Some Problems in Estimating Finite Population Total and Variance in Survey Sampling

(Thesis* submitted to the Indian Statistical Institute
for the degree of Doctor of Philosophy in Science)

By
SASWATI BHATTACHARYYA
INDIAN STATISTICAL INSTITUTE
Calcutta- 700035, INDIA
1996

* This is a revised version of the thesis submitted earlier.

Preface

This thesis is being submitted in fulfilment of the primary requirement for the degree of Doctor of Philosophy in Science of the Indian Statistical Institute(ISI). This is a revised version of the thesis with the same title, submitted in 1993, revision being made according to the suggestions of the examiners on that version.

No part of the thesis has ever been submitted for a degree of any university or a prize. However some of the results described in this dissertation are contained in five papers published respectively in the following journals : Statistics, Calcutta Statistical Association Bulletin, Metron and Journal of Statistical Planning and Inference.

The work leading to the thesis and its revision was carried out at the Computer Science Unit (CSU) of ISI and was supported by a scholarship offered by the Council of Scientific and Industrial Research, Government of India. The author remains deeply indebted to Dr. Parimal Mukhopadhyay of ISI, under whose consultation the work was carried out, for immense help, guidence and inspiration offered throughout the investigation.
The author acknowledges with gratitude the help and encouragement she received from Prof. Arijit Chaudhuri and Mr. Ajay Kumar Adhikari, CSU, ISI. The author is also greatly indebted to her parents Mr. Sohang Bhattacharyya and Mrs. Sukla Bhattacharyya and her husband Dr. Sankar Prasad Bhattacharya for the inspiration constantly offered by them.

Saswati Bhattacharyya

December, 1996
Computer Science Unit,
Indian Statistical Institute,
Calcutta-700035.

Contents

Chapter 1

INTRODUCTION AND SUMMARY

1.0. General Introduction

The problem of drawing inference concerning the parameters of a finite population of identifiable units has been increasingly engaging the attention of statisticians. The central problem here is to devise a suitable method of selecting a sample from the population and to employ an appropriate estimator to estimate the finite population total or mean. A considerable progress in this field of study has been made and many authors have contributed towards the development of the theory in this aspect of the problem of statistical inference.

Numerous papers have been written covering the first aspect of the problem, namely, method of selecting an appropriate sample from a given universe. It has been demonstrated that the unequal probability sampling provides more efficient estimator of population total than that obtained from an equal probability sampling.

Hansen and Hurwitz (1943) were the first to indicate the utility of the method of selection with varying probability. They gave a method of selecting a single unit with probability proportional to size which can be easily extended to select more than one unit if the selection be made with replacement – the probability proprotional to size with replacement (ppswr) scheme. Madow (1949) proposed the use of systematic sampling with unequal probabilities to avoid the possibility of units being selected more than once. Midzuno (1950), Narain (1951), among others, considered the problem of sampling with varying probabilities without replacement (wor). Closely following these authors Horvitz-Thompson (1952),Sen (1953), Yates and Grundy (1953) studied more general methods of sampling wor and with varying probabilities. The variance of the Horvitz-Thompson estimator (HTE) of population total is uniquely determined by the first order and second order inclusion-probabilities of units in a sample for a chosen design and reduces to zero if the variate values are exactly proportional to the corresponding inclusion-probabilities. As the values of the variable of interest are unknown it seems reasonable to choose an auxiliary variable (usually known as size-measure) which is believed to be closely related to the main variable and attempt has been made to develop fixed-size sampling designs with inclusion-probabilities proportional to size-measures. Such designs are

called IPPS (inclusion-probability proportional to size) designs or πps designs (Hanurav (1967)). It was desired to construct πps designs or designs which are approximately πps such that the variance of the HTE have variance less than that of the customary Hansen-Hurwitz (1943) estimator of population total in the ppswr sampling scheme. Apart from the estimator suggested by Horvitz and Thompson themselves an alternative expression for variance of HTE was derived independently by Sen (1953) and Yates and Grundy (1953) which is valid only if the number of units in the sample is fixed. The unbiased estimator of the variance of HTE proved to be disadvantageous since it is not always zero when the variance is zero. An alternative conditioanally unbiased estimator was suggested by Sen (1953) and by Yates and Grundy (1953) which possesses the particular property of being zero when the variance itself is zero. Both the estimators can assume negative values. However under some selection procedures it was demonstrated by Sen (1963), Raj (1956a), Rao and Singh (1963), Lanke (1974), Asok and Sukhatme (1974), among others, that the first estimator could take negative values for some of the pairs whereas the later one takes positive values for all the sample pairs. Many sampling designs such as designs due to Yates-Grundy (1953), Brewer (1963), Hanurav (1962), Fellegi (1963),Rao (1963), Hájek (1964), Carroll-Hartley (1964),Durbin (1967), Sampford (1967), Fuller (1971), Vijayan (1967), Mukhopadhyay (1972), Sinha (1973),Sengupta (1981), Gupta, Nigam and Kumar(1982), Saxena, Singh and Srivastava (1986), Sunter (1986)
, Hedayet, Rao and Stafker (1986) etc. were developed for using HTE. Schemes due to Yates-Grundy (1953), Narain (1951), Yates-Grundy (1953), Brewer and Undy (1969), Brewer (1963),Rao (1965), Hanurav (1967), Durbin (1967), Fuller (1971),Dodds and Fryer (1971), are applicable for $n = 2$ only. Apart from sampling strategies consisting of (approximately) πps design and corresponding HTE some very interesting special procedures like Rao-Hartley-Cochran strategy (1962), Midzuno (1950,52) procedure were suggested and developed. Schemes due to Rao-Hartley-Cochran (1962) and Chikkagouder (1967) use a special estimator other than HTE. Procedures derived by Midzuno (1950,52)-Lahiri (1951)-Sen (1952),
Sankarnarayanan (1969), Deshpande (1978) gave unbiased estimation for ratio estimator. Singh and Srivas-

tava's (1980) procedure was developed with regression estimator in view. Mukhopadhyay (1972) and Sinha (1973) attempted to obtain sampling designs realising a given set of second order inclusion-probabilities. This problem has also been considered by Herzel (1986). Das (1951), Raj (1956), Murthy (1975) have suggested certain other special estimators for use with Yates-Grundy's (1953) draw by draw procedure. Excellent review of sampling designs may be found in Brewer and Hanif (1980),Chaudhury and Vos (1988), Mukhopadhyay (1991).

Apart from the derivations of the above mentioned sampling designs the direction of theoretical research in survey sampling in recent years was primarily guided by two major findings, namely

(a) the non-existence of unbiased minimum variance estimation (Godambe (1955)) and (b) that the likelihood function is independent of the sampling design (Godambe (1966)).

Within the formal survey sampling model, which takes into account individual labels, the likelihood function was found to be independent of the mode of randomization. This implied that once the sample was drawn, any inference (estimation) consistent with the likelihood and conditionality principles should be independent of whether the sample was drawn at random or was drawn purposively.

The discovery that the likelihood function is independent of the mode of randomization gave rise to formal theories of (prior-probabilistic) model based inference and the related purposive selection. Given the data the problem of drawing inference about a population parameter using a predictor was foreshadowed and considered by Brewer (1963). Royall (1970,1971), Royall and Herson (1973) gave general formulations to the problem. After that different models and corresponding optimal estimators were examined by Royall (1975,1976),Cochran (1977), Sarndal (1980), Wright (1983), Isaki and Fuller (1982), Tam (1986) among others. Under exchangeable general linear models [Arnold (1979)] Mukhopadhyay (1988) found the UMVU predictor of the population total. Rodrigues et al (1985) extended the concepts to develop a general theory of prediction which covers both linear and quadratic functions of population values. Skinner (1983) considered the multivariate prediction of mean. Meanwhile question of robustness of the estimators (predictors) arose and Royall and Herson (1973) introduced the concept of balanced samples. Scott, Brewer and Ho (1978) generalised the concept of balanced sampling. Mukhopadhyay (1977, 1985) studied the robustness of some optimal predictors under a class of alternative models. Cumberland and Royall (1981) defined the π-balanced sample. Kott (1986 a,b),Pfefferman (1984), Pereira and Rodrigues (1983), Tallis (1978), among others, contributed some fruitful results in this area. As all types of balanced samples are really non-existent in practice, Royall and Herson (1973), Royall and Pfefferman (1982) recommended srs, approximately stratified random sampling as approximately balanced sampling. Royall and Cumberland (1981) proposed a sampling design to realise the concept of balanced sampling approximately. Iachan (1985) proposed a similar design with some modifications. However to take care of the brittleness of model-dependent predictors under departure from the assumed model, the genesis of the model-based predictors which combine both model-randomization and design-randomization, was evolved. Cassel, Sarndal and Wretman (1976) and Sarndal (1980) suggested the generalised regression predictor (GRP) of population total. Wright (1983) generalising GRP, introduced the QR-predictors. Montanari (1987) generalising Wright's result, defined an enlarged class of QR-predictors. Brewer, Samiuddin and Asad (1989) considered a linear design unbiased estimator, having ratio estimator property and some stability, specially for outliers. The strategies suggested here often do not possess any desirable properties (unbiasedness, attainment of a minimum variance bound etc.) in exact analysis, though in asymptotic analysis most of these properties hold. Brewer (1979) considered the class of predictors which are asymptotically design unbiased (ADU), the predictors being of a particular form suggested by a model and in this class the optimal strategy is one which minimises the asymptotic expected mse. His stand-point is somewhat between design and pure superpopulation as a basis for inference. Robinson and Tsui (1976,1982), Wright (1983), Robinson and Sarndal (1983), Fuller and Isaki (1981),Liu (1983) among others made valuable contributions in asymptotyic analysis. A review of different model-based and model-dependent strategies for drawing inference for a finite population total may be seen in Mukhopadhyay (1990,1992).

The problem of estimating a finite population variance has received little attention compared to a population total or mean. The problem of estimation of a finite population variance was first considered by Liu (1974a). He introduced Horvitz-Thompson (HT)-type estimator of variance and examined its forms under simple random sampling with and without replacement and probability proportional to size (ppswr) sampling procedures. He showed his estimator to be admissible in the class of all unbiased estimators for the population variance and also constructed an admissible general unbiased quartic estimator for the variance. Observing that Liu's estimator can sometimes take negative values, Chaudhury (1978) suggested non-negative alternative estimators and noted some of their properties. Das and Tripathi (1978) obtained the ratio-type and product-type estimators of variance under simple random sampling with replacement. Isaki (1983) considered multivariate ratio and regression estimators of variance following Olkin's (1958) multivariate estimators in case of population total. Following prediction-theory based works of Royall (1970, 1976), Royall and Herson (1973), Mukhopadhyay (1978, 1982, 1984), Chang and Lin (1985) obtained the optimal model-unbiased, design-unbiased and model-design-unbiased predictors of finite population variance. Minimax strategies for estimating the variance has also been obtained. Mukhopadhyay (1984) obtained optimal estimator of variance under generalised random permutation models. He suggested (1990) a predictor of finite population variance under probability sampling suggested by a multiple regression model and showed this to be asymptotically design unbiased and consistent. Valuable contributions in this direction were made by Sankarnarayanan (1980), Zacks and Solomon (1981), Skinner (1981, 1983), Strauss (1982), Ghosh and Meeden (1983), Liu and Thompson (1983), Singh (1983), Vardeman and Meeden (1983), Rodrigues et al (1985) and Sengupta (1988), among others.

1.1 Summary of the results

In the above context we have paid attention to some aspects of the problems of deriving suitable sampling strategies for estimation (prediction) of population total and variance under fixed population and obtained some results which are briefly described chapterwise as follows. In the description of the results of this thesis, references to earlier works have been omitted as they have been given in details in the different chapters deriving these results.

We have used design-based approach in chapter 2, section 4.3 and subsections 4.4.1-4.4.5 of chapter 4, model-based approach in chapter 3, remaining portion of chapter 4, chapters 5 and 6, where, of course, sampling designs have been invoked in some places to find optimal designs, robust designs etc. The work of chapter 7 is Bayesian estimation of finite population proportion.

Chapter 2 mainly covers the problems in estimating a finite population total. In section 2.2 the notations are described. In section 2.3 we have investigated the conditions under which the strategy consisting of Midzuno's scheme of sampling and ratio estimator would be superior to the Hansen-Hurwitz strategy consisting of ppswr sampling and the customary estimator, both the schemes using the same set of values of size-measures. Numerical examples have been given to show that such conditions are realised.

Section 2.4 shows the derivation of k-th order inclusion probability using the sampling design due to Singh and Srivastava (1980) and some investigation of the first and second order inclusion probabilities. It is shown that Yates-Grundy (1953) estimator of variance of Horvitz-Thompson estimator (HTE) of population total is non-negative for this design.

Chapter 3 also covers the problems in estimating a finite population total under superpopulation set-up.

In section 3.2 we considered the prediction of population total under a class of polynomial regression models with variance function given as a polynomial in the regressor variable under different balanced samples due to Royall and Herson (1973), Scott et al (1978), the predictors being some of Royall's (1970) optimal model-dependent predictors. The bias and mse of these predictors along with those of a Horvitz-Thompson predictor under π-balanced sample [Royall and Cumberland (1981)] have been compared. It has been shown that under a wide class of polynomial regression models, the Horvitz-Thompson predictor along with a πps-design (which is expected to provide π-balanced samples on an average) provides a better sampling strategy than model-dependent best linear unbiased predictors at balanced and over-balanced samples. Bias of two optimal predictors at an over balanced sample has also been examined.

Chapter 4 deals with some investigations in the problems of estimating a finite population variance. In section 4.2 we proved that for any given sampling design p with $\pi_{ij} > 0 \; \forall \; i \neq j$ there does not exist any uniformly minimum variance quadratic unbiased edtimator (UMVQUE) of the finite population variance S_y^2. It is also proved that for any non-census design there does not exist any UMVUE of S_y^2 in the class of all unbiased estimators.

We suggest in section 4.3 a non-negative unbiased estimator of a finite population variance S_y^2 which is applicable to any fixed size without replacement design. The variance and estimator of variance of this estimator have been obtained. We considered , in particular, estimation of S_y^2 using SRSWOR design, design due to Lahiri-Sen -Midzuno and design due to Singh and Srivastava (1980). Estimation of S_y^2 under a controlled sampling plan and unbiased estimator of S_y^2 under PPSWR sampling design have also been considered. The performance of several strategies for estimationg S_y^2 have been studied both numerically and under a superpopulation model.

Chapter 5 covers the optimal estimation of a finite population variance under some superpopulation models with exchangeable errors. The robustness of the optimal strategies under a class of alternative models has been examined. Sampling designs ensuring near unbiasedness under alternative models have been investigated.

In chapter 6 we have considered the estimation of finite population variance under measurement error model. A Bayesian approach for prediction of population parameters is considered in Bolfarine(1991) , Mukhopadhyay(1994 a, b, c) among others. In section 6.3 prediction of population variance is considered under uni-stage sampling and that under two-stage sampling is considered in section 6.4.

In chapter 7 a two-stage sampling design is used for estimating population proportion in Bayesian approach under two different priors. Section 7.2 contains the prior and posterior moments of the population proportion under multinomial setup and section 7.3 contains the same under hypergeometric setup.

Chapter 2

ON SOME SAMPLING DESIGNS FOR ESTIMATING A FINITE POPULATION TOTAL

2.1 Introduction and Some Review of Earlier Work

During the last five decades, starting from the work of Madow (1949), many unequal probability without replacement (upwor) sampling designs (s.d.) have been proposed in the literature. Horvitz and Thompson (1952) first suggested an unbiased estimator for any upwor sampling design. The Horvitz-Thompson estimator (HTE, e_{HT}) and its variance $[V(e_{HT})]$ are specified uniquely in terms of the inclusion-probabilities π_i's and π_{ij}'s, the first and second order inclusion probabilities, respectively. If the values y_i of the characteristic 'y' of interest are exactly proportional to π_i and the number of units in the sample is fixed, $V(e_{HT})$ reduces to zero. In practice, y values being unknown, one chooses an auxiliary variable 'x' whose values are (believed to be) closely related to 'y' and attempts have been made to develop fixed size sampling designs with $\pi_i \propto x_i$ $i = 1, 2, \cdots, N$. Such designs are called πps designs (Hanurav, 1967) or IPPS (inclusion- probability proportional to size) designs. Since, in general, x-values are not exactly proportional to y-values, one would be satisfied if π_i's are approximately proportional to x_i or $p_i = \frac{x_i}{X}$, a measure of size of unit i ($X = \sum x_k$).

As listed by Hanurav (1967) the following are some desirable properties of a sampling design to base e_{HT} (for notations, section 2.2 may be seen):

(i) $\pi_i = nx_i/X, x_i \leq \frac{X}{n} \;\forall\, i$,

(ii) $\nu(S) = n \;\forall\, s : p(S) > 0$. Here and subsequently, $p(S)$ denotes the probability of selecting the sample S, $\nu(S)$ the number of distinct units in S, n denoting the fixed sample size.

(iii) $\pi_{ij} > 0 \;\forall\, i \neq j$, where π_{ij} is the second order inclusion probability of the units i and j.

(iv) $\pi_{ij} \leq \pi_i\pi_j \;\forall\, i, j$.

(v) $\phi = \pi_{ij}/\pi_i\pi_j > \beta$, where β is not too close to zero.

(vi) π_{ij}'s should be computable from some simple formulae.

The properties (iii) to (iv) are required for ensuring the existence of an unbiased variance estimator $v(e_{HT})$ and the non-negativity of $v_{YG}(e_{HT})$, an estimator of $V(e_{HT})$ suggested by Yates and Grundy (1953). The property (v) is desirable to make the values of $v_{YG}(e_{HT})$ expectedly stable over different samples. Raj (1956) showed that in samples of size n = 2 if wor sampling is superior to ppswr sampling independently of y's then the condition (iv) is satisfied. Gabler (1984), however, improved upon these conditions and found that sufficient condition for a connected fixed size πps strategy to have $V(e_{HT}) \leq V(e_H)$ where e_H is the Hansen-Hurwitz estimator and $V(e_H)$ denotes its variance based on probability proportional to size with replacement (ppswr)-scheme, both the strategies using the same set $\{p_i, \ i = 1, 2, ..., N\}$ and same n is

$$\sum_i \min_j \frac{\pi_{ij}}{\pi_j} \geq n - 1$$

The inequality $V(e_{HT}) \leq V(e_H)$, is satisfied for the following schemes among others, for n=2 :

Scheme due to Yates-Grundy (1953),Narain (1951), Brewer-Undy (1962), Durbin (1953), Brewer (1963), Durbin (1967), Sampford(1967), Fellegi (1963) , Fuller (1971), Singh (1978).

For $n \geq 2$ the above inequality is asymptotically true under certain assumptions for the schemes due to Yates-Grundy (1953), Yates-Grundy's (1953) and Durbin's (1953) rejective procedure, Goodman and Kish (1950) and Hartley and Rao (1962), among others.

Another special type of sampling strategy is one due to Rao, Hartley and Cochran (1962).

Another estimator of prime inportance is the ratio estimator which is a biased estimator under SRSWOR, but is unbiased under the scheme due to Midzuno (1950, 52), Lahiri (1951)- Sen (1952) (henceforth referred to as Midzuno scheme). Sankarnarayanan's (1969) scheme, Deshpande's (1978) scheme and Deshpande-Ajigaonkar's (1969) schemes, among others are modification of Midzuno's scheme. We note that for both Midzuno's and

Sankarnarayanan's scheme, $p(S) = \alpha \sum_{i \in S} p_i + \beta$, where α, β are suitable constants. For Midzuno's scheme $\beta = 0$. Midzuno's scheme made πps has been studied among others, by Rao (1963), Chaudhury (1974), Mukhopadhyay (1974). Rao (1963) showed that for n = 2, Midzuno's scheme made πps coupled with HTE provides a better strategy than Hansen-Hurwitz (HH) scheme. Chaudhury (1974), Mukhopadhyay (1974) proved this proposition for any given n. We will compare Midzuno's original strategy with ppswr strategy, both using the same p_i and same n and find conditions under which Midzuno's strategy fares better than ppswr strategy in section 2.3.

The linear regression estimator which is biased under SRSWOR remains unbiased under the scheme developed by Singh and Srivastava (1980). We will investigate some properties of this scheme in section 2.4.

Some recently developed sampling designs are due to Chao (1982), Saxena, Singh and Srivastava (1986), Sunter (1977, 1986), Gabler (1987), among others.

2.2. Notations

We shall use the following notations in this chapter and also in the subsequent chapters.

$\mathcal{U}$ denotes a finite population of N identifiable units labelled $1, 2, \cdots, N$. y_k, a real quantity associated with unit k is the value of the variable 'y' of interest of which the population parameters like population total, population mean, variance are sought to be estimated . $\underset{\sim}{y} = (y_1, \cdots, y_k, \cdots, y_N)$ is a point in R_N, the N-dimensional Eucledian space.

$S = \{i_1, i_2, \cdots, i_{n(S)}\}$ is a set of $n(S)$ units in $\mathcal{U}$ denoting a sample of size $n(S)$ taken from $\mathcal{U}$. $\overline{S}$ denotes the complementary set of S, $\overline{S} = \mathcal{U} - S$.

$\mathcal{S} = \{S\}$ is the collection of all samples i.e., the sample space.

The combination $(\mathcal{S}, p)$, or simply p, where p is a probability measure defined on $(\mathcal{S}, \mathcal{P})$, $\mathcal{P}$ being the power set of $\mathcal{S}$ such that $p(S)$ is the probability of selecting S (hence having the property $p(S) \geq 0$, $\sum_{S \in \mathcal{S}} p(S) = 1$) will denote a sampling design. ρ_n will denote the class of all fixed size (n) [FS(n)] designs, $\rho_n : \{p \mid p(S) > 0 \Rightarrow n(S) = n \text{ for every } S \in \mathcal{S}\}$.

$\pi_{i_1, i_2, \cdots, i_k}$ will denote the probability of including units $i_1, i_2, \cdots, i_k$ in a sample. Hence

$$\pi_{i_1, i_2 \cdots, i_k} = \sum_{S \ni i_1, i_2, \ldots, i_k} p(S)$$

$d : \{(k, y_k); k \in S\}$, will denote the data obtained from the observatiuon of a set S and the associated y-values, $S \in \mathcal{S}, y_k \in R_1$.

We shall often have an access to the values of an auxiliary variable x closely related to the main variable y on all the units in the population , x_i being the value of x on unit i ; $p_i = \frac{x_i}{X}$ $(\sum_{i=1}^{N} x_i = X)$ will denote the size measure of unit i. (In the sequel $\underset{\sim}{x}$ will be a vector of real values).

Y =population total $= \sum_{i=1}^{N} y_i, \overline{Y}$ = population mean $= \frac{1}{N} \sum_{i=1}^{N} y_i$.

The estimator $e_H = \sum_{i \in s} \frac{y_i}{np_i}$, where s is a sequence of the units occuring in the sample drawn with probability proportional to size with replacement (ppswr), (where a unit may occur more than once), will denote

Hansen-Hurvitz (HH) estimator and $V(e_H)$ will denote variance of e_H, the combination $(e_H, ppswr)$, denoting the Hansen-Hurvitz (HH-) strategy or PPSWR strategy.

$e_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i}$: the Horvitz- Thompson estimator (HTE), $V(e_{HT})$ is its variance.

$v(T)$ will denote an unbiased estimator of variance of T, $V(T)$.

2.3 A Comparison between Midzuno strategy and PPSWR strategy

2.3.1 Introduction and formulation of the problem

Chaudhury (1974), Mukhopadhyay (1974) derived the superiority (in the smaller mse-sense) of the strategy consisting of Midzuno sampling scheme modified to a πps design and the ratio-estimator over Hansen-Hurwitz (HH) strategy for any n. The same problem had been considered by Rao (1963) for n=2. In this section following Gabler (1984) we investigate the conditions under which Midzuno's original strategy proves superior to HH strategy. The contents of this note was published in Mukhopadhyay and Bhattacharyya (1991).

Gabler (1984) investigated the conditions under which the strategy consisting of Horvitz-Thompson estimator and a fixed size (-n) πps design fares better (in the smaller variance sense) than the probability proportional to size with replacement (PPSWR) design combined with the customary estimator (Hansen-Hurwitz strategy) of population total, both the strategies using the same set of $\{p_i\}$ values. The results in this section are derived following his work.

In Midzuno scheme probability of drawing a sample $p(S) \propto \sum_{i\in S} x_i$ where x_i is the value of an auxiliary variable x on unit i, $p_i = \frac{x_i}{X}$, $\sum_{i=1}^{N} x_i = X$ and hence

$$p(S) = \frac{q_S}{M_1}$$

where

$$q_S = \textstyle\sum_{i\in S} p_i \ ,$$

$$M_i = \binom{N-i}{n-i}, i = 0,1,2$$

$$\pi_i = \tfrac{N-n}{N-1}p_i + \tfrac{n-1}{N-1}$$

$$\pi_{ij} = \tfrac{(n-1)(N-n)}{(N-1)(N-2)}(p_i + p_j) + \tfrac{(n-1)(n-2)}{(N-1)(N-2)}$$

An unbiased estimator of $Y = \sum_{i=1}^{N} y_i$ for Midzuno scheme is

$$e_R = \frac{\sum_{i\in S} y_i}{q_S} \qquad (2.3.1)$$

We denote this strategy by H_1. For PPSWR scheme using the same p_i values, the HH-estimator of Y is

$$e_H = \frac{1}{n}\sum_{i\in s}\frac{y_i}{p_i} \qquad (2.3.2)$$

This strategy is denoted by H_2. Now variance of H_1,

$$V(H_1) = \frac{1}{M_1}\sum_{S\in\mathcal{S}} \frac{\left(\sum_{i\in S} y_i\right)^2}{q_S} - Y^2,$$

$$= \sum_{i=1}^{N} y_i^2\beta_i + \sum\sum_{i\neq j=1}^{N} y_i y_j \beta_{ij},$$

where $\beta_i = \frac{1}{M_1}\sum_{S\ni i}\frac{1}{q_S} - 1$, $\beta_{ij} = \frac{1}{M_1}\sum_{S\ni i,j}\frac{1}{q_S} - 1$. Now $V(H_1)$ can also be written in the quadratic form as follows :

$$\begin{aligned}
V(H_1) &= \sum_{S\in\mathcal{S}}\left\{\frac{\sum_{i\in S} y_i}{q_S} - Y\right\}^2 p(S) \\
&= \sum_{S\in\mathcal{S}}\frac{1}{q_S^2}\left\{\sum_{i\in S} y_i - Yq_S\right\}^2 p(S) \\
&= \sum_{S\in\mathcal{S}}\frac{1}{M_1 q_S}\left\{\sum_{i\in S}(y_i - p_iY)\right\}^2 \\
&= \frac{1}{M_1}\left[\sum_{i=1}^{N} z_i^2(\beta_i+1) + \sum\sum_{i\neq j=1}^{N} z_i z_j(\beta_{ij}+1)\right]. \\
&= \underset{\sim}{Z}'\Lambda\underset{\sim}{Z} \text{ say} \qquad (2.3.3)
\end{aligned}$$

where $z_i = y_i - p_iY$, $\underset{\sim}{Z}' = (z_1,\cdots,z_N)$,

$\Lambda = ((\lambda_{ij}))$, an $N\times N$ matix with

$$\lambda_{ii} = \lambda_i = \beta_i + 1$$

$$\lambda_{ij} = \beta_{ij} + 1$$

We note

$$\sum_{i=1}^{N} z_i = \underset{\sim}{Z}'\underset{\sim}{1} = 0 \qquad (2.3.4)$$

where $\underset{\sim}{1} = (1,1,...,1)$.

Similarly,

$$\begin{aligned}
V(H_2) &= \frac{1}{n}\sum_{i=1}^{N} p_i\left(\frac{y_i}{p_i} - Y\right)^2 \\
&= \underset{\sim}{Z}'D\underset{\sim}{Z} \qquad (2.3.5)
\end{aligned}$$

where D = Diag $(d_{ii} = \frac{1}{\pi_i})$, an $N\times N$ matrix with $\pi_i = np_i$.

Following Gabler (1984) we find conditions under which

$$V(H_1) \leq V(H_2) \; \forall \; \underset{\sim}{y} \in R^N \tag{2.3.6}$$

2.3.2 Main Results

We consider the general eigenvalue problem

$$\Lambda \underset{\sim}{X} = \mu D \underset{\sim}{X} \tag{2.3.7}$$

where for the real eigenvalues $\mu_1 \geq \cdots \geq \mu_N$,

$$\mu_j = max \left\{ \frac{\underset{\sim}{X}' \Lambda \underset{\sim}{X}}{\underset{\sim}{X}' D \underset{\sim}{X}} | \underset{\sim}{X} \in R^N - \{0\}, \right.$$

$$\left. \underset{\sim}{X}' D \underset{\sim}{g}^i = 0, i = 1, 2, \cdots, j-1 \right\} \tag{2.3.8}$$

$\underset{\sim}{g}^i$ is the eigenvector corresponding to the eigenvalue $\mu_i (i = 1, 2, \cdots, N)$ [Rao, C.R. (1983), p.74].

Now (2.3.7) is equivalent to

$$C \underset{\sim}{X} = \mu \underset{\sim}{X} \tag{2.3.9}$$

where $C = ((C_{ij})) = D^{-1}\Lambda$

$$= \begin{pmatrix} \pi_1 & 0 & \cdots & 0 \\ 0 & \pi_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \pi_N \end{pmatrix} \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1N} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{N1} & \lambda_{N2} & \cdots & \lambda_{NN} \end{pmatrix}$$

$$= \begin{pmatrix} \pi_1\lambda_1 & \pi_1\lambda_{12} & \cdots & \pi_1\lambda_{1N} \\ \pi_2\lambda_{21} & \pi_2\lambda_2 & \cdots & \pi_2\lambda_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \pi_N\lambda_{N1} & \pi_N\lambda_{N2} & \cdots & \pi_N\lambda_N \end{pmatrix}$$

so that $C_{ii} = \pi_i\lambda_i, C_{ij} = \pi_i\lambda_{ij}$

Now,

$$\begin{aligned} \sum_{j=1}^{N} C_{ij} &= \sum_{j=1}^{N} \pi_i\lambda_{ij} = \pi_i \sum_{j=1}^{N} \lambda_{ij} = \pi_i \sum_{j=1}^{N} (\beta_{ij} + 1) \\ &= \pi_i \sum_{j=1}^{N} \left\{ \frac{1}{M_1} \sum_{s \ni ij} \frac{1}{q_s} \right\} \end{aligned}$$

$$
\begin{aligned}
&= \pi_i \left\{ \frac{1}{M_1} \sum_{S \ni i} \frac{1}{q_S} + \frac{1}{M_1} \sum_{j(\neq i)=1}^{N} \sum_{S \ni i,j} \frac{1}{q_S} \right\} \\
&= \pi_i \left\{ \frac{1}{M_1} \sum_{S \ni i} \frac{1}{q_S} + \frac{1}{M_1} \sum_{S \ni i} \sum_{j(\neq i) \in S} \frac{1}{q_S} \right\} \\
&= \pi_i \left\{ \frac{1}{M_1} \sum_{S \ni i} \frac{1}{q_S} + \frac{1}{M_1} \sum_{S \ni i} \frac{1}{q_S}(n-1) \right\} \\
&= \pi_i \frac{n}{M_1} \sum_{S \ni i} \frac{1}{q_S} \\
&= n\pi_i\lambda_i \qquad (2.3.10)
\end{aligned}
$$

$$
\begin{aligned}
\text{Again } \sum_{i=1}^{N} C_{ij} &= \pi_j\lambda_j + \sum_{i(\neq j)=1}^{N} \pi_i\lambda_{ij} \\
&= \pi_j\lambda_j + \sum_{i(\neq j)=1}^{N} np_i\{\frac{1}{M_1} \sum_{S \ni i,j} \frac{1}{q_S}\} \\
&= \pi_j\lambda_j + n \sum_{i(\neq j)=1}^{N} p_i \sum_{S \ni i,j} \frac{1}{M_1 q_S} \\
&= \pi_j\lambda_j + \frac{n}{M_1} \sum_{S \ni j} \frac{1}{q_S} \sum_{i(\neq j) \in S} p_i \\
&= \pi_j\lambda_j + \frac{n}{M_1} \sum_{S \ni j} \frac{q_S - p_j}{q_S} \\
&= \pi_j\lambda_j + n - np_j \frac{1}{M_1} \sum_{S \ni j} \frac{1}{q_S} \\
&= \pi_j\lambda_j + n - \pi_j\lambda_j \\
&= n \qquad (2.3.11)
\end{aligned}
$$

Thus C' is a generalised stochastic matrix where $C_{ij} \geq 0$ for all i, j and n is a simple eigenvalue of C. All other eigenvalue of C must be smaller than n. Now $(1, 1, \cdots, 1)$ is an eigenvector associated with the eigenvalue $\mu = n$.

Our problem is to find conditions under which $\frac{V(H_1)}{V(H_2)} = \frac{\underset{\sim}{Z}' \Lambda \underset{\sim}{Z}}{\underset{\sim}{Z}' D \underset{\sim}{Z}} \leq 1, \ \forall \ \underset{\sim}{Z} \in R^N$, subject to the condition $\underset{\sim}{Z}' \underset{\sim}{1} = 0$.

It follows that the eigenvectors $\underset{\sim}{g}^i$ corresponding to the eigenvalues μ_j $(j = 2, \cdots, N)$ must have co-ordinates with sum zero i.e, $\underset{\sim}{g}^i \underset{\sim}{1} = 0$, which satisfies the condition (2.3.4). Hence to find the maximum value of the ratio $\frac{\underset{\sim}{Z}'\Lambda\underset{\sim}{Z}}{\underset{\sim}{Z}'D\underset{\sim}{Z}}$, the second largest eigenvalue is to be considered.

We consider now the following theorem on eigenvalues of a stochastic matrix.

Theorem 2.1 [Brauer (1971), p.191]

Let m_ν be the minimum and M_ν the maximum of the elements of the v-th column of the generalised stochastic matrix A of order n with row sum s and t $=\sum_{\nu=1}^{n} m_\nu, T = \sum_{\nu=1}^{n} M_\nu$. Then each characteristic root μ of A, different from s satisfies

$$|\mu| \leq \min\,(s-t, T-s);$$

Using the theorem stated above on the generalised stochastic matrix C', we have

$$\mu_2 \leq \min\,\{n - \sum_i min_j\ C_{ij}, \sum_i max_j C_{ij} - n\} \qquad (2.3.12)$$

As $\lambda_{ij} \leq \lambda_i$ for all i, j, we have

$max_j\ C_{ij} = \pi_i\lambda_i$ for all i (since $C_{ij} = \pi_i\lambda_{ij}$)

$$\begin{aligned} \text{Now, } \pi_i\lambda_i &= \frac{np_i}{M_1}\sum_{S\ni i}\frac{1}{q_S} \geq \frac{np_iM_1}{\sum_{S\ni i} q_S} \\ &\quad \text{(since arithmatic mean } \geq \text{ harmonic mean)} \\ &= \frac{np_iM_1}{M_1p_i + M_2(1-p_i)} \\ &= \frac{np_i}{\frac{N-n}{N-1}p_i + \frac{n-1}{N-1}} \qquad (2.3.13) \\ \text{Hence } \quad \sum_i max_j\ C_{ij} &= \sum_i \pi_i\lambda_i \\ &= \frac{n}{M_1}\sum_i p_i \sum_{s\ni i}\frac{1}{q_S} \\ &= n\frac{M_0}{M_1} = N \geq n+1 \qquad (2.3.14) \end{aligned}$$

[1]

If $\mu_2 \leq 1$, then (2.3.6) holds. It, therefore, follows from (2.3.14) that a sufficient condition for (2.3.6) to hold good is

$$n - \sum_i min_j C_{ij} \leq 1$$

[1]The inequality in (2.3.14) was pointed out by a referee.

$$\text{i.e,} \sum_i min_j\ C_{ij} \geq n-1 \tag{2.3.15}$$

$$\text{i.e,} \frac{n}{M_1}\sum_i p_i\ min_j \sum_{s\ni ij} \frac{1}{q_S} \geq n-1 \tag{2.3.16}$$

Hence we have the following

Theorem 2.2 A sufficient condition for Midzuno strategy H_1 to be better than the PPSWR strategy H_2, in the smaller variance sense, both the strategies using the same n and $\{p_i, i = 1,2,\cdots,N\}$ values, is the condition (2.3.16).

Remark 2.3.1 Following Rao and Vijayan(1977), $V(H_1)$ can be written in the form

$$V(H_1) = \sum\sum_{i<j=1}^{N} p_i p_j \{1 - \frac{1}{M_1}\sum_{s\ni i,j}\frac{1}{q_S}\}(\frac{y_i}{p_i} - \frac{y_j}{p_j})^2$$

Also

$$V(H_2) = \frac{1}{n}\sum\sum_{i<j=1}^{N}(\frac{y_i}{p_i} - \frac{y_j}{p_j})^2.$$

Hence a sufficient condition for $V(H_1) \leq V(H_2) \forall \underset{\sim}{y} \in R_N$ is

$$1 - \frac{1}{M_1}\sum_{s\ni i,j}\frac{1}{q_S} \leq \frac{1}{n}\ \forall\, i \neq j$$

or

$$\frac{n}{M_1}\sum_{s\ni i,j}\frac{1}{q_S} \geq n-1\ \forall\ i \neq j = 1,...,N \tag{2.3.17}$$

The expression (2.3.17) gives a set of sufficient conditions for $V(H_1) \leq V(H_2)$. For large N it may be tedious to verify all these $\binom{N}{2}$ conditions. The condition (2.3.16) is easy to verify. The condition (2.3.16) may, therefore, be looked upon, as a referee has suggested, as a generation of conditions (2.3.17).

Example (2.1)

Consider a population with $N = 5, n = 3, \underset{\sim}{Y} = (2\ 4\ 5\ 9\ 10)$ and $\underset{\sim}{P} =$ (.16, .18, .20, .22, .24).

Here $\alpha \simeq .1583$

Thus $p_i \geq \alpha\ \forall\ i$ (i)

$$\Lambda = \begin{pmatrix} 1.73 & .89 & .87 & .85 & .83 \\ .89 & 1.70 & .85 & .83 & .82 \\ .87 & .85 & 1.67 & .82 & .80 \\ .85 & .83 & .82 & 1.64 & .78 \\ .83 & .82 & .80 & .78 & 1.62 \end{pmatrix}$$

$$C = \begin{pmatrix} .83 & .43 & .42 & .41 & .40 \\ .48 & .92 & .46 & .45 & .44 \\ .52 & .51 & 1.0 & .49 & .48 \\ .56 & .55 & .54 & 1.08 & .51 \\ .60 & .59 & .58 & .56 & 1.17 \end{pmatrix}$$

$$\underset{\sim}{Z} = (\text{-}2.8\ \text{-}1.4\ \text{-}1.0\ 2.4\ 2.8)$$

$$\underset{\sim}{\pi} = (.48\ .54\ .60\ .66\ .72)$$

$$D = diag.(d_{ii} = \tfrac{1}{\pi_i})$$

$$= \begin{bmatrix} 2.08 & & & & \\ & 1.85 & & & \\ & & 1.67 & & \\ & & & 1.52 & \\ & & & & 1.39 \end{bmatrix}$$

$$\sum_i min_j C_{ij} = 2.4105 > 2 (= n - 1) \qquad (ii)$$

(i) and (ii) show that the condition (2.3.16) is satisfied.

$$\text{Now } V(H_1) = \underset{\sim}{Z}' \pi \underset{\sim}{Z} = 20.56$$

$$\text{and } V(H_2) = \underset{\sim}{Z}' D \underset{\sim}{Z} = 41.25$$

Hence, $H_1 \succ H_2$.

<u>Example 2.2</u> Consider a population with $N = 6, n = 4$
$\underset{\sim}{Y} = (3\ 5\ 6\ 9\ 10\ 12)$ and $\underset{\sim}{P} = (.14\ .14\ .15\ .16\ .20\ .21)$.
Here $\alpha \simeq .13636$;

Hence $p_i > \alpha \ \ \forall\, i$

Also $\sum_i min_j C_{ij} = 3.48 \geq n - 1 = 3$

Thus the condition (2.3.16) is satisfied.

$$V(H_1) = 14.2684,$$

$$V(H_2) = 37.4553,$$

Hence $H_1 \succ H_2$

2.4 A note on sampling design due to Singh and Srivastava [SS(1980)]

In order mainly, to make the linear regression estimator

$$\overline{y}_{lr} = \overline{y} + b(\overline{X} - \overline{x}),$$

$$\text{where } b = \frac{\frac{1}{(n-1)}\sum_{i\in S}(x_i - \overline{x})(y_i - \overline{y})}{\frac{1}{(n-1)}\sum_{i\in S}(x_i - \overline{x})^2} = \frac{s_{xy}}{s_x^2} \text{ say },$$

unbiased for population mean $\overline{Y}$, SS(1980) suggested the following sampling procedure. The scheme is carried in two steps.

Step 1: Two units (i, j) are selected with probability proportional to $(x_i - x_j)^2$.

Step 2: $(n-2)$ units are selected from the remaining $(N-2)$ units in the population by SRSWOR:

It follows that

$$p(S) = \frac{s_x^2}{M_0 S_x^2} \qquad (2.4.1)$$

where

$$M_0 = \binom{N}{n}, s_x^2 = \frac{1}{(n-1)}\sum_{i\in S}(x_i - \overline{x})^2, \overline{x} = \frac{1}{n}\sum_{i\in S} x_i,$$

$$S_x^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \overline{X})^2, \overline{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Step 1 can also be achieved as follows:

Step 1 (a) : Select i at first draw with probability proportional to $(x_i - \overline{X})^2 + \frac{(N-1)S_x^2}{N}$.

Step 1 (b): Select j at second draw with conditional probability

$$p_{j|i} \alpha (x_j - x_i)^2$$

Under this scheme, $E(b) = B, E(\overline{y}_{lr}) = \overline{Y}$ where

$$B = \frac{S_{xy}}{S_x^2}, S_{xy} = \frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \overline{X})(y_i - \overline{Y}),$$

In section 4.4.2, we shall note that for this scheme $\widehat{S}_y(R) = \frac{s_y^2}{s_x^2} S_x^2$ is an unbi ased estimator of the population variance S_y^2, where $s_y^2 = \frac{1}{(n-1)} \sum_{i \in S} (y_i - \overline{y})^2$, and $S_y^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} (y_i - \overline{Y})^2$.

We give here an expression for the k^{th} order inclusion probability for the design.

Lemma 2.1 Under this scheme

$$\pi_{i_1 i_2 \cdots i_k} = 1 - \left(\frac{N-n}{N}\right) \sum_{q_1=1}^{k} \frac{S_x^2(i_q)}{S_x^2} + \frac{(N-n)^{(2)}}{N^{(2)}} \times \sum\sum_{q_1<q_2=1}^{k} \frac{S_x^2(i_{v_1}, i_{v_2})}{S_x^2} - \cdots + (-1)^k \frac{(N-n)^{(k)}}{N^{(k)}} \frac{S_x^2(i_1, \cdots, i_k)}{S_x^2} \qquad (2.4.2)$$

where $z^{(t)} = z(z-1)\cdots(z-t+1)$ and $S^2_{x(i_1,\cdots,i_q)}$ is the variance of x on $\mathcal{U} - (i_1, i_2, \cdots, i_q)$.

Proof: Under this scheme

$$\begin{aligned}
\pi_{i_1,i_2,\cdots,i_k} &= \sum_{S \ni i_1,i_2,\cdots,i_k} p(S) \\
&= \sum_{S \ni i_1,i_2,\cdots,i_k} \frac{s_x^2}{\binom{N}{n} S_x^2} \\
&= \frac{1}{\binom{N}{n} S_x^2} \sum_{S \ni i_1,i_2,\cdots,i_k} s_x^2 \\
&= \frac{1}{\binom{N}{n} S_x^2} \left[\sum_{S \in \mathcal{S}} s_x^2 - \sum_{S \not\ni i_1,i_2,\cdots,i_k} s_x^2 \right] \\
&= \frac{1}{\binom{N}{n} S_x^2} \left[\binom{N}{n} S_x^2 - \left\{ \bigcup_{p=1}^{k} \sum_{S \not\ni i_1,i_2,\cdots,i_p} s_x^2 \right\} \right] \\
&= \frac{1}{\binom{N}{n} S_x^2} \left[\binom{N}{n} S_x^2 - \left\{ \sum_{p_1=1}^{k} \left(\sum_{S \not\ni i_{p_1}} s_x^2 \right) \right. \right. \\
&\quad - \sum\sum_{p_1<p_2=1}^{k} \left(\sum_{S \not\ni i_{p_1}, i_{p_2}} s_x^2 \right) \\
&\quad \left. \left. + (-1)^{k-1} \left(\sum_{S \not\ni i_{p_1},\cdots,i_{p_k}} s_x^2 \right) \right\} \right] \qquad (2.4.3)
\end{aligned}$$

Now,

$$\sum_{S\not\ni i_1,i_2,\cdots,i_p} s_x^2 = \sum_{S\not\ni i_1,i_2,\cdots,i_p} \left\{ \frac{1}{n}\sum_{i\in S} x_i^2 - \frac{1}{n(n-1)}\sum\sum_{i\neq j\in S} x_i x_j \right\}$$

$$= \frac{1}{n} \sum_{i=1(\neq i_1,\cdots,i_p)}^{N} x_i^2 \sum_{S\ni i(\neq i_1,i_2,\cdots,i_p)} 1 - \frac{1}{n(n-1)} \sum\sum_{i\neq j=1(\neq i_1,i_2,\cdots,i_p)}^{N} x_i x_j \times$$

$$\sum_{S\ni i,j(S\not\ni i_1,\cdots,i_p)}^{N} 1$$

$$= \frac{1}{n} \sum_{i=1(\neq i_1,\cdots,i_p)}^{N} x_i^2 \binom{N-p-1}{n-1}$$

$$-\frac{1}{n(n-1)} \sum\sum_{i\neq j=1(\neq i_1,\cdots,i_p)}^{N} x_i x_j \binom{N-p-2}{n-2}$$

$$= \binom{N}{n} \frac{(N-n)\cdots(N-n-p+1)}{N(N-1)\cdots(N-p+1)} S^2_{x(i_1,\cdots,i_p)}$$

Hence

$$\pi_{i_1 i_2\cdots,i_k} = \frac{\binom{N}{n}}{\binom{N}{n} S_x^2} \left[S_x^2 - \left\{ \sum_{p_1=1}^{k} (\frac{N-n}{N}) S^2_{x(i_{p_1})} \right.\right.$$

$$- \sum\sum_{p_1<p_2=1}^{k} (\frac{N-n}{N})(\frac{N-n-1}{N-1}) S^2_{x(i_{p_1},i_{p_2})} + \cdots$$

$$\left.\left. +(-1)^{k-1} \frac{(N-n)\cdots(N-n-k+1)}{N(N-1)\cdots(N-k+1)} S^2_{x(i_1,\cdots,i_k)} \right\}\right]$$

$$= 1 - \frac{N-n}{N} \sum_{p_1=1}^{k} \frac{S^2_{x(i_{p_1})}}{S_x^2}$$

$$+ \frac{(N-n)(N-n-1)}{N(N-1)} \sum\sum_{p_1<p_2=1}^{k} \frac{S^2_{x(i_{p_1},i_{p_2})}}{S_x^2}$$

$$- \cdots + (-1)^k \frac{(N-n)\cdots(N-n-k+1)}{N(N-1)\cdots(N-k+1)} \frac{S^2_{x(i_1,\cdots,i_k)}}{S_x^2}$$

An elegant alternative proof using the notion of exclusion probability is suggested by a referee which is given below.

Let us denote the k-th order inclusion probability $P[S \not\ni (i_1,...,i_k)]$ as $\pi^c_{i_1,...,i_k}$. It follows from the definition (2.4.1) of the sampling design that

$$\begin{aligned}\pi^c_{i_1\ldots i_k} &= \sum_{S\not\ni i_1,\ldots,i_k} \frac{s_x^2}{\binom{N}{n} S_x^2} \\ &= \frac{1}{\binom{N}{n} S_x^2}[\sum_{S\not\ni i_1,\ldots,i_k}\{\frac{1}{n}\sum_{i\in S} x_i^2 - \frac{1}{n(n-1)}\sum\sum_{i\neq j\in S} x_i x_j\}] \\ &= \frac{\binom{N-k}{n}}{\binom{N}{n}} \frac{S^2_{x(i_1,\ldots,i_k)}}{S_x^2} = \frac{(N-n)^{(k)}}{N^{(k)}} \frac{S^2_{x(i_1,\ldots,i_k)}}{S_x^2} \qquad (2.4.4)\end{aligned}$$

Using the above result lemma 2.1 follows from an elementary result in probability theory viz.

$$\pi_{i_1,\ldots,i_k} = 1 - \sum_{j=1}^k \pi^c_{i_j} + \sum\sum_{j<j'=1}^k \pi^c_{i_j i_{j'}} + \ldots + (-1)^k \pi^c_{i_1,\ldots,i_k}$$

By lemma 2.1,

$$\begin{aligned}\pi_i &= 1 - \frac{N-n}{n}\frac{S^2_{x(i)}}{S_x^2} \\ &= \frac{n(N-1)-N}{N(N-2)} + \frac{(N-n)}{N-2} q_i, \qquad (2.4.5) \\ \text{where } q_i &= \frac{(x_i - \overline{X})^2}{\sum_{k=1}^N (x_k - \overline{X})^2} \\ \text{as } S^2_{x(i)} &= \frac{1}{(N-2)}\left[(N-1)S_x^2 - \frac{N}{N-1}(x_i - \overline{X})^2\right]\end{aligned}$$

Obviously, $\sum \pi_i = n$.

Similarly from (2.4.2) we have

$$\pi_{ij} = 1 - \frac{N-n}{N}\left\{\frac{S^2_{x(i)}}{S_x^2} + \frac{S^2_{x(j)}}{S_x^2}\right\}$$

$$+\frac{(N-n)(N-n-1)}{N(N-1)}\frac{S^2_{x(i,j)}}{S_x^2} \qquad (2.4.6)$$

As

$$S^2_{x(i,j)} = \frac{N-1}{N-3}S_x^2 - \frac{(N-1)}{(N-2)(N-3)}\left\{(x_i - \overline{X})^2 + (x_j - \overline{X})^2\right\}$$

$$-\frac{2}{(N-2)(N-3)}(x_i - \overline{X})(x_j - \overline{X}) \qquad (2.4.7)$$

We verify that $\sum_{j(\neq i)=1}^{N} \pi_{ij} = (n-1)\pi_i$;

We now investigate whether for this design $\psi_{ij} = \pi_i\pi_j - \pi_{ij} \geq 0 \; \forall \; i \neq j$. This condition ensures non-negativity of $v_{YG}(e_{HT})$ for the sampling design. The investigation is done following a suggestion of the reviewer.

Consider first $n = 2$. We have

$$\begin{aligned}
\pi_i\pi_j - \pi_{ij} &= \pi_i^c\pi_j^c - \pi_{ij}^c \\
&= (\frac{N-2}{N})^2\frac{S_{x(i)}^2 S_{x(j)}^2}{S_x^4} - \frac{(N-2)(N-3)}{N(N-1)}\frac{S_{x(ij)}^2}{S_x^2} \quad \text{by (2.4.4)} \\
&= \frac{(N-2)}{N^2(N-1)S_x^4}[(N-2)(N-1)S_{x(i)}^2 S_{x(j)}^2 \\
&\quad -N(N-3)S_x^2 S_{x(ij)}^2] \qquad (2.4.8)
\end{aligned}$$

Since

$$S_{x(i)}^2 = \frac{(N-1)}{(N-2)}[S_x^2 - \frac{N}{(N-1)^2}(x_i - \bar{X})^2]$$

and

$$\begin{aligned}
S_{x(ij)}^2 = \frac{(N-1)}{(N-3)}[S_x^2 - \frac{1}{N-2}\{(x_i - \bar{X})^2 + (x_j - \bar{X})^2\} \\
-\frac{2}{(N-1)(N-2)}(x_i - \bar{X})(x_j - \bar{X})]
\end{aligned}$$

therefore (2.4.8) simplifies to

$$\begin{aligned}
\frac{1}{N^2(N-1)^2 S_x^4}[(N-1)^2 S_x^4 + N^2(x_i - \bar{X})^2(x_j - \bar{X})^2 \\
+2N(N-1)S_x^2(x_i - \bar{X})(x_j - \bar{X})]
\end{aligned}$$

which is clearly non-negative. Therefore for $n = 2$, $\psi_{ij} \geq 0$, the result then follows from a general result of Lanke[(1975,Theorem 5.1); vide also Seth (1966)] which is stated as follows:

Theorem 2.3: Let p' be a s.d. of fixed size n' such that its first and second order inclusion probabilities, π_i' and π_{ij}' satisfy

$$\pi_{ij}' \leq \pi_i'\pi_j' \;\; \forall \;\; i,, j.$$

Further, let p be the sampling design which is performed as follows: first draw a sample S' according to p'. Then draw a sample S'' of size n'' with simple random sampling without replacement from $\mathcal{U} - S'$ and finally put $S = S' \cup S''$. Then the inclusion probabilities of p satisfy $\pi_{ij} \leq \pi_i\pi_j$ for all i and j for the sampling design p.

Chapter 3

ON SOME MODEL-BASED OPTIMAL SAMPLING STRATEGIES FOR PREDICTING A POPULATION TOTAL.

3.1 Introduction, Review of Some Earlier Work and Summary :

Brewer (1963), Royall (1970), Royall and Herson (1973) and their followers considered the prediction-theoretic approach for making inference about the population total Y, starting from an assumed super-population model. Here the population vector $\underset{\sim}{y} = (y_1, y_2, ..., y_N)$ is considered as a realisation of a random vector $\underset{\sim}{Y} = (Y_1, Y_2, ..., Y_N)$[$Y_i$ being the random variable corresponding to y_i] having a joint distribution η and the total $Y = \sum_{i=1}^{N} Y_i$, which is now a random variable, is predicted on the basis of the set of random variables $\{(k, Y_k),\ k \in S\}$ and where y_k must be substituted for Y_k after the data $d = \{(k, y_k),\ k \in S\}$ have been collected from the field. A statistic $t(\underset{\sim}{Y})$ is a predictor of Y if it is an estimator of $\mathcal{E}(Y)$, ie. if $\mathcal{E}\{t(\underset{\sim}{Y})\} = \mathcal{E}(Y)$, $\mathcal{E}$ *denoting expectation wrt model.* We note that we are using the same symbol Y to denote the sum of fixed population values $y_1, y_2, ..., y_N$ as well as the sum of N random variables $Y_1, Y_2, ..., Y_N$. For a given sample S, an optimal predictor $\hat{T}_S$ of Y is one which is unbiased $[\mathcal{E}(\hat{T}_S - Y) = 0\ \forall\ S\ :\ p(S) > 0]$ and for which η-variance of $\hat{T}_S - Y$ is minimum in the class of all unbiased predictors of Y $[\mathcal{E}(\hat{T}_S - Y)^2 \leq \mathcal{E}(\hat{T}'_S - Y)^2\ \forall\ S : p(S) > 0$ and $\forall\ \hat{T}'$, unbiased for Y]. If optimality is considered in the linear unbiased class, one gets the best linear unbiased predictor (BLUP) of Y.

Assuming that with each unit i there is available a real quantity x_i, the value of an auxiliary variable 'x' (closely related to main variable 'y') on i, Royall and Herson (1973) showed that under polynomial regression model $\xi(\delta_0, \delta_1, ..., \delta_J; v(x))$ where $Y_1, Y_2, ..., Y_N$ are independently distributed with

$$\begin{aligned} \mathcal{E}(Y_k \mid x_k) &= \sum_{j=0}^{J} \beta_j \delta_j x_k^j \\ \mathcal{V}(Y_k \mid x_k) &= \sigma^2 v(x_k) \qquad (3.1.1) \end{aligned}$$

where $\beta_0, ...\beta_J$ are unknown constants, $\delta_j = 1(0)$ according as the term x_k^j is present (absent) in $\mathcal{E}(Y_k|x_k)$ and $v(x_k)$ is a known function of x_k, the predictor

$$\hat{T}^{*}(\delta_0, ..., \delta_J; v(x)) = \sum_{k\in S} y_k + \sum_{k\in \overline{S}} \sum_{j=0}^{J} \hat{\beta}_j^{\;*} \delta_j x_k^j \tag{3.1.2}$$

where $\hat{\beta}_j^{\;*}$ is the generalised least square predictor of β_j under the model (3.1.1), is the BLUP of Y. Often $v(x_k)$ is taken as x_k^g, $g \in [0,2]$ for a wide class of socio-economic surveys [Jessen (1942), Mahalanobis (1946), Scott et al (1978)]. We shall denote $\hat{T}^{*}(0,1;x^g)$ as $\hat{T}_g^{*}$. thus

$$\begin{aligned}
\hat{T}_0^{\;*} &= \sum_S y_k + \frac{\sum_S x_k y_k}{\sum_S x_k^2} \sum_{\overline{S}} x_k \\
\hat{T}_1^{*} &= \sum_S y_k + \frac{\sum_S y_k}{\sum_S x_k} \sum_{\overline{S}} x_k \\
\hat{T}_2^{*} &= \sum_S y_k + \sum_S \frac{y_k}{x_k} \sum_{k\in\overline{S}} x_k
\end{aligned} \tag{3.1.3}$$

where $\sum_S$ denotes $\sum_{k\in S}$.

The predictors mentioned above do not depend on any sampling design (unlike, design-based estimators like HTE, etc.). The BLU-predictors have been shown to be subject to serious bias when the assumed superpopulation model is incorrect [Royall and Herson (1973), Mukhopadhyay (1977), Scott et al (1978), Hansen et al (1983)].

In section 3.2 taking a clue from Scott, Brewer and Ho (1978)' we consider the prediction of a population total under a class of polynomial regression models with variance function given as a polynomial in the regressor variable when the samples selected are balanced samples, such as those due to Royall and Herson (1973), Scott et al (1978), (who also defined overbalanced samples) and the predictors are chosen from the class of Royall's (1970) optimal model-dependent predictors. The bias and mse of these predictors and those of a Horvitz-Thompson predictor under π-balanced samples due to Royall and Cumberland (1981) have been compared. It has been shown that under a wide class of polynomial regression models , the Horvitz-Thompson predictor along with a πps-design (which is expeceted to provide π-balanced samples on an average) provides a better sampling straregy than model-dependent best linear unbiased predictors at balanced and over-balanced samples. Bias of two optimal predictors at an over-balanced sample has also been examined.

3.2 Predicting a Population Total Under Balanced Sample.

3.2.1 Review of the earlier work

We consider again the polynomial regression model $\xi = \xi(\delta_0, ..., \delta_J; v(x))$ as given in (3.1.1) and the BLUP of Y under this model as noted in (3.1.2).

For an arbitrary sample, $\hat{T}_1^*$ (in (3.1.3)) is ξ-biased for population total Y when ξ changes from $\xi(0,1;x)$ to any other model $\xi(\delta_0, ..., \delta_J; v'(x))$, $v'(x)$ being an arbitrary function of x. We note that the bias of $\hat{T}_g^*$ does not depend on the form of the variance function under the model $\xi(\delta_0, ..., \delta_J; v'(x))$ [Mukhopadhyay (1977)]. However, under the balanced sample $S_b(J)$ of order J [Royall and Herson (1973)] which satisfies

$$\overline{x}_S^{(j)} = \overline{x}_{\overline{S}}^{(j)} = \overline{X}^{(j)}, \ j = 1, 2, ..., J \quad (3.2.1)$$

where

$$\overline{x}_S{}^{(j)} = \frac{1}{n} \sum_S x_k^j, \quad \overline{x}_{\overline{S}}^{(j)} = \frac{1}{(N-n)} \sum_{\overline{S}} x_k^j$$

and $\overline{X}^{(j)} = \frac{1}{N} \sum_{i=1}^N x_k^j$, $\hat{T}_1^*$ is unbiased under $\xi(\delta_0, ..., \delta_J; v(x))$ for arbitrary $\delta_0, ..., \delta_J, v(x)$.

Similarly $\hat{T}_2^*$ is biased for Y under ξ in general, except on a over-balanced sample $S_o(J)$ of order J [Scott et al (1978)] where it is unbiased. An overbalanced sample $S_0(J)$ is defined as a sample S for which

$$\overline{x}_S{}^{(j-1)} = \frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}}, \ j = 0, 1, 2..., J \quad (3.2.2)$$

Scott et al showed that on a sample $S_{v(X)}(J)$ which we shall call "generalised balanced sample" or "v(x)-balanced sample" of order J and which satisfies

$$\frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}} = \frac{\sum_S \frac{x_k^{j+1}}{v(x_k)}}{\sum_S \frac{x_k^2}{v(x_k)}}, \ j = 0, 1, ..., J, \quad (3.2.3)$$

$\hat{T}_S^*(0,1;v(x))$ remains unbiased under $\xi(\delta_0, ..., \delta_J; V(x)) = \xi_0$ (say), V(x) being not necessarily identical with v(x). $S_{v(X)}(J)$ depends only on v(x) and not on $\delta_0, ..., \delta_J$. When $v(x) = x^g$ we may call $S_{v(X)}(J)$ as a g-balanced sample of order J (provided such a sample exists). $S_b(J)$ and $S_o(J)$ are $S_x(J)$ and $S_{x^2}(J)$ respectively. A x^0-balanced sample of order J will satisfy

$$\frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}} = \frac{\sum_S x_k^{j+1}}{\sum_S x_k^2}, \ j = 0, 1, ..., J$$

Any type of balanced sample are, however seldom available in practice.

Scott et al (1978) proved the following

Theorem 3.1 On $S_{v(X)}(J)$, $\hat{T_S}^*(0,1;v(x))$ is BLUP of Y under $\xi_0 = \xi(\delta_0,...,\delta_J;V(x))$ provided

$$V(x) = v(x)\sum_{j=0}^{J}\delta_j a_j x^{j-1} \qquad (3.2.4)$$

a_j's being arbitrary non-negative constants.

Thus on $S_{(v(x)}(J)$, $\hat{T}^*(0,1;v(x))$ is BLUP under $\xi(\delta_0,...,\delta_J;V(x))$ for arbitrary J, $\delta_0,...,\delta_J$ so long V(x) is of the form (3.2.4).

Royall and Cumberland (1981) defined a π-balanced sample of order J, $S_\pi(J)$ as one which satisfies

$$\Delta_j(S) = \overline{X}\overline{x}_S^{(j-1)} - \overline{X}^{(j)} = 0,\ j = 0,1,...,J \qquad (3.2.5)$$

For a πps sampling design ($\pi_i \propto \frac{x_i}{X}$, $i = 1,2,...,N$) π-balanced samples are met in expectation ie., $E\{\Delta_j(S)\} = 0$, $j = 1,2,...,J$, E denoting expectation wrt a πps-design p. The Horvitz-Thompson predictor $\hat{T}_{HT} = \frac{X}{n}\sum_S \frac{Y_k}{x_k}$ becomes model-unbiased for Y under $\xi(\delta_0,...,\delta_J;V(x))$ for arbitrary $\delta_0,...,\delta_J,V(x)$ on $S_\pi(J)$, because

$$\begin{aligned}
\mathcal{E}(\hat{T}_{HT}) &= \frac{X}{n}\sum_S\sum_{j=0}^{J}\delta_j\beta_j x_k^{j-1} \\
&= X\sum_{j=0}^{J}\delta_j\beta_j\overline{x}_S^{(j-1)} \\
&= N\sum_{j=0}^{J}\delta_j\beta_j\overline{X}^{(j)} \\
&= \sum_{j=0}^{J}\delta_j\beta_j\sum_{k=1}^{N}x_k^j = \mathcal{E}(\sum_{k=1}^{N}Y_k)
\end{aligned}$$

However, $\hat{T}_{HT}$ is not known to have any BLUP-property on $S_\pi(J)$under any subclass of models ξ.

In this section we shall consider purposive sampling designs like balanced sampling, π -balanced sampling etc., where the sample satisfying certain conditions (like (3.2.1), (3.2.2) etc.) is selected with certainty (and with probability $\frac{1}{K}$ if $K(>1)$ such samples exist).

Hence a sampling strategy will be here a combination $(\hat{T},S)$ S the purposively selected sample. We denote $(\hat{T}_1^*,S_x(J))$, $(\hat{T}_2^*,S_{x^2}(J)$, $(\hat{T}_{HT},S_\pi(J))$ as strategies H_1, H_2 and H_3 respectively. H_i will be a better strategy than H_j ($H_i \succ H_j$) if $\mathcal{V}(H_i) < \mathcal{V}(H_j)$.

3.2.2 Some specified subclasses of models ξ

We shall consider the subclasses of $\xi(\delta_0, ..., \delta_J; V(x))$ of the class of superpopulation models ξ as follows :

(a) The particular form

$$V(x) = \sum_{k}^{L} \lambda_j a_j x^j \qquad (3.2.6a)$$

will be denoted as $V'_{k,L}(x)$, where $\lambda_j = 1(0)$ if x^j is present (absent) in V(x). Here a_j's have been assumed as in theorem 3.1 to be arbitrary non-negative constants.

If further λ_j's satisfy the restriction

$$\delta_j = 0 \Rightarrow \lambda_j = 0, \ \lambda_{j+1} = 0 \quad j = 0, 1, ..., J-1 \qquad (3.2.6b)$$

$V'_{k,L}(x)$ will be denoted as $V_{k,L}(x)$.

The model $\xi(\delta_0, ..., \delta_J; V_{k,L}(x))$ will be denoted as $\xi_{k,L}$. Expectation operators and variance operators wrt this model will be denoted as $\mathcal{E}_{k,L}$ and $\mathcal{V}_{k,L}$ respectively.

Example 1. $\xi(0, 1, 1; V_{0,2}(x)) = \xi_{0,2}$ describes, according to the above definition,

$$\mathcal{E}(Y_k|x_k) = \beta_1 x_k + \beta_2 x_k^2$$

and

$$\mathcal{V}_{0,2}(x_k) = a_2 x_k^2$$

[since $\delta_0 = 0$, we have $\lambda_0 = 0, \ \lambda_1 = 0$ by (3.2.6b)].

Example 2. $\xi(1, 0, 1, 1; V_{0,3}(x)) = \xi_{0,3}$ describes

$$\mathcal{E}(Y_k|x_k) = \beta_0 + \beta_2 x_k^2 + \beta_3 x_k^3$$

and

$$\mathcal{V}_{0,3} = a_0 + a_3 x_k^3$$

[since $\delta_1 = 0$, we have $\lambda_1 = 0, \ \lambda_2 = 0$ by (3.2.6b)].

It follows from (3.2.4) that under $\xi_{1,J}$, $\hat{T}_1^*$ is BLUP at $S_x(J)$. This can be seen as follows :

Under $\xi_{1,J}$,

$$V_{1,J}(x) = \sum_{j=1}^{J} \lambda_j a_j x^j \quad (3.2.7)$$

where λ's satisfy (3.2.6b).

By theorem 3.1,

$\hat{T}_1^* = \hat{T}_S{}^*(0,1;x)$ is BLUP of Y under ξ_0 on $S_x(J)$ provided

$$V(x) = x\sum_{j=0}^{J}\delta_j a_j x^{j-1} = \sum_{j=0}^{J}\delta_j a_j x^j \qquad (3.2.8)$$

(3.2.7) is a particular case of (3.2.8) with $a_0 = 0$ [note that by (3.2.6b), $\delta_j = 0 \Rightarrow \lambda_j = 0$].

Similarly, under $\xi_{1,J}$, $\hat{T}_2^*$ is BLUP of Y at $S_{x^2}(J)$, because, by(3.2.4), $\hat{T}_2^*$ is BLUP under ξ_0 provided

$$V(x) = x^2\sum_{j=0}^{J}\delta_j a_j x^{j-1} = \sum_{j=0}^{J}\delta_j a_j x^{j+1} \qquad (3.2.9)$$

(3.2.7) is a particular case of (3.2.9) with $a_{J+1} = 0$, $a_0 = 0$ [note that by (3.2.6b) $\delta_j = 0 \Rightarrow \lambda_{j+1} = 0$].

It is seen that under $\xi_{1,J}$ (where $V_{1,J}(x) = \sum_{j=1}^{J}\lambda_j a_j x^j$), predictor $\hat{T}_1^*$ at $S_x(J)$ and predictor $\hat{T}_2^*$ at $S_{x^2}(J)$, are BLUP of Y. We shall, therefore work with the model $\xi_{1,J}$, since both (3.2.8) and (3.2.9) include $V_{1,J}(x)$.

We shall first examine the bias of $\hat{T}_0{}^*$ and $\hat{T}_1^*$ under general model ξ on over-balanced samples. We shall then compare the performance of the optimal strategies H_1, H_2 under models $\xi_{1,J-1}$. Subsequently we shall examine the performance of strategies H_1, H_2, H_3 under $\xi_{2,J-1}$ and observe that though $\hat{T}_{HT}$ is not known to have any BLUP-property on $S_\pi(J)$ under any subclass of models, H_3 fares better than both H_1 and H_2 under $\xi_{2,J-1}$ under very general conditions. This result seems to provide some justification for use of a suitable design-based strategy in preference to model-dependent strategies.

3.2.3. Some Preliminaries

(a) The following inequality will be used in the calculation :

Callebaut's (1965) inequality : If $(A_1, A_2, ..., A_M)$, $(B_1, B_2, ..., B_M)$ are two sets of positive quantities which are not proportional, the quantity

$$\phi = (\sum_{i=i}^{M} A_i^{w+z} B_i^{w-z})(\sum_{i=1}^{M} A_i^{w-z} B_i^{w+z}) \qquad (3.2.10)$$

increases monotonically with increasing value of $|z|$ for fixed w.

(b)**Lemma 3.1** [Mukhopadhyay (1977)] :

Let $\psi(x_1, ..., x_m; y_1, ..., y_m)$ be a polynomial in $y_1, ... y_m$ of order ≤ 2. Then under the model ξ

$$\mathcal{E}\left[\psi(x_1, ..., x_m; y_1, ..., y_m)\right]$$

$$= \psi[x_1, ..., x_m; \mathcal{E}(y_1), ..., \mathcal{E}(y_m)]$$
$$+\psi(x_1, ..., x_m; y_1, ..., y_m)_{\substack{y_i^2 = v_\xi(y_i) \\ \theta(y)=0}}$$

where $\theta(y)$ denotes terms containing y_i and $y_i y_j$ $(i \neq j = 1, 2, ..., m)$.

3.2.4 Bias of $\hat{T}_0^*$, $\hat{T}_1^*$ under ξ.

To observe the behaviour of biases of $\hat{T}_0^*$ and $\hat{T}_1^*$ under ξ algebraically, we consider the following section though it is not feasible to use the predictors $\hat{T}_0^*$ and $\hat{T}_1^*$ when one chooses an overbalanced sample.

We consider the following class of models

$$(i) \quad \xi_1 \ : \ \{\xi(\delta_0, ..., \delta_J; v(x))\}$$

with $\delta_0 = 0$ and at least one of $\delta_2, ..., \delta_J$ is not equal to zero;

$$(ii) \quad \xi_2 \ : \ \{\xi(\delta_0, ..., \delta_J; v(x))\},$$

with $\delta_o = 1$ and $\delta_2 = ... = \delta_J = 0$.

Next we consider bias of $\hat{T}_0^*$, $\hat{T}_1^*$ under ξ on a over-balanced sample $S_{x^2}(J)$. We have the bias of $\hat{T}_0^*$ under ξ,

$$\begin{aligned}
B_\xi(\hat{T}_0^*) &= \mathcal{E}_\xi(\hat{T}_0^* - Y) \\
&= \mathcal{E}_\xi\{\frac{\sum_S x_k Y_k}{\sum_S x_k^2}\sum_{\bar{S}} x_k - \sum_{\bar{S}} Y_k\} \\
&= \frac{\sum_{\bar{S}} x_k}{\sum_S x_k^2}\sum_S x_k(\sum_{j=0}^{J}\delta_j\beta_j x_k^j) \\
&\quad -\sum_{\bar{S}}(\sum_{j=0}^{J}\delta_j\beta_j x_k^j) \\
&= \sum_{j=0}^{J}\delta_j\beta_j\{\frac{\sum_S x_k^{j+1}}{\sum_S x_k^2} \\
&\quad -\frac{\sum_{\bar{S}} x_k^j}{\sum_{\bar{S}} x_k}\}\sum_{\bar{S}} x_k
\end{aligned}$$

At $S_{x^2}(J)$, using (3.2.2), this reduces to

$$B_\xi\{\hat{T}_0^*|S_{x^2}(J)\} = \sum_{\bar{S}} x_k \sum_{j=0}^{J}\delta_j\beta_j\{\frac{\sum_S x_k^{j+1}}{\sum_S x_k^2}$$

$$
\begin{aligned}
& -\frac{\sum_S x_k^{j-1}}{n}\} \\
= & \frac{\sum_{\overline{S}} x_k}{n \sum_S x_k^2} \sum_{j=0}^{J} \delta_j \beta_j \{n \sum_S x_k^{j+1} \\
& -\sum_S x_k^{j-1} \sum_S x_k^2\} \\
= & \frac{n \sum_{\overline{S}} x_k}{\sum_S x_k^2} \sum_{j=0}^{J} \delta_j \beta_j Cov.(x_k^2, x_k^{j-1}) \\
= & B_0 \quad \text{(say)}, \qquad (3.2.11)
\end{aligned}
$$

Cov. denoting sample covariance.

Similarly,

$$
\begin{aligned}
B_\xi(\hat{T}_1^* | S_{x^2}(J)) = & \mathcal{E}_\xi\{(\sum_{\overline{S}} x_k \frac{\sum_S Y_k}{\sum_S x_k} - \sum_{\overline{S}} Y_k) | S_{x^2}(J)\} \\
= & \sum_{\overline{S}} x_k \sum_{j=0}^{J} \delta_j \beta_j \frac{1}{\sum_S x_k} \{\sum_S x_k^j \\
& -\frac{1}{n} \sum_S x_k^{j-1} \sum_S x_k\} \\
= & \frac{\sum_{\overline{S}} x_k}{\overline{x}_S} \sum_{j=0}^{J} \delta_j \beta_j Cov.(x_k, x_k^{j-1}) \\
= & B_1 \quad \text{(say)} \qquad (3.2.12)
\end{aligned}
$$

Thus if

$$
x_k > 0 \ \forall \ k \text{ and } \beta_j \geq 0 \ \forall \ j = 0, 1, ..., J, \qquad (3.2.13)
$$

both (3.2.11) and (3.2.12) are positive under ξ_1. Under conditions (3.2.13) both (3.2.11) and (3.2.12) are negative under ξ_2.

Now coefficient of $\delta_j \beta_j (j \neq 0)$ in $(B_0 - B_1)$ is

$$
\begin{aligned}
& \frac{n \sum_{\overline{S}} x_k}{\sum_S x_k^2} Cov.(x_k^2, x_k^{j-1}) - \frac{\sum_{\overline{S}} x_k}{\overline{x}_S} Cov.(x_k, x_k^{j-1}) \\
& = (\sum_{\overline{S}} x_k) [\frac{n}{\sum_S x_k^2} \{\frac{1}{n} \sum_S x_k^{j+1} - (\frac{1}{n} \sum_S x_k^2)(\frac{1}{n} \sum_S x_k^{j-1})\} \\
& \qquad -\frac{1}{\overline{x}_S} \{\frac{1}{n} \sum_S x_k^j - (\frac{1}{n} \sum_S x_k)(\frac{1}{n} \sum_S x_k^{j-1})\}]
\end{aligned}
$$

$$= (\sum_{\bar{S}} x_k)[\frac{\sum_S x_k^{j+1}}{\sum_S x_k^2} - \frac{1}{n}\sum_S x_k^{j-1} - \frac{\sum_S x_k^j}{\sum_S x_k} + \frac{1}{n}\sum_S x_k^{j-1}]$$

$$= (\sum_{\bar{S}} x_k)\frac{[\sum_S x_k \sum_S x_k^{j+1} - \sum_S x_k^2 \sum_S x_k^j]}{(\sum_S x_k \sum_S x_k^2)} \qquad (3.2.14)$$

If $x_k > 0 \ \forall \ k$, (3.2.14) is a positive quantity.

This can be seen as follows :

Putting in Callebaut's inequality (3.2.10),

$$M = n, \ A_k = x_k, \ B_k = 1, \ \forall \ k, \ w = \frac{j+2}{2}$$

one gets

$$\phi = \sum_S x_k \sum_S x_k^{j+1} = \phi_1 \quad \text{(say) , for } \ z = \frac{j}{2}$$

and

$$\phi = \sum_S x_k^2 \sum_S x_k^j = \phi_2 \quad \text{(say), for } z = \frac{j-2}{2}.$$

Hence $\phi_1 > \phi_2$ by Callebaut's inequality.

Coefficient of $\delta_0\beta_0$ in $(B_0 - B_1)$ is

$$\frac{-n\sum_{\bar{S}} x_k \sum_S (x_k - \bar{x}_S)^2}{(\sum_S x_k)(\sum_S x_k^2)} \qquad (3.2.15)$$

Hence we have

Theorem 3.2 : Under the assumptions (3.2.13), both $\hat{T}_0^*$, $\hat{T}_1^*$ are positively biased on $S_{x^2}(J)$ under ξ_1. However, in this case

$$|B(\hat{T}_1^*)| \ < \ |B(\hat{T}_0^*)| \qquad (3.2.16)$$

Under (3.2.13), both $\hat{T}_0^*$, $\hat{T}_1^*$ are negatively biased on $S_{x^2}(J)$ under ξ_2. However here also (3.2.16) holds.

Thus under (3.2.13), $\hat{T}_1^*$ is a more robust estimator than $\hat{T}_0^*$ on $S_{x^2}(J)$ both under ξ_1 and ξ_2 in the sense of bias.

3.2.5 Model Variance of $\hat{T}^*(0,1;v(x))$:

Now we have under $\xi_0 = \xi(\delta_0, ..., \delta_J; V(x))$, variance of $\hat{T}^*(0,1;v(x))$ as

$$\mathcal{V}_{\xi_0}[\hat{T}^*(0.1;v(x)) - Y] \ = \ \mathcal{E}_{\xi_0}[\hat{T}^*(0,1;v(x)) - Y]^2 - [\mathcal{E}_{\xi_0}\{\hat{T}^*(0,1;v(x)) - Y\}]^2$$

$$= \mathcal{E}_{\xi_0}[\sum_{\overline{S}} x_k \frac{\sum_S \frac{x_k Y_k}{v(x_k)}}{\sum_S \frac{x_k^2}{v(x_k)}} - \sum_{\overline{S}} Y_k)]^2$$

$$= (\frac{\sum_{\overline{S}} x_k}{\sum_S \frac{x_k^2}{v(x_k)}})^2 \sum_S \frac{x_k^2}{(v(x_k))^2} V(x_k) + \sum_{\overline{S}} V(x_k) \qquad (3.2.17)$$

[using lemma 3.1]

When V(x) satisfies (3.2.4)

$$\mathcal{V}_{\xi_0}[\hat{T}^*(0,1;v(x)) - Y] = \{\frac{\sum_{\overline{S}} x_k}{\sum_S \frac{x_k^2}{v(x_k)}}\}^2 \sum_S \frac{x_k^2}{(v(x_k))^2} . v(x_k) \sum_{j=0}^{J} \delta_j \beta_j x_k^{j-1}$$

$$+ \sum_{\overline{S}} v(x_k) \sum_{j=0}^{J} \delta_j \beta_j x_k^{j-1}$$

$$= \sum_{j=0}^{J} \delta_j \beta_j [\{\frac{\sum_{\overline{S}} x_k}{\sum_S \frac{x_k^2}{v(x_k)}}\}^2 \sum_S \frac{x_k^{j+1}}{v(x_k)}$$

$$+ \sum_{\overline{S}} v(x_k) x_k^{j-1}] \qquad (3.2.18)$$

which reduces under $S_{v(x)}(J)$ [using (3.2.3)] to

$$\sum_{j=0}^{J} \delta_j \beta_j [\{\frac{\sum_{\overline{S}} x_k}{\sum_S \frac{x_k^2}{v(x_k)}}\}^2 \frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}} \sum_S \frac{x_k^2}{v(x_k)} + \sum_{\overline{S}} v(x_k) x_k^{j-1}]$$

$$[\text{ Because } \frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}} = \frac{\sum_S \frac{x_k^{j+1}}{v(x_k)}}{\sum_S \frac{x_k^2}{v(x_k)}},\ j = 0, 1, ..., J]$$

$$= \sum_{j=0}^{J} \delta_j \beta_j [\frac{(\sum_{\overline{S}} x_k)^2}{\sum_S \frac{x_k^2}{v(x_k)}} \frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}} + \sum_{\overline{S}} v(x_k) x_k^{j-1}]$$

$$= \sum_{j=0}^{J} \delta_j \beta_j [\frac{(\sum_{\overline{S}} x_k)^2}{\sum_S \frac{x_k^2}{v(x_k)}} \frac{\frac{1}{(N-n)} \sum_{\overline{S}} x_k^j}{\frac{1}{(N-n)} \sum_{\overline{S}} x_k} + \sum_{\overline{S}} v(x_k) x_k^{j-1}]$$

$$= \sum_{j=0}^{J} \delta_j \beta_j (\sum_{\overline{S}} x_k) [\frac{\sum_{\overline{S}} x_k^j}{\sum_S \frac{x_k^2}{v(x_k)}} + \frac{\sum_{\overline{S}} v(x_k) x_k^{j-1}}{\sum_{\overline{S}} x_k}] \qquad (3.2.19)$$

Thus (3.2.17) gives the variance of $\hat{T}^*(0,1;v(x))$ under ξ_0; (3.2.18) the variance of $\hat{T}^*(0,1;v(x))$ under ξ_0 when V(x) satisfies (3.2.4); (3.2.19) the variance of $\hat{T}^*(0,1;v(x))$ under ξ_0 when V(x) satisfies (3.2.4) and if further the sample is $S_{v(x)}(J)$ (which satisfies (3.2.3)).

From (3.2.17), putting $v(x) = x$, $V(x) = V_{0,J} = \sum_{j=0}^{J} \lambda_j a_j x^j$, we have variance of $\hat{T}_1^*$ under $\xi_{0,J}$ for the balanced sample $S_x(J)$,

$$\mathcal{V}_{\xi_{0,J}}(\hat{T}_1^* - Y|S_x(J)) = \mathcal{V}_{\xi_{0,J}}[\hat{T}(0,1;x) - Y|S_x(J)]$$

$$= [\{\frac{\sum_{\overline{S}} x_k}{\sum_S \frac{x_k^2}{x_k}}\}^2 \sum_S \frac{x_k^2}{(x_k)^2} \sum_{j=0}^{J} \lambda_j a_j x_k^j + \sum_{\overline{S}} \sum_{j=0}^{J} \lambda_j a_j x_k^j |S_x(J)]$$

$$= [\{\frac{\sum_{\overline{S}} x_k}{\sum_S x_k}\}^2 \sum_{j=0}^{J} \lambda_j a_j \sum_S x_k^j + \sum_{j=0}^{J} \lambda_j a_j \sum_{\overline{S}} x_k^j |S_x(J)]$$

$$= [\sum_{j=0}^{J} \lambda_j a_j \{(\frac{\sum_{\overline{S}} x_k}{\sum_S x_k})^2 \sum_S x_k^j + \sum_{\overline{S}} x_k^j\} |S_x(J)]$$

$$= \sum_{j=0}^{J} \lambda_j a_j \frac{\sum_{\overline{S}} x_k}{\sum_S x_k} (\frac{\sum_{\overline{S}} x_k \overline{X}^{(j)}}{\overline{X}} + \frac{\sum_S x_k \overline{X}^{(j)}}{\overline{X}})$$

[since under $S_x(J)$, $\overline{X}^{(j)} = \overline{x}_S^{(j)} = \overline{x}_{\overline{S}}^{(j)}$]

$$= \sum_{j=0}^{J} \lambda_j a_j \frac{(N-n)}{n} \frac{\overline{x}_{\overline{S}}}{\overline{x}_S} \frac{\overline{X}^{(j)}}{\overline{X}} (\sum_{\overline{S}} x_k + \sum_S x_k)$$

$$= \frac{N(N-n)}{n} \sum_{j=0}^{J} \lambda_j a_j \overline{X}^{(j)}$$

$$= M_1(0,J) \quad \text{(say)} \qquad (3.2.20)$$

We note that in calculating the expression for variance of $(\hat{T}_2^* - Y)$ at $S_{x^2}(J)$ under the model $\xi_{1,L}$ we can not take L beyond $(J-1)$ as in those cases the conditions (3.2.2) do not apply to some terms in the explicit expression of the variance. Also we can not take $k < 1$ as in that case the result in (3.2.4) will not apply to $\hat{T}_2^*$.

Thus, using (3.2.2), variance of $\hat{T}_2^*$ under $\xi_{1,J-1}$ on $S_{x^2}(J)$ is

$$\mathcal{V}_{\xi_{1,J-1}}(\hat{T}_2^* - Y|S_{x^2}(J)) = \mathcal{V}_{\xi_{1,J-1}}(\hat{T}^*(0,1:x^2) - Y|S_{x^2}(J))$$

$$= [\{\frac{\sum_{\overline{S}} x_k}{\sum_S (x_k^2)/(x_k^2)}\}^2 \sum_S \frac{x_k^2}{(x_k^2)^2} \sum_{j=1}^{J-1} \lambda_j a_j x_k^j + \sum_{\overline{S}} \sum_{j=1}^{J-1} \lambda_j a_j x_k^j |S_{x^2}(J)]$$

$$= \sum_{j=1}^{J-1} \lambda_j a_j \{(\frac{1}{n} \sum_{\overline{S}} x_k)^2 \sum_S x_k^{j-2} + \sum_{\overline{S}} x_k^j\} |S_{x^2}(J)]$$

$$= [\sum_{j=1}^{J-1} \lambda_j a_j (\frac{1}{n} \sum_{\overline{S}} x_k) \{ (\sum_{\overline{S}} x_k) \frac{1}{n} \sum_{S} x_k^{j-2} + \frac{\sum_{\overline{S}} x_k^j}{\frac{1}{n} \sum_{\overline{S}} x_k} \} | S_{x^2}(J)]$$

$$= \sum_{j=1}^{J-1} \lambda_j a_j (\frac{1}{n} \sum_{\overline{S}} x_k) \{ (N-n) \overline{x}_{\overline{S}}^{(j-1)} + n \overline{x}_S^{(j-1)} \}$$

$$[\text{since under } S_{x^2}(J),\ \overline{x}_S^{(j-1)} = \frac{\overline{x}_{\overline{S}}^{(j)}}{\overline{x}_{\overline{S}}}, \quad j = 0, 1, ..., J]$$

$$= \sum_{j=1}^{J-1} \lambda_j a_j \frac{(N-n)}{n} \overline{x}_{\overline{S}} (\sum_{i=1}^{N} x_i^{j-1})$$

$$= \frac{N(N-n)}{n} \overline{x}_{\overline{S}} \sum_{j=1}^{J-1} \lambda_j a_j \overline{X}^{(j-1)}$$

$$= M_2(1, J-1) \quad \text{(say)} \qquad (3.2.21)$$

Since under $S_{x^2}(J)$, $\overline{x}_{\overline{S}} \leq \overline{X}$ [follows from (3.2.2) for j=0, because, $\bar{x_S}^{(-1)} = \frac{1}{\bar{x_S}}$; again, as $A.M. \geq H.M.$, $\bar{x_S}^{(-1)} \geq \frac{1}{\bar{x_S}}$]

$$M_2 \leq \frac{N(N-n)}{n} \overline{X} \sum_{j=1}^{J-1} \lambda_j a_j \overline{X}^{(j-1)}$$

We shall now compare $M_1(1, J-1)$ and $M_2(1, J-1)$. Thus,

$$M_1(1, J-1) - M_2(1, J-1) \geq \frac{N(N-n)}{n} \sum_{j=1}^{J-1} \lambda_j a_j [\overline{X}^{(j)} - \overline{X}\overline{X}^{(j-1)}]$$
$$\geq 0 \text{ if } x_k \geq 0 \qquad (3.2.22)$$

Hence we have

Theorem 3.3 : Under $\xi_{1,J-1}$, $\hat{T}_1^*$ is BLUP at $S_x(J)$; Similarly under $\xi_{1,J-1}$, $\hat{T}_2^*$ is BLUP at $S_{x^2}(J)$. However, $\hat{T}_2^*$ at $S_{x^2}(J)$ is a better strategy than $\hat{T}_1^*$ at $S_x(J)$, $(H_2 \succ H_1)$, provided $x_k \geq 0 \ \forall\ k$.

Note : Scott, Brewer and Ho [SBH] (1978) showed that under ξ_0, where $V(x) = \sigma_1^2 x + \sigma_2^2 x^2$ both $(\hat{T}_1^*, S_x(J)) = H_1$ and $(\hat{T}_2^*, S_{x^2}(J)) = H_2$ are BLU-strategies. However

$$\mathcal{E}_\xi\{(\hat{T}_1^* - Y)^2 | S_x(J)\} = \frac{N(N-n)}{n} (\sigma_1^2 \overline{X} + \sigma_2^2 \overline{X}^{(2)})$$
$$= M_1' \quad \text{(say)}$$

$$
\begin{aligned}
\mathcal{E}_\xi\{(\hat{T}_2^* - Y)^2 | S_{x^2}(J)\} &= \frac{N(N-n)}{n}\overline{x}_{\overline{S}}(\sigma_1^2 + \sigma_2^2\overline{X}) \\
&= M_2' \quad \text{(say)} \\
&< \frac{N(N-n)}{n}\overline{X}(\sigma_1^2 + \sigma_2^2\overline{X}),
\end{aligned}
$$

using (3.2.2) for j=0.

Thus

$$
M_1' - M_2' \geq \frac{N(N-n)}{n}\sigma_2^2(\overline{X}^{(2)} - \overline{X}^2) \geq 0,
$$

showing H_1 is less efficient than H_2, the loss in efficiency will be small in general if σ_1^2 dominates σ_2^2 but can be substantial if σ_2^2 is relatively large.

The result in *theorem 3.3* pursues the result of SBH to polynomial regression models with variance function given as higher degree polynomials in the regressor variable.

Next we consider model variance of Horvitz-Thompson predictor $\hat{T}_{HT} = X\sum_S \frac{Y_k}{nx_k}$ where $\pi_i \propto x_i$, on a π-balanced sample under model $\xi_{1,J}$.

We have under $\xi_{0,J}$, for a πps-design on a sample $S_\pi(J)$,

$$
\begin{aligned}
\mathcal{V}_{\xi_{0,J}}(\hat{T}_{HT} - Y | S_\pi(J)) &= \mathcal{E}_{\xi_{0,J}}\{(\hat{T}_{HT} - Y)^2 | S_\pi(J)\} \\
&\quad -\{\mathcal{E}_{\xi_{0,J}}(\hat{T}_{HT} - Y)|S_\pi(J)\}^2 \\
&= [\sum_{j=0}^{J}\lambda_j a_j\{\frac{X^2}{n^2}\sum_S \frac{x_k^j}{x_k^2} + \sum_{k=1}^{N} x_k^j \\
&\quad -\frac{2}{n}\sum_{k=1}^{N} x_k \sum_S \frac{x_k^j}{x_k}\}|S_\pi(J)] \\
&= [\sum_{j=0}^{J}\lambda_j a_j\{\frac{X^2}{n^2}\sum_S x_k^{j-2} + \sum_{k=1}^{N} x_k^j \\
&\quad -\frac{2X}{n}\sum_S x_k^{j-1}\}|S_\pi(J)]
\end{aligned}
$$

Using the condition of π-balancing ie.,

$$
\frac{X}{n}\sum_S x_k^{j-1} = \sum_{k=1}^{N} x_k^j,
$$

in the above expression we get

$$
\mathcal{V}_{\xi_{0,J}}[(\hat{T}_{HT} - Y)|S_\pi(J)] = \sum_{j=0}^{J}\lambda_j a_j\{\frac{X}{n}\sum_{k=1}^{N} x_k^{j-1}
$$

$$
\begin{aligned}
&+\sum_{k=1}^{N} x_k^j - 2\sum_{k=1}^{N} x_k^j\} \\
&= \sum_{j=0}^{J} \lambda_j a_j \{\frac{X}{n}\sum_{k=1}^{N} x_k^{j-1} - \sum_{k=1}^{N} x_k^j\} \\
&= N^2 \sum_{j=0}^{J} \lambda_j a_j \{\frac{1}{n}\overline{X}\overline{X}^{(j-1)} - \frac{1}{N}\overline{X}^{(j)}\} \\
&= M_H(0, J) \quad \text{(say)} \qquad (3.2.23)
\end{aligned}
$$

Since the similar expressions as given in (3.2.21) for variance of $\hat{T}_2^*$ hold only if the variance function in the model $\xi(\delta_0, ..., \delta_J; V(x))$ is a polynomial of maximum order $(J-1)$, we shall consider $M_H(1, J-1)$ for comparing with $M_2(1, J-1)$.

We have therefore from (3.2.21) and (3.2.23)

$$
\begin{aligned}
M_H(1, J-1) - M_2(1, J-1) &= \sum_{j=1}^{J-1} \lambda_j a_j [\frac{N^2}{n}\overline{X}\overline{X}^{(j-1)} \\
&\quad - N\overline{X}^{(j)} - \frac{N(N-n)}{n}\overline{x}_{\overline{S}}\overline{X}^{(j)}] \\
&= \sum_{j=1}^{J-1} \lambda_j a_j [\sum_{k=1}^{N} x_k^{j-1} \frac{1}{n}\sum_{S} x_k \\
&\quad - \sum_{k=1}^{N} x_k^j] \qquad (3.2.24)
\end{aligned}
$$

where S refers to $S_{x^2}(J)$.

Now at $S_{x^2}(J)$,

$$
N\overline{X}^{(j)} = X\overline{x}_S^{(j-1)} + \sum_S x_k^j - \frac{1}{n}\sum_S x_k \sum_S x_k^{j-1}
$$

$$
[\text{since, } \sum_{\overline{S}} x_k^j = \sum_{\overline{S}} x_k . \overline{x}_S^{(j-1)}]
$$

Thus (3.2.24) simplifies to

$$
\sum_{j=1}^{J-1} \lambda_j a_j [\frac{1}{n}\sum_{\overline{S}} x_k \{\overline{x}_S \sum_S x_k^{j-2} - \sum_S x_k^{j-1}\} + \{\sum_S x_k^j - \overline{x}_S \sum_S x_k^{j-1}\}]
$$

$$
= \sum_{j=1}^{J-1} \lambda_j a_j [(\sum_{\overline{S}} x_k) Cov.(x_k, x_k^{j-2}) + nCov.(x_k, x_k^{j-1})] \qquad (3.2.25)
$$

If we consider, therefore, the model $\xi_{2,J-1}$ where $\lambda_1 = 0$ then under conditions $x_k \geq 0 \; \forall \; k = 1,...,N$, (3.2.25) is non-negative, since a_j's are non-negative constants.

Hence we have

Theorem 3.4 : Under $\xi_{2,J-1}$, $\hat{T}_2^*$ is BLUP at $S_{x^2}(J)$. However, $\hat{T}_{HT}$ at $S_\pi(J)$ is a better strategy than $\hat{T}_2^*$ at $S_{x^2}(J)$ in the sense of having smaller average variance, $(H_3 \succ H_2)$ provided $x_k \geq 0 \; \forall \; k$.

Combining theorem 3.3 and theorem 3.4 we have

Theorem 3.5 : Under $\xi_{2,J-1}$, $H_3 \succ H_2 \succ H_1$, provided $x_k \geq 0 \; \forall \; k$.

Remarks 3.1 We have compared above performance of the predictors $\hat{T}_1^*$, $\hat{T}_2^*$ and $\hat{T}_{HT}$ at the sampling designs where their performances are optimal, viz., Royall- Herson (1973) balanced sampling design for $\hat{T}_1^*$, Scott et al (1978) over-balanced sampling design for $\hat{T}_2^*$ and πps-balanced sampling design of Royall and Cumberland under the model $\xi_{2,J-1}$. The model condition $\xi_{2,J-1}$ have been invoked since under this model $\hat{T}_1^*$ is BLUP at $S_x(J)$ and $\hat{T}_2^*$ is BLUP at $S_{x^2}(J)$. Again on a π-balanced sample $S_\pi(J)$, Horvitz-Thompson predictor $\hat{T}_{HT}$ is model-unbiased under $\xi_{2,J-1}$. Thus under $\xi_{2,J-1}$, $(\hat{T}_1^*, S_x(J)) = H_1$, $(\hat{T}_2^*, S_{x^2}(J)) = H_2$, are optimal predictors. Also under $\xi_{2,J-1}$, $H_3 = (\hat{T}_{HT}, S_\pi(J))$ is model-unbiased though $\hat{T}_{HT}$ is not known to have any optimal property under $\xi_{2,J-1}$ or any subclass of it. The result in Theorem 3.5 shows that if $x_k \geq 0 \; \forall \; k$, a condition which generally holds , H_3 is a better strategy than H_2 and H_2 is again a better strategy than H_1. A πps design provides π-balanced samples on an average. The model $\xi_{2,J-1} = \xi(\delta_0, ..., \delta_J; V(x) = \sigma^2 \sum_2^{J-1} a_j \lambda_j x^j)$ where λ_j's satisfy (3.2.6b) is a very general polynomial regression model and covers a wide class of socio-economic situations. Thus under very general situations a $(\hat{T}_{HT}, \pi ps)$-design is expected to provide more precise prediction strategies than model-dependent optimum strategies $(\hat{T}_1^*, S_x(J))$ and $(\hat{T}_2^*, S_{x^2}(J))$. Moreover the samples $S_x(J)$ and $S_{x^2}(J)$ are often conspicuous by their absence.

3.2.6 A Suggested Study

The following extension to above investigations is incorporated as suggested by an examiner. Let $\xi(\delta_0, ..., \delta_J, V(x)) = \xi^{(J)}(V(x))$ (say).

Let for $g = 1, 2, ...$ the strategy H_g consist of $\hat{T}_g^* = \hat{T}^*(0, 1, x^g)$ based on a balanced sample satisfying (3.2.3) for $v(x) = x^g$ i.e., a g-balanced sample which satisfies

$$\frac{\bar{x_S}^{(j)}}{\bar{x_S}} = \frac{\sum_S x_k^{j-g+1}}{\sum_S x_k^{2-g}} \qquad (3.2.3a)$$

$$mboxwhere \quad \hat{T}_g^* = \hat{T}^*(0,1,x^g) = \sum_S y_k + \frac{\sum_S y_k x_k^{1-g}}{\sum_S x_k^{2-g}} \sum_{\bar{S}} x_k$$

Under model $\xi_{(h)} = \xi^{(J)}(x^h), h = g-1, ..., J+g-1$ it follows from (3.2.17) along with (3.2.3) that for $g-1 \leq J$, the model mean square error of H_g

$$\begin{aligned}
\mathcal{E}MSE(H_g) &= \mathcal{E}_{\xi_{(h)}}\{(\hat{T}_g^* - Y)^2 | S_{x^g}(J)\} \\
&= \frac{\sum_{\bar{S}} x_k}{\sum_S \frac{x_k^2}{x_k^g}} \sum_S \frac{x_k^2}{(x_k^g)^2} x_k^h + \sum_{\bar{S}} x_k^h \\
&= \{\frac{\sum_{\bar{S}} x_k}{\sum_S x_k^{2-g}}\}^2 \sum_S x_k^{h-2g+2} + \sum_{\bar{S}} x_k^h \\
&= \frac{\sum_{\bar{S}} x_k}{\sum_S x_k^{2-g}} [\sum_{\bar{S}} x_k^{h-g+1} + \sum_{\bar{S}} x_k^h \frac{\sum_S x_k^{2-g}}{\sum_{\bar{S}} x_k}] \text{using (3.2.3a)} \\
&= \frac{\sum_{\bar{S}} x_k}{\sum_S x_k^{2-g}} [\sum_{\bar{S}} x_k^{h-g+1} + \sum_S x_k^{h-g+1}] \text{using (3.2.3a)} \\
&= \frac{\sum_{\bar{S}} x_k}{\sum_S x_k^{2-g}} X^{(h-g+1)} \\
&= \frac{\sum_{\bar{S}} x_k^{g-1}}{n} X^{(h-g+1)} \text{(using (3.2.3a) for j=g-1)} \\
&= \frac{(N-n)}{n} N \bar{x_{\bar{S}}}^{(g-1)} \bar{X}^{h-g+1}
\end{aligned}$$

Now, for $j = g-1$, (3.2.3a) becomes

$$\frac{\bar{x_{\bar{S}}}^{(g-1)}}{\bar{x_{\bar{S}}}} = \frac{n}{\sum_S x_k^{2-g}} < \bar{x_S}^{(g-2)} \text{(since A.M. } \geq \text{ H.M.)}$$

Since $Cov(x_k, x_k^{g-2}) > 0$ for $g > 2$ (g is an integer), we have

$$\bar{x_S}^{(g-1)} > \bar{x_S}\bar{x_S}^{(g-2)}$$

and

$$\begin{aligned}
\bar{X}^{(g-1)} &= \frac{n}{N}\bar{x_S}^{(g-1)} + \frac{N-n}{N}\bar{x_{\bar{S}}}^{(g-1)} \\
&> \frac{n}{N}\bar{x_S}\bar{x_S}^{(g-2)} + \frac{N-n}{N}\bar{x_{\bar{S}}}^{(g-1)}
\end{aligned}$$

$$
\begin{aligned}
&> \frac{n}{N}\bar{x_S}\frac{\bar{x_{\bar{S}}}^{(g-1)}}{\bar{x_{\bar{S}}}} + \frac{N-n}{N}\bar{x_{\bar{S}}}^{(g-1)} \\
&> \bar{x_{\bar{S}}}^{(g-1)}\{\frac{n\bar{x_S} + (N-n)\bar{x_{\bar{S}}}}{N\bar{x_{\bar{S}}}}\} \\
&= \bar{x_{\bar{S}}}^{(g-1)}\frac{\bar{X}}{\bar{x_{\bar{S}}}} \qquad (3.2.26)
\end{aligned}
$$

For $j = 0$, (3.2.3a) becomes

$$\frac{1}{\bar{x_{\bar{S}}}} = \frac{\sum_S x_k^{1-g}}{\sum_S x_k^{2-g}}$$

Now, $Cov(x_k, x_k^{1-g})$ is negative for $g > 2$ and hence

$$\frac{1}{n}\sum_S x_k^{2-g} < \frac{1}{n}\sum_S x_k^{1-g}\frac{1}{n}\sum_S x_k$$

$$\text{or,} \frac{1}{\sum_S x_k^{2-g}} > \frac{1}{\frac{1}{n}\sum_S x_k^{1-g}\sum_S x_k}$$

$$\text{therefore, } \frac{1}{\bar{x_{\bar{S}}}} = \frac{\sum_S x_k^{1-g}}{\sum_S x_k^{2-g}} > \frac{\sum_S x_k^{1-g}}{\sum_S x_k^{1-g}.\bar{x_S}} = \frac{1}{\bar{x_S}}$$

$$\text{or, } \bar{x_S} > \bar{x_{\bar{S}}}$$

$$\Rightarrow \bar{X} > \bar{x_{\bar{S}}}$$

Using this in (3.2.26) we get, $X^{(\bar{g}-1)} \geq \bar{x_{\bar{S}}}^{(g-1)}$

Therefore under $\xi_{(h)}$ $(g-1 \leq h \leq J)$

$$
\begin{aligned}
\mathcal{E}MSE(H_g) - \mathcal{E}MSE(H_1) &= \frac{N(N-n)}{n}\bar{x_{\bar{S}}}^{(g-1)}\bar{X}^{(h-g+1)} \\
&\quad -\frac{N(N-n)}{n}\bar{X}^{(h)} \\
&\leq \frac{N(N-n)}{n}\{\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} \\
&\quad -\bar{X}^{(h)}\} \quad \text{(by (3.2.26))} \\
&= -\frac{N(N-n)}{n}Cov(x^{(g-1)}, x^{(h-g+1)}) \\
&\leq 0 \\
\text{Hence, } \mathcal{E}MSE(H_g) &\leq \mathcal{E}MSE(H_1)
\end{aligned}
$$

Since under $\xi^*_{K,L} = \xi^{(J)}(\sum_K^L a_h x^h)$ for non-negative constants a_h's,

$$\mathcal{E}_{\xi^*_{K,L}}MSE(H_g) = \sum_{h+K}^{L} a_h \mathcal{E}_{\xi_{(h)}}MSE(H_g)$$

it follows that under the model $\xi^*_{g-1,J}$,

$$\mathcal{E}MSE(H_g) \le \mathcal{E}MSE(H_1) \quad \text{for } g \ \ge \ 1$$

Next we consider the projection-type predictor

$$\hat{T}_g = \hat{\beta}X = \hat{T}(0,1,x^g) = X\frac{\sum_S y_i x_i/x_i^g}{\sum_S x_i^2/x_i^g} = X(\sum_S y_i x_i^{1-g})/(\sum_S x_i^{2-g})$$

Under the alternative model $\xi^{(J)}(x_h)$ to achieve unbiasedness $\hat{T}_g$ has to satisfy the condition of balancing i.e.,

$$\begin{aligned}
\mathcal{E}_{\xi^{(J)}}(\hat{T}_g - Y) &= X\frac{\sum_S \beta x_i^j . x_i^{1-g}}{\sum_S x_i^{2-g}} - \beta\sum_{i=1}^{N} x_i^j \\
&= \beta(\frac{X}{\sum_S x_i^{2-g}}\sum_S x_i^{j+1-g} - \sum_{i=1}^{N} x_i^j) \\
&= 0
\end{aligned}$$

$$\text{i.e.,} \ \frac{X}{\sum_S x_i^{2-g}}\sum_S x_i^{j-g+1} = \sum_{i=1}^{N} x_i^j$$

$$\text{For} \quad j = g-1, \quad \frac{X}{\sum_S x_i^{2-g}} = \frac{\sum_{i=1}^{N} x_i^{g-1}}{n}$$

and hence the predictor can be written in the form

$$\hat{T}_g = \frac{N}{n}\bar{X}^{(g-1)}\sum_S \frac{y_k}{x_k^{g-1}}$$

The strategy consisting of the predictor $\hat{T}_g$ at a properly balanced sample (i.e., a sample satisfying $\bar{X}^{(g-1)}\bar{x}_S{}^{(j-g+1)} = \bar{X}^{(j)}, (j = 0,1,...,J)$ is denoted by H_g^*. Now,

$$\begin{aligned}
\mathcal{E}MSE(H_g^*) &= \mathcal{E}_{\xi_{(h)}}\{(\hat{T}_g - Y)^2 | S^*_{x^g}(J)\} \\
&= \sum_S x_i^h(\frac{N}{n}\frac{\bar{X}^{(g-1)}}{x_i^{g-1}} - 1)^2 + \sum_{\bar{S}} x_i^h \\
&= \frac{N^2}{n^2}(\bar{X}^{(g-1)})^2\sum_S x_i^{h-2g+2} \\
&\quad -2\frac{N}{n}\bar{X}^{(g-1)}\sum_S x_i^{h-g+1} + \sum_{i=1}^{N} x_i^h \\
&= \frac{N^2}{n^2}(\bar{X}^{(g-1)})^2\sum_S x_i^{h-2g+2} - \sum_{i=1}^{N} x_i^h \\
&= \frac{N^2}{n}\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} - N\bar{X}^{(h)}
\end{aligned}$$

Hence,

$$\mathcal{E}MSE(H^*_{g'}) - \mathcal{E}MSE(H^*_g) = \frac{N^2}{n}(\bar{X}^{g'-1}\bar{X}^{h-g'+1} - \bar{X}^{g-1}\bar{X}^{h-g+1})$$

Suppose that $g' > g$. Using Callebauts' inequality for $A_i = x_i, B_i = 1, w = h/2$ and $z = g-1-h/2$ we see that the quantity $\phi = \bar{X}^{g-1}\bar{X}^{h-g-1}$ is increasing for increasing value of $|g-1-h/2|$. Now as $h > g' + g - 2$, therefore, $h > 2g-2$ (since $g' > g$), i.e., $h/2 > g-1 \Rightarrow h/2-g+1 > 0$ which implies that the value of ϕ is monotonically increasing with the increasing value of the difference $h/2-g+1$ Since h is a constant therefore $h/2-g+1$ increases as g decreases. Therefore $\bar{X}^{(g-1)}\bar{X}^{(h-g-1)}$ increases as g decreases and hence under the model $\xi_h(g'+g-2 \le h \le J+g-1)$and consequently under $\xi^*_{g'+g-2,J+g-1}$

$$\mathcal{E}MSE(H^*_{g'}) - \mathcal{E}MSE(H^*_g) < 0.$$

Again,

$$\begin{aligned}
\mathcal{E}MSE(H^*_g) - \mathcal{E}MSE(H_g) &= \frac{N^2}{n}\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} - N\bar{X}^{(h)} \\
&\quad -\frac{N(N-n)}{n}\bar{x_S}^{(g-1)}\bar{X}^{(h-g+1)} \\
&\le \frac{N^2}{n}\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} - N\bar{X}^{(h)} \\
&\quad -\frac{N(N-n)}{n}\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} \\
&= N(\bar{X}^{(g-1)}\bar{X}^{(h-g+1)} - \bar{X}^{(h)}) \\
&= -NCov(x^{g-1}, x^{h-g+1}) \\
&\le 0 \text{ for } 2(g-1) \le h \le J+g-1
\end{aligned}$$

$$\Rightarrow \mathcal{E}MSE(H^*_g) \le \mathcal{E}MSE(H_g)$$

Chapter 4

SAMPLING STRATEGIES FOR ESTIMATING A FINITE POPULATION VARIANCE

4.1. Introduction and Review of earlier work

In the theory of sampling from finite population, discussions have usually been centred on the estimation of population total or mean. The problem of estimating a finite population variance has received comparatively poor attention. The estimation of population variance is of considerable importance in many circumstances. The geneticsts often classify their population according to population variance [Thompson and Thoday (1979)]. In allocating sample size in a stratified random sampling according to optimum allocation rules, the stratum standard deviations are required to be estimated.

The problem of estimation of a finite population variance was first considered by Liu (1974a). He derived a Horvitz-Thompson (HT)-type estimator of variance and examined its forms under simple random sampling, with and without replacement and probability proportional to size with replacement (ppswr) sampling procedures. He showed his estimator to be admissible in the class of all unbiased estimators for the population variance and also constructed an admissible general unbiased quartic estimator for the variance of any unbiased quadratic estimator for the variance. Since a variance is a non-negative quantity, it is desirable that any estimator of the same should also be non-negative. Observing that Liu's estimator can sometimes take negative values, Chaudhury (1978) suggested non-negative alternative estimators and noted some of their properties. Das and Tripathi (1978) obtained the ratio-type and product-type estimators of variance, when the population mean or variance or co-efficient of variation of an auxiliary character is known, under simple random sampling with replacement (srswr) and studied their properties under large sample. Following Olkin (1958), Isaki (1983) considered multivariate ratio and regression estimators of variance under a multivariate normal set-up. Some other works on variance are due to Liu (1974b), Mukhopadhyay (1978, 1982, 1984, 1990), Chang and Lin (1985), Rodrigues et al (1985), Sankarnarayanan (1980), Zacks and Solomon (1981), Skinner (1981, 1983), Strauss (1982), Ghosh and Meeeden (1983), Liu and Thompson (1983), Singh (1983), Vardeman and Meeden (1983) and Sengupta (1988). Some of their works concentrate mainly on estimating (predicting) the variance under assumptions of some superpopulation models, following the prediction-theoretic approach of Royall (1970), Royall and Herson (1973).

4.2 Summary

Following Godambe(1955),Lanke(1975),and Basu(1971)we have proved some non -existence theorems for the population variance. We have proved that for any given sampling design p with second-order inclusion-probability $\pi_{ij} > 0, (\forall i \neq j = 1, ..., N)$,there does not exist any uniformly minimum variance quadratic unbiased estimator(UMVQUE) of the finite population variance $V(\underset{\sim}{y})$. It is also proved that for any non-census design there does not exist any UMVQUE of $V(\underset{\sim}{y})$ in the entire class of all unbiased estimators (section 4.3).

In section 4.4 we suggest a non-negative unbiased estimator of $V(\underset{\sim}{y})$ which is applicable to any fixed size without replacement sampling design $p \in \rho_n$. The variance and estimator of this variance have been obtained. We consider, in particular, estimator of $V(\underset{\sim}{y})$ using srswor design, design due to Lahiri (1951)-Sen(1952)-Midzuno(1951) [Midzuno scheme, in short] and design due to Singh and Srivastava(1980).Estimation of $V(\underset{\sim}{y})$ under controlled sampling plan and ppswr-sampling design have also been considered. The performance of several strategies have been studied empirically.

Finally we considered a superpopulation distribution for the main variable y and a gamma distribution for the auxiliary variable x and obtained expressions for the average variances of several strategies. These we compared for various values of the model parameters.

4.3. Some Non-existence Results

4.3.1 Some Preliminaries

Let as before $\mathcal{U}$ denote a finite population of N identifiable units labelled $1, 2, .., i, .., N$; y_i is the value of the characteristic y on unit i in the population . Our problem is that of estimating the finite population variance

$$\begin{aligned} V(\underset{\sim}{y}) &= \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{Y})^2. \\ &= a_1\sum_{i=1}^{N} y_i^2 - a_2\sum_{i\neq j=1}^{N}\sum y_i y_j \qquad (4.3.1) \end{aligned}$$

where

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$
$$a_1 = \frac{1}{N}(1-\frac{1}{N})$$
$$a_2 = \frac{1}{N^2}$$

by a sample survey where a sample S(s) is selected with probability p(S)[p(s)] according to a sampling plan and an estimator $v(S, \underset{\sim}{y})$ is employed for the purpose.

A statistic $v(S, \underset{\sim}{y})$ is a homogeneous quadratic unbiased estimator (QUE) of V if

$$v(S, \underset{\sim}{y}) = \sum_{i \in S} b_{Sii} y_i^2 + \sum\sum_{i \neq j \in S} b_{Sij} y_i y_j \qquad (4.3.2)$$

and

$$E(v(S, \underset{\sim}{y})) = V(\underset{\sim}{y}) \ \forall \ \underset{\sim}{y} \in R_N$$

i.e.,

$$\sum_{i=1}^{N} y_i^2 \sum_{S \ni i} b_{Sii} p(S) + \sum\sum_{i \neq j=1}^{N} \sum_{S \ni i,j} b_{sij} p(S) = a_1 \sum_{i=1}^{N} y_i^2 - a_2 \sum\sum_{i \neq j=1}^{N} y_i y_j$$
$$\forall \ \underset{\sim}{y} \in R_N$$

Here b_{Sii}, b_{Sij} are constants that may depend on S and units i and (ij), but not on y-values. A set of necessary and sufficient conditions for (4.3.2) to hold is :

$$\sum_{S \ni i} b_{Sii} p(S) = a_1 \qquad (4.3.3a)$$
$$\sum_{S \ni i,j} b_{Sij} p(S) = a_2 \qquad (4.3.3b)$$

As a special case of this , there is an estimator of Horvitz-Thompson(HT) type , suggested by Liu(1974),

$$v_{HT}(S, \underset{\sim}{y}) = a_1 \sum_{i \in S} \frac{y_i^2}{\pi_i} - a_2 \sum\sum_{i \neq j \in S} \frac{y_i y_j}{\pi_{ij}} \qquad (4.3.4)$$
$$= v_1 \text{ (say)}$$

which is unbiased for $V(\underset{\sim}{y})$ for any p. We note that there can not exist any non-homogeneous QUE

$$v' = b_0 + \sum_{i \in S} b_{Sii} y_i^2 + \sum_{i \neq j \in S}\sum b_{Sij} y_i y_j$$

because if $y_1 = \cdots = y_N = 0$, then $V(\underset{\sim}{y}) = 0$, $E(v') = b_0$ should be zero. Hence we consider only homogeneous QUE.

4.3.2 Main Result

Godambe(1955) first observed that in survay sampling no UMV-estimator in the class of all linear unbiased estimators of population total exists for any p in general. The proof was subsequently improved upon by Hege(1965), Hanurav(1966), Ericson(1974) and Lanke(1974).

In this section following Lanke's(1975) proof for non-existence of UMVU-estimator of population total we note a non-existence theorem (theorem 4.3) for estimating the population variance.

The following lemma follows from Hanurav (1966), Lanke (1975).

Lemma 4.1 : For any given design p, there is at least one unbiased estimator of $V(\underset{\sim}{y})$ iff $\pi_{ij} > 0$ $(i \neq j = 1, 2, \cdots N)$.

Following Liu (1974a) and Lanke (1975) we note the following theorem for estimating variance.

Theorem 4.1 : Let p be any sampling design with $\pi_{ij} > 0 \forall\, i \neq j = 1, 2, \cdots N$. If there exists a UMVQUE of $V(\underset{\sim}{y})$ then that must be v_1.

Following Lanke (1975) we prove the following.

Theorem 4.2 : Let p be any sampling design with $\pi_{ij} > 0 \ \forall\ i \neq j = 1, 2, \cdots, N$. Then v_1 is not UMV in the class of all homogeneous QUE of V.

Proof : We recall that a s.d. p is a unicluster design (UCD) if for any two samples

$$S_1,\ S_2,\ S_1 \neq S_2,\ p(S_1) > 0,\ p(S_2) > 0 \Rightarrow S_1 \bigcap S_2 = \phi.$$

Hence for a UCD, $\pi_{ij} = 0$ for many pairs (i,j). Therefore by lemma 4.1 if p is a unicluster design, no unbiased estimator of V is available. Hence p is a non-UCD. Therefore there must be at least two non-disjoint and non-identical samples S_1 and S_2 with $p(S_1) > 0$ and $p(S_2) > 0$. Suppose $i \in S_1 \bigcap S_2$ and $j \in S_1 \bigcap S_2^c$.

Let the estimator v_0 be defined as

$$v_0(S, \underset{\sim}{y}) = \begin{cases} v_1 - p(S_2) y_i^2 & \text{for} S = S_1 \\ v_1 + p(S_1) y_i^2 & \text{for } S = S_2 \\ v_1 & \text{for } S \neq S_1, S_2. \end{cases}$$

Hence v_0 is QUE of $V(\underset{\sim}{y})$.

Let us consider $\underset{\sim}{y} = \underset{\sim}{y}^{(2)} = (0, 0, ..., y_i, 0, ..., 0, y_j, 0, ..., 0)$.

For this point,

$$
\begin{aligned}
V(v_1) - V(v_0) &= E(v_1^2) - E(v_0^2) \\
&= -y_i^2 p(S_1)p(S_2)[y_i^2\{p(S_1) + p(S_2)\} \\
&\quad -2a_1\frac{y_j^2}{\pi_j} + 4a_2\frac{y_i y_j}{\pi_{ij}}]
\end{aligned} \tag{4.3.10}
$$

(4.3.10) can be made positive for some values of y_i, y_j. As for example, for $N = 10, y_i = 10, y_j = 100, \pi_j = .3, \pi_{ij} = .2, p(S_1) = .1, p(S_2) = .15, V(v_1) - V(v_0) = 8662.5 > 0$.

From (4.3.10) it can be observed that $V(v_0)$ can be smaller than $V(v_1)$ for certain values of y_i and y_j.

Hence v_1 is not UMV.

From theorem 4.1 and 4.2 we obtain the following non-existence theorem.

Theorem 4.3 : For any given p with $\pi_{ij} > 0 \ \forall \ i \neq j = 1, 2, \cdots, N$, there does not exist any uniformly minimum variance quadratic unbiased estimator (UMVQUE) of V.

In the next part the non-existence of a UMV estimator for V in C_u, the class of all unbiased estimators, is noted following Basu's(1971) result for population mean. The result is also contained in Liu and Thompson (1983).

Theorem 4.4 : Let p be any non-census design with $\pi_{ij} > 0$ $(i \neq j = 1, 2, \cdots, N)$. Then no UMV estimator exists in the class C_u of all unbiased estimators of V.

4.4 On Some Sampling Strategies for estimating the Population Variance

4.4.1 A non-negative unbiased estimator

In this section we try to develop suitable sampling strategies for estimati ng the finite population variance. Since a variance is a non-negative quantity , it is desirable that its estimator should also be non-negative. We assume that the values x_k of an auxiliary variable x are available for all the units in the population.

let

$$S_y^2 = \frac{N}{N-1}V(\underset{\sim}{y})$$

$$= \frac{1}{(N-1)}\sum_{i=1}^{N}(y_i-\overline{Y})^2$$

$$= a_1^{'}\sum_{i=1}^{N} y_i^2 - a_2^{'}\sum\sum_{i\neq j=1}^{N} y_i y_j \qquad (4.4.1)$$

where

$$a_1^{'} = \frac{1}{N}, a_2^{'} = \frac{1}{N(N-1)}, \overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i.$$

From (4.4.1), an unbiased estimator of S_y^2 is, writing $M_i = \binom{N-i}{n-i}$

$$t = a_1^{'}\sum_{i\in S}\frac{y_i^2}{M_1 p(S)} - a_2^{'}\sum\sum_{i\neq j\in S}\frac{y_i y_j}{M_2 p(S)}$$

$$= \frac{s_y^2}{M_0 p(S)} \qquad (4.4.2)$$

where

$$s_y^2 = \frac{1}{n}\sum_{i\in S}(y_i-\overline{y})^2,\ \overline{y} = \frac{1}{n}\sum_{i\in S} y_i.$$

(4.4.2) gives a non-negative unbiased (nnu) estimator of S_y^2 and is applicable to any fixed size(n-) without replacement design $p \in \rho_n$. For any fixed size (n-) without replacement sampling design $p \in \rho_n$, with positive probabilities of selection to all possible M_0 samples, the estimator (4.4.2) gives a non-negative unbiased estimator of S_y^2. The variance of t is

$$V(t) = \frac{1}{M_0^2}\sum_S \frac{s_y^4}{p(S)} - S_y^4 \qquad (4.4.3)$$

where $\sum_S$ denotes over all M_0-samples belonging to the sample space S and $s_y^4 = (s_y^2)^2$, $S_y^4 = (S_y^2)^2$. An unbiased variance estimator is

$$v(t) = t^2 - est.(S_y^4)$$

now,

$$S_y^4 = \{\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\overline{Y})^2\}^2$$

$$= \frac{1}{N^2}\sum y_i^4 + \frac{1}{N^2}(1+\frac{2}{(N-1)^2})\sum{}^{'} y_i^2 y_j^2$$

$$-\frac{4}{N^2(N-1)}\sum{}^{'} y_i^3 y_j - \frac{2(N-3)}{N^2(N-1)}\sum{}^{''} y_i^2 y_j y_k$$

$$+\frac{1}{N^2(N-1)^2}\sum{}^{'''} y_i y_j y_k y_l \qquad (4.4.4)$$

where $\sum, \sum', \sum'', \sum'''$ denote respectively $\sum_{i=1}^{N}$, $\sum\sum_{i\neq j=1}^{N}$, $\sum\sum\sum_{i\neq j\neq k=1}^{N}$, $\sum\sum\sum\sum_{i\neq j\neq k\neq l=1}^{N}$.

Hence

$$\begin{aligned} v(t) &= \frac{1}{[M_0 p(S)]^2} s_y^4 - \frac{1}{N^2}\sum_S \frac{y_i^4}{M_1 p(S)} \\ &-\frac{1}{N^2}\left(1+\frac{2}{(N-1)^2}\right)\sum_S{}' \frac{y_i^2 y_j^2}{M_2 p(S)} \\ &+\frac{4}{N^2(N-1)}\sum_S{}' \frac{y_i^3 y_j}{M_2 p(S)} \\ &+\frac{2(N-3)}{N^2(N-1)}\sum_S{}'' \frac{y_i^2 y_j y_k}{M_3 p(S)} - \frac{1}{N^2(N-1)^2}\sum_S{}''' \frac{y_i y_j y_k y_l}{M_4 p(S)} \qquad (4.4.5) \end{aligned}$$

where $\sum_S, \sum_S', \sum_S'', \sum_S'''$ denote respectively $\sum_{i\in S}$, $\sum\sum_{i\neq j\in S}$, $\sum\sum\sum_{i\neq j\neq k\in S}$, $\sum\sum\sum\sum_{i\neq j\neq k\neq l\in S}$.

A. general unbiased estimator of a polynomial function of population values was considered by Nanjamma et al (1959).

Chaudhury (1978) considered the problem of estimating finite population variance and gave several estimators. In particular he considered the estimator

$$e = \frac{1}{N^2}\frac{\alpha(S)}{P(S)} \qquad (4.4.5')$$

where

$$\alpha(S) = \sum_{i<j\in S} \frac{d_{ij}}{t_{ij}}, d_{ij} = (y_i - y_j)^2$$

$$t_{ij} = \sum_S m_{ij}(S) \text{ and } m_{ij}(S) = \begin{cases} 1 & \text{if } i,j \in S \\ 0 & \text{otherwise} \end{cases}$$

He proved the following results:

$$(1)\ \ Var(e) = \frac{1}{N^4}\{\sum \alpha^2(S)/P(S) - \theta^2\}$$

where $\theta = E(e) = V(\underset{\sim}{y})$. (2) An unbiased estimator for $Var(e)$ is

$$\frac{1}{N^4}[\alpha^2(S)(\frac{1}{P(S)} - 1) - \beta(S) + \gamma(S)]$$

where

$$\beta(S) = \sum_{i<j\in S} (d_{ij}^3/t_{ij}^2) l_{ij},\ l_{ij} = t_{ij}(t_{ij} - 1)$$

$$\gamma(S) = \sum_{i<j\in S}\sum_{k<l\in S} \frac{d_{ij}}{t_{ij}}\frac{d_{kl}}{t_{kl}}\frac{\gamma_{ijkl}}{f_{ijkl}}$$

where $\gamma_{ijkl} = \sum_S m_{ij}(S)m_{kl}(S)$ and $f_{ijkl} = \sum\sum_{S\neq S'} m_{ij}(S)m_{kl}(S')$.

The estimator (4.4.2) considered above is a special case of the estimator (4.4.5') (multiplied by $\frac{N}{N-1}$) and is applicable for any sampling design with a fixed size n and for which all the M_0 possible samples have positive probabilities of selection. The properties of the estimator (4.4.5') for different sampling designs have, however, not been studied. We consider here properties of the estimator (4.4.2) for the following sampling desogns, i.e., SRSWOR, Midzuno, Singh and Srivastava, PPSWR, etc. The estimator (4.4.2) was also considered by Isaki (1980).

4.4.2 Application to Different Without Replacement Sampling Designs

(a) For the simple random sampling without replacement(srswor) design, p_0 (say), $t = s_y^2$. From (4.4.5) an unbiased estimator of variance of s_y^2, $V(s_y^2)$ is

$$\begin{aligned} v(s_y^2) &= s_y^4 - \frac{1}{Nn}\sum_S y_i^4 - \frac{N^2-2N+3}{N(N-1)n(n-1)}\sum_S{}' y_i^2y_j^2 \\ &+\frac{4}{Nn(n-1)}\sum_S{}' y_i^3y_j + \frac{2(N-3)(N-2)}{Nn(n-1)(n-2)}\sum_S{}'' y_i^2y_jy_k \\ &-\frac{(N-2)(N-3)}{N(N-1)n(n-1)(n-2)(n-3)}\sum_S{}''' y_iy_jy_ky_l \qquad (4.4.6)\end{aligned}$$

We denote this strategy (p_0, s_y^2) by H_0.

(b) For Midzuno strategy, p_M(say), $p(S) = q_S/M_1$, where $q_S = \sum_{i\in S} p_i$, $p_i = x_i/X, X = \sum_{i=1}^N x_i$. Here the estimator (4.4.2) reduces to

$$t_M = \frac{ns_y^2}{Nq_S} \qquad (4.4.7)$$

with variance,

$$V(t_M) = \frac{n^2}{N^2M_1}\sum_S \frac{s_y^4}{q_S} - S_y^4 \qquad (4.4.8)$$

The strategy (p_M, t_M) is denoted by H_1.

(c) Under the scheme due to Singh and Srivastava(1980), p_R(say), $p(S) = \frac{s_x^2}{M_0S_x^2}$. Here (4.4.2) reduces to the ratio-type estimator

$$t_R = \frac{s_y^2}{s_x^2}S_x^2 \qquad (4.4.9)$$

with

$$V(t_R) = \frac{S_x^2}{M_0}\sum_{S}\frac{s_y^4}{s_x^4} - S_y^4 \qquad (4.4.10)$$

We denote the strategy (p_R, t_R) by H_2. The strategy was also mentioned by Nanjamma et al (1959).

Clearly t_R is biased under p_0. The bias of t_R, under p_0, to first order of approximation is, assuming $|\frac{s_x^2 - S_x^2}{s_x^2}| < 1$, following Cochran(1977,pp.160-162)

$$B(t_R) \approx [RV(s_x^2) - C(s_x^2, s_y^2)]/S_x^2 \qquad (4.4.11)$$

where $V(e)$, $C(e,e')$ denote variance of e and covariance of e and e' respectively and $R = \frac{S_y^2}{S_x^2}$. Calculations show that

$$\begin{aligned} C(s_x^2, s_y^2) &= \frac{N(N-n)\{(N-1)(N-1)-2\}}{n_{(2)}N_{(4)}}\sum(x_i - \overline{X})^2(y_i - \overline{Y})^2 \\ &\quad \frac{(N-1)(N-n)\{(n-1)(N-1)^2 - 2n\}}{n_{(2)}N_{(4)}}S_x^2 S_y^2 \\ &\quad + \frac{2(N-n)(N-n-1)}{n_{(2)}N_{(4)}}(N-1)^2 S_{xy}^2 \qquad (4.4.12) \end{aligned}$$

where

$S_{xy} = \frac{1}{(N-1)}\sum(x_i - \overline{X})(y_i - \overline{Y})$ and $u_{(v)} = u(u-1)...(u-v+1)$. Using expression for $V(s_x^2)$ [Sukhatme(1953)] and (4.4.12),from (4.4.11)

$$\begin{aligned} B(t_R) &\approx \frac{(N-n)}{n_{(2)}N_{(4)}S_x^2}[2(N-1)^2(N-n-1)\{S_x^2 S_y^2 - S_{xy}^2\} \\ &\quad -N\{(N-1)(n-1)-2\}\{\sum(x_i - \overline{X})^2(y_i - \overline{Y})^2 \\ &\quad -\frac{S_y^2}{S_x^2}\sum(x_i - \overline{X})^4\}] \qquad (4.4.13) \end{aligned}$$

Following Hartley and Rao(1954), an exact expression for bias of t_R under p_0 is given below.

We have

$$\begin{aligned} C(\frac{s_y^2}{s_x^2}, s_x^2) &= E(s_y^2) - E(\frac{s_y^2}{s_x^2})E(s_x^2) \\ \text{therefore } S_x^2(\frac{s_y^2}{s_x^2}) &= S_y^2 - C(\frac{s_y^2}{s_x^2}, s_x^2) \\ \text{or, } E(t_R) - S_y^2 &= -C(\frac{s_y^2}{s_x^2}, s_x^2) \end{aligned}$$

Hence bias in t_R is

$$B(t_R) = -\rho_{z_1,z_2}\sigma_{z_1}\sigma_{z_2} \tag{4.4.14}$$

where $z_1 = \frac{s_y^2}{s_x^2}, z_2 = s_x^2$. Thus

$$\frac{|B(t_R)|}{\sigma(t_R)} = |\rho_{z_1,z_2}|\frac{\sigma_{z_2}}{s_x^2} < cv(s_x^2) \tag{4.4.15}$$

where $cv(s_x^2)$ denotes coefficient of variation of s_x^2. Hence if the population is such that $cv(s_x^2) \leq .1$, then the bias may be regarded as negligible in comparison with the mse.

Now,

$$t_R - S_y^2 \simeq s_y^2 - Rs_x^2(\text{ assuming } \frac{S_x^2}{s_x^2} \simeq 1)$$

Hence mse of t_R is

$$E(t_R - S_y^2)^2 \simeq E(s_y^2 - Rs_x^2)^2$$

$$= \frac{(N-n)}{n_{(2)}N_{(4)}}[N\{(n-1)(N-1)-2\}\{\sum((y_i-\overline{Y})^2 - R(x_i-\overline{X})^2)^2\}$$
$$+4R(N-1)^2(N-n-1))S_x^2S_y^2 - S_{xy}^2)] \qquad (4.4.16)$$

The strategy (p_0, t_R) is denoted by H_3.

Das and Tripathi(1971) considered the class of ratio-type (of which t_R is a special case) and product-type estimators of $V(\underset{\sim}{y})$ in case of srswor and studied their properties for large samples.

4.4.3 Estimation under PPSWR

Under ppswr, an unbiased estimator of S_y^2 is from (4.4.1)

$$t_p = a_1' A_S - a_2' B_S \tag{4.4.17}$$

where

$$A_S = \sum_S \frac{y_i^2}{np_i},\ B_S = \frac{1}{n(n-1)}\sum_S{}' \frac{y_iy_j}{p_ip_j} - \frac{1}{2}(\frac{y_i^2}{p_i} + \frac{y_j^2}{p_j})$$

since,

$$E(A_S) = \sum y_i^2, E(B_S) = \sum{}' y_iy_j$$

This estimator was also contained in Murthy (1967, Chapter 6, Section 6.7). We have

$$V(t_p) = \frac{1}{(N-1)^2}V(\sum_S \frac{y_i^2}{np_i}) + \frac{a_2^2}{n^2(n-1)^2}V(\sum_S{}' \frac{y_iy_j}{p_ip_j})$$

$$-\frac{2a_2}{n(n-1)(N-1)}C(\sum_S \frac{y_i^2}{np_i}, \sum_S{}' \frac{y_iy_j}{p_ip_j})$$

Now,

$$\begin{aligned} V(\sum_S \frac{y_i^2}{np_i}) &= \sum \frac{y_i^4}{n^2p_i^2}V(r_i) + \sum{}' \frac{y_i^2y_j^2}{n^2p_ip_j}C(r_i,r_j) \\ &= \sum \frac{y_i^4}{n^2p_i^2}np_i(1-p_i) + \sum{}' \frac{y_i^2y_j^2}{n^2p_ip_j}(-np_ip_j) \\ &= \sum \frac{y_i^4}{np_i} - \frac{1}{n}(\sum y_i^2)^2 \qquad (4.4.18a) \end{aligned}$$

where r_i denotes the number of times i-th unit occurs in the sample S.

Denoting $\frac{y_i}{p_i}$ by z_i,

$$\begin{aligned} V(\sum_S{}' \frac{y_iy_j}{p_ip_j}) &= V(\sum_S{}' z_iz_j) \\ &= [2\sum_S{}' V(z_iz_j) + 4\sum_S{}'' C(z_iz_j, z_jz_k) + \sum_S{}''' C(z_iz_j, z_kz_l)] \end{aligned}$$

Now variance of the product of the two variables X_1 and X_2 is given by [Goodman (1960)]

$$V(X_1X_2) = (\overline{X}_1\overline{X}_2)^2[\frac{V(z_i)}{\overline{X}_1^2} + \frac{V(X_2)}{\overline{X}_2^2} + \frac{V(X_1)V(X_2)}{\overline{X}_1^2\overline{X}_2^2}]$$

where $\overline{X}_i = E(X_i), i = 1, 2$. Hence

$$\begin{aligned} V(z_iz_j) &= (\overline{z}^2)^2[\frac{V(z_i)}{\overline{z}^2} + \frac{V(z_j)}{\overline{z}^2} + \frac{V(z_i)V(z_j)}{\overline{z}^4}] \\ &= \overline{z}^4[2\frac{V(z_i)}{\overline{z}^2} + \frac{V(z_i)^2}{\overline{z}^4}], \end{aligned}$$

since z_i and z_j are identically distributed.

Using $\overline{z} = E(z_i) = Y$ and

$$V(z_i) = V(\frac{y_i}{p_i}) = \sum(\frac{y_i}{p_i} - Y)^2p_i = \alpha \text{ (say)}$$

$$V(z_iz_j) = Y^4[\frac{2\alpha}{Y^2} + \frac{\alpha^2}{Y^4}]$$

$$
\begin{aligned}
&= Y^2\alpha[2+\frac{\alpha}{Y^2}], \\
C(z_iz_j, z_iz_k) &= E(z_i^2 z_j z_k) - E(z_iz_j)E(z_iz_k) \\
&= Y^2\sum \frac{y_i^2}{p_i} - Y^4 \\
&= Y^2\alpha
\end{aligned}
$$

and $C(z_iz_j, z_kz_l) = 0$.

Therefore

$$
\begin{aligned}
V(\sum_S{}' \frac{y_iy_j}{p_ip_j}) &= 4\frac{n(n-1)}{2}Y^2\alpha(2+\frac{\alpha}{Y^2}) + 4n(n-1)(n-2)Y^2\alpha \\
&= 2n(n-1)Y^2\alpha[2+\frac{\alpha}{Y^2}+2n-4] \\
&= 2n(n-1)[(\sum\frac{y_i^2}{p_i})^2 + 2(n-2)Y^2\sum\frac{y_i^2}{p_i} \\
&\quad -(2n-3)Y^4] \qquad (4.4.18b)
\end{aligned}
$$

Again,

$$
\begin{aligned}
C(\sum_S \frac{y_i^2}{p_i}, \sum_S{}' \frac{y_iy_j}{p_ip_j}) &= 2\sum_S{}' C(\frac{y_i^2}{p_i}, \frac{y_iy_j}{p_ip_j}) \\
&\quad + \sum_S{}'' C(\frac{y_i^2}{np_i}, \frac{y_jy_k}{p_jp_k}) \\
&= 2\sum_S{}' \{E(\frac{y_i^3y_j}{p_ip_j}) - E(\frac{y_i^2}{p_i}E(\frac{y_iy_j}{p_ip_j})\} \\
&= 2(n-1)(Y\sum\frac{y_i^3}{p_i} - Y^2\sum y_i^2) \qquad (4.4.18c)
\end{aligned}
$$

Hence from (4.4.18a),(4.4.18b) and (4.4.18c)

$$
\begin{aligned}
V(t_p) &= \frac{1}{n(N-1)^2}[\{\sum\frac{y_i^4}{p_i} - (\sum y_i^2)^2\} \\
&\quad + \frac{2}{N^2(n-1)}\{(\sum\frac{y_i^2}{p_i})^2 + 2(n-2)Y^2\sum\frac{y_i^2}{p_i} - (2n-3)Y^4\} \\
&\quad - \frac{4}{N}\{Y\sum\frac{y_i^3}{p_i} - Y^2\sum y_i^2\}] \qquad (4.4.18)
\end{aligned}
$$

We denote the strategy (ppswr, t_p) as H_4.

Das and Tripathi(1978) also mentioned of the estimator t_p but they did not study any of its properties.

As noted by an examiner the estimator in (4.4.17) is not uniformly non-negative , but a non-negative unbisaed estimator of S_y^2 based on PPSWR may be considered, such as

$$\frac{1}{2N(N-1)n(n-1)}\sum_{s}{}' \frac{(y_i-y_j)^2}{p_ip_j} = t_p'(\text{say})$$

We denote the strategy (ppswr, t_p') as H_4'. Now variance of t_p' is

$$\begin{aligned}
V(t_p') &= \frac{1}{2n(n-1)N^2(N-1)^2}[-(2n-3)\{\sum\sum(y_i-y_j)^2\}^2 \\
&\quad +2(n-2)\{\sum_i \frac{1}{p_i}(\sum_{k\neq i}(y_i-y_k)^2)^2\} + \sum\sum\frac{(y_i-y_j)^4}{p_ip_j} \\
&= \frac{1}{n(n-1)N^2(N-1)^2}[\frac{y_i^4}{p_i}\{N(n-2)+\sum\frac{1}{p_i}\} \\
&\quad +4N(2n-3)\sum y_i^2(\sum y_i)^2 - 2(2n-3)(\sum y_i)^4 \\
&\quad +4(n-2)(\sum y_i)^2\sum\frac{y_i^2}{p_i} + 2(n-2)\{2N\sum\frac{y_i^2}{p_i}\sum y_i^2 \\
&\quad -4\sum\frac{y_i}{p_i}\sum y_i\sum y_i^2 + (\sum y_i^2)^2(\sum\frac{1}{p_i})\} + 3(\sum\frac{y_i^2}{p_i})^2 \\
&\quad -4\sum\frac{y_i^3}{p_i}(N(n-2)\sum y_i + \sum\frac{y_i}{p_i}) - 2N^2(2n-3)(\sum y_i^2)^2]
\end{aligned}$$

4.4.4 Ratio Estimator of Variance Under a Controlled Sampling Design

Another type of sampling of practical importance is controlled sampling. In SRSWOR, the number of possible samples $\binom{N}{n}$ is very large even for moderate N,n. In many field surveys, all the possible samples are not equally preferable from operational point of view - some samples may be inaccessible, inconvenient, expensinve, subject to insufficient or less efficient handling. It will, therefore, be advantageous if the sampling design is such that the total number of possible samples be much less than $\binom{N}{n}$, retaining at the same time unbiasedness property of the estimators of interest. Chakraborty (1963) first introduced the concept of using the incidence matrices of designs in sampling from finite populations in this respect. t-designs have been used in generating controlled sampling designs by Srivastava and Saleh (1985), Foody and Hedayet (1977), among others.

Some recent works on controlled sampling designs are due to Hedayet, Lin and Stufken (1989), Rao and Nigam (1989,1990), Mukhopadhyay and Vijayan (1992).

Controlled sampling has been used by several authors for estimating a finite population total [eg.Sukhatme and Avadhani(1965), Avadhani and Sukhatme(1967, 1968, 1973), Srivastava and Salah(1985), Rao and Nigam (1989)]. In this section we consider a controlled sampling design for estimating the population variance.

A balanced incomplete block design (BIBD) is considered with parameters v,b,r ,k,λ, where v is the total number of elements or units, b is the number of blocks, r is the number of replications of each unit, k is the size of a block and λ is the number of blocks in which every pair (i,j) of elements occur together, $i \neq j = 1,2,\cdots,v$. Each element is identified as a unit in the finite population and each block as a sample. Therefore $N = v,\ n = k$.

Samples(blocks) are selected with probability proportional to sample variance, s_x^2 ie.,$p(S) = \mu.s_x^2$, where μ is the constant of proportionality.

Using the relation $\sum_{S=1}^{b} p(S) = 1$ we get,

$$\mu \sum_{S=1}^{b}\{\frac{1}{n}\sum_{S} x_i^2 - \frac{1}{n(n-1)}\sum_{S}{}' x_i x_j\} = 1$$

$$\text{or, } \mu[\sum x_i^2.\frac{r}{n} - \frac{\lambda}{n(n-1)}\sum{}' x_i x_j] = 1$$

$$\text{or, } \frac{\mu r}{n}[\sum x_i^2 - \frac{1}{N-1}\sum{}' x_i x_j] = 1$$

[since $r(k-1) = \lambda(v-1) ie., r(n-1) = \lambda(N-1)$]

$$\Rightarrow \mu = \frac{n}{rNS_x^2}$$

$$\text{Hence } p(S) = \frac{ns_x^2}{rNS_x^2} \qquad (4.4.19)$$

Under this scheme, $t_R = \frac{s_y^2}{s_x^2}S_x^2$ becomes unbiased for S_y^2, since

$$\begin{aligned} E(t_R) &= \sum_{S=1}^{b} t_R p(S) \\ &= \sum_{S=1}^{b} \frac{s_y^2}{s_x^2} S_x^2 . \frac{ns_x^2}{rNS_x^2} \\ &= \frac{n}{rN}\sum_{S=1}^{b} s_y^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{n}{rN}\sum_{S=1}^{b}\{\frac{1}{n}\sum_{S} y_i^2 - \frac{1}{n(n-1)}\sum_{S}^{'} y_i y_j\} \\
&\doteq \frac{1}{rN}[r\sum y_i^2 - \frac{\lambda}{(n-1)}\sum^{'} y_i y_j] \\
&= S_y^2 \qquad (4.4.20) \\
V(t_R) &= E(t_R)^2 - \{E(t_R)\}^2 \\
&= \frac{n}{rN}\sum_{S=1}^{b}\frac{s_y^4}{s_x^2}S_x^2 - S_y^4 \qquad (4.4.21)
\end{aligned}
$$

We denote this strategy by H_5.

We note that the unbiasedness of the strategy H_5 follows simply from the fact that if for a design p, $\hat{S}_y^2$ and $\hat{S}_x^2$ are unbiased estimator of S_y^2 and S_x^2 respectively, the estimator $\hat{S}_y^2 S_x^2/\hat{S}_x^2$ is an unbiased estimator of S_y^2 under the modified design p^*, when $p^*(S) = (\hat{S}_x^2/S_x^2)p(S)$. This is also noted that H_5 is actually a family of sampling strategies and H_2 itself is a mamber of this family corresponding to the trivial BIBD. One may similarly think of controlled versions of H_0 and H_1.

4.4.5 An Empirical Study

To assess the performance of strategies H_i $(i = 0, 1, \cdots, 4)$ we considered 35 natural populations as listed in table 4.0 of chapter 4.

Along with the strategies $H_i(i = 0, ..., 4)$ we considered the following strategies

$$(a) \quad H_6 \equiv (p_M, e')$$

$$(b) \quad H_7 \equiv (p_R, e')$$

where e' is another non-negative unbiased estimator

$$e' = \frac{1}{N(N-1)}\sum\sum_{i \neq j \in S}\frac{(y_i - y_j)^2}{\pi_{ij}},$$

considered by Chaudhury (1974), Mukhopadhyay (1984), among others. For $n = 2$ strategies H_6 and H_7 coincide with the strategies H_1 and H_2 respectively.

Variances of H_i, $V(H_i)(i = 0(1)4, 6, 7)$ were computed for all the above mentioned 35 natural populations (excluding those populations for which variance of H_2 was not defined, because $p(S)$, which is a function of sample variance of the auxiliary variable x, becomes zero for some samples). The variances have been shown in tables 4.1, 4.2 and 4.3 for $n = 2, 3, 4$ respectively.

The strategies $H_i (i = 0, ..., 4)$ were ranked according to the magnitude of $V(H_i)$, one with the lowest magnitude of its variance being given rank 1, the next lowest rank 2, etc. The following table shows the average rank of the strategies over different populations for $n = 2, 3, 4$. If a strategy H_i has less average rank than a strategy H_j (on the basis of the populations considered) we will write $H_i \underset{\succ}{r} H_j$.

Table 4.A

		Average Rank						
	Sample size	H_0	H_1	H_2	H_3	H_4	H_6	H_7
	n = 2 (17)	4.59	1.53	3.29	3.18	2.41	–	–
Rank	n = 3 (32)	6.66	3.61	3.76	4.33	5.0	3.03	1.55
	n = 4 (32)	6.72	3.76	3.34	3.92	5.51	3.25	1.5

Figures in the parenthesis indicate the number of populations considered.

The above table shows that for the populations studied,

for n=2, $H_1 (\equiv H_6) \underset{\succ}{r} H_4 \underset{\succ}{r} H_3 \underset{\succ}{r} H_2 (\equiv H_7) \underset{\succ}{r} H_0$

for n=3, $H_7 \underset{\succ}{r} H_6 \underset{\succ}{r} H_1 \underset{\succ}{r} H_2 \underset{\succ}{r} H_3 \underset{\succ}{r} H_4 \underset{\succ}{r} H_0$

for n=4, $H_7 \underset{\succ}{r} H_6 \underset{\succ}{r} H_2 \underset{\succ}{r} H_1 \underset{\succ}{r} H_3 \underset{\succ}{r} H_4 \underset{\succ}{r} H_0$

on an average.

Hence, the following conclusions can be drawn from the above empirical study:

(i) for $n = 2$ the strategy H_1 is best among $H_i, i = 0, ...4$ and for $n = 3$ and 4, H_1 and H_2 are preferable among the strategies $H_i (i = 0, ..., 4)$. H_0 is the least preferable(having the largest average rank) for all n among the strategies $H_i (i = 0(1)4, 6, 7)$.

(ii) it is found that though in most cases H_7 and H_6 fared as the most stable estimator (in the sense of having lowest variance) the performance of H_1 and H_2 was quite satisfactory and quite close to that of H_6 and H_7 for almost all the populations and for all the values of n considered.

The investigations into the above natural populations though did not advocate vehemently in favour of the strategies $H_i (i = 0, ..., 4)$, it is felt that they, specially strategies H_1 and H_2, should give satisfactory estimators of finite population variance in practice. However, the purpose of our investigations was not for finding grounds for supporting any of our strategies $(H_0, ..., H_4)$ but also to find their position vis-a-vis the other rival strategies.

We now consider the performance of H_5 vis-a-vis the strategies $H_0, ..., H_4$.

We consider the following doubly balanced BIBD (8, 14, 7, 4, 3) [Chakraborty,M.C.(1963)]

1	2	3	4
5	6	7	8
1	2	7	8
3	4	5	6
1	3	6	8
2	4	5	7
1	4	6	7
2	3	5	8
1	2	5	6
3	4	7	8
1	3	5	7
2	4	6	8
1	2	5	8
2	3	6	7

We impose this design on the population numbered 1,2,3,4,8,13,14,15 and 16 of table 4.0 [of which the last three populations numbered 14,15 and 16 are confined to their first 8 units] and also on the population with the following x and y values [Rao and Singh(1973), table 1]

unit i	1	2	3	4	5	6	7	8
x_i	506	977	2252	2254	3802	4873	5542	7409
y_i	0	22	5	74	63	131	80	141

Again, we consider the following symmetrical BIBD (7, 7, 3, 3, 1) [Chakraborty,M.C. (1963)]

1	3	4
2	4	5
3	5	6
4	6	7
5	7	1
6	1	2
7	2	3

We impose this design on the population numbered 1-4 and 6-10 of table 4.0 and also on the population with the following x and y values [Rao and Singh(1973)]

unit i	1	2	3	4	5	6	7
x_i	25	20	18	17	16	24	22
y_i	121	110	99	85	82	116	108

The variances of the strategies $H_i(i = 0, ..., 5)$ of the above populations have been shown in tables 4.4 and 4.5.

The strategies were ranked as before and their average ranks are shown in the following table

Table 4.B

	Sample size	Popln. size	Average Rank H_0	H_1	H_2	H_3	H_4	H_5
Rank	n = 3	N = 7(10)	5.3	2.2	3.1	3.5	3.8	3.1
	n =4	N = 8(10)	4.7	2.2	3.3	2.9	4.5	3.5

Figures in the parenthesis indicate the number of populations considered.

Thus for n=3, $H_1 \underset{\succ}{r} H_2 = H_5 \underset{\succ}{r} H_3 \underset{\succ}{r} H_4 \underset{\succ}{r} H_0$

for n=4, $H_1 \underset{\succ}{r} H_3 \underset{\succ}{r} H_2 \underset{\succ}{r} H_5 \underset{\succ}{r} H_4 \underset{\succ}{r} H_0$

where '$H_i = H_j$' denotes the strategy H_i has the same performance as H_j in the sense of having equal value of average rank.

From the above table we infer that :

(i) the strategy H_1 is best among $H_i, i = 0, ..., 5$ and H_0 is the least preferable strategy. H_4 is the next superior.

(ii) For n=3 H_2 is superior to H_3 where for n=4 the position is reversed. The position of H_5 is somewhere around H_2 and H_3,- it is equivalent to H_2 for n=3 but is inferior to both H_3 and H_2 (but, ofcourse, superior to H_4,H_0) for n=4. Both the tables 4.A and 4.B show that H_1 is the best and H_0 is the worst strategy among $H_i, i = 0, ..., 5$.

4.4.6 Average Variances of the Strategies under a Superpopulation Model

To assess the relative performance of the strategies $H_i(i = 0, ..., 4)$ we compare their average variances under the following superpopulation madel:

$y_1, y_2, ..., y_N$ are considered as realisations of N independent normal variables $Y_1, Y_2, ..., Y_N$, each distributed with $(0, \sigma^2 x^g)$, σ^2, a constant (> 0) and g, a known quantity. We, further, assume that x's are independent gamma variables with parameter r. Such assumptions for comparison of average variances have been considered by several authors [eg. Rao and

Rao (1971a,b)] E_x and $\mathcal{E}$ will denote respectively expectation wrt x and the above superpopulation model.

We shall denote $E_x\mathcal{E}V(H_i)$ as γ_i. We have

$$\begin{aligned}
\gamma_0 &= E_x[\frac{1}{M_0}\sum_{S\in S}\{\frac{1}{n^2}\sum_{i\in S}3\sigma^4x_i^{2g} \\
&\quad +\frac{(n^2-2n+3)}{n^2(n-1)^2}\sum\sum_{i\neq j\in S}\sigma^4x_i^gx_j^g\} - \{\frac{1}{N^2}\sum_{i=1}^{N}3\sigma^4x_i^{2g} \\
&\quad +\frac{(N^2-2N+3)}{N^2(N-1)^2}\sum\sum_{i\neq j\in S}^{N}\sigma^4x_i^gx_j^g\}] \\
&= (\frac{(N-n)}{Nn}\sigma^4[3\frac{\Gamma(r+2g)}{\Gamma(r)} \\
&\quad -\frac{(Nn-3N-3n+3)}{(n-1)(N-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \qquad (4.4.22)
\end{aligned}$$

on simplification.

Again,

$$\begin{aligned}
\gamma_1 &= E_x[\frac{M_1}{M_0^2}\sum_{S\in S}\frac{1}{(\sum_{i\in S}p_i)}\{\frac{1}{n^2}\sum_{i\in S}3\sigma^4x_i^{2g} \\
&\quad +\frac{n^2-2n+3}{n^2(n-1)^2}\sum\sum_{i\neq j\in S}\sigma^4x_i^gx_j^g\} - \{\frac{1}{N^2}\sum_{i=1}^{N}3\sigma^4x_i^{2g} \\
&\quad +\frac{(N^2-2N+3)}{N^2(N-1)^2}\sum\sum_{i\neq j=1}^{N}\sigma^4x_i^gx_j^g\}] \\
&= \frac{M_1\sigma^4}{M_0^2}[\sum_{S\in S}E_x\{\frac{1}{n^2}\sum_{i\in S}3x_i^{2g} \\
&\quad +\frac{(n^2-2n+3)}{n^2(n-1)^2}\sum\sum_{i\neq j\in S}x_i^gx_j^g\} \\
&\quad +\sum_{S\in S}E_x\frac{(\sum_{i\in \bar{S}}x_i)}{\sum_{i\in S}x_i)}\{\frac{1}{n^2}\sum_{i\in S}3x_i^{2g} \\
&\quad +\frac{(n^2-2n+3)}{n^2(n-1)^2}\sum\sum_{i\neq j\in S}x_i^gx_j^g\}] \\
&\quad -\sigma^4[\frac{3}{N}\frac{r+2g}{r}+\frac{N^2-2N+3)}{N(N-1)}(\frac{r+g}{r})^2] \\
&\simeq \frac{M_1}{M_0^2}\sigma^4[\sum{}'\{\frac{3}{n}\frac{\Gamma(r+2g)}{\Gamma(r)}+\frac{(n^2-2n+3)}{n(n-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2\}
\end{aligned}$$

$$+\sum{}' \frac{(N-n)r}{nr}\{\frac{3}{n}\frac{\Gamma(r+2g)}{\Gamma(r)}$$
$$+\frac{(n^2-2n+3)}{n(n-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2\}]$$
$$-\sigma^4\{\frac{3}{N}\frac{\Gamma(r+2g)}{\Gamma(r)}+\frac{(N^2-2N+3)}{N(N-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2\}$$

[Using the approximation,

$$E_x(\frac{\phi_1(x)}{\phi_2(x)}) \simeq \frac{E_x(\phi_1(x))}{E_x(\phi_2(x))}, \quad (\star)$$

where $\phi_1(x)$ and $\phi_2(x)$ are rational functions of x with $\phi_2(x) \neq 0$.

Such approximations have been used extensively by various authors [eg. Rao (1978)]

$$= (\frac{N-n}{Nn})\sigma^4[3\frac{\Gamma(r+2g)}{\Gamma(r)} - \frac{(Nn-3N-3n+3)}{(n-1)(N-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \quad (4.4.23)$$

on simplification.

We note that $\gamma_0 = \gamma_1$, where however, the expression for γ_0 is exact and that of γ_1 is the approximated one.

$$E_x\mathcal{E}V(H_2) = E_x[\frac{1}{M_0}\sum{}'\frac{1}{s_x^2}\{k_1A(x,g,S)+k_2B(x,g,S)\}\{L_1(A(x,1,S)$$
$$+A(x,1,\bar{S}))-L_2(B((x,1,S)+B(x,1,\bar{S})$$
$$+2A(x,1/2,S)A(x,1/2,\bar{S}))\}]-\sigma^4[\frac{3}{N}\frac{\Gamma(r+2g)}{\Gamma(r)}$$
$$+\frac{(N^2-2N+3)}{(N(N-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \quad (4.4.24)$$

where

$$A(x,g,S) = \sum_{i\in S} x_i^{2g},$$
$$B(x,g,S) = \sum\sum_{i\neq j\in S} x_i^g x_j^g$$
$$k_1 = \frac{3\sigma^4}{n^2},$$

$$
\begin{aligned}
k_2 &= \frac{(n^2-2n+3)\sigma^4}{n^2(n-1)^2}, \\
L_1 &= \frac{1}{N}, \\
L_2 &= \frac{1}{N(N-1)}
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
A(x,g,\bar{S}) &= \sum_{i\in\bar{S}} x_i^2, \\
B(x,g,\bar{S}) &= \sum\sum_{i\neq j\in\bar{S}} x_i x_j \text{ etc.}
\end{aligned}
$$

now,

$$
E_x[A(x,g,S)\{A(x,1,S)+A(x,1,\bar{S})\}]
$$
$$
= n\frac{\Gamma(r+2g)}{\Gamma(r)}\{N(r^2+r)+4g^2+4rg+2g\} \qquad (4.4.25a)
$$
$$
E_x[A(x,g,S)\{B(x,1,S)+B(x,1,\bar{S})+2A(x,1/2,S)A(x,1/2,S)\}]
$$
$$
= n(N-1)r\frac{\Gamma(r+2g)}{\Gamma(r)}(Nr+4g) \qquad (4.4.25b)
$$
$$
E_x[B(x,g,S)\{A(x,1,S)+A(x,1,\bar{S})\}]
$$
$$
= n(n-1)(\frac{\Gamma(r+g)}{\Gamma(r)})^2\{N(r^2+r)+4rg+2g^2+2g\} \qquad (4.4.25c)
$$
$$
E_x[B(x,g,S)\{B(x,1,S)+B(x,1,\bar{S})+2A(x,1/2,S)A(x,1/2,\bar{S})\}]
$$
$$
= n(n-1)(\frac{\Gamma(r+g)}{\Gamma(r)})^2[N(N-1)r^2+4(N-1)rg] \qquad (4.4.25d)
$$
$$
\text{and } \; E_x(s_x^2) = r, \qquad (4.4.25e)
$$

on simplification.

Hence

$$
E_x \mathcal{E} V(H_2) = 3\sigma^4 \frac{\Gamma(r+2g)}{\Gamma(r)}\{\frac{(N-n)}{Nn} + \frac{4g^2+2g}{Nnr}\}
$$
$$
+\sigma^4(\frac{\Gamma(r+g)}{\Gamma(r)})^2\{-\frac{(N-n)(Nn-3N-3n+3)}{Nn(N-1)(n-1)}
$$
$$
+\frac{(n^2-2n+3)(2g^2+2g)}{Nn(n-1)r}\}, \qquad (4.4.26)
$$

on simplification using (4.4.25a)-(4.4.25e) and also the approximation $(\star)$ as used in the case of the strategy H_1.

Calculation of $E_X\mathcal{E}V(H_3)$ was found to be mathematically intractable and was omitted from comparison.

Lastly, from (4.4.18)

$$\begin{aligned}
E_x\mathcal{E}V(H_4) &= E_x\mathcal{E}(B_1+B_2+B_3)\\
\text{where, } B_1 &= \frac{1}{n(N-1)^2}\{\sum_{i=1}^{N}\frac{y_i^4}{p_i}-(\sum_{i=1}^{N}y_i^2)^2\}\\
B_2 &= \frac{2}{n(n-1)N^2(N-1)^2}\{(\sum_{i=1}^{N}\frac{y_i^2}{p_i})^2+2(n-2)Y^2\sum_{i=1}^{N}\frac{y_i^2}{p_i}\\
&\quad -(2n-3)Y^4\}\\
\text{and } B_3 &= -\frac{4}{nN(N-1)^2}\{Y\sum_{i=1}^{N}\frac{y_i^3}{p_i}-Y^2\sum_{i=1}^{N}y_i^2\}
\end{aligned}$$

Now,

$$\begin{aligned}
E_x\mathcal{E}(B_1) &= \frac{\sigma^4}{n(N-1)^2}E_x[3\sum_{i=1}^{N}\frac{x_i^{2g}}{p_i}-(3\sum_{i=1}^{N}x_i^{2g}+\sum\sum_{i\neq j=1}^{N}x_i^g x_j^g)]\\
&= \frac{N\sigma^4}{n(N-1)}[3r\frac{\Gamma(r+2g-1)}{\Gamma(r)}-(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \qquad (4.4.27)\\
E_x\mathcal{E}(B_2) &= \frac{2\sigma^4}{n(n-1)N(N-1)}[3\frac{\Gamma(r+2g-2)}{\Gamma(r)}\{(N-1)r^2+r\\
&\quad +2(n-1)r((r+2g-2)\}\\
&\quad +(\frac{\Gamma(r+g-1)}{\Gamma(r)})^2\{2(n-1)(r+g)(r+g-1)\\
&\quad +(N-2)^2r^2+(n-2)r-(4n-7)(r+g-1)^2\\
&\quad +2n(N-2)r(r+g-1)\}] \qquad (4.4.28)
\end{aligned}$$

and similarly,

$$E_x\mathcal{E}(B_3) = -\frac{4\sigma^4}{n(N-1)}[3r\frac{\Gamma(r+2g-1)}{\Gamma(r)}-(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \qquad (4.4.29)$$

Hence using (4.4.27)-(4.4.29) and on simplification

$$\gamma_4 = \frac{\sigma^4}{n(n-1)N(N-1)}[3r\frac{\Gamma(r+2g-2)}{\Gamma(r)}\{(n-1)(N-2)^2(r+2g-2)$$

$$+2(N-1)r+2\}+(\frac{\Gamma(r+g-1)}{\Gamma(r)})^2\{((n-1)(N-2)^2+6)(r+g-1)^2$$

$$+4(n-1)(r+g-1)+4n(N-2)r(r+g-1)+2(N-2)^2r^2$$
$$+2(N-2)r\}] \qquad (4.4.30)$$

No algebraic comparison between the average variances $\gamma_0(=\gamma_1)$, γ_2, γ_4 was possible.

Tables 4.6, 4.7 show that the values of average variances $E_x\mathcal{E}V(H_0)(= E_x\mathcal{E}V(H_1))$, $E_x\mathcal{E}V(H_2)$, $E_x\mathcal{E}V(H_4)$ (denoted respectively by ev1,ev2 and ev3 in the tables) for $r = 2(2)6$, $g = 0,1,2$, $f = \frac{n}{N} = .2,.4,.6$ for $N = 20$ and $N = 40$. It is seen that for large $r(=6)$ and high sampling fraction ($f = .6$) H_0 is the best among all these strategies for all values of g. Otherwise H_4 fares best followed by $H_0(\equiv H_1)$ and H_2.

The conclusions reached above is at variance with the conclusions reached about the performance of the strategies $H_0, ..., H_4$ on the basis of the empirical study based on first 30 natural populations as described in section 4.4.5. This is not surprising, because, the populations coinsidered in section 4.4.5 are expected to observe the model $\mathcal{E}(Y_i|x_i) = \beta x_i$, ($\beta$ a constant) where as the present model consideres $\mathcal{E}(Y_i|x_i) = \beta$ (a known constant). In the present model, the variance function $\mathcal{V}(Y_i|x_i)$ depends on x_i. The conclusions reached in this section should be applicable to populations where y-values remain constant on an average, though the conditional variance may depend on x-values.

4.4.7 A Suggested Study

The following investigation is considered as suggested by an examiner as follows :

We consider the strategy $H_* = (p,t)$ where p is a fixed size design with $p(S) \propto \mathcal{E}(s_y^4|x)$ and t is the estimator defined in (4.4.2). Assuming as before that x's are independent gamma variables with parameter r we have under the previous model

$$\begin{aligned}
E_x\mathcal{E}V(H_*) &= \sigma^4[3\frac{N-n}{Nn}\frac{\Gamma(r+2g)}{\Gamma(r)} \\
&\quad +\{\frac{n^2-2n+3}{n(n-1)} - \frac{N^2-2N+3}{N(N-1)}\}(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \\
&= (\frac{N-n}{Nn})\sigma^4[3\frac{\Gamma(r+2g)}{\Gamma(r)} - \frac{Nn-3(N-n)+3}{(n-1)(N-1)}(\frac{\Gamma(r+g)}{\Gamma(r)})^2] \\
&= \gamma_* \qquad \text{(say)}
\end{aligned}$$

which coincides with $\gamma_1 = E_x\mathcal{E}V(H_1)$. Therefore the suggested strategy H_* and our strategy H_1 are both equivalent in the sense of average variance.

We consider a model where $y_1, y_2, ..., y_N$ are assumed to be a realization of the variables $Y_1, Y_2, ..., Y_N$ each distributed independently as $N(\beta, \sigma^2 x_i^g)$. This model is considered because of the fact that the strategies involving the estimator in (4.4.2) is location invariant but the strategy H_4 is not. Hence, in order to compare the strategies under model it is justified to consider a model for which the mean is unknown but constant.

Under the new model

$$\begin{aligned}
\mathcal{E}V(H_4) &= \frac{1}{n(n-1)N^2(N-1)^2}[\sigma^4\{(\sum x_i^g)^2(N(n-1)(4-N) \\
&-6(2n-3)) + 4\sum \frac{x_i^{2g}}{p_i^2} + 2(\sum \frac{x_i^g}{p_i})^2 \\
&+4(n-2)\sum \frac{x_i^g}{p_i}\sum x_i^g - 4N(n-1)(4-N)\sum x_i^{2g} \\
&+\sum \frac{x_i^{2g}}{p_i}(8(n-2) + 3N(n-1)(4-N))\} \\
&+2\sigma^2\beta^2\{(-N(Nn+N+n-10) + 2\sum \frac{1}{p_i})\sum \frac{x_i^g}{p_i} \\
&+((Nn-N+2n+1) + 2(n-2)(\sum \frac{1}{p_i})\sum x_i^g \\
&+4\sum \frac{x_i^g}{p_i^2}\} + \beta^4(\sum \frac{1}{p_i} - N^2)(2\sum \frac{1}{p_i} + N^2(n-3))]
\end{aligned}$$

and

$$\begin{aligned}
E_x\mathcal{E}V(H_4) &= \frac{2}{n(n-1)N(N-1)}[\sigma^4\{2(N+2n+1)\frac{\Gamma(r+2g)}{\Gamma(r)} \\
&+(Nn(N-2) + 2(N-n) + 1)(\frac{\Gamma(r+g)}{\Gamma(r)})^2\} \\
&+\sigma^2\beta^2.4(N+n)\{\frac{\Gamma(r+2g)}{\Gamma(r)}\frac{\Gamma(r-g)}{\Gamma(r)} \\
&+(N-2)(\frac{\Gamma(r+g)}{\Gamma(r)})^2\frac{\Gamma(r-g)}{\Gamma(r)} - (N-1)\frac{\Gamma(r+g)}{\Gamma(r)}\} \\
&+\beta^4\{(N-2)(N-3)(\frac{\Gamma(r+g)}{\Gamma(r)})^2(\frac{\Gamma(r-g)}{\Gamma(r)})^2 \\
&-N(5N-4-Nn)\frac{\Gamma(r+g)}{\Gamma(r)}\frac{\Gamma(r-g)}{\Gamma(r)} \\
&+2\frac{\Gamma(r+2g)}{\Gamma(r)}\frac{\Gamma(r-2g)}{\Gamma(r)} \\
&+4(N-2)\frac{\Gamma(r+2g)}{\Gamma(r)}(\frac{\Gamma(r-g)}{\Gamma(r)})^2 - (3N-2-Nn)\}]
\end{aligned}$$

Furthermore it is to be noted that we can consider the mean of the previously assumed model to be zero for treating the estimator $t_p^{'}$ of section 4.4.3 as this one is invariant. Hence we get the following result.

$$
\begin{aligned}
\mathcal{E}V(t_p^{'}) &= \frac{\sigma^4}{2n(n-1)N^2(N-1)^2}[-(2n-3)\{12(N-1)^2\sum x_i^{2g} \\
&+4(N^2-2N+3)\sum{}' x_i^g x_j^g\} + 2(n-2)\{3(N-1)^2\sum \frac{x_i^{2g}}{p_i} \\
&+2(N+1)\sum{}' \frac{x_i^g x_j^g}{p_i} + 3\sum{}' \frac{x_j^{2g}}{p_i} + \sum{}'' \frac{1}{p_i} x_j^g x_k^g\} \\
&+6(\sum{}' \frac{x_i^g}{p_i p_j} + \sum{}' \frac{x_i^g x_j^g}{p_i p_j})]
\end{aligned}
$$

Putting $p_i = \frac{x_i^g}{\sum x_i^g}$ we get

$$
\begin{aligned}
\mathcal{E}V(t_p^{'}) &= \frac{\sigma^4}{4n^2(n-1)^2N^2(N-1)^2}[-2n(n-1)(2n-3) \\
&\{12(N-1)^2\sum x_i^{2g} + 4(N^2-2N+3)\sum{}' x_i^g x_j^g\} \\
&+4n(n-1)(n-2)\{3(N-1)^2\sum x_i^g \sum x_i^g \\
&+2(N+1)\sum x_i^g \sum_{j\neq i} x_j^g + 3(N-1)(\sum x_i^g)^2 \\
&+(\sum x_i^g)\sum{}'' x_i^{-g} x_j^g x_k^g\} + 12n(n-1)\{(\sum x_i^g)^2 \sum{}' x_i^g x_j^{-g} \\
&+N(N-1)(\sum x_i^g)^2\}
\end{aligned}
$$

and

$$
\begin{aligned}
E_x\mathcal{E}V(t_p^{'}) &= \frac{2\sigma^4}{2n(n-1)N(N-1)}[\frac{\Gamma(r+2g)}{\Gamma(r)}\{-14n(N-2)+22N-32\} \\
&+2(\frac{\Gamma(r+g)}{\Gamma(r)})^2\{n(8N^2-2N-14)-2(N^2+3N-17)\} \\
&+2(2n-1)\frac{\Gamma(r+3g)}{\Gamma(r)}\frac{\Gamma(r-g)}{\Gamma(r)} \\
&+6(2n+1)(N-2)\frac{\Gamma(r+2g)}{\Gamma(r)}\frac{\Gamma(r+g)}{\Gamma(r)}\frac{\Gamma(r-g)}{\Gamma(r)} \\
&+2(n+1)(N-2)(N-3)(\frac{\Gamma(r+g)}{\Gamma(r)})^3\frac{\Gamma(r-g)}{\Gamma(r)}]
\end{aligned}
$$

Table 4.0

pop. no.	Source	y	x	popln. size(N)	cv(x)	cv(y)	ρ
1	Murthy (1967) p.126 Vill. 1-8	No.of cultiv-ators in 1961	Area in 1951	8	0.449	0.517	0.69
2	do	Workers at household industry	Area in 1951	8	0.449	0.696	0.43
3	do, vill. 9-16	No. of cultiv-ators in 1961	No. of persons in 1961	8	0.634	0.530	0.92
4	do	Workers at household industry	do	8	0.634	0.777	0.89
5	do, p.132 bl. no.7	Timber volume	Strip length	13	0.368	0.351	0.95
6	do, p.178 vill. 1-10	Area under paddy	Geogra-phical area	10	0.085	0.344	0.25
7	do, p.180	Catch of fish	No. of boats landing	12	0.723	0.776	0.93
8	do, p.228	Output	No.of workers	8	0.056	0.308	0.82
9	Sukhatme & Sukhatme (1970)p.51	Area under rice	Total cultivated area	10	0.391	0.397	0.87
10	do, p.166	No.of banana bunches	No. of banana pits	10	0.248	0.194	0.85
11	Yates (1960)p.159	No.of absentees	Total no. of persons	12	0.503	0.650	0.81
12	do, p.163	Volume of timber	Eye-estimate	9	0.573	0.484	0.90

Table 4.0 (contd.)

pop. no.	Source	y	x	popln. size(N)	cv(x)	cv(y)	ρ
13	do, p.232	Length coloured	Length of line	8	0.562	1.609	0.47
14	Cochran (1977)p.182	No. of paralytic polio cases in placebo group	No. of 'placebo' children	10	0.423	1.077	0.22
15	Raj,D. (1972)p.70	No. of cattle	No.of farms	15	0.420	0.414	0.77
16	Sukhatme & Sukhatme (1970)p.185	Area under wheat (1973)	Area under wheat (1936)	20	0.723	0.744	0.99
17	R.K.Som (1973)p.255	Average no. of persons per h.h	Average monthly expenditure on cereals per h.h.	16	0.351	0.243	-0.35
18	Jessen (1978)p.151	Total no. of fish	No.of fish tagged	16	0.474	0.571	0.90
19	do, p.153	No.of h.h. in 1960	No.of h.h. in 1950	13	0.471	0.428	0.52
20	Konijn (1973)p.389	Measurement obtained in reinterview	Measurement obtained in first interview	10	0.202	0.152	0.77
21	do, p.49	Food expenditure	Total expenditure	16	0.078	0.111	0.95
22	Yamane (1967)p.334	Vaccancies	Apartment	10	0.353	0.344	0.98
23	Hanurav (1967)p.386	Artificial	Artificial	4	0.447	0.578	0.98
24	Horvitz & Thompson (1952)p.682	Eye-estimated no. of households in i-th block	No. of households in i-th block	20	0.426	0.389	0.87

Table 4.0 (contd.)

pop. no.	Source	y	x	popln. size(N)	cv(x)	cv(y)	ρ
25	Rao (1963) p.386	Acreage under grain corn in 1960	Acreage under grain corn in 1958	14	0.379	0.418	0.93
26	Hanurav (1967)p.386 (first 1-10 units)	Population in 1967	Population in 1957	10	0.171	0.205	0.92
27	do	do	do	10	0.481	0.487	0.96
28	do	do	do	10	0.733	0.739	0.99
29	do	do	do	10	0.328	0.306	1.00
30	I.S.I survey Report (1991)	Age at death	Age at which addiction started	7	0.371	0.157	0.30
31	Sarndal, Swensson & Wretman (1992)p.653	pop. in thousands 1985	pop. in thousands 1975	10	0.586	0.573	0.99
32	do			10	1.827	1.761	1.00
33	do			10	0.568	0.520	0.99
34	do			15	1.174	1.190	1.00
35	Fuller p.18	Yield	Soil nitrogen	11	0.236	0.092	0.64

Table 4.1 (n=2)

popln. no.	$V(H_0)$	$V(H_1)$	$V(H_2)$	$V(H_3)$	$V(H_4)$
1	0.129×10^{12}	0.159×10^{12}	0.373×10^{12}	0.235×10^{12}	0.169×10^{12}
2	0.101×10^{12}	0.174×10^{9}	0.379×10^{10}	0.269×10^{9}	0.157×10^{9}
3	0.257×10^{12}	0.191×10^{12}	0.354×10^{12}	0.317×10^{12}	0.263×10^{12}
4	0.207×10^{12}	0.856×10^{7}	0.320×10^{8}	0.166×10^{8}	0.451×10^{7}
9	0.410×10^{12}	0.113×10^{11}	0.150×10^{12}	0.146×10^{11}	0.181×10^{11}
10	0.408×10^{12}	0.416×10^{9}	0.286×10^{11}	0.725×10^{9}	0.782×10^{10}
11	0.334×10^{12}	0.639×10^{4}	0.583×10^{5}	0.108×10^{5}	0.689×10^{4}
13	0.525×10^{12}	0.227×10	0.159×10^{4}	0.399×10	0.202×10
18	0.249×10^{12}	0.309×10^{6}	0.753×10^{6}	0.557×10^{6}	0.247×10^{6}
19	0.311×10^{12}	0.743×10^{6}	0.159×10^{8}	0.935×10^{6}	0.145×10^{7}
23	0.124×10^{13}	0.134×10^{5}	0.724×10^{3}	0.263×10^{5}	0.140×10^{5}
25	0.287×10^{12}	0.376×10^{7}	0.372×10^{8}	0.638×10^{7}	0.902×10^{7}
26	0.744×10^{23}	0.671×10^{23}	0.191×10^{24}	0.972×10^{23}	0.437×10^{24}
27	0.211×10^{24}	0.283×10^{24}	0.275×10^{24}	0.170×10^{24}	0.204×10^{23}
28	0.205×10^{26}	0.911×10^{25}	0.669×10^{24}	0.130×10^{26}	0.124×10^{26}
29	0.181×10^{26}	0.135×10^{23}	0.496×10^{21}	0.199×10^{23}	0.730×10^{23}
32	0.181×10^{26}	0.627×10^{9}	0.824×10^{7}	0.252×10^{10}	0.554×10^{9}

Table 4.2 (n=3)

popln. no.	$V(H_0)$	$V(H_1)$	$V(H_2)$	$V(H_3)$
1	0.561×10^{11}	0.678×10^{11}	0.101×10^{12}	0.862×10^{11}
2	0.675×10^{11}	0.775×10^{8}	0.174×10^{9}	0.101×10^{9}
3	0.118×10^{12}	0.686×10^{11}	0.501×10^{11}	0.851×10^{11}
4	0.138×10^{12}	0.328×10^{7}	0.318×10^{7}	0.562×10^{7}
6	0.107×10^{12}	0.108×10^{6}	0.118×10^{7}	0.248×10^{6}
7	0.108×10^{12}	0.185×10^{11}	0.930×10^{10}	0.233×10^{11}
8	0.303×10^{12}	0.125×10^{12}	0.831×10^{12}	0.243×10^{12}
9	0.271×10^{12}	0.503×10^{10}	0.134×10^{11}	0.498×10^{10}
10	0.272×10^{12}	0.178×10^{9}	0.316×10^{9}	0.237×10^{9}
11	0.223×10^{12}	0.346×10^{4}	0.321×10^{4}	0.414×10^{4}
12	0.306×10^{12}	0.980×10^{7}	0.122×10^{8}	0.122×10^{8}
13	0.350×10^{12}	0.137×10	0.351×10	0.175×10
14	0.272×10^{12}	0.239×10^{2}	0.842×10^{2}	0.412×10^{2}
15	0.176×10^{12}	0.133×10^{10}	0.281×10^{10}	0.231×10^{10}
16	0.131×10^{12}	0.573×10^{9}	0.777×10^{8}	0.678×10^{9}
17	0.166×10^{12}	0.263×10	0.260×10^{2}	0.486×10
18	0.166×10^{12}	0.135×10^{6}	0.169×10^{6}	0.205×10^{6}
19	0.207×10^{12}	0.286×10^{6}	0.133×10^{7}	0.315×10^{6}
20	0.276×10^{12}	0.211×10^{7}	0.420×10^{6}	0.904×10^{6}
21	0.166×10^{12}	0.144×10^{7}	0.528×10^{6}	0.178×10^{7}
23	0.829×10^{12}	0.309×10^{4}	0.283×10^{3}	0.180×10^{3}
24	0.131×10^{12}	0.330×10^{4}	0.281×10^{4}	0.551×10^{4}
25	0.191×10^{12}	0.188×10^{7}	0.126×10^{7}	0.207×10^{7}
26	0.346×10^{23}	0.318×10^{23}	0.470×10^{23}	0.310×10^{23}
27	0.110×10^{24}	0.121×10^{24}	0.263×10^{23}	0.450×10^{23}
28	0.113×10^{26}	0.605×10^{25}	0.405×10^{24}	0.345×10^{25}
29	0.121×10^{26}	0.658×10^{22}	0.251×10^{21}	0.507×10^{22}
30	0.181×10^{26}	0.983×10^{4}	0.593×10^{5}	0.197×10^{5}
32	0.121×10^{26}	0.516×10^{9}	0.139×10^{7}	0.630×10^{9}
33	0.121×10^{26}	0.142×01^{5}	0.562×10^{3}	0.643×10^{4}
34	0.775×10^{25}	0.603×10^{6}	0.214×10^{4}	0.309×10^{6}
35	0.108×10^{26}	0.396×10^{4}	0.524×10^{5}	0.684×10^{4}

Table 4.2 (n=3) (contd.)

pop. no.	$V(H_4)$	$V(H_6)$	$V(H_7)$
1	0.106×10^{12}	0.606×10^{11}	0.509×10^{11}
2	0.889×10^{8}	0.719×10^{8}	0.669×10^{8}
3	0.173×10^{12}	0.611×10^{11}	0.190×10^{11}
4	0.282×10^{7}	0.288×10^{7}	0.177×10^{7}
6	0.145×10^{6}	0.108×10^{6}	0.191×10^{6}
7	0.212×10^{11}	0.161×10^{11}	0.609×10^{10}
8	0.178×10^{12}	0.127×10^{12}	0.867×10^{11}
9	0.120×10^{11}	0.462×10^{10}	0.249×10^{10}
10	0.518×10^{10}	0.175×10^{9}	0.107×10^{9}
11	0.444×10^{4}	0.328×10^{4}	0.192×10^{4}
12	0.536×10^{8}	0.931×10^{7}	0.353×10^{7}
13	0.114×10	0.130×10	0.130×10
14	0.235×10^{2}	0.232×10^{2}	0.337×10^{2}
15	0.591×10^{10}	0.122×10^{10}	0.126×10^{10}
16	0.595×10^{9}	0.520×10^{9}	0.111×10^{9}
17	0.664×10^{2}	0.239×10	0.268×10
18	0.162×10^{6}	0.123×10^{6}	0.918×10^{5}
19	0.848×10^{6}	0.244×10^{6}	0.146×10^{6}
20	0.195×10^{8}	0.210×10^{7}	0.411×10^{6}
21	0.360×10^{8}	0.143×10^{7}	0.430×10^{6}
23	0.933×10^{4}	0.285×10^{4}	0.990×10^{3}
24	0.101×10^{5}	0.318×10^{4}	0.174×10^{4}
25	0.599×10^{7}	0.190×10^{7}	0.347×10^{6}
26	0.291×10^{24}	0.322×10^{23}	0.144×10^{23}
27	0.135×10^{23}	0.111×10^{24}	0.537×10^{22}
28	0.824×10^{25}	0.595×10^{25}	0.835×10^{24}
29	0.487×10^{23}	0.660×10^{22}	0.630×10^{21}
30	0.115×10^{7}	0.947×10^{4}	0.186×10^{5}
32	0.370×10^{9}	0.522×10^{9}	0.919×10^{7}
33	0.330×10^{5}	0.143×10^{5}	0.404×10^{3}
34	0.440×10^{6}	0.609×10^{6}	0.287×10^{5}
35	0.211×10^{7}	0.402×10^{4}	0.356×10^{4}

Table 4.3 (n=4)

popln. no.	$V(H_0)$	$V(H_1)$	$V(H_2)$	$V(H_3)$
1	0.305×10^{11}	0.362×10^{11}	0.503×10^{11}	0.428×10^{11}
2	0.427×10^{11}	0.424×10^{8}	0.681×10^{8}	0.508×10^{8}
3	0.655×10^{11}	0.345×10^{11}	0.251×10^{11}	0.329×10^{11}
4	0.873×10^{11}	0.165×10^{7}	0.154×10^{7}	0.186×10^{7}
6	0.719×10^{11}	0.619×10^{5}	0.425×10^{6}	0.143×10^{6}
7	0.720×10^{11}	0.106×10^{11}	0.444×10^{10}	0.101×10^{11}
8	0.174×10^{12}	0.617×10^{11}	0.366×10^{12}	0.112×10^{12}
9	0.181×10^{12}	0.291×10^{10}	0.272×10^{10}	0.244×10^{10}
10	0.182×10^{12}	0.100×10^{9}	0.125×10^{9}	0.113×10^{9}
11	0.153×10^{12}	0.224×10^{4}	0.165×10^{4}	0.224×10^{4}
12	0.201×10^{12}	0.542×10^{7}	0.499×10^{7}	0.516×10^{7}
13	0.222×10^{12}	0.862	0.118×10	0.959
14	0.182×10^{12}	0.154×10^{2}	0.351×10^{2}	0.254×10^{2}
15	0.124×10^{12}	0.736×10^{9}	0.145×10^{10}	0.125×10^{10}
16	0.940×10^{11}	0.361×10^{9}	0.523×10^{8}	0.322×10^{9}
17	0.118×10^{12}	0.150×10	0.651×10	0.285×10
18	0.118×10^{12}	0.792×10^{5}	0.902×10^{5}	0.110×10^{6}
19	0.144×10^{12}	0.157×10^{6}	0.262×10^{6}	0.156×10^{6}
20	0.185×10^{12}	0.129×10^{7}	0.330×10^{6}	0.491×10^{6}
21	0.118×10^{12}	0.830×10^{6}	0.307×10^{6}	0.857×10^{6}
24	0.941×10^{11}	0.217×10^{4}	0.212×10^{4}	0.337×10^{4}
25	0.134×10^{12}	0.118×10^{7}	0.585×10^{6}	0.993×10^{6}
26	0.203×10^{23}	0.188×10^{23}	0.158×10^{23}	0.144×10^{23}
27	0.682×10^{23}	0.690×10^{23}	0.111×10^{23}	0.174×10^{23}
28	0.717×10^{25}	0.425×10^{25}	0.313×10^{24}	0.134×10^{25}
29	0.807×10^{25}	0.397×10^{22}	0.158×10^{21}	0.186×10^{22}
30	0.108×10^{26}	0.522×10^{4}	0.210×10^{5}	0.102×10^{5}
31	0.807×10^{25}	0.351×10^{5}	0.381×10^{4}	0.459×10^{5}
32	0.807×10^{25}	0.419×10^{9}	0.554×10^{6}	0.225×10^{9}
33	0.807×10^{25}	0.999×10^{4}	0.363×10^{3}	0.233×10^{4}
34	0.546×10^{25}	0.495×10^{6}	0.155×10^{4}	0.129×10^{6}
35	0.738×10^{25}	0.227×10^{4}	0.485×10^{4}	0.355×10^{4}

contd.

Table 4.3 (n=4) (contd.)

pop. no.	$V(H_4)$	$V(H_6)$	$V(H_7)$
1	0.780×10^{11}	0.314×10^{11}	0.277×10^{11}
2	0.628×10^{8}	0.392×10^{8}	0.380×10^{8}
3	0.129×10^{12}	0.265×10^{11}	0.108×10^{11}
4	0.207×10^{7}	0.138×10^{7}	0.950×10^{6}
6	0.967×10^{5}	0.614×10^{5}	0.878×10^{5}
7	0.158×10^{11}	0.909×10^{10}	0.473×10^{10}
8	0.118×10^{12}	0.624×10^{11}	0.426×10^{11}
9	0.898×10^{10}	0.254×10^{10}	0.166×10^{10}
10	0.388×10^{10}	0.996×10^{8}	0.616×10^{8}
11	0.329×10^{4}	0.213×10^{4}	0.141×10^{4}
12	0.399×10^{8}	0.519×10^{7}	0.242×10^{7}
13	0.801	0.812	0.797
14	0.168×10^{2}	0.149×10^{2}	0.190×10^{2}
15	0.440×10^{10}	0.707×10^{9}	0.684×10^{9}
16	0.446×10^{9}	0.328×10^{9}	0.117×10^{9}
17	0.428×10^{2}	0.131×10	0.131×10
18	0.121×10^{6}	0.706×10^{5}	0.582×10^{5}
19	0.606×10^{6}	0.123×10^{6}	0.755×10^{5}
20	0.146×10^{8}	0.129×10^{7}	0.328×10^{6}
21	0.270×10^{8}	0.830×10^{6}	0.356×10^{6}
24	0.759×10^{4}	209×10^{4}	0.131×10^{4}
25	0.449×10^{7}	0.124×10^{7}	0.238×10^{6}
26	0.218×10^{24}	0.193×10^{23}	0.103×10^{23}
27	0.101×10^{23}	0.581×10^{23}	0.453×10^{22}
28	0.618×10^{25}	0.424×10^{25}	0.857×10^{24}
29	0.365×10^{23}	0.407×10^{22}	0.770×10^{21}
30	0.826×10^{6}	0.493×10^{4}	0.727×10^{4}
31	0.182×10^{6}	0.316×10^{5}	0.117×10^{5}
32	0.277×10^{9}	0.424×10^{9}	0.814×10^{7}
33	0.248×10^{5}	0.102×10^{5}	0.470×10^{3}
34	0.330×10^{6}	0.504×10^{6}	0.334×10^{5}
35	0.157×10^{7}	0.238×10^{4}	0.193×10^{4}

Table 4.4 (N=7,n=3)

popln. no.	$V(H_0)$	$V(H_1)$	$V(H_2)$	$V(H_3)$
1	0.597×10^{11}	0.664×10^{11}	0.720×10^{11}	0.680×10^{11}
2	0.772×10^{11}	0.782×10^{8}	0.815×10^{8}	0.681×10^{8}
3	0.125×10^{12}	0.567×10^{11}	0.697×10^{11}	0.100×10^{12}
4	0.152×10^{12}	0.293×10^{7}	0.378×10^{7}	0.571×10^{7}
5	0.152×10^{12}	0.175×10^{5}	0.105×10^{5}	0.198×10^{5}
6	0.152×10^{12}	0.564×10^{5}	0.118×10^{7}	0.104×10^{6}
7	0.166×10^{12}	0.150×10^{11}	0.864×10^{10}	0.166×10^{11}
8	0.238×10^{12}	0.672×10^{11}	0.494×10^{12}	0.197×10^{12}
9	0.273×10^{12}	0.212×10^{10}	0.668×10^{9}	0.139×10^{10}
10	0.273×10^{12}	0.218×10^{9}	0.689×10^{9}	0.342×10^{9}

Table 4.4 (N=7,n=3) (contd.)

pop. no.	$V(H_4)$	$V(H_5)$
1	0.107×10^{12}	0.474×10^{11}
2	0.764×10^{8}	0.558×10^{8}
3	0.171×10^{12}	0.107×10^{12}
4	0.294×10^{7}	0.418×10^{7}
5	0.121×10^{7}	0.132×10^{5}
6	0.586×10^{5}	0.219×10^{7}
7	0.135×10^{11}	0.158×10^{11}
8	0.967×10^{11}	0.557×10^{12}
9	0.103×10^{11}	0.106×10^{10}
10	0.723×10^{10}	0.229×10^{10}

Table 4.5 (N=8,n=4)

popln. no.	$V(H_0)$	$V(H_1)$	$V(H_2)$	$V(H_3)$
1	0.151×10^7	0.163×10^7	0.156×10^7	0.186×10^7
2	0.305×10^{11}	0.362×10^{11}	0.503×10^{11}	0.428×10^{11}
3	427×10^{11}	0.424×10^8	0.681×10^8	0.508×10^8
4	0.655×10^{11}	0.345×10^{11}	0.251×10^{11}	0.329×10^{11}
5	0.873×10^{11}	0.165×10^7	0.154×10^7	0.186×10^7
6	0.150×10^{12}	0.617×10^{11}	0.366×10^{12}	0.112×10^{12}
7	0.194×10^{12}	0.862	0.118×10	0.959
8	0.194×10^{12}	0.171×10^2	0.214×10^2	0.186×10^2
9	0.194×10^{12}	0.832×10^9	0.145×10^{10}	0.122×10^{10}
10	0.195×10^{12}	0.389×10^9	0.558×10^8	0.174×10^9

Table 4.5 (N=8,n=4) (contd.)

pop. no.	$V(H_4)$	$V(H_5)$
1	0.218×10^7	184×10^7
2	0.780×10^{11}	0.504×10^{11}
3	0.628×10^8	0.589×10^8
4	0.129×10^{12}	0.263×10^{11}
5	0.207×10^7	0.148×10^7
6	0.118×10^{12}	0.462×10^{12}
7	0.801	0.145×10
8	0.191×10^2	0.247×10^2
9	$0.0.496 \times 10^{10}$	0.126×10^{10}
10	0.719×10^9	0.534×10^8

Table 4.6 (N=20)

n=4

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.56	0.56	*	0	0.56	0.56	1.48	0	0.56	0.56	1.31
1	3.45	4.38	4.36	1	11.38	13.15	17.39	1	23.81	26.73	39.10
2	70.61	118.42	42.92	2	488.56	663.48	458.15	2	1746.32	2194.63	1987.67

n=8

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.18	0.18	*	0	0.18	0.18	0.64	0	0.18	0.18	0.58
1	1.17	1.78	1.96	1	3.79	5.07	7.84	1	7.85	10.15	17.64
2	25.40	50.87	20.49	2	171.18	269.59	214.63	2	601.82	871.35	922.35

n=12

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.08	0.08	*	0	0.08	0.08	0.41	0	0.08	0.08	0.37
1	0.51	0.96	1.27	1	1.62	2.47	5.07	1	3.36	4.77	11.41
2	11.16	28.54	13.49	2	74.62	138.83	140.46	2	261.04	431.33	601.90

Table 4.7 (N=40)

n=8

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.23	0.23	*	0	0.23	0.23	0.66	0	0.23	0.23	0.60
1	1.54	1.93	2.05	1	4.95	6.11	8.21	1	10.24	12.69	18.46
2	33.64	47.68	21.77	2	225.77	291.43	226.79	2	791.54	1000.86	972.24

n=16

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.08	0.08	*	g	0.08	0.08	0.31	0	0.08	0.08	0.29
1	0.55	0.79	0.99	1	1.76	2.36	3.94	1	3.63	4.83	8.87
2	12.40	19.79	10.71	2	82.32	115.00	110.74	2	286.49	388.16	472.82

n=24

	r=2				r=4				r=6		
g	ev1	ev2	ev3	g	ev1	ev2	ev3	g	ev1	ev2	ev3
0	0.04	0.04	*	0	0.04	0.04	0.21	0	0.04	0.04	0.19
1	0.24	0.42	0.65	1	0.77	1.13	2.60	1	1.58	2.24	5.84
2	5.48	10.53	7.11	2	36.27	56.28	73.28	2	125.93	184.12	312.52

The values obtained in the previous version of this table have been thoroughly checked and are given above. In the above table $\star$ indicates undefined values.

Chapter 5

ESTIMATING A FINITE POPULATION VARIANCE UNDER SOME GENERAL LINEAR MODELS WITH EXCHANGEABLE ERRORS

5.1. Introduction and Summary

This chapter considers estimation of a finite population variance under certain class of superpopulation models. We have considered here superpopulation models with exchangeable errors. The models are depicted in (5.2.2) and (5.2.3). Optimal strategies have been obtained under these models and their robustness under a general class of alternative polynomial regression models have been considered. Sampling designs ensuring near robustness of these strategies have been investigated. Results of this chapter were published in Mukhopadhyay and Bhattacharyya (1990-91).

5.2 Superpopulation Models and Formulation of the Problem:

Let, as before, $\mathcal{U}$ denote a finite population of N identifiable units labelled $1,...,k,...,N$. Assciated with k are two real quantities y_k, x_k, values of main variable 'y' and a closely related variable 'x' respectively.

We assume that corresponding to each value y_i there is a random variable Y_i whose one particular value is y_i. $\underset{\sim}{y} = (y_1,...,y_N)$ can then be looked upon as a particular realisation of a random vector $\underset{\sim}{Y} = (Y_1,...,Y_N)$, [$Y_i$ being the random variable corresponding to y_i], having a probability distribution $\zeta_{\underset{\sim}{\theta}}$ indexed by a parameter vector $\underset{\sim}{\theta} \in \Theta$ the parameter space. It is required to predict

$$V(\underset{\sim}{Y}) = \frac{1}{N}\sum_{k=1}^{N}(Y_k-\overline{Y})^2$$

$$= a_1\sum_{k=1}^{N} Y_k^2 - a_2 \sum_{k\neq k'=1}^{N}\sum Y_k Y_{k'} \qquad (5.2.1)$$

where

$$a_1 = \frac{1}{N}(1-\frac{1}{N}),\ a_2 = \frac{1}{N^2},\ \overline{Y} = \frac{1}{N}\sum_{k=1}^{N} Y_k$$

which is a random variable, on the basis of the observed data $\{(k,y_k), k \in S\}$ and the assumed prior distribution of $\underset{\sim}{Y}$. We shall consider the following superpopulation models.

Let $\underset{\sim}{Y}$ follow a N-variate normal distribution with

$$\begin{aligned}\mathcal{E}(Y_k|x_k) &= \mu_k, \\ \mathcal{V}(Y_k|x_k) &= v_k, \\ \mathcal{C}(Y_k, Y_{k'}|x_k, x_{k'}) &= c_{kk'} = \rho\sqrt{(v_k v_{k'})}, \\ & -\frac{1}{N-1} \leq \rho \leq 1.\end{aligned}$$

where $\mathcal{E}, \mathcal{V}, \mathcal{C}$ denote respectively expectation, variance and covariance wrt superpopulation model. We denote the above model with exchangeable errors as $\xi(\mu_k, v_k, \rho) = \xi$ (say).

Following Royall and Herson (1973) we denote a model ξ with $\mu_k = \sum_0^J \delta_j \beta_j x_k^j, v_k = \sigma^2 \sum_0^L \delta_l^{'} \gamma_l x_k^l$ as $\xi(\delta_0, ..., \delta_J; \delta_0^{'}, ..., \delta_L^{'}; \rho)$ where $\delta_j = 1(0)$ if x_k^j is present (absent) in μ_k, $\delta_l^{'} = 1(0)$ if x_k^l is present (absent) in v_k, γ_l is a known non-negative constant (l=0,1,...,L), $\beta_j (j = 0, 1, ..., J), \sigma^2(> 0), \rho$ are unknown.We denote $\xi(1; 1; \rho) = \xi_1$ and $\xi(1, 1; 1; \rho) = \xi_2$ and shall consider prediction of V under these two superpopulation models with exchangeable errors.

Thus for ξ_1, Y_k's are independent normal with

$$\begin{aligned}\mathcal{E}(Y_k|x_k) &= \beta \\ \mathcal{V}(Y_k|x_k) &= \sigma^2, k = 1, 2, ..., N \qquad (5.2.2) \\ \mathcal{C}(Y_k, Y_{k'}|x_k, x_{k'}) &= \rho\sigma^2, k \neq k' = 1, 2, ..., N\end{aligned}$$

and for ξ_2, Y_k's are independent normal with

$$\begin{aligned}\mathcal{E}(Y_k|x_k) &= \alpha + \beta x_k \\ \mathcal{V}(Y_k|x_k) &= \sigma^2, k = 1, 2, ..., N \qquad (5.2.3) \\ \mathcal{C}(Y_k, Y_{k'}|x_k, x_{k'}) &= \rho\sigma^2, \ k \neq k' = 1, 2, ..., N\end{aligned}$$

Optimal design-(p-) unbiased, model-(m-) unbiased, design-model-(pm-) unbiasd estimation of V under model $\xi(0; \sigma^2 w(x); 0)$, w(x) a known function of x was considered by Mukhopadhyay (1978, 1982). In the sequel we shall consider the following purposive sampling designs :

(a) The design p_x^* is defined as

$$p_x^*(S) = \begin{cases} 1 & \text{for } S = S_0 \\ 0 & \text{otherwise} \end{cases} \qquad (5.2.4)$$

where S_0 is such that,

$$s_x^2(S_0) = \min_{S \in S_n} s_x^2(S)$$

, where

$$s_x^2(S) = \frac{1}{n-1}\sum_{i\in S}(x_i - \overline{x}_S)^2, \overline{x}_S = \frac{1}{n}\sum_{k\in S} x_k$$

(S_n being the set of all samples of size n), i.e., the sampling variance of x on S_0 is the minimum among sampling variance of x on all possible samples of size n.

(b) The design p_x^{**} is defined as

$$p_x^{**}(S) = \begin{cases} 1 & \text{for } S = S_1 \\ 0 & \text{otherwise} \end{cases} \qquad (5.2.5)$$

where S_1 is a sample such that $s_x^2(S_1) = max_{S\in S_n} s_x^2(S)$, i.e., the sample variance of x on S_1 is the maximum among sample variances of x on all possible sample of size n.

5.3. Prediction Under ξ_1 :

We recall a few definition.

Definition 5.1 A predictor $e(S, \underset{\sim}{Y})$ is p-unbiased for $V(\underset{\sim}{Y})$ if

$$E[e(S, \underset{\sim}{y})] = V(\underset{\sim}{Y})$$

i.e., if

$$E[e(S, \underset{\sim}{y})] = V(\underset{\sim}{y}) \quad \forall \ \underset{\sim}{y} \in R_N.$$

Definition 5.2 A predictor $e(S, \underset{\sim}{Y})$ is model(m-) unbiased for $V(\underset{\sim}{Y})$ if

$$\mathcal{E}[e(S, \underset{\sim}{Y})] = \mathcal{E}[V(\underset{\sim}{Y})] \ \ \forall \ S : p(S) > 0.$$

Definition 5.3 A predictor $e(S, \underset{\sim}{Y})$ is design-model (pm-) unbiased for $V(\underset{\sim}{Y})$ if

$$\mathcal{E}E[(e(S, \underset{\sim}{Y}) - V(\underset{\sim}{Y})] = 0.$$

5.3.1. Optimal prediction of V under ξ_1 :

Under ξ_1,

$$\begin{aligned} \mathcal{E}(V) &= a_1 N\sigma^2 - a_2 N(N-1)\rho\sigma^2 \\ &= \frac{N-1}{N}(1-\rho)\sigma^2 = \tau_1 \text{ (say)} \end{aligned}$$

It is known [Arnold (1979)] that under ξ_1,

$$s^2 = \frac{1}{n-1}\sum_{k\in S}(y_k - \overline{y}_S)^2, \overline{y}_S = \frac{1}{n}\sum_{k\in S} y_k,$$

is unbiased and complete sufficient for $(1-\rho)\sigma^2$.

Hence $\frac{N-1}{N}s^2 = e_1^*$ (say) is UMVU predictor (UMVUP) of V in the class of all unbiased predictors of V.

Under ξ_1, $(n-1)s^2 \sim \sigma^2(1-\rho)\chi^2_{(n-1)}$. Therefore

$$\begin{aligned} \mathcal{V}(e_1^*) &= (\frac{N-1}{N})^2\frac{2}{n-1}(1-\rho)^2\sigma^4 \\ &= \lambda_n \text{ (say)} \qquad (5.3.1) \end{aligned}$$

a constant dependent only on n.

Hence we have the following

Theorem 5.1. Under ξ_1, for any given $p \in \rho_n$

$$E_p\mathcal{E}(e_1^* - \tau_1)^2 \leq E_p\mathcal{E}(e-\tau_1)^2$$

for any m-unbiased predictor e. Again any $p \in \rho_n$ is optimal for using e_1^*.

Note 5.1. We note that the predictor e_1^* possesses the following minimax property.

$$max_{p\in\rho_n} R(e_1^*, p) = \underset{e\in\mathcal{M}}{Min}\ \underset{p\in\rho_n}{Max}\ R(e,p)$$

where $\mathcal{M}$ is the class of ξ_1-unbiased predictor of V and R(e,p) denotes $E_p\mathcal{E}(e-\tau_1)^2$.

Note 5.2. Although any $p \in \rho_n$ is equally efficient, considerations of robustness, following Godambe and Thompson (1977), with respect to an alternative class of superpopulation models suggests a purposive design $p^* \in \rho_n$ as optimal to use e_1^* as has been noted in section 5.3.3.

Note 5.3 We note that the results in this section are valid even if the normality assumption in the model (5.2.2) is relpaced by a more general asssumption that $Y_1, ..., Y_N$ have an exchangeable absolutely continuous distribution and no distributional assumption is made about $\xi(1,1;\rho)$.

5.3.2. Bias of e_1^* under a class of alternative models:

We consider the following general class of alternative models

$$\xi'(\delta_0, \cdots, \delta_J; \delta'_0, \cdots, \delta'_L; \rho)$$

Bias of e_1^* for a given sample S under $\xi'(\delta_0, \cdots, \delta_j; \delta'_0, \cdots, \delta'_L; 0)$

$$\begin{aligned} \mathcal{B}_{\xi'(\delta_0,\cdots,\delta_J;\delta'_0,\cdots,\delta'_L;0)}(e_1^*) &= \mathcal{E}_{\xi'(\delta_0,\cdots,\delta_J;\delta'_0,\cdots,\delta'_L;0)}(e_1^* - V) \\ &= \mathcal{E}_{\xi'}[(\frac{N-1}{N})(\frac{1}{n}\sum_S Y_k^2 - \frac{1}{n(n-1)}\sum_S{}' Y_kY_k') \end{aligned}$$

$$-(\frac{N-1}{N})(\frac{1}{N}\sum Y_k^2 - \frac{1}{N(N-1)}\sum' Y_k Y_k')]$$

$$= (\frac{N-1}{N})[\frac{1}{n}\sum_S\{\sigma^2\sum_{l=0}^{L}\delta_l'\gamma_l x_k^l + (\sum_{j=0}^{J}\delta_j\beta_j x_k^j)^2\}$$

$$-\frac{1}{n(n-1)}\sum_S'(\sum_{j=0}^{J}\delta_j\beta_j x_k^j)(\sum_{j=0}^{J}\delta_j\beta_j x_{k'}^j)$$

$$-\frac{1}{N}\sum\{\sigma^2\sum_{l=0}^{L}\delta_l'\gamma_l x_k^l + (\sum_{j=0}^{J}\delta_j\beta_j x_k^j)^2\}$$

$$+\frac{1}{N(N-1)}\sum'(\sum_{j=0}^{J}\delta_j\beta_j x_k^j)(\sum_{j=0}^{J}\delta_j\beta_j x_{k'}^j)]$$

$$= (\frac{N-1}{N})[\sigma^2\sum_{l=0}^{L}\delta_l'\gamma_l\{\frac{1}{n}\sum_S x_k^l - \frac{1}{N}\sum x_k^l\}$$

$$+\frac{1}{n-1}\sum_S\{\sum_{j=0}^{J}\delta_j\beta_j x_k^j - (\frac{1}{n}\sum_S\sum_{j=0}^{J}\delta_j\beta_j x_k^j)\}^2$$

$$-\frac{1}{N-1}\sum\{\sum_{j=0}^{J}\delta_j\beta_j x_k^j - (\frac{1}{N}\sum\sum_{j=0}^{J}\delta_j\beta_j x_k^j)\}^2]$$

$$= (\frac{N-1}{N})[\sigma^2\sum_{l=0}^{L}\delta_l'\gamma_l\{\frac{1}{n}\sum_S x_k^l - \frac{1}{N}\sum x_k^l\}$$

$$+\frac{1}{n-1}\sum_S\{\sum_{j=0}^{J}\delta_j\beta_j(x_k^j - \overline{x}_S^{(j)})\}^2$$

$$-\frac{1}{N-1}\sum\{\sum_{j=0}^{J}\delta_j\beta_j(x_k^j - \overline{X}^{(j)})\}^2,$$

$$\text{where } \overline{x}_S^{(j)} = \frac{1}{n}\sum_S x_k^j$$

$$\text{and } \overline{X}^{(j)} = \frac{1}{N}\sum x_k^j$$

$$= (\frac{N-1}{N})[\sigma^2\sum_{l=0}^{L}\delta_l'\gamma_l(\overline{x}_S^{(l)} - \overline{X}^{(l)}) + \sum_{j=0}^{J}\delta_j\beta_j^2(s_{x(j)}^2 - S_{x(j)}^2)$$

$$+\sum_{j\neq j'=0}^{J}\sum\delta_j\delta_{j'}\beta_j\beta_{j'}(s_{x(jj')} - S_{x(jj')})]$$

where

$$(n-1)s^2_{x(j)} = \sum_S (x^j_k - \overline{x}^{(j)}_S)^2 \qquad (5.3.2)$$

$$(N-1)S^2_{x(j)} = \sum (x^j_k - \overline{X}^{(j)})^2 \qquad (5.3.3)$$

$$(n-1)s_{x(jj')} = \sum_S (x^j_k - \overline{x}^{(j)}_S)(x^{j'}_k - \overline{x}^{(j')}_S) \qquad (5.3.4)$$

$$(N-1)S_{x(jj')} = \sum (x^j_k - \overline{X}^{(j)})(x^{j'}_k - \overline{X}^{(j')}) \qquad (5.3.5)$$

Thus

$$\mathcal{B}_{\xi'}(e^*_1) = (\frac{N-1}{N})[B(\delta_0,\cdots,\delta_J) + B'(\delta'_0,\cdots,\delta'_L)] \qquad (5.3.6)$$

where

$$B(\delta_0,\cdots,\delta_J) = \sum_{j=0}^{J} \delta_j \beta^2_j (s^2_{x(j)} - S^2_{x(j)})$$

$$+ \sum\sum_{j\neq j'=0}^{J} \delta_j \delta_{j'} \beta_j \beta_{j'} (s_{x(jj')} - S_{x(jj')}) \qquad (5.3.7)$$

and

$$B'(\delta'_0,\cdots,\delta'_L) = \sigma^2 \sum_{l=0}^{L} \delta'_l \gamma_l (\overline{x}^{(l)}_S - \overline{X}^{(l)}) \qquad (5.3.8)$$

Thus a set of suficient conditions for a sample S to preserve the unbiasedness of e^*_1 under alternative models $\xi'(\delta_0,\cdots,\delta_J;\delta'_0,\cdots,\delta'_L;0)$ is

$$\overline{x}^{(l)}_S = \overline{X}^{(l)} \ \forall\, l \text{ for which } \ \delta'_l = 1, l = 1,2,...,L \qquad (5.3.9)$$

$$s^2_{x(j)} = S^2_{x(j)} \ \forall\, j \text{ for which } \ \delta_j = 1, j = 1,2,...,J \qquad (5.3.10)$$

$$s_{x(jj')} = S_{x(jj')} \ \forall\, j,j' \text{ for which } \ \delta_j = 1, \delta_{j'} = 1, j \neq j' = 1,2,...,J \quad (5.3.11)$$

Condition (5.3.9) is similar to the conditions of Royall and Herson's (1973) balanced sampling design (as discussed in chapter 3 of this thesis). Conditions (5.3.10) and (5.3.11) are additional conditions of balancing required on the sample. We shall call conditions (5.3.10) and (5.3.11) as "variance-covariance balance of order J". A sample S satisfying conditions (5.3.9)-(5.3.11) may be called a strongly balanced (s.b.) sample. A s.b. sampling design p_B (say) may be defined as one for which $p_B(S) = \frac{1}{k}(0)$ for each of k s.b. samples $S = S_{B_i}, i = 1,2,...,k$ (otherwise).

Obviously a s.b. design will very often be conspicuous by its absence.

Note 5.3. We have

$$\mathcal{B}_{\xi'(\delta_0,\cdots,\delta_J;1;\rho)}(e_1^*) = \frac{N-1}{N}B(\delta_0,\cdots,\delta_J).$$

Note 5.4. Similar results can be obtained when

$$\mu_k = \sum_{j=0}^{J}\beta_j x_{kj}$$

ie., y has a multiple regression on (J+1) auxiliary variables $x_0, x_1, \cdots, x_J$, x_{kj} being the value (assumed known) of auxiliary variable x_j on $k, k = 1, 2, ..., N$.

5.3.3. Robustness of e_1^* under a class of alternative models:

As suggested by a referee we consider a measure of robustness (sensitivity) of an estimstor e^* which is optimum under ν, wrt alternative ν_1, for a given sample S as

$$\Delta_S = |\mathcal{E}_\nu(e^* - \tau)^2 - \mathcal{E}_{\nu_1}(e^* - \tau)^2| \qquad (5.3.12)$$

where τ is the parameter of interest. Less is the value of Δ_S more (less) is the robustness (sensitivity) of e^* wrt ν_1.

We shall consider robustness of e_1^* under the model $\xi'(\mu_k;1;\rho) = \xi'$ (say). e_1^* is biased under ξ'.

Now under $\xi'(\mu_k;1;\rho)$, $(n-1)s^2 \sim \sigma^2(1-\rho)\chi'^2_{(n-1),\lambda}$ where $\lambda = \frac{(n-1)s_\mu^2}{\sigma^2(1-\rho)}$, the non-centrality parameter. Hence,

$$\begin{aligned}\mathcal{E}_{\xi'}(e_1^* - \tau_1)^2 &= \mathcal{V}_{\xi'}(e_1^*) + [\mathcal{E}_{\xi'}(e_1^* - \tau_1)]^2 \\ &= (\frac{N-1}{N})^2[\frac{2}{n-1}\{\sigma^4(1-\rho)^2 + 2\sigma^2(1-\rho)s_\mu^2\} + s_\mu^4]\end{aligned}$$

Hence increase in the measure $\mathcal{E}(e_1^* - \tau_1)^2$ in using e_1^* from ξ_1 to ξ' is

$$\Delta_S = (\frac{N-1}{N})^2 s_\mu^2\{\frac{4}{n-1}\sigma^2(1-\rho) + s_\mu^2\} \qquad (5.3.13)$$

(5.3.13) shows that $\Delta_S(e_1^*|\xi_1,\xi')$ is minimised for the sample for which s_μ^2 is minimum. Thus an optimum design in the class ρ_n to use the (biased) estimator e_1^* under the general class of alternative models $\xi'(\mu_i;1;\rho)$, (optimum in the sense of robustness) is a purposive design, p_μ^* (say) where $p_\mu^*(\in \rho_n)$ is as defined in (5.2.4) ie.,

$$p_\mu^*(S) = \begin{cases} 1 & \text{for } S = S^* \\ 0 & \text{otherwise} \end{cases}$$

$S^\star$ being a sample such that

$$s_\mu^2(S^\star) = min_{S\in S_n} s_\mu^2(S),$$

Hence we conclude as follows : Suppose from the apriori information one believes his true model is ξ_1 and hence uses $e_1^\star$. Later the sampler for some reasons doubts the correctness of his assumed model and suspects it to be $\xi'(\mu_i; 1; \rho)$. under $\xi'(\mu_i; 1; \rho), v'(e_1^\star)$ will be minimum for $p_\mu^\star$. Again though any $p \in \rho_n$ (including $p_\mu^\star$) is equally efficient for $e_1^\star$ under $\xi_1, p_\mu^\star$ is the most robust one in ρ_n in the sense of minimising the absolute value of difference in model mean square errors, $\Delta(e_1^\star|\xi_1, \xi')$ and hence even under ξ_1 one should use $p_\mu^\star$ to take care of the contingencies due to the failure of the supposed model from ξ_1 to $\xi'(\mu_i; 1; \rho)$.

In particular if $\mu_i = \beta_0 + \beta_1 x_i$, a robust design to use $e_1^\star$ under ξ' is $p_x^\star$, as defined in (5.2.4).

We now consider the robustness of $e_1^\star$ under model $\xi'(\mu_i; g_i; 0) = \xi''$ (say).

$$\begin{aligned}\mathcal{E}_{\xi'(\mu_i;g_i;0)}(e_1^\star - \tau_1)^2 &= (\frac{N-1}{N})^2 \frac{2\sigma^4}{n^2(n-1)^2}[\frac{1}{n^2}\{n(n-2)\sum_S g_i^2 \\ &+(\sum_S g_i)^2\} + \frac{2n^2}{\sigma^2}\sum_S g_i(\mu_i - \overline{\mu}_S)^2 \\ &+\{s_\mu^2 + \sigma^2\frac{1}{n}\sum_S g_i - \sigma^2(1-\rho)\}^2]\end{aligned}$$

Hence a robust sampling design to use $e_1^\star$ under model ξ'' wrt change in value of model-variance of $e_1^\star$ is $p'^\star$ (say) where $p'^\star \in \rho_n$ and

$$p^{\star\prime}(S) = \begin{cases} 1 & \text{for } S = S^{\star\prime} \\ 0 & \text{otherwise} \end{cases} \qquad (5.3.28)$$

where $S^{\star\prime}$ is such that

$$(\sum_{S^{\star\prime}} g_i)^2 = min_{S\in S_n}(\sum_S g_i)^2$$

$$\sum_{S^{\star\prime}} g_i(\mu_i - \overline{\mu}_{S^{\star\prime}}) = min_{S\in S_n}\sum_S g_i(\mu_i - \overline{\mu}_S)^2$$

$$\sum_{S^{\star\prime}}(s_\mu^2 + \sigma^2\frac{1}{n}\sum_{S^{\star\prime}} g_i)^2 = min_{S\in S_n}\sum_S(s_\mu^2 + \sigma^2\frac{1}{n}\sum_S g_i)^2$$

We can also define, following Godambe and Thompson (1977) a measure of robustness (sensitivity) of an estimator e^* which is optimum under η, wrt, alternative η_1, for a given sample S as

$$\Delta_S = |\mathcal{V}_\eta(e^*) - \mathcal{V}_{\eta_1}(e^*)|$$

The designs which are optimum for the measure defined in (5.3.12) are also optimum under this measure.

5.4 Optimal prediction under ξ_2

We have under ξ_2

$$\begin{aligned} \mathcal{E}(V) &= (\frac{N-1}{N})\mathcal{E}\{\frac{1}{N}\sum Y_i^2 - \frac{1}{N(N-1)}\sum' Y_iY_j\} \\ &= (\frac{N-1}{N}\{\sigma^2(1-\rho) + \beta^2 V_x\} \\ &= \tau_2 \quad \text{(say)} \qquad\qquad (5.4.1) \end{aligned}$$

$$\text{where, } V_x = \frac{1}{N-1}\sum(x_i - \bar{X})^2, \bar{X} = \frac{1}{N}\sum x_i.$$

It is known [Arnold (1979)], the statistic ($\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}_y^2$) is unbiased and complete sufficient for (α, β, $(1-\rho)\sigma^2$), where

$$\hat{\alpha} = \bar{y}_S - \hat{\beta}\bar{x}_S, \quad \bar{y}_S = \frac{1}{n}\sum_S y_i, \quad \bar{x}_S = \frac{1}{n}\sum_S x_i$$

$$\hat{\beta} = \frac{\sum_S(x_i - \bar{x}_S)y_i}{\sum_S(x_i - \bar{x}_S)^2} = \frac{s_{xy}}{s_{xx}} \quad \text{(say)} \qquad (5.4.2)$$

and

$$\hat{\sigma}_y^2 = \frac{1}{n-2}\sum_S(y_i - \hat{\alpha} - \hat{\beta}x_i)^2$$

Thus $\hat{\alpha}, \hat{\beta}, \hat{\sigma}_y^2$ are minimum variance (model-)unbiased estimator of α, β and $(1-\rho)\sigma^2$ respectively. Also

$$\begin{aligned} \mathcal{V}(\hat{\beta}) &= \frac{1}{(s_{xx})^2}[\sum_S(x_i - \bar{x}_S)^2\mathcal{V}(Y_i) \\ &\quad + \sum_S{}'(x_i - \bar{x}_S)(x_j - \bar{x}_S)C(Y_i, Y_j) \\ &= \frac{1}{(s_{xx})^2}[\sum_S(x_i - \bar{x}_S)^2\sigma^2 \end{aligned}$$

$$+\rho\sigma^2 \sum_S (x_i - \overline{x}_S)(x_j - \overline{x}_S)]$$
$$= \frac{(1-\rho)\sigma^2}{s_{xx}^2} \sum_S (x_i - \overline{x}_S)$$
$$= \frac{(1-\rho)\sigma^2}{s_{xx}}$$

Hence an optimal estimator of β^2 is

$$\tilde{\beta}^2 = \hat{\beta}^2 - \frac{\delta_y^2}{s_{xx}} \tag{5.4.3}$$

Consequently an optimal estimator of τ_2 is given by

$$\frac{N-1}{N}(\delta_y^2 + V_x(\hat{\beta}^2 - \frac{\delta_y^2}{s_{xx}})) = e_2^* \quad \text{(say)} \tag{5.4.4}$$

Or,

$$e_2^* = (\frac{N-1}{N})(\delta_y^2(1 - \frac{V_x}{s_{xx}}) + V_x\hat{\beta}^2)$$

Under ξ_2, $\hat{\beta}$ and δ_y^2 are independently distributed with $\hat{\beta} \sim N(\beta, \frac{\sigma^2(1-\rho)}{s_{xx}})$ and $(n-2)\delta_y^2 \sim \sigma^2(1-\rho)\chi^2_{(n-2)}$. Therefore

$$\begin{aligned}\mathcal{V}_{\xi_2}(e_2^*) &= (\frac{N-1}{N})^2[(1 - \frac{V_x}{s_{xx}})^2\mathcal{V}(\delta_y^2) + V_x^2\mathcal{V}(\hat{\beta}^2)] \\ &= (\frac{N-1}{N})^2[(1 - \frac{V_x}{s_{xx}})^2\frac{2\sigma^4(1-\rho)^2}{(n-2)} \\ &\quad + V_x^2(2 + 4\frac{\beta^2 s_{xx}}{\sigma^2(1-\rho)}).\frac{\sigma^4(1-\rho)^2}{s_{xx}^2}]\end{aligned} \tag{5.4.5}$$

Here it seems that the purposive sampling design $p_x^{**} \in \rho_n$, as defined in (5.2.5) is likely to be optimum one in the sense of smaller variance. We can not, however, completely recommend the purposive design p_x^{**} as the optimal one, because of the term $(-\frac{V_x}{s_{xx}})$ appearing in the curly bracket in (5.4.5).

5.5 Some Approximately S.B. Designs :

In predicting population total under a super-population model Royall and Herson (1973) introduced the concept of balanced sampling designs in order to keep the ratio predictor unbiased under the class of alternative polynomial regression models as discussed in chapter 3. The conditions for balance have been given in section 3.3.1. As is obvious, balanced samples are, in general, conspicuous by their absence.

Royall and Cumberland (1981) proposed a restricted and purposive design to guarantee (first moment) nearly balanced samples and hence some robustness against model breakdown. Their scheme consists of the following two steps :

(i) a sample S is selected with simple random sampling out of $c = \binom{N}{n}$ possible samples ;

(ii) the sample S is accepted if $|\overline{x}_S - \overline{X}| < \delta$, otherwise (i) is repeated.

in this procedure probability of selecting S is

$$p(S) = \begin{cases} 0 & \text{if } |\overline{x}_S - \overline{X}| \geq \delta \\ \frac{1}{c_\delta} & \text{otherwise.} \end{cases}$$

where $c_\delta = \#\{S : |\overline{x}_S - \overline{X}| < \delta\}$.

This procedure, however, has some shortcoming as pointed out by Sarndal (1981), of having zero probabilities of inclusion for some units. Iachen (1985) proposed an alternative sampling design by assigning probability inversely proportional to $|\overline{x}_S - \overline{X}|$ to these samples S that do not pass the test $|\overline{x}_S - \overline{X}| < \delta$.

Thus

$$p(S) \propto \frac{\delta}{|\overline{x}_S - \overline{X}|} \quad \text{if} \quad |\overline{x}_S - \overline{X}| \geq \delta.$$

Thus their design becomes

$$p_1(S) = \begin{cases} \lambda_1 & \text{if } |\overline{x}_S - \overline{X}| < \delta \\ \lambda_1 \frac{\delta}{|\overline{x}_S - \overline{X}|} & \text{otherwise} \end{cases}$$

The constant $\lambda_1 = \lambda_1(\delta)$ has to be determined from the constraint $\sum_S p(S) = 1$.

Employing the similar technique we hereby derive some approximately strongly balanced designs applicable for certain models.

We have for a given S,

$$B_{\xi'(1,1;1;0)}(e_1^*) = \frac{N-1}{N} \beta^2 d_S(s_x^2) \tag{5.5.1}$$

where $d_S(s_x^2) = s_x^2 - S_x^2$

Similarly,

$$B_{\xi'}(1;1,1;0)(e_1^*) = \frac{N-1}{N} \sigma^2 \gamma_1 d_S(\overline{x}) \tag{5.5.2}$$

where $d_S(\overline{x}) = \overline{x}_S - \overline{X}$ and

$$B_{\xi'(1,1;1,1;0)}(e_1^*) = \frac{N-1}{N} [\sigma^2 \gamma_1 d_S(\overline{x}) + \beta^2 d_S(s_x^2)] \tag{5.5.3}$$

To make $B_{\xi'}(1,1;1;0)(e_1^*) \simeq 0$, we propose as in Iachan, a design p_1 such that

$$p_1(S) = \begin{cases} \lambda_1 & \text{if } d_S(s_x^2) < \delta_1 \\ \frac{\lambda_1\delta_1}{d_S(s_x^2)} & \text{otherwise} \end{cases}$$

Similarly to minimise (5.5.2) and (5.5.3) respectively we propose designs,

$$p_2(S) = \begin{cases} \lambda_2 & \text{if } d_S(s_x^2) < \delta_2 \\ \frac{\lambda_2\delta_2}{d_S(\overline{x})} & \text{otherwise} \end{cases}$$

and

$$p_3(S) = \begin{cases} \lambda_3 & \text{if } d_S(s_x^2) < \delta_1 \text{ and } d_S(\overline{x}) < \delta_2 \\ \frac{\lambda_3\delta_2}{d_S(\overline{x})} & \text{if } d_S(s_x^2) < \delta_1 \text{ and } d_S(\overline{x}) \geq \delta_2 \\ \frac{\lambda_3\delta_1}{d_S(s_x^2)} & \text{if } d_S(s_x^2) \geq \delta_1 \text{ and } d_S(\overline{x}) < \delta_2 \\ \frac{\lambda_3\delta_1\delta_2}{d_S(\overline{x})d_S(s_x^2)} & \text{if } d_S(s_x^2) \geq \delta_1 \text{ and } d_S(\overline{x}) \geq \delta_2 \end{cases}$$

The constants $\lambda_i = \lambda_i(\delta)$, $i = 1,2,3$ can be determined from the constraints

$$\sum_S p_i(S) = 1, \; i = 1,2,3 \tag{5.5.4}$$

Defining the sets D_1, D_2, D_{31}, D_{32}, D_{33} and D_{34} as

$$\begin{aligned} D_1 &= \{S : d_S(s_x^2) \geq \delta_1\}, \\ D_2 &= \{S : d_S(\overline{x}) \geq \delta_2\}, \\ D_{31} &= \{S : d_S(s_x^2) < \delta_1,\ d_S(\overline{x}) < \delta_2\}, \\ D_{32} &= \{S : d_S(s_x^2) < \delta_1,\ d_S(\overline{x}) \geq \delta_2\}, \\ D_{33} &= \{S : d_S(s_x^2) \geq \delta_1,\ d_S(\overline{x}) < \delta_2\}, \\ D_{34} &= \{S : d_S(s_x^2) \geq \delta_1,\ d_S(\overline{x}) \geq \delta_2\} \end{aligned}$$

$$\text{and } c_{\delta_1} = \#D_1^c,\ c_{\delta_2} = \#D_2^c,\ c_{\delta_1\delta_2} = \#D_{31} \tag{5.5.5}$$

we see that (5.5.4) is equivalent to, under p_1, p_2 and p_3 respectively,

$$\begin{aligned} \lambda_1 c_{\delta_1} + \lambda_1\delta_1 \sum_{S\in D_1} \{d_S(s_x^2)\}^{-1} &= 1 \\ \lambda_1 c_{\delta_2} + \lambda_2\delta_2 \sum_{S\in D_2} \{d_S(\overline{x})\}^{-1} &= 1 \end{aligned}$$

$$\lambda_3 c_{\delta_1\delta_2} + \lambda_3\delta_2 \sum_{S\in D_{32}} \{d_S(\overline{x})\}^{-1} + \lambda_3\delta_1 \sum_{S\in D_{33}} \{d_S(s_x^2)\}^{-1}$$

$$+\lambda_3\delta_1\delta_2 \sum_{S\in D_{34}} \{d_S(\overline{x}).d_S(s_x^2)\}^{-1} = 1$$

$$
\begin{aligned}
\text{ie., } \lambda_1 &= (c_{\delta_1} + \delta_1 d_1)^{-1} \\
\lambda_2 &= (c_{\delta_2} + \delta_2 d_2)^{-1} \qquad (5.5.6) \\
\lambda_3 &= (c_{\delta_1\delta_2} + \delta_2 d_{32} + \delta_1 d_{33} + \delta_1\delta_2 d_{34})^{-1}
\end{aligned}
$$

where

$$
\begin{aligned}
d_1 &= \sum_{S\in D_1} \{d_S(s_x^2)\}^{-1} \\
d_2 &= \sum_{S\in D_2} \{d_S(\overline{x})\}^{-1} \\
d_{32} &= \sum_{S\in D_{32}} \{d_S(\overline{x})\}^{-1} \\
d_{33} &= \sum_{S\in D_{33}} \{d_S(s_x^2)\}^{-1} \\
d_{34} &= \sum_{S\in D_{34}} \{d_S(\overline{x}).d_S(s_x^2)\}^{-1}.
\end{aligned}
$$

Clearly $\pi_{ij} > 0$ under p_k, $i \neq j = 1, 2, \cdots, N$, $k = 1, 2, 3$. Execution of designs p_k involves a preliminary choice of δ_1 and δ_2 and computation of $\overline{x}_S$ and s_x^2 for all samples S. One has then to consider equations (5.5.5) and (5.5.6).

Chapter 6

BAYESIAN ESTIMATION OF FINITE POPULATION VARIANCE UNDER MEASUREMENT ERROR MODEL

6.1. Review of Earlier Work :

We consider here the estimation of finite population variance under measurement error model. Here the variables involved may not be observed directly but only the values mixed with measurement errors. Sprent (1966) proposed a method based on the generalised least squares approach for estimating the linear regression coefficients under this model. Lindley (1966) and Lindley and El-sayad (1968) pioneered the Bayesian approach for that problem. Further Bayesian works can be found in Zellner (1971) and Reilly and Patino-Leal (1981). A general treatment for the inference problem in regression models with measurement errors is considered in Fuller (1987).

Normal theory Bayesian analysis in finite population sampling was elegantly caried out by Ericson (1969, 1988). A Bayesian approach for predicting finite population total and variance when the variables involved in the regression model are measured with error is considered in Bolfarine (1991), Mukhopadhyay (1994 a,b,c) among others. Bolfarine (1991) considered properties of predictors of finite population total under the location model with measurement errors and also under the simple regression model with measurement errors in both variables. Extensions are considered for the case of two-stage sampling. Bayes predictors of population variance are also derived .

Mukhopadhyay (1994 a,b ,c) obtained Bayes predictors of finite population total in unistage and two-stage sampling,domain total and finite population variance under regression models with measurement errors. A minimax predictor for population total was also obtained by him.

6.2. Introduction and Summary

In practical sample survey situations, the true value of the variables are seldom observed, but values mixed with measurement errors. We consider a finite population P of a known number N of identifiable units. Assosiated with i is the true value y_i of a study variable y. We assume that y_i can not be observed correctly, but a different value Y_i which is mixed with measurement error is observed.

We also assume that the true value y_i in the finite population is actually a realisation of a random variable $\mathcal{Y}_i$, the vector $\underset{\sim}{\mathcal{Y}} = (\mathcal{Y}_1, \mathcal{Y}_2, ..., \mathcal{Y}_N)$ obeying a super population model ς. However both y_i and $\mathcal{Y}_i$ are not observable and we do not make any notational distinction between them. The parameter of our interest is the population variance

$$S_y^2 = \frac{1}{N}\Sigma_{i=1}^N (y_i - \bar{y})^2, \bar{y} = \frac{1}{N}\sum_{i=1}^N y_i \qquad (6.2.1)$$

We intend to predict S_y^2 by drawing a sample S according to a fixed size(n)-sampling design p with selection probability $p(S)$ and observing the data $\underset{\sim}{Y_S} = (i, Y_i, i \in S)$ and using the model ξ.

In sub-section 6.3 we consider the problem in unistage sampling. The problem of estimation of S_y^2 in two-stage sampling has been considered in the next sub-section.

6.3. Prediction in unistage sampling

6.3.1 Description of Model

For the present discussion we consider the simple location model with measurement errors, which is described as follows :

$$\left.\begin{array}{lll} y_i = \mu + e_i, & E(e_i) = 0, & E(e_i^2) = \sigma_e^2, \\ & E(e_i e_{i'}) = 0, & (i \neq i') \\ Y_i = y_i + u_i, & E(u_i) = 0, & E(u_i^2) = \sigma_u^2 \\ & E(u_i u_{i'}) = 0, & (i \neq i') \\ & E(e_i u_j) = 0, & (i, j = 1, 2, ...) \end{array}\right\} \qquad (6.3.1)$$

Here e_i's are superpopulation model errors (also known as error in equation) and u_i's are measurement model errors. Assume that σ_e^2, σ_u^2 are known.

We further assume that the errors e_i and u_i are independently normally distributed. The model (6.3.1) was earlier considered by Bolfarine (1991), among others.

Under the model (6.3.1)

$$Y_i \sim N(\mu, \sigma^2) \qquad (6.3.2)$$

where

$$\sigma^2 = \sigma_u^2 + \sigma_e^2 \qquad (6.3.2')$$

As the distribution of a large number of variables including socio-economic variables is (at least approximately) normal in large samples, we assume a normal prior for μ with mean 0 and variance θ^2, i.e.,

$$\mu \sim N(0, \theta^2) \qquad (6.3.3)$$

Joint distribution of $\underset{\sim}{Y_S} = (Y_k, k \in S)$ and μ is, therefore,

$$L(\underset{\sim}{Y_S}|\mu).f(\mu) \propto e^{-\frac{n}{2\sigma^2}(\bar{Y}_S-\mu)^2}.e^{-\frac{\mu^2}{2\theta^2}}$$

where $\bar{Y}_S = \frac{1}{n}\sum_{i\in S} Y_i$.

The posterior density of μ is normal with mean

$$E(\mu|\underset{\sim}{Y_S}) = \frac{n\bar{Y}_S/\sigma^2}{\frac{n}{\sigma^2}+\frac{1}{\theta^2}} \qquad (6.3.4)$$

and variance

$$V(\mu|\underset{\sim}{Y_S}) = \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\theta^2}} \qquad (6.3.4')$$

Also, likelihood of $\underset{\sim}{Y_S}$ given $\underset{\sim}{y_S} = (y_i, i \in S)$ is

$$L(\underset{\sim}{Y_S}|\underset{\sim}{y_S}) \propto exp[-\frac{1}{2\sigma_u^2}\sum_{i\in S}(Y_i - y_i)^2]$$

Here posterior distribution of y_i's given $(\underset{\sim}{Y_S}, \mu)$ are independent with

$$y_i(i \in S) \sim N(\frac{Y_i\sigma_\epsilon^2 + \mu\sigma_u^2}{\sigma^2}, \sigma_0^2) \qquad (6.3.5)$$

and

$$y_i(i \in \bar{S}) \sim N(\mu, \sigma_\epsilon^2) \qquad (6.3.6)$$

where

$$\sigma_0^2 = \frac{\sigma_\epsilon^2\sigma_u^2}{\sigma^2} \qquad (6.3.6')$$

Under the above prior Mukhopadhyay (1994) considered Bayes estimation of population total. However Bayes estimator of population variance under the above prior was not considered by him.

6.3.2 Bayes Prediction of S_y^2

We shall assume a squared error loss function.
Therefore Bayes predictor of S_y^2 is

$$\hat{S_{yB}}^2 = E(S_y^2|\underset{\sim}{Y_S}) = E\{E(S_y^2|\underset{\sim}{Y_S}, \mu)|\underset{\sim}{Y_S}\} \qquad (6.3.7)$$

Consider the identity

$$S_y^2 = \frac{n}{N}s_y^2 + \frac{N-n}{N}s_{yr}^2 + \frac{n(N-n)}{N^2}(\bar{y}_r - \bar{y}_S)^2 \qquad (6.3.8)$$

(vide Bolfarine and Zacks (1991),(1.3.3)) where

$$s_y^2 = \frac{1}{n}\sum_{i\in S}(y_i - \bar{y}_S)^2, \quad \bar{y}_S = \frac{1}{n}\sum_{i\in S}y_i \qquad (6.3.9)$$

$$s_{yr}^2 = \frac{1}{N-n}\sum_{i\in \bar{S}}(y_i - \bar{y}_r)^2, \quad \bar{y}_r = \frac{1}{N-n}\sum_{i\in \bar{S}}y_i, \; \bar{S} = P - S \qquad (6.3.10)$$

We know that the distribution of the sum of squares

$$S = \sum_{j=1}^{n}(X_j - \bar{X})^2 \text{ where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n}X_j$$

when $X_1, X_2, ..., X_n$ are independent normal variables with X_j distributed normally with expected value ξ_j and standard deviation σ^2 (the same for all j) is σ^2 times a noncentral χ^2 with $(n-1)$ degrees of freedom and non-centrality parameter $\sum_{j=1}^{n}(\xi_j - \bar{\xi})^2/\sigma^2 = \lambda$ (say) with mean

$$E(\chi'^2) = \nu + \lambda$$

and variance

$$Var(\chi'^2) = 2(\nu + 2\lambda),$$

where ν is the degrees of freedom of the non-central χ^2 denoted by χ'^2.

From the posterior distribution of y_i $(i \in S)$ and y_i $(i \in \bar{S})$ given in (6.3.5) and (6.3.6) it follows that

$$\frac{ns_y^2}{\sigma_0^2} \sim \chi^2_{(n-1),\lambda} \qquad (6.3.11)$$

where

$$\lambda = \frac{\sigma_e^4}{\sigma^4}s_Y^2.\frac{n}{\sigma_0^2} \text{ and } s_Y^2 = \frac{1}{n}\sum_S(Y_i - \bar{Y})^2 \qquad (6.3.12)$$

Also,

$$(N-n)s_{yr}^2/\sigma_e^2 \sim \chi_{(N-n-1)}{}^2 \qquad (6.3.13)$$

Again,

$$\{(\bar{y}_r - \bar{y}_S)|\underset{\sim}{Y_S}, \mu\} \sim N(\mu_1, \sigma_1^2) \qquad (6.3.14)$$

where

$$\mu_1 = \frac{\sigma_e^2}{\sigma^2}(-\bar{Y}_S + \mu) \quad (6.3.15)$$

$$\sigma_1^2 = \frac{\sigma_e^2\sigma_u^2}{n\sigma^2} + \frac{\sigma_e^2}{(N-n)} = \frac{\sigma_e^2}{n(N-n)\sigma^2}[N\sigma^2 - (N-n)\sigma_e^2] \quad (6.3.16)$$

Hence, using (6.3.8)-(6.3.16)

$$\begin{aligned} E(S_y^2|\underset{\sim}{Y_S},\mu) &= \frac{n}{N}[\frac{n-1}{n}.\frac{\sigma_e^2\sigma_u^2}{\sigma^2} + \frac{\sigma_e^4}{\sigma^4}s_Y^2] + \frac{N-n}{N}\frac{N-n-1}{N-n}\sigma^2 \\ &+\frac{n(N-n)}{N^2}[\frac{\sigma_e^2\sigma_u^2}{n\sigma^2} + \frac{\sigma_e^2}{N-n} + \frac{\sigma_e^4}{\sigma^4}(\bar{Y}_S-\mu)^2] \\ &= \frac{n}{N}\frac{\sigma_e^4}{\sigma^4}[s_Y^2 + \frac{N-n}{N}(\bar{Y}_S-\mu)^2] + \frac{N-1}{N}\frac{\sigma_e^2}{\sigma^2}\{n\sigma_u^2 \\ &+(N-n)\sigma^2\} \quad (6.3.17) \end{aligned}$$

Therefore,

$$\begin{aligned} \hat{S_{yB}^2} &= E\{E(S_y^2|\underset{\sim}{Y_S},\mu)|\underset{\sim}{Y_S}\} \\ &= \frac{n}{N}\frac{\sigma_e^4}{\sigma^4}[s_y^2 + \frac{N-n}{N}\{\frac{\bar{Y}_S{}^2.\sigma^4}{(n\theta^2+\sigma^2)^2} + \frac{\theta^2\sigma^2}{(n\theta^2+\sigma^2)}\}] \\ &+\frac{N-1}{N^2}\sigma_e^2\sigma^2\{n\sigma_u^2 + (N-n)\sigma^2\} \qquad (\text{ using } (6.3.4),(6.3.4')) \\ &= \frac{n}{N}\frac{\sigma_e^4}{\sigma^4}[s_Y^2 + \frac{N-n}{N}\frac{\sigma^4\bar{Y}_S{}^2}{(n\theta^2+\sigma^2)^2} + \frac{n(N-n)}{N^2}\frac{\theta^2\sigma_e^4}{\sigma^2(n\theta^2+\sigma^2)} \\ &+\frac{N-1}{N^2}\frac{\sigma_e^2}{\sigma^2}\{n\sigma_u^2 + (N-n)\sigma^2\} \quad (6.3.18) \end{aligned}$$

If $\theta \to \infty$, the expression (6.3.18) becomes,

$$\begin{aligned} \underset{\theta\to\infty}{Lt}\ \hat{S}_{yB}^2 &= \frac{n}{N}\frac{\sigma_e^4}{\sigma^4}s_Y^2 + \frac{N-n}{N}\{\sigma_e^2 + \frac{n-1}{N-n}\sigma_0^2\} \quad (6.3.19) \\ &= \hat{S_{yB}^{\prime 2}} \quad (\text{say}) \end{aligned}$$

Bolfarine (1991) obtained the Bayes.predictor of S_y^2 [his expression (5)] under model (6.3.1) with normality assumption for e_i and u_i and with non-informative prior for μ [i.e., $p(\mu) \propto const.$]. His predictor, therefore should coincide with our predictor (6.3.18) when $\theta \to \infty$, that is, with $\hat{S_{yB}^{\prime 2}}$ in (6.3.19). This is found to be true. The predictor (6.3.19) coincides with Bolfarine's predictor.

Under model (6.3.1) (together with normality assumption of e_i's and u_i's) and Jeffrey's non-informative prior for the joint distribution of μ and $\frac{1}{\sigma_e^2}$, viz.

$$p(\mu, \sigma_e^{-2}) \propto \sigma_e^2, \quad \mu \in R_1, \quad \sigma_e^2 > 0$$

Mukhopadhyay (1994a) obtained Bayes predictor for S_y^2 [his expression (23)] as

$$\begin{aligned} \hat{S}_{yB}^{2''} &= \frac{k^2}{(k+1)^2}\frac{n(N+n-2)}{N(n-1)}\bar{Y}_S^{\,2} + \frac{k.s_Y^2}{(k+1)^2}[k + \frac{k(N-n)}{N(N-1)} \\ &+ \frac{2(n-1)}{N(n-3)}((N-n)(k+1)+1)] \end{aligned}$$

where $k = \frac{\sigma_e^2}{\sigma_u^2}$ and is assumed to be known.

However, we have considered here prior for μ only, assuming σ_e^2 and σ_u^2 to be known and therefore, our predictor (6.3.18) (and also (6.3.19)) differs from his expression.

We also note that if y_i is measured without error (i.e., $\sigma_u^2 = 0$) then $\hat{S}_{yB}^2$ reduces to the Bayes predictor

$$\begin{aligned} \hat{S}_B^2 &= \frac{n}{N}s_y^2 + \frac{N-n}{N}\sigma^2[1 - \frac{\sigma^2}{(N-n)\theta^2}(\frac{1}{n} - \frac{1}{N}).\frac{n\theta^2}{\sigma^2 + n\theta^2} \\ &+ \frac{n}{N}\frac{(\hat{\beta}_B - \bar{y}_S)^2}{\sigma^2}] \qquad (6.3.20) \end{aligned}$$

where $\hat{\beta}_B = \frac{\sum \frac{y_i}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{1}{\theta^2}} = \frac{\theta^2 \sum_S y_i}{n\theta^2 + \sigma^2}$. If further $\theta \to \infty$ (i.e., in the case of non-informative prior), the predictor $\hat{S}_B^{\,2}$ tends to

$$\lim_{\theta \to \infty} \hat{S}_B^{\,2} = \frac{n}{N}s_y^2 + \frac{N-n}{N}\sigma^2 = \hat{S}_B^{\prime 2}$$

which coincides with equation (3.1.27) in Bolfarine and Zacks (1991).

We now consider Bayes prediction risk of $\hat{S}_{yB}^{\,2}$. For a squared error loss function this is given by

$$E[Var(S_y^2 | \underset{\sim}{Y_S})] \qquad (6.3.21)$$

where the expression in (6.3.21) is taken with respect to predictive distribution of $\underset{\sim}{Y_S}$. Now the posterior variance of S_y^2 is

$$V(S_y^2 | \underset{\sim}{Y_S}) = E_\mu[V(S_y^2 | \mu, \underset{\sim}{Y_S}) | \underset{\sim}{Y_S}] + V_\mu[E(S_y^2 | \mu, \underset{\sim}{Y_S}) | \underset{\sim}{Y_S}]$$

$$= E_\mu[V\{\frac{n}{N}s_y^2 + \frac{N-n}{N}s_{yr}^2 + \frac{n(N-n)}{N^2}(\bar{y}_r - \bar{y}_S)^2 | \underset{\sim}{Y_S}, \mu\} | \underset{\sim}{Y_S}]$$
$$+V_\mu[\frac{n}{N}\frac{\sigma_e^4}{\sigma^4}(s_Y^2 + \frac{N-n}{N}(\bar{Y}_S - \mu)^2)$$
$$+\frac{N-1}{N^2}\frac{\sigma_e^2}{\sigma^2}(n\sigma_u^2 + \overline{N-n}\sigma^2) | \underset{\sim}{Y_S}] \qquad (6.3.22)$$

by (6.3.8) and (6.3.17))

Now,

$$E_\mu[\{V(S_y^2 | \underset{\sim}{Y_S}, \mu) | \underset{\sim}{Y_S}\}] = E_\mu[\{V(\frac{n}{N}s_y^2 | \underset{\sim}{Y_S}, \mu)$$
$$+V(\frac{N-n}{N}s_{yr}^2 | \underset{\sim}{Y_S}, \mu)$$
$$+V(\frac{n(N-n)}{N^2}(\bar{y}_r - \bar{y}_S)^2 | \underset{\sim}{Y_S}, \mu)\} | \underset{\sim}{Y_S}]$$

the covariance terms become zero because s_y^2, s_{yr}^2, $\bar{y}_S$ and $\bar{y}_r$ are all mutually independently distributed

$$= E_\mu[\frac{n^2}{N^2}\{2(n-1) + 4\frac{ns_Y^2.\frac{\sigma_e^4}{\sigma^4})}{\sigma_0^2}\}.(\sigma_0^2)^2.\frac{1}{n^2}$$
$$+\frac{(N-n)^2}{N^2}.2(N-n-1).\frac{\sigma_e^4}{(N-n)^2}$$
$$+\frac{(n^2(N-n)^2}{N^4}\{2 + 4\frac{(\frac{\sigma_e^4}{\sigma^4}(\bar{Y}_S-\mu)^2)}{\sigma_0^2}$$
$$+\frac{\sigma_e^2}{N-n}))\}.\{\sigma_0^2 + \frac{\sigma_e^2}{N-n}\}^2 | \underset{\sim}{Y_S}]$$
$$= \frac{2}{N^2}\frac{\sigma_e^4}{\sigma^4}\{(n-1)\sigma_u^2 + (N-n-1)\sigma^4$$
$$+\frac{1}{N^2}(\overline{N-n}\sigma_u^2 + n\sigma^2)^2\} + 4\frac{n}{N^2}\frac{\sigma_e^4}{\sigma^4}[\sigma_0^2 s_Y^2$$
$$+\frac{n(N-n)^2}{N^2}(\sigma_0^2 + \frac{\sigma_e^2}{N-n})\frac{\sigma^4\bar{Y}_S{}^2}{(n\theta^2+\sigma^2)^2}]$$
$$+4\frac{n^2(N-n)^2}{N^4}\frac{\sigma_e^4}{\sigma^4}(\sigma_0^2 + \frac{\sigma_e^2}{N-n}).\frac{\sigma^2\theta^2}{n\theta^2+\sigma^2} \qquad (6.3.23)$$

Again,

$$V_\mu[E(S_y^2 | \underset{\sim}{Y_S}, \mu) | \underset{\sim}{Y_S}] = \frac{n^2(N-n)^2}{N^4}\frac{\sigma_e^8}{\sigma^8}\{2\frac{\sigma^4\theta^4}{(n\theta^2+\sigma^2)^2}$$

$$+4\frac{\sigma^4\bar{Y_S}^2}{(n\theta^2+\sigma^2)^2}\frac{\sigma^2\theta^2}{(n\theta^2+\sigma^2)}\} \qquad (6.3.24)$$

Hence, by adding (6.3.23) and (6.3.24), and on simplification

$$\begin{aligned} V(S_y^2|\underset{\sim}{Y_S}) &= \frac{2}{N^2}\frac{\sigma_e^4}{\sigma^4}\{(n-1)\sigma_u^4+(N-n-1)\sigma^4\}+4.\frac{n}{N^2}\frac{\sigma_e^6}{\sigma^6}\sigma_u^2 s_Y^2 \\ &+\frac{2n^2(N-n)^2}{N^4}\{\frac{\sigma_e^2}{\sigma^2}(\frac{\sigma_u^2}{n}+\frac{\sigma^2}{N-n}+\frac{\sigma_e^2\theta^2}{(n\theta^2+\sigma^2)})\}^2 \\ &+\frac{4n^2(N-n)^2}{N^4}\frac{\sigma_e^4\bar{Y_S}^2}{(n\theta^2+\sigma^2)^2}\{\frac{\sigma_e^2}{\sigma^2}(\frac{\sigma_u^2}{n}+\frac{\sigma^2}{N-n} \\ &+\frac{\sigma_e^2\theta^2}{(n\theta^2+\sigma^2)})\} \qquad (6.3.25) \end{aligned}$$

The above expression shows that the posterior variance of S_y^2 given $\underset{\sim}{Y_S}$ is not independent of $\underset{\sim}{Y_S}$. Now the predictive distribution of $\bar{Y}_S$ under model (6.3.1) and prior (6.3.3) is $N(0, \frac{\sigma^2}{n}+\theta^2)$ [ref. Bolfarine and Zacks (1991), chap.3, p. 92]. Hence by (6.3.21), Bayes risk of $\hat{S_{yB}}^2$ is

$$\begin{aligned} E[V(S_y^2|\underset{\sim}{Y_S})] &= \frac{2}{N^2}\frac{\sigma_e^4}{\sigma^4}\{(n-1)\sigma_u^2+(N-n-1)\sigma^4+2(n-1)\sigma_e^2\sigma_u^2\} \\ & 2\frac{n^2(N-n)^2}{N^4}\{\frac{\sigma_e^2}{\sigma^2}(\frac{\sigma_u^2}{n}+\frac{\sigma^2}{N-n}+\frac{\sigma_e^2\theta^2}{(n\theta^2+\sigma^2)})\}^2 \\ &+4.\frac{n(N-n)^2}{N^4}\frac{\sigma_e^4}{(n\theta^2+\sigma^2)}\{\frac{\sigma_e^2}{\sigma^2}(\frac{\sigma_u^2}{n}+\frac{\sigma^2}{N-n} \\ &+\frac{\sigma_e^2\theta^2}{(n\theta^2+\sigma^2)})\} \qquad (6.3.26) \end{aligned}$$

Allowing $\theta \to \infty$ in (6.3.26), we note that limit of Bayes risk of S_{yB}^2 as

$$\frac{2}{N^2}\frac{\sigma_e^4}{\sigma^4}\{(N-1)\sigma^4-(n-1)\sigma_e^4\}$$

6.4. Extension to two-stage sampling

Following Bolfarine (1991) we consider the following superpopulation model applicable to two-stage sampling. The finite population is divided into k subpopulations (clusters) $\mathcal{P}_h$ of size M_h, $h = 1, 2, ..., k$. Let y_{hj} be the true value of the characteristic y associated with unit j in cluster h, $(j = 1, ..., M_h,\ h = 1, 2, ..., k)$. In the first stage, a sample S of n

clusters is selected from the k clusters in the population. In the second stage, a sample s_h of size m_h is selected from each cluster h in the sample S. However, as in the previous section the true value y_{hj} can not be observed, even when unit (hj) is in the sample, only their values Y_{hj} mixed with measurement errors are observed.

We assume that

$$y_{hj} = \mu_h + e_{hj} \qquad (6.4.1)$$

$$\mu_h = \mu + v_h \qquad (6.4.2)$$

$$Y_{hj} = y_{hj} + u_{hj} \qquad (6.4.3)$$

Here e_{hj}, v_h are soperpopulation model errors and u_{hj} is the measurement error. Assume that e_{hj}, v_h and u_{hj} are all independent,

$$e_{hj} \sim N(0, \sigma_{eh}^2)$$

$$v_h \sim N(0, \theta^2)$$

$$u_{hj} \sim N(0, \sigma_{uh}^2), j = 1, ..., M_h, h = 1, ..., k$$

Under the above model, Bolfarine (1991) considered the prediction of finite population total. Under a slightly different model which takes into account the finiteness of the survey population (and measurement errors), Mukhopadhyay (1995) considered prediction of finite population total in two-stage sampling. We shall here consider prediction of finite population variance under models (4.1)-(4.3) which are extensions of model due to Scott and Smith (1969) to measurement errors.

Now, $S_y^2 = \frac{1}{M}\sum_{h=1}^{k}\sum_{j=1}^{M_h}(y_{hj} - \bar{y})^2$ where $M = \sum_{h=1}^{k} M_h, \bar{y} = \sum_h \sum_j \frac{Y_{hj}}{M}$.
Here

$$Y_{hj}|\mu_h \sim N(\mu_h, \sigma_{eh}^2 + \sigma_{uh}^2 = \sigma_h^2) \qquad (6.4.4)$$

Hence following section (6.3.1) posterior distribution of y_{hj}'s given $(\underset{\sim}{Y_{hs}}, \mu_h)$ where $(\underset{\sim}{Y_{hS}} = (Y_{hj}, hj \in S_h)$ are independent with

$$y_{hj}(h \in S, i \in S_h) \sim N\left(\frac{Y_{hj}\sigma_{eh}^2 + \mu_h\sigma_{uh}^2}{\sigma_h^2}, \frac{\sigma_{eh}^2\sigma_{uh}^2}{\sigma_h^2}\right) \qquad (6.4.5a)$$

$$y_{hj}(h \in S, i \in \bar{S}_h) \sim N(\mu_h, \sigma_{eh}^2) \qquad (6.4.5b)$$

$$y_{hj}(h \in \bar{S}) \sim N(\mu_h, \sigma_{eh}^2) \qquad (6.4.5c)$$

where $\bar{S} = P - S,\ \bar{S}_h = P_h - S_h$.

After the sample S has been selected, we may write the population variance as

$$
\begin{aligned}
S_y^2 &= \frac{1}{(\sum_{h=1}^{k})}\sum_{h=1}^{k}\sum_{j=1}^{M_h}(y_{hj}-\bar{y})^2 \\
&= \frac{1}{M}[\sum_{h\in S}\sum_{j=1}^{M_h}(y_{hj}-\bar{y})^2+\sum_{h\in \bar{S}}\sum_{j=1}^{M_h}(y_{hj}-\bar{y})^2] \\
&= \frac{1}{M}[\{\sum_{h\in S}M_hS_h^2+\sum_{h\in S}M_h(\bar{y_h}-\bar{y})^2\}+\{\sum_{h\in \bar{S}}M_hS_h^2 \\
&\quad +\sum_{h\in /}M_h(\bar{y_h}-\bar{y})^2\}] \qquad (6.4.6)
\end{aligned}
$$

where $S_h^2=\frac{1}{M_h}\sum_{j=1}^{M_h}(y_{hj}-\bar{y_h})^2$, $\bar{y_h}=\frac{1}{M_h}\sum_{j=1}^{M_h}y_{hj}$.

Now, for $h\in\bar{S}$,

$$
\bar{y_h}-\bar{y}=\bar{y_h}-\frac{1}{M}[\sum_{h\in S}\{(m_h\bar{y_{hS}})+(M_h-m_h)\bar{y_{hr}}\}]-\frac{1}{M}\sum_{h\in\bar{S}}M_h\bar{y_h} \qquad (6.4.7)
$$

where $\bar{y_{hS}}=\frac{1}{m_h}\sum_{j\in S_h}y_{hj}, \bar{y_{hr}}=\frac{1}{M_h-m_h}\sum_{j\in\bar{S_h}}y_{hj}$. Again for $h\in S$,

$$
\begin{aligned}
\bar{y_h}-\bar{y} &= m_h(\frac{1}{M_h}-\frac{1}{M})\bar{y_{hS}}+(M_h-m_h)(\frac{1}{M_h}-\frac{1}{M})\bar{y_{hr}} \\
&\quad -\frac{1}{M}\sum_{l(\neq h)\in S}\{m_l\bar{y_{lS}}+(M_l-m_l)\bar{y_{lr}}\} \\
&\quad -\sum_{h\in\bar{S}}\frac{M_h}{M}\bar{y_h} \qquad (6.4.8)
\end{aligned}
$$

Hence we have

$$
E\{(\bar{y_h}-\bar{y})|\underset{\sim}{Y_S},\underset{\sim}{\mu}\}=\begin{cases}\frac{m_h}{M_h}\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}}-\mu_h)+\mu_{0h}, & h\in S\\ \mu_{0h}, & h\in\bar{S}\end{cases} \qquad (6.4.9)
$$

where

$$
\mu_{oh}=\mu_h-\sum_{h=1}^{k}\frac{M_h}{M}\mu_h-\sum_{h\in S}\frac{m_h}{M}\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}}-\mu_h) \qquad (6.4.10)
$$

and

$$
V\{(\bar{y_h}-\bar{y})|\underset{\sim}{Y_S},\underset{\sim}{\mu}\}=\begin{cases}-\frac{m_h}{M_h}(\frac{1}{M_h}-\frac{2}{M})\frac{\sigma_{eh}^4}{\sigma_h^2}+\sigma_{0h}^2, & h\in S\\ \sigma_{0h}^2, & h\in\overline{S}\end{cases} \qquad (6.4.11)
$$

where

$$\sigma_{0h}^2 = (\frac{1}{M_h} - \frac{2}{M})\sigma_{eh}^2 + \frac{1}{M^2}\sum_{h=1}^{k} M_h\sigma_{eh}^2 - \frac{1}{M^2}\sum_{h\in S} m_h\frac{\sigma_{eh}^4}{\sigma_h^2} \qquad (6.4.12)$$

Again, using the results of the uni-stage sampling given in (6.3.17) we can write

$$E(S_h^2|\underset{\sim}{Y_S},\underset{\sim}{\mu}) = \begin{cases} \frac{m_h}{M_h}\frac{\sigma_{eh}^4}{\sigma_h^4}[s_{Yh}^2 + \frac{M_h-m_h}{M_h}(\bar{Y_{Sh}} - \mu_h)^2] & \\ +\frac{M_h-1}{M_h^2}\frac{\sigma_{eh}^2}{\sigma_h^2}\{M_h\sigma_h^2 - m_h\sigma_{eh}^2\}, & h\in S \\ \sigma_{eh}^2.\frac{M_h-1}{M_h}, & h\in \bar{S} \end{cases}$$

Therefore using (6.4.6) - (6.4.12) we have

$$\begin{aligned} E(S_y^2|\underset{\sim}{Y_S},\underset{\sim}{\mu}) &= \frac{1}{M}[\sum_{h\in S} M_h\{\frac{m_h}{M_h}\frac{\sigma_{eh}^4}{\sigma_h^4}(s_{Yh}^2 + \frac{M_h-m_h}{M_h}(\bar{Y_{Sh}} - \mu_h)^2) \\ &+\frac{M_h-1}{M_h^2}\frac{\sigma_{eh}^2}{\sigma_h^2}(M_h\sigma_h^2 - m_h\sigma_{eh}^2) \\ &+\sum_{h\in S} M_h\{(\frac{m_h}{M_h}\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}} - \mu_h) + \mu_{0h})^2 \\ &-\frac{m_h}{M_h}(\frac{1}{M_h} - \frac{2}{M})\frac{\sigma_{eh}^2}{\sigma_h^2} + \sigma_{0h}^2\} \\ &+\sum_{h\in/S} M_h\{\frac{M_h-1}{M_h}\sigma_{eh}^2 + \mu_{0h}^2 + \sigma_{0h}^2\}] \\ &= \frac{1}{M}(1-\frac{1}{M})(\sum_{h=1}^{k} M_h\sigma_{eh}^2 - \sum_{h\in S} m_h\frac{\sigma_{eh}^4}{\sigma_h^2}) + \frac{1}{M}\sum_{h\in S} m_h\frac{\sigma_{eh}^4}{\sigma_h^4}s_{Yh}^2 \\ &+\frac{1}{M}\sum_{h\in S} m_h\frac{\sigma_{eh}^4}{\sigma_h^4}(\bar{Y_{Sh}} - \mu_h)^2 \\ &+\frac{2}{M}\sum_{h\in S} m_h\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}} - \mu_h)(\mu_h - \bar{\mu}) \\ &-\frac{1}{M^2}(\sum_{h\in S} m_h\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}} - \mu_h))^2 \\ &+\frac{1}{M}\sum_{h=1}^{k} M_h(\mu_h - \bar{\mu})^2 \qquad (6.4.13) \end{aligned}$$

Let us denote $\frac{\sigma_{eh}^2}{\sigma_h^2}(\bar{Y_{Sh}} - \mu_h)$ by z_h , then (6.4.13) becomes

$$E(S_y^2|\underset{\sim}{Y_S},\mu_{\sim}) = \frac{1}{M}(1-\frac{1}{M})(\sum_{h=1}^{k} M_h\sigma_{eh}^2 - \sum_{h\in S} m_h\frac{\sigma_{eh}^4}{\sigma_h^2})$$

$$+\frac{1}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^4} s_{Yh}^2 + \frac{1}{M}\sum_{h\in S} m_h z_h^2$$
$$+\frac{2}{M}\sum_{h\in S} m_h z_h(\mu_h - \bar{\mu}) - \frac{1}{M^2}(\sum_{h\in S} m_h z_h)^2$$
$$+\frac{1}{M}\sum_{h=1}^{k} M_h(\mu_h - \bar{\mu})^2 \qquad (6.4.14)$$

It follows from Scott and Smith (1969) that the posterior distribution of $\underset{\sim}{\mu} \,|\, \underset{\sim}{Y_S}$ is N-variate normal with mean

$$E(\mu_h \,|\, \underset{\sim}{Y_S}) = \lambda_h \bar{Y_{Sh}} + (1 - \lambda_h)\bar{Y_S} \qquad (6.4.15a)$$

where $\bar{Y_{Sh}} = \frac{1}{m_h}\sum_{j\in s_h} Y_{hj}, \bar{Y_S} = \frac{\sum_{h\in S}\lambda_h \bar{Y_{hS}}}{\sum_{h\in S}\lambda_h}$
and

$$\lambda_h = \begin{cases} \frac{\theta^2}{\theta^2 + \frac{\sigma_h^2}{m_h}}, & h \in S \\ 0 & h \in \bar{S} \end{cases} \qquad (6.4.15b)$$

and covariance matrix $((C_{hl}))$ where

$$C_{hl} = \begin{cases} (1-\lambda_h)^2\nu^2 + (1-\lambda_h)\theta^2, & h = l \\ (1-\lambda_h)(1-\lambda_l)\nu^2, & h \neq l \end{cases} \qquad (6.4.15c)$$

where

$$\nu^2 = [\sum_{h\in S}(\theta^2 + \frac{\sigma_h^2}{m_h})^{-1}]^{-1}.$$

Hence it is implied that $z_h | \underset{\sim}{Y_S}$ is N-variate normal with mean

$$\frac{\sigma_{eh}^2}{\sigma_h^2}(1-\lambda_h)(\bar{Y_{Sh}} - \bar{Y_S}) \qquad (6.4.16a)$$

and covariance matrix

$$C'_{hl} = \begin{cases} \frac{\sigma_{eh}^4}{\sigma_h^4}\{(1-\lambda_h)^2\nu^2 + (1-\lambda_h)\theta^2\}, & h = l \\ \frac{\sigma_{eh}^2}{\sigma_h^2}\frac{\sigma_{el}^2}{\sigma_l^2}(1-\lambda_h)(1-\lambda_l)\nu^2, & h \neq l \end{cases} \qquad (6.4.16b)$$

Using (6.4.15a)-(6.4.16b) in (6.4.14) we get

$$\begin{aligned} E(S_y^2 | \underset{\sim}{Y_S}) &= E_{\underset{\sim}{\mu}} E(S_y^2 | \underset{\sim}{Y_S}, \underset{\sim}{\mu}) \\ &= \frac{1}{M}(1 - \frac{1}{M})(\sum_{h=1}^{k} M_h \sigma_{eh}^2 - \sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^2}) \end{aligned}$$

$$+\frac{1}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^4} s_{Yh}^2$$
$$+\frac{2}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^2}{\sigma_h^2}(1-\lambda_h)(\bar{Y_{Sh}}-\bar{Y_S})\{\lambda_h \bar{Y_{Sh}}+(1-\lambda_h)\bar{Y_S}$$
$$-\sum_{h=1}^{k}\frac{M_h}{M}(\lambda_h \bar{Y_{Sh}}+\overline{1-\lambda_h}\bar{Y_S})\}$$
$$+\frac{1}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^4}(1-\lambda_h)^2(\bar{Y_{Sh}}-\bar{Y_S})^2-\frac{1}{M^2}(\sum_{h\in S} m_h \frac{\sigma_{eh}^2}{\sigma_h^2}(1-\lambda_h)(\bar{Y_{Sh}}$$
$$-\bar{Y_S}))^2+S^2_{(\lambda_h \bar{Y_{Sh}}-\overline{1-\lambda_h}\bar{Y_S})}$$
$$+\nu^2[-\frac{2}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^2}{\sigma_h^2}(1-\lambda_h)\{(1-\lambda_h)-\sum_{h=1}^{k}\frac{M_h}{M}(1-\lambda_h)\}$$
$$+\frac{1}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^4}(1-\lambda_h)^2-\frac{1}{M^2}(\sum_{h\in S} m_h \frac{\sigma_{eh}^2}{\sigma_h^2}(1-\lambda_h))^2+S^2_{(1-\lambda_h)}]$$
$$+\theta^2[-\frac{2}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^2}{\sigma_h^2}(1-\frac{M_h}{M})(1-\lambda_h)+\frac{1}{M}\sum_{h\in S} m_h \frac{\sigma_{eh}^4}{\sigma_h^4}(1-\lambda_h)$$
$$-\frac{1}{M^2}\sum_{h\in S} m_h^2 \frac{\sigma_{eh}^4}{\sigma_h^4}(1-\lambda_h)+\sum_{h=1}^{k}(\frac{M_h}{M}-\frac{M_h^2}{M^2})(1-\lambda_h)]$$

When $k=1$, the above expression becomes

$$\frac{1}{M}(1-\frac{1}{M})(M\sigma_e^2-m\frac{\sigma_e^2}{\sigma^2}+\frac{m}{M}(1-\frac{m}{M})\frac{\sigma_e^4}{\sigma^4}\frac{\sigma^2}{m}$$

When there are several clusters ($k>1$), prior of μ_h is taken as $\mu_h \sim N(\mu,\theta^2)$. In case of one cluster there is only one parameter μ and thus no question of variability occurs, hence the expression is independent of θ. hence the role of θ becomes irrevelent yielding the above θ-independent estimator.

Chapter 7

BAYESIAN ESTIMATION OF POPULATION PROPORTIONS IN A POLYTOMOUS POPULATION IN TWO-STAGE SAMPLING

7.1 Introduction

In practical situations data are frequently generated from polytomous processes with the parameter of interest being the respective frequencies of the attributes concerned. As for example if we consider rural indebtedeness for agriculture in the state of West Bengal, a farmer may fall into one of the following four categiories , viz. whether he has taken loan from nationalised banks, or from co-operative bank or from private money-lenders or for agricultural purposes he has not taken loan at all. One may be interested in estimating the proportion of farmers in these several categories in the rural areas of a district, which may be considered as a cluster. In practice , inferences about these frequencies are always based on the assumption that the data generating process is multinomial (in case the population is infinite) or hypergeometric (in case the population is finite). In this context one may wish to use Bayesian methods to predict finite population parameters using suitable priors. In the Bayesian approach to prediction problems the main role is played by the predictive density function which expresses the plausibility of the parameters in the light of the results of an informative experimant. The predictive distribution in its turn is influenced by the choice of the prior distribution.

In the present chapter we have considered a two-stage sampling design for the above problem and both the cases, where data follow the multinomial or the hypergeometric model are treated separately. Accordingly the priors are chosen from the class of Dirichlet distribution for the first case and from the ckass of Dirichlet-Multinomial (DM) distributions for the second case. Nandram and Sedransk (1993) made a similar study for a binary variable in two-stage cluster sampling and used a mixture of beta distributions as a prior for their analysis.

We consider the following set-up.

There are N clusters and t attributes are found to exist in each cluster. We assume that the number of units in k-th cluster having i-th characteristic is Y_{ki} and M_k is the size of cluster i.e., total number of units in the k-th cluster.

Hence $\sum_{i=1}^{t} Y_{ki} = M_k$.

Again, $\delta_{ki} = \frac{Y_{ki}}{M_k}$ denotes the proportion of units having i-th characteristic in cluster i.

We are interested in estimating the population proportions P_i with characteristic $i(i = 1, ..., k)$,

$$P_i = \frac{\sum_{k=1}^{N} Y_{ki}}{\sum_{k=1}^{N} M_k} = \sum_{k=1}^{N} (\frac{Y_{ki}}{M_k}).(\frac{M_k}{\sum_{k=1}^{N} M_k}) = \sum_{k=1}^{N} \delta_{ki}\rho_k \tag{7.1.1}$$

where $\rho_k = \frac{M_k}{\sum_{k=1}^{N} M_k} = \frac{M_k}{M}$ (denoting $\sum_{k=1}^{N} M_k$ by M) is the population proportion of units in the k-th cluster i.e., relative size of cluster k, through survey sampling. A two-stage sampling design is employed. At the first-stage a sample S of n clusters is selected by SRSWOR. At the second stage from each selected cluster, a sample of m_h elementary units is selected. We assume that the sampling design is a non-informative one. In this Bayesian approach our analysis will proceed on the basis of the data obtained from the final sample and the actual sampling design by which the sample have been selected will be considered irrelavent.

Let y_{ki} denote the number of units belonging to the i-th category among the sampled m_k units from the sampled k-th cluster. We denote $(y_{k1}, y_{k2}, ..., y_{kt})'$ by $\underset{\sim}{y_k}$.

7.2 Present Study Under Multinomial Setup

7.2.1 Prior Moments of P_i

We assume that $(Y_{k1}, Y_{k2}, ... Y_{kt})$ are jointly distributed as a multinomial distribution, i,e., there are t possible outcomes A_i for each of M_k trials, each observations representing a trial, i=1,2,...,t. Y_{ki} denotes the number of times event A_i occurs in M_k units in cluster k. Y_{ki}'s are discrete and integer valued with

$$P(Y_{k1}, Y_{k2}, ..., Y_{kt} | \theta_{k1}, \theta_{k2}, ..., \theta_{kt}) = \binom{M_k}{Y_{k1}, ..., Y_{kt}} \theta_{k1}^{Y_{ki}} ... \theta_{kt}^{Y_{kt}} \tag{7.2.1}$$

where $\theta_{k1}, ..., \theta_{kt}$ are the parameters of the distribution with $\sum_{i=1}^{t} \theta_{ki} = 1$.

We denote $(\theta_{k1}, ..., \theta_{kt})'$ by $\underset{\sim}{\theta_k}$.

We consider $\underset{\sim}{\theta_1}, ..., \underset{\sim}{\theta_N}$ to have independent Dirichlet distribution with parameters $(\alpha_1, ..., \alpha_t) = \underset{\sim}{\alpha}$ and density function

$$P(\theta_{k1}, ..., \theta_{kt}) = \frac{\Gamma(\alpha)}{\Pi_{j=1}^{t} \Gamma(\alpha_j)} \Pi_{j=1}^{t} \theta_{kj}^{\alpha_j - 1} \tag{7.2.2}$$

where $\alpha = \sum_{j=1}^{t} \alpha_j$, $0 \le \theta_{kj}, j = 1, 2, ..., t$, $\sum_{j=1}^{t} \theta_j \le 1$.
It is known that

$$E(\theta_{kj}) = \frac{\alpha_j}{\alpha} \qquad (7.2.2a)$$

$$V(\theta_{kj}) = \frac{\alpha_j(\alpha - \alpha_j)}{\alpha^2(\alpha+1)} \qquad (7.2.2b)$$

$$Cov.(\theta_{kj}, \theta_{kj'}) = -\frac{\alpha_j \alpha_{j'}}{\alpha^2(\alpha+1)} \qquad (7.2.2c)$$

(Johnson and Kotz, Continuous Multivariate Distributions, p.231)

Let $\underset{\sim}{\underset{\sim}{Y}} = (\underset{\sim}{Y_1}, \underset{\sim}{Y_2}, ..., \underset{\sim}{Y_N})_{t \times N}$, where $\underset{\sim}{Y_k} = (Y_{k1}, ..., Y_{kt})'$.

Considering (7.2.2) as a prior of $\underset{\sim}{\theta_k}$ we have using (7.2.1), the predictive distribution of the random vectors, $\underset{\sim}{Y_1}, \underset{\sim}{Y_2}, ..., \underset{\sim}{Y_N}$ as

$$\begin{aligned} P(\underset{\sim}{\underset{\sim}{Y}} \mid \underset{\sim}{\alpha}) &= \Pi_{k=1}^{N} \frac{M_k!}{\Pi_{j=1}^{t} Y_{kj}!} E_{\underset{\sim}{\theta_k}}(\Pi_{j=1}^{t} \theta_{kj}^{Y_{kj}}) \\ &= \Pi_{k=1}^{N} \frac{M_k!}{\Pi_{j=1}^{t} Y_{kj}!} [\frac{\alpha}{\Pi_{j=1}^{t} \Gamma \alpha_j} \frac{\Pi_{j=1}^{t} \Gamma(\alpha_j + Y_{kj})}{\Gamma(\alpha + M_k)}] \end{aligned} \qquad (7.2.3)$$

(Ref. Johnson and Kotz, Discrete Distributions, p. 183)

where $E_{\underset{\sim}{\theta_k}}(.)$ denotes expectation of (.) wrt distribution of $\underset{\sim}{\theta_k}$ as given in (7.2.2). It follows, therefore,

$$E(Y_{kj} \mid \underset{\sim}{\alpha}) = M_k p_j', \; p_j' = \frac{\alpha_j}{\alpha} \qquad (7.2.4)$$

i.e., $\quad E(\delta_{kj} \mid \underset{\sim}{\alpha}) = p_j', \;\; j = 1, 2, ..., t \qquad (7.2.5)$

$$V(\delta_{kj} \mid \underset{\sim}{\alpha}) = \frac{M_k + \alpha}{1 + \alpha} \frac{p_j'(1 - p_j')}{M_k} \qquad (7.2.6)$$

and $\quad Cov.(\delta_{ki}, \delta_{kj} \mid \underset{\sim}{\alpha}) = -(\frac{M_k + \alpha}{1 + \alpha})(\frac{1}{M_k}) p_i' p_j' \qquad (7.2.7)$

Hence,

$$E(P_i \mid \underset{\sim}{\alpha}) = \sum_{i=1}^{N} E(\delta_{kj} \rho_k \mid \underset{\sim}{\alpha}) = p_i' \sum_{k=1}^{N} \rho_k = p_i' \qquad (7.2.8)$$

and

$$V(P_i \mid \underset{\sim}{\alpha}) = \sum_{k=1}^{N} \rho_k^2 \frac{M_k + \alpha}{1 + \alpha} \frac{1}{M_k} \frac{\alpha_i}{\alpha}(1 - \frac{\alpha_i}{\alpha}) \qquad (7.2.9)$$

since $Cov.(\delta_{ki}, \delta_{li}) = 0 \ (k \neq l)$. ALso,

$$Cov.(P_i, P_j | \underset{\sim}{\alpha}) = p_i' p_j' \sum_{k=1}^{N} \rho_k^2 \{-\frac{M_k + \alpha}{M_k(1+\alpha)}\} \tag{7.2.10}$$

In particular, when $\alpha_j = \frac{\alpha}{t} \forall j = 1, 2, ..., t$, (7.2.8) and (7.2.9) reduce respectively to

$$E(P_i | \underset{\sim}{\alpha}) = \frac{1}{t}$$

and

$$V(P_i | \underset{\sim}{\alpha}) = \frac{1-t}{t^2} \sum_{k=1}^{N} \rho_k^2 \frac{M_k + \alpha}{M_k(1+\alpha)}.$$

7.2.2 Posterior Moments of P_i

We shall now find the posterior moments of P_i given the data.

We have already assumed that n clusters were sampled at the first stage and let y_{kj} denote the number of sampled units in cluster $k (\in S)$ having the characteristic A_j, $\sum_{j=1}^{t} y_{kj} = m_k$, Assume that the sampling design is a non-informative one. Since the sampling design at the second stage is simple random sampling with replacement, the likelihood of the sample observations for the k-th cluster is

$$P(y_{k1}, ..., y_{kt} | \theta_{k1}, ..., \theta_{kt}) = \binom{m_k}{y_{k1}, ..., y_{kt}} \theta_{k1}{}^{y_{k1}} ... \theta_{kt}{}^{y_{kt}} \tag{7.2.11}$$

Using the prior (7.2.2) of $\underset{\sim}{\theta_k}$, the joint distribution of $\underset{\sim}{\underset{\sim}{y}}$ and $\underset{\sim}{\underset{\sim}{\theta_S}}$ where,

$\underset{\sim}{\underset{\sim}{y}} = (\underset{\sim}{y_k}, k \in S)_{t \times n}$, and $\underset{\sim}{\underset{\sim}{\theta_S}} = (\underset{\sim}{\theta_k}, k \in S)_{t \times n}$ is

$$P(\underset{\sim}{\theta_S}, \underset{\sim}{\underset{\sim}{y}} \mid \underset{\sim}{\alpha}) = \Pi_{k \in S} \binom{m_k}{y_{k1}, ..., y_{kt}} \frac{\Gamma \alpha}{\Pi_{j=1}^{t} \Gamma(\alpha_j)} \Pi_{j=1}^{t} \theta_{kj}^{y_{kj} + \alpha_j - 1} \tag{7.2.12}$$

where,

$$\underset{\sim}{\theta_S} = (\underset{\sim}{\theta_k}, k \in S)_{t \times n}$$

$$\underset{\sim}{\underset{\sim}{y}} = (\underset{\sim}{y_k}, k \in S)_{t \times n},$$

$$\underset{\sim}{y_k} = (y_{k1}, ..., y_{kt})'.$$

Therefore marginal distribution of $\underset{\sim}{y}$ is

$$\phi(\underset{\sim}{y}) = \Pi_{k\in S}\binom{m_k}{y_{k1},...,y_{kt}}\frac{\Gamma(\alpha)}{\Pi_{j=1}^{t}\Gamma(\alpha_j)}\frac{\Pi_{j=1}^{t}\Gamma(y_{kj}+\alpha_j)}{\Gamma(\alpha+m_k)} \tag{7.2.13}$$

Hence posterior of $\underset{\sim}{\theta_S}$ given $(\underset{\sim}{y},\underset{\sim}{\alpha})$, is obtained by dividing (7.2.12) by (7.2.13),

$$P(\theta_{S\sim}\,|\,y_{\sim},\underset{\sim}{\alpha}) = \Pi_{k\in S}\frac{\Gamma(\alpha+m_k)}{\Pi_{j=1}^{t}\Gamma(\alpha_j+y_{kj})}\Pi_{j=1}^{t}\theta_{kj}^{y_{kj}+\alpha_j-1} \tag{7.2.14}$$

This posterior distribution will be used later (vide (7.2.20)) as prior of $\underset{\sim}{\theta_S}$ in finding a posterior distribution.

Therefore, for $k\in S$,

$$E(\underset{\sim}{\theta_k}\,|\,y_{\sim},\underset{\sim}{\alpha}) = \frac{1}{\alpha+m_k}(\underset{\sim}{y_k}+\underset{\sim}{\alpha})$$

$$V(\theta_{kj}|\,y_{\sim},\underset{\sim}{\alpha}) = \frac{(y_{kj}+\alpha_j)(\alpha+m_k-\alpha_j-y_{kj})}{(\alpha+m_k)^2(\alpha+m_k+1)}$$

$$Cov.(\underset{\sim}{\theta_k},\underset{\sim}{\theta_{k'}}\,|\,y_{\sim},\underset{\sim}{\alpha}) = \underset{\sim t\times t}{\underset{\sim}{0}}$$

where $\underset{\sim t\times t}{\underset{\sim}{0}}$ is the null matrix.

Now , given $\underset{\sim}{\theta_k}$ the probability that among the non-sampled units in cluster k there will be $(Y_{kj}-y_{kj})$ units of type j, is

$$P(Y_{k1}-y_{k1},...,Y_{kt}-y_{kt}|\,\underset{\sim}{\theta_k}) = \frac{(M_k-m_k)!}{\Pi_{j=1}^{t}(Y_{kj}-y_{kj})!}\Pi_{j=1}^{t}\theta_{kj}^{Y_{kj}-y_{kj}}$$

Hence likelihood of $(\underset{\sim}{Y_k}-I_k\underset{\sim}{y_k}), k=1,2,...,N$ where $I_k=1(0)$ if $k\in S$ (otherwise), is

$$P((\underset{\sim}{Y}-I_k\underset{\sim}{y}),|\,\underset{\sim}{\theta_k},D,k=1,2,...,N) = \Pi_{k=1}^{N}\frac{(M_k-I_km_k)!}{\Pi_{j=1}^{t}(Y_{kj}-I_ky_{kj})!}$$

$$\times\Pi_{j=1}^{t}\theta_{kj}{}^{Y_{kj}-I_ky_{kj}} \tag{7.2.15}$$

where

$$D=(m_k,y_{kj},j=1,2,...,t,k\in S)$$

Using (7.2.14) , (7.2.15) and (7.2.2), the joint distribution,

$$P((\underset{\sim}{Y_k} - I_k \underset{\sim}{y_k}), \underset{\sim}{\theta_k}, k = 1, 2, ..., N | \underset{\sim}{\alpha}, D)$$

$$= \Pi_{k=1}^{N} \frac{(M_k - I_k m_k)!}{\Pi_{j=1}^{t}(Y_{kj} - I_k y_{kj})!} \Pi_{j=1}^{t} \theta_{kj}{}^{Y_{kj} - I_k y_{kj}}$$

$$\times \frac{\Gamma(\alpha + I_k m_k)}{\Pi_{j=1}^{t}\Gamma(\alpha_j + I_k y_{kj})} \Pi_{j=1}^{t} \theta_{kj}{}^{I_k y_{kj} + \alpha_j - 1}$$

$$= \Pi_{k=1}^{N} \frac{(M_k - I_k m_k)!}{\Pi_{j=1}^{t}(Y_{kj} - I_k y_{kj})!}.$$

$$\frac{\Gamma(\alpha + I_k m_k)}{\Pi_{j=1}^{t}\Gamma(\alpha_j + I_k y_{kj})} \Pi_{j=1}^{t} \theta_{kj}{}^{Y_{kj} + \alpha_j - 1} \qquad (7.2.16)$$

Integrating out $\underset{\sim}{\theta_k}$ from (7.2.16) we get the marginal of $(\underset{\sim}{Y_k} - I_k \underset{\sim}{y_k}, k = 1, 2, ..., N)$ as

$$P(\underset{\sim}{Y_k} - I_k \underset{\sim}{y_k}, k = 1, 2, ..., N | \underset{\sim}{\alpha}, D)$$

$$= \Pi_{k=1}^{N} \{ \frac{(M_k - I_k m_k)!}{(Y_{kj} - I_k y_{kj})!} . \frac{\Gamma(\alpha + I_k m_k)}{\Pi_{j=1}^{t}\Gamma(\alpha_j + I_k y_{kj})} \}.$$

$$\frac{\Pi_{j=1}^{t}\Gamma(Y_{kj} + \alpha_j)}{\Gamma(\alpha + M_k)}$$

$$= \Pi_{k=1}^{N} \frac{\Pi_{j=1}^{t} \binom{Y_{kj} + \alpha_j - 1}{I_k y_{kj} + \alpha_j - 1}}{\binom{M_k + \alpha - 1}{M_k - I_k m_k}}$$

$$= \Pi_{k=1}^{N} \frac{\Pi_{j=1}^{t} \binom{Y_{kj} + \alpha_j - 1}{Y_{kj} - I_k y_{kj}}}{\binom{M_k + \alpha - 1}{M_k - I_k m_k}}$$

Hence, $E(Y_{kj} - I_k y_{kj} | \underset{\sim}{\alpha}) = \dfrac{(I_k y_{kj} + \alpha_j)(M_k - I_k m_k)}{\alpha + I_k m_k}$

i.e.,

$$\begin{aligned} E(Y_{kj} | \underset{\sim}{\alpha}) &= \frac{I_k y_{kj}(M_k + \alpha)}{I_k m_k + \alpha} + \frac{\alpha_j (M_k - I_k m_k)}{I_k m_k + \alpha} \\ &= M_k (I_k \frac{y_{kj}}{m_k}) \lambda_k + (1 - \lambda_k) \frac{\alpha_j}{\alpha} \qquad (7.2.17) \end{aligned}$$

$$\text{where} \lambda_k = \frac{1 + \frac{\alpha}{M_k}}{1 + \frac{\alpha}{m_k}} \qquad (7.2.18)$$

Therefore,

$$E(\frac{Y_{kj}}{M_k}|\underset{\sim}{\alpha}, D) = E(\delta_{kj}|\underset{\sim}{\alpha}) = \lambda_k.I_k\frac{y_{kj}}{m_k} + (1-\lambda_k)\frac{\alpha_j}{\alpha} \quad (7.2.19)$$

$$\Rightarrow \quad E(P_j|\underset{\sim}{\alpha}, D) = \sum_{k=1}^{N}\rho_k(\lambda_k I_k\frac{y_{kj}}{m_k} + (1-\lambda_k)\frac{\alpha_j}{\alpha}) \quad (7.2.20)$$

Again,

$$E\{(Y_{kj} - I_k y_{kj})(Y_{kj} - I_k y_{kj} - 1)|\underset{\sim}{\alpha}, D\}$$

$$= \frac{(I_k y_{kj} + \alpha_j)(I_k y_{kj} + \alpha_j + 1)(M_k - I_k m_k)(M_k - I_k m_k - 1)}{(I_k m_k + \alpha + 1)(I_k m_k + \alpha)} \quad (7.2.21)$$

$$E\{(Y_{kj} - I_k y_{kj})^2|\underset{\sim}{\alpha}, D\} = \frac{(I_k y_{kj} + \alpha_j)(M_k - I_k m_k)}{I_k m_k + \alpha)(I_k m_k + \alpha + 1)}\{M_k(I_k y_{kj} + \alpha_j + 1) - m_k(I_k y_{kj} + \alpha_j) - I_k y_{kj} - \alpha_j + \alpha\} \quad (7.2.22)$$

Using above results we get

$$V(Y_{kj}|\underset{\sim}{\alpha}, D) = M_k^2(1-\lambda_k)(\frac{1}{\alpha} + \frac{1}{M_k})\frac{(I_k y_{kj} + \alpha_j)(I_k m_k + \alpha - I_k y_{kj} - \alpha_j)}{(\alpha + I_k m_k)(I_k m_k + \alpha + 1)}$$

Hence,

$$V(P_j|\underset{\sim}{\alpha}) = \sum_{k=1}^{N}\rho_k^2 V(\delta_{kj}|\underset{\sim}{\alpha})$$

$$= \sum_{k=1}^{N}\rho_k^2(1-\lambda_k)(\frac{1}{\alpha} + \frac{1}{M_k})\frac{(i_k y_{kj} + \alpha_j)}{(I_k m_k + \alpha)}. \frac{(I_k m_k + \alpha - I_k y_{kj} - \alpha_j)}{(I_k m_k + \alpha + 1)} \quad (7.2.23)$$

7.3 Present Study Under Hypergeometric Setup

7.3.1 Moments of P_i

In a finite population set up , we usually consider a categorical data to have hyper geometric distribution . Here we choose the Dirichlet-Multinomial (DM) distribution as a prior of the parameters of Hypergeometric distribution. Hence our likelihood function becomes

$$L = \Pi_{k \in S}\Pi_{i=1}^{t}\binom{Y_{ki}}{y_{ki}} / \binom{M_k}{m_k} \quad (7.3.1)$$

Now the DM distribution is defined by

$$
\begin{aligned}
P(n_1, n_2, ..., n_k) &= \frac{N!}{\Pi_{j=1}^{k}} E[\Pi_{j=1}^{k} p_j^{n_j}] \\
&= \frac{N!}{(\sum_{j=1}^{k} \alpha_j)^{[N]}} \Pi_{j=1}^{k} \frac{\alpha_j^{[n_j]}}{n_j!}, \quad (n_j \geq 0, \ \sum_{j=1}^{N} n_j = N)
\end{aligned}
$$

where

$$h^{[j]} = h(h+1)...(h+j-1).$$

[Johnson & Kotz, Discrete Distributions, p.,308]

Here we consider that $\underset{\sim}{Y_k}$'s have independent DM distributions with parameters M_k and α_k. Therefore the joint prior distribution of $\underset{\approx}{Y}$ is

$$P(\underset{\sim}{y} \mid \underset{\sim}{\alpha}) = \Pi_{k=1}^{N} \frac{M_k!\Gamma(\alpha)}{\Gamma(\alpha + M_k)} \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + Y_{ki})}{Y_{ki}!\Gamma(\alpha_i)} \tag{7.3.2}$$

Hence the joint distribution of $\underset{\sim}{y_k}$'s and $\underset{\sim}{Y_k}$'s is

$$
\begin{aligned}
P(\underset{\sim}{y}, \underset{\sim}{Y} \mid \underset{\sim}{\alpha}) &= \Pi_{k=1}^{N} \frac{M_k!\Gamma(\alpha)}{\Gamma(\alpha + M_k)} \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + Y_{ki})}{(Y_{ki})!\Gamma(\alpha_i)} \frac{\Pi_{i=1}^{t} \binom{y_{ki}}{y_{ki}}}{\binom{M_k}{m_k}} \\
&= \Pi_{k=1}^{N} \frac{(M_k - I_k m_k)! I_k m_k! \Gamma(\alpha)}{\Gamma(\alpha + M_k)} \\
&\quad \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + Y_{ki})}{\Gamma(\alpha_i)(I_k y_{ki})!(Y_{ki} - I_k y_{ki})!} \\
&= \Pi_{k=1}^{N} \frac{(M_k - I_k m_k)!\Gamma(\alpha + I_k m_k)}{\Gamma(\alpha + M_k)} \\
&\quad \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + Y_{ki})}{\Gamma(\alpha_i + I_k y_{ki})(Y_{ki} - I_k y_{ki})!} \\
&\quad \Pi_{k=1}^{N} \frac{(I_k m_k)!\Gamma(\alpha)}{\Gamma(\alpha + I_k m_k)} \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + I_k y_{ki})}{\Gamma(\alpha_i)(I_k y_{ki})!}
\end{aligned} \tag{7.3.3}
$$

Integrating out $Z_{ki} = (Y_{ki} - I_k y_{ki})$ in (7.3.3) we get the marginal of $\underset{\sim}{y_k}$'s as

$$P(\underset{\sim}{y_k}, k = 1, 2, ..., N \mid \underset{\sim}{\alpha}) = \Pi_{k=1}^{N} \frac{(I_k m_k)!\Gamma(\alpha)}{\Gamma(\alpha + I_k m_k)} \Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + I_k y_{ki})}{\Gamma(\alpha_i)(I_k y_{ki})!} \tag{7.3.4}$$

and hence the posterior distribution of $Y_{ki} - I_k y_{ki}$, given $\underset{\sim}{y_k}$ is

$$P(\underset{\sim}{Y_k} - I_k \underset{\sim}{y_k}, k = 1, 2, ..., N | \underset{\sim}{y_k}, \underset{\sim}{\alpha})$$

$$= \Pi_{k=1}^{N} \frac{M_k - I_k m_k)!\Gamma(\alpha + I_k m_k)}{\Gamma(\alpha + M_k)}$$

$$\Pi_{i=1}^{t} \frac{\Gamma(\alpha_i + Y_{ki})}{\Gamma(\alpha_i + I_k y_{ki})(Y_{ki} - I_k y_{ki})!} \qquad (7.3.5)$$

which is also a Dirichlet-Multinomial (DM) distribution with parameters $(M_k - I_k m_k)$ and $(\alpha_i + I_k y_{ki}), i = 1, 2, ...t$.

Now , the posterior mean of Y_{ki} given y_{ki} is obtained as

$$\begin{aligned} E(Y_{ki} | \underset{\sim}{y_k}, \underset{\sim}{\alpha}) &= I_k y_{ki} + \frac{\alpha_i + I_k y_{ki}}{\alpha + I_k m_k}(M_k - I_k m_k) \\ &= M_k \lambda_k \frac{y_{ki}}{m_k} + (1 - \lambda_k)\frac{\alpha_i}{\alpha}.M_k \qquad (7.3.6) \end{aligned}$$

$$\text{where } \lambda_k = \left.\begin{matrix} \frac{1+\frac{\alpha}{M_k}}{1+\frac{\alpha}{m_k}} & \text{if } k \in S \\ 0 & \text{otherwise} \end{matrix}\right\}$$

$$\text{i.e., } E(\delta_{ki} | \underset{\sim}{y_k}, \underset{\sim}{\alpha}) = \lambda_k \frac{y_{ki}}{m_k} + (1 - \lambda_k)\frac{\alpha_i}{\alpha} \qquad (7.3.7)$$

and

$$Var(\delta_{ki} | \underset{\sim}{y_k}, \underset{\sim}{\alpha}) = (1 - \lambda_k)(\frac{1}{\alpha} + \frac{1}{M_k})\frac{\alpha_i + I_k y_{ki})}{(\alpha + I_k m_k)}$$

$$\frac{(\alpha + I_k m_k - \alpha_i - I_k y_{ki})}{(\alpha + I_k m_k + 1)} \qquad (7.3.8)$$

which implies

$$Var(P_i) = \sum_{k=1}^{N} \rho_k^2 (1 - \lambda_k)(\frac{1}{\alpha}$$

$$+ \frac{1}{M_k})\frac{\alpha_i + I_k y_{ki})(\alpha + I_k m_k - \alpha_i - I_k y_{ki})}{(\alpha + I_k m_k)(\alpha + i_k m_k + 1)} \qquad (7.3.9)$$

Hence we observe that both the cases considered in section (7.2) and (7.3) are yielding the same estimate of P_i and their variances are also same.

Bibliography:

1. Ajgaonkar, S. G. P. (1967) :Unbiased estimator of the variance of the Narain, Horvitz and Thompson estimator. *Sankhya A*, vol 29, pp 55-60.
2. Asok, C. and B.V. Sukhatme (1976) :On Sampford's procedure of unequal probability sampling without replacement. *Jour. Amer. Stat. Assoc.*, 71,912-918.
3. Arnold S.F. (1979) : Linear models with exchangeably distributed errors. *Jour. Amer. Stat. Assoc.*, 74, 194-199.
4. Avadhani,M.S. and B.V. Sukhatme (1967) : Controlled sampling with varying probability and without replacement. *Aust. Jour. of Stat.*, 9, 8-15.
5. ——————— (1968) : Stratified sampling for designing controlled simple random sampling. *Aust. Jour. of Stat.*, 10, 1-7.
6. ——————— (1973) : Controlled sampling with unequal probabilities and without replacement. *Int. Stat. Review.*,41(2), 175-182.
7. Bandyopadhyay S., A.K. Chattopadhyay & S. Kundu (1977) : On estimation of population total. *Sankhya C*, 39, 28-42.
8. Basu, D. (1971) : An essay on the logical foundation of survey sampling, part one. In V.P. Godambe and D.A. Sprott, Eds. *Foundations of Statistical Inference. Toronto : Holt, Rinehart & Winston.*, 203-242.
9. Bolfarine H. (1991) : Finite-population prediction under error-in-variables superpopulation models. *The Canadian Jour. of Stat.* vol.19, No.2,191-207.
10. Bolfarine H. and Zacks S. (1992) : Prediction theory for finite populations. *Springer-Verlag NewYork Inc.*
11. Brauer Alfred (1971) : A remark on the paper of H.H. Schaefer : Abschatzung der nichttrivalen Eigenwerte Stochastischer Matrizen. *Numer. Math.*
12. Brewer K.R.W (1963) : A model of systematic sampling with unequal probabilities. *Aust. Jour. of Stat.,* 5, 5-13
13. ——— (1979) : A class of robust sampling designs for large-scale surveys. *Jour. Amer. Stat. Assoc.,* 74, No. 368, 911-915.

14. ———, M.Samiuddin & H.Asad (1989) : Small sample performance of five unequal probability sampling estimators. (*Manuscript seen by courtsey of the authors*).

15. ——— & G.C. Undy (1969) : Samples of two units drawn with unequal probability without replacement. *Aust. Jour. of Stat.*, 4, 89-100.

16. Callebaut D.K. (1965) : Generalisation of Cauchy-Schwartz inequality. *Jour. Math. Analysis and Applications* 12, 491-494.

17. Carroll, J.L. and H.O. Hartley (1964) : The symmetric mathod of unequal
probability sampling without replacement. *Biometrika. (abstract only)*,20, 908-909.

18. Cassel C.M., C.E. Sarndal and J.H. Wretman (1976) : Some results on generalised difference estimator and generalised regression estimator for finite population. *Biometrika.* 63, 614-620.

19. ——————————— (1977) : Foundations of inference in survey sampling. *New York : John Wiley & Sons.*.

20. Chakraborty M.C (1963) : On the use of incidence matrices of designs in sampling from finite populations, *Jour. of the Ind. Stat. Assoc.*, 1, 78-85.

21. Chang, H.C. & T.H. Lin (1985) : Estimating variance of a finite population under a certain linear regression model. *Tamkang Jour. Management*, 6, 1-10.

22. Chao M.T (1982) : A general purpose unequal probability sampling plan. *Biometrika.*, 69, 653-656

23. Chaudhury A. (1974) : On some properties of the sampling scheme due to Midzuno. *Cal. Stat. Assoc. Bull.*, 23, 1-20.

24. ——— (1976) : A non-negativity criterion for a certain variance estimator. *Metrika*, 23, 201-205.

25. ——— (1978) : On estimating the variance of a finite population. *Metrika*, 25, 65-76.

26. ———(1981) : Non-negative unbiased variance estimators.*In: D.Krewski, R.Platek and J.N.K.Rao, eds. Current Topics in Survey Sampling. Academic Press, New-York.* 317-328.

27. ——— and J.W.E. Vos (1988) : Unified theory and strategies of survey sampling. *North Holland.*

28. Cochran W.G. (1977) : Sampling Techniques. 3rd Edn.*New York, Wiley.*

29. ———— (1984) : Sampling Techniques. *Wiley Eastern Ltd.*

30. Cramer H. (1946) : Mathematical Methods of Statistics. *Princeton N.J.: Princeton University Press.*

31. Das A.C. (1951) : Two phase sampling and sampling withh varying probabilities. *BISI*, 33, 105-112.

32. Das, A.K. & T.P. Tripathi (1978) : Use of auxiliary information in estimating the finite population variance. *Sankhya C* ,139-148.

33. Deshpande M.N. (1978) : A new sampling procedure with varying probabilities. *Jour. of the Ind. Soc. of Agr. Stat.*, 30, 110-4.

34. Dodds D.J. and J.G. Fryer (1971) : Some families of selection procedures of sampling with probability proportional to size. *Jour. Roy. Stat. Soc. B*, 33, 263-274.

35. Durbin J. (1953) : Some results in sampling theory when the unity are selected with unequal probabilities : *Jour. Roy. Stat. Soc. B*, 15, 262-9

36. Durbin J. (1967) : for the estimation of sampling errors (design of multistage surveys). *App. Stat.*, 16, 152-164.

37. Ericson, W.A. (1974) : A note on the non-existence of minimum variance unbiased estimator for a finite population total. *Sankhya C* 36, 181-184.

38. Espejo M.R. and Espejo M.M.R. (1992) : Equilibrated strategy for population variance estimation. *Unpublished manuscript seen by courtsey of the authors.*

39. Fellegi I. (1963) : Sampling with varying probabilities without replacement : rotating and non-rotating samples: *Jour. Amer. Stat. Assoc.*, 58, 183-201.

40. Foody. W & A. Hedayet (1977) : On theory and applications of BIB designs with repeated blocks : *Ann. Statist*, 5,932-

41. Fuller W.A. (1971) : A procedure for selecting nonreplacement unequal probability samples. *Unpublished manuscript seen by courtsey of the author.*

42. ———— (1987) : Measurement Error Models. *John Wiley. N.Y.*

43. ——— and C.T. Isaki (1981) : Survey designs under super-population models. *"Current topics in survey sampling", ed. D. Krewski, J.N.K. Rao and R. Plateck, New York, Academic Press.*

44. Gabler S. (1984) : On unequal probability sampling : sufficient conditions for the superiority of sampling without replacement. *Biometrika*, 71, 171-175.

45. Gabler S. (1987) : The nearest proportional to size sampling design. *Comm. Stat. Theory. Math.* , 16(4), 1117-1131.

46. Godambe V.P. (1955) : A unified theory of sampling from finite populations . *Jour. Roy. Stat. Soc. B* 17, 269-278.

47. ——— (1966) : A new approach to sampling from finite population, I,II. *Jour. Roy. Stat. Soc. B*, 28, 310-328.

48. ——— & V.M. Joshi (1965) : Admissibility and Bayes estimation in sampling from finite populations,I. *Ann. Math. Stat.* 36, 1707-1722.

49. ——— & M.E. Thompson (1977) : Robust near optimal estimation in survey practice. *Bull. Int. Stat. Inst.* 47(3), 129-146.

50. Ghosh, M. & G. Meeden (1983) : Estimation of the variance in finite population sampling. *Sankhya B* 45, 362-375.

51. Goodman L.A. (1960) : On the exact variance of products. *jour. Amer. Stat. Assoc.*,55.

52. Goodman R. & L. Kish (1950) : Controlled selection - a technique in probability sampling *Jour. Amer. Stat. Assoc.* , 45, 350-72

53. Gupta V.K. , A.K. Nigam & P. Kumar (1982) : On a family of sampling schemes with inclusion probability proportional to size. *Biometrika* , 69, 191-6.

54. Hájek J. (1959) : Optimum strategy and other problems in probability sampling. *Casopis. Pest. Math.* 84, 387-423.

55. Hanif M. & K.R.W. Brewer (1980) : Sampling with unequal probabilities without replacement: a review ; *Int. Stat. Rev.*, 48, 317-335.

56. Hansen M.H., W.G. Madow & B.J. Tepping (1983) : An evaluation of model-dependent and probability sampling inferences in sample surveys. *Jour. Amer. Stat. Assoc.*, 78, 776-793.

57. Hansen M.H & W.N. Hurwitz (1943) : On the theory of sampling from finite populations. *Ann. Math. Stat.* , 14, 333-362.

58. Hanurav T.V. (1962 a) : An existence theorem in sample survey. *Sankhya A*, 24, 327-330.

59. ——— (1962 b) : On Horvitz and Thompson estimator. *Sankhya A*, 24, 429-436.

60. ——— (1962 c) : Some sampling schemes in probability sampling.
Sankhya A, 24, 921-928.

61. ——— (1966) : Some aspects of unified sampling theory. *Sankhya A* 28, 175-204.

62. Hanurav T.V. (1967) : Optimum utilisation of auxiliary information :
πps sampling of two unity from a stratum. *Jour. Roy. Stat. Soc. B*, 23, 374-391

63. Hartley H.O. & J.N.K. Rao (1962) : Sampling with unequal probabilities and without replacement. *AMS*, 33, 350-74

64. Hartley H.O. & A. Ross (1954) : Unbiased ratio estimators. *Nature* 174, 270-271.

65. Hedayet A.S., B.Y. Lin & J. Stufken (1989) : The construction of πps sampling designs though a method of emptying boxes *Ann. Statist.* 17, 1986-1995.

66. ———, C.R.Rao & J. Stafker (1986) : Sampling plans excluding contaguous units. *Tech. Rep.*

67. Hege V.S. (1965) : Sampling designs which admit uniformly minimum variance unbiased estimates. *Cal. Statist. Assoc. Bull* 14, 160-162.

68. Herzel A. (1986) : Sampling without replacement with unequal probability sampledesigns with preassigned joint inclusion probabilities of any order. *Metron*, XLIV (1) 49-68.

69. Horvitz D.G & D.J. Thompson (1952) : A generalization of sampling without replacement from a finite universe. *Jour. Amer. Stat. Assoc.*, 47, 663-685.

70. Iachan Ronaldo (1985) : Robust designs for ratio and regression estimation. *J. of Stat. Plan. and Inf. II*, 149-161.

71. Isaki C.T. (1983) : Variance estimation using auxiliary information. *Jour. Amer. Stat. Assoc.* 78, *No.* 381, 117-123.

72. Isaki C.T. & W.A. Fuller (1982) : Survey design under the regression superpopulation model. *Jour. Amer. Stat. Assoc.*, 77, 89-96.

73. I.S.I. Survey Report : A survey of oral cancer patients in and arround calcutta, 1988-89. *Applied Statistics, Surveys and Computing Division, I.S.I. Calcutta.*

74. Jessen R.J. (1942) : Statistical investigations of a sample survey for obtaining farm facts. *Iowa Agr. Exp. Stat. Res. Bull.* 304.

75. ——— (1978) : Statistical Survey Techniques; *1st Ed. John Wiley & Sons.*

76. Konijn (1973) : Statistical Theory of Sample Survay Design and Analysis; *North Holland.*

77. Kott P.S. (1986 a) : Some asymptotic results for the systematic and stratified sampling of a finite population. *Biometrika*, 73,485-491.

78. ——— (1986 b) : When a mean of ratios is the best linear unbiased estimator under a model. *Am. Stat.*, 40 (3), 202-204.

79. Kumar P., V.K. Gupta & S.K. Agarwal (1985) : On variance estimation in unequal probability sampling. *Aust. J. Statist.*, 27(2), 195-201.

80. Lahiri D.B. (1951) : A method of sample selection providing unbiased ratio estimates. *Bull. Int. Statist. Inst.* 33(2), 133-140.

81. Lanke J. (1974) : On UMV estimators in survey sampling. *Metrika* 20, 196-202.

82. —— (1975) : Some contribution to the theory of survey sampling. *Lund. AV-Centralen*

83. Liu T.P. (1974a) : A general unbiased estimator for the variance of a finite population. *Sankhya* 36, 23-32.

84. —— (1974b) : Bayes estimation for the variance of a finite population. *Metrika* 21, 127-132.

85. —— & M.E. Thompson (1983) : Properties of estimators of quadratic finite populations functions : the batch approach. *Ann. Statist.* 11, 275-285.

86. Madow W.G. (1949) : On the theory of systimatic sampling II. *AMS*, 20, 333-54.

87. Mahalanobis P.C. (1946) : Recent experiments in Statistical sampling in the Indian Statistical Institute. *J. Roy. Stat. Soc. B* 109, 325-370.

88. Midzuno H. (1950) : An outline of the theory of sampling systems. *Ann. Inst. of Stat. Math.*, 1, 149-56.

89. ——— (1952) : On The sampling system with probability proportional to sum of sizes. *Ann. Inst. of Statist. Math.* 3, 99-107.

90. Montanari G.E. (1979) : Post-sample efficient QR-prediction in large sample surveys. *ISR*,55(2), 191-202.

91. Mukhopadhyay P. (1972) : A sampling scheme to realise a preassigned set of inclusion probabilities of first two orders. *Cal. Stat. Assoc. Bull.*, 21, 87-122.

92. ——— (1974) : *πps sampling schemes to base HTE. Cal. Stat. Assoc. Bull.*, 23, 21-44.

93. ——— (1977) : Robust estimators of finite population total under certain linear regression models. *Sankhya C*, 39, 71-87.

94. ——— (1978) : Estimating the variance of a finite population under a superpopulation model. *Metrika* 25, 115-122.

95. ——— (1982) : Optimum strategies for estimating the variance of a finite population under a superpopulation model. *Metrika* 29, 143-158.

96. ——— (1984) : Optimum estimation of finite population variance under generalised random permutation models. *Cal . Statist. Assoc. Bull.* 33, *No.*129-130, 93-104.

97. ——— (1990) : On asymptotic properties of a generalised predictor of finite population variance. *Sankhya,B*, 52,pt.3 343-346.

98. ——— (1991) : Varying probability without replacement sampling designs, a review. *ISI Tech. Rep. No. ASC/91/3.*

99. ——— (1992) : Some Inferential Problems in Survey Sampling. (*to be published by Wiley Eastern Ltd.*)

100. ———— (1994) (a) : Prediction in finite population under error-in-variables superpopulation models. *Jour. of Stat. Plan. and Inf.* 41, p. 151-161.

101. ———— (1994) (b) : Prediction of finite population total using multiauxiliary information under measurement error models. Festschrift in honour of Professor A.K. Bhattacharyya, eds. S.P. Mukherjee, A. Chaudhury, S.K. Basu, Dept. of Statistics, Presidency College.

102. ———— (1994) (c) : Bayes and minimax procedures for finite population sampling under measurement error models. *Comm. Stat.(T & M).* 23(7), 1953-1961.

103. ———— & S. Bhattacharyya (1990-91) : Estimating a finite population variance under some general linear models with exchangeable errors. *Cal. Stat. Assoc. Bull.*,40, 219-228.

104. ———— & ———— (1991) : On a comparison between Midzuno strategy and PPSWR strategy. *Statistics*, 22, 291-294.

105. ———— & K.Vijayan (1990) : On non-negative unbiased estimator of quadratic forms in finite population sampling. *University of Western Australia, Tech. Rep. No. 41.*

106. Murthy M.N. (1957) : Ordered and unordered estimators in sampling without replacement. *Sankhya*, 18, 379-390.

107. ———— (1967) : Sampling Theory and Methods. *Statistical Publishing Society, Calcutta.*

108. ———— (1977) : Sampling theory and methods. *Statistical Publishing Society, Calcutta.*

109. Nandram B. and Sedransk J.(1993) : Bayesian predictive inference for a finite population proportion: Two-stage cluster sampling. *Jour. Roy. Stat. Soc. B*,55,No.2,p.399-408.

110. Nanjamma N.S., Murthy M.N. and Sethi V.K. (1959) : Some sampling systems providing unbiased ratio estimators. *Sankhya* Vol. 21, 299-314.

111. Narain, R.D. (1951) : On sampling with varying probabilities without replacement, *Jour. of the Ind. Soc. of Agr. Stat.*, 3, 169-74.

112. Olkin I. (1958) : Multivariate ratio estimation for finite populations. *Biometrika* 45, 154-165.

113. Pathak P.K. and N.D. Sukla (1966) : Non-negativity of a variance estimator. *Sankhya*, 28, [illegible]

114. Pereira C.A.B. and J. Rodrig[illegible] ([illegible]) : Robust linear prediction in finite population. *ISR*, 51, 293-300.

115. Pfeffermann D. (1984) : A note on large sample properties of balanced samples. *Jour. Roy. Stat. Soc. B*, 46, 38-41.

116. Raghava Rao D. (1971) : Construction and combinatotial problems in design of experiments. John Wiley and Sons. N.Y.

117. Raj D. (1956) :Some estimators in sampling with varying probabilities without replacement. *Jour. Amer. Stat. Assoc.*, 51.

118. —— (1972) :The Design of Sample Surveys. *McGraw-Hill*

119. Ramakrishnan M.K. (1971) : Generalisation of Yates and Grundy variance estimators.
Research Report 92/MKR/1, Manchester- Sheffield.

120. Rao C.R. (1973) : Linear Statistical Inference and its Applications. *2nd edition, Wiley, New York.*

121. Rao, J.N.K. (1963) : On two systems of unequal probability sampling without replacement. *Ann. of Inst. of Stat. Math.*, 15, 67-72.

122. ———— (1969) : On three procedures of unequal probability sampling without replacement. *Jour. Amer. Stat. Assoc.*, 58, 208-295.

123. ———— (1965) : On two simple properties of unequal probability sampling without replacement. *Jour. of the Ind. Soc. of Agr. Stat.*, 3, 173-180.

124. ———— (1978) : Sampling design involving unequal probabilities of selection and robust estimation of a finite population total. *Contribution to Survey Sampling and Applied Statistics, edited by H.A. David. Academic Press,* 69-87.

125. ———— (1979) : On deriving mean square errors and their non-negative unbiased estimators in finite population sampling. *Jour. of the Ind. Stat. Assoc.*, 17, 125-136.

126. ————, Hartley, H.O. & Cochran, W.G. (1962) : On a simple procedure of unequal probability sampling without replacement. *Jour. Roy. Stat. Soc. B,* 24, 482-91.

127. ——— & A.K. Nigam (1989) : Controlled sampling with probability proportional to aggregate size. *Tech. Report No. 133. Laboratory for Research in Stat. and Prob. Carleton University, Ottawa*

128. ——— & K. Vijayan (1977) : On estimating the variance in sampling with probability proportional to aggragate size. *Jour. Amer. Stat. Assoc.*, 72, 579-584.

129. Rao P.S.R.S. & J.N.K. Rao (1971) : Small sample results for ratio estimators. *Biometrika* 58, 625-630.

130. Rao T.J. (1972) : On the variance of the ratio estimator. *Metrika*, 18, 209-215.

131. —— (1977) : Estimating the variance of the ratio estimator for the Midzuno-Sen sampling scheme. *Metrika*, 24, 203-208.

132. —— & M.P. Singh (1973) : On the choice of estimator in Survey Sampling. *Aust. Jour. of Stat.*, 93- .

133. Reilly P. and Patino-Leal H. : A Bayesian study of the error in variable models. *Technometrics*, 23(3), 221-231.

134. Robinson P.M. and C.E. Sarndal (1983) : Asymptotic properties of the generalised regression estimator in probability sampling. *Sankhya B*,45, 240-248.

135. ——— and K.W. Tsui (1979) : Optimal asymptotically design unbiased estimation of population total. *Tech. Rep. 634, Dept. of Stat., University of Wisconsin, Madison.*

136. Rodrigues J. H.Bolfarine & A.Rogatko (1985) : A general theory of prediction in finite populations. *Int. Statist. Rev.* 53, 239-254.

137. Royall R.M. (1970) : On finite population sampling theory under certain linear regtrssion models. *Biometrika* 57, 377-387.

138. ——— (1976) : The linear least square prediction approach to two-stage sampling. *Jour. Amer. Stat. Assoc.*,71, 657-664.

139. ——— & W.G. Cumberland (1981 a) : An empirical study of the ratio estimator and estimator of its variance. *Jour. Amer. Stat. Assoc.*, 76, 66-77.

140. ——————— (1981 b) : The finite population linear regression estimator and estimators of its variance – An empirical study. *Jour. Amer. Stat. Assoc.* 76, 924-930.

141. ―――――――― (1981 c) : Prediction models and unequal probability sampling. *Jour. Roy. Stat. Soc. B*, 49, 353-367.

142. ――― & J. Herson (1973) : Robust estimation in finite populations. *Jour. Amer. Stat. Assoc.* 68, 880-893.

143. ――― and D.Pfeffermann (1982) : Balanced samples and robust Bayesian inference in finite population sampling. *Biometrika.*, 69,401-409.

144. Sampford M.R. (1967) : On sampling without replacement with unequal probability of selection. *Biometrika,* 54, 499-513

145. Sankarnarayanan K. (1969) : An IPPS sampling scheme using Lahiri's mehtod of selection. *Jour. of the Ind. Soc. of Agr. Stat.,* 21(2), 59-66.

146. ――――― (1980) : A note on the admissibility of some non-negative quadratic estimators. *J. Roy. Stat. Soc. B* 42, 387-389.

147. Sarndal C.E. (1980) : On π-inverse weighting versus best linear weighting in probability sampling. *Biometrika*, 67(3), 639-650.

148. ――― (1981) : Comment to a paper by Royall and Cumberland. *Jour. Amer. Stat. Assoc.*, 76, 81-82.

149. ――― , B. Swensson & J. Wretman (1992) : Model Assisted Survey Sampling. *Springer-Verlag.*

150. Saxena P.R., P. Singh, & A.K. Srivastava (1986) : An unequal probability sampling scheme. *Biometrika,* 73, 761-763.

151. Scott A.J., K.R.W. Brewer & E.W.H. Ho (1978) : Finite population sampling and robust estimation. *Jour. Amer. Stat. Assoc.* 73, 359-361.

152. Scott A. and Smith T.M.F. (1969) : Estimation in multi-stage surveys. *Jour. Amer. Stat. Assoc.,* 64,830-840.

153. Sen A.R. (1952) : Present status of probability sampling and its use in the estimation of a characteristics (abstract). *Econometrika* 20, 103-

154. ―― (1953) : On the estimation of variance in sampling with varying probabilities. *Jour. Ind. Soc. Agr. Stat.*, 5, 119-127.

155. Sengupta S. (1988) : Optimality of design-unbiased strategy for estimating a finite population variance. *Sankhya B* 50, 149-152.

156. Seth G.R. (1966) : On estimates of variance of estimate of population total in varying probabilities. *Jour. Ind. Soc. Agr. Stat.*, 18, 52-66.

157. Sharma S.S. (1970) : On an estimator in T_3 - class of linear estimators in sampling with varying probabilities from a finite population. *Ann. Inst. Stat. Math.*, 22, 495-500.

158. Singh P. (1978) : A sampling scheme with inclusion - probability proportional to size. *Biomtrika*. 65, 450-54.

159. Singh P. & A.K. Srivastava (1980) : Sampling schemes providing unbiased regression estimators. *Biometrika*, 67, 205-9

160. Singh R.K. (1983) : Estimation of finite population variance using ratio and product methods of estimation. *Biometrical J.* 25, 193-200.

161. Sinha B.K. (1973) : On sampling schemes to realise preassigned sets of inclusion probabilities of first two orders. *CSAB*, 22, 69-110.

162. Skinner C.J. (1981) : Estimation of the variance of a finite population for cluster samples.*Sankhya B* 43, 392-398.

163. ——— (1983) : Multivariate prediction from selected samples . *Biometrika* 70, 289-292.

164. Som R.K. (1973) : A Manual of Sampling Techniques. *Heinemann Educational Books Ltd.*

165. Sprent P.(1966) : A generalized least squares approach to linear functional relationships. *Jour. Roy. Stat. Soc. Ser. B*, 28, 279-297.

166. Srivastava J.N. & F. Salah (1985) : Need of t-design in sampling theory. *Utilitas Mathematica*.,25, 5-17.

167. Strauss I. (1982) : On the admissibility of estimators for the finite population variance. *Metrika* 29, 195-202.

168. Sukhatme P.V. (1953) : Sampling theory of Surveys with applications . *Published by The Iowa State College Press, Ames, Iowa, U.S.A.*.

169. ——— and Sukhatme B.V. (1970) : Sampling Theory of Surveys with applications. *2nd revised Ed., Asia Publishing House.*

170. Sukhatme B.V. & M.S. Avadhani (1965) : Controlled selection – a technique in random sampling. *Ann. Inst. Stat. Math.*

171. Sunter A.B (1977) : List sequential sampling withequal or unequal probabilities without replacement, *App. Stat.*, 26, 261-8

172. ——— (1986) : Solutions to the problems of unequal probability sampling without replacement. *Int. Stat. Review*, 33-50.

173. Tallis G.M. (1978) : Note on robust estimation in finite populations. *Sankhya C*, 40, 136-138.

174. Tam.S. (1986) : Characterisation of best model-based predictors in survey sampling. *Biometrika.*,73, 232-235.

175. Thompson James, N. (Jr.) & J.M. Thoday (1979) : Quantitative genetic variation. *Academic Press, INC.*

176. Vardeman S. & G. Meeden (1983) : Admissible estimators in finite population sampling employing various types of prior information. *J. Stat. Plan. and Inf.* 7, 329-341.

177. Vijayan K. (1975) : On estimating variance in unequal probability sampling. *Jour. Amer. Stat. Assoc.*, 70, 713-716.

178. Wright R.L. (1983) : Sampling design with multivariate auxiliary information. *Jour. Amer. Stat. Assoc.* 78, 879-884.

179. Yamane T. (1967) : Elementary Sampling Theory. *Prentice Hall, Inc. New York.*

180. Yates F. & Grundy P.M. (1953) : Selection without replacement within strata with probability proportional to sizes. *Jour. Roy. Stat. Soc. B,* 15. 253-61.

181. Zacks S. & H. Solomon (1981) : Bayes equivariant estimators of the variance of a finite population for exponential priors. *Comm. Stat. Theory and Methods* 10, 427-437.

182. Zellner a, (1971) : An Introduction to Bayesian Inference in Econometrics. *Wiley, New York.*