37	30	42	46	48	48	46	53	57	36	54	40	49	25
4	52	54	24	34	43	33	40	18	39	29	40	44	43	25	60	32	65	35	45	33	48	14	40	60	34	36	45	47	34	32	39	45
5	34	35	23	25	22	26	27	22	26	25	7	5	30	20	43	24	16	14	10	22	21	23	18	27	16	32	27	35	30	17	33	27
6	34	48	32	32	42	38	29	26	40	56	28	57	28	35	35	35	39	41	29	35	44	33	42	32	39	36	38	26	34	65	33	22
7	72	52	36	51	41	26	28	62	55	46	63	46	54	27	59	31	50	12	20	24	30	30	16	3	46	24	11	23	48	18	38	41
8	15	4	8	50	44	21	26	56	6	12	10	49	48	22	39	61	35	34	33	38	60	40	28	37	31	29	33	36	18	38	29	38
9	64	41	39	43	30	17	41	35	53	49	32	43	4	12	36	26	38	30	47	27	7	28	48	34	39	37	41	15	4	48	36	38
10	38	30	9	9	17	53	34	23	46	18	5	18	20	61	25	30	28	6	13	44	47	39	24	18	6	26	22	45	26	35	23	23
11	54	69	54	42	22	36	23	30	41	42	48	20	26	48	33	21	56	43	52	22	21	34	15	6	39	34	27	16	12	58	10	1
12	38	20	20	23	22	34	13	4	34	32	33	17	7	29	14	9	30	16	43	30	40	54	27	15	19	36	20	31	36	35	41	15
13	34	23	44	38	44	41	36	17	28	50	31	38	27	41	34	8	54	29	43	43	59	39	37	9	25	33	43	43	55	44	14	16
14	29	59	35	0	22	13	5	4	41	23	8	24	14	9	10	24	52	58	25	19	8	12	22	58	49	48	20	12	13	22	17	16
15	33	60	3	5	11	21	34	23	29	1	4	21	37	45	41	37	1	12	18	41	41	45	23	34	6	24	31	48	40	41	29	22
16	19	43	32	44	40	36	24	29	25	49	30	38	49	26	29	14	19	55	25	47	38	27	19	72	37	40	46	43	6	20	46	50
17	6	13	34	18	32	23	47	39	25	3	5	26	26	25	27	29	11	13	7	21	36	35	32	34	28	14	27	25	41	23	37	37
18	26	29	34	24	46	25	28	19	25	30	35	43	26	25	29	20	36	25	36	9	21	41	32	34	46	31	3	12	38	34	46	27
19	25	7	2	51	35	32	29	32	2	29	36	27	42	33	23	31	8	66	49	41	36	39	23	34	24	79	43	38	48	47	36	30
20	22	72	46	32	52	30	22	39	10	40	32	32	51	42	37	28	13	27	47	33	44	37	27	43	8	29	39	51	38	36	38	50
21	6	18	62	48	32	40	36	40	19	33	37	31	36	44	37	46	15	42	22	18	20	30	51	39	20	16	20	21	39	51	35	30
22	7	36	25	10	38	54	36	43	7	3	3	13	16	26	29	29	30	45	35	26	32	41	26	26	23	24	14	31	33	29	35	11
23	38	35	47	45	28	43	35	34	32	34	31	39	35	50	55	36	21	19	35	51	35	22	48	59	18	10	25	35	36	41	31	11
24	17	12	17	8	21	14	19	18	35	26	46	38	30	21	34	34	52	22	55	36	35	17	23	25	50	34	52	21	47	35	42	49
25	57	34	28	39	34	39	47	29	25	37	47	41	30	37	40	27	35	29	21	6	28	43	53	40	21	11	12	25	6	14	11	40
26	5	23	26	25	9	40	41	33	39	34	35	20	14	31	32	20	30	25	28	20	24	27	40	35	43	38	59	29	31	31	40	35
27	29	39	15	45	33	32	33	30	27	40	42	38	42	43	48	35	31	38	29	56	57	30	57	37	29	78	56	56	39	64	36	39
28	30	45	62	58	70	40	78	54	57	85	35	73	49	69	34	36	65	61	59	90	54	73	66	57	36	73	37	67	58	68	35	32
29	46	46	70	44	49	33	23	27	56	34	61	55	28	31	21	25	56	37	30	21	21	42	43	30	41	47	31	17	19	30	41	28
30	51	24	42	35	44	30	25	45	17	30	22	41	41	25	36	86	42	27	22	12	45	21	21	7	48	26	35	45	11	12	7	20
31	54	57	22	12	9	6	14	40	85	0	19	10	21	25	21	27	37	54	26	44	33	24	22	30	82	51	40	35	51	57	41	30
32	66	84	81	30	29	12	14	29	32	58	70	55	71	57	47	70	58	65	34	56	76	60	24	22	44	60	61	91	86	60	48	29

Volume Unit: 0.1 m^3 based on 10 m $\times$ 10m

$\overset{32}{S}[\overset{32}{S}(y)]=34315$, $\overset{32}{S}[\overset{32}{S}(y^2)]=1395921$

In this case a strata means a half 80m×80m block, or 80m×40m block. Systematic sampling is taking one plot from each strata, with geometrically systematic in forestry.

Concerning heterogeneity in forest, we can recognize that such systematic sampling may be preferable and it may not have a very harmful difficulty accompanied with some systematic evidence, from the previous table of systematic sample means. If we use the formula for systematic sample variation,

$$V(\bar{y}_{sy})=\frac{\sigma^2}{n}\left\{\frac{N-n}{N}+(n-1)\rho_w\right\}$$

correction term of finite population will be adjusted with adding $(n-1)\ \rho_w$.

Here,

$$\rho_w=\frac{2}{kn(n-1)}\overset{k}{S}\underset{i<\mu}{\overset{n}{S}}\ \{(y_{ij}-\bar{Y})(y_{iu}-\bar{Y})/\sigma^2\}$$

Now we may check if ρ_w is negative or positive by use of this numerical example.

$$\overset{k}{S}\underset{j<\mu}{\overset{n}{S}}\left\{(y_{ij}-\bar{Y})(y_{iu}-\bar{Y})\right\}=\overset{k}{S}\underset{j<u}{\overset{n}{S}}(y_{ij}y_{iu})-\bar{Y}\overset{k}{S}\overset{n}{S}(y_j+y_u)_i+k\frac{(n-1)n}{2}\bar{Y}^2$$

$$=17783778-33.5107\times1063526+32\times49\times(33.51)^2$$

$$=-32734.9010$$

$$\rho_w=\frac{2}{32\times32\times31}\times(-32734.9010)/240.4691$$

$$=-0.0085766$$

$$V(\bar{y}_{sy})=\frac{240.4691}{32}\left\{\frac{1024-32}{1024}+(32-1)(-0.0085766)\right\}$$

$$=7.51466\times(0.9688-0.2660)$$

$$=7.51466\times0.70277$$

$$=5.2811$$

Variances due to different kind sampling will be shown as follows:

$$V_{st}=\overset{L}{S}\left\{\frac{W_h^2\ \sigma_h^2}{n_h}\right\}$$

$=0.0168+2.3781+0.3617=2.7566^{(1)}$ (from Table 2.1)

$$V_{sy}=5.2811^{(2)}$$

$V_{\text{ran}}=\frac{\sigma^2}{n}\frac{N-n}{N}=\frac{240.4691}{32}\times\frac{1024-32}{1024}=7.2799^{(4)}$ (when the fpc is considered)

$$V_{\text{ran}}=\frac{\sigma^2}{n}=\frac{240.4691}{32}=7.5147^{(3)}$$

From these results, it is recognized that $V_{st}{}^{(1)}$ is the smallest variance and $V_{sy}{}^{(2)}$ is a little smaller than $V_{\text{ran}}{}^{(3),(4)}$.

Chapter 6. Recent Developments of Plotless Sampling Estimate

1. Introduction

The angle gauge method of estimating the basal area per unit of a forest stand was first proposed by Dr. Walter Bitterlich, an Austrian forester, in 1948. This method requires no determination of plot boundary, nor any direct measurement of tree stems.

An observer selects in an unbiased manner of points within the stand, at each of which he sights a simple angle gauge at each stem in turn through a complete circle about his position. He counts all trees of which at breast height appear under than the width of the gauge. The estimate of basal area per hectare at each sample point in the stand is given by simply multiplying the total tree count per sweep by a known constant called the basal area factor. Basal area per hectare for stand is estimated as the average value per sweep. This simple estimation method of basal area per unit area is called by next various terms.

Angle count method (Winkelzählprobe=WZP): Bitterlich (1948)
A sample plot without a sample plot
(Probefläche ohne Probefläche): Wanner (1948)
Variable plot radius method: Grosenbaugh (1952)
Variable plot method: Grosenbaugh and Stover (1957)
Point sampling
Tree count method
Plotless sampling
Bitterlich's method

In 1955, Taneo Hirata (Faculty of Agr. Univ. of Tokyo, Japan) proposed the vertical angle count method of estimating the approximate value of the mean height of the forest stand. This method is theoretically as same as Bitterlich's one. He supposed horizontal circles, the centers of which lie at the toes of trees and their radius equal to their heights reduced or enlarged with constant ratio, for example, $1/\sqrt{\pi}$ in reducing. This mean height estimation method is called vertical angle count sampling (by Hirata) or cone method (by Essed). Bitterlich proposed the stand volume estimation method by the combination of Bitterlich's basal area estimation method and Hirata's mean height estimation method, and he called this method $B-H$ method ($B-H$ messung). Nowadays various measuring instruments used for Bitterlich's and Hirata's methods are deviced. We shall introduce briefly the principles of these methods, some instruments and notes in this chapter.

2. The principles of methods

The principles of the methods, which has been described by Finch (1957), are summarized as follows.

(i) Bitterlich's method

Let there be f_i trees per hectare radius r_i measured in meter and n diameter classes. Then the basal area per hectare of the stand will be

$$B=f_1\pi r_1^2+f_2\pi r_2^2+\cdots+f_n\pi r_n^2=\overset{n}{S}(f_i\pi r_i^2) \tag{9.1}$$

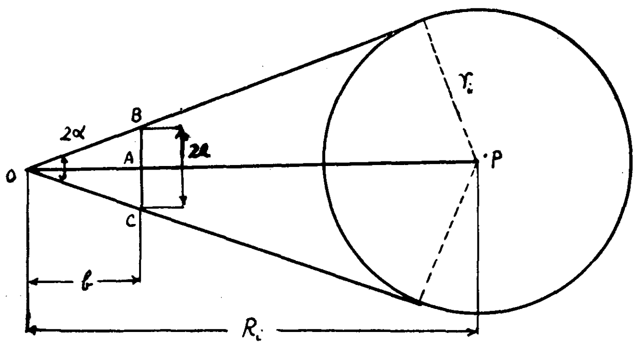


Fig. 3

In Fig. 3. let o be the position of the observer at one of the randomly selected sampling points.

Let 2a be the length of the cross-piece of the instrument and b be the length of the rod, both measured in meter.

Then 2α, the reference angle of the instrument, is given by

$$\tan\alpha = \frac{a}{b} \tag{9.2}$$

Also, if R_i is the maximum distance within which trees of radius r_i are counted

$$\sin\alpha = \frac{r_i}{R_i} \tag{9.3}$$

If z_i trees of radius r_i are counted from one sampling point, the estimate number per hectare of trees of this radius will be

$$f_i = \frac{z_i}{\pi R_i^2} \times (100)^2 \tag{9.4}$$

(since 100^2 m equals one hectare), and the estimated basal area of the trees of this radius will be

$$f_i \pi r_i^2 = 10{,}000 \frac{z_i}{\pi R_i^2} \pi r_i^2 = 10{,}000\ z_i \sin^2\alpha \tag{9.5}$$

If a and b or α of the instrument is constant, i. e.

$$k = 10{,}000 \sin^2\alpha \tag{9.6}$$

We have the total basal area per hectare estimated from one point

$$\begin{aligned} B &= \overset{n}{S}(f_i \pi r_i^2) = \overset{n}{S}(kz_i) \\ &= k\overset{n}{S}(z_i) \\ &= kZ \end{aligned} \tag{9.7}$$

where Z is the total number of trees counted in one sweep. For a number of sampling points, the estimated basal area per hectare of the stand will be k times the mean count. The constant k, which as has shown is determined by the reference angle 2α is known as the multiplication or basal area factor of the instrument. Basal area factors, 1, 2 and 4 are often used in Japanese forest.

(ii) Hirata's method

This procedure is as same as the basal area enumeration sweep. A reference angle, held in a vertical instead of a horizontal plane, is compared with the angles subtended at the sampling point by the total length of the trees surrounding the observer.

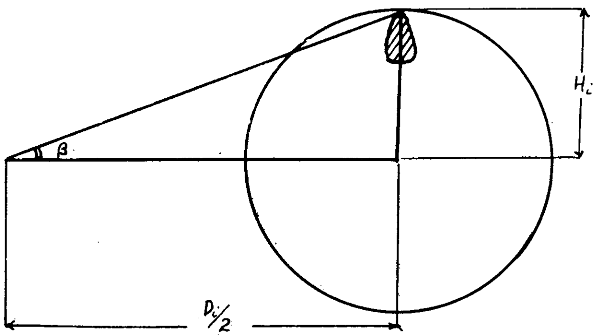


Fig. 4

In Fig. 4, $D_i/2$, is the maximum radius within which all stems of height H_i are counted. β is a reference angle in this case. In the height of each counted trees is imagined to represent the radius of a circle in the vertical plane. Then the area per hectare of all such circles (g_i) is given by the equation:

$$g_i = \frac{z_i \pi H_i^2}{\pi (D_i/2)^2} \times 10{,}000 \quad (\mathrm{m^2/ha})$$

But

$$H_i = (D_i/2) \tan\beta$$

Therefore

$$g_i = 10{,}000 \tan^2\beta \cdot z_i \ (\mathrm{m^2/ha})$$

If z_1, z_2, $\cdots\cdots$, z_n trees with respective heights $H_1, H_2, \cdots\cdots$ H_n, meter are counted, then

$$G = 10{,}000 \tan^2\beta (z_1 + z_2 + \cdots\cdots + z_n)$$
$$= 10{,}000 \tan^2\beta \cdot Z \ (\mathrm{m^2/ha}) \tag{9.8}$$

where G is total circle area per hectare.

G divided by n (the number of stems per hectare) estimates the area of the mean circle.

Therefore

$$\frac{G}{n} = \pi h_m^2 \qquad (\mathrm{m^2})$$

where h_m is mean height.

Substituting equation (9.8) for G, we have

$$h_m = \sqrt{\frac{10{,}000 \tan^2\beta \cdot Z}{n\pi}} \quad (\mathrm{m})$$

$$= \frac{100 \tan\beta}{\sqrt{\pi}} \sqrt{\frac{Z}{n}} \quad (\mathrm{m}) \tag{9.9}$$

If the instrument is constructed to define a limiting angle β of 60°34′, then $\tan\beta = \sqrt{\pi}$ and equation (9.9) can be reduced to its simplest form.

$$h_m = 100\sqrt{\frac{Z}{n}} \tag{9.10}$$

It should be noted that h_m is not the arithmetic mean height but the square root of the harmonic mean height of a stand.

3. Some instruments

1. Simple rod and blade

This is a simple instrument used by Bitterlich in the first place. The length of rod is one meter and the width of metal blade is 2 cm. Then basal area factor is 1.

After Grosenbaugh, it was firstly introduced into Japan by KINASHI in 1952. The size is 50 cm in rod length, 2 cm in metal slit width and the basal area factor is 4.

2. The Relascope

The simple rod and blade was improved by Bitterlich (1948) and automatic correction for slope was provided. This instrument is known as a relascope.

3. The Reflectorscope

This instrument was designed by Crown (1952) in Australia. It combines the principle of the relascope with the optical properties of a reflector gun sight.

4. The mirror Relascope (Das Spiegel Relaskop)

Bitterlich (1952) placed the mirror relascope on the market. This instrument has a much wider applications in the field of forest mensuration.

(a) The estimation of basal area per hectare by the angle count method. Basal area factor are 1/4, 1/2, 1, 2 and 4.

(b) A hypsometer, to measure total or intermediate tree height, in meter.

(c) A dendrometer, to determine the diameter of a stem in at any height above ground level.

(d) A range finder, to establish appropriate horizontal distance in meter.

(e) A clinometer, to measure slopes in degrees or percentages.

5. The tree count tube (Das Baumzählrohr)

The tree count tube, the optical properties of a thin, wedge-shaped prism, was introduced by Muller (1953) in Australia. When the stem of tree is viewed through a prism of this type, the image of the stem is displaced laterally by an amount depending on the distance to the tree and the refractive properties or strength of the prism. When making an enumeration sweep, a tree is counted only when the lateral displacement of the image appears less than the diameter of the stem at breast height.

6. Binocular wedge

Cromer (1954) found that with the tree count tube confusion occurred between images and actual trees, leading to difficulties in counting and considerable eye-strain. He devised an instrument fitted identical wedges to each objective of a low-powered binocular and it can reduce the number of borderline cases to a minimum.

7. Basal area measuring instrument

Senda and Maezawa (1955) in Japan developed a relascope which automatically corrects for slope. The width of the slit is adjustable, providing 1, 2 and 4 of basal area factors.

8. Caliples Bit L. I

Kaibara (1955) in Japan devised the optical instrument. We can see distinctly a stem of tree by this instrument, but no automatic correction for slope provided.

9. Tangent meter

Kurokawa (1959) in Japan devised the instrument of measuring both the mean height and the basal area.

This instrument can be used as:

(a) The direct measurement of a tree height.

(b) The measurement of constant horizontal distance 10, 20 and 30 m.

(c) The indirect measurement of a tree height from arbitrary distances by use of the angle graduation.

(d) The estimation of mean stand height.

(e) The estimation of basal area per hectare.

Then, the catch phrase of this instrument is five in one.

10. Conometer

Hirata (1956) devised the instrument of measuring the mean height of a forest stand and called it conometer. The reference angle is 68°15′ and we count stems whose tops appear above the critical elevation. Mean height (h_m) is given in meter by the equation.

$$h_m = 100\sqrt{\frac{2Z}{n}}$$

11. The new mirror relascope (Neues von Spiegel-Relaskop)

Bitterlich (1955) devised an improved mirror relascope with which we can put a various forest mensuration into practice. The use of this new relascope has pervaded all over the world. With a little practice the instrument can be used efficiently as:

(a) The estimation of basal area per hectare.

(b) The measurement of horizontal distance 15, 20, 25 and 30 m.

(c) The determination of the diameter of a stem in cm at any height above ground level from the above-mentioned constant distance.

(d) The sectional measurement of a tree volume.

(e) The direct measurement of form-height.

(f) The measurement of slopes in degrees or percentages.

12. Prism relascope

Bruce (1955) in America designed an instrument used the properties of wedged prism. To correct for effect of slope with the prismatic instrument, the principal face of the prism must be rotated above the line of sight by an amount corresponding to the angle of inclination. In effect, this lessens the horizontal displacement of the image to an extent which exacts by compensation for the increase in distance due to slope. Prism relascope accomplishes this automatically.

4. Notes

Now, we shall summarize results of the plotless sampling research which has been investigated till now in the various countries.

(1) Disadvantages of the plotless sampling estimate

(a) We cannot count the shaded trees.

(b) The form of the section at breast height are not really circle.

(c) On steepy inclined or broken ground, and especially where undergrowth obscures the base of trees, all but the most experienced observer will fall into error.

(d) It is necessary to obtain the knowledge of the stand area in advance.

(e) The measurement in the edge of forest is possible to arise the bias.

(f) As it is necessary to know the mean height and stand form class in order to estimate the stand volume, we must measure these items directly for some of trees.

(g) In the enumeration sweep, we have a tendency to obtain over or under-estimates according to observers.

(h) It is difficult to sight the position of breast height properly.

(i) As error will often arise in the measurement of slope, we cannot estimate the basal area per hectare properly.

(j) The estimate of mean height by Hirata's method is an approximate value, because it is the square root of the harmonic mean height of a stand.

(k) In Hirata's method, we must obtain the stem number per hectare.

(2) Advantages of the plotless sampling estimate

(a) Grosenbaugh (1952) pointed out that advantages of angle count method are fast and easy.

(b) It is recognized from the results of a number of field experiments that an appropriate estimate of basal area per hectare will be obtained by use of angle count method.

(c) It is expected that, as the counted stem numbers in angle count method are proportional to the size of basal area, v.i.z. plotless sampling is a kind of proportional sampling, an efficient estimate of basal area per hectare in a stand will be obtained.

(d) As the new mirror relascope have many proporties of estimating the basal area per hectare, mean height of stand, height of a tree and form height of a tree, we can estimate the stand volume with the instrument only by plotless sampling.

(e) Plotless sampling is a kind of the sampling technique, so we can obtain the sampling point numbers with the desired degree of precision in advance and evaluate the sampling error.

(f) It is possible to obtain an appropriate and efficient estimate of a stand value by use of double sampling design with ratio or regression, for instance, the combination of tree sampling or plot sampling (small sample) and plotless sampling (large sample).

Chapter 7. Numerical Examples of Simple Regression, used in Plotless and Two Occasion Sampling

1. Basal area and volume regression

The relationship of basal area and volume is well known in forestry from olden time. Bitterlich's method is due to this basis. Now we will calculate some useful factors on the population. Later we will show more complicate design with multiple regression. There is a basal area pattern which corresponds to the Shiragadake volume pattern. According to the stratification, chapter 5, the following numerical example may be shown in the whole population.

St	$\overset{n}{S}[(x-\bar{x})^2]$	$\overset{n}{S}[(x-\bar{x})(y-\bar{y})]$	$\overset{n}{S}[(y-y)^2]$	ρ
I	0.1500	0.8412	5.0252	0.9688
II	13.8912	96.3391	787.1955	0.9213
III	1.6571	15.6223	152.9691	0.9812
Sum	15.6983	112.8026	945.1898	0.9260

Then we can recognize that ρ—basal area and volume—is very high as the usual case. Example: Next 5 plots are selected at random from Fig. E (two variate's population), where x and y show basal area and volume respectively.

Case I

Sample	x	y
1	0.57	3.8
2	0.37	2.1
3	0.54	3.6
4	0.15	0.8
5	0.49	3.8
Sum	2.12	14.1
Mean	0.424	2.82

$\overset{5}{S}[(x-\bar{x})^2]=0.1171,\quad \overset{5}{S}[(y-\bar{y})^2]=7.128$

$\overset{5}{S}[(x-\bar{x})(y-\bar{y})]=0.8906$

Fig. E Shiragadake Cryptomeria Stand (1950)

Column Number

No. 1

Row Number

		1		2		3		4		5		6		7		8	
		x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
3	1	.55	4.1	.30	2.0	.43	3.3	.27	2.0	.50	3.6	.19	1.1	.47	3.4	.48	4.0
	2	.40	2.7	.42	2.9	.50	3.3	.49	3.5	.51	3 8	.30	2.5	.63	4.6	.29	1.9
	3	.63	5.5	.38	2.8	.30	2.1	.56	4.4	.19	1.3	.88	7.2	.63	4.4	.50	3.6
	4	.45	4.2	.27	2.0	.57	3.9	.73	6.0	.48	4.1	.56	3.4	.50	3.3	.41	3.1
	5	.28	1.7	.52	3.5	.52	3.0	.52	3.9	.46	2.2	.61	3.7	.60	4.1	.49	3.2
	6	.39	2.4	.40	2.4	.57	4.4	.64	4.6	.48	2.9	.79	5.3	.55	3.4	.45	2.8
	7	.40	3.0	.17	1.2	.48	3.6	.47	3.0	.48	3.1	.49	3.6	.35	2.2	.43	2.7
	8	.39	2.7	.27	1.8	.60	4.1	.64	4.5	.43	3.1	.39	3.1	.29	2.4	.14	0.8
4	9	.47	3.4	.59	4.7	.53	3.9	.40	3.0	.17	1.1	.20	1.2	.28	1.7	.35	2.8
	10	.50	2.9	.40	2.5	.42	2.5	.27	1.7	.35	2.1	.53	4.2	.61	4.9	.74	6.2
	11	.63	4.3	.66	4.7	.52	3.7	.46	3.0	.55	4.2	.57	4.6	.64	4.8	.62	4.8
	12	.65	4.6	.71	5.3	.76	5.7	.50	3.6	.71	5.4	.53	4.0	.69	4.9	.37	2.5
	13	.72	5.2	.72	5.4	.38	2.4	.50	3.4	.62	4.3	.48	3.3	.61	4.0	.30	1.8
	14	.56	3.9	.41	2.9	.57	4.0	.62	4 4	.64	4.3	.38	2.5	.86	6.0	.48	3.2
	15	.86	6.5	.49	3.5	.60	4.5	.48	3.3	.63	4.8	.20	1.4	.58	4.0	.84	6.0
	16	.46	3.4	.49	3.6	.65	4.5	.70	4.7	.50	3.4	.49	3.2	.60	3.9	.69	4.5
5	17	.43	3.4	.52	3.5	.39	2.3	.41	2.5	.35	2.2	.42	2.3	.45	2.7	.36	2.2
	18	.40	2.6	.41	2.5	.11	0.7	.08	0.5	.49	3.0	.35	2.0	.68	4.3	.36	2.4
	19	.25	1.6	.22	1.4	.16	1.0	.36	2.2	.34	2.1	.36	2.3	.29	1.8	.44	2.7
	20	.70	1.6	.48	3.2	.42	2.7	.55	3.5	.46	3.0	.26	1.7	.49	3.3	.42	2.7
	21	.50	3.4	.67	4.8	.47	3.2	.47	3.2	.54	4.2	.51	3.8	.39	2.9	.35	2.6
	22	.59	4.0	.75	5.6	.49	2.8	.75	5.7	.40	2.8	.51	3.5	.43	3.5	.52	3.5
	23	.57	3.9	.55	4.1	.42	2.9	.52	3.5	.67	4.4	.48	3.3	.51	4.2	.42	3.2
	24	.57	3.9	.50	3.6	.55	3.8	.36	2.6	.54	3.4	.93	6.5	.41	3.3	.30	2.2
6	25	1.05	7.2	.76	3.2	.55	3.6	.65	5.1	.55	4.1	.36	2.6	.39	2.8	.85	6.2
	26	.80	5.5	.64	4.6	.84	6.3	.72	4.6	.86	5.4	.40	2.7	.77	5.9	.44	3.1
	27	.79	5.0	.27	1.2	.32	2.0	.40	2.4	.49	3.0	.44	3.0	.24	1.6	.06	0.3
	28	.74	4.6	.40	2.4	.19	1.1	.27	2.3	.54	4.8	.20	1.8	.43	3.8	.52	4.1
	29	.26	1.5	.07	0.4	.11	0.8	.59	5.0	.51	4.4	.28	2.1	.34	2.6	.68	5.6
	30	.10	0.6	.20	1.2	.17	1.0	.61	4.9	.56	4.8	.29	2.2	.54	3.9	.80	6.1
	31	.54	3.5	.51	3.4	.44	3.3	.49	3.8	.75	6.0	.53	4.0	.36	2.8	.47	3.7
	32	.47	3.1	.44	2.9	.43	3.3	.49	3.6	.25	1.8	.50	3.8	.41	2.9	.52	3.8

		9		10		11		12		13		14		15		16	
		x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
3	1	.80	6.4	.54	4.1	.52	3.9	.64	4.3	.46	3.0	.30	1.7	.53	4.1	.49	3.5
	2	.81	5.3	.66	4.9	.45	3.2	.65	4.3	.07	0.4	.19	1.2	.48	3.6	.43	2.6
	3	.51	3.8	.43	3.0	.65	4.7	.38	2.7	.11	0.7	.40	2.8	.68	4.8	.55	3.4
	4	.57	3.9	.53	3.7	.59	4.1	.25	1.5	.06	0.4	.71	4.8	.54	3.6	.60	3.8
	5	.58	3.8	.49	3.0	.16	0.9	.16	0.9	.28	1.7	.76	5.3	.54	3.4	.38	2.3
	6	.69	4.6	.28	1.8	.09	0.5	.29	1.8	.31	2.0	.91	6.1	.40	2.5	.48	3.0
	7	.44	2.8	.10	0.6	.21	1.3	.59	4.4	.66	4.7	.62	3.9	.41	2.4	.30	1.8
	8	.10	0.6	.42	2.6	.31	2.2	.64	4.5	.40	2.6	.60	3.5	.38	2.3	.38	2.3
4	9	.68	5.4	.91	6.9	.78	5.4	.65	4.2	.35	2.2	.52	3.6	.35	2.3	.42	3.0
	10	.54	4.1	.58	4.2	.67	4.8	.32	2.0	.42	2.6	.71	4.8	.43	3.3	.35	2.1
	11	.76	5.6	.59	4.3	.49	5.2	.38	2.2	.35	2.1	.55	3.4	.22	1.5	.11	0.6
	12	.60	3.9	.54	3.4	.48	2.7	.28	1.6	.20	1.2	.77	5.8	.17	1.0	.01	1.1
	13	.59	3.8	.33	2.0	.33	2.0	.36	2.3	.38	2.2	.54	3.4	.21	1.3	.07	0.4
	14	.52	3.4	.48	3.2	.47	3.3	.24	1.7	.12	0.7	.44	2.9	.22	1.4	.15	0.9
	15	.44	3.0	.24	1.6	.61	4.3	.49	3.0	.56	4.0	.69	5.4	.37	2.7	.25	1.5
	16	.31	1.9	.54	3.6	.32	2.0	.49	3.1	.49	3.6	.48	3.5	.57	4.1	.25	1.5
5	17	.50	3.4	.34	2.3	.65	4.4	.55	3.8	.64	4.4	.58	4.1	.52	3.6	.28	1.7
	18	.39	2.8	.72	5.0	.47	3.1	.57	3.8	.39	2.7	.59	4.1	.58	3.4	.15	0.8
	19	.76	5.4	.41	2.9	.63	4.3	.62	4.3	.84	5.9	.54	3.9	.47	3.7	.15	0.9
	20	.38	2.5	.47	3.3	.59	4.3	.63	4.3	.79	5.5	.72	4.4	.22	4.1	.26	1.6
	21	.42	2.9	.83	5.9	.50	3.5	.00	0.0	.38	2.2	.21	1.3	.08	0.5	.08	0.4
	22	.63	4.1	.28	2.3	.15	0.8	.40	2.4	.24	1.4	.14	0.9	.15	1.0	.39	2.4
	23	.74	5.2	.68	5.8	.34	2.5	.31	1.9	.14	0.8	.21	1.2	.34	2.2	.87	5.8
	24	.62	4.9	.51	4.8	.26	2.0	.19	1.2	.21	1.3	.34	2.2	.26	1.7	.24	1.6
6	25	.47	3.3	.59	6.0	.05	0.3	.08	0.5	.19	1.1	.31	2.1	.51	2.4	.34	2.3
	26	.45	2.9	.03	0.1	.05	0.4	.32	2.1	.49	3.7	.60	4.5	.53	4.1	.46	3.7
	27	.01	0.1	.18	1.2	.27	1.8	.58	4.1	.53	4.1	.60	4.5	.35	2.3	.47	3.4
	28	.10	0.6	.35	2.4	.41	3.1	.60	4.8	.50	4.0	.52	4.1	.41	2.9	.32	2.2
	29	.27	1.9	.59	4.3	.44	3.2	.60	4.4	.54	4.0	.50	3.6	.37	2.4	.44	2.9
	30	.34	2.5	.69	4.9	.45	3.0	.57	3.8	.70	4.9	.41	2.6	.46	2.9	.22	1.4
	31	.26	1.9	.85	5.5	.36	2.5	.66	4.7	.56	3.8	.30	2.7	.18	1.9	.79	7.2
	32	.49	3.7	.55	4.0	.61	4.6	.59	4.3	.09	0.6	.31	2.0	.54	4.6	.60	5.0

Fig. E Shiragadake Cryptomeria Stand (1950)

Column Number

No. 2

Row Number	17		18		19		20		21		22		23		24		25		26		27		28		29		30		31		32	
	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
3 1	.10	0.6	.21	1.3	.52	3.4	.31	1.8	.54	3.2	.39	2.3	.66	4.7	.54	3.9	.83	5.7	.51	3.4	.42	2.8	.60	3.9	.52	3.4	.55	3.9	.63	4.7	.41	2.9
2	.55	2.5	.06	0.3	.10	0.5	.41	2.6	.46	2.6	.43	2.5	.44	2.7	.49	2.9	.37	2.5	.54	3.7	.65	4.7	.65	4.1	.47	3.0	.55	3.7	.57	4.0	.40	2.7
3	.18	1.1	.20	1.3	.15	0.7	.46	2.1	.62	3.6	.54	3.5	.53	3.2	.52	3.4	.48	3.5	.41	2.9	.28	2.1	.07	0.6	.38	2.8	.60	4.3	.76	5.3	.54	4.0
4	.46	2.8	.24	1.4	.48	2.7	.45	2.5	.64	4.1	.33	2.3	.57	3.7	.58	3.7	.32	2.1	.25	1.1	.16	1.2	.33	2.5	.09	0.6	.22	1.4	.16	1.1	.52	4.0
5	.46	2.6	.51	2.9	.55	3.4	.39	2.4	.76	4.6	.26	2.5	.38	2.8	.22	1.9	.08	0.5	.36	2.3	.38	2.6	.37	2.5	.13	0.9	.54	4.0	.56	4.1	.47	3.3
6	.42	2.5	.47	3.0	.49	3.5	.61	4.3	.32	2.6	.33	2.5	.40	2.9	.30	2.0	.53	3.9	.51	3.4	.55	3.5	.31	2.0	.22	1.4	.45	3.1	.46	3.2	.28	2.0
7	.55	3.6	.33	2.5	.35	3.6	.15	0.9	.34	2.1	.57	4.1	.50	3.2	.52	3.4	.50	3.0	.39	2.5	.44	2.8	.31	2.0	.36	2.4	.40	2.7	.60	4.0	.54	3.5
8	.62	4.6	.44	3.1	.06	0.3	.18	1.2	.52	3.8	.50	3.4	.73	4.6	.44	2.7	.67	4.3	.70	3.8	.83	5.9	.45	2.9	.48	3.1	.49	3.1	.63	4.0	.53	3.5
4 9	.53	2.5	.11	0.7	.03	0.2	.64	5.1	.50	3.5	.45	3.2	.43	2.9	.49	3.2	.44	2.9	.56	3.9	.24	1.5	.65	4.5	.51	3.3	.50	3.2	.52	3.3	.45	3.0
10	.04	0.2	.33	2.9	.43	3.6	.38	2.7	.60	4.2	.49	3.3	.34	2.3	.44	3.1	.37	2.7	.59	4.0	.55	4.2	.51	3.8	.60	4.2	.61	4.3	.71	4.8	.47	3.5
11	.13	0.8	.82	6.6	.64	4.9	.58	4.1	.53	3.6	.58	3.9	.34	2.3	.48	3.4	.39	3.1	.49	3.8	.37	2.9	.71	5.6	.69	5.7	.41	3.0	.75	5.7	.49	3.7
12	.29	2.4	1.01	7.9	.60	4.3	.57	3.8	.68	4.8	.65	4.7	.53	3.6	.40	3.0	.40	2.9	.91	7.8	.68	5.6	.62	5.6	.47	3.9	.79	6.4	.48	3.6	.52	3.9
13	.33	2.2	.94	7.2	.66	4.6	.47	3.2	.75	5.2	.42	3.0	.31	2.2	.51	3.9	.45	3.0	.54	4.5	.74	6.2	.61	5.8	.82	7.0	.51	4.0	.99	7.8	.69	5.4
14	.17	1.0	.60	4.0	.46	3.2	.46	3.2	.69	5.1	.57	4.2	.50	3.7	.39	2.8	.77	5.7	.97	8.5	.40	3.5	.82	7.3	.53	4.9	.80	6.9	.44	3.4	.46	3.6
15	.20	1.3	.40	2.7	.62	4.7	.44	3.3	.56	4.4	.49	3.7	.38	2.7	.55	4.3	.79	6.5	.65	6.1	.64	5.9	1.01	9.0	.64	5.4	.85	7.3	.81	6.6	.72	5.7
16	.12	0.8	.43	2.9	.52	3.9	.68	5.1	.52	3.8	.49	3.6	.52	3.8	.63	5.0	.45	3.6	.81	7.3	.42	3.7	.77	6.7	.66	5.8	.82	6.8	.47	3.5	.43	3.2
5 17	.09	0.6	.27	1.8	.86	6.2	.66	4.8	.43	3.2	.53	4.0	.57	3.6	.55	4.0	.57	4.6	.54	4.6	.80	7.0	.53	4.4	.60	4.9	.42	3.3	.32	2.3	.37	2.7
18	.31	1.9	.47	3.3	.52	3.7	.43	3.1	.50	3.6	.57	4.4	.51	3.7	.57	4.6	.66	5.6	.41	3.4	.74	6.1	.70	5.5	.37	2.8	.40	3.1	.27	2.1	.33	2.5
19	.23	1.5	.59	4.2	.33	2.2	.24	1.8	.30	2.0	.40	3.0	.68	5.1	.48	3.9	.69	5.6	.47	3.7	.41	3.0	.30	2.1	.28	2.1	.56	4.2	.62	4.3	.40	3.0
20	.32	2.0	.26	1.6	.29	2.0	.31	2.1	.59	3.9	.66	5.1	.59	3.5	.39	3.0	.47	4.1	.61	4.7	.43	3.1	.23	1.7	.30	1.9	.38	3.0	.50	4.1	.37	2.8
21	.10	0.7	.46	3.6	.39	2.5	.16	1.0	.49	3.8	.76	5.4	.55	3.6	.58	4.3	.68	5.1	.31	2.4	.56	4.2	.49	3.5	.57	4.4	.36	3.0	.32	2.5	.56	4.5
22	.11	0.7	.04	0.3	.05	0.3	.21	1.3	.23	1.6	.39	2.6	.40	2.9	.41	2.9	.25	1.7	.42	3.0	.32	2.2	.54	4.1	.52	4.1	.36	2.5	.44	3.6	1.01	8.6
23	.44	3.0	.65	4.5	.51	3.5	.39	2.6	.49	3.2	.55	4.1	.34	2.6	.36	2.6	.54	4.2	.36	2.7	.31	2.2	.16	1.2	.55	4.5	.24	2.1	.27	2.1	.10	0.7
24	.34	2.3	.35	2.4	.19	1.4	.46	3.1	.45	3.3	.40	2.9	.47	3.5	.16	1.1	.62	4.8	.33	2.6	.44	3.5	.62	4.5	.18	1.1	.18	1.2	.13	0.7	.33	2.0
6 25	.56	3.8	.47	3.5	.59	4.7	.61	4.5	.41	2.8	.55	4.3	.47	3.5	.44	3.4	.64	5.4	.61	5.7	.27	2.2	.21	1.2	.18	0.9	.11	0.6	.25	1.4	.62	4.0
26	.45	3.2	.48	3.4	.43	3.1	.54	3.9	.49	3.5	.67	5.0	.66	5.5	.43	3.6	.91	8.5	.00	0.0	.30	1.9	.19	1.0	.37	2.1	.45	2.5	.32	2.1	.41	2.7
27	.42	2.1	.28	1.9	.52	3.5	.72	5.1	.48	3.5	.29	2.2	.58	4.8	.67	5.9	.43	3.7	.67	5.4	.43	2.6	.73	4.4	.48	3.3	.39	2.4	.37	2.2	.48	3.0
28	.27	1.8	.16	1.0	.36	2.5	.52	3.5	.44	3.6	.49	4.1	.42	3.1	.17	1.1	.98	8.2	.66	5.1	.65	4.0	.48	3.5	.77	5.1	.76	5.7	.59	4.1	.50	3.0
29	.21	1.7	.16	1.2	.25	1.7	.15	0.8	.28	2.1	.16	1.4	.37	1.9	.26	1.8	.80	6.6	.99	8.4	1.00	8.1	.38	3.0	.29	2.9	.22	1.2	.23	1.4	.47	2.9
30	.42	3.5	.32	2.6	.63	4.6	.48	3.8	.37	3.0	.37	2.1	.54	3.4	.43	3.4	.42	3.2	.73	5.8	.84	7.0	.67	5.5	.94	7.1	.82	5.7	.69	4.7	.95	7.0
31	.62	5.2	.29	2.2	.77	5.5	.54	3.6	.56	3.5	.26	1.7	.36	2.3	.36	2.5	.78	5.8	.87	6.5	.45	3.4	.74	5.6	1.05	7.6	.84	6.0	.33	2.4	.32	2.2
32	.62	5.0	.49	3.4	.72	5.2	.34	2.1	.67	4.7	.48	3.5	.59	4.2	.66	4.9	.57	4.4	.79	6.0	.81	6.1	1.18	9.1	1.16	8.6	.88	6.0	.73	4.8	.43	2.9

(Corresponding to Fig. A)

x=Basal area (m^2), y=Volume (m^3)
Each based on 10 m×10 m

$$\rho = \frac{0.8906}{\sqrt{0.1171 \times 7.128}} = 0.9753$$

$$b = \frac{0.8906}{0.1171} = 7.6054$$

$$\begin{aligned} \bar{y}_{lr} &= 2.82 + 7.61(0.47 - 0.42) \\ &= 2.82 + 0.38 \\ &= 3.20 \end{aligned}$$

$$\begin{aligned} v(y_{lr}) &= \frac{7.128 - 7.61(0.891)}{5-2} \\ &= 0.1156 \end{aligned}$$

$$\begin{aligned} v(\bar{y}_{lr}) &= \left(\frac{1024-5}{1024}\right) \times (0.1156) \times \left\{\frac{1}{5} + \frac{(0.47-0.42)^2}{0.1171}\right\} \\ &= 0.02545 \end{aligned}$$

$$\sqrt{0.02545} = 0.1595$$

We may estimate the population volume by use of simple regression.

$$\begin{aligned} &1024\ (3.20 \pm 3.18 \times 0.16) \\ &= 3277 \pm 521 \end{aligned}$$

Then, error percentage: 15.9%

Now we may calculate n with the 10% precision.

We know that $C_v = 0.4625$, $\rho = 0.926$

Then we have next n from (6.6)

$$\begin{aligned} n_0 &= \left(\frac{2 \times 0.4625}{0.10}\right)^2 \{1 - (0.926)^2\} \\ &= (85.5625)(0.1425) \\ &= 12.1926 \end{aligned}$$

As $n_0/N = 12/1026$ is ignored. We have $n = 12$.

Sample size with desired degree of precision may be 12, enough. If we take $n = 12$, the following plots appeared.

Case II

Sample	x	y	
1	0.42	2.1	
2	0.25	1.8	
3	0.15	0.9	
4	0.49	2.9	
5	0.28	2.1	
6	0.25	1.1	
7	0.40	3.0	
8	0.55	4.0	
9	0.31	2.1	
10	0.47	3.2	
11	0.56	3.8	
12	0.55	3.7	
Sum	4.68	30.7	
Mean	0.39	2.5583	$\rho = 0.9474$

$$\bar{y}_{lr} = 2.56 + 7.03(0.47 - 0.39)$$
$$= 3.12$$
$$v(y_{lr}) = 0.1228$$
$$v(\bar{y}_{lr}) = 0.01376$$

Confidence limit : 3195 ± 268

Error percentage : 8.4%

Simple regression between basal area B and volume V.

(1) Without weight:

When we decide a and b so as to minimize

$$\overset{n}{S}\{[V - a - bB]^2\}$$
$$V = -0.36167 + 7.9252B$$
$$v(V) = 0.37[0.0715 + 0.0395 - 0.0918]$$
$$= 0.0071$$

(2) With weight:

$$\overset{n}{S}\left\{\frac{1}{B^2}(V - a - bB)^2\right\}$$
$$V = -0.0141 + 6.8631B, \quad \frac{V}{B} = 6.8631 - 0.0141\left(\frac{1}{B}\right)$$
$$v(V) = 1.08 \ [0.0001 + 0.0048 - 0.0007]$$
$$= 0.0046$$

If we use weight, large residual due to large tree and small residual due to small tree may be harmonized, consequently the regression may balance well. Usually the variance having weight will be less than the case of without weight. In Bitterlich's case, if we use weight $1/B^2$, the equation will be $V/B = b + a(1/B)$, where $V/B = R$. Then $R = b + a(1/B)$ shows that the ratio may be proportional to the reciprocal of B. In this mean, the equation $V = a + bB$ gives automatically the volume according to the volume-basal ratio with weight. This procedure may be resonable and from the older time we are using the ratio in this type, without thinking of it.

2. An example of two occasion sampling

The tree heights of Sugi (*Cryptomeria*) stand in 1954 and 1959 were measured for thirty stems on each occasion in 11th compartment of Ōdawara district forest office. The twenty stems of thirty which were measured in 1954 were remeasured and ten stems were sampled at random newly in 1959. The data on each occasion were as follows.

Sample tree Number	1	2	3	4	5	6	7	8	9	10
1954 (x'')(m)	24.8	28.8	26.7	27.5	23.2	24.5	31.6	24.5	26.4	26.5
1959	—	—	—	—	—	—	—	—	—	—

Sample tree Number	11	12	13	14	15	16	17	18	19	20
1954 (x')(m)	24.2	25.0	21.0	27.2	24.8	19.5	24.2	25.7	26.8	27.2
1959 (y')(m)	25.8	27.1	22.2	28.5	26.3	20.5	25.7	27.6	28.4	28.8

Sample tree Number	21	22	23	24	25	26	27	28	29	30
1954(x')(m)	25.9	23 7	24.8	26.9	20.8	22.0	25.6	23.2	21.0	22.9
1959(y')(m)	27.8	25.2	26.9	29.0	22.4	23.2	27.5	25.1	22.2	24.7

Sample tree Number	31	32	33	34	35	36	37	38	39	40
1954	—	—	—	—	—	—	—	—	—	—
1959(y'')(m)	28.3	22.8	21.5	19.2	18.8	19.4	26.2	23.2	27.9	26.7

We find

$$S(x')=482.4 \qquad S(y')=514.9$$

$$S(x'')=264.5 \qquad S(y'')=234.0$$

$$\bar{x}'=24.120 \qquad \bar{y}'=25.745 \qquad \bar{y}'-\bar{x}'=1.625$$

$$\bar{x}''=26.450 \qquad \bar{y}''=23.400 \qquad \bar{y}''-\bar{x}''=-3.050$$

$$\bar{x}=24.897 \qquad \bar{y}=24.963 \qquad \bar{y}-\bar{x}=0.066$$

$$S[(x'-\bar{x}')^2]=102.292$$

$$S[(x'-\bar{x}')(y'-\bar{y}')]=110.552$$

$$S[(y'-\bar{y}')^2]=120.909$$

Then we have

$$r=\frac{110.552}{\sqrt{102.292\times 120.909}}=0.994$$

$$b=\frac{110.552}{102.292}=1.08705$$

we also have

$$\lambda=\frac{2}{3} \qquad \mu=\frac{1}{3} \qquad \text{and thus}$$

$$c_1=\frac{\lambda}{1-\mu^2 r^2}=0.74888 \qquad \frac{\mu(1-\mu r^2)}{1-\mu^2 r^2}=1-c_1=0.25112$$

$$c_2=\frac{\lambda}{1-\mu r}=0.99701 \qquad \frac{\mu(1-r)}{1-\mu r}=1-c_2=0.00299$$

Then we have the next estimate of mean height of 1959 in this stand from (8.1)

$$\bar{y}_w=0.74888\{25.745+1.08705(24.897-24.120)\}+0.25112\times 23.400=25.78 \quad (\text{m})$$

Since $\qquad v(y)=\dfrac{279.370}{29}=9.633448 \qquad$ and

$$v(x)=\frac{192.590}{29}=6.641034$$

where $\qquad S[(y-\bar{y})^2]=279.370$ and $S[(x-\bar{x})^2]=192.590$

so $v(x)$ and $v(y)$ may reasonably be taken as equal. The pooled estimate, based on all the observations on each occasion (58 degrees of freedom) is $\dfrac{279.370+192.590}{58}=8.137241$

We then have from (8.3)

$$v(\bar{y}_w)=\frac{(1-1/3\times 0.994^2)\times 8.137241}{30(1-1/9\times 0.994^2)}=0.165701$$

then $\sqrt{v(\bar{y}_w)} = 0.407$

As t based on $2(n-1)=2(30-1)=58$ degrees of freedom equals about 2, the 95% confidence limit of mean height in 1959 is $25.78\pm2\times0.407=25.78\pm0.81$ (m), and from (8.5) and error percentage is $\dfrac{0.81}{25.78}\times100=3.1\%$.

Next, we may estimate the increment of mean height (change) from (8.9), (8.10) and (8.11).

From (8.9)

$$\begin{aligned}\bar{i}(1)&=c_2(\bar{y}'-\bar{x}')+(1-c_2)(\bar{y}''-\bar{x}'')\\&=0.99701\times1.625+0.00299\times(-3.050)\\&=1.61\ (\text{m})\end{aligned}$$

Using the pooled estimate $v(y) = 8.137241$, we have the variance of $\bar{i}(1)$ from (8.12)

$$v\{\bar{i}(1)\}=\frac{(1-0.994)\times2\times8.137241}{30(1-1/3\times0.994)}=0.004866$$

then $\sqrt{v\{\bar{i}(1)\}}=0.0698$

So we have next 95% confidence limit

$1.61\pm2\times0.0698=1.61\pm0.140$ (m) (error percentage is 8.7%)

From (8.10)

$$\bar{i}(2)=\bar{y}'-\bar{x}'=1.63$$

As same as the above calculation, we have the variance of $\bar{i}(2)$ from (8.13)

$$v\{\bar{i}(2)\}=\frac{(1-0.994)\times2\times8.137241}{2/3\times30}=0.004881$$

then $\sqrt{v\{\bar{i}(2)\}}=0.0699$

The 95% confidence limit is

$1.63\pm2\times0.0699=1.63\pm0.140$ (m) (error percentage is 8.6%)

From (8.11)

$$\bar{i}(3)=\bar{y}-\bar{x}=0.07$$

The variance of $\bar{i}(3)$ is obtained from (8.14)

$$v\{\bar{i}(3)\}=\frac{(1-2/3\times0.994)\times2\times8.137241}{30}=0.164638$$

Then

$$\sqrt{v\{\bar{i}(3)\}}=0.406$$

The 95% confidence limit is

$0.07\pm2\times0.406=0.07\pm0.812$ (m) (error percentage is 1160%)

As we pointed out in Chapter 4[(8) 5] (relative efficiencies between these 3 types), the relative efficiency between $\bar{i}(1)$ and $\bar{i}(2)$ is about equal in this case for λ is large and r is high value, so we had better use the simple estimate $\bar{i}(2)$ and $\bar{i}(3)$ is worest as we expected.

$r=0.994$, so we have next optimum λ from (8.6)

$$\lambda=\frac{\sqrt{1-0.994^2}}{1+\sqrt{1-0.994^2}}=0.10$$

Then it is efficient that only $n'=\lambda n=0.10\times 30=3$ is retained and $n''=\mu n=0.90\times 30=27$ is replaced in this case.

Chapter 8. Sampling Design based on Aerial Photograph

1. Two important aspects or aerial photograph in forestry

No other things is superior to aerial photograph to estimate extensive timber area. It has two great characteristics, one of which is the identification item and the other is the quantitative item. The identification item has already a wide application for making forest map, establishment of boundary of stand compartment and many other forestry applications. Japanese forest agency is now taking pictures for National forests, about 12 million ha, every year. Average expenses, 1952–1959, are about ¥51(¢14) per ha, for the mapping from photograph and about ¥36 (¢10) per ha, for taking new aerial photographs. Usually photo scale using in Japan is about 1/20,000~25,000 and enlarged photo scale is about 1/8,000. However, the second item, quantitative item, should be more necessary for inventory in forest survey. It is not sufficient to measure various quantitative item, in forestry with the photograph having scale 1/8,000. We hope that the photograph scale may be improved to about 1/5,000 or larger scale than 1/5,000. Consequently, concerning sampling design, two main procedures might be considered.

(A) Photograph is used for identified stratification only and field plot cruising is done within each stratum.

(B) Photograph is used for both identified stratification and rough estimation by interpretational measurement of quantitative item and the field plot is made as ground control or ground check with more accurate field measurement. Before showing the two main sampling procedures, we should analyse the data which came from interpretation and aerial photograph volume, especially these errors.

2. Error of aerial photo volume table and interpretation

We like to know how quantitative item measurements are precise. This is one examination which was done in Kirishima National Forests, *Cryptomeria* stand, 1960. Using photograph is enlarged scale in 1/8,000. (Kirishima Mt.-107 C2-5, 6. 58-10-28 JFTA Pacific)

Aerial volume equation or volume table was made on 30 plots (10 m×10 m) and all measurements were made on the ground. The number of trees, height of tree, and the crown diameter of tree were measured only as visible one by air photograph. It is quite difficult to correspond each tree on photo plot to one on field plot. So, here, the visible is not strict.

2.1. The error of aerial photo volume table

30 field data may give the following normals:

	1	N	$\overline{CD}$	$\overline{H}$	Y
1	30.000	216.000	86.470	615.700	163.748
N		1704.000	601.760	4392.500	1236.623
$\overline{CD}$			256.068	1784.876	469.980
$\overline{H}$				12705.070	3378.908
Y					950.838

In the following chapters, strict distinction between V and v will be neglected.

Analysis of variance due to each factor may be shown as follows:

Sources	ss	df	ms	F
Const	893.780	1		**
N	22.325	1		**
$\overline{CD}$	9.402	1		**
$\overline{H}$	13.785	1		**
$\overline{CD}\times\overline{H}$	0.580	1		not sig.
Errors	10.966	25	0.438	
Total	950.838	30		

The resulting volume equation is shown as in the following:

$$Y=-12.5607+0.6604N+0.9541\overline{CD}+0.5123\overline{H}$$

where Y: Timber volume in m^3

N: Number of trees based on $10\,m\times 10\,m$ plot

$\overline{CD}$: Average visible crown diameter in m

$\overline{H}$: Average visible height in m

$ms = 11.546/26 = 0.444$

standard error $\sqrt{0.444/30}=0.121$ error percentage $=2.21\%$

Variance of estimated volume Y is divided into two parts, i.e. regression variance and measurement variance in photo interpretation, expressing the following way,

$$V(Y)=s^2R+S_I^2$$

$$s^2R=s_{0.1234}[c_{11}+c_{22}N^2+c_{33}\overline{CD}^2+c_{44}\overline{H}^2 +2\{c_{12}N+c_{13}\overline{CD}+c_{14}\overline{H}+c_{23}N\cdot\overline{CD}+c_{24}N\cdot\overline{H}+c_{34}\overline{CD}\cdot\overline{H}\}]$$

c_{11}			=	7.15098
$c_{22}N^2$	=	0.01195 $(6.1071)^2$	=	0.44570
$c_{33}(\overline{CD})^2$	=	0.28083 $(2.786)^2$	=	2.17975
$c_{44}(\overline{H})^2$	=	0.01904 $(20.03)^2$	=	7.63887
Sum				17.41530
$c_{12}N$	=	0.22491 (6.1071)	=	1.37355
$c_{13}\overline{CD}$	=	(0.59542) (2.786)	=	−1.65884
$c_{14}\overline{H}$	=	(−0.34209) (20.03)	=	−6.85206
$c_{23}N.\overline{CD}$	=	(0.3328)(6.1071)(2.786)	=	0.56624
$c_{24}N.\overline{H}$	=	(0.00299)(6.1071)(20.03)	=	0.25566
$c_{34}\overline{CD}.\overline{H}$	=	(−0.02210)(20.03)(2.786)	=	−1.23326
Sum				−7.54871
		2×(−7.54871)	=	−15.09742
Total				2.31788
$s^2_{0.1234}$	=	0.4444		

s^2R = (0.4444)(2.31788) = 1.02913

$\sqrt{s^2R} = \sqrt{1.02913} = 1.014$ error percentage $= 1.014/5.458 = 0.186$

This error is rather large. It is the error per single plot. The resulting aerial photograph volume table may be shown as follows:

Volume per ha in m^3

Average Visible Height(m)	Average visible crown diameter in m				
	2	3	4	5	6
	number of trees 600/ha				
16	151	246	341	437	532
18	253	349	444	539	633
20	356	451	546	642	737
22	458	553	649	744	840
	number of trees 800/ha				
16	283	378	474	569	664
18	385	481	576	671	767
20	488	583	679	774	869
22	590	686	781	876	972
	number of trees 1000/ha				
16	415	560	606	701	796
18	517	663	708	804	899
20	620	765	811	906	1001
22	722	868	913	1008	1104
	number of trees 1200/ha				
16	547	642	738	833	929
18	649	745	840	936	1031
20	751	887	943	1038	1133
22	854	950	1045	1141	1236

2.2. Error of interpretation

On the same stand 7 plots were drawn systematically with 80 meters on a line, and 4 persons interpreted the number of trees, crown diameter and height of tree. Each variance and covariance may be calculated from within mean squares in analysis variances. Analysis of variance will be shown for between persons and within persons. There is only mean square.

Mean Square and Variance ratio

	df	N	CD	H	Y
Between	6	39090.7	0.653	15.344	19131.58
Within	21	13382.7	0.232	6.985	30267.60
Total	27				
F		2.92*	2.81*	2.19	0.63

$F^6_{21}(0.05)=2.57^*$, $F^6_{21}(0.01)=3.81^{**}$

This case shows significant differences in 0.05 level at number of trees and crown diameter interpretations. But height and volume are not significantly different. Similarly covariances may be analysed into between and within, and finally, variance and covariance are shown as follows:

	N	$\overline{CD}$	$\overline{H}$
N	13,382.7	122.8	−80.05
$\overline{CD}$		0.232	− 0.1157
$\overline{H}$			6.985

Then, error due to measurement in interpretation may be calculated as follows:

$$S_I^2=(0.6604)^2[V(N)]+(0.9541)^2[V(\overline{CD})]+(0.5123)^2[V(\overline{H})]$$
$$+2(0.6604)(0.9541)\,[Cov\,(N)\,(\overline{CD})]+2(0.6604)(0.5123)\,[Cov(N)(\overline{H})]$$
$$+2(0.9541)(0.5123)\,[Cov(\overline{CD})(\overline{H})]$$
$$=0.58366+0.21191+1.83301+2(0.41580-0.27083-0.05655)$$
$$=2.62858+2\times 0.08842$$
$$=2.79542$$

Interpretation variance in volume is about 2.5 times as compared with one of the regression of volume. It is very large. This point should be improved near future.

2.3. Total variance of volume equation

Then, total variance

$$V(Y)=s^2R+S_I^2$$
$$=1.02913+2.79543$$
$$=3.82456$$

Its square root: $\sqrt{3.82456}=\pm 1.96$

Error percentage to average volume$=1.96/5.458=0.359$

About 36% in error percentage is quite large.

The volume estimation per single plot has quite large error. Consequently we should use regression method, again. In this meaning, volume equation may be prefer in more simple type, for instance, one or two independent variables enough. Interpretation will be fast.

3. Regression between interpretation volume and ground check volume

There are 18 data including X interpretation value and Y ground checking volume.

Plot	X	Y
No. 1	650	471
〃	622	471
〃	534	471
No. 2	651	794
〃	765	794
No. 3	783	824
No. 4	489	487
〃	605	487
〃	383	487
〃	489	487
No. 5	424	644
〃	558	644
No. 6	357	413
〃	331	413
〃	335	413
〃	357	413
No. 7	569	544
〃	541	544
	9443	9801

$$b=0.7161$$

Regression equation $Y=544.5+0.7161(X-524.6)$

$$\rho=0.7348$$

Sometimes very high correlation coefficient, will be appeared,

$$\rho=0.959$$

Example of double samping case

$$\overline{X}_L=478.925$$
$$Y=596.57+1.0219(479-575)$$
$$=498.4676$$
$$V(\overline{X}_L)=179.8568$$
$$V(Y)=354.22+(479-575)^2\times 0.0179+(1.0219)^2\times 179.86$$
$$=707.0142 \qquad \sqrt{707.0142}=26.58$$

Volume estimate 498.47 ± 26.58 , Error percentage 5.33%

Consequently we can consider double sampling and triple sampling with aerial photograph in forestry.

If $C_v=0.5$ $E=0.10$ $t=2$

when $\rho=0.73$

$$n_0=(2\times 0.50/0.10)^2(1-0.73^2)=(1.00/0.1)^2(1-0.5329)$$
$$=100\times 0.4671=46.71$$

when $\rho=0.96$

$$n_0=100\times(1-0.9592)=100\times 0.0784=7.84$$
$$n_0/N\to 0$$

Consequently if $\rho=0.73$ and 0.96, we have $n=47$ and $n=8$, respectively.

4. Double sampling for stratification by aerial photograph Manual of Photographic interpretation (American society of photogrammetry 1960) is showing the following design of sampling inventory in forestry.

(1) The number of field plots in each volume class stratum, from (2.26) or (2.29)

$$n=\frac{t^2C_v^2}{E^2+\dfrac{t^2C_v^2}{N}}$$

(2) The total volume of any class or stratum

$$N_hy_h=(A_h/a)\bar{y}_h$$

where A_h=area in ha of the hth stratum

a=area of a field plot

From (2.2), we have the variance of $N_h\bar{y}_h$.

$$V(N_h\bar{y}_h)=\left(\frac{A_h}{a}\right)^2\frac{V(y_h)}{n_h}\left(1-\frac{n_ha}{A_h}\right)$$

(3) Aggregate volume of all "L" strata

$$\text{Aggregate}=\overset{L}{S}\left[\left(\frac{A_h}{a}\right)\cdot\bar{y}_h\right]$$

(4) The mean volume per plot through all-over strata

$$\bar{y}_{st}=\frac{\overset{L}{S}\left[\left(\dfrac{A_h}{a}\right)\bar{y}_h\right]}{\overset{L}{S}\left(\dfrac{A_h}{a}\right)}=\overset{I}{S}\left(\frac{\dfrac{A_h}{a}}{\overset{L}{S}\left(\dfrac{A_h}{a}\right)}\bar{y}_h\right)=\overset{L}{S}\left[\left(\frac{N_h}{N}\bar{y}_h\right)\right]=\overset{L}{S}(W_h\bar{y}_h)$$

where $W_h=N_h/N$ is the proportion of the area in the h class stratum.

(5) The variance of $\bar{y}_{st}$

$$V(\bar{y}_{st})=\overset{L}{S}\left[W_h{}^2\frac{V(y_h)}{n_h}\left(1-\frac{n_h}{N_h}\right)\right]$$

$$=\overset{L}{S}\left\{\frac{W_h{}^2V(y_h)}{n_h}\right\}-\overset{L}{S}\left\{\frac{W_hV(y_h)}{N}\right\}$$

This is the same formula as (2.5) shown already. If the fpc is ignored,

$$V(\bar{y}_{st})=\overset{L}{S}\left[\frac{W_h{}^2V(y_h)}{n_h}\right]$$

(6) If a is different in each stratum

$$\bar{y}_{st}=\frac{\overset{L}{S}\left[\frac{A_h}{a_h}\bar{y}_h\right]}{\overset{L}{S}\left(\frac{A_h}{a_h}\right)}=\frac{\overset{L}{S}(q_h\bar{y}_h)}{\overset{L}{S}(q_h)}$$

where $$\frac{A_h}{a_h}=q_h$$

$$V(\bar{y}_{st})=\frac{1}{[\overset{L}{S}(q_h)]^2}\overset{L}{S}[q_h{}^2V(\bar{y}_h)]$$

(7) If w_h is obtained by classifying n' sample photo plots by volume classes, $w_h=n_h'/n'$ is an estimate $W_h=N_h/N$, where n' is the size of the first sample and n_h' is fall in h stratum.

(8) The second sample with size n is taken independently from the n' size first sample, then double sampling may be performed.

$$\bar{y}_{st}=\overset{L}{S}(w_h\bar{y}_h) \quad \text{where} \quad w_h=\frac{n_h'}{n'}, \quad \bar{y}_h=\frac{\overset{n_h}{S}(y)}{n_h}$$

$$V(\bar{y}_{st})=\frac{n'}{n'-1}\overset{L}{S}\left\{\left(w_h{}^2-\frac{w_h}{n'}\right)\frac{V(y_h)}{n_h}+\frac{w_h(\bar{y}_h-\bar{y}_{st})^2}{n'}\right\}$$

where $$V(y_h)=\frac{1}{n_h-1}\overset{n_h}{S}[(y-\bar{y}_h)^2]$$

or for more practical use, the above formula will be available

$$V(\bar{y}_{st})=\overset{L}{S}\left\{\frac{w_h{}^2V(y_h)}{n_h}\right\}+\frac{1}{n'}\{\overset{L}{S}(w_h\bar{y}_h{}^2)-(\overset{L}{S}w_h\bar{y}_h)^2\}$$

n' is large relative to the n_h.

(9) Optimum allocation

$$n_h=\frac{W_h\sqrt{V(y_h)}}{\overset{L}{S}(W_h\sqrt{V(y_h)})}n$$

$$n_h=\frac{\frac{W_h\sqrt{V(y_h)}}{\sqrt{C_h}}}{\overset{L}{S}\left\{\frac{W_h\sqrt{V(y_h)}}{\sqrt{C_h}}\right\}}n$$

(10) The allocation is proportional to area

If the total number of samples is n, then

$$n_h=W_h(n)$$

$$V(\bar{y}_{st})=\overset{L}{S}\left\{\frac{W_hV(y_h)}{n}\right\}-\overset{L}{S}\left\{\frac{W_h^2V(y_h)}{N_h}\right\}$$

or

$$V(\bar{y}_{st})=\overset{L}{S}\left\{\frac{W_hV(y_h)}{n}\right\}-\overset{L}{S}\left\{\frac{W_hV(y_h)}{N}\right\}$$

where $$W_h=\frac{N_h}{N}$$

(11) Optimum allocation variance

$$n_h=\frac{W_h\sqrt{V(y_h)}}{\overset{L}{S}\{W_h\sqrt{V(y_h)}\}}n$$

$$V(\bar{y}_{st})_{\text{opt}}=\frac{(\overset{L}{S}W_h\sqrt{V(y_h)})^2}{n}-\frac{\overset{L}{S}(W_hV(y_h))}{N}$$

$$V(\bar{y}_{st})_{\text{opt}}=\frac{[\overset{L}{S}(W_h\sqrt{V(y_h)}\sqrt{C_h})]\left[\overset{L}{S}\left(\frac{W_h\sqrt{V(y_h)}}{\sqrt{C_h}}\right)\right]}{n}-\frac{\overset{L}{S}[W_hV(y_h)]}{N}$$

Plots are assigned to give minimum costs for a given error. C_h is the cost of a plot in h stratum.

(12) A more logical procedure is to select photo and field plots so that the error $V(\bar{y}_{st})_{\text{opt}}$ will be as small as possible for a given expenditure.

$$C=nC_n+n'C_{n'}$$

where

C_n is cost of a field plot

C_n' is cost of a photo plot

(13) If C_n for all strata and the cost of photo plot $C_{n'}$ is assumed to be constant.

$$V(\bar{y}_{st})=\frac{\{\overset{L}{S}(W_h\sigma_h)\}^2}{n}-\frac{\overset{L}{S}\{W_h(\mu_h-\mu)^2\}}{n'}$$

$$=\frac{X}{n}+\frac{Y}{n'}$$

$$C=nC_n+n'C_{n'} \tag{7.8}$$

$$n'=n\sqrt{YC_n}/\sqrt{XC_{n'}} \tag{7.9$'$}$$

(7.9)′ comes from minimum condition resulting from the combination of both (7.7) and (7.8) and then it is inserted into (7.7) and solve with regard to n,

$$n=(X\sqrt{C_n}+\sqrt{YXC_{n'}})/V(\bar{y}_{st})\sqrt{C_n}$$

$$n=\frac{(\overset{L}{S}(W_h\sigma_h))^2\sqrt{C_n}+\{\overset{L}{S}(W_h\sigma_h)\}\sqrt{\overset{L}{S}(W_h\mu_h^2)-\{\overset{L}{S}(W_h\mu_h)\}^2}\sqrt{C_{n'}}}{V(\bar{y}_{st})\sqrt{C_n}}$$

$$n'=\frac{\sqrt{\overset{L}{S}(W_h\mu_h^2)-\{\overset{L}{S}(W_h\mu_h)\}^2}\sqrt{C_n}}{\overset{L}{S}(W_h\sigma_h)\sqrt{C_{n'}}}n$$

5. Triple sampling by aerial photograph

(1) The first sample is a simple random sample of size n. It is photo plot.

$$w_h = \frac{n_h}{n} \quad \text{estimates} \quad W_h = \frac{N_h}{N}$$

2nd sample is a stratified random sample of size n_{Lh} drawn from stratum h in which plots interpretation should be made.

3rd sample is a random sample of size n_{sh} drawn from the second sample n_{Lh} of stratum h in which ground tree measurement should be done.

(2) mean

$$\bar{y}_{ts} = \overset{L}{S}\{w_h \bar{y}_{lrh}\}$$

where $\qquad \bar{y}_{lrh} = \bar{y}_{sh} + b_h(\bar{x}_{Lh} - \bar{x}_{sh})$

(3) variance of mean

$$V(\bar{y}_{ts}) = \overset{L}{S}\Bigg[w_h^2 \Big\{ \frac{N_h - n_{sh}}{N_h} V(y_{lrh}) \Big(\frac{1}{n_{sh}} + \frac{(\bar{x}_{Lh} - \bar{x}_{sh})^2}{\overset{n_s}{S}\{(x - \bar{x}_{sh})^2\}} \Big)$$

$$+ \Big(\frac{N_h - n_{Lh}}{N_h} \Big) \Big(\frac{V(y_h) - V(y_{lrh})}{n_{Lh}} \Big) \Big\} + \bar{y}_{lrh}^2 \Big(\frac{N_h - n_h}{N_h} \Big) V(w_h) \Bigg]$$

where $\quad V(y_{lrh})$ is given by (7.12)

$\qquad V(y_h)$ 〃 (7.13)

$\qquad V(w_h)$ 〃 (7.14)

(4) optimum allocation

(a) for fixed variance:

$$n_s = \frac{\sqrt{X}}{V(\bar{y}_{ts})\sqrt{c_s}} \{\sqrt{Xc_s} + \sqrt{Yc_L} + \sqrt{Zc_w}\}$$

(b) for fixed cost:

$$n_s = \Big(\frac{C}{c_s} \Big) \sqrt{\frac{c_s}{X}} \Bigg(\frac{1}{\sqrt{\frac{c_s}{X}} + \sqrt{\frac{c_L}{Y}} + \sqrt{\frac{c_w}{Z}}} \Bigg)$$

$$n_L = n_s \sqrt{\frac{Y}{X}} \sqrt{\frac{c_s}{c_L}}$$

$$n = n_s \sqrt{\frac{Z}{X}} \sqrt{\frac{c_s}{c_w}}$$

where

$$X = \overset{L}{S}[W_h^2 \sigma_{yh}^2 (1 - \rho_h^2)]$$

$$Y = \overset{L}{S}[W_h^2 \rho_h^2 \sigma_{yh}^2]$$

$$Z = \overset{L}{S}[\bar{y}_{lrh}^2 W_h (1 - W_h)]$$

6. Double sampling by photograph

There is no first sample in triple sampling.

(a) optimum sampling size:

$$n_L = n_s \sqrt{\frac{Y}{X}} \sqrt{\frac{c_s}{c_L}}, \quad n_s = \sqrt{\frac{\rho^2}{1 - \rho^2}} \sqrt{\frac{c_s}{c_L}}$$

(b) for fixed variance:

$$n_s = \frac{\sqrt{X}}{V(\bar{y}_{ds})\sqrt{c_s}} [\sqrt{Xc_s} + \sqrt{Yc_L}]$$

$$=\left(\frac{tC_v}{E}\right)^2\left[(1-\rho)^2+\rho\sqrt{1-\rho^2}\sqrt{\frac{c_L}{c_s}}\right]$$

(c) for fixed cost:

$$n_s=\frac{C}{c_s}\sqrt{\frac{c_s}{X}}\frac{1}{\sqrt{\frac{c_s}{X}}+\sqrt{\frac{c_L}{Y}}}=\frac{C}{c_s}\frac{\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}+\rho\sqrt{\frac{c_L}{c_s}}}$$

7. Some examples

At the first time, before cost consideration, we might denote the numerical values of each factor which are asummed here, in aerial photogrametric sampling. These assumming cost values were referred by the case of U.S.A.

(1) Cost (approximately assumed)

Field plot cost	c_f or c_n or c_s	$=$ 1,024$_{\text{yen}}$
Photo plot cost	c_p or $c_{n'}$ or c_w	$=$ 9$_{\text{yen}}$
	just for stratification only	
Photo plot cost	c_I or c_L	$=$ 36$_{\text{yen}}$
Field point by Bitterlich cost	c_b or c_{Lb}	$=$ 250$_{\text{yen}}$
Stratified by Bitterlich sampling cost	c_w	$=$ 100$_{\text{yen}}$
Total cost	C	$=$50,000$_{\text{yen}}$

(2) Assuming other factors

$$\rho=0.75$$
$$C_v=0.49$$
$$E=0.10$$

(A) Double sampling for regression estimates

(a) The case of fixed variance:

$$n_s=\left(\frac{2\times0.49}{0.10}\right)^2\{(1-0.75^2)+0.75\sqrt{1-0.75^2}\sqrt{\frac{36}{1024}}\}$$
$$=\left(\frac{0.98}{0.10}\right)^2\{(1-0.5625)+0.75\sqrt{1-0.5625}\times\frac{6}{32}\}$$
$$=(9.8)^2\{0.4375+0.75\sqrt{0.4375}\times0.1875\}$$
$$=96.04\times(0.4375+0.1406\times0.6614)$$
$$=96.04\times(0.4375+0.0929)$$
$$=96.04\times0.5304$$
$$=50.94\fallingdotseq50$$

(b) The case of fixed cost:

$$n_s=\frac{50{,}000}{1024}\frac{\sqrt{0.4376}}{\sqrt{0.4375}+0.75\times6/32}$$
$$=\frac{50{,}000}{1024}\times\frac{0.6614}{(0.6614+0.1406)}$$
$$=\frac{50{,}000}{1024}\times\frac{0.6614}{0.8020}$$
$$=\frac{33{,}070}{821}$$

$$=40.28 \fallingdotseq 40$$

$$n_L=\sqrt{\frac{(0.75)^2}{1-(0.75)^2}}\left(\frac{32}{6}\right) n_s$$

$$=171.45n_s$$

when $n_s=50$, $n_L=855$

when $n_s=40$, $n_L=684$

(B) Double sampling for stratification

h	W_h	$\bar{Y}_h$	$\sqrt{V(y_h)}$	$\sqrt{C_n}$	$W_h\sqrt{V(y_h)}$	$\bar{Y}_h^2$	$W_h\bar{Y}_h^2$	$W_h\bar{Y}_h$	$\sqrt{C_{n'}}$
1	0.3	30.0	15.0	32	4.50	900.00	270.000	9.00	3
2	0.6	50.0	20.0	32	12.00	2500.00	1500.000	30.00	3
3	0.1	10.0	6.0	32	0.60	100.00	10.00	1.00	3
Sum					17.10		1780.000	40.00	

$$n=\frac{(17.10)^2\times 32+(17.10)\sqrt{1780.00-(40.0)^2}\times 3}{32\times(2.0)^2}$$

$$=\frac{292.41\times 32+17.1\times\sqrt{180}\times 3}{128}$$

$$=\frac{6157.12+688.45}{128}$$

$$=\frac{6845.57}{128}$$

$$=53.5\fallingdotseq 54$$

$$n'=\frac{\sqrt{180}\times\sqrt{32}}{17.1\times 3}\times 54$$

$$=\frac{13.42\times\sqrt{32}}{51.3}\times 54$$

$$=8.37\times 54$$

$$=451.98\fallingdotseq 452$$

***n* and *n'* in chapter 8, section 4, (13).**

(C) Triple sampling design with aerial photograph

$\sqrt{c_s}=32$, $\sqrt{c_L}=6$, $\sqrt{c_w}=3$

h	W_h	W_h^2	σ_{yh}^2	ρ_h	ρ_h^2	$1-\rho_h^2$	$W_h^2\sigma_{yh}^2(1-\rho_h^2)$
1	0.2	0.04	(5) 25	0.65	0.4225	0.5775	0.5775
2	0.5	0.25	(10)100	0.75	0.5625	0.4375	10.9375
3	0.3	0.09	(20)400	0.85	0.7225	0.2775	9.9900
Sum							$X=11.5150$

$W_h^2\sigma_{yh}^2$	$W_h^2\sigma_{yh}^2\rho_h^2$	$1-W_h$	$W_h(1-W_h)$	$\bar{y}_{lrh}^2$	$\bar{y}_{lrh}^2 W_h(1-W_h)$
1.00	0.4225	0.8	0.16	100	16.00
25.00	14.0625	0.5	0.25	900	225.00
36.00	26.0100	0.7	0.21	2500	525.00
Sum	$Y=40.4950$		0.62		$Z=766.00$

$$\sqrt{X}=\sqrt{11.5150}=3.393\ ,\quad \sqrt{Y}=\sqrt{40.4950}=6.363\ ,\quad \sqrt{Z}=\sqrt{766}=27.67$$

$$n_s=\frac{\sqrt{11.5150}}{(0.8)^2\sqrt{1024}}(\sqrt{11.5150}\times 32+\sqrt{40.4950}\times 6+\sqrt{766.00}\times 3)$$

$$=\frac{3.393}{0.64\times 32}(3.393\times 32+6.368\times 6+27.67\times 3)$$

$$=\frac{3.393}{20.48}(108.578+38.178+83.01)$$

$$=\frac{3.393\times 229.766}{20.48}$$

$$=38.06=38 \qquad \text{(fixed variane)}$$

$$n_p=38.06\times\frac{6.363}{3.393}\times\frac{32}{6}=38.06\times 1.87\times 5.33=71.17\times 5.33=379.34\doteqdot 380$$

$$n=38.06\times\frac{27.67}{3.393}\times\frac{32}{3}=38.06\times 8.15\times 10.63=3288.01\doteqdot 3290$$

$n_p=n_L$ and n in chapter 8, section 5, (4).

Chapter 9. Combination of Aerial Photographs, Plotless Sampling and Plot Sampling

1. Characteristics of three methods

Before making design of combination, we must point out these characteristics, respectively.

(a) Aerial photograph has the following superior points;

(i) Broad magnificent view of whole forest

(ii) Superiority of identification in all points

(iii) Possibility of making map and stratification

(iv) Measurable of almost forest mensuration work

visible height, crown diameter, density, crown and number of trees.

Consequently, aerial photograph may be available as sampling flame, sampling list, stratification, and also measurement of almost forest mensuration items. However it is not necessarily easy to gain detail and precison of measurement given as well as on the ground.

(b) Plotless sampling has two great superiority. The greatest one is the plotless itself. This point may reduce the expense less than about half comparing to plot sampling. The second point is that plotless sampling is done by proportional probability of tree size. So strictly speaking, combinational use of plot sampling may be not theoretical. However, we may have some advantages in economic, that is, more samples, increasing sample size.

(c) Plot sampling is always essential and background of any forest mensuration work. It should be used for basic data where every measurement data will be checked.

These three sampling methods may be combined to gain more fruitful results. There may be many their combination types. Here, we will consider the following 3 types. These design, all of them, may be, what is called, triple sampling type which is defined by R. C. Wilson, in report of fifth World Forestry Congress Report, 1960. And this method also has been described by F.X. Schumacher and H.H. Chapman, Third Edition, their Sampling Textbook, Duke University, 1957.

I. Bitterlich or Photo stratification—Bitterlich's plotless—Field plot type
(Triple sampling)

1. Photo plot or Bitterlich's count may be used to classify stratum.
2. Next Bitterlich's plotless sampling on the field as the large sample.
3. Field plot may be used to make the regression as the small sample.

II. Photo stratification—Bitterlich—Field plot type
(Triple sampling)

1. Photo plot may be available to estimate the percentage of stratified area.
2. Bitterlich may be ground large sample to estimate the basal area.
3. Field plot sampling may be used to make the regression as the small sample.

III. Photo stratification—Photo interpretation—Field plot type
(Triple sampling)

1. Photo plot may be available to estimate the percentage of stratified area.
2. Moreover photo plot may be used to estimate volume by interpretation on large photo sample.
3. Field plot sampling may be used to make the regression as the small sample.

These construction will be clear because triple sampling already was discussed. These results will be compared with each other. Then, here, we are going to compare with each other by means of the examination of cost and precision sides.

2. Actual comparison of sampling design, mainly with variance and cost function

We will try the follwing types of the triple sampling design for Kyushu University forests, mainly from the view point of the variance-cost function.

(1) Construction of sampling design

The three types should be shown as follows;

(1) B–L–F type
(2) P–B–F type
(3) P–I–F type

These type may be clearly shown as the following table.

Type classification	First sample	Second sample	Third sample
B–L–F	Estimate of proportion of stratum by Bitterlich's counting	Estimate of volume by plotless sampling using angle gauge	Field plot measurement on the ground
P–B–F	Estimate of proportion of stratum by photo	Estimate of basal area by Bitterlich's count	Field plot measurement on the ground
P–I–F	Estimate of proportion of stratum by photo	Estimate of volume by identification measurement of photo	Field plot measurement on the ground

Remarks: P: stratification by photograph
B: Bitterlich's counting
F: Field plot
I: Photograph identification for measurement
L: Plotless sampling by angle gauge by Bitterlich

(2) Objective area

Hokkaido instruction Forests of Kyushu University is located at Ashoro Town in Hokkaido, situated on the left bank of the Toshibetsu Riverside, and the total area is

3727.70 ha. The forest type is broad-leaved tree stands on a relatively low hill which had belonged to the Military Horse Service before the end of the War. The length from south to north is 11.5 km, and the width from east to west is 8.0 km, latitude 43°17′–43°19′, longitude 143°29′–143°33′. The nature of geology is alluvial formation of the 3rd and the 4th stages. The slope is gentle and the average ground level is about 200–430 meters above the sea level. Everywhere the original type of forest has been disturbed. According to Dr. Tatewaki's report of forest plant investigation (1952), the standard type of vegetation is *Quercus crispula-Sasa*-colony. According to the revised management plan (1952), total forest area: 3696.88 ha, total volume: 460,447 cubic meters, timber growth: 3,477.9 cubic meters, average stand volume per ha 124.55 cubic meters.

For this total area, systematic line-plot sampling was carried out in 1954, July. The sampling intensity was about 0.08% with interval distance of 725 m and the size of sample and plot was 60 and 0.05 ha respectively. The results are shown as follows.

	Average per ha	Standard error per ha	Coefficient of variation
Volume	87.53 m^3	8.128 m^3	72.3%
Number of trees	466	43.0	71.5%

And main species are shown as follows:

Quercus sp.	21%,	*Sophora*	14%,
Acer sp.	14%,	*Tilia*	6%,
Fraxinus	5%,	*Betula* sp.	5%,
Salix sp.	5%,	and less items, no soft woods.	

(3) Pre-examination and materials of designs

So we have decided to classify by volume and vegetation type and we worked out according to the management list book of forest investigation issued in 1952 by Kyushu University Forests office. And this survey is limited to the 2nd and 4th sub-period in working plan, so total acreage area is 1756.04 ha. Now, the following classification may appear by checking of each compartment.

Volume class	Vegetation type A	Vegetation type B	total
I	409.57 ha	693.13 ha	1102.70 ha
II	123.42	291.91	415.33
III	61.34	176.67	238.01
total	594.33	1161.71	1756.04

Remarks: A: stand having mainly *Quercus* sp. more than 70%
B: stand having less *Quercus* sp. less than 60%
I: volume class, more than 101 m^3
II: volume class, between 51 and 100 m^3
III: volume class, less than 50 m^3

Here, in brief, we limited the stratification by the volume class only. They give numerical value of W_h, the proportion of stratum. Next we need the standard error of volume σ_{yh} in each stratum. According to Kenkichi KINASHI: Timber volume Estimation by Sampling Method in Hokkaido Instruction Forest of Kyushu University, 1955, (No. 25, the Bulletin of the Kyushu University Forests, page 72).

Strata	$\overline{Y}_{lrh}$ Average volume per ha m^3
I	137.42
II	60.26
III	16.32

3. **Optimum allocation with fixed variation in triple sampling design in the University Forest**

At the first time we should calculate

$$X=\overset{L}{S}[W_h^2\sigma_{yh}^2(1-\rho_h^2)]$$
$$Y=\overset{L}{S}[W_h^2\rho_h^2\sigma_{yh}^2]$$
$$Z=\overset{L}{S}[\overline{Y}_{lrh}^2W_h(1-W_h)]$$

where ρ_h may be assumed in both cases, photograph and Bitterlich, respectively. In the case of aerial photo, ρ may be correlation coefficient between interpretation volume and field plot volume. In the case of Bitterlich, it may be one between basal area of Bitterlich's way and field plot volume. The degree of correlation may be higher in Bitterlich case than aerial photograph case. Then X and Y will be estimated corresponding to both case, but Z will be the same, because Z does not include any ρ terms.

Stratum			Bitterlich		Aerial photograph	
	W_h^2	σ_{yh}^2	ρ_h	ρ_h^2	ρ_h	ρ_h^2
I	0.3943	21.4156	0.90	0.8100	0.80	0.6400
II	0.0559	3.0304	0.85	0.7225	0.75	0.5625
III	0.0184	1.2924	0.80	0.6400	0.70	0.4900

h	$W_h^2\sigma_{yh}^2\ (1-\rho_h^2)$	
	Bitterlich	Aerial photograph
I	1.6044	3.0399
II	0.0470	0.0741
III	0.0086	0.0121
	X_B=1.6600	X_P=3.1261

h	$W_h^2\ \rho_h^2\ \sigma_{yh}^2$	
	Bitterlich	Aerial photograph
I	6.8401	5.4053
II	0.1224	0.0952
III	0.0153	0.0116
	Y_B=6.9778	Y_P= 5.5121

h	$\overline{Y}_{lrh}$	$\overline{Y}_{lrh}^2$	$W_h(1-W_h)$	$\overline{Y}_{lrh}^2W_h(1-W_h)$
I	13.74	188.7876	0.2336	44.1008
II	6.03	36.3609	0.1806	6.5668
III	1.63	2.6569	0.1171	0.3111
				Z=50.9787

Assuming

$$V(\bar{y}_{ts})=\left(\frac{0.875}{2}\right)^2=(0.437)^2=0.1910$$

Above calculations are based on 0.1 ha plot and 0.875 is 10% of average volume 87.5 m^3 per ha in the last sampling result.

Calculation of small sample size n_s :

Type	$\sqrt{Xc_s}+\sqrt{Yc_L}+\sqrt{Zc_w}$	$\frac{\sqrt{X}}{V(\bar{y}_{st})\sqrt{c_s}}$	n_s	Number of field plot (small sample)
(1)	41.2288+52.8300+71.4000=165.4588	0.2108	34.8787	35
(2)	41.2288+52.8300+21.4200=115.4788	0.2108	24.3429	24
(3)	56.5792+14.0868+21.4200= 92.0860	0.2893	26.6405	27

Calculation of large sample size n_L:

Type	$\sqrt{Y/X}$	$\sqrt{c_s/c_L}$	$n_L=\sqrt{\frac{Y}{X}}\sqrt{\frac{c_s}{c_L}}$	Number of large sample
(1)	2.0502	1.6000	114.4126	114
(2)	2.0502	1.6000	79.8520	80
(3)	1.3279	5.333	188.6707	189

Calculation of first sample size n:

Type	$\sqrt{\frac{Z}{X}}$	$\sqrt{\frac{c_s}{c_w}}$	$\sqrt{\frac{Z}{X}}\sqrt{\frac{c_s}{c_w}}$	Number of first sample
(1)	1.6445	3.2000	183.5457	184
(2)	1.6445	10.6667	427.0085	427
(3)	1.1983	10.6667	340.5162	341

All these design aims at about 10% error in 95% probability. Based on each number of the different stage sampling plot and the corresponding expenditures, we may be able to compare total cost of inventory respectively in each case.

Total cost will be shown as in the following table:

		Design (1) B-L-F type		Design (2) S-B-F type		Design (3) S-I-F type	
		Number of plots	Cost	Number of plots	Cost	Number of plots	Cost
First sample	n	184	¥ 100	427	¥ 9	341	¥ 9
Second sample	n_L	114	400	80	400	189	36
Third sample	n_s	35	1,024	24	1,024	27	1,024
Total Cost			¥ 99,840		¥ 60,419		¥ 37,521

Conclusion

Design (1) is the most expensive. There is no use of aerial photograph. Even if Bitterlich's method is fully used, cost is expensively high.

Design (2) is next and aerial photograph is used only as stratification. But even though for stratification only, aerial photograph improves so economical that cost down about one third of the first case.

Design (3) is the most economical way and this is due to aerial photograph only, except using ground plot as subsidiary check plot only. It's cost down so low that is about one third of the Bitterlich's method. In any way increasing of the degree of mixing with aerial photograph brings the results of the decreasing of cost. In present time, aerial interpretation is improving more and more, finally we should make complete procedure of interpretation of aerial photograph which means high value.

Even slight modification or any idea of development of interpretation of aerial photograph will help to improve the increasing correlation coefficient.

Remaining problem

It may be important remaining problem to decide the value of plot size so as to having high correlation between ground plot value and crown density of photo interpretation, or basal area or value by Bitterlich's plotless sampling.

Chapter 10. Utilization of Multiple Regression System of Forest Sampling Design for Management Plan

1. Introductory study, how we apply sampling data to management plan

The information of forest resources may be gained much more by means of the recently improved sampling design which we have discussed through several chapters. Only big problem which we couldn't have a solution is the estimation of each stand, especially in investigation for management plan. We have always the policy of estimation of individual stands. There are, however, the resulting information construction, what is called and known as Yield Table, in our old traditional way in forestry. The problems of yield table so frequently have been discussed by many persons. Recently some excellent treatments by F. X. Schumacher has appeared. Here we should learn the practical application to a case study in their book. (Schumacher and Coile: Growth and Yields of Natural Stands of the Southern Pines, 1960, page 24.)

"The preparation of an inventory of land and timber resources of a 50,000-acre property in Alabama is this. Soil-site maps were first made of the entire property. Then stand condition-classes that were recognizable on aerial photographs were sketched thereon. These areas were later examined in the field while running a line of plots through the center of each tier of 'fortys' of each square mile, with a tenth-acre sample plot at 10-chain intervals.

It is the policy of the owners of this property to keep detailed records of each section (640 acres) as the unit. Consequently, as condition-classes were finally delineated in the field, they were numbered separately within sections and described by certain characteristics: covertype; age, determined by increment boring; site class, from the initial soil-site survey; density of understory which might affect reproduction or site preparation; presence of special products, such as poles or piling; and finally, an ocular estimate of stocking percentage of the main stand, in which 100 represents the well-stocked area.

With but little practice, stocking percentage can be readily estimated by eye, since it is proportional to basal area per acre. For purpose of forest management it is necessary to have for each classification within each section an estimate of present volume in standard cords and in board feet, and the predicted volume for 10 years hence. If, for each classification within each section, the total basal area and total number of trees per acre are known, then the required estimates can be based upon this method which will be shown.

Because of the nature of the inventory design, site index, age of stand, and ocular estimate of stocking percentage apply to the entire area of each classification within its parent section. On the other hand, the number of trees and basal area measured directly and thus apply only to the sample plots. Since the 0.1-acre plots were located at 10-chain intervals along parallel lined 20 chain apart, they supply direct measurements on only about one-half of one per cent of the forest area. Furthermore, many of the classifications did not contain sample plot in every section of occurrence, particularly if small in area and off the cruise lines; hence they required indirect estimates of their basal areas and number of trees.

Altogether, 855 sample plots of the loblolly-pine type were measured in the field survey of the entire property. When these were sorted by site-index, age, and ocular stocking percentage, it is deemed that the best estimate of number of trees and basal area per acre of each class was not the directly measured averages of the plot data but their corresponding indirectly measured average. These may be smoothed either by a free-hand curve process, or by fitting suitable forms of regression equations to the plot data so as to express (a) basal area per acre, and (b) number of trees per acre in terms of site-index (S), age (A), and ocular stocking percentage (SP) of the class.

(a) $B=-10.48+0.4956(S)+0.0273(SP)-0.5383(\mathrm{A}-5)$

(b) $N=-26.43+0.9577(S)+1.6388(SP)-0.8091(\mathrm{A})$

Since these equations were derived from data of the loblolly-pine type of the entire property, they should apply equally well to that type throughout the property".

This inventory idea is much approach to the management plan—we were looking forward to such approach too much. Solution may be complete and we like to try these examinations.

Here we have some case study. We like to introduce our example. Before our illustration, we should make clear Prof. Schumacher's idea and the principle of this procedure.

2. Principle

Schumacher presented multiple regression methods for yield table construction which consists of four multiple regressions and one normal approach equation.

(1) Height regression

The first regression is a height curve which is well known in his text book, Bruce and Schumacher: Forest Mensuration, 1935–1950.

$$\mathrm{Log}\ H=b_0+b_1\ (1/A) \tag{10.1}$$

Where H: height of stand

A: age of stand

b_0, b_1: regression coefficient

The site curve may be derived from this height equation. If site index will be defined as the height of trees at 40 years in age, in Japanese Sugi (*Cryptomeria Japonica* D. Don) plantations, the tree height of the same site index in any age may be calculated from the following regression:

$$\mathrm{Log}\ H=\mathrm{Log}\ SI+b_1(1/A-1/40) \tag{10.2}$$

where SI: site index which means the height of stand in 40 years.

(2) Stocking percentage regression

This formula may be quite a peculiar form in comparison with other stocking equations. It, however, is a reasonable one. They give the following definition—by stocking percentage is meant the ground area that an even-aged stand or sample plot of given age, dominant tree height of stand, and *dbh* distribution, would have utilized in a well-stocked stand relative to the actual area of the stand or sample plot. It is the percentage ratio of the calculated ground area to actual ground area. Thus if B represents basal area per ha,

$$S=B_0[b_0+b_1(H/10)+b_2(10/A)+b_3(H/A)] \tag{10.3}$$

The stocking percentage equation is a property of its basic data. If the calculated stocking percentage, obtained by inserting into the equation the numerical value of B, H and A of an actual stand should turn out to be 100, one does not assert that the stand is normally stocked, but merely that its stocking is about the average of the data that provided the equation; and this average may, or may not correspond to the ideal of normal stocking.

(3) Normal approach assumption

Normal approach will be assumed in linear type in logarithms showing as the following formula. This assumption may have some forest biological background, but here we can recognize some evidence by the change of stands.

After showing this assumption formula, one of authors can give some demonstration by Cryptomeria stands.

$$\frac{\text{Log } s-2}{\text{Log } s_0-2}=\frac{A_0}{A} \tag{10.4}$$

Where A_0 is an initial age and A is expected age. s_0 and s are stocking percentage corresponding age A_0 and A respectively.

Since $(\text{Log } s-2)$ means difference between $\text{Log } s$ and Log 100, the above formula shows that the difference of the stocking may change and approach to normal proportionally to the reciprocal of age. This evidence may be shown by the following Japanese Cryptomeria data. Using stocking formula in metric system, while proceeding equations are based on Japanese traditional measurement unit;

$$S=B[1.0226+0.0272H+45.9591(1/A)-2.6180(H/A)]$$

			Stocking percentage at A_1		Estimate discrepancy
Plot	A_0	A_1	Actual	Calculation	[Log Act-Log Cal]
1	26	43	105.60	77.55	0.133
2	36	41	80.41	88.27	−0.041
3	51	57	68.80	78.18	−0.055
4	51	57	59.75	57.50	0.016
5	15	24	101.19	89.83	0.051
6	27	41	77.91	70.52	0.044
7	18	28	72.99	78.71	−0.033
8	27	37	53.98	64.21	−0.076
9	65	75	106.37	95.96	0.083
					0.083

Then $SE(0.083)=\sqrt{9(0.004403)}=0.184$

where 0.004403 is the sum of square of estimate discrepancy in logarithmics

$$t=\frac{0.083}{0.184}=0.45 \text{ not significant } (df=8).$$

Per average stand, the discrepancy in logarithm which shows the reliability of this assumption formula is 0.0092 ± 0.0204. It equivalents $2.1\%\pm4.7\%$ in usual scale. Also Schumacher verified this point concerning Southern Pines in U.S.A. So this normal assumption may be satisfied in Japan.

(4) Number of tree equation

If N is number of trees per ha,

$$\text{Log } N=b_0+b_1 \text{ Log } H+b_2 \text{ Log } B+b_3(1/A) \tag{10.5}$$

(5) Volume equation

If V is volume per ha,

$$\text{Log } V = b_0 + b_1 \text{ Log } B + b_2 \text{ Log } H + b_3 \text{ Log } N \tag{10.6}$$

All these multiple regression will be fitted well by the least square solution and analysed by analysis variance.

Then if B and N may be expressed by SI, S and A, we can estimate each stand volume and growth by just looking ocular estimate of stocking or by simple Bitterlich's count estimate of stocking percentage. These idea will be checked by the following analysis of variance. This shows that the stand volume depended upon stocking, site, age, main 3 effects and 5 interactions. One of them will be examined as shown in the following analysis of variance of stand volume:

Source of variation	*df*	*ss*	*ms*	*F*
Replication	1	15105.9	15105.9	0.1927
Stocking percentage	2	4335768.1	2167884.1	27.6486
Site index (SI)	2	720311.3	360155.7	4.5933
Age	3	12163313.1	4054437.7	51.7091
$S\times SI$	4	67247.7	16811.9	0.2144
$S\times A$	6	1167441.3	194573.6	2.4815
$SI\times A$	6	860536.7	143422.8	1.8292
$S\times SI\times A$	12	849458.3	70788.2	0.9028
Errors	35	2744299.6	78408.5	
Total	71	22923482		

This table may indicate that all of three main effects, stocking percentage, age and site, are significant for volume variation, but interactions are not significant except (stocking$\times$age) interaction. It is noticeable that site effect is significant at 5% level and all interactions with site are not significant, regardless of the importance of site for estimation of stand volume. Also we understand that site may be independent factor, while stocking percentage may connect with age, much reasonable.

3. Some illnstration of Japanese Cryptomeria data

The multiple regressions may be shown from Cryptomeria data.

(1) Well-stocked strata;

$$S = B[0.0939 + 0.0045H + 4.2203(1/A) - 0.4371(H/A)]$$
$$\text{Log } H = 1.1396 - 8.2459(1/A)$$
$$\text{Log } H = \text{Log } SI - 8.2459(1/A - 1/40)$$
$$\text{Log } N = 3.7967 - 0.6947 \text{ Log } H + 2.7458(1/A)$$
$$\text{Log } N = 3.6653 - 0.6751 \text{ Log } H + 1.1496(1/A)^{1/2}$$
$$\text{Log } V = 1.5103 \text{ Log } B - 0.3290 \text{ Log } N$$

(2) Average-stocked strata;

$$S = B[0.1454 + 0.0036H + 4.4809(1/A) - 0.4779(H/A)]$$
$$\text{Log } H = 1.0793 - 7.2137(1/A)$$
$$\text{Log } H = \text{Log } SI - 7.2137(1/A - 1/40)$$
$$\text{Log } N = 2.9764 + 6.0916(1/A)$$
$$\text{Log } V = 1.5401 \text{ Log } B - 0.3458 \text{ Log } N$$

(3) Unrestricted data, no stratification;

$$S = B[0.650 + 4.1208(1/A)]$$

$$\text{Log } H = 1.0699 - 6.3641(1/A)$$
$$\text{Log } H = \text{Log } SI - 6.3641(1/A - 1/40)$$
$$\text{Log } N = 2.7733 - 1.0237 \text{ Log } H + 0.4901 \text{ Log } B$$
$$\text{Log } V = -0.9430 + 1.0344 \text{ Log } B + 0.9077 \text{ Log } H + 0.1280 \text{ Log } N$$

and the difference between well and average are

$$S_w = 0.0885B + 0.0502BH + 0.4494B/A - 0.4674BH/A$$
$$S_A = 0.1313B + 0.0502BH + 0.4494B/A - 0.4674BH/A$$
$$\text{Diff.} = S_w - S_A = (0.0885 - 0.1313)B = -0.0428B$$

This shows that the same basal area gives lower stocking percentage by well stock equation than by average stocking equation. So always we must determine what base of stocking one taking in the problem.

If $S_w = S_A = 100$,

$$\frac{B_A}{B_w} = \frac{0.0885}{0.1313} = 0.67$$

means that basal area of average will be lower about 33% than well stocked stand. (Kenkichi KINASHI: Yield Table Study of Cryptomeria Growing in Northern Kyushu, The reports of the Kyushu University Forests, No. 12, March, 1959.)

One noticeable thing is that random data will not supply good stocking equation but it may give superior volume equation. In Japan, the investigation of forest management plan is adopting two systems in parallel, one is based on sampling and the other is old traditional way. Sampling is giving the total volume only, while the old method can give the evaluation of each stand volume. Then, they have two kind of total estimates, one of which comes from sampling results, and another one of which comes from the accumulation total of each stand estimation. And then the latter is slide up or down toward the lowest confidence limit of sampling result. Each stand volume might be adjusted in this direction. This procedure is not agreeable, because even some stand which was underestimated may be more decreased downward, if the whole accumulated result was higher than the total of sampling result. In the other hand, even some stand which was overestimated may be more increased upward, if the whole accumulated results were lower than the total of sampling estimate. So it might be better that reasonable stand factors in each stand are given by means of regression equations in which the number of trees and the basal area per ha will be given by stocking percentage (obtained by ocular estimate or some simple devices like Bitterlich's method), site index and age of stand.

Finally one of authers, K. KINASHI is grateful to H.R. JOSEPHSON, Director of Economic Research, U. S. Forest Service who encouraged our work and gave us much help.

References

Bitterlich, W.: Optishe Zählmessung in der Bestands aufnahme. Fahrbuch, Hochschule für Bodenkultur in Wien. (1948)

Bitterlich, W.: Was ist die Hirata-Höhe? Allgemine Forstzeitung. 67. Jahrgang, November, (1956)

Bruce and Schumacher: Forest Mensuration, McGraw Hill, (1935)

Cochran, W. G.: Sampling techniques, John Wiley & Son. (1953)

Cromer, D. A. N.: Basal area per acre. Austral. For. XVI

Cromer, D. A. N.: Techniques and instruments for the determination of basal area and volume per acre. Austral. For. XVIII 2, (1954)

Finch, S.: Plotless enumeration with angle gauges. Forest. Vol. XXX, No. 2 (1957)

Grosenbaugh, L. R.: Plotless timber estimates. New, Fast, Easy. Jour. Forestry. 50 (1952)

Hasen, M. H., Hurwitz, W. N. & Madow, W. G.: Sample survey methods and theory. I, II. John Wiley & Sons (1953)

Hirata, T. : Height estimation through Bitterlich's method vertical angle count sampling. Journal of the Japanese Forestry Soc. 37. (1955)

Hush, B.: Results of an investigation of the variable plot method of cruising. Jour. Forestry. 54(8) (1955)

Kendall, M. G. and Buckland, W. R.: A Dictionary of Statistical terms. International Statistical Institute (1957)

Kinashi, K.: Forest Inventory by Sampling Methods, Bulletin of the Kyushu University Forests, No. 23, March, (1954)

Kinashi K.: Yield Table study of Cryptomeria growing in Northern Kyushu, The Reports of the Kyushu University Forests, No. 12, (1959)

Matérn, S.: Forest surveys and the statistical theory of sampling some recent developments. (Fifth world Forestry Congress) (1960)

Matérn B.: Spatial Variation—Stocastic models and their application to some problems in forest surveys and other sampling investigations. (1960)

Meyer, H. A.: Forest Mensuration, Penns. Valley, (1953)

Nishizawa, M.: Research on the estimation and prediction methods of the stand increments. Bul. of the Gov. Forest Exp. No. 129, March (1961)

Patterson, H. D.: Sampling on successive occasion with partial replacements. Jour. Roy. Stat. Soc. Series B 12, (1950)

Schumacher, F. X. and Chapman, R. A.: Sampling Methods in Forestry and Range Management, Duke University, (1948)

Schumacher and Coile: Growth and Yields of Natural Stands of the Southern Pines, T.S. Coile, Inc. (1960)

Sukhatme, P. V.: Sampling theory of surveys with applications. Iowa State College Press (1954)

Warner, J.: Die Probefläche ohne Probefläche. Allg. Forst-und Holz w. 2tg. lix. (1948)

Wilson, R. C.: Surveys particularly applicable to extensive forest areas. (Fifth world Forestry Congress, 1960)

Yates, F: Sampling methods for censuses and surveys, Hafner publishing company (1953)