CONTRIBUTIONS TO THE THEORY OF SAMPLING STRATEGIES

T. J. RAO

A thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Calcutta
1967

ACKNOWLEDGEMENTS

It is a great pleasure to express my deep gratitude to Prof. C. R. Rao, F.R.S., Director of the Research and Training School of the Indian Statistical Institute for providing facilities to carry out this research work and to Prof. M. N. Murthy, Head of the Design, Research and Training Division of the National Sample Survey Department for his valuable guidance and constant encouragement. I am grateful to Prof. B.P.Adhikari who introduced me to the field of Sampling Theory.

Early association with Dr. T.V.Hanurav and Dr. K.Vijayan was very helpful to me. Thanks are due to Dr. P.K.Pathak for allowing me to include some of the results of our joint paper 'Inadmissibility of customary estimators in sampling over two occasions'.

Mr. C.H.Sastry has taken great pains in writing the programmes while organising the calculations on the IBM 1401 Computer. To him are my heartfelt thanks.

Mr. M.Muraleedharan and Mr. K.Viswanath rendered valuable help during the final stages in the preparation of the thesis.

Finally, I wish to thank Sri Gour Mohon Das for his elegant and excellent typing.

T. J. RAO

C O N T E N T S

GO ON TO THE NEXT PAGE

* * *

CHAPTER I

INTRODUCTION

Sample survey as a technique for collection of data has been widely accepted since quite a long time and it is only during the thirties and forties that a more systematic development has taken place owing to the most significant and remarkable works of Cochran, Hansen, Hurwitz, Mahalanobis, Neyman, Sukhatme, Yates and others. During this period, there were many important advances and tremendous progress in this field was achieved in India under the direction of Professor P. C. Mahalanobis. Besides the theoretical developments, stress was laid on the practical techniques and large-scale sample surveys played a major rôle in this context (Mahalanobis [1944], [1946]).

Towards the beginning of the second half of this period, Neyman [1934], Hansen and Hurwitz [1943] and Cochran [1942] considered the problems of stratification and the question of utilisation of auxiliary information for selection purposes and for estimation, which gave rise to the varying probability selection method and the theory of ratio and regression estimates. Madow and Madow [1944] developed the theory of systematic sampling.

These earlier developments in sampling theory of finite populations, concerned with the techniques of sampling, suited to the situations that arose in practice and mainly dealt with the problems of estimation of population parameters such as totals, means or ratios and their errors. Reviews of the developments of theory of sampling from finite populations were published by Yates [1946], Stephan [1948], Seng [1951] and others. It was Horvitz and Thompson in 1952 who first recognized the need for dealing systematically with the theory of sampling from finite populations and besides formulating the theory neatly, they defined three classes of estimators. Later in 1955, Godambe proposed a unified theory of sampling from finite populations, with a view to discuss the fundamental problems of sampling within this framework and also formulated the definition of linearity with a general theory of sampling. Godambe [1955] has established that for any sample design there does not exist a uniformly minimum variance unbiased estimator of the population total in the class of all linear unbiased estimators (with some exceptions, characterized by Hanurav [1965] later). This leads to the choice of estimators from the class of admissible estimators and various other criteria have been put forward to arrive at an optimum choice. Hanurav [1965, 1966] recently introduced the concept of hyper admissibility and proved that the estimator due to Horvitz and Thompson [1952] is hyper admissible.

It was first shown by Cochran [1946] that whenever we have auxiliary information on a characteristic closely related to the study variable, we can utilise that information by considering a stochastic model, which he called as the super-population concept, the idea being taken from the Bayesian inference. Godambe [1955] has proved that in the class of all sampling strategies (sampling design together with an estimator being called a strategy (Hájek [1958])) with ν distinct units, a strategy which has (i) the same number of distinct units for every sample, (ii) the inclusion probability of any unit is proportional to the auxiliary information on that unit and (iii) the estimator is the corresponding Horvitz-Thompson estimator is the best in Bayesian sense. These optimum designs are termed by Hanurav [1965, 1966] as πPS designs and the sampling schemes which construct such designs are called πPS sampling schemes. Various applications of πPS sampling schemes are also given in this thesis.

We shall now give a brief summary of the author's contributions.

The thesis is divided into eight chapters. After the first introductory chapter, we explain in Chapter II the basic concepts and definitions which will be used in this thesis.

In Chapter III we study the method of ratio estimation for the estimation of population mean or total, the simplest method of utilisation of auxiliary information, and show how the various unbiased ratio estimators suggested by different authors can be constructed by the combination of two well-known biased ratio estimators. Relation between ratio estimators and convexity is briefly mentioned. A new concept 'stochastic convexity' is introduced which figures in the construction of unbiased ratio estimators.

Chapter IV deals with sampling schemes which provide unbiased ratio estimators. An explicit and exact expression for the variance of the unbiased ratio estimator, obtained by modification of the usual sampling scheme, due to Midzuno [1952] and Sen [1952], is obtained and a number of interesting relations among the coefficients involved in this variance expression, which depend on the auxiliary information, are established. Finally, exact expressions for the Bias and Mean Square Error of the ratio estimator for simple random sampling using these coefficients are given. Numerical examples are provided at the end of the chapter as an illustration.

Chapter V aims at choosing an optimum strategy for the method of ratio estimation. Three strategies are suggested for this purpose and a comparison between these strategies is presented. An application to cluster sampling is also given. The part played by πPS sampling schemes in choosing a suitable strategy here is emphasised and a general problem of choice between πPS schemes is presented towards the end. Illustrative examples are added in the last section.

In Chapter VI the problem of optimum allocation of sample size to strata is examined in the light of a priori distributions and further applications of πPS sampling schemes in stratification are studied. Allocation of sample size to strata minimizing the expected variance of πPS strategy with Horvitz-Thompson estimator is derived. It is shown that πPS sampling scheme for unstratified sampling is inferior to πPS stratified sampling scheme with the above mentioned allocation, under a general super-population model.

Chapter VII deals with the general problem of choice of an optimum sampling strategy. Accepting the criteria of unbiasedness and hyper admissibility the problem reduces to the question of possibility of construction of a sampling design D_1, for a given design D_2, with the same inclusion probabilities (so as to maintain the costs of the designs at the same level), but with uniformly smaller (or equal) joint inclusion probabilities π_{ij}'s. Barring certain trivialities, this is answered in the affirmative and a constructive proof is presented. An illustrative example at the end makes things clear.

Finally, in Chapter VIII we demonstrate the inadmissibility of customary estimators of population total in sampling over two occasions, by providing improved estimators under (i) equal probability sampling scheme and (ii) varying probability sampling scheme. Expressions for the gain in efficiency of the improved estimators are also derived.

CHAPTER II

CONCEPTS AND DEFINITIONS

In this chapter we explain the basic concepts and definitions which will be used in this thesis.

A collection of units $U_1, U_2, \ldots, U_N$ is known as a 'finite population' $\mathcal{U}$, where N is a known finite number and $U_1, U_2, \ldots, U_N$ are distinguishable. We denote the finite population by

$$\mathcal{U} : \left\{ U_1, U_2, \ldots, U_N \right\}. \tag{2.1}$$

A list of such units as (2.1) is termed as a 'sampling frame' and N is called the 'size of the population'.

Any finite ordered sequence of units from $\mathcal{U}$ is called a 'sample' from $\mathcal{U}$ and is denoted by

$$s = (u_1, u_2, \ldots, u_{n_s}) \tag{2.2}$$

which is a particular realization of units from $\mathcal{U}$, viz.,

$$s = (U_{i_1}, U_{i_2}, \ldots, U_{i_{n_s}}) \tag{2.3}$$

where $1 \leq i_k \leq N$ for $1 \leq k \leq n_s$.

We call n_s the 'sample size' and if ν_s stands for the number of distinct units in s we call ν_s the 'effective sample size'.

A collection **S** of samples s from $\mathcal{U}$ with a probability measure P defined on it such that corresponding to every $s \in S$ is a probability P_s attached such that $P_s \geq 0$ and $\sum_{s \in S} P_s = 1$, is called a 'Sample Design' and is denoted by

$$D = D(S, P) \qquad (2.4)$$

Thus the definition of the sample design gives us a method of choice of a sample, which needs the listing down of all the possible samples and choosing one from the list with the corresponding probability. But, in practice, it is very difficult to list down all possible samples and select one out of them. To draw n units with replacement from a population of N units, it needs listing down of N^n possible n-tuples and then we have to select one of the n-tuples with probability N^{-n}. It can be seen how unmanageable and difficult would this process be if we have to work with large values of N and n. As an alternative one can draw a unit from the population of N units with probability $\frac{1}{N}$ to each unit, replace it and repeat the method n times. We can thus work with reference to a single sampling frame consisting of N units and the process will be much simpler. This leads to the one-by-one drawing mechanisms which consist of drawing units from $\mathcal{U}$ one-by-one with replacement with probabilities which depend on the previous draws.

A 'drawing mechanism' (or a 'sampling mechanism') is a function

$$q\ (u,\ k,\ s_{k-1}) \tag{2.5}$$

in $u \in \mathcal{U}$, k, a positive integer, and s_{k-1} a sample of size $k-1$ such that

$$q(u,\ k,\ s_{k-1}) \geq 0 \quad \text{for all} \quad u,\ k \quad \text{and} \quad s_{k-1}$$

and $$\sum_{u \in \mathcal{U}} q(u,\ k,\ s_{k-1}) = 1 \quad \text{for all} \quad k \quad \text{and} \quad s_{k-1}\ . \tag{2.6}$$

Thus q denotes the probability of drawing the unit $u \in \mathcal{U}$ in the k^{th} draw, which depends on u and k and also on the outcome s_{k-1} of the previous $(k-1)$ draws. In this connection the following theorem is proved by Hanurav.

Theorem 2.1 (Hanurav [1962a]): To any given design $D(S,P)$ there corresponds a unique drawing mechanism M such that sampling according to M results in the design $D(S,P)$ and conversely.

This theorem enables us to work within the unified frame work of designs, while searching for optimum designs.

Considering the problem of estimation, let $\mathcal{Y}$ be a real-valued characteristic taking values Y_i on U_i of (2.1), $i = 1,2,\ldots, N$. Any function of $Y_1,Y_2,\ldots, Y_N$ is called a 'parametric function' and any function defined over a design such that for samples $s \in D$ the function depends only on the values of $\mathcal{Y}$ for the units belonging to the sample is called a 'statistic'. A statistic t when used to estimate a parametric function $T(Y_1,Y_2,\ldots, Y_N)$ is called an 'estimator' of T. An estimator t of T is called an 'Unbiased estimator' if and only if

$$E(t) = \sum_{s \in D} t_s P_s = T(Y_1,Y_2,\ldots,Y_N) \quad \text{for all values of } \underset{\sim}{Y} . \qquad (2.7)$$

An estimator t which is not unbiased for T is said to be 'Biased'.

The deviation of t_s from T is taken as the 'error' on the basis of the sample s, in the estimation of T. Any convex function $f(t_s - T)$ is called the 'loss function' and $E(f)$ is called the 'expected loss'. A oft-used loss function is the 'Mean Square Error' given by

$$M(t) = E(t_s - T)^2 \qquad (2.8)$$

where t_s is used as an estimate of T based on a sample s.

When t_s is an unbiased estimate of T then $E(t_s - T)^2$ is also the 'Variance' of t_s. It is easy to see that

$$V(t) = \sum_{s \subset S} t_s^2 \, p_s - T^2. \qquad (2.9)$$

For a given design $D(S,P)$ the 'inclusion probability' of a unit U_i is given by

$$\pi_i = \sum_{s \ni i} p_s \qquad (2.10)$$

summation being taken over all samples that contain U_i at least once. The 'joint inclusion probability' of a pair (U_i, U_j), $i \neq j$ is given by

$$\pi_{ij} = \sum_{s \ni i,j} p_s \qquad (2.11)$$

summation being taken over all samples which contain both U_i and U_j. Though it is more often used, a simple, rigorous proof was first given by Hanurav [1965] for the result that a Necessary and Sufficient Condition for the estimability of the population total Y is that $\pi_i > 0$, $i = 1,2,\ldots, N$ and that a Necessary and Sufficient Condition for the estimability of the variance of an unbiased estimator t of Y is that $\pi_{ij} > 0$ for $1 \leq i \neq j \leq N$.

A design $D(S,P)$ plus an estimator t of T, defined over D is called a 'Sampling Strategy' for the estimation of T. This definition is due to Hájek [1958] and bears relation to the Game theory and the importance of this terminology is stressed by Hanurav [1965,1966]. Thus we denote a strategy by

$$H(D, t) = H(S, P, t) . \tag{2.12}$$

A strategy $H(S, P, t)$ used for the estimation of T is called an 'Unbiased strategy' if t is an unbiased estimator of T. Otherwise, it is called a 'Biased strategy'. The expectation, variance or mean square error of a strategy are defined as the expectation, variance or mean square error of the estimator t over D.

Of the estimators t_1 and t_2 of a parametric function $T(Y_1, Y_2, \ldots, Y_N)$ both defined over a design $D(S, P)$, t_1 is said to be 'Uniformly better' than t_2 if and only if

$$M(t_1) \leq M(t_2) \quad \text{for all } \underset{\sim}{Y} , \tag{2.13}$$

with strict inequality at least once. Thus for a given design $D(S, P)$ and a class $\mathcal{C}$ of estimators of a parametric function $T(Y_1, Y_2, \ldots, Y_N)$, defined over D, a $t_1 \in \mathcal{C}$ is the 'best estimator' in $\mathcal{C}$ if

$$M(t_1) \leq M(t_2) \quad \text{for all } \underset{\sim}{Y} , \tag{2.14}$$

and for all t_2 different from t_1 and belonging to $\mathcal{C}$.

Thus we say that a strategy H_1 is better than another strategy H_2 iff

$$M(H_1) \leq M(H_2) \quad \text{for all } \underset{\sim}{Y}, \qquad (2.15)$$

with strict inequality at least once and that H_1 is best in a class $\mathcal{C}$ iff

$$M(H_1) \leq M(H_2) \quad \text{for all } \underset{\sim}{Y} \qquad (2.16)$$

and for all $H_2 \neq H_1$ and belonging to $\mathcal{C}$.

As mentioned before, we shall be concerned with finite populations alone and with reference to the estimation theory there is a natural deviation from the classical theory of estimation for the case of infinite populations, wherein a linear estimator based on a sample of n units is defined as $\sum_{i=1}^{n} w_i y_i$, w_i's being constants, independent of the observations y_i and $\bar{y} = \frac{1}{n} \sum_{1}^{n} y_i$ is the minimum variance unbiased estimator of $E(y_i)$ in the class of all linear unbiased estimators. The major difference between the theory of infinite and finite populations is that of 'distinguishability of units' which plays an important rôle in the case of finite populations. Thus it is important to know whether the same values of y belonged to the same unit repeated, or to two different units. This fact was first observed by Des Raj and Khamis [1958] and Basu [1958] who found that for a simple random sample of size n taken with replacement, the mean of the 'effective sample', viz., $\frac{1}{\nu_s}\sum_1^{\nu_s} y_i'$, is better than $\bar{y}$. While Des Raj and Khamis derived the result algebraically, Basu proved this by introducing first into sampling theory the concept of 'sufficient statistic'. Thus the estimator $\frac{1}{\nu_s}\Sigma y_i'$ has coefficients attached to y_i' which unlike the w_i's considered above, depend on not only ν_s but also on the sample selected till the (i-1) draws.

This necessitates a general definition of linear estimators and Horvitz and Thompson [1952] formulated the problem in accordance with a unified approach and defined 3 classes of linear estimators:

$t_1 = \sum_{i \in s} \beta_i y_i$, ~~where the coefficient~~ β_i is a constant to be used as weight for the element selected at the i^{th} draw;

$$t_2 = \sum_{i=1}^{n} c_i y_i ,$$

where the given constant c_i $(i = 1,2,\ldots, N)$ is attached to the i^{th} element whenever it is selected for the sample and

$$t_3 = \nu_s \sum_{i \in s} y_i ,$$

where the estimators have a single coefficient ν_s attached, whenever the s^{th} sample is selected. (2.17)

Ratio estimator falls under t_3 class and we study this in further detail in this thesis. Throught this thesis, the symbol $\sum_{i \in s}$ indicates that the summation is over all units U_i contained in the sample s. Godambe (1955) generalized and gave $t_s = \sum_{i \in s} \beta_{si} y_i$ as the most general estimator. We have the condition for

$$t_s = \sum_{i \in s} \beta_{si} y_i \tag{2.18}$$

to be unbiased for the population total Y, given by

$$\sum_{s \ni i} \beta_{si} P_s = 1, \quad 1 \leq i \leq N,$$

and we further have

$$V(t_s) = \sum_{i=1}^{N} Y_i^2 \left(\sum_{s \ni i} \beta_{si}^2 P_s - 1 \right) + \sum_{i \neq j}^{N} \sum^{N} Y_i Y_j \left(\sum_{s \ni i, j} \beta_{si} \beta_{sj} P_s - 1 \right). \tag{2.19}$$

Koop [1963] has proposed seven classes of linear estimators by making the coefficients of the sample observations depend on, what he calls as 'axioms of sample formation', based on (i) order of selection of the unit, (ii) its occurrence in the sample and (iii) the sample as a whole. Murthy [1963] has developed a 'technique of generating estimators' for any sampling design in which the class of parameters that can be expressed as sum of single valued set functions defined over a class of sets of units belonging to the finite population is considered and a number of possibly different unbiased estimators are generated.

After giving a generalized estimator Godambe [1955] proved the following theorem.

Theorem 2.2 (Godambe [1955]): There does not exist a minimum variance unbiased estimator for any general sampling design D. Later, Hanurav [1966] pointed out some exceptions to this and gave some non-trivial designs (called uni-cluster designs) where a best estimator exists.

This leads to the search of optimum estimators in a subclass of designs. The criterion 'admissibility' though helps in eliminating bad (inadmissible) estimators, does not help much in restricting to optimum estimators. Murthy's [1957], Des Raj and Khamis' [1958], Basu's [1958] and Roy and Chakravarty's [1960] results in this direction eliminate inadmissible estimators which depend on the order in which the units appear or which have a unit in the sample repeated. While in the uni-stage sampling Pathak [1964] has modified the usual estimators providing better estimators, extensions o two-stage and two-phase sampling to demonstrate the inadmissibility of customary estimators are given by

Ghosh [1963] and Pathak and Rao [1967] respectively. Recently, Godambe and Joshi [1965] and Joshi [1965a], [1965b] considered admissibility removing the restriction of linearity and later relaxing the criterion of unbiasedness. Other criteria studied are 'linear invariance' and 'regular estimators' by Roy and Chakravarty [1960] and the most recent one, 'hyperadmissibility' due to Hanurav [1966]. We shall discuss in a later chapter the various criteria of optimality.

The search for an optimum should be considered amongst designs that are equally costly. We consider the simple linear 'cost function',

$$C_s = C_0 + C_1 \nu_s, \qquad (2.20)$$

where C_s is the cost for a sample, consisting of the overhead cost C_0; the cost of collecting the data which is assumed to be proportional to ν_s; and C_1 standing for the cost of collection of data on a single unit. The cost of a strategy $C(H(S,P,t))$ is the expected cost of samples of H, given by

$$C(H) = C_0 + C_1 \nu(H). \qquad (2.21)$$

We consider the class of all strategies where $\nu(H) = \nu$, a given number, while discussing the criteria of an 'optimum strategy'.

If optimality is judged from uniform minimization of the variance of a strategy, then we can see that there does not exist such a one. But, whenever some auxiliary information on a characteristic $\mathcal{X}$ which takes values X_i on U_i, $i = 1, 2, \ldots, N$ is available, closely related to the study variable $\mathcal{Y}$, taking values Y_i on U_i, $i = 1, 2, \ldots, N$, it is possible to use this information in setting up a criterion of optimality. We have mentioned this briefly in the introductory chapter and here we shall explain it further.

The information on $\mathcal{X}$, known before hand can be used to assume a reasonable a priori distribution over $\underset{\sim}{Y}$. According to this 'super-population concept', as termed by Cochran [1946], $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_N)$ is assumed to be a realisation of a random N-vector with certain distribution depending on $\underset{\sim}{X} = (X_1, X_2, \ldots, X_N)$ and some unknown parameters. This distribution is denoted by ∂ and we can talk of expectations, Variances and Covariances taken over ∂. We now minimize the expected variance over ∂, namely

$$\int V(H)\, d\, \partial \qquad (2.22)$$

for H varying over $H(\nu)$ the class of all equi-cost strategies. A H_o which minimizes (2.22), uniformly with respect to all the parameters of the distribution ∂ is called a '∂-optimum strategy' in $H(\nu)$.

Two classes of *a priori* distributions Δ_1 due to Godambe [1955] and Δ_2 due to Hájek [1959] were introduced to study this problem.

Δ_1 is the class of all *a priori* distributions ∂_1 with the properties:

$$\left.\begin{array}{ll} \text{i)} & \mathcal{E}_{\partial_1}(Y_i \mid X_i) = aX_i \\ \text{ii)} & \mathcal{V}_{\partial_1}(Y_i \mid X_i) = \sigma^2 X_i^2 \\ \text{and iii)} & \mathcal{C}_{\partial_1}(Y_i, Y_j \mid X_i, X_j) = 0, \text{ where} \end{array}\right\} \quad (2.23)$$

$\mathcal{C}$ denotes the Covariance and Δ_2 is the class of all *a priori* distributions ∂_2 satisfying

$$\left.\begin{array}{ll} \text{i)} & \mathcal{E}_{\partial_2}(Y_i \mid X_i) = aX_i \\ \text{ii)} & \mathcal{V}_{\partial_2}(Y_i \mid X_i) = \sigma^2 X_i^2 \\ \text{and iii)} & \mathcal{C}_{\partial_2}(Y_i, Y_j \mid X_i, X_j) = \omega(\mid j-i \mid), \end{array}\right\} \quad (2.24)$$

where $\mathcal{C}$ denotes the Covariance and ω is a single valued convex function. It is easy to see that Δ_2 is wider than Δ_1.

Godambe [1955] proved that if ν is an integer $\left(\nu < \frac{X}{\text{Max. } X_i}\right)$ any strategy $H_o\ (D_o,\ t_o)$ in $H\ (\nu)$ with

$$\left.\begin{aligned} &(a)\ \pi_i = \nu \frac{X_i}{\sum_{i=1}^{N} X_i}, \quad i = 1, 2, \ldots, N \\ &(b)\ \nu_s = \nu \quad \text{for all } s \text{ with } P_s > 0, \\ \text{and}\quad &(c)\quad t_o = \hat{Y}_{HT}\ (D_o) = \sum_{i \in s} \frac{y_i}{\pi_i(D_o)} \end{aligned}\right\} \quad (2.25)$$

is ∂_1-optimum for any $\partial_1 \in \Delta_1$.

Hájek [1959] proved that the strategy $H_o'\ (D_o',\ t_o')$ given by the varying probability linear systematic sampling scheme (Goodman and Kish [1950]) with

$$\left.\begin{aligned} &(a)\quad \pi_i = \nu X_i \Big/ \sum_{1}^{N} X_i, \quad i = 1, 2, \ldots, N \\ &(b)\ \nu_s = \nu \quad \text{for all } s \text{ with } P_s > 0 \text{ and} \\ &(c)\quad t_o' = \hat{Y}_{HT}\ (D_o') = \sum_{i \in s} \frac{y_i}{\pi_i(D_o')} \end{aligned}\right\} (2.26)$$

is Δ_2 -optimum for any $\partial_2 \in \Delta_2$. The sampling design D_o' has the most important disadvantage that the variance of the optimum estimator is not estimable. So, we shall restrict ourselves to the model Δ_1. But in Δ_1 we have taken $V_{\partial_1}(Y_i \mid X_i) = \sigma^2 X_i^2$ and there are many situations where this is not so; in fact, it was shown that $V(Y_i \mid X_i) = \sigma^2 X_i^g$, where g lies in many practical situations between 1 and 2 (Mahalanobis [1944], Smith [1938], Jessen [1942]). So, it is really of interest to study our problems under such a realistic general model. We consider the class Δ_g of all prior distributions ∂_g (wider than Δ_1) satisfying

$$\left.\begin{array}{ll} \text{i)} & \mathcal{E}_{\partial_g}(Y_i \mid X_i) = a\, X_i \\ \text{ii)} & \mathcal{V}_{\partial_g}(Y_i \mid X_i) = \sigma^2 X_i^g \\ \text{and iii)} & \mathcal{C}_{\partial_g}(Y_i, Y_j \mid X_i, X_j) = 0 \end{array}\right\} \quad (2.27)$$

where $\mathcal{C}$ denotes the covariance and study the problems in this thesis under Δ_g

CHAPTER III

UNBIASED RATIO ESTIMATORS

3.0 Summary

In this chapter, after a brief introduction to biased ratio estimators, we study the problem of construction of unbiased ratio estimators and show how unbiased ratio estimators due to Murthy and Nanjamma [1959], Hartley and Ross [1954] and Nieto de Pascual [1961], for the estimation of the population mean $\bar{Y}$ of a characteristic $\mathcal{Y}$ taking values Y_i on the unit U_i, can be obtained as linear combinations of the two biased estimators $\frac{\bar{y}}{\bar{x}}\bar{X}$ and $\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i}$, where $\bar{X}$ is the mean of the auxiliary characteristic $\mathcal{X}$, closely related to $\mathcal{Y}$, taking values X_i on the unit U_i. Some interesting relations among these estimators have been mentioned. The relation between ratio estimators and convexity is briefly given and a new concept 'stochastic convexity' is introduced which figures in the construction of unbiased ratio estimators.

3.1 Introduction

While estimating the population parameters, such as population total or population mean of a characteristic of interest, say $\mathcal{Y}$, taking values Y_i on the units U_i, $i = 1,2,\ldots, N$, of a finite population of size N, it is often advantageous to make use of information, readily available on an auxiliary variate $\mathcal{X}$, taking values X_i on the units U_i, $i = 1,2,\ldots, N$, closely related to $\mathcal{Y}$, for increasing the precision of the estimates. One such method where auxiliary information is made use of is the method of ratio estimation which is the simplest and is found to be more efficient than those which are not based on such information. The method of ratio estimation for estimating the population mean $\bar{Y} = \sum_1^N Y_i/N$ (or equivalently the population total Y) of the characteristic $\mathcal{Y}$, consists in getting an estimator $\hat{R}$ of the population ratio

$$R = \frac{Y}{X} = \frac{\sum_1^N Y_i}{\sum_1^N X_i}$$

and multiplying this estimator by the known population mean $\bar{X} = \frac{\sum_1^N X_i}{N}$, of the $\mathcal{X}$-variate (or X). It is well known that the ratio estimator would be more efficient than the estimator based on sample mean, if the correlation coefficient between $\hat{Y}$ and $\hat{X}$ is greater than half the ratio of the coefficient of variation of $\hat{X}$ to the coefficient of variation of $\hat{Y}$.

3.2 Bias of the ratio estimator

It is easily seen that the usual estimator $\hat{R} = \frac{\hat{Y}}{\hat{X}}$, where $\hat{Y}$ and $\hat{X}$ are unbiased estimators of Y and X respectively, is biased. Approximate expressions for the bias and variance of the estimator were given by Hansen, Hurwitz and Madow [1953], Koop [1951], Cochran [1953], Sukhatme [1954], Murthy and Nanjamma [1959], and other authors with varying degree of approximation. The validity of the method depends on the assumption that $\left| \frac{\hat{X} - X}{X} \right| < 1$ and it was observed by Murthy and Nanjamma [1959] that if the sample size is fairly large this assumption is valid and for simple random sampling many empirical studies have shown the validity of the assumption for samples of size not less than 30. Koop's method [1951] applies an ingenious device which does not use this assumption and gives an expression for the bias of the ratio estimator to the order of n^{-4}. Exact expressions for the bias and the mean square error of the ratio estimator are, however, given in a later chapter which use coefficients based on the auxiliary information.

Quenouille [1956] has proposed the estimator $\hat{R}_Q = g\,\hat{R} - \frac{g-1}{g} \sum_{1}^{g} \hat{R}_j$ which has bias of order n^{-2} at most, where $\hat{R}$ is an estimator based on n observations whose bias is $an^{-1} + O(n^{-2})$, a being a constant and g the number of groups each of size k, into which the sample is randomly divided such that $n = kg$ and $\hat{R}_j$ is the estimator obtained from the sample after omitting the j^{th} group. Durbin [1959] has made use of this result in the case of ratio estimators of the type $\hat{R} = \frac{\hat{Y}}{\hat{X}}$, when the regression of $\hat{Y}$ on $\hat{X}$ is linear and $\hat{X}$ is normally distributed and has shown that $\hat{R}_Q$ has asymptotically smaller variance, when $g = 2$. J.N.K.Rao [1965] gets an extension of this result, obtaining expressions to $O(n^{-3})$ of the bias and variance of $\hat{R}_Q$ for any g under the same assumptions as of Durbin and shows that $g = n$ would be the optimum choice of the number of groups.

3.3 Construction of unbiased ratio estimators

While, on one hand, there is an attempt to reduce the bias, on the other hand, there was an attempt by many authors to construct unbiased ratio estimators, during the last decade or so. The usual estimators suggested for estimating the population ratio ar given by

$$R_1' = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} \qquad (3.3.1)$$

and

$$R_n' = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i} \quad . \qquad (3.3.2)$$

Both these estimators are biased and in most of the cases the bias of one of these estimators is estimated and then the estimator is corrected for its bias. Estimators of this type were given by Murthy and Nanjamma [1959], Hartley and Ross [1954] and Nieto de Pascual [1961]. In the next sections we briefly review these estimators.

3.3.1 Murthy and Nanjamma's estimator:

Consider (3.3.1) and (3.3.2) suggested for the estimation of $R = \frac{Y}{X}$. Murthy and Nanjamma [1959] established the following

Lemma 3.3.1: The bias of R_n' is n times that of R_1', to the second degree of approximation.

Also, since

$$E(R_1') = R + B(R_1')$$

and

$$E(R_n') = R + B(R_n'),$$

where B stands for the bias of the estimator, follows that

$$E(R_n' - R_1') = B(R_n') - B(R_1')$$

$$= (n-1)\ B(R_1'), \quad \text{by Lemma 3.3.1.}$$

$$\text{So,} \quad B(R_1') = \frac{R_n' - R_1'}{n - 1} \text{.}$$

Having estimated the bias of R_1', R_1' is then corrected for its bias to get an (almost) unbiased ratio estimator. Thus we have

$$\hat{R}_{MN} = R_1' - \hat{B}\ (R_1') = \frac{nR_1' - R_n'}{n - 1} \text{.} \qquad (3.3.3)$$

It can be seen that an (almost) unbiased estimate of the population mean $\bar{Y}$ (or equivalently the population total Y) is immediate from (3.3.3) by multiplying it by $\bar{X}$ (or X). Thus, we have

$$\hat{\bar{Y}}_{MN} = \frac{nR_1 - R_n}{n - 1} \text{,} \qquad (3.3.4)$$

where

$$R_1 = \bar{X}\ \frac{\Sigma\, y_i}{\Sigma\, x_i} \text{ and } R_n = \frac{\bar{X}}{n}\ \Sigma\, \frac{y_i}{x_i} \text{.}$$

3.3.2 Hartley and Ross' estimator:

Hartley and Ross [1954] considered the estimator

$$\frac{\bar{X}}{n} \sum_{1}^{n} \frac{y_i}{x_i}$$

and have shown that

$$B\left(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}\right) = - \text{cov.} \left(\frac{y}{x}, x\right)$$

and they unbiasedly estimated this bias assuming that the sample is drawn by Simple Random Sampling Without Replacement (SRSWOR) and corrected the estimator for its bias. Thus the unbiased estimator due to Hartley and Ross takes the form

$$\hat{Y}_{HR} = \frac{\bar{X}}{n} \sum_{1}^{n} \frac{y_i}{x_i} + \frac{N-1}{N} \frac{n}{n-1}\left(\bar{y} - \frac{\bar{x}}{n} \sum_{1}^{n} \frac{y_i}{x_i}\right). \qquad (3.3.5)$$

3.3.3 Nieto de Pascual's estimator:

Nieto de Pascual [1961] considered the other estimator, viz.,

$$\bar{X} \frac{\Sigma y_i}{\Sigma x_i}$$

and got an approximate expression for its bias, using Lemma 3.3.1 and when the estimator is corrected for its bias he gets the corrected estimator as

$$\hat{\bar{Y}}_p = \frac{\bar{y}}{\bar{x}} \bar{X} + \frac{N-1}{N} \frac{1}{n-1} \left(\bar{y} - \frac{\bar{x}}{n} \sum_{1}^{n} \frac{y_i}{x_i}\right). \qquad (3.3.6)$$

3.4 General method of constructing Unbiased ratio estimators:

We now consider the two ratio estimators $\frac{\bar{y}}{\bar{x}}\bar{X}$ and $\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i}$ for estimation of the population mean $\bar{Y}$ and present a general method of deriving unbiased ratio estimators by considering linear combinations of these two estimators and determining the coefficients for which the combined estimators are unbiased. For particular values of the coefficients, we show that the estimators (3.3.4), (3.3.5) and (3.3.6) are obtained as particular cases.

Theorem 3.4.1: The linear combination $\theta\frac{\bar{y}}{\bar{x}}\bar{X} + (1-\theta)\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i}$, where θ is a constant is unbiased (almost) with $\theta = \frac{n}{n-1}$, for $\bar{Y}$.

Proof: For the linear combination

$$\theta\frac{\bar{y}}{\bar{x}}\bar{X} + (1-\theta)\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i} \qquad (3.4.1)$$

to be unbiased for $\bar{Y}$, we have the condition that

$$E\left(\theta\frac{\bar{y}}{\bar{x}}\bar{X} + (1-\theta)\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i}\right) = \bar{Y},$$

i.e.

$$\theta\bar{Y} + \theta B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right) + (1-\theta)\bar{Y} + (1-\theta)B\left(\frac{\bar{X}}{n}\Sigma\frac{y_i}{x_i}\right) = \bar{Y}$$

or $$\frac{B\left(\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i}\right)}{B\left(\frac{\bar{y}}{\bar{x}}\;\bar{X}\right)} = -\,\frac{\theta}{1-\theta}\,.$$

From Lemma 3.3.1, it is clear that this ratio is always positive, and hence follows that θ does not belong to the open interval (0, 1). Further more, we have

$$-\,\frac{\theta}{1-\theta} = n$$

and hence

$$\theta = \frac{n}{n-1}\,. \tag{3.4.2}$$

Thus the linear combination considered in (3.4.1) yields the estimator

$$\hat{\bar{Y}} = \frac{n}{n-1}\,\frac{\bar{y}}{\bar{x}}\,\bar{X} - \frac{1}{n-1}\,\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i} \tag{3.4.3}$$

which is nothing but the Murthy-Nanjamma's estimator given in (3.3.4).

<u>Remark.</u> A similar approach by Hartley (unpublished lecture notes, 1954) of combining these two types of ratio estimators has been pointed out by Nieto de Pascual [1961], which results in (3.3.4).

Theorem 3.4.2: The linear combination $\theta \frac{\bar{\bar{y}}}{\bar{x}} \bar{X} + (1-\theta) \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}$,

where θ is a random variable is unbiased with

$\theta = \frac{n}{n-1} \frac{N-1}{N} \frac{\bar{\bar{x}}}{\bar{X}}$ for a Simple Random Sample of size n taken With Out Replacement.

Proof: Consider a combined estimator of the type

$$\hat{\bar{Y}}' = \theta \frac{\bar{\bar{y}}}{\bar{x}} \bar{X} + (1-\theta) \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}, \qquad (3.4.4)$$

where θ is a random variable.

We have for the unbiasedness of $\hat{\bar{Y}}'$

$$E(\theta \frac{\bar{\bar{y}}}{\bar{x}} \bar{X} + (1-\theta) \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) = \bar{Y}$$

i.e.

$$E(\theta \frac{\bar{\bar{y}}}{\bar{x}} \bar{X} - \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) = -B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}). \qquad (3.4.5)$$

Having assumed that the sample is a SRSWOR, we have

$$B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) = -\text{cov.}(\frac{y}{x}, x)$$

$$= -\text{cov.}(\frac{1}{n} \Sigma \frac{y_i}{x_i}, \bar{x}) \frac{Nn}{N-n} \frac{N-1}{N}$$

$$= \frac{-n(N-1)}{N-n} \left\{ E(\frac{\bar{x}}{n} \Sigma \frac{y_i}{x_i}) - \bar{Y} - B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) \right\},$$

which on simplification gives that

$$B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) = \frac{n}{n-1} \frac{N-1}{N} \left\{ E(\frac{\bar{x}}{n} \Sigma \frac{y_i}{x_i} - \bar{y}) \right\} . \tag{3.4.6}$$

Using (3.4.6) in (3.4.5) we get

$$E(\theta \frac{\bar{X}}{\bar{x}} \left\{ \bar{y} - \frac{\bar{x}}{n} \Sigma \frac{y_i}{x_i} \right\}) = E(\frac{n}{n-1} \frac{N-1}{N} \left\{ \bar{y} - \frac{\bar{x}}{n} \Sigma \frac{y_i}{x_i} \right\})$$

for which

$$\theta = \frac{n}{n-1} \frac{N-1}{N} \frac{\bar{x}}{\bar{X}} \tag{3.4.7}$$

is a solution and substituting this value of θ in (3.4.4) the combined estimator takes the form

$$\hat{\bar{Y}}' = \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i} + \frac{n}{n-1} \frac{N-1}{N} (\bar{y} - \frac{\bar{x}}{n} \Sigma \frac{y_i}{x_i}) \tag{3.4.8}$$

which is just the Hartley-Ross' estimator (3.3.5).

<u>Corollary 3.4.2a</u>: When θ in (3.4.4) is a constant, it follows from (3.4.5) that

$$\theta = \theta_M = \frac{B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i})}{B(\frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}) - B(\frac{\bar{y}}{\bar{x}} \bar{X})} \simeq \frac{n}{n-1}$$

and thus from this (3.3.4) is immediate.

Denote θ of (3.4.7) by θ_H.

<u>Corollary 3.4.2b:</u> The expected value of θ_H is greater than unity while θ_M is greater than unity.

<u>Proof:</u> $\theta_M = \frac{n}{n-1} > 1$ and

$$E(\theta_H) = E\left(\frac{n}{n-1} \cdot \frac{N-1}{N} \cdot \frac{\bar{\bar{x}}}{\bar{\bar{X}}}\right) = \frac{n}{n-1} \cdot \frac{N-1}{N} > 1.$$

<u>Theorem 3.4.3:</u> The linear combination $\theta \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i} + (1-\theta)\frac{\bar{y}}{\bar{x}} \bar{X}$, where θ is a random variable is unbiased with

$$\theta = -\frac{N-1}{N(n-1)} \frac{\bar{\bar{x}}}{\bar{\bar{X}}}$$

for a simple random sample of size n taken without replacement.

<u>Proof:</u> Consider a combined estimator of the type

$$\hat{\bar{Y}}'' = (1-\theta) \frac{\bar{y}}{\bar{\bar{x}}} \bar{X} + \theta \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i} \qquad (3.4.9)$$

where θ is a random variable.

For $\hat{\bar{Y}}''$ to be unbiased for $\bar{Y}$, we have

$$E(\hat{\bar{Y}}'') = E[(1-\theta) \frac{\bar{y}}{\bar{\bar{x}}} \bar{X} + \theta \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}] = \bar{Y},$$

i.e. $$E(\theta \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i} - \frac{\bar{y}}{\bar{\bar{x}}} \bar{X}) = -B(\frac{\bar{y}}{\bar{x}} \bar{X}) . \qquad (3.4.10)$$

We have

$$B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right) = -\,\text{cov.}\left(\frac{\bar{y}}{\bar{x}}\,,\ \bar{x}\right) \simeq -\frac{1}{n}\,\text{cov.}\left(\frac{y}{x}\,,\ x\right).$$

Using (3.4.6) we have

$$B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right) \simeq \frac{1}{n-1}\,\frac{N-1}{N}\left(E\left\{\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i} - \bar{y}\right\}\right) \qquad (3.4.11)$$

From (3.4.10) and (3.4.11) it follows that

$$E\left(\theta\left\{\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\}\right) = E\left(-\frac{1}{n-1}\,\frac{N-1}{N}\,\frac{\bar{x}}{\bar{X}}\left\{\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\}\right).$$

A solution of this equation gives

$$\theta = -\frac{N-1}{N(n-1)}\,\frac{\bar{x}}{\bar{X}} \qquad (3.4.12)$$

which on substitution in the combined estimator (3.4.9) gives

$$\hat{\bar{Y}}'' = \frac{\bar{y}}{\bar{x}}\bar{X} + \frac{N-1}{N}\,\frac{1}{n-1}\left(\bar{y} - \frac{\bar{x}}{n}\,\Sigma\,\frac{y_i}{x_i}\right) \qquad (3.4.13)$$

which is due to Nieto de Pascual [1961].

<u>Corollary 3.4.3a:</u> If θ of theorem (3.4.3) is a constant, then from (3.4.10) follows that

$$\theta = \theta'_M = \frac{-\,B\left(\frac{\bar{y}}{\bar{x}}\,\bar{X}\right)}{B\left(\frac{\bar{X}}{n}\,\Sigma\,\frac{y_i}{x_i}\right) - B\left(\frac{\bar{y}}{\bar{x}}\,\bar{X}\right)} \simeq -\frac{1}{n-1}$$

which on substitution in (3.4.9) gives the Murthy-Nanjamma estimator (3.3.4).

Denote θ of (3.4.12) by θ_p.

<u>Corollary 3.4.3b:</u> The expected value of θ_p is less than zero while θ'_M itself is less than zero.

<u>Proof:</u> $\theta'_M = -\frac{1}{n-1} < 0$ and

$$E(\theta_p) = E(-\frac{1}{n-1} \cdot \frac{N-1}{N} \cdot \frac{\bar{x}}{\bar{X}}) = -\frac{1}{n-1}\frac{N-1}{N} < 0.$$

<u>3.5 Ratio estimators and convexity:</u>

Let R_1 and R_2 be two ratio estimators with bias in the same direction. Then any convex combination of R_1 and R_2 would also be biased. For, consider a convex combination, say,

$$\lambda R_1 + (1-\lambda)R_2, \quad \text{where } \lambda \in \text{ open } (0, 1).$$

Then

$$E(\lambda R_1 + (1-\lambda) R_2) = \lambda R + \lambda B(R_1) + (1-\lambda)R + (1-\lambda)B(R_2)$$

$$= R + \lambda B(R_1) + (1-\lambda)B(R_2)$$

$$= R + B,$$

where $B \gtrless 0$, since $B(R_1)$ and $B(R_2)$ are either both positive or both negative, and therefore

$$E(\lambda R_1 + (1-\lambda) R_2) \neq R.$$

If the combination $\lambda R_1 + (1-\lambda) R_2$ for any real λ (constant) were to be unbiased, we should have

$$\lambda B(R_1) + (1 - \lambda)B(R_2) = 0,$$

which implies that

$$\frac{\lambda}{\lambda - 1} = \frac{B(R_2)}{B(R_1)} .$$

Let $\frac{B(R_2)}{B(R_1)} = \rho$, positive and not equal to 1.

We distinguish two cases:

i) $\rho > 1$:

Here $\frac{\lambda}{\lambda-1} = \rho > 1$, giving

$$\lambda = \frac{\rho}{\rho - 1} > 1.$$

ii) $0 < \rho < 1$:

Here $\frac{\lambda}{\lambda-1} = \rho < 1$, which gives

$$\lambda < 0.$$

Thus, inorder that the combined estimator

$$\lambda R_1 + (1 - \lambda) R_2$$

be an unbiased estimator, we should necessarily have that λ should not belong to the open interval (0, 1). Thus any convex combination of R_1 and R_2 cannot yield unbiased ratio estimators and if a linear combination of R_1 and R_2

considered above should be unbiased, then it should not be a convex combination.

Recalling our method of construction of unbiased ratio estimators, we see that the two estimators considered have bias in the same direction and further the combination $\theta R_1 + (1 - \theta)R_n$, where R_1 and R_n are defined in (3.3.4), is unbiased with $\theta = \frac{n}{n-1} > 1$ and the combination $(1 - \theta)R_1 + \theta R_n$ is unbiased with $\theta = -\frac{1}{n-1} < 0$.

Next when we consider the linear combination

$$\theta \frac{\bar{y}}{\bar{x}} \bar{X} + (1 - \theta) \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i}$$

where θ is a random variable (Theorem 3.4.2); we observe that this is unbiased with $\theta = \frac{n}{n-1} \frac{N-1}{N} \frac{\bar{x}}{\bar{X}}$. Here it is quite interesting to see from Corollary 3.4.2b, that $E(\theta) > 1$. Also, considering the linear combination

$$(1 - \theta) \frac{\bar{y}}{\bar{x}} \bar{X} + \theta \frac{\bar{X}}{n} \Sigma \frac{y_i}{x_i} ,$$

being a random variable (Theorem 3.4.3), we find that it is unbiased with $\theta = -\frac{1}{n-1} \frac{N-1}{N} \frac{\bar{x}}{\bar{X}}$. From Corollary 4.3b, it now follows that $E(\theta) < 0$. Thus it is clear that he linear combinations considered here are unbiased when the ombining factors are such that their expectations are greater than unity in one case and less than zero in the other case.

This leads us to a new definition.

<u>Definition 3.5.1:</u> A set $\mathcal{C}$ of random variables is said to be Stochastic Convex if $x_1 \in \mathcal{C}$, $x_2 \in \mathcal{C}$ implies that $\lambda x_1 + (1 - \lambda) x_2 \in \mathcal{C}$ for any random variable λ, where $E(\lambda) \in$ open $(0, 1)$.

Thus we see that, while the unbiased ratio estimators derived in Corollaries (3.4.2a) and (3.4.3a) are not convex combinations of the estimators R_1 and R_n considered there the unbiased ratio estimators of Theorems 3.4.2 and 3.4.3 are not stochastic convex combinations (in the sense of definition 3.5.1) of the estimators R_1 and R_n.

CHAPTER IV

SAMPLING SCHEMES PROVIDING UNBIASED RATIO ESTIMATORS

4.0 Summary:

After reviewing the sampling schemes which provide unbiased ratio estimators, we give an exact expression for the variance of the ratio estimator for one such simple scheme which is due to Midzuno [1952] and Sen [1952]. This expression involves coefficients based on the auxiliary information and many interesting properties of these coefficients are proved. Three different estimators of this variance are studied. Finally, exact expressions for the Bias and Mean Square Error of the ratio estimator for simple random sampling which make use of the above mentioned coefficients are obtained. As an illustration, numerical examples are worked out at the end.

4.1 Introduction

We have seen in the previous chapter how unbiased ratio estimators are constructed. The other aspect of the problem is to get sampling schemes which provide unbiased ratio estimators. Consider the usual estimator $\sum_{i \in s} y_i / \sum_{i \in s} x_i$ for the estimation of the population ratio $R = \frac{Y}{X}$. This is in general biased and here we must make the estimator unbiased by properly choosing the sampling design.

It is shown by Hanurav [1962a] that any sampling design D is generated by a drawing mechanism Q(see def. 2.5). If D is a partially specified design, we can choose additional specifications consistent with the initial specifications and which together with the initial specifications completely specify the design D and then we can have a drawing mechanism Q' which generates the design satisfying the initial specifications. To make the drawing mechanism Q' as simple as possible we must choose the additional specifications in an efficient manner. This occasion arises while considering the present problem of making the estimator $\sum_{i \in s} y_i / \sum_{i \in s} x_i$ unbiased. Midzuno [1952] and Sen [1952] gave a partial specification of the design D(S,P) by demanding that each sample $s \in S$ should contain n distinct units. Then a probability measure P was found such that $\sum_{i\in s} y_i / \sum_{i\in s} x_i$ is unbiased for $R = \frac{Y}{X}$. It can be easily verified that this will be unbiased if the probability of selecting the sample is given by

$$P_s = \sum_{i\in s} x_i \Big/ \binom{N-1}{n-1} X \tag{4.1.1}$$

4.2 Lahiri's method of selection:

Lahiri [1951] suggested that the estimator $\sum_{i\in s} y_i / \sum_{i\in s} x_i$ will be unbiased if the sample is selected with probability proportional to its total size $\sum_{i\in s} x_i$, since P_s, the probability of selecting the sample given by (4.1.1) is nothing but $\sum_{i\in s} x_i \Big/ \sum_{s\in S} \left(\sum_{i\in s} x_i \right)$. Thus his procedure involves finding out the totals for every possible sample and then to select one sample with probability proportional to the total. Listing down all possible samples, finding out their totals of size and selecting the sample would be time consuming and not so easy for larger values of n and N.

4.3 Midzuno [1952] - Sen [1952] sampling scheme:

Midzuno [1952] and Sen [1952] have independently given a very simple procedure which makes the ratio estimator unbiased. We have already mentioned in the introduction the ingenuity of their method. For further discussion we refer to Hanurav [1966]. Their method consists in drawing the first unit with probability proportional to size and the rest of the $(n-1)$ units by SRSWOR from the $(N-1)$ units of the population. It can be easily derived that the probability of selection of the sample under this scheme is given by

$$P_s = \sum_{i\in s} x_i \Big/ \binom{N-1}{n-1} X$$

which is the same as (4.1.1).

4.4 Variance of the ratio estimator:

In this section we derive an exact expression of the variance of the ratio estimator $\sum_{i\in s} y_i \Big/ \sum_{i\in s} x_i$ under the Midzuno-Sen sampling scheme of (4.3). Use is made of certain coefficients which are based on the auxiliary information available.

Firstly, it is easily verified that with the probability of selection of the sample given by (4.1.1) the estimator is unbiased. For,

$$E(\hat{R}) = \sum_{s \in S} \left\{ \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i} \; P_s \right\}$$

$$= \sum_{s \in S} \left\{ \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i} \; \frac{\sum_{i \in s} x_i}{\binom{N-1}{n-1} X} \right\}$$

$$= \frac{\binom{N-1}{n-1} Y}{\binom{N-1}{n-1} X} = \frac{Y}{X} = R .$$

Next, we have

$$V(\hat{R}) = E(\hat{R}^2) - R^2$$

$$= E \left\{ \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i} \right\}^2 - R^2$$

$$= \sum_{s \in S} \left\{ \frac{(\sum_{i \in s} y_i)^2}{(\sum_{i \in s} x_i)^2} \; P_s \right\} - R^2$$

$$= \frac{1}{\binom{N-1}{n-1} X} \sum_{s \in S} \frac{(\sum_{i \in s} y_i)^2}{(\sum_{i \in s} x_i)} - R^2 . \qquad (4.4.1)$$

t X_g^i be the sum of the g^{th} set of (n-1) distinct X's other than X_i and let G be the class of all such $\binom{N-1}{n-1}$ sets and let $X_{g'}^{ij}$ be the sum of the g'^{th} set of (n-2) distinct X's other than X_i, X_j and let G' be the class of all such $\binom{N-2}{n-2}$ sets. Further let

$$\left.\begin{aligned} J_i &= \sum_{g \in G} \frac{1}{X_i + X_g^i} \\ \text{and} \quad J_{ij} &= \sum_{g' \in G'} \frac{1}{X_i + X_j + X_{g'}^{ij}} \,. \end{aligned}\right\} \qquad (4.4.2)$$

Then, collecting the coefficients of Y_i^2 and Y_iY_j from the r.h.s. of (4.4.1) we get

$$V(\hat{R}) = \frac{1}{\binom{N-1}{n-1}X}\left(\sum_i^N Y_i^2 J_i + \sum_{i \neq j}^N \sum^N Y_iY_j J_{ij}\right) - R^2 \qquad (4.4.3)$$

$$\left.\begin{aligned} \text{Define} \quad \lambda_i &= \frac{J_i}{\binom{N-1}{n-1}X} - \frac{1}{X^2} \\ \text{and} \quad \lambda_{ij} &= \frac{J_{ij}}{\binom{N-1}{n-1}X} - \frac{1}{X^2} \,. \end{aligned}\right\} \qquad (4.4.4)$$

Then

$$V(\hat{R}) = \sum_{i=1}^N \lambda_i Y_i^2 + \sum_{i \neq j}^N \sum^N \lambda_{ij} Y_iY_j \,. \qquad (4.4.5)$$

Thus we have the following theorem.

<u>Theorem 4.4.1:</u> The variance of the ratio estimator $\sum_{i\in s} y_i \big/ \sum_{i\in s} x_i$ under the Midzuno-Sen sampling scheme is given by

$$\sum_{i=1}^{N} \lambda_i Y_i^2 + \sum_{i\neq j}^{N}\sum^{N} \lambda_{ij}\, Y_i Y_j,$$

where λ_i, λ_{ij} are as defined above.

<u>Corollary 4.4.1:</u> For the estimator of the population total Y given by $\left(\sum_{i\in s} y_i \big/ \sum_{i\in s} x_i\right)X$, we have

$$V(\hat{Y}) = \sum_{i=1}^{N} (T_i - 1)\, Y_i^2 + \sum_{i\neq j}^{N}\sum^{N} (T_{ij} - 1) Y_i Y_j \tag{4.4.6}$$

where

$$T_i = \frac{X}{\binom{N-1}{n-1}} \sum_{g\in G} \frac{1}{X_i + X_g^i}$$

and

$$T_{ij} = \frac{X}{\binom{N-1}{n-1}} \sum_{g'\in G'} \frac{1}{X_i + X_j + X_{g'}^{ij}} . \tag{4.4.7}$$

Proof: Follows from (4.4.5) by multiplying it by X^2.

4.5 Estimation of variance:

In this section, we consider some of the estimators of the variance function (4.4.5) and study their merits and demerits.

First we prove a Lemma.

Lemma 4.5.1. (Sen [1953]): The probability of inclusion of the i^{th} unit in the sample for the Midzuno-Sen sampling scheme is given by

$$\pi_i' = \frac{n-1}{N-1} + \frac{N-n}{N-1} P_i$$

and the probability of joint inclusion of the units i and j in the sample is given by

$$\pi_{ij}' = \frac{n-1}{N-1} \frac{n-2}{N-2} + \frac{n-1}{N-1} \frac{N-n}{N-2} (P_i + P_j),$$

where $P_i = \frac{X_i}{X}$.

Proof: A direct proof is as follows.

Since, $P_s = \sum_{i \in s} x_i \Big/ \binom{N-1}{n-1} X$

$$\pi_i' = \sum_{s \ni i} P_s \quad \text{(see def. 2.10)}$$

$$= \sum_{s \ni i} \frac{\sum_{i \in s} x_i}{\binom{N-1}{n-1} X} = \frac{X_i \binom{N-1}{n-1} + (X - X_i)\binom{N-2}{n-2}}{\binom{N-1}{n-1} X}$$

$$= \frac{n-1}{N-1} + \frac{N-n}{N-1} \cdot P_i \tag{4.5.1}$$

$$\pi'_{ij} = \sum_{s \ni i,j} P_s \qquad \text{(see def. 2.11)}$$

$$= \sum_{s \ni i,j} \frac{\sum_{i \in s} x_i}{\binom{N-1}{n-1} X} = \frac{(X_i + X_j)\binom{N-2}{n-2} + (X - X_i - X_j)\binom{N-3}{n-3}}{\binom{N-1}{n-1} X}$$

$$= \frac{n-1}{N-1}\,\frac{n-2}{N-2} + \frac{n-1}{N-1}\,\frac{N-n}{N-2}\,(P_i + P_j) . \qquad (4.5.2)$$

__Theorem 4.5.1:__ An unbiased estimate of the variance of $\hat{R}$ is given by

$$\hat{V}(\hat{R}) = \sum_{i \in s} \lambda_i\, y_i^2 / \pi'_i + \sum_{(i \neq j) \in s}\sum \lambda_{ij}\, y_i y_j / \pi'_{ij}, \text{ where}$$

π'_i and π'_{ij} are given by (4.5.1) and (4.5.2).

__Proof:__ $E\,\hat{V}(\hat{R}) = E\left(\sum_{i \in s} \lambda_i\, y_i^2 / \pi'_i + \sum_{(i \neq j) \in s}\sum \lambda_{ij} y_i y_j / \pi'_{ij}\right)$

$$= \sum_{1}^{N} \lambda_i Y_i^2 + \sum_{i \neq j}^{N}\sum^{N} \lambda_{ij} Y_i Y_j = V(\hat{R}) \text{ and hence etc.}$$

(λ_i, λ_{ij} are known quantities.)

__Lemma 4.5.2:__ λ_i is positive for all i.

__Proof:__ From (4.4.7),

$$T_i = \frac{X}{\binom{N-1}{n-1}} \left(\frac{1}{X_i + X_1 + \dots + X_{n-1}} + \frac{1}{X_i + X_2 + \dots + X_n} + \dots \binom{N-1}{n-1} \text{ terms} \right)$$

= inverse of the harmonic mean of

$$\frac{X_i + X_1 + \dots + X_{n-1}}{X}, \ \frac{X_i + X_2 + \dots + X_n}{X}, \dots \binom{N-1}{n-1} \text{ terms}$$

inverse of the arithmetic mean of these $\binom{N-1}{n-1}$ quantities

$$= \left\{ \frac{\binom{N-1}{n-1} X_i + \binom{N-2}{n-2}(X - X_i)}{\binom{N-1}{n-1} X} \right\}^{-1}$$

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

$$= \frac{1}{\pi_i'}$$

$$> 1 .$$

$\therefore \quad T_i - 1 > 0$ (4.5.3)

and from (4.4.4) follows that

$\lambda_i X^2 > 0$ which implies that

$\lambda_i > 0$ for all i.

Theorem 4.5.2: A sufficient condition for $\hat{V}(\hat{R})$ to be positive is that $\lambda_{ij} \geq 0$ (or equivalently $T_{ij} \geq 1$) for all i, j.

Proof: By Lemma 4.5.2, since $\lambda_i > 0$ for all i, follows that a sufficient condition for $\hat{V}(\hat{R})$ to be positive is that $\lambda'_{ij} \geq 0$ or equivalently $T_{ij} \geq 1$ (follows from (4.4.4) and (4.4.7)).

Theorem 4.5.3. (Des Raj [1955] and Sen [1955]): An estimator of $V(\hat{R})$ is given by

$$\hat{V}_D(\hat{R}) = \hat{R}^2 - \frac{\sum_{i\in s} y_i^2 + 2\frac{N-1}{n-1}\sum_{\substack{i>j\\ \in s}} y_i y_j}{N\,n\,\bar{\bar{x}}\,\bar{\bar{X}}}.$$

Proof: $E\hat{V}_D(\hat{R}) = E(\hat{R}^2) - E\widehat{R^2} = V(\hat{R})$ and hence etc.

Remark (a). Both $\hat{V}_D(\hat{R})$ and $\hat{V}(\hat{R})$ take sometimes negative values and since the variance function is always positive it is desirable to have an estimator of the variance which is always positive definite. An attempt in this direction, in the context of the present estimator was due to Roy Choudhury [1956] and for the sake of completeness we briefly review his work.

Theorem 4.5.4. (Roy Chaudhury [1956]):

$\hat{V}_R(\hat{R}) = \frac{1}{n\,\bar{x}^2}\left\{m_2(y) - \frac{2\,\bar{y}}{\bar{x}}\,m_{11}(x,y) + \left(\frac{\bar{y}}{\bar{x}}\right)^2 m_2(x)\right\}$ is a positive definite biased estimate of the variance of $\hat{R}$, where

$$m_2(y) = \frac{N-n}{N-1}\frac{1}{n}\sum_1^n (y_i - \bar{y})^2, \quad m_{11}(x,y) = \frac{N-n}{N-1}\frac{1}{n}\sum_1^n (x_i - \bar{x})(y_i - \bar{y})$$

and

$$m_2(x) = \frac{N-n}{N-1}\;\frac{1}{n}\;\sum_1^n (x_i - \bar{x})^2.$$

Proof: From the finite bivariate population, draw n units with the probability of selecting the i^{th} unit in the first draw being P_i, $i = 1,2,\ldots, N$; $P_i > 0$ and $\sum_{i=1}^{N} P_i = 1$ and draw the remaining units with equal probability without replacement. For the two-dimensional samples (x_i, y_i) $i = 1,2,\ldots, n$ we have

$$E\,(\hat{V}_R\,(\hat{R})) = V\,(\frac{\bar{y}}{\bar{x}}) + 0\,(n^{-3/2})$$

and

$$V(\hat{V}_R\,(\hat{R})) = 0(n^{-3})$$

Since $\quad \hat{V}_R\,(\hat{R}) = \dfrac{G}{n\;\bar{X}^2}$,

where

$$G = m_2(y) - \frac{2\,\bar{y}}{\bar{x}}\,m_{11}(x,y) + (\frac{\bar{y}}{\bar{x}})^2\,m_2(x)$$

(i) possesses continuous derivatives of first and second order and (ii) G and its derivatives of first and second order are bounded throughout the range of variation of (x_i, y_i), $i = 1, 2,\ldots, n$.

For further details of the proof we refer to Roy Chaudhury [1956].

Remark (b): $\hat{V}(\hat{R})$ gives an unbiased estimate of $V(\hat{R})$ but $\hat{V}_R(\hat{R})$ gives a biased estimate, but still it has the advantage of being positive definite always.

4.6 Some properties of the coefficients T_i and T_{ij}:

We prove below some interesting properties of the coefficients T_i and T_{ij} which occur in the variance expression (vide corollary 4.4.1 and definitions (4.4.6) and (4.4.7)).

Theorem 4.6.1: $\sum_{j \neq i}^{N} T_{ij} = (n-1)\, T_i$.

Proof:

$$\sum_{j \neq i}^{N} T_{ij} = \sum_{j \neq i}^{N} \frac{X}{\binom{N-1}{n-1}} \sum_{g' \in G'} \frac{1}{X_i + X_j + X_{g'}^{ij}}$$

$$= \frac{X}{\binom{N-1}{n-1}} \left\{ \sum_{g' \in G'} \frac{1}{X_i + X_1 + X_{g'}^{i1}} + \sum_{g' \in G'} \frac{1}{X_i + X_2 + X_{g'}^{i2}} + \cdots \right.$$

$$\left. + \sum_{g' \in G'} \frac{1}{X_i + X_N + X_{g'}^{iN}} \right\},$$

each summation being taken over the $\binom{N-2}{n-2}$ sets.

Now, it can be verified that by proper rearrangement, similar terms can be collected and the r.h.s. takes the form

$$= \frac{X}{\binom{N-1}{n-1}} \sum_{g \in G} (n-1) \frac{1}{X_i + X_g^i} = (n-1)\, T_i \, .$$

Theorem 4.6.2: $\sum_{i=1}^{N} T_i X_i = \frac{NX}{n}$, where $X = \sum_{i=1}^{N} X_i$.

Proof: $\sum_{i=1}^{N} T_i X_i = \frac{X}{\binom{N-1}{n-1}} \sum_{i=1}^{N} \sum_{g \in G} \frac{X_i}{X_i + X_g^i}$

$$= \frac{X}{\binom{N-1}{n-1}} \left\{ \frac{X_1}{X_1 + (X_{g_1}^1)} + \frac{X_1}{X_1 + (X_{g_2}^1)} + \ldots \binom{N-1}{n-1} \text{ terms} \right.$$

$$+ \frac{X_2}{X_2 + (X_{g_1}^2)} + \frac{X_2}{X_2 + (X_{g_2}^2)} + \ldots \binom{N-1}{n-1} \text{ terms}$$

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

$$\left. + \frac{X_N}{X_N + (X_{g_1}^N)} + \frac{X_N}{X_N + (X_{g_2}^N)} + \ldots \binom{N-1}{n-1} \text{ terms} \right\} .$$

Properly combining the similar terms we have r.h.s.

$$= \frac{X}{\binom{N-1}{n-1}} \left(1 + 1 + 1 + \ldots \binom{N}{n} \text{ terms} \right) = \frac{NX}{n} \, .$$

Theorem 4.6.3:

$$\sum_{i=1}^{N} T_i X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij} X_i X_j = X^2 .$$

Proof: We can give a direct proof by verification as given in theorem 4.6.2. But a more elegant proof is as follows: We have from (4.4.6)

$$V(\frac{\bar{y}}{\bar{x}} X) = \sum_{i=1}^{N} T_i Y_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij} Y_i Y_j - Y^2.$$

Putting $y_i = x_i$, for all i, we have

$$0 = \sum_{i=1}^{N} T_i X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij} X_i X_j - X^2 .$$

$$\sum_{i=1}^{N} T_i X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij} X_i X_j = X^2 .$$

Theorem 4.6.4: $T_i \geq \frac{1}{\pi_i'}$, where $\pi_i' = \frac{n-1}{N-1} + \frac{N-n}{N-1} \frac{X_i}{X}$

Proof: See lemma 4.5.2.

Theorem 4.6.5: $T_{ij} \geq \frac{1}{\pi_{ij}'} (\frac{n-1}{N-1})^2$,

where $\pi_{ij}' = \frac{n-1}{N-1} \frac{n-2}{N-2} + \frac{n-1}{N-1} \frac{N-n}{N-2} (P_i + P_j)$.

Proof:

$$T_{ij} = \frac{X}{\binom{N-1}{n-1}} \sum_{g'} \frac{1}{X_i + X_j + X^{ij}_{g'}}$$

$$= \frac{1}{\frac{N-1}{n-1}\binom{N-2}{n-2}} \sum_{g'} \frac{1}{\frac{X_i + X_j + X^{ij}_{g'}}{X}} .$$

$\therefore$ $\frac{N-1}{n-1} T_{ij}$ = inverse of the harmonic mean of $\frac{X_i+X_j+X^{ij}_{g'}}{X}$, g' varying over $\binom{N-2}{n-2}$ quantities,

$\geq$ inverse of the arithmetic mean of $\frac{X_i+X_j+X^{ij}_{g'}}{X}$, g' varying over $\binom{N-2}{n-2}$ terms

$$= \left[\frac{\binom{N-2}{n-2}(X_i + X_j) + \binom{N-3}{n-3}(X - X_i - X_j)}{\binom{N-2}{n-2} X}\right]^{-1}$$

$$= [(P_i + P_j) + \frac{n-2}{N-2}(1 - P_i - P_j)]^{-1}$$

$$= [\frac{N-1}{n-1} \pi'_{ij}]^{-1} \qquad \text{from (4.5.2).}$$

$\therefore$ $T_{ij} \geq (\frac{n-1}{N-1})^2 \frac{1}{\pi'_{ij}}$.

orem 4.6.6: $T_i X_i \geq T_j X_j$ if and only if $X_i \geq X_j$.

Proof:

$$T_i X_i - T_j X_j = \frac{X}{\binom{N-1}{n-1}} \left\{ \sum_{g} \frac{X_i}{X_i + X_g^i} - \sum_{g} \frac{X_j}{X_j + X_g^i} \right\}$$

$$= \frac{X}{\binom{N-1}{n-1}} \left\{ \sum_{\substack{g \\ g \ni j}} \frac{X_i}{X_i + X_g^i} + \sum_{g} \frac{X_i}{X_i + X_g^{ij}} - \sum_{\substack{g \\ g \ni i}} \frac{X_j}{X_j + X_g^j} - \sum_{g} \frac{X_j}{X_j + X_g^{ji}} \right\}$$

$$= \frac{X}{\binom{N-1}{n-1}} \left\{ (X_i - X_j) \Big(\sum_{\substack{g \\ g \ni j}} \frac{1}{X_i + X_g^i} \Big) + \sum_{g} \Big(\frac{1}{1 + \frac{X_g^{ij}}{X_i}} - \frac{1}{1 + \frac{X_g^{ji}}{X_j}} \Big) \right\} \quad (4.6.1)$$

≥ 0 if $X_i \geq X_j$.

$T_i X_i - T_j X_j \geq 0$ and let if possible $X_i < X_j$. Then first and second terms of (4.6.1) are negative, thereby ding to a contradiction to our assumption that $X_i - T_j X_j \geq 0$ and so it follows that $X_i \geq X_j$. This pletes the proof of the theorem.

. An exact expression for the Bias and Mean Square Error the ratio estimator:

We have already seen in 3.2 that the ratio estimator $\frac{\hat{Y}}{\hat{X}}$, where $\hat{Y}$ and $\hat{X}$ are unbiased estimates of Y and respectively, for the estimation of the population ratio $\frac{Y}{X}$, is biased and that approximate expressions for the

Bias and Mean Square Error of the estimator given by Hansen, Hurwitz and Madow [1953], Cochran [1953], Sukhatme [1954], Murthy and Nanjamma [1959], and others use certain assumptions which are mostly valid for large samples. A different derivation is due to Koop [1951] who gets an expression for the bias of the ratio estimator to the order of n^{-4}. In this section, we use the knowledge of auxiliary information and construct T_i and T_{ij} coefficients of higher order and show that they are quite useful in getting exact expressions for the Bias and Mean Square Error, which are given in a compact form.

Analagous to (4.4.7) we define

$$T_i^{(\alpha)} = \frac{X}{\binom{N-1}{n-1}} \sum_{g \in G} \frac{1}{(X_i + X_g^i)^\alpha}$$

and

$$T_{ij}^{(\alpha)} = \frac{X}{\binom{N-1}{n-1}} \sum_{g' \in G'} \frac{1}{(X_i + X_j + X_{g'}^{ij})^\alpha} \qquad (4.7.1)$$

The explanation of the symbols is already given before. Also we denote $T_i^{(1)}$ and $T_{ij}^{(1)}$ by T_i and T_{ij} of (4.4.7) respectively.

<u>Theorem 4.7.1:</u> For a Simple Random Sample of size n taken from a population of size N, Without Replacement,

$$B\left(\frac{\bar{y}}{\bar{x}}\right) = \frac{n}{N\bar{X}}\left(\sum_{i=1}^{N} T_i Y_i\right) - R$$

and

$$M.S.E.\left(\frac{\bar{y}}{\bar{x}}\right) = \frac{n}{N\bar{X}}\left(\sum_{i=1}^{N} T_i^{(2)} Y_i^2 + \sum_{i \neq j}\sum^{N}\sum^{N} T_{ij}^{(2)} Y_i Y_j\right) - 2R\frac{n}{N\bar{X}}\sum_{i=1}^{N} T_i Y_i + R^2.$$

Proof:

$$B\left(\frac{\bar{y}}{\bar{x}}\right) = E\left(\frac{\bar{y}}{\bar{x}}\right) - R$$

$$= E\left(\frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i}\right) - R$$

$$= \sum_{s\in S}\left(\frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i}\right) P_s - R$$

$$= \binom{N}{n}^{-1} \sum_{s\in S}\left(\frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i}\right) - R$$

$$= \binom{N}{n}^{-1} \sum_{i=1}^{N} Y_i \left(\sum_{g\in G} \frac{1}{X_i + X_g^i}\right) - R$$

$$= \frac{n}{N\bar{X}} \sum_{i=1}^{N} T_i Y_i - R. \qquad (4.7.2)$$

$$M.S.E.\left(\frac{\bar{y}}{\bar{x}}\right) = E\left(\frac{\bar{y}}{\bar{x}} - R\right)^2$$

$$= \sum_{s\in S}\left(\frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i} - R\right)^2 P_s$$

$$= \binom{N}{n}^{-1} \left\{ \sum_{s\in S} \frac{(\sum_{i\in s} y_i)^2}{(\sum_{i\in s} x_i)^2} - 2R \sum_{s\in S} \left(\frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i}\right) + \sum_{s\in S} R^2 \right\}$$

$$= \binom{N}{n}^{-1} \left\{ \sum_{1}^{N} Y_i^2 \left(\sum_{g\in G} \frac{1}{(X_i + X_g^i)^2} \right) + \sum_{i\neq j}^{N} \sum^{N} Y_i Y_j \left(\sum_{g'\in G'} \frac{1}{(X_i + X_j + X_{g'}^{ij})^2} \right) \right.$$

$$\left. - 2R \sum_{i=1}^{N} Y_i \left(\sum_{g\in G} \frac{1}{(X_i + X_g^i)} \right) + \binom{N}{n} R^2 \right\}$$

$$= \frac{n}{NX} \left(\sum_{i=1}^{N} T_i^{(2)} Y_i^2 + \sum_{i\neq j}^{N} \sum^{N} T_{ij}^{(2)} Y_i Y_j \right) - 2R \frac{n}{NX} \sum_{i=1}^{N} T_i Y_i + R^2. \tag{4.7.3}$$

We mention below some of the directly useful properties of the $T^{(2)}$ coefficients.

__Theorem 4.7.2:__ $\sum_{j\neq i}^{N} T_{ij}^{(2)} = (n-1) T_i^{(2)}$.

__Proof:__ $$\sum_{j\neq i}^{N} T_{ij}^{(2)} = \sum_{j\neq i}^{N} \frac{X}{\binom{N-1}{n-1}} \sum_{g'\in G'} \frac{1}{(X_i + X_j + X_{g'}^{ij})^2}$$

$$= \frac{X}{\binom{N-1}{n-1}} \left\{ \sum_{g'\in G'} \frac{1}{(X_i + X_1 + X_{g'}^{i1})^2} + \sum_{g'\in G'} \frac{1}{(X_i + X_2 + X_{g'}^{i2})^2} + \right.$$

$$\left. \dots + \sum_{g'\in G'} \frac{1}{(X_i + X_N + X_{g'}^{iN})^2} \right\}$$

each summation being taken over $\binom{N-2}{n-2}$ sets.

By proper rearrangement and by collecting the similar terms, we have the r.h.s.

$$= \frac{X}{\binom{N-1}{n-1}} \sum_{s \in G} (n-1) \frac{1}{(X_i + X_s^i)^2}$$

$$= (n - 1)\, T_i^{(2)} .$$

Corollary 4.7.2.a: One can similarly prove that

$$\sum_{j \neq i}^{N} T_{ij}^{(\alpha)} = (n-1)\, T_i^{(\alpha)}.$$

Theorem 4.7.3:

$$\sum_{i=1}^{N} T_i^{(2)} X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij}^{(2)} X_i X_j = \frac{NX}{n} .$$

Proof: As in theorem 4.7.2 we can give direct proof by verification, but we give below an elegant and simple proof.

We have from (4.7.3)

$$M.S.E.(\frac{\bar{Y}}{\bar{x}}) = \frac{n}{NX}\left(\sum_{i=1}^{N} T_i^{(2)} Y_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij}^{(2)} Y_i Y_j\right) - 2R\, \frac{n}{NX} \sum_{i=1}^{N} T_i Y_i + R^2.$$

Putting $y_i = x_i$ for all i, we get

$$= \frac{n}{NX}\left(\sum_{i=1}^{N} T_i^{(2)} X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij}^{(2)} X_i X_j\right) - 2\frac{n}{NX} \sum_{i=1}^{N} T_i Y_i + 1,$$

$$\frac{n}{NX}\left(\sum_{i=1}^{N} T_i^{(2)} X_i^2 + \sum_{i \neq j}^{N} \sum^{N} T_{ij}^{(2)} X_i X_j\right) = 1, \text{ using theorem 4.6.2.}$$

$$\sum_{i=1}^{N} T_i^{(2)} X_i^2 + \sum_{i\neq j}^{N}\sum^{N} T_{ij}^{(2)} X_i X_j = \frac{N\bar{X}}{n} .$$

<u>Remark.</u> An alternate form of the Bias and Mean Square Error of the ratio estimator using theorems 4.6.2 and 4.7.3, is given by

$$B\left(\frac{\bar{y}}{\bar{x}}\right) = \frac{\sum_{i=1}^{N} T_i Y_i}{\sum_{i=1}^{N} T_i X_i} - R \qquad (4.7.4)$$

and

$$\text{M.S.E.}\left(\frac{\bar{y}}{\bar{x}}\right) = \frac{\sum_{i=1}^{N} T_i^{(2)} Y_i^2 + \sum_{i\neq j}^{N}\sum^{N} T_{ij}^{(2)} Y_i Y_j}{\sum_{i=1}^{N} T_i^{(2)} X_i^2 + \sum_{i\neq j}^{N}\sum^{N} T_{ij}^{(2)} X_i X_j} - 2R\,\frac{\sum_{i=1}^{N} T_i Y_i}{\sum_{i=1}^{N} T_i X_i} + R^2 \qquad (4.7.5)$$

4.8 Rôle of the coefficients T_i and T_{ij}:

The expression for the variance of the ratio-type estimator $\hat{Y}_R = \frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i} X$ under the Midzuno-Sen sampling scheme takes the familiar form $\sum_{1}^{N} T_i Y_i^2 + \sum_{i\neq j}^{N}\sum^{N} T_{ij} Y_i Y_j - Y^2$.

This has resemblance to the variance expression of the Horvitz-Thompson [1952] estimator $\hat{Y}_{HT} = \sum_{i\in s} \frac{y_i}{\pi_i}$, which is given by

$$\sum_{i=1}^{N} \frac{1}{\pi_i} Y_i^2 + \sum_{i \neq j}^{N} \sum^{N} \frac{\pi_{ij}}{\pi_i \pi_j} \cdot Y_i Y_j - Y^2,$$

where π_i and π_{ij} are the inclusion probability of the i^{th} unit and the joint inclusion probability of the units i and j in the sample respectively, for any design D. This enables us to make a comparison between these two estimators and this will be taken up in the next chapter. It is also interesting to observe that T_i and $\frac{1}{\pi_i}$ are respectively the inverses of the harmonic mean and the arithmetic mean of the same quantities which gives us the relation $T_i > \frac{1}{\pi_i}$ for all i, π_i, the probability of inclusion of the i^{th} unit in the sample for the Midzuno-Sen sampling scheme being equal to $\frac{n-1}{N-1} + \frac{N-n}{N-1} P_i$; while with $\pi_i = nP_i$, we have T_i and π_i connected by the relation $\sum_{i=1}^{N} T_i X_i = \sum_{i=1}^{N} \frac{1}{\pi_i} X_i = \frac{N X}{n}$. Another interesting observation is that $\sum_{j \neq i}^{N} T_{ij} = (n-1) T_i$ which reminds us of the result $\sum_{j \neq i} \pi_{ij} = (n-1) \pi_i$ for a fixed sample size design. We have also seen that the higher order coefficients $T_i^{(2)}$ and $T_{ij}^{(2)}$, which are straight-forward extensions of T_i and T_{ij}, help us in deriving exact expressions for the Bias and Mean Square Error of the ratio estimator for SRSWOR designs.

4.9 Numerical examples:

We illustrate the results obtained in the above sections by considering 5 populations of which Yates' population is one of them. Data for the other four live examples consists of the district-wise population figures of four of the states of India, viz., Kerala, Andhra Pradesh, Gujarat and West Bengal, the first two representing the Southern part and the latter the Northern. The population figures as per 1961 Census of India are taken as the y- variable (study variable) and those as per 1951 Census (rounded off to thousands) are taken as the x -variable (auxiliary information). Yates' population is used to check up the programs whenever the computations are organised on the electronic computer IBM 1401. Owing to the repeated use we present below the data for these four states.

1. Yates Population

Serial No.	X_i	Y_i
1	1	5
2	2	12
3	3	21
4	4	32
	10	70

2. Kerala

Sl. No.	Name of the District	X_i	Y_i
1	Cannanore	1375	1,780294
2	Kozhikode	2065	2,617,189
3	Palghat	1565	1,776,566
4	Trichur	1363	1,639,862
5	Ernakulam	1530	1,859,913
6	Kottayam	1328	1,732,880
7	Alleppey	1521	1,811,252
8	Quilon	1474	1,941,228
9	Trivandrum	1328	1,744,531
		13549	16,903,715

3. Andhra Pradesh

Sl. No.	Name of the district	X_i	Y_i
1	Srikakulam	2123	2,340,878
2	Visakhapatnam	2073	2,290,759
3	East Godavari	2302	2,608,375
4	West Godavari	1698	1,978,257
5	Krishna	1736	2,076,956
6	Guntur	2560	3,009,900
7	Nellore	1795	2,033,679
8	Chittor	1666	1,914,639
9	Cuddapah	1163	1,342,015

contd.

3. Andhra Pradesh (contd.)

Sl. No.	Name of the district	X_i	Y_i
10	Anantapur	1464	1,767,464
11	Kurnool	1617	1,908,740
12	Mahbubnagar	1447	1,590,686
13	Hyderabad	1822	2,062,995
14	Medak	1110	1,227,361
15	Nizamabad	835	1,022,013
16	Adilabad	832	1,009,292
17	Karimnagar	1428	1,621,515
18	Warangal	1330	1,545,435
19	Khammam	808	1,057,542
20	Nalgonda	1287	1,574,946
		31,116	35,983,447

4. Gujarat

Sl. No.	Name of the district	X_i	Y_i
1	Jamnagar	617	828,419
2	Rajkot	930	1,208,519
3	Surendranagar	506	663,206
4	Bhavnagar	886	1,119,435
5	Amreli	539	667,823
6	Junagadh	988	1,245,643

contd.

4. Gujarat (contd.)

Sl. No.	Name of the district	X_i	Y_i
7	Kutch	568	696,440
8	Banaskantha	774	996,144
9	Sabarkantha	684	918,587
10	Mehsana	1394	1,689,963
11	Ahmedabad	1676	2,210,199
12	Kaira	1612	1,977,540
13	Panchmahals	1131	1,468,946
14	Baroda	1212	1,527,326
15	Broach	718	891,969
16	Surat	1982	2,451,624
17	Dangs	47	71,567
		16264	20,633,350

5. West Bengal

Sl. No.	Name of the district	X_i	Y_i
1	Darjeeling	460	624,640
2	Jalpaiguri	915	1,359,292
3	Cooch Behar	671	1,019,806
4	West Dinajpur	979	1,323,797
5	Malda	938	1,221,923
6	Murshidabad	1716	2,290,010
7	Nadia	1145	1,713,324
8	24-Parganas	4459	6,280,915
9	Calcutta	2698	2,927,289

contd.

5. West Bengal (contd.)

Sl. No.	Name of the district	X_i	Y_i
10	Howrah	1611	2,038,477
11	Hooghly	1604	2,231,418
12	Burdwan	2192	3,082,846
13	Birbhum	1067	1,446,158
14	Bankura	1319	1,664,513
15	Midnapur	3359	4,341,855
16	Purulia	1169	1,360,016
		26302	34,926,279

The following tables give the variance of the estimator of the total of y-variate under the Midzuno-Sen sampling scheme (viz., (4.4.6)) and the exact Mean Square Error of the ratio estimator of the total (i.e. M.S.E.($\frac{\bar{y}}{\bar{x}}$ X)). We also present the exact Bias and Mean Square Error of the ratio estimator $\hat{R}$ for SRSWOR scheme (viz.,(4.7.4) and (4.7.5)). Only summary tables are given here and the T_{ij}, $T_{ij}^{(2)}$, T_i, $T_i^{(2)}$ values for the 5 populations are given in the Appendix A1 through A 10 for the sample sizes n = 2 and 3.

TABLE 4.9.1. Sample size n = 2

Population	$V_{MS}(\hat{Y})$	$M_R(\hat{Y})$	$B(\frac{\bar{Y}}{\bar{X}})$	$M(\frac{\bar{Y}}{\bar{X}})$
Yates	$.36286 \times 10^2$	$.46424 \times 10^2$	-.15476	.464237
Kerala	$.29536 \times 10^{12}$	$.29991 \times 10^{12}$	.57250	1633.657475
Andhra Pradesh	$.82651 \times 10^{12}$	$.94894 \times 10^{12}$	3.21806	980.122812
Gujarat	$.22784 \times 10^{12}$	$.24931 \times 10^{12}$	3.78768	942.544291
West Bengal	$.48716 \times 10^{13}$	$.49090 \times 10^{13}$	5.24390	7095.992340

TABLE 4.9.2. Sample size n = 3

Population	$V_{MS}(\hat{Y})$	$M_R(\hat{Y})$	$B(\frac{\bar{Y}}{\bar{X}})$	$M(\frac{\bar{Y}}{\bar{X}})$
ates	$.12037 \times 10^2$	$.13908 \times 10^2$	-.04861	.139082
Kerala	$.16810 \times 10^{12}$	$.16981 \times 10^{12}$	.29221	925.028627
ndhra radesh	$.50695 \times 10^{12}$	$.54965 \times 10^{12}$	1.89103	567.743800
jarat	$.14664 \times 10^{12}$	$.15445 \times 10^{12}$	2.27980	583.909687
est ngal	$.32142 \times 10^{13}$	$.32701 \times 10^{13}$	2.37740	4727.199569

CHAPTER V

πPS SAMPLING SCHEMES AND RATIO METHOD OF ESTIMATION

5.0 Summary

Three sampling strategies are considered: (i) $H(M, \hat{T}_1)$ consisting of the Midzuno [1952] - Sen [1952] sampling scheme and the estimator $\hat{T}_1 = (\sum_{i \in s} y_i \mid \sum_{i \in s} x_i) X$ of the population total Y, where the symbol $\sum_{i \in s}$ indicates that the summation is over all the units U_i contained in the sample s; (ii) $H(M, \hat{Y}_{HT})$ consisting of the Midzuno-Sen scheme and the estimator $\hat{Y}_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i}$, where π_i is the probability of inclusion of the i^{th} unit in the sample, (iii) $H(\pi PS, \hat{T}_2)$ consisting of the πPS sampling scheme (Hanurav [1965],[1967]) and the estimator $\hat{T}_2 = \frac{X}{n} \sum_{i \in s} \frac{y_i}{x_i}$ of Y. It is shown in this chapter that the strategy $H(\pi PS, \hat{T}_2)$ is suitable for the method of ratio estimation, under a general super-population model. A direct application to cluster sampling is also given. A brief resume on πPS sampling schemes is given and a general problem of choice between πPS schemes is presented towards the end. Illustrative examples are added in the last section.

5.1 Introduction:

It is already observed that the method of ratio estimation is used for estimating the population total Y of a characteristic $\mathcal{Y}$, when we have auxiliary information on a characteristic $\mathcal{X}$ related to it. It consists of getting an estimator of the population ratio $\frac{Y}{X}$ and multiplying this estimator by the known population total X. We have seen earlier that two estimators are usually suggested for estimating the population ratio, viz., $\sum_{i\in s} y_i / \sum_{i\in s} x_i$ and $\frac{1}{n}\sum_{i\in s} y_i/x_i$ and these at once give the two estimators of the population total

$$\hat{T}_1 = \frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i} X \qquad (5.1.1)$$

and

$$\hat{T}_2 = \frac{X}{n}\sum_{i\in s}\frac{y_i}{x_i}. \qquad (5.1.2)$$

It is clear that both (5.1.1) and (5.1.2) are biased and from (4.1) we see that $\hat{T}_1$ will be unbiased if the probability of selecting the sample is given by (4.1.1), i.e. $p_s = \sum_{i\in s} x_i \Big/ \binom{N-1}{n-1} X$, which can be easily achieved by using the Midzuno-Sen sampling scheme, described in 4.3. Following Hájek [1958], and Hanurav [(1965, 1967)] we denote the

Midzuno-Sen sampling scheme and the estimator $\hat{T}_1$ of Y (given by (5.1.1)) by the strategy $H(M, \hat{T}_1)$.

We consider next the estimator $\hat{T}_2$ of Y given by (5.1.2). This can be thrown into the form $\sum_{i \in s} y_i \Big/ \left(\frac{nX_i}{X}\right)$ and when we use a $\pi P S$ (π_i, the probability of inclusion of the i^{th} unit, Proportional to Size) sampling scheme (Vide Hanurav [1965], [1967]), the denominator of this estimator is equal to π_i, the probability of inclusion of the i^{th} unit in the sample ($\mathcal{X}$ being the measure of size). The estimator therefore reduces to $\sum_{i \in s} \frac{y_i}{\pi_i}$, the Horvitz-Thompson [1952] estimator, where $\pi_i = np_i$ ($p_i = \frac{X_i}{X}$ assumed less than n^{-1}). We denote the $\pi P S$ design and the estimator $\hat{T}_2$ of Y by the strategy $H(\pi P S, \hat{T}_2)$.

Let π_i' be the probability of inclusion of the i^{th} unit in the sample for the Midzuno-Sen sampling scheme. Then we have, at once, another choice of an estimator,

$$\hat{Y}_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i'} , \qquad (5.1.3)$$

which again is the corresponding Horvitz-Thompson estimator. The Midzuno-Sen sampling scheme and the Horvitz-Thompson estimator (5.1.3) together constitute the strategy $H(M, \hat{Y}_{HT})$. It can be observed that this is a non-$\pi P S$ strategy, since

π_i' is not proportional to P_i. In fact, π_i' is equal to $\frac{n-1}{N-1} + \frac{N-n}{N-1} P_i$ (refer (4.5.1)).

5.2. Comparison of the strategies:

In this section, we present a discussion on the choice of a suitable strategy under certain super-population set-up. This concept taken from Bayesian inference was introduced by Cochran [1946]. When we take uniform minimization of the variance of the strategy as the criterion of optimality there does not exist such a one. So under other criteria of optimality, Godambe [1955] and Hájek [1959] have shown the existence of an optimum strategy in particular situations. Here, an a priori distribution ∂ over $\underset{\sim}{Y}$ is assumed and $\int V(H) d\partial$ is minimized for H varying over $\mathcal{H}(\nu)$ the class of all strategies with the same effective sample size. Such a H_0 minimizing $\int V(H) d\partial$, uniformly with respect to all the parameters of the distribution ∂ is called a ∂-optimum strategy in $\mathcal{H}(\nu)$.

Let Δ_g be the class of prior distributions ∂ satisfying

$$\left.\begin{array}{ll} \text{(i)} & \mathcal{E}_\partial (Y_i \mid X_i) = a\, X_i \\ \text{(ii)} & \mathcal{V}_\partial (Y_i \mid X_i) = \sigma^2 X_i^g , \; g > 1 \\ \text{and (iii)} & \mathcal{C}_\partial (Y_i, Y_j \mid X_i, X_j) = 0 \end{array}\right\} \qquad (5.2.1)$$

where $\mathcal{C}$ denotes covariance.

In most of the situations met in practice, the parameter g is found to lie between 1 and 2.

We know that

$$V\left[\sum_{i\in s}\frac{y_i}{\pi_i}\right] = \sum_{i=1}^{N}\frac{Y_i^2}{\pi_i} + \sum_{i\neq j}^{N}\sum^{N}\frac{Y_i\;Y_j}{\pi_i\;\pi_j}\pi_{ij} - Y^2 , \qquad (5.2.2)$$

We also have from (4.4.6)

$$V[\hat{T}_1] = \sum_{i=1}^{N} T_i\,Y_i^2 + \sum_{i\neq j}^{N}\sum^{N} T_{ij}\,Y_iY_j - Y^2, \qquad (5.2.3)$$

where the coefficients T_i and T_{ij} in the r.h.s of (5.2.3) are as defined in (4.4.7).

Theorem 5.2.1: Under the model (5.2.1),

$$\mathcal{E}_{\vartheta}\,V(\hat{T}_1) = \sigma^2\sum_{1}^{N} X_i^g(T_i - 1).$$

Proof:

$$\mathcal{E}_{\vartheta}\,V(\hat{T}_1) = \sum_{i=1}^{N}(T_i-1)\,\mathcal{E}(Y_i^2 \mid X_i) + \sum_{i\neq j}^{N}\sum^{N}(T_{ij}-1)\,\mathcal{E}(Y_iY_j\,|X_i,X_j)$$

$$= \sum_{i=1}^{N}(T_i-1)(a^2X_i^2 + \sigma^2X_i^g) + \sum_{i\neq j}^{N}\sum^{N}(T_{ij}-1)a^2X_iX_j$$

$$= \sigma^2\sum_{i=1}^{N} X_i^g(T_i-1) + a^2\left[\sum_{i=1}^{N}(T_i-1)X_i^2 + \sum_{i\neq j}^{N}\sum^{N}(T_{ij}-1)X_iX_j\right]$$

$$= \sigma^2\sum_{i=1}^{N} X_i^g\,(T_i-1), \text{ by Theorem 4.6.3.}$$

Theorem 5.2.2: Under the model (5.2.1)

$$\xi_\partial V(\sum_{i\in s} \frac{y_i}{\pi_i}) = \sigma^2 \sum_{i=1}^{N} X_i^g (\frac{1}{\pi_i} -1) + a^2 V(\sum_{i\in s} \frac{X_i}{\pi_i}),$$

where π_i is the probability of inclusion of the i^{th} unit in the sample, not proportional to X_i and when π_i is proportional to X_i, the expected variance equals $\sigma^2 \sum_{i=1}^{N} X_i^g(\frac{1}{nP_i} -1)$, where $P_i = \frac{X_i}{X}$.

Proof:

$$\xi_\partial V(\sum_{i\in s} \frac{y_i}{\pi_i}) = \sum_{i=1}^{N} (\frac{1}{\pi_i} -1)\xi(Y_i^2|X_i) + \sum_{i\neq j}^{N}\sum^{N} (\frac{\pi_{ij}}{\pi_i\pi_j} -1)\xi(Y_iY_j|X_i,X_j)$$

$$= \sum_{i=1}^{N} (\frac{1}{\pi_i} -1)(\sigma^2 X_i^g + a^2X_i^2) + \sum_{i\neq j}^{N}\sum^{N} (\frac{\pi_{ij}}{\pi_i\pi_j} -1)a^2X_iX_j$$

$$= \sigma^2 \sum_{i=1}^{N} X_i^g(\frac{1}{\pi_i}-1) + a^2 \left\{ \sum_{i=1}^{N} (\frac{1}{\pi_i} -1)X_i^2 + \sum_{i\neq j}^{N}\sum^{N}(\frac{\pi_{ij}}{\pi_i\pi_j} -1)X_iX_j \right\}$$

$$= \sigma^2 \sum_{i=1}^{N} X_i^g(\frac{1}{\pi_i} -1) + a^2 V(\sum_{i\in s} \frac{X_i}{\pi_i}).$$

When $\pi_i = nP_i$, then $\sum_{i\in s} \frac{X_i}{\pi_i}$ is a constant and $V(\sum_{i\in s} \frac{X_i}{\pi_i})$ equals zero. Thus for a πPS scheme

$$\xi_\partial V(\sum_{i\in s} \frac{y_i}{\pi_i}) = \sigma^2 \sum_{i=1}^{N} X_i^g (\frac{1}{nP_i} - 1).$$

Thus we have

$$\mathcal{E}_{\partial}\, V(\hat{Y}_{HT}) = \sigma^2 \sum_{i=1}^{N} X_i^g\left(\frac{1}{\pi_i'} - 1\right) + n^2\, V\left(\sum_{i\in s} \frac{X_i}{\pi_i'}\right),$$

where $\hat{Y}_{HT}$ is as defined in (5.1.3)

and

$$\mathcal{E}_{\partial}\, V(\hat{T}_2) = \sigma^2 \sum_{i=1}^{N} X_i^g\left(\frac{1}{\pi_i} - 1\right),$$

where $\pi_i = \frac{nX_i}{X}$.

Now we have the following theorem which compares the strategies $H(\pi PS, \hat{T}_2)$ and $H(M, \hat{T}_1)$.

<u>Theorem 5.2.3:</u> The sampling strategy $H(\pi PS, \hat{T}_2)$ is superior to the strategy $H(M, \hat{T}_1)$ in the Δ_g-sense.

<u>Proof:</u>

$$\mathcal{E}_{\partial}\left[\frac{V(\hat{T}_1) - V(\hat{T}_2)}{\sigma^2}\right] = \sum_{i=1}^{N} X_i^g(T_i - 1) - \sum_{i=1}^{N} X_i^g\left(\frac{1}{\frac{n\,X_i}{X}} - 1\right)$$

$$= \sum_{i=1}^{N} X_i^g(T_i - 1) - \sum_{i=1}^{N}\left(\frac{X_i^{g-1}\, X}{n} - X_i^g\right)$$

$$= \sum_{i=1}^{N} X_i^{g-1}\left(T_i X_i - \frac{\bar{X}}{n}\right)$$

$$= \sum_{i=1}^{N} X_i^{g-1}\left(T_i X_i - \frac{\sum_{i=1}^{N} T_i X_i}{N}\right), \text{ by Theorem 4.6.2.}$$

$$= N \text{ cov. } (T_i X_i,\, X_i^{g-1})$$

$$> 0, \quad \text{for } g > 1, \text{ since } T_i X_i \geq T_j X_j$$

if and only if $X_i \geq X_j$ (by Theorem 4.6.6).

This establishes the superiority of $H(\pi PS, \hat{T}_2)$ over $H(M, \hat{T}_1)$ in the Δ_g-sense.

We now present a comparison between the two strategies $H(\pi PS, \hat{T}_2)$ and $H(M, \hat{Y}_{HT})$ which are in fact based on a πPS sampling scheme and a non-πPS sampling scheme respectively. In the latter scheme we have $\pi_i' = \alpha + \beta P_i$, where $P_i = X_i / X$ and $\alpha = \frac{n-1}{N-1}$, $\beta = \frac{N-n}{N-1}$ (Refer (4.5.1)). Using a πPS scheme, the probability of inclusion of the i^{th} unit in the sample is $\pi_i = nP_i$.

We first prove a lemma.

<u>Lemma 5.2.1:</u>

$$\frac{P_i^{g-1}}{P_i + \theta} \text{ is an increasing function of } P_i \text{ if}$$

$P_i < \frac{\theta(g-1)}{2-g}$, where $0 < P_i < 1$, $1 < g < 2$ and θ is positive.

<u>Proof:</u>

$$\frac{\partial}{\partial P_i}\left(\frac{P_i^{g-1}}{P_i + \theta}\right) = \frac{(P_i + \theta)(g-1)P_i^{g-2} - P_i^{g-1}}{(P_i + \theta)^2}$$

$$= \frac{P_i^{g-2}}{(P_i + \theta)^2}\left\{(P_i + \theta)(g-1) - P_i\right\} > 0$$

$$\text{if } P_i(g-2) > \theta(1-g)$$

$$\text{or if } P_i < \frac{\theta(g-1)}{2-g}.$$

<u>Theorem 5.2.4:</u> The sampling strategy $H(\pi PS, \hat{T}_g)$ is superior to the strategy $H(M, \hat{Y}_{HT})$ in the Δ_g - sense if the sample size is greater than

$$n_o = \frac{(2g-3) + \sqrt{4(N-1)(g-1)(2-g) + 1}}{2(g-1)}, \text{ where } 1 < g < 2.$$

When $g = 2$, $H(\pi PS, \hat{T}_g)$ is always superior to $H(M, \hat{Y}_{HT})$.

<u>Proof:</u> We have

$$D = \mathcal{E}_\delta\left[\frac{V(\hat{Y}_{HT}) - V(\hat{T}_g)}{\sigma^2}\right] = \sum_{i=1}^{N} X_i^g\left(\frac{1}{\pi_i'} - 1\right) + \frac{a^2}{\sigma^2} V\left(\sum_{i \in s} \frac{X_i}{\pi_i'}\right)$$

$$- \sum_{i=1}^{N} X_i^g\left(\frac{1}{\pi_i} - 1\right)$$

$$= \sum_{i=1}^{N}\left(\frac{1}{\pi_i'} - \frac{1}{\pi_i}\right) X_i^g + \frac{a^2}{\sigma^2} V\left(\sum_{i \in s} \frac{X_i}{\pi_i'}\right)$$

$$> \sum_{i=1}^{N}\left(\frac{1}{\pi_i'} - \frac{1}{\pi_i}\right) X_i^g$$

$$= \frac{N^2}{n}\left(\frac{n}{N^2}\sum_{1}^{N}\frac{X_i^g}{\pi_i'} - \frac{n}{N^2}\sum_{1}^{N}\frac{X_i^g}{\pi_i}\right)$$

$$= \frac{N^2}{n}\left(\frac{1}{N^2}\sum_{1}^{N}\pi_i' \sum_{1}^{N}\frac{X_i^g}{\pi_i'} - \frac{\sum_{1}^{N} X_i^g}{N} - \frac{1}{N^2}\sum_{1}^{N}\pi_i \sum_{1}^{N}\frac{X_i^g}{\pi_i} + \frac{\sum_{1}^{N} X_i^g}{N}\right)$$

$$\left(\because \sum_{1}^{N}\pi_i' = \sum_{1}^{N}\pi_i = n\right)$$

$$= \frac{N^2}{n}\left\{\text{cov.}\left(\frac{X_i^g}{\pi_i}, \pi_i\right) - \text{cov.}\left(\frac{X_i^g}{\pi_i'}, \pi_i'\right)\right\}$$

$$= \frac{N^2}{n} \left\{ \text{cov.}\left(\frac{X_i^g}{nP_i}, nP_i\right) - \text{cov.}\left(\frac{X_i^g}{\alpha + \beta P_i}, \alpha + \beta P_i\right) \right\}$$

$$- \frac{N^2}{n} \left\{ \text{cov.}\left(\frac{X_i^g}{P_i}, P_i\right) - \text{cov.}\left(\frac{X_i^g}{\theta + P_i}, P_i\right) \right\}$$

where $\theta = \alpha / \beta$.

$$= \frac{N^2}{n} \left\{ \text{cov.}\left(P_i, X_i^g\left(\frac{1}{P_i} - \frac{1}{P_i + \theta}\right)\right) \right\}.$$

Consequently

$$D > \frac{N^2}{n} X^g \theta \;\; \text{cov.}\left(P_i, \frac{P_i^{g-1}}{P_i + \theta}\right). \qquad (5.2.4)$$

We distinguish two cases.

(i) $1 < g < 2$.

Since $n > n_o$, we have

$$n^2(g-1) - n(2g-3) - N(2-g) > 0, \text{ so that}$$

$$\frac{n\,\theta(g-1)}{2-g} > 1 > \pi_i = nP_i. \qquad (5.2.5)$$

Consequently,

$$P_i < \frac{\theta(g-1)}{2-g}.$$

By Lemma 5.2.1, it follows that $\frac{P_i^{g-1}}{P_i + \theta}$ is an increasing function of P_i and thus we have

$$\text{cov.}\left(P_i, \frac{P_i^{g-1}}{P_i + \theta}\right)$$

is positive.

(ii) $g = 2$.

In this case,

$$\text{cov.}(P_i, \frac{P_i^{g-1}}{P_i + Q}) = \text{cov.}(P_i, \frac{P_i}{P_i + Q}) > 0$$

and this completes the proof of the theorem.

Remark (a). The condition $n > n_o$ is sufficient for the covariance to be positive. (This condition is due to the fact that we are omitting some positive contribution in the derivation of (5.2.4). For $g = 1.5$ we find that $n_o = \sqrt{N}$ and for $1.5 < g < 1.9$ we have $n_o < \sqrt{N}$. This is not a serious restriction in practice. We present below a table giving the minimum sample size $[n_o] + 1$, given by the sufficient condition for the πPS strategy to be better than the non-πPS strategy considered, the range of values of g being 1.1(.1)1.9 and of N being 10(10) 100(20) 500(50) 1000.

From table 5.2.1 we find that although the restriction seems to be severe for the initial values of g, it is a mild one for the latter values of g. In practice, the values of g are found to lie above 1.3 in some cases (Smith [1938] and Mahalanobis [1944]). Moreover, we are using 1 as the upper bound for π_i in (5.2.5), but in fact $\pi_i << 1$, which means that the restriction on minimum sample size will be still milder.

TABLE 5.2.1

Minimum sample size required for the πPS strategy to be better than the non-πPS strategy

N / [illegible]	10	20	30	40	50	60	70	80	90	100
1.1	7	10	13	16	18	20	22	24	25	27
1.2	5	8	10	12	13	15	16	17	18	19
1.3	5	7	8	10	11	12	13	13	14	15
1.4	4	6	7	8	9	10	10	11	12	12
1.5	4	5	6	7	8	8	9	9	10	10
1.6	3	4	5	6	6	7	7	8	8	9
1.7	3	4	4	5	5	6	6	7	7	7
1.8	2	3	4	4	4	5	5	5	6	6
1.9	2	2	3	3	3	4	4	4	4	4

TABLE 5.2.1 (contd.)

g \ N	120	140	160	180	200	220	240	260	280	300
1.1	30	32	35	37	39	41	43	45	47	49
1.2	21	23	24	26	27	29	30	31	33	34
1.3	17	18	19	20	21	23	24	24	25	26
1.4	14	15	16	17	18	18	19	20	21	21
1.5	11	12	13	14	15	15	16	17	17	18
1.6	10	10	11	12	12	13	13	14	14	15
1.7	8	9	9	10	10	11	11	11	12	12
1.8	6	7	7	8	8	8	9	9	9	10
1.9	5	5	5	5	6	6	6	6	7	7

TABLE 5.2.1. (contd.)

N / g	320	340	360	380	400	420	440	460	480	500
1.1	50	52	54	55	57	58	60	61	62	64
1.2	35	36	37	38	39	40	41	42	43	44
1.3	27	28	29	30	30	31	32	33	33	34
1.4	22	23	23	24	25	25	26	27	27	28
1.5	18	19	19	20	20	21	21	22	22	23
1.6	15	16	16	17	17	17	18	18	19	19
1.7	13	13	13	14	14	14	15	15	15	15
1.8	10	10	10	11	11	11	11	12	12	12
1.9	7	7	7	7	8	8	8	8	8	8

TABLE 5.2.1 (contd.)

N / g	550	600	650	700	750	800	850	900	950	1000
1.1	67	70	73	76	79	81	84	87	89	91
1.2	46	48	50	52	54	56	57	59	61	62
1.3	36	37	39	40	42	43	44	46	47	48
1.4	29	30	31	33	34	35	36	37	38	39
1.5	24	25	26	27	28	29	30	30	31	32
1.6	20	21	21	22	23	24	24	25	26	26
1.7	16	17	17	18	19	19	20	20	21	21
1.8	13	13	14	14	15	15	15	16	16	17
1.9	9	9	9	10	10	10	11	11	11	11

Thus theorems 5.2.3 and 5.2.4 lead to the choice of a strategy suitable for ratio method of estimation, since the πPS strategy fares better than both the non-πPS strategy considered here (in most of the practical situations) and the Midzuno-Sen strategy.

<u>R mark (b)</u>: After obtaining (5.2.4) which gives

$$D > \frac{N^2}{n} X^g \theta \quad \text{cov.} \left(P_i, \frac{P_i^{g-1}}{P_i+\theta}\right)$$

we distinguished the two cases, $1 < g < 2$ and $g = 2$. It is clear that for $g = 2$, $\text{cov.}\left(P_i, \frac{P_i}{P_i+\theta}\right) > 0$ and hence that the strategy $H(\pi PS, \hat{T}_2)$ is superior to $H(M, \hat{Y}_{HT})$.

Let $1 < g < 2$.

If $g > g_0 = 1 + \frac{N - n}{n^2 - 2n + N}$,

then $g > g_0 = \frac{2(N-n) + n(n-1)}{(N-n) + n(n-1)}$

$$= \frac{2 + n\theta}{1 + n\theta}$$

i.e. $g(1 + n\theta) > 2 + n\theta$

or $\frac{n\theta(g-1)}{(2-g)} > 1 > \pi_i = nP_i$

and consequently, $P_i < \frac{\theta(g-1)}{2-g}$.

Now by Lemma 5.2.1, it follows that $\frac{P_i^{g-1}}{P_i+\theta}$ is an increasing function of P_i and that

$$\text{cov.}\left(P_i, \frac{P_i^{g-1}}{P_i + \theta}\right) \text{ is positive.}$$

Thus we have

Theorem 5.2.5: The sampling strategy $H(\pi PS, \hat{T}_2)$ is superior to the strategy $H(N, \hat{Y}_{HT})$ in the Δ_g-sense, for a fixed sample size n, if the super-population parameter

$$g > g_0 = 1 + \frac{N - n}{n^2 - 2n + N} .$$

It is easy to see that for $n = \sqrt{N}$, $g_0 = 1.5$. A table like the one above can be constructed. Theorem 5.2.5 has similarity with the theorem of Vijayan [1966] where he proves the existence of a g_0 such that for $g > g_0$ the Horvitz-Thompson strategy is better than the Symmetrized Des Raj strategy (Des Raj [1956], Murthy [1957]), while for $g < g_0$ it is the other way and for $g = g_0$ both are equally efficient. Theorem 5.2.4 is of more practical interest, while theorem 5.2.5 is of general theoretical interest.

5.3 Cluster sampling as viewed through ratio method of estimation:

Suppose we have a population, grouped in N clusters. If M_i is the size of the i^{th} cluster and $\bar{y}_i$ is the mean of the population characteristic $\mathcal{Y}$ in the i^{th} cluster, the population mean of $\mathcal{Y}$ is

$$\bar{Y} = \frac{\sum_{i=1}^{N} M_j \, \bar{y}_i}{\sum_{i=1}^{N} M_i} . \qquad (5.3.1)$$

If we want to estimate $\bar{Y}$ from a sample of n clusters, we have at our disposal the biased estimators

$$\hat{\bar{Y}}_1 = \frac{\sum_{i \in s} M_i \, \bar{y}_i}{\sum_{i \in s} M_i} \qquad (5.3.2)$$

and

$$\hat{\bar{Y}}_2 = \frac{1}{n} \sum_{i \in s} \bar{y}_i . \qquad (5.3.3)$$

We can use the methods of Chapter III and construct unbiased estimators by modifying (5.3.2) and (5.3.3) suitably. Corresponding to (3.3.4) we have an estimator of the type (almost unbiased)

$$\hat{\bar{Y}}_{HR} = \frac{n\left(\frac{\sum_1^n M_i \, \bar{y}_i}{\sum_1^n M_i}\right) - \frac{1}{n}\sum \bar{y}_i}{(n-1)}$$

$$= \frac{n \sum_1^n M_i \, \bar{y}_i - \sum_1^n \bar{y}_i \sum_1^n M_i/n}{(n-1) \sum_1^n M_i} . \qquad (5.3.4)$$

Corresponding to (3.3.5) we have a Hartley-Ross type unbiased estimator

$$\hat{\bar{Y}}_{HR} = \frac{1}{n}\sum_{1}^{n}\bar{y}_i + \frac{N-1}{N}\,\frac{1}{n-1}\,\frac{1}{\bar{M}}\left(\sum_{1}^{n} M_i\bar{y}_i - \frac{\sum_{1}^{n} M_i \sum_{1}^{n}\bar{y}_i}{n}\right), \quad (5.3.5)$$

where $\bar{M} = \sum_{1}^{N} M_i/N$.

Analagous to (3.3.6) we have an almost unbiased estimator

$$\hat{\bar{Y}}_P = \frac{\sum_{1}^{n} M_i\,\bar{y}_i}{\sum_{1}^{n} M_i} + \frac{N-1}{Nn\,(n-1)}\;\frac{1}{\bar{M}}\left(\sum_{1}^{n} M_i\,\bar{y}_i - \frac{\sum_{1}^{n} M_i \sum_{1}^{n}\bar{y}_i}{n}\right). \quad (5.3.6)$$

Though we can construct unbiased estimators of the above type, computation of error estimates would not be quite easy. So, by modifying the sampling schemes we can make (5.3.2) or (5.3.3) unbiased and we can utilize the theorems proved in 5.2 towards a search of a suitable strategy in this case.

Since we may write $\hat{\bar{Y}}_2$((5.3.3)) in the form $\hat{\bar{Y}}_2 = n^{-1}\sum_{i\in s} M_i\bar{y}_i/M_i$, we may define a sampling scheme, denoted by πPM, in which the probability that the i^{th} cluster is included in the sample is proportional to M_i, and this together with the estimator $\hat{\bar{Y}}_2$, may be called the strategy $H(\pi PM, \hat{\bar{Y}}_2)$. Taking cluster size as our auxiliary characteristic, this is clearly a πPS strategy in Hanurav's sense, and is therefore superior to the Midzuno-Sen scheme (with clusters as the sampling units) and $\hat{\bar{Y}}_1$ as the estimator.

5.4 A brief review on πPS sampling schemes:

We have seen in 5.1, 5.2 and 5.3 how πPS sampling schemes can be used for the choice of a suitable strategy in the ratio method of estimation and in cluster sampling. It is not just enough to have sampling schemes in which

(i) $\pi_i = \frac{n X_i}{X}$ and (ii) $n_s = n$ for all samples s with positive probability, but we should also have same optimality properties like (iii) $\pi_{ij} > 0$, for the estimation of the variance, (iv) $\pi_{ij} \leq \pi_i \pi_j$ for a positive definite estimate of the variance when the Yates and Grundy [1953] estimator of variance is used, (v) $\frac{\pi_{ij}}{\pi_i \pi_j} > c$ (not very near to zero) and (vi) π_{ij}'s should be computable from compact formulae. The attempts by a host of research workers such as Goodman and Kish [1950], Horvitz and Thompson [1952], Yates and Grundy [1953], Durbin [1953], Grundy [1954], Des Raj [1956], Hájek [1959], Hanurav [1962], Hartley and Rao [1962], Rao, Hartley and Cochran [1962], Felligi [1963], Stuart [1964] and others do not satisfy one or other of the above properties. For a full discussion on this we refer to Hanurav [1965,1967]. Hanurav [1965, 1967] gave a scheme for πPS sampling of two units from a stratum and later generalised (Hanurav [1966])his method. Vijayan [1967] gave another method of πPS sampling for the sample size n. Hanurav, under the assumption that $P_i \leq P_{i+1}$ (without loss of generality) and $P_{N-n+1} = P_N$ gives the selection probabilities at the t^{th} draw equal to

$$P_{it} = \frac{P_{i,t-1} - Q_{i,t-1}}{1 - \sum_i Q_{i,t-1}} ,$$

where $Q_{i,t-1}$ is defined recursively in terms of P_{ij}'s, and then selects a PPS sample of size n with probabilities P_{it} at the t^{th} step for the i^{th} unit. If all the units are distinct the sample is accepted, otherwise the sample is rejected and the process is repeated. This sequential sampling scheme terminates quite rapidly and has the above mentioned optimum properties. He gives a slight modification for the case $P_{N-n+1} \neq P_N$. Vijayan also assumes, without loss of generality, that $P_i \leq P_{i+1}$ and $P_{N-n+1} = P_N$. He selects the k^{th} unit with a specified probability, having selected $U_{i_1}, U_{i_2}, \ldots, U_{i_{k-1}}$. If $P_{N-n+1} < P_N$, he chooses a number from $1, 2, \ldots, n$ with certain probability and if k is selected he chooses a sample of size k from the first $N-n+k$ units with size measures properly revised and adjoins the last $(n-k)$ units. This scheme also satisfies the optimality properties.

Given a sample size n, it is possible to get a meaningful set of inclusion probabilities if and only if $n P_i < 1$, since $\pi_i < 1$ for all i, which led us to the assumption $P_i < n^{-1}$ (cf. 5.1). The condition expressed in the form

$\frac{n\,X_{max}}{X} < 1$ is mostly satisfied in the practical examples we consider. Koop [1966] expressed this condition in a slightly different way as $\frac{X_{max}}{X - X_{max}} < \frac{1}{(n-1)}$.

5.5 Choice between πPS strategies:

Considering the Horvitz-Thompson estimator for the estimation of the population total $Y = \sum_{1}^{N} Y_i$, given by $\hat{Y} = \sum_{i \in s} y_i/\pi_i$ (vide 5.1), we recall from (5.2.2) that

$$V\left(\sum_{i \in s} \frac{y_i}{\pi_i}\right) = \sum_{1}^{N}\left(\frac{1-\pi_i}{\pi_i}\right)Y_i^2 + \sum_{i \neq j}^{N}\sum^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}\right)Y_iY_j \quad (5.5.1)$$

Under the model Δ_g where

$$\left.\begin{array}{l} \text{i) } \mathcal{E}_\partial (Y_i \mid X_i) = aX_i \\ \text{ii) } \mathcal{V}_\partial (Y_i \mid X_i) = \sigma^2 X_i^g \\ \text{iii) } \mathcal{C}_\partial (Y_i, Y_j \mid X_i, X_j) = 0, \text{ where } \mathcal{C} \text{ denotes} \\ \qquad\qquad \text{the covariance,} \end{array}\right\} \quad (5.5.2)$$

we have from theorem 5.2.2 that

$$\mathcal{E}_{\partial(g)} V\left(\sum_{i \in s} \frac{y_i}{\pi_i}\right) = \sigma^2 \sum_{1}^{N} X_i^g\left(\frac{1}{\pi_i} - 1\right) + a^2 V\left(\sum_{i \in s} \frac{X_i}{\pi_i}\right). \quad (5.5.3)$$

Under the particular model Δ_2((5.5.2) with $g = 2$), Godambe [1955] has proved that there exists a Δ_2-optimum strategy for which

i) inclusion probability of a unit is proportional to the value taken by the auxiliary characteristic on that unit

ii) every sample has n distinct units and

iii) the estimator used being the corresponding Horvitz-Thompson estimator,

in the class of all unbiased strategies with n distinct units. Further $\mathcal{E} V(\sum_{i \in s} \frac{y_i}{\pi_i})$ attains the minimum value for this Δ_2-optimum strategy and the minimum value is equal to

$$\sigma^2 \sum_{1}^{N} \left(\frac{1}{n P_i} - 1 \right) X_i^2 , \qquad (5.5.4)$$

since the second term $V(\sum \frac{X_i}{\pi_i})$ vanishes. It is also seen that there do not exist Δ_g-optimum strategies for $g \neq 2$.

From (5.5.3), one can easily see that, the second term $a^2 V(\sum_{i \in s} \frac{X_i}{\pi_i})$ can be minimised when $\pi_i \propto X_i$ and infact it then takes the value zero. Similarly, considering the first term we can find out for what values of π_i it is minimised and in this connection we prove the following theorem.

<u>Theorem 5.5.1:</u> $\sigma^2 \sum_{1}^{N} \left(\frac{1-\pi_i}{\pi_i}\right) X_i^g$ is minimized if $\pi_i \propto X_i^{g/2}$

<u>Proof</u>: We have to minimize $\sigma^2 \sum_{1}^{N} \left(\frac{1}{\pi_i} - 1\right) X_i^g$ subject to the condition that $\sum_{i=1}^{N} \pi_i = n$.

Introducing Lagrange's multiplier λ, we minimize

$$\sigma^2 \left(\sum_1^N \left(\frac{1}{\pi_i} - 1\right) X_i^g\right) + \lambda\left(\sum_1^N \pi_i - n\right)$$

which on minimization gives

$$-\frac{\sigma^2 X_i^g}{\pi_i^2} + \lambda = 0$$

or

$$\pi_i = \frac{\sigma X_i^{g/2}}{\sqrt{\lambda}} .$$

But $\sum_1^N \pi_i = \frac{\sigma \sum X_i^{g/2}}{\sqrt{\lambda}} = n,$

so that $\frac{1}{\sqrt{\lambda}} = \frac{n}{\sigma \sum_1^N X_i^{g/2}}$ and consequently,

$$\pi_i = \frac{n X_i^{g/2}}{\sum_1^N X_i^{g/2}} \tag{5.5.5}$$

or $\pi_i \propto X_i^{g/2}$ and hence the theorem.

Thus we see from theorem 5.5.1 that when π_i's are taken proportional to $X_i^{g/2}$ we minimize the first term of (5.5.3) and on the other hand when π_i's are proportional to X_i's the second term of (5.5.3) is minimized and this leads to the question of choice between various πPS sampling schemes.

From (5.5.3) we have (theorem 5.2.2)

$$\mathcal{E}_{\partial(g)} \; V\left(\Sigma \frac{y_i}{\pi_i}\right) = \sigma^2 \sum_{1}^{N} \left(\frac{1}{\pi_i} - 1\right) X_i^g$$

when $\pi_i = nP_i$

or $$\mathcal{E}_{\partial(g)} \; V\left(\Sigma \frac{y_i}{n P_i}\right) = \sigma^2 \sum_{1}^{N} \left(\frac{1}{n P_i} - 1\right) X_i^g \qquad (5.5.6)$$

and when $\pi_i \propto X_i^\alpha$, where α is real, we have

$$\pi_i = \frac{n X_i^\alpha}{\sum_{1}^{N} X_i^\alpha} = n P_i', \quad \text{where} \quad P_i' = X_i^\alpha \Big/ \sum_{1}^{N} X_i^\alpha$$

and from (5.5.3) and theorem 5.2.2 it follows that

$$\mathcal{E}_{\partial(g)} \; V\left(\Sigma \frac{y_i}{n P_i'}\right) > \sigma^2 \sum_{1}^{N} \left(\frac{1}{n P_i'} - 1\right) X_i^g . \qquad (5.5.7)$$

Comparison of (5.5.6) and (5.5.7) leads to the following result.

<u>Theorem 5.5.2:</u>

$$\mathcal{E}_{\partial(g)} \left[\frac{V\left(\Sigma \frac{y_i}{n P_i'}\right) - V\left(\Sigma \frac{y_i}{n P_i}\right)}{\sigma^2} \right] > \frac{1}{n} f(\alpha),$$

where

$$f(\alpha) = \sum_{1}^{N} X_i^{g-\alpha} \sum_{1}^{N} X_i^\alpha - \sum_{1}^{N} X_i^{g-1} \sum_{1}^{N} X_i .$$

Proof

$$\mathcal{E}_{a(g)}\left\{\frac{V(\Sigma \frac{y_i}{nP_i'}) - V(\Sigma \frac{y_i}{nP_i})}{\sigma^2}\right\} > \sum_1^N \frac{X_i^g}{n\,P_i'} - \sum_1^N \frac{X_i^g}{n\,P_i}$$

$$= \sum_1^N \frac{X_i^g}{nX_i^\alpha} \sum_1^N X_i^\alpha - (\sum_1^N X_i^{g-1} \sum_1^N X_i \;/n)$$

$$= \frac{1}{n}\,(\sum_1^N X_i^{g-\alpha} \sum_1^N X_i^\alpha - \sum_1^N X_i^{g-1} \sum_1^N X_i)$$

$$= \frac{f(\alpha)}{n},$$

where $$f(\alpha) = \sum_1^N X_i^{g-\alpha} \sum_1^N X_i^\alpha - \sum_1^N X_i^{g-1} \sum_1^N X_i . \qquad (5.5.8)$$

Lemma 5.5.1 (Callebaut (1965))*: For $\underset{\sim}{a} = (a_1, a_2, \ldots, a_N)$ and $\underset{\sim}{b} = (b_1, b_2, \ldots, b_N)$, positive vectors which are not proportional, the expression

$$(\Sigma\, a_i^{r+x}\, b_i^{r-x})(\Sigma\, a_i^{r-x}\, b_i^{r+x})$$

increases with increasing $|x|$ for any real number r.

Taking $a_i = X_i$, $b_i = 1$, $r = g/2$ and $x = g/2 - \alpha$, where α is a real number, we have

$$\sum_1^N a_i^{r+x}\, b_i^{r-x} \sum_1^N a_i^{r-x}\, b_i^{r+x} = \sum_1^N X_i^{g-\alpha} \sum_1^N X_i^\alpha \qquad (5.5.9)$$

and in view of Lemma 5.5.1, it follows that

---

*The author is grateful to his colleague Mr. K.Viswanath for bringing into his notice this reference.

$$\Sigma X_i^{g-\alpha} \quad \Sigma X_i^{\alpha}$$

increases with $|\frac{g}{2} - \alpha|$, for any real g (in fact g lies between 0 and 2 and more often between 1 and 2).

Thus

$$\Sigma X_i^{g-\alpha} \ \Sigma X_i^{\alpha} > \Sigma X_i^{g-1} \ \Sigma X_i \qquad (5.5.10)$$

if $$|\frac{g}{2} - \alpha| > |\frac{g}{2} - 1|. \qquad (5.5.11)$$

Thus it can be easily shown that (5.5.10) holds if

$$\alpha > 1 \ \text{ or } \ \alpha < g-1. \qquad (5.5.12)$$

Thus we have the following theorem.

<u>Theorem 5.5.3:</u> $f(\alpha)$ of (5.5.8) is positive for α not lying between $g-1$ and 1.

<u>Remark (a):</u> To illustrate the above result, we consider the graph of $f(\alpha) = \sum_1^N X_i^{g-\alpha} \sum_1^N X_i^{\alpha} - \sum_1^N X_i^{g-1} \sum_1^N X_i$. We distinguish here three cases (i) $2 > g > 1$, (ii) $0 < g < 1$ and (iii) $g = 1$. The following tables give the values of $f(\alpha)$ corresponding to specified α's.

Table 5.5.1

(i) $2 > g > 1$

α	$f(\alpha)$
0	$N \Sigma X_i^g - \Sigma X_i \Sigma X_i^{g-1} > 0$
$g-1$	0
$g/2$	$(\Sigma X_i^{g/2})^2 - \Sigma X_i \Sigma X_i^{g-1} < 0$
1	0
g	$N \Sigma X_i^g - \Sigma X_i \Sigma X_i^{g-1} > 0$

Table 5.5.2

(ii) $0 < g < 1$

α	$f(\alpha)$
$g-1(< 0)$	0
0	$N \Sigma X_i^g - \Sigma X_i \Sigma X_i^{g-1} < 0$
$g/2$	$(\Sigma X_i^{g/2})^2 - \Sigma X_i \Sigma X_i^{g-1} < 0$
g	$N \Sigma X_i^g - \Sigma X_i \Sigma X_i^{g-1} < 0$
1	0

Table 5.5.3

(iii) $g = 1$

α	$f(\alpha)$
0	0
$g/2$	< 0
1	0

In the figures 5.5.1, 5.5.2 and 5.5.3 are shown the values of $f(\alpha)$ plotted against certain specified values of α.

Case (i) $2 > g > 1$

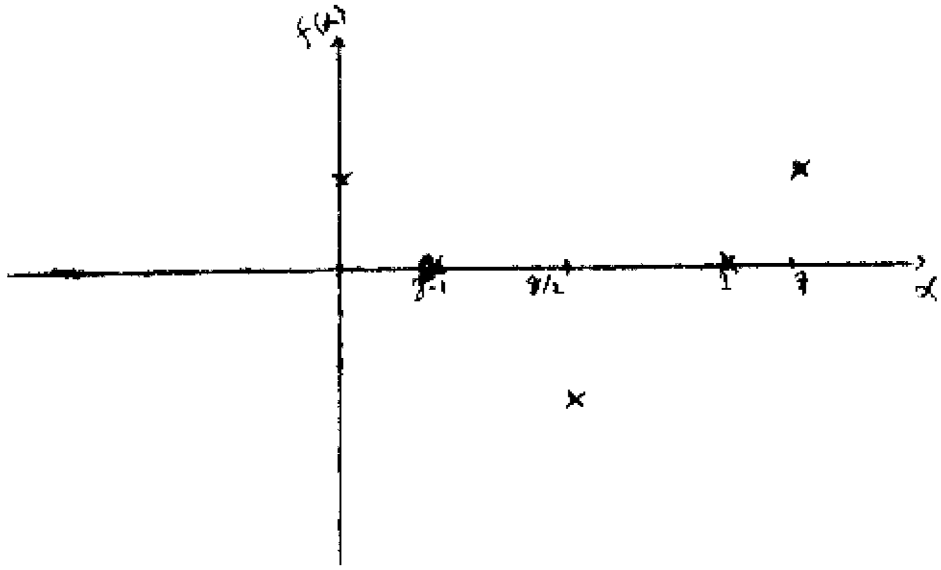

FIGURE 5.5.1

Case (ii) $0 < g < 1$.

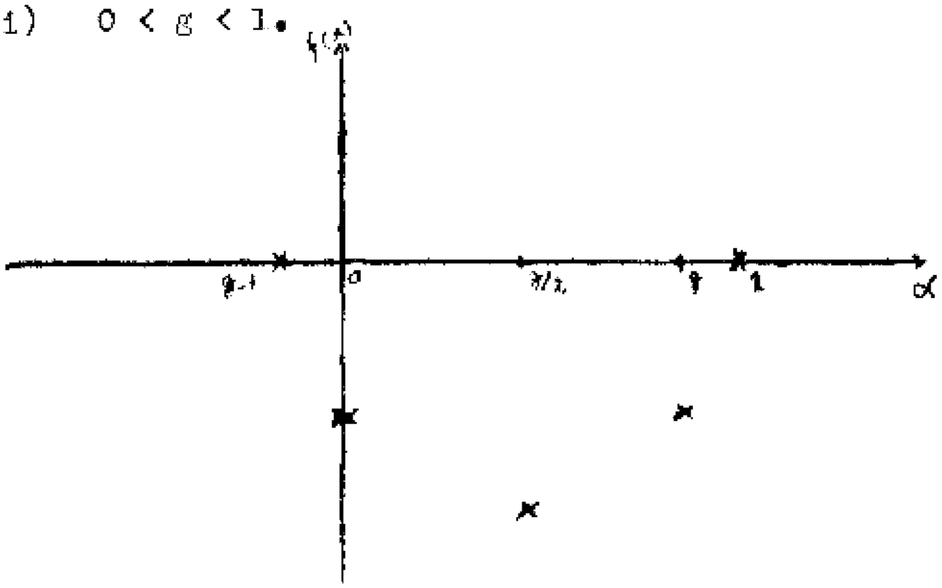

FIGURE 5.5.2

Case (iii) $g = 1$

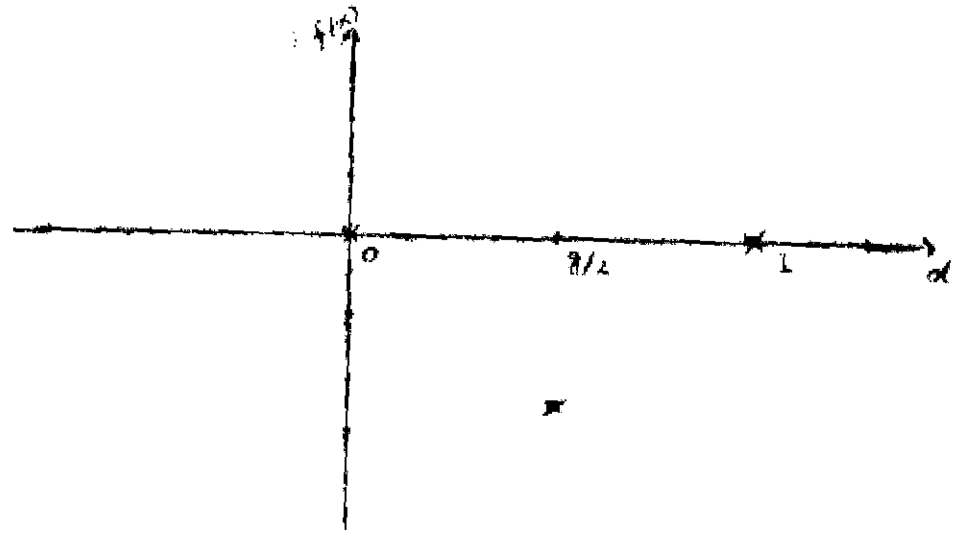

FIGURE 5.5.3

Theorems 5.5.2 and 5.5.3 imply that πPX strategy is better than the πPX^{α} strategies if $\alpha \geq 1$ or $\alpha \leq g-1$ for any given g. (g lying between 0 and 2 and more often between 1 and 2).

Remark(b). Since $f(\alpha) = \Sigma X_i^{g-\alpha} \Sigma X_i^{\alpha} - \Sigma X_i^{g-1} \Sigma X_i$, for a given $g = g_0$,

$$f(g_0) = \Sigma X_i^{g_0-g_0} \Sigma X_i^{g_0} - \Sigma X_i^{g_0-1} \Sigma X_i$$

$$= N \Sigma X_i^{g_0} - \Sigma X_i^{g_0-1} \Sigma X_i$$

$$> 0, \text{ if } g_0 > 1,$$

and

$$f(g_0-1) = \Sigma X_i^{g_0-(g_0-1)} \Sigma X_i^{g_0-1} - \Sigma X_i^{g_0-1} \Sigma X_i$$

$$= 0.$$

Thus it follows that for a given $g_0 > 1$, πPX is better than πPX^{g_0} and $\pi PX^{g_0 - 1}$. Nothing can be said about $\pi PX^{g_0/2}$, since $f(\frac{g_0}{2}) < 0$.

Remark (c): We have seen in theorem 5.2.2. that

$$\mathcal{E}_a \; V\Big(\sum_{i \in s} \frac{y_i}{\pi_i}\Big) = \sigma^2 \sum_{i=1}^{N} X_i^g \Big(\frac{1}{\pi_i} - 1\Big) + a^2 V\Big(\sum_{i \in s} \frac{X_i}{\pi_i}\Big). \quad (5.5.13)$$

Towards the beginning of this section we have also noticed that the second term can be minimized (made zero) by choosing π_i's proportional to X_i and the first term can be minimized by choosing π_i's proportional to $X_i^{g/2}$. We shall now discuss the minimization of the expression (5.5.13) subject to certain internal consistency restrictions. (5.5.13) can be put in the form

$$\mathcal{E}_a \, V\Big(\sum_{i \in s} \frac{y_i}{\pi_i}\Big) = \sigma^2 \sum_{1}^{N} \Big(\frac{1}{\pi_i} - 1\Big) X_i^g + a^2 \Big[\sum_{1}^{N} \Big(\frac{1}{\pi_i} - 1\Big) X_i^2 + \sum_{i \neq j}^{N} \sum^{N} \Big(\frac{\pi_{ij}}{\pi_i \pi_j} - 1\Big) X_i X_j \Big]$$

$$\sum_{1}^{N} \frac{(\sigma^2 X_i^g + a^2 X_i^2)}{\pi_i} + \sum_{i \neq j}^{N} \sum^{N} a^2 \Big(\frac{\pi_{ij}}{\pi_i \pi_j} X_i X_j\Big) - C, \quad \text{where}$$

$$C = \sigma^2 \sum_{1}^{N} X_i^g + a^2 X^2,$$

$$a^2 \Big[\sum_{1}^{N} \frac{\Big(\frac{\sigma^2}{a^2} X_i^g + X_i^2\Big)}{\pi_i} + \sum_{i \neq j}^{N} \sum^{N} \frac{\pi_{ij}}{\pi_i \pi_j} X_i X_j - \frac{C}{a^2} \Big] \quad (5.5.14).$$

To give the problem a familiar look, let

$$\sigma^2/a^2 = \alpha, \quad X_i = m_i, \quad \pi_i = Z_i \quad \text{and} \quad \pi_{ij} = Z_{ij} \quad \text{and} \quad -\frac{C}{a^2} = K.$$

Then (5.5.14) becomes

$$a^2 \left[\sum_{1}^{N} \frac{\alpha\, m_i^g + m_i^2}{Z_i} + \sum_{i \neq j}^{N} \sum^{N} \frac{Z_{ij}}{Z_i\, Z_j}\, m_i m_j + K \right].$$

Thus we have the problem of the following type:

$$\text{Minimize} \quad f(Z) = \sum_{i}^{N} (\alpha\, m_i^g + m_i^2) \frac{1}{Z_i} + \sum_{i \neq j}^{N} \sum^{N} \frac{Z_{ij}}{Z_i Z_j}\, m_i m_j \qquad (5.5.15)$$

where m_i, α and g are constants, subject to the restrictions

$$\left.\begin{aligned} &1 > Z_i > 0 \quad \text{for} \quad 1 \leq i \leq N \\ &\text{Min.}\ (Z_i, Z_j) \geq Z_{ij} > 0, \quad \text{for} \quad 1 \leq i \neq j \leq N \\ &Z_{ij} \geq Z_i + Z_j - 1, \quad \text{for all} \quad i \neq j \\ &\sum_{i=1}^{N} Z_i = n, \end{aligned}\right\} \qquad (5.5.16)$$

subject to (5.5.16) together with $\sum_{\neq i} Z_{ij} = (n-1)Z_i$.

The objective function (5.5.15) is non-linear and we have to attack the problem by finding approximations to $1/Z_i$ and $\frac{Z_{ij}}{Z_i Z_j}$ and by using the techniques like separable objective functions and the techniques of linear piecewise approximations to functions. However, the size of the problem is too big and simpler techniques of attempting general non-linear programming problems are yet to be developed which can be fruitfully used in solving the problem posed in this section.

5.6. Illustrative Examples:

We first illustrate the results obtained in Section 5.2 by considering the data on the four states of India, given in Chapter IV (pp. 63-67).

For the three stategies considered here, we have under the model (5.2.1)

$$E_1 = \mathcal{E}_\partial \; V(\hat{T}_1)/\sigma^2 = \sum_1^N X_i^g \,(T_i - 1),$$

$$E_2 = \mathcal{E}_\partial \; V(\hat{Y}_{HT})/\sigma^2 = \sum_1^N X_i^g \left(\frac{1}{\pi_i'} - 1\right) + K_p \, V\left(\sum_{i \in s} \frac{X_i}{\pi_i'}\right)$$

and $$E_3 = \mathcal{E}_\partial \; V(\hat{T}_o)/\sigma^2 = \sum_1^N X_i^g \left(\frac{1}{nP_i} - 1\right),$$

where T_i and π_i' are as defined in (4.4.7) and (4.5.1) respectively and K_p is a constant for the population P (equal to a^2/σ^2 for that population).

Tables 5.6.1 and 5.6.2 give these expected variances for he four populations for the values of g given by 1.5, 1.75 and and for the sample sizes $n = 2$ and $n = 3$.

Table 5.6.1

Sample size $n = 2$

tate	g-value	E_1	E_2	E_3
dhra adesh	1.5	$.11163 \times 10^8$	$.10884 \times 10^8 + K_1(.11682 \times 10^8)$	$.10856 \times 10^8$
	1.75	$.76585 \times 10^8$	$.74701 \times 10^8 + K_1(.11682 \times 10^8)$	$.73539 \times 10^8$
	2	$.45474 \times 10^9$	$.44384 \times 10^9 + K_1(.11682 \times 10^8)$	$.43123 \times 10^9$
erala	1.5	$.18422 \times 10^7$	$.18515 \times 10^7 + K_2(.46305 \times 10^6)$	$.18306 \times 10^7$
	1.75	$.11503 \times 10^8$	$.11437 \times 10^8 + K_2(.46305 \times 10^6)$	$.11392 \times 10^8$
	2	$.71917 \times 10^8$	$.71513 \times 10^8 + K_2(.46305 \times 10^6)$	$.70972 \times 10^8$
'arat	1.5	$.39023 \times 10^7$	$.36360 \times 10^7 + K_3(.87834 \times 10^7)$	$.35560 \times 10^7$
	1.75	$.22472 \times 10^8$	$.21010 \times 10^8 + K_3(.87834 \times 10^7)$	$.19872 \times 10^8$
	2	$.13082 \times 10^9$	$.12271 \times 10^9 + K_3(.87834 \times 10^7)$	$.11277 \times 10^9$
st Bengal	1.5	$.76059 \times 10^7$	$.69990 \times 10^7 + K_4(.26586 \times 10^8)$	$.69642 \times 10^7$
	1.75	$.50384 \times 10^8$	$.46648 \times 10^8 + K_4(.26586 \times 10^8)$	$.44143 \times 10^8$
	2	$.34047 \times 10^9$	$.31711 \times 10^9 + K_4(.26586 \times 10^8)$	$.28564 \times 10^9$

Table 5.6.2

Sample size n = 3

State	g-value	E_1	E_2	E_3
Andhra Pradesh	1.5	$.70858 \times 10^7$	$.69289 \times 10^7 + K_1(.12874 \times 10^8)$	$.68142 \times 10^7$
	1.75	$.48762 \times 10^8$	$.47690 \times 10^8 + K_1(.12874 \times 10^8)$	$.46099 \times 10^8$
	2	$.29080 \times 10^9$	$.28453 \times 10^9 + K_1(.12874 \times 10^8)$	$.26986 \times 10^9$
Kerala	1.5	$.10546 \times 10^7$	$.10490 \times 10^7 + K_2(.48577 \times 10^6)$	$.10439 \times 10^7$
	1.75	$.65919 \times 10^7$	$.65571 \times 10^7 + K_2(.48577 \times 10^6)$	$.64886 \times 10^7$
	2	$.41257 \times 10^8$	$.41042 \times 10^8 + K_2(.48577 \times 10^6)$	$.40376 \times 10^8$
Gujarat	1.5	$.24762 \times 10^7$	$.23298 \times 10^7 + K_3(.92305 \times 10^7)$	$.21869 \times 10^7$
	1.75	$.14388 \times 10^8$	$.13564 \times 10^8 + K_3(.92305 \times 10^7)$	$.12161 \times 10^8$
	2	$.84476 \times 10^8$	$.79792 \times 10^8 + K_3(.92305 \times 10^8)$	$.68685 \times 10^8$
West Bengal	1.5	$.48572 \times 10^7$	$.44897 \times 10^7 + K_4(.30391 \times 10^8)$	$.42391 \times 10^7$
	1.75	$.32676 \times 10^8$	$.30345 \times 10^8 + K_4(.30391 \times 10^8)$	$.26607 \times 10^8$
	2	$.22414 \times 10^9$	$.20911 \times 10^9 + K_4(.30391 \times 10^8)$	$.17034 \times 10^9$

It is easily seen that $E_3 < E_1$ in all the cases and it is also interesting to see that $E_3 < E_2$, even though we neglect the positive contribution in the second term of E_2. These examples show that a sample of size 2 or 3 itself would suffice to make $E_3 < E_2$ and this is what we have been stressing in remark (a) on the theorem 5.2.4.

Next, we illustrate the results of section 5.5, using the data on the 4 populations considered in Chapter IV (pp. 63-67). While comparing the two strategies for one of which $\pi_i \propto X_i$ and for the other $\pi_i \propto X_i^\alpha$, where α is real, we proved in theorem 5.5.2 that

$$\mathcal{E}_{\partial(g)} \left\{ \frac{V\left(\Sigma \dfrac{y_i}{(n X_i^\alpha / \Sigma X_i^\alpha)}\right) - V\left(\Sigma \dfrac{y_i}{(n X_i | \Sigma X_i)}\right)}{\sigma^2} \right\} > \frac{1}{n} f(\alpha),$$

where $f(\alpha) = \sum_1^N X_i^{g-\alpha} \sum_1^N X_i^\alpha - \sum_1^N X_i^{g-1} \sum_1^N X_i$.

The three tables given below present the values of $f(\alpha)$ for various values of α and for specified values of g given by 1.5, 0.5 and 1, for the 4 populations.

Table 5.6.3

$g = 1.5$

$f(\alpha)$ / α	Andhra Pradesh	Kerala	Gujarat	West Bengal
0	1145750.48	46026.40	1158671.18	3024900.22
.5	0	0	0	0
.75	- 139478.87	-5709.56	-121638.86	-347713.70
1.	0	0	0	0
1.5	1145750.48	46026.40	1158671.18	3024900.22

Table 5.6.4

$g = .5$

$f(\alpha)$ / α	Andhra Pradesh	Kerala	Gujarat	West Bengal
-.5	0	0	0	0
0	-801.64	-27.80	-2511.20	-1803.37
.25	-899.16	-31.26	-2741.60	-2009.98
.5	-801.64	-27.80	-2511.20	-1803.37
1	0	0	0	0

Table 5.6.5

$g = 1.$

$f(\alpha)$ / α	Andhra Pradesh	Kerala	Gujarat	West Bengal
0	0	0	0	0
.5	-14830.70	-562.63	-21370.07	-34404.41
1	0	0	0	0

It is quite clear from the above tables that the value of $f(\alpha)$ is positive for α not lying between $g-1$ and 1, thus illustrating that πPX strategy is better than πPX^{α} strategy if $\alpha \notin (0, 1)$ for any g.

CHAPTER VI

ALLOCATION OF SAMPLE SIZE IN STRATIFIED SAMPLING

6.0 Summary:

In this chapter, the problem of optimum allocation of sample size to strata is examined in the light of a priori distributions. In this context, the justification for the assumption that the unknown proportionate values of σ_i^2's can be replaced by the proportionate values of the known α_i^2's which are estimates of σ_i^2's is discussed. Allocation of sample size to strata which minimizes the expected variance of the strategy consisting of πPS sampling scheme and the Horvitz-Thompson estimator under a general super-population model is derived. It is further shown that, in the sense of expected variance, πPS sampling for unstratified sampling is inferior to πPS stratified sampling with this type of allocation, unless the super-population parameter g attains the value two, in which case both the schemes are equally precise.

6.1 Introduction:

In the case of Neyman's ootimum allocation (Neyman [1934]) of sample size to strata, we have the allocation given by

$$n_{i,opt} = n \frac{N_i \sigma_i}{\sum_{i=1}^{h} N_i \sigma_i} , \qquad (6.1.1)$$

where N_i is the size of the i^{th} stratum, $i = 1,2,\ldots, h$; n is the total sample size and σ_i^2 is the within variance for the i^{th} stratum. Computation of $n_{i,opt.}$'s requires at least the proportionate values of σ_i^2's, which are unknown. In practice, some estimates α_i^2's of σ_i^2's are substituted. These estimates, usually, are the σ_i^2's of some auxiliary information closely related to the characteristic under study. In most of the cases, the values of the same characteristic studied in a previous occasion are treated as the auxiliary information. The justification for the assumption that the unknown proportionate values of σ_i^2's are usually not quite different from the proportionate values of the known α_i^2's is examined in the light of an *a priori* distribution by Hanurav [1965] and a further development is given in this chapter.

6.2 Optimum allocation of sample size and prior distributions:

We consider the same super-population set-up as was considered earlier (refer 5.2.1). Let Δ_g be the class of all prior distributions ∂ for which

$$\left.\begin{aligned} &\text{i)}\quad \mathcal{E}_\partial (Y_i \mid X_i) = aX_i, \\ &\text{ii)}\quad \mathcal{V}_\partial (Y_i \mid X_i) = \sigma^2 X_i^g, \quad g > 1 \\ &\text{and iii)}\quad \mathcal{C}_\partial (Y_i, Y_j \mid X_i, X_j) = 0 \end{aligned}\right\} \qquad (6.2.1)$$

where $\mathcal{C}$ denotes covariance.

We have

$$\sigma_i^2 = \frac{1}{N_i} \left(\sum_{j=1}^{N_i} Y_{ij}^2 - \frac{Y_i^2}{N_i} \right), \qquad (6.2.2)$$

where Y_{ij} is the value of the characteristic $\mathcal{Y}$ for the j^{th} unit of the i^{th} stratum ($j = 1,2,\ldots, N_i$; $i = 1,2,\ldots,h$) We then have

$$\mathcal{E}_\partial (\sigma_i^2) = \frac{1}{N_i} \left[\sum_j (a^2 X_{ij}^2 + \sigma^2 X_{ij}^g) - \frac{1}{N_i} (a^2 X_i^2 + \sigma^2 \sum_j X_{ij}^g) \right]$$

$$= \frac{1}{N_i} \left[a^2 \left(\sum_j X_{ij}^2 - \frac{X_i^2}{N_i} \right) + \sigma^2 \left(1 - \frac{1}{N_i}\right) \sum_j X_{ij}^g \right]$$

$$= a^2 \alpha_i^2 + \sigma^2 \frac{N_i - 1}{N_i} \frac{\sum_j X_{ij}^g}{N_i} ,$$

where $\alpha_i^2 = \frac{1}{N_i} \left(\sum_j X_{ij}^2 - \frac{X_i^2}{N_i} \right)$

$\simeq a^2 \alpha_i^2 + \sigma^2 \sum_j X_{ij}^g / N_i$, assuming $\frac{N_i - 1}{N_i}$'s to be nearly equal to unity.

Consequently,

$$\mathcal{E}_\partial(\sigma_i^2) = a^2\alpha_i^2 + \sigma^2[(\sum_j X_{ij}^g/N_i) - (\sum_j X_{ij}^2/N_i) + \alpha_i^2 + \bar{X}_i^2]$$

$$= (a^2 + \sigma^2)\alpha_i^2 + \sigma^2[\bar{X}_i^2 + \frac{1}{N_i} (\sum_j X_{ij}^g - \sum_j X_{ij}^2)] . \tag{6.2.3}$$

Thus σ_i^2 's can be expected to be in the same proportion as α_i^2 's if α_i^2 's are proportional to

$$\bar{X}_i^2 + \frac{1}{N_i} (\sum_j X_{ij}^g - \sum_j X_{ij}^2);$$

i.e., if the 'corrected coefficients of variation' (c.c.v.) of the $\mathcal{X}$-characteristic, defined by

$$\frac{\alpha_i^2}{\bar{X}_i^2 - \partial_i/N_i^2} ,$$

where $\partial_i = N_i (\sum_j X_{ij}^2 - \sum_j X_{ij}^g)$ are equal for all the strata and when this condition is satisfied the allocation is given by

$$n_i = n \frac{N_i \sqrt{\bar{X}_i^2 - \frac{\partial_i}{N_i^2}}}{\sum_{i=1}^{h} N_i \sqrt{\bar{X}^2 - \frac{\partial_i}{N_i^2}}}$$

$$= n \frac{\sqrt{T_i^2 - \partial_i}}{\sum_{i=1}^{h} \sqrt{T_i^2 - \partial_i}} , \quad \text{where} \quad T_i = \sum_{j=1}^{N_i} X_{ij} .$$

(6.2.4)

Thus we have the following theorem.

__Theorem 6.2.1:__ Under the super-population model Δ_g of (6.2.1), Neyman's optimum allocation reduces to allocation proportional to $\sqrt{T_i^2 - \partial_i}$, where $T_i = \sum_{j=1}^{N_i} X_{ij}$ is the total of the auxiliary variate for the i^{th} stratum and

$$\partial_i = N_i \left(\sum_{j=1}^{N_i} X_{ij}^2 - \sum_{j=1}^{N_i} X_{ij}^g \right),$$

when the corrected coefficients of variation of the $\mathcal{X}$ characteristic are equal for all the strata.

From the above theorem follows immediately

Theorem 6.2.2 (Hanurav [1965]): Neyman's optimum allocation reduces to allocation proportional to the stratum totals of the auxiliary variate $\mathcal{X}$, under Δ_2, when the coefficients of variation of the $\mathcal{X}$-characteristic are equal for all the strata.

Proof: Under Δ_2, we have $a_i = 0$ and hence

$$n_i = n\, T_i \Big/ \sum_{i=1}^{h} T_i = \frac{n\, T_i}{T} \quad \text{and hence etc.}$$

Remark: Since, in the derivation of (6.2.3) we are assuming that $\frac{N_i - 1}{N_i}$'s are approximately equal to unity, it can be easily seen that the same algebra works out for Simple Random Sampling With Out Replacement case also, where $n_{i,opt.}$'s are given by

$$n_{i,opt.} = n \frac{N_i S_i}{\sum_{i=1}^{h} N_i S_i}.$$

6.3. Stratified πPS sampling:

Consider a population consisting of h strata with N_i as the size of the i^{th} stratum, $i = 1, 2, \ldots, h$. Let Y_{ij} and X_{ij} be the values of the $\mathcal{Y}$-characteristic and the $\mathcal{X}$-characteristic (auxiliary information) respectively, for the j^{th} unit of the i^{th} stratum. Let a πPS sample (vide 5.4) of size n_i be taken from the i^{th} stratum, such that $\sum_{i=1}^{h} n_i = n$. Let π_{ij} denote the probability of inclusion of the j^{th} unit of the i^{th} stratum in the sample (not to be confused with the usual notation of π_{ij} for the joint inclusion probability for the pair of units (i,j)). As an estimator of the population total, consider the Horvitz-Thompson [1952] estimator

$$\hat{Y} = \sum_{i=1}^{h} \sum_{j=1}^{n_i} y_{ij} / \pi_{ij}. \qquad (6.3.1)$$

Then we have the following theorem.

<u>Theorem 6.3.1</u>: Under the super-population model Δ_g, allocation of the sample size to the strata which minimizes the expected variance of (6.3.1) is given by

$$n_i = \frac{n \sqrt{T_i \sum_{j=1}^{N_i} X_{ij}^{g-1}}}{\sum_{i=1}^{h} \sqrt{T_i \sum_{j=1}^{N_i} X_{ij}^{g-1}}},$$

where T_i is the total of X-values for the i^{th} stratum.

Proof: We have

$$\mathcal{E}_\partial V(\hat{Y}) = \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left(\frac{1}{\pi_{ij}} - 1\right)\sigma_{ij}^2 \qquad (6.3.2)$$

where $\sigma_{ij}^2 = \sigma^2 X_{ij}^g$ (Godambe [1955], Hanurav [1962a]).

We now minimize (6.3.2) subject to the condition $\sum_{i=1}^{h} n_i = n$.

Introducing the Lagrange's multiplier λ, consider

$$\mathcal{E}_\partial V(\hat{Y}) + \lambda\left(\sum_{i=1}^{h} n_i - n\right)$$

$$= \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left(\frac{1}{n_i P_{ij}} - 1\right) \sigma_{ij}^2 + \lambda\left(\sum_{i=1}^{h} n_i - n\right)$$

ere $P_{ij} = \dfrac{X_{ij}}{T_i}$.

Differentiating with respect to n_i and equating to zero, we have

$$-\frac{1}{n_i^2} \sum_j \frac{\sigma_{ij}^2}{P_{ij}} + \lambda = 0, \quad \text{which gives}$$

$$n = \frac{1}{\sqrt{\lambda}} \sqrt{\sum_j \frac{\sigma_{ij}^2}{P_{ij}}} \quad \text{and thus}$$

$$n_i = n\sqrt{\sum_j (\sigma_{ij}^2 / P_{ij})} \Big/ \sum_{i=1}^{h} \sqrt{\sum_j \sigma_{ij}^2 / P_{ij}} .$$

$$= n\sqrt{T_i \sum_j X_{ij}^{g-1}} \Big/ \sum_{i=1}^{h} \sqrt{T_i \sum_j X_{ij}^{g-1}} \qquad (6.3.3)$$

and hence the theorem.

Definition 6.3.1: The allocation of sample size to the strata given in (6.3.3) is called the Δ_g-optimum allocation.

Theorem 6.3.2: Under the super-population model Δ_2, allocation of the sample size to the strata proportional to the stratum totals of the $\mathcal{X}$ -variate minimizes the expected variance of (6.3.1).

Proof: Follows directly from theorem 6.3.1, by putting $g=2$.

Let n_i be the Δ_g-optimum sample size for the i^{th} stratum and let the sample be selected by using a πPS sampling scheme within each stratum. Then as an estimator of the population total consider the estimator (6.3.1)

$$\hat{Y} = \sum_{i=1}^{h} \sum_{j=1}^{n_i} \frac{y_{ij}}{\pi_{ij}} = \sum_{i=1}^{h} \sum_{j=1}^{n_i} \frac{y_{ij}}{(n_i x_{ij}/T_i)} .$$

We now have the following theorem.

<u>Theorem 6.3.3:</u> In the sense of expected variance, under Δ_g, πPS unstratified sampling is inferior to πPS stratified sampling with Δ_g optimum allocation and for $g = 2$, both the schemes are equivalent.

<u>Proof:</u> We have

$$\mathcal{E}_\partial \, V(\hat{Y}_{st.}) = \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left[\frac{1}{\pi_{ij}} - 1\right] \sigma^2 X_{ij}^g$$

$$\text{where} \quad \pi_{ij} = \frac{n\mu_i X_{ij}}{X_i}, \quad \text{where} \quad \mu_i = \frac{\sqrt{X_i \sum_j X_{ij}^{g-1}}}{\sum_{i=1}^{h} \sqrt{X_i \sum_j X_{ij}^{g-1}}},$$

and

$$\mathcal{E}_\partial \, V(\hat{Y}_{unst.}) = \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left[\frac{1}{\pi'_{ij}} - 1\right] \sigma^2 X_{ij}^g ,$$

$$\text{where} \quad \pi'_{ij} = \frac{n X_{ij}}{X} .$$

Therefore

$$\mathcal{E}_\partial \left[\frac{V(\hat{Y}_{unst.}) - V(\hat{Y}_{st.})}{\sigma^2}\right] = \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left[\frac{X_{ij}^g}{\pi'_{ij}} - \frac{X_{ij}^g}{\pi_{ij}}\right]$$

$$= \sum_{i=1}^{h} \sum_{j=1}^{N_i} \left[\frac{X_{ij}^{g}}{\left(\frac{nX_{ij}}{X}\right)} - \frac{X_{ij}^{g}}{\left(\frac{n\mu_i X_{ij}}{X_i}\right)} \right]$$

$$= \frac{X}{n} \sum_i \sum_j X_{ij}^{g-1} - \sum_i \frac{X_i}{n\mu_i} \sum_j X_{ij}^{g-1}$$

$$= \frac{X}{n} \sum_i \sum_j X_{ij}^{g-1} - \sum_i \frac{X_i \sum_j X_{ij}^{g-1}}{n\left(\frac{\sqrt{X_i \sum_j X_{ij}^{g-1}}}{\sum_i \sqrt{X_i \sum_j X_{ij}^{g-1}}}\right)}$$

$$= \frac{X}{n} \sum_i \sum_j X_{ij}^{g-1} - \left(\sum_i \sqrt{X_i \sum_j X_{ij}^{g-1}}\right)^2 / n$$

$$= \frac{1}{n} \left[\sum_{i=1}^{h} X_i \sum_{i=1}^{h} \left(\sum_{j=1}^{N_i} X_{ij}^{g-1} \right) - \left(\sum_{i=1}^{h} \sqrt{X_i} \sqrt{\sum_{j=1}^{N_i} X_{ij}^{g-1}} \right)^2 \right] \geq 0$$

by Cauchy-Shwartz's inequality.

It is easily seen that when $g = 2$, equality occurs. Hence the theorem.

CHAPTER VII

CHOICE OF AN OPTIMUM SAMPLING STRATEGY

7.0 Summary:

We first discuss briefly the criteria of optimality of sampling strategies and show that under the criteria of unbiasedness and hyper admissibility the problem reduces to that of constructing a sampling design D_1, such that D_1 has the same inclusion probabilities as the given design D_2 so that they are equally costly and D_1 has uniformly smaller (or equal) joint inclusion probabilities π_{ij}'s, than those of D_2. It is shown in this chapter that it is possible to construct such designs in general and a constructive proof is given. An illustrative example explains the algorithm used.

7.1 *Introduction:*

Of the various criteria of optimality of sampling strategies, we have already seen how the Bayesian approach can be utilised as a tool for searching optimum strategies. Here, we assume a stochastic model for $\underset{\sim}{Y}$, the values of the characteristic, parameteric functions of which we are estimating, and treat the actual $\underset{\sim}{Y}$- values as a realisation of a random vector whose distribution depends on $\underset{\sim}{X}$, the auxiliary characteristic, highly correlated with $\underset{\sim}{Y}$ and known before hand. Under this apriori distribution, say ∂, we minimise the expected $V(H(S, P, T))$ uniformly in $\underset{\sim}{Y}$, for variations of H over the class of all equally-costly strategies. Thus, here we minimise the expected loss

$$\underset{\partial}{E}\, V(H) = \int V(H)\, d\partial \quad \text{over} \quad \partial. \qquad (7.1.1)$$

In chapter V, we have already noticed various applications of this criterion which helps in arriving at optimal sampling strategies.

Another concept, is the concept of 'linear invariance' introduced by Roy and Chakravarty [1960] wherein they demand that the estimator should be invariant under linear transformation of $\underset{\sim}{Y}$. But this criterion is not of much use leading us to an optimum estimator, because most of the own estimators of varying probability sampling scheme do t fall under this category.

Roy and Chakravarty [1960] also introduced a second oncept which deals with a class of estimators $\{ t \}$, known 'regular estimators', which are such that

$$V(t) = k.\sigma^2, \tag{7.1.2}$$

where k is a constant and $\sigma^2 = \frac{1}{N} (\sum_{1}^{N} Y_i^2 - N \bar{Y}^2)$.

(7.1.2) is satisfied by designs termed by Roy and Chakravarty as 'balanced designs' and they have shown that a best estimator exists in this class of designs. But this being a severe demand, it does not help us much in choosing an optimum.

We say that an estimator t_1 of an estimable parameteric function T is 'admissible' if and only if there oes not exist another estimator t_2 of T which is better than t_1. In other words, given any other $t_2 \neq t_1$, there xists at least one value of $\underset{\sim}{Y}^o$ such that

$$M(t_1)_{\text{at } \underset{\sim}{Y} = \underset{\sim}{Y}^o} < M(t_2)_{\text{at } \underset{\sim}{Y} = \underset{\sim}{Y}^o}. \tag{7.1.3}$$

This criterion can be used to weed out all the inadmissible estimators, but does not lead to an optimum.

Hanurav [1965, 1966] introduced a criterion which is stronger than admissibility, but weaker than uniform minimum variance. While admissibility of t_1 needs $M(t_1) < M(t_2)$ for any other t_2, at at least one $\underset{\sim}{Y}^o$ in R^N, the new concept termed by Hanurav as 'hyperadmissibility', demands for an estimator t_1 of a parametric function T to be hyper admissible that for any other estimator t_2 belonging to the class $\mathcal{C}$ of unbiased estimators of the parametric function T, in every hyper plane of R^N, there should exist at least one point at which

$$M(t_1) < M(t_2). \qquad (7.1.4)$$

One obvious, practical implication of this criterion, which is highly useful in large scale surveys is that, the sub-totals, means of sub-populations etc., would all be estimated by admissible estimators, derived from the single estimator t which is hyper admissible, by setting $Y_i = 0$, if u_i does not belong to the sub-population in which we are interested. This concept also leads to the following theorem of

Hanurav [1966], which focuses the attention on a single estimator during the search of optimum estimators.

Theorem 7.1.1. (Hanurav [1966]): For any design D(S, P) which is not a unicluster design* the class of all polynomial unbiased estimators of $Y = \sum_{1}^{N} Y_i$ admits just one estimator which is hyper admissible and this optimum estimator is the Horvitz-Thompson estimator $\sum_{i \in s} \frac{y_i}{\pi_i}$.

This pins down the optimum estimator uniquely to the Horvitz-Thompson estimator and thus helps us in tiding over the difficulties involved in the first step of the choice of an optimum strategy. Thus, the next step in the problem of choice of an optimum strategy is the problem of choice of optimum designs. In the next section we formulate the problem, consider the open problem posed by Hanurav [1966] and give a method of constructing optimum designs.

7.2 Problem of choice of sampling design:

If we start with the criteria of unbiasedness and minimum variance, we can say that a design D_1 is better than another design D_2, if given any estimator t_2, there

* A design D(S,P) is said to be a 'unicluster design' if for any two samples s_i and s_j of D, either both contain the same units of the population or they do not have any common unit.

exists an estimator t_1 such that

$$V(t_1) \leq V(t_2) \quad \text{for all } \underset{\sim}{Y} \in R^N \qquad (7.2.1)$$

where t_1 and t_2 are any admissible estimators.

Though this is the natural and correct way of choosing the design, we do not have a result which leads to an optimum. We have utilised the Bayesian approach of 7.1 to arrive at Δ-optimum designs, but the results 2.25 and 2.26 tell us the restricted nature of these optimum designs. But, instead of unbiasedness and minimum variance, if we accept unbiasedness and hyper admissibility as the criteria for choosing the optimum, then we have just to show that

$$V(\hat{Y}_{HT}(D_1)) \leq V(\hat{Y}_{HT}(D_2)), \text{ for all } \underset{\sim}{Y} \in R^N \qquad (7.2.2)$$

with strict inequality for at least one $\underset{\sim}{Y}$, to establish the superiority of D_1 over D_2.

For any design D with π_i, $i = 1,2,\ldots, N$ as the probability of inclusion of the i^{th} unit and π_{ij}, $(i \neq j) = 1,2,\ldots, N$ as the probability of joint inclusion of the pair of units (i,j) we recall that

$$V(\hat{Y}_{HT}(D)) = \sum_{1}^{N} \frac{Y_i^2}{\pi_i(D)} + \sum_{i \neq j}^{N} \sum^{N} \frac{Y_i Y_j}{\pi_i(D)\pi_j(D)} \pi_{ij}(D) - Y^2. \tag{7.2.3}$$

Thus from (7.2.2) and (7.2.3) we have the condition for D_1 to be superior to D_2 as

$$\sum_{1}^{N} Y_i^2 \left\{ \frac{1}{\pi_i(D_1)} - \frac{1}{\pi_i(D_2)} \right\} + \sum_{i \neq j}^{N} \sum^{N} Y_i Y_j \left\{ \frac{\pi_{ij}(D_1)}{\pi_i(D_1)\pi_j(D_1)} - \frac{\pi_{ij}(D_2)}{\pi_i(D_2)\pi_j(D_2)} \right\}$$

$$\leq 0, \quad \text{for all } \underset{\sim}{Y} \subset R^N. \tag{7.2.4}$$

A necessary condition for (7.2.4) to hold is that

$$\pi_i(D_1) \geq \pi_i(D_2). \tag{7.2.5}$$

A meaningful comparison of the designs can be effected when they are equally costly, i.e., the cost involved for one design should be the same as that of the other and this restriction, when the cost is assumed to be an increasing function of π_i, such as expected effective sample size ν, together with (7.2.5) implies that

$$\pi_i(D_1) = \pi_i(D_2) = \pi_i(\text{say}) \quad \text{for } i=1,2,\ldots, N. \tag{7.2.6}$$

Thus from (7.2.4) and (7.2.6) we get

$$\sum_{i \neq j}^{N}\sum^{N} \frac{Y_i Y_j}{\pi_i \pi_j} \left\{ \pi_{ij}(D_1) - \pi_{ij}(D_2) \right\} \leq 0, \quad \text{for all } \underset{\sim}{Y} \varepsilon R^N \tag{7.2.7}$$

as a set of necessary and sufficient conditions for D_1 to be superior to D_2. A set of sufficient conditions for (7.2.7) to be true is that

$$\pi_{ij}(D_1) \leq \pi_{ij}(D_2) \quad \text{for} \quad i = 1,2,\ldots,N;$$
$$j = 1,2,\ldots, N \quad \text{and} \quad i \neq j. \tag{7.2.8}$$

Given a design D_2 with $\pi_i(D_2)$ as the probability of inclusion of the i^{th} unit in the sample, $i = 1,2,\ldots, N$; and π_{ij} (D_2) as the joint inclusion probability of the pair of units (i,j), $i = 1,2,\ldots, N$; $j = 1,2,\ldots, N$; $i \neq j$; the problem thus reduces to that of constructing a design D_1, with $\pi_i(D_1)$ equal to $\pi_i(D_2)$, $i = 1,2,\ldots, N$, and with π_{ij} (D_1) uniformly smaller or equal to $\pi_{ij}(D_2)$, $i = 1,2,\ldots, N$; $j = 1,2,\ldots, N$ and $i \neq j$, in which case D_1 would be superior to D_2. Hereafter owing to the repeated use of this, we denote this property by $\wp$.

Lemma 7.2.1. (Hanurav [1966]): For any design $D(S, P)$

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} \geq \nu(\nu - 1) + f(1 - f)$$

where ν is expected effective sample size and f is the fractional part of ν given by $\nu - [\nu]$, where $[\nu]$ denotes the greatest integer not exceeding ν.

Proof: It is easy to see that

$$V(\nu(s)) \geq f(1 - f) \qquad (7.2.9)$$

since we should have

$$\nu(s) = [\nu] \text{ with probability } f$$
$$= [\nu] + 1 \text{ with probability } 1 - f.$$

Also we have from Hanurav [1962a]

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} = \nu(\nu - 1) + V(\nu(s)). \qquad (7.2.10)$$

From (7.2.9) and (7.2.10) follows that

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} \geq \nu(\nu - 1) + f(1 - f). \qquad (7.2.11)$$

Theorem 7.2.1: It is not possible to construct a design D_1 with the required property $\wp$, when $\sum_{i \neq j}^{N} \sum^{N} \pi_{ij}(D_2)$ attains its lower bound.

Proof: Let, if possible, there exist a design D_1 with the required property $\wp$. Then follows that

$$\pi_{ij}(D_1) \leq \pi_{ij}(D_2), \text{ with strict inequality for}$$

at least one pair (i, j), which gives that

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij}(D_1) < \sum_{i \neq j}^{N} \sum^{N} \pi_{ij}(D_2) = \nu(\nu - 1) + f(1 - f). \tag{7.2.12}$$

But from (7.2.11) we have

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij}(D_1) \geq \nu(\nu - 1) + f(1 - f) \tag{7.2.13}$$

and thus (7.2.12) contradicts (7.2.13) and hence it follows that there does not exist a design D_1 with the property $\wp$.

Theorem 7.2.2: If the design D_2 is such that

$$\pi_{ij}(D_2) = \pi_i(D_2) + \pi_j(D_2) - 1,$$

then there does not exist a D_1 with the property $\wp$.

Proof: Let, if possible, there exist a design D_1 with the required property $\wp$. Then we have $\pi_{ij}(D_1) \leq \pi_{ij}(D_2)$, with strict inequality for at least one pair (i,j), from which it follows that

$$\pi_{ij}(D_1) \leq \pi_{ij}(D_2)$$
$$= \pi_i(D_2) + \pi_j(D_2) - 1$$
$$= \pi_i(D_1) + \pi_j(D_1) - 1, \qquad (7.2.14)$$

with strict inequality for at least one pair (i,j).

But we know from the consistency relations between π_{ij}'s and π_i's that for any design D (Hanurav [1966])

$$\pi_{ij}(D) \geq \pi_i(D) + \pi_j(D) - 1. \qquad (7.2.15)$$

Thus for the design D_1, (7.2.14) is contradictory to the consistency relation (7.2.15) which does not allow the construction of the design D_1 with the required property.

Remark: For the design D_2 of theorem 7.2.2, the probability that neither U_i nor U_j is included in a random sample is zero. This implies that the minimum size of the samples in the design is $N-1$ and mostly we are not interested in such designs.

Lemma 7.2.2: If $\sum_{i=1}^{\partial} a_i > \frac{\Delta}{2} > 0$, a_i being non-negative and not all zero, then there exists a partition of $\frac{\Delta}{2}$ such hat

$$\frac{\Delta}{2} = \Delta_1 + \Delta_2 + \dots + \Delta_\partial$$

such that $a_i - \Delta_i \geq 0$ for $i = 1,2,\dots,\partial$ with strict inequality at least once.

oof: Given non-negative a_i, not all zero such that

$$\sum_{i=1}^{\partial} a_i > \frac{\Delta}{2} > 0 . \qquad (7.2.16)$$

Choose

$$\begin{aligned} \Delta_1 &= a_1 \\ \Delta_2 &= a_2 \\ &\vdots \\ \Delta_{\partial-1} &= a_{\partial-1} \end{aligned}$$

and

$$\Delta_\partial = \frac{\Delta}{2} - \sum_{i=1}^{\partial-1} a_i .$$

en $\sum_{i=1}^{\partial} \Delta_i = \sum_{i=1}^{\partial-1} a_i + \frac{\Delta}{2} - \sum_{i=1}^{\partial-1} a_i$

$$= \frac{\Delta}{2}$$

d further $a_i - \Delta_i = 0$, for $i = 1,2,\ldots, \partial-1$

d $a_\partial - \Delta_\partial = a_\partial - (\frac{\Delta}{2} - \sum_{i=1}^{\partial-1} a_i)$

$$= \sum_{i=1}^{\partial} a_i - \frac{\Delta}{2} > 0, \text{ in view of (7.2.16).}$$

This constructive procedure shows that it is possible to have a partition of $\frac{\Delta}{2}$ into $\Delta_1, \Delta_2, \ldots, \Delta_\partial$ such that $a_i - \Delta_i \geq 0$ with strict inequality at least once, provided $\sum_1^{\partial} a_i > \Delta/2$.

Now, let us consider the design D_2 where the lower bound of

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} \text{ (Lemma 7.2.1)}$$

is not attained. Let

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} (D_2) = \nu(\nu - 1) + f(1- f) + \Delta , \tag{7.2.17}$$

where $\Delta > 0$.

Let

$$\pi_{ij}(D_2) = \pi_i(D_2) + \pi_j(D_2) - 1 + \Delta_{ij}(D_2), \qquad (7.2.18)$$

where $\Delta_{ij}(D_2) \geq 0$ with strict inequality at least once (since, we know from the consistency relation (7.2.15) that $\pi_{ij}(D_2) \geq \pi_i(D_2) + \pi_j(D_2) - 1$ with strict inequality at least once (the case of equality for all (i,j) has been already dealt with in Theorem 7.2.2)). Let ∂ be the number of non-zero elements in the matrix

$$\left(\left(\begin{matrix} D2 \\ \Delta_{ij} \\ i < j \end{matrix} \right)\right).$$

__Theorem 7.2.3:__

$$\sum_{\substack{\text{over } \partial \\ \text{elements} \\ i<j}} \sum \Delta_{ij}(D_2) > \frac{\Delta}{2}.$$

__Proof:__

$$2 \sum_{\substack{\text{over } \partial \\ \text{elements} \\ i<j}} \sum \Delta_{ij}(D_2) = \sum^{N}_{i \neq j} \sum^{N} \Delta_{ij}(D_2)$$

$$= \sum^{N}_{i \neq j} \sum^{N} (\pi_{ij}(D_2) - \pi_i(D_2) - \pi_j(D_2) + 1), \quad \text{from (7.2.18)}$$

$$= \nu(\nu - 1) + f(1-f) + \Delta - 2\nu(N-1) + N(N-1)$$

$$= \Delta + f(1-f) + (\nu - N)(\nu - N + 1)$$

$$> \Delta, \text{ since } \nu \ll N,$$

so that

$$\sum_{\substack{\text{over } \partial \\ \text{elements} \\ i<j}} \sum \Delta_{ij} (D_2) > \frac{\Delta}{2} . \qquad (7.2.19)$$

Theorem 7.2.4: There exists a partition of $\frac{\Delta}{2}$ such that $\sum_{i=1}^{\partial} \Delta_i = \frac{\Delta}{2}$, with the property that $\underset{i<j}{\Delta_{ij}} (D_1)$ got by subtracting $\Delta_1, \Delta_2, \ldots, \Delta_\partial$ from the ∂ non-zero elements of $\Delta_{ij} (D_2)$ are non-negative with at least one positive.

Proof: Follows as a direct consequence of Theorem 7.2.3 and Lemma 7.2.2.

Note that the zero elements of $\underset{i<j}{\Delta_{ij}} (D_2)$ are unchanged while constructing $\underset{i<j}{\Delta_{ij}} (D_1)$.

Construct the matrix $\left(\left(\Delta_{ij} (D_1) \right)\right)$ with

$$\Delta_{ij} (D_1) = \Delta_{ji} (D_1), \quad i \neq j. \qquad (7.2.20)$$

Theorem 7.2.5: The matrix $\left(\left(\pi_{ij} (D_1) \right)\right)$, where

$$\pi_{ij} (D_1) = \pi_i + \pi_j - 1 + \Delta_{ij} (D_1)$$

satisfies the property $\wp$.

<u>Proof</u>: Using $\Delta_{ij}(D_1)$ of (7.2.20) construct the matrix

$$((\pi_{ij}(D_1))) = ((\pi_i(D_2) + \pi_j(D_2) - 1 + \Delta_{ij}(D_1)))$$

$$= ((\pi_i(D_1) + \pi_j(D_1) - 1 + \Delta_{ij}(D_1)))$$

$$= ((\pi_i + \pi_j - 1 + \Delta_{ij}(D_1))). \qquad (7.2.21)$$

Obviously,

$\pi_{ij}(D_1) \leq \pi_{ij}(D_2)$ with strict inequality at least once.

Also, $\pi_{ij}(D_1) > 0$ and $\pi_{ij}(D_1) = \pi_{ji}(D_1)$, by construction and

$\pi_{ij}(D_1) \geq \pi_i(D_1) + \pi_j(D_1) - 1$, by construction, with strict inequality at least once.

Thus $((\pi_{ij}(D_1)))$ satisfy the property $\mathcal{P}$, besides the internal consistency relations. The question of possibility of constructing a design D_1 with the same π_i's but with uniformly smaller (or equal) π_{ij}'s of a given design D_2, when $\sum\sum_{i \neq j} \pi_{ij}(D_2)$ is greater than its lower bound is answered in the affirmative and the theorems 7.2.3, 7.2.4 and 7.2.5 actually give a constructive proof.

7.3. An Illustrative example:

In this section to explain the algorithm given above, we consider a simple hypothetical example. Consider a population of size $N = 4$ and let the sampling design $D_2(S, P)$ be as follows:

sample s	probability P_s
1	.1
1,2	.1
1,2,3	.2
2,4	.1
1,2,4	.2
1,3,4	.1
3,4	.2
	$\sum_{s \in S} P_s = 1$.

We have corresponding to this design D_2,

$$\pi_1 (D_2) = .7$$

$$\pi_2 (D_2) = .6$$

$$\pi_3 (D_2) = .5$$

$$\pi_4 (D_2) = .6$$

and $\sum_{i=1}^{4} \pi_i (D_2) = \nu = 2.4$

and

$$\pi_{12}\ (D_2) = .5 = .7 + .6 - 1 + (.2)$$

$$\pi_{13}\ (D_2) = .3 = .7 + .5 - 1 + (.1)$$

$$\pi_{14}\ (D_2) = .3 = .7 + .6 - 1 + (0)$$

$$\pi_{23}\ (D_2) = .2 = .6 + .5 - 1 + (.1)$$

$$\pi_{24}\ (D_2) = .3 = .6 + .6 - 1 + (.1)$$

$$\pi_{34}\ (D_2) = .3 = .5 + .6 - 1 + (.2)$$

and $\sum_{i \neq j}^{4} \sum^{4} \pi_{ij}\ (D_2) = 3.8.$

Min. $\sum_{i \neq j}^{4} \sum^{4} \pi_{ij} \doteq (2.4)(1.4) - (.6)(.4) = 3.6$, so that

$$\Delta = \text{excess} = 0.2.$$

We now have

$$((\Delta_{ij}(D_2)))_{i<j} = \begin{pmatrix} * & .2 & .1 & 0 \\ & * & .1 & .1 \\ & & * & .2 \\ & & & * \end{pmatrix}$$

and $\partial = 5$ and $\frac{\Delta}{2} = .1.$

Let $\Delta_i = .02, \quad i = 1,2,\ldots, 5.$

en

$$((\Delta_{ij}(D_1)))_{i<j} = \begin{pmatrix} * & .18 & .08 & 0 \\ & * & .08 & .08 \\ & & * & .18 \\ & & & * \end{pmatrix}$$

o that

$$((\Delta_{ij}(D_1))) = \begin{pmatrix} * & .18 & .08 & 0 \\ .18 & * & .08 & .08 \\ .08 & .08 & * & .18 \\ 0 & .08 & .18 & * \end{pmatrix}$$

and consequently

$$((\pi_{ij}(D_1))) = \begin{pmatrix} * & .48 & .28 & .30 \\ .48 & * & .18 & .28 \\ .28 & .18 & * & .28 \\ .30 & .28 & .28 & * \end{pmatrix}$$

Also here we observe that

$$\sum_{i \neq j}^{4} \sum^{4} \pi_{ij}(D_1) = 3.6 = \text{lower bound, } \nu(\nu - 1) + f(1-f).$$

Thus $\pi_{ij}(D_1)$ satisfies the required properties.

We can see that $((\pi_{ij}(D_1)))$ is not unique, because for any onsistent partition of $\Delta/2$ we have a corresponding $(\pi_{ij}(D_1)))$, as for example in the above case with

$\Delta_1 = .01$, $\Delta_2 = .02 = \Delta_3$, $\Delta_4 = .01$ and $\Delta_5 = .04$, we have

$$((\pi^*_{ij}(D_1))) = \begin{pmatrix} * & .49 & .28 & .30 \\ .49 & * & .18 & .29 \\ .28 & .18 & * & .26 \\ .30 & .29 & .26 & * \end{pmatrix}.$$

This has still the property $\sum\limits_{i \neq j}^{4} \sum^{4} \pi^*_{ij} (D_1) = 3.6$, the lower bound.

<u>Remark</u>: We have an upper bound for $\sum\limits_{i \neq j}^{N} \sum^{N} \pi_{ij}$ given by (Hanurav [1966])

$$\sum_{i \neq j}^{N} \sum^{N} \pi_{ij} \leq N\,(\nu - 1). \tag{7.3.1}$$

In the above example this upper bound is equal to $4\,(2.4-1) = 5.6$ and we have seen that the lower bound of $\sum\limits_{i \neq j} \sum \pi_{ij}$ which is attained for the design D_1 constructed here, is equal to 3.6 and thus there will be considerable increase in the precision.

CHAPTER VIII

INADMISSIBILITY OF CUSTOMARY ESTIMATORS IN SAMPLING OVER TWO OCCASIONS

8.0 Summary

In this chapter we demonstrate the inadmissibility of customary estimators of population total in sampling over two occasions, by providing improved estimators under (i) equal probability sampling scheme and (ii) varying probability sampling scheme. Expressions for the gain in efficiency of the improved estimators are derived.

1 Introduction:

The idea of replacement of the sample for obtaining current estimates has been studied by Jessen [1942], Patterson [1950], Cochran [1963] and Yates [1960] among others. Estimates were given mostly under the scheme of simple random sampling of the units. However, Eckler [1961] has shown instances where the rotation policy could be used under probability proportional to size sampling schemes and Des Raj [1965] obtained the theory for the estimation of population parameters on the current occasion. We consider here both the schemes of sampling over two occasions: (i) when simple random sampling (srs) without replacement is used on both the occasions and (ii) when probability proportional to size (pps) sampling is used on both the occasions and show that the customary estimators of the population total are inadmissible by providing more efficient estimators.

We have already seen before that the criterion 'admissibility' helps us in eliminating bad estimators while searching for an optimum and from the theorems of Murthy [1957], Des Raj and Khamis [1958], Basu [1958] and Roy and Chakravarty [1960] it follows that for any sampling design, an unbiased estimator of the population mean which is either an ordered estimator or depends on repetition of a unit in the sample, s inadmissible. Pathak [1961] has modified the customary estimators in the case of uni-stage sampling by providing better estimators and extensions to two-stage sampling have been given by Ghosh [1963]. Here we consider the case of two-phase sampling and demonstrate the inadmissibility of customary estimators of population total.

8.2 Improved estimators for SRS scheme:

Under this scheme, a sample of size n is selected by simple random sampling without replacement from the population of N units on the first occasion. Then, on the second occasion, m units of the first sample are retained and $(n-m)$ units are selected independently by simple random sampling (without replacement) from the whole population. For estimating the total of the population on the second (current) occasion, the customary estimator (Cochran [1963]) of the population total is given by

$$t_2 = \phi\,(N\,\bar{y}_{2u}) + (1-\phi)\,N\,\bar{y}'_{2m} \qquad (8.2.1)$$

where $\bar{y}_{hu}$ is the mean of the unmatched portion on the occasion h ($h = 1,2$), $\bar{y}_{hm}$ is the mean of the matched portion on the occasion h $(h = 1,2)$, $\bar{y}_h$ is the mean of the whole population on the occasion h $(h = 1,2)$,

$$\bar{y}'_{2m} = \bar{y}_{2m} + b\,(\bar{y}_1 - \bar{y}_{1m}) \qquad (8.2.2).$$

is the regression estimate based on the matched portion and ϕ is determined such that the variance of t_2 is minimised.

Now, let the (n- m) units selected independently on the second occasion be represented by $s = (s_1, s_2)$, where $s_1 = (u_{11}, u_{12}, \ldots, u_{1m_2})$ denotes the sample units of s which come from the matched portion and $s_2 = (u_{21}, u_{22}, \ldots, u_{2\lambda_2})$ denotes the remaining units.

The following theorem now provides an estimator more precise than t_2 and thus proves the inadmissibility of t_2.

Theorem 8.2.1: Let

$$t_2^* = N\,\phi\,\left(\frac{m_2\bar{y}_{2m} + \lambda_2\,\bar{y}_{2\lambda_2}}{m_2 + \lambda_2}\right) + N(1 - \phi)\,\bar{y}'_{2m}\,,$$

where $\bar{y}_{2\lambda_2}$ is the mean of the units based on the sample s_2. Then $E(t_2^*) = E(t_2)$ and for any convex loss function t_2^* does not have greater expected loss than t_2.

Proof: We have

$$t_2 = \frac{N\,\phi}{\lambda_2 + m_2}\,(m_2\,\bar{y}_{2m_2} + \lambda_2\,\bar{y}_{2\lambda_2}) + N(1-\phi)\,\bar{y}'_{2m}\,,$$

where $\bar{y}_{2m_2}$ is the mean of the units based on the sample s_1.

Then

$$E(t_2 \mid s_2, s_m) = \frac{N\,\phi}{\lambda_2 + m_2}\left\{m_2 E(\bar{y}_{2m_2} \mid s_m) + \lambda_2 \bar{y}_{2\lambda_2}\right\} + N(1-\phi)\,\bar{y}'_{2m}$$

$$= \frac{N\,\phi}{\lambda_2 + m_2}\left\{m_2 \bar{y}_{2m} + \lambda_2 \bar{y}_{2\lambda_2}\right\} + N(1-\phi)\bar{y}'_{2m} = t_2^* \,. \tag{8.2.3}$$

Therefore,

$$E(t_2^*) = E(t_2), \text{ by } (8.2.3).$$

As a consequence of Jensen's inequality, it follows from (8.2.3) that t_2^* does not have greater expected loss than t_2.

<u>Corollary 8.2.1:</u> If squared error is the loss function, the gain in efficiency on using t_2^* is given by

$$E(t_2^* - t_2)^2 = N^2\,\phi^2\,S^2\,\frac{m-1}{N-1}\left(\frac{1}{n-m} - \frac{1}{N}\right).$$

Proof: $E(t_2^* - t_2)^2 = \frac{N^2 \phi^2}{(n-m)^2} E(m_2^2(\bar{y}_{2m_2} - \bar{y}_{2m})^2)$,

since $\ell_2 + m_2 = n - m$,

$$= \frac{N^2 \phi^2}{(n-m)^2} E(E(m_2^2(\bar{y}_{2m_2} - \bar{y}_{2m})^2 | m_2; s_m))$$

$$= \frac{N^2 \phi^2}{(n-m)^2} E(m_2^2 \frac{m-m_2}{m\, m_2} S_m^2)$$

$$= \frac{N^2 \phi^2 S^2}{(n-m)^2} E(m_2 (1 - \frac{m_2}{m}))$$

$$= \frac{N^2 \phi^2 S^2}{(n-m)^2} E(m_2 - \frac{1}{m} E(m_2^2)). \qquad (8.2.4)$$

We have

$$E(m_2) = (n-m)\frac{m}{N}$$

and

$$V(m_2) = (n-m)\frac{m}{N}\frac{N-m}{N}\frac{N-n+m}{N-1}. \qquad (8.2.5)$$

Using (8.2.5) in (8.2.4) follows that

$$E(t_2^* - t_2)^2 = \frac{N^2 \phi^2 S^2}{(n-m)^2} [(n-m)\frac{m}{N} - \frac{1}{m}(\frac{m^2(n-m)^2}{N^2} + \frac{(n-m)m(N-m)N-n+m)}{N^2 (N-1)})]$$

$$= \frac{N^2 \phi^2 S^2}{(n-m)^2} [(n-m)\frac{m}{N} - \frac{m(n-m)^2}{N^2} - \frac{(n-m)}{N^2}\frac{(N-m)}{(N-1)}(N-n+m)]$$

$$= \frac{N^2 \phi^2 S^2}{(n-m)^2} [\frac{m}{N}(n-m)^2(\frac{1}{n-m} - \frac{1}{N})(1 - \frac{N-m}{m(N-1)})]$$

$$= N^2 \phi^2 S^2 (\frac{1}{n-m} - \frac{1}{N}) \frac{m-1}{N-1}, \qquad (8.2.6)$$

on simplification.

Remark 8.2.1a:

We know that an expression for the variance of t_2 is given by

$$V(t_2) = N^2 S^2 (\frac{n-u\rho^2}{m\,n} - \frac{1}{N})(\frac{1}{u} - \frac{1}{N})/(\frac{n^2-u^2\rho^2}{mnu} - \frac{2}{N})$$

(8.2.7)

where ρ is the coefficient of correlation between the units over two occasions and $u = (n-m)$. Suppose we define the relative gain in efficiency of t_2^* over t_2 as

$$G = \frac{E(t_2^* - t_2)^2}{V(t_2^*)}.$$

Then, we can derive an expression for G, getting an expression for $V(t_2^*)$ using

$$V(t_2^*) = V(t_2) - E(t_2 - t_2^*)^2$$

and by virtue of (8.2.6) it follows that

$$G = \frac{\phi \frac{m-1}{N-1}}{1 - \phi \frac{m-1}{N-1}} \qquad (8.2.8)$$

where

$$\phi = \left(\frac{n - u\,\rho^2}{m\,n} - \frac{1}{N} \right) \Big/ \left(\frac{n^2 - u^2\,\rho^2}{m\,n\,u} - \frac{2}{N} \right). \qquad (8.2.9)$$

In practice, for large N,

$$\phi = \frac{(n - u\,\rho^2)u}{n^2 - u^2\,\rho^2} \qquad \text{(see Cochran [1963])}$$

and

$$\phi - \frac{u}{n} = \frac{u^2\,\rho^2(u - n)}{n(n^2 - u^2\,\rho^2)} < 0, \quad \text{since} \quad n - u\rho > 0.$$

Thus $\phi < \frac{u}{n}$ in practice and hence it follows that

$$G < \frac{u}{n} \cdot \frac{m-1}{N-1} \Big/ 1 - \frac{u}{n} \cdot \frac{m-1}{N-1} . \qquad (8.2.10)$$

(8.2.10) will be a rough upper bound for the relative gain in efficiency of t_2^* over t_2.

Remark 8.2.1b. If instead of simple random sampling without replacement we use simple random sampling with replacement scheme on both the occasions then we have

$$E(t_2^* - t_2)^2 = \frac{N^2 \phi^2}{(n-m)^2} \; E(m_2^2 \; \frac{\sigma_m^2}{m_2})$$

$$= \frac{N^2 \phi^2}{(n-m)^2} \; \frac{m-1}{m} \sigma^2 \; E(m_2)$$

$$= N^2 \phi^2 \sigma^2 \left(\frac{m-1}{(n-m)N}\right). \qquad (8.2.11)$$

8.3 Improved estimators for PPS scheme:

We now consider the other sampling scheme over two occasions where probability proportional to size (pps) sampling is used on both the occasions (Des Raj [1965]). In this scheme, a sample s_f of size n is selected by pps with replacement on the first occasion and on the second occasion a simple random sample of size m is retained from s_f and an independent sample of $(n-m)$ units is selected by pps with replacement from the whole population.

As an estimator of the population total on the second occasion, Des Raj [1965] considers the following estimator:

$$z_2 = \phi\, z_{2u} + (1 - \phi)\, z'_{2m} \qquad (8.3.1)$$

where

$$z_{hu} = \frac{1}{n - m} \Sigma \frac{y_{hi}}{p_i} \qquad (8.3.2)$$

is the estimate of the population total based on the unmatched sample, on occasion $h (h = 1,2)$,

$$z_{hm} = \frac{1}{m} \Sigma \frac{y_{hi}}{p_i} \qquad (8.3.3)$$

is the estimate of the population total based on the matched sample, on occasion h $(h = 1,2)$,

$$z_h = \frac{1}{n} \Sigma \frac{y_{hi}}{p_i} \qquad (8.3.4)$$

is the estimate of the population total based on the whole sample, on occasion h $(h = 1,2)$, and

$$z'_{2m} = (z_{2m} - z_{1m}) + z_{1n} \qquad (8.3.5)$$

is the difference estimate of the population total on the second occasion, based on the matched portion.

Let s_m denote the matched sample and represent the unmatched sample on the second occasion by $s = (s_1, s_2)$, where $s_1 = (u_{11}, u_{12}, \ldots, u_{1m_2})$ denotes the sample units which come from the matched portion and $s_2 = (u_{21}, u_{22}, \ldots, u_{2\lambda_2})$ denotes the remaining units. We then have the following theorem which demonstrates the inadmissibility of the customary estimator under this scheme.

<u>Theorem 8.3.1:</u>

Let
$$z_2^* = \frac{\phi}{n-m} \left[\sum_2 \frac{y_{2i}}{p_i} + \frac{m_2 \left(\sum_{i=1}^{m} y_{2i} \right)}{\sum_{i=1}^{m} p_i} \right] + (1-\phi) z_{2m}' ,$$

where $\sum_2$ denotes the summation over units in s_2 and $\sum_{i=1}^{m}$ denotes the summation over the matched portion. Then $E(z_2^*) = E(z_2)$ and for any convex loss function, z_2^* does not have greater expected loss than z_2.

<u>Proof:</u> We have

$$z_2 = \frac{\phi}{n-m} \left[\sum_2 \frac{y_{2i}}{p_i} + \sum_{s_1} \frac{y_{2i}}{p_i} \right] + (1-\phi) z_{2m}' \tag{8.3.6}$$

and

$$E(z_2 \mid s_2, s_m) = \frac{\phi}{n-m}\left[\sum_{2}\frac{y_{2i}}{p_i} + E\sum_{s_1}\frac{y_{2i}}{p_i}\Big|_{s_m}\right] + (1-\phi)z'_{2m}$$

$$= z_2^* . \tag{8.3.7}$$

Therefore,

$$E(z_2^*) = E(z_2). \tag{8.3.8}$$

That z_2^* does not have greater expected loss than z_2 for any convex loss function follows from Jensen's inequality.

Corollary 8.3.1: If squared error is the loss function then the gain in efficiency in using z_2^* is given by

$$E(z_2 - z_2^*)^2 = \frac{\phi^2(m-1)}{m(n-m)}\Big[\sum_1^N (Y_i - \sum_1^N Y_i P_i)^2$$

$$+ (m-1)\sum_1^N P_i^2\Big(\sum_1^N \frac{Y_i^2}{P_i} - Y^2\Big) - N\Big(\sum_1^N Y_i P_i\Big)^2 + Y^2 \sum_1^N P_i^2\Big].$$

Proof:

$$E(z_2 - z_2^*)^2 = \frac{\phi^2}{(n-m)^2}\,E\Big(\sum_1^{m_2}\Big(\frac{y_{2i}}{p_i} - \frac{\sum_1^m y_{2i}}{\sum_1^m p_i}\Big)\Big)^2$$

$$= \frac{\phi^2}{(n-m)^2} E\Big(E \frac{1}{(\sum_1^m p_i)^2} \Big(\sum_1^{m_2} \Big(\frac{y_{2i}}{p_i \big| \Sigma p_i} - \sum_1^m y_{2i}\Big)\Big)^2 \Big| m_2, s_m \Big)$$

$$= \frac{\phi^2}{(n-m)^2} E\Big(\frac{m_2}{(\Sigma p_i)^2} V\Big(\frac{y_{2i}}{p_i \big| \Sigma p_i}\Big)\Big)$$

$$= \frac{\phi^2}{(n-m)^2} E\Big(\frac{m_2}{(\Sigma p_i)^2} \frac{1}{2m^2} \sum_{i \neq j}^m \sum^m \Big(\frac{y_{2i}}{p_i \big| \Sigma p_i} - \frac{y_{2j}}{p_j \big| \Sigma p_j}\Big)^2\Big)$$

$$= \frac{\phi^2}{2(n-m)^2 m^2} E\Big((n-m)\Big(\sum_1^m p_i\Big) \sum_{i \neq j}^m \sum^m \Big(\frac{y_{2i}}{p_i} - \frac{y_{2j}}{p_j}\Big)^2\Big)$$

$$= \frac{\phi^2}{2(n-m)m^2} E\Big(\Big(\sum_1^m p_i\Big) \Big\{2(m-1) \sum_1^m \frac{y_i^2}{p_i} - 2\sum_{i \neq j}^m \sum^m \frac{y_i\, y_j}{p_i\, p_j}\Big\}\Big)$$

$$= \frac{\phi^2}{m^2(n-m)} \Big[(m-1)\Big\{m \sum_1^N Y_i^2 + m(m-1)\sum_1^N P_i^2 \sum_1^N \frac{Y_i^2}{P_i}\Big\}$$

$$- 2m(m-1)\sum_1^N Y_i P_i \sum_1^N Y_i - m(m-1)(m-2)\sum_1^N P_i^2 \Big(\sum_1^N Y_i\Big)^2\Big]$$

$$= \frac{\phi^2(m-1)}{m(n-m)} \Big[\sum_1^N \big(Y_i - \sum_1^N Y_i P_i\big)^2 + (m-1) \sum_1^N P_i^2 \Big(\sum_1^N \frac{Y_i^2}{P_i} - Y^2 \Big)$$

$$- N\Big(\sum_1^N Y_i P_i \Big)^2 + Y^2 \sum_1^N P_i^2 \Big]. \qquad (8.3.9)$$

Remark: When $P_i = \frac{1}{N}$ for all i, the above expression (8.3.9) reduces to

$$\frac{\phi^2(m-1)}{m(n-m)} \big[N\sigma^2 + (m-1) N\sigma^2 \big] = N^2 \phi^2 \sigma^2 \Big(\frac{m-1}{(n-m)N} \Big), \quad (8.3.10)$$

which is exactly the same as (8.2.11), the corresponding expression for the SRS with replacement case.

APPENDIX

A1. YATE'S POPULATION (n=2)

A1.1 Table of T_{ij}

i \ j	1	2	3	4
1	-	1.1111	0.8333	0.6667
2		-	0.6667	0.5556
3			-	0.4762
4				-

A1.2 Table of $T_{ij}^{(2)}$

i \ j	1	2	3	4
1	-	0.3704	0.2083	0.1333
2		-	0.1333	0.0926
3			-	0.0680
4				-

A1.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$
1	2.6111	0.7120
2	2.3333	0.5963
3	1.9762	0.4097
4	1.6984	0.2940

A2. KERALA (n=2)

A2.1. Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9
1	-	.4923	.5761	.6186	.5830	.6266	.5848	.5945	.6266
2		-	.4666	.4941	.4711	.4992	.4723	.4786	.4992
3			-	.5784	.5472	.5854	.5488	.5573	.5854
4				-	.5854	.6294	.5873	.5970	.6294
5					-	.5926	.5551	.5638	.5926
6						-	.5945	.6044	.6377
7							-	.5655	.5945
8								-	.6044
9									-

A2.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

j / i	1	2	3	4	5	6	7	8	9
1	-	.1431	.1959	.2259	.2007	.2318	.2019	.2087	.2318
2		-	.1285	.1441	.1310	.1471	.1317	.1352	.1471
3			-	.1975	.1768	.2024	.1778	.1834	.3024
4				-	.2024	.2339	.2036	.2104	.2339
5					-	.2073	.1819	.1877	.2073
6						-	.2087	.2157	.2401
7							-	.1888	.2087
8								-	.2157
9									-

A2.3 Table of T_i and $T_i^{(2)}$

i	T_i	T_i^2 (X 10^{-2})
1	4.7024	.1640
2	3.8732	.1108
3	4.4452	.1465
4	4.7194	.1652
5	4.4908	.1495
6	4.7697	.1687
7	4.5027	.1503
8	4.5654	.1546
9	4.7697	.1687

A3. ANDHRA PRADESH (n= 2)

A3.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9	10
1	-	.3903	.3701	.4286	.4244	.3497	.4180	.4322	.4984	.4540
2		-	.3743	.4343	.4300	.3535	.4234	.4380	.5061	.4604
3			-	.4094	.4056	.3368	.3997	.4127	.4726	.4326
4				-	.4769	.3846	.4688	.4868	.5724	.5147
5					-	.3812	.4638	.4814	.5649	.5086
6						-	.3760	.3875	.4399	.4050
7							-	.4732	.5536	.4994
8								-	.5789	.5199
9									-	.6187
10										-

contd.

Table A3.1 (contd.)

j / i	11	12	13	14	15	16	17	18	19	20
1	.4379	.4587	.4151	.5066	.5536	.5542	.4612	.4743	.5587	.4803
2	.4438	.4653	.4205	.5145	.5632	.5637	.4678	.4812	.5684	.4874
3	.4179	.4368	.3971	.4800	.5221	.5226	.4391	.4509	.5266	.4563
4	.4940	.5207	.4653	.5832	.6465	.6473	.5239	.5408	.6335	.5486
5	.4884	.5145	.4603	.5754	.6370	.6377	.5176	.5341	.6437	.5417
6	.3921	.4087	.3737	.4462	.4824	.4828	.4107	.4210	.4862	.4257
7	.4800	.5051	.4528	.5637	.6227	.6234	.5081	.5241	.6292	.5314
8	.4988	.5261	.4695	.5899	.6548	.6556	.5293	.5466	.6620	.5546
9	.5891	.6275	.5486	.7205	.8197	.8209	.6321	.6569	.8309	.6684
10	.5281	.5587	.4954	.6313	.7062	.7071	5624	.5820	.7145	.5910
11	-	.5345	.4762	.6005	.6679	.6687	.5378	.5557	.6753	.5639
12		-	.5010	.6405	.7177	.7186	.5696	.5897	.7262	.5990
13			-	.5586	.6164	.6171	.5039	.5196	.6227	.5268
14				-	.8420	.8433	.6453	.6712	.8538	.6832
15					-	.9824	.7237	.7564	.9968	.7718
16						-	.7246	.7575	.9986	.7729
17							-	.5938	.7324	.6032
18								-	.7660	.6258
19									-	.7817
20										-

A3.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

j	1	2	3	4	5	6	7	8	9	10
1	-	.0930	.0836	.1122	.1100	.0747	.1067	.1141	.1517	.1259
2		-	.0856	.1152	.1129	.0763	.1095	.1171	.1564	.1294
3			-	.1024	.1004	.0693	.0976	.1040	.1364	.1143
4				-	.1389	.0903	.1342	.1447	.2001	.1617
5					-	.0887	.1314	.1415	.1949	.1579
6						-	.0863	.0917	.1182	.1001
7							-	.1367	.1872	.1523
8								-	.2046	.1650
9									-	.2337
10										-

Table A3.2 (contd.)

i \ j	11	12	13	14	15	16	17	18	19	20
1	.1171	.1285	.1052	.1567	.1872	.1875	.1299	.1373	.1906	.1408
2	.1203	.1322	.1079	.1616	.1937	.1941	.1336	.1414	.1973	.1451
3	.1066	.1165	.0963	.1407	.1664	.1667	.1177	.1241	.1693	.1271
4	.1490	.1656	.1322	.2077	.2552	.2559	.1676	.1786	.2608	.1838
5	.1457	.1616	.1294	.2022	.2478	.2483	.1636	.1742	.2530	.1792
6	.0939	.1020	.0853	.1216	.1421	.1423	.1030	.1082	.1444	.1107
7	.1407	.1558	.1252	.1941	.2368	.2373	.1577	.1677	.2417	.1724
8	.1519	.1690	.1346	.2125	.2618	.2624	.1711	.1825	.2676	.1878
9	.2119	.2404	.1838	.3170	.4102	.4115	.2439	.2635	.4216	.2728
10	.1703	.1906	.1498	.2434	.3045	.3053	.1931	.2068	.3117	.2133
11	-	.1744	.1385	.2202	.2724	.2731	.1766	.1886	.2785	.1942
12		-	.1532	.2505	.3145	.3153	.1981	.2124	.3221	.2191
13			-	1905	.2320	.2325	.1550	.1648	.2368	.1694
14				-	.4329	.4342	.2542	.2751	.4452	.2850
15					-	.5893	.3198	.3494	.6067	.3637
16						-	.3206	.3504	.6089	.3647
17							-	.2153	.3276	.2222
18								-	.3583	.2391
19									-	.3731
20										-

A3.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (X 10^{-2})
1	8.6663	.2453
2	8.7860	.2523
3	8.2632	.2225
4	9.8005	.3156
5	9.6873	.3082
6	7.7438	.1949
7	9.5165	.2971
8	9.8979	.3221
9	11.7201	.4560
10	10.4901	.3629
11	10.0508	.3324
12	10.6190	.3722
13	9.4403	.2923
14	11.9498	.4745
15	13.2831	.5886
16	13.2990	.5900
17	10.6864	.3771
18	11.0477	.4038
19	13.4274	.6015
20	11.2137	.4164

A4. GUJARAT (n = 2)

A4.1 Table of T_{ij}

1	2	3	4	5	6	7	8	9	10
-	.6571	.9052	.6763	.8793	.6333	.8578	.7308	.7813	.5055
	-	.7079	.5597	.6920	.5300	.6786	.5965	.6298	.4374
		-	.7302	.9727	.6804	.9465	.7941	.8542	.5350
			-	.7133	.5424	.6991	.6123	.6475	.4458
				-	.6657	.9182	.7742	.8312	.5259
					-	.6533	.5769	.6080	.4267
						-	.7575	.8119	.5181
							-	.6972	.4689
								-	.4892
									-

contd.

A4.1 Table of T_{ij} (contd.)

j	11	12	13	14	15	16	17
1	.4433	.4560	.5816	.5558	.7614	.3911	1.5309
2	.3901	.3999	.4932	.4746	.6168	.3491	1.0404
3	.4659	.4799	.6210	.5917	.8305	.4086	1.8382
4	.3968	.4069	.5040	.4845	.6337	.3544	1.0895
5	.4589	.4728	.6087	.5805	.8087	.4032	1.7346
6	.3816	.3910	.4797	.4630	.5958	.3423	.9821
7	.4530	.4663	.5983	.5711	.7904	.3986	1.6528
8	.4149	.4260	.5336	.5118	.6813	.3688	1.2381
9	.4307	.4427	.5601	.5361	.7250	.3813	1.3906
10	.3311	.3382	.4026	.3901	.4813	.3011	.7054
11	-	.3092	.3621	.3520	.4246	.2779	.5900
12		-	.3706	.3600	.4363	.2828	.6127
13			-	.4338	.5498	.3265	.8629
14				-	.5267	.3183	.8074
15					-	.3765	1.3288
16						-	.5010
17							-

A4.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

j / i		2	3	4	5	6	7	8	9
1	-	.4247	.8060	.4500	.7607	.3946	.7239	.5254	.6006
2		-	.4929	.3082	.4710	.2763	.4530	.3501	.3902
3			-	.5246	.9308	.4554	.8812	.6204	.7178
4				-	.5006	.2894	.4808	.3689	.4124
5					-	.4359	.8295	.5896	.6796
6						-	.4198	.3274	.3636
7							-	.5644	.6485
8								-	.4782
9									-
⋮									

contd.

A1.2 Table of $T_{ij}^{(2)}$ (X 10^{-3}) (contd.)

j / i	10	11	12	13	14	15	16	17
1	.2514	.1933	.2046	.3327	.3039	.5704	.1505	2.3055
2	.1882	.1497	.1573	.2393	.2215	.3743	.1199	1.0649
3	.2816	.2135	.2260	.3793	.3444	.6785	.1642	3.3240
4	.1955	.1549	.1629	.2499	.2309	.3951	.1236	1.1677
5	.2720	.2072	.2197	.3645	.3315	.6433	.1599	2.9601
6	.1792	.1432	.1504	.2264	.2100	.3493	.1152	.9489
7	.2641	.2019	.2139	.3521	.3208	.6146	.1563	2.6876
8	.2163	.1693	.1786	.2801	.2577	.4566	.1338	1.5081
9	.2354	.1825	.1928	.3086	.2828	.5171	.1430	1.9023
10	-	.1079	.1125	.1594	.1497	.2279	.0892	.4895
11		-	.0940	.1290	.1219	.1774	.0760	.3424
12			-	.1351	.1275	.1872	.0787	.3693
13				-	.1852	.2973	.1049	.7325
14					-	.2729	.0996	.6413
15						-	.1394	.1737
16							-	.2469
17								-

A4.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (X 10^{-2})
1	11.3466	.8998
2	9.2530	.5682
3	12.3618	1.1041
4	9.4966	.6015
5	12.0397	1.0356
6	8.9512	.5285
7	11.7715	.9812
8	10.1830	.7025
9	10.8167	.8055
10	7.3021	.3420
11	6.4819	.2664
12	6.6510	.2811
13	8.2883	.4476
14	7.9563	.4102
15	10.5676	.7638
16	5.7815	.2101
17	17.9054	2.2428

A5. WEST BENGAL ($n = 2$)

A5.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9
1	-	1.2752	1.5504	1.2185	1.2543	.8058	1.0925	.3565	.5552
2		-	1.1056	.9258	.9463	.6665	.8512	.3263	.4853
3			-	1.0627	1.0898	.7346	.9656	.3418	.5205
4				-	.9147	.6506	.8255	.3224	.4769
5					-	.6607	.8418	.3249	.4823
6						-	.6129	.2840	.3973
7							-	.3129	.4563
8								-	.2450
9									-
⋮									

contd.

A5.1 Table of T_{ij} (contd.)

j	10	11	12	13	14	15	16
1	.8467	.8495	.6612	1.1483	.9856	.4591	1.0764
2	.6942	.6961	.5644	.8847	.7849	.4103	.8414
3	.7684	.7708	.6125	1.0089	.8811	.4351	.9530
4	.6770	.6788	.5530	.8570	.7630	.4042	.8163
5	.6879	.6898	.5602	.8745	.7769	.4081	.8322
6	.5270	.5282	.4487	.6301	.5777	.3455	.6078
7	.6362	.6379	.5255	.7929	.7116	.3893	.7578
8	.2889	.2892	.2636	.3173	.3035	.2243	.3116
9	.4069	.4076	.3586	.4657	.4365	.2895	.4534
10	-	.5454	.4611	.6548	.5985	.3528	.6307
11		-	.4619	.6565	.5999	.3533	.6523
12			-	.5380	.4994	.3159	.5217
13				-	.7349	.3962	.7842
14					-	.3748	.7048
15						-	.3872
16							-

A5.2 Table of $T_{ij}^{(2)}$ ($X\ 10^{-3}$)

i \ j	1	2	3	4	5	6	7	8	9
1	-	.9275	1.3708	.8468	.8972	.3703	.6807	.0725	.1758
2		-	.6971	.4888	.5107	.2533	.4132	.0607	.1343
3			-	.6441	.6773	.3077	.5317	.0666	.1545
4				-	.4771	.2414	.3887	.0593	.1297
5					-	.2489	.4041	.0602	.1326
6						-	.2142	.0460	.0900
7							-	.0558	.1187
8								-	.0342
9									-

contd.

A5.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

j	10	11	12	13	14	15	16
1	.4088	.4116	.2493	.7520	.5540	.1202	.6608
2	.2748	.2763	.1816	.4464	.3513	.0960	.4037
3	.3367	.3388	.2139	.5805	.4428	.1080	.5179
4	.2614	.2628	.1744	.4189	.3320	.0932	.3800
5	.2699	.2714	.1790	.4362	.3442	.0950	.3950
6	.1584	.1591	.1148	.2264	.1904	.0681	.2107
7	.2309	.2320	.1575	.3584	.2888	.0864	.3275
8	.0476	.0477	.0396	.0574	.0525	.0287	.0554
9	.0944	.0947	.0733	.1237	.1087	.0478	.1173
10	-	.1696	.1212	.2445	.2043	.0710	.2269
11		-	.1217	.2458	.2052	.0712	.2280
12			-	.1651	.1422	.0569	.1552
13				-	.3080	.0895	.3507
14					-	.0801	.2833
15						-	.0855
16							-

A5.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (X 10^{-2})
1	14.1354	.8498
2	11.4581	.5516
3	12.8006	.6968
4	11.1467	.5199
5	11.3443	.5399
6	8.4773	.2900
7	10.4096	.4489
8	4.5121	.0784
9	6.4370	.1630
10	8.7765	.3120
11	8.7972	.3136
12	7.3456	.2146
13	10.7438	.4803
14	9.7333	.3888
15	5.5456	.1198
16	10.3109	.4398

A6. YATE'S POPULATION (n = 3)

A6.1 Table of T_{ij}

i \ j	1	2	3	4
1	-	1.0317	.9722	.8929
2		-	.9259	.8466
3			-	.7870
4				-

A6.2 Table of $T_{ij}^{(2)}$

i \ j	1	2	3	4
1	-	.1606	.1447	.1201
2		-	.1337	.1092
3			-	.0932
4				-

A6.3 Table of T_i and $T_{ij}^{(2)}$

i	T_i	$T_i^{(2)}$
1	1.4484	.2127
2	1.4021	.2018
3	1.3426	.1858
4	1.2632	.1613

A7. KERALA (n = 3)

A7.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9
1	-	.6938	.7622	.7931	.7675	.7986	.7688	.7759	.7986
2		-	.6713	.6952	.6754	.6995	.6764	.6819	.6995
3			-	.7640	.7401	.7691	.7414	.7480	.7691
4				-	.7692	.8006	.7706	.7777	.8006
5					-	.7745	.7463	.7530	.7745
6						-	.7758	.7831	.8062
7							-	.7543	.7758
8								-	.7831
9									-

A7.2 Table of $T_{ij}^{(2)} (X\ 10^{-3})$

i \ j	1	2	3	4	5	6	7	8	9
1	-	.1422	.1719	.1862	.1743	.1868	.1749	.1782	.1888
2		-	.1331	.1438	.1347	.1445	.1351	.1373	.1445
3			-	.1727	.1621	.1751	.1627	.1656	.1751
4				-	.1751	.1897	.1757	.1790	.1897
5					-	.1775	.1648	.1678	.1775
6						-	.1781	.1815	.1923
7							-	.1684	.1781
8								-	.1815
9									-

A7.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ ($\times 10^{-3}$)
1	3.0793	.7026
2	2.7465	.5571
3	2.9827	.6591
4	3.0855	.7054
5	3.0002	.6670
6	3.1037	.7137
7	3.0047	.6690
8	3.0285	.6797
9	3.1037	.7137

A8. ANDHRA PRADESH (n = 3)

A8.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9	10
1	-	.5792	.5574	.6185	.6143	.5348	.6078	.6221	.6845	.6433
2		-	.5620	.6242	.6199	.5390	.6133	.6278	.6912	.6495
3			-	.5990	.5950	.5202	.5890	.6023	.6604	.6222
4				-	.6652	.5729	.6577	.6744	.7483	.6995
5					-	.5693	.6529	.6694	.7421	.6941
6						-	.5637	.5759	.6287	.5940
7							-	.6618	.7327	.6859
8								-	.7536	.7041
9									-	.7851
10										-

contd.

Table A8.1 (contd.)

i \ j	11	12	13	14	15	16	17	18	19	20
1	.6277	.6478	.6049	.6915	.7313	.7318	.6501	.6624	.7355	.6679
2	.6335	.6540	.6104	.6986	.7393	.7398	.6564	.6689	.7435	.6745
3	.6076	.6264	.5863	.6671	.7041	.7045	.6286	.6400	.7080	.6451
4	.6810	.7048	.6543	.7570	.8050	.8056	.7076	.7222	.8100	.7287
5	.6759	.6993	.6495	.7506	.7976	.7984	.7001	.7164	.8028	.7229
6	.5807	.5979	.5612	.6348	.6682	.6686	.5998	.6102	.6717	.6149
7	.6681	.6910	.6423	.7410	.7870	.7875	.6936	.7076	.7917	.7139
8	.6854	.7095	.6583	.7624	.8111	.8117	.7123	.7271	.8162	.7337
9	.7618	.7918	.7286	.8584	.9206	.9213	.7953	.8138	.9271	.8222
10	.7113	.7374	.6822	.7947	.8478	.8484	.7404	.7564	.8533	.7636
11	-	.7168	.6646	.7708	.8207	.8212	.7197	.7348	.8259	.7410
12		-	.6872	.8016	.8556	.8562	.7464	.7626	.8613	.7699
13			-	.7367	.7821	.7826	.6898	.7037	.7868	.7099
14				-	.9338	.9345	.8051	.8241	.9405	.8327
15					-	1.0085	.8597	.8813	1.0154	.8911
16						-	.8603	.8820	1.0163	.8918
17							-	.7659	.8654	.7733
18								-	.8873	.7907
19									-	.8973
20										-

A8.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

i \ j	1	2	3	4	5	6	7	8	9	10
1	-	.1031	.0954	.1177	.1161	.0877	.1136	.1191	.1443	.1274
2		-	.0970	.1199	.1182	.0891	.1157	.1213	.1473	.1299
3			-	.1103	.1088	.0829	.1066	.1115	.1343	.1191
4				-	.1364	.1008	.1333	.1403	.1731	.1510
5					-	.0995	.1313	.1382	.1702	.1487
6						-	.0976	.1019	.1215	.1084
7							-	.1350	.1658	.1451
8								-	.1756	.1531
9									-	.1908
10										-

contd.

Table A8.2 (contd.)

j	11	12	13	14	15	16	17	18	19	20
1	.1212	.1292	.1125	.1474	.1649	.1651	.1302	.1352	.1667	.1374
2	.1236	.1318	.1146	.1505	.1686	.1686	.1327	.1379	.1705	.1402
3	.1135	.1207	.1056	.1370	.1527	.1528	.1216	.1261	.1543	.1281
4	.1430	.1534	.1319	.1771	.2004	.2007	.1546	.1611	.2029	.1640
5	.1409	.1509	.1300	.1741	.1968	.1970	.1521	.1585	.1992	.1614
6	.1036	.1098	.0967	.1239	.1373	.1374	.1106	.1144	.1387	.1162
7	.1376	.1473	.1271	.1696	.1914	.1916	.1484	.1546	.1937	.1573
8	.1449	.1554	.1335	.1797	.2035	.2038	.1567	.1633	.2060	.1663
9	.1795	.1941	.1639	.2286	.2631	.2635	.1959	.2052	.2668	.2095
10	.1562	.1680	.1435	.1955	.2226	.2229	.1694	.1769	.2255	.1804
11	-	.1587	.1361	.1838	.2084	.2087	.1600	.1668	.2110	.1700
12		-	.1457	.1990	.2268	.2271	.1722	.1799	.2298	.1834
13			-	.1676	.1889	.1892	.1468	.1528	.1912	.1556
14				-	.2707	.2711	.2008	.2105	.2746	.2149
15					-	.3158	.2290	.2408	.3201	.2463
16						-	.2293	.2412	.3206	.2467
17							-	.1815	.2320	.1850
18								-	.2441	.1936
19									-	.2497
20										-

A8.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (X 10^{-2})
1	6.1063	.1217
2	6.1626	.1240
3	5.9126	.1139
4	6.6179	.1436
5	6.5689	.1414
6	5.6534	.1039
7	6.4942	.1381
8	6.6597	.1455
9	7.3838	.1796
10	6.9066	.1567
11	6.7245	.1484
12	6.9588	.1592
13	6.4606	.1367
14	7.4679	.1838
15	7.9303	.2074
16	7.9355	.2077
17	6.9859	.1604
18	7.1286	.1672
19	7.9780	.2099
20	7.1929	.1703

A9. GUJARAT (n= 3)

A9.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9
1	-	.8374	1.0089	.8525	.9933	.8183	.9798	.8936	.9294
2		-	.8757	.7566	.8640	.7298	.8539	.7884	.8158
3			-	.8923	1.0484	.8547	1.0334	.9375	.9772
4				-	.8801	.7411	.8697	.8017	.8301
5					-	.8436	1.0169	.9240	.9625
6						-	.8340	.7715	.7977
7							-	.9125	.9500
8								-	.8688
9									-
.									
.									
.									
17									

contd.

A9.1 Table of T_{ij} (contd.)

i \ j	10	11	12	13	14	15	16	17
1	.7054	.6437	.6567	.7746	.7519	.9155	.5880	1.2682
2	.6392	.5884	.5992	.6951	.6769	.8052	.5417	1.0668
3	.7320	.6657	.6797	.8070	.7823	.9618	.6062	1.3565
4	.6478	.5956	.6067	.7053	.6866	.8192	.5478	1.0914
5	.7239	.6590	.6727	.7972	.7731	.9476	.6007	1.3292
6	.6282	.5791	.5896	.6821	.6645	.7876	.5338	1.0361
7	.7169	.6533	.6667	.7886	.7651	.9354	.5959	1.3060
8	.6707	.6148	.6267	.7327	.7124	.8568	.5639	1.1591
9	.6902	.6311	.6436	.7562	.7346	.8896	.5775	1.2195
10	-	.5213	.5297	.6026	.5890	.6827	.4845	.8616
11		-	.4947	.5574	.5458	.6248	.4552	.7713
12			-	.5671	.5550	.6371	.4615	.7901
13				-	.6359	.7472	.5154	.9672
14					-	.7260	.5055	.9320
15						-	.5723	1.1960
16							-	.6926
17								-

A9.2 Table of $T_{ij}^{(2)}$ ($X\ 10^{-3}$)

j / i	1	2	3	4	5	6	7	8	9	10
1	-	.3596	.5332	.3734	.5157	.3426	.5011	.4122	.4480	.2514
2		-	.3946	.2912	.3837	.2702	3745	.3172	.3406	.2054
3			-	.4106	.5789	.3750	5613	.4559	.4980	.2712
4				-	.3990	.2789	.3892	.3285	.3532	.2111
5					-	.3649	.5424	.4423	.4823	.2651
6						-	.3564	.3032	.3249	.1982
7							-	.4308	.4691	.2599
8								-	.3887	.2267
9									-	.2404
10										-
.										
.										
.										
.										
17										-

contd.

A9.2 Table of $T_{ij}^{(2)}$ (X 10^{-3}) (contd.)

i \ j	11	12	13	14	15	16	17
1	.2081	.2169	.3055	.2871	.4340	.1727	.8353
2	.1732	.1798	.2443	.2312	.3315	.1462	5832
3	.2228	.2325	.3324	.3115	.4814	.1837	.9585
4	.1776	.1845	.2517	.2381	.3436	.1496	.6115
5	.2183	.2277	.3241	.3039	.4666	.1803	.9199
6	.1677	.1740	.2349	.2226	.3164	.1419	5488
7	.2144	.2236	.3170	.2975	.4541	.1774	.8873
8	.1895	.1971	.2723	.2569	.3775	.1586	.6931
9	.1998	.2081	.2907	.2736	.4084	.1665	.7702
10	.1353	.1398	.1820	.1737	.2351	.1166	.3745
11	-	.1216	.1551	.1486	.1958	.1027	.2982
12		-	.1607	.1538	.2038	.1056	.3133
13			-	.2033	.2835	.1322	.4757
14				-	.2671	.1270	.4406
15					-	.1635	.7398
16						-	.2392
17							-

A9.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (X 10^{-2})
1	6.8087	.3098
2	6.0670	.2413
3	7.1097	.3401
4	6.1622	.2496
5	7.0181	.3308
6	5.9458	.2310
7	6.9391	.3228
8	6.4175	.2725
9	6.6368	.2931
10	5.2128	.1743
11	4.8007	.1464
12	4.8884	.1521
13	5.6658	.2083
14	5.5183	.1968
15	6.5524	.2851
16	4.4213	.1232
17	8.5218	.4845

A10. WEST BENGAL (n = 3)

A 10.1 Table of T_{ij}

i \ j	1	2	3	4	5	6	7	8	9
1	-	1.2072	1.3097	1.1823	1.1981	.9523	1.1220	.5503	.7546
2		-	1.1322	1.0347	1.0469	.8542	.9881	.5164	.6920
3			-	1.1103	1.1243	.9047	1.0566	.5342	.7245
4				-	1.0281	.8417	.9713	.5120	.6839
5					-	.8497	.9820	.5148	.6891
6						-	.8110	.4651	.6018
7							-	.5007	.6637
8								-	.4139
9									-
.									
.									
.									
10									

contd.

A.10.1 Table of T_{ij} (contd.)

i \ j	10	11	12	13	14	15	16
1	.9796	.9815	.8450	1.1496	1.0645	.6622	1.1137
2	.8761	.8776	.7671	1.0095	.9432	.6136	.9816
3	.9293	.9310	.8073	1.0812	1.0054	.6389	1.0493
4	.8630	.8645	.7571	.9920	.9280	.6073	.9651
5	.8713	.8728	.7635	1.0032	.9377	.6113	.9757
6	.7348	.7359	.6573	.8252	.7808	.5420	.8067
7	.8306	.8320	.7322	.9491	.8905	.5913	.9246
8	.4713	.4717	.4389	.5059	.4893	.3852	.4991
9	.6124	.6131	.5581	.6730	.6436	.4734	.6608
10	-	.7519	.6700	.8456	.7990	.5505	.8261
11		-	.6708	.8469	.8003	.5511	.8274
12			-	.7437	.7077	.5065	.7288
13				-	.9078	.5987	.9432
14					-	.5754	.8953
15						-	.5891
16							-

A 10.2 Table of $T_{ij}^{(2)}$ (X 10^{-3})

i \ j	-	2	3	4	5	6	7	8	9
1	-	.4424	.5218	.4238	.4356	.2701	.3801	.0873	.1668
2		-	.3895	.3234	.3314	.2168	.2938	.0769	.1401
3			-	.3739	.3838	.2437	.3372	.0822	.1537
4				-	.3191	.2104	.2836	.0755	.1368
5					-	.2145	.2900	.0764	.1389
6						-	.1949	.0622	.1054
7							-	.0722	.1287
8								-	.0491
9									-
.									
.									
.									
16									

contd.

A 10.2 Table of $T_{ij}^{(2)}$ (X 10^{-3}) (contd.)

j / i	10	11	12	13	14	15	16
1	.2865	.2877	.2107	.3999	.3407	.1274	.3743
2	.2286	.2294	.1734	.3072	.2666	.1094	.2898
3	.2578	.2588	.1923	.3538	.3039	.1187	.3323
4	.2216	.2224	.1688	.2965	.2578	.1071	.2798
5	.2260	.2268	.1717	.3032	.2634	.1056	.2861
6	.1590	.1595	.1263	.2019	.1802	.0850	.1927
7	.2048	.2055	.1577	.2703	.2367	.1015	.2559
8	.0639	.0640	.0553	.0737	.0689	.0425	.0717
9	.1092	.1095	.0903	.1324	.1209	.0646	.1276
10	-	.1667	.1314	.2125	.1890	.0878	.2025
11		-	11317	.2132	.1896	.0880	.2032
12			-	.1628	.1471	.0741	.1561
13				-	.2463	.1041	.2668
14					-	.0960	.2338
15						-	.1007
16							-

A 10.3 Table of T_i and $T_i^{(2)}$

i	T_i	$T_i^{(2)}$ (x 10^{-2})
1	7.5364	.2378
2	6.7702	.1909
3	7.1695	.2152
4	6.6706	.1850
5	6.7342	.1888
6	5.6815	.1311
7	6.4228	.1706
8	3.6343	.0511
9	4.7289	.0887
10	5.8057	.1374
11	5.8142	.1378
12	5.1770	.1075
13	6.5374	.1772
14	6.1792	.1571
15	4.2482	.0708
16	6.3882	.1687

B I B L I O G R A P H Y

BASU, D. (1958): On sampling with and without replacement, Sankhya, 20, 287-294.

CALLEBAUT, D.K. (1965): Generalization of the Cauchy-Schwartz inequality, Jour. Math. Anal. Appl. 12, 491-494.

COCHRAN, W.G. (1942): Sampling theory when the sampling units are of unequal sizes, Jour. Amer. Stat. Assn., 37, 199-212.

COCHRAN, W.G. (1946): Relative accuracy of systematic and stratified random samples for a certain class of populations, Ann. Math. Stat., 17, 164-177.

COCHRAN, W.G. (1953, 1963): Sampling Techniques, Wiley, New York, Ed. 1 and Ed. 2.

DES RAJ (1955): On sampling with probabilities propoftionate to size, Ganita, 6 , 175-182.

DES RAJ (1956) Some estimators in sampling with varying probabilities without replacement, Jour. Amer. Stat. Assn., 51, 269-284.

DES RAJ (1956): A note on determination of optimum probabilities in sampling without replacement, Sankhya, 17, 197-200.

DES RAJ (1965): On sampling over two occasions with probability proportionate to size, Ann. Math. Stat. 36, 327-330.

DES RAJ AND KHAMIS, S.H. (1958): Some remarks on sampling with replacement, Ann. Math. Stat., 39, 550-557.

DURBIN, J. (1953): Some results in sampling theory when the units are selected with unequal probabilities, Jour. Roy. Stat. Soc., Ser B, 15, 262-269.

DURBIN, J. (1959): A note on the application of Quenouille's method of bias reduction to the estimation of ratios., Biometrika, 46, 477-480.

ECKLER, A. R. (1961): The continuous population and labour force survey in the United States. Family living studies, Int. Lab. Office, Geneva.

FAIRFIELD SMITH, H. (1938): An empirical law govering soil heterogeneity, Jour. Agr. Sci., 28, 1-23.

FELLIGI, I.M. (1963): Sampling with varying probabilities without replacement: Rotating and non-rotating samples., Jour. Amer. Stat. Assn. 58, 183-201.

GHOSH, S. P. (1963): Estimating the mean by two-stage sampling with replacement, Cal. Stat. Assoc. Bull., 12, 97 - 103.

GODAMBE, V.P. (1955): A unified theory of sampling from finite populations, Jour. Roy. Stat. Soc., B, 17, 269-278.

GODAMBE, V.P. AND JOSHI, V.M. (1965): Admissibility and Bayes estimation in sampling finite populations-I, Ann. Math. Stat., 36, 1707-1722.

GOODMAN, R. AND KISH, L. (1950): Controlled selection - a technique in probability sampling, Jour. Amer. Stat. Assn., 45, 350-372.

GRUNDY, P.M. (1954): A method of sampling with probability exactly proportional to size, Jour. Roy. Stat. Soc., Ser. B, 16, 236-238.

HÁJEK, J. (1958): Some contributions to the theory of probability sampling, Bull. Int. Stat. Inst. 36 (3), 127-134.

HÁJEK, J. (1959): Optimum strategy and other problems in probability sampling, Casipis Pest. Math. 84, 387-428.

HENSEN, M. H. AND HURWITZ, W.N. (1943): On the theory of sampling from finite populations, Ann. Math. Stat. 14, 333-362.

HANSEN, M.H., HURWITZ, W.N. AND MADOW, W.G. (1953): Sample survey methods and theory, Wiley, New York.

HANURAV, T.V. (1962a): An existence theorem in sampling, Sankhya, A, 24, 227-230.

HANURAV, T.V. (1962b): Some sampling schemes in probability sampling, Sankhya A, 24, 421-428.

HANURAV, T.V. (1962c): On Horvitz-Thompson's estimator, Sankhya, A, 24, 429-436.

HANURAV, T.V. (1965): Optimum sampling strategies and some related problems, Ph.D. Thesis submitted to the Indian Statistical Institute.

HANURAV, T.V. (1966):Some aspects of unified sampling theory, Sankhya, A,28, 175-204.

HANURAV, T.V. (1966): An optimum πPS sampling, (an abstract) Ann. Math. Stat., 37, 1859.

HANURAV, T.V. (1967): Optimum utilization of auxiliary information-πPS sampling of two units from a stratum, Jour. Roy. Stat. Soc., B, 29.

HARTLEY, H.O. AND RAO, J.N.K.(1962): Sampling with unequal probabilities and without replacement, Ann. Math. Stat., 33, 350-374.

HARTLEY, H.O. AND ROSS, A. (1954): Unbiased ratio estimators, Nature, 174, 270-271.

HORVITZ, D.G. AND THOMPSON, D.J. (1952): A generalization of sampling without replacement from a finite universe, Jour. Amer. Stat. Assn., 47, 663-685.

JESSEN, R. J. (1942): Statistical investigation of a sample survey for obtaining farm facts, Iowa Agr. Exp. Sta. Res. Bull. 304.

JOSHI, V. M. (1965a): Admissibility and Bayes estimation in sampling finite populations-II, Ann. Math. Stat., 36, 1723-1729.

JOSHI, V.M. (1965b): Admissibility and Bayes estimation in sampling finite populations-III, Ann. Math. Stat., 36, 1730-1742.

KOOP, J.C. (1951): A note on the bias of the ratio estimate, Bull. Int. Stat. Inst., 33, 141-146.

KOOP, J.C. (1963): On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite universes, I II, III, Metrika, 7, 81-114 and 165-204.

KOOP, J.C. (1966): Letter to the editor, The American Statistician, 20, 48.

LAHIRI, D.B. (1951): A method of sample selection providing unbiased ratio estimates, Bull. Int. Stat. Inst., 33(2), 133-140.

MADOW, W.G. AND MADOW, L.H. (1944): On the theory of systematic sampling-I, Ann. Math. Stat., 15, 1-24.

MAHALANOBIS, P.C. (1944) On large scale sample surveys, Roy. Soc. Phil. Trans., B 231, 329-451.

MAHALANOBIS, P.C. (1946): Recent experiments in statistical sampling in the Indian Statistical Institute Jour. Roy. Stat. Soc., 109, 325-378.

MIDZUNO, H. (1952): On the sampling system with probability proportional to the sum of sizes, Ann. Inst. Stat. Math., 3, 99-108.

MURTHY, M.N. (1957): Ordered and unordered estimators in sampling without replacement, Sankhya, 18, 378-390.

MURTHY, M.N. (1963): Generalized unbiased estimation in sampling from finite populations, Sankhya Ser. B, 25, 245-262.

MURTHY, M.N. AND NANJAMMA, N.S. (1959): Almost unbiased ratio estimates based on interpenetrating sub-sample estimates, Sankhya, 21, 381-392.

MURTHY, M.N., NANJAMMA, N.S. AND SETHI, V.K. (1959): Some sampling systems providing unbiased ratio estimators, Sankhya, 21, 299- 304.

NEYMAN, J. (1934): On two different aspects of the representative method, the method of stratified sampling and the method of purposive selection, Jour. Roy. Stat. Soc., 97, 558-606.

NIETO DE PASCUAL, J. (1961): Unbiased ratio estimators in stratified sampling, Jour. Amer. Stat. Assn. 56, 70-87.

PATHAK, P.K. (1961): Some contributions to the theory of sampling, Ph.D. Thesis submitted to the Indian Statistical Institute.

PATHAK, P.K. (1964): Sufficiency in sampling theory, Ann. Math. Stat., 35, 795-808.

PATTERSON, H.D. (1950): Sampling on successive occasions with partial replacement of units, Jour. Roy. Stat. Soc., B, 12, 241-255.

QUENOUILLE, M.H. (1956): Notes on bias in estimation, Biometrika, 43, 353-360.

RAO, J.N.K. (1965): A note on estimation of ratios by Quenouille's method, Biometrika, 52, 647-649.

RAO, J.N.K., HARTLEY, H.O. AND COCHRAN, W.G. (1962): A simple procedure of unequal probability sampling without replacement, Jour. Roy. Stat. Soc. Ser. B, 24, 482-491.

RAO, T.J. (1966a): On certain unbiased ratio estimators, Ann. Inst. Stat. Math., 18, 117-121.

RAO, T.J. (1966b): On the variance of the ratio estimator for Midzuno-Sen sampling scheme, Metrika, 10, 89-91.

RAO, T.J. (1967): On the choice of a strategy for ratio method of estimation, Jour. Roy. Stat. Soc. B, 29.

RAO, T.J. (1967): On the allocation of sample size in stratified sampling (an abstract) Sankhya, 29, 120.

RAO, T.J. (with PATHAK, P.K.) (1967): Inadmissibility of customary estimators in sampling over two occasions, Sankhya, Ser. A, 29, 49-54.

ROY, J. AND CHAKRAVARTY, I.M. (1960): Estimating the mean of a finite population, Ann. Math. Stat., 31 392-398.

ROY CHAUDHURY, D.K. (1956): Sampling with varying probabilities - thesis submitted to the Indian Statistical Institute for the award of AISI.

SEN, A.R. (1952): Present status of probability sampling and its use in estimation of farm characteristics (abstract), Econometrica, 20, 103.

SEN, A.R. (1953): On the estimate of variance in sampling with varying probabilities, Jour. Ind. Soc. Agr. Stat. 5, 119-127.

SEN, A.R. (1955): On the selection of n primary sampling units from a stratum structure (n > 2), Ann. Math. Stat., 26, 744-751.

SENG, YON POH (1951): Historical survey of the development of sampling theories and practice, Jour. Roy. Stat. Soc., Ser. A, 114, 214-231.

SINGH, M.P. (1965): On the estimation of ratio and product of population parameters, Sankhya, B, 27, 321-328.

STEPHAN, F.F. (1948): History of the use of modern sampling procedures, Jour. Amer. Stat. Assn., 43, 12-39.

STUART, A. (1964): Multistage sampling with preliminary random stratification of first stage units, Rev. Int. Stat. Inst., 32, 193-201.

SUKHATME, P.V. (1954): Sampling theory of surveys with applications, Ind. Soc. of Agr. Stat. and Iowa State Coll. Press.

VIJAYAN, K. (1966): On Horvitz-Thompson and Des Raj estimators, Sankhya, Ser. A, 28, 87-92.

VIJAYAN, K. (1967): On an exact πPS sampling (an abstract), Sankhya, A, 29, 113.

YATES, F. (1946): A review of recent statistical developments in sampling and sampling surveys, Jour. Roy. Stat. Soc., 109, 12-43.

YATES, F. (1960): Sampling Methods for Censuses and Surveys, 3rd edn. Griffin, London.

YATES, F. AND GRUNDY, P.M. (1953): Selection without replacement from within strata with probability proportional to size, Jour. Roy. Stat. Soc., B, 15, 253-261.
