http://dx.doi.org/10.5455/sf.71109

Science Forum (Journal of Pure and Applied Sciences)

journal homepage: www.atbuscienceforum.com

Regression-cum-exponential ratio imputation class of estimators of population mean in the presence of non-response

Ahmed Audu[1], Ishaq Olatunji Olawoyin[2], Yahaya Zakari[3], Daniel Dauda Wisdom[1], Jamiu Muili[1], Abdulaziz Mohammed Ndatsu[1]

[1] *Department of Mathematics, Usmanu Danfidiyo University, Sokoto, Nigeria*

[2] *Department of Statistics, Kano State Uni. Sci. and Tech, Wudil, Nigeria*

[3] *Department of Statistics, Ahmadu Bello University, Zaria, Nigeria*

ABSTRACT

Experiments, for example, medical and social science surveys, conducted by human are often characterized by the problem of non-response of missing observations. In this study, an alternative imputation method to Singh and Horn and Singh et al. compromised imputation methods has been suggested due to decrease in their efficiency when the value of unknown weights tend to unity. The properties (bias and mean squared error) of the proposed estimators were derived up to first-order approximation using Taylor series approach. Conditions for which the proposed estimator more efficient than other estimators considered in the study were also established. Numerical illustration was conducted and the results revealed that the proposed estimator is more efficient.

ARTICLE INFO

Article history:
Received 25 October 2019
Received in revised form
08 December 2019
Accepted 09 December 2019
Published 21 February 2020
Available online 21 February 2020

KEYWORDS

Imputation
Non-response
Estimator
Population mean
Mean squared error

1. Introduction

Experiments that include humans (for example, medical and social science surveys) are often confronted with the problem of non-response. These lost values in turn create complications in the handling and analysis of data. Over time, many ways have been developed to deal with the problem of estimating unknown parameters in the presence of missing values. Imputation is a common technique used to treat situations in which data are missing. Missing values can be completed with specific alternatives and the data can be analyzed using the standard methods. Information on imputation methods help to improve the accuracy of estimates of population parameters Hansen and Hurwitz (1946) were the first to consider the problem of non-response. Several authors also proposed estimates to estimate unknown parameters when missing values. Kadilar and Cingi (2008) suggested estimating the average population of missing data values. Diana and Perri (2010) used additional information about the lost values. Singh et al. (2014) suggested exponential type of the projection method. Other authors, including Mishra et al. (2017), Prasad (2017), Singh and Gogoi, Singh et al. (2010), etc also proposed estimators of population mean under imputation methods. However, the efficiency of the estimators proposed by Singh and Horn (2000) and Singh et al. (2014) decreases as the value of unknown weight λ $(0 < \lambda < 1)$and v $(0 < v < 1)$, respectively, approach unity due to the loss of auxiliary information attributed

*** Corresponding author** Ahmed Audu ✉ ahmed.audu@udusok.edu.ng ✉ Department of Mathematics, Usmanu Danfidiyo University, Sokoto, Nigeria.

to such situation. In the current study, we suggest an alternative exponential-type compromised imputation method to estimate population mean $\bar{Y}$ of variable y.

Consider a population Ω_N of size N from which a random sample S_n of size n is drawn without replacement. Let r be the number of responding units out of n samples. Also, let denote the set of units responded by R and the units missing by R^c. For each $i \in R$, the value of y_i is observed. However, for unit $i \in R^c$, y_i is missing and the calculated values are derived using different methods. The computation is done with the help of the auxiliary variable x so that the values of x_i is known and positive.

The Mean Method of imputation, values found missing are to be replaced by the mean of the rest of observed values. The study variable thereafter, takes the form given as

$$y_{.i} = \begin{cases} y_i & \text{if } i \in R \\ \bar{y}_r & \text{if } i \in R^c \end{cases} \tag{1.1}$$

Under the method of imputation, sample mean denoted by $\hat{\bar{\theta}}_{\text{mean}}$ can be derived as

$$\hat{\theta}_{\text{mean}} = \frac{1}{n}\left(\sum_{i\in R} y_i + \sum_{i \in R^c} \bar{y}_r\right) = \frac{1}{n}\left(r\bar{y}_r + (n-r)\bar{y}_r\right) = \bar{y}_r \tag{1.2}$$

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

The bias and mean square error (MSE) of $\hat{\bar{\theta}}_{\text{mean}}$ are given by Eqs. (1.3) and (1.4), respectively.

$$\text{Bias}(\hat{\theta}_{\text{mean}}) = 0 \tag{1.3}$$

$$\text{MSE}(\hat{\theta}_{\text{mean}}) = \varphi_{r,N}\bar{Y}^2 C_Y^2 \tag{1.4}$$

where $\varphi_{r,N} = \frac{1}{r} - \frac{1}{N}$, $C_Y = \frac{S_Y}{\bar{Y}}$,

$$S_Y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2},\ \bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

Under ratio method of imputation, values found missing in the study variable are to be replaced by values obtained using the expression $\hat{\beta} = \sum_{i=1}^{r} y_i / \sum_{i=1}^{r} x_i = \bar{y}_r / \bar{x}_r$. The study variable thereafter, takes the form given as

$$y_{.i} = \begin{cases} y_i & i \in R \\ \hat{\beta}x_i & i \in R^c \end{cases} \tag{1.5}$$

Under the method of imputation, estimator of population mean denoted by $\hat{\bar{\theta}}_{\text{ratio}}$ can be derived as

$$\hat{\theta}_{\text{ratio}} = \frac{n^{-1}\left(\sum_{i\in R} y_i + \sum_{i\in R^c}\hat{\beta}x_i\right)}{n^{-1}\left(\sum_{i\in R} x_i + \sum_{i \in R^c} x_i\right)} N^{-1}\left(\sum_{i\in S_n} x_i + \sum_{i \in S_{N-n}} x_i\right) = \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r} \tag{1.6}$$

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

The bias and MSE of $\hat{\bar{\theta}}_{\text{ratio}}$ up first order approximation are given by Eqs. (1.7) and (1.8), respectively as

$$\text{Bias}(\hat{\theta}_{\text{ratio}}) = \varphi_{r,n}\bar{Y}\left(C_X^2 - C_{YX}\right) \tag{1.7}$$

$$\text{MSE}(\hat{\theta}_{\text{ratio}}) = \varphi_{r,N}\bar{Y}^2 C_Y^2 + \varphi_{r,n}\bar{Y}^2\left(C_X^2 - 2C_{YX}\right) \tag{1.8}$$

where $C_{YX} = \rho_{YX}C_Y C_X$, $\rho_{YX} = \frac{S_{YX}}{S_Y S_X}$, $C_X = \frac{S_X}{\bar{X}}$,

$$S_X = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2}$$

$$S_{YX} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(x_i - \bar{X}\right)},\ \bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i,\ \varphi_{r,n} = \frac{1}{r} - \frac{1}{n}$$

Singh and Horn (2000) utilized information from imputed values for responding and non-responding units as well, thereafter giving study variable the form given by Eq. (1.9).

$$y_{.i} = \begin{cases} \lambda \frac{n}{r} y_i + (1-\lambda)\hat{\beta}x_i & i \in R \\ (1-\lambda)\hat{\beta}x_i & i \in R^c \end{cases} \tag{1.9}$$

Under this method of imputation, estimator of population mean denoted by $\hat{\bar{\theta}}_{\text{cmp}}$ can be derived as

$$\hat{\theta}_{\text{cmp}} = \frac{1}{n}\left(\sum_{i\in R}\left(\frac{n\lambda}{r} y_i + \left(1-\lambda\right)\hat{\beta}x_i\right) + \left(1-\lambda\right)\hat{\beta}\sum_{i\in R^c} x_i\right) = \bar{y}_r\left(\lambda + \left(1-\lambda\right)\frac{\bar{x}_n}{\bar{x}_r}\right) \tag{1.10}$$

The bias and MSE of $\hat{\bar{\theta}}_{\text{cmp}}$ up first-order approximation are given by Eqs. (1.11) and (1.12), respectively as

$$\text{Bias}\left(\hat{\theta}_{\text{cmp}}\right) = (1-\lambda)\varphi_{r,n}\bar{Y}\left(C_X^2 - C_{YX}\right) \tag{1.11}$$

$$\mathrm{MSE}\left(\hat{\theta}_{\mathrm{cmp}}\right)=\varphi_{r,N}\bar{Y}^2C_Y^2+\varphi_{r,n}\bar{Y}^2\left(\left(1-\lambda\right)^2C_X^2-2\left(1-\lambda\right)C_{YX}\right) \quad (1.12)$$

$\hat{\theta}_{\mathrm{cmp}}$ attained optimality when $\lambda=1-C_{YX}/C_X^2$ and the minimum MSE of θ_{cmp} denoted by $\mathrm{MSE}\left(\hat{\theta}_{\mathrm{cmp}}\right)_{\min}$ is given by

$$\mathrm{MSE}\left(\hat{\theta}_{cmp}\right)_{\min}=\bar{Y}^2C_Y^2\left(\varphi_{r,N}-\varphi_{r,n}\rho_{YX}^2\right) \quad (1.13)$$

The efficiency of the estimator $\hat{\theta}_{\mathrm{cmp}}$ proposed by Singh and Horn (2000) reduces as the value of λ reduces to unity due to loss of auxiliary information attributed to such situation.

Singh et al. (2014) proposed Exponential-Type Compromised Imputation method as

$$y_{.i}=\begin{cases} \nu\dfrac{n}{r}y_i+(1-\nu)\bar{y}_r\exp\left(\dfrac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right) & \text{if } i\in R \\ (1-\nu)\bar{y}_r\exp\left(\dfrac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right) & \text{if } i\in R^c \end{cases} \quad (1.14)$$

The point estimator for the population mean $\bar{Y}$ in the case of Exponential-Type Compromised Imputation method is given as

$$\hat{\theta}_{\mathrm{ExpCmp}}=\frac{1}{n}\left(\nu\frac{n}{r}\sum_{i\in R}y_i+r(1-\nu)\bar{y}_r\exp\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right)+\left(n-r\right)(1-\nu)\bar{y}_r\exp\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right)\right) \quad (1.15)$$

$$\hat{\theta}_{\mathrm{ExpCmp}}=\nu\bar{y}_r+(1-\nu)\bar{y}_r\exp\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right) \quad (1.16)$$

The bias and MSE of $\hat{\theta}_{cmp}$ up first order approximation are given by Eqs. (1.17) and (1.18), respectively, as

$$\mathrm{Bias}\left(\hat{\theta}_{\mathrm{ExpCmp}}\right)=(1-\nu)\varphi_{r,N}\bar{Y}\left(\frac{3}{8}C_X^2-\frac{1}{2}C_{YX}\right) \quad (1.17)$$

$$\mathrm{MSE}\left(\hat{\theta}_{\mathrm{ExpCmp}}\right)=\varphi_{r,N}\bar{Y}^2\left(C_Y^2+\frac{(1-\nu)^2}{4}C_X^2-(1-\nu)C_{YX}\right) \quad (1.18)$$

$\hat{\theta}_{\mathrm{ExpCmp}}$ attained optimality when $\nu=1-2C_{YX}/C_X^2$ and the minimum MSE of $\hat{\theta}_{\mathrm{ExpCmp}}$ denoted by $\mathrm{MSE}\left(\hat{\theta}_{\mathrm{ExpCmp}}\right)_{\min}$ is given by

$$\mathrm{MSE}\left(\hat{\theta}_{\mathrm{ExpCmp}}\right)_{\min}=\varphi_{r,N}\bar{Y}^2C_Y^2\left(1-\rho_{YX}^2\right) \quad (1.19)$$

The efficiency of the estimator $\hat{\theta}_{\mathrm{ExpCmp}}$ proposed by Singh et al. (2014) reduces as the value of ν reduces to unity due to loss of auxiliary information attributed to such situation.

Estimators defined in Eqs. (1.6), (1.10), and (1.16) are ratio-based estimators and this implies that their efficiency depend on the strong correlation between the study and auxiliary variables unlike regression approach that is less sensitive to the level of correlation.

2. Proposed Estimator under Imputation

In this study, we have proposed an exponential type imputation method using regression approach defined as

$$y_{.i}=\begin{cases} y_i & \text{if } i\in R \\ \bar{y}_r\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)\exp\left(\dfrac{\lambda\left(\bar{X}-\bar{x}_r\right)}{\lambda\left(\bar{X}+\bar{x}_r\right)+2\eta}\right) & \text{if } i\in R^c \end{cases} \quad (2.1)$$

The point estimator of the population mean under this method of imputation is given by

$$\hat{\theta}_{\mathrm{PExp}}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\bar{y}_r\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)\exp\left(\frac{\lambda\left(\bar{X}-\bar{x}_r\right)}{\lambda\left(\bar{X}+\bar{x}_r\right)+2\eta}\right) \quad (2.2)$$

By varying λ and η, the following estimators are obtained

$$\hat{\theta}_{\mathrm{PExp}}\left(\lambda=1,\eta=1\right)\rightarrow\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\bar{y}_r\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)\exp\left(\frac{\left(\bar{X}-\bar{x}_r\right)}{\left(\bar{X}+\bar{x}_r\right)+2}\right)=\hat{\theta}_{\mathrm{PExp1}} \quad (2.3)$$

$$\hat{\theta}_{\mathrm{PExp}}\left(\lambda=1,\eta=-1\right)\rightarrow\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\bar{y}_r\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)\exp\left(\frac{\left(\bar{X}-\bar{x}_r\right)}{\left(\bar{X}+\bar{x}_r\right)-2}\right)=\hat{\theta}_{\mathrm{PExp2}} \quad (2.4)$$

$$\hat{\theta}_{\mathrm{PExp}}\left(\lambda=1,\eta=0\right)\rightarrow\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\bar{y}_r\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)\exp\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right)=\hat{\theta}_{\mathrm{PExp3}} \quad (2.5)$$

$$\hat{\theta}_{\text{PExp}}\left(\lambda=0,\eta\in\Re\right)\rightarrow\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\bar{y}_r$$

$$\left(w_1+w_2\left(\bar{X}-\bar{x}_r\right)\right)=\hat{\theta}_{\text{PExp4}} \quad (2.6)$$

To obtain bias and MSE of the proposed estimator up to the first order of approximation, we use following transformations as follows:

$$\bar{y}_r=\bar{Y}\left(1+\varepsilon_0\right)\Rightarrow\varepsilon_o=\left(\frac{\bar{y}_r-\bar{Y}}{\bar{Y}}\right),$$

$$\bar{x}_r=\bar{X}\left(1+\varepsilon_1\right)\Rightarrow\varepsilon_1=\left(\frac{\bar{x}_r-\bar{X}}{\bar{X}}\right)$$

such that $\left|\varepsilon_i\right|<1, \quad \forall \ \mathrm{i}=\left(0,\ 1\right).$

The expected values of ε_i up to first-order approximation are given by Eq. (2.7) as

$$\left.\begin{array}{l}E\left(\varepsilon_0\right)=E\left(\varepsilon_1\right)=0,E\left(\varepsilon_0^2\right)=\theta_{r,N}C_Y^2\\ E\left(\varepsilon_1^2\right)=\theta_{r,N}C_X^2,\mathrm{E}\left(\varepsilon_0\varepsilon_1\right)=\theta_{r,N}\rho\, C_XC_Y\end{array}\right\} \quad (2.7)$$

Under the above transformations Eq. (2.2) is defined as

$$\hat{\theta}_{\Pr Exp}=\bar{Y}\left(\frac{r}{n}\left(1+\varepsilon_0\right)+\left(1-\frac{r}{n}\right)\left(\begin{array}{l}w_1-\left(\mu w_1+w_2\bar{X}\right)\varepsilon_1+w_1\varepsilon_0+\\ \left(\frac{3}{2}\mu^2w_1+\mu w_2\bar{X}\right)\varepsilon_1^2-\left(\mu w_1+w_2\bar{X}\right)\varepsilon_0\varepsilon_1\end{array}\right)\right) \quad (2.8)$$

where $\mu=\dfrac{\lambda\bar{X}}{2\left(\lambda\bar{X}+\eta\right)}$

Subtract $\bar{Y}$ from both sides of Eq. (2.8), we obtained

$$\hat{\theta}_{\Pr Exp}-\bar{Y}=\bar{Y}\left(\left(\frac{r}{n}-1\right)+\left(\frac{r}{n}+w_1\frac{r}{n}\right)\varepsilon_0+\left(1-\frac{r}{n}\right)\left(\begin{array}{l}w_1-\left(\mu w_1+w_2\bar{X}\right)\varepsilon_1+\left(\frac{3}{2}\mu^2w_1+\mu w_2\bar{X}\right)\varepsilon_1^2\\ -\left(\mu w_1+w_2\bar{X}\right)\varepsilon_0\varepsilon_1\end{array}\right)\right) \quad (2.9)$$

Now, taking expectation both the sides of Eq. (2.9) and apply the results of Eq. (2.7), we get the bias of proposed estimator $\hat{\theta}_{\text{PrExp}}$ as

$$\mathrm{Bias}\left(\hat{\theta}_{\text{PrExp}}\right)=\bar{Y}\left(\left(\frac{r}{n}-1\right)+\left(1-\frac{r}{n}\right)\varphi_{r,N}\left(\begin{array}{l}\left(\frac{3}{2}\mu^2w_1+\mu w_2\bar{X}\right)C_X^2\\ +\left(\mu w_1+w_2\bar{X}\right)C_{YX}\end{array}\right)\right) \quad (2.10)$$

Square and take expectation both the sides of Eq. (2.9) and then apply the results of Eq. (2.7), we get the MSE of proposed estimator $\hat{\theta}_{\text{PrExp}}$ as

$$\mathrm{MSE}\left(\hat{\theta}_{\text{PrExp}}\right)=\bar{Y}^2\left(1-\frac{2r}{n}+\frac{r^2}{n^2}\left(1+\varphi_{r,N}C_Y^2\right)+w_1^2A+w_2^2B+2w_1C+2w_2D+2w_1w_2E\right) \quad (2.11)$$

where

$$A=\left(1-\frac{r}{n}\right)^2\left(1+\theta_{r,N}\left(C_Y^2+4\mu^2C_X^2-4\mu C_{YX}\right)\right) \quad (2.12)$$

$$B=\left(1-\frac{r}{n}\right)^2\varphi_{r,N}\bar{X}C_X^2 \quad (2.13)$$

$$C=\left(1-\frac{r}{n}\right)\left(\frac{r}{n}\left\{1+\varphi_{r,N}\left(\frac{3}{2}\mu^2C_X^2-2\mu C_{YX}\right)\right\}-\left\{1+\varphi_{r,N}\left(\frac{3}{2}\mu^2C_X^2-\mu C_{YX}\right)\right\}\right) \quad (2.14)$$

$$D=\left(1-\frac{r}{n}\right)\left(\frac{r}{n}\varphi_{r,N}\bar{X}\left(\mu C_X^2-2C_{YX}\right)+\varphi_{r,N}\bar{X}\left(\mu C_X^2-C_{YX}\right)\right) \quad (2.15)$$

$$E=2\left(1-\frac{r}{n}\right)^2\varphi_{r,N}\bar{X}\left(\mu C_X^2-C_{YX}\right) \quad (2.16)$$

To obtain the expression for the optimum value of w_1 and w_2, we partially differentiate $\mathrm{MSE}\left(\hat{\theta}_{\text{PrExp}}\right)$ with respect to w_1 and w_2, and then equate the results to zero as

$$w_1=\frac{-\left(C+w_2E\right)}{A} \quad (2.17)$$

$$w_2 = \frac{-\left(D + w_1 E\right)}{B} \quad (2.18)$$

Solve Eqs. (2.17) and (2.18) simultaneously, the expressions for optimum values of $w_i, i = 1,2$ denoted by $w_i^{\text{opt}}, i = 1,2$ are obtained as

$$w_1^{\text{opt}} = \frac{DE - BC}{AB - E^2} \quad (2.19)$$

$$w_2^{\text{opt}} = \frac{CE - AD}{AB - E^2} \quad (2.20)$$

Substituting Eqs. (2.15) and (2.16) in Eq. (2.7), we get the minimum MSE of proposed estimator $\hat{\theta}_{\text{PrExp}}$ denoted by $\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right)_{\min}$ as

$$\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right)_{\min} = \bar{Y}^2\left(1 - \frac{2r}{n} + \frac{r^2}{n^2}\left(1 + \varphi_{r,N} C_Y^2\right) - \frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right) \quad (2.21)$$

3. Efficiency Comparison

1.) $\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right) < \text{MSE}\left(\hat{\theta}_{\text{mean}}\right)$ if (3.1) is satisfied

$$\left[\left\{1 - \frac{2r}{n} + \frac{r^2}{n^2}\left(1 + \varphi_{r,N} C_Y^2\right)\right\} - \frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right] - \varphi_{r,N} C_Y^2 < 0 \quad (3.1)$$

2.) $\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right) < \text{MSE}\left(\hat{\theta}_{\text{ratio}}\right)$ if Eq. (3.2) is satisfied

$$\left[\left\{1 - \frac{2r}{n} + \frac{r^2}{n^2}\left(1 + \varphi_{r,N} C_Y^2\right)\right\} - \frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right] - \varphi_{r,N} C_Y^2 + \varphi_{r,n}\left(C_X^2 - 2C_{YX}\right) < 0 \quad (3.2)$$

3.) $\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right) < \text{MSE}\left(\hat{\theta}_{\text{cmp}}\right)$ if Eq. (3.3) is satisfied

$$\left[\left\{1 - \frac{2r}{n} + \frac{r^2}{n^2}\left(1 + \varphi_{r,N} C_Y^2\right)\right\} - \frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right] - C_Y^2\left(\varphi_{r,N} - \varphi_{r,n}\rho_{YX}^2\right) < 0 \quad (3.3)$$

4.) $\text{MSE}\left(\hat{\theta}_{\text{PrExp}}\right) < \text{MSE}\left(\hat{\theta}_{\text{Exp}}\right)$ if (3.4) is satisfied

$$\left(1 - \frac{2r}{n} + \frac{r^2}{n^2}\left(1 + \varphi_{r,N} C_Y^2\right) - \frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right) - \varphi_{r,N} C_Y^2\left(1 - \rho_{XY}^2\right) < 0 \quad (3.4)$$

4. Numerical Illustration

For the empirical justification of the results, we consider four sets of real data. The performance of the proposed estimator is justified by comparing its percentage relative efficiency (PRE) to those of some existing estimators considered in the study. PRE is computed using the formula

$$\text{PRE}(\hat{\theta}_j) = \frac{\text{MSE}\left(\hat{\theta}_{\text{mean}}\right)}{\text{MSE}\left(\hat{\theta}_j\right)} \times 100 \quad (4.1)$$

where $\hat{\theta}_j = \hat{\theta}_{\text{mean}}, \hat{\theta}_{\text{ratio}}, \hat{\theta}_{\text{comp}}, \hat{\theta}_{\text{Exp}}, \hat{\theta}_{\text{PrExp}}$

Population 1: [**Source:** Singh (2000)]

X: the number of laborers (in thousands), Y: quantity of raw materials required (in lakhs of bales).

$N = 3055, n = 611, r = 520, \bar{Y} = 308582.4,$

$\bar{X} = 56.5, S_y = 425312.8, S_x = 72.3$

$S_{yx} = 20817828.5, \rho = 0.677$

Population 2: (Source: [1])

X: the number of household income earners, Y: the household net disposal income

$N = 8011, n = 400, r = 360, \bar{Y} = 28229.43,$

$\bar{X} = 1.69, S_y = 22216.56, S_x = 0.78$

$S_{yx} = 7971.302, \rho = 0.46$

Population 3: [Source: Mukhopadhyaya (2000)]

$N = 20, n = 7, r = 5, \bar{Y} = 41.50, \bar{X} = 441.95,$

$S_y^2 = 95.937, S_x^2 = 10215.21$

$S_{yx} = 644.8782906, \rho = 0.6521$

Tables 1 and 2 show the numerical results of (MSE and PRE) of $\hat{\theta}_{\text{mean}}, \hat{\theta}_{\text{ratio}}, \hat{\theta}_{\text{comp}}, \hat{\theta}_{\text{Exp}}, \hat{\theta}_{\text{PrExp}}$ estimators using population sets 1, 2, and 3. Of all the estimators considered in the study, the proposed estimator has minimum MSE and maximum PRE for all

Table 1. MSE of $\hat{\theta}_{mean}, \hat{\theta}_{ratio}, \hat{\theta}_{comp}, \hat{\theta}_{Exp}, \hat{\theta}_{PrExp}$ using population 1, 2, and 3.

Estimators	MSE		
	Population 1	Population 2	Population 3
$\hat{\theta}_{mean}$	288655816	1309431	14.39055
$\hat{\theta}_{ratio}$	268185247	1282612	12.60977
$\hat{\theta}_{cmp}$	264909782	1280420	12.05937
$\hat{\theta}_{Exp}$	156356484	1032355	8.271193
$\hat{\theta}_{PrExp1}$	83007982	796537.3	2.391907
$\hat{\theta}_{PrExp2}$	83017378	796655.5	2.391997
$\hat{\theta}_{PrExp3}$	83012626	796585	2.391952
$\hat{\theta}_{PrExp4}$	82662777	796433.1	2.358922

Table 2. PRE of $\hat{\theta}_{mean}, \hat{\theta}_{ratio}, \hat{\theta}_{comp}, \hat{\theta}_{Exp}, \hat{\theta}_{PrExp}$ using population 1, 2, and 3.

Estimators	PRE		
	Population 1	Population 2	Population 3
$\hat{\theta}_{mean}$	100	100	100
$\hat{\theta}_{ratio}$	107.633	102.091	114.1222
$\hat{\theta}_{cmp}$	108.9638	102.2658	119.3309
$\hat{\theta}_{Exp}$	184.6139	126.8392	173.984
$\hat{\theta}_{PrExp1}$	347.7447	164.3904	601.635
$\hat{\theta}_{PrExp2}$	347.7053	164.366	601.6125
$\hat{\theta}_{PrExp3}$	347.7252	164.3806	601.6238
$\hat{\theta}_{PrExp4}$	349.1969	164.4119	610.0477

the population sets. This implies that the proposed method demonstrates high level of efficiency over others and can produce better estimate of population mean in the presence of non-response on the average.

5. Conclusion

From the results of the empirical study, it was obtained that the proposed estimator is more efficient than other estimators considered in the study and therefore, it is recommended for use for estimating population mean when certain values of study variables are missing.

References

Diana G, Perri PF. Improved estimators of the population mean for missing data. Commun Stat Theory Methods 2010; 39:3245–51.

Hansen MN, Hurwitz WN. The problem of non-response in sample surveys. J Am Stat Assoc 1946; 41:517–29.

Kadilar C, Cingi H. Estimators for the population mean in the case of missing data. Commun Stat Theory Methods 2008; 37:2226–36.

Mishra P, Singh P, Singh R. A generalized class of estimators for estimating population mean using imputation technique. J Reliab Stat Stud 2017; 10(2):33–41.

Mukhopadhyaya P. Theory and methods of survey sampling. Prentice Hall of India Pvt. Ltd., New Delhi, India, 2000.

Prasad S. An exponential imputation in the case of missing data. J Stat Manag Syst 2017; 20(6):1127–40.

Singh S. A new method of imputation in survey sampling. Statistics 2009; 43:499–511.

Singh S, Horn S. Compromised imputation in survey sampling. Metrika 2000; 51:266–76.

Singh BK, Gogoi U. Estimation of population mean using exponential dual to ratio type compromised imputation for missing data in survey sampling. J Stat Appl Pro 6(3):515–22.

Singh AK, Singh P, Singh VK. Exponential-type compromised imputation in survey sampling. J Stat Appl 2014; 3(2):211–7.

Singh GN, Priyankaa K, Kim J-M, Singh S. Estimation of population mean using imputation techniques in sample surveys. J Korean Stat Soc 2010; 39:67–74.