Abstract

The theory of sampling has its origin way back in the history of mankind. In general, people are interested to study totality, called population, to decide about its nature. For such purpose, it is necessary to collect information regarding the population with respect to some characteristics. e.g. in agricultural surveys, to estimate the production of food, the data are collected on some portion of land under different crops. Most of government and non-government bodies collect information regularly about the total population, its distribution by area, sex, age etc. for future planning. In business, information is also required regarding the role and character of wholesale, retail and service trades etc. Such information is collected either by complete enumeration or by sample survey.

The sample survey method is the most important tool of collection of such information because of its efficiency, accuracy, speed and some constraints over the others. The main problem in survey sampling is to develop an appropriate procedure for selecting sample from the population containing required information about the population under the given constraints and constructing a formulation based on the sample selected for estimating the population parameters of interest.

The thesis has been divided into six chapters.

Chapter 1 gives the general introduction and review of literature relating to my research topic ‘Construction of Efficient Sampling Strategies in Survey Sampling’ for estimating the population parameters.

In chapter 2, we propose a generalized ratio and product type estimator of population mean under stratified random sampling using known information on parameters $X_h$, $\theta_h$ and $\phi_h$ of auxiliary variable $x$ based upon each stratum. It has been shown that the estimators $\bar{y}_{stSK}$, $\bar{y}_{stSD}$, $\bar{y}_{stUS1}$ and $\bar{y}_{stUS2}$ defined by Kadilar & Cingi (2003) are particular cases of proposed estimator. The expressions for mean square errors of the proposed estimator and Kadilar and Cingi (2005) estimator $\bar{y}_{stp}$ and their minimum mean square errors have been obtained. It has been shown theoretically that the optimum estimator of the proposed estimator is always better than the optimum estimator of $\bar{y}_{stp}$. It has also been shown numerically that proposed estimator
is always efficient than the estimator $\bar{y}_{xUS1}$, $\bar{y}_{xUS2}$, $\bar{y}_{xSK}$, $\bar{y}_{xSD}$ and $\bar{y}_{xsp}$ when the constant $\theta_h$ is taken as regression coefficient of $y$ on $x$ in the $h^{th}$ stratum. The results have also been illustrated graphically.

In chapter 3, we propose a generalized difference-cum-ratio type estimator based on two unknown constants $\alpha$ and $\theta$ for estimation of population mean using double sampling technique. The expressions for bias and mean square error of the proposed estimator have been obtained. The MSE of proposed estimator has been minimized w.r.t. one of the constant $\alpha$ while keeping the value of second constant $\theta$ fixed. The comparison of the proposed estimator has been made with the mean per unit estimator and regression estimator w.r.t. their MSE’s theoretically and numerically. By comparing the MSE’s, we find that proposed estimator is more efficient than the regression estimator under the same sampling design for $0 < \theta < 2$ and it is also more efficient than the mean per unit estimator for $1 - \frac{1}{\sqrt{1-\rho^2}} < \theta < 1 + \frac{1}{\sqrt{1-\rho^2}}$ at moderate value of correlation coefficient $\rho$. The results have also been verified by generating data using simulation study which also show the same efficiency.

Following chapter 3, we have proposed a generalized difference-cum-ratio type estimator based on two unknown constants $\alpha$ and $\theta$ for the estimation of population variance under double sampling design in chapter 4. The expressions for bias and mean square error of the proposed estimator have been obtained up to first order of approximation. While keeping the value of the one of constant $\theta$ fixed, the MSE of the proposed estimator has been minimized w.r.t. other constant. The comparison of the proposed estimator has been made with the conventional sample variance and regression type estimators w.r.t. their MSE’s. By comparing the MSE’s, we find that proposed estimator is more efficient than the regression type estimator under the same sampling design for $0 < \theta < 2$ and it is also more efficient than the sample variance for $1 - \frac{1}{\sqrt{1-\rho^2}} < \theta < 1 + \frac{1}{\sqrt{1-\rho^2}}$ at moderate value of correlation coefficient $\rho$. The results have also been illustrated numerically and graphically.

In chapter 5, we have proposed a generalized class of estimators of population variance of stratified sample mean using known information of some parameters based upon either auxiliary variable $x$ only or both study variable $y$ and auxiliary variable $x$. The expressions for
mean square error of the proposed estimator and its minimum value have been obtained. It has been shown that the optimum estimator of the proposed estimator is always better than the usual variance estimator and customary stratified ratio estimator of variance theoretically. By taking some data from literature, it has also been shown numerically that the proposed estimator is almost six times efficient than the usual variance estimator and two times efficient than the customary stratified ratio estimator of variance. Efforts have also been made to illustrate the results graphically.

In a situation when population is heterogeneous, we have divided the population w.r.t two characteristics which give rise to two-way stratification of data. Then using general sampling design for selection of samples from each sub-strata, sampling strategies for estimating the population mean have been proposed in chapter 6. The expressions for the biases and mean square errors of the proposed sampling strategies have been obtained. The comparison of the proposed sampling strategies with the existing ones have been made theoretically and the conditions under which the proposed sampling strategies are better than the existing ones have been obtained. For comparing the results numerically, we have considered simple random sampling design for selection of samples. So, by considering hypothetical population, it has been shown that the efficiency of product estimator is almost double under proposed sampling design where as the increase in efficiency of separate ratio estimator is small as there is almost negative correlation between the study variable y and auxiliary variable x in the considered hypothetical population. Efforts have also been made to illustrate the results graphically.