Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 6): pp. 1096-1113
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.06.116

Study Of Some General Classes Of Estımator For Estımatıng Populatıon Mean In Compromısed Imputatıon Under Two-Phase Samplıng Scheme

Krishnajyothi Nath, "Department of Mathematics, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh - 791109, India, ORCID:0000-0003-3601-7106 krishnath10@gmail.com,

B.K. Singh, Department of Mathematics, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh - 791109, India, ORCID:0000-0002-7463-8862 bksinghstat@gmail.com ,"

Abstract. In this paper, authors have proposed some general classes of estimators for estimating population mean in compromised imputation under the framework of two-phase sampling design in presence of missing values. Two different sampling designs in two-phase sampling are compared under imputed data and their Biases, mean square error (MSE) expressions and percentage relative efficiency (PRE) are obtained. Further, theoretical results stating superiority of the proposed estimators, over the existing estimators have been verified through empirical illustrations based on different data sets from the classical statistical literature.

Keywords: Imputation methods, Bias, Mean Square error, Missing data, Non-response, Simple Random Sampling without Replacement (SRSWOR), Efficiency

1. INTRODUCTION

Missing data occurs frequently in almost every data collection activity and imputation method is used in replacing missing values in the sample with a reasonable set of values. Kalton et al. (1981) and Sande (1979) suggested imputation methods that make an incomplete data set structurally complete and its analysis simple. Lee et al. (1994, 1995) used the information on an auxiliary variable for the purpose of imputation. Later Singh and Horn (2000) introduced a compromised method of imputation based on auxiliary variables. Ahmed et al. (2006) discussed several new imputation based estimators that used the information on an auxiliary variate and compared their performance with the mean method of imputation. For more about missing data and imputation based methods one can refer to Shukla and Thakur (2008), Rueda and Gonzalez (2008), Kadilar and Cingi (2008), Shukla et al. (2009), Shukla et al. (2009a), Singh et al.(2010), Diana and Francesco Perri (2010), Baraldi and Enders (2010), Thakur et al. (2011), Shukla et al. (2011), Thakur et al. (2013), Shukla et al.(2013), Singh et al. (2014).

Recently, Singh et al. (2015, 2016,2017), Thakur et al. (2016), B. K. Singh et al. (2017) discussed some imputation methods of missing data for estimating the population mean using two-phase sampling scheme. The objective of the present research work is to derive some imputation methods for mean estimation and to carry out numerical illustrations to show their efficieny.

2. NOTATIONS

Let $U = (U_1, U_2, U_3, ..., U_N)$ be a finite population of size N and character under study be y. 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 as: a sub-sample from sample S' (denoted by design I) or as independent to sample S' (denoted by design II) 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 i^{th} 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 $\overline{x} = (n)^{-1}\sum_{i=1}^{n} x_i$ and $\overline{x}' = (n')^{-1}\sum_{i=1}^{n'} x_i$ respectively.

Remark 1: $\overline{y}_r = \overline{Y}(1+e_1)$; $\overline{x}_r = \overline{X}(1+e_2)$; $\overline{x} = \overline{X}(1+e_3)$; 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 (1995) and the mechanism of Missing Completely at Random (MCAR), for given r, n and n', we have:

(i) Under design F₁ [Case I]:

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

(ii) Under design F₂ [Case II]:

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

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

$M_1 = \left(\frac{1}{r} - \frac{1}{N}\right)$, $M_2 = \left(\frac{1}{n} - \frac{1}{N}\right)$, $M_3 = M_1 - M_2$, $f_1 = \frac{r}{n}$, $\lambda = \frac{a\overline{x}}{a\overline{x} + b}$.

3. REVIEWING EXISTING IMPUTATION METHODS AND CORRESPONDING ESTIMATORS

Let $\overline{Y} = N^{-1}\sum_{i=1}^{N} y_i$ 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 i^{th} 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:

3.1. 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 $y_{oi} = \begin{cases} y_i & if \quad i \in R \\ \overline{y_r} & if \quad i \in R^C \end{cases}$

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

(i) $B(\bar{y}_m) = 0$

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

3.2. RATIO METHOD OF IMPUTATION

Following the notations of Lee, et al. (1994), in the case of single imputation method, if the i[th] unit requires imputation, the value $\hat{b}x_i$ is imputed.

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

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 $\bar{y}_r = \frac{1}{r}\sum_{i \in R} y_i$, $\bar{x}_r = \frac{1}{r}\sum_{i \in R} x_i$ 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

(i) $B(\bar{y}_{RAT}) = \bar{Y}\left(\frac{1}{r} - \frac{1}{n}\right)\left(C_x^2 - \rho C_y C_x\right)$

(ii) $M(\bar{y}_{RAT}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\left[S_Y^2 + R_1^2 S_X^2 - 2R_1 S_{XY}\right]$ $\quad$ *where* $R_1 = \frac{\bar{Y}}{\bar{X}}$

3.3. COMPROMISED METHOD OF IMPUTATION

Singh and Horn (2000) 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

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

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

(i) $B(\bar{y}_{COMP}) = \bar{Y}(1-\alpha)\left(\frac{1}{r} - \frac{1}{n}\right)(C_X^2 - \rho C_Y C_X)$

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

(iii) $M(\bar{y}_{COMP})_{\min} = \left[\left(\frac{1}{r} - \frac{1}{N}\right) - \left(\frac{1}{r} - \frac{1}{n}\right)\rho^2\right]S_Y^2$

3.4. AHMED METHODS

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

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

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

Lemma: 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

$$\text{(i) } B(t_1) = \bar{Y}\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{\beta_1(\beta_1+1)C_X^2}{2} - \beta_1\rho C_Y C_X\right)$$

$$\text{(ii) } M(t_1) = \bar{Y}^2\left[\left(\frac{1}{r}-\frac{1}{N}\right)C_Y^2 + \beta_1^2\left(\frac{1}{n}-\frac{1}{N}\right)C_X^2 - 2\beta_1\left(\frac{1}{n}-\frac{1}{N}\right)\rho C_Y C_X\right]$$

$$\text{(iii) } M(t_1)_{\min} = \left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2 - \left(\frac{1}{n}-\frac{1}{N}\right)\frac{S_{XY}^2}{S_X^2}$$

(2) $$y_{2i} = \begin{cases} y_i & if \quad i \in R \\ \frac{1}{(n-r)}\left[n\bar{y}_r\left(\frac{\bar{x}}{\bar{x}_r}\right)^{\beta_2} - r\bar{y}_r\right] & if \quad i \in R^C \end{cases}$$

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 $t_2 = \bar{y}_r\left(\frac{\bar{x}}{\bar{x}_r}\right)^{\beta_2}$

Lemma: 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

$$\text{(i) } B(t_2) = \left(\frac{1}{r}-\frac{1}{n}\right)\bar{Y}\left(\frac{\beta_2(\beta_2+1)}{2}C_X^2 - \beta_2\rho C_Y C_X\right)$$

$$\text{(ii) } M(t_2) = \bar{Y}^2\left[\left(\frac{1}{r}-\frac{1}{N}\right)C_Y^2 + \beta_2^2\left(\frac{1}{r}-\frac{1}{n}\right)C_X^2 - 2\beta_2\left(\frac{1}{r}-\frac{1}{n}\right)\rho C_Y C_X\right]$$

$$\text{(iii) } M(t_2)_{\min} = \left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2 - \left(\frac{1}{r}-\frac{1}{n}\right)\frac{S_{XY}^2}{S_X^2}$$

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

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

Lemma: 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

$$\text{(i) } B(t_3) = \left(\frac{1}{r} - \frac{1}{N} \right) \bar{Y} \left(\frac{\beta_2(\beta_2 + 1)}{2} C_X^2 - \beta_2 \rho C_Y C_X \right)$$

$$\text{(ii) } M(t_3) = \left(\frac{1}{r} - \frac{1}{N} \right) \bar{Y}^2 \left[C_Y^2 + \beta_3^2 C_X^2 - 2\beta_3 \rho C_Y C_X \right]$$

$$\text{(iii) } M(t_3)_{\min} = \left(\frac{1}{r} - \frac{1}{N} \right) S_Y^2 (1 - \rho^2)$$

3.5 PANDEY ET AL. (2015) SUGGESTED IMPUTATION STRATEGIES

For the case where y_{ji} denotes the i^{th} available observation for the j^{th} imputation method, the imputation method y_{1i} is given as follows:

(1)
$$y_{1i} = \begin{cases} y_i & if \quad i \in R \\ \frac{\bar{y}_r}{(1 - f_1)} \left[\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) - f_1 \right] & if \ i \in R^C \end{cases}$$

Point estimator of $\bar{Y}$ is $t_4 = \bar{y}_r \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$

The Bias and MSE of t_4 respectively are given by

$$\text{(i) } B(t_4) = \bar{Y} M_1 \left(\frac{3}{8} C_X^2 - \frac{1}{2} \rho C_y C_x \right) \text{ and}$$

$$\text{(ii) } M(t_4) = \bar{Y}^2 \left[M_1 C_y^2 + M_2 \left(\frac{1}{4} C_x^2 - \rho C_y C_x \right) \right]$$

(2)
$$y_{2i} = \begin{cases} y_i & if \quad i \in R \\ \frac{\bar{y}_r}{(1 - f_1)} \left[\exp\left(\frac{\bar{x}_r - \bar{x}}{\bar{x}_r + \bar{x}} \right) - f_1 \right] & if \ i \in R^C \end{cases}$$

Point estimator of $\bar{Y}$ is $t_5 = \bar{y}_r \exp\left(\frac{\bar{x}_r - \bar{x}}{\bar{x}_r + \bar{x}} \right)$

The Bias and MSE of t_4 respectively are given by

$$\text{(i) } B(t_5) = \frac{\bar{Y}}{2} \left[(M_1 - M_2) \rho C_y C_x - \frac{1}{4} (M_1 - 2M_2) C_x^2 \right] \text{ and}$$

(ii) $M(t_5) = \frac{\overline{Y}^2}{4}\left[4M_1C_y^2 + M_1C_x^2 + (M_1 - M_2)\rho C_y C_x\right]$

(3) $$y_{3i} = \begin{cases} y_i & if \quad i \in R \\ \frac{\overline{y}_r}{(1-f_1)}\left[\exp\left(\frac{\overline{X} - \overline{x}_r}{\overline{X} + \overline{x}_r}\right) - f_1\right] & if \quad i \in R^C \end{cases}$$

Point estimator of $\overline{Y}$ is $t_6 = \overline{y}_r \exp\left(\frac{\overline{X} - \overline{x}_r}{\overline{X} + \overline{x}_r}\right)$

The Bias and MSE of t_6 respectively are given by

(i) $B(t_6) = \overline{Y}M_1\left(\frac{3}{8}C_X^2 - \frac{1}{2}\rho C_y C_x\right)$ and

(ii) $M(t_6) = \overline{Y}^2 \frac{M_1}{4}\left[4C_y^2 + C_x^2 - 4\rho C_y C_x\right]$

4. SOME PROPOSED METHODS OF IMPUTATION AND THEIR ESTIMATORS

For estimating the population mean of the study variate, Khoshnevisan et al. (2007) has considered a family of estimators given by

$$t = \overline{Y}\left[\frac{a\overline{X} + b}{\alpha(a\overline{x} + b) + (1-\alpha)(a\overline{X} + b)}\right]^g$$

Where $a(\neq 0)$ and b are real numbers.

Motivated by Khoshnevisan et al. (2007), we here propose the following three generalized methods of imputation under two-phase sampling.

Let y'_{ji} denote the i^{th} available observation for the j^{th} imputation method.

The **first proposed** method:

$$y'_{1i} = \begin{cases} y_i & if \quad i \in R \\ \frac{1}{n-r}\left[n\overline{y}_r\left\{\frac{a\overline{x}' + b}{\alpha(\overline{x} + b) + (1-\alpha)(a\overline{x}' + b)}\right\}^g - r\overline{y}_r\right] & if \quad i \in R^c \end{cases}$$

where a and b are free parameters and α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Using above, the imputation-based estimator of population mean $\overline{Y}$ is

$$t'_7 = \overline{y}_r\left[\frac{a\overline{x}' + b}{\alpha(a\overline{x} + b) + (1-\alpha)(a\overline{x}' + b)}\right]^g \quad (1)$$

Theorem 4.1:

The estimator, Bias, M.S.E. and minimum M.S.E of t'_7 in terms of e_i; i=1, 2, 3 and e'_3 under design $t = I, II$ upto first order of approximation are given by

$$t_7^{'}=\bar{Y}\left[1+e_1-g\lambda\alpha\left(e_3-\lambda e_3e_3^{'}-e_3^{'}+\lambda e_3^{'2}\right)+\frac{g(g+1)}{2}\lambda^2\alpha^2\left(e_3^2+e_3^{'2}-2e_3e_3^{'}\right)-g\lambda\alpha\left(e_1e_3-e_1e_3^{'}\right)\right] \quad (2)$$

$$B\left(t_7^{'}\right)_I=\bar{Y}\left[(\delta_2-\delta_3)\left\{\frac{g(g+1)}{2}\lambda^2\alpha^2C_x^2-g\lambda\alpha\rho C_yC_x\right\}\right] \quad (3)$$

$$B\left(t_7^{'}\right)_{II}=\bar{Y}\left[\left\{-g\lambda\alpha\left(\lambda\delta_3+\delta_5\rho\frac{C_y}{C_x}\right)+\frac{g(g+1)}{2}\lambda^2\alpha^2(\delta_3+\delta_5)\right\}C_x^2\right] \quad (4)$$

$$M\left(t_7^{'}\right)_I=\bar{Y}^2\left[\delta_1C_y^2+(\delta_2-\delta_3)\left\{g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\rho C_yC_x\right\}\right] \quad (5)$$

$$M\left(t_7^{'}\right)_{II}=\bar{Y}^2\left[\delta_4C_y^2+(\delta_3+\delta_5)g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\delta_5\rho C_yC_x\right] \quad (6)$$

and $\left[M\left(t_7^{'}\right)_I\right]_{\min}=\bar{Y}^2\left[\delta_1-(\delta_2-\delta_3)\rho^2\right]C_y^2$ when $\alpha=\dfrac{\rho C_y}{g\lambda C_x}$ (7)

$$\left[M\left(t_7^{'}\right)_{II}\right]_{\min}=\bar{Y}^2\left[\delta_4-\delta_5^2(\delta_3+\delta_5)^{-1}\rho^2\right]C_y^2 \quad \text{when } \alpha=\frac{\delta_4\rho C_y}{(\delta_3+\delta_4)g\lambda C_x} \quad (8)$$

The **second proposed** method:

$$y_{2i}^{'}=\begin{cases} y_i & if \quad i\in R \\ \dfrac{1}{n-r}\left[n\bar{y}_r\left\{\dfrac{a\bar{x}+b}{\alpha\left(\bar{x}_r+b\right)+(1-\alpha)\left(a\bar{x}+b\right)}\right\}^g-r\bar{y}_r\right] & if \quad i\in R^c \end{cases}$$

where a *and* b are free parameters and α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Using above, the imputation-based estimator of population mean $\bar{Y}$ is

$$t_8^{'}=\bar{y}_r\left[\frac{a\bar{x}+b}{\alpha\left(a\bar{x}_r+b\right)+(1-\alpha)\left(a\bar{x}+b\right)}\right]^g \quad (9)$$

Theorem 4.2:

The estimator, Bias, M.S.E. and minimum M.S.E of $t_8^{'}$ in terms of e_i; i=1, 2, 3 and $e_3^{'}$ under design $t=I,II$ upto first order of approximation are given by

$$t_8^{'}=\bar{Y}\left[1+e_1-g\lambda\alpha\left(e_2-\lambda e_2e_3-e_3+\lambda e_3^{2}\right)+\frac{g(g+1)}{2}\lambda^2\alpha^2\left(e_2^2+e_3^{2}-2e_2e_3\right)-g\lambda\alpha\left(e_1e_2-e_1e_3\right)\right] \quad (10)$$

$$B\left(t_8^{'}\right)_I=\bar{Y}\left[(\delta_1-\delta_2)\left\{\frac{g(g+1)}{2}\lambda^2\alpha^2C_x^2-g\lambda\alpha\rho C_yC_x\right\}\right] \quad (11)$$

$$B\left(t_8^{'}\right)_{II}=\bar{Y}\left[(\delta_4-\delta_5)\left\{\frac{g(g+1)}{2}\lambda^2\alpha^2C_x^2-g\lambda\alpha\rho C_yC_x\right\}\right] \quad (12)$$

$$M\left(t_8^{'}\right)_I=\bar{Y}^2\left[\delta_1C_y^2+(\delta_1-\delta_2)\left\{g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\rho C_yC_x\right\}\right] \quad (13)$$

$$M\left(t_8^{'}\right)_{II}=\bar{Y}^2\left[\delta_4 C_y^2+(\delta_4-\delta_5)\left\{g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\rho C_yC_x\right\}\right] \quad (14)$$

and $$\left[M\left(t_8^{'}\right)_I\right]_{\min}=\bar{Y}^2\left[\delta_1-(\delta_1-\delta_2)\rho^2\right]C_y^2 \qquad when\ \alpha=\frac{\rho C_y}{g\lambda C_x} \quad (15)$$

$$\left[M\left(t_8^{'}\right)_{II}\right]_{\min}=\bar{Y}^2\left[\delta_4-(\delta_4-\delta_5)\rho^2\right]C_y^2 \qquad when\ \alpha=\frac{\rho C_y}{g\lambda C_x} \quad (16)$$

The **third proposed** method:

$$y_{3i}^{'}=\begin{cases} y_i & if \quad i\in R \\ \dfrac{1}{n-r}\left[n\bar{y}_r\left\{\dfrac{a\bar{x}^{'}+b}{\alpha\left(\bar{x}_r+b\right)+(1-\alpha)\left(a\bar{x}^{'}+b\right)}\right\}^g-r\bar{y}_r\right] & if \quad i\in R^c \end{cases}$$

where $a\ and\ b$ are free parameters and α is a suitably chosen constant, such that the resultant variance of the estimator is minimum.

Using above, the imputation-based estimator of population mean $\bar{Y}$ is

$$t_9^{'}=\bar{y}_r\left[\frac{a\bar{x}^{'}+b}{\alpha\left(a\bar{x}_r+b\right)+(1-\alpha)\left(a\bar{x}^{'}+b\right)}\right]^g \quad (17)$$

Theorem 4.3:

The estimator, Bias, M.S.E. and minimum M.S.E of $t_9^{'}$ in terms of e_i; i=1, 2, 3 and $e_3^{'}$ under design $t=I,II$ upto first order of approximation are given by

$$t_9^{'}=\bar{Y}\left[1+e_1-g\lambda\alpha\left(e_2-\lambda e_2e_3^{'}-e_3^{'}+\lambda e_3^{'2}\right)+\frac{g(g+1)}{2}\lambda^2\alpha^2\left(e_2^2+e_3^{'2}-2e_2e_3^{'}\right)-g\lambda\alpha\left(e_1e_2-e_1e_3^{'}\right)\right] \quad (18)$$

$$B\left(t_9^{'}\right)_I=\bar{Y}\left[(\delta_1-\delta_3)\left\{\frac{g(g+1)}{2}\lambda^2\alpha^2C_x^2-g\lambda\alpha\rho C_yC_x\right\}\right] \quad (19)$$

$$B\left(t_9^{'}\right)_{II}=\bar{Y}\left[\left\{-g\lambda\alpha\left(\lambda\delta_3+\delta_4\rho\frac{C_y}{C_x}\right)+\frac{g(g+1)}{2}\lambda^2\alpha^2(\delta_3+\delta_4)\right\}C_x^2\right] \quad (20)$$

$$M\left(t_9^{'}\right)_I=\bar{Y}^2\left[\delta_1C_y^2+(\delta_1-\delta_3)\left\{g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\rho C_yC_x\right\}\right] \quad (21)$$

$$M\left(t_9^{'}\right)_{II}=\bar{Y}^2\left[\delta_4C_y^2+(\delta_3+\delta_4)g^2\lambda^2\alpha^2C_x^2-2g\lambda\alpha\delta_4\rho C_yC_x\right] \quad (22)$$

and $$\left[M\left(t_9^{'}\right)_I\right]_{\min}=\bar{Y}^2\left[\delta_1-(\delta_1-\delta_3)\rho^2\right]C_y^2 \qquad when\ \alpha=\frac{\rho C_y}{g\lambda C_x} \quad (23)$$

$$\left[M\left(t_9^{'}\right)_{II}\right]_{\min}=\bar{Y}^2\left[\delta_4-\delta_4^2(\delta_3+\delta_4)^{-1}\rho^2\right]C_y^2 \qquad when\ \alpha=\frac{\delta_4\rho C_y}{(\delta_3+\delta_4)g\lambda C_x} \quad (24)$$

5. COMPARISON

In this section, we derive the conditions under which the suggested estimators are superior to the existing estimators in design I and II.

(1) $D_1 = V(\bar{y}_m) - [M(t_7')_I]_{\min} = \bar{Y}^2\left[\left\{\left(\frac{1}{n'}-\frac{1}{N}\right)+\left(\frac{1}{n}-\frac{2}{n'}+\frac{1}{N}\right)\rho^2\right\}C_y^2\right]$

$(t_7')_I$ is better than $\bar{y}_m$, when $D_1 > 0 \Rightarrow n < \frac{Nn'}{2N-n'}$

(2) $D_2 = V(\bar{y}_m) - [M(t_7')_{II}]_{\min} = \bar{Y}^2\left[\left\{\left(\frac{1}{N-n'}-\frac{1}{N}\right)+\frac{\left(\frac{1}{n}-\frac{1}{N-n'}\right)^2\rho^2}{\left[\left(\frac{1}{n'}-\frac{1}{N}\right)+\left(\frac{1}{n}-\frac{1}{N-n'}\right)\right]}\right\}C_y^2\right]$

Which is positive. Thus $(t_7')_{II}$ is better than $\bar{y}_m$

(3) $D_3 = M(\bar{y}_{RAT}) - [M(t_7')_I]_{\min}$

$= \bar{Y}^2\left[\left(\frac{1}{n'}-\frac{1}{N}\right)C_Y^2+\left(\frac{1}{r}-\frac{1}{n}\right)\left(C_X^2-2\rho C_Y C_X\right)+\left(\frac{1}{n}-\frac{2}{n'}+\frac{1}{N}\right)\rho^2 C_Y^2\right]$

$(t_7')_I$ is better than $\bar{y}_{RAT}$, when $D_3 > 0$

This generates two conditions,

(i) When $\left(C_X^2-2\rho C_Y C_X\right)>0 \Rightarrow \rho\frac{C_Y}{C_X}<\frac{1}{2}$

and (ii) When $\left(\frac{1}{n}-\frac{2}{n'}+\frac{1}{N}\right)>0 \Rightarrow n<\frac{Nn'}{2N-n'}$

(4) $D_4 = M(\bar{y}_{RAT}) - [M(t_7')_{II}]_{\min}$

$= \bar{Y}^2\left[\left(\frac{1}{N-n'}-\frac{1}{N}\right)C_Y^2+\left(\frac{1}{r}-\frac{1}{n}\right)\left(C_X^2-2\rho C_Y C_X\right)+\frac{\left(\frac{1}{n}-\frac{1}{N-n'}\right)^2\rho^2 C_Y^2}{\left[\left(\frac{1}{n'}-\frac{1}{N}\right)+\left(\frac{1}{n}-\frac{1}{N-n'}\right)\right]}\right]$

$(t_7')_{II}$ is better than $\bar{y}_{RAT}$, when $D_4 > 0$

This condition holds,

When $\left(C_X^2-2\rho C_Y C_X\right)>0 \Rightarrow \rho\frac{C_Y}{C_X}<\frac{1}{2}$

(5) $D_5 = [M(\bar{y}_{COMP})]_{\min} - [M(t_7')_I]_{\min} = \left[\left(\frac{1}{n'}-\frac{1}{N}\right)S_Y^2+\left(\frac{2}{n}-\frac{2}{n'}+\frac{1}{N}-\frac{1}{r}\right)\rho^2 S_Y^2\right]$

$(t_7')_I$ is better than $\bar{y}_{COMP}$, when $D_5 > 0 \Rightarrow r < \frac{Nnn'}{2Nn'-2Nn+nn'}$

(6) $D_6 = [M(\bar{y}_{COMP})]_{\min} - [M(t_7')_{II}]_{\min} = (M_1-\delta_4)S_Y^2 - \left[\frac{M_3(\delta_3+\delta_5)-\delta_5^2}{\delta_3+\delta_5}\right]\rho^2 S_Y^2$

$(t_7^{'})_{II}$ is better than $\bar{y}_{COMP}$, when $D_6 > 0 \Rightarrow -\sqrt{P} < \rho < \sqrt{P}$

Where $P = \dfrac{(M_1 - \delta_4)(\delta_3 + \delta_5)}{M_3(\delta_3 + \delta_5) - \delta_5^2}$, $M_1 = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)$, $M_3 = \left(\dfrac{1}{r} - \dfrac{1}{n}\right)$

(7) $D_7 = V(\bar{y}_m) - [M(t_8^{'})_I]_{\min} = \bar{Y}^2\left[\left\{\left(\dfrac{1}{n'} - \dfrac{1}{N}\right) + \left(\dfrac{1}{r} - \dfrac{1}{n}\right)\rho^2\right\}C_Y^2\right] > 0$

Which is always true.

(8) $D_8 = V(\bar{y}_m) - [M(t_8^{'})_{II}]_{\min} = \bar{Y}^2\left[\left\{\left(\dfrac{1}{N-n'} - \dfrac{1}{N}\right) + \left(\dfrac{1}{r} - \dfrac{1}{n}\right)\rho^2\right\}C_Y^2\right] > 0$

Which is always true.

(9) $D_9 = M(\bar{y}_{RAT}) - [M(t_8^{'})_I]_{\min}$

$$= \bar{Y}^2\left[\left(\frac{1}{n'} - \frac{1}{N}\right)C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)(C_X^2 - 2\rho C_Y C_X) + \left(\frac{1}{r} - \frac{1}{n}\right)\rho^2 C_Y^2\right]$$

$(t_8^{'})_I$ is better than $\bar{y}_{RAT}$, when $D_9 > 0 \Rightarrow \rho \dfrac{C_Y}{C_X} < \dfrac{1}{2}$

(10) $D_{10} = M(\bar{y}_{RAT}) - [M(t_8^{'})_{II}]_{\min}$

$$= \bar{Y}^2\left[\left(\frac{1}{N-n'} - \frac{1}{N}\right)C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)(C_X^2 - 2\rho C_Y C_X) + \left(\frac{1}{r} - \frac{1}{n}\right)\rho^2 C_Y^2\right]$$

$(t_8^{'})_{II}$ is better than $\bar{y}_{RAT}$, when $D_{10} > 0 \Rightarrow \rho \dfrac{C_Y}{C_X} < \dfrac{1}{2}$

(11) $D_{11} = [M(\bar{y}_{COMP})]_{\min} - [M(t_8^{'})_I]_{\min} = \bar{Y}^2\left(\dfrac{1}{n'} - \dfrac{1}{N}\right)C_Y^2 > 0$

Which is always true.

(12) $D_{12} = [M(\bar{y}_{COMP})]_{\min} - [M(t_8^{'})_{II}]_{\min} = \bar{Y}^2\left(\dfrac{1}{N-n'} - \dfrac{1}{N}\right)C_Y^2 > 0$

Which is always true.

(13) $D_{13} = V(\bar{y}_m) - [M(t_9^{'})_I]_{\min} = \bar{Y}^2\left[\left\{\left(\dfrac{1}{n'} - \dfrac{1}{N}\right) + \left(\dfrac{1}{r} - \dfrac{2}{n'} + \dfrac{1}{N}\right)\rho^2\right\}C_Y^2\right]$

$(t_9^{'})_I$ is better than $\bar{y}_m$ when $D_{13} > 0 \Rightarrow r < \dfrac{Nn'}{2N - n'}$

(14) $D_{14} = V(\bar{y}_m) - [M(t_9^{'})_{II}]_{\min} = \bar{Y}^2\left[\left\{\left(\dfrac{1}{N-n'} - \dfrac{1}{N}\right) + \dfrac{\left(\dfrac{1}{r} - \dfrac{1}{N-n'}\right)^2\rho^2}{\left[\left(\dfrac{1}{n'} - \dfrac{1}{N}\right) + \left(\dfrac{1}{r} - \dfrac{1}{N-n'}\right)\right]}\right\}C_y^2\right] > 0$

Which is always true.

(15) $D_{15} = M(\bar{y}_{RAT}) - [M(t_9^{'})_I]_{\min}$

$$= \bar{Y}^2\left[\left(\frac{1}{n'} - \frac{1}{N}\right)C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)(C_X^2 - 2\rho C_Y C_X) + \left(\frac{1}{r} - \frac{2}{n'} + \frac{1}{N}\right)\rho^2 C_Y^2\right]$$

$\left(t_9'\right)_I$ is better than $\bar{y}_{RAT}$, when $D_{15} > 0$

This generates two conditions,

(i) When $\left(C_X^2 - 2\rho C_Y C_X\right) > 0 \Rightarrow \rho \dfrac{C_Y}{C_X} < \dfrac{1}{2}$

and (ii) When $\left(\dfrac{1}{r} - \dfrac{2}{n'} + \dfrac{1}{N}\right) > 0 \Rightarrow r < \dfrac{Nn'}{2N - n'}$

(16) $D_{16} = M\left(\bar{y}_{RAT}\right) - \left[M\left(t_9'\right)_{II}\right]_{\min}$

$$= \bar{Y}^2 \left[\left(\frac{1}{N-n'} - \frac{1}{N}\right) C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\left(C_X^2 - 2\rho C_Y C_X\right) + \frac{\left(\frac{1}{r} - \frac{1}{N-n'}\right)^2 \rho^2 C_Y^2}{\left[\left(\frac{1}{n'} - \frac{1}{N}\right) + \left(\frac{1}{r} - \frac{1}{N-n'}\right)\right]} \right]$$

$\left(t_9'\right)_{II}$ is better than, when $D_{16} > 0$

This condition holds,

When $\left(C_X^2 - 2\rho C_Y C_X\right) > 0 \Rightarrow \rho \dfrac{C_Y}{C_X} < \dfrac{1}{2}$

(17) $D_{17} = \left[M\left(\bar{y}_{COMP}\right)\right]_{\min} - \left[M\left(t_9'\right)_I\right]_{\min} = \bar{Y}^2 \left[\left\{ \left(\dfrac{1}{n'} - \dfrac{1}{N}\right) + \left(\dfrac{1}{n} - \dfrac{2}{n'} + \dfrac{1}{N}\right)\rho^2 \right\} C_Y^2 \right]$

$\left(t_9'\right)_I$ is better than $\bar{y}_{COMP}$, when $D_{17} > 0 \Rightarrow n < \dfrac{Nn'}{2N - n'}$

(18) $D_{18} = \left[M\left(\bar{y}_{COMP}\right)\right]_{\min} - \left[M\left(t_9'\right)_{II}\right]_{\min} = \left(M_1 - \delta_4\right) S_Y^2 - \left[\dfrac{M_3\left(\delta_3 + \delta_4\right) - \delta_4^2}{\delta_3 + \delta_4}\right] \rho^2 S_Y^2$

$\left(t_9'\right)_{II}$ is better than $\bar{y}_{COMP}$, when $D_{18} > 0 \Rightarrow -\sqrt{Q} < \rho < \sqrt{Q}$

Where $Q = \dfrac{\left(M_1 - \delta_4\right)\left(\delta_3 + \delta_4\right)}{M_3\left(\delta_3 + \delta_4\right) - \delta_4^2}$, $M_1 = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)$, $M_3 = \left(\dfrac{1}{r} - \dfrac{1}{n}\right)$

(19) $D_{19} = \left[M\left(t_1\right)\right]_{\min} - \left[M\left(t_7'\right)_I\right]_{\min} = \bar{Y}^2 \left[\left\{ \left[\dfrac{1}{n'} - \dfrac{1}{N}\right] - 2\left[\dfrac{1}{n'} - \dfrac{1}{N}\right]\rho^2 \right\} C_Y^2 \right]$

$\left(t_7'\right)_I$ is better than t_1, when $D_{19} > 0 \Rightarrow -\dfrac{1}{\sqrt{2}} < \rho < \dfrac{1}{\sqrt{2}}$

(20) $D_{20} = \left[M\left(t_1\right)\right]_{\min} - \left[M\left(t_7'\right)_{II}\right]_{\min} = \left[M_1 - \delta_4\right] S_Y^2 - \left[\dfrac{M_2\left(\delta_3 + \delta_5\right) - \delta_5^2}{\left(\delta_3 + \delta_5\right)}\right] \rho^2 S_Y^2$

$\left(t_7'\right)_{II}$ is better than t_1, when $D_{20} > 0$

$$\Rightarrow \rho^2 < \frac{\left(M_1 - \delta_4\right)\left(\delta_3 + \delta_5\right)}{\left[M_2\left(\delta_3 + \delta_5\right) - \delta_5^2\right]} \Rightarrow -\sqrt{R} < \rho < \sqrt{R}$$

Where, $R = \dfrac{\left(M_1 - \delta_4\right)\left(\delta_3 + \delta_5\right)}{\left[M_2\left(\delta_3 + \delta_5\right) - \delta_5^2\right]}$; $M_1 = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)$; $M_2 = \left(\dfrac{1}{n} - \dfrac{1}{N}\right)$

(21) $D_{21} = [M(t_2)]_{\min} - [M(t_8^{'})_I]_{\min} = \bar{Y}^2\left(\frac{1}{n'} - \frac{1}{N}\right)C_Y^2 > 0$

Which is always true.

(22) $D_{22} = [M(t_2)]_{\min} - [M(t_8^{'})_{II}]_{\min} = \bar{Y}^2\left[\frac{1}{N-n'} - \frac{1}{N}\right]C_Y^2 > 0$

Which is always true.

(23) $D_{23} = [M(t_3)]_{\min} - [M(t_9^{'})_I]_{\min} = \bar{Y}^2\left[\left\{\left[\frac{1}{n'} - \frac{1}{N}\right] - 2\left[\frac{1}{n'} - \frac{1}{N}\right]\rho^2\right\}C_Y^2\right]$

$(t_9^{'})_I$ is better than t_3, when $D_{23} > 0 \Rightarrow -\frac{1}{\sqrt{2}} < \rho < \frac{1}{\sqrt{2}}$

(24) $D_{24} = [M(t_3)]_{\min} - [M(t_9^{'})_{II}]_{\min} = [M_1 - \delta_4]S_y^2 - \left[M_1 - (\delta_3 + \delta_5)^{-1}\delta_4^2\right]\rho^2 S_Y^2$

$(t_9^{'})_{II}$ is better than t_3, when $D_{24} > 0$

$$\Rightarrow \rho^2 < \frac{(M_1 - \delta_4)(\delta_3 + \delta_5)}{[M_1(\delta_3 + \delta_4) - \delta_4^2]} \Rightarrow -\sqrt{T} < \rho < \sqrt{T}$$

Where $T = \frac{(M_1 - \delta_4)(\delta_3 + \delta_5)}{[M_1(\delta_3 + \delta_4) - \delta_4^2]}; M_1 = \left(\frac{1}{r} - \frac{1}{N}\right)$

(25) $D_{25} = [M(t_4)]_{\min} - [M(t_7^{'})_I]_{\min}$

$$= \bar{Y}^2\left[\left(\frac{1}{n^{'}} - \frac{1}{N}\right)C_y^2 + \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{1}{4}C_x^2 - \rho C_y C_x\right) + \left(\frac{1}{n} - \frac{2}{n^{'}} + \frac{1}{N}\right)\rho^2 C_y^2\right]$$

$(t_7^{'})_I$ is better than t_4, when $D_{25} > 0$

This generates two conditions,

(i) When $\left(\frac{1}{4}C_X^2 - \rho C_Y C_X\right) > 0 \Rightarrow \rho\frac{C_Y}{C_X} < \frac{1}{4}$

and (ii) When $\left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N}\right) > 0 \Rightarrow n < \frac{Nn'}{2N - n'}$

(26) $D_{26} = [M(t_4)]_{\min} - [M(t_7^{'})_{II}]_{\min}$

$$= \bar{Y}^2\left[\left(\frac{1}{n^{'}} - \frac{1}{N}\right)C_y^2 + \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{1}{4}C_x^2 - \rho C_y C_x\right) + \frac{\left(\frac{1}{n} - \frac{1}{N-n'}\right)^2\rho^2 C_y^2}{\left(\frac{1}{n'} - \frac{1}{N}\right) + \left(\frac{1}{n} - \frac{1}{N-n'}\right)}\right]$$

$(t_7^{'})_{II}$ is better than t_4, when $D_{26} > 0$

This condition holds,

When $\left(\frac{1}{4}C_X^2-\rho C_Y C_X\right)>0 \Rightarrow \rho\frac{C_Y}{C_X}<\frac{1}{4}$

(27) $D_{27}=\left[M(t_5)\right]_{\min}-\left[M(t_8^{'})_I\right]_{\min}$

$$=\bar{Y}^2\left[\left(\frac{1}{r}-\frac{1}{n}\right)C_y^2+\left(\frac{1}{r}-\frac{1}{N}\right)\frac{1}{4}C_x^2+\left(\frac{1}{r}-\frac{1}{n}\right)\left(\frac{1}{4}\rho C_y C_x+\rho^2C_y^2\right)\right]>0$$

This is always true.

(28) $D_{28}=\left[M(t_5)\right]_{\min}-\left[M(t_8^{'})_{II}\right]_{\min}$

$$=\bar{Y}^2\left[\left(\frac{1}{N-n'}-\frac{1}{N}\right)C_y^2+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{1}{4}C_x^2+\left(\frac{1}{r}-\frac{1}{n}\right)\left(\frac{1}{4}\rho C_y C_x+\rho^2C_y^2\right)\right]>0$$

Which is always true.

(29) $D_{29}=\left[M(t_6)\right]_{\min}-\left[M(t_9^{'})_I\right]_{\min}$

$$=\bar{Y}^2\left[\left(\frac{1}{n^{'}}-\frac{1}{N}\right)C_y^2+\left(\frac{1}{r}-\frac{1}{N}\right)\left(\frac{1}{4}C_x^2-\rho C_y C_x\right)+\left(\frac{1}{r}-\frac{2}{n^{'}}+\frac{1}{N}\right)\rho^2C_y^2\right]$$

$(t_9^{'})_I$ is better than t_6, when $D_{29}>0$

This generates two conditions,

(i) When $\left(\frac{1}{4}C_X^2-\rho C_Y C_X\right)>0 \Rightarrow \rho\frac{C_Y}{C_X}<\frac{1}{4}$

and (ii) When $\left(\frac{1}{r}-\frac{2}{n'}+\frac{1}{N}\right)>0\Rightarrow r<\frac{Nn'}{2N-n'}$

(30) $D_{30}=\left[M(t_6)\right]_{\min}-\left[M(t_9^{'})_{II}\right]_{\min}$

$$=\bar{Y}^2\left[\left(\frac{1}{N-n^{'}}-\frac{1}{N}\right)C_y^2+\left(\frac{1}{r}-\frac{1}{N}\right)\left(\frac{1}{4}C_x^2-\rho C_y C_x\right)+\frac{\left(\frac{1}{r}-\frac{1}{N-n'}\right)^2\rho^2C_y^2}{\left(\frac{1}{n'}-\frac{1}{N}\right)+\left(\frac{1}{r}-\frac{1}{N-n'}\right)}\right]$$

$(t_9^{'})_{II}$ is better than t_6, when $D_{30}>0$

This condition holds,

When $\left(\frac{1}{4}C_X^2-\rho C_Y C_X\right)>0 \Rightarrow \rho\frac{C_Y}{C_X}<\frac{1}{4}$

6. ILLUSTRATİVE EXAMPLES

Population **A [Source: Kadilar and Cingi (2006)]:** Y represents apple production and X represents number of apple trees. Other related information to population are:

$N = 106$ $n = 20$ $\overline{Y} = 2212.59$ $\overline{X} = 27421.70$ $\rho = 0.86$ $S_y = 11551.53$

$C_y = 5.22$ $S_x = 57460.61$ $C_x = 2.10$ $n' = 60$ $r = 15$

Population **B [Source: Koyuncu and Kadilar (2009)]:** Y is number of teachers and X is number of students. Other related information to population are:

$N = 923$ $n = 180$ $\overline{Y} = 436.4345$ $\overline{X} = 11440.498$ $\rho = 0.9543$ $S_y = 749.9394$

$C_y = 1.7183$ $S_x = 21331.131$ $C_x = 1.8645$ $n' = 360$ $r = 175$

Population **C [Source: Murthy (1967)]:** Y is area under winter paddy and X is corresponding geographical area. The following data is based on the given population:

$N = 108$ $n = 30$ $\overline{Y} = 172.3704$ $\overline{X} = 461.3981$ $\rho = 0.7896$ $S_y = 134.3567$

$C_y = 0.7795$ $S_x = 318.5022$ $C_x = 0.6903$ $n' = 70$ $r = 20$

Then, the M.S.E. and Percentage Relative Efficiencies (P.RE.'s) of the estimators $\overline{y}_m, \overline{y}_{RAT}$, $\overline{y}_{COMP}$, Ahmed's et al. (2006), Pandey et al. (2015), t_7', t_8' *and* t_9' for population A, B and C with respect to $\overline{y}_m$ is computed by using the formula

$$PRE\left(*, \overline{y}_m\right) = \left[\frac{MSE\left(\overline{y}_m\right)}{MSE(*)}\right] \times 100$$ where (*) represents the respective estimator.

Table 6.1: **MSE and Percentage Relative Efficiencies (PRE) of Estimators**

Estimators	Population A		Population B		Population C	
	MSE	PRE	MSE	PRE	MSE	PRE
$\overline{y}_m$	7637008.759	100	2604.439	100	735.4405599	100
$\overline{y}_{RAT}$	6456040.034	118.2924629	2524.573355	103.1635304	550.6829307	133.5506367
$\overline{y}_{COMP}$	5990287.043	127.4898632	2523.046473	103.2259623	547.9118047	134.2260841
t_1	5992164.918	127.4499094	2523.140668	103.2221086	547.8626529	134.2381263
t_2	3633520.921	210.18205	313.9047128	829.6909523	464.4946943	158.3313155
t_3	1988677.081	384.0245776	232.6063798	1119.676512	276.9167872	265.5817898
t_4	7634615.07	100.0313531	2604.331984	100.0041092	516.8127906	142.303088
t_5	8135820.108	93.86894815	3394.043796	76.73557435	932.3068638	78.88395854
t_6	5302118.856	144.0369212	674.1498099	386.3294126	365.4094254	201.2648029
t_7'	4094721.103	186.5086429	1096.570493	237.5076675	486.9444334	151.0317214

$t_8^{'}$	5025473.008	151.9659691	1570.161843	165.8707356	457.1674913	160.8689537
$t_9^{'}$	2450392.738	311.664683	1015.275195	256.5254242	299.3496978	245.6794062

7. CONCLUSION

Imputation is a mechanism which provides nearby value corresponding to the non-respondent units in a survey. In literature, several strategies for imputation of population mean under missing group of data exists and estimators under each strategy observe an inherent deviation. The main of the research in imputation is to reduce this gap and improve the estimator for achieving higher precision.

Table 6.1 gives an overview on the percentage Relative Efficiency (P.R.E.) of various estimators with respect to $\bar{y}_m$. In this paper, the PRE of the suggested estimators viz; $t_7^{'}, t_8^{'}$ *and* $t_9^{'}$ has been compared with estimators $\bar{y}_m, \bar{y}_{RAT}, \bar{y}_{COMP}$, Ahmed's et al.(2006) and Pandey et al. (2015). The performance of the proposed estimators is justified theoretically and numerically. From table 6.1, it is observed that the proposed estimators $t_7^{'}, t_8^{'}$ *and* $t_9^{'}$ in its optimality are performing better than the estimators taken for comparisons in the presence of missing data in population A, population B and population C with exception of the estimators suggested by Ahmed's et al.(2006), where proposed estimator $t_7^{'}$ is only better. Also, the proposed estimator $t_9^{'}$ is performing better than the proposed estimators $t_7^{'}$ *and* $t_8^{'}$.

We, therefore, conclude that the proposed general classes of estimators for estimating population mean in compromised imputation strategies are superior to the existing mean, ratio, compromised, Ahmed's method (t_1) and Pandey et al. The present research paper is therefore an important contribution in the area of missing data as it offers efficient estimators.

REFERENCES:

1. Ahmed, M. S., Al-Titi, O., Al-Rawi, Z., & Abu-Dayyeh, W. (2006). Estimation of a population mean using different imputation methods, Statsitics in Transition, 7(6), p. 1247-1264.
2. Arnab, R. & Singh, S. (2006). A new method for estimating variance from data imputed with ratio method of imputation. Statistics & probability letters, 76(5):513-519.
3. Baraldi, A. N. & Enders, C. K. (2010). An introduction to modern missing data analyses. Journal of school psychology, 48(1):5-37.
4. Diana, G. & Francesco Perri, P. (2010). Improved estimators of the population mean for missing data. Communications in StatisticsTheory and Methods, 39(18):3245-3251.
5. Gonz_alez, S., Rueda, M., & Arcos, A. (2008). An improved estimator to analyse missing data. Statistical Papers, 49(4):791-796.
6. Kadilar, C. & Cingi, H. (2008). Estimators for the population mean in the case of missing data. Communications in StatisticsTheory and Methods, 37(14):2226-2236.
7. Kalton, G., Kasprzyk, D., & Santos, R. (1981). Issues of non-response and imputation of income and participation. Current topics in survey sampling. In D. Krevoski, R. Platek and J N. K. Rao(Eds.), (pp.455-480). New York: Acad. Press.
8. Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N. and Smarandache, F., (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(2):181-191.

9. Lee, H., Rancourt, E., & Sarndal, C. E. (1994). Experiments with variance estimation from survey data with imputed values. Journal of Official Statistics, 10(3). P. 231-243.

10. Lee, H., Rancourt, E., & Sarndal, C. E. (1995). Variance estimation in the presence of imputed data for the generalized estimation system. Proc. of the American Statist. Assoc. (Social Survey Research Methods Section), p. 384-389.

11. Pandey, R., Thakur, N. S. and Yadav, K. (2015).Estimation of population mean using exponential ratio type imputation method under survey non-response. Journal of Indian Statistical Association, Vol.52 No. 1&2, 89-107.

12. Rueda, M. & Gonz_alez, S. (2008). A new ratio-type imputation with random disturbance. Applied Mathematics Letters, 21(9):978-982.

13. Sande, I. G. (1979). A personal view of hot deck approach to automatic edit and imputation. Journal Imputation Procedures. Survey Methodology, 5, 238-246.

14. Shukla, D., Thakur, N. S. (2008). Estimation of mean with imputation of missing data using factor-type estimator, Statistics in Transition, 9(1), p. 33-48.

15. Shukla, D., Thakur, N. S., Pathak, S. & Rajput D. S. (2009). Estimation of mean with imputation of missing data using factor-type estimator in two-phase sampling, Statistics in Transition, 10(3), p. 397-414.

16. Shukla, D., Thakur, N. S. & Thakur, D. S. (2009a). Utilization of non-response auxiliary population mean in imputation for missing observations, Journal of Reliability and Statistical Studies, 2(1), p. 28-40.

17. Shukla, D., Thakur, N. S., Thakur, D. S. & Pathak, S. (2011). Linear combination based imputation method for missing data in sample, International Journal of Modern Engineering Research (IJMER), 1(2). p. 580-596.

18. Shukla, D., Thakur, N., and Pathak, S. (2013). Some new aspects on imputation in sampling. African Journal of Mathematics and Computer Science Research, 6(1):5-15.

19. Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. Metrika, 51(3):267-276.

20. Singh, G., Priyanka, K., Kim, J.-M., and Singh, S. (2010). Estimation of population mean using imputation techniques in sample surveys. Journal of the Korean Statistical Society, 39(1):67-74.

21. Singh, A. K., Singh, P. & Singh, V. K. (2014). Imputation methods of missing data for estimating the population mean using simple random sampling, Global Journal of Advance Research, 1(2), p.253-263.

22. Singh, Priyanka, Singh, A. K. and Singh, V. K. (2015). On the use of compromised imputation for missing data using Factor-Type estimators, Journal of Statistics Applications and Probability Letters,2(2), 1-9.

23. Singh, A. K., Singh, P., and Singh V.K. (2016). Estimation of Mean under Imputation of Missing data using Exponential-type Estimator in Two-phase sampling, International Journal of Statistics and Economics, 17(1):73-81.

24. Singh, P., Singh A. K., and Singh, V.K. (2017). On the Use of Chain –type estimator in Compromised Imputation under two phase sampling, International Journal of Statistics and Economics, 18(1):39-51

25. Thakur, N. S. & Yadav, K. (2011). Estimation of mean in presence of missing data under two-phase sampling scheme, Journal of Reliability and Statistical Studies, 4(2), p. 93-104.

26. Thakur, N. S., Yadav, K., Pathak, S. (2013). On mean estimation with imputation in two-phase sampling design, Research Journal of Mathematical and Statistical Sciences, 1(3), p. 1-9 .
27. Thakur , N. S. and yadav, K and Pandey, R. (2016). A Family of Factor-type Estimators for estimation of population variance, Journal of Reliability and Statistical Studies, 9(2), 71-79.