Math. J. Interdiscip. Sci. Vol. 8, No. 2 (2020), pp.49–56

CHITKARA

Mathematical Journal of Interdisciplinary Sciences

Journal homepage: https://mjis.chitkara.edu.in/

Two-Phase Stratified Sampling Estimator for Population Mean in the Presence of Nonresponse Using One Auxiliary Variable

A.E. Anieting[1*] and E. I. Enang[2]

[1*]*Department of Statistics, University of Uyo, Uyo, Nigeria*

[2]*Department of Statistics, University of Calabar, Nigeria*

[*]Email: akaninyeneanieting@uniuyo.edu.ng

ARTICLE INFORMATION

Received: November 13, 2019
Revised: January 27, 2020
Accepted: February 03, 2020
Published Online: March 30, 2020

Keywords:
Stratified Two-phase sampling, Nonresponse, Auxiliary variable

DOI: 10.15415/mjis.2020.82006

ABSTRACT

In this article, a new estimator for the mean of population in stratified two-phase sampling in the presence of nonresponse using one auxiliary variable is been suggested. The Mean Squared Error (MSE) and the bias of the suggested estimator have been given using large sample approximation. The empirical study shows that the MSE of the suggested estimator is better than existing estimators in terms of efficiency. The optimal values of first and second phase sample have been determined.

1. Introduction

Double Sampling for stratification is one of those sample survey designs that uses auxiliary information in the process of estimation. It was introduced by Neyman (1938). The efficacy of the procedure of estimation using auxiliary information relies on technique whereby the estimator is been suggested. Research works abound whereby the accuracy of the estimators increases by using auxiliary information. Shabbir and Gupta (2005) and Kadilar and Cingi (2003) extended Singh (1999) estimators in evaluating the mean of the population that has been stratified. Although stratified double sampling is useful, the nonresponse problem is intrinsic in every survey which may lead to incorrect evaluation of the parameters. Hansen and Hurwitz (1946) suggested a method of sub-sampling of non-respondents so as to modify the non-response in their mail surveys. Khoshnevisan et al. (2007), Chaudhary et al. (2009), Chaudhary and Singh (2013) and Chaudhary and Kumar (2015) have all suggested different estimators in two-phase stratified sampling under nonresponse.

In this article, an efficient ratio-product estimator in stratified two-phase sampling under nonresponse using one auxiliary variable is been suggested. The characteristics of the estimator suggested have been given.

2. Sampling Design

Consider a population of size N divided into k strata. Let the size of the i^{th} stratum be N_i(i = 1,2,….k) such that $\sum_i^k N_i = N$. A large first sample of size $n_i^{'}$ is drawn from N_i units by (SRSWOR) scheme for the i^{th} stratum and auxiliary variable $\bar{x}_i^{'}$ is observed to estimate the population mean $\bar{X}$, which is unknown. Secondly, a smaller second phase sample of size n_i is drawn from $n_i^{'}$ unit by SRSWOR. Let Y be the study variable with population mean $\bar{Y} = \sum_i^k p_i \bar{Y}_i$ and assume that at the second phase, non-response is examined on the study variable while the auxiliary variable does not undergo nonresponse. Also, assume that at the second phase, there are n_{i1} respondent units and n_{i2} non-respondent units in n_i units. Using the method of subsampling the non-respondent introduced by Hansen and Hurwitz (1946), $h_{i2} = \frac{n_{i2}}{L_i}$, $L_i \geq 1$ being a sub-sample is drawn from the sample of n_{i2} non-respondents units and the needed data is obtained from them all.

3. Some Existing Estimators

Some existing estimators for the mean of the population in stratified two-phase sampling with one auxiliary variable under non-response shall be presented in this section.

ISSN No.: 2278-9561 (Print) ISSN No.: 2278-957X (Online) Registration No. : CHAENG/2013/49583 Periodicity: Bi-Annually

3.1. Rao (1991) Difference- Type Estimator

Rao (1991) gave a difference-type estimator given as

$$\bar{y}_{p_{10}st} = d_6 \bar{y}^*_{st} + d_7(\bar{x}'_{st} - \bar{x}^*_{st}) \tag{1}$$

Where d_6 and d_7 are constants

$$\bar{x}^*_{st} = \sum_i^k p_i \bar{x}^*_i ; \bar{x}^*_i = \frac{n_{i1}\bar{x}_{n_{i1}} + n_{i2}\bar{x}_{h_{i2}}}{n_i} ; \bar{x}'_{st} = \sum_i^k p_i \bar{x}'_i$$

With Mean Square Error (MSE),

$$\text{MSE}\left(\bar{y}_{p_{10}st}\right) = \frac{\bar{Y}^2\left(\frac{1}{\bar{\bar{Y}}^2}\Sigma_i^k p_i^2\left\{\left(\frac{1}{n_i}-\frac{1}{N_i}\right)S^2_{y_i} + \frac{(L_i-1)}{n_i}S^2_{y_{i2}}\right\}\left(1-\rho^2_{yx(st)}\right)\right)}{1+\left(\frac{1}{\bar{\bar{Y}}^2}\Sigma_i^k P_i^2\left\{\left(\frac{1}{n_i}-\frac{1}{N_i}\right)S^2_{y_i} + \frac{(L_i-1)}{n_i}S^2_{y_{i2}}\right\}\left(1-\rho^2_{yx(st)}\right)\right)} \tag{2}$$

3.2. Khare and Srivastava (1993) Estimator

Khare and Srivastava suggested a ratio estimator of the form

$$\hat{\bar{Y}}_{RD} = \bar{y}^*_{st}\frac{\bar{x}'_{st}}{\bar{x}^*_{st}} \tag{3}$$

With mean square error,

$$\begin{aligned}\text{MSE}\left(\hat{\bar{Y}}_{RD}\right) = \bar{Y}^2\Bigg[&\frac{1}{\bar{Y}^2}\Sigma_i^k p_i^2\left(\left\langle\frac{1}{n_i}-\frac{1}{N_i}\right\rangle S^2_{y_i} + \frac{(L_i-1)}{n_i}W_{i2}S^2_{Y_{i2}}\right) \\ &+\frac{1}{\bar{X}^2}\Sigma_i^k p_i^2\left(\left\langle\frac{1}{n'_i}-\frac{1}{N_i}\right\rangle S^2_{X_i} + \frac{(L_i-1)}{n_i}W_{i2}S^2_{X_{i2}}\right) \\ &-\frac{2}{\bar{X}\bar{Y}}\Sigma_i^k p_i^2\left(\left\langle\frac{1}{n_i}-\frac{1}{N_i}\right\rangle S_{X_i} + \frac{(L_i-1)}{n_i}W_{i2}S_{yx_{i2}}\right)\Bigg]\end{aligned} \tag{4}$$

3.3. Khare and Srivastava (1997) Estimator

Khare and Srivastava (1997) proposed two ratio estimators given as

$$T_1^* = \bar{y}^*_{st}\frac{\bar{x}'_{st}}{\bar{x}_{st}} \tag{4}$$

With mean square error,

$$\begin{aligned}\text{MSE}\left(T_1^*\right) = &\sum_i^k f'_i p_i^2 S^2_{Y_i} + \sum_i^k f^*_i p_i^2\left(S^2_{Y_i} + R^2S^2_{X_i} - 2R\rho_i S_{X_i}S_{Y_i}\right) \\ &+\sum_i^k \frac{(L_i-1)}{n_i}W_{i2}P_i^2S^2_{Y_{i2}}\end{aligned} \tag{5}$$

Where $f'_i = \left(\frac{1}{n'_i}-\frac{1}{N_i}\right), f^*_i = \left(\frac{1}{n_i}-\frac{1}{n'_i}\right)$ and

$$T_2^* = \bar{y}^*_{st}\frac{\bar{x}_{st}}{\bar{x}'_{st}} \tag{6}$$

With mean square error,

$$\begin{aligned}\text{MSE}\left(T_1^*\right) = &\sum_i^k f'_i\ p_i^2 S^2_{Y_i} + \sum_i^k f^*_i p_i^2\left(S^2_{Y_i} + R^2S^2_{X_i} - 2R\rho_i S_{X_i}S_{Y_i}\right) \\ &+\sum_i^k \frac{(L_i-1)}{n_i}W_{i2}P_i^2S^2_{Y_{i2}}\end{aligned} \tag{7}$$

Where $f'_i = \left(\frac{1}{n'_i}-\frac{1}{N_i}\right), f^*_i = \left(\frac{1}{n_i}-\frac{1}{n'_i}\right),$

3.4.Chaudhary et al. (2009) Estimator

Chaudhary et. al. (2009) suggested a family of combined-type estimators given as

$$T_c = \bar{y}^*_{st}\left[\frac{a\bar{X}+b}{\alpha\left(a\bar{X}_{st}+b\right)+(1-\alpha)\left(a\bar{X}+b\right)}\right]^g \tag{8}$$

Where $\bar{X}_{st} = \sum_i^k p_i\bar{x}_i$ and $\bar{X} = \Sigma_i^k p_i\bar{X}_i; \bar{x}_i$ and $\bar{X}_i$ are mean based on n_i units and mean based on N_i individually in each stratum for auxiliary variable.
The mean squared error (MSE) was given as

$$\begin{aligned}\text{MSE}\left(T_C\right) = &\Sigma_i^k f_i p_i^2\left[S^2_{Y_i} + \alpha^2\lambda^2g^2R^2S^2_{X_i} - 2\alpha\lambda gR\rho_i S_{X_i}S_{Y_i}\right] \\ &+\Sigma_i^k\frac{(L_i-1)W_{i2}P_i^2S^2_{Y_{i2}}}{n_i}\end{aligned} \tag{9}$$

Where $f_i = \left(\frac{1}{n_i}-\frac{1}{N_i}\right), \lambda = \frac{a\bar{X}}{a\bar{X}+b},$

3.5.Chaudhary et al. (2012) Estimator

Chaudhary et al. (2012) gave a combined type estimator given as

$$T_{Fc}(\alpha) = \bar{y}^*_{st}\left[\frac{(A+C)\bar{X} + fB\bar{X}_{st}}{(A+B)\bar{X} + C\bar{X}_{st}}\right] \tag{10}$$

Where $\text{A} = (\alpha-1)(\alpha-2), \text{B} = (\alpha-1)(\alpha-4),$

$\text{C} = (\alpha-3)(\alpha-2)(\alpha-4), \alpha > 0, \text{f} = \frac{n'_i}{N_i}$

The mean square error as obtained by them is

$$\begin{aligned}\text{MSE}\left(T_{Fc}(\alpha)\right) = &\Sigma_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right) \\ &P_i^2\left[S^2_{Y_i} + \phi(\alpha)^2RS^2_{X_i} - 2\phi(\alpha)R\rho_i S_{X_i}S_{Y_i}\right]\end{aligned} \tag{11}$$

Where $\phi(\alpha) = \frac{C-fB}{A+fB+C}$

3.6. Chaudhary And Kumar (2015) Estimator

A combined type estimator was proposed by Chaudhary and Kumar (2015) as follows:

$$T_c' = \bar{y}_{st}^*\left[\frac{a\bar{x}_{st}' + b}{\alpha\left(a\bar{x}_{st} + b\right) + (1-\alpha)\left(a\bar{x}_{st}' + b\right)}\right]^g \quad (12)$$

With mean square error given as

$$\mathrm{MSE}\left(T_c'\right) = \Sigma_i^k f_i' \, p_i^2 S_{y_i}^2 + \Sigma_i^k f_i^* p_i^2 (S_{y_i}^2 + g^2\lambda^2 R^2 \alpha^2 S_{x_i}^2 - 2g\lambda R\alpha p_i S_{x_i} S_{y_i}) + \Sigma_i^k \left(\frac{L_i - 1}{n_i}\right) W_{i2} p_i^2 S_{y_i}^2 \quad (13)$$

4. Proposed Estimator and its Properties

The proposed estimator for population mean in two-phase stratified in the presence of nonresponse using single auxiliary variable is given as,

$$T_{ae} = \bar{y}_{st}^*\left(\frac{\bar{x}_{st}' + \varphi}{\bar{x}_{st} - \varphi}\right) \quad (14)$$

Where $\varphi = \Sigma_i^k C_{x_i}$ where C_{x_i} is the coefficient of variation of the auxiliary variable.

$$\bar{y}_{st}^* = \Sigma_i^k p_i \bar{y}_i^*; \bar{y}_i^* = \frac{n_{i1}\bar{y}_{n_{i1}} + n_{i2}\bar{y}_{h_{i2}}}{n_i}, p_i = \frac{N_i}{N}$$

$\bar{y}_{n_{i1}}$ and $\bar{y}_{h_{i2}}$ are the means based on n_{i1} respondent units and h_{i2} subsampled non-respondent units respectively for the study variable and $\bar{x}_{st}^{'} = \Sigma_i^k p_i \bar{x}_i^{'}, \bar{x}_{st} = \Sigma_i^k p_i \bar{x}_i$

4.1. The Mean Squared Error (MSE) and the Bias of the Estimator Proposed

To derive the expression for the mean squared error and the bias of the estimator proposed for the mean of population in stratified two-phase sampling under nonresponse, the following symbolization shall be used.

This method is adopted from Chaudhary and Kumar (2015).

$\bar{y}_{st}^* = \bar{Y}(1+e_0)$, $\bar{x}_{st} = \bar{X}(1+e_1)$, $\bar{x}_{st}^{'} = \bar{X}(1+e_2)$, where $e_i s$ are the relative error terms and are defined as

$$e_0 = \frac{\bar{y}_{st}^* - \bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x}_{st} - \bar{X}}{\bar{X}}, e_2 = \frac{\bar{x}_{st}' - \bar{X}}{\bar{X}}$$

Such that the following expectations are applied

$\mathrm{E}(e_0) = (e_1) = (e_2) = 0$

$$E\left(e_0^2\right) = \frac{1}{\bar{Y}^2}\Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{Y_i}^2 + \frac{(L_i - 1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]$$

$$\mathrm{E}\left(e_1^2\right) = \frac{1}{\bar{X}^2}\Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right]$$

$$\mathrm{E}\left(e_2^2\right) = \frac{1}{\bar{X}^2}\Sigma_i^k\left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] = \mathrm{E}(e_1 e_2)$$

$$\mathrm{E}\left(e_0 e_1\right) = \frac{1}{\bar{X}\bar{Y}}\Sigma_i^k\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i}$$

$$\mathrm{E}\left(e_0 e_2\right) = \frac{1}{\bar{X}\bar{Y}}\Sigma_i^k\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i}$$

Using the large sample approximation used by Chaudhary and Kumar (2015) the estimator proposed has its bias stated as

$$\begin{aligned}\mathrm{Bias}\left(T_{ae}\right) = \bar{Y}\Bigg[& b^2\frac{1}{\bar{X}^2}\Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] \\ & - ab\frac{1}{\bar{X}^2}\Sigma_i^k\left[\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] \\ & - b\frac{1}{\bar{Y}\bar{X}}\Sigma_i^k\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i} \\ & + a\frac{1}{\bar{Y}\bar{X}}\Sigma_i^k\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i}\Bigg]\end{aligned}$$

Hence, the Mean Squared Error (MSE) of the estimator proposed is stated as

$$\begin{aligned}\mathrm{MSE}\left(T_{ae}\right) = \Bigg[& b^2 R^2 \Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] \\ & - 2abR^2\Sigma_i^k\left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] \\ & + a^2 R^2 \Sigma_i^k\left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] \\ & - 2bR\,\Sigma_i^k\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i} \\ & + 2aR\,\Sigma_i^k\left(\frac{1}{n_i'} - \frac{1}{N_i}\right)p_i^2 \rho_i S_{X_i} S_{Y_i} \\ & + \Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2 S_{Y_i}^2 + \frac{(L_i - 1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]\Bigg]\end{aligned}$$

Where $\mathrm{R} = \frac{\bar{Y}}{\bar{X}}, \mathrm{a} = \frac{\bar{X}}{\bar{X} + \varphi}$ and $\mathrm{b} = \frac{\bar{X}}{\bar{X} - \varphi}$

4.2. The Survey Cost and Obtaining Optimal Values of n_i, $n_i^{'}$ and L_i for T_{ae}

Let $c_i^{'}$ represent the cost in each unit related to the sample size of the first phase $n_i^{'}$ also let c_{i0} represent the cost of each unit of the first attempt on the study variable having the sample size of second phase, n_i. Consider c_{i1} and c_{i2} to represent the cost in each unit of computing the n_{i1} units that responded and h_{i2} units that did not respond. Hence the overall cost for each stratum is stated as

$$C_i = c_i^{'}n_i^{'} + c_{i0}n_i + c_{i1}n_{i1} + c_{i2}h_{i2} \tag{15}$$

The expected cost for each stratum is given by

$$\mathrm{E}(C_i) = c_i^{'}n_i^{'} + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right). \tag{16}$$

Where W_{i1} is the response rate in the i^{th} stratum, W_{i2} represents the rate of non-response rate in the strata and L_i is the inverse sampling rate.
Hence the overall cost over the entire strata is given as

$$\begin{aligned} C_0 &= \Sigma_i^k \mathrm{E}(C_i) \\ &= \Sigma_i^k\left[c_i^{'}n_i^{'} + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)\right] \end{aligned} \tag{17}$$

Consider the Lagrange function

$$\begin{aligned} \phi &= MSE(T_{ae}) + \lambda C_0 \\ &= \left[b^2R^2\Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2S_{X_i}^2\right] - 2abR^2\Sigma_i^k\left[\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2S_{X_i}^2\right]\right. \\ &\quad + a^2R^2\Sigma_i^k\left[\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2S_{X_i}^2\right] - 2bR\Sigma_i^k\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2\rho_iS_{X_i}S_{Y_i} \\ &\quad + 2aR\Sigma_i^k\left(\frac{1}{n_i^{'}} - \frac{1}{N_i}\right)p_i^2\rho_iS_{X_i}S_{Y_i} \\ &\quad \left. + \Sigma_i^k\left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)p_i^2S_{Y_i}^2 + \frac{(L_i-1)}{n_i}p_i^2W_{i2}S_{Y_{i2}}^2\right]\right] \\ &\quad + \lambda\Sigma_i^k\left[c_i^{'}n_i^{'} + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)\right] \end{aligned}$$

Where λ is the multiplier of the Lagrange function
To derive the optimum values of n_i, $n_i^{'}$, and L_i, ϕ is differentiated with regard to n_i, $n_i^{'}$, and L_i individually and the derivatives equated to zero. Hence, for stratum i, we get

$$\begin{aligned} \frac{\partial\phi}{\partial n_i} &= \frac{-P_i^2}{n_i^2}\left[S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_iS_{Y_i}S_{X_i}\right] \\ &\quad - \frac{(L_i-1)W_{i2}S_{Y_{i2}}^2P_i^2}{n_i^2} \\ &\quad + \lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right) = 0 \end{aligned} \tag{18}$$

$$\frac{\partial\phi}{\partial n_i^{'}} = \frac{P_i^2}{n_i^{'2}}\left[2abR^2S_{X_i}^2 - a^2R^2S_{X_i}^2 - 2aR\rho_iS_{Y_i}S_{X_i}\right] + \lambda c_i^{'} = 0 \tag{19}$$

$$\frac{\partial\phi}{\partial L_i} = \frac{P_i^2}{n_i}\left[W_{i2}S_{Y_{i2}}^2\right] - \lambda n_ic_{i2}\frac{W_{i2}}{L_i^2} = 0 \tag{20}$$

From (18)

$$\begin{aligned} &\lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right) \\ &= \frac{P_i^2}{n_i^2}\left[S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_iS_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2\right] \\ n_i &= \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_iS_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2}}{\sqrt{\lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)}} \end{aligned} \tag{21}$$

From (19)

$$\begin{aligned} \lambda c_i^{'} &= \frac{P_i^2}{n_i^{'2}}\left[a^2R^2S_{X_i}^2 + 2aR\rho_iS_{Y_i}S_{X_i} - 2abR^2S_{X_i}^2\right] \\ n_i^{'} &= \frac{P_i\sqrt{Q}}{\sqrt{\lambda c_i^{'}}} \end{aligned} \tag{22}$$

Where $Q = a^2R^2S_{X_i}^2 + 2aR\rho_iS_{Y_i}S_{X_i} - 2abR^2S_{X_i}^2$
From (20)

$$\lambda = \frac{L_i^2P_i^2S_{Y_{i2}}^2}{n_i^2c_{i2}} \text{ so that } \sqrt{\lambda} = \frac{L_iP_iS_{Y_{i2}}}{n_i\sqrt{C_{i2}}} \tag{23}$$

putting the value of $\sqrt{\lambda}$ in (23) into (21), we get

$$n_i = \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_iS_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2}}{\frac{L_iP_iS_{Y_{i2}}}{n_i\sqrt{C_{i2}}}\sqrt{\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)}}$$

$$L_i S_{Y_{i2}} \sqrt{\left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i}\right)}$$
$$= \sqrt{(S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} + (L_i - 1) W_{i2} S_{Y_{i2}}^2) c_{i2}}$$

Squaring both sides, we have

$$L_i^2 S_{Y_{i2}}^2 \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i}\right)$$
$$= (S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} + (L_i - 1) W_{i2} S_{Y_{i2}}^2) c_{i2}$$
$$L_i^2 S_{Y_{i2}}^2 c_{i0} + L_i^2 S_{Y_{i2}}^2 c_{i1} W_{i1} + S_{Y_{i2}}^2 c_{i2} L_i W_{i2} - S_{Y_{i2}}^2 c_{i2} L_i W_{i2}$$
$$= \left(S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} - W_{i2} S_{Y_{i2}}^2\right) c_{i2}$$
$$L_i^2 S_{Y_{i2}}^2 \left(c_{i0} + c_{i1} W_{i1}\right)$$
$$= \left(S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} - W_{i2} S_{Y_{i2}}^2\right) c_{i2}$$

$$L_{i(opt)} = \frac{\sqrt{c_{i2}} B_i}{S_{Y_{i2}} A_i} \tag{24}$$

Where $A_i = \sqrt{c_{i0} + c_{i1} W_{i1}}$;

$$B_i = \sqrt{S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} - W_{i2} S_{Y_{i2}}^2}$$

On substituting the value of $L_{i(opt)}$ from (57) into (54), we can express n_i as

$$n_i = \frac{P_i \sqrt{S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} + \left(\frac{\sqrt{c_{i2}} B_i}{S_{Y_{i2}} A_i} - 1\right) W_{i2} S_{Y_{i2}}^2}}{\sqrt{\lambda \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{\frac{\sqrt{c_{i2}} B_i}{S_{Y_{i2}} A_i}}\right)}}$$

$$n_i = \frac{P_i \sqrt{S_{Y_i}^2 + b^2 R^2 S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} - W_{i2} S_{Y_{i2}}^2 + \left(\frac{\sqrt{c_{i2}} B_i S_{Y_{i2}} W_{i2}}{A_i}\right)}}{\sqrt{\lambda \left(c_{i0} + c_{i1} W_{i1} + \sqrt{c_{i2}} \frac{W_{i2} S_{Y_{i2}} A_i}{B_i}\right)}}$$

$$n_i = \frac{P_i \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda} \sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}}} \tag{25}$$

To obtain the value of $\sqrt{\lambda}$ with regard to the total cost C_0, the value of n_i, $n_i^{'}$, and $L_{i(opt)}$ are put into (17), hence we have,

$$C_0 = \Sigma_i^k \left[c_i^{'} n_i^{'} + n_i \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{L_i}\right)\right]$$
$$= \Sigma_i^k \left[c_i^{'} \frac{p_i \sqrt{Q}}{\sqrt{\lambda c_i^{'}}} + \frac{p_i \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} s_{Y_{i2}}}{A_i}}}{\sqrt{\lambda} \sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}}} \left(c_{i0} + c_{i1} W_{i1} + c_{i2} \frac{W_{i2}}{\frac{\sqrt{c_{i2}} B_i}{S_{Y_{i2}} A_i}}\right)\right]$$
$$= \Sigma_i^k \left[\sqrt{c_i^{'}} \frac{p_i \sqrt{Q}}{\sqrt{\lambda}} + \frac{p_i \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} s_{Y_{i2}}}{A_i}}}{\sqrt{\lambda} \sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}}} \left(c_{i0} + c_{i1} W_{i1} + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}\right)\right]$$
$$= \Sigma_i^k \left[\sqrt{c_i^{'}} \frac{P_i \sqrt{Q}}{\sqrt{\lambda}} + \frac{P_i \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda} \sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}}} \left(A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}\right)\right]$$
$$\sqrt{\lambda} = \frac{1}{C_0} \Sigma_i^k \left[\sqrt{c_i^{'}} P_i \sqrt{Q} + P_i \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} S_{Y_{i2}}}{A_i}} \sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}}\right]$$

$$=\frac{1}{C_0}\Sigma_i^k\left[\sqrt{c_i^{'}}P_i\sqrt{Q}+P_i\sqrt{\left(B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}\right)\left(A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}\right)}\right]$$

$$=\frac{1}{C_0}\Sigma_i^k\left[\sqrt{c_i^{'}}P_i\sqrt{Q}+P_i\sqrt{\frac{B_i^3A_i^3+2\sqrt{c_{i2}}W_{i2}B_i^2A_i^2S_{Y_{i2}}+c_{i2}B_iA_iW_{i2}^2S_{Y_{i2}}^2}{B_iA_i}}\right]$$

$$=\frac{1}{C_0}\Sigma_i^k\left[\sqrt{c_i^{'}}P_i\sqrt{Q}+P_i\sqrt{B_i^2A_i^2+2\sqrt{c_{i2}}B_iA_iW_{i2}S_{Y_{i2}}+c_{i2}W_{i2}^2S_{Y_{i2}}^2}\right]$$

$$=\frac{1}{C_0}\Sigma_i^k\left[\sqrt{c_i^{'}}P_i\sqrt{Q}+P_i\sqrt{\left(B_iA_i+\sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)^2}\right]$$

$$\sqrt{\lambda}=\frac{1}{C_0}P_i\Sigma_i^k\left[\sqrt{c_i^{'}}\sqrt{Q}+\left(B_iA_i+\sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)\right] \tag{26}$$

Substituting the value of $\sqrt{\lambda}$ from (26) into (25) and (22), the optimum values of n_i and $n_i^{'}$ are derived as

$$n_{i(opt)}=\frac{C_0\sqrt{B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}}{\sqrt{A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}\sum_i^k\left[\sqrt{c_i^{'}}\sqrt{Q}+\left(B_iA_i+\sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)\right]}$$

$$n_{i(opt)}^{'}=\frac{C_0\sqrt{Q}}{\sqrt{c_i^{'}}\sum_i^k\left[\sqrt{c_i^{'}}\sqrt{Q}+\left(B_iA_i+\sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)\right]}$$

5. Empirical Study

Using the data set used by Chaudhary and Kumar (2015) shown on Table 1 below

Table 1:

Stratum no	N_i	$n_i^{'}$	n_i	$\bar{Y}_i$	$\bar{X}_i$	$S_{Y_i}^2$	$S_{X_i}^2$	ρ_i	$S_{Y_{i2}}^2$
1	73	65	26	40.85	39.56	6369.1	6624.44	0.999	618.88
2	70	25	10	27.57	27.57	1051.07	1147.01	0.998	240.91
3	97	48	19	25.44	25.44	2014.97	2205.4	0.999	265.52
4	44	11	5	20.36	20.36	538.47	485.27	0.997	83.69

Table 2: below shows the Mean Squared Error (MSE) and Relative Efficiency percentage (PRE) of different estimators together with the estimator proposed being compared tofor different choices of.

W_{i2}	L_i	$V\,\bar{y}_{st}^{*}$	MSE(T_1^{*})	MSE$(T_c^{'})$	MSE(T_{ae})	PRE(T_1^{*})	PRE$(T_c^{'})$	PRE(T_{ae})
0.1	2	34.42	23.62	6.28	4.66	145.72	548.09	738.6
	2.5	34.67	24.01	6.54	4.92	144.4	530.12	704.7
	3	34.92	24.41	6.79	5.17	143.06	514.29	675.4
	3.5	35.18	24.8	7.04	5.31	141.85	499.72	662.5
0.2	2	34.92	24.41	6.79	5.26	143.06	514.29	663.9
	2.5	35.43	25.2	7.3	5.66	140.6	485.34	625.97
	3	35.94	25.99	7.8	6.18	138.28	460.77	581.6
	3.5	36.44	26.78	8.31	6.70	136.07	438.51	543.9

W_{i2}	L_i	$V\,\bar{y}^*_{st}$	$MSE(T_1^*)$	$MSE(T_c')$	$MSE(T_{ae})$	$PRE(T_1^*)$	$PRE(T_c')$	$PRE(T_{ae})$
0.3	2	35.43	25.2	7.3	5.66	140.6	485.34	625.97
	2.5	36.19	26.38	8.06	6.44	137.19	449.01	561.96
	3	36.95	27.57	8.82	7.22	134.02	418.93	511.8
	3.5	37.71	28.75	9.57	7.98	131.17	394.04	472.6
0.4	2	35.94	25.99	7.8	6.18	138.28	460.77	581.6
	2.5	36.95	27.57	8.82	7.22	134.02	418.93	511.8
	3	37.96	29.14	9.83	8.26	130.27	386.16	459.6
	3.5	38.97	30.72	10.84	9.29	126.86	359.5	419.5

Conclusion

The optimum values of n_i', n_i and L_i using the proposed estimator had been ascertained under the cost of the survey. Also, Table 2 shows that the proposed estimator is better than other estimators in terms of efficiency.

References

Chaudhary, M. K. and Kumar, A.: Use of double sample scheme in estimating the mean of Stratified population Under nonresponse. STATISTICA, anno, Lxxv(4) (2015).

Chaudhary, M. K. and Singh, V. K.: Estimating the population mean in two-stage sampling with equal size clusters under non-response using auxiliary characteristic. Mathematical Journal of Interdisciplinary Sciences 2(1), 43–55 (2013).

Chaudhary, M. K., Singh, R., Shukla, R. K., Kumar, M.: 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 (2009).

Chaudhary, M. K., Singh, V. K. and Shukla, R. K.: 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 (2012).

Hansen, M. H. and Hurwitz, W. N.: The problem of non-response in sample surveys. Journal of The American Statistical Association 41(236), 517–529 (1946). DOI: 10.1080/01621459.1946.10501894

Kadilar, C. and Cingi, H.: Ratio estimators in stratified random sampling. Biometrical journal 2(45), 218–225 (2003).

Khare, B. B. and Sinha, R. R.: Estimation of finite population ratio using two-phase sampling scheme in the presence of non-response. Aligarh J. Stat. **24,** 43-56 (2004).

Khoshnevisan, M., Singh, Chauhan, R., P., Sawan, N. and Smarandache, F.: 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 (2007).

Shabbir, J. and Gupta, S.: Improved ratio estimators in stratified sampling. American Journal of Mathematical and Management Sciences 25(3-4), 293–311 (2005).

Singh, R., Malik, S., Chaudhary, M. K., Verma, H. K. and Adewara, A. A.: A general family of ratio-type estimators in systematic sampling. Journal of Reliability and Statistical Studies 5(1), 73–82 (2012).

Upadhyaya, L. N. and Singh, H. P.: Use of transformed auxiliary variable in estimating the finite population mean. Biometrical Journal 41(5), 627–636 (1999).

Mathematical Journal of Interdisciplinary Sciences

Chitkara University, Saraswati Kendra, SCO 160-161, Sector 9-C,
Chandigarh, 160009, India

CHITKARA

Volume 8, Issue 2 **March 2020** **ISSN 2278-9561**