CHAPTER I
CHAP. I

INTRODUCTION

1.0. Introduction.

The application of sampling methodology plays an important role in Sample Surveys. It aids in profit making of business organisations. It is also used as a research tool in the behavioral, medical, social and economic sciences. The main problem of survey sampling from finite populations consists of devising an appropriate sampling procedure for selecting a representative sample from a given population and developing an appropriate procedure for estimation of the population parameters of interest such as population mean or total or population ratio with a view to maximising the precision of the estimator within the available resources of time and cost or alternatively minimizing the cost for achieving a given level of precision. Systematic interest in the use of sampling methods appeared towards the end of last century when Kiaer used the "representative method" for collecting data independently of the census. Sampling then had a somewhat different meaning from today. In line with the present thinking the representative method used by Kiaer and others dealt with a part of the population and had its aim in drawing the inference from the sample that would be valid for
the populations. The primary difference in the earlier use of sampling method lay in the selections of the sample; it was thought that a purposive selection based on sampler's knowledge of population with regard to related characteristics was the best way of getting a sample that could be considered representative of the population. This type of selection dominated the earlier history of sampling methods, although clear traces can be found of early discussions of the merits of a random selection which was used for survey purposes by Tschuprow as early as the First World War (Romanovsky, 1946) and many others immediately thereafter (Hubback, 1926, Gini and Galvani, 1929; Clapham, 1930; Anderson, 1949 etc.). Professor Neyman's 1934 paper was in this respect the final link in a relatively long chain of developments, practices and experiences. Once for all his paper made it clear that random selection, unlike purposive selection, has in its basis a sound scientific theory which definitely gave to sample surveys the character of an objective research tool and made it possible to predict the validity of the survey estimates. Purposive selection nevertheless remained in use mainly because of its appeal as a cheap method, particularly in the form of the quota method as practiced by public opinion survey organisations, although it has been crystal-clear that there is no theory which would permit determining realiability of estimates based thereon.
What is more important is that Neyman's 1934 paper opened up numerous new avenues for fruitful research in the theory and philosophy of sample surveys, and in this sense it can be said to mark the beginning of a new era. The observations contained in this paper on such basic problems of surveys as the choice of sampling units, methods of estimation and the use of supplementary information provide the illustrations of this new approach. In earlier practices, units used for survey purposes were large clusters such as communes or villages. To illustrate the consequences of the use of large clusters, Neyman discussed the case presented by Gini and Galvani (1929). In order to get sample estimates of a census Gini and Galvani selected at random a sample of 29 counties out of 214 and found that the estimates obtained were not representative. With the clear knowledge of the relationships involved, Neyman explained that 10 percent of the elementary units grouped in a large number of smaller clusters would lead to highly reliable estimates. In subsequent developments this idea has become the basis for a considerable body of research on the efficiency of sample surveys. Likewise, Neyman pointed out that the use of supplementary information and the use of appropriate methods of estimation have a considerable bearing on the efficiency of surveys. His optimum allocation of units in sampling from a stratified population has been a clear demonstration of the possibilities that could be found along these lines. This is what exerted a significant influence on developments which have
taken place since 1934 in the theory of sample surveys. Not only considerable research followed with the aim of developing new and more efficient methods of estimation, but also the whole philosophy of sample surveys was built on this basis. The essence of this philosophy is expressed by the term "efficiency". Within the framework of sampling history prior to Neyman's paper this philosophy did not exist. Since his 1934 paper, it has become the goal, stimulating continuously new research in the theory of sample surveys and its practical applications.

1.1. Nature of the problem.

From times immemorial the concept of generalizing from a "part" of the population to the "whole" has been used more or less subjectively in daily life. But not until the later half of the nineteenth century, objective methods of generalizing from a part to the whole seems to have received much attention. In this case two questions arise (i) how to select the "part" from the "whole" and (ii) how to generalize from the selected part to the whole. The problem is one of finding that combination of selection and estimation procedure which would minimize the risk involved in generalizing from a part to the whole per unit of cost. Alternatively, the problem may be viewed as that of finding that combination of selection and estimation procedure which would minimize the cost, ensuring at the same time a specified precision for the inference from a part to the whole.
The earlier developments in this field relate to the second question posed above and the result has been a fairly well developed theory of estimation and statistical inference based on the simplest of selection procedures, namely, equal probability sampling with replacement. Improvements in selection procedures may be considered to have been initiated by Bowley (1926) who used stratified simple random sampling with proportional allocation. Neyman (1934) and Sukhatme (1935) considered the question of optimum allocation in stratified sampling.

During the decade 1940-50, considerable developments in sampling theory have taken place and the works of Cochran (1942) Hansen and Hurwitz (1943, 1946), Madow and Madow (1944), and Mahalanobis (1940, 1944, 1946) need special mention. Cochran considered the question of utilizing supplementary information at the estimation stage by using ratio and regression estimators. Hansen and Hurwitz considered the question of the supplementary information for selecting the units with probability proportional to a suitable measure of size and also the problem of non-response in surveys. Mahalanobis realized the importance of assessment of non-sampling errors as early as 1938 and developed the technique of interpenetrating sub-samples to assess the errors in surveys. Madow and Madow developed the technique of selecting units systematically. Cochran (1963), Deming (1950), Hansen, Hurwitz and
Madow (1953), Sukhatme (1953), and Yates (1962) have covered in detail the earlier developments in the theory of sampling in their respective books.

1.2. **Historical Developments.**

The name "random sampling" has often been used synonymously with "scientific sampling" or "probability sampling". Random selection implies a known probability of an element or unit being included in the sample. The most common procedure of random selection is that with "equal probabilities". In this case all units in the population have equal chances of being selected and a probability of selecting a unit at any draw is constant. Generally the sampling units differ much in their sizes. The selection of sampling units by simple random sampling method may not be always a very efficient procedure. To improve the efficiency of the sampling design, the technique of selection of units with unequal probability was developed. A theoretical base for the technique was given by Hansen and Hurwitz (1943). Hansen and Hurwitz (1943) introduced the selection of primary units (in a sub-sampling scheme) with probability proportional to some measure of their size for sampling of one primary sampling unit per stratum. When the sizes of primary sampling units or estimates of them are available,
Hansen and Hurwitz (1949) gave a method of determining optimum probability of selection which minimized the variance of the estimate for a given cost, and also some approximation to the optimum. Midzuno (1950) generalized the Hansen and Hurwitz approach to sampling a combination of \( n \) elements of the universe with probability proportional to some measure of size of the combination. Narain (1951) investigated the theory of sampling with variable probabilities when two units are selected from each stratum without-replacement. Sen (1952) developed formulae for a system of selection in which the first unit is chosen with probability proportional to size and subsequent units are chosen with equal probabilities and without replacement.

Lahiri (1951) was concerned with the bias in the ratio estimate and he set out to find an estimator which while retaining the character of a ratio estimate was also to be free from bias. The usual ratio estimate of the population total for single stage sampling is

\[
\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} X
\]
where $X$ is the sum of $X$ values ($\sum_{i=1}^{N} X_i = X$) for all units in the population. He saw that if a sampling scheme could be derived in which the probability of obtaining a given group of $n$ elements is proportional to the total measure of their auxiliary characteristics, then the estimate would be unbiased. He devised such a scheme. It must be said his approach was much more direct than that of Midzuno of whose work he was unaware, and he arrived at his sampling scheme, as he confessed, by geometrical intuition.

Sen (1952) worked out the consequences of the ideas contained in Midzuno's 1950 paper, and among other results he obtained independently the same type of ratio estimate as Lahiri's in the case of sampling without replacement and in two stages. Dalenius (1962) review paper reveals the fact that Hajek (1949) also discovered the ratio estimator of the same type as the one proposed by these three authors.

A more general approach to the problem of sampling finite populations with unequal probabilities and without replacement of elements was formulated by Horvitz and Thompson (1952). In their theory the probabilities of selection were completely arbitrary at each draw. They formulated three classes of linear estimators of the population total with the coefficient for each class depending on the order of draw, the presence or
absence of an element in a sample and the particular sample in question. Subsequently the general estimators corresponding to each of these features of sample formation would typify what has been designated as class one, class two and class three respectively. However, the logical consequences of these ideas were not explored except in respect of the estimator with coefficients depending on the appearance or non-appearance of the elements in the sample. These classes of linear estimators were scrutinized by Koop (1957,1963), Prabhu-Ajgaonkar (1962, 1965, 1967 a,b,c, 1970 a,b and 1971 a) and they formulated seven classes of linear estimators. Godambe (1955) defined the most general class of linear estimators for the population total and proved that it is not possible to construct a selection procedure with an associated unbiased linear estimator which is uniformly best for all populations excepting for some special cases completely characterized by Hanurav (1965), later termed as "unicluster designs."

The non-existence of the uniformly minimum variance unbiased estimator of the population total in the class of all linear unbiased estimators for any sample design has led to the choice of estimators from a sub-class of admissible estimators. Alternatively, various criteria have been put forward by many others to arrive at an optimum choice, namely (i) Bayesness
(Godambe 1955, Hajek 1958), (ii) invariance and regular class
(Rey and Chakravarti 1960) and (iii) hyper-admissibility
(Hanurav 1968) to quote a few.

Murthy (1957) showed that corresponding to any ordered
unbiased estimator there exists an unordered estimator which
is more efficient than the ordered one. He applied his theory
to Des Raj's ordered estimator to obtain a more efficient
estimator.

Rao, Hartley and Cochran (1962) modified Durbin's (1953)
grouping method by considering \( n \) groups instead of two and
spliting the population at random into \( n \) groups of sizes
\( N_1, \ldots, N_n (N_1 + \ldots + N_n = N) \). They proposed an unbiased estimator
of the population total and proved that the variance of their
estimator is always smaller than that of the customary estimator
in unequal probability sampling with replacement.

Using Bayes approach in sampling theory, it was first shown
by Cochran (1946) that whenever auxiliary information on a
supplementary variate closely related to the study variate
is available, we can utilize this information to formulate a
prior distribution for the variate under study. This is now well
known as the "super-population" concept. As stated earlier
with the criteria of unbiasedness and minimum variance, in
general, there does not exist a best estimator in the class of linear estimators. However, this desperate situation prevails only when no apriori knowledge is available on the study variate. When information on a positive valued supplementary variate is available on all units of the population in advance, Godambe (1955) has shown that in the class of all sampling strategies with a given expected number of distinct units, any strategy such that

(i) every sample has same number of distinct units,

(ii) probability of including any unit in the sample is proportional to the value of the supplementary variate for that unit and

(iii) the estimator is the Horvitz-Thompson estimator is best in the well defined Bayesian sense.

These optimum designs are termed by Hanurav (1965, 1967) as \( \alpha PS \) designs and the sampling scheme to construct such designs as \( \alpha PS \) schemes. The problem of constructing \( \alpha PS \) schemes has attracted the attention of many and notable contributions in this direction have been made by Goodman and Kish (1950), Midzuno (1952), Horvitz and Thompson (1952), Durbin (1953), Hajek (1959), Hanurav (1962b, 1965, 1967), Rao, Hartley and Cochran (1962), Rao (1967 a), Foreman and Brewer (1971) and many others.
It is well known that the technique of stratified sampling consists in classifying the population into a certain number of homogeneous groups called strata and then selecting samples independently from each group or stratum. Broadly, the important points that need careful consideration in the stratified sampling are:

i) choice of sampling design within strata,
ii) choice of stratification variable,
iii) allocation of sample size to strata,
iv) number of strata, and
v) determination of strata.

Prabhu-Ajgaonkar (1965), in search of a "serviceable" estimator introduced the "necessary bestness" criterion and proved the "necessary bestness" of the Horvitz-Thompson estimator $\hat{y}_{HT}$ in a sub-class of linear unbiased estimators of $Y$. Hege (1967) generalized his result to the class of all linear unbiased estimators and Ougus (1969) simplified Hege's proof. Hanurav (1968) introduced the "hyper-admissibility" criterion and established the unique hyper-admissibility of Horvitz-Thompson estimator $\hat{y}_{HT}$ in the class of all polynomial unbiased estimators of $Y$. Rao and Singh (1973) have shown that the principle hyper-surfaces (phs's) of dimension one plays the key role in arriving at the unique choice through these criteria and the
"necessary bestness" of Horvitz-Thompson estimator $\hat{\gamma}_{HT}$ implies that $\hat{\gamma}_{HT}$ is the only possible "hyper-admissible" estimator in a wide class of unbiased estimators of $Y$.

1.3. Notations and Definitions.

In this section we explain the notations and definitions which will be used in the thesis.

1.3.1. The Unified Theory: A "finite population" $\mathcal{U}$ is defined as a collection $U_1, U_2, \ldots, U_N$ where $N$ is a known finite number and $U_1, U_2, \ldots, U_N$ are distinguishable. We denote this population by

$$\mathcal{U} = (U_1, U_2, \ldots, U_N) \quad (1.3.1)$$

A list of such as (1.3.1) is called a "sampling frame" and $N$ is called the "size of the population."

A sample space

$$S = \left\{ s \right\} \quad (1.3.2)$$

is the collection of all samples $s$ from $\mathcal{U}$ where
\[ S = \{ s \} \text{ is the collection of all finite ordered sequences } s \text{ of units from } U \text{ in which case} \]

\[
\text{either } s = (U_{i_1}, U_{i_2}, \ldots, U_{i_n}) \quad (1.3.3)
\]

is a finite ordered sequence of (not necessarily distinct) units from \( U \)

\[
\text{Or } S = \{ s \} \text{ is the collection of all nonempty subsets } s \text{ of } U \quad (1.3.4)
\]

The number of units in a sample \( s \) is called the sample size of \( s \) and is denoted by \( n(s) \). Thus if \( s \) is a sequence of units from \( U \), \( n(s) \) denotes length of the sequence \( s \) and if \( s \) is a subset of \( U \), \( v(s) \) is the number of units in \( s \).

We consider a real valued variable \( \underline{y} \) (study variable \( \underline{y} \)) defined over \( U \) and taking value \( Y_i \) on the unit \( U_i \), \( 1 \leq i \leq N \). Let \( \underline{y} \) denote the vector

\[
\underline{y} = (Y_1, Y_2, \ldots, Y_N) \quad (1.3.5)
\]
The \( Y_i \)'s are unknown \textit{apriori} and the problem in general is to estimate real valued function of \( Y \), the parameter of interest which is a point in \( \mathbb{R}^n \) (the \( N \)-dimensional Euclidean space), called parametric function, on the basis of the observations \( Y_i \) for if's where \( s \) is a sample drawn with given probability \( p_s \) from the totality of all possible samples \( S \). We call the function \( p \) on \( S \) defining the probabilities \( p_s \) for all samples in \( S \), such that for every \( s \in S \), \( p_s \geq 0 \) and \( \sum_{s \in S} p_s = 1 \), a "sampling design" (or simply a "design" or "sampling plan") and denote it by

\[
D = D(S, p) = (S, p)
\]  

Thus the definition of the design gives us a method of selecting a sample which requires the listing down of all the possible samples and choosing one from the list with the corresponding probability. But, in practice especially in large-scale surveys, it is very difficult to list down all possible samples and follow this procedure. Alternatively, Hanurav (1965) defined a "sampling mechanism" of drawing units from \( U \) one-by-one with probabilities which depend on the previous draws. A "sampling mechanism" (or a "drawing mechanism") is a function

\[
q(u, r, s_{r-1})
\]  

(1.3.7)
where \( u \in \mathbb{U} \), \( r \) is a positive integer and \( s_{r-1} \) is a sample of size \( (r-1) \) such that

\[
q(u, r, s_{r-1}) \geq 0 \text{ for all } u, r \text{ and } s_{r-1}
\]

and

\[
\sum_{u \in \mathbb{U}} q(u, r, s_{r-1}) = 1 \text{ for all } r \text{ and } s_{r-1}. \tag{1.3.8}
\]

In this connection the following result is obtained by Hanurav.

**Theorem 1.3.1.** (Hanurav 1962a): There exists one to one correspondance between sampling designs and sampling mechanisms.

Corresponding to any given design \( D(S, P) \) the "inclusion probability" of a unit \( U_i \) is defined as

\[
\pi_i = \sum_{s \supset i} p_s \tag{1.3.9}
\]

where the summation is taken over all the samples that contain \( U_i \) at least once. The "joint inclusion probability" of a pair \((U_i, U_j), i \neq j\), is defined as

\[
\pi_{ij} = \sum_{s \supset i, j} p_s \tag{1.3.10}
\]
the summation being taken over all samples which contain both $U_i$ and $U_j$.

Considering the problem of estimation in finite populations, any function $t$ defined over a design $D(S,p)$ such that for samples $s \in D$ the function $t$ depends only on the values of $\bar{Y}$ for the units belonging to the sample is called a "statistic". A statistic $t$ when used to estimate a parametric function $T(\bar{Y})$ is called an "estimator" of $T$. An estimator $t$ is called an "unbiased estimator" of $T$ iff

$$E(t) = \sum_{s \in S} t_sp_s = T(\bar{Y}), \text{ for all values of } \bar{Y} \quad (1.3.11)$$

An estimator $t$ which is not unbiased for $T$ is said to be "biased".

In the estimation of $T$, the deviation $(t_s - T)$ is taken as the "error" on the basis of sample $s$. Any convex function $f(t_s - T)$ is taken to be the "loss function" and $E(f)$ is called the "expected loss". A loss function which is used quite often is the "mean square error" (m.s.e.) given by

$$M(t) = E(t - T)^2 = \sum_{s \in S} (t_s - T)^2 p_s \quad (1.3.12)$$
where $t_s$ used as an estimator of $T$ based on the sample $s$ and $p_s$ is the probability of selecting the sample $s$. If $t$ is an unbiased estimator of $T$ then $E(t-T)^2$ is called the "variance" of $t$ which may also be written as

$$\text{Var}(t) = \sum_{s \in S} t_s^2 p_s - T^2.$$  \hfill (1.3.13)

A design $D(S,p)$ together with an estimator $t$ of $T$ defined over $D$ is called a "sampling strategy" for the estimation of $T$ and is denoted by

$$H = H(D,t) = H(S,p,t).$$  \hfill (1.3.14)

This definition is due to Hajek (1958). A strategy $H$ when used for estimation of $T$ is called an "unbiased strategy" if $t$ is an unbiased estimator of $T$. Otherwise, it is called a "biased strategy". The expectation and mean square error (variance) of a strategy $H$ are defined as the expectation and mean square error (variance) respectively of the estimator $t$ over $D$.

Of the two estimators $t_1$ and $t_2$ of a parametric function $T(Y)$, both defined over a design $D(S,p)$, $t_1$ is said to be "uniformly better" than $t_2$ if and only if
\[ M(t_1) \leq M(t_2) \quad \text{for all } Y \quad (1.3.15) \]

with a strict inequality at least for one \( Y \). Also in a class \( C \) of estimators of \( T \) defined over a given design \( D(S,p) \), an estimator \( t_1 \in C \) is the "best estimator" if and only if \( t_1 \) is uniformly better than \( t_2 \) for all \( t_2 \) different from \( t_1 \) and belonging to \( C \).

It is well known that in the case of infinite populations the sample mean \( \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i \) is the minimum variance unbiased estimator of \( \mu \) in the class of all linear unbiased estimators \( \frac{1}{n} \sum_{i=1}^{n} w_i Y_i \) (\( w_i \)'s being constants independent of the observations \( Y_i \) and \( \sum_{i=1}^{n} w_i = 1 \)) where \( E(Y_i) = \mu \) for all \( i \). The ''distinguishability of units'' which plays an important role in the theory of finite population is the main difference between the theory of finite and infinite populations. Thus it is important to know whether two equal values of \( Y \) in the sample belong to the same unit repeated, or two different units. This fact was first observed by Des Raj and Khamis (1958) and Basu (1958) who found that in simple random sampling with replacement (SRSWR) of sample size \( n \), the mean of the
"effective sample" i.e. the average of the $y_i$-values corresponding to the distinct units in the sample is better than the usual sample mean $\bar{y}$.

This necessitated a general definition of linear estimator and Horvitz and Thompson (1952) defined following three classes of estimators:

$$t_1 = \sum_{r=1}^{n} \alpha_r y_r$$

where $\alpha_r$ for $r=1,2,\ldots,n$ is the weight associated with the element turned-up at the $r$th draw in the sample. Thus all weights $(\alpha_1,\alpha_2,\ldots,\alpha_n)$ occurring in the estimator are fixed in advance.

$$t_2 = \sum_{i=1}^{N} w_i y_i$$

where $w_i$ is the weight attached to the $i$th unit whenever it is selected in the sample and

$$t_3 = v_s \sum_{i=1}^{n} y_i$$

where estimator has a single coefficient $v_s$ attached to the $s$th sample total.
Subsequently these classes of linear estimators were scrutinized by Koop (1957, 1963), Prabhu–Ajgaonkar (1962, 1965, 1967 a,b,c, 1970 a,b and 1971 a) and they developed seven classes of linear estimators. Prabhu–Ajgaonkar (1962) studied the $T_1$-class in detail and introduced the criteria of empty and non-empty classes.

If for a given sampling system there does not exist an unbiased estimator in the class, that class is termed empty for that sampling scheme.

For a given sampling system a class is designated as non-empty whenever the class contains an unbiased estimator.

Godambe (1955) generalized the above classes of estimators by considering

$$t_s = \sum_{i \in s} \beta_{si} y_i$$  \hspace{1cm} (1.3.14)

where $\beta_{si}$ is a constant used as a weight attached to the $i$th population unit whenever it is included in the $s$th sample and proved the following theorem.

1.3.2. \textbf{Theorem.} Godambe (1955): There does not exist a uniformly minimum variance unbiased estimator for any sampling design $D$. 
Later Hanurav (1966) pointed out some exceptions to this result by constructing some non-trivial designs (called "unicluster designs") where a best estimator exists. This leads to the search for optimum estimators in a sub-class of designs. Even though the criterion of "admissibility" helps in eliminating bad (inadmissible) estimators, it does not help much in obtaining optimum estimators. We say that an estimator $t_1$ of $T$ is admissible if and only if there does not exist another estimator $t_2 \perp t_1$ of $T$ which is better than $t_1$.

An estimator is said to be "inadmissible" if it is not admissible. The results of Murthy (1957), Des Raj and Khamis (1958) and Basu (1958) showed that estimators which depend on the order in which the units appear or which have a unit in the sample repeated are inadmissible. Godambe and Joshi (1965) and Joshi (1965) and Joshi (1965a, 1965b) considered admissibility removing the restriction of linearity and later relaxing the criterion of unbiasedness. Three other criteria, namely "linear invariance" and "regular estimators" by Roy and Chakravarti (1960) and "hyper-admissibility" due to Hanurav (1966) are defined as follows:

An estimator $t$ of $T$ is called "regular estimator" if

$$\text{Var}(t) = K \lambda^2$$

(1.3.15)
where $K$ is a constant and $\hat{\sigma}^2 = \frac{1}{N} \left( \sum_{i=1}^{N} Y_i^2 - NY^2 \right)$.

Roy and Chakravorti have shown that a best estimator exists in the class of regular estimators.

An estimator $t$ of $T$ is said to be possessing the property of "linear invariance" if it is invariant under linear transformation of $\bar{y}$.

In a class $\mathcal{C}$ of unbiased estimators of $T(y)$, an estimator $t_1 \in \mathcal{C}$ is said to be "hyper-admissible" if given any other estimator $t_2 \in \mathcal{C}$ in every hyper plane of $R^N$, there exists at least one point $\bar{y}$ at which

$$\text{Var}(t_1) < \text{Var}(t_2)$$

(1.3.16)

Prabhu-Ajgaonkar (1965), in search of a "servicable estimator" introduced the criteria of "necessary bestness" and proved the necessary bestness of the Horvitz-Thompson (1952) estimator in a subclass of linear unbiased estimator of $Y$.

An estimator $t_1$ is said to be the best estimator for a class $\mathcal{C}$, of unbiased estimators if for every other estimator $t$ belonging to class $\mathcal{C}$

$$\text{Var}(t_1) \leq \text{Var}(t)$$

(1.3.17)
for all \( Y \), with inequality holding for at least one \( Y \).

Let \( \mathcal{Q} \) be the class of linear unbiased estimators, the variance of which can be expressed as

\[
\text{Var}(\mathcal{Q}) = \sum_{i=1}^{N} A_{ii}Y_i^2 + \sum_{i \neq j=1}^{N} A_{ij}Y_iY_j \tag{1.3.18}
\]

where the quantities \( A_{ij} \)'s involve the known functions of probabilities and sample and population sizes.

Further let \( t' \) be an unbiased estimator belonging to the class \( \Omega \) and the variance of \( t' \) be given by

\[
\text{Var}(t') = \sum_{i=1}^{N} B_{ii}Y_i^2 + \sum_{i \neq j=1}^{N} B_{ij}Y_iY_j \tag{1.3.19}
\]

where \( B_{ij} \)'s as before, involve the known functions of probabilities and sample and population sizes.

Consider the quantity

\[
Q = \text{Var}(\mathcal{Q}) - \text{Var}(t')
= \sum_{i=1}^{N} (A_{ii}-B_{ii})Y_i^2 + \sum_{i \neq j=1}^{N} (A_{ij}-B_{ij})Y_iY_j \tag{1.3.20}
\]

The estimator \( t' \) is necessary best estimator of order \( r \) for the class \( \mathcal{Q} \), if all the leading principal minors of \( Q \) up to the order \( r \), are positive.
It must now be stressed that a search for optimum should be considered amongst strategies which are equally costly. We consider a simple linear "cost-function".

\[ G(s) = a_0 + b_0 v(s) \]  \hspace{1cm} (1.3.21)

where \( a_0 \) is the overhead cost, \( b_0 \) is the cost for collecting data on one sample unit and \( v(s) \) is the effective size of sample \( s \). The expected cost of a strategy \( H(S,p,t) \) or equivalently a design \( D(S,p) \) is defined as

\[ C(H) = C(D) = \sum_{s \in S} C(s)p_s = a_0 + b_0 \mu(d) \]  \hspace{1cm} (1.3.22)

where \( \mu(d) = \sum_{s \in S} v(s)p_s \), the expected effective sample size of \( D \). Hence, under this cost set up, two designs (or strategies) are equally costly iff they have the same expected effective sample size.

If optimality is judged from uniform minimization of the variance of a strategy, then we can see that there does not exist such a one. But, whenever some auxiliary information on a characteristic \( X \) which takes value \( X_i \) on unit \( U_i, i=1,2,...,N \) is available closely related to the study variate \( Y \), taking value \( Y_i \) on \( U_i, i=1,2,...,N \), it is possible to use this information in setting up a criterion of optimality.
The information $\mathbf{X}$, known beforehand can be used to assume a reasonable a priori distribution over $\mathbf{Y}$. According to this "super-population concept" as termed by Cochran (1946), $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$ is assumed to be a realisation of a random $N$-vector with certain distribution depending upon $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ and some unknown parameters. This distribution is denoted by $\delta$ and we can talk of expectations, variances and covariances taken over $\delta$. We now minimize the expected variance over $\delta$ namely

$$\int \text{Var}(H) d\delta$$

(1.3.23)

for $H$ varying over $\mathcal{G}(H)$, the class of all equi-cost-strategies. A $H_0$ which minimizes (1.3.23) uniformly with respect to all the parameters of the distribution $\delta$ is called "$\delta$-optimum strategy" in $\mathcal{G}(H)$.

Let $\Delta_\delta$ be the class of all prior distributions $\delta$ satisfying

$$\mathbb{E}(Y_1 | X_1) = aX_1$$

$$\text{Var}(Y_1 | X_1) = b^2 X_1^g, \quad 0 < g < 2 \quad \text{and}$$

$$\text{Cov}(Y_1, Y_j | X_1, X_j) = 0$$

(1.3.24)
where $\mathcal{E}$, $\mathcal{V}$ and $\mathcal{C}$ denote respectively the expectation, variance and covariance, $a$, $s^2$ and $g$ are unknown constants. This $\Delta_g$ class of prior distributions has been widely used in survey sampling.

1.4. Results Obtained.

In addition to this introductory chapter the present thesis contains eight chapters.

Various PPS sampling strategies when Rao, Hartley and Cochran's (1962) estimator is employed are described in Chapter 2 and are compared under the general superpopulation model. It has been shown that the generalized $\pi$PS sampling strategy, consisting of the design with $\pi_i$, the probability of inclusion of the $i$th population unit in the sample, proportional to the modified size and the Horvitz-Thompson estimator is superior to the generalized PPS sampling strategy consisting of the design with $p_i$, the selection probability of the $i$th population unit, proportional to the modified size and the Rao, Hartley and Cochran's estimator under general super population model. It is also demonstrated that the generalized $\pi$PS sampling strategy is superior to the Rao, Hartley and Cochran's sampling strategy under a general super population model. In the subsequent section an attempt has been made to compare the efficiencies of Horvitz-Thompson-Ikeda-Sen's sampling strategy (Horvitz-Thompson's estimator and Ikeda-Sen's sampling scheme), with Rao, Hartley and Cochran's
sampling strategy. It is demonstrated that under certain conditions Rao, Hartley and Cochran’s sampling strategy is better than Horvitz-Thompson-Ikeda-Sen’s sampling strategy.

In Chapter 3 the three unbiased estimators belonging to the three classes of linear estimators formulated by Horvitz and Thompson (1952) for Ikeda-Sen’s sampling procedure are presented. Using the Bayesian estimation theory an attempt is made to select an estimator with the help of the criterion of necessary bestness. The optimum sampling procedures corresponding to the three optimum estimators in each class under super population concept are determined, which give rise to three sampling strategies. It is noted that there does not exist a necessary best sampling strategy under the super-population concept for Ikeda-Sen’s sampling procedure.

Chapter 4 deals with the $T_1$-class of linear estimators. Prabhu-Ajgaonkar (1962) studied the $T_1$-class of linear estimators in detail and introduced the criteria of empty and non-empty classes. In this chapter we present a sampling procedure consisting of selecting first $m$ elements with varying probabilities with replacement, the $m+1$ th element is selected with varying probabilities without replacement and on the subsequent draws the $n-m-1$ elements are selected with equal probability without replacement. The estimators engendered by $T_1$-class of linear estimators are considered. It is demonstrated
that for this sampling procedure the $T_1$ class is non-empty and interestingly enough that there exists independent minimum-variance-linear unbiased estimator.

Problem of optimum probability set for Ikeda-Sen's sampling strategy is considered in Chapter 5. Midzuno (1950) suggested an estimator which belongs to Horvitz and Thompson's (1952) $T_3$-class of linear estimators. This $T_3$-class of linear estimators has been examined by Prabhu-Ajgaonkar (1967a) who noted that a best estimator does not exist for the $T_3$-class. Godambe (1966) considering the new approach to sampling from finite populations, demonstrated that for any fixed sample size design, the only unbiased estimator of the population total which is linearly sufficient and which satisfies the principal of censoring belongs to the $T_3$-class. An attempt has been made to determine an optimum probability set for Ikeda-Sen's sampling procedure when Midzuno's estimator belonging to the $T_3$-class is employed. For this purpose the Bayesian technique is used.

Problem of Bayes estimation is considered in Chapter 6. Prabhu-Ajgaonkar (1965) derived the best estimator for the $T_1$-class when Ikeda-Sen's sampling procedure is employed. He (1971b) demonstrated that the variance function of a linear estimator can be expressed into a quadratic form and accordingly
he noted that there exists nine principal classes of estimators out of which one principal class he examined in detail. He obtained expression for a unique estimator variance of the best estimator in the $T_1$-class for the Ikeda-Sen's sampling procedure. We have derived a lower bound for the Bayes risk of an unbiased variance estimator for the variance of the best estimator in the $T_1$-class.

In Chapter 7 it is shown that an admissible general unbiased quadratic estimator for the population variance is hyper-admissible and general unbiased estimator of the variance of the hyper-admissible estimator is hyper-admissible.

Chapter 8 deals with the problem of optimum allocation of sample size to the strata under the finite population model furnished by Avadhani and Sukhatme (1970).

Chapter 9 deals with the scope for further work.

A list of references used in this thesis is given at the end. The concepts of Chapter 2 and Chapter 3 are published (Metrika, 25, 1978 and Jour. Ind. Stat. Assoc. 16, 1978). The results of Chapter 5 were presented at the 29th Annual Conference of the Indian Society of Agricultural Statistics in the year 1975 (Jour. Ind. Soc. Agri. Stat. 27, 1975). The contents of Chapter 6 are accepted for publication (Jour. Ind. Stat. Assoc.). The contents of Chapter 7 and 8 have been submitted for publication.