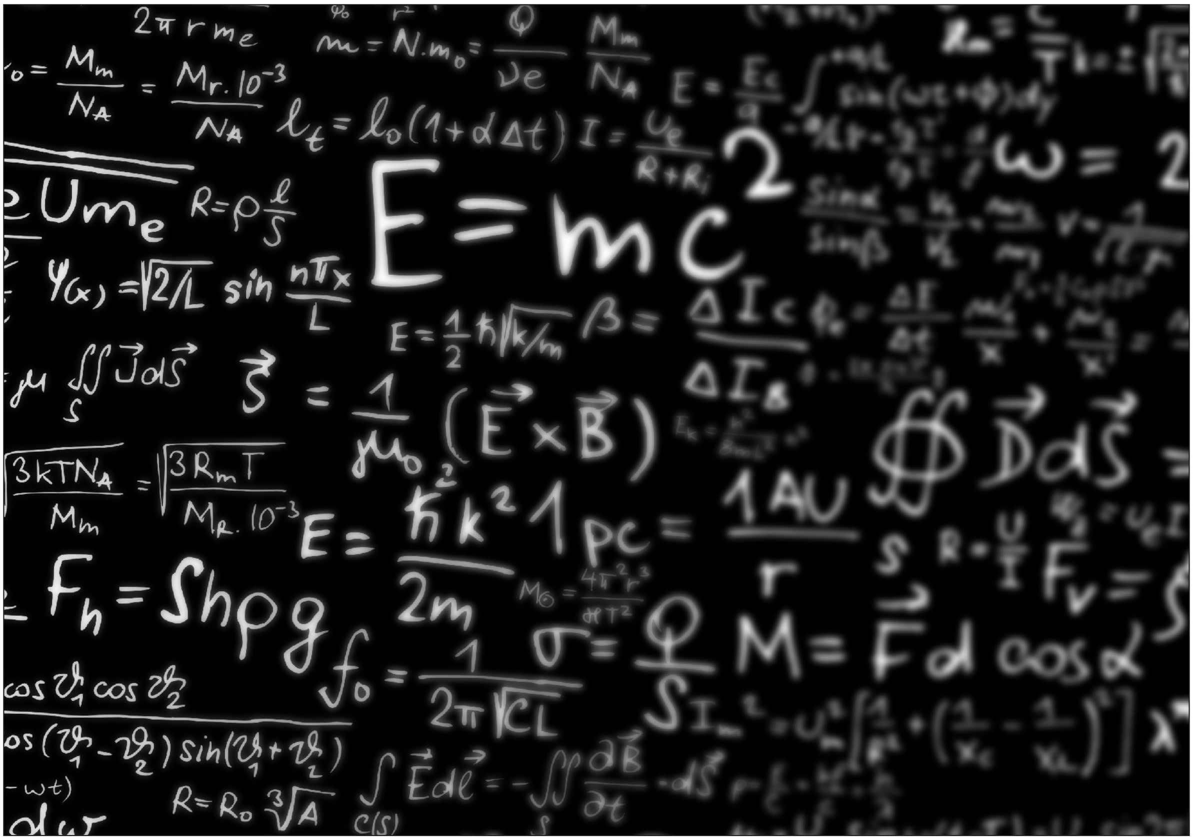

Asian Journal of Mathematics & Statistics

ISSN 1994-5418

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS
Asian Journal of Mathematics and Statistics

ISSN 1994-5418
DOI: 10.3923/ajms.2018.27.39

Research Article

Population Mean Estimation Using Ratio-cum Product Compromised-method of Imputation in Two-phase Sampling Scheme

Krishnajyothi Nath and B.K. Singh

Department of Mathematics, North Eastern Regional Institute of Science and Technology (NERIST) Nirjuli, 791109, Arunachal Pradesh, India

Abstract

Background and Objective: In literature there has been a study on ratio cum product estimator of a finite population mean in two-phase sampling in sample surveys, but it lacks study when there is non-response on sample observations. So the main objective of this paper was to propose three generalized classes of ratio cum product compromised imputation techniques in presence of missing values in two-phase sampling design and its properties have been studied. **Materials and Methods:** The estimators were compared with other existing estimators in two different designs. The bias and M.S.E. of suggested estimators were derived in the form of population parameters using the concept of large sample approximations. **Results:** The results showed the superiority of the proposed estimators over the existing methods. Numerical studies are performed over two population data sets using the expressions of bias and M.S.E. and their efficiencies are compared with other existing estimators. **Conclusion:** It was observed that the proposed estimators were performing better than the estimators taken for comparisons in the presence of missing data.

Key words: Missing data, bias, mean squared error, two-phase sampling, ratio, product, large sample approximation, simple random sampling

Citation: Krishnajyothi Nath and B.K. Singh, 2018. Population mean estimation using ratio-cum product compromised-method of imputation in two-phase sampling scheme. Asian J. Math. Stat., 11: 27-39.

Corresponding Author: B.K. Singh, Department of Mathematics, North Eastern Regional Institute of Science and Technology (NERIST) Nirjuli, 791109 Arunachal Pradesh, India Tel: +919436270096

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Missing data is a problem encountered in almost every data collection activity but particularly in sample survey. To overcome the problem of missing observations or non-response in sample surveys, the technique of imputation is frequently used to replace the missing data. In literature, several imputation techniques are described, some of them are better over others. To deal with missing values effectively Kalton *et al.*[1] and Sande[2] suggested imputation that make an incomplete data set structurally complete and its analysis simple. Lee *et al.*[3,4] used the information on an auxiliary variate if it is available. Later Singh and Horn[5] suggested a compromised method of imputation. Shukla[6], Singh and Deo[7], Ahmed *et al.*[8], Arnab and Singh[9], Rueda and Gonzalez[10], Gonzalez *et al.*[11], Bouza[12], Kadilar and Cingi[13], Shukla and Thakur[14], Shukla *et al.*[15,16], Baraldi and Enders[17], Diana and Perri[18], Singh *et al.*[19], Shukla *et al.*[20], Thakur *et al.*[21], Shukla *et al.*[22], Thakur *et al.*[23], Singh *et al.*[24,25], Thakur *et al.*[26], Singh and Gogoi[27], Nath and Singh[28] and Singh and Espejo[29] suggested several new imputation based estimators that use information on an auxiliary variate under single and double sampling scheme. The objective of the present research work was to derive some imputation methods for mean estimation in case population parameter of auxiliary information is missing or unknown.

NOTATIONS

Let W = {1, 2,..., N} be a finite population with Y_i as a variable of main interest and X_i (i = 1, 2,..., N) an auxiliary variable. As usual:

$$\overline{Y} = N^{-1}\sum_{i=1}^{N} Y_i,\ \overline{X} = N^{-1}\sum_{i=1}^{N} X_i$$

are population means, $\overline{X}$ is unknown and $\overline{Y}$ under investigation.

Consider a preliminary large sample S′ of size n′ is drawn from population Ω by Simple Random Sampling Without Replacement (SRSWOR) and a secondary sample S of size n(n<n′) is drawn in either of the following manners:

- **Case-I:** as a sub-sample from sample S′ (denoted by design I) as in Fig. 1a
- **Case-II:** Independent to sample S′ (denoted by design II) as in Fig. 1b, without replacing S′

Let sample size S of n units contains r responding units (r<n) forming a sub-space R and (n-r) non-responding with sub-space R^C in $S = R \cup R^C$. For every $i \in R$, y_i is observed available. For $i \in R^C$, the y_i values are missing and imputed values are computed. The ith value x_i of auxiliary variate is used as a source of imputation for missing data when $i \in R^C$. Assume for S, the data $x_s = \{x_i: i \in S\}$ and for $I' \in S'$, the data $\{x_i: I' \in S'\}$ are known with mean:

$$\overline{x} = (n)^{-1}\sum_{i=1}^{n} x_i \ \text{and}\ \overline{x}' = (n')^{-1}\sum_{i=1}^{n'} x_i$$

respectively. The following symbols are used hereafter:

$\overline{X}, \overline{Y}$	=	The population mean of X and Y
$\overline{x}, \overline{y}$	=	The sample mean of X and Y
$\overline{x}_r, \overline{y}_r$	=	Sample mean of X and Y for corresponding responding units
ρ_{xy}	=	The correlation co-efficient between X and Y
S_x^2, S_y^2	=	The population mean squares of X and Y
C_x, C_y	=	The co-efficient of variation of X and Y

$$\delta_1 = \left(\frac{1}{r} - \frac{1}{n'}\right);\ \delta_2 = \left(\frac{1}{n} - \frac{1}{n'}\right);\ \delta_3 = \left(\frac{1}{n'} - \frac{1}{N}\right);$$
$$\delta_4 = \left(\frac{1}{r} - \frac{1}{N-n'}\right);\ \delta_5 = \left(\frac{1}{n} - \frac{1}{N-n'}\right)$$

LARGE SAMPLE APPROXIMATION

Let, $\overline{y}_r = \overline{Y}(1+e_1)$; $\overline{x}_r = \overline{X}(1+e_2)$; $\overline{x} = \overline{X}(1+e_3)$ and $\overline{x}' = \overline{X}(1+e_3')$ which implies the results $e_1 = \frac{\overline{y}_r}{\overline{Y}} - 1$; $e_2 = \frac{\overline{x}_r}{\overline{X}} - 1$; $e_3 = \frac{\overline{x}}{\overline{X}} - 1$ and $e_3' = \frac{\overline{x}'}{\overline{X}} - 1$. Now by using the concept of two-phase sampling, following Rao and Sitter[30] and the mechanism of missing completely at random (MCAR), for given r, n and n′, we have:

- **Under design F_1 [Case I]:** Under design F_1 is shown in Fig. 1a:

$$E(e_1) = E(e_2) = E(e_3) = E(e_3') = 0;\ E(e_1^2) = \delta_1 C_Y^2;\ E(e_2^2) = \delta_1 C_X^2;$$
$$E(e_3^2) = \delta_2 C_X^2;\ E(e_3'^2) = \delta_3 C_X^2;\ E(e_1 e_2) = \delta_1 \rho C_Y C_X;$$
$$E(e_1 e_3) = \delta_2 \rho C_Y C_X;\ E(e_1 e_3') = \delta_3 \rho C_Y C_X;\ E(e_2 e_3) = \delta_2 C_X^2;$$
$$E(e_2 e_3') = \delta_3 C_X^2;\ E(e_3 e_3') = \delta_3 C_X^2$$

- **Under design F_2 [Case II]:** Under design F_2 is shown in Fig.1b:

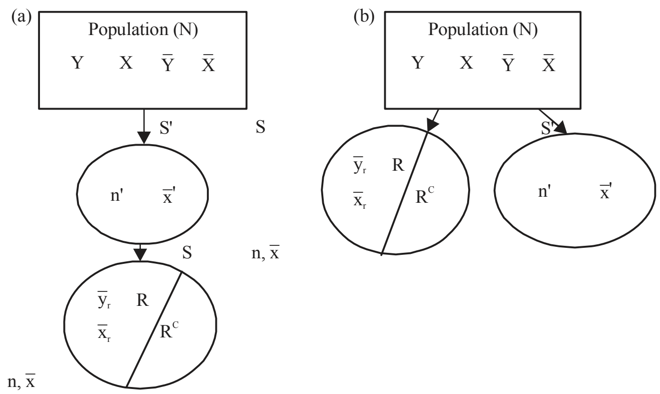

Fig. 1(a-b): (a) Design I, F_1 and (b) Design II, F_2

$$E(e_1) = E(e_2) = E(e_3) = E(e_3') = 0;\ E(e_1^2) = \delta_4 C_Y^2;\ E(e_2^2) = \delta_4 C_X^2;$$
$$E(e_3^2) = \delta_5 C_X^2;\ E(e_3'^2) = \delta_3 C_X^2;\ E(e_1 e_2) = \delta_4 \rho C_Y C_X;\ E(e_1 e_3) = \delta_5 \rho C_Y C_X;$$
$$E(e_1 e_3') = 0;\ E(e_2 e_3) = \delta_5 C_X^2;\ E(e_2 e_3') = 0;\ E(e_3 e_3') = 0$$

SOME EXISTING IMPUTATION TECHNIQUES

Let, $\bar{Y} = N^{-1}\sum_{i=1}^{N} y_i$ be the mean of the finite population under consideration. A simple random sampling without replacement (SRSWOR) S of size n is drawn from $\Omega = \{1, 2,..., N\}$ to estimate the population mean $\bar{Y}$. Let the number of responding units out of sampled n units be denoted by r(r<n), the set of responding units by R and that of non-responding units by R^C. For every unit $i \in R$ the value y_i is observed, but for the units $i \in R^C$, the observations y_i are missing and instead imputed values are derived. The ith value x_i of auxiliary variate is used as a source of imputation for missing data when $i \in R^C$. Assume for S, the data $x_s = \{x_i : i \in S\}$ are known with mean $\bar{x} = (n)^{-1}\sum_{i'=1}^{n} x_i$. Under this setup, some well known imputation methods are given below:

Mean methods of imputation: The mean imputation method is to replace each missing datum with the mean of the observed value. The data after imputation becomes:

For y_i define y_{oi} as:

$$y_{oi} = \begin{cases} y_i & \text{if} \quad i \in R \\ \bar{y}_r & \text{if} \quad i \in R^C \end{cases}$$

Using above, the imputation-based estimators of population mean $\bar{Y}$ is $\bar{y}_m = \frac{1}{r}\sum_{i \in R} y_i = \bar{y}_r$.

The bias and mean square error is given by:

- $B(\bar{y}_m) = 0$
- $V(\bar{y}_m) = \left(\frac{1}{r} - \frac{1}{N}\right) S_Y{}^2$

Ratio method of imputation: Following the notations of Lee *et al.*[3] in the case of single imputation method, if the ith unit requires imputation, the value $\hat{b} x_i$ is imputed.

For y_i and x_i, define y_{oi} as:

$$y_{oi} = \begin{cases} y_i & \text{if} \quad i \in R \\ \hat{b} x_i & \text{if} \quad i \in R^C \end{cases} \quad \text{where,} \quad \hat{b} = \sum_{i \in R} y_i \Big/ \sum_{i \in R} x_i$$

Using above, the imputation-based estimator is:

$$\bar{y}_s = \frac{1}{n}\sum_{i \in S} y_{oi} = \bar{y}_r \left(\frac{\bar{x}_n}{\bar{x}_r}\right) = \bar{y}_{RAT}$$

Where:

$$\bar{y}_r = \frac{1}{r}\sum_{i \in R} y_i,\ \bar{x}_r = \frac{1}{r}\sum_{i \in R} x_i \ \text{and}\ \bar{x}_n = \frac{1}{n}\sum_{i \in S} x_i$$

The bias and mean square error of $\bar{y}_{RAT}$ is given by:

- $B(\bar{y}_{RAT}) = \bar{Y}\left(\frac{1}{r} - \frac{1}{n}\right)\left(C_x^2 - \rho C_y C_x\right)$
- $M(\bar{y}_{RAT}) = \left(\frac{1}{n} - \frac{1}{N}\right) S_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\left[S_Y^2 + R_1^2 S_X^2 - 2R_1 S_{XY}\right]$

Where:

$$R_1 = \frac{\overline{Y}}{\overline{X}}$$

Compromised method of imputation: Singh and Horn[5] suggested a compromised method of imputation. It based on using information from imputed values for the responding units in addition to non-responding units. In case of compromised imputation procedures, the data take the form:

$$y_{oi} = \begin{cases} (\alpha n / r) y_i + (1-\alpha)\hat{b} x_i & \text{if} \quad i \in R \\ (1-\alpha)\hat{b} x_i & \text{if} \quad i \in R^C \end{cases}$$

where, α is a suitably chosen constant, such that the resultant variance of the estimator is optimum. The imputation-based estimator, for this case, is:

$$\bar{y}_{COMP} = \left[\alpha \bar{y}_r + (1-\alpha)\bar{y}_r \frac{\bar{x}}{\bar{x}_r}\right]$$

The bias, mean square error and minimum mean square error at $\alpha = 1 - \rho\frac{C_Y}{C_X}$ of $\bar{y}_{COMP}$ are given by:

- $B(\bar{y}_{COMP}) = \overline{Y}(1-\alpha)\left(\frac{1}{r}-\frac{1}{n}\right)(C_X^2 - \rho C_Y C_X)$
- $M(\bar{y}_{COMP}) = \left\{\left(\frac{1}{n}-\frac{1}{N}\right)S_Y^2 + \left(\frac{1}{r}-\frac{1}{n}\right)\left[S_Y^2 + R_1^2 - 2R_1 S_{XY}\right]\right\} - \left(\frac{1}{r}-\frac{1}{n}\right)\alpha^2 \overline{Y}^2 C_X^2$
- $M(\bar{y}_{COMP})_{min} = \left[\left(\frac{1}{r}-\frac{1}{N}\right) - \left(\frac{1}{r}-\frac{1}{n}\right)\rho^2\right] S_Y^2$

Ahmed methods: For the case where, y_{ji} denotes the ith available observation for the jth imputation method, the three imputation methods y_{1i}, y_{2i} and y_{3i} are given as follows:

(1) $$y_{1i} = \begin{cases} y_i & \text{if} \quad i \in R \\ \frac{1}{(n-r)}\left[n\bar{y}_r\left(\frac{\overline{X}}{\bar{x}}\right)^{\beta_1} - r\bar{y}_r\right] & \text{if} \quad i \in R^C \end{cases}$$

where, β_1 is a suitably chosen constant, such that the variance of the resultant estimator is minimum. Under this method, point estimator of y_{1i} is:

$$t_1 = \bar{y}_r\left(\frac{\overline{X}}{\bar{x}}\right)^{\beta_1}$$

The bias, mean square error and minimum mean square error at $\beta_1 = \rho\frac{C_y}{C_x}$ of t_1 are given by:

- $B(t_1) = \overline{Y}\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{\beta_1(\beta_1+1)C_X^2}{2} - \beta_1\rho C_Y C_X\right)$
- $M(\bar{y}_{COMP}) = \left\{\left(\frac{1}{n}-\frac{1}{N}\right)S_Y^2 + \left(\frac{1}{r}-\frac{1}{n}\right)\left[S_Y^2 + R_1^2 - 2R_1 S_{XY}\right]\right\} - \left(\frac{1}{r}-\frac{1}{n}\right)\alpha^2 \overline{Y}^2 C_X^2$
- $M(\bar{y}_{COMP})_{min} = \left[\left(\frac{1}{r}-\frac{1}{N}\right) - \left(\frac{1}{r}-\frac{1}{n}\right)\rho^2\right] S_Y^2$

(2) $$y_{1i} = \begin{cases} y_i & \text{if} \quad i \in R \\ \frac{1}{(n-r)}\left[n\bar{y}_r\left(\frac{\overline{X}}{\bar{x}}\right)^{\beta_1} - r\bar{y}_r\right] & \text{if} \quad i \in R^C \end{cases}$$

where, β_2 is a suitably chosen constant, such that the variance of the resultant estimator is minimum. Under this method, the point estimator of y_{2i} is:

$$t_1 = \bar{y}_r\left(\frac{\overline{X}}{\bar{x}}\right)^{\beta_1}$$

The bias, mean square error and minimum mean square error at $\beta_2 = \rho\frac{C_Y}{C_X}$ of t_2 are given by:

- $B(t_2) = \left(\frac{1}{r}-\frac{1}{n}\right)\overline{Y}\left(\frac{\beta_2(\beta_2+1)}{2}C_X^2 - \beta_2\rho C_Y C_X\right)$
- $M(t_2) = \overline{Y}^2\left[\left(\frac{1}{r}-\frac{1}{N}\right)C_Y^2 + \beta_2^2\left(\frac{1}{r}-\frac{1}{n}\right)C_X^2 - 2\beta_2\left(\frac{1}{r}-\frac{1}{n}\right)\rho C_Y C_X\right]$
- $M(t_2)_{min} = \left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2 - \left(\frac{1}{r}-\frac{1}{n}\right)\frac{S_{XY}^2}{S_X^2}$

(3) $$y_{3i} = \begin{cases} y_i & \text{if} \quad i \in R \\ \frac{1}{(n-r)}\left[n\bar{y}_r\left(\frac{\overline{X}}{\bar{x}_r}\right)^{\beta_3} - r\bar{y}_r\right] & \text{if} \quad i \in R^C \end{cases}$$

where, β_3 is a suitably chosen constant, such that the variance of the resultant estimator is minimum. Under this method, the point estimator of y_{3i} is:

$$t_3 = \bar{y}_r\left(\frac{\overline{X}}{\bar{x}_r}\right)^{\beta_3}$$

The bias, mean square error and minimum mean square error at $\beta_3 = \rho\frac{C_Y}{C_X}$ of t_3 are given by:

- $B(t_3) = \left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}\left(\frac{\beta_2(\beta_2+1)}{2}C_X^2 - \beta_2\rho C_Y C_X\right)$
- $M(t_3) = \left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2\left[C_Y^2 + \beta_3^2 C_X^2 - 2\beta_3\rho C_Y C_X\right]$
- $M(t_3)_{min} = \left(\frac{1}{r} - \frac{1}{N}\right)S_Y^2(1-\rho^2)$

Singh and Gogoi method of imputation: For the case where, y_{ji} denotes the ith available observation for the jth imputation method, the imputation method y_{1i}, y_{2i}, y_{3i} is given as follows:

$$(1)\quad y_{1i} = \begin{cases} \alpha\frac{n}{r}y_i\frac{\bar{X}}{\bar{x}} + (1-\alpha)\bar{y}_r\frac{\bar{x}}{\bar{X}} & i \in R \\ (1-\alpha)\bar{y}_r\frac{\bar{x}}{\bar{X}} & i \in R^C \end{cases}$$

The point estimator of the population mean $\bar{Y}$ under proposed method of imputation is:

$$\bar{y}_{RPE1} = \bar{y}_r\left\{\alpha\frac{\bar{X}}{\bar{x}} + (1-\alpha)\frac{\bar{x}}{\bar{X}}\right\}$$

The bias, mean square error and minimum mean square error of $\bar{y}_{RPE1}$, respectively are given by:

- $B(\bar{y}_{RPE1}) = \bar{Y}\lambda_{n,N}\{\alpha + (1-2\alpha)\varphi_{YX}\}C_X^2$
- $MSE(\bar{y}_{RPE1}) = \bar{Y}^2[\lambda_{r,N}C_Y^2 + \lambda_{n,N}\{1 + 4\alpha(\alpha-1) + 2(1-2\alpha)\varphi_{YX}\}C_X^2]$
- $MSE(\bar{y}_{RPE1})_{min} = \bar{Y}^2\left(\lambda_{r,N} - \lambda_{n,N}\rho_{YX}^2\right)C_Y^2$

$$(2)\quad y_{2i} = \begin{cases} \alpha\frac{n}{r}y_i\frac{\bar{x}}{\bar{x}_r} + (1-\alpha)\bar{y}_r\frac{\bar{x}_r}{\bar{x}} & i \in R \\ (1-\alpha)\bar{y}_r\frac{\bar{x}_r}{\bar{x}} & i \in R^C \end{cases}$$

The point estimator of the population mean $\bar{Y}$ under proposed method of imputation is:

$$\bar{y}_{RPE2} = \bar{y}_r\left\{\alpha\frac{\bar{x}}{\bar{x}_r} + (1-\alpha)\frac{\bar{x}_r}{\bar{x}}\right\}$$

The bias, mean square error and minimum mean square error of $\bar{y}_{RPE2}$, respectively are given by:

- $B(\bar{y}_{RPE2}) = \bar{Y}\lambda_{r,n}\{\alpha + (1-2\alpha)\varphi_{YX}\}C_X^2$
- $MSE(\bar{y}_{RPE2}) = \bar{Y}^2[\lambda_{r,N}C_Y^2 + \lambda_{r,n}\{(1-2\alpha)^2 + 2(1-2\alpha)\varphi_{YX}\}C_X^2]$
- $MSE(\bar{y}_{RPE2})_{min} = \bar{Y}^2\left(\lambda_{r,N} - \lambda_{r,n}\rho_{YX}^2\right)C_Y^2$

$$(3)\quad y_{3i} = \begin{cases} \alpha\frac{n}{r}y_i\frac{\bar{X}}{\bar{x}_r} + (1-\alpha)\bar{y}_r\frac{\bar{x}_r}{\bar{X}} & i \in R \\ (1-\alpha)\bar{y}_r\frac{\bar{x}_r}{\bar{X}} & i \in R^C \end{cases}$$

The point estimator of the population mean $\bar{Y}$ under proposed method of imputation is:

$$\bar{y}_{RPE3} = \bar{y}_r\left\{\alpha\frac{\bar{X}}{\bar{x}_r} + (1-\alpha)\frac{\bar{x}_r}{\bar{X}}\right\}$$

The bias, mean square error and minimum mean square error of $\bar{y}_{RPE3}$, respectively are given by:

- $B(\bar{y}_{RPE3}) = \bar{Y}\lambda_{r,N}\{\alpha + (1-2\alpha)\varphi_{YX}\}C_X^2$
- $MSE(\bar{y}_{RPE3}) = \bar{Y}^2\lambda_{r,N}[C_Y^2 + \{1 + 4\alpha(\alpha-1) + 2(1-2\alpha)\varphi_{YX}\}C_X^2]$
- $MSE(\bar{y}_{RPE3})_{min} = \bar{Y}^2\lambda_{r,N}\left(1 - \rho_{YX}^2\right)C_Y^2$

MATERIALS AND METHODS

For estimating the population mean of the study variate y, Singh and Espejo[29] considered an estimator of the ratio-product type in two phase sampling given by:

$$\bar{y}_{RP} = \bar{y}\left\{k\frac{\bar{x}'}{\bar{x}} + (1-k)\frac{\bar{x}}{\bar{x}'}\right\}$$

where, $\bar{y}$ and $\bar{x}$ are the sample means of y and x, respectively based on a sample size n out of the population of N units. $\bar{x}'$ is the sample mean of sample size n' and k is suitably chosen constant.

Motivated by Singh and Espejo[29] ratio-product estimator of a finite population mean and for the case where y_{ji} denotes the ith available observation for the jth imputation method, we here proposed the following three ratio cum product type methods of imputation in two phase sampling:

(1) $$y_{1i} = \begin{cases} \alpha \dfrac{n}{r} y_i \dfrac{\bar{x}'}{\bar{x}} + (1-\alpha)\bar{y}_r \dfrac{\bar{x}}{\bar{x}'} & i \in R \\ (1-\alpha)\bar{y}_r \dfrac{\bar{x}}{\bar{x}'} & i \in R^C \end{cases}$$

where, α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Under this strategy, the point estimator of $\bar{Y}$ is:

$$\bar{y}^{(d)}_{IRP1} = \bar{y}_r\left[\alpha \frac{\bar{x}'}{\bar{x}} + (1-\alpha)\frac{\bar{x}}{\bar{x}'}\right] \tag{1}$$

(2) $$y_{2i} = \begin{cases} \alpha \dfrac{n}{r} y_i \dfrac{\bar{x}}{\bar{x}_r} + (1-\alpha)\bar{y}_r \dfrac{\bar{x}_r}{\bar{x}} & i \in R \\ (1-\alpha)\bar{y}_r \dfrac{\bar{x}_r}{\bar{x}} & i \in R^C \end{cases}$$

where, α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Under this strategy, the point estimator of $\bar{Y}$ is:

$$\bar{y}^{(d)}_{IRP2} = \bar{y}_r\left[\alpha \frac{\bar{x}}{\bar{x}_r} + (1-\alpha)\frac{\bar{x}_r}{\bar{x}}\right] \tag{2}$$

(3) $$y_{3i} = \begin{cases} \alpha \dfrac{n}{r} y_i \dfrac{\bar{x}'}{\bar{x}_r} + (1-\alpha)\bar{y}_r \dfrac{\bar{x}_r}{\bar{x}'} & i \in R \\ (1-\alpha)\bar{y}_r \dfrac{\bar{x}_r}{\bar{x}'} & i \in R^C \end{cases}$$

where, α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Under this strategy, the point estimator of $\bar{Y}$ is:

$$\bar{y}^{(d)}_{IRP3} = \bar{y}_r\left[\alpha \frac{\bar{x}'}{\bar{x}_r} + (1-\alpha)\frac{\bar{x}_r}{\bar{x}'}\right] \tag{3}$$

PROPERTIES OF PROPOSED ESTIMATORS

Let $B(.)_t$ and $M(.)_t$ denote the bias and Mean Square Error (MSE) of an estimator under the given sampling design t = I, II. The properties of $\bar{y}^{(d)}_{IRP1}, \bar{y}^{(d)}_{IRP2}$ and $\bar{y}^{(d)}_{IRP3}$ are derived in the following theorems respectively.

Theorem 1: Estimators $\bar{y}^{(d)}_{IRP1}, \bar{y}^{(d)}_{IRP2}$ and $\bar{y}^{(d)}_{IRP3}$ in terms of e_i; I = 1, 2, 3 and e_3' can be expressed upto first order of approximation as:

- $\bar{y}^{(d)}_{IRP1} = \bar{Y}\left[1 + e_1 + \left(e_3 - e_3' + e_3'^2 - e_3 e_3'\right) + \alpha\left(2e_3' - 2e_3 + e_3^2 - e_3'^2\right) + \left(e_1 e_3 - e_1 e_3'\right) + 2\alpha\left(e_1 e_3' - e_1 e_3\right)\right]$
- $\bar{y}^{(d)}_{IRP2} = \bar{Y}\left[1 + e_1 + \left(e_2 - e_3 + e_3^2 - e_2 e_3\right) + \alpha\left(2e_3 - 2e_2 + e_2^2 - e_3^2\right) + \left(e_1 e_2 - e_1 e_3'\right) + 2\alpha\left(e_1 e_3 - e_1 e_2\right)\right]$
- $\bar{y}^{(d)}_{IRPE3} = \bar{Y}\left[1 + e_1 + \left(e_2 - e_3' + e_3'^2 - e_2 e_3'\right) + \alpha\left(2e_3' - 2e_2 + e_2^2 - e_3'^2\right) + \left(e_1 e_2 - e_1 e_3'\right) + 2\alpha\left(e_1 e_3' - e_1 e_2\right)\right]$

by ignoring the terms $E\left[e_i^r, e_j^s\right], E\left[e_i^r, (e_j')^s\right]$ for r+s>2, where, r, s = 0.1, 2,... and i = 1, 2, 3; j = 2, 3 which is first approximation.

Proof:

- $$\begin{aligned} \bar{y}^{(d)}_{IRP1} &= \bar{y}_r\left[\alpha \frac{\bar{x}'}{\bar{x}} + (1-\alpha)\frac{\bar{x}}{\bar{x}'}\right] \\ &= \bar{Y}(1+e_1)\left[\alpha(1+e_3')(1+e_3)^{-1} + (1-\alpha)(1+e_3)(1+e_3')^{-1}\right] \\ &= \bar{Y}\left[1 + e_1 + \left(e_3 - e_3' + e_3'^2 - e_3 e_3'\right) + \alpha\left(2e_3' - 2e_3 + e_3^2 - e_3'^2\right) + \left(e_1 e_3 - e_1 e_3'\right) + 2\alpha\left(e_1 e_3' - e_1 e_3\right)\right] \end{aligned}$$
- $$\begin{aligned} \bar{y}^{(d)}_{IRP2} &= \bar{y}_r\left[\alpha \frac{\bar{x}}{\bar{x}_r} + (1-\alpha)\frac{\bar{x}_r}{\bar{x}}\right] \\ &= \bar{Y}(1+e_1)\left[\alpha(1+e_3)(1+e_2)^{-1} + (1-\alpha)(1+e_2)(1+e_3)^{-1}\right] \\ &= \bar{Y}\left[1 + e_1 + \left(e_2 - e_3 + e_3^2 - e_2 e_3\right) + \alpha\left(2e_3 - 2e_2 + e_2^2 - e_3^2\right) + \left(e_1 e_2 - e_1 e_3'\right) + 2\alpha\left(e_1 e_3 - e_1 e_2\right)\right] \end{aligned}$$

- $$\begin{aligned}\bar{y}_{\text{IRP3}}^{(d)} &= \bar{y}_r\left[\alpha\frac{\bar{x}'}{\bar{x}_r}+(1-\alpha)\frac{\bar{x}_r}{\bar{x}'}\right]\\ &= \bar{Y}(1+e_1)\left[\alpha\left(1+e_3'\right)(1+e_2)^{-1}+(1-\alpha)(1+e_2)\left(1+e_3'\right)^{-1}\right]\\ &= \bar{Y}\left[1+e_1+\left(e_2-e_3'+e_3'^2-e_2e_3'\right)+\alpha\left(2e_3'-2e_2+e_2^2-e_3'^2\right)+\left(e_1e_2-e_1e_3'\right)+2\alpha\left(e_1e_3'-e_1e_2\right)\right]\end{aligned}$$

Theorem 2: Biases of $\bar{y}_{\text{IRP1}}^{(d)}$, $\bar{y}_{\text{IRP2}}^{(d)}$ and $\bar{y}_{\text{IRP3}}^{(d)}$ under design I and II, upto first order of approximation are:

$$B\left(\bar{y}_{\text{IRP1}}^{(d)}\right)_{\text{I}} = \bar{Y}\left[(\delta_2-\delta_3)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{4}$$

$$B\left(\bar{y}_{\text{IRP1}}^{(d)}\right)_{\text{II}} = \bar{Y}\left[\left\{\delta_3-\alpha(\delta_3-\delta_5)+(1-2\alpha)\delta_5\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{5}$$

$$B\left(\bar{y}_{\text{IRP2}}^{(d)}\right)_{\text{I}} = \bar{Y}\left[(\delta_1-\delta_2)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{6}$$

$$B\left(\bar{y}_{\text{IRP2}}^{(d)}\right)_{\text{II}} = \bar{Y}\left[(\delta_4-\delta_5)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{7}$$

$$B\left(\bar{y}_{\text{IRP3}}^{(d)}\right)_{\text{I}} = \bar{Y}\left[(\delta_1-\delta_3)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{8}$$

$$B\left(\bar{y}_{\text{IRP3}}^{(d)}\right)_{\text{II}} = \bar{Y}\left[\left\{\delta_3-\alpha(\delta_3-\delta_4)+(1-2\alpha)\delta_4\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{9}$$

Proof:

- $$\begin{aligned}B\left(\bar{y}_{\text{IRP1}}^{(d)}\right)_{\text{I}} &= E\left[\bar{y}_{\text{IRP1}}^{(d)}-\bar{Y}\right]\\ &= \bar{Y}E\left[1+e_1+\left(e_3-e_3'+e_3'^2-e_3e_3'\right)+\alpha\left(2e_3'-2e_3+e_3^2-e_3'^2\right)+\left(e_1e_3-e_1e_3'\right)+2\alpha\left(e_1e_3'-e_1e_3\right)-1\right]\\ &= \bar{Y}\left[(\delta_2-\delta_3)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right]\end{aligned}$$

- $$\begin{aligned}B\left(\bar{y}_{\text{IRP1}}^{(d)}\right)_{\text{II}} &= E\left[\bar{y}_{\text{IRP1}}^{(d)}-\bar{Y}\right]\\ &= \bar{Y}E\left[1+e_1+\left(e_3-e_3'+e_3'^2-e_3e_3'\right)+\alpha\left(2e_3'-2e_3+e_3^2-e_3'^2\right)+\left(e_1e_3-e_1e_3'\right)+2\alpha\left(e_1e_3'-e_1e_3\right)-1\right]\\ &= \bar{Y}\left[\left\{\delta_3-\alpha(\delta_3-\delta_5)+(1-2\alpha)\delta_5\rho\frac{C_Y}{C_X}\right\}C_X^2\right]\end{aligned}$$

- $$\begin{aligned}B\left(\bar{y}_{\text{IRP2}}^{(d)}\right)_{\text{I}} &= E\left[\bar{y}_{\text{IRP2}}^{(d)}-\bar{Y}\right]\\ &= \bar{Y}E\left[1+e_1+\left(e_2-e_3+e_3^2-e_2e_3\right)+\alpha\left(2e_3-2e_2+e_2^2-e_3^2\right)+\left(e_1e_2-e_1e_3'\right)+2\alpha\left(e_1e_3-e_1e_2\right)-1\right]\\ &= \bar{Y}\left[(\delta_1-\delta_2)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right]\end{aligned}$$

- $$\begin{aligned}B\left(\bar{y}_{\text{IRP2}}^{(d)}\right)_{\text{II}} &= E\left[\bar{y}_{\text{IRP2}}^{(d)}-\bar{Y}\right]\\ &= \bar{Y}E\left[1+e_1+\left(e_2-e_3+e_3^2-e_2e_3\right)+\alpha\left(2e_3-2e_2+e_2^2-e_3^2\right)+\left(e_1e_2-e_1e_3'\right)+2\alpha\left(e_1e_3-e_1e_2\right)-1\right]\\ &= \bar{Y}\left[(\delta_4-\delta_5)\left\{\alpha+(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right]\end{aligned}$$

- $$\begin{aligned} B\left(\overline{y}^{(d)}_{IRP3}\right)_{I} &= E\left[\overline{y}^{(d)}_{IRP3} - \overline{Y}\right] \\ &= \overline{Y}E\left[1 + e_1 + \left(e_2 - e_3^{'} + e_3^{'2} - e_2 e_3^{'}\right) + \alpha\left(2e_3^{'} - 2e_2 + e_2^2 - e_3^{'2}\right) + \left(e_1 e_2 - e_1 e_3^{'}\right) + 2\alpha\left(e_1 e_3^{'} - e_1 e_2\right) - 1\right] \\ &= \overline{Y}\left[(\delta_1 - \delta_3)\left\{\alpha + (1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} B\left(\overline{y}^{(d)}_{IRP3}\right)_{II} &= E\left[\overline{y}^{(d)}_{IRP3} - \overline{Y}\right] \\ &= \overline{Y}E\left[1 + e_1 + \left(e_2 - e_3^{'} + e_3^{'2} - e_2 e_3^{'}\right) + \alpha\left(2e_3^{'} - 2e_2 + e_2^2 - e_3^{'2}\right) + \left(e_1 e_2 - e_1 e_3^{'}\right) + 2\alpha\left(e_1 e_3^{'} - e_1 e_2\right) - 1\right] \\ &= \overline{Y}\left[\left\{\delta_3 - \alpha(\delta_3 - \delta_4) + (1-2\alpha)\delta_4\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

Theorem 3: Mean squared errors of $\overline{y}^{(d)}_{IRP1}$, $\overline{y}^{(d)}_{IRP2}$ and $\overline{y}^{(d)}_{IRP3}$ under the design I and II, upto first order of approximation can be written as:

$$M(\overline{y}^{(d)}_{IRP1})_I = \overline{Y}^2\left[\delta_1 C_y^2 + (\delta_2 - \delta_3)\left\{4\alpha^2 + (1-4\alpha) + 2(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{10}$$

$$M(\overline{y}^{(d)}_{IRP1})_{II} = \overline{Y}^2\left[\delta_4 C_Y^2 + \left\{(\delta_3 + \delta_5)\left(1 + 4\alpha^2 - 4\alpha\right) + 2(1-2\alpha)\delta_5\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{11}$$

$$M(\overline{y}^{(d)}_{IRP2})_I = \overline{Y}^2\left[\delta_1 C_Y^2 + (\delta_1 - \delta_2)\left\{4\alpha^2 + (1-4\alpha) + 2(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{12}$$

$$M(\overline{y}^{(d)}_{IRP2})_{II} = \overline{Y}^2\left[\delta_4 C_Y^2 + (\delta_4 - \delta_5)\left\{4\alpha^2 + (1-4\alpha) + 2(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{13}$$

$$M(\overline{y}^{(d)}_{IRP3})_I = \overline{Y}^2\left[\delta_1 C_Y^2 + (\delta_1 - \delta_3)\left\{4\alpha^2 + (1-4\alpha) + 2(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{14}$$

$$M(\overline{y}^{(d)}_{IRP3})_{II} = \overline{Y}^2\left[\delta_4 C_Y^2 + \left\{(\delta_3 + \delta_4)\left(1 + 4\alpha^2 - 4\alpha\right) + 2(1-2\alpha)\delta_4\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \tag{15}$$

Proof:

- $$\begin{aligned} M(\overline{y}^{(d)}_{IRP1})_I &= E\left[\overline{y}^{(d)}_{IRP1} - \overline{Y}\right]^2 = \overline{Y}^2 E\left[e_1 + \left(e_3 - e_3^{'}\right) + 2\alpha\left(e_3^{'} - e_3\right)\right]^2 \\ &= \overline{Y}^2 E\left[e_1^2 + \left(e_3^2 + e_3^{'2} - 2e_3 e_3^{'} + 2e_1 e_3 - 2e_1 e_3^{'}\right) + 4\alpha^2\left(e_3^{'2} + e_3^2 - 2e_3 e_3^{'}\right) + 4\alpha\left(e_3 e_3^{'} - e_3^2 - e_3^{'2} + e_3 e_3^{'} + e_1 e_3^{'} - e_1 e_3\right)\right] \\ &= \overline{Y}^2\left[\delta_1 C_y^2 + (\delta_2 - \delta_3)\left\{4\alpha^2 + (1-4\alpha) + 2(1-2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} M(\overline{y}^{(d)}_{IRP1})_{II} &= E\left[\overline{y}^{(d)}_{IRP1} - \overline{Y}\right]^2 = \overline{Y}^2 E\left[e_1 + \left(e_3 - e_3^{'}\right) + 2\alpha\left(e_3^{'} - e_3\right)\right]^2 \\ &= \overline{Y}^2 E\left[e_1^2 + \left(e_3^2 + e_3^{'2} - 2e_3 e_3^{'} + 2e_1 e_3 - 2e_1 e_3^{'}\right) + 4\alpha^2\left(e_3^{'2} + e_3^2 - 2e_3 e_3^{'}\right) + 4\alpha\left(e_3 e_3^{'} - e_3^2 - e_3^{'2} + e_3 e_3^{'} + e_1 e_3^{'} - e_1 e_3\right)\right] \\ &= \overline{Y}^2\left[\delta_4 C_Y^2 + \left\{(\delta_3 + \delta_5)\left(1 + 4\alpha^2 - 4\alpha\right) + 2(1-2\alpha)\delta_5\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} M\left(\bar{y}^{(d)}_{IRP2}\right)_I &= E\left[\bar{y}^{(d)}_{IRP2} - \bar{Y}\right]^2 = \bar{Y}^2 E\left[e_1 + \left(e_2 - e_3\right) + 2\alpha\left(e_3 - e_2\right)\right]^2 \\ &= \bar{Y}^2 E\left[e_1^2 + \left(e_2^2 + e_3^2 - 2e_2e_3 + 2e_1e_2 - 2e_1e_3\right) + 4\alpha^2\left(e_3^2 + e_2^2 - 2e_2e_3\right) + 4\alpha\left(e_2e_3 - e_2^2 - e_3^2 + e_2e_3 + e_1e_3 - e_1e_2\right)\right] \\ &= \bar{Y}^2\left[\delta_1 C_Y^2 + (\delta_1 - \delta_2)\left\{4\alpha^2 + (1 - 4\alpha) + 2(1 - 2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} M\left(\bar{y}^{(d)}_{IRP2}\right)_{II} &= E\left[\bar{y}^{(d)}_{IRP2} - \bar{Y}\right]^2 = \bar{Y}^2 E\left[e_1 + \left(e_2 - e_3\right) + 2\alpha\left(e_3 - e_2\right)\right]^2 \\ &= \bar{Y}^2 E\left[e_1^2 + \left(e_2^2 + e_3^2 - 2e_2e_3 + 2e_1e_2 - 2e_1e_3\right) + 4\alpha^2\left(e_3^2 + e_2^2 - 2e_2e_3\right) + 4\alpha\left(e_2e_3 - e_2^2 - e_3^2 + e_2e_3 + e_1e_3 - e_1e_2\right)\right] \\ &= \bar{Y}^2\left[\delta_4 C_Y^2 + (\delta_4 - \delta_5)\left\{4\alpha^2 + (1 - 4\alpha) + 2(1 - 2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} M\left(\bar{y}^{(d)}_{IRP3}\right)_I &= \bar{Y}^2 E\left[\bar{y}^{(d)}_{IRP3} - \bar{Y}\right]^2 = \bar{Y}^2 E\left[e_1 + \left(e_2 - e_3'\right) + 2\alpha\left(e_3' - e_2\right)\right]^2 \\ &= \bar{Y}^2 E\left[e_1^2 + \left(e_2^2 + e_3'^2 - 2e_2e_3' + 2e_1e_2 - 2e_1e_3'\right) + 4\alpha^2\left(e_3'^2 + e_2^2 - 2e_2e_3'\right) + 4\alpha\left(e_2e_3' - e_2^2 - e_3'^2 + e_2e_3' + e_1e_3' - e_1e_2\right)\right] \\ &= \bar{Y}^2\left[\delta_1 C_Y^2 + (\delta_1 - \delta_3)\left\{4\alpha^2 + (1 - 4\alpha) + 2(1 - 2\alpha)\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

- $$\begin{aligned} M\left(\bar{y}^{(d)}_{IRP3}\right)_{II} &= \bar{Y}^2 E\left[\bar{y}^{(d)}_{IRP3} - \bar{Y}\right]^2 = \bar{Y}^2 E\left[e_1 + \left(e_2 - e_3'\right) + 2\alpha\left(e_3' - e_2\right)\right]^2 \\ &= \bar{Y}^2 E\left[e_1^2 + \left(e_2^2 + e_3'^2 - 2e_2e_3' + 2e_1e_2 - 2e_1e_3'\right) + 4\alpha^2\left(e_3'^2 + e_2^2 - 2e_2e_3'\right) + 4\alpha\left(e_2e_3' - e_2^2 - e_3'^2 + e_2e_3' + e_1e_3' - e_1e_2\right)\right] \\ &= \bar{Y}^2\left[\delta_4 C_Y^2 + \left\{(\delta_3 + \delta_4)\left(1 + 4\alpha^2 - 4\alpha\right) + 2(1 - 2\alpha)\delta_4\rho\frac{C_Y}{C_X}\right\}C_X^2\right] \end{aligned}$$

Theorem 4: Minimum mean squared error of $\bar{y}^{(d)}_{IRP1}$, $\bar{y}^{(d)}_{IRP2}$ and $\bar{y}^{(d)}_{IRP3}$ under the design I and II are given as:

- $\left[M(\bar{y}^{(d)}_{IRP1})_I\right]_{min} = \left[\delta_1 - (\delta_2 - \delta_3)\rho^2\right]S_Y^2$
- $\left[M(\bar{y}^{(d)}_{IRP1})_{II}\right]_{min} = \left[\delta_4 - \delta_5\rho^2(\delta_3 + \delta_5)^{-1}\right]S_Y^2$
- $\left[M(\bar{y}^{(d)}_{IRP2})_I\right]_{min} = \left[\delta_1 - (\delta_1 - \delta_2)\rho^2\right]S_Y^2$
- $\left[M(\bar{y}^{(d)}_{IRP2})_{II}\right]_{min} = \left[\delta_5\rho^2 + \delta_4\left(1 - \rho^2\right)\right]S_Y^2$
- $\left[M(\bar{y}^{(d)}_{IRP3})_I\right]_{min} = \left[(1 - \rho^2)\delta_1 + \delta_3\rho^2\right]S_Y^2$
- $\left[M(\bar{y}^{(d)}_{IRP3})_{II}\right]_{min} = \left[\left\{1 - \delta_4\rho^2(\delta_3 + \delta_4)^{-1}\right\}\delta_4\right]S_Y^2$

Proof:

- Differentiating (Eq. 10) with respect to α and equating it to zero, we get:

$$\frac{d}{d\alpha}\left[M\left(\bar{y}^{(d)}_{IRP1}\right)_I\right] = 0 \Rightarrow \alpha = \frac{1}{2}\left(1 + \rho\frac{C_Y}{C_X}\right)$$

Putting the value of α in Eq. 10, we obtain:

$$\left[M(\bar{y}^{(d)}_{IRP1})_I\right]_{min} = \left[\delta_1 - (\delta_2 - \delta_3)\rho^2\right]S_Y^2$$

- Similarly proceeding for Eq. 11-15, we have:

$$\alpha = \frac{1}{2}\left(1 + \frac{\delta_5}{\delta_3 + \delta_5}\rho\frac{C_Y}{C_X}\right)$$
$$\left[M(\bar{y}^{(d)}_{IRP1})_{II}\right]_{min} = \left[\delta_4 - \delta_5\rho^2(\delta_3 + \delta_5)^{-1}\right]S_Y^2$$

- $$\alpha = \frac{1}{2}\left(1 + \rho\frac{C_Y}{C_X}\right)$$
$$\left[M(\bar{y}^{(d)}_{IRP2})_{II}\right]_{min} = \left[\delta_4 - \delta_5\rho^2(\delta_3 + \delta_5)^{-1}\right]S_Y^2$$

- $$\alpha = \frac{1}{2}\left(1 + \rho\frac{C_Y}{C_X}\right)$$
$$\left[M(\bar{y}^{(d)}_{IRP2})_{II}\right]_{min} = \left[\delta_5\rho^2 + \delta_4\left(1 - \rho^2\right)\right]S_Y^2$$

- $$\alpha = \frac{1}{2}\left(1 + \rho\frac{C_Y}{C_X}\right)$$
$$\left[M(\bar{y}^{(d)}_{IRP3})_I\right]_{min} = \left[(1 - \rho^2)\delta_1 + \delta_3\rho^2\right]S_Y^2$$

- $$\alpha = \frac{1}{2}\left(1 + \frac{\delta_4}{\delta_3 + \delta_4}\rho\frac{C_Y}{C_X}\right)$$
  $$\left[M(\bar{y}^{(d)}_{IRP3})_{II}\right]_{min} = \left[\left\{1 - \delta_4\rho^2(\delta_3 + \delta_4)^{-1}\right\}\delta_4\right]S_Y^2$$

COMPARISON

It derived the conditions under which the suggested estimators are superior to Ahmed *et al.*[8] estimators in design I and II:

- $$\Delta_1 = \left[M(t_1)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP1}\right)_I\right]_{min} = \left[\frac{1}{n'} - \frac{1}{N}\right]S_Y^2 - 2\left[\frac{1}{n'} - \frac{1}{N}\right]\rho^2 S_Y^2$$

$\left(\bar{y}^{(d)}_{IRP1}\right)_I$ is better than t_1 when:

$$\Delta_1 > 0 \Rightarrow -\frac{1}{\sqrt{2}} < \rho < \frac{1}{\sqrt{2}}$$

- $$\Delta_2 = \left[M(t_1)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP1}\right)_{II}\right]_{min} = [\delta_6 - \delta_4]S_Y^2 - \left[\frac{\delta_7(\delta_3 + \delta_5) - \delta_5^2}{(\delta_3 + \delta_5)}\right]\rho^2 S_Y^2$$

$\left(\bar{y}^{(d)}_{IRP1}\right)_{II}$ is better than t_1 when:

$$\Delta_2 > 0 \Rightarrow \rho^2 < \frac{(\delta_6 - \delta_4)(\delta_3 + \delta_5)}{\left[\delta_7(\delta_3 + \delta_5) - \delta_5^2\right]} \Rightarrow -\sqrt{P} < \rho < \sqrt{P}$$

Where:

$$P = \frac{(\delta_6 - \delta_4)(\delta_3 + \delta_5)}{\left[\delta_7(\delta_3 + \delta_5) - \delta_5^2\right]}; \delta_6 = \left(\frac{1}{r} - \frac{1}{N}\right); \delta_7 = \left(\frac{1}{n} - \frac{1}{N}\right)$$

- $$\Delta_3 = \left[M(t_2)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP2}\right)_I\right]_{min} = \left(\frac{1}{n'} - \frac{1}{N}\right)S_Y^2 > 0$$

which is always true.

- $$\Delta_4 = \left[M(t_2)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP2}\right)_{II}\right]_{min} = \left[\frac{1}{N - n'} - \frac{1}{N}\right]S_Y^2 > 0$$

which is always true.

- $$\Delta_5 = \left[M(t_3)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP3}\right)_I\right]_{min} = \left[\frac{1}{n'} - \frac{1}{N}\right]S_Y^2 - 2\left[\frac{1}{n'} - \frac{1}{N}\right]\rho^2 S_Y^2$$

$\left(\bar{y}^{(d)}_{IRP3}\right)_I$ is better than t_3 when:

$$\Delta_5 > 0 \Rightarrow -\frac{1}{\sqrt{2}} < \rho < \frac{1}{\sqrt{2}}$$

- $$\Delta_6 = \left[M(t_3)\right]_{min} - \left[M\left(\bar{y}^{(d)}_{IRP3}\right)_{II}\right]_{min} = [\delta_6 - \delta_4]S_y^2 - \left[\delta_6 - (\delta_3 + \delta_5)^{-1}\delta_4^2\right]\rho^2 S_Y^2$$

$\left(\bar{y}^{(d)}_{IRP3}\right)_{II}$ is better than t_3 when:

$$\Delta_6 > 0 \Rightarrow \rho^2 < \frac{(\delta_6 - \delta_4)(\delta_3 + \delta_5)}{\left[\delta_6(\delta_3 + \delta_4) - \delta_4^2\right]} \Rightarrow -\sqrt{Q} < \rho < \sqrt{Q}$$

Where:

$$Q = \frac{(\delta_6 - \delta_4)(\delta_3 + \delta_5)}{\left[\delta_6(\delta_3 + \delta_4) - \delta_4^2\right]}; \delta_6 = \left(\frac{1}{r} - \frac{1}{N}\right)$$

NUMERICAL ILLUSTRATIONS

It consider two populations A and B, first one is artificial population of size[15,16] N = 200 and another one is from Ahmed *et al.*[8] with the following parameters as given in Table 1.

Let there be n′ = 60, n = 40, r = 5 for population A and n′ = 2000, n = 500, r = 450 for population B, then the bias and the MSE of the suggested estimators under design I and II and that of Ahmed *et al.*[8] methods for population A and B respectively are given in Table 2-4. The sampling efficiency of suggested estimators under design I and II over Ahmed *et al.*[8] is defined as:

$$E_i = \frac{\text{Opt}\left[M\left(\bar{y}^{(d)}_{IRPi}\right)_j\right]}{\text{Opt}\left[M(T_i)\right]}; \quad i = 1,2,3; \quad j = I, II, \quad T = t, \bar{y}_{RPE}$$

RESULTS

The present work proposed some imputed estimators for estimating population mean in two phase sampling under two different designs. The study of the estimators had been made taking two different population data sets (Table 1) for their efficiencies. The results of the proposed estimators certainly showed the superiority of it over the traditional estimators viz., the mean method of imputation, ratio method of imputation, compromised method of imputation and feel better or little bit inferior depending on sampling design and population data sets over Singh and Gogoi and three Ahmed methods of

Table 1: Parameters of population A and B

Population	N	$\bar{Y}$	$\bar{X}$	S_Y^2	S_X^2	ρ	C_X	C_Y
A	200	42.48518	18.515	199.0598	48.5375	0.8652	0.3763	0.3321
B	8306	253.75	343.316	338006	862017	0.522231	2.70436	2.29116

Table 2: Bias and MSE of suggested estimators for population A

	Design I		Design II	
Estimators	Bias	MSE	Bias	MSE
$\bar{y}_{IRP1}^{(d)}$	-0.072618419	36.990998	0.043315139	36.780688
$\bar{y}_{IRP2}^{(d)}$	-0.489371285	10.417476	-0.489371285	12.313284
$\bar{y}_{IRP3}^{(d)}$	0.308521799	10.914177	0.369750827	11.291671

Table 3: Bias and MSE of suggested estimators for population B

	Design I		Design II	
Estimators	Bias	MSE	Bias	MSE
$\bar{y}_{IRP1}^{(d)}$	0.398611502	478.8387296	1.874647334	556.7868378
$\bar{y}_{IRP2}^{(d)}$	0.034241055	561.634485	0.034241055	677.0363965
$\bar{y}_{IRP3}^{(d)}$	1.309280875	458.3537412	2.2178586	536.845824

Table 4: Bias and MSE for population A and B for estimators $\bar{y}_m$, $\bar{y}_{RAT}$, $\bar{y}_{COMP}$

	Population A		Population B	
Estimators	Bias	MSE	Bias	MSE
$\bar{y}_m$	0	38.81893	0	710.4302
$\bar{y}_{RAT}$	0.248907	15.24075	0.22994	768.7752
$\bar{y}_{COMP}$	0.190058	12.74059	0.050411	689.9429
t_1	0.010856	35.83645	0.43025	537.1631
t_2	0.094991	12.73984	0.050868	689.9452
t_3	0.105847	9.759633	0.481117	516.678
$\bar{y}_{RPE1}$	0.253841	35.83855	0.250621	537.16124
$\bar{y}_{RPE2}$	2.221107	12.74058	0.029630	689.94285
$\bar{y}_{RPE3}$	2.474948	9.760204	0.280251	516.67625

Source: Ahmed *et al.*[8] and Singh and Gogoi[27]

Table 5: Efficiency of suggested estimators for population A and B

	Population A		Population B	
Efficiency	Design I	Design II	Design I	Design II
E_1	1.032217	1.026349	0.891421	1.036532
E_2	0.817708	0.966518	0.814027	0.981290
E_3	1.118297	1.156977	0.887116	1.039033
E_4	1.032156	1.026348	0.891424	1.036535
E_5	0.817661	0.966461	0.814030	0.981293
E_6	1.118232	1.156909	0.887229	1.039037

Source: Ahmed *et al.*[8] and Singh and Gogoi[27]

imputation (Table 2-4) as far as MSEs are concerned. The Table 5 gives the relative efficiency of the proposed estimators. All the three proposed estimators have been found to be better in their performances over the estimators proposed by Singh and Gogoi and Ahmed's estimators in design I of population B. The second proposed estimator is showing better efficiency in both the designs of the population A and population B and other proposed estimators are either little bit inferior or having close proximity with estimators under comparisons.

DISCUSSION

The ratio method of estimation (or the product method of estimation[31,32], yields a more efficient estimator than the simple unbiased estimator provided the correlation coefficient between study variate y and auxiliary variate x has high positive value (or high negative value). Further, the ratio estimator is most effective and is as efficient as the regression estimator, when the relationship between the study variate, y and the auxiliary variate, x, is linear through the origin and the variate of y is proportional to x. However, in many practical situations, the line does not pass through the vicinity of the origin. Keeping this deed in view, an attempt has been made to improve these estimators in compromised imputation taking three sets of linear combination of ratio and product estimators when there is non-response either on study variable or auxiliary variable or both in two-phase sampling. Kalton *et al.*[1] and Sande[2] were the first to discussed imputation technique that make an incomplete data set structurally complete and its analysis simple. Lee *et al.*[3,4] used the information on an auxiliary variate for the purpose of imputation. Later Singh and Horn [5] suggested a compromised method of imputation. Ahmed *et al.*[8] suggested several new imputation based estimators that use the information on an auxiliary variate.

CONCLUSION

The present work proposed some imputed estimators for estimating population mean in two phase sampling scheme in sample survey. The results demonstrated the superiority of these proposed estimators over the traditional estimators' viz., the mean method of imputation, ratio method of imputation, compromised method of imputation and the second proposed estimator is showing better efficiency in both the designs of the population A and population B.

SIGNIFICANCE STATEMENT

This study discovers that the estimators under considerations are better in comparison to traditional estimators viz., the mean method of imputation, ratio method of imputation, compromised method of imputation and better than Singh and Gogoi and three Ahmed methods of imputation depending on sampling design in estimating the population mean. This study will help the researcher to uncover the area of formulation of imputation methods experimenting different sampling designs.

REFERENCES

1. Kalton, G., D. Kasprzyk and R. Santos, 1981. Issues of Non-Response and Imputation of Income and Participation. In: Current Topics in Survey Sampling, Krevoski, D., R. Platek and J.N.K. Rao (Eds.)., Acadmic Press, New York, USA., pp: 455-480.
2. Sande, I.G., 1979. A personal view of hot deck approach to automatic edit and imputation. J. Imput. Procedures Surv. Methodol., 5: 238-246.
3. Lee, H., E. Rancourt and C.E. Sarndal, 1994. Experiments with variance estimation from survey data with imputed values. J. Official Stat., 10: 231-243.
4. Lee, H., E. Rancourt and C.E. Sarndal, 1995. Variance estimation in the presence of imputed data for the generalized estimation system. Proceedings of the Section on Survey Research Methods, August 13-17, 1995, American Statistical Association, pp: 384-389.
5. Singh, S. and S. Horn, 2000. Compromised imputation in survey sampling. Metrika, 51: 267-276.
6. Shukla, D., 2002. FT estimator under two-phase sampling. Metron, 59: 253-263.
7. Singh, S. and B. Deo, 2003. Imputation by power transformation. Stat. Papers, 44: 555-579.
8. Ahmed, M.S., O. Al-Titi, Z. Al-Rawi and W. Abu-Dayyeh, 2006. Estimation of a population mean using different imputation methods. Stat. Trans., 7: 1247-1264.
9. Arnab, R. and S. Singh, 2006. A new method for estimating variance from data imputed with ratio method of imputation. Stat. Prob. Lett., 76: 513-519.
10. Rueda, M. and S. Gonzalez, 2008. A new ratio-type imputation with random disturbance. Applied Math. Lett., 21: 978-982.
11. Gonzalez, S., M. Rueda and A. Arcos, 2008. An improved estimator to analyse missing data. Stat. Papers, Vol. 49. 10.1007/s00362-007-0045-8.
12. Bouza, C.N., 2013. Estimation of the population mean wit missing observations using product type estimators. Invest. Operacional, 29: 207-222.
13. Kadilar, C. and H. Cingi, 2008. Estimators for the population mean in the case of missing data. Commun. Stat.: Theory Methods, 37: 2226-2236.
14. Shukla, D. and N.S. Thakur, 2008. Estimation of mean with imputation of missing data using factor-type estimator. Stat. Trans., 9: 33-48.
15. Shukla, D., N.S. Thakur, S. Pathak and D.S. Rajput, 2009. Estimation of mean under imputation of missing data using factor-type estimator in two-phase sampling. Stat. Trans., 10: 397-414.
16. Shukla, D., N.S. Thakur and D.S. Thakur, 2009. Utilization of non-response auxiliary population mean in imputation for missing observations. J. Reliab. Stat. Stud., 2: 28-40.
17. Baraldi, A.N. and C.K. Enders, 2010. An introduction to modern missing data analyses. J. School Psychol., 48: 58-37.
18. Diana, G. and P.F. Perri, 2010. Improved estimators of the population mean for missing data. Commun. Stat.-Theory Meth., 39: 3245-3251.
19. Singh, G.N., K. Priyanka, J.M. Kim and S. Singh, 2010. Estimation of population mean using imputation techniques in sample surveys. J. Korean Statistical Soc., 39: 67-74.
20. Shukla, D., N.S. Thakur, D.S. Thakur and S. Pathak, 2011. Linear combination based imputation method for missing data in sample. Int. J. Modern Eng. Res., 1: 580-596.

21. Thakur, N.S., K. Yadav and S. Pathak, 2011. Estimation of mean in presence of missing data under two-phase sampling scheme. J. Reliab. Stat. Stud., 4: 93-104.
22. Shukla, D., N.S. Thakur and S. Pathak, 2013. Some new aspects on imputation in sampling. Afr. J. Math. Comput. Sci. Res., 6: 5-15.
23. Thakur, N.S., K. Yadav and S. Pathak, 2013. On mean estimation with imputation in two-phase sampling design Res. J. Math. Stat. Sci., 1: 1-9.
24. Singh, A.K., P. Singh and V.K. Singh, 2014. Imputation methods of missing data for estimating the population mean using simple random sampling. Global J. Adv. Res., 1: 253-263.
25. Singh, P., A.K. Singh and V.K. Singh, 2015. On the use of compromised imputation for missing data using factor-type estimators. J. Stat. Applic. Prob. Lett., 24: 105-113.
26. Thakur, N.S., K. Yadav and R. Singh, 2016. A family of factor-type estimators for estimation of population variance. J. Reliab. Stat. Stud., 9: 71-79.
27. Singh, B.K. and U. Gogoi, 2018. Estimation of population mean using ratio cum product imputation techniques in sample survey. Res. Rev.: J. Stat., 7: 38-49.
28. Nath, K. and B.K. Singh, 2018. Estimation of population mean using imputation techniques in sample survey under two-phase sampling scheme. Int. J. Math. Stat., 19: 26-46.
29. Singh, H.P. and M.R. Espejo, 2007. Double sampling ratio-product estimator of a finite population mean in sample surveys. J. Applied Stat., 34: 71-85.
30. Rao, J.N. and R.R. Sitter, 1995. Variance estimation under two-phase sampling with application to imputation for missing data. Biometrika, 82: 453-460.
31. Robson, D.S., 1957. Applications of multivariate polykays to the theory of unbiased ratio-type estimation. J. Am. Stat. Assoc., 52: 511-522.
32. Murthy, M.N., 1964. Product method of estimation. Sankhya: Indian J. Stat. Ser. A, 26: 69-74.