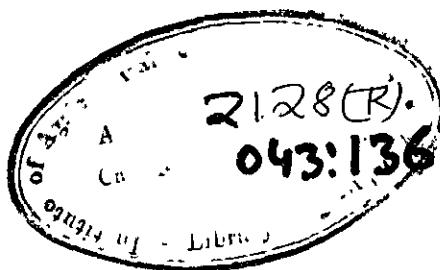


S O U R E S T U D I E S C P L A 1

O F

F R O M I N S O N V A X S



✓ GERALD STINGA

PK-69
9

INSTITUTE OF AGRICULTURAL RESEARCH STATISTICS
(I.C.A.R.)
MADRAS AVENUE, NEW DELHI-12.
1966

SOME STUDIES OF PRACTICE OF
PLANT SURVEY

BY

JOSEPH SINGH

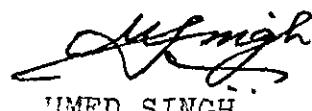
DISSENTATION.

SO MITTER IS FULL LENGTH OF THE
REQUISITES FOR THE AWARD OF DIPLOMA
IN AGRICULTURE AND ANIMAL HUSBANDRY STATISTICS
OF THE INSTITUTE OF AGRICULTURAL AND RURAL STATISTICS
(I.C.A.R.), NEW DELHI

A C K N O W L E D G E M E N T

I wish to express my deep sense of gratitude to Dr. B.V. Sukhatme, Senior Professor of Statistics, Indian Council of Agricultural Research, for his inspiring guidance and constant encouragement throughout the course of investigation and in preparation of thesis.

I am also deeply indebted to Dr. E.G. R. Seth, Statistical Adviser, Indian Council of Agricultural Research, for providing me with adequate research facilities and allowing me to use data collected in the lime survey.



UMESH SINGH

C O N T E N T S

CHAPTER

page.

I	Acknowledgement	(i)
ii	Contents	(ii)
(I)	Introduction and summary	1
(II)	Stratification	6
(III)	Method of sampling	19
(IV)	Probability scheme of sampling	29
(V)	Estimation of average yield	42
(VI)	Estimation of total production	68
(VII)	Estimation procedure in stratified multi-stage sampling	95
	Bibliography	227

CHAPTER ONE

I N T R O D U C T I O N A N D S U M M A R Y

The importance of increasing the production of fruit crops hardly needs to be emphasized especially in view of the fact that the per capita production and availability of fruits is very low. It is, therefore, necessary to formulate suitable development programmes for increasing the production of fruit crops. However, this is not possible in the absence of comprehensive and reliable statistics regarding area, yield rates and cultivation practices of fruits. The present position in regard to statistics of acreage under lime, like other fruit crops is not very satisfactory. So far as yield statistics are concerned, no reliable information is available regarding the average yield per tree. The total production is estimated by the official agency on the basis of normal yield of 12,000 lbs per acre and is entirely based on subjective judgement and has hardly any relevance so far as the yield for a particular year or tract under varied soil and climatic conditions is concerned. In order to fill up the lacuna in agricultural statistics and supply sound bases for the planning of development in horticulture, the Indian Council of Agricultural Research initiated a series of sampling investigations on a coordinated basis on important fruit crops. The main objectives of these sampling investigations were

- (1) to evolve a suitable sampling technique for estimating with a desired degree of precision, the average yield rate and annual production.
- (2) to estimate the area under fruit crop with a given degree of precision.
- (3) to collect reliable data on manurial and cultivation practices of the crops as practised by the cultivators.

One such investigation was carried out in 1963-64, to

study yield rates, production and cultivation practices of lime crop in Nellore district of A.P. Nellore is one of the most important centres in Andhra Pradesh growing lime crop in systematic and well-cared orchards. Out of 13 Talukas in the district, lime is grown in five southern talukas viz, Venkatagiri, Rapur, Sulurpal, Kovur and Atmakur. The sampling design adopted for the lime survey may be described as stratified two phase three stage random sampling with talukas forming main strata, villages as primary units of sampling, orchards as secondary units of sampling and cluster of trees as ultimate units of sampling. The survey was designed to collect information concerning:

- (a) Extent of lime cultivation as indicated by
 - (i) area under the crop
 - (ii) Total number of bearing and young orchards &
 - (iii) total number of trees under bearing and non-bearing categories
- (b) Yield per tree both in terms of weight and number of fruits during the entire harvesting period
- (c) Cultivation practices such as spacing between adjacent trees, method of planting, frequency of irrigation, manuring and other cultural operations as followed by the cultivators, incidence of pests and diseases of lime crop.

From the point of view of inspection and supervision of field work, availability of sampling frame and administrative considerations, geographical proximity was adopted as the basis of stratification. Talukas were, therefore, taken as strata. Allocation of sampling units to different strata was made roughly in proportional to the reported area under lime in the stratum subject to a minimum of four units per stratum. Table below

gives the area reported under lime, the number of villages reported under lime and the allocation of the sampled villages to the different strata for the purpose of enumeration & yield study.

Taluka	Area under lime (in ha)	No. of villages in the Taluka	No. of villages selected for the study of extent of cultivation	No. of villages selected for the study of yield
Venka-tagiri	836	141	33	16
Rapur	438	70	20	10
Sulur-pat	214	26	12	5
Kovur	63	25	10	7
Atma-kur	89	42	11	7
Total	1613	304	86	45

The five talukas under survey together comprise of 732 villages of which 304 villages were reported to be growing lime. A sample of 86 villages was selected and the selection of villages within each stratum was made with probabilities proportional to area reported under lime during the year 1961-62. All the selected villages were completely enumerated to obtain information regarding area, number of orchards and total number of bearing trees. Out of the 86 selected villages a sub-sample of 45 villages was retained for yield study. Four orchards were selected at random in each of the selected villages and in each of these selected orchards a sample of twelve trees was selected for the purpose of yield study. The selection at second and third stages was made with equal probabilities and without replacement.

The main object in this investigation is to make a critical study of the data collected in the survey on lime crop with a view to obtain some guidance in the planning of surveys of fruit crops. Some of the most important problems requiring considerable thought and study in the planning of surveys are:

- (i) Principle of stratification
- ✓(ii) Determining optimum number of strata
- ✓(iii) Determination of optimum points of stratification
- ✓(iv) Allocation of sample size to different strata
- (v) Choice of sampling unit
- (vi) Determination of number of stages and choice
- (vii) Probability scheme of sampling at each stage ,
- (viii) Double sampling in relation to cost ,
- (ix) Best estimation procedures for various characters under study.
- (x) Determination of optimum sample size etc.

This will naturally involve the best possible use of all the ancillary information available from past surveys. The idea is to utilize all this information taking into consideration the various factors mentioned above in such a manner that it is possible to estimate the various characters under study with maximum precision.

Chapter II deals with the determination of optimum points of stratification and four different methods for determining the strata boundaries have been compared. The problem of determination of the optimum number of strata has been studied here. Also the problem of allocation of sample sizes to the different strata with a view to estimate the total production and the number of bearing trees with maximum precision is considered in this chapter. In chapter III the relative efficiency of the two sub-sampling designs,

one with villages as the primary unit of sampling, orchards as the secondary unit of sampling and a cluster of trees as the ultimate unit of sampling and the other villages as the primary unit of sampling and a cluster of trees as the secondary and ultimate unit of sampling, has been discussed.] Investigation of the efficiency of adopting varying probability sampling at the first stage as compared to sampling with equal probability for estimating simultaneously several characters such as yield rate, total production, number of bearing trees etc. has been made in chapter IV.] In chapter V theory has been developed for a generalized ratio-type estimate of average yield per-tree.

The estimate considered is of the form

$$\bar{y} = \frac{\hat{y}}{\hat{B}}$$

where the numerator represents the estimate of the total production and the denominator represents the estimate of the total number of bearing trees. Ten particular cases of this generalized ratio-type estimate have been considered for determining the best estimation procedure for average yield per tree among them.

Theory for a Generalized Double Sampling Estimate of total production has been developed in chapter VI. Efficiency of double sampling singly and in relation to cost has also been studied here. The important problem of finding the best weights for obtaining an over-all estimate of average yield per tree for a stratified population has been considered in the last chapter.

CHAPTER TWO

S T R A T I F I C A T I O N

1. DETERMINATION OF OPTIMUM POINTS OF STRATIFICATION

Stratification can be used to increase the reliability of sample results. The amount of increase in precision of sample estimates accomplished by stratification will depend on the degree of homogeneity that is achieved within strata, or, saying the same thing in another way, on how much of the variability in the characteristic being estimated is reflected in the differences among the strata. This in turn depends on how effectively the strata have been defined.

The determination of the strata is a matter in which effective use can be made of prior knowledge, personal intuition, and judgement, as well as of objective statistical information that may be available. Whether objective information is available or not, the final determination of the strata is a subjective matter, in which the decisions must be judgements. Statistical theory does not provide a general series of procedures or steps for determining the one best set of strata. It does provide some guiding principles and gives a method for comparing and choosing among alternatives.

The most effective variable on which to stratify would be the characteristic to be measured and since in practice this is not feasible, stratification on the basis of variable highly correlated with the one under study is expected to result in substantial reduction of the variance. However, the cost of accomplishing the stratification must be taken into account, and this may lead to adoption of simpler stratification techniques.

On the other hand, stratification might be carried out on the basis of a casual external inspection of the different



units in the population. Another source of information for stratification might result from records of one sort or another that might be available on individual units. Examples of more objective approaches to the problem of stratification are the methods for setting up boundaries of strata designed by Dalenius, Gurney and others (1953, 1959, 1959, 1948).

Let y_0, y_1, \dots, y_k be the points of stratification, the strata being numbered 1, 2, ..., k and let μ_t, σ_t be the mean and standard deviation of the t-th stratum and $p_t = \frac{N_t}{N}$ the proportion of sampling units in the t-th stratum.

With stratified sampling the usual estimate of the population mean is

where \bar{y}_{nt} is the sample mean in the t -th stratum.

and its variance is given by

$$V(\bar{y}_{st}) = \sum_{t=1}^k p_t \left(\frac{1}{N_t} - \frac{1}{N} \right) \sigma_t^2 \quad \dots \dots \quad 2$$

Dalenium (1951) has shown that under proportional allocation when n_t is proportional to p_t , the optimum points of stratification are obtained by solving the set of simultaneous equations

$$y_t = \frac{1}{2}(\mu_t + \mu_{t-1}) \quad \dots \quad (t=1,2,\dots,k-1) \quad \dots \quad 3$$

As μ_{ϵ_0} and μ_{ϵ_1} are themselves functions of the stratification points, the optimum points of stratification will have to be found by some iterative method, such as the one suggested by Sethi (1963). This method is perfectly general and can be used directly in actual problems.

SETHI'S ITERATIVE METHOD

(i) Start with some arbitrary set of points

$$y_1^i, y_2^i, \dots, y_{k-1}^i$$

(ii) Calculate the proportions and the means for the strata defined by this set.

(iii) Replace the initial set by

$$y_t^n = \frac{1}{2} (\mu_t + \mu_{t-1})$$

(iv) Repeat the steps (ii) and (iii) till two consecutive sets are either identical or differ by negligible quantities.

It has also been shown by Dalenius (1951) that under Neyman allocation where n_t is proportional to $p_t \sigma_t^2$ points of stratification satisfy the equations

$$\sigma_t^2 + \frac{(y_t - \mu_t)^2}{\sigma_t^2} = \sigma_{t+1}^2 + \frac{(y_{t+1} - \mu_{t+1})^2}{\sigma_{t+1}^2} \dots \dots \dots \quad 4$$

(t=1,2, ... k-1)

These equations cannot usually be solved unless the number of strata is small. Attempts have therefore been made to find approximations which are given below:

(a) Equalization of Strata Totals

Hansen, Hurwitz and Madow (1953) suggested that under certain conditions a very simple rule may give optimum points of stratification. This rule is to make strata totals equal. The condition under which it gives optimum points of stratification is that within strata coefficients of variation are equal and remain about the same on adjusting the strata sizes. This rule cannot be of general validity, since the boundaries which it gives are changed if the origin of the scale is changed, whereas the correct solution is invariant under change of origin.

(b) APPROXIMATION SUGGESTED BY DALENIUS & HODGES (1959)

Let $f(y)$ be the density function of the variable under study. According to this approach, the values of $\sqrt{f(y)}$ are cumulated. If the cumulated total over the whole range of the

variable is H, the first approximates to the optimum points of stratification are given by

$$y_t = \frac{tH}{k} \quad t=1, 2, \dots, k-1$$

A proof justifying this approach and its illustration is given by Dalenius & Hodges (1959).

In practice, we come across discrete distribution with unequal class-intervals, the intervals becoming wider towards the upper end of the Y scale. When the interval changes from one of length d to one of length ud , the value of $\sqrt{f(y)}$ in the longer interval must be multiplied by \sqrt{u} when forming cumulative \sqrt{f} .

(c) EKMAN'S RULE

Ekman (1959) suggested making

$$p_t (y_t - y_{t-1}) = \text{Constant}$$

In this method we equalize the product of the cumulative frequency within the stratum and width of the stratum. This method is a little troublesome to apply, because the value of

$\sum_{t=1}^k p_t (y_t - y_{t-1})$ is not constant, depending both on k and on the positions of the boundaries. Hence for given k , it is not at first clear as to what figure we should try to equate $p_t (y_t - y_{t-1})$ in the individual strata. A rough guide suggested by Ekman is to compute the product for $k=1$. For k strata, the constant value per stratum is approximately $\frac{Q}{\frac{k}{2}}$, where $Q=(\text{cum } f)$.
 $(y_k - y_0)$.

(d) DURBIN'S RULE

Let $F(y)$ be the cumulative distribution of the variable under study. Form a rectangular distribution

$$r(y) = \frac{F(y_k)}{y_k - y_0} \quad \text{over the same range}$$

where y_k and y_0 are the highest and lowest values of

the variate in the population. The stratum boundaries are then obtained by taking equal intervals on the cumulative of

$$\frac{1}{2} \left\{ r(y) + f(y) \right\}$$

This rule amounts to forming the strata by taking equal areas under a frequency distribution with density half-way between the original distribution and a rectangular distribution.

1.2 DETERMINATION OF OPTIMUM NUMBER OF STRATA

There is a practical limit, often reached very early, as to the number of strata that can be introduced, since at least one unit must be included in the sample from each stratum.

Although some (and sometimes very small) reduction in variance may be expected by stratifying deeper and deeper, the first few strata, if they are well chosen, are likely to be most effective and further stratification may not yield any significant gains.

From the point of view of sampling theory, if stratification is carried out to the point of including in the sample only one unit from a stratum, then there is no way of obtaining a consistent estimate of the variance of the sample itself.

However, this may not be regarded disadvantage in any particular problem as methods are available for estimating approximately the sampling variance.

It should be pointed out that restriction to smaller number of strata puts a strain on the methods of determining the optimum points of stratification for optimum and proportional allocations, most of which depend more or less on the assumption that the number of strata is reasonably large so that within a stratum the frequency function $f(y)$ can be assumed to be rectangular. But some of the methods do vary satisfactorily even with small number of strata. Although the number of strata used in practice will vary with the conditions of the survey, there

are practical advantages in keeping this number small. Dalenius (1957) has showed that with the best boundaries, the rate at which the variance $V(\bar{y}_{st})$ decreases with increasing number of strata agrees well with the formula $\frac{V_k}{V_{k-1}} = \left(\frac{k-1}{K}\right)^2$, where V_{k-1} is the variance when the number of strata is $(k-1)$.

1.3 ALLOCATION OF SAMPLING UNITS AMONG THE STRATA

It is not true that any stratified random sample gives a smaller variance for the estimated mean or total than is given by a comparable simple random sample. If the distribution of the sample among the different strata is far from optimum, stratified sampling may not prove efficient. Hence proper allocation of sampling units is very essential.

Tschupro (1923) and Neyman (1934) established an important result that in stratified random sampling, the variance of the estimated mean \bar{Y}_{st} is minimum for a fixed total size of the sample, if the number of sampling units n_t allocated to the t -th stratum is proportional to $N_t^{-\alpha}$. This requires knowledge of the population standard deviations which is rarely available. The non-availability of this information forces one to utilize information on some other auxiliary character closely correlated with the one under study or information collected from some previous survey. Thus instead of optimum allocation one may, in practice, only hope to get close approximation to it.

In some studies there are several items under considerations and since the allocation that is best for one item will not in general be best for another, some compromise must be reached in a survey with numerous items. If the correlation among the items is high, then optimum allocations may differ relatively little. If however this is not the case, it may be desireble to adopt

proportional allocation.

1.4 NUMERICAL ILLUSTRATION WITH SPECIAL REFERENCE TO PLANNING OF SURVEYS ON LIME CROPS

For the purpose of illustration, we shall consider the problem of stratification in surveys designed to obtain reliable estimates of the extent of cultivation, yield rate, and total production of lime. The only information available for this purpose is village-wise area under lime as reported by the revenue agency for the year 1961-62 for all the 304 villages in the tract under survey. Since the area under lime as reported by the revenue agency is expected to have high correlation with the actual area, it is desirable to study it with special reference to planning of future surveys.

Using the data on reported area under lime we shall first consider the problem of determining the optimum number of strata under Neyman and proportional allocations. For this purpose we shall obtain variances of the estimated mean or total for different number of strata when the strata boundaries are determined in an optimum fashion by using the approximations suggested by Dalenius and Hodges (1959) and Sethi (1963) under Neyman and proportional allocations respectively. The tables below give the percentage points of stratification and the corresponding variances for various values of k under Neyman and proportional allocations.

TABLE II. Percentage points of stratification and corresponding variances under Neyman allocation for various values of K.

K	% points of Stratification	nV(\bar{Y}_{st})
3	0-10 10-30 30-100	: 32.4
4	0- 5 5-15 15- 35 35-100	: 21.0
5	0- 5 5-10 10-25 25- 45 45-100	: 17.3
6	0- 5 5-10 10- 15 15- 30 30-45 45-100	: 15.1

TABLE III. Percentage points of stratification and corresponding variances under proportional allocation for various values of K.

K	% points of Stratification						nV (\bar{Y}_{st})
3	0-10	10-40	40-100				37.8
4	0-10	10-20	20-50	50-100			24.6
5	0-5	5-15	15-30	30-50	50-100		19.4
6	0-5	5-10	10-20	20-30	30-50	50-100	16.7

It is clearly seen from the above tables that the variance of the estimate decreases with increasing number of strata. The reduction in variance is considerable as K increases from 3 to 4 . However, the reduction in variance is hardly appreciable for values of K exceeding 5. Thus it does not pay to have more than five strata that one should have for estimating the area under lime.

Next, we shall consider the problem of formation of strata. Having seen that about four or five is the optimum value of the number of strata, one should have and from other considerations we shall assume that it has been decided to form only four strata. Using the data on reported area under lime, the optimum points of stratification as determined by using the various procedures outlined in section 1.1 with corresponding variances are given in the table below:

TABLE IV. OPTIMUM POINTS OF STRATIFICATION.

Type of allocation	Optimum Strata Boundaries				nV (\bar{Y}_{st})
1. Proportional allocation	0-10%	10-20%	20-50%	50-100%	24.6
2. Neyman Allocation Hansen-Hurwitz- Madow approximation	0-10%	10-20%	20-35%	35-100%	19.7
3. Neyman Allocation Dalenius-Hodges Approximation	0- 5%	5-15%	15-35%	35-100%	21.0
4. Neyman Allocation Ekman's approximation	0-10%	10-20%	20-40%	40-100%	17.5
5. Neyman Allocation Durbin's approximation	0- 5%	5-20%	20-50%	50-100%	20.8

As is to be expected, the optimum strata as determined in the case of Neyman allocation always give a more efficient estimate than in the case of proportional allocation. Among the four approximations used for determining the optimum strata under Neyman allocation, the one suggested by Ekman's seems to be the most efficient one. Since however, the strata have been determined on the basis of area under lime as reported in 1961-62 and the distribution of area is likely to change, the question arises whether the stratification as determined on the basis of reported area under lime will actually result in substantial gains. For this purpose we will utilize the data collected in pilot sample survey on lime carried out in Nellore District, Andhra Pradesh, during the year 1963-64. In this survey, a sample of 86 villages was selected to determine the extent of cultivation of lime. Post stratification was carried out and estimates of variances are obtained for

- (i) Optimum strata determined under proportional allocation
- (ii) Optimum strata determined under Neyman allocation using Ekman's approximation and
- (iii) Strata determined on the basis of geographical proximity.

These results are given below:

	n Est $V(\bar{y}_{st})$
(i) Optimum strata determined under proportional allocation:	978,650
(ii) Strata determined on the basis of geographical proximity and proportional allocation:	3,337,832
(iii) Optimum strata determined under Neyman allocation:	596,447
(iv) Strata determined on the basis of geographical proximity and Neyman Allocation:	2,223,677

It is observed that the use of information on reported area under lime in forming strata has proved efficient. The optimum stratification as determined on the basis of reported area under lime has resulted in substantial gains over the stratification based on geographical proximity for both Neyman and proportional allocations. The gains are of the order of more than three hundred percent. The percentage efficiency of optimum stratification over stratification based on geographical proximity is 373 under Neyman allocation and 341 under proportional allocation. Therefore, it follows that if practical considerations permit, it is worthwhile to stratify the population of villages on the basis of the reported area under lime for obtaining more reliable estimates of the extent of cultivation of lime.

Next we consider the gain due to stratification based on geographical proximity for estimating the average lime yield and the total lime production. The relevant formulae for estimating the gain due to stratification for estimating the average yield and the total production are given in the last chapter. It is found that the percentage gain in efficiency due to stratification for estimating the average yield is 33.45 while the percentage gain in efficiency due to stratification for estimating the total production is 47.03. This shows that the device of stratification has proved efficient. The gains are more if the basis for stratification is the information on some correlated character instead of geographical proximity. The gain due to stratification is more for estimating the total production as compared to the gain due to stratification for estimating the average yield. The present survey covered only a district where we cannot expect much variation in soil and climatic conditions

even than the gain due to stratification is appreciable and we can expect still higher gain due to stratification for large scale surveys.

Having determined the optimum number of strata and the best way of determining strata boundaries, we shall consider the problem of allocation of sample sized to the different strata with a view to estimate the total production and the number of bearing trees with maximum precision. For the purpose of this study, we shall consider the following allocations;

(i) allocation as adopted in the survey. In the survey the allocation of sampling units to different strata was made roughly in proportional to the reported area under lime in the different strata subject to a minimum of four units per stratum.

(ii) Proportional allocation: $n_t \sim N_t$, where N_t is the total no. of villages is the t-th stratum.

(iii) Proportional allocation: $n_t \sim A_t$, A_t being the area under lime in the t-th stratum

(iv) Neyman allocation : $n_t \sim N_t \sigma_t$, σ_t being the standard deviation of the t-th stratum.

In each of these cases, the actual allocation was determined by using the data on the auxiliary variable, reported area under lime, instead of the actual variable under study. Table (V) below gives the allocation of sample size n among the different strata for different types of allocation: TABLE (V)

Stratum	Allocation as adopted in the survey	Proportional allocation $n_t \sim N_t$	Proportional allocation $n_t \sim A_t$	Neyman allocation $n_t \sim N_t \sigma_t$
1.	16	21	22	20
2.	10	10	11	10
3.	5	4	6	8
4.	7	4	1	1
5.	7	6	5	6
Total	45	45	45	45

For estimating the total production we shall consider the ratio type estimate, using the number of bearing trees as the auxiliary variable, Goswami (1961), while for the number of bearing trees we shall take the simple unbiased estimate for the study of allocation of sample sizes to the different strata. The estimates of the variances of the estimated total production and the number of bearing trees say $\hat{V}(T)$ and $\hat{V}(B_n)$ respectively are obtained from the lime survey data for the year 1963-64 for various allocation procedures.

Table VI below gives estimates of variances of the estimates of the total production and the number of bearing trees under various allocations and their percentage efficiencies over actual allocation.

Allocation	$\hat{V}(T)$ in 10^6	% Efficiency over actual allocation	$\hat{V}(B_n)$	% Efficiency
Actual	2,090,122	100	434,198,683	100
Proportional to N_t	1,753,135	119	366,435,200	119
Proportional to A_t	1,652,523	127	343,739,218	126
Neyman	1,782,190	117	370,395,137	117

It is seen from the above table that the Neyman allocation determined by using the data on the auxiliary variable reported under lime, instead of the actual variable under study has not proved efficient over proportional allocation for the estimation of the total production and the number of bearing trees. The percentage efficiency of Neyman allocation determined by $n_t \sim N_t \sigma_t$ over actual allocation is only 117 for both the cases. On the other hand percentage efficiency of proportional allocation is slightly higher. Allocation determined by $n_t \sim A_t$ seems

to be the most efficient in both the cases and its efficiency over actual allocation is about 127.

Next we shall consider the Neyman allocation based on the estimates of strata standard deviations available from the previous survey. It was found that the Neyman allocation for the study of yield hardly differed from that for the number of bearing trees. It follows that in surveys to estimate the number of bearing trees and the total production, it will suffice to adopt the Neyman allocation based on the number of bearing trees since this allocation will also be optimum for estimating the total production. More-over the cost involved in estimating additional villages to determine the number of bearing trees is much less as compared to the cost required for determining the total production. Thus more efficient estimates of strata standard deviations for the character 'the number of bearing trees' can be obtained by taking a larger sample for complete enumeration of villages and thereby obtaining a more efficient allocation for the yield study.

Finally we shall find the percentage efficiency of these allocations over the actual one both for the study of yield and the number of bearing trees. These are given below:

<u>Character under study</u>	<u>% Efficiency</u>
(i) Total Production.	139
(ii) Total number of bearing trees.	136

It will be seen that the use of estimated standard deviations for the number of bearing trees available from the previous survey in determining the allocation of the sample size to the different strata has resulted in further gain in precision in the case of both the characters, the gain in precision being about 10 to 12%.

C H A P T E R III

M E T H O D O F S A M P L I N G

3.1 INTRODUCTION:

In fruit surveys, the units which suggest themselves are a village, an orchard or a cluster of trees. Direct selection of trees may not be possible due to lack of a sampling frame giving the list of all the sampling units with proper identification particulars. A list of all the bearing trees in the population under study may not be available, whereas a list of orchards or villages may be easily available. Direct selection of orchards may be scattered all over the region increasing considerably the expenditure on travel. Because of increased travel it may not be possible for the field assistants to record the yield of all the selected orchards. Also it may not be the most efficient procedure to select orchards or trees directly, especially when data concerning the character under study or some related characters is available at the village level. This is especially the case in Andhra Pradesh where data concerning reported area under fruits is available at the village level from the revenue records. A list of villages growing fruits is also readily available for each taluka or district. The villages as the primary unit of sampling, therefore, appears to be the best solution and also convenient from the practical point of view.

Since village as the primary unit of sampling seems to be most suitable from various considerations and it is neither practicable nor worthwhile to record the yield of the entire village from the view point of cost considerations. It is thus essential to adopt a multistage sampling design.

Various possibilities suggest themselves. One possibility is to select a certain number of orchards on each of the selected villages and observe them for the purpose of yield study. This is however hardly a practical proposition considering the size of the orchard and the resources available by way of field staff to record the yield of the crop. It therefore seems absolutely necessary to resort to sub-sampling even at the orchard level. The design that suggests itself naturally is therefore a three stage design with village as the primary unit of sampling, orchard as the secondary unit of sampling and a cluster of trees as the ultimate unit of sampling.

Alternatively, we can select clusters of trees directly from each of the selected villages, thus resulting in a two stage design with village as the primary unit of sampling and a cluster of trees as the ultimate unit of sampling. We shall therefore consider two sub-sampling designs one with villages as the primary unit of sampling, orchards as the secondary unit of sampling and a cluster of trees as the ultimate unit of sampling and the other with villages as the primary unit of sampling and a cluster of trees as the secondary and ultimate unit of sampling. The former is desirable on account of operational convenience. It is also likely to be more efficient than the latter from considerations of cost although may not be efficient from the point of view of precision if cost is not taken into consideration.

In order to discuss the relative efficiency of the two designs, we shall once again consider the data collected in the lime survey conducted in Nellore district, Andhra Pradesh and estimate the sampling variances of the estimated yield rate and

the total production of lime. As already mentioned before, the design adopted for the lime survey was stratified-multistage random sampling. We shall consider the data collected in one stratum only, namely Venkatagiri. In this stratum, the villages were selected with probabilities proportional to reported area under lime in a previous year while the orchards and clusters of trees were selected with equal probability and without replacement.

3.2 TWO STAGE SUB-SAMPLING AGAINST THREE STAGE SUB-SAMPLING

Assume that a sample of n villages is selected out of N with replacement and with probabilities p_i for the purpose of observing the characters under study while an additional sample of $(n-n)$ villages is selected with the same scheme for the purpose of complete enumeration of trees. From each of the n selected villages a sample of m_j orchards is selected at random while from each of the selected orchards, a sample of b_{ij} trees is selected at random for the purpose of yield study.

Let

N = total number of villages in the stratum

M_i = total number of orchards in the i -th village

($i=1, 2, \dots, N$)

B_{ij} = total number of bearing trees in the j -th orchard
of the i -th village ($j=1, 2, \dots, M_i$; $i=1, 2, \dots, N$)

$$B_i = \sum_{j=1}^{M_i} B_{ij}; \bar{B}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} B_{ij}$$

p_i = the probability of selecting i -th primary unit at
each draw ($i=1, 2, \dots, N$; $\sum_{i=1}^N p_i = 1$)

$$U_{ij} = B_{ij} / \bar{B}_i.$$

y_{ijk} = the value of the character under study for the
 k -th tree of the j -th orchard in the i -th village
($k=1, 2, \dots, B_{ij}$)

$$y_{ij} = \sum_{k=1}^{B_{ij}} y_{ijk}$$

$$\bar{y}_{ij.} = y_{ij.} / B_{ij}$$

$$y_{i..} = \sum_{j=1}^{m_i} \sum_{k=1}^{B_{ij}} y_{ijk}$$

$$\bar{y}_{i..} = y_{i..} / B_i$$

$$y_{\dots} = \sum_{i=1}^n y_{i\dots}$$

$$\bar{y}_i(M_i) = y_{i..} / M_i$$

$$y_i(M_i) = \frac{1}{M_i} \sum_{j=1}^{m_i} \bar{y}_{ij}.$$

$$\bar{y}_{ij} (b_{ij}) = \frac{1}{b_{ij}} \sum_k b_{kj} y_{ijk} ; \quad b_i = \sum_j b_{ij}$$

Estimation of Total Production.

It can now be shown that an unbiased estimate of total production is given by

while its variance is given by

$$V(u) = \frac{1}{n} \sigma_y^2 + \frac{1}{n} \sum_{i=1}^n \frac{B_i^2}{p_i} \left\{ \left(\frac{1}{m_i} - \frac{1}{m} \right) S_{iy}^2 + \frac{m_i}{m} \sum_{j=1}^{m_i} u_{ij}^2 \left(\frac{1}{c_{ij}} - \frac{1}{B_{ij}} \right) S_{yj}^2 \right\} \quad .2$$

where $\hat{\sigma}_j^2 = \sum_{i=1}^n p_i \left(\frac{y_{ei}}{p_i} - y \dots \right)^2$

$$S_{iy}^2 = \frac{1}{M_{i-1}} \sum_{j=1}^{M_i} (u_{ij} - \bar{y}_{ij} - \bar{y}_{...})^2$$

$$S_{ijy}^2 = \frac{1}{B_{ijk}-1} \sum_{k=1}^{B_{ijk}} (y_{ijk} - \bar{y}_{ij.})^2$$

Defining $s_{\bar{y}}^2 = \frac{1}{n-1} \sum_i^n \left(\frac{m_i}{\rho_i} \frac{1}{m_i} \sum_j m_i B_{ij} \bar{Y}_{ij}(e_{ij}) - u \right)^2$

it can be shown that an unbased estimate of the variance is given by

If the sample is selected in two stages then we shall assume for the sake of comparison that an equivalent number of trees, namely $b_i = \sum_j^m b_{ij}$, is selected at random from the i -th selected village. Clearly an unbiased estimate of the total production

is given by

$$u' = \frac{1}{n} \sum_i^n \frac{B_i}{p_i} \sum_j^m \sum_k^r y_{ijk} \quad \dots \dots \dots 5$$

while its sampling variance is given by

$$V(u') = \frac{\sigma_y^2}{n} + \frac{1}{n} \sum_i^n \frac{B_i}{p_i} \left(\frac{1}{B_i} - \frac{1}{B_{i-1}} \right) s_{iy}^2 \quad \dots \dots \dots 6$$

$$\text{Where } s_{iy}^2 = \frac{1}{B_{i-1}} \sum_j^m \sum_k^r (y_{ijk} - \bar{y}_{ij..})^2 \quad \dots \dots \dots 7$$

Defining

$$s_{iy}^2 = \frac{1}{m_{i-1}} \sum_j^m \{ u_{ij} \bar{y}_{ij} (b_{ij}), - \sum_j^m u_{ij} \bar{y}_{ij} (b_{ij}) \}^2 \quad \dots \dots \dots 8$$

$$s_{ijy}^2 = \frac{1}{b_{ij-1}} \sum_k^r (y_{ijk} - \bar{y}_{ij} (b_{ij}))^2 \quad \dots \dots \dots 9$$

$$\text{so that } E s_{ijy}^2 = s_{ijy}^2 \quad \dots \dots \dots 10$$

$$\& E s_{iy}^2 = s_{iy}^2 + \frac{1}{M_i} \sum_j^m u_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \quad \dots \dots \dots 11$$

Obviously

$$\text{Est } s_{ijy}^2 = s_{ijy}^2$$

$$\text{Est } s_{iy}^2 = s_{iy}^2 - \frac{1}{m_i} \sum_j^m u_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \quad \dots \dots \dots 13$$

An unbiased estimate of σ_y^2 is given by

$$\text{Est } \sigma_y^2 = \sigma_y^2 - \frac{1}{n} \sum_i^n \frac{B_i^2}{p_i^2} \left\{ \left(\frac{1}{m_i} - \frac{1}{M_i} \right) s_{iy}^2 + \frac{1}{m_i M_i} \sum_j^m u_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \right\}$$

To estimate the 2nd component of $V(u')$, we first

observe that

$$(B_{i-1} - 1) s_{iy}^2 = \sum_{j=1}^{m_i} (B_{ij} - 1) s_{ijy}^2 + \sum_{j=1}^{m_i} B_{ij} (\bar{y}_{ij..} - \bar{y}_{i..})^2$$

$$= \sum_{j=1}^{m_i} (B_{ij-1}) s_{ijy}^2 + \sum_{j=1}^{m_i} B_{ij} \bar{y}_{ij..}^2 - B_{i..} \bar{y}_{i..}^2 \quad \dots \dots \dots 15$$

And

$$\text{Est } \bar{y}_{ij..}^2 = \bar{y}_{ij}^2 (b_{ij}) - \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \quad \dots \dots \dots 16$$

$$\text{Est } \bar{y}_{i..}^2 = \text{Est} \left\{ \frac{1}{M_i} \left\{ \sum_{j=1}^{m_i} u_{ij}^2 \bar{y}_{ij}^2 - (M_{i-1}) s_{iy}^2 \right\} \right\}$$

$$\frac{1}{n} \sum_j m_j^2 = \frac{1}{n} \sum_j b_{ij}^2 = \left(1 - \frac{1}{n} \right) \frac{1}{n} \\ = \frac{1}{n} \sum_j m_j^2 \left(\frac{1}{n} - \frac{1}{n^2} \right) \dots 17$$

$i_j = i - 1 + c - 2$

$$i_j = \frac{1}{n-1} \left[\frac{1}{n} \left(1 - \frac{1}{n} \right)^2 - \frac{1}{n} \sum_j m_j^2 \bar{y}_{ij}^2 \right] \\ - \frac{1}{n} \sum_j i_j^2 - i_j (i-1) - \frac{1}{n} \sum_j \left(1 - \frac{1}{n} \right) i_j \bar{y}_{ij}^2 \\ + \frac{1}{n} \sum_j m_j^2 \left(\frac{1}{n} - \frac{1}{n^2} \right) \dots 18$$

$$n - 2 \cdot i_j = \frac{1}{n} \sigma_y^2 + \frac{1}{n} \sum_i \frac{i_j}{n^2} \left(\frac{1}{n} - \frac{1}{n^2} \right) \dots 19$$

$$r_{ij} = \frac{\sigma_y^2}{n} + \frac{n}{4} \dots 14 \quad 17 \text{ no } \dots$$

obtaining $r_{ij} = r_{ij} + r_{ij} + r_{ij} + r_{ij}$

$$= \frac{1}{n} \sum_i r_{ij} + \frac{1}{n} \sum_i r_{ij} + \dots$$

$$= \frac{1}{n} \sum_i \frac{1}{x_i} - \frac{1}{n} \sum_j m_j^2 - \frac{1}{n} \sum_i \frac{1}{x_i} \dots 20$$

$$= \frac{1}{n} \sum_i \frac{1}{x_i} - \frac{1}{n} \sum_j m_j^2$$

$$= \frac{1}{n} \sum_i \frac{1}{x_i} \dots 21$$

$$\bar{v}(i) \cong v(i) + \sum_{j \neq i} \left\{ \bar{v}(v_j) - \bar{v}(v_i) \right\} - 2 \sum_{j \neq i} \left\{ \text{cov}(u, v_j) - \text{cov}(u, v_i) \right\} \dots 22$$

To correct for $\bar{v}(i)$

$$\text{Defn } s_p^2 = \frac{1}{n-1} \sum_i \left(\frac{\bar{v}_i}{p_i} - \frac{1}{n} \sum_i \frac{\bar{v}_i}{p_i} \right)^2$$

$$s^2 = \frac{1}{n-1} \sum_i \left(\frac{\bar{v}_i}{p_i} - \frac{1}{n} \sum_j \bar{v}_{ij} - \bar{v} \right)^2$$

$$s_{EJ}^2 = \frac{1}{n-1} \sum_i \left(\frac{\bar{v}_i}{p_i} - \frac{1}{n} \sum_j \bar{v}_{ij} - \bar{v}_{ij} b_{ij} - u \left(\frac{\bar{v}_i}{p_i} - \frac{1}{n} \sum_i \frac{\bar{v}_i}{p_i} \right) \right)^2 \dots 23$$

$$s^2 = \frac{1}{n-1} \sum_i \left(\frac{\bar{v}_i}{p_i} - \frac{1}{m_i} \sum_j \bar{v}_{ij} - \bar{v}_{ij} b_{ij} - u \left(\frac{\bar{v}_i}{p_i} - \frac{1}{m_i} \sum_j \bar{v}_{ij} \right) \right)^2$$

It can thus be easily seen that unbiased estimate of $v(i)$

$$\cdot v(v_i); \text{cov}(u, v_i) \text{ and cov}(v, v_i) \text{ are provided by}$$

$$\text{Est } v(w) = \frac{s_p^2}{n}, \quad \text{Est } v(v_i) = \frac{s_{EJ}^2}{n}$$

$$\text{Est cov}(u, v_i) = \frac{s_{uv}}{n}$$

$$\text{Est cov}(u, v_i) = \frac{s_{uv}}{n}$$

to obtain $v(i)$

$$\text{Est } v(i) = \frac{u}{n}$$

Using the above results, a consistent estimate of the

variance of T is thus given by

$$\text{Est } v(T) = \frac{s_p^2}{n} + \left(\frac{u}{n} \right)^2 \left(\frac{s_{EJ}^2}{n} - \frac{s_{uv}^2}{n^2} \right) - \frac{2u}{n} \left(\frac{s_{uv}}{n} - \frac{s_{uv}}{n} \right) \dots 25$$

If we take $v(i)$ as the unbiased estimate of $v(i)$, then the corresponding estimate of the

$$\begin{aligned} v(i) &= \frac{1}{n} \sum_i \frac{\bar{v}_i}{p_i} - \frac{1}{n} \sum_i \frac{\bar{v}_i}{p_i} \sum_k \bar{v}_{ijk} - \frac{1}{n} \sum_i \frac{\bar{v}_i}{p_i} \\ &= \frac{u}{n} \times v' \left(\text{cov} \right) \dots 26 \end{aligned}$$

$$\text{where } u' = \frac{1}{n} \sum_i^B \frac{\beta_i}{p_i} \frac{1}{b_i} \sum_j^m \sum_k^r y_{ijk}, \quad v' = \frac{1}{n} \sum_i^B \frac{\beta_i}{p_i}$$

$$v(T) \cong v(u') + \bar{y}^2 \dots \{v(v') - v(w)\} - 2\bar{y} \dots \{\text{cov}(u', v') - \text{cov}(u', w)\} \dots 27$$

A consistent estimate of the variance of T' is given by

$$\text{Est } v(T') = \text{Est } v(u') + \left(\frac{u}{w}\right)^2 \left(\frac{1}{n} - \frac{1}{n'}\right) s_B^2 - \frac{2u}{w} \left(\frac{1}{n} - \frac{1}{n'}\right) s_y^2 \dots 28$$

where $\text{Est } v(u')$ is given by 19.

ESTIMATE OF AVERAGE YIELD PER FARM.

To estimate the average yield per tree, we shall consider the estimate

$$\bar{y} = \frac{\frac{1}{n} \sum_i^B \frac{\beta_i}{p_i} \frac{1}{b_i} \sum_j^m \sum_k^r y_{ijk}}{\frac{1}{n} \sum_i^B \frac{\beta_i}{p_i} \frac{1}{b_i} \sum_j^m \sum_k^r y_{ijk}} = \frac{u}{w} \dots 29$$

Using the result of example, we have

$$v(\bar{y}) \cong \frac{1}{B} \left\{ v(u) + \bar{y}^2 \dots v(w) - 2\bar{y} \dots \text{cov}(u, w) \right\} \dots 30$$

The estimate $v(\frac{u}{w})$ is given by

$$\text{Est } v\left(\frac{u}{w}\right) = \frac{1}{B} \left\{ \frac{1}{n} + \left(\frac{u}{w}\right)^2 \frac{s_B^2}{n} - \frac{2u}{w} \frac{s_y^2}{n} \right\} \dots 31$$

If the sample is selected without replacement, then the covariance estimate

of average yield per tree is given by

$$\bar{y} = \frac{\frac{1}{n} \sum_i^B \frac{\beta_i}{p_i} \frac{1}{b_i} \sum_j^m \sum_k^r y_{ijk}}{\frac{1}{n} \sum_i^B \frac{\beta_i}{p_i}}$$

and its variance is given by

$$v\left(\frac{u}{w}\right) \cong \frac{1}{B} \left\{ v(u') + \bar{y}^2 \dots v(w') - 2\bar{y} \dots \text{cov}(u', w') \right\}$$

This variance $v\left(\frac{u}{w}\right)$ may be estimated by

$$\text{Est } v\left(\frac{u}{w}\right) = \frac{1}{B} \left\{ 1 + v(u') + \left(\frac{u}{w}\right)^2 \frac{s_B^2}{n} - \frac{2u}{w} \frac{s_y^2}{n} \right\}$$

where $\text{Est } v(u')$ is given by (13).

3.3 ESTIMATES OF THE MEAN.

In this section, estimation of the two means is discussed

i.e., adjustment of the mean in a simple random sampling.

and the average yield per tree, with special reference to the data collected in Venkatagiri Taluka alone. Venkatagiri Taluka accounts for a fairly large portion of the population under study. A sample of 16 villages was selected for the study of yield purposes while a sample of 33 villages was selected for the complete enumeration of villages to record the data on the number of bearing trees. The relative efficiencies of the two sub-sampling designs for estimating the total production and yield per tree are obtained. The results are presented in the Table VII below, which gives the percentage relative efficiencies of various estimates of two stage sampling over three stage sampling.

TABLE VII

Estimate	Three-stage sampling estimated variance	Two-stage sampling estimated variance	
1. Total produc- tion simple unbiased estimate	$1,544,695 \times 10^6$	$1,035,680 \times 10^6$	149
2. Total produc- tion ratio estimate	$822,113 \times 10^6$	$398,661 \times 10^6$	206
3. Yield per tree ratio estimate	21.605	14.775	146

It is seen from the above table that the sampling variances of all the estimates considered here are considerably smaller in the case of two stage sub-sampling design as compared to those for a three stage sub-sampling design. The percentage relative efficiency of the former is 146 for the purpose of yield study

and is still higher for the estimates of the total production. For ratio-type estimate of the total production the gain in efficiency is 106%. As is to be expected, the two stage sampling design turns out to be efficient over the three stage design for the estimation of characteristics such as yield rate and the total production. Thus it seems that it will pay substantially from the point of view of precision if we adopt the two stage sampling design in preference to the three stage sampling design provided it is operationally convenient and does not substantially increase the cost of the survey.

The three-stage sampling design is likely to be more convenient from the operational point of view especially when yield data is to be collected. Also the cost factor is not likely to be substantially different in the two cases. Unless, therefore, there is appreciable gain in precision, it may be desirable to adopt a three-stage sampling. If however the gain in precision is very considerable as is the case here, it seems desirable to adopt a two-stage sampling design even though it may not be very convenient from operational point of view.

CHAPTER IV

PROBABILITY SCHEME OF SAMPLING

4.1 The probability scheme of sampling is an important aspect in the planning of surveys and has to be decided after taking into account all the available information. This information can be effectively used in sample surveys in several ways. It can be made use of in selection of units or in estimation or in both. Since a large unit contributes more to the population total or mean, it is natural to expect that a scheme of selection which gives more weight to larger units than to smaller units would provide more efficient estimates than simple random sampling. Sampling with probability proportional to size assigns a larger probability to the larger units of being included in the sample. This technique suggested by Hansen and Hurwitz (1943) has found its principle use in surveys which employ sub-sampling. For instance, in agricultural surveys, the sizes of the units differ markedly and the selection of sampling units in proportion to the size of unit has generally proved efficient. In the absence of knowledge of the actual sizes of the units, the value of a highly correlated variable are used as measures of the sizes of the units. In case when information on a highly correlated variable is available even the simple estimators of total are improved by the use of the information on the highly correlated variable.

Here we shall investigate the efficiency of adopting varying probability sampling at the first stage as compared to sampling with equal probability for estimating simultaneously several characters such as yield rate, total production, number of bearing trees etc.

4.2 EFFICIENCY OF VARYING PROBABILITY SAMPLING OVER SIMPLE RANDOM SAMPLING

Assume that a sample of n villages is selected out of N with replacement and with probabilities p_i for the purpose of observing the characters under study while an additional sample of $(n-n)$ villages is selected with the same scheme for the purpose of complete enumeration of villages to record data on the number of bearing trees and the area under lime. From each of the n selected villages a sample of m_i orchards is selected at random while from each of the selected orchards a sample of b_{ij} trees is selected at random for the purpose of yield study.

ESTIMATION OF NUMBER OF BEARING TREES:

It can now be shown that an unbiased estimate of the total number of bearing trees is given by

$$\hat{B} = \frac{1}{n} \sum_{i=1}^{n'} \frac{B_i}{p_i} \dots \dots \dots \dots \dots \quad 1$$

while its sampling variance is given by

$$V(\hat{B}) = \frac{\sigma_B^2}{n} \dots \dots \dots \dots \dots \quad 2$$

$$\text{where } \sigma_B^2 = \sum_{i=1}^{n'} p_i \left(\frac{B_i}{p_i} - \hat{B} \right)^2$$

Defining

$$\hat{s}_B^2 = \frac{1}{n'-1} \sum_{i=1}^{n'} \left(\frac{B_i}{p_i} - \hat{B} \right)^2 \dots \dots \dots \quad 3$$

it can be shown that an unbiased estimate of the variance

is given by

$$\text{Est } V(\hat{B}) = \frac{\hat{s}_B^2}{n} \dots \dots \dots \dots \dots \quad 4$$

If the sample is selected with equal probability and without replacement then we assume for the sake of comparison that an equivalent number of units are selected at random. The corresponding estimate takes the form

while its sampling variance is given by

$$V(\hat{B}_e) = N^2 \left(\frac{1}{n} - \frac{1}{N} \right) S_B^2 \dots \dots \dots \quad 6$$

$$\text{where } S_B^2 = \frac{1}{N-1} \sum_{i=1}^N (B_i - \bar{B}_N)^2$$

Now S_B^2 is to be estimated from an equivalent sample drawn with varying probability. For this we proceed as

$$(N-1) \quad S_B^2 = \sum_{i=1}^N B_i^2 - \frac{B^2}{N}$$

$$\text{and } \sigma_{\theta}^2 = \sum_{i=1}^n \frac{B_i^2}{p_i} - B^2$$

$$\text{So } (N-1) S_B^2 = \sum_{i=1}^N B_i^2 - \frac{1}{N} \sum_{i=1}^N \frac{B_i^2}{p_i} + \frac{\sigma_B^2}{N}$$

Clearly its unbiased estimate is given by

$$\text{Est } (N-1) \quad S_B^2 = \frac{1}{n'} \sum_{i=1}^{n'} \frac{B_{i.}^2}{p_{i.}} - \frac{1}{n} \sum_{j=1}^n \frac{B_{.j}^2}{N p_{.j}} + \frac{B^2}{N} \quad ...7$$

From 6 and 7 it follows that

$$Est\ V(\hat{B}_e) = N^2 \left(\frac{1}{N}, -\frac{1}{N} \right) \sum_{i=1}^N \left[\frac{1}{N} \sum_i \frac{B_i^2}{p_i} - \frac{1}{N} \sum_i \frac{B_i}{N p_i} + \frac{\bar{B}_e^2}{N} \right] \quad \dots \dots \quad 8$$

Hence an estimate of the relative gain in precision is given by

$$\frac{\hat{V}(\hat{B}_\theta) - \hat{V}(\hat{B})}{\hat{V}(\hat{B})} = \frac{N(N-n)}{N-1} \left[\frac{1}{n} \sum_i^n \frac{\hat{B}_{i\cdot}^2}{\hat{p}_i} - \frac{1}{n} \sum_i^n \frac{\hat{B}_{i\cdot}^2}{N\hat{p}_i^2} + \frac{\hat{\beta}_\theta^2}{N} \right] - \hat{\beta}_\theta^2 \quad \dots\dots 9$$

It is clear that the estimate \hat{B}_e of the number of bearing trees is not likely to be efficient since no use has been made of the available information on the area under lime. Naturally it is expected that area under lime will be correlated with the number of bearing trees and the use of this information is likely to result in increasing the precision of the estimate.

We will, therefore, consider the ratio estimate

$$\hat{B}_R = \frac{\frac{N}{n'} \sum_{i=1}^{n'} P_i}{\frac{N}{n'} \sum_{i=1}^{n'} A_i} \quad A = \frac{n'}{a_{n'}} \quad A \quad (\text{say}) \dots\dots\dots 10$$

where A_i = Area under line in the i -th stratum

A = total area under line in the stratum

$$\bar{A}_N = \frac{1}{n'} \sum_{i=1}^{n'} A_i$$

$$b_{n'} = \frac{N}{n'} \sum_{i=1}^{n'} P_i$$

$$a_{n'} = \frac{N}{n'} \sum_{i=1}^{n'} A_i$$

Using the results of large sample theory, we have

$$V(\hat{B}_R) \cong V(b_{n'}) - 2 \frac{B}{A} \text{Cov}(a_{n'}, b_{n'}) + \frac{B^2}{A^2} V(a_{n'}) \dots\dots\dots 11$$

$$n' V(a_{n'}) = \frac{\pi^2}{(n')} \left(\frac{1}{n'} - \frac{1}{N} \right) S_A^2 \quad \} \dots\dots\dots 12$$

$$\text{and } V(b_{n'}) = \frac{\pi^2}{(n')} \left(\frac{1}{n'} - \frac{1}{N} \right) S_P^2 \quad \}$$

$$\text{Cov}(a_{n'}, b_{n'}) = \frac{\pi^2}{(n')} \left(\frac{1}{n'} - \frac{1}{N} \right) S_{AB} \quad \}$$

$$\text{where } S_A^2 = \frac{1}{N-1} \sum_{i=1}^{N-1} (A_i - \bar{A}_N)^2$$

$$S_{AB} = \frac{1}{N-1} \sum_{i=1}^{N-1} (A_i - \bar{A}_N) (P_i - \bar{B}_N)$$

Substituting the expression for variance in covariance given

in (12) in (11) and simplifying, we obtain

$$V(\hat{B}_R) \cong N^2 \left(\frac{1}{n'} - \frac{1}{N} \right) (S_B^2 - 2 \frac{B}{A} S_{AB} + \frac{B^2}{A^2} S_A^2) \dots\dots\dots 13$$

To estimate $V(\hat{B}_R)$ we first observe that unbiased estimates

of S_B^2 ; S_A^2 and S_{AB} are given by

$$\begin{aligned}
 E_{S^2} S_n^2 &= \frac{1}{n-1} \left\{ \frac{1}{n} \sum_i^n \frac{x_i^2}{p_i} - \frac{1}{n} \sum_i^n \frac{x_i^2}{N p_i} + \frac{\sigma^2}{n} \right\} \\
 E_{S^2} S_i^2 &= \frac{1}{n-1} \left\{ \frac{1}{n} \sum_i^n x_i^2 - \frac{A}{n} \right\} \\
 \text{Let } S_{B_i}^2 &= \frac{A}{n-1} \left(\frac{1}{n} \sum_i^n x_i^2 - \frac{1}{n} \sum_i^n \frac{B_i}{N p_i} \right) \\
 \text{Also Let } \frac{B}{A} &= \frac{v}{A}
 \end{aligned}
 \quad \dots 14$$

Using these results, it follows that a consistent estimate of $V(\hat{B}_R)$ is given by

$$\begin{aligned}
 \text{Est } V(\hat{B}_R) &\cong \sigma^2 \left(\frac{1}{n} - \frac{1}{N} \right) \frac{1}{N-1} \left\{ \frac{1}{n} \sum_i^n \frac{B_i^2}{p_i} - \frac{1}{n} \sum_i^n \frac{x_i^2}{N p_i} + \frac{\sigma^2}{n} \right\} \\
 &+ \frac{\sigma^2}{A} \frac{1}{n} \sum_i^n x_i^2 + \frac{\sigma^2}{n} - 2 \frac{\sigma^2}{n} \sum_i^n B_i
 \end{aligned}
 \quad \dots 15$$

and it is a consistent estimate in probability of the true probabilities.

Similarly for \hat{B}_j , this is given by

$$\frac{E_{S^2} V(\hat{B}_R) - E_{S^2} V(\hat{B}_j)}{\text{Est } V(\hat{B}_j)} \quad \dots 16$$

Estimation of average yield per tree:

$$\begin{aligned}
 \text{To estimate the average yield per tree, we take another} \\
 \text{estimate} &= \frac{\frac{1}{n} \sum_i^n \frac{x_i}{p_i} - \frac{1}{n} \sum_j^m \bar{x}_{ij} (b_{ij})}{\frac{1}{n} \sum_i^n \frac{x_i}{p_i} - \frac{1}{n} \sum_j^m \bar{x}_{ij}} \quad \dots 17
 \end{aligned}$$

variance $V(\hat{y})$ is a function of $V(\hat{B}_j)$ and is given by the previous equation (30 or 31). If the yield is let y is equal probability + the error term = c , then the variance of \hat{y} is $c^2 + V(\hat{B}_j)$.

\sum_i

$$\frac{\frac{n}{r} \sum_i^r \frac{1}{m_i} \sum_j^{m_i} \bar{x}_{ij} (\bar{u}_{ij})}{\frac{n}{r} \sum_i^r \frac{1}{m_i} \sum_j^{m_i} \bar{x}_{ij}} = \frac{1}{n} \sum_i^r \bar{x}_i \dots 1$$

$$w_e = \frac{n}{r} \sum_i^r \frac{m_i}{m} \sum_j^{m_i} B_{ij}$$

$$w_{ij} = \frac{1}{n} \sum_i^r \frac{1}{m_i} \sum_j^{m_i} \bar{x}_{ij} \bar{y}_{ij} (e_y)$$

$$w(\bar{x}) \approx \frac{1}{n} \left\{ w(u) - \dots w(v(1, \dots)) - \dots w(v(r, \dots)) \right\} \dots 10$$

$$\text{re } V(u) = \left\{ \left(\frac{1}{n} - \frac{1}{r} \right) \bar{x}_i + \frac{1}{n} \sum_{i=1}^r \bar{x}_i \left(\frac{1}{n} - \frac{1}{M_i} \right) \bar{s}_{iy} \right\} \\ + \frac{1}{r} \sum_{i=1}^r \frac{1}{m_i} \sum_{j=1}^{m_i} \bar{x}_{ij} \left(\frac{1}{r} - \frac{1}{m_i} \right) \bar{s}_{ij} \dots \\ w(v) = r \left\{ \left(1 - \frac{1}{r} \right) \bar{x}_i + \frac{1}{r} \sum_{i=1}^r \bar{x}_i \left(\frac{1}{n} - \frac{1}{M_i} \right) \bar{s}_{iy} \right\} \\ w(v(1, \dots)) = r \left\{ \left(1 - \frac{1}{r} \right) \bar{x}_i + \frac{1}{r} \sum_{i=1}^r \bar{x}_i \left(\frac{1}{n} - \frac{1}{M_i} \right) \bar{s}_{iy} \right\} \dots 21$$

$$w(v(r, \dots)) = \frac{1}{r-1} \sum_{i=1}^r (\bar{x}_{ij} - \bar{x}_i)^2$$

$$= \frac{1}{r-1} \sum_{j=1}^{m_i} (\bar{x}_{ij} - \bar{x}_i)^2$$

$$\bar{s}_{iy} = \frac{1}{r-1} \sum_{i=1}^r (\bar{x}_{ij} - \bar{x}_i) (\bar{x}_{ij} - \bar{x}_i)$$

$$S_{iy} = \frac{1}{r-1} \sum_{j=1}^{m_i} (\bar{x}_{ij} - \bar{y}_{i(m_i)})(\bar{x}_{ij} - \bar{x}_i)$$

$$S_{ij} = \frac{1}{M_i-1} \sum_{i=1}^r (\bar{x}_{ij} - \bar{x}_i)^2$$

$$S_{ij}^2 = \frac{1}{M_i-1} \sum_{j=1}^{m_i} (\bar{x}_{ij} - \bar{x}_i)^2; S_{ij}^2 = \frac{1}{M_i-1} \sum_{k=1}^{m_i} (\bar{x}_{ij} - \bar{x}_i)^2$$

substituting and simplifying

$$\begin{aligned}
 v(\frac{\Delta}{y})_e &\cong \frac{n^2}{B^2} \left\{ \left(\frac{1}{n} - \frac{1}{r} \right) (s_y^2 + \bar{y}^2 \dots s_n^2 - 2\bar{y} \dots s_{yB}^2) \right. \\
 &+ \frac{1}{m_i} \sum_{j=1}^{m_i} r_i^2 \left(\frac{1}{r_i} - \frac{1}{b_{ij}} \right) (s_{ijy}^2 + \bar{y}^2 \dots s_{iB}^2 - 2\bar{y} \dots s_{iyB}^2) \\
 &+ \left. \frac{1}{n^2} \sum_{i=1}^n \frac{r_i}{m_i} \sum_{j=1}^{m_i} s_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \right\} \dots \dots \dots 22
 \end{aligned}$$

For the estimation of the variance of \hat{Y}_e , we define

$$\begin{aligned}
 s_{iy}^2 &= \frac{1}{m_{i-1}} \sum_j^m \left(s_{ij} \bar{y}_{ij(bij)} - \frac{1}{m_i} \sum_j^m s_{ij} \bar{y}_{ij(bij)} \right)^2 \quad \dots \dots \dots 23 \\
 s_{iyB}^2 &= \frac{1}{m_{i-1}} \sum_j^m \left(s_{ij} \bar{y}_{ij(bij)} - \frac{1}{m_i} \sum_j^m s_{ij} \bar{y}_{ij(bij)} \right) \left(s_{ij} - \frac{1}{m_i} \sum_j^m s_{ij} \right)^2 \\
 s_{iB}^2 &= \frac{1}{m_{i-1}} \sum_j^m \left(s_{ij} - \frac{1}{m_i} \sum_j^m s_{ij} \right)^2
 \end{aligned}$$

in a ratio to those s_{ijy}^2 ; s_y^2 ; s_{yB}^2 ; s_B^2 ; & s_B^2

defined in previous chapter.

Then it can be seen that

$$\bar{s}_{iy}^2 = s_{iy}^2 + \frac{1}{m_i} \sum_{j=1}^{m_i} s_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2$$

$$E s_{iy}^2 = s_{iyB}^2$$

$$\bar{s}_{iB}^2 = s_{iB}^2$$

It follows that unbiased estimates of s_{ijy}^2 ; s_{iy}^2

$$s_{iy^n}^2; s_{iy}^2; s_{iy}^2; s_{yB}^2; s_B^2; \bar{s}_{iy}^2; \bar{s}_{iy}^2; \text{ & } \bar{s}_{iyB}^2 \text{ are given by}$$

$$+ s_{ijy}^2 = s_{ijy}^2$$

$$\text{Est } \hat{s}_{ij}^2 = \frac{1}{n} \sum_{j=1}^{m_i} \left(\frac{1}{r_{ij}} - \frac{1}{r_i} \right) \hat{s}_{ij}^2$$

$$\text{Est } \hat{s}_{ij}^2 = \hat{s}_{ij}^2$$

$$\text{Est } \hat{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ij}^2 + \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ijy}^2$$

$$\text{Est } \hat{\sigma}_{y'}^2 = \frac{1}{n} \sum_{i=1}^n \frac{r_i}{2} \left(\frac{1}{r_i} - \frac{1}{r_i} \right) \hat{s}_{ijy}^2$$

....24

$$\text{Est } \hat{s}_i^2 = \frac{1}{n} \sum_{j=1}^{m_i} \left(\frac{1}{r_{ij}} - \frac{1}{r_i} \right) \hat{s}_{ij}^2$$

$$\text{Est } \hat{s}_y^2 = \frac{\text{Est } \hat{s}_i^2}{n(n-1)} + \frac{1}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \right.$$

$$\left\{ \frac{1}{n} \sum_{i=1}^n \frac{r_i}{2} \sum_{j=1}^{m_i} \hat{s}_{ij}^2 - \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \right.$$

$$\left. \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ij}^2 + \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ijy}^2 \right\}$$

$$+ \text{Est } \hat{s}_y^2 = \frac{\text{Est } \hat{s}_i^2}{n(n-1)} + \frac{1}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \right.$$

$$\left. \left\{ \frac{r_i}{2} \sum_{j=1}^{m_i} \hat{s}_{ij}^2 - r_i \hat{s}_{ijy}^2 - \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ijy}^2 \right\} \right\}$$

$$\text{Est } \hat{s}_y^2 = \frac{\text{Est } \hat{s}_i^2}{n(n-1)} + \frac{1}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{r_i} - \frac{1}{r} \right) \right.$$

$$\left. \left\{ \frac{r_i}{2} \sum_{j=1}^{m_i} \hat{s}_{ij}^2 - r_i \left(\frac{1}{r_i} - \frac{1}{r} \right) \hat{s}_{ijy}^2 \right\} \right\}$$

" i, i bias

$$\text{Est } \bar{Y} = \frac{u}{w}$$

Using these estimates, it follows that

$$\begin{aligned} \text{Est } V\left(\frac{\hat{y}}{e}\right) &\approx \frac{n^2}{w^2} \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) \left(\text{Est } S_y^2 + \frac{u^2}{w^2} \text{Est } S_B^2 - \frac{2u}{w} \text{Est } S_{By} \right) \right. \\ &+ \frac{1}{n^2 N} \sum_i \frac{m_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{n_i} \right) \left(\text{Est } S_{iy}^2 + \frac{u^2}{w^2} \text{Est } S_{iB}^2 \right. \\ &- \frac{2u}{w} \text{Est } S_{iyB} + \frac{1}{n^2 N} \sum_i \frac{m_i^2}{p_i} \frac{1}{m_i^2} \sum_j p_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right), \\ &\quad \left. \text{Est } S_{ijy}^2 \right\} \end{aligned} \quad \dots\dots\dots 25$$

An estimate of the relative gain in precision of varying probability sampling over equal probability sampling is given by

$$\frac{\text{Est } V\left(\frac{\hat{y}}{e}\right) - \text{Est } V\left(\hat{y}\right)}{\text{Est } V\left(\hat{y}\right)} \quad \dots\dots\dots 26$$

Estimation of total production :-

We shall estimate the total production by using ratio type estimates, we shall take the estimate namely

$$T = \frac{\frac{1}{n} \sum_i \frac{m_i}{p_i} \frac{1}{m_i} \sum_j m_{ij} \bar{y}_{ij}(b_{ij})}{\frac{1}{n} \sum_i \frac{m_i}{p_i} \frac{1}{m_i} \sum_j B_{ij}} \quad \frac{1}{n'} \sum_i \frac{m_i}{p_i} \quad \dots\dots\dots 27$$

The variance estimate of the variance has already been given in the previous chapter. Now the corresponding estimate for equal probability sampling scheme will take the form

$$\begin{aligned}
 T_e &= \frac{\frac{N}{n} \sum_i \frac{l_i}{m_i} \sum_j b_{ij} \bar{y}_{ij} (bij)}{\frac{N}{n} \sum_i \frac{l_i}{m_i} \sum_j b_{ij}} \times \frac{\frac{N}{n} \sum_i l_i}{\dots 28} \\
 &= \frac{u_e}{b_n} \text{ say} \\
 \text{where } u_e &= \frac{N}{n} \sum_i \frac{l_i}{m_i} \sum_j b_{ij} \bar{y}_{ij} (bij) \\
 w_e &= \frac{N}{n} \sum_i \frac{l_i}{m_i} \sum_j b_{ij}
 \end{aligned}$$

and its variance is given by

$$\begin{aligned}
 v(T_e) &\cong v(u_e) + \bar{y}^2 v(w_e) - 2\bar{y} \text{Cov}(u_e, w_e) + 2\bar{y} \dots \text{Cov}(u_e, b_n) \\
 &\quad - \bar{y}^2 \dots v(b_n) \dots 29
 \end{aligned}$$

Substituting for various variance and covariance terms in 29

and simplifying, it follows that

$$\begin{aligned}
 v(T_e) &\cong l^2 \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) (s_y^2 + \bar{y}^2 \dots s_p^2 - 2\bar{y} \dots s_B) \right. \\
 &\quad + \left(\frac{1}{n} - \frac{1}{N} \right) (2\bar{y} \dots s_{yB} - \bar{y}^2 \dots s_B^2) \\
 &\quad + \frac{1}{nN} \sum_{i=1}^N \frac{l_i^2}{m_i} \left(\frac{1}{m_i} - \frac{1}{N} \right) (s_{iy}^2 + \bar{y}^2 \dots s_{iB}^2 - 2\bar{y} \dots s_{iyB}) \\
 &\quad \left. + \frac{1}{nN} \sum_{i=1}^N \frac{l_i^2}{m_i} \sum_{j=1}^{m_i} b_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) s_{ijy}^2 \right\} \dots 30
 \end{aligned}$$

and its estimate is given by

$$\begin{aligned}
 \text{Est } v(T_e) &= l^2 \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) (2s_y^2 s_{yB}^2 + \frac{u^2}{w^2} \text{Est } s_B^2 - 2\frac{u}{w} \text{Est } s_{yB}) \right. \\
 &\quad + \left(\frac{1}{n} - \frac{1}{N} \right) (\frac{2u}{w} \text{Est } s_{yB} - \frac{u^2}{w^2} \text{Est } s_B^2) + \frac{1}{n^2} \sum_i \frac{l_i^2}{m_i} \\
 &\quad \left. \left(\frac{1}{m_i} - \frac{1}{N} \right) (2\text{Est } s_{iy}^2 s_{iB}^2 + \frac{u^2}{w^2} \text{Est } s_{iB}^2 - 2\frac{u}{w} \text{Est } s_{iyB}) + \frac{1}{n^2} \sum_i \frac{l_i^2}{m_i} \right. \\
 &\quad \left. \left(\frac{1}{m_i} - \frac{1}{N} \right) (2\text{Est } s_{iy}^2 s_{yB}^2 + \frac{u^2}{w^2} \text{Est } s_{yB}^2 - 2\frac{u}{w} \text{Est } s_{iyB}) + \frac{1}{n^2} \sum_i \frac{l_i^2}{m_i} \right\}
 \end{aligned}$$

$$+ \frac{1}{N^2} \sum_i \frac{\sum_j p_{ij}^2}{p_i^2} \sum_j \left(\frac{1}{h_{ij}} - \frac{1}{s_{ij}} \right) \text{Est } s_{ijy}^2 \quad \} \dots 31$$

The estimate of the relative precision in precision of varying probability sampling over equal probability sampling is given by

$$\frac{\text{Est } V(T_e)}{\text{Est } V(T)}$$

4.3 NUMERICAL ILLUSTRATION.

For the purpose of illustration, we shall consider the problem of probability sampling in surveys designed to obtain reliable estimates of the extent of cultivation, yield rate and total production of lime. The only ancillary information available is village-wise area under lime as reported by the revenue agency for the year 1961-62 for all the 141 villages in Vetalgiri Taluka under study. These area figures were considered as measures of sizes of the units and with a view to obtain more reliable estimates, the villages were selected with probability proportional to area under lime as reported by the revenue agency for the year 1961-62. The table below gives the percentage relative efficiencies for various estimates in case of varying probability sampling over corresponding estimates in case of simple random sampling.

Table VIII.

Estimate	Est. Variance Varying Prob	Est. Variance Equal Prob ? without replacement	% Rel Efficiency
1. Total number of bearing trees		958,507,016	360
(i) Simple unbiased estimate	265,744,263		
(ii) Ratio estimate		244,608,674	92
2. Total production	$822,113 \times 10^6$	$1,576,584 \times 10^6$	192
3. Average Yield.	21.605	35.156	160

It is seen from the above table that for the purpose of estimation of the characteristics such as yield rate, and total production in surveys which employ sub-sampling designs, use of varying probability at the first stage proves to be efficient over the use of equal probability and without replacement. The gains in efficiency are substantially high and range from 60% in case of average yield to 92% in the case of total production. For estimating the number of bearing trees, it is seen that when simple unbiased estimates are used for both the systems of sampling, the percentage relative efficiency of varying probability sampling over equal probability sampling is 360. The two systems are however almost equally efficient if the ratio estimate is used. Also it is seen that in most of the cases the gains in precision are very considerable.

It seems that when information on some related variable

is readily available in advance and is known to be highly correlated with the character under study, as in the present case where the correlation coefficient between the reported area under lime and the number of bearing trees was found to be 0.63, it may be desirable to select the primary units with probability proportional to their size rather than selecting them with equal probability and without replacement. Further it suggests that it may not be advisable to use equal probability sampling when the character under study is either the yield or the total production since, as seen in this section, the gains in precision are very considerable for such characters. If the prime object of the survey is to estimate the number of bearing trees only, it may suffice if the units are selected with equal probability and without replacement. The accuracy in estimation of the related character is also used at the estimation stage instead of at the selection stage.

Normally, however, the object of such surveys is not merely to obtain reliable data on the extent of cultivation but also reliable information on characteristics such as yield rate and total production. Since these characteristics can be estimated more efficiently by using the ancillary information on the correlated character at the selection stage rather than at the estimation stage, there is hardly any loss of precision even in respect of extent of cultivation, it seems desirable to select villages with probability proportional to their size such as reported areas and the figures for the previous year if this information is available.

....

C. A P T A R I.

ESTIMATION OF AVERAGE YIELD.

5.1 INTRODUCTION.

So far we discussed the various important aspects which require thought and consideration in the planning of surveys, especially in the field of agriculture. It was seen that a sub-sample design in which the first stage units were selected with probabilities proportional to some measure of size of the units, had been found to bring about improvement in precision compared with sub-sampling systems using equal probabilities.

Having decided on the various aspects for the planning of agricultural surveys, it is necessary to evolve reliable estimation procedures for the estimation of the parameter like population mean & total of the character under study. If Y be the character under study and $\bar{Y} \dots$, the population mean for Y to be estimated we shall consider a two stages sampling design where primary units are selected with varying probabilities p_i ($\sum_{i=1}^N p_i = 1$) in replacement. The estimates of $\bar{Y} \dots$ will be considered in the present work as ratio-type estimates. All these estimates are found to be particular cases of a generalized ratio-type estimate of average yield for tree whose theory has been developed in this chapter. The estimate considered is of the form

$$\hat{\bar{Y}} \dots = \frac{\hat{Y}}{\hat{S}}$$

where the numerator represents the estimate of the total production considered by Ravindra Singh (1962) and the denominator represents the estimate of the total no. of bearing trees.

5.2. METRIC AND SAMPLE SCHEME.

Let

n = total number of villages in the population

i = total number of orchards in the i -th village

($i = 1, 2, \dots, n$)

B_{ij} = total number of trees in the j -th orchard of i -th village.

($j = 1, 2, \dots, k_i$; $i = 1, 2, \dots, n$)

y_{ijk} = the value of variable under study for the k -th tree of the j -th orchard in the i -th village.

($k = 1, 2, \dots, B_{ij}$)

Z_{ijk} = the value of auxiliary variable for the k -th tree of the j -th orchard in the i -th village

p_i = the probability of the selected i -th village at each draw. ($\sum_{i=1}^n p_i = 1$)

A sample of n villages is selected out of n with varying probabilities p_i ($i = 1, 2, \dots, n$) with replacement. Then we select m_i orchards out of k_i , in the i -th selected village, with equal probabilities and without replacement for the purpose of observing the number of bearing trees and retain r_i for observing both the character under study and the auxiliary variable. At the third stage, out of B_{ij} trees contained in the j -th selected orchard of the i -th selected village, we select ... 44..

b_{ij} trees with equal root heights and without replacement.

5.3 GENERALIZED RATIO-TYPE ESTIMATES OF THE AVERAGE YIELD PER TREE.

Depending upon the availability of ancillary information, several ratio-type estimates of the average yield per tree can be built up. We shall shortly see that all these estimates are particular cases of a generalized ratio-type estimate which may be defined as

$$\hat{Y}_{ij..} = \frac{\frac{1}{n} \sum_i \frac{\hat{y}_{ij..}}{p_i}}{\frac{1}{n} \sum_i \frac{\hat{B}_{ij..}}{p_i}} = \frac{\frac{1}{n} \sum_i \frac{a_i}{p_i} \left(\frac{\sum_{j=1}^{m_i} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{j=1}^{m_i} l_{ij} \bar{z}_{ij}(b_{ij})} \right)}{\frac{1}{n} \sum_i \frac{a_i}{p_i} \frac{1}{m'_i} \sum_j \hat{B}_{ij}}$$

$$= \frac{u}{w} \quad \text{say ... 1}$$

$$\text{where } \hat{Y}_{ij..} = \frac{\sum_{j=1}^{m_i} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{j=1}^{m_i} l_{ij} \bar{z}_{ij}(b_{ij})}; \quad \hat{B}_{ij..} = \frac{M_i}{m'_i} \sum_{j=1}^{m'_i} B_{ij};$$

$$u = \frac{1}{n} \sum_i \frac{\hat{y}_{ij..}}{p_i}, \quad w = \frac{1}{n} \sum_i \frac{\hat{B}_{ij..}}{p_i}$$

and a_i , t_{ij} and l_{ij} are some fixed constants.

By suitably choosing the constants and the values Z_{ijk} of the variable Z we can get the various estimates.

For example

Case (I) If $a_i = b_i$; $l_{ij} = 1$, $t_{ij} = B_{ij}$ and $Z_{ijk} = 1$ for all i, j, k we get the estimates

$$\hat{Y}_{ij..} = \frac{\frac{1}{n} \sum_i \frac{M_i}{p_i} \frac{1}{m_i} \sum_{j=1}^{m_i} B_{ij} \bar{y}_{ij}(b_{ij})}{\frac{1}{n} \sum_i \frac{b_i}{p_i} \sum_{j=1}^{m_i} \frac{1}{m_i} B_{ij}} \quad \text{if } a_i = b_i$$

$$\Delta_{12} = \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m B_{ij} \bar{y}_{ij}(b_{ij})}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i}} \quad \text{if } m'_i = h_i$$

Case II. If $a_i = r_i$; $l_{ij} = t_{ij} = 1$; $z_{ijk} = 1$ for all i, j, k
we get the estimates

$$\begin{aligned} \hat{y}_{21} &= \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m \bar{y}_{ij}(b_{ij})}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m B_{ij}} \quad \text{if } r'_i = r_i \\ \hat{y}_{22} &= \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m \bar{y}_{ij}(b_{ij})}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i}} \quad r'_i = u_i \end{aligned}$$

Case III. If $a_i = r_i$; $l_{ij} = t_{ij} = p_{ij}$; $z_{ijk} = 1$ for all i, j, k
we get the estimates

$$\begin{aligned} \hat{y}_{31} &= \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{\sum_j^m B_{ij} \bar{y}_{ij}(b_{ij})}{\sum_j^m r_{ij}}}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m r_{ij}} \quad \text{if } r'_i = n_i \\ \hat{y}_{32} &= \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{\sum_j^m r_{ij} \bar{y}_{ij}(b_{ij})}{\sum_j^m r_{ij}}}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i}} \quad \text{if } r'_i = M_i \end{aligned}$$

Case IV. If $a_i = r_i$; $t_{ij} = r_{ij}$; $l_{ij} = t_{ij}$; $z_{ijk} = 1$ for all i, j, k ,

we get the estimates

$$\begin{aligned} \hat{y}_{41} &= \frac{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{\sum_j^m r_{ij} \bar{y}_{ij}(b_{ij})}{\sum_j^m r_{ij}}}{\frac{1}{n} \sum_i^m \frac{r_i}{p_i} - \frac{1}{r_i} \sum_j^m r_{ij}} \quad \text{if } r'_i = r_i \end{aligned}$$

$$\hat{\Sigma}_{42} = \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \frac{\sum_{j=1}^{m_i} \bar{v}_{ij} v_{ij}(p_{ij})}{\sum_{j=1}^{m_i} \bar{v}_{ij}} / \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \dots \text{if } z_{ij} = 1$$

$\hat{\Sigma}_{42}$ is the i th estimate of the i th v_{ij} .

also we have $(\sum_{j=1}^{m_i} \bar{v}_{ij}) = z_{ij}$

$$z_{ij} = \bar{v}_{ij} \dots \text{if } z_{ij} = 1 \text{ or } z_{ij} = 0$$

$\hat{\Sigma}_{42}$ is the i th estimate of the i th v_{ij} .

case 2: if $p_i = 0$:

$$\hat{\Sigma}_{41} = \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \frac{\sum_{j=1}^{m_i} \bar{v}_{ij} v_{ij}(p_{ij})}{\sum_{j=1}^{m_i} \bar{v}_{ij}} / \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \sum_{j=1}^{m_i} \bar{v}_{ij} \dots \text{if } p_i = 0$$

$$\hat{\Sigma}_{43} = \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \frac{\sum_{j=1}^{m_i} \bar{v}_{ij} v_{ij}(p_{ij})}{\sum_{j=1}^{m_i} \bar{v}_{ij}} / \frac{1}{n} \sum_i^n \frac{z_{ij}}{p_i} \dots \text{if } p_i = 0$$

then $\hat{\Sigma}_{41}$ is the i th estimate of v_{ij} , $\hat{\Sigma}_{43}$ is the i th estimate of v_{ij} .

at this time $\hat{\Sigma}_{41}$ is the i th estimate of v_{ij} and $\hat{\Sigma}_{43}$ is the i th estimate of v_{ij} .

thus the estimates of v_{ij} are estimates of v_{ij} .

5.4 EXPLANATION: VARIANCE & ESTIMATED VARIANCE OF $\hat{\Sigma}_{41}$

$$\text{Let } u = v + e \dots$$

$$v^2 = v + e' \dots$$

$$v + e = E(e) = E(v') = 0$$

$$T(u, v) = v + \dots + u$$

$v + \dots + u$ is the i th estimate of v_{ij} by

$$\begin{aligned} u &= E(u) - y \dots = \sum_{i=1}^n R_i \left[1 + \left(\frac{1}{p_i} - \frac{1}{z_{ij}} \right) \left(\frac{s_{ij}^2}{p_i} - \frac{s_{ij}^2}{z_{ij}} \right) \right] \\ &\quad + \frac{1}{p_i} \sum_{j=1}^{m_i} \left(\frac{1}{p_i} - \frac{1}{z_{ij}} \right) \left(\frac{1}{p_i} - \frac{1}{z_{ij}} \right) \left(\frac{s_{ij}^2}{p_i} - \frac{s_{ij}^2}{z_{ij}} \right) \dots \end{aligned}$$

... 47..

$$= \sum_{i=1}^n U_i - \dots - y \\ e^{n+2} = \frac{G_{ij}}{\frac{n}{m} \bar{U}_{ij}}$$

$$G_{ij} = \left(\frac{1}{\frac{n}{m}} \sum_{j=1}^m t_{ij} - \bar{y}_{ij}(b_{ij}) \right)$$

$$\bar{U}_{ij} = \left(\frac{1}{\frac{n}{m}} \sum_{j=1}^m l_{ij} - \bar{y}_{ij}(b_{ij}) \right)$$

$$S_{ij} = \frac{1}{k-1} \sum_{j=1}^m (t_{ij} - \bar{y}_{ij}) \left(l_{ij} - \bar{y}_{ij} - \frac{1}{k} \sum_{j=1}^m l_{ij} \bar{y}_{ij} \right)$$

$$S_{ijyz} = \frac{1}{B_{ij}-1} \sum_{k=1}^{B_{ij}} (y_{ijk} - \bar{y}_{ij}) (z_{ijk} - \bar{z}_{ij})$$

$$S_{ij}^2 = \frac{1}{k-1} \sum_{j=1}^m (l_{ij} - \bar{l}_{ij})^2$$

$$S_{ijz}^2 = \frac{1}{B_{ij}-1} \sum_{k=1}^{B_{ij}} (z_{ijk} - \bar{z}_{ij})^2$$

Substituting in 2, $\beta_{ij} = 1 + \alpha_{ij} S_{ij}^2$ for all j in row i .

to n order we get $\beta = 0$.

$$\hat{y} = \frac{U + \alpha}{U + \alpha'} = \frac{U}{\beta} \left(1 + \frac{\alpha}{\beta} - \frac{\alpha'}{\beta} + \frac{\alpha'^2}{\beta^2} - \frac{\alpha \alpha'}{\beta^2} \right)$$

$$|\frac{\alpha}{\beta}| < 1 \text{ so } \beta \neq 0 \text{ if } |\alpha| < \beta$$

$$(1 + \frac{\alpha}{\beta})^{-1} \text{ is valid}$$

making exact iteration not like, so α

$$\hat{y}(\beta) = \frac{U}{\beta} \left(1 + \frac{\alpha}{\beta} - \frac{\alpha (\alpha')}{\beta^2} \right)$$

$\hat{y}_i = \hat{y}(\beta)$ is given by

$$\frac{\frac{1}{x_1} \dots \frac{1}{x_n}}{\dots} = \dots = \frac{\tau(\cdot)}{\pi} = \frac{u'(\cdot)}{\pi} = \dots$$

4. 4.

$$\tau(\cdot) = \sum_{i=1}^N x_i \left(\frac{1}{x_1} - \dots \right) + \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} \left(\frac{1}{x_1} - \dots \right) = \frac{1}{x_1} \dots$$

$$w_i = \frac{1}{x_1^{-1}} \sum_{j=1}^{m_i} (x_{i,j}^{-1})$$

4. 5.

$$w_{ij} = \frac{1}{x_1} \sum_{j=1}^{m_i} t_{ij} \bar{v}_{ij}(v_{ij})$$

$$w_{ij} = \frac{1}{x_1} \sum_{j=1}^{m_i} t_{ij} \bar{v}_{ij}(v_{ij})$$

$$w_{ij} = \frac{1}{x_1} \sum_{j=1}^{m_i} v_{ij}$$

4. 6.

$$\int \tau_v(u, w) = u \left(\frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} - \frac{x_1}{x_1 - x_i} \right), \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} v_i \neq 0$$

$$= u \left\{ \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} - \frac{x_1}{x_1 - x_i} \right\}^{1/2}, \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} v_i \neq 0$$

$$\left\{ \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} - \left(\frac{x_1}{x_1 - x_i}, v_{1,i} \right)^{1/2} \right\}.$$

$$\frac{1}{x_1} \overline{\sigma_{uv}} = \frac{1}{x_1} \sum_{i=1}^N \frac{1}{x_i} v_i \left(\frac{x_1}{x_1 - x_i}, v_{1,i} \right)^{1/2} \dots$$

$$\therefore \overline{\sigma_{uv}} = \sum_{i=1}^N \frac{1}{x_i} \left(\frac{x_1}{x_1 - x_i} - \tau \right) \left(- \frac{x_1}{x_1 - x_i} - \tau \right)$$

$$\text{Cov} \left(\frac{e_{im_i}}{h_{im_i}} \right), \frac{s_{im_i}}{v_{im_i}} = E \left(\frac{e_{im_i}}{h_{im_i}} \right) v_{im_i} - E \left(\frac{e_{im_i}}{h_{im_i}} \right) E \left(\frac{s_{im_i}}{v_{im_i}} \right)$$

$$\cong \frac{G_{iy}}{E_{iz}} \bar{s}_{i.} \left\{ \text{Cov} \left(\frac{e_{im_i}}{h_{im_i}}, \frac{s_{im_i}}{v_{im_i}} \right) - \frac{\text{Cov}(h_{im_i}, s_{im_i})}{E_{iz} \bar{s}_{i.}} \right\}$$

$$\cong \frac{1}{E_{iz}} \left\{ \left(\frac{1}{p_i} - \frac{1}{F_i} \right) (s_{ij} - R_i S_{izB}) \right\} \dots \dots \dots 9$$

$$s_{ijB} = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (t_{ij} - \bar{t}_{ij.}) \left(\bar{t}_{ij.} - \bar{s}_{i.} \right)$$

$$s_{izB} = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (t_{ij} - \bar{t}_{ij.}) \left(\bar{t}_{ij.} - \bar{s}_{i.} \right)$$

We see on substitution in 7 that

$$\text{Cov}(u, w) = \frac{\sigma_{uw}}{n} + \frac{1}{n} \sum_{i=1}^n \frac{a_i}{E_{iz}} \left\{ \left(\frac{1}{p_i} - \frac{1}{F_i} \right) (S_{iyB} - R_i S_{izB}) \right\} \dots \dots \dots 9$$

Substitution from 16 in 5, we find that the r.v. is 5 in the

e.r. to $\hat{y} \dots i.$ given by

$$y = \frac{U - y \dots}{\dots} + \frac{1}{n} \left(\frac{\sigma_y^2}{B^2} - \frac{\sigma_{uy}^2}{U^2} \right) + \frac{1}{n^2} \sum_{i=1}^n \frac{R_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{F_i} \right) s_{ijB}^2 - \frac{1}{n U^2} \sum_{i=1}^n \frac{a_i R_i}{p_i F_{iz}} \left\{ \left(\frac{1}{m_i} - \frac{1}{F_i} \right) (S_{iyB} - R_i S_{izB}) \right\} \dots \dots \dots 10$$

Further, we have

$$V(\hat{y} \dots) \cong \frac{1}{B} \left\{ V(u) + \gamma^2 I(u) - 2R \text{Cov}(u, w) \right\} \dots \dots \dots 11$$

$$\text{and } R = \frac{U}{B}$$

It can be seen that

$$V(u) = \frac{\sigma_u^2}{n} + \frac{1}{n} \sum_{i=1}^n \frac{a_i^2}{E_{iz}} \left\{ \left(\frac{1}{p_i} - \frac{1}{F_i} \right) s_i^2 + \frac{1}{F_i} \sum_{j=1}^{M_i} \left(\frac{1}{b_{ij}} - \frac{1}{v_{ij}} \right) s_{ij}^2 \right\} \dots \dots \dots 12$$

$$\text{where } S_i^2 = \frac{1}{l_i - 1} \sum_{j=1}^{m_i} \left\{ t_{ij} \bar{s}_{ij} - l_{ij} \bar{s}_{ij} - \frac{1}{l_i} \sum_{j=1}^{m_i} (t_{ij} \bar{s}_{ij} - l_{ij} \bar{s}_{ij}) \right\}^2$$

$$S_{ij}^2 = \frac{1}{l_{ij} - 1} \sum_{k=1}^{B_{ij}} \left\{ t_{ijk} \bar{s}_{ijk} - l_{ijk} \bar{s}_{ijk} - (t_{ij} \bar{s}_{ij} - l_{ij} R_i \bar{s}_{ij}) \right\}^2$$

" i ~ 12, " 6 in 11 & similarly for other indices.

$$V(\bar{y} \dots) = \frac{1}{n^2} \left(\frac{1}{n} (\sigma_u^2 + \sigma_B^2 - 2 \sigma_{ub}) \right)$$

$$\begin{aligned} &+ \frac{1}{n} \sum_{i=1}^n \frac{1}{l_i^2} \left\{ \left(\frac{1}{m_i} - \frac{1}{l_i} \right) S_i^2 + \frac{1}{l_i} \sum_{j=1}^{m_i} \left(\frac{1}{l_{ij}} - \frac{1}{l_i} \right) S_{ij}^2 \right\} \\ &+ \frac{R^2}{n} \sum_{i=1}^n \frac{1}{p_i} \left(\frac{1}{l_i} - \frac{1}{l_i} \right) S_i^2 \\ &- \frac{2R}{n} \sum_{i=1}^n \frac{1}{p_i} \left(\frac{1}{l_i} - \frac{1}{l_i} \right) (S_i - R_i S_{iB}) \end{aligned} \quad \dots \dots .13$$

Let the estimate of the variance of $V(\hat{y} \dots)$ be $\hat{V}(\hat{y} \dots)$.

estimate of $\bar{y} \dots$ can be improved by taking the estimate of $\hat{y} \dots$ = \hat{y} .

by using (11), the estimate of $V(\hat{y} \dots)$ is given by

$$\text{Let } V(\hat{y} \dots) = \frac{1}{2} \left(\text{est } V(u) + \left(\frac{n}{n} \right)^2 \text{est } V(w) - \frac{2uw}{n} - \text{cov}(u, w) \right) \quad \dots \dots .14$$

If we define

$$S_{ij}^2 = \frac{1}{l_{ij} - 1} \sum_j \left(\frac{t_{ij}}{l_{ij}} - \frac{\sum_j t_{ij} \bar{s}_{ij} (l_{ij})}{\sum_j l_{ij} \bar{s}_{ij} (l_{ij})} - 1 \right)^2$$

$$S_i^2 = \frac{1}{l_i - 1} \sum_j \left(\frac{t_{ij}}{l_i} - \frac{\sum_j t_{ij} \bar{s}_{ij} (l_{ij})}{\sum_j l_{ij} \bar{s}_{ij} (l_{ij})} - 1 \right)^2 \quad \dots \dots .15$$

$$S_{uw}^2 = \frac{1}{n-1} \sum_i \left(\frac{t_i}{l_i} - \frac{\sum_j t_{ij} \bar{s}_{ij} (l_{ij})}{\sum_j l_{ij} \bar{s}_{ij} (l_{ij})} - 1 \right) \left(\frac{1}{p_i} - \frac{1}{l_i} \sum_j l_{ij} \bar{s}_{ij} (l_{ij}) \right)$$

on multiplying by $\frac{1}{n}$ to both sides it can be shown that

$$E \frac{S_u^2}{n} = V(u); \quad E \frac{S_w^2}{n} = V(w); \quad E \frac{S_{uw}}{n} = \text{cov}(u, w)$$

It follows that whatever is true in such cases is also true in general.

That is, covariance and variance $V(\hat{y} \dots)$ are minimum. 51..

$$\text{est } v(\hat{y} \dots) = \frac{1}{n} \left(\frac{1}{n} \left(\frac{s^2}{s^2 + \frac{u^2}{n}} s^2 - \frac{u^2}{n} s_{uw} \right) \right) \dots \dots 15$$

5.5 COMPARISON OF DIFFERENT ESTIMATES.

From eq. 5.1, the variance of the linearized ratio type estimate \hat{y}_{ij} is $s_{\hat{y}_{ij}}^2$, corresponding result for the estimate \hat{y}_{ijk} is $s_{\hat{y}_{ijk}}^2$, $s_{\hat{y}_{ij}}^2$, $s_{\hat{y}_{ijk}}^2$ giving successive variances of a_i , l_{ij} , l_{ijk} and univariate Z_{ijk} .

Case I

For $t^2 = 1$, $a_i = l_i$; $l_{ij} = 1$, $t_{ij} = r_{ij}$ and $Z_{ijk} = 1$ for all i, j, k .

$$E(u) = y \dots; E_{ij} = \bar{y}_{i(l_i)}; E_{ij} = 1 \& R_i = \bar{y}_i (l_i)$$

Estimate \hat{y}_{11} :

Using 10, we obtain

$$P_{11} = \frac{1}{n} \left(\frac{\sigma_B^2}{s^2} - \frac{\sigma_{y\dots}^2}{y \dots^2} \right) + \frac{1}{n} \sum_{i=1}^m \frac{s_{ij}^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{l_i} \right) \left(\frac{s'_{iB}^2}{B^2} - \frac{s'_{iyB}^2}{y \dots^2} \right) \dots \dots 17$$

$$\text{where } S'_{iyB} = \frac{1}{m_i-1} \sum_{j=1}^{m_i} (r_{ij} \bar{y}_{ij} - \bar{y}_{i\dots}) (u_{ij} - 1) \dots \dots 18$$

Take R_{ij} , R_i , and t to be unity, so that finite correction factors can be ignored further letting $m_i = m$, $n_j = b$

$$B = \frac{1}{n} \left(\frac{\sigma_B^2}{s^2} - \frac{\sigma_{y\dots}^2}{y \dots^2} \right) + \frac{1}{nm} \left(\frac{s'_{iB}^2}{B^2} - \frac{s'_{iyB}^2}{y \dots^2} \right) \dots \dots 19$$

$$\text{where } S'_{iyB}^2 = \sum_{i=1}^m \frac{s_{ij}^2}{p_i} S'_{iB}^2$$

$$\frac{z_i}{w_i} = \sum_{l=1}^n \frac{\frac{z_i}{x_l}}{x_l} - z_i$$

Therefore $\sigma_{z_i}^2 = \sum_{l=1}^n \frac{1}{x_l^2} \sigma_{x_l}^2 + z_i^2$

$$\frac{z_i}{w_i} = \frac{\sigma_{x_i}^2}{\sigma_w^2} = \frac{1}{\sum_{l=1}^n x_l^2}$$

$$\text{Var}(z_i) = \text{Var}\left(\frac{z_i}{w_i}\right)$$

$$\text{Var}\left(\frac{z_i}{w_i}\right) = \frac{1}{\sum_{l=1}^n x_l^2} (\sigma_{x_i}^2 + \sum_{l=1}^n x_l^2 \sigma_{x_l}^2)$$

$$= \frac{1}{\sum_{l=1}^n x_l^2} \left(\frac{1}{x_i^2} - \frac{1}{\sum_{l=1}^n x_l^2} \right) \sum_{j=1}^{M_i} x_{ij}^2 + \sum_{j=1}^{M_i} \left(\frac{x_{ij}^2}{x_i^2} \sum_{l=1}^n x_l^2 \right)$$

$$= \frac{1}{\sum_{l=1}^n x_l^2} \frac{1}{x_i^2} \sum_{j=1}^{M_i} x_{ij}^2 \left(\frac{1}{x_i^2} - \frac{1}{\sum_{l=1}^n x_l^2} \right) \sum_{l=1}^n x_l^2 \quad \dots \dots \dots$$

$$\frac{\Delta}{x_1}$$

$$= \frac{1}{\sum_{l=1}^n x_l^2} \left(\frac{\sigma_{x_i}^2}{x_i^2} - \frac{\sigma_{x_i}^2}{\sum_{l=1}^n x_l^2} \right) \quad \dots \dots \dots$$

$$\rho_{y_B} \quad \frac{\sigma_B/B}{\sigma_y/y}$$

$$= \frac{i_B}{\sum_{l=1}^n x_l^2}$$

R_B $=$ Un.

$$\begin{aligned}
 v(\hat{y}_{12}) &= \frac{1}{n^2} \left[\frac{1}{n} (\sigma_y^2 + \sum_{i=1}^n \sigma_{iy}^2 - \sum_{i=1}^n \sigma_{iyB}^2) \right. \\
 &\quad + \frac{1}{n} \sum_{i=1}^n \frac{\nu_i}{\nu_i} \left(\left(\frac{1}{m_1} - \frac{1}{k_1} \right) s_{iy}^2 \right. \\
 &\quad \left. \left. + \frac{1}{m_1} \sum_{j=1}^m u_{ij}^2 \left(\frac{1}{v_{ij}} - \frac{1}{B_{ij}} \right) s_{ijB}^2 \right) \right] \dots \dots 23
 \end{aligned}$$

From 21 and 23 it easily follows that

$$v(\hat{y}_{11}) - v(\hat{y}_{12}) = \sum_{i=1}^n \frac{\nu_i}{\nu_i} \left(\frac{1}{m_1} - \frac{1}{k_1} \right) \left(\bar{s}_y^2 - \bar{s}_{iyB}^2 \right) \dots \dots 24$$

To make \hat{y}_{11} to be larger we have to p. c. c. i. the \bar{s}_{iyB} in
for. i.e. $s_{iyB} = 0$; $s_{ijB} = 0$; $s_{ijB} = 0$;

$$\sum_i \nu_i \bar{s}_{iyB}^2 = \bar{s}_y^2; \sum_i \nu_i s_{iyB}^2 = \bar{s}_{iyB}^2 \text{ and } \sum_i \nu_i s_{ijB}^2 = \bar{s}_{ijB}^2 = \bar{\rho}_w \bar{s}_{ijB}^2;$$

i.e. if we want \hat{y}_{11} will be efficient than \hat{y}_{12} i.e.

$$\bar{\rho}_w > \frac{1}{2} \bar{s}_y^2 - \frac{\bar{s}_{ijB}^2}{\bar{s}_{ijB}^2} \dots \dots 25$$

See II. Here we have $a_i = B_i$; $l_{ij} = t_{ij} = 1$; $Z_{ijk} = 1$ for all i, j ,

$$\text{Therefore } E(\lambda) = y_{1\dots} = B_i \bar{y}_{i(l_i)}$$

From 10 the relative bias B_{21} in the est. of \hat{y}_{21} is given by

$$\begin{aligned}
 B_{21} &\equiv \frac{\bar{y}_{1\dots} - y_{1\dots}}{y_{1\dots}} + \frac{1}{r} \left(\frac{\sigma_y^2}{\bar{s}_y^2} - \frac{\sigma_{yB}^2}{\bar{s}_{yB}^2} \right) \\
 &\quad + \frac{1}{r} \sum_{i=1}^n \frac{\nu_i}{\nu_i} \left(\frac{1}{m_1} - \frac{1}{k_1} \right) \left(\frac{\bar{s}_{iyB}^2}{\bar{s}_y^2} - \frac{\bar{s}_{ijB}^2}{\bar{s}_{ijB}^2} \right) \dots \dots 26
 \end{aligned}$$

$$\therefore \sigma_{yB}^2 = \sum_{i=1}^n \nu_i \left(\frac{B_i}{\nu_i} - B \right) \left(\frac{\bar{s}_{iyB}^2}{B_i} - \bar{s}_{iyB}^2 \right)$$

$$\begin{aligned}
 \bar{s}_{L_{ij}}^2 &= \frac{1}{n_i - 1} \sum_{j=1}^{m_i} (\bar{y}_{ij.} - \bar{\bar{y}}_{i(\bar{L}_i)}) (u_{ij} - 1) \quad \dots\dots 27 \\
 &= -\frac{1}{n_i - 1} \sum_{j=1}^{m_i} (\bar{y}_{ij.} - \bar{\bar{y}}_{i.}) (\bar{y}_{ij.} - \bar{\bar{y}}_{i(\bar{L}_i)}) + \frac{1}{n_i - 1} \left(\frac{\sigma_w^2}{n_i} - \frac{\sigma_{y_{ij.}}^2}{n_i - 1} \right) \\
 &\quad + \frac{1}{n_i - 1} \sum_{j=1}^{m_i} \frac{\bar{y}_{ij.}}{n_i} \left(\frac{1}{n_i} - \frac{1}{n_i} \right) \left(\frac{\bar{y}_{ij.}^2}{n_i^2} - \frac{\bar{y}_{ij.}^2}{n_i(n_i - 1)} \right) \\
 &= -\frac{1}{n_i - 1} \sum_{j=1}^{m_i} \bar{y}_{ij.} (\bar{y}_{ij.} - 1) + \frac{1}{n_i - 1} \left(\frac{\sigma_w^2}{n_i} - \frac{\sigma_{y_{ij.}}^2}{n_i - 1} \right) \\
 &\quad + \frac{1}{n_i - 1} \sum_{j=1}^{m_i} \frac{\bar{y}_{ij.}}{n_i} \left(\frac{1}{n_i} - \frac{1}{n_i} \right) \left(\frac{\bar{y}_{ij.}^2}{n_i^2} - \frac{\bar{y}_{ij.}^2}{n_i(n_i - 1)} \right) \\
 \end{aligned}$$

.....27

The $\bar{y}_{ij.}$ + $\bar{y}_{ij.}^2$ = $\sigma_w^2 + s_{ij.}^2$ can be used

$$\bar{y}_{ij.}^2 - 1 = n_i \bar{y}_{ij.} = n_i - 1 = \bar{y}_{ij.}^2 - 1$$

$$\begin{aligned}
 \bar{s}_{L_{ij}}^2 &\approx \frac{\bar{y}_{ij.}^2}{n_i - 1} + \frac{1}{n_i - 1} \left(\frac{\sigma_w^2}{n_i} - \frac{\sigma_{y_{ij.}}^2}{n_i - 1} \right) + \frac{1}{n_i - 1} \left(\frac{\bar{y}_{ij.}^2}{n_i^2} - \frac{\bar{y}_{ij.}^2}{n_i(n_i - 1)} \right) \\
 \end{aligned}$$

.....27

+ $\bar{y}_{ij.}^2$ & $\bar{y}_{L_{ij}}^2$ = $s_{ij.}^2$ \therefore It follows $s_{ij.}^2 = s_{ij.}^2$

It follows

$$s_{ij.}^2 = \bar{y}_{ij.}^2 / \bar{y}_{ij.}^2 = \bar{y}_{ij.}^2 / \bar{y}_{ij.}^2$$

$$\begin{aligned}
 \bar{y}_{ij.}^2 &\approx \frac{1}{n_i} \left(\frac{1}{n_i} (\sigma_{ij.}^2 + \bar{y}_{ij.}^2) - \frac{1}{n_i} \sigma_{y_{ij.}}^2 \right) \\
 &\quad + \frac{1}{n_i} \sum_{j=1}^{m_i} \frac{\bar{y}_{ij.}}{n_i} \left(\frac{1}{n_i} - \frac{1}{n_i} \right) s_{ij.}^2 + \dots\dots 55.
 \end{aligned}$$

$$+ \frac{1}{i} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{x_k} \left(\frac{1}{r_1} - \frac{1}{r_k} \right) \left(\frac{y_1}{x_1} + \frac{y_2}{x_2} + \dots + \frac{y_k}{x_k} - 2\bar{y}_1 - \dots - 2\bar{y}_{k-1} \right) \dots \dots \dots \quad (29)$$

$$j_{ij_1}^2 = \frac{1}{n-1} \sum_{j=1}^{n-1} (j_{ij_1} - \bar{j}_{ij_1})^2$$

20 10 + $\frac{1}{2}$ + π x^2 y^2 z^2 B_{22} \ln $+ \pi$ b^2 $\frac{\Delta}{\delta}$ a^2 $- 1$

$$S_{ij} \cong -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m \left(\bar{x}_{ij} - \bar{\bar{x}}_i \right) \left(\bar{x}_{ij} - \bar{x}_j \right) + \frac{1}{n} \left(\bar{x} - \frac{\bar{x}_1 + \bar{x}_2}{2} \right) \dots \dots 30$$

III 4. *at* *if* *is* *in* *is* *to* *3* *to* *r* *is* *-AG*

(+) - 25 ° 35 ... x + 7 c = 2 x - x - 4

$$(i^z) \quad j_1, \dots = \sigma_{j_1}^{-1} / \sigma^{j_2} \quad \dots, \quad r_{j_1}^{-1} = \frac{c_{j_1}}{c_{j_2}}$$

$$U = \{x^2 - y^2\} \cap V_{(y_1, y_2)} = V(\hat{y}_1) \cap V(\hat{y}_2)$$

Fri., 20 & 31 : 10 'o - n

$$\tau(\hat{\binom{z}{j-1}}) - \tau(\hat{\binom{z}{j+1}}) = \frac{1}{\pi} \sum_{i=1}^{\infty} \frac{\hat{\binom{z}{i}}}{\hat{\mu}_i} \left(\frac{1}{z_i} - \frac{1}{\hat{z}_{i+1}} \right) (\hat{\beta}_1^{i-1}, \dots, \hat{\beta}_{i-1}^{i-1}, -\hat{\beta}_i^i, \dots, \hat{\beta}_{i+j-1}^i, B) \quad \dots \dots 32$$

$$x_1 = 2; x_1 = \frac{2}{R},$$

$$\sum_i p_i \cdot \frac{1}{i} = \frac{1}{\bar{S}_B}; \quad \sum_i p_i \cdot \frac{1}{i^2} = \frac{1}{\bar{S}_{B_1}}; \quad \sum_i p_i \cdot \frac{1}{i^3} = \frac{1}{\bar{S}_{B_2}}.$$

it can be seen that \hat{y}_{20} will be less efficient than \hat{y}_{21} if

$$\bar{P}, \quad \leftarrow \frac{1}{2} \bar{y}, \dots, \frac{\bar{z}_{n+1}}{\bar{z}_n} \dots 56..$$

Case III

For the case $v_i = 0$

$$a_{ij} = b_{ij}, \quad l_{ij} = t_{ij} = \bar{v}_{ij}; \quad z_{ijk} = 1 \text{ for all } i, j, k$$

$$\bar{v}_{ij} = \bar{y}_i(\bar{v}_j); \quad H_{iz} = \bar{v}_i. \quad \text{then } v_i = \bar{y}_i.$$

$$\sum_{c=1}^n s_{ic..} \left\{ 1 + \left(\frac{1}{a_{ii}} - \frac{1}{l_{ii}} \right) \left(\bar{v}_{iB}^2 - \frac{\sigma_{vB}^2}{\bar{y}_{i..}} \right) \right\} = s_{i..} \dots \dots 34$$

$$\begin{aligned} \bar{v}_{31} &= \sum_{c=1}^n \frac{s_{ic..}}{s_{i..}} \left(\frac{1}{a_{ii}} - \frac{1}{l_{ii}} \right) \left(\bar{v}_{iB}^2 - \frac{\sigma_{vB}^2}{\bar{y}_{i..}} \right) + \frac{1}{n} \left(\frac{\sigma_z^2}{2} - \frac{\sigma_{vB}^2}{\bar{y}_{2..}} \right) \\ &\quad + \frac{1}{nB^2} \sum_{c=1}^n \frac{s_{ic..}}{a_{ii}} \left(\frac{1}{a_{ii}} - \frac{1}{l_{ii}} \right) \bar{v}_{iB}^2 - \frac{1}{nB^2} \sum_{c=1}^n \frac{s_{ic..}}{a_{ii}} \left(\frac{1}{a_{ii}} - \frac{1}{l_{ii}} \right) (\bar{v}_{i..}^2 - \bar{y}_{i..}^2) \end{aligned} \dots \dots 35$$

If $v_1, v_2 \rightarrow 0$ or $v_i = 0$ ($i = 1, 2, \dots, n$) the regression of

y_{ij} , or \bar{v}_{ij} is linear and the regression line passes through the origin then

$$\bar{v}_{31} \equiv \frac{1}{n} \left(\frac{\sigma_z^2}{3^2} - \frac{\sigma_{vB}^2}{\bar{y}_{2..}} \right) + \frac{1}{nB^2} \sum_{c=1}^n \frac{s_{ic..}^2}{a_{ii}} \left(\frac{1}{a_{ii}} - \frac{1}{l_{ii}} \right) s_{iB}^2 \dots \dots 36$$

Now, let's let $m_i = m$ and F_i to be linear so that the p.c can be

independent

$$\bar{v}_{31} = \frac{1}{n} \left(\frac{\sigma_z^2}{E^2} - \frac{\sigma_{vB}^2}{\bar{y}_{2..}^2} \right) + \frac{\sigma_z^2}{m} \dots \dots 37$$

and will be satisfied if the sample size is sufficiently large.

Using 13 and 36 we get in

$$\bar{v}(\bar{v}_{31}) = \frac{1}{n} \left(\frac{1}{n} \left(\frac{\sigma_z^2}{E^2} + \frac{\sigma_z^2}{\bar{y}_{2..}^2} \right) - 2\bar{v}_{2..} \sigma_{vB}^2 \right) +$$

.... 57 ..

$$\begin{aligned}
 & + \frac{\sum_{i=1}^n}{n} \sum_{i=1}^n \frac{s_i^2}{p_i} \left(\frac{1}{r_i} - \frac{1}{m_i} \right) \bar{y}_{i..} s_{iB}^2 \\
 & + \frac{1}{n} \sum_{i=1}^n \frac{s_i^2}{p_i} \left(\frac{1}{r_i} - \frac{1}{r_i} \right) \left(\bar{y}_{i..}^2 + \sum_{j=1}^{m_i} u_{ij}^2 - 2\bar{y}_{i..} s_{iB}^2 \right) \\
 & + \frac{1}{n} \sum_{i=1}^n \frac{s_i^2}{p_i} \left(\frac{1}{r_i} - \frac{1}{r_i} \right) \sum_{j=1}^{m_i} u_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{s_{ij}} \right) s_{ij..}^2 \quad \dots \dots .36
 \end{aligned}$$

$$w \text{ var } \hat{y}_{i..}^2 = \frac{1}{i-1} \sum_{j=1}^{m_i} u_{ij}^2 (\bar{y}_{ij..} - \bar{y}_{i..})^2$$

The above $\hat{y}_{i..}$ in relation to \hat{y}_{32} is given by

$$\hat{y}_{32} = \sum_{i=1}^n \frac{\bar{y}_{i..}}{p_i} \left(\left(\frac{1}{r_i} - \frac{1}{m_i} \right) \left(\hat{y}_{i..}^2 - \frac{s_{iB}^2}{p_i} \right) + \frac{1}{r_i} \left(\frac{\sigma_y^2}{r_i} - \frac{\sigma_y^2}{s_{i..}} \right) \right). \quad \dots .39$$

If $u_{ij} \neq 0$, y_{ij} is on line r_i then $\hat{y}_{ij..}$ is the mean of

y_{ij} on B_{ij} i.e. line r_i proves that $\hat{y}_{ij..}$ is on line r_i then

on B_{ij}

$$\hat{y}(u) = y_{i..}$$

$$\hat{y}_{32} = \frac{1}{n} \left(\frac{\sigma_y^2}{r^2} - \frac{\sigma_y^2}{s^2} \right)$$

if i is on r then $\hat{y}_{i..}$ is on $\frac{1}{p_i}$

$$\frac{1}{p_i}$$

so $\frac{1}{p_i}$ is line r so \hat{y}_{32} is on r i.e. \hat{y}_{32} is on r .

It can be seen that $\hat{y}(y_{32})$ is given by

$$\begin{aligned}
 \hat{y}(\hat{y}_{32}) & = \frac{1}{r^2} \left\{ \frac{1}{n} \left(\frac{\sigma_y^2}{r^2} + \sum_{i=1}^n \frac{\sigma_y^2}{p_i^2} - \sum_{i=1}^n \frac{\sigma_y^2}{s_{i..}^2} \right) \right. \\
 & \left. + \frac{1}{n} \sum_{i=1}^n \frac{s_i^2}{p_i} \left(\left(\frac{1}{r_i} - \frac{1}{r_i} \right) \hat{y}_{i..}^2 + \frac{1}{r_i} \sum_{j=1}^{m_i} u_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{s_{ij}} \right) s_{ij..}^2 \right) \right\} \dots \dots .40
 \end{aligned}$$

From 33 it follows that \hat{y}_{32} will be a coefficient in the equation

of \hat{y}_{31} if

$$\sum_{i=1}^N \frac{s_{i..}^2}{k_i} \left(\frac{1}{k_i} - \frac{1}{M_i} \right) s_{i..}^{12} > 2 \sum_{i=1}^N \frac{s_{i..}^2}{k_i} \left(\frac{1}{k_i} - \frac{1}{M_i} \right) (s_{i..}^{1..} - \bar{y}_{i..}) s_{i..}^{12}$$

.....41

always true. Therefore the sufficient

L.H.S of 41 is a condition for \hat{y}_{32} to be a coefficient over \hat{y}_{31} in the

equation (i = 1, 2, .. N)

$$-\frac{s_{i..}^{12}}{s_{i..}^{1..}} \leq P_{i..}^{12} \leq \frac{s_{i..}^{12}}{s_{i..}^{1..}}$$

.....42

$$\text{where } P_{i..}^{12} = \frac{s_{i..}^{12}}{s_{i..}^{1..}} \quad \& \quad P_{i..}^{1..} = \frac{s_{i..}^{1..}}{s_{i..}^{1..}}$$

Stage IV.

It is now to solve

$$i = i, \quad t_{ij} = \tau_{ij}, \quad l_{ij} = A_{ij}; \quad r_{ij} = 1 \text{ for all } i, j, k.$$

$$G_{iy} = \bar{y}_{i(\cdot i)}; \quad E_{iz} = \bar{k}_i \quad \text{therefore } R_i = s_{i..} / \bar{k}_i$$

$$\therefore \phi(u) = \sum_{i=1}^N y_{i..} \left\{ 1 + \left(\frac{1}{m_i} - \frac{1}{k_i} \right) \left(\frac{s_{i..}^2}{s_{i..}^{1..}} - \frac{s_{i..}^{1..}}{\bar{k}_i \bar{y}_{i(\cdot i)}} \right) \right\} = y_{3..} \dots$$

.....43

$$\text{where } s_{i..}^2 = \frac{1}{k_i-1} \sum_{j=1}^{M_i} (A_{ij} - \bar{k}_i)^2$$

$$s_{i..}^{1..} = \frac{1}{k_i-1} \sum_{j=1}^{M_i} (A_{ij} - \bar{k}_i) (y_{ij} - \bar{y}_{i(\cdot i)})$$

From 10 it can be seen that the relative bias B_{41} in the estimate

\hat{y}_{41} is given by

.....59..

$$\begin{aligned}
 & \sum_{i=1}^n \frac{x_i}{x_i} \left\{ \left(\frac{1}{x_i} - \frac{1}{\bar{x}_i} \right) \left(\frac{\bar{x}_i}{x_i} - \frac{\bar{x}_{i+1}}{\bar{x}_i} \right) \right\} + \frac{1}{n} \left(\frac{\bar{x}_1^2}{x_1} - \frac{\bar{x}_{n+1}^2}{x_n} \right) \\
 & = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{x_i} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_i} \right) \frac{\bar{x}_i^2}{x_i} = \\
 & = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{x_i} \left\{ \left(\frac{1}{x_i} - \frac{1}{\bar{x}_i} \right) \left(\frac{\bar{x}_{i+1}}{x_i} - \frac{\bar{x}_{i+1}}{\bar{x}_i} \right) \right\} \dots \dots 4 \\
 & \sigma_{j_2} = \sum_{i=1}^n x_i \left(\frac{\bar{x}_{i+1}}{x_i} - \frac{\bar{x}_i}{x_i} \right) \left(\frac{x_i}{x_i} - \frac{\bar{x}_i}{x_i} \right) \\
 & = -\frac{x_{i+1}}{x_i} \left\{ 1 + \left(\frac{1}{x_i} - \frac{1}{\bar{x}_i} \right) \left(\frac{\bar{x}_i}{x_i} - \frac{\bar{x}_i}{\bar{x}_{i+1}} \right) \right\} \\
 & = \sum_{i=1}^n x_i \frac{\bar{x}_i}{x_i} \\
 & = \frac{1}{n-1} \sum_{j=1}^{n-1} (\bar{x}_{i+j} - \bar{x}_i) (\bar{x}_{i+j} - \bar{x}_i) \\
 & = \frac{1}{n-1} \sum_{j=1}^{n-1} (x_{i+j} - \bar{x}_i) (x_{i+j} - \bar{x}_i) \\
 & \therefore \hat{\sigma}_{j_2} = \frac{1}{n-1} \left(\sigma_{j_2}^2 + \bar{x}_2^2 \sigma_{j_2}^2 - \bar{x}_j \sigma_{j_2}^2 \right) \sum_{i=1}^n \frac{x_i}{x_i} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_i} \right) x_i
 \end{aligned}$$

$$+ \sum_{i=1}^N \frac{i}{x_i} \left(\frac{1}{z} - \frac{1}{i} \right) \left(\omega_{i,j_1}^{(1)} - \omega_{j_1, i}^{(2)} - \omega_{j_2, i}^{(3)} - \dots - \omega_{j_N, i}^{(N)} \right)$$

$$+ \sum_{i=1}^N \frac{i}{x_i} \left(\frac{1}{z} - \frac{1}{i} \right) \sum_{j=1}^{M_i} \omega_{i,j}^{(2)} \left(\frac{1}{z} - \frac{1}{x_j} \right) \omega_{i,j}^{(3)}, \dots 45..$$

$$e^{S_{ijz}} = \frac{1}{i-1} \sum_{j=1}^{i-1} \frac{z^j}{j} \left(\frac{i^{j+1}}{i-j} - \frac{1}{j+1} \right)$$

$$B_{42} = \frac{1}{\gamma} \left(\frac{\sigma_3^2}{\beta^2} - \frac{\sigma_{y_{20}}}{y_{3...}} \right) + \sum_{i=1}^N \frac{y_{i..}}{y...} \left(\frac{1}{y_i} - \frac{1}{M_i} \right) \left(\frac{s_{iA}}{y_i^2} - \frac{s_{i..}}{A_i y_i (K_i)} \right)$$

The relative abundance index

$$(i) \quad P_{3^q} = \frac{C_q}{3^q}$$

$$(\text{iii}) \quad P_{i,j} = \frac{c_{ij}}{c_{i,j}} \quad \text{or ev}_j \cdot i \quad (i = 1, 2, \dots)$$

Clearly the variance $V(\hat{y}_{12})$ of the estimate \hat{y}_{12} is given by

$$\nabla \left(\frac{\hat{y}_i}{y_{iB}} \right) = \frac{1}{g^2} \left\{ \frac{1}{n} \left(\frac{s_i^2}{v_{iz}} + \frac{v_{iz}^2}{y_{iz}} \frac{s_i^2}{v_{izB}} - 2 \bar{y}_{iz} \frac{s_i^2}{v_{izB}} \right) \right. \\ \left. + \frac{1}{n} \sum_{i=1}^n \frac{s_i^2}{p_i} \left\{ \left(\frac{1}{n_i} - \frac{1}{m_i} \right) s_{ijy}^2 \right. \right. \\ \left. \left. + \frac{1}{n_i m_i} \sum_{j=1}^{m_i} u_{ij}^2 \left(\frac{1}{v_{ij}} - \frac{1}{v_{ijB}} \right) s_{ijy}^2 \right\} \right\} \quad 47$$

I⁺ at once follows from 47 to 1 the estimate

\hat{y}_{42} will be efficient if and only if \hat{y}_{41} is.

The sufficient condition for α to be a good estimate for θ is

i ($i = 1, 2, \dots, n$)

$$P_{ij3} \leq s_{i_0} \geq \frac{y_{i..}}{A_i} P_i \quad i = i_0$$

$$\text{i.e. } \rho_{iy^3} \geq \rho_{iAB}$$

...61...

Case I.

In this case we have

$$a_i = z_{i..}, \quad t_{ij} = l_{ij} = p_{ij}; \quad g_{ij} = \bar{z}_i(l_i); \quad r_i = \bar{z}_i(l_i) \text{ whence}$$

$$\begin{aligned} \frac{s_{i..}}{r_{i..}} &= R_i \quad \text{so } E(u) = \sum_{j=1}^m p_{ij} v_j = \bar{v} \quad \text{so } \\ \frac{s'_{i..}}{r'_{i..}} &= \frac{y_{i..}}{p_i} \left\{ 1 + \left(\frac{1}{l_i} - \frac{1}{r_i} \right) \left(\frac{s_i^2}{l_i(l_i)} - \frac{s_{ij}^2}{l_i(l_i) \bar{z}_i(l_i)} \right) \right. \\ &\quad \left. + \frac{1}{r_i} \sum_{j=1}^m b_{ij}^2 \left(\frac{1}{b_{ij}} - \frac{1}{r_i} \right) \left(\frac{s_{ij}^2}{l_i(l_i)} - \frac{s_{ij} y_{ij}}{l_i(l_i) \bar{z}_i(l_i)} \right) \right\} \dots \dots 50 \end{aligned}$$

$$\frac{s'_{i..}}{r'_{i..}} = \frac{1}{r_{i..}} \sum_{j=1}^m (z_{ij} - \bar{z}_i(l_i))^2$$

$$\frac{s'_{ij}}{r'_{ij}} = \frac{1}{r_{ij}} \sum_{j=1}^m (y_{ij} - \bar{z}_i(l_i)) (z_{ij} - \bar{z}_i(l_i))$$

If we assume the condition of consistency of $y_{i..}$;

The total error will be negligible if it is less than $\frac{3}{2}$

and since $y_{i..}$ is linear in $y_{i..}/p_i$ and r_i/p_i is linear

and since it is less than the ratio. Furthermore, the relative error is less than $\frac{\Delta}{y_{51}}$ by

$$\begin{aligned} \frac{s_{51}}{r_{51}} &= \frac{1}{n} \left(\frac{\sigma_B^2}{n} - \frac{\sigma_{1..}^2}{n} \right) + \frac{(u) - y}{\dots} \\ &\quad + \frac{1}{n^2} \sum_{i=1}^n \frac{n^2}{k_i} \left(\frac{1}{k_i} - \frac{1}{n} \right) \frac{s_i^2}{\dots} - \end{aligned}$$

$$= \frac{1}{\sigma} \sum_{i=1}^N \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right) \left(z_i - \frac{\bar{z}_i}{z_i} \bar{z}_i \right) \dots$$

estimates $\hat{\theta}_1, \hat{\theta}_2, \dots$ given by

$$\hat{\theta}_1 = \frac{1}{\sigma^2} \left\{ \frac{1}{z_1} + \frac{1}{z_2} + \dots + \frac{1}{z_N} \right\}$$

$$+ \dots + \sum_{i=1}^N \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right)$$

$$- \frac{1}{\sigma^2} \sum_{i=1}^N \left\{ \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right)^2 + \frac{1}{z_i} \sum_{j=1}^{M_i} \left(\frac{1}{z_j} - \frac{1}{\bar{z}_j} \right) \right\}$$

$$- \frac{1}{\sigma^2} \sum_{i=1}^N \frac{B_i}{\mu_i} \left\{ \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right) \left(\bar{z}_i - \frac{1}{z_i} \bar{z}_i \right) \right\} \dots$$

$$\hat{\theta}_2 = \frac{1}{\sigma^2} \left\{ \frac{1}{z_1} + \frac{1}{z_2} + \dots + \frac{1}{z_N} \right\}$$

$$+ \frac{1}{\sigma^2} \sum_{i=1}^N \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right)^2 + \frac{1}{\sigma^2} \sum_{i=1}^N \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right) \dots$$

$$= \frac{1}{\sigma^2} \dots$$

$$\sigma^2 = \sum_{i=1}^N \left(\frac{1}{z_i} - \frac{1}{\bar{z}_i} \right)^2$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{M_i} \left\{ \left(\bar{x}_{i+1} - \bar{x}_i \right)^2 - \left(\bar{x}_i - \bar{x}_{i+1} \right)^2 \right\}$$

$$\sigma^2 = \frac{1}{n-1} \sum_{k=1}^{B_i} \left\{ x_{ik} - x_{ik} \left(\bar{x}_{ik} - \bar{x}_{ik} \right)^2 \right\} \dots$$

From Eq. 53 we have $\hat{s}_{ij}^2 = \frac{\Delta}{y_{ij}} \hat{y}_{ij}$ will be strictly less than the
error $\sigma = \sqrt{\frac{2}{n-1}} \hat{y}_{ij}$

$$\left[\frac{\hat{s}_{ij}^2 - \frac{1}{n} \sum_{i=1}^n \frac{\hat{s}_{ij}^2}{p_i} \left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) \hat{y}_{ij}^2}{\sigma^2} - \frac{2\bar{y}^2 \dots - \frac{1}{n} \sum_{i=1}^n \frac{\hat{s}_{ij}^2}{p_i} \left(\left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) \left(s_{ij}^2 - \frac{s_{ij}^2}{p_i} s_{ij}^2 \right) \right)}{s^2} \right] > 0 \quad \dots \dots 55$$

Therefore the precision condition is satisfied for \hat{s}_{ij}^2 to be strictly less than
the error σ_{ij} ($i = 1, 2, \dots$)

$$P_{ij} \hat{s}_{ij}^2 s_{ij}^2 \geq P_{ij} \frac{\hat{y}_{ij}^2}{p_i} s_{ij}^2 s_{ij}^2$$

$$\dots P_{ij} \hat{s}_{ij}^2 \geq P_{ij} \frac{\frac{\hat{s}_{ij}^2}{p_i}}{s_{ij}^2} \quad \dots \dots \dots 56$$

Let's take $\frac{\Delta}{y_{ij}}$ instead of \hat{y}_{ij}

The left side of Eq. 56 will be sufficient if $\frac{\Delta}{y_{ij}}$ is

$$\begin{aligned} \sum_{i=1}^n \frac{\hat{s}_{ij}^2}{p_i} \left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) \left(s_{ij}^2 - \frac{s_{ij}^2}{p_i} \right) &\leq 0 \\ + \sum_{i=1}^n \frac{1}{p_i} \frac{1}{\bar{p}_i} \sum_{j=1}^m \left(1 - \frac{s_{ij}^2}{p_i} \right) \left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) s_{ij}^2 &> 0 \quad \dots \dots \dots 57 \end{aligned}$$

To prove it we need to show Eq. 57 to hold up to re

$$(i) \sum_{j=1}^m \left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) s_{ij}^2 > \sum_{j=1}^m \frac{1}{p_i} \left(\frac{1}{p_i} - \frac{1}{\bar{p}_i} \right) s_{ij}^2$$

$$(ii) s_{ij}^2 > s_{ij}^2 \quad \text{for every } i \quad (i = 1, 2, \dots)$$

Condition (i) will hold just if for every i there exist a correle.

in tr. P_{ij} and S_{ijy} i.e. (ii) stand true i.e. \hat{y}_{ij} is efficient over $\hat{y}_{ij'}$

between \hat{y}_{ij} and B

$j.$ $i.j$

$$\hat{y}_{ij} = \frac{\hat{y}_{12}}{y_{12}} + \frac{\hat{y}_{32}}{y_{32}}$$

The st. \hat{y}_{12} will be efficient over \hat{y}_{32} if

$$v(\hat{y}_{32}) - v(\hat{y}_{12}) > 0 \quad \dots \dots \dots \quad 58$$

" " 23 and 36 we have

$$v(\hat{y}_{32}) - v(\hat{y}_{12}) = \frac{1}{nB^2} \sum_{i=1}^m \frac{i}{p_i} \left(\frac{1}{\hat{y}_{12}} - \frac{1}{\hat{y}_{32}} \right) (S_{ij2}^2 - S_{ij'}^2)$$

then if \hat{y}_{12} is to be efficient over \hat{y}_{32} , then

$$\frac{1}{n} \sum_{i=1}^m \frac{i}{p_i} \left(\frac{1}{\hat{y}_{12}} - \frac{1}{\hat{y}_{32}} \right) (S_{ij2}^2 - S_{ij'}^2) > 0 \quad \dots \dots \quad 59$$

By referring to the sufficient conditions to 59 we see that

$$\text{if } S_{ij2}^2 > S_{ij'}^2 \quad \text{for every } i \quad (i = 1, 2, \dots, n)$$

$$\text{i.e. } \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_{i..})^2 > \sum_{j=1}^{m_i} \left(\frac{y_{ij}}{p_i} - \frac{\bar{y}_{i..}}{p_i} \right)^2$$

$$\text{or if } \sum_{j=1}^{m_i} \left\{ \left(\frac{y_{ij}}{p_i} - 1 \right) \bar{y}_{i..} \right\} \left\{ \left(\frac{y_{ij}}{p_i} - \bar{y}_{i..} \right) - \left(\frac{y_{ij}}{p_i} - 1 \right) \bar{y}_{i..} \right\} \leq 0$$

$$\text{or if } \frac{P_{ijy}}{S_{ijy}} < \frac{1}{2} \frac{c_i}{c_{i'}}$$

$$\text{or if } P_{ijy} < \frac{1}{2} \frac{c_i}{c_{i'}} \quad \dots \dots \quad 60$$

$$\text{If } c_i = c_{i'} \text{ then } P_{ijy} \text{ is equal to } P_{ijy} < \frac{1}{2}$$

$$\text{dist} \rightarrow \hat{s}_{12} \quad i \rightarrow \hat{s}_{52}$$

$T_{ij} + \hat{s}_{12}$ sufficient $\hat{s}_{52} \leq 0$

$$v(\hat{s}_{42}) - v(\hat{s}_{12}) > 0 \quad \dots \dots \dots \quad 62$$

Uniqueness of solution can be found from

$$27 \quad \hat{s}_{11} = 62 \quad \dots \dots$$

$$\frac{1}{n^{32}} \sum_{i=1}^n \frac{\hat{s}_{ii}}{p_i} \left\{ \left(\frac{1}{a_i} - \frac{1}{b_i} \right) (s_{ij}^2 - s_{ij}^1) \right\} > 0 \quad \dots \dots \quad 63$$

The solution obtained for 63 is unique.

$$s_{ijy}^2 > s_{ijy}^1$$

Therefore if $s_{ijy}^2 > s_{ijy}^1$ then s_{ijy}^2 is unique solution.

$$p_{ijy} < \frac{1}{2} \frac{\hat{s}_{ii}}{s_{ijy}} \quad \dots \dots \quad 64$$

Further if $s_{ijy}^2 = s_{ijy}^1$ and previous equal to the value of s_{ijy} consider it for all i

$$p_{ijy} < \frac{1}{2}$$

But $\hat{s}_{12} \rightarrow \hat{s}_{52}$
 ~~$\hat{s}_{12} \rightarrow \hat{s}_{52}$~~ is sufficient if $\hat{s}_{52} \leq 0$

$$\frac{1}{n^{32}} \sum_{i=1}^n \frac{\hat{s}_{ii}}{p_i} \left\{ \left(\frac{1}{a_i} - \frac{1}{b_i} \right) (s_{ijy}^2 - s_{ijy}^1) \right\}$$

$$+ \frac{1}{n^{32}} \sum_{i=1}^n u_{ij}^2 \left(\frac{1}{a_i} - \frac{1}{b_i} \right) (s_{ijy1}^2 - u_{ij}^2 - s_{ijy}^2) \} > 0$$

The sufficient condition for $\hat{s}_{12} \rightarrow \hat{s}_{52}$ are

$$(i) \quad s_{ijz}^2 > s_{ijy}^2$$

$$(ii) \quad s_{ijz^2}^2 > s_{ijy}^2 + u_{ij}^2$$

for $j = 1, 2, \dots, l_i$

$i = 1, 2, \dots, n$.

On Simplification the two give

$$P_{ijyz} < \frac{y_{i..}}{z_{i..}} \frac{s_{ijz}}{s_{ijy}}$$

$$P_{izy} < \frac{y_{i..}}{z_{i..}} \frac{s_i}{s_{iy}}$$

Similar comparisons can be made for other counts and rest estimates. It is however important to observe that conditions given above are only sufficient to draw conclusions. In order to have a reliable conclusion appropriate to the procedure, we shall consider in addition to individual results referred to in the previous paragraph, the significance in this regard. For this purpose we shall use the data of Venkataraman and co-workers on estimate of average yield per tree in respect of relative biomass differences.

The table below gives the various estimates of average yield, their relative bias and variances.

Estimate.	Estimated average yield (in lbs.).	Relative bias.	Quantity of water
\hat{y}_{11}	53.25	-0.0027%	21.6032
\hat{y}_{12}	15.14	0.2142%	10.7930
\hat{y}_{13}	81.52	-0.0177%	52.2575
\hat{y}_{20}	50.30	-0.00592	27.7651
\hat{y}_{31}	95.21	+0.0323%	137.8919
\hat{y}_{32}	60.13	-0.00007	32.3257
\hat{y}_4	135.00	-0.0155%	338.7371
\hat{y}_{50}	65.19	-0.00245	133.893

It is observed from figure that estimate is small.
 Also it is taken to compare with the true value.
 The true amount of the estimate \hat{y}_{11} is the sum of
 of area of the three plots which is 53.25 lbs. it
 is 10.793.

C H A P T E R E.

Part. 1. Total Product.

Introduction.

Let us first consider the total product by the method of successive approximation. We start with the initial value \bar{y}_1 which is the value of the function at the origin. This is obtained by the usual procedure of differentiation, i.e., by the formula

$$\bar{y}_1 = \frac{1}{2} \left(\bar{y}(0) + \bar{y}'(0) \right).$$

We then calculate the value of the function at the point $(1, \bar{y}_1)$. It consists of two parts, one due to the linear term and the other due to the quadratic term. The quadratic term is given by

$$(\bar{y}_1)^2 - \bar{y}(0).$$

We repeat the process for the next iteration. Thus we have

$$\bar{y}_2 = \frac{1}{2} \left(\bar{y}(0) + \bar{y}'(0) + \frac{1}{2} (\bar{y}_1)^2 \right).$$

For the third iteration, we have

$$\bar{y}_3 = \frac{1}{2} \left(\bar{y}(0) + \bar{y}'(0) + \frac{1}{2} (\bar{y}_2)^2 \right).$$

For the fourth iteration, we have

$$\bar{y}_4 = \frac{1}{2} \left(\bar{y}(0) + \bar{y}'(0) + \frac{1}{2} (\bar{y}_3)^2 \right).$$

$$\text{Total product} = \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \dots$$

Let

the total product be denoted by T .

\bar{y}_i = total product secondary to i -th iteration.

\bar{y}_{ij} = $\bar{y}_i + \bar{y}_{i+1} + \dots + \bar{y}_j$ ($i, j = 1, 2, \dots$)

x_{ij} = the value of variable for the i -th unit
at the j -th second, i is the i -th unit.

x_{ij} = are of the j -th second, i is the i -th unit.

x_{ij} = the value of variable at the i -th unit
at the j -th unit.

x_{ij} = the value of variable at the i -th unit.
 j is the j -th unit.

x_i = the probability of selecting the unit i from

$$\left(\sum_{i=1}^n x_i = 1 \right)$$

the number of units in the population.

The probability x_i is measured by both the characters
and visible units, while x_{ij} is measured by
($n - n$) units in respect of the second to the j -th
character. From each of the n units, one is selected
as a unit for observing the character under study, while an additional sample size $n - n$ units is
selected in case to observe the character at long.
From each of the selected n units, a j -th x_{ij} unit is selected to measure of observing the character j only.

GENERALIZED PRACTICAL TEST.

General & Detailed Instruction.

Describes various methods of collecting information,
surveys, etc., and the method of title registration to be built up.
Detailed rules in the states are particular case of
general rules, so far as the definition

$$\begin{aligned}
 &= \frac{\frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_{i..}}{x_{i..}}}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_{i..}}} \quad ; \quad \frac{1}{n} \sum_{i=1}^n \frac{\hat{z}_{i..}}{x_{i..}} \quad \checkmark \\
 &= \frac{\frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_{i..}}{x_{i..}} \left(\frac{\sum_{j=1}^{m_i} t_{ij} \bar{z}_{j..} \binom{x_{i..}}{x_{j..}}}{\sum_{j=1}^{m_i} l_{ij} \bar{z}_{j..} \binom{x_{i..}}{x_{j..}}} \right)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_{i..}} - \frac{1}{n} \sum_{j=1}^{m_i} l_{ij} \bar{z}_{j..}} - \frac{1}{n} \sum_{i=1}^n \frac{\hat{z}_{i..}}{p_i} \\
 &= \frac{\frac{y_n}{x_n}}{x_n} \quad \text{--- ①} \\
 &\hat{t}_{ij..} = \frac{\sum_{j=1}^{m_i} t_{ij} \bar{z}_{j..} \binom{x_{i..}}{x_{j..}}}{\sum_{j=1}^{m_i} l_{ij} \bar{z}_{j..} \binom{x_{i..}}{x_{j..}}} ; \quad \hat{r}_{i..} = \frac{1}{\sum_{j=1}^{m_i} l_{ij} \bar{z}_{j..}} \\
 &\hat{r}_i = \frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_{i..}}{x_{i..}} , \quad \hat{r}_n = \frac{1}{n} \sum_{i=1}^n \frac{\hat{z}_{i..}}{x_{i..}} , \quad \hat{r}_{i..} = \frac{1}{n} \sum_{i=1}^n \frac{\hat{z}_{i..}}{x_{i..}}
 \end{aligned}$$

t_{ij} and l_{ij} are free const. s.

By suitable choosing the constant and the value of t_{ij} , l_{ij}

we get $\hat{y}_{i..} = \hat{z}_{i..}$ which is called the fit of the model in t.

Ex. 1.

Ex. 1. $\hat{r}_i = 1$, $\hat{r}_j = 1$, $t_{ij} = r_{ij}$, $l_{ij} = L_{ij}$, $r_{ij} = 1$

$\hat{y}_{i..} = \hat{z}_{i..}$

and $\hat{r}_i = \hat{r}_j$

$$T_{11} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{j}{p_i} b_{ij}(v_j - v_i)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{1}{p_i} b_{ij}} \times \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}$$

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} = n$$

$$T_{11} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{j}{p_i} b_{ij}(v_j - v_i)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}} \times \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}$$

Case III. If $v_i = v_{i..}$; $b_{ij} = t_{ij} = 1$, $v_{ij} = 1$ for $i, j \neq i..$

$$b_{ij} = v_{ij}$$

Get $v_{i..}$ as

$$T_{21} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{j}{p_i} b_{ij}(v_{i..} - v_i)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{1}{p_i} b_{ij}}, \quad v_{i..} = n$$

$$T_{21} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{j}{p_i} b_{ij}(v_{i..} - v_i)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}} \times \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}$$

Case III. If $v_i = v_{i..}$; $b_{ij} = v_{ij} = v_{i..} - v_{ij}$, $v_{ij} = 1$ for $i, j \neq i..$

$$T_{31} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{j}{p_i} b_{ij}(v_{i..} - v_i)}{\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} - \frac{1}{n} \sum_{i,j}^m \frac{1}{p_i} b_{ij}} \times \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i}$$

$$T_{z2} = \frac{\frac{1}{n} \sum_i^m \frac{i}{p_i} \frac{\sum_j^m z_{ij} b_{ij}}{\sum_j^m z_{ij}}}{\frac{1}{n} \sum_i^m \frac{i}{p_i}} \times \frac{1}{n'} \sum_i^m \frac{i}{p_i} i - n_i^2 / i$$

Case IV. $I_{ij} = A_{ij}$, $t_{ij} = z_{ij}$; $l_{ij} = l_{ij}$, $T_{ij} = 1$ for i, j, k

We get b_{ij} easily

$$T_{z1} = \frac{\frac{1}{n} \sum_i^m \frac{i}{p_i} \frac{\sum_j^m z_{ij} b_{ij}}{\sum_j^m z_{ij}}}{\frac{1}{n} \sum_i^m \frac{i}{p_i} - \frac{1}{n} \sum_i^m B_{ij}} \times \frac{1}{n'} \sum_i^m \frac{i}{p_i} i - n_i^2 / i$$

$$T_{z2} = \frac{\frac{1}{n} \sum_i^m \frac{i}{p_i} \frac{\sum_j^m z_{ij} b_{ij}}{\sum_j^m z_{ij}}}{\frac{1}{n} \sum_i^m \frac{i}{p_i} - \frac{1}{n} \sum_i^m B_{ij}} \times \frac{1}{n'} \sum_i^m \frac{i}{p_i} i - I_4$$

Case V. $I_{ij} = z_{ij}$, $t_{ij} = l_{ij} = B_{ij}$, $x_{ij} = B_{ij}$, $x_{ij} \neq 0$, $z_{ij} =$

yield of $k-i$ tree $j-i$ is $\frac{1}{n} \sum_i^m \frac{i}{p_i}$ if $r_i = i$

value of i tree $j-i$ survival time is $\frac{1}{n} \sum_i^m \frac{i}{p_i}$

$$T_{z1} = \frac{\frac{1}{n} \sum_i^m \frac{z_{ij..}}{p_i} \frac{\sum_j^m z_{ij} b_{ij}}{\sum_j^m z_{ij}}}{\frac{1}{n} \sum_i^m \frac{z_{ij..}}{p_i} - \frac{1}{n} \sum_i^m \frac{n_{ij}}{p_i}} \times \frac{1}{n'} \sum_i^m \frac{n_{ij}}{p_i} i - n_i^2 / i$$

$$\begin{aligned}
 & \frac{1}{n} \sum_{i=1}^n \frac{i..}{F_i} = \frac{\sum_{j=1}^{m_1} \frac{x_j i_o(b_{ij})}{x_j i_o(b_{ij})}}{\sum_{j=1}^{m_1} 1} + \frac{1}{n^2} \sum_{i=1}^n \frac{i..}{F_i} \text{ if } x_i = 1 \\
 & \text{and } \frac{1}{n} \sum_{i=1}^n \frac{3_{ij}}{F_i} = \frac{1}{n^2} \sum_{i=1}^n \frac{3_{ij}}{F_i} \text{ if } x_i = 1
 \end{aligned}$$

Since x_i is a random variable, $x_i = 1$ is st. if $x_i = 1$

and all now left in the expression, we get \hat{U}_1

$\hat{U}_1 = f(x) + \text{bias double sum} + \text{var}(\hat{U}_1)$

6.4 $\hat{U}_1 = \hat{U}_2 + \text{bias} + \text{var}(\hat{U}_1)$

$$\text{Let } x_1 = 1 + \varepsilon, \dots, 0$$

$$d x_1 = \varepsilon + e' \dots 0$$

$$d x_{ij} = \varepsilon + e'' \dots 0$$

$$\hat{U}_1 = U + \varepsilon + 3_{ij}$$

$\hat{U}_2 = \frac{U}{n} + \text{bias} + \text{var}(\hat{U}_2) \text{ given } \hat{U}_2 = \text{double sum}$

Substituting in \hat{U}_2 , we get $\hat{U}_2 = \frac{U}{n} + \text{var}(\hat{U}_2)$

$\hat{U}_1 = \hat{U}_2 + \text{bias} + \text{var}(\hat{U}_1)$

$$\hat{U}_1 = \frac{U + \varepsilon + e''}{n} + \text{var}(\hat{U}_1) = U \left(1 + \frac{\varepsilon}{n} + \frac{e''}{n} + \frac{\varepsilon^2}{n^2} + \frac{e''^2}{n^2} + \frac{\varepsilon e''}{n^2} - \frac{\varepsilon^3}{n^3} \right)$$

$$\left| \frac{1}{n} \right| < 1 \text{ so that } \left| \frac{\varepsilon}{n} \right| < 1, \left| \frac{e''}{n} \right| < 1, \left| \frac{\varepsilon^2}{n^2} \right| < 1, \left| \frac{e''^2}{n^2} \right| < 1, \left| \frac{\varepsilon e''}{n^2} \right| < 1$$

$$\hat{U}_1 = U + \text{var}(\hat{U}_1)$$

$$\hat{U}_1 = U + \frac{U^2}{n^2} = \frac{U^2}{U_x} + \frac{U^2}{U_x} = \frac{U^2}{U_x} = \frac{U^2}{\varepsilon^2}$$

Now $\hat{U}_1 = U + \text{var}(\hat{U}_1)$

$$\begin{aligned} \hat{\gamma}_T &= \frac{\mathbb{E}(Y_{i_1}) - \bar{y}_{\text{avg}}}{\sigma_{\epsilon}} = \frac{n - \bar{y}_{\text{avg}}}{\sigma_{\epsilon}} \\ &\quad + \left(\frac{\sigma_{\epsilon}}{\sigma_{\epsilon}} - \frac{\text{cov}(Y_{i_1}, \bar{y}_{\text{avg}})}{\sigma_{\epsilon}} \right) - \frac{\text{cov}(Y_{i_1}, \bar{y}_{\text{avg}})}{\sigma_{\epsilon}} - \frac{\text{var}(Y_{i_1})}{\sigma_{\epsilon}^2} \dots \end{aligned} \quad \dots \dots \quad 7$$

$$\begin{aligned} \bar{y}_{\text{avg}} &= \frac{1}{n} \sum_{i=1}^n \hat{x}_{i_1} \left(\frac{\hat{x}_{i_2}}{p_{i_1}} - \bar{x} \right) + \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1}}{x_{i_1}} \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) \bar{x} \quad \dots \dots \\ \text{var } \hat{s}_{i_1}^2 &= \frac{1}{n-1} \sum_{j=1}^{n-1} (\hat{x}_{i_1 j} - \bar{x}_{i_1})^2 \\ &= \frac{\sigma_{\epsilon}^2}{n} + \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1}}{x_{i_1}} \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) \bar{x} \quad \dots \dots \quad 7 \end{aligned}$$

$$\text{cov}(Y_{i_1}, \bar{y}_{\text{avg}}) = \frac{\sigma_{\epsilon}^2}{n} + \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1} \bar{x}_{i_1}}{x_{i_1}} \left\{ \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) (s_{i_1}^2 - s_{\bar{x}}^2) \right\} \quad \dots \dots \quad 8$$

$$\begin{aligned} \text{cov}(Y_{i_1}, \bar{x}_{i_1}) &= \frac{\sigma_{\epsilon}^2}{n} - \frac{\sigma_{\epsilon}^2}{n-1} \quad \dots \dots \\ \text{var } \left(\frac{\hat{x}_{i_1}}{x_{i_1}} \right) &= \frac{\sigma_{\epsilon}^2}{n} \quad \dots \dots \quad 10 \end{aligned}$$

由上式可得 $\hat{\gamma}_T$ 的表达式，即 $(1), (2), (3), (4)$ 中的 $\hat{\gamma}_T$

即 $\hat{\gamma}_T = \frac{n - \bar{y}_{\text{avg}}}{\sigma_{\epsilon}} + \left(\frac{1}{p_{\bar{x}}} - \frac{1}{p_{\bar{x}}} \right) \left(\frac{\sigma_{\epsilon}^2}{n} - \frac{\sigma_{\epsilon}^2}{n-1} \right) + \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1}}{x_{i_1}} \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) s_{i_1}^2$

由上式可知 $\hat{\gamma}_T$ 为 $(1), (2), (3), (4)$ 中的 $\hat{\gamma}_T$ 之和。

由上式可知 $\hat{\gamma}_T$ 为 $(1), (2), (3), (4)$ 中的 $\hat{\gamma}_T$ 之和。

由上式可知 $\hat{\gamma}_T$ 为 $(1), (2), (3), (4)$ 中的 $\hat{\gamma}_T$ 之和。

$$\begin{aligned} \hat{\gamma}_T &= \frac{n - \bar{y}_{\text{avg}}}{\sigma_{\epsilon}} + \left(\frac{1}{p_{\bar{x}}} - \frac{1}{p_{\bar{x}}} \right) \left(\frac{\sigma_{\epsilon}^2}{n} - \frac{\sigma_{\epsilon}^2}{n-1} \right) + \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1}}{x_{i_1}} \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) s_{i_1}^2 \\ &\quad - \frac{1}{n} \sum_{i=1}^n \frac{\hat{x}_{i_1} \bar{x}_{i_1}}{x_{i_1}^2} \left\{ \left(\frac{1}{p_{i_1}} - \frac{1}{p_{\bar{x}}} \right) (s_{i_1}^2 - s_{\bar{x}}^2) \right\} \quad \dots \dots \quad 11 \end{aligned}$$

$\beta_{ij} = \sigma_{ij}$

$$V(\hat{y}_{...}) = \left\{ \sigma^2 (y_{..} - \bar{y}_{..})^2 + \sigma^2 (\beta_{..} + 2R\text{cov}(y_{..}, \beta_{..}) - 2\bar{y}_{..}\text{cov}(y_{..}, \beta_{..}) \right. \\ \left. - 2\bar{y}_{..}\text{cov}(\beta_{..}, \beta_{..}) - 2\bar{\beta}_{..}^2) \right\} \quad12$$

$$\lambda = \frac{y}{r}$$

$\beta_{..}$ is given by

$$\beta_{..} = \frac{\sigma_u^2}{n} + \frac{1}{n} \sum_{i=1}^N \frac{a_i^2}{p_i} \left\{ \left(\frac{1}{a_i} - \frac{1}{b_{ii}} \right) s_{ii}^2 \right. \\ \left. - \frac{1}{a_i b_{ii}} \sum_{j=1}^{M_i} \left(\frac{1}{a_j} - \frac{1}{b_{ij}} \right) s_{ij}^2 \right\} \quad13$$

(Refer to chapter 17)

$y_{..} = (1^2), (2^2), (3^2), (4^2), (5^2)$ in (12) are given below

so obtain $s_{..}$ as follows,

$$V(\hat{y}_{...}) = \left[\frac{1}{n} (\sigma_u^2 + \bar{x}^2 \sigma_v^2) + \frac{1}{n} \left(\frac{2R\sigma_{uv}}{n} - R^2 \sigma_x^2 \right) - 2 \frac{R\sigma_{uv}}{n} \right. \\ + \frac{1}{n} \sum_{i=1}^N \frac{a_i^2}{p_i} \left\{ \left(\frac{1}{a_i} - \frac{1}{b_{ii}} \right) s_{ii}^2 + \frac{1}{a_i b_{ii}} \sum_{j=1}^{M_i} \left(\frac{1}{a_j} - \frac{1}{b_{ij}} \right) s_{ij}^2 \right\} \\ + \frac{\bar{x}^2}{n} \sum_{i=1}^N \frac{a_i^2}{p_i} \left(\frac{1}{a_i} - \frac{1}{b_{ii}} \right) s_{ii}^2 - \\ \left. - \frac{2R}{n} \sum_{i=1}^N \frac{a_i^2}{p_i} \left\{ \left(\frac{1}{a_i} - \frac{1}{b_{ii}} \right) (s_{ii}^2 - R s_{ii} v) \right\} \right] \quad14$$

Let us consider the variance $V(\hat{y}_{...})$ we get the following

In the case of one-dimensional time series t , $\hat{R} = \frac{\hat{J}_{xx}}{\tau_n}$

From (12), constant τ_n is given by

$$\text{Est } \tau(\hat{y}_{...}) = \left\{ \omega \tau(y_{..}) - \left(\frac{y_n}{n} \right)^2 \text{Est } \tau(y_{n'}) + \left(\frac{y_n}{n} \right)^2 \text{Est } \tau(y_0) \right. \\ \left. + 2 \frac{y_n}{x_n} \text{Est } \text{Cov}(y_{..}, x_{n'}) - 2 \frac{y_n}{x_n} \text{Est } \text{Cov}(y_n, x_n) \right\} \dots\dots 15$$

$$\begin{aligned} & \text{If } \tau^{(1,2)} = \frac{\sum_{ij} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{ij} l_{ij} \bar{z}_{ij}(b_{ij})} = \dots \}^2 \\ & = \frac{1}{n-1} \sum_i \left(\frac{x_i}{p_i} - \frac{\sum_{ij} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{ij} l_{ij} \bar{z}_{ij}(b_{ij})} \right)^2 \\ & = \frac{1}{n-1} \sum_i \left(\frac{x_i}{p_i} - \frac{1}{m'_i} \sum_j t_{ij} - \dots \right)^2 \\ & s_y^2 = \frac{1}{n-1} \sum_i \left(\frac{a_i}{p_i} - \frac{\sum_{ij} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{ij} l_{ij} \bar{z}_{ij}(b_{ij})} - y_n \right) \left(\frac{1}{p_i} - \frac{1}{m'_i} \sum_j t_{ij} - n \right) \\ & s_x^2 = \frac{1}{n-1} \sum_i \left(\frac{x_{i..}}{x_i} - \frac{1}{n} \sum_i \frac{x_{i..}}{x_i} \right)^2 \\ & s_{xy}^2 = \frac{1}{n-1} \sum_i \left(\frac{e_i}{x_i} - \frac{\sum_{ij} t_{ij} \bar{y}_{ij}(b_{ij})}{\sum_{ij} l_{ij} \bar{z}_{ij}(b_{ij})} - y_n \right) \left(\frac{e_{i..}}{x_i} - \frac{1}{n} \sum_i \frac{e_{i..}}{x_i} \right) \quad \dots 16 \end{aligned}$$

then $\tau(\hat{y}_{...}) = \tau(y_{..})$; $\tau(\hat{x}_{..}) = \tau(x_{..})$; $\text{Cov}(\hat{y}_{..}, \hat{x}_{..}) = \text{Cov}(y_{..}, x_{..})$; $\text{Cov}(\hat{x}_{..}, \hat{x}_{..}) = \tau(x_{..})$;

$$\text{Cov}(\hat{y}_{..}, \hat{x}_{..}) = \text{Cov}(y_{..}, x_{..})$$

It follows that $\hat{y}_{...}$ is unbiased if and only if variance $V(\hat{y}_{...})$ is given by

$$\begin{aligned} \text{Est. } V(\hat{y}_{...}) &= \left(\frac{1}{n} \left(s_y^2 + \frac{s_x^2}{2} - \frac{s_{xy}^2}{n} \right) - \frac{s_{yy}}{n} \right) \\ &\quad + \frac{1}{n} \left(2 \frac{s_{xy}}{n} s_y^2 - \frac{s_{yy}}{n} s_x^2 \right) \} \end{aligned} \quad \dots\dots .17$$

EFFICIENT ESTIMATES

The variance of the unbiased $\hat{y}_{...}$ can be written as

$$\begin{aligned} V(\hat{y}_{...}) &= V(y_1) + \left(\frac{1}{n} - \frac{1}{n'} \right) (R\bar{v}_y^2 - 2R\bar{v}_x) \\ &\quad + \frac{1}{n} \sum_{i=1}^{N-1} \frac{s_i^2}{E_i} \left(\frac{1}{E_i} - \frac{1}{L_i} \right) (R\bar{s}_{ix}^2 - \frac{R\bar{s}_{ii}^2}{E_i}) \\ &\quad + \frac{2R}{n} \sum_{i=1}^{N-1} \frac{s_i}{E_i} \frac{\bar{v}_i}{L_i} \left(\frac{1}{E_i} - \frac{1}{L_i} \right) s_i \\ &= V(y_1) + \left(\frac{1}{n} - \frac{1}{n'} \right) 2R\bar{v}_x^2 (\frac{\sigma^2}{\bar{v}_y^2} - R^2) \\ &\quad - \frac{1}{n} \sum_{i=1}^{N-1} \frac{s_i^2}{E_i} \left(\frac{1}{E_i} - \frac{1}{L_i} \right) 2R\bar{s}_{ix}^2 - \frac{1}{n} \left(\frac{2R}{\sum_{i=1}^{N-1} \frac{s_i^2}{E_i}} - \frac{2R}{\sum_{i=1}^{N-1} \frac{s_i^2}{L_i}} \right) R^2 \\ &\quad + \frac{2R}{n} \sum_{i=1}^{N-1} \frac{s_i}{E_i} \frac{\bar{v}_i}{L_i} \left(\frac{1}{E_i} - \frac{1}{L_i} \right) s_i \end{aligned} \quad \dots\dots .18$$

First we note that $V(\hat{y}_{...})$ is minimum by 3rd measure

$\sigma \pi(\hat{\gamma}_{ij})$ is a convex function of γ_{ij} .

$\therefore t - t' = t\gamma_{j_1}, \gamma_{j_1} \geq 0 \Rightarrow t + t' = (1^n)$

All the sum. Also the constraint is $t = t'$.

Thus

$$(+) P_t > \frac{1}{2} - \frac{1}{n}$$

$$(+) P_{t_i} > \frac{1}{2} - \frac{1}{\frac{n-i}{n}} \quad i = 1, 2, \dots, n$$

\therefore $t = \frac{1}{2}n$, $t_i = \frac{n-i}{n}$ is the solution.

t is the estimate of γ_{ij} .

6.6

PART II: 3.330

By direct substitution of γ_{ij} in the objective function, we get

$\gamma_{ij} = \hat{\gamma}_{ij}$ is called the optimal solution of the problem.

and substitution of $\gamma_{ij} = \hat{\gamma}_{ij}$ in the constraint, we get

variables t_{ij} in $t_{ij} = \sum_{k=1}^n \gamma_{ik}$ are non-negative integers, and $t_{ij} = \sum_{k=1}^n \gamma_{kj}$ are also non-negative integers. Thus t_{ij} is a non-negative integer variable. Also $t_{ij} \leq t$ and $t_{ij} \geq 0$. This is the constraint of the problem. Now we have to find the total probability $P_{t_{ij}}$, $t_{ij} = t$ and $t_{ij} \geq 0$.

Let $t = P_{11}$

$$\text{For } i = 1, \gamma_{i1} = l_{11}, l_{1j} = 1, l_{ij} = \gamma_{ij}, \gamma_j = \sum_{i=1}^n l_{ij}$$

$$i_j = 1 \quad \text{if } i, j, \dots, i_r = \dots; G_i \quad \bar{i}_{ij}, v_{ij} = i_r;$$

$$1 \quad \bar{i}_j = \bar{i}_j(\dots)$$

$\alpha = 11 \dots$

$$\bar{\gamma}(T_{11}) \approx \left(\frac{1}{\alpha} - \frac{1}{r} \right) \left(\frac{\sigma_1^2}{\alpha} + \frac{\sigma_r^2}{r} \dots \right) + \sum_{k=2}^{\infty} \left(\frac{1}{\alpha} - \frac{1}{r} \right) \left(\frac{j^2}{\alpha^2} - \frac{j^2}{r^2} \dots \right)$$

..... 10

$$1 \quad \bar{i}_j = \frac{p_{ij}}{p_i} \quad \text{if } i = 1, \dots, r; \quad \bar{i}_{ij} = \bar{i}_j \quad \text{if } i > r.$$

$$(1) \quad \bar{i}_j = \frac{p_{ij}}{p_i} \quad \text{if } i = 1, \dots, r; \quad \bar{i}_{ij} = \frac{v_{ij}}{v_i} \quad \text{if } i > r.$$

vari 1.

$$(1) \quad \bar{i}_j = \frac{p_{ij}}{p_i} \quad \text{if } i = 1, 2, \dots, r; \quad \bar{i}_{ij} = \frac{v_{ij}}{v_i} \quad \text{if } i > r.$$

li $\gamma = p_{ij} + v_{ij} \quad \text{if } i = 1, \dots, r$

For (1a), we will use $\bar{\gamma}(T_{11})$ in (1).

$$\bar{\gamma}(T_{11}) = \left(\frac{1}{\alpha} \left(\sigma_1^2 + \dots + \sigma_r^2 \right) - \bar{\sigma}_B^2 \dots \bar{\sigma}_r^2 \right) - \frac{1}{\alpha^2} \dots \left(\bar{\sigma}_B^2 \dots \bar{\sigma}_r^2 \right)$$

$$+ \frac{1}{\alpha} \sum_{i=2}^r \frac{p_{ij}}{p_i} \left(\frac{1}{\alpha} - \frac{1}{r} \right) \left(\frac{\sigma_1^2}{\alpha} + \dots + \frac{\sigma_{i-1}^2}{\alpha} - \frac{\sigma_i^2}{r} \dots \frac{\sigma_r^2}{r} \right)$$

$$+ \frac{1}{\alpha} \sum_{i=2}^r \frac{p_{ij}}{p_i} \frac{1}{\sum_{j=1}^r p_{ij}} \left(\frac{1}{\alpha} - \frac{1}{r} \right) \left(\frac{\sigma_1^2}{\alpha} + \dots + \frac{\sigma_{i-1}^2}{\alpha} - \frac{\sigma_i^2}{r} \dots \frac{\sigma_r^2}{r} \right) \dots \dots \dots$$

If $v_{ij} = 0$, $\bar{i}_{ij} = 0$, $\bar{i}_j = \frac{p_{ij}}{p_i}$, $\bar{i}_{ij} = \frac{v_{ij}}{v_i}$, $\bar{i}_j = \frac{p_{ij}}{p_i}$ finite

or $v_{ij} \neq 0$ obtain $v_{ij} = 0$, $\bar{i}_{ij} = 0$, $\bar{i}_j = \frac{p_{ij}}{p_i}$, $\bar{i}_{ij} = \frac{v_{ij}}{v_i}$

....80..

$$V(\gamma_{ij}) = \left\{ \frac{1}{n} (\sigma_j^2 - \bar{\sigma}_{ij} \sigma_i^2) - \bar{\sigma}_i \sigma_j + \frac{\bar{\sigma}_{ij}}{n}; \bar{\sigma}_{ij} = \frac{\sigma_{ij}^2}{n} \right\}$$

$$+ \frac{1}{n} \sum_{i=1}^N - \frac{1}{n} - \left(\frac{1}{n} - \frac{1}{n} \right)$$

$$- \frac{1}{n} \sum_{i=1}^N \frac{\sigma_i^2}{n} = - \sum_{j=1}^{M_i} \frac{\sigma_{ij}^2}{n} \quad \dots \dots \quad$$

$$\bar{\sigma}_i^2 = \frac{\sigma_i^2}{n} + \frac{\sigma_{ij}^2}{n} - \frac{1}{n} \sum_{i=1}^N \frac{\sigma_i^2}{n} \quad !$$

$$+ \frac{1}{n} \sum_{i=1}^N \frac{\sigma_i^2}{n} - \frac{1}{n} \sum_{j=1}^{M_i} \frac{\sigma_{ij}^2}{n} \quad \dots \dots \quad$$

$$\bar{\sigma}_i^2 = \bar{\sigma}_i^2 + \frac{1}{n} \sum_{j=1}^{M_i} \frac{\sigma_{ij}^2}{n} - \bar{\sigma}_i^2 = \sigma_{ij}^2$$

$$\bar{\sigma}_i^2 = \frac{1}{n} \sum_{j=1}^{M_i} \frac{\sigma_{ij}^2}{n} = \sigma_{ij}^2$$

$$\frac{a_1}{a} + \frac{a_2}{a} + \frac{a_3}{a} + \frac{a_4}{a-b}$$

$$\bar{\sigma}_i^2 = \sigma_{ij}^2 - \sigma^2$$

$$\bar{\sigma}^2$$

$$\sum_{i=1}^N \frac{\sigma_i^2}{n}$$

$$\sum_{i=1}^N \frac{\sigma_i^2}{n} = \frac{1}{n} \sum_{j=1}^{M_i} \frac{\sigma_{ij}^2}{n}$$

$$\frac{n}{15}$$

$$(I_1) = \left(\frac{1}{n} - \frac{1}{n} \right) \left(\frac{\sigma^2}{n} - \frac{\sigma^2}{n} \right)$$

..... 13

... 71 ..

$$\text{From } (11), \quad P_1 = \frac{\sigma_3}{y}$$

From (11), $i = 1, 2, 3$

$$\begin{aligned} r(T_{11}) &= \left(\frac{1}{n} \bar{y}^2 + \bar{y}^2 \sigma_{\bar{y}}^2 - \bar{y}^2 \sigma_{\bar{y}}^2 \right) + \left(2\sigma_{\bar{y}}^2 - \bar{y}^2 \sigma_{\bar{y}}^2 \right) * \frac{\sigma_3^2}{n}, \\ &+ \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} \left(\frac{1}{x_i} - \frac{1}{\bar{x}} \right) S_{ij}^2 \\ &+ \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} \left(\frac{1}{x_i} - \frac{1}{\bar{x}} \right) \sum_{j=1}^m \frac{1}{x_{ij}} \left(\frac{1}{x_{ij}} - \frac{1}{\bar{x}_{ij}} \right)^2 S_{ij}^2 \quad \dots\dots 24 \end{aligned}$$

Similar to T_{11}

$$S_{ij} = 1, \quad x_i = \bar{x}_i, \quad 1_{ij} = v_{ij} = 1, \quad \bar{x}_{ij} = 1 \text{ for all } i, j, r$$

$$r(T_{22}) = r_{1\dots} = \bar{y}^2 - \bar{y}_i^2 (1_i)$$

From (11), $r(T_{22})$ gives $r(T_{22}) = \dots$

$$r(T_{22}) = \sum_{i=1}^n \frac{i \cdot (\bar{y}_i (1_i) - \bar{y}_{i\dots})}{y_{i\dots}} + \left(\frac{1}{n} - \frac{1}{n} \right) \left(\frac{\sigma_1^2}{\bar{y}_{1\dots}} - \frac{\sigma_{\bar{y}}^2}{\bar{y}^2} \right) \quad \dots\dots 25$$

Theorem 2.2.2 is proved by induction. Now $r(T_{22}) = \dots$

Now \bar{y}_i is equal to unit, i.e., $\bar{y}_i = 1$, then

$$r(T_{22}) = \sum_{i=1}^n \frac{i \cdot (\bar{y}_i (1_i) - \bar{y}_{i\dots})}{y_{i\dots}} + \left(\frac{1}{n} - \frac{1}{n} \right) \left(\frac{\sigma_1^2}{\bar{y}^2} - \frac{\sigma_{\bar{y}}^2}{\bar{y}^2} \right)$$

Therefore $r(T_{22}) = \dots$ $\therefore r(T_{22}) = \dots$

$$(i) \quad r(T_{22}) = \bar{y}_{ij}, \quad \text{and } r(T_{22}) = r(T_{11}) P_1 = \frac{\sigma_3}{y}$$

Using (14) we get $r(T_{22}) = r(T_{11}) P_1$

$$v(T_{32}) = \left(\frac{1}{n} (\sigma_{y_1}^2 + \bar{y}_1^2 \dots \sigma_{y_n}^2) - \bar{y}_1 \dots \bar{y}_n \right) + \bar{y}_1 \dots \frac{1}{n} (2\sigma_{y_1}^2 - \bar{y}_1 \dots \bar{y}_n)$$

$$+ \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} \left(\frac{1}{k_i} - \frac{1}{\bar{k}_i} \right) s_{i..}^2 + \frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} \frac{1}{k_i} \frac{1}{\bar{k}_i} \sum_{j=1}^{M_i} \left(\frac{1}{k_{ij}} - \frac{1}{\bar{k}_{ij}} \right) s_{ij..}^2$$

.....25

Estimate T_{32}

For t^i i.e.

$$t^i = \bar{B}_{i..}, \quad l_{ij} = t_{ij} = \bar{t}_{ij}, \quad r_{ij} = \bar{r}_{ij}; \quad s_{ij..}^2 = 1 \quad \text{for all } i, j, k$$

$$g_{ij} = \bar{y}_i(\bar{k}_i); \quad \bar{B}_{i..} = \bar{B}_i \quad \text{therefore } \bar{R}_i = \bar{y}_{i..} \quad .$$

$$\text{and } v(y_{ij}) = \sum_{i=1}^n y_{i..} \left(1 + \left(\frac{1}{m_i} - \frac{1}{\bar{k}_i} \right) \left(s_{i..}^2 + \frac{s_{i..}^2 B}{\bar{y}_{i..}} \right) \right) = y_{i..} \dots$$

From (12) the r.l. of $v(y_{ij})$ in the estimate T_{32} is $v(y_{ij})$

$$v(T_{32}) = \sum_{i=1}^n \frac{y_{i..}}{p_i} \left(\left(\frac{1}{k_i} - \frac{1}{\bar{k}_i} \right) \left(s_{i..}^2 + \frac{s_{i..}^2 B}{\bar{y}_{i..}} \right) \right. \\ \left. + \left(\frac{1}{n} - \frac{1}{n} \right) \left(\frac{\sigma_{y_i}^2}{k_i^2} - \frac{\sigma_{y_i}^2 B}{\bar{y}_{i..}^2} \right) \right) \quad \dots \dots 27$$

It can be seen that for every i ($i=1, 2, \dots, n$) the regression

of y_{ij} on $\bar{y}_{i..}$ will be $\sigma_{y_i}^2$ times the $\bar{y}_{i..}$ + error term, then it turns out that $v(y_{ij}) = y_{i..}$ & $v(T_{32}) = \left(\frac{1}{n} - \frac{1}{n} \right) \left(\frac{\sigma_{y_i}^2}{B^2} - \frac{\sigma_{y_i}^2 B}{\bar{y}_{i..}^2} \right)$

and it will vanish if the regression $\frac{y_{i..}}{p_i}$ on $\frac{y_{i..}}{p_i}$ is linear and passes through the origin.

.....83..

It can be seen that variance $V(T_{32})$ is given by

$$V(T_{32}) = \left[\frac{1}{n} (\sigma_{y_2}^2 + \sum_{i=1}^m \sigma_{x_i}^2 - 2\bar{y}_2 \sigma_{x_2}) + \sum_{i=1}^m \frac{1}{p_i} (2\sigma_{y_2}^2 - \bar{y}_2 \sum_{j=1}^{n_i} \frac{\sigma_{x_j}^2}{p_{ij}}) \right. \\ \left. + \frac{1}{n} \sum_{i=1}^m \frac{p_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{p_i} \right) s_{iy_2}^2 + \frac{1}{\sum_{i=1}^m p_i} \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{s_{ij}^2}{p_{ij}} \left(\frac{1}{p_{ij}} - \frac{1}{p_i} \right) s_{ij}^2 \right] \dots\dots 28$$

6.7

COMPARISON OF VARIANCES.

Estimate T_{12} against estimate T_{11}

The estimate T_{12} will be more efficient than T_{11} if

$$V(T_{12}) - V(T_{11}) > 0$$

or if $\frac{1}{n} \sum_{i=1}^m \frac{p_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{p_i} \right) (2s_{iy_2}^2 - \bar{y}_2 \sum_{j=1}^{n_i} s_{ij}^2) > 0$
(from 20 & 21)

Having p_i to be large enough so that first correction to s_{iy_2} can be ignored, let $m_i = 1$ $p_i = \frac{B_i}{3}$

$$\text{or if } 2 \sum_{i=1}^m p_i s_{iy_2}^2 > \bar{y}_2 \sum_{i=1}^m p_i s_{ij}^2$$

$$\text{or if } \frac{\bar{s}_{iy_2}^2}{2} > \bar{y}_2 \sum_{i=1}^m p_i s_{ij}^2$$

or if $\bar{\rho}_w > \frac{1}{2} \frac{\bar{s}_{iy_2}^2}{\sum_{i=1}^m p_i s_{ij}^2} = \frac{1}{2} \frac{\bar{s}_{iy_2}^2}{\bar{s}_{ij}^2}$ 29

Estimate T_{12} thus. estimate T_{11}

The estimate T_{11} is better than the estimate T_{12}

$$\text{if } v(T_{\gamma}) - v(T_{\gamma_1}) > 0 \quad \dots \dots \dots$$

and this can be done if $v(T) < v(T_{\gamma_1})$ for $\gamma < 26.8$ in (30)

22

that is to say the b_{ij} 's estimate of y_{ij} , i.e. population

to be so that it is not far than the inequality (30) will

be true.

$$\text{if } \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{b_{ij}} \right)^2 s_{ij}^2 + \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \times \frac{1}{b_{ij}} \times \frac{1}{p_i} \sum_{j=1}^{m_i} \left(\frac{1}{r_{ij}} - \frac{1}{b_{ij}} \right) s_{ij}^2$$

$$+ \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{b_{ij}} \right)^2 (s_{ij}^2 + s_{ij}^2 s_{ij}^2 + s_{ij}^2 s_{ij}^2)$$

$$- \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \times \frac{1}{b_{ij}} \times \frac{1}{p_i} \sum_{j=1}^{m_i} (1 - u_{ij}^2) \left(\frac{1}{r_{ij}} - \frac{1}{b_{ij}} \right) s_{ij}^2 > 0 \quad \dots \dots \dots$$

$$\text{or } \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{b_{ij}} \right) (s_{ijy_1}^2 - s_{ijy_1}^2 s_{ijy_1}^2 + s_{ijy_1}^2 s_{ijy_1}^2)$$

$$- \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \times \frac{1}{b_{ij}} \times \frac{1}{p_i} \sum_{j=1}^{m_i} (1 - u_{ij}^2) \left(\frac{1}{r_{ij}} - \frac{1}{b_{ij}} \right) s_{ijy_1}^2 > 0$$

To refer to the first one T_{γ_1} , T_{γ_1} is the merit

$$(i) \quad \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \left(\frac{1}{m_i} - \frac{1}{b_{ij}} \right) (s_{ijy_1}^2 - s_{ijy_1}^2 s_{ijy_1}^2 + s_{ijy_1}^2 s_{ijy_1}^2)$$

$$(ii) \quad \frac{1}{n} \sum_{i=1}^n \frac{v_i^2}{p_i} \times \frac{1}{b_{ij}} \times \frac{1}{p_i} \sum_{j=1}^{m_i} (1 - u_{ij}^2) \left(\frac{1}{r_{ij}} - \frac{1}{b_{ij}} \right) s_{ijy_1}^2$$

If u_{ij}^2 is very small, much less than i ($i=1, 2, \dots, n$),

then the second one is probably larger than the first one.

So we can get the following conclusions if b_{ij} is large, well

as the merit of the third one is better than the first one.

∴ T₁₁ will be efficient over T₃₂ if for all i

$$\text{or } \rho_{i,j} > \frac{s_{ij}^2}{s_{i,j}^2 + s_{i,i}^2 - 2s_{i,j}} \quad \text{or } \rho_{i,j} = \frac{s_{ij}^2}{s_{i,i}^2 - s_{i,j}^2}$$

$$\underline{\text{SITR } \neq \text{ T}_{32} \text{ GTSR SITR } \neq \text{T}_{11}}$$

The estimate T₁₁ will be more efficient than T₃₂

$$\text{if } V(T_{32}) - V(T_{11}) > 0$$

Substituting the expression of V(T₃₂) and V(T₁₁) in (29) and (30)

in (32), we find that condition (32) is equivalent to

possible if but if it is not, is not possible then T₁₁ is (32)

will hold if

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{p_i} \left(\frac{1}{s_{i,i}^2} - \frac{1}{s_{i,j}^2} \right) \left(s_{i,j}^2 - (s_{i,j}^2 + s_{i,i}^2 - 2s_{i,j}) \right) > 0$$

.....35

If condition (32) is don't hold for every i (i = 1, 2, .. n)

then the estimate T₁₁ will be more efficient than T₃₂ for each i

$$\rho_{i,j} > \frac{s_{ij}^2}{s_{i,i}^2 + s_{i,j}^2 - 2s_{i,j}}$$

....86..

similar companies can also be referred to other activities. It is however important to observe that the following above analysis sufficiant to no error conclusion intended. In order to have some guidance in the choice of an appropriate sampling procedure, we shall once again consider the double line. It is referred to in the previous chapters on give some indication of the two purposes which will use the data of Venkateswara.
For the various estimates of total reduction in sampling error, standard errors, variances and the efficiency of double sampling or random sampling.

The table below gives the various estimates of total error, i.e., standard errors and their variances and the efficiencies of double sampling and random sampling.

Estimates	Relative bias.	Estimated variance (in 10^2)	Efficiency	Estimate of total error (in Rs.)
\bar{m}_{11}	-0.2057	8,221,126,540	17.7	13,501,354
\bar{m}_{12}	-0.0419	11,462,72,557	13.7	2,712,512
T_{20}	1.77	2,420,425,177	211.7	12,52,327
T_{30}	-0.4026	2,771,45,505	21.	13,071,113

Total sample size = 3 relative bias = 1.77 estimated error = 13,501,354

small. This is a little more than $\frac{1}{n}$ of the cost of T_{11} .

Another method of obtaining the tree branching trees as an auxiliary variable is described in the paper by G. R. Wilson.

Suppose x_1, x_2, \dots, x_n are the n nodes of T_{11} to T_{11}^* .

First introduce the auxiliary variables

$v_i = \sum_{j=1}^n p_j x_j$ for $i = 1, 2, \dots, n$

where p_i is the probability of node i being the root of T_{11} .

Maximize v_i over all possible T_{11} to T_{11}^* subject to $v_i \leq v$.

It is obvious that the maximum value of v_i is $\frac{n}{n+1}$.

Similarly, $v_1 + v_2$ is the maximum value of $v_1 + v_2$ for T_{11} .

And so on. In other words, we can obtain the auxiliary variables v_1, v_2, \dots, v_n by solving the equations

$$v_i = v_1 T_{11} + v_2 T_{11}^* + \dots + v_n T_{11}^{n-1}$$

$$\text{where } T_{11} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{p_i}{x_i} - \frac{1}{n} \sum_{j=1}^m \frac{p_j}{x_j} v_{ij} (v_{ij})'}{\frac{1}{n} \sum_{i=1}^n \frac{p_i}{x_i} - \frac{1}{n} \sum_{j=1}^m v_{ij}}$$

$$v_i = \frac{\eta(v_{11}) - \text{cov}(T_{11}, v_{11})}{\eta(T_{11}) - \text{cov}(v_1, v_{11}) + \eta(v_{11})}$$

$$v_0 = 1 - v_1$$

In order to obtain the equations, we must consider the following two types of situations. In the first case, we wish to estimate the total number of elements in the set of auxiliary variables.

Est. No.	Estimate of to 1,700 (K.s.)	Standard error	Rel. eff.
J	1,174,532	0.0000	100
T_{11}	10,774,751	~0.157	1.7
T'_{11}	21,601,51	~0.157	1.7
T	15,121,425	~0.1567	2.3

$$_1 = 0.5513$$

$$w_2 = 0.3127$$

It is seen above that we are likely to gain efficiency if two auxiliary variables are used in comparison to the total production. However, the relative efficiency when there are two auxiliary variables is only 23% as compared to 87% when the number of trees as an auxiliary variable. It is seen, both the auxiliary variables, are used there is appreciable gain in efficiency which is of the order of 38%.

6.8

EFFICIENCY OF DOUBLE SAMPLING IN RELATION TO COST.

It is of considerable interest to know the various estimates corresponding to the cost of survey. This will enable us to find out the most economic way of observing the auxiliary variable is worth the cost of survey.

In this case the cost function in relation to the survey cost is $C = T \cdot J \cdot N$. In this case we have to take into account the cost of surveying the auxiliary variable.

of w_i is selected at second and third stage of sampling to observe the character y under study. We shall now consider the various items that make up the total cost of the survey.

Denote by

C_1 , the cost of collecting all basic information selected primary stage unit, including cost of travel.

C_2 , the additional cost incurred in each primary stage unit, for travelling and collecting information necessary for observing the character y .

C_3 , the cost of preliminary operations involved in 2nd stage units including travel from orchard to orchard.

C_4 , the cost of making actual observations in each tertiary unit for the character, cost of tabulation and analysis per tertiary unit.

The total cost of the survey can then be expressed in the form

$$C = C_1 n' + C_2 n + C_3 n m + C_4 n \cdot b \quad \dots \dots 1$$

Further if we assume that σ_{ij} and r_i are large enough so that finite population correction factors can be ignored, then from (2)

$c = t \cdot V$.

$$\begin{aligned} V(T_{11}) &= \frac{1}{n'} (\bar{\sigma}_y^2 - \bar{\sigma}_i^2) + \frac{1}{n} \bar{\sigma}_i^2 + \frac{1}{m} \sum_{i=1}^m \frac{n_i^2}{r_i} - \frac{1}{r_i} \\ &+ \frac{1}{n} \sum_{i=1}^m \frac{B_i^2}{r_i} - \frac{1}{n} \sum_{i=1}^m u_i^2 - \frac{1}{n} \sum_{i=1}^m v_i^2 \quad \dots \dots 2 \\ &= \frac{A_1}{n'} + \frac{A_2}{n} + \frac{A_3}{m} + \frac{A_4}{n \cdot m} \end{aligned}$$

where A_1, A_2, A_3, A_4 are independent of n' , n , m & b

$$u = \sqrt{\frac{1}{q_1}} + \sqrt{\frac{1}{q_2}} + \sqrt{\frac{1}{q_3}} + \sqrt{\frac{1}{q_4}}$$

$$\sqrt{-\frac{1}{q_1}}$$

$$\sqrt{-\frac{1}{q_2}}$$

$$v = \sqrt{\frac{1}{q_1} + \sqrt{\frac{1}{q_2} + \sqrt{\frac{1}{q_3} + \sqrt{\frac{1}{q_4}}}}}$$

$$= (q_1^{-1/2} - q_2^{-1/2} - q_3^{-1/2} - q_4^{-1/2}) \dots$$

$$r(v) = \frac{\sigma^2}{n} - \frac{1}{nb} \sum_{i=1}^N \frac{y_i^2}{x_i} - \frac{1}{n} \sum_{j=1}^{Mc} \beta_j^2 + \frac{1}{n} \sum_{i=1}^N \frac{\beta_i^2 S_{ij}^2}{p_i} \dots$$

$$= \frac{\sigma^2}{n} - \frac{1}{n} \sum_{i=1}^N \frac{y_i^2}{x_i} + \dots$$

$$\frac{\sigma^2}{n}$$

$$= \sum_{i=1}^N \frac{y_i^2}{x_i} - \frac{1}{n}$$

To find the best estimate of $\theta_1, \theta_2, \theta_3$

minimise $\sum \rho_i^2$ w.r.t $\theta_1, \theta_2, \theta_3$ given θ_0

$$\begin{aligned} \text{Ansatz: } & \frac{\partial}{\partial \theta_1} \sum \rho_i^2 = \sqrt{(\rho_1')^2 + (\rho_2')^2 + (\rho_3')^2} \\ & = \sqrt{\frac{\theta_1^2}{\theta_0^2} - \frac{\theta_1 \theta_2}{\theta_0^2}} \quad \text{Ansatz: } \frac{\partial}{\partial \theta_2} \sum \rho_i^2 = \sqrt{\frac{\theta_2^2}{\theta_0^2} - \frac{\theta_1 \theta_2}{\theta_0^2}} \\ & \quad \vdots \quad \vdots \quad \vdots \end{aligned} \quad \dots \dots \dots$$

on differentiating w.r.t θ_1, θ_2 and from (2) & (3),

we find that the optimum values are given by

$$\theta_{1,2} = \frac{\left(\sqrt{(\rho_1')^2 + (\rho_2')^2} + \sqrt{\rho_3' \rho_2} + \sqrt{\rho_1' \rho_3} \right)}{\rho_0} \quad \dots \dots \dots$$

To find the optimum value of θ_3 we have

ρ_3' is proportional to $\theta_1 + \theta_2 + \theta_3$

$$\theta_3 = \frac{\left(\sqrt{(\rho_1')^2 + (\rho_2')^2} + \sqrt{\rho_1' \rho_2} + \sqrt{\rho_1' \rho_3} \right)}{\sqrt{\rho_1' + \sqrt{\rho_1' \rho_2} + \sqrt{\rho_1' \rho_3} + \sqrt{\rho_2' \rho_3}}} \quad \dots \dots \dots 10$$

To find the optimum value of θ_0

it is required to find the optimum value of $\rho_1', \rho_2', \rho_3'$ which is obtained

by substituting the values of θ_1, θ_2 in the observation equation

on eliminating $\theta_1 + \theta_2 + \theta_3$. In this case, we consider

a case where $\theta_1 = \theta_2 = \theta_3$

$$\theta = \theta_1 + \theta_2 + \theta_3 \quad \text{where } \theta_3 = \theta_1 = \theta_2 \quad \dots \dots \dots 11$$

To find the optimum values of $\theta_1, \theta_2, \theta_3$ we minimize the variance $\sigma^2(\theta_1)$ i.e. the total sum of the squares is fixed to be ρ_0

Consider

$$\phi = \frac{1}{r^1} + \frac{2}{r^2} + \frac{2}{r^3} + \frac{1}{r^4} + \lambda' C_1 n^1 + C_2 n^2 + C_3 n^3 + C_4 n^4 \dots \dots 12$$

Differentiating w.r.t. r^1 , we get $C_1 n^1$

Like wise for r^2, r^3, r^4

$$\frac{\partial \phi}{\partial r^1} = \text{div } C_1 n^1 = \sqrt{\frac{A_1 C_1}{\lambda}} \dots \dots 13$$

$$\frac{\partial \phi}{\partial r^2} = 0 \quad \text{div } C_2 n^2 = \sqrt{\frac{A_2 C_2}{\lambda}} \dots \dots 14$$

$$\frac{\partial \phi}{\partial r^3} = 0 \quad \text{div } C_3 n^3 = \sqrt{\frac{A_3 C_3}{\lambda}} \dots \dots 15$$

From (13), (14) and (15) we get

$$\lambda = \frac{\{\sqrt{A_1} + \sqrt{A_2} + \sqrt{A_3}\}^2}{C_0^2} \dots \dots 16$$

Substituting value of λ , in eqn. (12), (13), (14) and (15) in the expression for the variance, $V(T_{11})$ we get

$c = \frac{1}{n}$ & 1.

clearly

$V(T_{11})$ is minimized when n is maximum subject to the conditions

$$C_1 n^1 = \sqrt{\frac{A_1}{\lambda}} ; \quad C_2 n^2 = \sqrt{\frac{A_2}{\lambda}} ; \quad C_3 n^3 = \sqrt{\frac{A_3}{\lambda}}$$

where n is such that $n^1 = 1$

Therefore

$$\begin{aligned} n^1 &= \sqrt{\frac{A_1}{\lambda}} - \frac{1}{I} & \text{where } L = \sqrt{A_1} n^1 + \sqrt{A_2} n^2 + \sqrt{A_3} n^3 \\ n^2 &= \sqrt{\frac{A_2}{\lambda}} - \frac{C_2}{I} \\ n^3 &= 1 - \frac{C_3}{I} \\ n &= I - \frac{C_0}{I} \\ b &= \sqrt{\frac{A_1 A_2}{A_2 C_4}} \end{aligned} \quad \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \dots \dots 17$$

and the objective function is given by

$$V(T_{11})_{opt} = \frac{(\sqrt{\beta_1} + \sqrt{\beta_2} + \sqrt{\beta_3}) (\sqrt{\alpha_1} + (\frac{\beta_1}{\alpha_1} + \frac{\beta_2}{\alpha_2} + \sqrt{\beta_3} + \sqrt{\beta_4})}{C_1} \quad \dots \dots 17$$

Now we have to find the value of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ so that $V(T_{11})_{opt}$ is minimum.

$$C_1 = C_1 n + C_2 n^2 + C_3 n^3 \quad \dots \dots 18$$

Since x_1, x_2, x_3, x_4 are variables in $V(T_{11})_{opt}$ it is given by

$$V(y_a) = \frac{x_1}{n} + \frac{x_2}{n^2} + \frac{x_3}{n^3} \quad \dots \dots 19$$

If x_1, x_2, x_3, x_4 are the values of x_1, x_2, x_3, x_4 respectively, then x_1, x_2, x_3, x_4 are given by

$$x_1 = \frac{\beta_0}{L'} + \sqrt{\frac{\beta_2}{C_1}} \quad \dots \dots 20$$

$$x_2 = \sqrt{\frac{\beta_1 \beta_2}{A_1 A_2}} - \frac{\beta_1}{A_2} \quad x_3 = \sqrt{\frac{\beta_1 \beta_2}{A_1 A_2}} - \frac{\beta_2}{A_1} \quad x_4 = \sqrt{\frac{\beta_1 \beta_2}{A_1 A_2}} \quad \dots \dots 21$$

$$x_1 + x_2 + x_3 + x_4 = \sqrt{C_1 A_1} + \sqrt{C_2 A_2} + \sqrt{C_3 A_3} + \sqrt{C_4 A_4}$$

and the optimum value is given by

$$V(T_{11})_{opt} = \frac{(\sqrt{\beta_1} + \sqrt{\beta_2} + \sqrt{\beta_3} + \sqrt{\beta_4})^2}{C_0} \quad \dots \dots 22$$

Therefore the objective function is given by

subject to $x_1 + x_2 + x_3 + x_4 = b$

$$R.E. = \frac{(\sqrt{\beta_1} + \sqrt{\beta_2} + \sqrt{\beta_3} + \sqrt{\beta_4})^2}{(\sqrt{C_1} A_1 + \sqrt{C_2} A_2 + \sqrt{C_3} A_3 + \sqrt{C_4} A_4)} \quad \dots \dots 23$$

The table below gives the value of T_{11} subject to $x_1 + x_2 + x_3 + x_4 = b$

different value $c^e \frac{C_1}{C_2}$, $\frac{C_3}{C_2}$ and $\frac{C_4}{C_2}$

Table: Ratio of efficiency of T_{11} to T_3 , to fixed cost of survey $\frac{C_4}{C_2}$

$$\frac{C_4}{C_2}$$

			0.01			0.04			0.09		
			$\frac{C_1}{C_2}$			$\frac{C_3}{C_2}$			$\frac{C_4}{C_2}$		
			0.25	0.3	0.	0.25	0.36	0.5	0.25	0.76	0.19
$\frac{C_1}{C_2}$	0.15	27	22	22	2	22	22	225	221	276	270
	0.27	197	21	11	22	22	22	255	253	251	251
	0.36	247	15	43	242	20	229	233	236	235	235

It is seen from the above table that the efficiency decreases

with increasing values of $c^e \frac{C_1}{C_2}$, $\frac{C_3}{C_2}$, $\frac{C_4}{C_2}$.

It is interesting to observe that even a slight increase in $\frac{C_1}{C_2}$ results

in appreciable increase in efficiency as is to be expected since the number of parallel units to be selected is inversely proportional to C_1 .

Since the ratio between the values of $\frac{C_3}{C_2}$ or $\frac{C_4}{C_2}$ is less in terms of efficiency, the ratio is really appreciable.

C H A P T E R VII.

Multi-stage Procedure in Stratified Lattice Sampling.

INTRODUCTION:

The principle of stratification is to divide the population into homogeneous groups called strata. In different strata the sampling procedure may be different. Stratification is particularly suitable when the strata are homogeneous and the strata are large, so that the estimate λ' is more accurate than λ . It is difficult to estimate λ if the strata are small. It is important to ensure the equality of strata.

7.1. DEFINITION OF THE SCHEME:-

Let

$$t = t(x^k) \text{ where } x^k = 1 + x^{k-1} + \dots + 1.$$

$$n_r = t(x^r) \text{ number of primary units in the } r\text{-th stratum}$$

$$\sum_{i=1}^k n_i = n; \quad Y = \text{number of terms in the polynomial } t$$

$$r_i = \text{the } n_i \text{ primary units in the } i\text{-th stratum of }$$

r_{ij} , the j th number of units in the (i,j) th primary unit

of the i th stratum

$$i = 1, 2, \dots, n, \quad j = 1, 2, \dots$$

$$\sum_{j=1}^{m_{hi}} 1$$

$$x_{n_1, i_k} = \text{var}_1 + \text{var}_2 + \dots + \text{var}_n$$

$$x_{rj} = \sum_{i=1}^r x_{ri} + \dots + x_{rj} + \dots + x_{rs}$$

$$x_r = \sum_{i=1}^r x_{ri} + \dots + x_{rs} + \dots + x_{rd}$$

$$(\sum_{i=1}^{N_r} 1 - 1, \quad r = 1, 2, \dots)$$

$$= \sum_{A=1}^K ; \quad (i = 1, 2, \dots)$$

in relation to the variables of our study. Both characteristics under study are measured in pairs, so the value of x_{ij} is the sum of the values of x_{ri} and x_{sj} for each of the i and j categories of x_{ij} . It follows from this that all the variables of our study. From the definition of x_{ij} it is evident that x_{ij} is a measure of the character under study.

¹ O' Reilly et al., 1977, p. 104.

$$\hat{y}_j = \frac{\sum_{i=1}^{m_{ji}} y_{ij}}{m_{ji}}$$

$$\widehat{\text{xi}}(x^i) = \frac{1}{n} \sum_j x_{ji}$$

$$n_i = \sum_j m_{ij}$$

73

37-1075-2-A-17

It is true that on 11

ij *l* = *x* *y* *t* *v* *e*^{*a*} : *ab* *t* *bc*

$$\sum_{r=1}^n \frac{1}{\tau_r} = \frac{1}{\tau_1} + \dots + \frac{1}{\tau_n}$$

תְּהִלָּה, תְּבוּנָה וְבַרְכָה בְּעֵדוֹתָיו וְבְמִזְרָחָיו

-th stratum and in -3 are best combining

परं तु न विजयते तद्वारा एवं तदा

... will be used. Several types have been developed.

10. The following table shows the number of hours worked by 1000 workers in a certain industry.

17. $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{32}$

$$z = \sum_{n=1}^k \frac{1}{\frac{\pi_n}{\sum_{i=1}^k \pi_i}} = \frac{1}{\frac{\pi_1}{\sum_{i=1}^k \pi_i}} + \dots + \frac{1}{\frac{\pi_k}{\sum_{i=1}^k \pi_i}} \quad \dots \dots \dots \quad (1)$$

$$\frac{1}{n} \sum_{i=1}^n \hat{\nu}_{x_i}, \quad \frac{1}{n} \sum_{i=1}^n \hat{\nu}_{z_i}$$

$$\frac{1}{n} \sum_{i=1}^n \hat{\nu}_{\hat{x}_i}$$

Probabilistic ν is ν like ν

so get the 1

Q.S. I. $T_f =$

$$\hat{\nu}_1 = \sum_{n=1}^K \frac{1}{r} \frac{\sum_{i=1}^{n'} \hat{\nu}_{x_i}}{\sum_{n=1}^K \frac{1}{r} \sum_{i=1}^{n'} \hat{\nu}_{x_i}}$$

Q.S. II. $T_f =$

$$\hat{\nu}_n = \sum_{n=1}^K \frac{1}{r} \frac{\sum_{i=1}^{n'} \hat{\nu}_{x_i}}{\sum_{n=1}^K \frac{1}{r} \sum_{i=1}^{n'} \hat{\nu}_{x_i}}$$

Q.S. III. $T_f =$

$$\hat{\nu}_n = \sum_{n=1}^K \frac{1}{r} \frac{\sum_{i=1}^{n'} \hat{\nu}_{x_i}}{\sum_{n=1}^K \frac{1}{r} \sum_{i=1}^{n'} \hat{\nu}_{x_i}}$$

Q.S. IV. $T_f = 1^n$

so

$$\hat{\nu}_1 = \frac{1}{K} \sum_{n=1}^K \hat{\nu}_n$$

$$\sum_{n=1}^k \frac{r_n}{B_n} = \frac{1}{B_1} + \frac{r_2}{B_2} + \dots + \frac{r_k}{B_k}$$

$$\sum_{n=1}^k \frac{r_n}{B_n} \Delta$$

$$c_1 \dots c_m \frac{1}{B_1} + \frac{r_2}{B_2} + \dots + \frac{r_k}{B_k} \Delta$$

$$\frac{c_1 \dots c_m}{B_1} + \frac{r_2}{B_2} + \dots + \frac{r_k}{B_k} \Delta$$

$$T_{\alpha_1} \dots T_{\alpha_m} \dots$$

$$= \dots + T_{\alpha_m} \dots$$

$$\dots + T_{\alpha_m} z \dots$$

$$T'_{\alpha_{m+1}} = T_{\alpha_{m+1}}^{-1} = \tau \left(\frac{1}{z} \right) = z$$

$$\tau \left(\frac{1}{z} \right) = \tau \left(\frac{1}{z} \right) - \tau' \left(\frac{1}{z} \right) - \tau \left(\frac{1}{z} \right) \tau'_{\alpha_{m+1}} \tau_{\alpha_{m+1}}$$

Substituting $\tau'_{\alpha_{m+1}} = 0$, $\tau_{\alpha_{m+1}} = 0$ in (1) we get $T'_{\alpha_{m+1}} = 0$.

Ignoring higher order terms.

$$\sum_{n=1}^k \frac{r_n \dots + r_n}{z + z} \dots + z \frac{1}{z} \frac{1}{z} \dots$$

$$\approx \sum_{n=1}^k \frac{r_n \dots + r_n}{B_n} \frac{1}{z} \frac{-1}{z} \frac{r^2}{z} + \dots = \frac{\sum r_n^2}{B_n} + \frac{1}{B_1} \frac{r^2}{z} + \dots$$

$$\frac{\sigma_{r^2}}{r} + \frac{\sigma_{x^2}}{x^2} = \frac{\gamma(\sum e_{n_3})}{\sum r_{n_3} x} + \frac{(\sum e_{n_3})}{\sum r_{n_3}} \frac{\gamma(\sum e_{n_3})}{\sum r_{n_3}} \frac{(\sum e_{n_3})}{\sum r_{n_3}}$$

$$| \frac{e_{n_3}}{r} | < 1 ; | \frac{e_{n_3}}{x} | < 1 ; | \frac{x^2}{r} | < 1 ;$$

so $\text{the error} = f(1 - \frac{c}{B})^{-1}$

$$\text{and } (1 + \frac{c}{B})^{-1} \approx v^{-1}.$$

so $t_1 = 0.5 \cdot \text{dist} \cdot v \approx 0$

$$E(\hat{s}_{n_3}) = \sum_{n=1}^k \dots \frac{r_{n_3}}{r} \left\{ 1 + \frac{\text{cov}(u_r, v_r)}{r_{n_3} v_r} - \frac{\text{cov}(u_r, u_r)}{r_{n_3} r_{n_3}} \right.$$

$$- \frac{\text{cov}(u_r, v_r)}{r_{n_3} r_{n_3}} + \frac{\sum v_r}{r^2} + \frac{\pi(r)}{r^2} - \frac{\pi(u_r)}{r_{n_3} v_r}$$

$$- \frac{\pi(v_r)}{r_{n_3} r_{n_3}} + \frac{\text{cov}(u_r, v_r)}{r^2} \}$$

we see $t_1 = \text{dist} \cdot v \approx E(\hat{s}_{n_3}) \approx 0$

$$E(\hat{s}_{n_3}) \approx \frac{E(\hat{s}_{n_3}) - \bar{s}_{n_3}}{\bar{s}_{n_3}} = \frac{\sum_{n=1}^k \frac{y_{n_3}}{r_{n_3}} - \bar{s}_{n_3}}{\bar{s}_{n_3}}$$

$$+ \sum_{n=1}^k \frac{r_{n_3} v_{n_3}}{r_{n_3} r_{n_3}} \left(\frac{\text{cov}(u_r, v_r)}{r_{n_3} v_r} - \frac{\text{cov}(u_r, u_r)}{r_{n_3} r_{n_3}} \right)$$

$$= \frac{\text{Cov}(u_r, v_r)}{r \dots} + \frac{\sum_{n=1}^k r(n)}{2} + \frac{v(\dots)}{B_r^2} - \frac{\text{Cov}(u_r, v_r)}{\dots v_r \dots}$$

U.S. ECONOMIC POLICY

$$V\left(\frac{r}{n}\right) = \frac{1}{n} \sum_{i=1}^n \frac{X_i}{nB} + \frac{1}{n} \sum_{i=1}^n \frac{N_n - i}{n} \left(\frac{1}{1 - \frac{1}{n}} - \frac{1}{1 - \frac{i}{n}} \right)$$

$$\tau(\cdot) = -\frac{1}{\gamma_1} \frac{\sigma_{nx}^2}{\gamma_1} + \frac{1}{\gamma_1} \sum_{i=1}^{N_n} \frac{1}{\gamma_i} \left(\frac{1}{\gamma_1} - \frac{1}{\gamma_i} \right)$$

$$h(v(u_n, v_n)) = \frac{1}{n^2} \sigma_+ + \frac{1}{n} \sum_{i=1}^{N_n} \frac{\tau_i}{\rho_i} \left(\frac{1}{x_i} - \frac{1}{y_i} \right),$$

$$\text{Cov}(u_i, u_j) = \frac{1}{n} \sigma_u^2 + \frac{1}{r} \sum_{i=1}^{Nr} \left(\frac{1}{x_{ij}} / \left(\frac{1}{x_{ij}} - \frac{1}{r} \right) \right) \approx \frac{1}{r^2 - r}$$

$$v(r, \theta) = \frac{1}{n_1} \sigma_{m=0} + \frac{1}{n_2} \sum_{i=1}^{N_2} \frac{r_i^2}{p_{i,i}} \left(\frac{1}{r_i} - \frac{1}{R_i} \right) e^{i\theta},$$

More than one month in (6) resulting from disease.

Sub. in (4). (6) in (5) gives the relation

$\text{d}z + \mathcal{B}\left(\frac{\hat{y}_t}{y_t}\right) dz = -\alpha \ln y_t dy_t$, in $t \in [0, T]$ varies with y_t .

CO. 107 IS THE 7TH TO THE 2ND SECOND FLOOR.

"sin" - 1.0 + 2.7 i - 1.7 sin π z - 3.6 e $^{2\pi z}$ + 5.2 e $^{-2\pi z}$ + 6.3 e $^{4\pi z}$ + 7.5 e $^{-4\pi z}$

$$\hat{\eta}(v_t) \equiv \sum_{r=1}^k \left(\frac{r \dots v_r \dots}{\dots r \dots} \right)^{\alpha} + \left(\frac{\eta(v_r)}{v_r} + \frac{\eta(v)}{v} + \frac{\eta(v)}{v^2} \right)$$

T. expression for $\text{cov}(u_i, v_j)$, $\text{cov}(u_i, u_j)$, $\text{cov}(v_i, v_j)$, $\text{var}(u_i)$

$\pi(\tau_f)$ vs β is given in Fig. 1.

$$C_0 = \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\sigma d\tau = \frac{1}{2} \pi^2.$$

$$\begin{aligned}
 & \text{I order} \\
 & \hat{y}_r = \frac{1}{n-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - u_r \right)^2, \quad w = \frac{1}{n-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - v_r \right)^2 \\
 & s_{v_r}^2 = \frac{1}{n_r-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - \frac{1}{n_r} \sum_{i=1}^{n_r} \frac{\hat{y}_{ri}}{x_{ri}} \right)^2 \\
 & u_r = \frac{1}{n-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - u_r \right) \left(\frac{\hat{y}_{ri}}{x_{ri}} - \frac{1}{n_r} \sum_{i=1}^{n_r} \frac{\hat{y}_{ri}}{x_{ri}} \right) \\
 & u_r = \frac{1}{n-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - u_r \right) \left(\frac{\hat{y}_{ri}}{x_{ri}} - v_r \right) \\
 & v_r = \frac{1}{n_r-1} \sum_{i=1}^{n_r} \left(\frac{\hat{y}_{ri}}{x_{ri}} - v_r \right) \left(\frac{\hat{y}_{ri}}{x_{ri}} - v_r \right) \\
 & \frac{i}{r} = \eta\left(\frac{i}{r}\right), \quad \frac{v}{r} = \eta\left(\frac{v}{r}\right); \quad \frac{v}{r} = \eta\left(\frac{v}{r}\right)
 \end{aligned}$$

$$\begin{aligned} \therefore \frac{\bar{u}'v}{n} &= \text{cov}(u, v); \quad \therefore \frac{\bar{u}'^2}{n} = \text{var}(u, v) \\ \therefore \frac{\bar{x}'v}{n} &= \text{cov}(x, v) \end{aligned} \quad \left. \begin{array}{l} \vdots \\ \vdots \\ \vdots \end{array} \right\} \dots \dots \dots$$

It is also the variance of the sum of the individual observations.

Thus, the covariance of x and v is $\text{cov}(\hat{x}, v)$.

$$\begin{aligned} \text{cov}(\hat{x}, v) &= \sum_{n=1}^k \left(\frac{\bar{u}_n}{r(\sum_{n=1}^k)} \right) \left\{ \frac{v_n}{u_n^2} + \frac{v_n}{r(\sum_{n=1}^k)} + \frac{v_n}{r(\sum_{n=1}^k)} + \sum_{n=1}^k \frac{v_n}{r(\sum_{n=1}^k)} \cdot \frac{1}{r(\sum_{n=1}^k)} \right. \\ &\quad \left. + \frac{v_n^2}{r(u_n v_n)} - \frac{v_n^2}{r^2 u_n^2} - \frac{v_n^2}{r^2 (\sum_{n=1}^k)^2} \right\} \end{aligned}$$

$$= \frac{v_n^2}{r(u_n v_n)} - \frac{v_n^2}{r^2 u_n^2} - \frac{v_n^2}{r^2 (\sum_{n=1}^k)^2} + \frac{v_n^2}{r^2 (\sum_{n=1}^k)^2} \quad \left. \begin{array}{l} \vdots \\ \vdots \\ \vdots \end{array} \right\} \dots \dots \dots 10$$

or $\text{cov}(x, v) = r(u_n v_n) - \frac{v_n^2}{r^2 u_n^2}$

giving the corresponding expression for \hat{x} .

7.5

Gain Due To Strategic Solution

The strategic solution is the one which maximizes the total gain.

Strategic solution is the average of all the n strategies.

Gain due to strategic solution is given by

$$\begin{aligned} \hat{s}_j &= \sum_{n=1}^k \frac{\frac{1}{n} \sum_{i=1}^{n-1} p_{ij} \frac{r_i}{r_{ij}} \sum_{i=1}^{m_{ji}} \frac{m_{ji}}{r_{ij}} - \frac{1}{n} \sum_{i=1}^{n-1} p_{ij} \frac{r_i}{r_{ij}} \sum_{i=1}^{m_{ji}} \frac{m_{ji}}{r_{ij}}}{\frac{1}{n} \sum_{i=1}^{n-1} p_{ij} \frac{r_i}{r_{ij}} - \frac{1}{n} \sum_{i=1}^{n-1} p_{ij} \frac{r_i}{r_{ij}}} \end{aligned}$$

$$= \sum_{n=1}^K \frac{u_r}{n_r} \cdot \frac{1}{\sum_{i=1}^d} \quad \dots\dots .11$$

$$= \frac{1}{n_r} \sum_c \frac{m_c}{p_{ri}} - \frac{1}{\sum_j m_{ri}} \sum_j \tilde{\sigma}_{rij} (\tilde{\sigma}_{rij})$$

$$= \frac{1}{n_r} \sum_c \frac{1}{p_{ri}} - \frac{1}{m_{ri}} \sum_d \tilde{\sigma}_{rij}$$

$$v_r = \frac{1}{n_r} \sum_c \frac{B_{ri}}{p_{ri}}$$

$\sigma_{xy} = \sigma_{y^2} = \sigma_{x^2} = v_r$

$$\pi(\hat{x}) = \frac{1}{n^2} \left(\sum_{n=1}^K \frac{1}{n_r} (\sigma_{xy}^2 + \tilde{y}_{ri}^2 - \tilde{\sigma}_{rij} \tilde{\sigma}_{xij}) \right)$$

$$+ \left[\frac{1}{n_r^2} \sum_c \frac{m_c}{p_{ri}} \left(\frac{1}{p_{ri}} - \frac{1}{m_{ri}} \right) \times (\tilde{\sigma}_{xij}^2 \tilde{\sigma}_{yij}^2 \tilde{\sigma}_{rij}^2 \tilde{\sigma}_{xij}^2) \right] \dots\dots .12$$

$$+ \sum_{n=1}^K \frac{1}{n_r} \sum_c \frac{1}{p_{ri}} \frac{1}{m_{ri}} \frac{1}{l_{ri}} \sum_{j=1}^{m_{ri}} \frac{2}{l_{rij}} \left(\frac{1}{p_{ri}} - \frac{1}{m_{ri}} \right) \tilde{\sigma}_{xij}^2 \dots\dots .12$$

$$+ \frac{1}{n_r^2} \left(\sum_{n=1}^K \left(\frac{\sigma_{xy}^2}{n_r} + \frac{\tilde{y}_{ri}^2}{n_r} - \frac{\tilde{\sigma}_{rij} \tilde{\sigma}_{xij}}{n_r} \right) \right) \dots\dots .12$$

$$- \frac{1}{n_r^2} \left(\sum_{n=1}^K \frac{1}{n_r} \left(\sigma_{xy}^2 + \tilde{y}_{ri}^2 - \tilde{\sigma}_{rij} \tilde{\sigma}_{xij} \right) \right) \dots\dots .12$$

then \hat{x} is the true value of x .

If $\sigma_{xy} = 0$ and $\tilde{\sigma}_{rij} = 0$, then \hat{x} is also 0.

\hat{x} is also unbiased.

$$\hat{x}_{13} = \frac{\frac{1}{n} \sum_c \frac{1}{p_{ci}} \frac{m_c}{p_{ci}} \sum_j \tilde{\sigma}_{cij} (\tilde{\sigma}_{cij})}{\frac{1}{n} \sum_c \frac{1}{p_{ci}} \frac{1}{m_{ci}} \sum_j \tilde{\sigma}_{cij}} = \frac{u}{q_u} \quad \dots\dots .13$$

From the results of last section, we have

$$V(\hat{u}_{ij}) = \frac{1}{n} \left\{ V(u) + \bar{j}^2 V(\cdot) - \sum_{k=1}^N \alpha_k (u_{ik}, \cdot) \right\} \quad \dots \dots 14$$

where $V(u) = \frac{1}{n} \sigma_y^2 + \frac{1}{n} \sum_{e=1}^N \frac{1}{x_e} \left\{ \left(\frac{1}{x_e} - \frac{1}{\bar{x}} \right) \sigma_y^2 \right.$

$$\left. + \frac{1}{\epsilon_e} \sum_{j=1}^{M_e} \left(\frac{1}{x_j} - \frac{1}{\bar{x}_e} \right) \alpha_{ej} \right\}$$

$$V(\cdot) = \frac{1}{n} \sigma_y^2 + \frac{1}{n} \sum_{e=1}^N \frac{1}{x_e} \left(\frac{1}{x_e} - \frac{1}{\bar{x}} \right) \sigma_y^2 \quad \dots \dots 15$$

$$\text{Cov}(u_{ij}, \cdot) = \frac{\sigma_y^2}{n} + \frac{1}{n} \sum_{e=1}^N \frac{1}{x_e} \left(\frac{1}{x_e} - \frac{1}{\bar{x}} \right)$$

$\pi_C = 1 - r$

$$V(\hat{u}_{ij}) = \frac{1}{n} \left\{ \frac{1}{x_i} (\sigma_y^2 + \bar{j}^2 \sigma_y^2 - \bar{j}^2 \sigma_y^2) + \frac{1}{n} \sum_{e=1}^N \frac{1}{x_e} \left(\frac{1}{x_e} - \frac{1}{\bar{x}} \right) \cdot \right.$$

$$\left. (\sigma_{ey}^2 + \bar{j}^2 \sigma_{ey}^2 - \bar{j}^2 \sigma_{ey}^2) + \frac{1}{n} \sum_{e=1}^N \frac{1}{x_e} \sum_{j=1}^{M_e} \alpha_{ej} \left(\frac{1}{x_j} - \frac{1}{\bar{x}_e} \right) \sigma_{ey}^2 \right\}$$

.....1

$$\pi_C = 1 - r \quad (10) \quad (10) \text{ gives } \pi_C = 1 - r^2.$$

Let us now consider the variance of \hat{u}_{ij} given by (10).

Suppose $t_i = t_j = t$. Then $\hat{u}_{ij} = \hat{u}_i$ and $\hat{u}_{ij} = \hat{u}_j$.

Now $\hat{u}_i = \bar{u} - \hat{u}_i$, where $\bar{u} = \bar{u}_i$.

$$\hat{u}_{ij} = \frac{\hat{u}_i}{\bar{u}} - \hat{u}_i \quad \sum_{e=1}^{N_e} \hat{u}_e = \hat{u}_i$$

Since \hat{u}_i is unbiased for \bar{u} and \hat{u}_e is unbiased for \bar{u}_e for all i, j, e ,

we have $\hat{u}_{ij} = \bar{u}$.

....107..

$$\sigma^2 = \sum_{n=1}^N \left(\frac{x_n - \bar{x}}{x} \right)^2 = \sum_{n=1}^N \sum_{i=1}^{N_n} \left(\frac{x_{ni} - \bar{x}_i}{x_i} \right)^2$$

$$= \sum_{n=1}^N \sum_{i=1}^{N_n} \left[\frac{\frac{\partial}{\partial x_i} x_{ni}}{x_i} - \frac{\frac{\partial}{\partial x_i} \bar{x}_i}{x_i} \right]^2 = \dots$$

$$= \sum_{n=1}^N \sum_{i=1}^{N_n} \left[\left\{ \left(\frac{\frac{\partial}{\partial x_i} x_{ni}}{x_i} - \frac{\frac{\partial}{\partial x_i} \bar{x}_i}{x_i} \right) \left(\frac{\frac{\partial}{\partial x_i} x_{ni}}{x_i} - \frac{\frac{\partial}{\partial x_i} \bar{x}_i}{x_i} \right) \right\} \right]$$

$$\sum_{n=1}^N \sum_{i=1}^{N_n} \left(\frac{\frac{\partial}{\partial x_i} x_{ni}}{x_i} - \dots \right) \sum_{n=1}^N \left(\frac{\frac{\partial}{\partial x_i} x_{ni}}{x_i} - \dots \right)$$

$$= \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i} + \sum_{n=1}^N \frac{\sigma_{ii}}{x_i} \dots$$

$$\sigma^2 = \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i} + \sum_{n=1}^N \frac{\sigma_{ii}}{x_i} \dots$$

$$\sigma_{ij}^2 = \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i} + \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_j} \dots$$

$$\sigma_{ij}^2, \sigma_{ji}^2 = \sigma^2$$

$$\eta(\hat{\sigma}) = \frac{1}{k} \left\{ \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i} + \frac{\sigma_{ii}^2}{x_j} - \frac{\sigma_{ij}^2}{x_i x_j} \right\}$$

$$= \left(\sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i} - \dots \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_j} - \sum_{n=1}^N \frac{\sigma_{ij}^2}{x_i x_j} \right)$$

$$+ \frac{1}{\dots} \sum_{k=1}^{\infty} \frac{1}{x_k} \sum_{i=1}^{n_k} \frac{1}{x_{ri}} \left(\frac{1}{x_{ri}} - \frac{1}{x_i} \right) (x_i^2 + x_{i+1}^2 + \dots + x_{ri}^2)$$

It is not evident to

$$\tau_j^{\frac{1}{r}} : \quad \frac{1}{2} \left\{ \sum_{k=1}^K \left(\tau_{(1)} + \tau_{(2)} + \dots + \tau_{(r)} - \omega_{(1)} - \omega_{(2)} - \dots - \omega_{(r)} \right) \right\} \dots .21$$

$$= \sum_{k=1}^{\infty} \frac{1}{k^2} \left(1 - \dots + \frac{(-1)^{k-1}}{k} - \frac{(-1)^k}{k+1} + \dots \right)$$

$$B = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x} \right)^2 - \frac{1}{d} \sum_{j=1}^m \left(\frac{y_j - \bar{y}}{\sigma_y} \right)^2$$

$$= \frac{1}{n-1} \sum_{c} \left(\frac{m_{ci}}{x_{ci}} - \frac{1}{n-1} \sum_j \frac{m_{ji}}{x_{ji}} \right)$$

$$S_{u_i, r} = \frac{1}{n-1} \sum_{c=1}^m \left(\frac{r_i}{p_{ri}} - \frac{1}{x_1} \sum_{j=1}^{m_{rc}} i_{ij} \left(i_{ij} - u \right) \times \left(\frac{r_i}{p_{ri}} - \frac{1}{x_1} \sum_{j=1}^{m_{rc}} i_{ij} - \cdot \cdot \cdot \right) \right)$$

Then estimate $\hat{y}_{1,1}$ in $V(\hat{y}_{1,1})$ will be given by

$$\hat{v}(j) = \frac{1}{(\sum_{r'}^k)} \left\{ \sum_{r=1}^k \frac{1}{4p} \left(-\frac{v_r}{x_r} + \frac{\hat{v}_r}{x_r} - \frac{v_r}{x_r} - \frac{2\hat{v}_r}{x_r} \right) + \sum_{r=1}^k \frac{1}{4p} \left(\frac{v_r}{x_r} - \frac{\hat{v}_r}{x_r} \right)^2 \right\} \quad \dots \dots 22$$

$$47 \quad f_1 + f_2 = 0^\circ \quad \eta(\frac{\Delta}{\pi}) \quad \text{is even}$$

$$\begin{aligned} \gamma_{\perp}^2 \cdot \pi(\hat{s}_r) &= \frac{1}{(\sum s_r)^2} \left(\sum_{n=1}^k \frac{1}{n_r} (s_{nr}^2 + \frac{\hat{s}_r^2}{n_r} s_r^2) - 2 \frac{u_r}{n_r} s_{nr} u_r \right) \\ &\quad + \sum_{n=1}^k \frac{1}{n_r} \left(\frac{\hat{s}_r^2}{n_r} s_r^2 - 2 \frac{\hat{s}_r}{n_r} s_{nr} \right) - \sum_{n=1}^k \frac{1}{n_r} \left(\frac{u_r^2}{n_r} s_r^2 - \frac{u_r^2}{n_r} s_{nr} u_r \right) \end{aligned}$$

$$d_{\perp}^2 = \frac{1}{n-1} \sum_c \frac{n_r}{p_{rc}} \left(\frac{v_r}{p_{rc}} - \frac{v_r'}{p_{rc}} \right)^2, \quad v_r' = \frac{1}{n-1} \sum_c \frac{u_r}{p_{rc}} \quad \dots \dots 23$$

$$\gamma_{\perp}^2 u_r' = \frac{1}{n-1} \sum_c \left(\frac{1_{ri}}{p_{rc}} - \frac{1}{m_{rc}} \sum_j \frac{m_{rc}}{m_{rij}} \tilde{y}_{rij} (u_{rij} - v_r') \right) \left(\frac{v_r'}{p_{rc}} - v_r' \right)$$

($r = 1, 2, \dots, k$)

The γ_{\perp}^2 is called variance of s_{nr} .

Definition of γ_{\perp}^2 is given below:

Def. γ_{\perp}^2 is the variance of s_{nr} = $E(s_{nr}^2) - E(s_{nr})^2$:

$$\begin{aligned} \text{def. } \gamma_{\perp}^2 &= \frac{1}{n-1} \sum_c \frac{u_r}{p_{rc}} \cdot v_r' \quad T_1 = \frac{1}{n-1} \sum_c \frac{u_r}{p_{rc}} \\ T_1 &= \sum_{n=1}^k \frac{\frac{1}{n_r} \sum_c \frac{u_{ri}}{p_{rc}} - \frac{1}{n_r} \sum_{n=1}^k \frac{m_{rc}}{m_{rij}} \tilde{y}_{rij} (u_{rij} - v_r')}{\frac{1}{n-1} \sum_c \frac{u_r}{p_{rc}} - \frac{1}{m_{rc}} \sum_{n=1}^k \frac{m_{rc}}{m_{rij}} \tilde{y}_{rij}} - \frac{1}{n-1} \sum_c \frac{u_r}{p_{rc}} \\ &= \sum_{n=1}^k \frac{u_r}{n_r} \cdot v_r' \quad \dots \dots 24 \end{aligned}$$

and it is calculated by

$$\begin{aligned} V(T_1) &= \sum_{n=1}^k \left\{ V(u_r) + \frac{\hat{s}_r^2}{n_r} V(v_r) + \frac{\hat{s}_r^2}{n_r} V(v_r') + \frac{2\hat{s}_r}{n_r} \text{cov}(u_r, v_r') \right. \\ &\quad \left. - 2\hat{s}_r \text{cov}(u_r, v_r) - \frac{\hat{s}_r^2}{n_r} \text{cov}(v_r, v_r') \right\} \quad \dots \dots 25 \end{aligned}$$

$$= \sum_{k=1}^K \left[\frac{1}{r} (\sigma_{\epsilon_k}^2 + \frac{\sigma_{\epsilon_k}^2}{n} \sigma_{\epsilon_k}^2) - \bar{\sigma}_{\epsilon_k}^2 + \frac{1}{r} (2 \frac{\bar{\sigma}_{\epsilon_k}^2}{r} - \frac{\sigma_{\epsilon_k}^2}{r}) \right]$$

$$\begin{aligned} &+ \sum_{k=1}^K \frac{1}{n} \sum_{i=1}^{m_k} \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \left(\frac{\bar{\sigma}_{\epsilon_k}^2}{r_i} - \frac{\sigma_{\epsilon_k}^2}{r_i} \right) \\ &+ \sum_{k=1}^K \frac{1}{n} \sum_{i=1}^{m_k} \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \sum_{j=1}^{m_k} \frac{1}{r_j} \left(\frac{1}{r_j} - \frac{1}{r} \right) \bar{\sigma}_{\epsilon_k \epsilon_j}^2 \} \dots \dots 26 \end{aligned}$$

$T = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \bar{\sigma}_{\epsilon_i \epsilon_j}^2$

$$T = \frac{\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \bar{\sigma}_{\epsilon_i \epsilon_j}^2}{\frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} - \frac{1}{r}} \cdot \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \dots \dots 27$$

$\bar{\sigma}_{\epsilon_i \epsilon_j}^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \bar{\sigma}_{\epsilon_i \epsilon_j}^2$

$$\begin{aligned} \bar{\sigma}_{\epsilon_i \epsilon_j}^2 &= \left(\frac{1}{n} (\sigma_{\epsilon_i}^2 + \frac{\sigma_{\epsilon_i}^2}{n} \sigma_{\epsilon_i}^2) - \bar{\sigma}_{\epsilon_i}^2 \right) + \frac{1}{n} \left(\bar{\sigma}_{\epsilon_i}^2 - \frac{\sigma_{\epsilon_i}^2}{n} \right) \\ &+ \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \left(\bar{\sigma}_{\epsilon_i}^2 + \frac{\sigma_{\epsilon_i}^2}{r_i} \sigma_{\epsilon_i}^2 - \frac{\bar{\sigma}_{\epsilon_i}^2}{r_i} \bar{\sigma}_{\epsilon_i}^2 \right) \\ &+ \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \left(\frac{1}{r_i} - \frac{1}{r} \right) \frac{1}{n} \sum_{j=1}^m \frac{1}{r_j} \left(\frac{1}{r_j} - \frac{1}{r} \right) \bar{\sigma}_{\epsilon_i \epsilon_j}^2 \} \dots \dots 28 \end{aligned}$$

To determine $\bar{\sigma}_{\epsilon_i \epsilon_j}^2$, we need to calculate $\bar{\sigma}_{\epsilon_i}^2$ and $\bar{\sigma}_{\epsilon_j}^2$.

This is done by calculating the mean of the squared error terms.

Since $\bar{\sigma}_{\epsilon_i}^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_i^2$, we can add to $\bar{\sigma}_{\epsilon_i}^2$

$\bar{\sigma}_{\epsilon_i}^2 + r - t + b$, so $\bar{\sigma}_{\epsilon_i}^2 =$

$$\frac{\omega_0}{\omega_r} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cos(n\theta)$$

17. 09. 22. 5 in. fc. + 7 - 2. 5 ft.
at 40 ± 10 sec.
Max. 1. 5.

$$\begin{aligned} \sigma_y^2 &= \sum_{k=1}^K \left(\frac{\hat{y}_k - y_k}{\hat{s}_k} \right)^2 = \sum_{k=1}^K \sum_{i=1}^{N_k} \left(\frac{\hat{y}_{ki} - y_{ki}}{\hat{s}_{ki}} \right)^2 \\ &= \sum_{n=1}^N \alpha_n \left\{ \sum_{i=1}^{N_n} \left(\frac{1}{\hat{s}_{ni}} \left(\hat{y}_{ni} - \bar{y}_{ni} \right) + \frac{\bar{y}_{ni} - y_{ni}}{\hat{s}_{ni}} \right)^2 \right\} \\ &= \sum_{n=1}^N \alpha_n \sum_{i=1}^{N_n} \left[\left(\frac{\hat{y}_{ni} - \bar{y}_{ni}}{\hat{s}_{ni}} \right)^2 + \left(\frac{\bar{y}_{ni} - y_{ni}}{\hat{s}_{ni}} \right)^2 \right] \end{aligned}$$

"*it* vanishes

$$= \sum_{n=1}^k \frac{1}{\zeta_n} \sum_{i=1}^{N_n} \left(\frac{\zeta_{r_{1...}}}{p_{i_1}} - \dots \right) + \sum_{n=1}^k \left(\frac{\zeta_{r_{1...}}}{p_{i_n}} - \dots \right)$$

$$= \sum_{n=1}^k \frac{\sigma_n^2}{\tau_n} - \sum_{n=1}^k \frac{\tau_n^2}{\tau_n} = ? \dots$$

$$\text{S.E. of } \bar{x} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n(n-1)}} + \sqrt{\frac{\sum_{i=1}^n \frac{x_i^2}{n}}{\frac{n}{n-1}}} - \bar{x}^2 \quad \dots \dots 30$$

$$\bar{O}_T = \sum_{n=1}^k \frac{\bar{O}_{x_n, B}}{x_n} + \sum_{n=1}^k \frac{\text{...}}{x_n} = \bar{y} \dots \dots \dots$$

Also we have

$$r_e = r_{ri} ; \quad e = \frac{e}{r_{ri}} ; \quad e_j = \frac{e_j}{r_{rj}} = r_{rj}$$

...111..

$\sigma_{\text{H}_2} \sim \sigma_{\text{J}}^2, \sigma_{\text{H}_2} \sim \sigma_{\text{B}}^2$ and $\sigma_{\text{H}_2} \sim \sigma_{\text{J}}^2 + \sigma_{\text{B}}^2$

irrevitableness.

$$\begin{aligned} \pi_{(j)} &= \left(\frac{1}{\sum_{k=1}^K} - \frac{1}{x} \right) \sum_{k=1}^K \frac{1}{x} \left(\frac{\sigma_k^2}{x} + \frac{x^2 \sigma_k^2}{x} - 2 \frac{\sigma_k}{x} \bar{G}_{jk} \right) + \left(\frac{1}{x} - \frac{1}{x} \right) x \\ &= \left(\frac{\sum_{k=1}^K \frac{1}{x}}{x} + \frac{-2}{x} \sum_{k=1}^K \frac{\sigma_k}{x} \right) - \frac{1}{x} \left(\sum_{k=1}^K \frac{\sigma_k^2}{x} + \sum_{k=1}^K \frac{x^2 \sigma_k^2}{x} - 2 \sum_{k=1}^K \frac{\sigma_k}{x} \bar{G}_{jk} \right) \\ &+ \frac{1}{x} \sum_{k=1}^K \sum_{l=1}^{N_k} \frac{1}{x} \left(\frac{1}{x} - \frac{1}{x} \right) \left(\bar{G}_{jk} - \frac{\sigma_k^2}{x} - \frac{x^2 \sigma_k^2}{x} + 2 \frac{\sigma_k}{x} \right) \\ &= \frac{1}{x} \sum_{k=1}^K \sum_{l=1}^{N_k} \frac{1}{x} \left(\frac{1}{x} - \frac{1}{x} \right) \sum_{j=1}^{M_{kl}} \left(\frac{1}{x} - \frac{1}{x} - \frac{x^2}{x} \right) \dots \end{aligned}$$

$$\begin{aligned} \pi(x^2, x^1, \dots, 0) &= \sum_{n \geq 1} \left\{ \frac{\lambda_n}{x_n} (\pi(u_n) + \frac{1}{x_n} \pi') - \pi(u_n), \quad \right\} \\ &\quad - \frac{1}{x_1} \left\{ \frac{1}{x_1} \pi(v_1) - \frac{1}{x_1} \pi(u_1, x^1) \right\} \\ &\quad + \left(\frac{1}{n_{k_p}} - \frac{1}{n^1 x_p} \right) \left(\dots - \frac{x^1}{x_p} \dots - \frac{x^{k_p}}{x_p} - \frac{x^{k_p+1}}{x_p} \right). \end{aligned}$$

$$+ \tau(n) = \sum_{k=1}^K \left(\frac{1}{\frac{r^2}{x_2} + \frac{a^2}{x_1}} - \frac{1}{\frac{r^2}{x_2} + \frac{a^2}{x_1} + \frac{1}{r^2}} \right)^2 + \frac{1}{r^2} \left(\frac{\frac{1}{x_1} - \frac{1}{x_2}}{\frac{1}{r^2} - \frac{1}{x_1}} \right)^2 + \frac{1}{r^2} \left(\frac{\frac{1}{x_1} - \frac{1}{x_2}}{\frac{1}{r^2} - \frac{1}{x_2}} \right)^2$$

$\pi^2 \pi^{-2} = \pi^2 - \pi^2 + \pi^2 - \pi^2 = \pi^2(\frac{1}{1})$ 34

$$\sum_{k=1}^{\infty} \left(\frac{1}{\frac{1}{n}} + \frac{1}{\frac{2}{n}} \left(\frac{\frac{S_2^2}{n}}{\frac{n-1}{n}} \right)^{n-1} - \frac{u}{n!} - \frac{2u}{n!} S_{n-1} \right) \dots \dots \dots 35$$

... 112..

7.6

NUMERICAL ILLUSTRATION

For the purpose of illustrating the results developed in this chapter we will consider the data of lime survey conducted in the year 1963-64 in Andhra Pradesh. A total number of 86 villages was selected from five talukas with replacement with probabilities proportional to reported area, out of which 45 villages were retained for yield study. All the orchards in each of these 86 villages were completely enumerated for area and number of bearing trees. In each selected village retained for yield study a maximum number of 4 orchards and in each of these orchards 12 trees were selected for observing the yield.

The table below gives taluka-wise estimates of average yield per tree, total number of bearing trees and total production along with their percentage standard errors.

Taluka	Av yield per tree (in kgs.)	No. of bear- ing trees	Total production (in Kgs.)
Venkatagiri	63.25 (7.3)	214,926 (7.6)	13,594,284 (6.6)
Rapur	64.38 (16.3)	56,866 (34.6)	3,660,805 (15.1)
Sulurpet	52.70 (11.9)	34,169 (15.8)	1,802,321 (10.5)
Kovur	30.02 (38.8)	3,813 (26.3)	114,470 (37.0)
Atmakur	49.41 (20.2)	4,904 (20.6)	242,302 (25.6)
District	61.69 (4.4)	314,708 (8.3)	19,414,182 (6.3)

As regards the overall estimates there is no difficulty as far as the characters 'total number of bearing trees' & 'total production' are concerned. These are obtained by pooling together the results in the different talukas and are given in the last-row of the table.

The problem of obtaining an overall estimate of average yield per tree is however not so simple. We shall consider the various weighted estimates enumerated in section and then discuss their efficiency. These alongwith their relative biases and percentage standard errors are given in the table below.

Estimate	Overall estimate of average yield per tree	% Relative bias	Relative Efficiency w.r.t. \hat{y}_2
\hat{y}_1	61.64	-0.4581	70.51
\hat{y}_2	61.69	-0.2197	100.00
\hat{y}_3	60.81	0.3715	52.05
\hat{y}_4	51.69	-0.3128	43.06
\hat{y}_5	60.10	0.7649	47.83

As is to be expected the estimate \hat{y}_2 is seen to be the most efficient followed by \hat{y}_1 . The estimates \hat{y}_4 & \hat{y}_5 are

no doubt less efficient but may be desirable at times from the point of view of computational convenience and simplicity. The low precision of the estimate \hat{y}_5 explains the fact that the strata averages differ markedly and to obtain efficient over-all estimate of the average yield it is necessary to have reliable weights while the low precision of the estimate \hat{y}_4 explains that the area under lime and the number of bearing trees do not have high correlation. It is also seen from the table that the percentage bias is almost negligible in all the cases. Using the estimate \hat{y}_2 , which has been found to be most efficient both from theoretical point of view as also survey data, the overall estimate of average yield per tree was found to be 61.69 kgs. with 4.4% standard error.

Since \hat{y}_2 has been found to be the most efficient estimate of average yield and T_1 , the most efficient estimate of total production. We shall consider these two estimates for the purpose of determining the sample sizes required to estimate these characters with specified precision. Assuming proportional allocation, it can be seen that n , the number of villages required to estimate the total number of bearing trees with 6% and 7% standard errors are 158 & 116 respectively. For these values of n and different values of m (no of orchards) the table below gives n_1 the number of villages required to estimate the average yield per tree for given percentage standard errors. The allocation is assumed to be proportional.



Table: III

Estimated values of n under proportional allocation for estimating average yield per tree for different values of m and given degree of precision.

$$n' = 116$$

m	4%		5%
	S.E	S.E	
3	102	66	
4	75	55	
5	74	48	
6	67	43	
7	62	40	

The next table gives the number of villages required to estimate the total production for different values of m and given degree of precision. The allocation is again assumed to be proportional table:

Table: IV

Table showing estimated values of n' for different values of m and different percentage standard errors. Allocation being proportional.

m	$n' = 158$		$n' = 116$	
	4% S.E	5% S.E	4% S.E	5% S.E
3	83	46	100	50
4	69	38	82	41
5	50	30	65	36
6	54	30	65	33
7	50	28	60	30

Table III & IV reflect that for a given percentage standard error the number of villages required to estimate the average yield per tree is considerably less than that required to estimate the total number of bearing trees. Also from table IV it is observed that that is considerable variation between villages in regard to the number of trees. As the contribution to the standard error of the estimates of average yield per tree arises from three different sources:

- (i) Variation between villages
- (ii) Variation between orchards within villages
- & (iii) Variation between trees within orchards.

of these the variation between villages is the most important one as it has been reflected in table which gives for varying number of orchards the number of villages required to be selected estimate the average yield per tree with given percentage standard error.

B I B L I O G R A F E

1. Cochran, W.G. (1953) Sampling techniques John Wiley & Sons Ltd., New York
2. Cochran, W.G. (1961) Comparison of methods for determining stratum boundaries. Bulletin of International Statistical Institute Vol.38 No.2.
3. Dalenius, T. (1950) The problem of optimum stratification Skandinavik Aktuarietidskrift
4. " (1957) Sampling in Sweden. Chapter 8. Almqvist & Wiksell, Stockholm.
5. Dalenius, T. & Gurney, M. (1951) The problem of optimum stratification II. Skandinavisk Aktuarietidskrift,
6. Dalenius, T. & Hodges, J.L. (1957) The choice of stratification points. Skandinavisk Aktuarietidskrift.
7. " (1959) Minimum variance stratification Journal of American Statistical Association
8. Durbin, J. (1959) Review of sampling in Sweden Journal of the Royal Statistical Society
9. Edman, G. (1959) An approximation useful in univariate stratification. The Annals of Mathematical Statistics
10. Goswami, J.N. (1961) "On some extensions of double sampling" unpublished thesis for Diploma of I.C.A.R., New Delhi.
11. Goswami, J.N. & Sukhatme (B.V.) (1965) Ratio method of Estimation in multi-phase sampling with several auxiliary variables. Journal of Indian Society of Agricultural Statistics Vol.1
12. Hansen, M.H. & Hurwitz, W.N. (1943) "On theory of sampling from finite populations" Annals of Mathematical Statistics Vol.14
13. Hansen, M.H., Hurwitz, W.N. & Madow, W.G. (1953) Sample survey methods & theory Vol.I & Vol.II. John Wiley & Sons. Ltd., New York.

14. Neyman, J. (1934) "On the two different aspects of the representative method: The method of stratified sampling and the method of Purposive selection". "Journal of Royal Statistical Society Vol.109.
15. Ravindra Singh (1962) "Some Estimation procedures in Fruit Surveys"
Unpublished Thesis for Diploma of I.C.A.R., New Delhi.
16. Sethi, V.K. (1963) "A note on optimum stratification of populations for estimating the populations means". Australian Journal of Statistics (April)
17. Sukhatme, B.V. (1959) "A contribution to double sampling"
Journal of Indian Society of Agricultural Statistics Vol.II
18. Sukhatme (P.V.) Sampling theory of Surveys with applications.
The Iowa State College Press, Ames, Iowa, U.S.A.
19. Tschuprow, A.A. (1923) "On the mathematical expectation of the moments of Frequency distributions in the case of correlated observations". Metron, 2.