# SOME ASPECTS OF ESTIMATION IN SAMPLING FROM FINITE POPULATIONS



By

M. P. SINGH



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

ACKNOWLEDGEMENTS

I have great pleasure in expressing my deep gratitude to Dr. M. N. Murthy, Head of the Sampling Division of the National Sample Survey Department for his guidance and valuable suggestions in the preparation of this thesis. I am also grateful to Dr. J. N. K. Rao, Dr. D. Basu and Prof D. B. Lahiri for their encouragement and helpful discussions I have had with them and to Prof. C. R. Rao, F.R.S., Director of the Research and Training School of Indian Statistical Institute for the keen interest he has taken in my research work.

Thanks are also due to Shri T. J. Rao and Shri A.S.Roy for their comments on the draft of a part of the thesis and to Shri C. M. Das for his elegant typing of the manuscript.

M. P. SINGH

# C O N T E N T S

# CHAPTER I

## INTRODUCTION

Interest in the use of sampling methods for obtaining the statistical data was discernible towards the end of the last century itself, but sampling then had a somewhat different meaning from to-day, the primary difference being the absence of mechanism of randomisation in the sample selection and the probabilistic interpretation of the data collected. The outstanding contributions of Mahalanobis, Neyman and Sukhatme during the thirties marked a turning point in the history of the sampling theory and opened-up new avenues for valuable researches in the theory and philosophy of the sample surveys. It was not until the advent of large-scale surveys carried out under the guidance of Professor Mahalanobis during the late thirties that the philosophy of 'efficiency per unit cost' came to be realised and in its train were explored many new methods of selection and estimation procedures.

During the next decade followed some significant developments in the sampling theory of finite populations, mainly relating to the use of supplementary information at different stages, through the notable contributions of Cochran, Hansen, Hurwitz, Yates, Madow, Dalenius and others. These developments gave rise to a number of techniques of sampling and estimation procedures appropriate to various situations in practice to estimate the population total of a real-valued characteristic defined for the units in the population.

With the availability of various selection and the corresponding estimation procedures, the need was felt to evolve a unified theory of sampling and steps in this direction were taken by Horvitz and Thompson (1952), Godambe (1955), Koop (1963) and Murthy (1963), and consequently generalised linear estimators were also proposed. Under this unified set-up, search for an optimum estimator was made by several authors and the following results are available in the literature for the homogeneous linear and the entire class of unbiased estimators, for any design P:

(a) non-existence of a uniformly minimum variance (best) estimator in the class of all homogeneous linear unbiased estimators $L_u^*$ of the population total (Godambe, 1955; Koop, 1963)

(b) admissibility of the Horvitz and Thompson estimator (the H.T. estimator) of the population total Y in $L_u^*$ (Godambe, 1960; Roy and Chakravarti, 1960),

(c) extension of the results in (a) and (b) to the entire class of unbiased estimators (Godambe and Joshi, 1965) of the population total.

(d) admissibility of the variance estimator (of the H.T. estimator) provided by Horvitz and Thompson (1952) in the entire class of unbiased estimators (Godambe and Joshi, 1965),

(e) non-existence of the best estimator in the entire class of unbiased estimators of the variance of the H.T. estimator (Godambe and Joshi, 1965).

Joshi (1965a, 1966) removed the unbiasedness condition of estimators and proved that the sample mean and the usual ratio estimators are admissible in the class of all estimators for any design with respect to variance as loss function and a generalisation of this result is given by him (Joshi, 1968) showing the admissibility of these estimators for any convex-loss function. Further, Joshi (1965b) has also shown an estimator, which includes the H. T. estimator, to be admissible for fixed sample size designs in the class of all estimators of the population total.

Apart from the above results in this direction, Murthy (1957) and Basu (1958), respectively proved that any estimator which depends on the order of selection and the repetition of units in the sample, is inadmissible. Basu also introduced the concept of 'sufficiency' in sampling theory, which was later developed by Pathak (1962, 1964) and others.

The result mentioned in (a) led to the choice of estimators from the class of admissible estimators and various criteria were then put forward to arrive at an optimum choice, namely (i) Bayesness (Godambe, 1955), (ii) invariance and regular class (Roy and Chakravarti, 1960) (iii) hyper-admissibility (Hanurav, 1965, 66, 68) and (iv) necessary bestness (Prabhu Ajgaonkar, 1965). We shall briefly mention below these criteria and reserve a detailed discussion on them for Chapter III.

Bayes approach, introduced in sampling theory by Cochran (1946), assumes the existence of some knowledge about the population prior to construction of the design and that with the help of this knowledge it is possible to formulate a prior distribution for the character under study. With the help of a specific form of an apriori distribution

Godambe (1955) proved that the strategy consisting of the H. T. estimator and a design with constant effective sample size having the inclusion probability of a unit proportionate to the corresponding value of the supplementary character, known apriori, is optimum in $L_u^*$. Afterwards Godambe and Joshi (1965) did away with the linearity restriction on the estimators, and Hanurav (1962) and Vijayan (1966) showed that the result is true even for designs with expected effective sample size being constant for the linear and entire class, respectively. As regard the second criterion no further work is traceable.

The criterion of hyper-admissibility (h-admissibility, for short), which is based on the concept of admissibility, requires an estimator to be admissible not only in the whole space $R_N$ but also in each of its principal hyper-surfaces (phs's), which are $(2^N-1)$ in number. It was shown by Hanurav (1968) that the H. T. estimator is the unique h-admissible estimator for any non-unicluster design in the class of all polynomial unbiased estimators of Y.

The necessary best criterion for choice of an estimator takes into account only the first part of the variance of an estimator t, namely $\sum_i A_i Y_i^2$ (leaving aside the second part $\sum_{i \neq j} A_{ij} Y_i Y_j$) and an estimator t is said to be the necessary best in a class of unbiased estimators if the coefficient $A_i$ $(i = 1, 2, \ldots, N)$ of $Y_i^2$ is least among the coefficients (of $Y_i^2$) for any other estimator in that class. Prabhu Ajgaonkar proved that the H. T. estimator is the necessary best in a sub-class $T_5$ of homogeneous linear unbiased estimator $L_u^*$ of $Y$ and Hege (1967) has extended this result to the class $L_u^*$ itself.

Another side of development, as pointed out earlier, was with regard to the use of information on supplementary character at the estimation and selection stages. Cochran (1942) developed ratio method of estimation using information on a single supplementary variable which was later extended to an estimator using two or more such variables by Olkin (1958). The product method of estimation, complementary to the ratio method, was considered by Murthy (1963) and conditions for choosing any of the unbiased, ratio or product estimator was given for a specific class of designs. Koop (1964) obtained similar conditions for ratio estimator applicable to any design. Recently, Srivastava (1967,1968) has proposed a generalisation of the usual ratio estimator, Olkin's multivariate ratio estimator and the ratio cum product estimator, given by the author (1967b, discussed in Chapter VII). J.N.K.Rao (1968a) has however given another generalisation of the usual ratio estimator which is obtained as a linear combination of the unbiased estimator and the ratio estimator and hence simple to compute. These estimators are discussed in Chapter VIII and compared with an estimator suggested there in.

Attempts were also made to make the ratio estimator unbiased (or almost unbiased), by modifying the sampling schemes by Lahiri (1951), Midzuno (1952), Sen (1952), Murthy, Nanjamma and Sethi (1959) and others, or by adjusting for its bias by Hartley and Ross (1954), Quenouille (1956), Durbin (1959), Murthy and Nanjamma (1959), Rao (1966) and others.

Hansen and Hurwitz (1943) introduced probability proportionate to size (pps) sampling with-replacement, the size being the value of the supplementary variable, which was later extended to pps - without replacement and to inclusion probabilities proportionate to size schemes by Horvitz and Thompson (1952), Durbin (1953), Des Raj (1956), Murthy (1957), Stevens (1958), Hajek (1959), Hartley and Rao (1962), Hartley and Cochran (1962), Fellogi (1963), Brewer (1963), Seth (1966), Hanurav (1967), Sampford (1967) and others.

Saito (1958), Des Raj (1964) and D. Singh and B.D. Singh (1965) considered pps with replacement selection of the second-phase sample using the information on the supplementary variable collected in the first-phase sample in two-phase sampling.

Comprehensive reviews of the developments in sampling theory have been given by Yates (1946), Sukhatme (1959), Seth (1961), Dalenius (1962) and Murthy (1963), besides the author (1966, in collaboration with Murthy).

We shall now give a brief summary of the author's contributions to the theory of sampling from finite populations.

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

In Chapter III we examine critically the concepts and definitions of bestness and admissibility as applied to sampling theory of finite populations and give some basic results in this direction. Definitions are given for a best, the best and the uniformly best estimators and for admissible and essentially admissible estimators, and the possible inexactitude in the use of these definitions in the current literature have been pointed out. It is, however, shown that for the H. T. estimator (and some other estimators) the two definitions of admissibility are identical and that this is not so in general is also established. Then removing the unbiasedness condition, as has been done by Joshi, it is shown that any 'constant' is essentially admissible for estimating the population total Y. Sufficient condition for the non-existence of a best estimator for a class of estimators is obtained and it is shown that there does not exist a best estimator and hence the best and the uniformly best estimator of the population total in the class of all linear estimators and the class of all estimators of Y.

Further, some aspects of the optimality criteria, mentioned earlier have been studied in this chapter, giving the earlier developments on them. As regards Bayesness it is noted that for a more realistic apriori distribution than that considered by Godambe, there does not exist an optimum strategy. Choice among some strategies have been given in the next chapter. In connection with h-admissibility of an estimator, Hanurav's result of unique h-admissibility of the H.T. estimator has been extended to a wider class of unbiased estimators of Y. The concept of h-admissibility is then extended to estimation of the variance of the H. T. estimator and it is shown that the variance estimator ($v_{ht}$) proposed by Horvitz and Thompson (referred in (d) of page 3) is the unique h-admissible estimator. The necessary bestness of the H. T. estimator and its variance estimator $v_{ht}$ have also been established and it is then shown that the necessary best estimator of Y and the variance of the H. T. estimator are unique h-admissible estimators for the corresponding parameters in a wide class of estimators. Importance of the vectors $y$ belonging to the principal hyper-surfaces of one dimension in proving the above results has been emphasised and some suggestions for the modification of the existing criteria and use of cost function and stability of the estimator and the variance estimator for the choice of reasonably good estimator(s) have been suggested.

Chapter IV is devoted to extension of some well-known uni-phase pps without-replacement schemes to two-phase sampling schemes, where data on the size measure is not at hand and is collected in a large first-phase sample and later utilised for pps selection of the second-phase sample which is a sub-sample of the first-phase sample. These schemes are then compared among themselves under an appropriate super-population model.

In Chapters V to VIII we consider the use of detailed information on one (or more) supplementary character $x$ in the estimation procedure for estimating the parametric function $\theta$ and some non-linear for parametric function $\eta(\theta)$. Chapter V deals with usual ratio and product estimators. More general conditions than Murthy's, for choosing either ratio, unbiased or product estimator are obtained and their use in systematic sampling is discussed. Some empirical studies are included for illustration purpose.

Chapter VI extends the univariate product method of estimation to a multivariate product method for estimating $\theta$ and compares this estimator with Olkin's multivariate ratio estimator demarcating the regions for their preference. An extension of this method to two-phase sampling is given and the manner in which multi-supplementary information may be used is explained. An empirical study is also included.

Chapter VII deals with estimation of a non-linear parametric function $\eta(\theta)$ in general and in particular with the estimation of ratio (R) and product (P) of the parameters $\theta_0$ and $\theta_1$. Two estimators for each R and P have been proposed, which utilise information on a supplementary character, and compared with the usual estimators for ratios and products. Configurational representations for the regions of preference for these estimators are also given. Two combinations of the proposed estimators, one of which gives the usual double ratio estimator as a particular case, are then considered for estimating R and P, and compared with other estimators. Some almost unbiased estimators are also suggested and a method of using information on several characters is also considered.

Later, these estimators are extended to estimate the parametric function $\theta_0$ itself and compared with other known estimators of $\theta_0$ such as unbiased, ratio, product, multivariate ratio, product estimators etc. Some empirical studies are also included for illustration.

In Chapter VIII a generalised ratio estimator, as an alternative to the estimator suggested by Srivastava, has been proposed. This estimator is quite simple to compute, unbiased for large sample sizes and has variance equal to that of the usual regression estimator. Two multivariate estimators corresponding to the suggested estimator have been proposed and compared with other known estimators, which use same amount of information.

Lastly we consider the situation where the detailed information is not available about any related character and is costly to collect as well but on the other hand an apriori value ($\theta_0$) of the parameter $\theta$ is known from previous census or surveys or even from expert guesses. And in Chapter IX we suggest an estimator of $\theta$ which utilize $\theta_0$ and is given by the weighted average of $\theta_0$ and the unbiased estimator $t$ of $\theta$. Bias and m.s.e of the estimator are obtained. It is noted that the optimum weight is a function of the relative difference between $\theta$ and $\theta_0$ and relative standard error of $t$ which may not be known in practice. Hence the suggested estimator is modified to use an approximate optimum weight. This estimator is then compared with the usual unbiased estimator and table showing its efficiency is given. Some special cases of this estimator are then pointed.

# CHAPTER II

## PRELIMINARIES

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

### 2.1 Sample Design:

A collection of known finite number $N$ of identifiable and distinct units $U_1, U_2, \ldots, U_i, \ldots, U_N$ is called the 'finite population' and it will be denoted by

$$U: \{U_1, U_2, \ldots, U_i, \ldots, U_N\} \qquad (2.1.1)$$

where $U_i$ corresponds to the i-th unit of the population $U$ and it will be sometimes denoted by its subscript $i$ only, such that, $U$ may be represented as a set of integers $1, 2, \ldots, i, \ldots, N$. A list of units in $U$ is termed the 'sampling frame' and the number $N$ is called the 'population size'.

A 'sample' from $U$ is an ordered finite sequence of units from $U$ and is denoted by $s$. Thus

$$s = U_{i_1}, U_{i_2}, \ldots, U_{i_{n(s)}} \qquad (2.1.2)$$

where $n(s) < \infty$ and

$$1 \leq i_t \leq N \quad \text{for} \quad 1 \leq t \leq n(s).$$

The $i_t$'s need not necessarily be distinct and the interchange of $U_{i_t}$ and $U_{i_{t'}}$ for $i_t \neq i_{t'}$ results in a new sample. That is in such a sample s repetition of units is allowed and different samples are obtained corresponding to different orders in which units can be arranged. For a sample s, n(s) denotes the sample size and $\mu(s)$, the number of 'distinct' units in sample s, denotes the effective size of s.

We define S, the collection of all possible samples s from U as the basic 'sample space':

$$S = \{ s \}, \qquad (2.1.3)$$

Evidently, S contains a countably infinite number of samples.

The sample space S with the probability measure P defined on it, such that corresponding to every $s \in S$ is a probability P(s) attached, where

$$P(s) \geq 0 \quad \text{and} \quad \sum_{s \in S} P(s) = 1, \qquad (2.1.4)$$

is called the 'sample design' and is denoted by D(U,S,P) or briefly the 'design P'.

The design P is completely specified by a list of all possible samples including all permutations and repetitions of the units in s with their respective probabilities of selections. In practice, however, samples are not drawn by listing all possible samples and corresponding $P(s)$'s due to the fact that for large values of $N$ and $n$ it becomes quite a difficult and unmanageable task. Instead they are drawn by some workable procedure termed as sampling scheme. Any sampling scheme in which the samples are ordered samples as in (2.2) gives rise to a unique design. Elaborate sampling schemes to implement a given design $D(U,S,P)$ have been suggested by Lahiri (1951), Midzuno (1952), Horvitz and Thompson (1952), Durbin (1953) and many others.

Among these sampling schemes the unit drawing mechanism, which consists in selecting units from $U$ one by one, is of special interest. The drawing mechanism $q: q(u,k,s_{k-1})$ denotes the probability of drawing unit $u$ from $U$ at 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 an important result due to Hanurav (1962) and Subrahmanya (1965) is that of one-one correspondence between a completely specified design and the drawing mechanism $q$. Because of this one can always work within the unified frame work of the designs for search of an optimum estimator. However, the situation with partially specified designs is different as in such cases there does not exist a unique mechanism. Examples of such designs are (i) specification of samples without considering the permutations and repetitions of units, (ii) specification of only the inclusion probabilities.

Given a design $P$ the inclusion probability $\pi_i$ of the unit $U_i \in U$ in the sample $s$ is

$$\pi_i = \sum_{s \supset i} P(s) \qquad (2.1.5)$$

where summation is taken over all samples containing $U_i$. Similarly the joint inclusion probability of units $U_i$ and $U_j$ is

$$\pi_{ij} = \sum_{s \supset i,j} P(s) \qquad (2.1.6)$$

where summation is over $s$ containing the pair $U_i$ and $U_j$.

For a given design $P$ both $\pi_i$ and $\pi_{ij}$ are constants, some inter-relationships between them are mentioned below:

(i) Yates and Grundy (1953): For a constant effective sample size design $P_\mu$, for which

$$P(s) = 0 \text{ if } \mu(s) \neq \mu \text{ for all } s \in S, \qquad (2.1.7)$$

where $\mu$ is a constant,

we have

$$\sum_{j \neq i} \pi_{ij} = (\mu - 1)\pi_i$$

(ii) Godambe (1955): For any design

$$\sum_{i \in U} \pi_i = \mu$$

and (iii) Hanurav (1962): For any design P

$$\sum_{i \neq j} \sum \pi_{ij} = \mu(\mu-1) + V(\mu(s))$$

$$= \mu(\mu-1), \text{ if } P \text{ is } P_{\mu},$$

where $\mu = \sum_{s \in S} \mu(s) p(s)$.

For details of internal consistency of the inclusion probabilities reference may be made to Hanurav (1966).

## 2.2 Estimation:

Let $\mathcal{Y}$ be the real-valued variable defined on the units (integers) of U taking value $Y_i$ on $U_i$ $(i = 1, 2, \ldots, N)$. Then the space of all possible vectors

$$y = (Y_1, Y_2, \ldots, Y_N) \quad (2.2.1)$$

of the variate $\mathcal{Y}$ is the N-dimensional Euclidean space $R_N$ and any function

$$\theta = \theta(y) = \theta(Y_1, Y_2, \ldots, Y_N) \quad (2.2.2)$$

of 'y is called the parametric function (pf) defined on $R_N$. A parametric function of particular interest is the 'population total, a function defined on $R_N$, given by

$$T(y) = \sum_{i \in U} Y_i \quad \ldots \quad \ldots \qquad (2.2.3)$$

for every $y \in R_N$. It is sometimes denoted by just $Y$. A general problem in sampling is to estimate $\theta(y)$ by observing the values of $Y_i$ for just those units which belong to the specified sample $s$ from $U$ selected through the design $P$.

We define an <u>estimator</u> $t(s,y)$ as a real-valued function $t$ defined on the space $S \times R_N$, depending on $y$ through only those $Y_i$'s for which $i \in s$. Obviously the estimator $t(s,y)$ need not be defined over those $s$ for which $P(s) = 0$ and hence we tacitly assume that the sample space $S$ is such that $P(s) > 0$ for all $s$. Also we note that, given the design, for any two vectors $y$ and $y'$ for which $Y_i = Y_i'$ for $i \in s$, $t(s,y) = t(s,y')$.

The value realised by an estimator is called an <u>estimate</u> of the parameter. Although there is no means of ascertaining the error in an individual estimate, the average error over all possible estimates may be determined with the help of the probability distribution, usually termed as its sampling distribution. Specifically, the degree of concentration of the sampling distribution about the parameter to be estimated represents a probabilistic measure of the degree of precision. The more the concentration is, the greater the probability of the estimate being nearer to the parameter, that is the more precise is the estimator. Thus, error resulting from a particular estimate $t(s,y)$ will be given by the difference $(t(s,y) - \theta(y))$. A convex function L of this error is called the loss-function and E(L), the expected value of L, is called the expected loss. A commonly used loss-function is the mean-square error and will be denoted by mse or simply M. Thus by definition, we have

$$M(t) = M(t(s,y)) = E(t(s,y) - \theta)^2$$

$$= \sum_{s \in S} (t(s,y) - \theta)^2 P(s) \qquad (2.2.4)$$

for every $y \in R_N$. The M(t) is a real-valued non-negative function and is taken as the criterion for the choice of estimators.

Now we explain below the concepts of at least as good as (denoted by the symbol $\succ$ ) and better than (denoted by $\not\succ$ ) as applied to two estimators $t_1(s,y)$ and $t_2(s,y)$ for a given design P. It may be mentioned that these concepts play an important role in choice of estimators and help in pointing out the possible inexactitude present in the liturature on sampling from finite populations regarding the formulation and application of the definitions of best and admissible estimators, a detail discussion about which is given in the next chapter.

For a given design P, an estimator $t_1$ is said to be at least as good as an other estimator $t_2$ if

$$M(t_1) \leq M(t_2), \text{ for all } y. \qquad (2.2.5)$$

Suppose the estimator $t_2$ is also at least as good as the estimator $t_1$, then

$$M(t_2) \leq M(t_1), \quad \text{for all } y. \qquad (2.2.6)$$

Now if (2.2.5) and (2.2.6) hold simultaneously then they imply that

$$M(t_1) = M(t_2), \quad \text{for all } y. \qquad (2.2.7)$$

Hence, we have, $t_1 \succ t_2$ and $t_2 \succ t_1$ if and only if (iff), $M(t_1) = M(t_2)$ for all $y \in R_N$. Again, from (2.2.5), $t_1$ is <u>not</u> at least as good as ( $\not\succ$ ) $t_2$ if

$$M(t_1) > M(t_2), \quad \text{for at least one } y, \qquad (2.2.8)$$

and similarly from (2.2.6) $t_2 \not\succ t_1$ if

$$M(t_2) > M(t_1), \quad \text{for at least one } y. \qquad (2.2.9)$$

Thus from (2.2.8) and (2.2.9) $t_1 \not\succ t_2$ and $t_2 \not\succ t_1$, that is neither $t_1$ nor $t_2$ is at least as good as the other if

$$M(t_1) \gtrless M(t_2) \qquad (2.2.10)$$

for all $y \in R_N$ with each inequalities ( >, < ) holding true for at least one $y$.

The estimator $t_1$ is said to be <u>better</u> than ( $\succ$ ) another estimator $t_2$ if

$$t_1 \succ t_2 \quad \text{but} \quad t_2 \not\succ t_1.$$

other words from (2.2.5) and (2.2.9), $t_1 \succ t_2$ if

$$M(t_1) \leq M(t_2), \quad \text{for all } y \qquad (2.2.11)$$

and

$$M(t_1) < M(t_2), \quad \text{for at least one } y. \qquad \text{~~(2.2.12)~~}$$

Similarly, $t_2 \succ t_1$ if

$$t_2 \succ t_1 \quad \text{but} \quad t_1 \not\succ t_2.$$

and the expression similar to (2.2.11) is obtained from (2.2.6) and (2.2.8). Thus $t_1$ is <u>not better than</u> ($\not\succ$) $t_2$ if

either $M(t_1) = M(t_2)$, for all $y$

or $M(t_1) > M(t_2)$, for at least one $y$. (2.2.12)

The estimator $t_1$ is said to be uniformly <u>better than</u> another estimator $t_2$ if

$$M(t_1) < M(t_2), \quad \text{for all } y. \qquad (2.2.13)$$

The following remarks, which are outcomes of the above concepts are to be noted to facilitate the discussion in the next chapter.

<u>Remark 2.2.1:</u> $t_1 \succ t_2$ does not necessarily imply that $t_2 \not\succ t_1$.

<u>Remark 2.2.2:</u> From (2.2.5) and (2.2.11), $t_1 \succ^{*} t_2$ is stronger criterion than $t_1 \succ t_2$ for choosing an estimator $t_1$, since in the former case the inequality, $M(t_1) < M(t_2)$, must necessarily hold for at least one $y \in R_N$.

<u>Remark 2.2.3:</u> From (2.2.8) and (2.2.12) it is obvious that $t_1 \not\succ t_2$ is stronger criterion than $t_1 \not\succ^{*} t_2$ for choosing the estimator $t_1$ since in the former case $t_1$ is discarded (i.e. $t_1$ is bad) only if (2.2.8)(which the second part of (2.2.12)) holds while in the latter case $t_1$ is discarded if either of the two conditions in (2.2.12) is satisfied.

<u>Remark 2.2.4:</u> The two criteria $t_1 \not\succ t_2$ and $t_1 \not\succ^{*} t_2$ are equivalent iff $M(t_1) = M(t_2)$, for all $y \in R_N$ implies $t_1$ and $t_2$ are identical for all $y$. That this is not so, in general, will be shown in the next chapter.

Another criterion for choosing an estimator is its unbiasedness which is nearly always taken for granted in the field of sample surveys, due to its intuitive appeal and statistical interpretibility.

For a given design $P$, an estimator $t$ will be said to be <u>unbiased</u> for the pf $\theta$ if

$$E(t) = \sum_{s \in S} t(s,y)\, P(s) = \theta \qquad (2.2.14)$$

for all $y \in R_N$. The estimator $t$ will be <u>biased</u> if (2.2.14) is not satisfied.

For a given design P, the class of all estimators will be denoted by A and the class of all such estimators satisfying (2.2.14) will be denoted by $A_{\mu}$. A parametric function $\theta$ will be said to be <u>estimable</u>, with respect to a given design P, if there exists an estimator which is unbiased.

If an estimator t is unbiased for $\theta$ (i.e., $t \in A_{\mu}$) then M(t) in (2.2.4) is equal to the variance of t (denoted by V(t)), given by

$$V(t) = \sum_{s \in S} (t(s,y) - E(t))^2 P(s)$$

$$= \sum_{s \in S} t^2(s,y) P(s) - \theta^2 \qquad (2.2.15)$$

for every $y \in R_N$.

<u>Remark 2.2.5:</u> If we are considering the two estimators $t_1$ and $t_2$ both belonging to $A_{\mu}$ then we can replace mean square error by variance in the above discussion.

<u>Remark 2.2.6:</u> A design P together with an estimator t defined over P is called a sampling <u>strategy</u> for the estimation of $\theta$ and is denoted by H(P,t). Hanurav (1966) brought into light the necessity of this definition which is due to Hájek (1958). Unbiasedness of a strategy depends upon the unbiasedness of t. The expectation, mean square error or variance of a strategy are defined as the expectation, mean square error or variance of the estimator t over the design P. And the choice between the

two strategies $H_1$ and $H_2$, in both of which $\theta$ is estimable should be based on the joint consideration of the variance and the cost per unit.

Now we shall consider a class of relatively simple estimators what are called as homogeneous linear estimators devised as linear functions of the sample observations and therefore quite prevalent in practice. The theory of linear estimation in classical theory of estimation for the case of infinite populations is quite different from the one which we use in sample surveys. The essential difference of the populations encountered in sample surveys from those considered in classical theory is that the survey populations (finite) are composed of units which are 'identifiable' and in studying the population it matters whether a set of identical values relate to the same units repeated, or to different units of U. This fact was brought to light by the results of Basu (1958) and Des Raj and Khamis (1958). This necessitated a fundamental change in the formulation of general linear estimators. The first attempt in giving generalised estimators was made by Horvitz and Thompson (1952) who defined three classes of linear estimators:

$$T_1 = \sum_{i \in s} \beta_i Y_i$$

where $\beta_i$ is a constant to be used as weight for the unit selected at the i-th draw,

$$T_2 = \sum_{i=1}^{n} C_i Y_i$$

where $C_i$ is constant for a given design $P$ and is attached to the i-th unit $(i = 1,2,\ldots, N)$ whenever it is selected in the sample and

$$T_3 = \gamma_s \sum_{i \in s} Y_i ,$$

where $\gamma_s$ is the coefficient to be used as a weight whenever $s^{th}$ sample is selected.

Godambe (1955) generalised these classes and defined the homogeneous linear estimators by

$$t(s,y) = \sum_{i \in s} \beta(s,i)\, Y_i \qquad (2.2.16)$$

where the coefficients $\beta$'s depend both on the sample and the units to which they are attached, but do not depend on the variate values $Y_i$'s. The condition for $t(s,y)$ in (2.2.16), to be unbiased for $T(y)$ is given by

$$\sum_{s \ni i} \beta(s,i)P(s) = 1, \quad 1 \leq i \leq N. \qquad (2.2.17)$$

We shall call the estimators satisfying (2.2.16) as the homogeneous linear estimators and the class of such estimators will be denoted by $L^*$. The corresponding unbiased estimators (satisfying (2.2.17)) will be denoted by $L_u^*$. The variance of an estimator $t \in L_u^*$ is given by

$$V(t) = \sum_{i=1}^{N} Y_i^2 \left( \sum_{s \supset i} \beta^2(s,i)P(s) - 1 \right) + \sum_{i \neq j}^{N} \sum^{N} Y_i Y_j \left( \sum_{s \supset i,j} \beta(s,i)\beta(s,j)P(s) - 1 \right)$$

(2.2.18)

It is pertinent to mention in this connection that Murthy (1963) has developed a technique of generating estimators for any design for the class of parameters that can be expressed as sum of single valued set-function defined over a class of sets of units belonging to the finite population and a number of possibly different estimators are generated. Further Koop (1963) has proposed seven classes of linear estimators by making the coefficients of sample observations depend upon, what he calls as the '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. However, recalling that the estimator in (2.2.16) is defined for the ordered sequence, it can be seen that the most general homogeneous linear estimator proposed by Koop is, infact, identical with $t$ in (2.2.16).

Further, in the class $A$ of all estimators of $T(y)$, any estimator $t$ expressible as

$$t(s,y) = \alpha(s) + \sum_{i \in s} \beta(s,i)\, Y_i,$$

where $\alpha(s)$ and $\beta(s,i)$ do not depend on $y$, is said to belong to the class of <u>linear estimators</u> and this class of estimators will be denoted by $L$ and the corresponding class of unbiased estimator by $L_u$. Obviously, if

$$\alpha(s) = 0, \quad \text{for all } s \in S$$

the class $L^*$ and $L$ (or $L_u^*$ and $L_u$) are identical.

Higher order polynomial estimators can be defined similarly.

# CHAPTER III

## ON BESTNESS, ADMISSIBILITY AND OPTIMALITY OF ESTIMATORS IN SAMPLING FINITE POPULATIONS

3.0 Summary: In this chapter some aspects of the concepts and definitions of bestness and admissibility as applied to the theory of sampling from finite populations and the practical utility of some criteria put forward for the choice of an optimum estimator from among the class of admissible estimators, have been critically examined and some basic results have been given. In section 3.1 a brief historical back-ground for the origin of admissibility concept in statistical theory and developments in this direction in the field of sampling theory are mentioned.

In section 3.2 the definitions of best and admissible estimators are enumerated. Definitions are given for a best, the best and the uniformly best estimators and for admissible and essentially admissible estimators using the concepts of 'at least as good as' and 'better than', developed in the previous chapter. The possible inexactitude in the use of these definitions in the current literature have been pointed out in section 3.3. It is noted in this section that the two definitions (3.2.4 and 3.2.5) of admissibility are equivalent iff the equality of mean square errors of two estimators $t_1$ and $t_2$, for all vectors y implies that $t_1$ and $t_2$ are identical for all y. And in this connection it is shown in theorem 3.3.1 that the two definitions are equivalent if either $t_1$ or $t_2$ is the estimator proposed by Horvitz and Thompson (the H.T.estimator, for short) for estimating the population total Y. Certain other estimators are also shown to possess this property. And

that this is not so, in general has been pointed out on the basis of a simple illustration. Further, it is shown, in theorem 3.3.2, that if we remove the unbiasedness condition, as has been done by Joshi in recent papers, any 'constant' becomes admissible for estimating Y.

In section 3.4 the sufficient condition for the non-existence of a best estimator in a class of estimators has been obtained and it is shown in theorem 3.4.3 that there does not exist a best and hence the best and the uniformly best estimator for the class of all linear estimators and all estimators of Y for any design. Some specialised designs and the classes of unbiased estimators are then recalled where the best estimator doesexist. The discussion upto this section also based on results obtained by the author in collaboration with Murthy (1968).

Sections 3.5 to 3.8 are devoted to some aspects of three optimality criteria namely Bayesness, hyper-admissibility (h-admissibility, for short) and necessary bestness put forward in the literature. It is noted in section 3.5 that the H.T. estimator ceases to be Bayes solution (a-optimum) if the parameter g of the apriori distribution deviates even slightly from 2, and that no optimum solution exists in such cases.

Section 3.6 deals with the h-admissibility of an estimator (the criterion introduced by Hanurav). In section 3.6.1, which is based on the results obtained by the author in collaboration with Ramakrishna, we have (in theorem 3.6.2) extended the result (mentioned in theorem 3.6.1) obtained by Hanurav regarding unique h-admissibility of the H.T. estimator to a wider class of unbiased estimators for any non-unicluster design. We follow an alternative method for proof which sim-pin-points the vital role played by the vectors $y$ belonging to the principal hyper-surface's (phs) of one dimension. In section 3.6.2 we extend the criterion of h-admissibility to the estimation of the variance of the H.T. estimator and following the same approach, we prove that the variance estimator proposed by Horvitz and Thompson ($v_{ht}$) is uniquely h-admissible in wide classes of unbiased estimators (theorems 3.6.3 and 3.6.4) for any design in which the variance is estimable.

In section 3.7 we discuss the criterion of necessary bestness, introduced by Prabhu Ajgaonkar. It is noted that this criterion is equivalent to the criterion of the best estimator when the parameter is restricted to the phs's of one dimension, the phs which played the vital role in the proof h-admissibility of the H.T. estimator and the variance estimator $v_{ht}$. It is further shown on the basis of lemmas 3.6.1-3.6.4 that these estimators are necessary best (theorem 3.7.1-3.7.2) for a wide class of non-homogeneous estimators. That a necessary best estimator is h-admissible is also shown. Some practical significance of these optimality criteria are then mentioned in the last section and some suggestions have also been made for choice of reasonably good estimators under different situations commonly met with in practice. The discussion from section 3.6.2 onwards are based on the results obtained by the author in collaboration with J.N.K. Rao.

## 3.1 Introduction:

During the last two decades the concept of admissibility has been increasingly applied to the problems of estimation, testing of hypothesis, etc., in statistical theory. Lehmann (1947) introduced the concept of admissibility in the field of statistical testing of hypothesis. Wald (1947) has also used this concept for defining an admissible decision function in relation to the decision theory. Since then several contributions have been made in this direction in these fields.

In the field of sampling theory for finite populations, however, the work in this direction started much latter and at first the steps were taken by several authors more or less at the same time to characterise the class of estimators which were inadmissible. Murthy (1957) proved that when the design is one generated by the customary 'probability proportionate to size (pps)' sampling without replacement, estimators that take into account the <u>order</u> in which the units occur in the sample are inadmissible and hence can be uniformly improved upon. He also furnished a method of getting uniformly better estimators in such cases.

Another important contribution in this direction was due to Des Raj and Khamis (1958) and Basu (1958). They proved that in simple random sampling with-replacement the sample mean which takes into account the repetition of units in the sample is inadmissible being inferior to the mean over <u>distinct</u> units of the sample. Des Raj and Khamis obtained the result by direct calculation of the variances of the two estimators. Basu, extended the result to pps sampling with-replacement and proved that the customary estimator of the population total for this design also is inadmissible. He introduced the fruitful notion of 'sufficient statistic' in the field of sampling theory and proved his results by using Rao-Blackwell theorem.

A statistic (any function of sample values of the variable under study) is called sufficient statistic if it yields all the information in the sample concerning the parameter. A sufficient statistic for estimating the population total was thus defined to be the unordered set of distinct units in the sample $s$ obtained from a design $P$, together with corresponding real-valued variable. Evidently it is sufficient because any two 'effectively equivalent' samples, together with the real-valued variable yields the same information about the population parameter. For a given design $P$, two samples $s_1$ and $s_2$ are said to be 'effectively equivalent', in symbols

$$s_1 \sim s_2, \quad P(s_1) > 0, \quad P(s_2) > 0$$

iff every unit belonging to $s_1$ belongs to $s_2$ and conversely. Thus given any one of such samples one can always construct the class of effectively equivalent samples by repetition or arrangement of the units in it. Thus the value of the sufficient statistic is same for any sample belonging to the class of effectively equivalent samples and the conditional distribution of any other statistic given the sufficient statistic, will tell us nothing further about the parameter to be estimated. Pathak (1962, 1964) has carried out a series of investigations in this connection for different sampling schemes.

Roy and Chakravarti (1960) proved that for any design an unbiased estimator of the population mean which is either ordered or depends on the repetition of the units in the sample is inadmissible. In fact, as pointed out by Hanurav (1966, 1968), Basu gave the first clues to the generality of this result while introducing the notion of sufficiency in this theory and proved the following.

Theorem 3.1.1 (Basu). Given a design P in which $T(y)$ is estimable, if $t$ is an unbiased estimator of $T(y)$ then the estimator $t^*$ defined by

$$t^*(s,y) = \frac{\Sigma\, t(s_o,y)\, P(s_o)}{\Sigma\, P(s_o)}, \quad \text{if } \Sigma\, P(s_o) > 0$$

$$= 0 \quad , \text{otherwise} \qquad (3.1.1)$$

where $\Sigma$ summation ($\Sigma$) is taken over all samples $s_o$ effectively equivalent to $s$, is unbiased for $T(y)$ and for any convex loss function

$$a(t^*) \leq a(t), \quad \text{for all } y \in R_N, \qquad (3.1.2)$$

with strict inequality holding for at least one $y$ iff

$$P(s_1 \sim s_2, \ t(s_1) \neq t(s_2)) > 0.$$

Thus it is necessary for an estimator $t$ of $T(y)$ to be admissible in any class of unbiased estimator, that

$$t(s_1, y) = t(s_2, y)$$

for all $s_1$ and $s_2$ for which

$$P(s_1) > 0, \ P(s_2) > 0 \ \text{and} \ s_1 \sim s_2. \qquad (3.1.3)$$

In fact this is true for any estimable parametric function. We shall consider this aspect further in section 3.6 in connection with hyper-admissibility of an estimator.

Besides weeding out certain inadmissible estimators attempts were also made to provide admissible estimators due to non-existence of the best estimator (Godambe, 1955, Koop 1963). Godambe (1960) and Roy and Chakravarti (1960), independently, applied the admissibility concept to the sampling theory by defining and searching for admissible estimators. Since then contributions in this direction have been made, among others by Godambe and Joshi (1965), Hanurav (1966, 1968) and Joshi (1965a,b, 1966, 1968).

It is interesting to note that the definitions of best and admissible estimators used in the papers relating to sampling from finite populations are not always exactly the same and that there appears to be some possibility of inexactitude about their formulation and application. Further, the definitions, that are used, have been sometimes treated as equivalent, and even in cases where the difference in the definitions has been noted, the significance of this difference is not fully comprehended. In the next section we enumerate the definitions of best and admissible estimators and then certain basic results are obtained in the sections 3.3 and 3.4.

Sections 3.5 to 3.8 are devoted to some aspects of the problem of choice of optimum estimator.

## 3.2 Bestness and Admissibility:

Using the basic concepts, particularly at least as good as ( $\succeq$ ) and better than ( $\succ$ ), given in the earlier chapter we now give the definitions of best and admissible estimators for a given design P.

Definition 3.2.1: For a given design P, in a class C(P) of estimators of T(y), a member $t_1 \in C(P)$ is said to be the uniformly best estimator if it is uniformly better than any other member $t\ (\neq t_1) \in C(P)$, that is if for every other $t \in C(P)$

$$M(t_1) < M(t), \quad \text{for all } y. \qquad (3.2.1)$$

Definition 3.2.2: For a given design $P$, a member $t_1 \varepsilon C(P)$ is termed the best estimator in $C(P)$ if $t_1 \succ t$ for every other $t \varepsilon C(P)$, that is if for every other estimator $t \varepsilon C(P)$

$$M(t_1) \leq M(t), \text{ for all } y. \qquad (3.2.2)$$

with inequality holding true for at least one $y$.

Definition 3.2.3: For a given design $P$, a member $t_1 \varepsilon C(P)$ is said to be a best estimator in $C(P)$ if for every other estimator $t$ in $C(P)$ $t_1 \succ t$, that is if for every other $t \varepsilon C(P)$

$$M(t_1) \leq M(t), \text{ for all } y. \qquad (3.2.3)$$

Remark 3.2.1: The definitions 3.2.1 to 3.2.3 are practically the same, and the distinction made out are mainly of academic interest, especially since there does not exist any estimator, which is best according to any of the above diffinitions, even in restricted class of homogeneous linear unbiased estimators of $T(y)$.

Now, as the next step, it is logical to attempt to reduce the class $C(P)$ of estimators without loss of relevant information and in this context we have the following criterion.

Definition 3.2.4: For a given design $P$, in a class $C(P)$ of estimators of $T(y)$, an estimator $t_1 \varepsilon C(P)$ is said to be admissible if there does not exist any other estimator $t(\neq t_1) \varepsilon C(P)$ which is $\succ t_1$, that is if there is no other $t$ in $C(P)$ for which

$$M(t) \leq M(t_1), \quad \text{for all } y, \qquad (3.2.4)$$

with inequality holding for at least one $y$.

A sub-class $C_e(P) \subset C(P)$ is said to be <u>complete</u> in $C(P)$ iff for every estimator $t \in C(P) - C_e(P)$ there exist an estimator $t' \in C_e(P)$ which is better than $t$, that is iff for every $t \notin C_e(P)$ there exist a $t' \in C_e(P)$ such that

$$M(t') \leq M(t), \quad \text{for all } y,$$

with inequality holding for at least one $y$. Evidently every unbiased estimator in the complete class satisfies the necessary condition, given in (3.1.3), for being admissible.

A complete class is <u>minimal complete</u> if it does not contain a complete sub-class. Thus if a minimal complete class exists, as is usually the case, it exactly concides with the totality of the admissible estimators in the class $C(P)$.

Now, we define the following variant of the complete class notion using the concept of <u>at least as good as</u> instead of <u>better than</u> used above.

A sub-class $C_e(P) \subset C(P)$ is said to be <u>essentially complete</u> in $C(P)$ iff for any estimator $t \in C(P) - C_e(P)$ there exist an estimator $t' \in C_e(P)$ which is at least as good as $t$, that is, iff for every $t \notin C_e(P)$ there exist a $t' \in C_e(P)$ such that

$$M(t') \leq M(t), \quad \text{for all } y.$$

Clearly a complete class is necessarily essentially complete, that is $C_e(P) \supseteq C_c(P)$ since 'at least as good as' is less restrictive criterion than 'better than' for choosing an estimator. The minimal essentially complete class is also defined similarly and this class will concide with the class of all estimators what may be termed as essentially admissible estimators, defined as follows:

Definition 3.2.5: For a given design P and the class C(P) of estimators an estimator $t_1 \in C(P)$ is termed as essentially admissible if there does not exist any other estimator $t(\neq t_1) \in C(P)$ which is $\succ t_1$, that is there exist no other estimator $t \in C(P)$ for which

$$M(t) \leq M(t_1), \text{ for all } y. \qquad (3.2.5)$$

Further, it may be noted that if $t_1$ and $t_2$ are two estimators of T(y) both belonging to the minimal complete class then $t_1 \not\succ t_2$ and $t_2 \not\succ t_1$, that is either $M(t_1) \gtrless M(t_2)$, for all $y \in R_N$ with each inequalities (>, < ) holding true for at least one $y \in R_N$, or $M(t_1) = M(t_2)$, for all $y \in R_N$. Similarly, if $t_1$ and $t_2$ are two estimators both belonging to the minimal essentially complete class then $t_1 \not\succ t_2$ and $t_2 \not\succ t_1$, that is $M(t_1) \gtrless M(t_2)$ holds for all y with each inequalities holding true for at least one $y \in R_N$. Thus it is seen that the difference between the minimal complete and minimal essentially complete classes is due to the estimators having equal mean square errors for all $y \in R_N$. Thus if an estimator $t_1$ belongs to the minimal complete class then all other estimators having mean square error equal to that of $t_1$ for all $y$, must also belong to this class but a minimal essentially complete class admits only one member from among the estimators having equal mean square error. And hence the essentially minimal complete class is further reduction of the minimal complete class in the sense that the former is included in the latter. In order to make these concepts clearer, we shall express the definitions 3.2.4 and 3.2.5 in a positive manner as the definitions 3.2.1 to 3.2.3.

Definition 3.2.4': An estimator $t_1 \epsilon C(P)$ is admissible in $C(P)$ if for every other $t \epsilon C(P)$

$$\text{either} \quad M(t_1) = M(t), \quad \text{for all } y$$

$$\text{or} \quad M(t_1) < M(t), \quad \text{for at least one } y. \qquad 3.2.6)$$

Definition 3.2.5': An estimator $t_1 \epsilon C(P)$ is essentially admissible in $C(P)$ if for every other $t \epsilon C(P)$

$$M(t_1) < M(t), \text{ for at least one } y. \qquad (3.2.7)$$

The point $y$ in the above inequalities may however depend on $t_1$ and $t$ and possibly on the design $P$. From the definitions 3.2.4' and 3.2.5' it is clear that the only difference between the definitions 3.2.4 and 3.2.5 is that of the former allows equality of the mean square errors of the estimators where as the latter does not allow this.

<u>Remark 3.2.2:</u> If an estimator is essentially admissible, it is necessarily admissible, but the reverse is not always true. This is clear from the following example. Let us consider a class $C_o(P)$ containing four estimators $t_1, t_2, t_3$ and $t_4$, and let $M(t_1) = M(t_2)$ for all $y$ and $M(t_1) \gtrless M(t_i)$, $i = 3,4$ for all $y$ with inequalities holding for at least some $y$'s. In this situation both $t_1$ and $t_2$ are admissible according to definition 3.2.4, but none of them is essentially admissible (definition 3.2.5). However, if $t_2$ is excluded from $C_e(P)$ then $t_1$ becomes essentially admissible and if $t_2$ is excluded $t_1$ becomes essentially admissible. Thus only one of $t_1$ or $t_2$ can belong to the minimal essentially complete class whereas both $t_1$ and $t_2$ belong to the minimal complete class.

<u>Remark 3.2.3:</u> If we are considering only the class of unbiased estimators, then we can replace mean square error by variance in the above discussion. On the other hand the loss function need not be restricted to the mean square error only we can replace any loss function $\delta$ in its place and any parametric function $\theta(y)$ in the above discussion.

<u>Remark 3.2.4:</u> The definitions given above may be extended directly to other branches of statistical theory. For instance by replacing the estimator by test or by decision function and using the corresponding loss function (or power of test) we get admissibility and bestness of tests and decision functions.

Remark 3.2.5: It is of interest to note that the definitions 3.2.4 and 3.2.5 have been arrived at by (i) adopting the definition of an estimator $t_1$ being better than $t$, namely $t_1 \succ t$ but $t \not\succ t_1$ and (ii) taking the first part in (i), namely $t_1 \succ t$, to be valid. By considering the second part in (i) namely $t \not\succ t_1$, is valid we get an other definition, which is stronger than definitions 3.2.4 and 3.2.5 and in fact reduces to the definition of a best estimator (definition 3.2.3) when expressed in its positive form.

## 3.3 Admissibility of Estimators

In this section we first point out the significance of the difference between the two concepts and then prove that any constant is essentially admissible estimator for estimating the population total. It may be pointed out that Wald (1947) has noted the difference between the definitions 3.2.4 and 3.2.5, that is between admissible and essentially admissible concepts while introducing them in decision theory in connection with choice of a decision function. Lehmann (1947) has also observed this distinction while formulating the concept of minimal complete and minimal essentially complete classes of tests. To cite more recent reference, Burkholder (1960) and others have also used these concepts recognising the difference. The term complete class has been used by Wald (1947), Hodges and Lehmann (1951) and some others in the sense of essentially complete.

As mentioned earlier the difference between definitions 3.2.4 and 3.2.5 is however not quite apparent in the papers relating to sampling theory applied to finite populations. Godambe (1960) and Hanurav (1966) have used essentially admissible in the sense of admissible estimator while Roy and Chakravarti (1960), Godambe and Joshi (1965), Joshi (1965a,b, 1966, 1968) have used definition 3.2.4. The two definitions have been some times considered to be equivalent (Godambe and Joshi, 1965 ; Hanurav, 1968) without pointing out the significance of this equivalence.

It is, however, clear from (3.2.6) and (3.2.7) that the definitions 3.2.4' and 3.2.5' are equivalent iff $M(t_1) = M(t_2)$, for all $y$, implies $t_1$ and $t_2$ are identical for all $y \in R_N$. In the following we give instances where equality of mean square errors implies identity of estimators for all $y$.

Consider the estimators $t_1$ and $t_2$ both belonging to $A_u$, the class of all unbiased estimators, and let

$$t_1(s,y) = t_2(s,y) + e(s,y) \qquad (3.3.1)$$

where $e(s,y)$ is a function defined on $S \times R_N$ and depending on $y = (Y_1, Y_2 \ldots, Y_N)$ only through those $Y_i$'s for which $i \in s$. Then from the condition of unbiasedness of $t_1$ and $t_2$

$$\sum_{s \in S} e(s,y)\, P(s) = 0, \quad \text{for all } y. \qquad (3.3.2)$$

Now the problem is to examine whether

$$e(s,y) = 0, \quad \text{for all} \quad s \in S \quad \text{and} \quad y \in R_N \tag{3.3.3}$$

having given that

$$V(t_1) = V(t_2), \quad \text{for all} \quad y \in R_N. \tag{3.3.4}$$

Since $E(t_1) = E(t_2) = T(y)$, (3.3.4) implies

$$\sum_{s \in S} t_1^2(s,y)\, P(s) = \sum_{s \in S} t_2^2(s,y)\, P(s)$$

for all $y$, that is,

$$\sum_{s \in S} e^2(s,y)\, P(s) = -2 \sum_{s \in S} t_2(s,y) e(s,y) P(s) \tag{3.3.5}$$

for all $y$.

Thus having given (3.3.5) if (3.3.3) holds then the two definitions are equivalent. It seems that in general (3.3.4) and (3.3.2) together may not result in (3.3.3). However, following the proof of the theorem 4.1 of Godambe and Joshi (1965) it is easily seen that (3.3.4) and (3.3.2) together imply (3.3.3) if $t_1(s,y)$ (or $t_2(s,y)$) in (3.3.1) is an estimator given by

$$t_{ht} = \sum_{i \in s} \frac{Y_i}{\pi_i} \tag{3.3.6}$$

where $\pi_i (> 0)$ is the inclusion probability for $i^{th}$ unit of the population and summation is taken over the distinct units in the sample. Hence we have the following

Theorem 3.3.1: For a given design P, if $t_{ht}$ is the estimator given by (3.3.6) and $t$ is any estimator of Y belonging to $A_u$ then

$$V(t) = V(t_{ht}), \quad \text{for all } y \varepsilon R_N \tag{3.3.7}$$

implies that $t$ and $t_{ht}$ are identical.

It is pertinent to note here that the estimator $t_{ht}$ was suggested by Horvitz and Thompson (1952), as the unique unbiased estimator in his $T_2$ class, while formulating three classes of linear estimators mentioned in the previous chapter. This estimator, which we shall call (as is usually called) as the H.T. estimator for brevity, has drawn considerable attention in sampling theory. The sampling variance of $t_{ht}$ is given by

$$V(t_{ht}) = \sum_{i=1}^{N} \left(\frac{1-\pi_i}{\pi_i}\right) Y_i^2 + \sum_{i} \sum_{j \neq i} (\pi_{ij} - \pi_i \pi_j) \frac{Y_i Y_j}{\pi_i \pi_j} \tag{3.3.8}$$

As regards admissibility criterion, the H.T. estimator has been shown to be admissible in $L_u^*$, the class of homogeneous linear unbiased estimators, by Godambe (1960) and Roy and Chakravarti (1960), and also in the class of all unbiased estimators $A_u$ by Godambe and Joshi (1965) for any design, and in the class of all measurable estimators for fixed sample size designs (Joshi, 1965b). Some other results regarding Bayesness and hyper-admissibility of the HT estimator will be discussed in section 3.5 and 3.6.

Another situation where equality of mse's implies identity of estimators, is given by Joshi (1965,b). He has considered an estimator

$$t_b = \sum_{i \in s} b_i Y_i \tag{3.3.9}$$

where $\sum_{i \in U} b_i^{-1} = n$ and $b_i > 0, \; i = 1,2,\ldots, N$ (3.3.10)

and has proved it to be admissible in the class of regular estimators for fixed sample size designs and further it is shown that $M(t_b) = M(t)$ implies $t$ is identical with $t_b$ for all $y$. It may be mentioned that the usual ratio estimator has also been proved to be admissible by Joshi (1966) in the class of all estimators for any design but no such result has been shown to hold for this estimator.

These examples, however, do not prove that equality of mean square error implies always the identity of estimators. We give below a simple illustration to support this assertion.

Let us consider two estimators $t_1$ and $t_2$ such that

$$t_1(s,y) = t_1'(s,y) + e(s) \tag{3.3.11}$$

and

$$t_2(s,y) = t_1'(s\ y) - e(s) \tag{3.3.12}$$

where $t_1'$ is an unbiased estimator of the population total $T(y)$. The mean square error of $t_1$ and $t_2$ are, respectively,

$$M(t_1) = V(t_1) + \sum_{s\in S} e^2(s)P(s) + 2\sum_{s\in S}(t_1'(s,y)-T(y))e(s)P(s)$$

and $$M(t_2) = V(t_2) + \sum_{s\in S} e^2(s)P(s) - 2\sum_{s\in S}(t_1'(s,y)-T(y))e(s)P(s).$$

Now it is observed that

$$M(t_1) = M(t_2), \quad \text{for all } y \tag{3.3.13}$$

if $$\sum_{s\in S}(t_1'(s,y) - T(y))e(s)P(s) = 0, \quad \text{for all } y. \tag{3.3.14}$$

Supposing that (3.3.14) is satisfied even then $t_1$ and $t_2$ need not necessarily be identical. Obviously (3.3.14) is satisfied whenever

$$e(s) = e, \quad \text{a constant.} \tag{3.3.15}$$

But even then $t_1$ and $t_2$ are not identical. Hence the assersion is true in general. However, it is of interest to find exactly the situations under which the definitions 3.2.4 and 3.2.5 are equivalent.

On the basis of above discussion and the theorem 3.3.1 we have the following remarks to make:

Remark 3.3.1: The result obtained by Godambe (1960) and Roy and Chakravarti (1960) regarding the admissibility of the H.T. estimator are equivalent. That is the H.T. estimator is essentially admissible.

Remark 3.3.2: Hege's (1966) proof regarding admissibility of certain estimators in the class $L_u^*$ infact establishes its essential admissibility.

It may be mentioned here that recently Joshi (1965,66,68) has investigated the admissibility of estimators by removing the criterion of unbiasedness and has proved sample mean and the usual ratio estimator to be admissible for any design. In the following we show that the criterion of admissibility leads one astray when the unbiasedness condition is removed by showing any 'constant' to be admissible for the population total. We prove the following

Theorem 3.3.2: For any given design P, any constant is essentially admissible in the class L of linear estimators of the population total, except in a trivial case.

Proof: Consider a design P and an estimator $t_0$, belonging to the class L, given by

$$t_0 = \lambda, \quad \text{for all } y \in R_N \text{ and } s \in S, \tag{3.3.16}$$

where $\lambda$ is a constant not equal to zero. Further, let $t_1$ be any other estimator belonging to L, then

$$t_1 = \alpha(s) + \sum_{i \in s} \beta(s,i) Y_i, \quad \text{for all } y. \tag{3.3.17}$$

Suppose $\mathcal{Y}_{(\lambda)}$ is the class of vectors in $R_N$ such that

$$T(y) = \lambda, \quad \text{for all } y \in \mathcal{Y}_{(\lambda)}. \tag{3.3.18}$$

Then the mean square error of $t_o$

$$M(t_o) = (\lambda - T(y))^2$$

$$= 0 , \qquad \text{for all } y \in \mathcal{Y}_{(\lambda)} \tag{3.3.19}$$

whereas

$$M(t_1) = \sum_{s \in S} (t_1 - t(y))^2 \, P(s)$$

$$\geq 0 , \qquad \text{for all } y \in R_N \,. \tag{3.3.20}$$

Thus from (3.3.19) and (3.3.20), $t_o$ is essentially admissible unless

$$M(t_1) = 0, \quad \text{for all } y \in \mathcal{Y}_{(\lambda)}. \tag{3.3.21}$$

But it is observed that (3.3.21) is true only when

$$t_1 = t_o, \quad \text{for all } y \in \mathcal{Y}_{(\lambda)}. \tag{3.3.22}$$

Now the theorem will follow if we can show that (3.3.22) is true for all $y \in R_N$, that is $t_o$ and $t_1$ are identical estimators.

To show this let $y^{(i)}_{(\lambda)}$ denote vectors belonging to $y_{(\lambda)}$ given by

$$Y_i = \lambda, \quad Y_j = 0, \quad j(\neq i) = 1,2,\ldots, N. \qquad (3.3.23)$$

Then for all $y \in y^{(i)}_{(\lambda)}$,

$$t_1 = \lambda = \alpha(s) + \beta(s,i)\lambda, \quad \text{for} \quad s \in S_i$$

$$= \alpha(s), \quad \text{for} \quad s \in S_i^* \qquad (3.3.24)$$

where

$$S = S_i \cup S_i^* \quad \text{and} \quad S_i = \{s,\ i \in s\}. \qquad (3.3.25)$$

Since $i$ is arbitrary, (3.3.24) holds for $1 \leq i \leq N$ and this gives

$$\alpha(s) = \lambda = t_0 = t_1, \quad \text{for all} \quad s \in S, \qquad (3.3.26)$$

except for $s_N$, the sample containing all the units of the population and (3.3.26) implies from (3.3.24) that

$$\beta(s,i) = 0 \ , \quad \text{for all} \quad i.$$

Hence $t_1$ is identically equal to $t_0$ when $s_N$ is excluded from the effective sample space.

Hence the theorem.

## 3.4 Non-existence of a Best Estimator

We first summarise the work done in this direction and point out the difference in the definitions used by the authors and then give some new results in this section.

Godambe (1955) proved the non-existence of the uniformly best estimator in $L_u^*$, the class of homogeneous linear unbiased estimators of the population total $T(y)$, by attempting to minimise the variance of an estimator $t \in L_u^*$. The linearity restriction has been removed by Godambe and Joshi (1965), who have proved non-existence of the best estimator in the entire class of unbiased estimators $A_u$.

It is pertinent to note here that the non-existence of the uniformly best estimator in $L_u^*$ has been interpreted as non-existence of a best estimator (definition 3.23) by Godambe (1955) and as the non-existence of the best estimator (definition 3.2.2) by Roy and Chakravarti (1960). However, it appears to be reasonable to be satisfied for all practical purposes with a best estimator (if it exist) in a given class provided mean square error is the only criterion for choice of an estimator. In the following theorems, however, we demonstrate its non-existence for some classes of estimators of the population total. We first prove the following.

Theorem 3.4.1: For a given design P, if an estimator $t_1$ is admissible in class C(P) and if there exist another estimator $t_2$ in C(P) such that $t_1 \not> t_2$, then $t_1$ is not a best estimator in that class and this implies non-existence of <u>a best</u> estimator in the class C(P).

<u>Proof</u>: Since $t_1$ is an admissible estimator, the relation (3.2.6) holds and further as $t_1 \not> t_2$,

$$M(t_1) > M(t_2), \quad \text{for at least one } y$$

also holds which contradicts the condition (3.2.3) required for $t_1$ to be a best estimator. This also evidently implies non-existence of a best estimator in the class $C(P)$.

Hence the Theorem.

<u>Theorem 3.4.2</u>: If there exist two (or more) admissible estimators in a class C(P), with unequal mean square errors for at least one y, for a given design P, then that class does not contain a best estimator.

<u>Proof</u>: Let $t_1$ and $t_2$ be two admissible estimators in C(P) then from definition $t_1 \not> t_2$ and $t_2 \not> t_1$, that is

$$\text{either} \quad M(t_1) = M(t_2), \quad \text{for all } y. \tag{4.3.1}$$

or

$$M(t_1) \gtrless M(t_2), \quad \text{for all } y \tag{4.3.2}$$

with inequalities (>, < ) being satisfied for at least one y.

But under the assumption the relation (4.3.1) does not hold and whenever (4.3.2) is satisfied, neither $t_1$ nor $t_2$ is a best estimator. Hence the class does not contain a best estimator.

This completes the proof.

Remark 4.3.1: If there are at least two essentially admissible estimators in a class then there does not exist a best estimator in that class.

Remark 4.3.2: If $t_1$ is an admissible estimator in a class $C(P)$ and if it is not a best in $C(P)$ or even in a class $C'(P) \subseteq C(P)$ then there does not exist a best estimator in $C(P)$.

Remark 4.3.3: Non-existence of a best estimator in a class implies non-existence of the best or uniformly best estimator in that class.

Theorem 4.3.2: There does not exist a best estimator, and hence the best and the uniformly best estimator in the class of all linear estimators (L) and all estimators (A) of the population total for any design P.

Proof: Noting that any constant belongs to the class L and that it is essentially admissible (Theorem 3.3.2) in L, the non-existence of a best estimator in L follows from Remark 4.3.1.

As regards the class A, the two admissible estimator with unequal mean square errors for at least some y for this class are

$$t_2 = \frac{N}{n(s)} \sum_{i \in s} Y_i$$

and
$$t_3 = \frac{\sum_{i \in s} Y_i}{\sum_{i \in s} X_i} X$$

where $X_i > 0$, $i = 1,2,\ldots, N$, $X_i \neq X_j$ for at least one pair $(i,j)$ and $X = \sum_{i \in U} X_i$, $X_i$'s being constants (Joshi, 1968), and hence the non-existence of a best estimator follows from Theorem 3.4.2.

Hence the theorem.

The non-existence of the uniformly best estimator in $L_u^*$ (Godambe 1955, Koop, 1963) and $A_u^*$ Godambe and Joshi (1965) is now well known. Though non-existence of a best estimator and hence that of the best and uniformly best estimators is established in very wide classes of estimators for any design, there may be some special classes of estimators and/or designs (generally restricted in nature) which give rise to the best estimator. Examples of such classes of estimators and/or designs are those which result in a unique unbiased estimator. For instance, the H.T. estimator is the unique unbiased estimator, whatever may be the design, in Horvitz and Thompson's $T_2$ class, referred in Chapter II, of linear estimators.

The unicluster design considered by Hege (1965) and Hanurav (1966) is an other example where the H.T. estimator turns out to be the unique unbiased estimator in the class of homogeneous linear unbiased estimators $L_u^*$. It may be mentioned that no best estimator exist even for the unicluster designs when the class of unbiased estimators considered is linear unbiased estimators $L_u$ instead of $L_u^*$. This follows easily from Theorem 3.4.2 noting that there are more than one admissible estimators in this wider class (Hanurav, 1966, p. 1963). The example given by

Godambe and Joshi (1965, p. 1712) or a generalised form of it is another instance of restricted class having minimum variance estimator. Roy and Chakravarti's (1960) balanced design admitting the best estimator in the restricted class of regular estimators is also an example of very restricted design.

Due to non-existence of the best estimators in general, and the largeness of the class of admissible estimators, certain criteria for reducing the class of admissible estimators (by use of our apriori knowledge or otherwise) to the point that it gives a unique estimator for a given class have been introduced in the literature. We shall discuss now in the remaining part of this chapter the theoretical significance and practical utility of three such criteria, viz., (i) Bayesness (Godambe 1955) (ii) Hyper-admissibility (Hanurav, 1966, 65, 68) and (iii) Necessary bestness (Prabhu Ajgaonkar, 1965), which have gained considerable importance in current researches in the theory of sampling from finite populations. We also give some basic results for each of these criteria.

### 3.5 Bayesness:

Bayes approach to choose an optimum strategy for estimating the population total, is applicable when the same population U presents itself repeatedly and independently with a fixed but unknown apriori distribution of the parameter. Thus a finite population under study is by itself a random sample from this sequence of populations usually termed as the super-population, the concept originally introduced by Cochran (1946) and thereafter used by many others. However, not all populations in practice come to us imbedded in a sequence but when they do the Bayes approach offers certain advantages over any approach which ignores the fact that the parameter itself is a random variable, as well as over any approach which assumes a conventional distribution of the parameter not subjected to change with experience. Thus if $y = (Y_1, Y_2, \ldots, Y_N)$ is the random vector corresponding to the variable under study, having fixed but unknown apriori distribution then based on our experience and knowledge about a correlated vector $x = (X_1, X_2, \ldots, X_N)$, which in many cases is the value of $y$ on some previous occasion, it would be possible for us to impose a realistic probability distribution of $y$ given $x$ and call this distribution as an apriori distribution. This we shall denote by $\partial$.

The formulation of such an apriori distribution is a point on which opinions differ and since any statistical interpretation finds itself in and out of fashion as the doctrine bulows, we shall not discuss these differences here, however, it has its own importance when one considers the prior knowledge about the variable under study, at least in the present set-up, to be formulated in some sort of a prior distribution $\partial$, at least partially. It is important to note that the assumption of the existence of a prior distribution $\partial$ is used only for choosing an optimum strategy and the ultimate inference about the parameter would exclusively depend upon the observed sample $s$ and the variate values $Y_i$ for $i \in s$.

Now for the determination of an optimum sampling strategy the criterion of bestness is slightly modified and instead of minimising the actual loss function $V(H)$, we minimise the expected loss $E_{\partial}V(H)$ over the distribution $\partial$ for $H$ varying over $H_{\mu}$, the class of all equi-cost strategies.

A $H_o$ which minimises $E_{\partial}\ V(H)$, uniformly with respect to all the parameters of the distribution $\partial$ is called a '$\partial$-optimum strategy' in $H_{\mu}$.

Two important results in this direction with regard to the class of homogeneous linear unbiased estimators of the population total are due to Godambe (1955) and Hajek (1959). Godambe considered the class $\triangle_1$ of all apriori distributions $\partial_1$ for which

i) $E_{\partial_1}\ (Y_i|X_i) = a\ X_i$

ii) $V_{\partial_1}\ (Y_i|X_i) = \sigma^2 X_i^2$

and iii) $C_{\partial_1}\ (Y_i, Y_{i'}|X_i, X_{i'}) = 0$ (3.5.1)

where $E_{\partial_1}$, $V_{\partial_1}$ and $C_{\partial_1}$ respectively denote the expectation, variance and covariance over $\partial_1$, and proved that any strategy $H_o$ in $H'_{\mu}$ with

i) $\pi_i = \mu X_i/X, \qquad i = 1, 2, \ldots, N,$

ii) $\mu_{(s)} = \mu$ for all $s \in S,$

iii) $t_o = t_{ht} = \sum_{i \in s} Y_i/\pi_i, \qquad (3.5.2)$

is $\partial_1$-optimum for any $\partial_1 \in \triangle_1$, where $H_\mu'$ belongs to $H_\mu$ and contains estimators belonging to the homogeneous linear class $(L_u^*)$ of unbiased estimators. This result has been extended to the class $A_u$ of entire unbiased estimators by Godambe and Joshi (1965). Hanurav (1962) and Vijayan (1966) showed that the result is true even for designs with expected effective sample size being constant for $L_u^*$ and $A_u$ classes respectively.

Hájek (1959) considered the class $\triangle_2$ of distributions $\partial_2$ for which (iii) in (3.5.1) is replaced by

iii) $C_{\partial_2}(Y_i, Y_j | X_i, X_j) = w(|j - i|) \qquad (3.5.3)$

where $w$ is single-valued convex function, and proved that any strategy $H_o$ in $H_\mu$ given by unequal probability systematic sampling with $\pi_i \propto X_i$ and HT estimator is $\partial_2$- optimum for any $\partial_2 \in \triangle_2$. Thus the scheme of selection is also specified in this case, however, the draw back of the scheme is that the variance of the estimator is not estimable. Therefore our choice falls on $\triangle_1$ but this model is of merely theoretical interest as in most practical situations the model where (ii) of (3.5.1) is replaced by

ii) ~~$V_{\partial_g}(Y_i | X_i) = \sigma^2 X_i^g, \quad g > 0$~~ (3.5.4)

$$\text{(ii)}' \quad V_{\partial_g}(Y_i|X_i) = \sigma^2 X_i^g, \quad g > 0 \tag{3.5.4}$$

is appropriate, where $g$ lies between 1 and 2 (Mahalanobis (1944), Jessen (1942)). The class of such apriori distributions $\partial_g$ will be denoted by $\triangle_g$. For this class of distributions we note the following result which has, however, been mentioned by Hanurav (1966, 68).

For the class of distributions $\triangle_g$ (with $g$ usually lying between 1 and 2) no $\delta_g$ - optimum sets of $\pi_i$'s exist if $g \neq 2$. And in this connection it is easily seen from Godambe and Joshi (1965) that for any estimator $t \in A_u$

$$E_{\partial_g} V(T) = \sigma^2 \sum_{i=1}^{N} X_i^g \left(\frac{1-\pi_i}{\pi_i}\right) + \phi \tag{3.5.5}$$

where $\phi$ is non-negative quantity and becomes zero only if the design is of constant sample size with $\pi_i \propto X_i$ and the estimator $t$ is the HT estimator. That is

$$E_{\partial_g} V(t_{ht}) = \sigma^2 \sum_{i=1}^{N} X_i^g \left(\frac{1-\pi_i}{\pi_i}\right) \tag{3.5.6}$$

for constant sample size designs with $\pi_i \propto X_i$. But minimising the RHS of (3.5.6) under the condition $\sum_{i=1}^{N} \pi_i = n$, we find that this term is minimised only if

$$\pi_i \propto X_i^{g/2} \tag{3.5.7}$$

which is contradictory to the condition for $\phi$ to be zero unless $g = 2$.

This leads to the choice of the strategies for different values of $g$. This is amply illustrated by the results of empirical studies given by Rao and Bayless (1969). Comparison of some strategies for two phase scheme under the model $\triangle_g$ is given in the next chapter.

Linear invariance: This concept introduced by Roy and Chakravarti (1960) requires an estimator to remain invariant under linear transformations of the vector y.

Regular class: This class is also discussed by Roy and Chakravarti. An estimator was said to belong to this class if its variance is proportional to the variance of the character under study in the population.

It is pertinent here to note that the linear invariance, though a desirable criterion for an estimator, does not lead, to an optimum estimator and the demand for an estimator to belong to the regular class seems to have no justification and no further work is traceable in this regard.

## 3.6 Hyper-admissibility:

This criterion for the choice of an optimum estimator was recently introduced by Hanurav (1966, 68). It is based on the concept of admissibility of an estimator and requires an estimator to be admissible not only in the whole space $R_N$ but also in each of its principal hyper-surfaces (phs), which are $(2^N-1)$ in number.

Let us denote the totality of the $(2^N-1)$ phs's by $\mathcal{R}_r$ and the phs of dimension r by $R_r$. There are $\binom{N}{r}$ such phs's of dimension r, for $1 \leq r \leq N$. The hyper-admissibility is then defined as follows:

Definition 3.6.1: In a class C(P) of unbiased estimators of Y, $t_1 \in C(P)$ is said to be hyper-admissible (h-admissible, for brevity) if it is admissible when the vector $y = (Y_1, Y_2, \ldots, Y_N)$ is restricted to any phs $R_r$ (not including any point of lower dimension) of $\mathcal{R}_r$.

Replacing the population total Y by any estimable parametric function $\theta(y)$ in the above definition we get the corresponding definition of h-admissibility of an estimator for estimating $\theta(y)$. We shall consider the h-admissibility of a variance estimator in section 3.6.2. This criterion is evidently stronger than admissibility but is weaker than uniform minimum variance. We define below an estimator belonging to the class of polynomial unbiased estimators of y before giving the result obtained by Hanurav (1968).

A <u>polynomial</u> estimator of $n^{th}$ degree is of the form

$$t_p^*(sy) = t^{(1)}(s,y) + t^{(2)}(s,y) + \dots + t^{(n)}(s,y)$$

(3.6.1)

where $t^{(i)}(s,y)$ is a homogeneous polynomial of degree $i$ (which may vanish identically for some or all $s$) in its arguments which are the $y$- values of only those units that occur in $s$, and $t^{(n)}(s,y)$ is non-zero for at least one sample $s$ with $P(s) > 0$. Further $t_p^*(s,y)$ will be unbiased for $Y$ if

$$E(t^{(i)}(s,y)) = 0\,, \quad \text{for } i \neq 1$$

and $$E(t^{(1)}(s,y)) = Y\,, \quad \text{for all } y. \qquad (3.6.1')$$

Further a <u>general polynomial</u> estimator $t_p(s,y)$ is defined as

$$t_p(s,y) = \alpha(s) + t_p^*(s,y) \qquad (3.6.2)$$

where $\alpha(s)$ are some constants depending only on $s$ and $t_p^*(s,y)$ is defined in (3.6.1). Thus $t_p(s,y)$ will be unbiased if, in addition to the condition (3.6.1'), we have

$$\sum_{s \in S} \alpha(s)\, P(s) = 0. \qquad (3.6.2')$$

The class of unbiased estimators defined as $t_p^*(s,y)$ and $t_p(s,y)$ will be denoted by $P_u^*$ and $P_u$, respectively.

Hanurav (1968) considered the class $P_u$ and proved the following theorem.

Theorem 3.6.1 (Hanurav): For any non-unicluster design P for which Y is estimable, the class $P_u$ contains a unique h-admissible estimator, which is the H.T. estimator given in (3.3.6).

The definition of uni-cluster design (which is due to Hanurav, 1966) is given below for ready reference.

Definition 3.6.2: A design P is said to be unicluster design if any two samples with positive probabilities are either disjoint or effectively equivalent, that is

$$s_1 \cap s_2 = \emptyset \quad \text{or} \quad s_1 \sim s_2. \qquad (3.6.3)$$

$P(s_1) > 0, \; P(s_2) > 0.$

Remark 3.6.1: Proof of the Theorem 3.6.1 may be divided into two main parts, namely,

i) For any non-unicluster design P, the H.T. estimator is the only possible h-admissible estimator of Y in the class $P_u$ and

ii) that in fact this estimator is h-admissible in $P_u$. Hanurav provided the proof of the first part by showing that for any non-unicluster design the H.T. estimator is the only estimator in the class $P_u$ which satisfies the <u>necessary condition</u>, mentioned in (3.1.3), for being admissible in each of the phs's of $\mathcal{R}_r$. And for the proof of the second part it is pointed out that the proof of Godambe and Joshi (1965, Theorem 4.1) for admissibility of the H.T. estimator in class $A_u$ for any given design P, in the whole parameter space $R_N$, can be trivially modified to show that the H.T. estimator is admissible in each of the phs's (not including points of lower dimensions) of $\mathcal{R}_r$ which implies h-admissibility of the H.T. estimator in the class $A_u$ and hence in $P_u$. This completes his proof.

<u>Remark 3.6.2:</u> It is important to note here that the proof of the first part may alternatively be accomplished by showing that for a given design P the H.T. estimator is <u>the best</u> estimator in the class C(P), in any <u>one</u> of the $(2^N-1)$ phs's, as this will exclude the possibility of any other estimator being admissible in that phs and hence from being h-admissible in the class C(P).

In this connection it may be seen that Godambe (1960), while proving admissibility of the H.T. estimator in $L_u^*$, the class of linear homogeneous unbiased estimators of Y, considered all vectors $y \in R_1$, the phs of dimension one, and in fact proved that

$$V(t_{ht}) < V(t), \quad \text{for all } y \in R_1, \qquad (3.6.4)$$

where $t_{ht}$ denote the H.T. estimator given in (3.5.6) and t is any other estimator belonging to the class $L_u^*$.

Thus Godambe's proof (as such) of admissibility of the H.T. estimator in fact establishes a more general result, that is the H. T. estimator is the best estimator in all the phs's, of dimension one, which are N in number. And hence from Remark 3.6.1 and 3.6.2 the unique h-admissibility of the H.T. estimator in the class $L_u^*$ follows for any design P. This result, however, is very much restricted as compared to Hanurav's result in theorem 3.6.1 in the sense that the class $P_u$ is much wider than $L_u^*$.

In the following section 3.6.1 we shall, however, extend Hanurav's result of theorem 3.6.1 to a wider class of estimators than $P_u$ following the alternative approach for proving the first part of the theorem, as mentioned in Remark 3.6.1. Then in section 3.6.2 we prove a similar result for a variance estimator of the H.T. estimator for a wide class of estimators.

3.6.1 Hyper-admissibility of the H.T. Estimator:

Let $t_h(s,y)$ be an estimator of the population total, continuous in its arguments which are the $\gamma$-values of only those units that occur in s, such that

$$t_h(s,y) = 0, \text{ if } Y_i = 0 \text{ for all } i \in s \qquad (3.6.5)$$

and let the class of these unbiased estimators be denoted by $H_u$. Evidently, this class is wide enough to include the class of homogeneous linear, quadratic, polynomial etc., estimators of Y. That is $H_u \supset P_u^* \supset L_u^*$ where $P_u^*$ is the class of polynomial unbiased estimators given in (3.6.1).

The class $H_u$ is however, restricted in the sense that it does not include all estimators of the class $P_u$ (or even $L_u$). Therefore, we consider a wider class of estimators than $H_u$ and denote it by $A_u^*$. An estimator t belonging to $A_u^*$ may be expressed as

$$t(s,y) = \alpha(s) + t_h(s,y) \qquad (3.6.6)$$

where $\alpha(s)$ do not depend on y but on s and $t_h(s,y)$ is an estimator belonging to $H_u$, given in (3.6.5).

The condition for unbiasedness of $t(s,y)$ in (3.6.6) is

$$\sum_{s \in S} \alpha(s)\, P(s) = 0$$

and

$$\sum_{s \in S} t_h(s,y)\, P(s) = Y \qquad (3.6.7)$$

for every $y \in R_N$.

Evidently, the class $A_u^*$ of unbiased estimators of Y is wider than $P_u$, the class of general polynomial unbiased estimators considered by Hanurav while proving theorem 3.6.1, since $A_u^* \supset P_u^*$. Now we prove the following theorem for the class $A_u^*$.

Theorem 3.6.2: For any non-unicluster design P in which Y is estimable, the H.T. estimator is uniquely h-admissible in $A_u^*$, the class of unbiased estimators given by (3.6.6).

Proof: Let us consider vectors $y \in y^{(i)}$, where

$$y^{(i)} = (0, 0, \ldots, Y_i, 0 \ldots 0) \qquad (3.6.8)$$

for $1 \leq i \leq N$, then evidently every $y \in y^{(i)}$ is a point in $R_1$, the phs of dimension one. Thus the population total

$$Y = Y_i, \quad \text{for all} \quad y \in Y^{(i)},$$

and the unbiasedness condition in (3.6.7) for an estimator $t(s,y) \in A_u$ becomes

$$\sum_{s \in S} \alpha(s) \, P(s) = 0 \qquad (3.6.9)$$

and
$$\sum_{s \in S_i} t_h(s,y) \, P(s) = Y_i, \quad \text{for all} \quad y \in y^{(i)}, \qquad (3.6.10)$$

since $t_h(s,y) = 0$ for samples $s \in S_i^*$, where

$$S = S_i \cup S_i^* \quad \text{and} \quad S_i = \{s, \; i \in s\}. \qquad (3.6.11)$$

We now first establish Lemma (3.6.1) and (3.6.2) given below.

Lemma 3.6.1: For any design P, the H.T. estimator is the best estimator in $H_u$, the class of unbiased estimators given by (3.6.8), for estimating the population total, for all $y \in y^{(i)}$.

Proof: We have for all $y \in y^{(i)}$,

$$V(t_{ht}) = \frac{Y_i^2}{\pi_i} - Y_i^2 \qquad (3.6.12)$$

from (3.3.8) and

$$V[t_h(s,y)] = \sum_{s \in S_i} t_h^2(s,y)\, P(s) - Y_i^2 \qquad (3.6.13)$$

since from the definition

$$t_h(s,y) = 0, \text{ for all } s \in \bar{S}_i \text{ and } y \in y^{(i)}. \qquad (3.6.14)$$

Now using the condition of unbiasedness in (3.6.10), we get

$$V[t_h(s,y)] - V(t_{ht}) = \sum_{s \in S_i} P(s)\,[t_h(s,y) - \frac{Y_i}{\pi_i}]^2 > 0$$

(3.6.14 )

since $t_{ht} \neq t_h(s,y)$, for at least one $y$.

Hence the Lemma.

Lemma 3.6.2: For any non-unicluster design in which $Y$ is estimable and $\pi_i < 1$, the class $H_u$ is complete in $A_u^*$ for estimating $Y$, for all $y \in y^{(i)}$.

<u>Proof:</u> Suppose $H_u$ is not complete in $A_u^*$ for $y\epsilon y^{(i)}$ then evidently there exist an estimator $t(s,y) \in (A_u^* - H_u)$ which for $y\epsilon y^{(i)}$ can be expressed as

$$t(s,y^{(i)}) = \alpha(s) + t_h(s,y^{(i)}), \quad \text{if } s\in S_i$$
$$= \alpha(s), \quad \text{if } s\in S_i^* \tag{3.6.15}$$

and which satisfies the necessary conditions mentioned in 3.1.3 for being admissible.

Further, noting that $S_i$ and $S_i^*$ are effectively equivalent samples for all $y\in y^{(i)}$, we get from (3.6.15) and (3.1.3)

$$t_s(i) = \alpha_1(s^i) + t_h(s^i,y^{(i)}), \quad \text{if } s\in S_i$$
$$= \alpha_2(i), \quad \text{if } s\in S_i^* \tag{3.6.16}$$

where $\alpha_2(i)$ is independent of $s$ and $s^i$ stands for samples in $S_i$.

Since $i$ is arbitrary, (3.6.16) is true for $1 \leq i \leq N$. Now we note that $\alpha_1(s^i) = \alpha_1(i)$, that is $\alpha(s^i)$ is also independent of $s\in S_i$ since every $s\in S$ also belongs to $\bigcup_{i=1}^{N} S_i^*$, $\pi_i$ being less than one, and that $\alpha(s)$ is independent of $s\in S_i^*$, $1\leq i \leq N$ from (3.6.16). Further from the condition of unbiasedness of $t_s(i)$ as well it is easily observed that $\alpha_1(s^i) = \alpha_1(i)$ and that infact

$$\alpha_1(i) = \alpha_2(i)\left(\frac{1-\pi_i}{\pi_i}\right). \tag{3.6.17}$$

Hence, for all samples $s$ with $P(s) > 0$, we have

$$\alpha(s) = \alpha_1(i), \quad \text{if } s \in S_i$$
$$= \alpha_2(i), \quad \text{if } s \in S_i^c$$

and similarly

$$\alpha(s) = \alpha_1(j), \quad \text{if } s \in S_j$$
$$= \alpha_2(j), \quad \text{if } s \in S_j^c. \qquad (3.6.18)$$

Now the lemma will follow if

$$\alpha_1(i) = \alpha_2(i) = 0, \quad 1 \leq i \leq N. \qquad (3.6.19)$$

To show this we follow the argument given by Hanurav (1968, p. 629), that is, use the property of non-unicluster design in which there exist at least one pair $(i_o, j_o)$ such that

$$0 < \pi_{i_o j_o} < \pi_{i_o}. \qquad (3.6.20)$$

The right side of (3.6.20) implies that there is one sample $s \in S_i S_j^c$, for which $P(s) > 0$, and this gives

$$\alpha_2(j) = \alpha_1(i). \qquad (3.6.21)$$

Similarly, the left side of (3.6.20) implies that there is a sample $s \in S_i S_j$, for which $P(s) > 0$, and this gives

$$\alpha_1(i) = \alpha_1(j). \tag{3.6.22}$$

Hence from (3.6.21) and (3.6.22), we get

$$\alpha_1(j) = \alpha_2(j).$$

This implies that

$$\alpha(s) = \alpha, \text{ say, for } P(s) > 0$$

which from the condition of unbiasedness in (3.6.9) implies that

$$\alpha(s) = 0, \quad \text{for } P(s) > 0.$$

Hence the lemma.

Now the lemma 3.6.1 and 3.6.2 implies that for any non-unicluster design the H.T. estimator is the best estimator of the population total in the class $A_u^*$ for all $y \in y^{(1)}$. Further from Remark 3.6.2 it follows that the H.T. estimator is the <u>only possible</u> h-admissible estimator in $A_u^*$ and that in fact it h-admissible follows from argument in Remark 3.6.1.

This complete the proof of our theorem.

## 3.6.2 <u>Hyper-admissibility of a Variance Estimator:</u>

The sampling variance of the estimator $t_{ht}$ is

$$V(t_{ht}) = \sum_{i \in U} \left(\frac{1-\pi_i}{\pi_i}\right) Y_i^2 + \sum_{i} \sum_{j(\neq i) \in U} \frac{Y_i Y_j}{\pi_i \pi_j} (\pi_{ij} - \pi_i \pi_j). \qquad (3.6.26)$$

Two unbiased estimators of $V(t_{ht})$ well-known in the literature are due to Horvitz and Thompson (1952) and Yates and Grundy (1953) and Sen (1953). The estimator due to former authors is given by

$$v_{ht} = \sum_{i \in s} \left(\frac{1-\pi_i}{\pi_i^2}\right) Y_i^2 + \sum_{i (\neq j) \in s} \sum \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}}\right) \frac{Y_i Y_j}{\pi_i \pi_j} \qquad (3.6.27)$$

where $\pi_i$ and $\pi_{ij}$ are inclusion probabilities of $i$ and $(i,j)$ defined earlier.

Later Yates and Grundy (1953) criticised $v_{ht}$ on three grounds viz., that it is not reducible to a linear function of squares of the difference between the $(Y_i/\pi_i)$'s, (ii) it does not vanish when all $(Y_i/\pi_i)$'s are equal and (iii) it takes negative values rather often. They further noted that for fixed sample size the variance of $t_{ht}$ in (3.6.26) can be expressed as

$$V(t_{ht}) = \sum_{i \neq j \in U} \sum (\pi_i \pi_j - \pi_{ij}) \left(\frac{Y_i}{\pi_i} - \frac{Y_j}{\pi_j}\right)^2 \qquad (3.6.28)$$

and consequently provided an unbiased estimator of $V$ in (3.6.28), for fixed sample size designs, which is given by

$$v_{yg} = \sum_{i \neq j \in s} \sum \left(\frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}}\right) \left(\frac{Y_i}{\pi_i} - \frac{Y_j}{\pi_j}\right)^2, \qquad (3.6.29)$$

which assumes negative values less often.

As the variance is a non-negative quantity it is very desirable to demand that its estimator should also be non-negative. It is easily seen that no sufficient condition can be given to make $v_{ht}$ a non-negative estimator, while for $v_{yg}$ we have a simple condition under which this estimator is non-negative namely

$$\pi_{ij} \leq \pi_i \pi_j, \quad \text{for all } i \text{ and } j. \qquad (3.6.30)$$

It is pertinent to point out that several sampling schemes have also been proposed by different authors (Sen, 1953; Des Raj 1956; Rao, 1961; Brewer, 1963; Seth, 1966; Hanurav, 1967, and others) for which $v_{yg}$ is always positive.

Thus $v_{yg}$ has several advantages over $v_{ht}$ as regards its practical utility. The estimator $v_{ht}$ is, however, known to be admissible (Godambe and Joshi, 1965, Theorem 4.3) in the entire class of unbiased estimators of $V(t_{ht})$.

It is also known from Godambe and Joshi (1965, corollary 4.2) that there does not exist a best estimator of $V(t_{ht})$ in the class of all unbiased estimators of $V(t_{ht})$, hence in the following we proceed to examine whether there exist a unique h-admissible variance estimator for estimating $V(t_{ht})$, like $t_{ht}$ for Y, in wide classes of unbiased estimators of $V(t_{ht})$. We shall first consider the class of unbiased estimators of $V(t_{ht})$ denoted by $H_{uv}$ (v for variance). An unbiased estimator of $V(t_{ht})$ denoted by $v_h(s,y)$, continuous in its arguments which are the $\mathcal{Y}$-values of only those units that occur in $s$, belonging to the class $H_{uv}$ is such that

$$v_h(s,y) = 0, \text{ if } Y_i = 0 \text{ for all } i \in s, \tag{3.6.31}$$

and for unbiasedness

$$\sum_{s \in S} v_h(s,y)\, P(s) = V(t_{ht}), \text{ for all } y \in R_N. \tag{3.6.32}$$

Using the definition of h-admissibility of a variance estimator as given in definition 3.6.1 with Y replaced by $V(t_{ht})$, we prove the following theorem for the class $H_{uv}$.

<u>Theorem 3.6.3:</u> For any design P in which $V(t_{ht})$ is estimable, the variance estimator $v_{ht}$ in (3.6.27) is <u>uniquely h-admissible</u> in $H_{uv}$, the class of unbiased estimator of $V(t_{ht})$ given by (3.6.31).

We first establish the following.

<u>Lemma 3.6.3:</u> For any design P, in which $V(t_{ht})$ is estimable, the variance estimator $v_{ht}$ is <u>the best</u> estimator of $V(t_{ht})$ in the class $H_{uv}$, for all $y \in y^{(1)}$.

<u>Proof:</u> We note the following,

$$V(t_{ht}) = \sum_{i=1}^{N} a_i Y_i^2 + \sum_{i \neq j}\sum a_{ij} Y_i Y_j \tag{3.6.33}$$

where $a_i = \frac{1-\pi_i}{\pi_i}$ and $a_{ij} = (\frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j})$.

from (3.6.26) and

$$V(v_{ht}) = E(v_{ht}^2) - [V(t_{ht})]^2$$

$$= \sum_{i=1}^{N} \frac{a_i^2 Y_i^4}{\pi_i} + \phi_{ijk}(y) - [V(t_{ht})]^2 \qquad (3.6.34)$$

where $\phi_{ijk}(y)$ is function of $(Y_iY_j), (Y_jY_k)$ etc., such that it is zero for $y \in y^{(i)}$ and

$$V[v_h(s,y)] = \sum_{s \in S} v_h^2(s,y)\ P(s) - [V(t_{ht})]^2. \qquad (3.6.35)$$

Hence for the vectors $y \in y^{(i)} = (0,0,\ldots,Y_i,\ldots 0)$, we get from above,

$$V(t_{ht}) = a_i Y_i^2 \qquad (3.6.36)$$

$$V(v_{ht}) = \frac{a_i^2 Y_i^4}{\pi_i} - a_i^2 Y_i^4 \qquad (3.6.37)$$

and $$V[v_h(s,y)] = \sum_{s \in S_i} v_h^2(s,y)\ P(s) - a_i^2 Y_i^4 ; \qquad (3.6.38)$$

and from the unbiasedness condition of $v_h(s,y)$ in (3.6.32) we have

$$\sum_{s \in S_i} v_h(s,y)\ P(s) = a_i Y_i^2 , \qquad (3.6.39)$$

since $v_h(s,y) = 0$ for all $s \in S_1^*$ from the definition.

Thus for all $y \in y^{(1)}$, we have from (3.6.37) and (3.6.38)

$$V[v_h(s,y)] - V[v_{ht}] = \sum_{s \in S_1} [v_h^2(s,y) - \frac{a_1^2 Y_1^4}{\pi_1^2}]P(s)$$

$$= \sum_{s \in S_1} [v_h(s,y) - \frac{a_1 Y_1^2}{\pi_1}]^2 \quad \text{from } 3.6.39$$

$$\geq 0, \tag{3.6.40}$$

since $v_h(s,y) \neq v_{ht}$ for at least one $s$.

Hence the lemma.

<u>Proof of the Theorem</u>: From the result in the above lemma it follows (from Remark 3.6.2) that for any design P, the variance estimator $v_{ht}$ is the only possible h-admissible estimator in the class $H_{uv}$.

That in fact $v_{ht}$ is h-admissible in $H_{uv}$ easily follows by trivially modifying the proof of Godambe and Joshi (1965) given for admissibility of $v_{ht}$ in the entire class of unbiased estimators of $V(t_{ht})$.

Since this is all that is required for the h-admissibility of $v_{ht}$, we conclude that $v_{ht}$ is the unique h-admissible estimator in $H_{uv}$.

This complete the proof of our theorem.

Next, we consider a wider class of estimators than $H_{uv}$, denoted by $A_{uv}^*$ (v for variance), which is parallel to $A_u^*$

considered in (3.6.6) for estimating $Y$. An estimator belonging $A^*_{uv}$ may be expressed as

$$v(s,y) = \alpha(s) + v_h(s,y) \qquad (3.6.41)$$

and the unbiasedness condition of $v(s,y)$ is

$$\sum_{s \in S} \alpha(s)\, P(s) = 0$$

and

$$\sum_{s \in S} v_h(s,y)\, P(s) = V(t_{ht}) \qquad (3.6.42)$$

for every $y \in R_N$, where $\alpha(s)$ is independent of $y$-values and $v_h(s,y)$ is an estimator of $V(t_{ht})$ belonging to $H_{uv}$. We now prove the following.

<u>Theorem 3.6.4:</u> For any design $P$ in which $V(t_{ht})$ is estimable, the variance estimator $v_{ht}$ is uniquely h-admissible in $A^*_{uv}$ the class of unbiased estimators of $V(t_{ht})$ given in (3.6.41).

We first establish the following.

<u>Lemma 3.6.4:</u> For any design $P$ in which $V(t_{ht})$ is estimable, the class $H_{uv}$ is <u>complete</u> in $A^*_{uv}$ for estimating $V(t_{ht})$, for all $y \in y^{(1)}$.

<u>Proof:</u> Suppose $H_{uv}$ is not complete in $A^*_{uv}$ for $y \in y^{(1)}$ then there exist an estimator $v(s,y)$ in $(A^*_{uv} - H_{uv})$ which for $y \in y^{(1)}$ can be expressed as

$$\begin{aligned} v(s,y) &= \alpha(s) + v_h(s,y), \quad \text{if } s \in S_1 \\ &= \alpha(s), \quad \text{if } s \in S_1^* \end{aligned} \qquad (3.6.43)$$

for all $y \epsilon y^{(i)}$ and which satisfies the necessary condition for being admissible.

Further, since $S_i$ and $S_i^*$ are effectively equivalent samples for $y \epsilon y^{(i)}$, we get from above (and (3.6.17)) using the condition of unbiasedness

$$v_s(i) = \alpha_1(i) + v_h(s^i, y^{(i)}), \text{ if } s \epsilon S_i$$
$$= \alpha_2(i), \qquad \text{ if } s \epsilon S_i^* \tag{3.6.44}$$

where $\alpha_1(i)$ and $\alpha_2(i)$ are independent of $s$ and $y$.

Now proof of the Lemma will be complete if

$$\alpha_1(i) = \alpha_2(i) = 0, \quad 1 \leq i \leq N. \tag{3.6.45}$$

But this condition is same as (3.6.19) and hence remembering that $V(t_{ht})$ is estimable only for non-unicluster designs that is only if

$$0 < \pi_{ij} < \pi_i, \tag{3.6.46}$$

which is same as (3.6.20), the proof of this condition (3.6.45) follows from the proof given in Lemma 3.6.2.

Hence the Lemma.

<u>Proof of the theorem</u>: The above Lemma 3.6.4 together with Lemma 3.6.3 establishes that the variance estimator $v_{ht}$ is the best estimator of $V(t_{ht})$, for all $y \epsilon y^{(i)}$, in the class $A^*_{uv}$ for any design in which $V(t_{ht})$ is estimable. Hence these

lemmas together with Remark 3.6.2 shows that $v_{ht}$ is the only possible h-admissible estimator of $V(t_{ht})$ in $A^*_{uv}$. And that this estimator is in fact h-admissible follows from the argument given the proof of Theorem 3.6.3.

Since this is all that is required for the h-admissibility of $v_{ht}$, we conclude that $v_{ht}$ is unique h-admissible in $A^*_{uv}$.

This complete the proof of theorem.

Remark 3.6.3. The Theorems 3.6.2 to 3.6.4 may also be proved following the approach, mentioned in Remark 3.6.1, used by Hanurav while establishing his result in Theorem 3.6.1.

Remark 3.6.4: From the results in this sub-section it follows that if we decide to choose an estimator of $V(t_{ht})$, restricting ourselves to the class $A^*_{uv}$, on the criteria of unbiasedness and h-admissibility alone (with respect to variance as a loss function) then we have to discard all other estimators belonging to $A^*_{uv}$ (including $v_{yg}$) save the estimator $v_{ht}$.

## 3.7 Necessary-Best Estimator:

This criterion for choice of an estimator has been introduced by Prabhu Ajgaonkar (1965). It is defined as follows:

Between two unbiased estimators $t$ and $t'$ with variances

$$V(t) = \sum_{i=1}^{N} A_i Y_i^2 + \sum_{i \neq j} A_{ij} Y_i Y_j \qquad (3.7.1)$$

and $V(t') = \sum_{i=1}^{N} B_i Y_i^2 + \sum_{i \neq j} B_{ij} Y_i Y_j$ (3.6.2)

the estimator $t$ is said to be a <u>necessary better estimator</u> than $t'$ if

$$A_i < B_i, \quad \text{for } 1 \leq i \leq N. \qquad (3.7.3)$$

Now if, for a given design $P$, an unbiased estimator $t_1$, in a class $C(P)$ of unbiased estimators, is a necessary better estimator than every other estimator in $C(P)$, then $t_1$ is termed as the necessary best estimator for the class $C(P)$.

Prabhu Ajgaonkar (1965) considered a sub-class of linear estimators termed $T_5$-class based on the features inherent in the $T_2$ and $T_3$ classes (given in Chapter II):

$$T_5 = \phi_{ij \ldots 1 \ldots m}\left(\frac{Y_i}{\lambda_i} + \frac{Y_j}{\lambda_j} + \ldots + \frac{Y_m}{\lambda_m}\right)$$

where $\phi_{ij \ldots 1 \ldots m}(i \neq j \neq \ldots \neq m = 1,2,\ldots N)$ is a constant used as weight associated with the sample consisting of the $i^{th}, j^{th} \ldots m^{th}$ elements and $\lambda_i$ $(i = 1,2,\ldots,N)$ is a constant used as weight whenever the $i^{th}$ unit of the population appears in the sample. And proved that for this class of estimators the H.T. estimator $(t_{ht})$ is the necessary best estimator of the population total.

Recently, Hege (1967) extended this result by showing that the H.T. estimator is the necessary best estimator in $L_u^*$, the class of all homogeneous linear unbiased estimators of the population total.

Remark 3.7.1: It is important to note that for proving necessary bestness of an estimator $t$ in a given class of unbiased estimators it is enough to show that the estimator $t$ is the best estimator for that class for all vectors $y \epsilon y^{(1)} = (0, 0, \ldots, Y_i, 0 \ldots 0)$, $1 \leq i \leq N$, since for such vectors the second term in rhs of $V(t)$ in (3.7.1) becomes zero giving rise to the condition for the bestness of $t$ to be identical with the requirement given in (3.7.3) for its necessary bestness. That is the criterion of necessary bestness is equivalent to the criterion of the best estimator when the parameter is restricted to the phs's of one dimension, the phs which played the vital role in the proof of unique h-admissibility of the H.T.estimator and the variance estimator $v_{ht}$.

Remark 3.7.2: In view of the above remark, it is pertinent to point out that the result established by Hege (and hence by Prabhu Ajagaonkar) follows from the proof given by Godambe (1960) itself.

In view of the above remarks and on the basis of the results obtained for the vector $y \epsilon y^{(1)}$ in the earlier section we have the following.

Theorem 3.7.1: For any design $P$ in which $Y$ is estimable, the H.T. estimator is the necessary best estimator in $H_u$, the class of unbiased estimators of $Y$ given by (3.6.5).

Proof of the Theorem follows from Lemma 3.6.1 and the above Remark 3.7.1.

In fact, from Lemma 3.6.1, 3.6.2 and the Remark 3.7.1 it follows that for any non-unicluster design the H.T. estimator is the necessary best estimator in the wider class $A^*_{uv}$.

Now if the criterion of necessary bestness is extended to the estimation of $V(t_{ht})$, retaining the essential feature of the criterion, namely, the coefficient of $f(Y_i) = Y_i^4$ being least for the necessary best estimator from among the estimators in a given class of unbiased estimator of $V(t_{ht})$, then the Remark 3.7.1 remains valid even for estimating $V(t_{ht})$. And from the Lemma 3.6.3 we get the following.

<u>Theorem 3.7.2</u>: For any design P, in which $V(t_{ht})$ is estimable the variance estimator $v_{ht}$ is the necessary best estimator in $H_{uv}$, the class of unbiased estimators of $V(t_{ht})$ given by (3.6.31).

The extension of this result for the class $A^*_{uv}$ follows immediately from Lemma 3.6.4.

Having recognised the importance of the part played by the vectors belonging to the phs of dimension one, that is $y\epsilon R_1 (y \epsilon y^{(1)})$ in both the criterion, namely h-admissibility and the necessary bestness, together with unbiasedness we give below the following.

<u>Theorem 3.7.3</u>: For any design P in which Y is estimable the necessary bestness of an estimator implies unique h-admissibility of that estimator for the class $H_u$ of estimators of Y and that it is the H.T. estimator which uniquely satisfies both the criterion for the class $H_u$.

Proof of the theorem follows from Lemma 3.6.1 Theorem 3.7.1 and Remarks 3.6.1, 3.6.2 and 3.7.1.

In fact the truth of the above theorem follows for a wider class of estimators $A^*_u$ for any non-unicluster design from Lemma 3.6.2 and the above Theorem. And similar result is easily seen to hold for $v_{ht}$ from among the estimators of the variance of the H.T. estimator belonging to $A^*_{uv}$.

3.8 Concluding Remarks:

Since a best estimator does not generally exist for wide classes of estimators it is logical, as the first step, to reduce the class of estimators without loss of information. That is, the class can be reduced to the extent that for each estimator outside the reduced class there is one inside it which is at least as good as (or better than) that estimator and further in the reduced class no estimator is not at least as good as the other. In this sense, the minimal essentially complete class which coincides with the class of essentially admissible estimators, provides the maximum possible reduction of the class of estimators, without loss of information, with the mean square error as the loss function. However, as the criterion of unbiasedness is generally taken for granted in the field of sample surveys and is very appealing in the sense that without this even a 'constant' (Theorem 3.3.2) is also included in the minimal essentially complete class, we may further reduce this class by confining ourselves to the unbiased estimators only. Introduction of any criterion to shrink this class further, without using any apriori information about the parameter, with the aim to pin point a unique unbiased estimator of the parameter can not lead us far in estimating Y since such unique estimators will be preferred over all other estimators only for the purpose for which the criterion has been introduced provided, however, it has a practically justifiable purpose.

Hanurav (1965, 1968 , p. 629) while introducing the criterion of h-admissibility, which gives rise to the unique estimator (Theorem 3.6.1 and 3.6.2) for estimating the population total, has clearly provided the justification for the introduction of this criterion whereas no such justification is available in the literature for the introduction of necessary bestness except that of getting a unique estimator. Hence if the purpose is two-fold namely estimation of the population total as well as estimation of all linear parametric functions (such as totals for each of the sub-populations) the H.T. estimator, is naturally the best (unique) choice, when we decide to use an estimator which is unbiased and admissible for each of the linear parametric functions (sub-population totals) with respect to variance as the loss function.

However, limitations of our choice of an estimator based on unbiasedness and h-admissibility, with respect to variance as the loss function, becomes apparent from the fact that the estimator $v_{ht}$ (which is known to be a bad estimator for reasons mentioned in section 3.6) is uniquely h-admissible in a wide class of unbiased estimators of $V(t_{ht})$; and also from the vital role played by vectors $y$ belonging to the phs's of dimension one. Further, it is felt that in actual practice we rarely need estimation of sub-population totals for <u>each</u> of the sub-populations ($2^N-1$ in number) and particularly the sub-populations consisting of single units. Hence in this connection it is of considerable interest to modify the h-admissibility criterion so as to exclude the vectors $y$ belonging to the phs's of dimension one (the phs's which played vital role in proving unique h-admissibility) (or more) and to examine whether H.T. estimator remains uniquely h-admissible (as per modified definition) for estimating $Y$ and to see what happens to $v_{ht}$ in that case.

As regards Bayes solution, where we use apriori information about the parameter, it is pertinent to note that the solution heavily depends on the apriori distribution, based on prior information which however may not be usually reliable, if available at all. And the estimator may not remain optimum if the guessed apriori distribution, departs even slightly from the actual ([illegible]) that is if $\sigma_i^2 \neq \sigma^2 X_i^2$ then even if $\sigma_i^2 = \sigma^2 X_i^g$, $g \geq 0$ (that is (3.5.4) is valid), the H.T.estimator ceases to be the optimum and in fact no optimum exist unless $g$ is exactly equal to 2. This is amply illustrated by the results of empirical studies given by Rao and Bayless (1968a).

From what has been said above, it appears that restriction of the class of estimators in a meaningful way is possible only if other considerations in addition to the concept of variance are taken into account. For instance the cost aspect is a very important consideration in practical situations and hence the choice of an estimator may be made to depend more on efficiency per unit cost than just on the sampling efficiency (Mahalanobis, 1944, United Nations, 1948). Further, since in practice the interest lies not only in getting stable estimator but also in having stable variance estimator, another consideration for the choice of estimators may be efficiency of the variance estimator (Rao and Bayless, 1968a, b). Further, whatever be the criterion adopted for the choice of estimators it would be necessary to carry out extensive empirical studies on natural populations and types of distributions which they conform to with a view to spotting out reasonably good estimator(s) for different situations commonly met with in practice.

# CHAPTER IV

## SOME TWO-PHASE VARYING PROBABILITY SCHEMES

4.0 Summary. In this chapter some well-known pps without replacement schemes, originally given for uni-phase sampling, have been considered for selection of a second-phase sample in two-phase sampling. Estimators and the corresponding variances are obtained in Section 4.2 and in Section 4.3 these schemes are compared under an appropriate super-population model.

4.1 Three Two-phase Schemes. In the method of two-phase sampling, two samples are selected one of which is a sub-sample of the other. The larger sample is called the first-phase sample (fps) and the smaller one the second-phase sample (sps). In some cases, however, the sps is not a sub-sample of the fps but it is selected independently from the population. Information on certain supplementary character x which can be observed at lower cost per unit than the main character y under study is collected from all the units in the fps and utilised either in the selection of sps or in building up estimators which improve upon the usual unbiased estimator of the study character, based on sps. This method, therefore, is suited to the situations where collection of information on y is costly and the population values for highly correlated character x are not known beforehand.

Neyman (1938) has considered the use of such information x collected for the fps in stratification where as Cochran (1942) considered its use in ratio and regression methods of estimation. These methods are now well-known and the theoretical developments can be found in the standard text-books (Cochran (1963), Murthy (1967) etc.). Use of fps information in selection of the sps has been considered by Saito (1958), Des Raj (1964) and D. Singh and B. D. Singh (1965), using probability proportionate to size (pps) with-replacement scheme, the size being information on x collected from fps.

In this chapter we consider the three pps without-replacement schemes, (with associated estimators) for the selection of sps, due to (i) Des Raj (1956), (ii) Hartley and Rao (1962), (iii) Rao, Hartley and Cochran (1962).

These schemes, as is well-known, have been considered by the respective authors for uni-phase sampling only. In the present case we consider selection of fps by simple random sampling without-replacement, for collecting data on x and sampling of n of these N' units in the second-phase following any of the three schemes mentioned above.

An unbiased estimator T of the population total $Y = \Sigma Y_j$ in two-phase sampling will be of the form

$$T = (N/N')t \qquad (4.1.1)$$

where $t$ is an unbiased estimator of the fps total $Y' = \Sigma_1 Y_j$, where $\Sigma$ and $\Sigma_1$ denote the summations over all the units in the population and the fps respectively. It is easy to verify that if $t$ is unbiased for $Y'$ then $T$ is so for $Y$.

The sampling variance for $T$ can be written as

$$V(T) = V_1 E_2(T) + E_1 V_2(T) \qquad (4.1.2)$$

where $E_1$ and $V_1$ are unconditional and $E_2$ and $V_2$ are the conditional (given the fps) expectation and variance respectively (Des Raj, 1956).

It may be mentioned that for the schemes considered here, the first term in (4.1.2) is the same and is given by

$$V_1 E_2(T) = N(N - N')S_y^2 / N' \qquad (4.1.3)$$

where

$$S_y^2 = (N-1)^{-1} \Sigma (Y_j - \frac{Y}{N})^2. \qquad (4.1.4)$$

Again for these schemes, $V(T)$ in (4.1.2) can be expressed as

$$V(T) = V(T_{d'}) - V' \qquad (4.1.5)$$

where

$$V(T_{d'}) = \left(\frac{N-N'}{N'}\right)S_y^2 + \frac{N}{N-1}\,\frac{N'-1}{N'}\,\frac{1}{n}\,\Sigma\, P_j\left(\frac{Y_j}{P_j} - Y\right)^2 \qquad (4.1.6)$$

is the sampling variance of an unbiased estimator

$$T_{d'} = \left(\frac{N}{N'}\right)\frac{1}{n}\,\Sigma_2\,\frac{Y_j}{P'_j}\,. \qquad (4.1.7)$$

given by Des Raj (1964), when the second-phase sample is selected as a sub-sample of the first-phase sample following pps with-replacement procedure of selection. $\Sigma_2$ denotes the summation over sps units and $P'_j = X_j / \Sigma_1 X_j$ is the probability of selection of the j-th fps unit and $P_j$ is defined as $X_j/\Sigma X_j$. Thus efficiency of these three schemes depend upon the magnitude of $V'$ in (4.1.5). It may be mentioned that the sampling variance obtained by the respective authors in the case of uni-phase sampling represent the conditional variance of $t$ in the present situation and hence we evaluate $E_1V_2(T)$ for different schemes in the following section and express the corresponding variance in the form (4.1.5) for easy comparisons.

## 4.2 Estimators and Variances:

(i) Des Raj strategy. For a pps without replacement sps of size $n$ from fps an unbiased estimator of $Y$ from Des Raj (1956), is

$$T_d = \left(\frac{N}{N'}\right)\frac{1}{n}\,\Sigma_2\, t_j \qquad (4.2.1)$$

where $t_1 = \frac{y_1}{P_1'}$ and

$$t_j = (y_1+y_2+\dots+y_{j-1}+y_j/P_j' \,(1-P_1'\dots P_{j-1}')$$

for $j = 2,3,\dots, n.$

From Pathak (1967), an upper bound of the conditional variance of $T_d$ can be expressed as

$$V_2(T_d) \leq \frac{N^2}{nN'^2} \Sigma_1 (\frac{Y_j}{P_j'} - Y')^2 P_j' - \frac{N^2}{N'^2}(\frac{n-1}{2n}) [ \Sigma_1 P_j'^2 \Sigma_1 P_j' (\frac{Y_j}{P_j'} - Y')^2$$

$$+ \Sigma_1 P_j'^2 (\frac{Y_j}{P_j'} - Y')^2 ] \qquad (4.2.2)$$

to the order of $O(N'^{-1})$, assuming that $\max P_j' = O(N'^{-1})$. Evaluating $E_1$ of $V_2(T_d)$ term by term, we get,

$$E_1(t_1) = (\frac{N}{N'})^2 \frac{1}{n} E_1(X' \ \Sigma_1 \frac{Y_j^2}{X_j} - Y'^2)$$

$$= \frac{N}{N'} \ \frac{N'-1}{N-1} V_{pps} \qquad (4.2.3)$$

where

$$V_{pps} = \frac{1}{n} \Sigma P_j' \ (\frac{Y_j}{P_j'} - Y)^2$$

is the variance based on a pps sample of size $n$ selected directly from the population.

$$E_1(t_2) = \frac{N^2(n-1)}{N'^2 2n} E_1[\ \Sigma_1 Y_j^2 P_j' - Y'^2 \Sigma_1 P_j'^2 + \sum_{j \neq j'} \frac{Y_1^2}{P_{j'}'} P_{j'}'^2\ ]$$

$$= \frac{N(N'-1)}{N'(N-1)}\left(\frac{n-1}{2n}\right)\left[\frac{N}{N'} \Sigma P_j^2 \Sigma P_j\left(\frac{Y_1}{P_j} - Y\right)^2 - \left(\frac{N}{N'}-1\right)\Sigma Y_j^2 - Y^2 \Sigma P_j^2\right] \qquad (4.2.4)$$

to the order of $O(N^{-1})$

$$E_1(t_3) = \frac{N}{N'} \frac{n-1}{2n} \Sigma P_j^2\left(\frac{Y_1}{P_j} - Y\right)^2 \qquad (4.2.5)$$

to the order of $O(N'^{-1})$, where $t_1, t_2$ and $t_3$ are the first, second and third terms in (4.2.2).

Hence from (4.2.3) to 4.2.5) and (4.1.2), we get

$$V(T_d) \leq V(T_{d'}) - \frac{n-1}{2n} \frac{N}{N'}\left[\Sigma P_j^2 \Sigma P_j\left(\frac{Y_1}{P_j} - Y\right)^2 + \frac{N-N'}{N'} \Sigma Y_j^2 P_j + \Sigma P_j^2\left(\frac{Y_1}{P_j} - Y\right)^2\right)\right] \qquad (4.2.6)$$

to the order of $O(N'^{-1})$.

(ii) Hartley and Rao strategy. In this scheme for selecting a pps sample, the population units are listed in random order, their sizes are cumulated and a systematic selection of units from a random start is made on cumulation. However, for the present situation this method will consist in just cumulating the sizes of the fps units and selecting a pps sample systematically with a random start. It is assumed that $nP_j' \leq 1$ to avoid any repetition of units. An unbiased estimator of the population total $Y$ is

$$T_{hr} = \frac{N}{N'} \sum_{j=1}^{n} \frac{y_j}{\pi_j'} \qquad (4.2.7)$$

where $\pi_j'$ is the inclusion probability of the j-th fps unit. The estimator $T_{hr}$ is essentially the HT estimator considered earlier.

From Hartley and Rao the conditional variance of $T_{hr}$ is given by

$$V_2(T_{hr}) = (\frac{N}{N'})^2 \frac{1}{n}[\ \Sigma_1 P_j'(\frac{Y_j}{P_j'} - Y')^2 - (n-1)\Sigma_1 P_j'^2(\frac{Y_j}{P_j'} - Y')^2 \qquad (4.2.8)$$

assuming max $P_j' = O(N'^{-1})$. This formula is valid for large values of $N'$ and values of $n$ relatively smaller.

We get from (4.1.2), (4.1.3) and (4.2.8)

$$V(T_{hr}) = V(T_{d'}) - \frac{n-1}{n} \frac{N}{N'} \Sigma P_j^2(\frac{Y_j}{P_j} - Y)^2 \qquad (4.2.9)$$

where $V(T_{d'})$ is defined in (4.1.6).

(iii) Rao, Hartley and Cochran strategy. This scheme, for selecting a sps of units, in the present situation, will consist of

i) forming $n$ groups of the fps units (as they occur) having sizes $N_1', N_2' \ldots N_n'$, such that $\sum_{i=1}^{n} N_i' = N'$ and

ii) drawing one unit from each group with the probability proportionate to the X-values of units collected in fps.

An unbiased estimator of the population total $Y$ is then given by

$$T_{rhc} = \frac{N}{N'} \Sigma_2 \frac{Y_i}{X_j / \phi'_j} \tag{4.2.10}$$

where $\phi'_j$ is defined as the sum of all $X_j$ for which $U_i$ $(i = 1,2,\ldots, N')$ belongs to the $j$-th group.

The conditional variance of the estimator is

$$V_2(T_{rhc}) = (1 - \frac{n-1}{N'-1}) \frac{N^2}{nN'^2} \Sigma_1 P'_j (\frac{Y_i}{P'_j} - Y')^2 \tag{4.2.11}$$

if $N'$ is a multiple of $n$ and

$$= [1 - \frac{n-1}{N'-1} + \frac{k(n-k)}{N'(N'-1)}] \frac{N^2}{nN'^2} \Sigma_1 P'_j (\frac{Y_i}{P'_j} - Y')^2 \tag{4.2.12}$$

if $N' = Qn + q$ and $N'_1 = \ldots = N'_k = Q+1$, $N_{k+1} = \ldots N_n = Q$
$0 < q < n$, $Q > 0$.

The unconditional expectation for $V_2(T_{rhc})$, in (4.2.11) is

$$E_1 V_2(T_{rhc}) = (1 - \frac{n-1}{N'-1}) \frac{N}{N-1} \frac{N'-1}{N'} \frac{1}{n} \Sigma P_j (\frac{Y_i}{P_j} - Y)^2 \tag{4.2.13}$$

and for $V_2(T_{rhc})$ in (4.2.12), the unconditional expectation will be given by replacing the factor $(1 - \frac{n-1}{N'-1})$ by $[1 - \frac{n-1}{N'-1} + \frac{k(n+k)}{N'(N'-1)}]$. However, we shall consider the case when

N' is multiple of n as it gives optimum solution and we got,

$$V(T_{rhc}) = V(T_{d'}) - \frac{n-1}{n} \frac{N}{N-1} \frac{1}{N'} \Sigma P_j \left(\frac{Y_j}{P_j} - Y\right)^2. \qquad (4.2.14)$$

It may be noted that $V(T_{d'})$ and $V(T_{rhc})$ are exact where $V(T_d)$ and $V(T_{hr})$ are approximate expressions.

(iv) Ratio Estimator: As has been pointed out earlier that the information on x collected from fps can also be used in ratio method of estimation instead of using for the selection of the sps. Usually in such cases the sps is also selected as a sps also to be a simple random sample from the fps. The fps s ample in such cases are meant for estimating $\bar{X}$ on a large sample and to use it in building up the ratio estimation given by

$$T_r = \frac{\Sigma_2 Y_j}{\Sigma_2 X_j} \Sigma_1 X_j.$$

The bias and sampling variance of $T_r$ are given in Cochran (1963) and other books hence we shall not consider it here.

## 4.3 Efficiency Comparison of the Estimators:

For making efficiency comparison of the above estimators, we regard the finite population as being drawn from an infinite super-population in which y is correlated with x. Let $\triangle_g$ be the class of apriori distributions $\theta_g$ satisfying

$$E_{\partial_g}(Y_j \mid X_j) = aX_j, \quad a > 0$$

$$V_{\partial_g}(Y_j \mid X_j) = \sigma^2 X_j^g, \quad g \geq 0 \tag{4.3.1}$$

$$Cov_{\partial_g}(Y_j, Y_{j'} \mid X_j, X_{j'}) = 0.$$

It may be mentioned however that in most situations met with in practice, the parameter 'g' is found to lie between 1 and 2.

Under the above model we have the expected variances of the estimators as

$$E_{\partial_g} V(T_{d'}) = \frac{N}{N-1}\,\frac{N'-1}{N'}\,\frac{\sigma^2}{n}\,(X\,\Sigma\,X_j^{g-1} - \Sigma\,X_j^g) +$$

$$\frac{N(N-N')}{N-1}\,\frac{\sigma^2}{N'}\,(\,\Sigma\,X_j^g - \frac{1}{N}\,\Sigma\,X_j^g) \tag{4.3.2}$$

$$E_{\partial_g}(VT_d) = E_{\partial_g}(V(T_{d'})) - \frac{n-1}{2n}\,\frac{N}{N'}\,\sigma^2[\,\Sigma\,X_j^{g-1}(X\Sigma\,P_j^2 + X_j) + O(N^{-2})] \tag{4.3.3}$$

$$E_{\partial_g}(V(T_{HT})) = E_{\partial_g}(V(T_{d'})) - \frac{n-1}{n}\,\frac{N}{N'}\,\sigma^2[\,\Sigma\,X_j^g + O(N^{-2})] \tag{4.3.4}$$

$$E_{\partial_g}(V(T_{rhc})) = E_{\partial_g}(V(T_{d'})) - \frac{n-1}{n}\,\frac{\sigma^2}{N'}[\,\Sigma\,X_j^{g-1}(X - X_j) + O(N^{-2})]. \tag{4.3.5}$$

It may be mentioned that pps with replacement estimator, which obviously is less efficient than pps without replacement estimators considered here, has been shown to be more efficient than ratio estimator $T_r$, whenever $g > 1$, by

Des Raj (1964). Hence in the following theorems we compare the three schemes among thems olves.

Theorem 4.3.1 Des Raj strategy is superior to the Hartley and Rao strategy under the above model $\Delta_g$ in (4.3.1).

Proof: $E_{\delta_g} V(T_{hr}) - E_{\delta_g} V(T_d) = \frac{n-1}{Nn} \frac{N}{N'} \sigma^2 [X \Sigma P_j^2 \Sigma X_j^{g-1} - \Sigma X_j^g]$

$$= \frac{n-1}{Nn} \frac{\sigma^2}{N'X} [\Sigma X_j^2 \Sigma X_j^{g-1} - \Sigma X_j \Sigma X_j^g]$$

$$= \frac{n-1}{Nn} \frac{\sigma^2}{N'X} [\Sigma\Sigma X_j^{g-1} X_{j'} (X_{j'} - X_j)]$$

$> 0$ if all $X_j$ are not same

i.e. if $X_{j'} > X_j$ for at least one $j'$ and $g > 1$. Hence the theorem.

Theorem 4.3.2 Des Raj strategy is superior to the Rao, Hartley and Cochran strategy under the above model $\Delta_g$.

Proof: $E_{\delta_g}(V(T_{rhc})) - E_{\delta_g}(V(T_d)) = \frac{n-1}{n} \frac{\sigma^2}{N'}$

$$[(\frac{N}{2} \Sigma P_j^2 - 1) X \Sigma X_j^{g-1} + (\frac{N}{2} - 1) \Sigma X_j^g]$$

$$= \frac{n-1}{n} \frac{\sigma^2}{N'} [(N + 2) \Sigma X_j^g - X \Sigma X_j^{g-1}] \qquad (4.3.6)$$

since max. $P_j = O(N^{-1})$;

$$\Sigma P_j^2 > N^{-1} \quad \text{if some } P_j \neq N^{-1}. \qquad (4.3.7)$$

Also that all $X_j'^{*}$ are positive, we have from an elementary inequality

$$X \Sigma X_j^{g-1} \lesseqgtr N \Sigma X_j^g \qquad (4.3.8)$$

according as $g \gtrless 1$. From (4.3.6) and (4.3.8) we get

$$E_{\partial_g} V(T_d) < E_{\partial_g} V(T_{rhc}) \quad \text{whenever} \quad g > 1.$$

Hence the theorem.

<u>Theorem 4.3.3.</u> Hartley and Rao strategy is superior to the Rao, Hartley and Cochran strategy under the model $\triangle_g$ in (4.3.1).

<u>Proof:</u> $E_{\partial_g}(V(T_{rhc})) - E_{\partial_g}(V(T_{hr})) = \frac{n-1}{n}\frac{\sigma^2}{N'}[N \Sigma X_j^g - X \Sigma X_j^{g-1} + \Sigma X_j^g]$

with the help of inequality (4.3.8), it is easily seen that

$$E_{\partial_g}(V(T_{hr})) \lessgtr E_{\partial_g}(V(T_{rhc}))$$

according as $g > 1$ or $< 1$ and hence the theorem. We can now summarise our result as follows:

If (i) $N$ and $N'$ are sufficiently large,

(ii) max. $P_j' = O(N'^{-1})$ and max. $P_j = O(N^{-1})$ with at least one $P_j \neq N^{-1}$

and (iii) the parameter '$g$' lies between 1 and 2, then

$$E_{\delta_g} V(T_d) < E_{\delta_g} V(T_{hr}) < E_{\delta_g} V(T_{rhc}) < E_{\delta_g} V(T_{d'}). \qquad (4.3.9)$$

However for $g = 2$, the optimality of the HT estimator is easily seen to hold.

Remark 4.3.1. It may be mentioned that in Hartley and Rao scheme when used in uniphase sampling, it is necessary to arrange the units in random order before selection of sample and in Rao, Hartley and Cochran scheme the units are required to be grouped into n random groups before selection of the sample. This randomisation in both the cases is a difficult task, especially when number of units in the population is quite large. However, the application of these schemes at the second phase in two-phase sampling becomes quite simple, for the first-phase sample is a random sample and the units within it are automatically arranged in random order. Thus application of these schemes for selection of sps, in many cases, may be even simpler than pps with-replacement scheme.

Remark 4.3.2 By considering Murthy's strategy (pps non-sampling using unordered Des Raj estimator) to two-phase sampling it will be observed that this strategy fares better than Des Raj strategy, however, the gain in efficiency achieved by using Murthy's strategy over Des Raj's strategy is expected to be very small for large N'.

# CHAPTER V

# RATIO AND PRODUCT METHODS OF ESTIMATION

5.0 Summary: Various methods by which information on supplementary character may be used have been classified in section 5.1. In section 5.2 & 5.4 the situations in which the ratio and product estimators of the parameter fare better than the usual unbiased estimator is given on the basis of approximate expressions for their mse's for a class of designs. The exact mse have been considered in section 5.3 and preference regions are obtained for any design. And in section 5.4 product estimator has been made unbiased for the case of simple random sampling.

## 5.1 Use of Supplementary Information:

With complete lack of supplementary information, although this would hardly be the case in practice, the only possible methods for getting valid estimates of the population parameters (totals means etc.) is to use a sampling scheme which assigns equal probabilities to all possible samples along with the usual unbiased estimator involving the character under study alone. There are, however, a number of methods proposed in sampling theory to improve upon this unbiased estimator, which utilise information on one or more supplementary characters. The major uses of such information may be classified under the following heads:

(a) designing estimation formulae,

(b) determining the probabilities of selection of the sampling units,

(c) suitable arrangement of the units in combination with systematic or cluster sampling,

(d) stratification of the population and

(e) allocation problems.

Another instance of the methods is two-phase sampling, considered in the previous chapter. In this and the subsequent chapters we shall confine ourselves to discussion of use of supplementary information in the method (a).

5.2 Ratio and Product Methods of Estimation:

In practice we come across several situations in which the ratio of the character under study (y) to the supplementary character (x) is less variable than y's themselves. It is better in such situations to estimate the ratio of y to x in the population, based the values of y's and x's for the units selected in the sample, and then multiply it by the known parameter $\theta_1$ to estimate the parameter $\theta_o$ where $\theta_o$ and $\theta_1$ are the functions of y and x respectively. This procedure is called ratio method of estimation. In some other situations, however, the product of y and x may be believed to be less variable than y's, in which case it is better to estimate the product of y and x in the population, from the sample and then divide it by the known parameter $\theta_1$ to estimate $\theta_o$. Let $T_o$ and $T_1$ denote the usual unbiased estimators of the parameters $\theta_0$ and $\theta_1$ respectively, then the ratio and product estimators of the parameter $\theta_0$, are given by

$$T_r = r\,\theta_1 \quad \text{and} \quad T_p = p/\theta_1 \qquad (5.2.1)$$

where $\theta_1$ is assumed to be known and $r$ and $p$ are given by

$$r = T_0 / T_1 \quad \text{and} \quad p = T_0 T_1 \qquad (5.2.2)$$

respectively. Both the estimators $T_r$ and $T_p$ are biased but consistent. Their bias and mean square error (mse) to the order $n^{-1}$, where $n$ is the sample size, may be obtained as follows:

Writing

$$T_i = \theta_i\,(1 + e_i), \qquad i = 0, 1 \qquad (5.2.3)$$

where $E(e_i) = 0$ and assuming that for large $n$, $|e_1| < 1$ (which is necessary for derivation in the case of ratio estimator), we get

$$B(T_r) = E\,[\,r\,\theta_1 - \theta_0]$$

$$= \theta_0\,E[(1 + e_0)(1 + e_1)^{-1} - 1]$$

$$= \theta_0\,E[(e_0 - e_1) + (e_1^2 - e_0 e_1)\ldots]$$

$$\doteq \theta_0\,(V_1^2 - V_{01}^{11}) \qquad (5.2.4)$$

and $B(T_p) = E[p/\theta_1 - \theta_0]$

$$= \theta_0 E[(\epsilon_0 + \epsilon_1) + \epsilon_0\epsilon_1]$$

$$= \theta_0 \, V_{01}^{11} . \qquad (5.2.5)$$

Similarly, the mse of $T_r$ and $T_p$ are

$$M(T_r) = E(r\,\theta_1 - \theta_0)^2$$

$$= \theta_0^2 E[(\epsilon_0^2 - 2\epsilon_0\epsilon_1 + \epsilon_1^2) - (2\epsilon_0^2\epsilon_1 - 2\epsilon_0\epsilon_1^2 + \epsilon_1^3) + \ldots)$$

$$\doteq \theta_0^2 \,(V_0^2 - 2V_{01}^{11} + V_1^2) \qquad (5.2.6)$$

and $M(T_p) = E(p/\theta_1 - \theta_0)^2$

$$= \theta_0^2 \, E(\epsilon_0 + \epsilon_1 + \epsilon_0\epsilon_1)^2$$

$$= \theta_0^2(V_0^2 + V_1^2 + 2V_{01}^{11} + 2V_{01}^{12} + V_{01}^{21} + V_{01}^{22}) \qquad (5.2.7)$$

$$\doteq \theta_0^2(V_0^2 + V_1^2 + 2V_{01}^{11}) \qquad (5.2.8)$$

where

$$V_{01}^{rs} = E(\epsilon_0^r \cdot \epsilon_1^s) = \frac{E[(T_0 - \theta_0)^r (T_1 - \theta_1)^s]}{\theta_0^r \, \theta_1^s} . \qquad (5.2.9)$$

It may be mentioned that the expression in (5.2.7) gives exact value of $M(T_p)$ where (5.2.8) is its approximation. Further Murthy (1964) has compared ratio and product estimators with that of conventional unbiased estimator for estimating the population total Y (that is $T_r$, $T_p$ and $T_0$ in the present set-up) and obtained the following result.

Theorem 5.2.1 (Murthy): For any design obtained by selecting the units with equal probability or varying probability sampling with replacement or any other sampling scheme involving selection of independent sub-samples, the preference regions for either product, unbiased or ratio estimator, considering the mse to order $n^{-1}$, are given by

$$-1 \leq \rho_{01} \leq -\frac{1}{2}\left(\frac{C_1}{C_o}\right)$$

$$-\frac{1}{2}\left(\frac{C_1}{C_o}\right) \leq \rho_{01} \leq +\frac{1}{2}\left(\frac{C_1}{C_o}\right)$$

and

$$+\frac{1}{2}\left(\frac{C_1}{C_o}\right) \leq \rho_{01} \leq +1 \qquad (5.2.10)$$

respectively, where $C_o^2 = V_o^2$, $C_1^2 = V_1^1$, $C_{01} = V_{01}^{11}$ and $\rho_{01}$ is the correlation coefficient between $T_o$ and $T_1$.

Remark 5.2.1: It is easily seen that the above preference regions are valid only if both the parameters $\theta_o$ and $\theta_1$ are either positive or negative. Otherwise, the preference regions for $T_p$ and $T_r$ get interchanged.

Remark 5.2.2: The Theorem holds only for restricted sampling schemes as indicated. For these sampling schemes

$$V_{01}^{rs} = Q(V_{01}^{rs})^* \qquad (5.2.11)$$

where $(V_{01}^{rs})^*$ stands for $V_{01}^{rs}$ where $n = 1$ and $Q$ is some constant inversely proportional to $n$, the sample size or the number of sub-samples as the case may be.

It is pertinent to note that the preference regions given in Theorem 5.2.1 is based on approximate expression of the mse for ratio and product estimators and hence an alternative criterion to choose any of the three estimators, which is based on the exact mse and is valid for any design is presented in the following section.

5.3 <u>A General Criterion for Preference</u>:

We shall prove the following theorem.

<u>Theorem 5.3.1</u>: For a design $P$, where the units at any given step may be selected with or without replacement and with equal and unequal probabilities, the product estimator $T_p$ will be more efficient than the linear unbiased estimator $T_o$ for estimating $\theta_o$ if

$$\alpha + \beta + \gamma \geq 0 \tag{5.3.1}$$

where

$$\alpha = \frac{\text{Cov}\left[(T_o T_1)^2, \frac{1}{T_1^2}\right]}{E\left(\frac{1}{T_1}\right)^2 V(T_o T_1)}$$

$$\beta = C_1/T_1(1 - \delta^2)$$

$$\gamma = \frac{2C_1/T_1}{C_{T_o T_1}}\left[\frac{C_1/T_1}{2C_{T_o T_1}} - \delta\right]$$

and

$$\rho = \frac{\mathrm{Cov}[(T_o T_1),\ 1/T_1]}{\sqrt{V(T_o T_1)\ V(1/T_1)}} \tag{5.3.2}$$

__Proof:__ Let us consider the identity

$$T_o = \frac{(T_o T_1)}{T_1}. \tag{5.3.3}$$

Then we have

$$V(T_o) = V\Big[\frac{(T_o\ T_1)}{T_1}\Big]$$

$$= E\Big[\frac{(T_o T_1)}{T_1}\Big]^2 - \Big[E\Big(\frac{T_o\ T_1}{T_1}\Big)\Big]^2, \tag{5.3.4}$$

Noting that

$$\mathrm{Cov}[(T_o T_1)^a, T_1^{-a}] = E[(T_o T_1)^a\ T_1^{-a}] - E(T_1^{-a})E(T_o T_1)^a$$

for $a = 1, 2, 3 \ldots$ and using this relation we get from (5.3.4)

$$V(T_o) = \mathrm{Cov}[(T_o^2\ T_1^2),\ 1/T_1^2] + E(1/T_1^2)\ E(T_o^2\ T_1^2)$$

$$- \{\mathrm{Cov}[(T_o T_1),\ 1/T_1] + E(1/T_1)\ E(T_o T_1)\}^2 \tag{5.3.5}$$

Next, since

$$V(1/T_1) = E(1/T_1^2) - [E(1/T_1)]^2$$

and

$$V(T_o^2\ T_1^2) = E(T_o^2\ T_1^2) - [E(T_o\ T_1)]^2$$

from (5.3.5), we get

$$V(T_0) = \text{Cov}[(T_0^2\ T_1^2),\ 1/T_1^2] - \{\text{Cov}[(T_0\ T_1),\ 1/T_1]\}^2$$

$$- 2\text{Cov}\ [(T_0\ T_1),\ 1/T_1]\ E(1/T_1) E\ (T_0\ T_1) +$$

$$+ V(1/T_1)\ V(T_0\ T_1) + E(T_0^2\ T_1^2)\ V(1/T_1)$$

$$+ E(1/T_1^2)\ V(T_0\ T_1). \qquad (5.3.6)$$

After some simplification $V(T_0)$ can also be expressed as

$$V(T_0) = [1 + \alpha + \beta + \gamma\ ]\ V(T_0\ T_1)\ E(1/T_1^2) \qquad (5.3.7)$$

where $\alpha$, $\beta$, and $\gamma$ are defined in (5.3.2).

Again noting that

$$E(1/T_1^2) = V(1/T_1) + [E(1/T_1)]^2$$

$$= [E(1/T_1)\ ]^2$$

$$\geq [\frac{1}{E(T_1)}]^2 = \frac{1}{\theta_1^2} \qquad (5.3.8)$$

Substituting the value of $E(1/T_1^2)$ from (5.3.8) in (5.3.7) we get

$$V(T_0) \geq [\ 1 + \alpha + \beta + \gamma\ ]\ \frac{V(T_0\ T_1)}{\theta_1^2}$$

where $V(T_0\ T_1)/\theta_1^2$ is same as $V(T_p)$.

Hence the theorem:

Remark 5.3.1: From equation (5.3.6) it is observed that whenever $T_o$ is inversely proportional to $T_1$, the covariance terms, that is the first three terms in (5.3.6), vanished automatically and hence in such cases we get

$$V(T_o) = V(1/T_1)V(T_o\ T_1) + E(T_o^2\ T_1^2)V(1/T_1)$$
$$+ E(1/T_1^2)V(T_o\ T_1). \qquad (5.3.9)$$

Noting that the first and the second terms in (5.3.9) are always positive

$$V(T_o) > V(T_o\ T_1)\ E(1/T_1^2)$$

and from (5.3.8), therefore we have

$$V(T_o) > \frac{V(T_o.\ T_1)}{\theta_1^2}.$$

Thus if $T_o$ is inversely proportional to $T_1$

$$V(T_o) > V(T_p).$$

This is also seen from the form (5.3.7) since in this situation $\alpha$, $\delta$ are zero, which implies that $\beta + \gamma$ is always greater than zero.

Remark 5.3.2: The condition (5.3.1) is satisfied in a variety of other situations as well. For example it is easily seen that $\delta$ will never attain value $\pm 1$, and hence $\beta > 0$ always.

Thus (5.3.1) is satisfied whenever $\alpha + \gamma \geq 0$.

But $\gamma > 0$ if and only if

$$\delta < \frac{C_{1/T_1}}{2C_{T_0}\, T_1}$$

and this condition is likely to be satisfied for large values of $C_{1/T_1}$, in which case (5.3.1) may be realised even when $\alpha < 0$.

On the other hand if

$$\delta > \frac{C_{1/T_1}}{2C_{T_0}\, T_1},$$

then $\alpha + \gamma$ will be positive if $\alpha$ is much larger.

It may be mentioned that Koop (1964) has obtained similar conditions for the preference of ratio estimator in comparison to the linear unbiased estimators. These conditions are as follows.

Theorem 5.3.2 (Koop): For any sampling design, where the units at any given step may be selected with or without replacement and with equal or unequal probabilities, the ratio estimator $T_r$ will be more efficient than the linear unbiased estimator

$T_e$ for estimating $\theta_0$ if

$$\alpha' + \beta' + \gamma' > 0 \tag{5.3.9}$$

where

$$\alpha' = \frac{Cov[T_1^2, (T_0/T_1)^2]}{[E(T_1)]^2 \; V(T_0/T_1)}$$

$$\beta' = C_{T_1}(1 - \delta'^2)$$

$$\gamma' = \frac{2C_{T_1}}{C_{T_0/T_1}} \left[ \frac{C_{T_1}}{2C_{T_0/T_1}} - \delta' \right] \tag{5.3.10}$$

and

$$\delta' = \frac{Cov[T_1, T_0/T_1]}{(V(T_1) \; V(T_0/T_1))^{1/2}} \; .$$

Remark 5.3.3: From (5.3.9) and (5.3.10) it is observed that whenever the linear estimators $T_0$ and $T_1$ are directly proportional then $\alpha'$, $\delta'$, becomes zero leaving the conditions (5.3.9) as

$$\beta' + \gamma' > 0$$

which is in this situation always true. However, there exist other situations as has been pointed out by Koop where the condition (5.3.9) is satisfied.

Thus on the basis of remarks (5.3.1) and (5.3.3) we get clearly demarcated regions of preference for unbiased, ratio or product estimator based on the proportionality of the estimator $T_0$ with that of $T_1$.

An Emperical Study:

For the purpose of the present study reference is made to an investigation undertaken by the Biometry Research Unit of the Indian Statistical Institute in connection with multivariate investigation of blood chemistry. Such investigation entails the collection of multiple measurements on each individual examined, to study the blood chemistry, and it was carried out on three groups of persons namely (i) Urban non-vegeterian males (Group A), (ii) Urban non-vegeterian females (Group B) and (iii) Urban vegeterian males (Group C). Data on 32 variables was collected, the details of findings have been given by Das (1966). Considering 'height' as the supplementary variable we give below the relative efficiency of the unbiased, product and ratio methods of estimation for estimating the remaining 31 characteristics based on data for Group B.

Table 5.3.1: Comparison of the estimators $T_o$, $T_r$ and $T_p$.

| sl. no. | characters | mean | S.D. | corr-coeff. with height | efficiency of estimator | | |
|---|---|---|---|---|---|---|---|
| | | | | | unbiased | product | ratio |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1 | Age | 34.96 | 11.89 | 0.1932 | 100 | 103.3 | 94.2 |
| 2 | Weight | 47.75 | 10.76 | 0.2270 | 100 | 89.1 | 105.1 |
| 3 | Amylase activity | 104.57 | 33.34 | 0.0219 | 100 | 97.7 | 98.8 |
| 4 | Calcium | 8.65 | 1.25 | 0.1518 | 100 | 85.5 | 100.1 |
| 5 | Chloride | 518.16 | 34.62 | -0.1289 | 100 | 81.9 | 84.2 |
| 6 | Total Cholestrol | 153.13 | 37.91 | -0.0279 | 100 | 98.2 | 96.3 |
| 7 | Free Cholestrol | 60.03 | 14.88 | -0.0334 | 100 | 98.5 | 95.8 |
| 8 | Cholestrol ester | 93.09 | 23.37 | -0.0237 | 100 | 97.9 | 96.4 |
| 9 | Creatinine | 0.99 | 0.20 | 0.0697 | 100 | 94.3 | 98.7 |
| 10 | Glucose | 67.78 | 37.31 | -0.2029 | 100 | 102.8 | 96.4 |
| 11 | Non-protein Nitrogen | 35.90 | 7.28 | -0.0834 | 100 | 99.2 | 93.2 |
| 12 | Acid phosphate | 2.77 | 2.26 | -0.1279 | 100 | 101.2 | 98.4 |
| 13 | Alkaline phosphate | 8.77 | 5.19 | -0.1599 | 100 | 101.9 | 97.3 |
| 14 | Total protein | 7.29 | 0.79 | 0.0660 | 100 | 83.6 | 90.6 |
| 15 | Serum pH | 7.33 | 0.13 | -0.0741 | 100 | 48.2 | 70.2 |
| 16 | Urea | 18.21 | 5.37 | -0.2116 | 100 | 104.6 | 89.6 |
| 17. | Uric acid | 1.80 | 0.74 | -0.0332 | 100 | 99.5 | 98.3 |
| 18 | Erythrocytes | 3.61 | 0.34 | 0.2822 | 100 | 68.9 | 104.5 |
| 19 | Leukocytes | 859.57 | 1172.96 | -0.0713 | 100 | 97.8 | 91.7 |
| 20 | Neutrophiles | 60.43 | 6.61 | 0.0673 | 100 | 83.8 | 91.5 |

continued ...

contd..

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|
| 21 | Lymphocytes | 31.89 | 5.67 | -0.0062 | 100 | 95.2 | 94.5 |
| 22 | Eosinophils | 6.22 | 4.47 | -0.1185 | 100 | 101.3 | 98.2 |
| 23 | Homoglobin | 11.16 | 1.14 | 0.2804 | 100 | 71.3 | 107.2 |
| 24 | Sedimentation rate | 28.75 | 20.27 | -0.1742 | 100 | 102.0 | 97.6 |
| 25 | Pulse (right) | 75.46 | 11.58 | -0.1326 | 100 | 100.2 | 86.0 |
| 26 | Pulse (left) | 74.43 | 9.45 | -0.0929 | 100 | 95.7 | 88.4 |
| 27 | Systolic pressure(right) | 117.39 | 21.69 | 0.0614 | 100 | 92.6 | 95.2 |
| 28 | Systolic pressure(left) | 116.41 | 22.72 | -0.0591 | 100 | 98.1 | 93.3 |
| 29 | Dystolic pressure(right) | 71.62 | 12.03 | 0.0065 | 100 | 93.9 | 93.8 |
| 30 | Dystolic pressure(left) | 72.24 | 12.39 | 0.0144 | 100 | 93.8 | 94.2 |
| 31 | Oral temperature | 99.06 | 0.73 | -0.1245 | 100 | 62.6 | 47.2 |

From this study it is observed that out of 31 cases product estimator fare better in 8 cases and the ratio estimator in 4 cases only and in the remaining the unbiased estimator is more efficient. The gain in efficiency by using ratio or product estimator is in no case substantial due to the fact that the magnitude of the correlation coefficient ($\rho_{01}$) is quite small in all the cases. However, this study gives an indication of the situations in which product estimator can be efficiently used.

## 5.4 Estimators in Systematic Sampling :

Significant contributions of Madow (1944), Griffith (1946) and others describe various situations under which systematic sampling possesses an edge over other schemes both from the view point of its efficiency and practical convenience. However, work in this field has been concentrated to the conventional unbiased estimator only. In this section was shall consider the use of ratio and product estimators in uni-stage and two-stage systematic sampling and also present a discussion for the general case.(Singh, 1967a)

i) Uni-stage systematic sampling: Consider a population $U = (U_1, U_2, \ldots, U_N)$ of $N$ identifiable units. The systematic sampling consists in drawing a number, say $r$, between 1 and I, and selecting the units bearing serial numbers

$$r,\ r+I,\ r+2I, \ldots,\ r+(n-1)I$$

which constitute the (linear) systematic sample. I is called the sampling interval and it is taken to be ratio of N and n. That $I = N/n$.

We define $Y_{ij}$ and $X_{ij}$ as the value of the unit bearing $r+(j-1)I$ number in the population for the character under study $y$ and the supplementary character $x$ respectively. Then the conventional unbiased estimators of the population mean

$$\bar{\bar{Y}} = \frac{1}{N} \sum_{i=1}^{N} Y_i \quad \text{and} \quad \bar{\bar{X}} = \frac{1}{N} \sum_{i=1}^{N} X_i \tag{5.4.1}$$

are given by

$$\bar{y}_s = \bar{y}_i = \frac{1}{n} \sum_{j=1}^{n} y_{ij} \quad \text{and} \quad \bar{x}_s = \bar{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_{ij}. \tag{5.4.2}$$

respectively.

The sampling variance of the estimators $\bar{y}_s$ and $\bar{x}_s$ are

$$\sigma^2_{sy} = \frac{\sigma^2_y}{n} \left(1 + \overline{n-1}\, \rho_y\right) \tag{5.4.3}$$

and

$$\sigma^2_{sx} = \frac{\sigma^2_x}{n} \left(1 + \overline{n-1}\, \rho_x\right) \tag{5.4.4}$$

where $\sigma^2_y$ and $\sigma^2_x$ are the variances and $\rho_y$ and $\rho_x$ are the intra-class correlation coefficients for the character $y$ and $x$ respectively. We now prove the following Theorem.

<u>Theorem 5.4.1</u>: The region of preference, mentioned in Theorem 5.2.1, for the ratio and product estimators given by

$$\bar{y}_{rs} = (\bar{y}_s / \bar{x}_s)\bar{\bar{X}} \quad \text{and} \quad \bar{y}_{ps} = (\bar{y}_s\, \bar{x}_s) / \bar{\bar{X}} \tag{5.4.5}$$

respectively, remains unchanged even for the systematic sampling scheme whenever the condition

$$\rho_y = \rho_x = \rho \qquad (5.4.6)$$

is satisfied, for large $N$.

__Proof:__ For the present case, we get from (5.2.8) and (5.2.9) the mse of $\bar{y}_{rs}$ and $\bar{y}_{ps}$ as

$$M(\bar{y}_{rs}) = \frac{\bar{Y}^2}{n} [C_y^2 (1 + \overline{n-1}\, \rho_y) + C_x^2(1 + \overline{n-1}\, \rho_x) - 2C_xC_y\rho_{yx}$$

$$(1 + \overline{n-1}\rho_y)^{\frac{1}{2}} . (1 + \overline{n-1}\, \rho_x)^{\frac{1}{2}}]$$

and

$$M(\bar{y}_{ps}) = \frac{\bar{Y}^2}{n} [C_y^2(1 + \overline{n-1}\, \rho_y) + C_x^2(1 + \overline{n-1}\, \rho_x) + 2C_xC_y\rho_{yx}$$

$$(1 + \overline{n-1}\, \rho_y)^{\frac{1}{2}} . (1 + \overline{n-1}\, \rho_x)^{\frac{1}{2}}]$$

which under the condition (5.4.6) can be written as

$$M(\bar{y}_{rs}) = M(\bar{y}_r)_{ran} (1 + \overline{n-1}\, \rho) \qquad (5.4.7)$$

(Swain, 1964) and

$$M(\bar{y}_{ps}) = M(\bar{y}_p)_{ran} (1 + \overline{n-1}\, \rho) \qquad (5.4.8)$$

respectively, where

$$M(\bar{y}_r)_{ran} = \frac{\bar{Y}^2}{n} (C_y^2 + C_x^2 - 2C_xC_y\rho_{yx}) \qquad (5.4.9)$$

and

$$M(\bar{y}_p)_{rsn} = \frac{\bar{Y}^2}{n}(C_y^2 + C_x^2 + 2C_xC_y\rho_{yx}) \qquad (5.4.10)$$

are the mse of ratio and product estimators in simple random sampling.

Now comparison of (5.4.7) and (5.4.8) establishes the truth of the Theorem.

ii) Two-stage sampling: We consider the scheme suggested by Sukhatme (1953) which consists in selecting the first-stage units with replacement and with varying probabilities $P_1, P_2, \ldots, P_N$, where $N$ denotes the number of first-stage units in the population, and the second-stage units within the selected first-stage units are sampled by the method of systematic sampling. Let $n$ denote the number of first-stage units selected in the sample and $m_i$ be the number of second stage units selected from the $i^{th}$ unit containing $M_i$ second stage units. Let $m_i/M_i$ be an integer and $\sum_{i=1}^{n} m_i = m_o$. Both the characters are observed for $m_o$ units in the sample.

An unbiased estimator for $\bar{Y}$ is

$$\bar{\bar{y}}_s = \sum_{i=1}^{n} M_i\bar{y}_{si}/n\, M_o P_i$$

where $\bar{y}_{si} = \frac{1}{m_i}\sum_{j=1}^{m_i} y_{ij}$ and $M_o = \sum_{i=1}^{N} M_i$. The estimator of $\bar{X}$ denoted by $\bar{x}_{si}$ will also have similar definition. The corresponding ratio and product estimators $\bar{y}_{rs}$ and $\bar{y}_{ps}$ may also be defined similarly.

<u>Theorem 5.4.2:</u> Under the assumption

$$\rho_{ix} = \rho_{iy} = \rho', \tag{5.4.11}$$

the mse of $\bar{\bar{y}}_{rs}$ for this scheme, is

$$M(\bar{\bar{y}}_{rs}) = V_{1r} + \frac{1}{n} \Sigma \frac{M_i}{M_0^2 P_i} \frac{S_i^2}{m_i} (1 + \rho'(m_i-1)) \tag{5.4.12}$$

where $V_{1r}$ is the contribution of sampling of first stage units and is given by

$$V_{1r} = \frac{1}{n} \sum_{i=1}^{N} \frac{M_i}{M_0^2 P_i} (\bar{Y}_i - R\, \bar{X}_i)^2 \tag{5.4.13}$$

and

$$S_i^2 = (\sigma_{iy}^2 + R^2 \sigma_{ix}^2 - 2R\sigma_{ix}\, \sigma_{iy}\, \rho_{iyx}) \tag{5.4.14}$$

<u>Proof:</u> Proof of the theorem follows by noting that

$$V(\bar{\bar{y}}_s) = \frac{1}{n}\left( \sum_{1}^{N} \frac{M_i^2\, \bar{Y}_i^2}{M_0^2 P_i} - \bar{\bar{Y}}^2 \right) + \frac{1}{n} \sum_{1}^{N} \frac{M_i^2}{M_0^2 P_i} \frac{\sigma_{iy}^2}{m_i} (1+\rho_{iy}(m_i-1))$$

with a similar definition for $V(\bar{\bar{x}}_s)$ and

$$\text{Cov}\,(\bar{x}_s, \bar{y}_s) = \frac{1}{n} \sum_{1}^{N} \frac{M_i^2 Y_i X_i}{M_i^2 P_i} - \bar{Y}\bar{X}) + \frac{1}{n} \sum_{1}^{N} \frac{M_i^2}{M_o^2 P_i} \sigma_{ix}\sigma_{iy}\rho_{iyx}.$$

$$\frac{1}{m_i}(1+ \overline{m_i - 1}\, \rho_{iy})^{\frac{1}{2}} (1+ \overline{m_i - 1}\, \rho_{ix})^{\frac{1}{2}}$$

where $\sigma_{iy}^2$ and $\sigma_{ix}^2$ are variances corresponding to the $i^{th}$ first stage units and $\rho_{iy}$, $\rho_{ix}$ are the corresponding intra class correlation coefficients for the character $y$ and $x$ respectively.

The expression for the mse of the product estimator can also be obtained for this scheme and compared with the $M(\bar{y}_{rs})$.

iii) <u>General case:</u>

The general form of the variances of these estimators in r-stage sampling is given by

$$V(T) = E_1 V_2 \ldots E_{r-1} E_r(T) + E_1 E_2 V_3 \ldots E_r(T) + \ldots + E_1 E_2 \ldots E_{r-1} V_r(T) \qquad (5.4.15)$$

where $T$ is any estimator.

This shows that total variance consists of $r$ parts each part giving variation between units of particular stage within units of the previous stage. This last part of the variance, in case where the systematic sampling has been adopted at the ultimate-stage with any scheme of sampling at the previous stages, for ratio and product estimators will consist of two parts, namely, (i) contribution due to the scheme of sampling at previous stages which differs from one procedure of estimation to the other and (ii) the contribution of systematic sampling, unaffected by the produre of estimation. The contribution of systematic sampling at the $r^{th}$ stage is given by $(1+\rho_r(n_r-1))$, where $\rho_r$ is the intra-class correlation coefficient between the ultimate stage units belonging to the individual penultimate-stage units and $n_r$ is the number of ultimate stage units selected from the sampled penultimate stage units and this may vary from one penultimate stage unit to another according to the scheme of sampling. The assumption here, however, is that $\rho_{rx} = \rho_{ry} = \rho_r$, where $\rho_{rx}$ and $\rho_{xy}$ have definition similar to $\rho_r$.

The contribution of the last term, thus makes it clear that efficiency of systematic selection at the $r^{th}$ stage depends on the value of $\rho_r$. If $\rho_r$ is negative or due to some arrangement of the ultimate stage units it could be made negative, then these estimators in systematic sampling will be more efficient than the corresponding estimators in simple random sampling. Hence it may be concluded that systematic sampling can be efficiently used in building-up either of the three estimators only when both the conditions are satisfied, namely, (i) $\rho_r$ is less than zero and (ii) $\rho_{yx}$ satisfies the preference regions mentioned in theorem 5.2.1.

Remark: The comparison of a ratio and product estimators have been made using the usual approximation for their variances, which has been used by many authors in past (Swain, 1964, Sukhatme 1963 etc.), this approximation however, does not include all the terms upto order $n^{-1}$ in general.

5.4. Unbiased Estimators:

For simple random sampling without replacement, Hartley and Ross (1954) has obtained ratio type estimator which is unbiased for estimating $\bar{Y}$. The estimator is given by

$$\bar{y}'_r = \bar{r}\,\bar{X} + \frac{N-1}{N}\,\frac{n}{n-1}\,(\bar{y} - \bar{r}\,\bar{x})$$

where $\bar{r} = \frac{1}{n}\sum_{1}^{n} y_i/x_i$.

An unbiased product estimator for estimating $\bar{Y}$, similar to $\bar{y}'_r$ may also be obtained by subtracting unbiased estimator of its bias given by

$$b(\bar{y}_p) = \frac{1}{\bar{X}}\,\frac{N-n}{Nn}\,\frac{1}{n-1}\left(\sum_{1}^{n} y_i x_i - n\bar{y}\,\bar{x}\right).$$

The unbiased product estimator is thus given by

$$\bar{y}'_p = \frac{1}{\bar{X}}\left[\,\bar{y}\,\bar{x} - \frac{N-n}{Nn}\,\frac{1}{n-1}\left(\sum_{1}^{n} y_i x_i - n\bar{y}\,\bar{x}\right)\right]$$

Some other estimators of this type have been given Chapter VII.

# CHAPTER VI

# MULTIVARIATE RATIO AND PRODUCT METHODS OF ESTIMATION

6.0 Summary: In this chapter the uni-variate product estimator has been extended to multivariate product method of estimation, which utilise information on several supplementary characters. This estimator has been considered in Sections 6.1 and 6.2 in a general manner. This estimator is then compared with the multivariate ratio estimator in Section 6.3. The case of two-phase sampling is included in the last section and an empirical study is also included for illustration.

## 6.1 Multivariate Product Method of Estimation:

Whenever there is supplementary information available from any source, the sampler tends to utilise it in the method of estimation which gives maximum efficiency. We have already considered in the previous chapter the ordinary ratio and product methods of estimation which make use of information on just one supplementary characteristic and provide more efficient estimators than the usual unbiased estimator, under certain situations commonly met with in practice.

Quite often information on more than one such characteristic is available in the survey which can be utilised to increase the efficiency of the estimator. Olkin (1958), in connection with the estimation of population total, considered the use of information on multi-supplementary variables in building-up multivariate ratio estimator and this estimator was found to be more efficient than the ordinary ratio estimator in most situations. Use of information on multi-supplementary variables in suitable manners have been considered by Des Raj (1965a), Shukla (1966), Srivastava (1967) and others. Author's contribution in this direction have been presented in this and the following chapters. We give below an extension of the ordinary product estimator to the multivariate case. The multivariate product estimator, using information on two or more supplementary characters, is being introduced here in a quite general form.

Let $\theta_0$ be the parametric function to be estimated and let $\theta_1, \theta_2 \ldots \theta_k$ be the parametric functions corresponding to $k$ supplementary characters. Let $T_i$ $(i = 0,1,\ldots, k)$ denote the linear unbiased estimators of $\theta_i$ based on any design P. Then the estimator considered is given by

$$T_{pk} = \sum_{i=1}^{k} \frac{w_i p_i}{\theta_i} \qquad (6.1.1)$$

where $p_i = T_0 . T_i$ and $w_i$'s are weights such that

$$\sum_{i=1}^{k} w_i = 1. \qquad (6.1.2)$$

It is assumed that $\theta_1, \theta_2 \ldots \theta_k$ are known in advance. The discussion for unknown values of those parameters has been presented in the last section. We prove below a theorem regarding the bias and m.s.e. of $T_{pk}$.

<u>Theorem 6.1.1</u> For any design $D(U,S,P)$, exact expressions for the bias and m.s.e. of $T_{pk}$ are given by

$$B(T_{pk}) = \theta_o w b' \tag{6.1.3}$$

and $$M(T_{pk}) = \theta_o^2 \, w \, (A + B + C + bb')w' \tag{6.1.4}$$

respectively, where $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$, defined below, are matrices of order $k \times k$ each, $b$ and $w$ are vectors represented by $(b_1, b_2, \ldots, b_k)$ and $(w_1, w_2, \ldots, w_k)$ respectively, and $b'$ and $w'$ are transpose of $b$ and $w$.

<u>Proof:</u> Let us define

$$T_i = \theta_i \, (1 + \varepsilon_i) \tag{6.1.5}$$

where $E(\varepsilon_i) = 0$ for all $i = 0, 1, 2, \ldots, k$.

We get,

$$E(T_{pk}) = \theta_o \sum_i w_i \, E\left(\frac{p_i}{P_i}\right) \tag{6.1.6}$$

and

$$M(T_{pk}) = \theta_o^2 \sum_i \sum_j w_i w_j \operatorname{cov}(p_i, p_j)/P_i P_j \tag{6.1.7}$$

where $P_i = \theta_o \, \theta_i$, $i = 1, 2, \ldots, k$.

Now substituting $T_i$ from (6.1.5) in $p_i$, we get

$$p_i = P_i(1 + e_o)(1 + e_i)$$

$$= P_i(1 + \alpha_i + \beta_i),$$

where $\alpha_i = (e_o + e_i)$ and $\beta_i = (e_o\, e_i)$, which gives

$$E\left(\frac{p_i}{P_i}\right) = 1 + E(\beta_i)$$

$$= 1 + b_i. \qquad (6.1.8)$$

Further,

$$\frac{\text{Cov}\ (p_i\ ,\ p_j)}{P_i\ P_j} = E(p_i - E(p_i))(p_j - E(p_j))/P_iP_j$$

$$= E\left(\frac{p_i}{P_i} - 1 - b_i\right)\left(\frac{p_j}{P_j} - 1 - b_j\right)$$

$$= E[(1 + \alpha_i + \beta_i)(1 + \alpha_j + \beta_j) - (1 + \beta_i + \beta_j) + b_i b_j]$$

$$= E(\alpha_i(\alpha_j + \beta_j)) + E(\alpha_j\beta_i) + E(\beta_i\beta_j) + b_i b_j$$

$$= E[(e_o^2 + e_o e_i + e_o e_j + e_{ij}) + (e_o^2 e_i + 2 e_o e_i e_j$$

$$+ e_o^2 e_j) + e_o^2 e_i e_j + b_i b_j]$$

$$= a_{ij} + b_{ij} + c_{ij} + b_i b_j \qquad \ldots (6.1.9)$$

where

$$b_i = V_{oi}^{11}$$

$$a_{ij} = (V_o^2 + V_{oi}^{11} + V_{oj}^{11} + V_{ij}^{11})$$

$$b_{ij} = (V_{oi}^{21} + 2V_{oij}^{111} + V_{oj}^{21})$$

$$c_{ij} = V_{oij}^{211} \qquad (6.1.10)$$

and

$$V_{oij}^{rst} = \frac{E(T_o - \theta_o)^r (T_i - \theta_i)^s (T_j - \theta_j)^t}{\theta_o^r \, \theta_i^s \, \theta_j^t} \qquad (6.1.11)$$

Substitution of $E(p_i / P_i)$ and $Cov\ (p_i, p_j)/P_iP_j$ in (6.1.6) and (6.1.7) confirms the truth of the theorem.

<u>Remark 6.1.1</u> It is easily seen that the estimator will be unbiased if $b' = 0$, which will happen if $E(e_o e_i) = 0$, that is if the correlation coefficient between $T_o$ and $T_i$ is zero for all $i$.

<u>Remark 6.1.2</u> The matrices A, B, and C are atleast semi-positive definite. Further writing $V_{ij}^{11} = C_{ij} = C_iC_j\rho_{ij}$ where $C_i$, $C_j$ and $C_{ij}$ are relative variances (coefficient of variations) and covariances of the estimators and $\rho_{ij}$ correlation coefficient between them, we get,

$$b_i = C_o \, C_i \, \rho_{oi}$$

and
$$a_{ij} = (C_o^2 + C_oC_i\rho_{oi} + C_oC_j\rho_{oj} + C_iC_j\rho_{ij}). \qquad (6.1.12)$$

Next, in order to compare this estimator with the multivariate ratio estimator we obtain approximate expressions for the bias and m.s.e. of $T_{pk}$. Since for sampling schemes, such as simple random sampling, varying probability schemes with replacement, etc., $E(p_i / P_i)$ and $Cov(p_i, p_j)/P_iP_j$ can be expressed as

$$E\left(\frac{p_i}{P_i}\right) = 1 + \frac{b_i}{n}$$

and

$$\frac{Cov(p_i, p_j)}{P_iP_j} = \frac{a_{ij}}{n} + \frac{b_{ij}}{n^2} + \frac{c_{ij}}{n^3}$$

where $n$ is the sample size and $a_{ij}$, $b_{ij}$ and $c_{ij}$ have similar meaning as the corresponding terms in (6.1.10), except for $n$.

This shows that the contribution of the terms involving $n^{-2}$ and $n^{-3}$ in the m.s.e. may be neglected for large values of $n$, in which case (6.1.6) reduces to

$$M(T_{pk}) = e_o^2 \; w \; A \; w'. \qquad (6.1.13)$$

Now arises the question of the choice of weights. Of all the vectors $w$ satisfying the condition (6.1.2), we shall choose that one which minimises $M(T_{pk})$ in (6.1.13). We prove the following lemma.

<u>Lemma (6.1.1)</u> The estimator $T_{pk}$ with its m.s.e. given in (6.1.13) attains minimum value when each

$$w_i = \frac{\text{sum of elements in } i^{th} \text{ column of } A^{-1}}{\text{sum of } k^2 \text{ elements of } A^{-1}} .$$

<u>Proof</u>: We minimise $M(T_{pk})$ under the condition $\sum_{i=1}^{k} w_i = 1$. In other words the function

$$\phi = \sigma_o^2 w A w' - \lambda(we' - 1)$$

is minimised, where $\lambda$ is a lagrangian multiplier and vector $e = (1, 1, ..., 1)$.

Differentiating $\phi$ with respect to $w$ and equating to zero, we get

$$wA - \lambda e = 0.$$

Assuming $A^{-1}$ exists,

$$w = \lambda e A^{-1} \tag{6.1.14}$$

that is, $we' = \lambda e A^{-1} e'$, hence

$$\lambda = \frac{1}{e A^{-1} e'}, \text{ since } we' = 1 \tag{6.1.15}$$

which gives from (6.1.14), the optimum

$$w = \frac{e A^{-1}}{e A^{-1} e'} . \tag{6.1.16}$$

Hence the lemma.

Assuming the weights for all the k supplementary characters to be uniform, which will happen only when the sums of each column matrix A are equal, the optimum weight w is given by (e/k) and the corresponding bias and m.s.e.

$$B(T_{pk}) = \frac{\theta_o\ e\ b'}{k}$$

and

$$M(T_{pk}) = \theta_o^2\ m/k \qquad (6.1.17)$$

where m is scalar such that $eA = em$, $(m \neq 0)$ and $m = e$ implies that A is singular.

Theorem 6.1.2 Assuming the weights to be uniform, the bias and m.s.e. of $T_{pk}$, to order $n^{-1}$, is given by

$$B(T_{pk}) = \theta_o\ C\ C_o \rho_o \qquad (6.1.18)$$

and

$$M(T_{pk}) = \frac{\theta_o^2}{k}[(C^2(1-\rho) + k(C_o^2 + 2C\ C_o \rho_o + \rho C^2)] \qquad (6.1.19)$$

where $C_i$, $C_j$ etc., are as defined before.

Proof: Let us consider

$$\left.\begin{array}{l} C_i = C, \quad \rho_{oi} = \rho_o \\ \text{and } \rho_{ij} = \rho \text{ for all } i, j\ (1,2,\ldots,k) \end{array}\right| \qquad (6.1.20)$$

as an example of uniform weights. Then we get

$$b_i = \rho_o \sigma_o \sigma$$

$$a_{ii} = (\sigma_o^2 + 2\sigma_o \sigma \rho_o + \sigma^2)$$

and $$a_{ij} = (\sigma_o^2 + 2\sigma\sigma_o \rho_o + \rho\sigma^2)$$

which gives

$$M(T_{pk}) = \frac{\theta_o^2}{k} [k a_{ii} + k(k-1) a_{ij}]$$

$$= \frac{\theta_o^2}{k} [\sigma^2(1-\rho) + k(\sigma_o^2 + 2\sigma\sigma_o \rho_o + \rho\sigma^2)]$$

and also $B(T_{pk})$ as in (6.1.18). Hence the theorem. Further, if in addition, we assume

$$\rho_o = \rho \quad \text{and} \quad \sigma_o = \sigma$$

we get,

$$B(T_{pk}) = \theta_o \rho \sigma^2$$

and

$$M(T_{pk}) = \frac{\theta_o^2}{k} [(k+1) - \rho(1-3k)]\sigma^2. \qquad (6.1.22)$$

## 6.2 Multivariate Ratio Method of Estimation:

Olkin's (1958) multivariate ratio estimator for estimating the parametric function $\theta_o$ can be expressed as

$$T_{rk} = \sum_{i=1}^{k} w_i r_i \theta_i \qquad (6.2.1)$$

where

$$w_i = \bar{X}_0 / \bar{X}_i \quad \text{and} \quad \sum_{i=1}^{k} w_i = 1. \qquad (6.2.2)$$

Here again we assume that $\theta_1, \theta_2, \ldots, \theta_k$ are known in advance. We have the following theorem 6.2.1 regarding the bias and m.s.e. of the estimator $T_{rk}$ in (6.2.1).

<u>Theorem 6.2.1</u> (Olkin): For any design P, with the procedure of selection mentioned in Lemma 6.1.1, the bias and m.s.e of $T_{rk}$, to order $n^{-1}$, are given by

$$B(T_{rk}) = \theta_0 \, w^{*\prime} b^*,$$

and

$$M(T_{rk}) = \theta_0^2 \, w^{*\prime} A^* w^*, \qquad (6.2.3)$$

respectively, where $b^*$ and $w^*$ are vectors represented by $(b_1^*, b_2^*, \ldots, b_k^*)$ and $(w_1^*, w_2^*, \ldots, w_k^*)$ respectively and $b_i^*$ is defined as

$$b_i^* = C_i^2 - C_0 C_i \rho_{0i}.$$

Further, $A^*$ in (6.2.3) is a matrix of order $k \times k$, $A^* = (a_{ij}^*)$, where

$$a_{ij}^* = (C_0^2 - C_0 C_i \rho_{0i} - C_0 C_j \rho_{0j} + C_i C_j \rho_{ij})$$

$C_1$ etc., being defined earlier.

Proof of the theorem is a similar to the proof given in Theorem 6.1.1 and is omitted.

The optimum weights can be obtained in a similar way by minimising $M(T_{rk})$ and it will be observed that these weights turn out to be same as in (6.1.16) with matrix $A$ replaced by the matrix $A^*$.

Theorem 6.2.2 Under the condition (6.1.20), the bias and m.s.e. of $T_{rk}$, to order $n^{-1}$, are given by

$$B(T_{rk}) = \theta_o C^2(1-\rho_o)$$

and $$M(T_{rk}) = \frac{\theta_o^2}{k}[C^2(1+\rho)+k(C_o^2-2CC_o\rho_o+\rho C^2)]. \qquad (6.2.4)$$

Proof is straightforward.

If in addition (6.1.20), $C = C_o$ and $\rho = \rho_o$, then

$$B(T_{rk}) = \theta_o C^2(1-\rho)$$

and

$$M(T_{rk}) = \frac{\theta_o^2}{k} C^2(1+k)(1-\rho). \qquad (6.2.5)$$

6.3 Comparison of Estimators:

We now compare the estimators $T_{pk}$ and $T_{rk}$ with $T_o$, the linear unbiased estimator, under the condition of (6.1.20) uniform weights and prove the following.

<u>Theorem 6.3.1</u> For any design $D(A,S,P)$, with selection procedures mentioned in Lemma 6.1.1, under the condition (6.1.20), considering the approximation to order $n^{-1}$, the preference regions for the estimator $T_{pk}$, $T_o$ and $T_{rk}$ are given by

$$-1 \leq \frac{k\,\rho_o}{1+(k-1)\rho} \leq -\frac{1}{2}\left(\frac{C}{C_o}\right)$$

$$-\frac{1}{2}\left(\frac{C}{C_o}\right) \leq \frac{k\,\rho_o}{1+(k-1)\rho} \leq +\frac{1}{2}\left(\frac{C}{C_o}\right)$$

$$+\frac{1}{2}\left(\frac{C}{C_o}\right) \leq \frac{k\,\rho_o}{1+(k-1)\rho} \leq +1 \qquad (6.3.1)$$

respectively.

<u>Proof.</u> The sampling variance of the unbiased estimator $T_o$ is

$$V(T_o) = \theta_o^2\, C_o^2. \qquad (6.3.2)$$

Comparing the $M(T_{pk})$ in (6.1.19) and $V(T_o)$ in (6.3.2), it is seen that $T_{pk}$ will be more efficient than $T_o$ if

$$\frac{C(1-\rho)}{k} + 2\rho_o C_o + \rho C < 0$$

and similarly comparing $M(T_{rk})$ in (6.2.4) with $V(T_o)$, $T_{rk}$ is found to be more efficient whenever

$$\frac{C(1-\rho)}{k} - 2\rho_o C_o + \rho C < 0,$$

which in turn gives the regions of preference mentioned in

(6.3.1). Hence the theorem.

Remark 6.3.1 It is easily seen that the Theorem 6.3.1 is a generalisation of Theorem 5.2.1 since for $k = 1$, the former gives the results obtained in the latter.

Remark 6.3.2. The preference regions in (6.3.1) hold good only if all $\theta_i$'s are positive and also when the number of positive and negative parameters are same, in case, in case otherwise, the conditions for $T_{pk}$ and $T_{rk}$ gets interchanged.

Remark 6.3.3 If in addition to (6.1.20), $C = C_o$ and $\rho = \rho_o$ then $T_{pk}$ will be more efficient if

$$\rho < -\frac{1}{(3k - 1)}$$

and $T_{rk}$ will be more efficient if

$$\rho > \frac{1}{k + 1} .$$

Further, it is pertinent to compare the univariate product estimator $(T_{p1})$ considered in the previous chapter and the multivariate product estimator. We prove the following.

Theorem 6.3.2 Let $T_{pk'}$ and $T_{pk}$ be the multivariate product estimators and $T_{rk'}$ and $T_{rk}$ be the multivariate ratio estimators of $\theta_o$, with optimum weights based on $k'$ and $k$ supplementary real-valued characteristics, where $k$ is greater than $k'$, then

$$M(T_{pk}) \leq M(T_{pk'}) \qquad (6.3.3)$$

and

$$M(T_{rk}) \leq M(T_{rk'}). \qquad (6.3.4)$$

The inequality (6.3.4) is a straightforward extension of Olkin's result and (6.3.3) can also be obtained in a similar manner.

As a special case of the above theorem, when weights are uniform and (6.1.20) is satisfied then

$$M(T_{pk'}) - M(T_{pk}) = \theta_o^2 C^2 (1-\theta)\left(\frac{k-k'}{kk'}\right)$$

which is always positive. Thus if $k' = 1$ and $k = 2$ then this implies that $T_{p2}$ will be more efficient than $T_{p1}$. Similar results hold for ratio estimator also.

An Empirical Study

For the purpose of this study we again refer to the multivariate investigation of blood chemistry undertaken by the Biometry Research Unit of the Indian Statistical Institute. In the present case we compare the univariate product estimator $T_{p1}$ with that of two-variate product estimator $T_{p2}$ for estimating the 'Eosinophil' content based on data for Group C. Let us denote by $y$ the variable under study and $x_1$ and $x_2$ the supplementary variables, height and weight respectively. The problem is to estimate $\bar{Y}$ based on data for 69 individuals belonging to Group C for a simple random sample of size n. We compare below the unbiased estimator $T_0$ with the univariate and two-variate estimators $T_{p1}$ and $T_{p2}$ respectively.

For two-variate product estimator, we have $M(T_{p2}) = \bar{Y}^2 \, wAw'$ where $w = (w_1, w_2)$, $w' = \binom{w_1}{w_2}$ and

$$A = (a_{ij})_{2\times 2} = \begin{pmatrix} C_0^2 + 2C_{01} + C_1^2 & C_0^2 + C_{01} + C_{02} + C_{12} \\ C_0^2 + C_{02} + C_{01} + C_{21} & C_0^2 + 2C_{02} + C_2^2 \end{pmatrix}.$$

The optimum weight

$$w_1 = \frac{2C_0^2 + 3C_{02} + C_2^2 + C_{01} + C_{12}}{4(C_0^2 + C_{01} + C_{02}) + C_1^2 + C_2^2 + 2C_{12}} = 1 - w_2.$$

It may be mentioned that $C_i$ and $\rho_{ij}$ $(i \neq j = 0,1,2)$ are the coefficient of variation and correlation coefficient respectively for the sample means and for the present scheme $C_i^2 = QC_i'^2$, where $Q = \frac{N-n}{(N-1)n}$ and $C_i'$ are defined for the corresponding variables. We have the following values for this population,

$$C_0' = 0.60 \qquad \rho_{01} = -0.1752$$

$$C_1' = 0.033 \qquad \rho_{02} = -0.2505$$

$$C_2' = 0.28 \qquad \rho_{12} = 0.0099$$

which gives $w_1 = 0.48$ and $w_2 = 0.52$ and the corresponding m.s.e. as

$$M(T_{p1}) = \bar{\bar{Y}}^2 Q(0.3541) \quad \text{and} \quad M(T_{p2}) = \bar{\bar{Y}}^2 Q\,(0.3343)$$

showing that the relative efficiency of unbiased, univariate product and two-variate product estimator is 100, 104 and 110 respectively. It may be noted that this example is being given here by a way of illustration. Here the gain in efficiency is marginal, due to the fact that $\rho_{01}$ and $\rho_{02}$, although negative and quite small in magnitude.

6.4 Estimators in Two-Phase Sampling:

Two-phase sampling scheme is applied when the information on the real-valued supplementary characteristics is not available in advance. A large initial sample usually known as first-phase sample is selected to observe the values of the supplementary characteristics and in the second-phase sample, which can be either a sub-sample of this large sample or an independent sample of smaller size data on both the characteristic under study and the supplementary characteristics are collected. It is further assumed that cost of collecting data on initial sample is very small compared to that for the second-phase sample.

We shall further assume, for the sake of simplicity, that simple random sampling without-replacement has been applied at both the phases and that the parameter under consideration is the population mean $\bar{\bar{Y}}$. [However, some complex two-phase schemes have been considered in the last chapter of this thesis.]

Case 1. Sub-sample of the initial sample:

Following (6.1.1) an estimator of the population mean can be defined as

$$\hat{\bar{Y}}_{pkt} = \sum_{i=1}^{k} \frac{w_i p_i}{\bar{X}_i'} \qquad (6.4.1)$$

where $p_i = \bar{\bar{y}}\,\bar{\bar{x}}_i$, $\bar{\bar{y}}$ and $\bar{\bar{x}}_i$ being the sample means based on the second-phase sample of size $n_2$ and $\bar{X}_i'$ denote the mean for the i-th supplementary characteristic based on the first phase sample of size $n_1$. We get the m.s.e as

$$M(\hat{\bar{Y}}_{pkt}) = \bar{Y}^2 \sum_i \sum_j w_i w_j \,\mathrm{Cov}\,(p_i, p_j)/P_i' P_j'$$

$$= \bar{Y}^2 \sum_i \sum_j w_i w_j d_{ij}$$

$$= \bar{Y}^2 \, wDw' \qquad (6.4.2)$$

where

$$d_{ij} = \frac{1}{n_2}(1-f_2)C_0^2 + \frac{1}{n_2}(1-f_2)(C_i C_j \rho_{ij} + \rho_{0i} C_0 C_i + \rho_{0j} C_0 C_j), \qquad (6.4.3)$$

$D = (d_{ij})_{k \times k}$, $P_i' = \bar{Y}.\bar{X}_i'$ and $f_2 = (\frac{n_2}{N})$.

Optimum values of $w_i$'s can be defined in a similar way by replacing matrix A with D.

Multivariate ratio estimator for this scheme can be defined from (6.2.1) as

$$\hat{\bar{Y}}_{rkt} = \sum_{i=1}^{k} w_i r_i \bar{X}_i' \tag{6.4.4}$$

where $r_i = \bar{y} / \bar{x}_i$ and we get

$$M(\hat{\bar{Y}}_{rkt}) = \bar{Y}^2 \sum_i \sum_j w_i w_j \text{ Cov } (r_i, r_j)/R_i' R_j'$$

$$= \bar{Y}^2 \sum_i \sum_j w_i w_j d^*_{ij}$$

$$= \bar{Y}^2 \, wD^*w'$$

where

$$d^*_{ij} = \frac{1}{n_2}(1 - f_2) C_0^2 + \frac{1}{n_2}(1 - f_2)(C_i C_j \rho_{ij} - \rho_{oi} C_o C_i - \rho_{oj} C_o C_j),$$

$$D^* = (d^*_{ij}), \quad R_i' = \bar{Y} / \bar{X}_i. \tag{6.4.5}$$

Optimum weights $w_i$'s are obtained by replacing the matrix $A^*$ by $D^*$.

The m.s.e of univariate product and ratio estimator when the weights are uniform are given by

$$M(\hat{\bar{Y}}_{pt}) = \frac{\bar{Y}^2}{n_2}[(1 - f_2) C_o^2 + (1 - f_1)(C^2 + 2\rho_o C_o C)] \tag{6.4.6}$$

and

$$M(\hat{\bar{Y}}_{rt}) = \frac{\bar{Y}^2}{n_2}[(1 - f_2) C_o^2 + (1 - f_1)(C^2 - 2\rho_o C_o C)] \tag{6.4.7}$$

where $f_1 = \frac{n_2}{n_1}$ and $f_2 = \frac{n_2}{N}$.

Ignoring the terms of $O(N^{-1})$,

$$M(\hat{\bar{Y}}_{pt}) = \bar{Y}^2\left[\frac{C_o^2 + C^2 + 2\varrho_o C_o C}{n_2} + \frac{-2\varrho_o C_o C - C_o^2}{n_1}\right] \qquad (6.4.9)$$

and

$$M(\hat{\bar{Y}}_{rt}) = \bar{Y}^2\left[\frac{C_o^2 + C^2 - 2\varrho_o C_o C}{n_2} + \frac{2\varrho_o C_o C - C_o^2}{n_1}\right]. \qquad (6.4.9)$$

Case II. Independent sub-sample:-

The form of the estimators $\hat{\bar{Y}}_{pkt}$ and $\hat{\bar{Y}}_{rkt}$ will remain the same as (6.4.1) and (6.4.4) respectively. The m.s.e in this case is given by

$$M(\hat{\bar{Y}}_{pkt}) = \bar{Y}^2\, wGw'$$

and $\qquad (6.4.10)$

$$M(\hat{\bar{Y}}_{rkt}) = \bar{Y}^2\, w^*G^*w^{*\prime}$$

where $G = (g_{ij})_{k \times k}$ and $G^* = (g^*_{ij})_{k \times k}$

and

$$g_{ij} = \frac{1}{n_2}(1-f_2)(C_o^2 + \varrho_{oi}C_oC_i + \varrho_{oj}C_oC_j + \varrho_{ij}C_iC_j) +$$
$$+ \frac{1}{n_1}\left(1-\frac{n_1}{N}\right)\varrho_{ij}C_iC_j$$

$$g^*_{ij} = \frac{1}{n_2}(1-f_2)(C_o^2 - \varrho_{oi}C_oC_i - \varrho_{oj}C_oC_j + \varrho_{ij}C_iC_j) +$$
$$+ \frac{1}{n_1}\left(1-\frac{n_1}{N}\right)\varrho_{ij}C_iC_j.$$

Again the m.s.e of ratio and product estimators (assuming weights to be uniform) and considering term upto $O(N^{-1})$, are given by

$$M(\hat{\bar{Y}}_{pt}) = \bar{Y}^2 \left[\frac{C_o^2 + 2\rho_o C_o C + C^2}{n_2} + \frac{C^2}{n_1}\right] \qquad (6.4.11)$$

and

$$M(\hat{\bar{Y}}_{rt}) = \bar{Y}^2 \left[\frac{C_o^2 - 2\rho_o C_o C + C^2}{n_2} + \frac{C^2}{n_1}\right]. \qquad (6.4.12)$$

It may be noted that the expressions (6.4.9) and (6.4.12) are the same as given by Cochran (1963) and a comparison between (6.4.12) and (6.4.11) will lead to the same preference regions mentioned in Theorem (5.2.1).

# CHAPTER VII

## SOME RATIO CUM PRODUCT ESTIMATORS

7.0 Summary: Estimation of a non-linear parametric function (nlf) of the parameters, along with the parameters themselves, have been considered in this Chapter. Section 7.1 deals with estimation of a nlf and in Section 7.2 some estimators for the ratio and products (special cases of the nlf) of the parameter have been proposed. Sections 7.3 and 7.4 gives comparison of these estimators with the usual ratio and product estimators. As these estimators are biased some almost unbiased estimators are suggested in Section 7.5 and in Section 7.6 two combinations of the estimators, suggested above, have been proposed and compared. Illustrations for the gain in efficiency of these estimators are also given on the basis of a live data.

Lastly these estimators have been considered for estimating the parameters themselves and some comparisons with other relevant estimators are also made in Sections 7.7 and 7.8.

7.1 The Non-Linear Parametric Function:

Estimation of a non-linear function (nlf) of the parameters, apart from the parameters themselves, are of considerable interest in practice. Let us denote $\eta(\theta)$ as a nlf of the parameters $\theta_0, \theta_1, \ldots, \theta_k$ where $\theta_0$ is a function of the character $y$ and $\theta_i$'s are functions defined for the characters $x_i$ $(i = 1, 2, \ldots, k)$.

Then we can express $\eta(\theta)$ as

$$\eta(\theta) = f(\theta_0, \theta_1, \ldots, \theta_k).$$

Two important examples of $\eta(\theta)$, for $k = 1$, are the population ratio $R$ and the product $P$, given by

$$R = \theta_0 / \theta_1 \quad \text{and} \quad P = \theta_0 . \theta_1 \tag{7.1.1}$$

respectively. The estimation of ratio of ratios, that is

$$\eta(\theta) = R_1 / R_2$$

where $R_1 = \theta_0/\theta_1$ and $R_2 = \theta_2/\theta_3$ are also of considerable interest in some situations(Keyfitz, see Yates 1960; Kish 1960, Rao 1957).

These nlf besides $\theta_i$'s themselves are often required to be estimated. For instance, it is often required to estimate the proportion of population under different means of livelihood or the total crop production, which is product of cultivated area and the yield rate, besides the total population or the total cultivated area.

For estimating such nlf of the parameters, each of which can be unbiasedly estimated, usually the same function of the unbiased estimators of the parameters is taken as an estimator. For instance in estimating the ratio $R$ and the product $P$ of two parameters $\theta_0$ and $\theta_1$, the commonly adopted practice is to take the ratio and product respectively, of the unbiased estimators of the parameters, as an estimator. However, in this case, whenever, the true value of one of the parameters is available one may feel that it is sufficient to estimate the other parameter only and obtain an estimator of the true ratio or product by using the known value of the parameter. It is interesting to note that even if true value of either of the parameter is known, the ratio or product of unbiased estimators proves to be better, in many cases, than the estimator obtained by using the actual value of the parameter, mainly because of the correlation between the estimators. That is, for instance if the ratio $R$ is of interest and it is believed that the ratio of $y$ to $x_1$ is less variable than $y$'s themselves then it would be better to estimate $R$ from the sample itself, instead of estimating $\theta_0$ from the sample and then dividing it by the known parameter $\theta_1$. Similarly if the product $P$ is to be estimated and it is believed that the product of $y$ and $x_1$ is less variable then it is better to estimate $P$ on the basis of sample estimates of $\theta_0$ and $\theta_1$ as compared to estimating simply $\theta_0$ and multiplying it by known parameter $\theta_1$.

Let us denote $G(T)$ as an estimator of $\eta(\theta)$, then usually $G(T)$ is a function of $T_0, T_1, \ldots, T_k$, where $T_i$'s $(i = 1,2,\ldots,k)$ are unbiased estimators of $\theta_i$'s for any design $D(U,S,P)$.

Next, let us suppose that information on a supplementary character $x_{k+1}$ related to $y$ and $x_i$ $(i = 1,\ldots, k)$ is available in the survey and we are interested in using this information if it is found to be beneficial. Let us denote the parameter relating to $x_{k+1}$ by $\theta_{k+1}$ and suppose that it is known and let $T_{k+1}$ be its unbiased estimator from the sample. Then an estimator $G^*(T)$, which utilises supplementary information, is given by

$$G^*(T) = G(T)\left(\frac{T_{k+1}}{\theta_{k+1}}\right)^{\xi} \quad \ldots \quad (7.1.2)$$

where $\xi$ is a constant and its value may be obtained by minimising the mean square error of $G^*(T)$. Obviously $\xi$ will be a function of the coefficient of variations of $T_i$ and the correlation coefficient of between $T_i$ and $T_j$ for $i(\neq j) = 0, 1, \ldots, k+1$.

In case $\theta_{k+1}$ is not known in advance and if it is found easier and cheaper to observe $x_{k+1}$ than $y$ and $x_i$'s then a two-phase sample may be used. This procedure has already been discussed in a earlier chapter hence this discussion is omitted here.

It is well-known that if $\eta(\theta)$ is a linear function of $\theta_i$'s then $G(T)$ will be unbiased, for instance, if $\eta(\theta) = \theta_0 - \theta_1$ then $G(T) = T_0 - T_1$ is unbiased for $\eta(\theta)$. But as $\eta(\theta)$ is nlf, the corresponding estimator $G(T)$ and also the proposed estimator $G^*(T)$ will be biased. The approximate expressions for the bias and mean square error of these estimators can be obtained by using Taylor's series. In this chapter we consider the estimation problems in relation to the ratio $R$ and product $P$ given in (7.1.1) as a special case of $\eta(\theta)$.

7.2 Estimators for Ratio and Product:

The commonly adopted method of estimating the ratio R and the product P, as indicated earlier, is to find the ratio

$$r = \frac{\bar{y}_0}{\bar{x}_1} \tag{7.2.1}$$

and the product

$$p = \bar{y}_0\bar{x}_1 \tag{7.2.2}$$

Further, it is well-known that r and p, respectively, estimate the true ratio R and product P efficiently only if the study characters y and $x_1$ are highly positively correlated in the former case and highly negatively correlated in the latter. In practice, however, we come across several situations where the above condition is not satisfied and the question then arises what estimator to use in such cases.

In other words, while estimating R, we come across situations in which the ratio of y to x is believed to be highly variable along with y's themselves or on the other hand while estimating P, the product of y and x is believed to be highly variable along with y's themselves. In such situations both the estimators r and $T_0/\hat{\theta}_1$ of R are likely to have large errors and similarly the estimators p and $T_0\hat{\theta}_1$ are likely to have large error for estimation of P. We suggest below an estimator, on the basis of the estimator $G^*(T)$ for estimating $\eta(\theta)$, which may be efficiently utilised for the above situations for estimating R and P. For estimating the ratio of ratios, $R_1/R_2$, Keyfitz (see Yates 1960), Rao (1957) and Kish (1960) have considered double ratio estimators. The estimators suggested here utilise the knowledge of a supplementary character and are given by

$$R^* = r\left(\frac{T_2}{\theta_2}\right)^{\xi_2} \qquad (7.2.3)$$

and

$$P^* = P\left(\frac{T_2}{\theta_2}\right)^{\delta_2} \qquad (7.2.4)$$

respectively, where $\theta_2$ is assumed to be known and $\xi_2$ and $\delta_2$ are constants and may be determined by minimising the mean square error of the estimators $R^*$ and $P^*$ respectively.

Thus writing $T_i = \theta_i(1 + e_i)$ for $i = 0,1,2$ such that $E(e_i) = 0$ and assuming $|e_i| < 1$ the m.s.e of $R^*$ and $P^*$ are given by

$$M(R^*) = E(R^* - R)^2$$

$$= R^2E(e_0^2 + e_1^2 + \xi_2^2e_2^2 - 2e_0e_1 + 2\xi_2e_0e_2 - 2\xi_2e_1e_2 \ldots)$$

$$\doteq M(r) + R^2(\xi_2^2C_2^2 + 2\xi_2C_{02} - 2\xi_2C_{12})\ldots \qquad (7.2.5)$$

and $M(P^*) = E(P^* - P)^2$

$$= P^2E(e_0^2 + e_1^2 + \delta_2^2e_2^2 + 2e_0e_1 + 2\delta_2e_0e_2 + 2\delta_2e_1e_2 \ldots)$$

$$\doteq M(p) + P^2(\delta_2^2C_2^2 + 2\delta_2C_{02} + 2\delta_2C_{12}). \qquad (7.2.6)$$

Minimising $M(R^*)$ and $M(P^*)$ for values of $\xi_2$ and $\delta_2$

respectively, we get their optimum value as

$$\xi_2^* = \left(\frac{C_1}{C_2}\right)\rho_{12} - \left(\frac{C_0}{C_2}\right)\rho_{02} \tag{7.2.7}$$

and

$$\delta_2^* = -\left(\frac{C_1}{C_2}\right)\rho_{12} - \left(\frac{C_0}{C_2}\right)\rho_{02} \tag{7.2.8}$$

respectively.

Thus these constants involve the values of coefficients of variation of $T_0$, $T_1$ and $T_2$ and the correlation coefficients between them, and sometimes it may be difficult in practice to guess the value of $\xi_2$ and $\delta_2$. For the sake of simplicity let us consider the value of $\xi_2$ and $\delta_2$ as 1,0 and -1. Then for $\xi_2$ and $\delta_2$ to be zero $R^*$ and $P^*$ are same as $r$ and $p$ respectively and for the values 1 and -1, we get,

$$R_1^* = (rT_2)/\theta_2 \quad \text{and} \quad R_2^* = (r\,\theta_2)/T_2 \tag{7.2.9}$$

as the two estimators of $R$ and

$$P_1^* = (p\,T_2)/\theta_2 \quad \text{and} \quad P_2^* = (p\,\theta_2)/T_2 \tag{7.2.10}$$

as the two estimators of $P$.

We obtain in the following, approximate expressions (to the second degree) for the bias and mse of these estimators. Writing $T_i = \theta_i(1 + e_i)$, $i = 0,1,2$ where $E(e_i) = 0$ for all $i$ and $|e_i| < 1$ for large samples (whenever the estimator is in the denominator) so as to make ~~the~~ expansion valid. We shall denote $B(r)$, $B(R_1^*)$, ~~$B(R_1^*)$~~, $B(R_2^*)$, $B(p)$, $B(P_1^*)$ and $B(P_2^*)$ as the biases and $M(r)$, $M(R_1^*)$, $M(R_2^*)$, $M(p)$, $M(P_1^*)$ and $M(P_2^*)$ as the mean square errors of the corresponding estimators. Then the biases are

$$
\begin{aligned}
B(r) &= E(r-R) \\
&= RE[(1+e_0)(1-e_1+e_1^2 \dots)-1] \\
&= RE[(e_0-e_1)+(e_1^2-e_0e_1)\dots] \\
&\doteq R(C_1^2-C_{01})
\end{aligned}
\qquad (7.2.11)
$$

$$
\begin{aligned}
B(R_1^*) &= RE[(1+e_0)(1+e_2)(1-e_1+e_1^2 \dots)-1] \\
&= RE[(e_0+e_2-e_1)+(e_0e_2-e_0e_1-e_1e_2+e_1^2)+\dots] \\
&\doteq B(r)-R(C_{12}-C_{02})
\end{aligned}
\qquad (7.2.12)
$$

$$
\begin{aligned}
B(R_2^*) &= RE[(1+e_0)(1-e_2+e_2^2\dots)(1-e_1+e_1^2 \dots)-1] \\
&= RE[(e_0-e_1-e_2)+(e_1^2+e_2^2-e_0e_1-e_0e_2+e_1e_2)+\dots] \\
&\doteq B(r)+R(C_2^2-C_{01}-C_{02}+C_{12})
\end{aligned}
\qquad (7.2.13)
$$

$$
\begin{aligned}
B(p) &= PE[(e_0+e_1)+e_0e_1] \\
&= P\,C_{01}.
\end{aligned}
\qquad (7.2.14)
$$

$$B(P_1^*) = PE[(1+e_0)(1+e_1)(1+e_2) - 1]$$

$$= PE[(e_0 + e_1 + e_2)+(e_0e_1 + e_0e_2 + e_1e_2)]$$

$$\doteq B(p) + P(C_{02} + C_{12}) \qquad (7.2.15)$$

$$B(P_2^*) = PE[(1 + e_0)(1+ e_1)(1- e_2 + e_2^2 + \dots ) - 1]$$

$$= PE[(e_0 + e_1 - e_2)+(e_0e_1 - e_0e_2 - e_1e_2 + e_2^2)+ \dots]$$

$$\doteq B(p) + P(C_2^2 - C_{02} - C_{12}). \qquad (7.2.16)$$

Similarly the mean square errors are

$$M(r) = E(r - R)^2$$

$$\doteq R^2(C_0^2 + C_1^2 - 2C_{01}). \qquad (7.2.17)$$

$$M(R_1^*) = R^2E(e_0^2 + e_1^2 - 2e_0e_1 + e_2^2 + 2e_0e_2 - 2e_1e_2 + \dots )$$

$$\doteq M(r) + R^2(C_2^2 + 2C_{02} - 2C_{12}) \qquad (7.2.18)$$

$$M(R_2^*) = R^2E(e_0^2 + e_1^2 + e_2^2 - 2e_0e_1 - 2e_0e_2 + 2e_1e_2 + \dots )$$

$$\doteq M(r) + R^2(C_2^2 - 2C_{02} + 2C_{12}). \qquad (7.2.19)$$

$$M(p) = E(p - P)^2$$

$$\doteq P^2(C_0^2 + C_1^2 + 2C_{01}) \qquad (7.2.20)$$

$$M(P_1^*) = P^2E(e_0^2 + e_1^2 + e_2^2 + 2e_0e_1 + 2e_0e_2 + 2e_1e_2 \dots )$$

$$\doteq M(p) + P^2(C_2^2 + 2C_{02} + 2C_{12}). \qquad (7.2.21)$$

$$M(P_2^*) = P^2E(e_0^2 + e_1^2 + e_2^2 + 2e_0e_1 - 2e_0e_2 - 2e_1e_2 \dots)$$

$$\doteq M(p) + P^2(C_2^2 - 2C_{02} - 2C_{12}) \qquad (7.2.22)$$

where $C_i = V(T_i)/\theta_i$ and $C_{ij} = C_iC_j\rho_{ij}$ for $i,j = 0,1,2$ and $\rho_{ij}$ $(i \neq j)$ are the correlation coefficients between $T_i$ and $T_j$. Estimators of the mse's may be obtained by substituting the estimates of $C_i$ and $C_{ij}$ in the corresponding expressions.

### 7.3 Comparison of Estimators:

In this section we shall compare the proposed estimators with those of the usual estimators for their mse's and get the following two Theorems.

Theorem 7.3.1: For a given design $D(U,S,P)$, the estimators $R_1^*$ and $R_2^*$ will be more efficient than the usual estimator $r$ for estimating the ratio $R$ if the conditions

$$\rho_{02}\left(\frac{C_0}{C_2}\right) - \rho_{12}\left(\frac{C_1}{C_2}\right) < -\frac{1}{2} \qquad (7.3.1)$$

and

$$\rho_{02}\left(\frac{C_0}{C_2}\right) - \rho_{12}\left(\frac{C_1}{C_2}\right) > \frac{1}{2} \qquad (7.3.2)$$

respectively, are satisfied.

<u>Proof:</u> Comparing $M(r)$ in (7.2.17) with those of $M(R_1^*)$ and $M(R_2^*)$ in (7.2.18) and (7.2.19) respectively, it is observed that $R_1^*$ and $R_2^*$ to be more efficient, the terms in the bracket should be less than zero, which on rearrangement gives the corresponding conditions stated in the above Theorem.

<u>Theorem 7.3.2.</u> For a given design $D(U,S,P)$ the estimators $P_1^*$ and $P_2^*$ will be more efficient than the usual estimator $p$ for estimating the product $P$ if the conditions

$$\rho_{02}\left(\frac{C_0}{C_2}\right) + \rho_{12}\left(\frac{C_1}{C_2}\right) < -\frac{1}{2} \tag{7.3.3}$$

and

$$\rho_{02}\left(\frac{C_0}{C_2}\right) + \rho_{12}\left(\frac{C_1}{C_2}\right) > \frac{1}{2} \tag{7.3.4}$$

respectively, are satisfied.

Proof of the theorem follows by comparing (7.2.20) to (7.2.22) and rearranging the terms.

<u>Remark 7.3.1:</u> In the above comparisons the LHS of equations (7.3.1) and (7.3.2) is same as $-\xi_2^*$ and the LHS of equation (7.3.3) and (7.3.4) is expressible as $-\delta_2^*$.

<u>Remark 7.3.2:</u> It may be mentioned that the conditions for efficiency of the estimators have been derived under the assumption that all the parameters $\theta_0$, $\theta_1$ and $\theta_2$ are positive or all are negative. Otherwise, the conditions will get interchanged. These changes are shown in the following table.

Table: Showing the interchange of conditions in the Theorems 7.3.1 and 7.3.2 depending on the sign of $\theta_i$'s.

| sign of | | | conditions for preference of | | | |
|---|---|---|---|---|---|---|
| $\theta_0$ | $\theta_1$ | $\theta_2$ | $R_1^*$ | $R_2^*$ | $R_1^*$ | $P_2^*$ |
| + | + | + | (7.3.1) | (7.3.2) | (7.3.3) | (7.3.4) |
| - | - | - | (7.3.1) | (7.3.2) | (7.3.3) | (7.3.4) |
| + | + | - | (7.3.2) | (7.3.1) | (7.3.4) | (7.3.3) |
| - | - | + | (7.3.2) | (7.3.1) | (7.3.4) | (7.3.3) |
| + | - | + | (7.3.3) | (7.3.4) | (7.3.1) | (7.3.2) |
| - | + | - | (7.3.3) | (7.3.4) | (7.3.1) | (7.3.2) |
| - | + | + | (7.3.4) | (7.3.3) | (7.3.2) | (7.3.1) |
| + | - | - | (7.3.4) | (7.3.3) | (7.3.2) | (7.3.1) |

Remark 7.3.3 The conditions for efficiency of the estimators are valid regardless of the value of the correlation coefficient between $T_0$ and $T_1$, ($\rho_{01}$). Hence these estimators may be more efficient than r (or p) even if $T_0$ and $T_1$ are positively (or negatively for p) correlated. However, in case $\rho_{01}$ takes a value +1, the usual ratio estimator r is the only estimator to be used provided the coefficient of variations are equal.

## 7.4 Configurational Representation:

The purpose of this section is to provide better appreciation of the preference regions obtained in the previous section. We shall assume

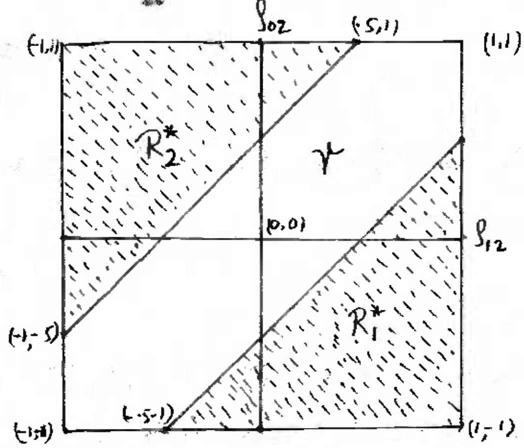
$$C_0 = C_1 = C_2 = C \qquad (7.4.1)$$

in which case the conditions for $R_1^*$ and $R_2^*$ to be more efficient than $r$ will turn out be

$$(\rho_{02} - \rho_{12}) < -\frac{1}{2}$$

and
$$(\rho_{02} - \rho_{12}) > \frac{1}{2}$$

respectively. These conditions may be represented with the help of the configuration $F_R$ given below.




$F_R$

Thus whenever it is possible to choose a supplementary character $x_2$ such that the pair $(\rho_{02}, \rho_{12})$ lies in either of the regions $R_1^*$ and $R_2^*$, then under the assumption (7.4.1), the corresponding estimator is an improvement over the usual ratio estimator. It may be mentioned that the relative increase in efficiency is higher when $y$ and $x_1$ are negatively correlated. For illustration let us consider two extreme situations

i) $\rho_{01} = -1.0$, $\rho_{02} = -1.0$ and $\rho_{12} = +1.0$

ii) $\rho_{01} = -1.0$, $\rho_{02} = +1.0$ and $\rho_{12} = -1.0$.

In the case (i) the point $(\rho_{02}, \rho_{12})$ lies in the region of $R_1^*$ whereas in case (ii) it lies in the region $R_2^*$ and we get

$$B(r) = 2RC^2 \quad \text{and} \quad M(r) = 4R^2C^2$$

in both the cases, while

$$B(R_1^*) = 0 \quad \text{and} \quad M(R_1^*) = (RC)^2$$

in the first case and

$$B(R_2^*) = B(r)/2 \quad \text{and} \quad M(R_2^*) = (RC)^2$$

in the second case.

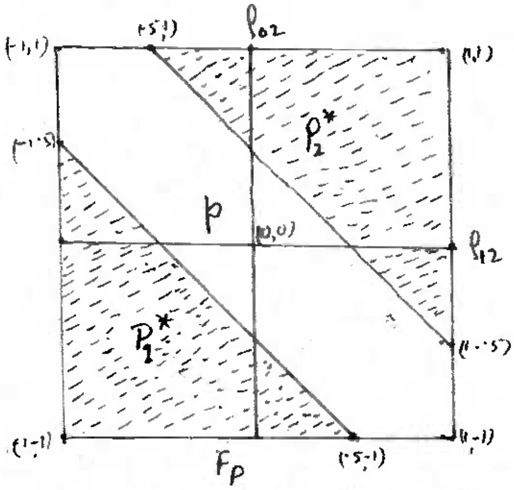
Similarly, the set of conditions obtained in Theorem 7.3.2 can also be given a configurational representation. Under the assumption (7.4.1), the conditions (7.3.3) and (7.3.4) take the form

$$(\rho_{02} + \rho_{12}) < -\frac{1}{2}$$

and

$$(\rho_{02} + \rho_{12}) > +\frac{1}{2}$$

respectively and we get the following configuration $F_p$.




Thus when the supplementary character $x_2$ is such that the pain ($\rho_{02}$, $\rho_{12}$) lies in either of the regions of $P_1^*$ and $P_2^*$, then the corresponding estimator is an improvement over the usual product estimator provided (7.4.1) is satisfied. It is again seen that the gain in efficiency is higher if $y_0$ and $x_1$ are found to have positive correlation. For illustration if (i) $\rho_{01} = +1.0$, $\rho_{02} = +1.0$ and $\rho_{12} = +1.0$ and (ii) $\rho_{01} = +1.0$, $\rho_{02} = -1.0$ and $\rho_{12} = -1.0$. Then under the assumption (7.4.1) $B(p) = PC^2$, $M(p) = 4P^2C^2$ in both the cases and $M(P_2^*) = M(p)/4$ with $B(P_2^*) = 0$ in the first case, and $M(P_1^*) = M(p)/4$ with $B(P_1^*) = 0$ in the second case.

An Empirical Study:

The data for all 61 blocks of Ahemedabad city Ward No. 1 (Khadia I) taken from 1961 Population Census have been considered for the purpose of the present study. It is intended to find the ratio R of total female workers (Y) to the total female population ($X_1$). The supplementary characters chosen for this purpose are (i) educated female ( $X_2$ ) and (ii) female population in services ( $X_3$ ) (group IX of the Population Census).

For this population we get,

$\bar{Y} = 7.46$ $\quad C_0'^2 = 0.5046$ $\quad \rho_{01} = 0.0388$

$\bar{X}_1 = 265.54$ $\quad C_1'^2 = 0.0379$ $\quad \rho_{02} = -0.2070$

$\bar{X}_2 = 179.00$ $\quad C_2'^2 = 0.0633$ $\quad \rho_{03} = 0.7731$

$\bar{X}_3 = 5.31$ $\quad C_3'^2 = 0.5737$ $\quad \rho_{12} = 0.7373$

and $\quad \rho_{13} = -0.0474$

...re $\bar{Y}$, $\bar{X}_1$, $\bar{X}_2$ and $\bar{X}_3$ denote mean for the corresponding ...aracters, $C_i'^2$ ($i = 0, 1, 2, 3$) stand for the square of their coefficient of variations of the characters and $\rho_{ij}$ ($i \neq j = 0, 1, 2, 3$) for the corresponding correlation coefficients.

Let us assume that a simple random sample of $n$ blocks is selected from this population to estimate the ratio $R$, in which case $C_i^2 = QC_i'^2$ where $Q = (N-n)/(N-1)n$.

For this population it is observed that we can use the characters $x_2$ and $x_3$ in constructing the proposed estimator

$$R_1^* = \left(\frac{\bar{y}}{\bar{x}_1}\right)\left(\frac{\hat{X}_2}{\bar{x}_2}\right) \quad \text{and} \quad R_2^* = \left(\frac{\bar{y}}{\bar{x}_1}\right)\left(\frac{\bar{x}_3}{\hat{X}_3}\right)$$

respectively. Further, we find from (7.2.7) the optimum value of $\xi_2$ (using $x_2$) and $\xi_3$ (using $x_3$) as

$$\xi_2 = 1.178 \qquad \xi_3 = -0.737.$$

The efficiency of the estimators is given in the following table.

Table 7.4.1: Showing efficiency of estimators.

| Estimators | (MSE)/ $R^2 Q$ | % Relative efficiency |
|---|---|---|
| $P$ | .5318 | 100 |
| $R_1^*(x_2)$ | .4491 | 118 |
| $R^*(x_2)$ | .4477 | 119 |
| $R_2^*(x_3)$ | .2593 | 205 |
| $R^*(x_3)$ | .2191 | 242 |

The supplementary character used in a particular estimator is indicated in the bracket. It is observed that $x_3$ is more suitable than $x_2$. Further, the gain in using the exact value of $\xi_2$ and $\xi_3$ is not much and therefore $R_1^*$ and $R_2^*$ which does not require them may safely be used.

### 7.5 Almost Unbiased Estimators:

We have seen in Section 7.2 that the proposed estimators, like the usual ratio and product estimators, are biased to order $n^{-1}$, where $n$ is the size of the sample, for estimating the ratio $R$ and product $P$. In this section some estimators corresponding to $R^*$ and $P^*$ are obtained following Quenouille's (1956) and Murthy and Nanjamma's (1959) techniques of bias reduction.

Quenouille's method of bias reduction, from order $n^{-1}$ to $n^{-2}$, consists in dividing the sample in to $g$ random groups each of size $m$, such that $mg = n$, and thus the estimators

$$t_r = gr - \frac{g-1}{g} \Sigma r_j' \qquad \ldots \qquad (7.5.1)$$

$$t_p = gp - \frac{g-1}{g} \Sigma p_j' \qquad \ldots \qquad (7.5.2)$$

$$t_{R^*} = gR^* - \frac{g-1}{g} \Sigma R_j^{*\prime} \qquad \ldots \qquad (7.5.3)$$

$$t_{P^*} = gP^* - \frac{g-1}{g} \Sigma P_j^{*\prime} \qquad (7.5.4)$$

corresponding $r$, $p$, $R^*$ and $P^*$ have bias of order $n^{-2}$ almost; where $r_j'$, $p_j'$, $R_j^{*\prime}$ and $P^{*\prime}$ are the corresponding estimators computed from the sample omitting the $j^{th}$ group. Durbin (1959) and J. N. K. Rao (1965) have studied some properties of $t_r$ assuming different models for the variates.

Murthy and Nanjamma have suggested a technique for getting almost unbiased (unbiased to order $n^{-2}$) estimators on the basis of interpenetrating sub-samples. The technique consists in drawing the sample in the form of $g$ independent interpenetrating sub-samples, and using the relation between the biases of the estimators computed from the sub-samples and that from the sample as a whole, to get almost unbiased estimators. Accordingly the estimators

$$t_r' = \frac{g}{g-1} r - \frac{1}{g(g-1)} \sum_1^g r_j \qquad (7.5.5)$$

$$t_p' = \frac{g}{g-1} p - \frac{1}{g(g-1)} \sum_1^g p_j \qquad (7.5.6)$$

$$t'_{R*} = \frac{g}{g-1} R* - \frac{1}{g(g-1)} \sum_{1}^{g} R^*_j \qquad (7.5.7)$$

$$t'_{P*} = \frac{g}{g-1} P* - \frac{1}{g(g-1)} \sum_{1}^{g} P^*_j \qquad (7.5.8)$$

are almost unbiased corresponding to $r$, $p$, $R^*$ and $P^*$, where $r_j$, $p_j$, $R^*_j$ and $P^*_j$ are the corresponding estimators computed from the sub-samples. The estimator $t'_r$ is given by Murthy and Nanjamma (1959) and $t'_p$ by Murthy (1964). It may be mentioned that $t_p$ and $t'_p$ are unbiased.

These estimators can also be obtained with the help of the technique developed by T. J. Rao (1966). Estimators corresponding to $R^*_1$, $R^*_2$ and $P^*_1$, $P^*_2$ may be obtained from the above expressions. Such estimators may also be obtained by substracting an unbiased estimator of the bias for the corresponding estimators. Efficiencies of these estimators are not being studied here.

## 7.6 Use of Two or More Supplementary Characters:

In the previous sections of this chapter we have considered the estimator $R^*$ and $P^*$ for estimating R and P, and two important special forms of $R^*$, namely $R^*_1$ and $R^*_2$ in (7.2.5), and that of $P^*$, namely $P^*_1$ and $P^*_2$ in (7.2.6), and compared them in detail. These estimators, as mentioned earlier, utilised only one supplementary character $x_2$ (or $x_3$). In this section we consider more general estimators using two or more such characters. For the sake of simplicity, we first consider the case of two characters which suggests an immediate extension.

Let us consider two supplementary characters $x_2$ and $x_3$, information on which is available, and denote the corresponding estimators of the parameters $\theta_2$ and $\theta_3$ by $T_2$ and $T_3$ respectively. Then the proposed estimators is the linear combination of $R^*$'s and $P^*$'s obtained by using $T_2$ and $T_3$. Thus an estimator for $R$ is

$$R_c^* = w_1 r\left(\frac{T_2}{\theta_2}\right)^{\xi_2} + w_2 r\left(\frac{T_3}{\theta_3}\right)^{\xi_3} \qquad (7.6.1)$$

where $w_1$ and $w_2$ are weights such that $w_1 + w_2 = 1$, and $\xi_2$ and $\xi_3$ are constants to be suitably chosen.

Next, writting $T_i = \theta_i(1+e_i)$ for $i = 0,1,2,3$ and assuming $|e_i| < 1$ for large samples, the bias and mse of $R_c^*$ will be given by

$$B(R_c^*) = RE[w_1(1+e_0)(1+e_1)^{-1}(1+e_2)^{\xi_2} + w_2(1+e_0)(1+e_1)^{-1}(1+e_3)^{\xi_3} - 1]$$

$$= RE[(e_0 - e_1 + e_1^2 - e_0 e_1) + w_1\xi_2 e_2 + w_1\xi_3 e_3 + w_1(\xi_2 e_0 e_2 - \xi_2 e_1 e_2 + \frac{\xi_2(\xi_2-1)}{2} e_2^2) + w_2(\xi_3 e_0 e_3 - \xi_3 e_1 e_3 + \frac{\xi_3(\xi_3-1)}{2} e_3^2) \ldots]$$

$$\doteq B(r) + R[w_1(\xi_2 C_{02} - \xi_2 C_{12} + \frac{\xi_2(\xi_2-1)}{2} C_2^2) + w_2(\xi_3 C_{03} - \xi_3 C_{13} + \frac{\xi_3(\xi_3-1)}{2} C_3^2)] \qquad (7.6.4)$$

and $\quad M(R_c^*) \doteq R^2E[e_o-e_1+w_1\xi_2e_2+w_2\xi_3e_3]^2$

$$= R^2E[(e_o-e_1)^2+w_1^2\xi_2^2e_2^2+w_2^2\xi_3^2+2w_1(\xi_2e_2e_o-\xi_2e_1e_2)$$

$$+2w_2(\xi_3e_oe_3-\xi_3e_1e_3)+2w_1w_2\xi_2\xi_3e_2e_3]$$

$$= M(r)+R^2[w_1^2\xi_2^2C_2^2+w_2^2\xi_3^2C_3^2+2w_1(\xi_2C_{o2}-\xi_2C_{12})$$

$$+2w_2(\xi_3C_{o3}-\xi_3C_{13})+2w_1w_2\xi_2\xi_3C_{23}] \qquad (7.6.5)$$

The optimum values of $w_1$ and $w_2$ may be determined, in the usual way by minimising (7.6.5) under the condition $w_1 + w_2=1$, and we get,

$$w_1 = \frac{\xi_3^2\, C_3^2 - \xi_2(C_{o2} - C_{12}) + \xi_3(C_{o3} - C_{13}) - \xi_2\xi_3\, C_{23}}{\xi_2^2C_2^2 + \xi_3^2\, C_3^2 - 2\xi_2\xi_3C_{23}}$$

$$= 1 - w_2. \qquad (7.6.6)$$

The optimum values of $\xi_2$ and $\xi_3$ in (7.6.1) may be obtained by minimising $M(R_c^*)$ using optimum weights. But since they will be complicated, in general, we shall consider the simple case of $\xi_2 = 1$ and $\xi_3 = -1$ as in the earlier case. Then the estimator $R_c^*$ may be written as linear combination of $R_1^*$ and $R_2^*$. That is

$$R_{c'}^* = w_1R_1^* + w_2R_2^*. \qquad (7.6.7)$$

The bias of this estimator may be directly obtained from $B(R_c^*)$ in (7.6.4). Similarly

$$M(R_{c'}^*) = M(r) + R^2[w_1^2C_2^2 + w_2^2C_3^2 + 2w_1(C_{o2}-C_{12}) - 2w_2(C_{o3}-C_{13})$$

$$- 2w_1w_2C_{23}]. \qquad (7.6.8)$$

It is interesting to note that $w_1 = w_2 = 1/2$ when

$$C_j = C, \quad \rho_{o2} = \rho_{13} \quad \text{and} \quad \rho_{o3} = \rho_{12},$$

in which case

$$M(R_{c'}^*) = \frac{R^2C^2}{2}[(1-\rho_{23}) - 4(\rho_{o1} + \rho_{o3} - \rho_{o2} - 1)]. \qquad (7.6.9)$$

Remark 7.6.1: A direct comparison of $R_c^*$ with $R^*$ or $R_{c'}^*$, with $R_1^*$ and $R_2^*$, will not lead to any usable conclusions. However, in general the weighted estimators are likely to fare better than the unweighted ones. But, the major draw-back with the weighted estimators is that of determination of optimum weights. It may happen that in some situations the values of the optimum weights are known in advance from some source and then no difficulty arises, however in such cases one has to be quite certain that the weights used are atleast near optimum as otherwise these estimators may fare worse than the unweighted ones. On the other hand, if the weights are not known and they are to be estimated from the sample data itself, then the total mse gets automatically increased.

We, therefore, suggest below an estimator, free from the weights ($w_i$'s), and which utilise information on both the supplementary character $x_2$ and $x_3$, for estimating the ratio $R$. The estimator is

$$R_c^{*\prime} = r\left(\frac{T_2}{\theta_2}\right)^{\xi_2}\left(\frac{T_3}{\theta_3}\right)^{\xi_3} \tag{7.6.10}$$

where $\xi_2$ and $\xi_3$ are constants similar to $\xi_2$ in (7.2.7). The bias and mse of this estimator is

$$B(R_c^{*\prime}) \doteq B(r)+R[\xi_2(C_{02}-C_{12})+\xi_3(C_{03}+C_{13})+\xi_2\xi_3C_{23} + \frac{\xi_2(\xi_2-1)}{2}C_2^2 + \frac{\xi_3(\xi_3-1)}{2}C_3^2] \tag{7.6.11}$$

and

$$M(R_c^{*\prime}) = M(r)+R^2[\xi_2^2C_2^2+\xi_3^2C_3^2+2\xi_2(C_{02}-C_{12})+2\xi_3(C_{03}-C_{13}) + 2\xi_2\xi_3C_{23}]. \tag{7.6.12}$$

Again the optimum values of $\xi_2$ and $\xi_3$ may be obtained by minimising $M(R_c^{*\prime})$ but as in the earlier case these optimum values are complicated to use in practice. We therefore consider $\xi_2 = 1$ and $\xi_3 = -1$ which gives the estimator

$$R_c^{*\prime} = r\left(\frac{T_2}{T_3}\right)\left(\frac{\theta_3}{\theta_2}\right) \tag{7.6.13}$$

the mse of which is

$$M(R_c^{*''}) = M(r) + R^2[C_2^2 + C_3^2 + 2(C_{02} - C_{12}) - 2(C_{03} - C_{13}) - 2C_{23}].$$

(7.6.14)

It may be mentioned that $R_c^{*''}$ is the usual double ratio estimator. Because of the reasons pointed out in the remark 7.6.1 we do not compare $R_c^{*'}$ and $R_c^{*''}$ directly for their mse. However, in the empirical study given later both the estimators are considered. In the following we shall compare $R_c^{*''}$ in (7.6.13) with that of $R_1^*$ and $R_2^*$ using $x_2$ and $x_3$ respectively and get the following.

<u>Theorem 7.6.1:</u> For a design D(U,S,P) the estimator $R_c^{*''}$ will be more efficient than $R_1^*$ and $R_2^*$ respectively for estimating the ratio R if the conditions

$$\left(\frac{C_0}{C_3}\right)\rho_{03} - \left(\frac{C_1}{C_3}\right)\rho_{13} + \left(\frac{C_2}{C_3}\right)\rho_{23} > \frac{1}{2} \qquad (7.6.15)$$

and

$$\left(\frac{C_0}{C_2}\right)\rho_{02} - \left(\frac{C_1}{C_2}\right)\rho_{12} - \left(\frac{C_3}{C_3}\right)\rho_{23} < -\frac{1}{2} \qquad (7.6.16)$$

respectively, are satisfied.

Proof is omitted.

There are various situations in which the above conditions are satisfied. Broadly speaking, if $\rho_{03}$ is positive and $\rho_{02}$ negative, conditions will hold.

In case, however, good guess values of $\xi_2^*$ and $\xi_3^*$ (defined as $\xi_2^*$ for character $x_3$) are available they may be used in getting $R_e^{*'}$. In that case it is observed that $R_e^{*'}$ is more efficient than $R^*(\xi_2^*)$ if

$$\frac{\xi_2^* \, C_2}{\xi_3^* \, C_3} \, \rho_{23} < \frac{1}{2}$$

and that it is more efficient than $R^*(\xi_3^*)$ if

$$\frac{\xi_3^* \, C_3}{\xi_3^* \, C_2} \, \rho_{23} < \frac{1}{2}$$

where $R^*(\xi_2^*)$ denotes the estimator $R^*$ using $x_2$ with $\xi_2^*$ and $R^*(\xi_3^*)$ denotes the estimator $R^*$ using $x_3$ with $\xi_3^*$. Thus it is expected that $R_e^{*'}$ will be more efficient than $R^*(\xi_2^*)$ and $R^*(\xi_3^*)$ if the magnitude of $\rho_{23}$ is quite small.

Before giving the empirical study, we consider below the estimators of products, which utilise information on both $x_2$ and $x_3$. These estimators, similar to $R_e^*$ and $R_e^{*'}$, are given by

$$P_e^* = w_1' p\left(\frac{T_2}{\theta_2}\right)^{\delta_2} + w_2' p\left(\frac{T_3}{\theta_3}\right)^{\delta_3} \qquad (7.6.17)$$

and

$$P_e^{*'} = p\left(\frac{T_2}{\theta_2}\right)^{\delta_2}\left(\frac{T_3}{\theta_3}\right)^{\delta_3} \qquad (7.6.18)$$

where $w_1'$ and $w_2'$ are weights such that $w_1' + w_2' = 1$ and

$\partial_2$ and $\partial_3$ are constants to be determined as $\partial_2$ in (7.2.8). The approximate bias and mse of these estimators are given by

$$B(P_c^*) = B(p) + P[w_1' \partial_2 (C_{02}+C_{12}) + w_2' \partial_3 (C_{03}+C_{13})$$

$$+ w_1' \frac{\partial_2(\partial_2-1)}{2} C_2^2 + w_2' \frac{\partial_3(\partial_3-1)}{2} C_3^2],$$

$$B(P_c^{*\prime}) = B(p) + P[\partial_2(C_{02}+C_{12}) + \partial_3(C_{03}+C_{13}) + \partial_2\partial_3 C_{23} +$$

$$\frac{\partial_2(\partial_2-1)}{2} C_2^2 + \frac{\partial_3(\partial_3-1)}{2} C_3^2],$$

$$M(P_c^*) = M(p) + P^2 [w_1'^2 \partial_2^2 C_2^2 + w_2'^2 \partial_3^2 C_3^2 + 2w_1' \partial_2 (C_{02}+C_{12})$$

$$+ 2w_2' \partial_3 (C_{03}+C_{13}) + 2w_1' w_2' \partial_2 \partial_3 C_{23}],$$

$$M(P_c^{*\prime}) = M(p) + P^2 [\partial_2^2 C_2^2 + \partial_3^2 C_3^2 + 2\partial_2 (C_{02}+C_{12}) + 2\partial_3 (C_{03}+C_{13})$$

$$+ 2\partial_2 \partial_3 C_{23}].$$

Further the optimum weights $w_1'$ and $w_2'$ used in $P_c^*$ are given by

$$w_1' = \frac{\partial_3 C_3^2 - \partial_2 (C_{02}+C_{12}) + \partial_3 (C_{03}+C_{13}) - \partial_2 \partial_3 C_{23}}{C_2^2 + C_3^2 - 2\partial_2 \partial_3 C_{23}}$$

$$= 1 - w_2'.$$

Considering $\partial_2 = 1$ and $\partial_3 = -1$, we get the mse of

$P^{*''}_{c} = p(T_2/T_3)(\theta_3/\theta_2)$ as

$$M(P^{*''}_{c}) = M(p) + P^2[C_2^2 + C_3^2 + 2(C_{o2}+C_{12}) - 2(C_{o3}+C_{13}) - 2C_{23}].$$

Again due to remark 7.6.1 we do not compare $P^{*}_{c}$ with other estimators, however, a comparison of $P^{*''}_{c}$ with those of $P^{*}_{1}$ and $P^{*}_{2}$ yields the following Theorem.

Theorem 7.6.2: For any design D(U,S,P), the estimator $P^{*'}$, in (7.6.16) with $\partial_2 = -1$, $\partial_3 = 1$), will be more efficient than $P^{*}_{1}$ and $P^{*}_{2}$ for estimating the product P if the conditions

$$(\frac{C_o}{C_3})\rho_{o3} + (\frac{C_1}{C_3})\rho_{13} + (\frac{C_2}{C_3})\rho_{23} > \frac{1}{2}$$

and

$$(\frac{C_o}{C_2})\rho_{o2} + (\frac{C_1}{C_2})\rho_{12} - (\frac{C_3}{C_2})\rho_{23} < -\frac{1}{2}$$

respectively, are satisfied.

Proof is straight forward and hence it is omitted.

An Empirical Study

Here again we refer to the same data considered in Section 7.4 and consider the same scheme of sample selection which has been used to compare the efficiencies of different estimators (shown in Table 7.4.1) using single supplementary character $x_2$ or $x_3$. The four estimators $R^{*}_{c}$, $R^{*}_{c''}$, $R^{*'}_{c}$ and $R^{*''}_{c}$ using both $x_2$ and $x_3$, developed in Section 7.6 will be compared with that of $r$. Using the values of $\xi_2$ and $\xi_3$ given by

$$\xi_2 = 1.178 \quad \text{and} \quad \xi_3 = -0.737,$$

obtained earlier, we get the optimum weights $w_1$ and $w_2$, for the estimator $R^*_e$, from (7.6.6) as

$$w_1 = 0.2137 \quad \text{and} \quad w_2 = 0.7863.$$

Similarly the optimum weights for the estimator $R^*_{e'}$ (with $\xi_2 = 1$ and $\xi_3 = -1$), are

$$w_1 = 0.3506 \quad \text{and} \quad w_2 = 0.6494.$$

Thus using these values of $\xi_i$'s and $w_i$'s, we get the following table giving the efficiency of different estimators.

Table 7.6.1: Showing efficiency of different estimators using both $(X_2)$ and $(X_3)$.

| Estimator | $MSE/R_0^2$ | % Relative efficiency |
|---|---|---|
| $r$ | 0.5318 | 100 |
| $R^*_e$ | 0.2021 | 263 |
| $R^*_{e'}$ | 0.1812 | 293 |
| $R^{**}_e$ | 0.1390 | 382 |
| $R^*_{e''}$ | 0.1800 | 296 |

This study reveals that $R_c^{*'}$ which does not use $w_i$'s is most efficient. However, in some situations $R_c^*$ is expected to fare better. But the estimator $R_c^{*'}$ which does not use either $w_i$ or $\xi_i$ seems to be quite usable in practice and its construction is also simplest as it requires only a rough knowledg of $\theta_{o2}$ and $\theta_{o3}$.

In the following for the sake of completeness, we consider the estimators $R_{kc}^*$ and $R_{kc}^{*'}$ as an extension $R_c^*$ and $R_c^{*'}$, using k suitable supplementary characters. We define

$$R_{kc}^* = \sum_{i=2}^{k+1} w_i \, r\left(\frac{T_i}{\theta_i}\right)^{\xi_i} \qquad (7.6.19)$$

where $\sum_{i=2}^{k+1} w_i = 1$, $w_i$'s being weights to be obtained by minimising $M(R_{kc}^*)$ and

$$R_{kc}^{*'} = r \prod_{i=2}^{k+1} \left(\frac{T_i}{\theta_i}\right)^{\xi_i} \qquad (7.6.20)$$

where $\xi_i$'s are constants (as in 7.2.7). Similar estimator for estimating P will be given by

$$P_{kc}^* = \sum_{i=2}^{k+1} w_i' \, p\left(\frac{T_i}{\theta_i}\right)^{\partial_i} \quad \text{and} \quad P_{kc}^{*'} = p \prod_{2}^{k+1} \left(\frac{T_i}{\theta_i}\right)^{\partial_i} \qquad (7.6.21)$$

where $\sum_{i=2}^{k+1} w_i' = 1$ and $\partial_i$'s similar to $\xi_i$'s.

For obtaining the mse we consider the case where for $q$ characters which are positively related with $y$'s, $\xi_i = \delta_i = -1$ and for the remaining $k-q$ characters negatively correlated with $y$'s, $\xi_i = \delta_i = +1$, we get

$$M(R^{*}_{ke}) = M(r) + R^2[M_k - 2\sum_{i=1}^{q} C_1 C_i \rho_{1i} + 2\sum_{j=q+1}^{k+1} C_1 C_j \rho_{1j}]$$

and $$M(P^{*\prime}_{ke}) = M(p) + P^2[M_k + 2\sum_{i=1}^{q} C_1 C_i \rho_{1i} - 2\sum_{j=q+1}^{k+1} C_1 C_j \rho_{1j}]$$

where

$$M_k = \sum_{i=2}^{k+1} C_i^2 + 2\sum\sum_{i<i'}^{q} C_i C_{i'} \rho_{ii'} + 2\sum\sum_{j<j'} C_j C_{j'} \rho_{jj'}$$

$$+ 2C_0 \sum_{i=1}^{q} C_i \rho_{0i} - 2C_0 \sum_{j=q+1}^{k+1} C_j \rho_{0j} + 2\sum_i \sum_j C_i C_j \rho_{ij}.$$

Expressions for the general estimators may be obtained in a similar way. Comparison of these estimators is not attempted here. In the next section we proceed to extend these estimators for estimating the parameters themselves.

### 7.7 Estimators of the Parametric Function:

In this section we consider some estimators obtained as combination of well known ratio and product estimators for estimating the parametric function $\theta_0$ itself. For this purpose let us suppose that information on two or more supplementary characters is available in the survey and for simplicity we decide to use any two of them in the estimation procedure. We immediately meet with the problem of choice between the two and how to utilise them as to yield more efficient estimators of $\theta_o$ than those which do not use information on any such character or utilise only one of them. Let $x_1$ and $x_2$ be two such chosen characters. Then we may have the following situation.

(A) $x_1$ and $x_2$ are such that one of them, say $x_1$, is highly positively correlated with $y$ and $x_2$ is negatively correlated such that $x_1$ and $x_2$ can be used in constructing ratio and product estimators respectively.

(B) $x_1$ and $x_2$ both are not highly correlated (positive or negative) with $y$ such that usual unbiased estimator $T_o$ is quite efficient.

(C) Both $x_1$ and $x_2$ are either highly positively or negatively correlated with $y$ in which case we use two variate ratio or product estimator as the case may be.

(D) Only $x_1$ (or $x_2$) could be efficiently used in ratio (or product) method of estimation.

We consider below three estimators $t_1, t_2$ and $t_3$, which are direct extensions of the estimators considered in Section 7.2, and compare them with other known estimators for the situations mentioned in A to D. The estimators considered are

$$t_1 = R_1^* \theta_1 = (\frac{T_0}{T_1} T_2)(\frac{\theta_1}{\theta_2}), \qquad (7.7.1)$$

$$t_2 = R_2^* \theta_1 = (\frac{T_0}{T_1 T_2})\theta_2\theta_1 \qquad (7.7.2)$$

and $$t_3 = R_1^* / \theta_1 = \frac{T_0 T_1 T_2}{\theta_1 \theta_2} \qquad (7.7.3)$$

It may be pointed out here that the parameter $\theta_1$ corresponds to the character $x_1$ which is the supplementary character in the present case. The bias and mse of these estimators may be directly obtained from the corresponding estimators of Section 7.2 by multiplying them with $\theta_1^2$. We give below these expressions and compare them for their mse's $M_1, M_2$ and $M_3$ respectively. We get,

$$M_1 = \theta_0^2(C_0^2 + C_1^2 + C_2^2 + 2C_{02} - 2C_{01} - 2C_{12}) \qquad (7.7.4)$$

$$M^2 = \theta_2^2(C_0^2 + C_1^2 + C_2^2 - 2C_{02} - 2C_{01} + 2C_{12}) \qquad (7.7.5)$$

$$M_3 = \theta_0^2(C_0^2 + C_1^2 + C_2^2 + 2C_{01} + 2C_{02} + 2C_{12}) \qquad (7.7.6)$$

where $C_i = V(T_i)/\theta_i$ and $C_{ij} = C_i C_j \rho_{ij}$.

## 7.8 Comparison of Estimators:

Here we compare unbiased, ratio, product and multivariate ratio and product estimators with the estimators suggested in the previous section under the situations (A) to (D), mentioned therein. We have the following theorems.

Theorem 7.8.1: For any sampling design $D(U,S,P)$, the estimators $t_1$ and $t_2$ under the situation (A) of Section 7.7, will be more efficient than the ratio estimator $T_r$, which utilize information on $x_1$, for estimating $\theta_o$, if the conditions (7.3.1) and (7.3.2) respectively, mentioned in Theorem 7.3.1, are satisfied.

Proof is omitted.

Theorem 7.8.2: For any design $D(U,S,P)$, the estimators $t_1$ and $t_3$, under the situation (A) in Section 7.7, will be more efficient than the product estimator $T_p$ which utilizes information on $x_2$, for estimating $\theta_o$ if the conditions (7.3.3) and (7.3.4) respectively, mentioned in Theorem 7.3.2, are satisfied. We omit the proof.

It may be noted that for

$$C_1 = C_2 = C_o = C \qquad (7.8.1)$$

we get the same configurational representation $F_R$ and $F_P$ as given in Section 7.4. However, in the present situation $\rho_{o1}$, the correlation coefficient between $T_o$ and $T_1$, is supposed to be greater than half whereas in the comparisons of $R_1^*$ and $R_2^*$ with r no such restriction was imposed on $\rho_{o1}$. Due to this restriction, we find that some points in the region of $R_1^*$ (which in the present case will mean the preference region for $t_1$) can not be realised due to restriction of the positivity of the correlation matrix of $\rho_{o1}$, $\rho_{o2}$ and $\rho_{12}$.

Similarly, under the condition (A) the correlation coefficient $\rho_{o2}$ should be less than minus half, whereas no such condition was imposed on $\rho_{o1}$ (which now corresponds to $\rho_{o2}$) for the configuration $F_p$. Owing to this condition we again find that some of the points, specially towards the corner of the region for $P_1^*$ (which correspond to the region of $t_3$) and $P_2^*$ (which correspond to the region of $t_1$) are not realised in practice due to positivity restriction of the correlation matrix.

However, it can easily be shown that there do exist a number of cases where the point located by the pair lies in the preference regions of the proposed estimators satisfying the condition of the positivity of the correlation matrix. For example, consider the following pair of values of $\rho_{o1}$, $\rho_{o2}$ and $\rho_{12}$ in the table below.

Table V.8.1.

| sl.no. | $\rho_{o1}$ | $\rho_{o2}$ | $\rho_{12}$ |
|---|---|---|---|
| 1 | + 0.6 | − 0.8 | − 0.2 |
| 2 | + 0.7 | + 0.2 | − 0.6 |
| 3 | 0.0 | − 0.7 | − 0.7 |
| 4 | − 0.2 | − 0.3 | − 0.7 |
| 5 | + 0.3 | + 0.3 | − 0.7 |
| 6 | + 0.4 | + 0.4 | + 0.6 |

It may be noted that the pair $\rho_{o1}$ and $\rho_{o2}$ in the above table are so chosen that they satisfy the different situations mentioned in Section 7.7. For instance in case I situation (A) is satisfied and in this case $t_1$ is better than both $T_r$ and $T_p$. Similarly, in cases 4,5 and 6 unbiased estimator is preferred but from configuration $F_{rp}$ (to follow) we get more efficient estimators and so on.

Now we consider the situation (B) and compare the proposed estimators with that of the usual unbiased estimator $T_o$ which happens to be more efficient than $T_r$ and $T_p$. Noting that

$$V(T_o) = \theta^2 C_o^2, \qquad (7.8.2)$$

we have the following

<u>Theorem 7.8.3</u>: For any design $D(U,S,P)$, under the situation (B), the proposed estimators $t_1$, $t_2$ and $t_3$ will be more efficient than the usual unbiased estimator $T_o$ if the conditions

$$\left(\frac{C_1^2 + C_2^2}{C_1 C_2}\right) + 2\left(\frac{C_o}{C_1}\right)\rho_{o2} - 2\left(\frac{C_o}{C_2}\right)\rho_{o1} - 2\rho_{12} < 0 \qquad (7.8.3)$$

$$\left(\frac{C_1^2 + C_2^2}{C_1 C_2}\right) - 2\left(\frac{C_o}{C_1}\right)\rho_{o2} - 2\left(\frac{C_o}{C_2}\right)\rho_{o1} + 2\rho_{12} < 0 \qquad (7.8.4)$$

and

$$\left(\frac{C_1}{C_2}\right) + \left(\frac{C_2}{C_1}\right) + 2\left(\frac{C_o}{C_1}\right)\rho_{o2} + 2\left(\frac{C_o}{C_2}\right)\rho_{o1} + 2\rho_{12} < 0 \qquad (7.8.5)$$

are respectively satisfied.

Proof follows by comparing (7.8.2) with (7.7.4) to (7.7.6). Under the condition of equality of the coefficient of variations, we get the respective conditions (7.8.3) to (7.8.5) as

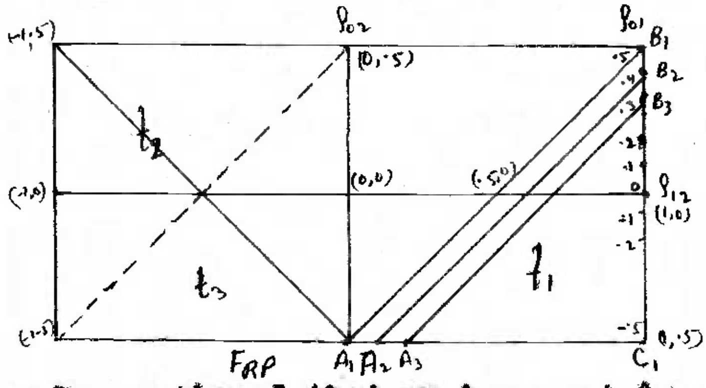
$$(\rho_{o1} - \rho_{o2}) > 1 - \rho_{12} \qquad (7.8.6)$$

$$(\rho_{o1} - \rho_{o2}) > 1 + \rho_{12} \qquad (7.8.7)$$

and

$$(\rho_{o1} + \rho_{o2}) < -(1 + \rho_{12}). \qquad (7.8.8)$$

Assigning different plausible values to $\rho_{o1}$, $\rho_{o2}$ and $\rho_{12}$, with the condition (B), that is, $\rho_{o1}$ and $\rho_{o2}$ lies between -.5 to +.5, we get the following configuration $F_{RP}$.




It may be mentioned that we have not imposed any restriction on $\rho_{12}$. It is clear from above that $t_1$ will be more efficient than $T_e$ in the triangular region $A_1B_1C_1$ for $\rho_{o1} \leq 0.5$ and in the region $A_2B_2C_1$ if $\rho_{o1} \leq 0.4$ and so on. The regions of preference for $t_2$ and $t_3$ are also indicated above.

Next we compare these estimators with that of Olkin's multivariate ratio estimator which has been discussed in the previous chapter, in a general form. The mse of this estimator for uniform optimum weights, under the condition

$$C_1 = C_2 = C \qquad (7.8.9)$$

for $k = 2$ is given by

$$M(T_{r2}) = e_o^2[\frac{C^2}{2}(1 + \rho_{12}) + C_o^2 - C_oC(\rho_{12} + \rho_{o2})]. \qquad (7.8.10)$$

Under the condition (7.8.9), we get from (7.7.4)

$$M_1 = e_o^2[C_o^2 + C^2(2 - \rho_{12}) + 2CC_o(\rho_{o2} - \rho_{o1})]. \qquad (7.8.11)$$

<u>Theorem 7.8.4:</u> For any design $D(U,S,P)$, the estimator $t_1$, $t_2$ and $t_3$ under the condition (7.8.9), will be more efficient than $T_{r2}$ whenever the conditions

$$\frac{\rho_{o1} - 3\rho_{o2}}{3 - 3\rho_{o2}} > \frac{1}{2}(\frac{C}{C_o}) \qquad (7.8.12)$$

$$\frac{\rho_{o1} + \rho_{o2}}{1 + \rho_{12}} > \frac{3}{2}(\frac{C}{C_o}) \qquad (7.8.13)$$

and

$$\frac{\rho_{o1} + \rho_{o2}}{1 + \rho_{12}} < -\frac{1}{2}(\frac{C}{C_o}) \qquad (7.8.14)$$

respectively, are satisfied.

Lastly, we compare these estimators with the multivariate product estimator $T_{pk}$ proposed in the previous chapter. The mse of $T_{pk}$, for $k = 2$ with uniform weight under (7.8.9), is

$$M(T_{p2}) = e_0^2\left[\frac{C^2}{2}(1 + \rho_{12}) + C_0^2 + CC_0(\rho_{01} + \rho_{02})\right]. \quad (7.8.15)$$

<u>Theorem 7.8.5:</u> For any design $D(U,S,P)$ the estimators $t_1$, $t_2$ and $t_3$, under the condition (7.8.9), will be more efficient than $T_{p2}$ whenever the conditions

$$\frac{3\rho_{01} - \rho_{02}}{3 - 5\rho_{12}} > \frac{1}{2}\left(\frac{C}{C_0}\right) \quad (7.8.16)$$

$$\frac{\rho_{01} + \rho_{02}}{1 + \rho_{12}} > \frac{1}{2}\left(\frac{C}{C_0}\right) \quad (7.8.17)$$

and

$$\frac{\rho_{01} + \rho_{02}}{1 + \rho_{12}} < -\frac{3}{2}\left(\frac{C}{C_0}\right) \quad (7.8.18)$$

respectively, are satisfied.

Proofs of the theorems easily follow by comparing the mse of the estimators under the condition (7.8.1).

<u>Remark 7.8.1:</u> It may be noted, however, that the estimator $t_{r2}$ has been suggested for the situation where both $x_1$ and $x_2$ are highly positively correlated and $t_{p2}$ for the situation where both $x_1$ and $x_2$ are highly negatively correlated hence

it is $t_2$ which is comparable with $t_{r2}$ and $t_3$ with that $t_{p2}$ since they satisfy the situation mentioned in (c).

It is easily verified that these sets of conditions, mentioned in Theorem 7.8.4 and 7.8.5, are met with in a variety of situations. Further $t_1$ will in general be more often preferred than any other estimator considered here, whenever, one of the character is positively correlated and the other is negatively correlated with $y$. However, for this situation, it is worthwhile to consider another estimator based on the estimator $R^{*'}$ suggested for estimating the ratio $R$ in the previous chapter. The estimator in this case will be of the form

$$t_1^* = w_1 r \theta_1 + w_2 p \theta_2 \qquad (7.8.19)$$

where $w_1$ and $w_2$ are weights such that

$$w_1 + w_2 = 1 \text{ and } r = (T_o / T_1), \; p = (T_o T_2).$$

This estimator like the earlier ones is biased. Bias and mse are easily seen to be (from $R_2^*$)

$$B(t_1^*) = w_1 B(r) + w_2 B(p)$$

where $B(r)$ and $B(p)$ are bias of ratio and product estimators $T_r$ and $T_p$ respectively.

$$M(t_1^*) = \theta^2[C_0^2 + w_1^2 C_1^2 + w_2^2 C_2^2 - 2w_1 \rho_{01} C_0 C_1 + 2w_2 \rho_{02} C_0 C_2$$

$$- 2w_1 w_2 \rho_{12} C_1 C_2]. \qquad (7.8.20)$$

The optimum weights $w_1$ and $w_2$ are

$$w_1 = \frac{C_2^2 + \rho_{01} C_0 C_1 + \rho_{02} C_0 C_2 + \rho_{12} C_1 C_2}{C_1^2 + C_2^2 + 2\rho_{12} C_1 C_2} = 1 - w_2. \qquad (7.8.21)$$

Next, comparing $t_1$ with $t_1^*$, we find $t_1$ will be more efficient, if

$$\frac{\rho_{01} - \rho_{02}}{1 - \rho_{12}} > + \frac{1}{2}\left(\frac{C}{C_0}\right) \qquad (7.8.22)$$

It is observed that this condition is encountered in many cases. For instance, under the condition (A), $\rho_{01} - \rho_{02}$ will always be positive and greater than unity making the rhs greater than half which is the requirement whenever $C = C_0$. Hence $t_1$ is to be preferred to $t_1^*$ in such cases from the view point of its mse as well as its simplicity.

# CHAPTER VIII

## SOME GENERALISED RATIO ESTIMATORS

8.0 Summary: In this chapter we consider some estimators which are generalisations of the usual ratio estimator. Srivastava (1967) considered an estimator $\bar{y}'_r = \bar{y}(\bar{X}/\bar{x})^{\alpha}$ and Rao (1968a) has suggested $\bar{y}''_r = \alpha\bar{y}_r + (1-\alpha)\bar{y}$ as an alternative to $\bar{y}'_r$ where $\alpha$ is some constant to be suitably chosen. These generalised estimators are biased and have mean square error equal to that of the usual regression estimator for optimum value of $\alpha$. We suggest in section 8.2 an estimator $\bar{y}^*_r = (\bar{y}/t_1)\bar{X}$, where $t_1 = \alpha\bar{x} + (1-\alpha)\bar{X}$, which is almost unbiased and has same mean square error for large samples. In section 8.3 two types of extensions of $\bar{y}^*_r$ are suggested and comparisons are made with other known estimators which use the same amount of information.

### 8.1 Introduction:

Let $(Y_i, X_i)$ denote the value of the $i^{th}$ unit $(i = 1,2,...N)$ in the population consisting of $N$ distinct units for the character under study (y) and the supplementary character (x) respectively. We shall consider in this chapter, for the sake of simplicity, estimation of the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ on the basis of a simple random sample of $n$ units selected without replacement, and assume that $X_i$'s are known constants. Let $\bar{y}$ and $\bar{x}$ denote the usual unbiased estimators (the sample means) of the population mean $\bar{Y}$ and $\bar{X}$ respectively where $\bar{X}$ is the population mean of $x$ and is known. The information on $x$ may now be used in getting the ratio estimator $\bar{y}_r = (\bar{y}/\bar{x})\bar{X}$ or the product $\bar{y}_p = (\bar{y}\,\bar{x})/\bar{X}$ on the basis of preference regions given in an earlier chapter. Srivastava (1967) has suggested an estimator, which is a generalisation of $\bar{y}_r$, given by

$$\bar{y}_r' = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha} \qquad (8.1.1)$$

where $\alpha$ is a constant to be suitably chosen. Evidently, for $\alpha = +1$ and $-1$ the estimator $\bar{y}_r'$ is same as $\bar{y}_r$ and $\bar{y}_p$ respectively. The optimum value of $\alpha$ obtained by minimising the mean square of $\bar{y}_r'$, to order $n^{-1}$, is given by

$$\alpha = \frac{C_o}{C_1}\,\rho_{01} \qquad (8.1.2)$$

where $C_o = \frac{\sigma_o}{\bar{Y}}$ and $C_1 = \frac{\sigma_1}{\bar{X}}$

are the coefficient of variations for $y$ and $x$ respectively, $\sigma_o^2$ and $\sigma_1^2$ are the variances of $y$ and $x$ and $\rho_{01}$ is the correlation coefficient between $y$ and $x$.

The estimator $\bar{y}_r'$, like usual ratio estimator, is biased and has the mean square, to order $n^{-1}$, equal to that of usual regression estimator when optimum value of $\alpha$ is used. Hence in case a good guess value of $\alpha$ is available from some previous census or surveys, the use of estimator $\bar{y}_r'$ will improve over $\bar{y}_r$ (or $\bar{y}_p$).

It may be noted here that the computation of $\bar{y}_r'$, in general, will be complicated in the sense that it will require computation of log and anti-log of the estimators. Rao (1968) has, however, suggested an estimator, as an alternative to $\bar{y}_r'$, which is given by

$$\bar{y}_r'' = \alpha\, \bar{y}_r + (1 - \alpha)\, \bar{y}. \qquad (8.1.3)$$

This estimator being a linear combination of the usual unbiased estimator and the ratio estimator, is quite simple to compute and utilise exactly the same information as that of $\bar{y}_r'$. The estimator $\bar{y}_r''$, like $\bar{y}_r'$, is biased and has the same mean square error to order $n^{-1}$, as that of $\bar{y}_r'$ or the usual regression estimator for the optimum value of $\alpha$. The optimum value of $\alpha$ obtained by minimising the mse of $\bar{y}_r''$, to order $n^{-1}$, is same as given in (8.1.2).

In the following section we propose an estimator, as an alternative to $\bar{y}_r'$ (or $\bar{y}_r''$), which is unbiased, to order $n^{-1}$, and has the same mean square error. Extensions of this estimator to utilise information on several characters have been considered in section 8.3. Some comparisons with other estimators are also made in this section. Results of this chapter are based on a paper (Singh 1968d) by the author in collaboration with K. B. Pathak.

8.2 The Estimator, its Bias and Mean Square Error:

The proposed estimator is given by

$$\bar{y}_R = \left(\frac{\bar{y}}{t_1}\right) \bar{X} \qquad (8.2.1)$$

where

$$t_1 = \alpha \bar{x} + (1 - \alpha) \bar{X} \qquad (8.2.2)$$

and $\alpha$ is some constant to be suitably chosen. The estimator $\bar{y}_r^*$ may also be expressed as

$$\bar{y}_r^* = \bar{y} (1 + \alpha e_1)^{-1} \qquad (8.2.3)$$

where

$$e_1 = (\frac{\bar{x} - \bar{X}}{\bar{X}}) \text{ and } E(e_1) = 0.$$

Further, writing $e_o = (\frac{\bar{y} - \bar{Y}}{\bar{Y}})$, such that $E(e_o) = 0$, and assuming that $|\alpha e_1| < 1$ for large sample sizes, the bias and mean square error of $\bar{y}_r^*$, to order $n^{-1}$, are respectively given by

$$\begin{aligned} B(\bar{y}_r^*) &= E[\bar{y}(1 + \alpha e_1)^{-1} - \bar{Y}] \\ &= \bar{Y} E[(1 + e_o)(1 - \alpha e_1 + \alpha^2 e_1^2 - \ldots) - 1] \\ &= \bar{Y} [(e_o - \alpha e_1) + (\alpha^2 e_1^2 - \alpha e_o e_1) + \ldots ] \\ &\doteq \frac{f\bar{Y}}{n}( \alpha^2 C_1^2 - \alpha C_o C_1 \rho_{01}) \end{aligned} \qquad (8.2.4)$$

and

$$\begin{aligned} M(\bar{y}_r^*) &= E[\bar{y}(1 + \alpha e_1)^{-1} - \bar{Y}]^2 \\ &= \bar{Y}^2 E[e_o - \alpha e_1) + (\alpha^2 e_1^2 - \alpha e_o e_1) + \ldots ]^2 \\ &\doteq \frac{f\bar{Y}^2}{n}(C_o^2 + \alpha^2 C_1^2 - 2\alpha C_o C_1 \rho_{01}) \end{aligned} \qquad (8.2.5)$$

where

$$f = \frac{N - n}{(N-1)}.$$

The optimum value of $\alpha$ which minimises the mean square error of $\bar{y}_r^*$ can be obtained by differentiating $M(\bar{y}_r^*)$ with respect to $\alpha$ and setting the derivative equal to zero. It is observed that the optimum value of $\alpha$ is same as given in (8.1.2). Now substituting the optimum value of $\alpha$ in the approximate expression for bias and mean square error of $\bar{y}_r^*$, we get

$$B(\bar{y}_r^*) = \frac{f\,\bar{Y}}{n}\,(C_o^2 \rho_{01}^2 - C_o^2 \rho_{01}^2) = 0 \qquad (8.2.5)$$

and

$$M(\bar{y}_r^*) = \frac{f\,\bar{Y}^2\,C_o^2}{n}\,(1 - \rho_{01}^2)$$

$$= \frac{f}{n}\,\sigma_o^2\,(1 - \rho_{01}^2). \qquad (8.2.6)$$

Thus it is observed that the proposed estimator $\bar{y}_r^*$ is unbiased, to order $n^{-1}$, and has the same mean square error as that of the usual regression estimator, when the optimum value of $\alpha$ is used. But the bias of $\bar{y}_r'$, to order $n^{-1}$, is easily seen to be

$$B(\bar{y}_r') = \frac{f}{2n}\,\bar{Y}(C_o C_1 \rho_{01} - C_o^2 \rho_{01}^2). \qquad (8.2.7)$$

<u>Remark 8.3.1</u> The estimator $\bar{y}_r^*$ is unbiased, to order $n^{-1}$, only if the optimum value of $\alpha$ is used in the estimator. In actual practice, however, this optimum value will not be available in most situations and hence this estimator will also be biased to that extent. The bias is expected to be negligible if the guessed value used is very near to the true optimum value of $\alpha$. The estimator $\bar{y}_r^*$ is however preferable to other estimators in practice

since it is quite simple to compute and is expected to be almost unbiased for a good guess of $\alpha$. It may however be mentioned that any of these generalised estimators should be used in practice only if very good guess of $\alpha$ is available as otherwise they may lead to loss in efficiency.

## 8.3 Generalised Multivariate Ratio Estimators:

In this section we shall assume that the information is available on $k$ supplementary characters $x_1, x_2, \ldots, x_k$ some of which are positively correlated with $y$ and others negatively correlated. Let $\bar{x}_i$ denote the usual unbiased estimator of the population mean $\bar{X}_i$ $(i=1,\ldots,k)$ corresponding to the variable $x_i$ and we assume that $\bar{X}_i$ is known. We shall consider the following two estimators as extensions of $\bar{y}_r$, given by

$$\bar{y}_{rk} = \bar{y} \sum_{i=1}^{k} w_i (1 + \alpha_i e_i)^{-1} \tag{8.3.1}$$

and

$$\bar{y}'_{rk} = \bar{y} \prod_{i=1}^{k} (1 + \alpha_i e_i)^{-1} \tag{8.3.2}$$

where $w_i$'s in (8.3.1) are weights, such that $\sum_{i=1}^{k} w_i = 1$, $\alpha_i$'s are constants to be suitably chosen and

$$e_i = \frac{\bar{x}_i - \bar{X}_i}{\bar{X}}, \quad i = 0, 1, 2, \ldots k$$

($i = 0$ stands for the character $y$).

We shall now obtain the bias and mean square errors of these estimators, for large sample size, and compare them with some other estimators known in the literature. Assuming that $|\alpha_i e_i| < 1$ for large samples, the bias of $\bar{y}^*_{rk}$ is given by

$$B(\bar{y}^*_{rk}) = \bar{Y}E\left[\sum_i w_i(1+e_o)(1+\alpha_i e_i)^{-1}-1\right]$$

$$= \bar{Y}\sum_i w_i E[(e_o-\alpha_i e_i)+(\alpha_i^2 e_i^2-\alpha_i e_i e_o)+\ldots]$$

$$\doteq \frac{\bar{Y}f}{n}\sum_i w_i b_i^*$$

$$= \frac{\bar{Y}f}{n} w\, b^{*\prime} \qquad (8.3.4)$$

where $b^{*\prime} = (b_i^*)_{1\times k^*}$ $w = (w_i)_{k\times 1}$ and

$$b_i^* = \alpha_i^2 C_i^2 - \alpha_i C_o C_i \rho_{oi}. \qquad (8.3.5)$$

And the mean square error is

$$M(\bar{y}^*_{rk}) = \bar{Y}^2 E\left[\sum_i w_i(1+e_o)(1+\alpha_i e_i)^{-1}-1\right]^2$$

$$= \bar{Y}^2 E\left[\sum_i w_i(1+e_o-\alpha_i e_i+\ldots)-1\right]^2$$

$$\doteq \bar{Y}^2 E\left[\sum_i w_i(e_o-\alpha_i e_i)\right]^2$$

$$= \bar{Y}^2 E\sum_i\sum_j w_i w_j (e_o-\alpha_i e_i)(e_o-\alpha_j e_j)$$

$$= \frac{\bar{Y}^2 f}{n}\sum_i\sum_j w_i w_j a^*_{\alpha ij}$$

$$= \frac{\bar{Y}^2 f}{n} w\, A^*(\alpha) w' \qquad (8.3.6)$$

where

$$a^{*}_{\alpha ij} = (C_o^2 - \alpha_i C_o C_i \rho_{oi} - \alpha_j C_o C_j \rho_{oj} + \alpha_i \alpha_j C_i C_j \rho_{ij})$$

(8.3.7)

and $A^{*}(\alpha)$ is matrix of order $k \times k$, $A^{*}(\alpha) = (a^{*}_{\alpha ij})$.

Similarly, the bias and mean square error of $\bar{y}^{*\prime}_{rk}$ are given by

$$B(\bar{y}^{*\prime}_{rk}) = \bar{Y}E[(1+e_o)\prod_{i=1}^{k}(1+\alpha_i e_i)^{-1} - 1]$$

$$= \bar{Y}E[(e_o - \sum_i \alpha_i e_i) + \sum_i \alpha_i^2 e_i^2 - \sum_i \alpha_i e_o e_i$$

$$+ \sum_{i<j}\sum \alpha_i \alpha_j e_i e_j + \ldots]$$

$$\doteq \frac{\bar{Y}f}{n}(\sum_i \alpha_i^2 C_i^2 - \sum_i \alpha_i C_o C_i \rho_{oi} + \sum_{i<j}\sum \alpha_i \alpha_j C_i C_j \rho_{ij})$$

$$= \frac{\bar{Y}f}{n}(b^{*\prime} + \sum_{i<j}\sum \alpha_i \alpha_j C_i C_j \rho_{ij}) \qquad (8.3.7)$$

where $b^{*} = (b^{*}_i)_{1k}$ given in (8.3.6) and

$$M(\bar{y}^{*\prime}_{rk}) = \bar{Y}^2 E[(1+e_o)\prod_{i=1}^{k}(1+\alpha_i e_i)^{-1} - 1]^2$$

$$= \bar{Y}^2 E\,[(e_o + \sum_i \alpha_i e_i) + \ldots]^2$$

$$\doteq \frac{\bar{Y}f}{n}[C_o^2 + \sum_i \alpha_i^2 C_i^2 - 2\sum_i \alpha_i C_o C_i \rho_{oi} + \sum_{i \neq j}\sum \alpha_i \alpha_j C_i C_j \rho_{ij}]$$

$$= \frac{\bar{Y}^2 f}{n} \cdot A^{*}(\alpha) \qquad (8.3.8)$$

where $A^*(\alpha)$ is matrix $(a^*_{\alpha ij})_{k \times k}$ and $a^*_{\alpha ij}$ is given in (8.3.6), $\rho_{ij}$ is the correlation coefficient between $x_i$ and $x_j$.

Next, we shall compare the proposed estimators $\bar{y}^{**}_{rk}$ and $\bar{y}^{*'}_{rk}$ within themselves and with some other estimators known in the literature using exactly the same amount of information. In this connection we consider the following two types of estimators given by

$$\bar{y}_{rk} = \bar{y} \sum_{i=1}^{k} w_i \left(\frac{\bar{X}_i}{\bar{x}_i}\right)^{\alpha_i}, \quad \sum_i w_i = 1 \qquad (8.3.9)$$

and

$$\bar{y}'_{rk} = \bar{y} \prod_{i=1}^{k} \left(\frac{\bar{X}_i}{\bar{x}_i}\right)^{\alpha_i} \qquad (8.3.10)$$

where $\alpha_i$'s are constants to be suitably chosen.

It is pertinent to note that the estimator $\bar{y}'_{rk}$ with $\alpha_i = +1$ is the multivariate ratio estimator which was suggested by Olkin (1958) for situations where $\rho_{oi}$ is positive and high for all $i$ and this estimator with $\alpha_i = -1$ is the multivariate product estimator (Singh, 1967d) considered earlier in Chapter VI for the situations where $\rho_{oi}$ is negative and high. The other estimator $\bar{y}'_{rk}$ is the ratio cum product estimator considered in Chapter VII (Singh, 1967) for the situations where $\rho_{oi}$ is positive for some $x_i$'s and negative for others with $\alpha_i = +1$ in the former case and $-1$ in the latter. This estimator may be considered as a generalisation of the usual double ratio estimator given by Keyfitz (see, Yates, 1960).

Noting that $\bar{\bar{y}}_{rk}$ and $\bar{\bar{y}}'_{rk}$ may be expressed as

$$\bar{\bar{y}}_{rk} = \bar{\bar{y}} \sum_{i=1}^{k} w_i(1 + e_i)^{-\alpha_i}$$

and

$$\bar{\bar{y}}'_{rk} = \bar{\bar{y}} \prod_{i=1}^{k} (1 + e_i)^{-\alpha_i}$$

their bias, to order $n^{-1}$, are easily seen to be

$$B(\bar{\bar{y}}_{rk}) = \bar{\bar{Y}} E[\sum_i w_i(1+e_0)(1-\alpha_i e_i + \frac{\alpha_i(\alpha_i+1)}{2} e_i^2 + \ldots) - 1]$$

$$= \bar{\bar{Y}} \sum_i w_i E[(e_0 - \alpha_i e_i) + (\frac{\alpha_i(\alpha_i+1)}{2} e_i^2 - \alpha_i e_i e_0) + \ldots]$$

$$\doteq \frac{\bar{\bar{Y}} f}{n} \sum_i w_i b_i$$

$$= \frac{\bar{\bar{Y}} f}{n} wb' \qquad (8.3.11)$$

where $b' = (b_1, b_2, \ldots b_k)$ and

$$b_i = (\frac{\alpha_i(\alpha_i+1)}{2} C_i^2 - \alpha_i C_0 C_i \rho_{0i}), \qquad (8.3.12)$$

and $$B(\bar{\bar{y}}'_{rk}) = \bar{\bar{Y}} E[(1+e_0)\prod_{i=1}^{k} (1-\alpha_i e_i + \frac{\alpha_i(\alpha_i+1)}{2} e_i^2 \ldots) - 1]$$

$$= \bar{\bar{Y}} E[(e_0 - \sum_i \alpha_i e_i) + \sum_i \frac{\alpha_i(\alpha_i+1)}{2} e_i^2$$

$$- \sum_i \alpha_i e_0 e_i + \sum_{i>j} \alpha_i \alpha_j e_i e_j + \ldots ]$$

$$= \frac{f}{n} \bar{\bar{Y}} (b' + \sum_{i<j} \alpha_i \alpha_j C_i C_j \rho_{ij}). \qquad (8.3.13)$$

As regards the mean square error of these estimators, it is observed that, to order $n^{-1}$,

$$M(\bar{y}'_{rk}) = M(\bar{y}''_{rk})$$

and

$$M(\bar{y}'_{rk}) = M(\bar{y}''^{\prime}_{rk}). \tag{8.3.14}$$

It may be mentioned that Srivastava (1967, 1969) has considered the use of optimum value of $\alpha_i$'s in the estimator $\bar{\bar{y}}_{rk}$ and $\bar{y}'_{rk}$. The optimum values of $\alpha_i$'s are, however, quite complicated to use in practice as they involve many unknown parameters. In case good guess values of the coefficient of variations and correlation coefficients are available from previous census or surveys then a suitable choice of $\alpha_i$, based on $\alpha$ in (8.1.2), is

$$\alpha_i = \left(\frac{C_o}{C_i}\right)\rho_{oi} \tag{8.3.15}$$

Substituting this value of $\alpha_i$ in the expression for the bias of the above estimators, we get

$$B(\bar{y}''_{rk}) = 0 \tag{8.3.16}$$

$$B(\bar{y}''^{\prime}_{rk}) = \frac{f}{n}\bar{Y}\,C_o^2 \sum_{i<j} \rho_{oi}\rho_{oj}\rho_{ij} \tag{8.3.17}$$

$$B(\bar{\bar{y}}_{rk}) = \frac{f}{kn}\bar{\bar{Y}}\,C_o \sum_i W_i\rho_{oi}(C_i - C_o\rho_{oi}) \tag{8.3.18}$$

$$B(\bar{y}'_{rk}) = B(\bar{y}''^{\prime}_{rk}) + \frac{f}{kn}\bar{\bar{Y}}C_o \sum_i \rho_{oi}(C_i - C_o\rho_{oi}) \tag{8.3.19}$$

Remark 8.3.1. The estimator $\bar{\bar{y}}^{*\prime}_{rk}$ and $\bar{\bar{y}}'_{rk}$ utilise the same amount of information i.e. knowledge of $\alpha_i$'s and $\bar{X}_i$'s where as $\bar{\bar{y}}^{*}_{rk}$ and $\bar{\bar{y}}_{rk}$ require in addition the optimum weights $w_i$'s and hence they can be used efficiently only if $w_i$'s are known beforehand. As regards mse of these estimators there is no choice between $\bar{\bar{y}}^{*}_{rk}$ and $\bar{\bar{y}}_{rk}$ and between $\bar{\bar{y}}^{*\prime}_{rk}$ and $\bar{\bar{y}}'_{rk}$ on the basis of the approximation considered.

Comparison between $\bar{\bar{y}}^{*}_{rk}$ and $\bar{\bar{y}}_{rk}$ shows that the former is an obvious choice over the latter for the reasons that (i) $\bar{\bar{y}}^{*}_{rk}$, the suggested estimator is the ratio of linear estimators of $\bar{\bar{Y}}$ and $\bar{\bar{X}}$ and hence quite simple to compute whereas $\bar{\bar{y}}_{rk}$ would in general be complicated since it will require calculation of log and antilog of the estimate, (ii) $\bar{\bar{y}}^{*}_{rk}$ is unbiased, to order $n^{-1}$, when $\alpha_i$ in (8.3.15) is used where $\bar{\bar{y}}_{rk}$ is biased.

Comparing $\bar{\bar{y}}^{*\prime}_{rk}$ with $\bar{\bar{y}}'_{rk}$ it is seen from (8.2.21) that both are equally preferable as regards their mse. However, the suggested estimator $\bar{\bar{y}}^{*\prime}_{rk}$ is much simpler to compute than that of $\bar{\bar{y}}'_{rk}$ which again requires calculation of log and antilog of the estimate. And in fact this will be the simplest of all the four estimators in general as it does not use $w_i$'s. Further, it is easily observed from (8.3.17) and (8.3.19) that the bias of $\bar{\bar{y}}^{*\prime}_{rk}$ would be quite small for large sample size and it will be smaller than the bias of $\bar{\bar{y}}'_{rk}$ for the situations where the second term in rhs of (8.3.19) is positive ($\rho_{ij}$ being greater than zero), That is, when $\alpha_i < 1$ for almost all $i$ ($i = 1,2,\ldots, k$)

the bias of $\bar{y}^{*\prime}_{rk}$ will be less than the bias of $\bar{y}^{\prime}_{rk}$. Assuming that this condition is satisfied then the choice rests between the two estimators $\bar{y}^{*}_{rk}$ and $\bar{y}^{*\prime}_{rk}$ suggested in this section. As pointed out earlier $\bar{y}^{*\prime}_{rk}$ will be simpler to compute and since it does not require any weights $w_i$'s this estimator is an obvious choice when $w_i$'s are not available. Since in such cases determination of $w_i$'s from the sample will be complicated and that the estimator may not retain its efficiency with these estimated weights. In case however $w_i$'s are available which is very rarely the situation $\bar{y}^{*}_{rk}$ may be a better choice since it is almost unbiased estimator when $\alpha$ in (8.3.15) is used.

## CHAPTER IX

## ON A METHOD OF USING AN APRIORI VALUE OF THE PARAMETER IN THE ESTIMATION PROCEDURE

9.0 Summary: In this chapter an estimation procedure is suggested which utilises the knowledge of an apriori value of the population parameter $\theta$. The apriori value may be available to the experimentor from previous census or surveys or even expert guesses. The proposed estimator $t_o$ is essentially a weighted average of $t$, the usual unbiased estimator of $\theta$, and the apriori value $\theta_o$ of $\theta$. That is, $t_o = kt + (1-k)\theta_o$, where $k$ is some constant to be suitably choosen. The optimum value of $k$ which minimises the mse of $t_o$ is found to be $k_o = \delta^2/(\delta^2 + e^2)$, where $|\delta| = |1 - \theta_o/\theta|$ and $e$ is the relative standard error of $t$. In many cases $e$ may be known in practice, especially when the survey has been planned to achieve a specified precision but $\delta$ is always unknown. For such situations, using $\delta_1$ as an apriori value of $\delta$ an approximately optimum estimator $t_{01}$ is obtained where $k_{01} = \delta_1^2/(\delta_1^2 + e^2)$. $t_{01}$ is then compared with $t$ for estimating $\theta$ and a table showing the efficiency of $t_{01}$ as compared to $t$ has been given for various values of $\delta$, $e$ and $\delta_1$ in section 9.3. The situation where approximate

values of both $\theta$ and $\theta$ are used has also been discussed. Some special cases of $t_{\theta}$ have been mentioned & the situation where $t$ is biased estimator of $\theta$ has been discussed briefly. Lastly, an estimator of $\theta$ which utilise information on a supplementary character besides $\theta_0$ is also suggested.

9.1 Introduction:

Several methods by which supplementary information may be utilised have already been mentioned in an earlier chapter and it is well-known in this connection that the use of supplementary information in a suitable manner generally improves the estimator of the population parameters. The usual techniques of using the supplementary information assume that values of one or more supplementary character related to the character of interest are known for each unit of the population. A variant of the procedure is that of two-phase sampling discussed earlier where the data on supplementary character may be collected from a large first-phase sample, if found economical, and utilised in a suitable manner.

In many cases, however, such detailed apriori information may not be available or may be quite costly to collect. On the other hand, some summary information, for instance, an apriori value of the parameter $\theta$, quite close to it may be known to the experimenter. Knowledge of this apriori value may be available from previous census or surveys or even from expert guesses by the specialists in the concerned field. It may also happen that the upper and lower limits of the parameter may be known (Dalenius, 1965 and in that case a simple or modified average depending on the expected skewness of the distribution of $\theta$ may provide a good approximation of $\theta$.

It seems worthwhile, therefore, to develop an estimator which utilises this apriori value of $\theta$ and achieves a mean square error considerably smaller than the variance of the usual unbiased estimator of the parameter. In section 9.2 of this chapter we propose such an estimator. This estimator utilises an apriori value of $\theta$ and an estimator of the correction factor which involves the usual estimate of $\theta$ as well as the relative difference from its apriori value and the relative standard error (rse) of the usual estimator. As the exact value of the relative difference is not likely to be available in practice, the suggested estimator is modified so as to use the approximate value of the relative difference and then the modified estimator is compared with the usual unbiased estimator in section 9.3. A table is also given showing the efficiency of this estimator for different values of approximate relative difference and rse. The value of rse may be known in many cases especially when the survey is planned to achieve a prespecified precision. However, the effect of using an approximate value is also considered. Some special cases of the suggested estimator are then pointed out.

In section 9.4 the case where an unbiased estimator of $\theta$ is not available (such as rates, products etc.) is considered. Lastly a ratio estimator which utilises the apriori value and the detailed information on a supplementary character is briefly discussed. Results of this chapter are based on a paper by the author in collaboration with A.S.Roy.

## 9.2 The Estimator, its Bias and Mean Square Error:

Let $\theta_o$ be an apriori value of the parameter $\theta$, which the statistician believes to be quite close to $\theta$ and let $D$ be the difference between $\theta$ and $\theta_o$, that is

$$D = (\theta - \theta_o).$$

Then we proceed to estimate $\theta$ in two step process: In step 1, an unbiased estimate $d$ of $D$ is obtained on the basis of a random sample $s$ from the population $U$, following a design $P$, and it is weighted by a factor $k$ suitably chosen. In step 2, the weighted estimate 'kd' is applied as a correction factor to improve upon the apriori value $\theta_o$. The proposed estimator of $\theta$ thus becomes

$$t_o = \theta_o + kd \qquad (9.2.1)$$

where $d = (t - \theta_o)$, $t$ being an unbiased estimator of $\theta$. Alternatively, $t_o$ may be looked upon as a weighted estimate of $\theta_o$ and $t$, that is

$$t_o = kt + (1-k)\theta_o. \qquad (9.2.2)$$

The factor $k$ may be obtained by minimising the mean square error of $t_o$, which will give its optimum value. It may also be chosen to satisfy certain other criteria. For instance $k$ may be taken as a random variable, instead of a fixed constant, and if its expectation is unity then the estimator $t_o$ will remain unbiased for $\theta$.

The estimator $t_o$, in general, is biased. We have

$$E(t_o) = k\theta + (1-k)\theta_o$$

hence the bias of $t_o$ is

$$B(t_o) = E(t_o) - \theta$$

$$= (k-1)(\theta - \theta_o) = (k-1)D. \qquad (9.2.3)$$

The sampling variance of $t_o$ is given by

$$V(t_o) = E(t_o^2) - E^2(t_o)$$

$$= k^2 V(t),$$

where

$$V(t) = E(t^2) - E^2(t)$$

is the variance of the unbiased estimator $t$.

The mean square error (mse) of $t_o$ thus becomes

$$M(t_o) = V(t_o) + B^2(t_o)$$

$$= k^2 V(t) + (k-1)^2 D^2. \qquad (9.2.4)$$

The optimum value of $k$ which minimises this mse can be obtained by differentiating $M(t_o)$ with respect to $k$ and setting the derivative equal to zero. This gives

$$k_o = \frac{D^2}{D^2 + V(t)}. \qquad (9.2.5)$$

Suppose $\delta = \left(\frac{\theta - \theta_0}{\theta}\right)$ is the relative difference between $\theta$ and $\theta_0$, and $e(t) = \sqrt{V}(t)/\theta$, denoted by simply $e$, is the relative standard error (rse) of the estimator $t$. Then the optimum weight $k_0$ can be expressed as

$$k_0 = \frac{\delta^2}{\delta^2 + e^2} . \tag{9.2.6}$$

Since $e^2 \geq 0$, we have $0 \leq k_0 \leq 1$. Now putting this value of $k_0$, the minimum value of $M(t_0)$ is seen to be

$$M_0(t) = \left(\frac{\delta^2}{\delta^2 + e^2}\right)^2 V(t) + \left(\frac{e^2}{\delta^2 + e^2}\right)^2 D^2$$

$$= \frac{1}{(\delta^2 + e^2)^2} (\delta^4 V(t) + e^4 D^2)$$

$$= \frac{\delta^2(\delta^2 + e^2)}{(\delta^2 + e^2)^2} V(t)$$

$$= k_0 V(t), \tag{9.2.7}$$

since $V(t) = \theta^2 e^2$ and $D^2 = \delta^2 \theta^2$.

The relative efficiency of $t_0$ as compared to the usual unbiased estimator $t$ is given by

$$\text{Eff.} (t_0) = \frac{V(t)}{M_0(t_0)}$$

$$= \frac{1}{k_0} . \tag{9.2.8}$$

Evidently $t_o$ will be more efficient than $t$ since $k_o$ is less than unity. But the difficulty in construction of the estimator $t_o$, using optimum weight $k_o$, is that $k_o$ involves the parameter $\theta$ itself and hence in practice we can use only the approximate optimum weight. Next section is devoted to use of such approximate weight.

## 9.3. Use of Approximate Optimum Weight :

The optimum value $k_o$ requires an exact knowledge of $c$ and $\delta$. Of these two quantities, the value of $c$ may be known in many cases especially when the survey is planned to achieve a prespecified degree of precision. The value of $c$ is needed also at other stages of sample selection, for instance in determination of the sample size etc. But the exact value of $\delta$ is always unknown in practice. Hence we can obtain only an approximate value of the optimum weight $k_o$, using some idea about the magnitude of $\delta$, and also of $c$, if it is known. In this connection we discuss below two possible cases viz.,

Case 1: An approximate value of $\delta_1$ is used in place of the exact value $\delta$ and exact value of $c$ is known.

Case 2: Approximate values $\delta_1$ and $c_1$ are used in place of the unknown values $\delta$ and $c$.

Case 1: The proposed estimator in this case becomes

$$t_{01} = k_{01}t + (1 - k_{01})\theta_o \qquad (9.3.1)$$

where

$$k_{01} = \frac{\delta_1^2}{\delta_1^2 + e^2} \cdot \qquad (9.3.2)$$

The mse of $t_{01}$ may be obtained by substituting the value of $k_{01}$ for $k$ in the expression for $M(t)$ in (9.2.4). This gives

$$\frac{M(t_{01})}{e^2} = \left(\frac{\delta_1^2}{\delta_1^2 + e^2}\right)^2 e^2 + \left(\frac{e^2}{\delta_1^2 + e^2}\right)^2 \delta^2$$

$$= \frac{e^2\ \delta_1^2}{(\delta_1^2 + e^2)^2} \left(\delta_1^2 + \frac{e^2\ \delta^2}{\delta_1^2}\right)$$

$$= \frac{e^2\ \delta_1^2(\delta_1^2 + e'^2)}{(\delta_1^2 + e^2)^2} \qquad (9.3.3)$$

where $e' = e\delta / \delta_1$.

Now $t_{01}$ will be more efficient than $t_o$ if

$$\frac{M(t_{01})}{e^2} < e^2,$$

that is, if

$$\delta_1^2\ (\delta_1^2 + e'^2) < (\delta_1^2 + e^2)^2$$

or $$(\delta^2 - e^2) < 2\ \delta_1^2$$

or $$\delta_1^2 > \frac{(\delta^2 - e^2)}{2} \ . \qquad (9.3.4)$$

Thus if

$$e^2 \geq \partial^2$$

which implies

$$\frac{V(e)}{D^2} \geq 1,$$

for any value of $|\partial_1| \neq 0$, the condition (9.3.4) is satisfied and in which case the estimator $t_{01}$ becomes more efficient than $t$. Further, since $e \geq 0$ always, a slightly stronger condition for $t_{01}$ to be more efficient than $t$ may be written as

$$|\partial_1| > \frac{|\partial|}{\sqrt{2}} . \qquad (9.3.5)$$

Thus even if $\partial_1$ differs from $\partial$, $t_{01}$ will be more efficient than $t$ as long as (9.3.4) is satisfied; however, too much departure of $\partial_1$ from $\partial$ will reduce the gain in efficiency of the estimator. Evidently a small value of $|\partial_1|$ which violates (9.3.4) will make $t_{01}$ less efficient than $t_0$. The expression for the efficiency of $t_{01}$ with respect to $t$ is given by

$$\text{Eff}\ (t_{01}) = \frac{(\partial_1^2 + e^2)}{\partial_1^2} \cdot \frac{(\partial_1^2 + e^2)}{(\partial_1^2 + e'^2)} \qquad (9.3.6)$$

which is greater than

$$\frac{\partial_1^2 + e^2}{\partial_1^2} = \frac{1}{K_{01}}$$

if $e > e'$, that is, if $|\partial_1| > |\partial|$.

Thus it is interesting to note that in case $|\delta_1| > |\delta|$, the efficiency of $t_{01}$ is more than what it would have been if $|\delta_1|$ were the true value and the optimum value $k_{01}$ in that case, were used. The values of Eff $(t_{01})$ have been given in the tables in the end of this chapter.

Case 2: In this case the proposed estimator becomes

$$t_{02} = k_{02}t + (1 - k_{02})\,\theta_o$$

where

$$k_{02} = \frac{\delta_1^2}{\delta_1^2 + c_1^2}\,.$$

As in (9.3.3), we get

$$\frac{M(t_{02})}{c^2} = \left(\frac{\delta_1^2}{\delta_1^2 + c_1^2}\right)^2 c^2 + \left(\frac{c_1^2}{\delta_1^2 + c_1^2}\right)^2 \delta^2$$

$$= \frac{c^2\,\delta_1^2}{(\delta_1^2 + c_1^2)^2}\left(\delta_1^2 + \frac{c_1^4\,\delta^2}{c^2\,\delta_1^2}\right)$$

$$= \frac{c^2\,\delta_1^2(\delta_1^2 + c_1^2\,e''^2)}{(\delta_1^2 + c_1^2)^2}$$

where $e'' = (\frac{c_1}{c})(\frac{\delta}{\delta_1})$.

The condition for $t_{02}$ to be more efficient than $t$ becomes $\frac{M(t_{02})}{\theta^2} < c^2$. That is

$$\delta_1^2 (\delta_1^2 + c_1^2 c''^2) < (\delta_1^2 + c_1^2)^2$$

or $$\delta_1^2 c''^2 - c_1^2 < 2\delta_1^2$$

or $$\delta_1^2 > \left(\frac{\delta^2 - c^2}{2}\right)\left(\frac{c_1}{c}\right)^2.$$

The above condition will be satisfied if $\delta_1^2 > (\delta^2 - c^2)/2$ provided $c_1 \leq c$. Thus when anticipated values of $|\delta|$ and $c$ are to be used in determination of $k$, it would be safer to take a slightly higher value for $|\delta|$ and smaller value for $c$ for calculation of $k_{02}$. It may be mentioned, however, that as the distance between the anticipated values and the true values increases the gain in efficiency of the proposed estimator decreases. Below we give two special cases of $c$ and $\delta_1$.

i) <u>$c$ is zero:</u> This implies that $k_0 = 1$ and that the proposed estimator and the usual unbiased estimator $t$ are identical. It is quite logical, since $c$ zero means $V(t)$ is zero which is the minimum value that an estimator can attain and hence $t = t_0$.

ii) <u>$\delta_1$ is unity:</u> This implies that $(1 - \frac{\theta_0}{\theta}) = 1$ that is $\theta_0 = 0$. In this case, we get

$$k_0 = \frac{1}{1 + e^2} \qquad (9.3.7)$$

and the estimator $t_0$ becomes

$$t_{00} = (1 + e^2)^{-1}\, t. \qquad (9.3.8)$$

The mean square error of $t_{00}$ then becomes

$$M(t_{00}) = \frac{V(t)}{(1 + e^2)} \;. \qquad (9.3.9)$$

The relative efficiency of $t_{00}$ as compared to the usual estimator $t$ is thus given by

$$\text{Eff}\;(t_{00}) = (1 + e^2) \qquad (9.3.10)$$

which is greater than unity. However, the relative efficiency of $t_{00}$ as compared to $t_{01}$ is given by

$$\text{Eff}\;(t_{00}) = \left(\frac{\delta_1^2 + e'^2}{\delta_1^2 + e^2}\right)\left(\frac{\delta_1^2 + e^2\delta_1^2}{\delta_2^2 + e^2}\right). \qquad (9.3.11)$$

~~But~~ The term in second bracket is less than unity since $|\delta_1| < 1$ and a sufficient condition for $t_{00}$ to be less efficient than $t_{01}$ is that $e' < e$ that is $|\delta_1| > |\delta|$.

<u>Remark</u>: Although in the above discussion we have assumed that the value of $k_0$ <u>can not</u> be known exactly in practice, but situations exist where we can get the value of $k_0$ even though

$\partial$ and $c$ may not be known separately. For instance, since $k_0$ can also be expressed as

$$k_0 = \frac{1}{1 + \lambda^2} \tag{9.3.12}$$

where $\lambda = c/\partial$, and in case the value of $\lambda$ is known we can get value of $k_0$. In some cases, however, the upper or lower bound of $\lambda$ may be specified which helps in determining the value of $k_0$.

Simple random sampling: Let us consider the parameter to be estimated is $\bar{Y}$, the population mean and let $\bar{y}$ denote the usual unbiased estimator, the sample mean, when sample is selected with-replacement and with equal probability, then $t_{00}$ in (9.3.8), where $\partial$ is assumed to unity, becomes

$$\bar{y}_{00} = k_0\, \bar{y}$$

$$= \frac{1}{1 + c_y^2/n}\, \bar{y}$$

$$= \frac{n}{n + c_y^2}\, \bar{y} \tag{9.3.13}$$

where $c_y^2$ is the population coefficient of variation. This estimator was suggested by Searls (1964). Efficiency of this estimator is given by the entry corresponding to $|\partial| = |\partial_1| = 1.0$ in tables 9.1.1 – 9.1.4.

The estimator $t_{00}$ for sampling without-replacement with equal probability becomes

$$\bar{y}'_{00} = \frac{n}{n + fc_y^2}\ \bar{y}. \tag{9.3.14}$$

where $f = \frac{N-n}{N-1}$ .

However, when the apriori value $\bar{y}_o$ of $\bar{Y}$ is not taken as zero (i.e. when $\delta_1$ is not unity) the estimators corresponding to $t_{01}$ for the sampling with and without replacement with equal probabilities will be given by

$$\bar{y}_{01} = \bar{y}_0 + \frac{n}{n + \lambda^2} (\bar{y} - \bar{y}_0) \tag{9.3.15}$$

and

$$\bar{y}'_{01} = \bar{y}_0 + \frac{n}{n + f\,\lambda^2} (\bar{y} - \bar{y}_0) \tag{9.3.16}$$

respectively where $\lambda^2 = c_y^2 / \delta_1^2$. Now, whenever an approximate value of $\lambda$ is known these estimators may be utilised efficiently. A lower bound of $\lambda$ which is usually assumed to be the population size $N$ it-self, may also be used in building up these estimates.

## 9.4 Use of Biased Estimator:

So far we have assumed $t$ to be unbiased estimator of $\theta$. However, in many situations, a simple unbiased estimator of $\theta$ may not be available in general. For instance, the parameter $\theta$ may be birth rate, death rate, per capita consumer expenditure, total crop production etc., when the usual estimator of $\theta$ is biased. In such cases the suggested estimator, denoted by

$t_0'$, is of the same form as $t_0$, that is

$$t_0' = \theta_0 + k(t - \theta_0) \qquad (9.4.1)$$

but its bias and mean square error are respectively given by

$$B(t_0') = (k - 1)D + kB \qquad (9.4.2)$$

and $$M(t_0') = k^2 V(t) + (k - 1)D^2 + 2k(k - 1)BD \qquad (9.4.3)$$

where

$$B = E(t) - \theta$$

is the bias in $t$.

Differentiating $M(t_0')$ with respect to $k$ and equating the derivative to zero, we get optimum $k$ as

$$k_0 = \frac{D(B + D)}{V(t) + (B + D)^2} \qquad (9.4.4)$$

which on substitution in $M(t_0')$ gives the optimum mean square error as

$$M_0(t_0') = \frac{D^2 \quad V(t)}{V(t) + (B + D)^2} . \qquad (9.4.5)$$

Thus the proposed estimator will be more efficient than $t$ if $M_0(t_0')$ is less than $M(t) = V(t) + B^2$. That is if

$$V(t)[1 - \frac{B^2}{(B + D)^2 + V(t)}] + B^2 > 0 \qquad (9.4.6)$$

which is always true.

Here again the exact value of $k_0$ will not be known in practice. The efficiency of this estimator using approximate optimum weight may be obtained as in the earlier section. It is expected that with reasonably good approximation of the optimum weight the proposed estimator will be more efficient than $t$ as in unbiased case.

Next, we consider briefly the situation where both the type of information namely, the apriori value $\theta_0$ and detail information on the supplementary character $x$ are available and we are interested in using these informations. Let $\theta_1$ denote the parameter based on the character $x$ then assuming that $\theta_1$ is known, as usual, we suggest a ratio estimator of $\theta$, given by

$$t_{0r} = (\frac{t_0}{t_1})\, \theta_1 \qquad (9.4.7)$$

where $t_0$ is the estimator suggested in (9.2.2) for $\theta$ and $t_1$ is an estimator of $\theta_1$ similar to $t_0$ given by

$$t_1 = \alpha\, t' + (1 - \alpha)\theta_1 \qquad (9.4.8)$$

where $t'$ is the usual unbiased estimator of $\theta_1$ and $\alpha$ is a constant to be suitably chosen. A suitable choice of $\alpha$, as in the chapter VIII, is

$$\alpha = \left(\frac{c_o}{c_1}\right) \rho_{o1} \qquad (9.4.9)$$

where $c_o$ and $c_1$ are the coefficient of variation of $y$ and $x$ and $\rho_{o1}$ is the correlation between them.

If $n$ denote the size of the sample selected from the population then writing $t = \theta(1 + e_o)$ and $t' = \theta(1 + e_1)$ where $E(e_o) = E(e_1) = 0$ and assuming $n$ is large so that $|E\, e_1| < 1$, the bias and mse of $t_{Or}$, to order $n^{-1}$, are given by

$$B_{Or} = B_0 + (1 - k)\theta_o\, \alpha^2 c_1^2 - \alpha k\, \theta\, (c_{o1} - \alpha\, c_1^2) \qquad (9.4.10)$$

and

$$M_{Or} = B_0^2 + (1 - k)^2 \theta_o^2 \alpha^2 c_1^2 - 2k(1 - k)\alpha\theta_o\theta(c_{o1} - \alpha\, c_1^2)$$

$$+ k\theta_o^2(c_o^2 + \alpha^2 c_1^2 - 2\alpha c_o c_1 \rho_{o1}) \qquad (9.4.11)$$

where $B_0$ is the bias of $t_0$.

Substitution of the value of $\alpha$ we get

$$B_{Or} = B_0 + (1 - k)\theta_o c_o^2 \rho_{o1}^2 \qquad (9.4.12)$$

and

$$M_{Or} = B_0^2 + (1 - k)^2 \theta_o^2 c_o^2 \rho_{o1}^2 + k\theta^2 c_o^2 (1 - \rho_{o1}^2) \qquad (9.4.13)$$

It may be mentioned that the estimator $\bar{y}_r^*$ suggested in the previous chapter turns out to be a special case of $t_{0r}$ for $k = 1$, since in that case

$$t_{0r} = t_r^* = \left(\frac{\bar{y}}{t_1}\right) \theta_1 \qquad (9.4.14)$$

Thus for $\theta_1 = \bar{Y}$, $t_r^*$ is same as $\bar{y}_r^*$ in (8.2.1).

Table (1): Showing the efficiency of $t_{01}$ as compared to $t$ for different values of $\partial$, $\partial_1$ and $c$. (all expressed in percentage).

Table (1.1): c = 15 per cent

| $\partial_1$ \ $\partial$ | 100 | 50 | 20 | 15 | 10 | 5 |
|---|---|---|---|---|---|---|
| (0) | (1) | (2) | (3) | (4) | (5) | (6) |
| 100 | 102.2 | 104.0 | 104.4 | 104.5 | 104.5 | 104.5 |
| 50 | 87.4 | 109.0 | 117.1 | 117.8 | 118.4 | 118.7 |
| 20 | 16.2 | 54.1 | 156.2 | 185.4 | 214.0 | 235.8 |
| 15 | 8.8 | 33.0 | 144.0 | 200.0 | 276.9 | 360.0 |
| 10 | 4.7 | 18.4 | 105.6 | 174.2 | 325.0 | 676.0 |
| 5 | 2.8 | 11.1 | 69.0 | 122.0 | 270.3 | 1000.0 |

Table (1.2): c = 12 per cent

| $\partial_1$ \ $\partial$ | 100 | 50 | 20 | 15 | 10 | 5 |
|---|---|---|---|---|---|---|
| (0) | (1) | (2) | (3) | (4) | (5) | (6) |
| 100 | 101.5 | 102.5 | 103.0 | 103.1 | 103.1 | 103.2 |
| 50 | 89.3 | 105.8 | 110.7 | 112.1 | 112.1 | 113.0 |
| 20 | 18.4 | 56.8 | 136.1 | 158.6 | 169.8 | 184.9 |
| 15 | 9.2 | 34.1 | 126.6 | 186.5 | 209.2 | 262.1 |
| 10 | 4.2 | 17.2 | 102.2 | 151.2 | 244.0 | 437.8 |
| 5 | 2.6 | 8.4 | 48.1 | 108.6 | 190.3 | 676.0 |

Table (1.3): e = 10 per cent

| $\theta$ / $\theta_1$ | 100 | 50 | 20 | 15 | 10 | 5 |
|---|---|---|---|---|---|---|
| (0) | (1) | (2) | (3) | (4) | (5) | (6) |
| 100 | 101.0 | 101.8 | 102.0 | 102.0 | 102.0 | 102.0 |
| 50 | 93.2 | 104.0 | 107.5 | 107.8 | 108.0 | 108.1 |
| 20 | 21.6 | 61.0 | 125.0 | 137.0 | 147.0 | 153.8 |
| 15 | 10.0 | 35.2 | 116.6 | 144.4 | 174.2 | 198.8 |
| 10 | 4.0 | 15.4 | 80.0 | 123.1 | 200.0 | 320.0 |
| 5 | 1.6 | 6.2 | 38.5 | 67.6 | 147.0 | 500.0 |

Table (1.4): e = 5 per cent

| $\theta$ / $\theta_1$ | 100 | 50 | 20 | 15 | 10 | 5 |
|---|---|---|---|---|---|---|
| (0) | (1) | (2) | (3) | (4) | (5) | (6) |
| 100 | 100.2 | 100.4 | 100.5 | 100.5 | 100.5 | 100.5 |
| 50 | 98.1 | 101.0 | 101.8 | 101.9 | 102.0 | 102.0 |
| 20 | 44.1 | 81.2 | 106.2 | 109.7 | 112.2 | 112.5 |
| 15 | 20.8 | 55.1 | 103.7 | 111.1 | 117.6 | 121.9 |
| 10 | 6.0 | 21.6 | 78.1 | 100.0 | 125.0 | 147.0 |
| 5 | 1.0 | 4.0 | 23.5 | 40.0 | 80.0 | 200.0 |

# R E F E R E N C E S

Basu, D. (1958): On sampling with and without replacement; Sankhya, 20, 287-294.

Brewer, K.R.W. (1963): A model of systematic sampling with unequal probabilities; Aust. Jour. Stat., 5, 5-13.

Burkholder, D.L. (1960): Effect on the minimal complete class of changes in testing problems; Ann. Math. Stat., 31, 323-331.

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 efficiency of systematic and stratified random samples for a certain class of population; Ann. Math. Stat. 17, 164-177.

Cochran, W.G. (1963): Sampling Techniques; 2nd Edn. John Wiley and Sons, New York.

Dalenius, T. (1950): The problem of optimum stratification I; Skand. Akt. 33, 203-213.

Dalenius, T. (1962): Recent advances in sample survey theory and methods; Ann. Math. Stat., 33, 325-349.

Dalenius, T. (1965): Current trends in the developments of sample survey theory and methods; Unpublished.

Das, B.C. (1966): Multivariate investigation of blood chemistry and Morphology; Proc. Sympo. Adop. Phy., Defence Institute of Phisiology and Allied Science, Madras.

Deming, W.E. (1956): On simplification of sampling design through replication with equal probabilities and without stages; Jour. Amer. Stat. Assn. 51, 24-253.

Des Raj (1956): Some estimators in sampling with varying probabilities without replacement; Jour. Amer. Stat. Assn., 51, 269-284.

Des Raj and Khamis, S.H. (1958): Some remarks on sampling with replacement, Ann. Math. Stat., 39, 550-557.

Des Raj (1964a): On double sampling for pps estimation; Ann. Math. Stat., 35, 900-902.

Des Raj (1964b): A note on the variance of the ratio estimate; Jour. Amer. Stat. Assn., 59, 481-486.

Des Raj (1965a): On a method of using multi-auxiliary information in sample surveys; Jour. Amer. Stat. Assn. 60, 270-277.

Des Raj (1965b): Variance estimation in randomised systematic sampling with probability proportionate to size; Jour. Amer. Stat. Assn., 60, 278-284.

Des Raj (1968): Sampling Theory; McGraw-Hill, London.

Durbin, J. (1953): Some results in sampling theory when the units are selected with unequal probabilities; Jour. Roy. Stat. Soc., B, 15, 262-269.

Durbin, J. (1959): A note on the application of Quenouille's method of bias reduction to the estimation of ratio's, Biometrika, 46, 477-480.

Fellegi, I.P. (1963): Sampling with varying probabilities without replacement - rotating and non-rotating samples; Jour. Amer. Stat. Assn., 58, 183-201.

Godambe, V.P. (1955): A unified theory of sampling from finite populations; Jour. Roy. Stat. Soc., B, 17, 269-278.

Godambe, V.P. (1960): An admissible estimate for any sampling design, Sankhya, 22, 285-288.

Godambe, V.P. and Joshi, V.M. (1965): Admissibility and Bayes estimation in sampling finite populations I; Ann. Math. Stat., 36, 1707-1722.

Griffith, A.L. (1946): The efficiency of enumerations; Forest Research Institute, Dehradun, Indian Forest Leaflets, 83-91, 93-96.

Hajek, J. (1958): Some contributions to the theory of probability sampling; Bull. Int. Stat. Inst., 36(3), 127-134.

Hajek, J. (1959): Optimum strategy and other problems in probability sampling; Casipis Pest. Math. 84, 387-423.

Hansen, 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; Vol. I and II, John Wiley and Sons, New York.

Hanurav, T.V. (1962a): An existence theorem in sampling; Sankhyā, A, 24, 327-330.

Hanurav, T.V. (1962b): On Horvitz and Thompson's estimator; Sankhyā, A, 24, 429-436.

Hanurav, T.V. (1965): Optimum 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. (1967): Optimum utilisation of auxiliary information - π PS sampling of two units from a stratum; Jour. Roy. Stat., Soc., B, 29

Hanurav, T.V. (1968): Hyper-admissibility and optimum estimators for sampling from finite populations; Ann. Math. Stat., 39, 621-642.

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.

Hoge, V. (1965): Sampling designs which admit uniformly minimum variance unbiased estimates; Bull. Cal. Stat. Assn., 14, 160-162.

Hoge, V. (1966): Admissible estimates for any sampling design; Bull. Cal. Stat., 15, 120-126.

Hoge, V. (1967): An optimum property of the Horvitz and Thompson estimate; Jour. Amer. Stat. Assn., 62, 1013-1017.

Hodges, J.L. (Jr.) and Lehmann, E.L. (1951): Some applications of the Cramer-Rao inequality; Proce. Second Berk. Symp. of Math. Stat. and Prob., 13-22.

Horvitz, D.G. and Thompson, D.J. (1952): A generalisation 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. Expt. Sta. Res. Bull. No. 304.

Joshi, V.M. (1965a,b): Admissibility and Bayes estimation in sampling from finite populations II and III; Ann. Math. Stat., 36, 1723-1742.

Joshi, V.M. (1966): Admissibility and Bayes estimation in sampling from finite populations IV, Ann. Math. Stat., 37, 1658-1670.

Joshi, V.M. (1968): Admissibility of the sample mean as an estimate of the mean of finite population; Ann. Math. Stat., 37, 1658-1670.

Keyfitz, N. (1957): Estimates of sampling variance when two units are selected from each stratum; Jour. Amer. Stat. Assn., 52, 503-510.

Kish, L. (1965): Survey Sampling; John Wiley and Sons, New York.

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 populations, I, II, III, Metrika, 7, 81-114 and 165-204.

Koop, J.C. (1964): On an identity for the variance of a ratio of two random variables; Jour. Roy. Stat. Soc., B, 26, 484-486.

Koop, J.C. (1968): An exercise in ratio estimation; Amer. Statistician, 22, 29-30.

Lahiri, D.B. (1951): A method of sample selection providing unbiased ratio estimators, Bull. Int. Stat. Inst., 33 (2), 133-140.

Lehmann, E.L. (1947): On families of admissible tests; Ann. Math. Stat., 18, 97-104.

Madow, W.G. and Madow, L.H. (1944): On the theory of systematic sampling I, Ann. Math. Stat., 15, 1-24.

Mahalanobis, P.C. (1940): A sample survey of acreage under jute in Bengal; Sankhyā, 4, 511-530.

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 proportionate to the sum of sizes; Ann. Inst. Stat. Math., 3, 99-108.

Murthy, M.N. (1957): Ordered and unordered estimators in sampling without replacement; Sankhyā, 18, 379-390.

Murthy, M.N. (1962): Almost unbiased estimators based on interpenetrating sub-samples; Sankhyā, A, 24, 303-314.

Murthy, M.N. (1963a): Generalised unbiased estimation in sampling from finite populations; Sankhyā, B, 25, 245-262.

Murthy, M.N. (1963b): Some recent advances in sampling theory; Jour. Amer. Stat. Assn., 58, 737-755.

Murthy, M.N. (1964): Product method of estimation, Sankhyā, A 26, 69-74.

Murthy, M.N. (1967): Sampling Theory and Methods, Statistical Publishing Society, India.

Murthy, M.N. and Nanjamma (1959): Almost unbiased ratio estimates based on interpenetrating sub-samples; Sankhyā, 21, 381-392.

Murthy, M.N. and Singh M.P. (1966): Some aspects of recent developments in sampling theory; Symp. Proc., BHU Varanasi.

Murthy, M.N and Singh M.P. (1968b): On the concepts of best and admissible estimators in sampling theory; to appear in Sankhya, 31

Neyman, J. (1934): On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection; Jour. Roy. Stat. Soc., 97, 558-625.

Neyman, J. (1938): Contributions to the theory of sampling human populations; Jour. Amer. Stat. Assn. 33, 101-116.

Olkin, I. (1958): Multivariate ratio estimation for finite populations; Biometrika, 45, 154-165.

Pathak, P.K. (1962): On sampling units with unequal probabilities; Sankhya, A, 24, 315-326.

Pathak, P.K. (1964): Sufficiency in sampling theory; Ann. Math. Stat., 35, 795-808.

Prabhu Ajgaonkar, S.G. (1965): On a class of linear estimators in sampling with varying probabilities without replacement; Jour.Amer. Stat. Assn., 60, 637-642.

Quenouille, M.H. (1956): Notes on bias in estimation; Biometrika, 43, 353-360.

Rao, J.N.K. (1957): Double ratio estimate in forest surveys; Jour. Ind. Soc. Agr. Stat., 9, 191-204.

Rao, J.N.K., Hartley, H.O. and Cochran, W.G. (1962): On a simple procedure of unequal probability sampling without replacement; Jour. Roy. Stat. Soc., B, 24, 482-491.

Rao, J.N.K. (1963): On two systems of unequal probability sampling without replacement; Ann. Inst. Stat. Math., 15, 67-72.

Rao, J.N.K. (1965): A note on estimation of ratios by Quenouille's method. Biometrika, 52, 647-649.

Rao, J.N.K. (1966): On the relative efficiency of some estimators in pps sampling for multiple characteristics; Sankhyā, A, 28, 61-70.

Rao, J.N.K. and Bayless, D.L. (1968a): An empirical study of the stabilities of estimators and variance estimators in pps sampling n=2; to appear in Jour. Amer. Stat. Assn.

Rao, J.N.K. and Bayless, D.L. (1968b): An empirical study of the stabilities of estimators and variance estimators in pps. sampling (n > 2), to be published.

Rao, J.N.K. (1968c): Ratio and regression estimators: Proc. Symp. Foundations of Survey Sampling, Chapel Hill (in press).

Rao, T.J. (1966): On certain unbiased ratio estimators; Ann. Inst. Stat. Math., 18, 117-121.

Rao, T.J. (1967): On the choice of a strategy for ratio method of estimation; Jour. Roy. Stat. Soc., B, 29, 392-398.

Rao, T.J. (1967): <u>Contributions to the theory of sampling strategies</u>; Ph.D. Thesis submitted to the Indian Statistical Institute.

Rao, S.R.S. and Mudholkar, G.S. (1967): Generalised multivariate estimator for the mean of finite populations; Jour. Amer. Stat. Assn., 62, 1009-1012.

Roy, J. and Chakravarti, I.M. (1960): Estimating the mean of a finite population; Ann. Math. Stat., 31, 392-398.

Saito, K. (1958): Theoretical considerations on the use of supplementary information in sample survey designs III, Sophia, Econ. Rev. 86.

Sampford, M.R. (1967): On sampling without replacement with unequal probabilities of selection; Biometrika, 54, 499-514.

Searls, Donald T. (1964): The utilization of a known coefficient of variation in the estimation procedure; Jour. Amer. Stats Assn. 59, 1225-26.

Sen, A.R. (1952): Present status of probability sampling its use in estimation of farm characteristics; (abstract), Econometrica, 20, 103.

Sen, A.R. (1953): On the estimate of the variance in sampling with varying probabilies; Jour. Ind. Soc. Agr. Stat., 5, 119-127.

Seth, G.R. (1961): Some aspects of sampling theory and practice; Presidential address, Proceedings of Indian Science Congress, 48th Session.

Seth, G.R. (1966): On estimates of variance of estimate of population total in varying probabilities; Jour. Ind. Soc. Agr. Stat., 18, 52-56.

Singh, D. (1956): On the efficiency of cluster sampling; Jour. Ind. Soc. Agr. Stat., 8, 45-55.

Singh, D. and Singh, S.D. (1965): Some contributions to two-phase sampling; Aust. Jour. Stat., 7, 45-47.

Singh, M.P. (1965): On the estimation of ratio and product of population parameters; Sankhyā, B, 27, 321-328.

Singh, M.P. (1966): Some aspects of recent developments in sampling theory; Symp. Proc. Banaras Hindu University (with M.N.Murthy).

Singh, M.P. (1967a): On the efficiency of systematic sampling in ratio and product estimation; Metrika, 11 (3), 199-205.

Singh, M.P. (1967b): Ratio cum product method of estimation; Metrika, 12(1), 34-42.

Singh, M.P. (1967c): The relative efficiency of some two-phase sampling schemes, Ann. Math. Stat., 38, 937-940.

Singh, M.P. (1967d): Multivariate product method of estimation for finite populations; Jour. Ind. Soc. Agr. Stat. 19, 2. 1-10.

Singh, M.P. (1968a): Comparison of some ratio cum product estimators; (Tech. paper No.24, NSS, I.S.I.) presented at the 56th session of Indian Science Congress.

Singh, M.P. (1968b): On the concepts of best and admissible estimators in sampling theory; (with M.N. Murthy, to appear in Sankhyā, 31), presented at the 56th Session of Indian Science Congress.

Singh, M.P. (1969): On the utilization of an apriori value of the parameter in the estimation procedure; (with A.S.Roy, Tech. paper No.30 NSS-I.S.I.) presented at the 56th Session of Indian Science Congress .

Singh, M.P. (1969): Some estimators using supplementary information in sampling finite populations, (submitted for publication, with K. B. Pathak).

Srivastava, S.K. (1965): An estimate of the mean of a finite population using several auxiliary variables; Jour. Ind. Stat. Assn., 3(4), 189-194.

Srivastava, S.K. (1967): An estimate using auxiliary information in sample surveys; Cal. Stat. Assn. Bull., 16, 121-132.

Srivastava, S.K. (1969): A generalised estimator for the mean of a finite population using multi-auxiliary information. (abstract, Ann. Math. Stat.

Stevens, W.L. (1958): Sampling without replacement with probability proportional to size; Jour. Roy. Stat. Soc., B, 20, 393-397.

Subrahmanya, M.T. (1965): On generalised estimators in sampling from finite populations; Metrika, 9, 234-239.

Sukhatme, P.V. (1935): Contributions to the theory of representative method; Jour. Roy. Stat. Soc. Supp. 2, 253-268.

Sukhatme, P.V. (1953): Sampling Theory of Surveys with applications; Ind. Soc. Agr. Stat. and Iowa state college press.

Sukhatme, P.V. (1959): Major developments in the theory and applications of sampling during the last twenty-five years, Estadistica, 17, 652-679.

Supla, G. K. (1966): An alternative multivariate ratio estimate for finite populations; Bull. Cal. Stat. Assn. 15, 127-134

Swain, A.K.P.C. (1964): The use of systematic sampling in rati estimate; Jour. Ind. Stat. Assn., 11, 160-164

Tin, M. (1965): Comparison of some ratio estimators; Jour Amer. Stat. Assn., 60, 294-307.

United Nations (1947-51): Reports of the sub-commission on statistical sampling to the Statistical Commission, New York.

Vijayan, K. (1966): On Horvitz - Thompson and Des Raj estimator; Sankhyā, A, 28, 87-92.

Vijayan, K. (1966): Some contributions to Sampling Theory; Ph. D. Thesis submitted to Indian Statistical Institute.

Wald, A. (1947): An essentially complete class of admissible decision function; Ann. Math. Stat. 18.

Wald, A. (1950): Statistical Decision Functions; John Wiley and Sons. New York.

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.

---