World Applied Sciences Journal 33 (6): 874-882, 2015
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2015.33.06.146

Estimating the Population Mean in Stratified Random Sampling Using Two-Phase Sampling in the Presence of Non-Response

Manoj K. Chaudhary and Amit Kumar

Department of Statistics, Banaras Hindu University, Varanasi, India

Abstract: The present paper presents the salient feature on estimating the population mean in stratified random sampling using two-phase sampling scheme in the presence of non-response. We have used an auxiliary variable to estimate the population mean assuming that the non-response is observed on only study variable. A general family of estimators of population mean has been introduced under the situation in which the population mean of auxiliary variable is not known. The properties of the suggested family have been conferred and the optimum estimator has been determined. A study of cost survey has been carried out. The theoretical results have also been comprehended with the support of empirical study.

Key words: Two-phase sampling scheme · Stratified random sampling · Auxiliary variable · Population mean and non-response

INTRODUCTION

In a sample survey, it is expected that the information would be collected from all the selected units in the sample, but practically, it is generally not possible because of the non-response. Some of the units may not respond or these may not be contacted during the survey period. The non-response introduces an increment in sampling variance of estimates as the effective sample size is reduced from that originally required and it introduces to bias of the estimates also when the non-respondents differ from respondents in the characteristics observed. In the very beginning, Hansen and Hurwitz [1] have tackled the problem of non-response while conducting the mail surveys and introduced a technique of sub-sampling of non-respondents in order to estimate the population mean in the presence of non-response.

The stratified random sampling is used to estimate the parameters if the population under consideration consists of heterogeneous units. Here, first the population is divided into different homogeneous groups (strata). Secondly, a sample of specified size is selected from each group by simple random sampling. If the non-response is inherent in all strata of the population, one can utilize the Hansen and Hurwitz [1] technique of sub-sampling of non-respondents to estimate the population mean. Khare [2] has proposed the estimator of population mean under optimum allocation in the presence of non-response. Khan *et al.* [3] have discussed the method of optimum allocation while conducting a mail survey in multivariate stratified random sampling under non-response. Chaudhary *et al*. [4] have suggested a family of estimators of population mean in stratified random sampling whenever the non-response is observed on study variable. Chaudhary *et al*. [5] have proposed some new allocation schemes based on response or/and non-response rates in stratified random sampling and compared them with proportion and Neyman allocation schemes.

There is an important role of auxiliary information in estimating the parameters of a population. If the information about the auxiliary variable(s) in the population is available then it can be easily utilized in order to improve the precision of the estimator of a parameter. But, the use of auxiliary information is doubtful if the population parameters of auxiliary variable(s) are not known. In such situations, the double sampling (two-phase sampling) scheme may be utilized to remove the doubt. Double sampling scheme is a very effective technique in terms of cost and its applications. Two-phase sampling is a procedure in which we obtain the information about auxiliary variable(s) reasonably from a larger sample at first phase and relatively small sample at the second phase. Khare and Sinha [6] have proposed

Corresponding Author: Manoj K. Chaudhary, Department of Statistics, Banaras Hindu University, Varanasi, India.

some estimators for estimating the population ratio adopting the two-phase sampling in the presence of non-response. Singh and Kumar [7] have suggested a general class of estimators of population mean using two-phase sampling under non-response.

In the above circumstances, we have suggested a general class of estimators for estimating the population mean in stratified random sampling using two-phase sampling if the population mean of auxiliary variable is not known and non-response is observed on study variable. The suggested class of estimators has been discussed along with its optimum property. A comparative study of the optimum estimator with the usual estimators has also been carried out through numerical data.

Let us suppose that a population consists of N units, is divided into k strata. Let the size of the ith stratum be $N_i (i = 1,2,...,k)$. A sample of size n is selected from the entire population assuming that N_i units should be selected from the ith stratum. Thus, we have $\sum_{i=1}^{k} N_i = N$ and $\sum_{i=1}^{k} n_i = n$. Let Y and X be the study and auxiliary variables respectively with their respective population means $\bar{Y} = \sum_{i=1}^{k} p_i \bar{Y}_i$ and $\bar{X} = \sum_{i=1}^{k} p_i \bar{X}_i$ ($\bar{Y}_i$ and $\bar{X}_i$ being the means based on N_i units for study and auxiliary variables respectively and $p_i = N_i / N$). Let us assume that the non-response is observed on study variable and auxiliary variable is free from non-response. Under the above assumptions, it is noted that out of n_i units, n_{i1} units respond and n_{i2} units do not respond on Y. Using Hansen and Hurwitz [1] technique of sub-sampling of non-respondents, a sub-sample of h_{i2} units is selected from n_{i2} units of non-respondents such that $n_{i2} = L_i h_{i2}$, $L_i \geq 1$ and the information are gathered from all h_{i2} units. Thus, the Hansen and Hurwitz [1] estimator of population mean $\bar{Y}$ in stratified random sampling under non-response without using auxiliary information is given by

$$\bar{y}_{st}^{*} = \sum_{i=1}^{k} p_i \bar{y}_i^{*} \tag{1.1}$$

where $\bar{y}_i^{*} = \frac{n_{i1}\bar{y}_{ni1} + n_{i2}\bar{y}_{hi2}}{n_i}$, $\bar{y}_{ni1}$ and $\bar{y}_{hi2}$ are the means based on n_{i1} respondent units and h_{i2} sub-sampled non-respondent units respectively for study variable. The variance of $\bar{y}_{st}^{*}$ is given as

$$V\left(\bar{y}_{st}^{*}\right) = \sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Yi}^2 + \sum_{i=1}^{k}\frac{(L_i - 1)}{n_i} W_{i2} p_i^2 S_{Yi2}^2 \tag{1.2}$$

where S_{Yi}^2 and S_{Yi2}^2 are the population mean squares of the entire group and non-response group respectively in the ith stratum for study variable. W_{i2} is the non-response rate in the ith stratum.

Chaudhary *et al.* [4] have suggested a family of combined-type estimators of population mean $\bar{Y}$ in stratified random sampling using the information of an auxiliary variable under non-response considering Khoshnevisan *et al.* [8] as

$$T_C = \bar{y}_{st}^{*}\left[\frac{a\bar{X} + b}{\alpha\left(a\bar{x}_{st} + b\right) + (1-\alpha)\left(a\bar{X} + b\right)}\right]^{g} \tag{1.3}$$

where $\bar{x}_{st} = \sum_{i=1}^{k} p_i \bar{x}_i$, $\bar{x}_i$ is the mean based on n_i units for the auxiliary variable. $a \neq 0$ and b are either real numbers or functions of known parameters of auxiliary variable. α and g are the constants and to be determined. The bias and mean square error (MSE) of T_C up to the first order of approximation are respectively given by

$$B(T_C) = \frac{1}{\bar{Y}} \sum_{i=1}^{k} f_i p_i^2 \left[\frac{g(g+1)}{2}\alpha^2\lambda^2 R^2 S_{Xi}^2 - \alpha\lambda g R \rho_i S_{Xi} S_{Yi}^2\right] \tag{1.4}$$

and

$$MSE(T_C) = \sum_{i=1}^{k} f_i p_i^2 \left[S_{Yi}^2 + \alpha^2\lambda^2 g^2 R^2 S_{Xi}^2 - 2\alpha\lambda g R \rho_i S_{Xi} S_{Yi}\right] + \sum_{i=1}^{k}\frac{(L_i - 1)}{n_i} W_{i2} p_i^2 S_{Yi2}^2 \tag{1.5}$$

where $f_i = \left(\frac{1}{n_i} - \frac{1}{N_i}\right)$, $\lambda = \frac{a\bar{X}}{a\bar{X}+b}$, $R = \frac{\bar{Y}}{\bar{X}}$, S_{Xi}^2 is the population mean squire of auxiliary variable in the ith stratum. ρ^i is the population correlation coefficient between Y and X in the ith stratum.

Proposed Family of Estimators: The two-phase sampling scheme may be utilized in estimating the population mean $\bar{Y}$ of study variable if the population mean $\bar{X}$ of auxiliary variable is not known. Under this sampling scheme, first an estimate of the population mean $\bar{X}$ can be obtained from a larger first phase sample of size n'_i selected from N_i units by simple random sampling without replacement (SRSWOR). Secondly, a smaller sample of size n_i is selected from n'_i by SRSWOR. Thus, we have $\sum_{i=1}^{k} n_i' = n'$ and $\sum_{i=1}^{k} n_i = n$. At the second phase sample, it is observed that out of n_i units, n_{i1} units respond and n_{i2} units do not respond on Y. Using Hansen and Hurwitz [1] technique of sub-sampling of non-respondents, we select a sub-sample of $h_{i2}\left(=n_{i2}/L_i, L_i \geq 1\right)$ units from the n_{i2} non-respondent units by SRSWOR and collect the information from all the h_{i2} units.

Assuming that the complete information are observed on auxiliary variable X at the first phase sample as well as second phase sample, the usual combined ratio and usual combined product estimators of population mean $\bar{Y}$ using two-phase sampling in stratified random sampling under non-response are respectively given by

$$T_1^* = \frac{\bar{y}_{st}^*}{\bar{x}_{st}^*}\bar{x}_{st}' \tag{2.1}$$

and

$$T_2^* = \frac{\bar{y}_{st}^*}{\bar{x}_{st}'}\bar{x}_{st}^* \tag{2.2}$$

where $\bar{x}_{st}^* = \sum_{i=1}^{k} p_i \bar{x}_i^*$, $\bar{x}_{st}' = \sum_{i=1}^{k} p_i \bar{x}_i'$, $\bar{x}_i^* = \frac{n_{i1}\bar{x}_{ni1} + n_{i2}\bar{x}_{hi2}}{n_i}$, $\bar{x}_{ni1}$ and $\bar{x}_{hi2}$ are the means based on n_{i1} units and h_{i2} units respectively for auxiliary variable. $\bar{x}_i'$ is the mean based on n'_i units of first phase sample for auxiliary variable.

The mean square errors of T_1^* and T_2^* are respectively represented as

$$MSE\left(T_1^*\right) = \sum_{i=1}^{k} f_i' p_i^2 S_{Yi}^2 + \sum_{i=1}^{k} f_i^* p_i^2 \left(S_{Yi}^2 + R^2 S_{Xi}^2 - 2R\rho_i S_{Xi} S_{Yi}\right)$$
$$+\sum_{i=1}^{k} \frac{\left(L_i - 1\right)}{n_i} W_{i2} p_i^2 \left(S_{Yi2}^2 + R^2 S_{Xi2}^2 - 2R\rho_{XYi2} S_{Xi2} S_{Yi2}\right) \tag{2.3}$$

and

$$MSE\left(T_2^*\right) = \sum_{i=1}^{k} f_i' p_i^2 S_{Yi}^2 + \sum_{i=1}^{k} f_i^* p_i^2 \left(S_{Yi}^2 + R^2 S_{Xi}^2 + 2R\rho_i S_{Xi} S_{Yi}\right)$$
$$+\sum_{i=1}^{k} \frac{\left(L_i - 1\right)}{n_i} W_{i2} p_i^2 \left(S_{Yi2}^2 + R^2 S_{Xi2}^2 + 2R\rho_{XYi2} S_{Xi2} S_{Yi2}\right) \tag{2.4}$$

where $f_i^{'} = \left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)$, $f_i^{*} = \left(\frac{1}{n_i} - \frac{1}{n_i^{'}}\right)$, S_{Xi2}^2 is the population mean square of the non-response group for auxiliary variable in the ith stratum and ρ_{XYi2} is the population correlation coefficient between Y and X for the non-response group in the ith stratum.

If the population mean $\bar{X}$ of auxiliary variable is known, one can easily use the family of combined-type estimators given in equation (1.3) in estimating the population mean $\bar{Y}$ of the study variable. But the situations in which the population mean $\bar{X}$ is unknown, it is very difficult to consider the present form of the given family of estimators and it is advisable to adopt the two-phase sampling scheme in order to estimate the population mean $\bar{Y}$. Thus, the family of combined-type estimators of population mean $\bar{Y}$ in stratified random sampling using two-phase sampling under non-response is given by

$$T_C^* = \bar{y}_{st}^* \left[\frac{a\bar{x}_{st}^{'} + b}{\alpha\left(a\bar{x}_{st}^{*} + b\right) + (1-\alpha)\left(a\bar{x}_{st}^{'} + b\right)} \right]^g \tag{2.5}$$

Using large sample approximation, we can obtain the bias and MSE of T_C^*. Let us assume that

$$\bar{y}_{st}^* = \bar{Y}(1+e_0),\ \bar{x}_{st}^* = \bar{X}(1+e_1),\ \bar{x}_{st}^{'} = \bar{X}\left(1+e_1^{'}\right)$$

such that $E(e_0)=E(e_1)=E(e'_2)=0$,

$$E\left(e_0^2\right) = \frac{1}{\bar{Y}^2}\sum_{i=1}^{k}\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{Yi}^2 + \frac{(L_i-1)}{n_i}W_{i2}p_i^2 S_{Yi2}^2\right],$$

$$E\left(e_1^2\right) = \frac{1}{\bar{X}^2}\sum_{i=1}^{k}\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{Xi}^2 + \frac{(L_i-1)}{n_i}W_{i2}p_i^2 S_{Xi2}^2\right],\ E\left(e_1^{'2}\right) = \frac{1}{\bar{X}^2}\sum_{i=1}^{k}\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 S_{Xi}^2,$$

$$E\left(e_0 e_1\right) = \frac{1}{\bar{X}\bar{Y}}\sum_{i=1}^{k}\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{Xi} S_{Yi} + \frac{(L_i-1)}{n_i}W_{i2}p_i^2 \rho_{XYi2} S_{Xi2} S_{Yi2}\right],$$

$$E\left(e_0 e_1^{'}\right) = \frac{1}{\bar{X}\bar{Y}}\sum_{i=1}^{k}\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{Xi} S_{Yi} \text{ and } E\left(e_1 e_1^{'}\right) = \frac{1}{\bar{X}^2}\sum_{i=1}^{k}\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 S_{Xi}^2.$$

We now express T_C^* in terms of e_0, e_1 and e'_1. The resulting expression is given as

$$T_C^* = \bar{Y}(1+e_0)\left(1+\lambda e_1^{'}\right)^g \left[1 + \lambda\left\{\alpha e_1 + (1-\alpha)e_1^{'}\right\}\right]^{-g} \tag{2.6}$$

Expanding the above equation and neglecting the terms of e_0, e_1 and e'_1 having power greater than two, we get

$$T_C^* - \bar{Y} = \bar{Y}\left[-g\lambda\left\{\alpha e_1 + (1-\alpha)e_1^{'}\right\} + \frac{g(g+1)}{2}\lambda^2\left\{\alpha^2 e_1^2 + (1-\alpha)^2 e_1^{'2} + 2\alpha(1-\alpha)e_1 e_1^{'}\right\} + \right.$$

$$g\lambda e_1' - g^2\lambda^2 e_1'\left\{\alpha e_1 + (1-\alpha)e_1'\right\} + \frac{g(g-1)}{2}\lambda^2 e_1'^2 + e_0 - g\lambda e_0\left\{\alpha e_1 + (1-\alpha)e_1'\right\} + g\lambda e_1' e_0\Bigg] \tag{2.7}$$

Taking expectation both the sides of equation (2.7), we get the bias of T_C^* up to the first order of approximation as

$$B\left(T_C^*\right) = \frac{1}{\bar{Y}} g\lambda \sum_{i=1}^{k} p_i^2 \Bigg[f_i \left\{ \frac{(g+1)}{2}\lambda\alpha^2 R^2 S_{Xi}^2 - 2R\rho_i S_{Xi} S_{Yi} \right\} +$$

$$f_i' \left\{ \left(\frac{(g+1)}{2}\lambda(1-\alpha)^2 R^2 - g\lambda(1-\alpha)R^2 + (g+1)\lambda\alpha(1-\alpha)R^2 - g\lambda\alpha R^2 + \frac{(g-1)}{2}\lambda R^2 \right) S_{Xi}^2 + R\alpha\rho_i S_{Xi} S_{Yi} \right\}$$

$$+ \frac{(L_i - 1)}{n_i} W_{i2} \left(\frac{(g+1)}{2}\lambda\alpha^2 R^2 S_{Xi2}^2 - R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2} \right) \Bigg] \tag{2.8}$$

Squaring both the sides of equation (2.7) and then taking expectation by neglecting the terms of e_0, e_1 and e'_1 having power greater than two, we get the MSE of T_C^* up to the first order of approximation as

$$MSE\left(T_C^*\right) = \sum_{i=1}^{k} f_i' p_i^2 S_{Yi}^2 + \sum_{i=1}^{k} f_i^* p_i^2 \left(S_{Yi}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi}^2 - 2g\lambda R\alpha\rho_i S_{Xi} S_{Yi} \right)$$

$$+ \sum_{i=1}^{k} \frac{(L_i - 1)}{n_i} W_{i2} p_i^2 \left(S_{Yi2}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2} \right) \tag{2.9}$$

Optimum Choice of *α*: To choose the optimum value of α, we differentiate $MSE(T_C^*)$ with respect to α and equating the derivative to zero

$$\frac{\partial MSE\left(T_C^*\right)}{\partial\alpha} = 2g^2\lambda^2 R^2 \alpha \sum_{i=1}^{k} f_i^* p_i^2 S_{Xi}^2 - 2g\lambda R \sum_{i}^{k} f_i^* p_i^2 \rho_i S_{Xi} S_{Yi} +$$

$$2g^2\lambda^2 R^2 \alpha \sum_{i=1}^{k} \frac{(L_i - 1)}{n_i} W_{i2} p_i^2 S_{Xi2}^2 - 2g\lambda R \sum_{i}^{k} \frac{(L_i - 1)}{n_i} W_{i2} p_i^2 \rho_{XYi2} S_{Xi2} S_{Yi2} = 0 \tag{2.10}$$

$$\Rightarrow \qquad \alpha_{opt} = \frac{\sum_{i}^{k} f_i^* p_i^2 \rho_i S_{Xi} S_{Yi} + \sum_{i}^{k} \frac{(L_i - 1)}{n_i} W_{i2} p_i^2 \rho_{XYi2} S_{Xi2} S_{Yi2}}{g\lambda R \left[\sum_{i=1}^{k} f_i^* p_i^2 S_{Xi}^2 + \sum_{i=1}^{k} \frac{(L_i - 1)}{n_i} W_{i2} p_i^2 S_{Xi2}^2 \right]} \tag{2.11}$$

For α_{opt} given in equation (2.11), the MSE of T_C^* would attain its minimum.

Cost of the Survey and Optimum *n_i*, *n'_i*, *L_i*: Let c'_i be the unit cost connected with the first phase sample of size n'_i and c_{i0} be the unit cost of first attempt on study variable with second phase sample of size n_i. Let c_{i1} and c_{i2} be respectively the unit cost associated with enumerating the n_{i1} respondent units and h_{i2} non-respondent units at the second phase. Thus, the total cost for the ith stratum is given by

$$C_i = c_i^{'} n_i^{'} + c_{i0} n_i + c_{i1} n_{i1} + c_{i2} h_{i2} \quad \forall\ i = 1,2,...,k$$

The expected average cost per stratum is represented as

$$E\left(C_i\right) = c_i^{'} n_i^{'} + n_i \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i} \right).$$

Thus, the total cost over all the strata is given by

$$C_0 = \sum_{i=1}^{k} E\left(C_i\right) = \sum_{i=1}^{k} \left[c_i^{'} n_i^{'} + n_i \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i} \right) \right]. \tag{2.12}$$

Let us define the Lagrange function

$$\phi = MSE\left(T_C^*\right) + \mu C_0 \tag{2.13}$$

where μ is Lagrange's multiplier.

Differentiating equation (2.13) with respect to n_i, n'_i and L_i respectively and equating the derivatives to zero, we get

$$\frac{\partial \phi}{\partial n_i} = -\frac{p_i^2}{n_i^2}\left(S_{Yi}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi}^2 - 2g\lambda R\alpha\rho_i S_{Xi} S_{Yi}\right) - \frac{\left(L_i - 1\right)}{n_i^2} W_{i2} p_i^2 \left(S_{Yi2}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2}\right) + \mu\left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i} \right) = 0 \tag{2.14}$$

$$\frac{\partial \phi}{\partial n_i^{'}} = \frac{p_i^2}{n_i^{'2}}\left(g^2\lambda^2 R^2 \alpha^2 S_{Xi}^2 - 2g\lambda R\alpha\rho_i S_{Xi} S_{Yi}\right) + \mu c_i^{'} = 0 \tag{2.15}$$

and

$$\frac{\partial \phi}{\partial L_i} = \frac{p_i^2}{n_i} W_{i2}\left(S_{Yi2}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2}\right) - \mu n_i c_{i2} \frac{W_{i2}}{L_i^2} = 0 \tag{2.16}$$

From equations (2.14), (2.15) and (2.16), we respectively get

$$n_i = \frac{p_i \sqrt{S_{Yi}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi}^2 - 2g\lambda R\alpha\rho_i S_{Xi} S_{Yi} + \left(L_i - 1\right) W_{i2}\left(S_{Yi2}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2}\right)}}{\sqrt{\mu\left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i} \right)}} \tag{2.17}$$

$$n_i^{'} = \frac{p_i \sqrt{2g\lambda R\alpha\rho_i S_{Xi} S_{Yi} - g^2\lambda^2 R^2 \alpha^2 S_{Xi}^2}}{\sqrt{\mu c_i^{'}}} \tag{2.18}$$

$$\sqrt{\mu} = \frac{p_i L_i \sqrt{S_{Yi2}^2 + g^2\lambda^2 R^2 \alpha^2 S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2} S_{Xi2} S_{Yi2}}}{n_i \sqrt{c_{i2}}} \tag{2.19}$$

Substituting the value of $\sqrt{\mu}$ from equation (2.19) into equation (2.17), we get

$$L_{i(opt)} = \frac{\sqrt{c_{i2}}B_i}{D_i A_i} \tag{2.20}$$

where $A_i = \sqrt{c_{i0} + c_{i1}W_{i1}}$,

$$B_i = \sqrt{S_{Yi}^2 + g^2\lambda^2R^2\alpha^2S_{Xi}^2 - 2g\lambda R\alpha\rho_i S_{Xi}S_{Yi} - W_{i2}\left(S_{Yi2}^2 + g^2\lambda^2R^2\alpha^2S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2}S_{Xi2}S_{Yi2}\right)}$$

and

$$D_i = \sqrt{S_{Yi2}^2 + g^2\lambda^2R^2\alpha^2S_{Xi2}^2 - 2g\lambda R\alpha\rho_{XYi2}S_{Xi2}S_{Yi2}}$$

Putting the value of $L_{i(opt)}$ from equation (2.20) into equation (2.17), n_i can be expressed as

$$n_i = \frac{p_i\sqrt{B_i^2 + \frac{\sqrt{c_{i2}}B_iW_{i2}D_i}{A_i}}}{\sqrt{\mu}\sqrt{A_i^2 + \frac{\sqrt{c_{i2}}A_iW_{i2}D_i}{B_i}}} \tag{2.21}$$

Substituting the values of n'_i, $L_{i(opt)}$ and n_i respectively from equations (2.18), (2.20) and (2.21) into equation (2.12), we get the value of $\sqrt{\mu}$ in terms of total cost C_0

$$\sqrt{\mu} = \frac{1}{C_0}\sum_{i=1}^{k}\left[p_i\sqrt{c_i'}\sqrt{2g\lambda R\alpha\rho_i S_{Xi}S_{Yi} - g^2\lambda^2R^2\alpha^2S_{Xi}^2} + p_i\left(A_iB_i + \sqrt{c_{i2}}W_{i2}D_i\right)\right] \tag{2.22}$$

Putting the value of $\sqrt{\mu}$ from equation (2.22) into equations (2.21) and (2.18), we respectively get the optimum values of n_i and n'_i

$$n_{i(opt)} = \frac{C_0p_i\sqrt{B_i^2 + \frac{\sqrt{c_{i2}}B_iW_{i2}D_i}{A_i}}}{\sqrt{A_i^2 + \frac{\sqrt{c_{i2}}A_iW_{i2}D_i}{B_i}}\sum_{i=1}^{k}\left[p_i\sqrt{c_i'}\sqrt{2g\lambda R\alpha\rho_i S_{Xi}S_{Yi} - g^2\lambda^2R^2\alpha^2S_{Xi}^2} + p_i\left(A_iB_i + \sqrt{c_{i2}}W_{i2}D_i\right)\right]} \tag{2.23}$$

$$n'_{i(opt)} = \frac{C_0p_i\sqrt{2g\lambda R\alpha\rho_i S_{Xi}S_{Yi} - g^2\lambda^2R^2\alpha^2S_{Xi}^2}}{\sqrt{c_i'}\sum_{i=1}^{k}\left[p_i\sqrt{c_i'}\sqrt{2g\lambda R\alpha\rho_i S_{Xi}S_{Yi} - g^2\lambda^2R^2\alpha^2S_{Xi}^2} + p_i\left(A_iB_i + \sqrt{c_{i2}}W_{i2}D_i\right)\right]} \tag{2.24}$$

Table 1: Particulars of Parameters

St. No. (i)	N_i	n'_i	n_i	$\bar{Y}_i$	$\bar{X}_i$	S^2_{Yi}	S^2_{Xi}	ρ_i	S^2_{Yi2}	S^2_{Xi2}	ρ_{XYi2}
1	73	65	26	40.85	39.56	6369.10	6624.44	0.999	5095.28	5299.55	0.799
2	70	25	10	27.83	27.57	1051.07	1147.01	0.998	840.86	917.61	0.798
3	97	48	19	25.79	25.44	2014.97	2205.40	0.999	1611.97	1764.32	0.799
4	44	11	5	20.64	20.36	538.47	485.27	0.997	430.78	388.21	0.797

Table 2: MSE and PRE of T_1^*, T_2^* and T_C^* with respect to $\bar{y}_{st}^*$

W_{i2}	L_i	$V\left(\bar{y}_{st}^*\right)$	$MSE\left(T_1^*\right)$	$MSE\left(T_2^*\right)$	$MSE\left(T_C^*\right)$	$PRE\left(T_1^*\right)$	$PRE\left(T_2^*\right)$	$PRE\left(T_C^*\right)$
0.1	2	30.53	24.77	391.63	6.98	125.71	7.95	446.13
	2.5	31.74	25.74	400.25	7.58	126.81	8.15	430.61
	3	32.93	26.71	408.87	8.19	127.81	8.35	416.85
	3.5	34.14	27.68	417.49	8.8	128.76	8.54	405.00
0.2	2	32.94	26.71	408.87	8.19	127.82	8.35	416.85
	2.5	35.34	28.65	426.12	9.4	129.63	8.71	395.11
	3	37.74	30.59	443.36	10.6	131.22	9.05	378.68
	3.5	40.14	32.53	460.66	11.8	132.62	9.36	365.60
0.3	2	35.34	28.65	426.12	9.4	129.63	8.71	395.11
	2.5	38.94	31.56	451.98	11.2	131.94	9.21	371.80
	3	42.54	34.47	477.85	13.01	133.88	9.66	354.73
	3.5	46.15	37.38	503.72	14.82	135.50	10.05	341.77
0.4	2	37.74	30.59	443.61	10.6	131.22	9.04	378.68
	2.5	42.54	34.47	477.85	13.01	133.85	9.65	354.65
	3	47.35	38.35	512.34	15.42	135.98	10.18	338.20
	3.5	52.15	42.22	546.83	17.83	137.73	10.63	326.13

Empirical Study: In order to support the theoretical results, we have considered the data used by Chaudhary *et al*. [9]. The data relate to 284 municipalities in Sweden, varying considerably in size and other characteristics. The municipalities have been divided into four strata having sizes 73, 70, 97 and 44. The population (in thousands) in the year 1985 is considered as study variable Y and that in year 1975 as auxiliary variable X. The particulars of the parameters are given below:

To compare the efficiency of the proposed family at the optimum choice of α with the usual combined ratio and usual combined product estimators using two phase sampling in the presence of non-response, we represent the MSE and percentage relative efficiency (PRE) of the estimators. Table 2 depicts the MSE and PRE of T_1^*, T_2^* and T_C^* (at $\alpha_{(opt)}$, $a = 1$, $b = 1$ and $g = 1$) with respect to $\bar{y}_{st}^*$ for the different choices of W_{i2} and L_i.

Concluding Remark: In the present article, we have suggested a combined-type family of estimators of population mean in stratified random sampling using the information of an auxiliary variable in the presence of non-response when the information about population mean of the auxiliary variable is not available. The optimum estimator of the suggested family has been pioneered out and it is compared with the usual combined ratio and usual combined product estimators under the non-response whenever population mean of auxiliary variable is not known. The optimum values of first phase sample size n'_i, second phase sample size n_i and inverse sampling rate L_i for the different strata have been obtained in the terms of cost of the survey through the suggested family. From the table 2, we have seen that the optimum estimator of the proposed family T_C^* provides better estimates than the estimators $\bar{y}_{st}^*$, T_1^* and T_2^*. It is also seen that the precision of all the estimators decreases with the increase in the non-response rate W_{i2} as well as with increase in inverse sampling rate L_i.

REFERENCES

1. Hansen, M.H. and W.N. Hurwitz, 1946. The problem of non-response in sample surveys. Journal of the American Statistical Association, 41: 517-529.
2. Khare, B.B., 1987. Allocation In stratified sampling in presence of non-response. Metron, 45 (I/II): 213-221.

3. Khan, M.G.M., E.A. Khan and M.J. Ahsan, 2008. Optimum allocation in multivariate stratified sampling in presence of non-response. J. Ind. Soc. Agril. Statist., 62(1): 42-48.
4. Chaudhary, M.K., R. Singh, R.K. Shukla and M. Kumar, 2009. A family of estimators for estimating population mean in stratified sampling under non-response. Pakistan Journal of Statistics and Operation Research, 5(1): 47-54.
5. Chaudhary, M.K., V.K. Singh, R. Singh and F. Smarandache, 2011. On some new allocation schemes in stratified random sampling under non-response. Studies in Sampling Techniques and Time series, Zip Publishing 1313 Chesapeake Avenue Columbus, Ohio 43212, USA: 40-54.
6. Khare, B.B. and R.R. Sinha, 2004. Estimation of finite population ratio using two-phase sampling scheme in the presence of non-response. Aligarh J. Stat., 24: 43-56.
7. Singh, H.P. and S. Kumar, 2009. A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information. SORT, 33(1): 71-84.
8. Khoshnevisan, M., R. Singh, P. Chauhan, N. Sawan and F. Smarandache, 2007. A general family of estimators for estimating population mean using known value of some population parameter(s). Far East Journal of Theoretical Statistics, 22: 181-191.
9. Chaudhary, M.K., V.K. Singh and R.K. Shukla, 2012. Combined-type family of estimators of population mean in stratified random sampling under non-response. Journal of Reliability and Statistical Studies (JRSS), 5(2): 133-142.