STATISTICS IN TRANSITION new series, March 2019
Vol. 20 No. 1, pp. 21–40, DOI 10.21307/stattrans-2019-002

IMPUTATION OF MISSING VALUES BY USING RAW MOMENTS

Muhammed Umair Sohail[1] Javid Shabbir[2], Fariha Sohil[3]

ABSTRACT

The estimation of population parameters might be quite laborious and inefficient, when the sample data have missing values. In comparison follow-up visits, the method of imputation has been found to be a cheaper procedure from a cost point of view. In the present study, we can enhance the performance of imputation procedures by utilizing the raw moments of the auxiliary information rather than their ranks, especially, when the ranking of the auxiliary variable is expensive or difficult to do so. Equations for bias and mean squared error are obtained by large sample approximation. Through the numerical and simulation studies it can be easily understood that the proposed method of imputation can outperform their counterparts.

Key words: non-response, imputation, raw moments, relative efficiency.
Mathematical classification: 62D05.

1. Introduction

In survey sampling, the common problem which is faced by most of social sciences, economic and scientific studies is the item or unit non-response or missing values. The main reason of the non-response is the sensitive or embarrassing nature of the questions which are relevant to the variable of interest. Usually respondents hesitate to respond to questions related to the sensitive issues, such as age, income, tax returns etc., or due to summer vocations remain a problematic issues in survey sampling. The best available sources need to be utilized for reducing the non-response rate as much as possible. In most of social studies, item or unit non-response mislead the researchers about the effective inference about the problem of interest. Usually the missing values can create a problem, when the follow-up visits are expensive, population is highly dispersed over the frame or difficult to reach. Alternatively, imputation is the most cheapest and easiest procedure to impute the non-responses by appropriate use of the auxiliary information, which is correlated with the variable of interest.
In the last few decades, several methods of imputation have been proposed to handle out such problems in an effective manner. Among them Rubin (1976) was the first who considered a comprehensive examination of non-response and explain

[1]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan. E-mail: umairsohailch@gmail.com. ORCID ID: https://orcid.org/0000-0002-5440-126X.
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
[3]Department of Education, Government College University, Faisalabad, Pakistan.

the different models under which it would occur, such as missing at random (MAR), observed at random (OAR) and if the prior distributions are specified (PDS). Heitjan and Basu (1996) and Ahmed et al. (2006) provided different imputation procedures by correct use of the auxiliary information after Rubin (1976). The problem of non-response under ranked set sampling, when the ranking of observation units is inexpensive was discussed by Herrera and Al-Omari (2011). They consider the problem of missing values under the hot deck (HD) imputation strategy by the significant use of supplementary information. Grover and Kaur (2014) provide an alternative estimation procedure by combination the features of the proposed estimators by Rao (1991) and Bahl and Tuteja (1991) to provide better results than existing ones. An extensive discussion on item and unit non-response was considered by Little and Rubin (2014) in their text. They explained a different method of imputation in significant manners with suitable real life examples. Recently, Mohamed et al. (2016) provided an efficient model for handling the problem of non-response by using multi auxiliary information. Haq et al. (2017) suggested an estimation procedure for the estimation of population mean by using the ranks of the supplementary information. Sohail et al. (2017) considered the problem of scrambled non-response for the estimation of population mean and suggested a class of estimators by modifying the existing ones.
Motivated by Mohamed et al. (2016) and Sohail et al. (2017), in the present study, we appraise the problem of missing completely at random (MCAR), i.e. the probability of obtaining the response from i^{th} unit does not depend on x_i, y_i or survey design and the respondents units are representative of the selected sample for the estimation of population mean. The objective of the study is to provide an alternative procedure for those situations where the ranking of the auxiliary information is expensive or difficult to create. The proposed model not only provides more better results in terms of efficiency than Grover and Kaur (2014) and Haq et al. (2017) estimators but is also easier to understand than others.

The rest of article is structured as follows: In Section 3, we discuss some existing estimators in the literature for the imputation of missing values. In Section 4, we propose an estimator by utilizing the second raw moment of the auxiliary variable for imputing the missing values. The numerical and empirical studies are considered in Section 6. We conclude our study in Section 7.

2. Notations

Let r^* be the total number of the respondents (individuals or items) who belong to group G in sample (s) and $(n-r^*)$ are those who do not provide the respond, are belong to group G^c. So, $s = G \cup G^c$, and it is also assumed that $\hat{\bar{Y}}_r^* = \frac{1}{r^*}\sum_{j=1}^{r^*}\hat{Y}_j$ is the sample mean of the study variable obtained from respondent units in group G.

Let $\bar{X} = \sum_{j=1}^{N} X_j/N, \bar{R} = \sum_{j=1}^{N} R_j/N$ and $\bar{U} = \sum_{j=1}^{N} U_j/N$ be the population mean of the auxiliary variable, rank variable and second raw moment, respectively, and also let $\bar{x}_{r^*} = \sum_{j=1}^{r^*} x_j/r^*, \bar{r}_{r^*} = \sum_{j=1}^{r^*} r_j/r^*$ and $\bar{u}_{r^*} = \sum_{j=1}^{r^*} u_j/r^*$ be the sample mean of the

auxiliary variable, ranked variable and second raw moment, respectively, from the respondent group.
For evaluating the mathematical expressions for bias and mean square error of the existing and proposed estimators, we defined some useful notations as follows:
Let

$$\begin{aligned} e_0 &= \frac{\hat{\bar{Y}}_{r^*}}{\bar{Y}} - 1, \quad e_1 = \frac{\bar{x}_{r^*}}{\bar{X}} - 1, \quad e_3 = \frac{\bar{r}_{r^*}}{\bar{R}} - 1, \quad e_5 = \frac{\bar{u}_{r^*}}{\bar{U}} - 1, \text{ such that} \\ E(e_i) &= 0 \quad \text{for} \quad i = 0,1,3,5. \end{aligned}$$

and

$$\begin{aligned} E\left(e_0^2\right) &= \theta_{r^*,N} C_y^2, \quad E\left(e_1^2\right) = \theta_{r^*,N} C_x^2,, \quad E\left(e_3^2\right) = \theta_{r^*,N} C_r^2, \quad E\left(e_5^2\right) = \theta_{r^*,N} C_u^2, \\ E\left(e_0 e_1\right) &= \theta_{r^*,N} \rho_{xy} C_x C_y, \quad E\left(e_0 e_3\right) = \theta_{r^*,N} \rho_{ry} C_y C_r, \quad E\left(e_0 e_5\right) = \theta_{r^*,N} \rho_{uy} C_u C_y, \\ E\left(e_1 e_3\right) &= \theta_{r^*,N} \rho_{xr} C_x C_r, \quad E\left(e_1 e_5\right) = \theta_{r^*,N} \rho_{xu} C_u C_x, \quad E\left(e_3 e_5\right) = \theta_{r^*,N} \rho_{ru} C_u C_r, \end{aligned}$$

where

$$\begin{aligned} \tau &= \frac{1}{N} \sum_{j=1}^{N} \tau_j, \quad C_\tau^2 = \frac{\sigma_\tau^2}{\bar{\tau}^2}, \quad \rho_{\tau\psi} = \frac{S_{\tau\psi}}{S_\tau S_\psi}, \quad \theta_{r^*,N} = \left(\frac{1}{r^*} - \frac{1}{N}\right), \\ S_{\tau\psi} &= \frac{1}{N-1} \sum_{j=1}^{N} (\tau_j - \bar{\tau})(\psi_j - \bar{\psi}), \quad \text{where} \quad \tau, \psi = R, U, X, Y. \end{aligned}$$

3. Some Existing Methods of Imputation

In this section, we discuss some existing methods of imputation, which are available in the literature and commonly used for estimation of the population mean. These are defined below.
• Under mean imputation approach

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j \varepsilon G \\ \hat{\bar{Y}}_{r^*} & \text{if } j \varepsilon G^c, \end{cases} \tag{1}$$

The point estimator for population mean ($\bar{Y}$) is given by

$$\hat{\bar{Y}}_M = \frac{1}{n} \left[\sum_{j=1}^{r^*} \hat{Y}_j + \sum_{j=1}^{n-r^*} \hat{Y}_j \right] = \hat{\bar{Y}}_{r^*} \tag{2}$$

The variance of the mean estimator is given by:

$$Var(\hat{\bar{Y}}_M) = \theta_{r^*,N} \bar{Y}^2 C_y^2. \tag{3}$$

• Cochran (1940) suggested the ratio estimator for the estimation of the population mean. We can rewrite it for imputing missing values as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j\varepsilon G \\ \frac{1}{1-f_1}\left[\frac{\hat{\bar{Y}}_{r^*}}{\bar{x}_{r^*}}\bar{X} - f_1\hat{\bar{Y}}_{r^*}\right] & \text{if } j\varepsilon G^c, \end{cases} \tag{4}$$

where $f_1 = \frac{r^*}{n}$ and $\bar{X}$ are the population mean of the auxiliary variable. The point estimator is given as:

$$\hat{\bar{Y}}_R = \hat{\bar{Y}}_{r^*}\frac{\bar{X}}{\bar{x}_{r^*}}. \tag{5}$$

The ratio estimator is conditionally more efficient as compared to the mean estimator when the correlation between y and x is positive. The bias and the mean square error are given by

$$Bias(\hat{\bar{Y}}_R) \cong \theta_{r^*,N}\bar{Y}\left(C_x^2 - \rho_{yx}C_yC_x\right) \tag{6}$$

and

$$MSE(\hat{\bar{Y}}_R) \cong \theta_{r^*N}\bar{Y}^2\left(C_y^2 + C_x^2 - 2\rho_{yx}C_yC_x\right). \tag{7}$$

• Bahl and Tuteja (1991) proposed the ratio-exponential type estimator for imputing non-response, is expressed as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j\varepsilon G \\ \frac{1}{1-f_1}\left[\hat{\bar{Y}}_{r^*}\exp\left(\frac{\bar{X}-\bar{x}_{r^*}}{\bar{X}+\bar{x}_{r^*}}\right) - f_1\hat{\bar{Y}}_{r^*}\right] & \text{if } j\varepsilon G^c, \end{cases} \tag{8}$$

The point estimator is given by:

$$\hat{\bar{Y}}_{B.T-R} = \hat{\bar{Y}}_{r^*}\exp\left(\frac{\bar{X}-\bar{x}_{r^*}}{\bar{X}+\bar{x}_{r^*}}\right), \tag{9}$$

with bias and mean squared error

$$Bias(\hat{\bar{Y}}_{B.T-R}) \cong \theta_{r^*,N}\bar{Y}\left(\frac{3}{8}C_x^2 - \frac{1}{2}\rho_{yx}C_yC_x\right) \tag{10}$$

and

$$MSE(\hat{\bar{Y}}_{B.T-R}) \cong \frac{1}{4}\theta_{r^*N}\bar{Y}^2\left(4C_y^2 + C_x^2 - 4\rho_{yx}C_yC_x\right). \tag{11}$$

The product-exponential type estimator for imputing the missing values is given by

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j \varepsilon G \\ \frac{1}{1-f_1}\left[\hat{\bar{Y}}_{r^*} \exp\left(\frac{\bar{x}_{r^*}-\bar{X}}{\bar{x}_{r^*}+\bar{X}}\right) - f_1 \hat{\bar{Y}}_{r^*}\right] & \text{if } j \varepsilon G^c, \end{cases} \tag{12}$$

The point estimator for the population mean is given as:

$$\hat{\bar{Y}}_{B.T-P} = \hat{\bar{Y}}_{r^*} \exp\left(\frac{\bar{x}_{r^*}-\bar{X}}{\bar{x}_{r^*}+\bar{X}}\right). \tag{13}$$

The bias and mean squared error of $\hat{\bar{Y}}_{B.T-P}$ are

$$Bias(\hat{\bar{Y}}_{B.T-P}) \cong \theta_{r^*,N}\bar{Y}\left(\frac{1}{2}\rho_{yx}C_yC_x - \frac{3}{8}C_x^2\right). \tag{14}$$

and

$$MSE(\hat{\bar{Y}}_{B.T-P}) \cong \frac{1}{4}\theta_{r^*,N}\bar{Y}^2\left(4C_y^2 + C_x^2 + 4\rho_{yx}C_yC_x\right). \tag{15}$$

• The conventional difference estimator is defined as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j \varepsilon G \\ \frac{1}{1-f_1}\left[\hat{\bar{Y}}_{r^*} + k(\bar{X}-\bar{x}_{r^*}) - f_1 \hat{\bar{Y}}_{r^*}\right] & \text{if } j \varepsilon G^c, \end{cases} \tag{16}$$

where k is an un-known constant. The point estimator for the population mean is defined as:

$$\hat{\bar{Y}}_D = \hat{\bar{Y}}_{r^*} + k(\bar{X}-\bar{x}_{r^*}). \tag{17}$$

The optimum value of k i.e. $k_{opt.} = \rho_{yx}(S_y/S_x)$. The minimum MSE($\hat{\bar{Y}}_D$) is given by

$$MSE(\hat{\bar{Y}}_D)_{min.} \cong \theta_{r^*,N}\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right). \tag{18}$$

• Rao (1991) difference type estimator can be reformulated for imputing the missing values, as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j \varepsilon G \\ \frac{1}{1-f_1}\left[v_1\hat{\bar{Y}}_{r^*} + v_2(\bar{X}-\bar{x}_{r^*}) - f_1 \hat{\bar{Y}}_{r^*}\right] & \text{if } j \varepsilon G^c, \end{cases} \tag{19}$$

where v_1 and v_2 are unknown, which are to be determined. The point estimator $\hat{Y}_j$ is given by:

$$\hat{\bar{Y}}_{R.D} = v_1\bar{y}_{r^*} + v_2(\bar{X}-\bar{x}_{r^*}). \tag{20}$$

The optimum values of v_1 and v_2 are

$$v_{1(opt.)} = \frac{1}{1+\theta_{r^*,N}C_y^2\left(1-\rho_{yx}^2\right)} \quad \text{and} \quad v_{2(opt.)} = \frac{\bar{Y}C_y^2\rho_{yx}}{\bar{X}C_x\left(1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)\right)}.$$

The bias and MSE$(\hat{\bar{Y}}_{R.D})_{min.}$ are given by

$$Bias(\hat{\bar{Y}}_{R.D}) \cong \theta_{r^*N}\bar{Y}\left(k_1-1\right) \tag{21}$$

and

$$MSE(\hat{\bar{Y}}_{R.D})_{min.} \cong \frac{\theta_{r^*N}\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right)}{1+\theta_{r^*N}C_y^2\left(1-\rho_{yx}^2\right)}. \tag{22}$$

• In line with Grover and Kaur (2014), we can reformulate the given procedure for the imputation of missing values, as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j\varepsilon G \\ \frac{1}{(n-f_1)}\left[\left(\alpha_1\hat{\bar{Y}}_{r^*}+\alpha_2(\bar{X}-\bar{x}_{r^*})\right)\times \right. & \\ \left. \exp\left(\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\right)-f_1\hat{\bar{Y}}_{r^*}\right] & \text{if } j\varepsilon G^c, \end{cases} \tag{23}$$

where α_1 and α_2 are the suitably chosen constants, where a and b are known parameters of the auxiliary variable, see Table 1, which is described below. The point estimator for the population mean is given as:

$$\hat{\bar{Y}}_{GK}^* = \left[\alpha_1\hat{\bar{Y}}_{r^*}+\alpha_2(\bar{X}-\bar{x}_{r^*})\right]\exp\left[\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\right]. \tag{24}$$

The optimum values of α_1 and α_2 are defined as:

$$\alpha_{1(opt.)} = \frac{8-\theta_{r^*,N}\theta^2C_x^2}{8[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]}$$

and

$$\alpha_{2(opt.)} = \frac{\bar{Y}[\theta_{r^*,N}\theta^3C_x^3+8C_y\rho_{yx}-\theta_{r^*,N}\theta^2C_x^2C_y\rho_{yx}-4\theta C_x\{1-\theta_{r^*,N}C_y^2(1-\rho_{yx})\}]}{8\bar{X}C_x[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]}.$$

where $\theta = \frac{a\bar{X}}{a\bar{X}-b}$. The bias of $\hat{\hat{Y}}^*_{G.K}$ is given as:

$$Bias(\hat{\hat{Y}}^*_{GK}) \cong \theta_{r^*,N}\bar{Y}\left[(\alpha_1 - 1) + \theta_{r^*,N}\theta\alpha_1 C_x\left(\frac{3}{2}C_x - \rho_{yx}C_y\right)\right] + \theta_{r^*,N}\theta\alpha_2\bar{X}C_x^2. \quad (25)$$

Substituting the optimum values of α_1 and α_2, we get the minimum mean squared error of $\hat{\hat{Y}}^*_{GK}$ as follows:

$$MSE(\hat{\hat{Y}}^*_{GK})_{min.} \cong \frac{\theta_{r^*,N}\bar{Y}^2\left[\begin{array}{l}64C_y^2(1-\rho_{yx}^2) - \theta_{r^*,N}\theta^4 C_x^4 \\ \quad - 16\,\theta_{r^*,N}\theta^2 C_x^2 C_y^2(1-\rho_{yx}^2)\end{array}\right]}{64[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]}. \quad (26)$$

• Following Haq et al. (2017), the imputation procedure for imputing the missing values is defined as:

$$\hat{Y}_j = \begin{cases} \hat{Y}_j & \text{if } j\varepsilon G \\ \frac{1}{(n-f_1)}\left[\left(\beta_1\hat{\hat{Y}}_{r^*} + \beta_2(\bar{X}-\bar{x}_{r^*}) + \beta_3(\bar{R}-\bar{r}_{r^*})\right)\right. & \\ \left.\exp\left(\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\right) - f_1\hat{\hat{Y}}_{r^*}\right] & \text{if } j\varepsilon G^c, \end{cases} \quad (27)$$

where β_1, β_2 and β_3 are the unknown constants, these constant values are determined by minimizing the resultant mean squared error. The point estimator for procedure given in (27) is given as:

$$\hat{\hat{Y}}^*_{Haq.} = \left\{\beta_1\bar{y}_{r^*} + \beta_2(\bar{X}-\bar{x}_{r^*}) + \beta_3(\bar{R}-\bar{r}_{r^*})\right\}\exp\left\{\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\right\}. \quad (28)$$

The optimum values of β_1, β_2 and β_3 are given by:

$$\beta_{1(opt.)} = \frac{8-\theta_{r^*,N}\theta^2 C_x^2}{8[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]},$$

$$\beta_{2(opt.)} = \frac{\bar{Y}\left[\begin{array}{l}\theta_{r^*,N}\theta^3 C_x^3(-1+\rho_{xr_x}^2) + (-8C_y + \theta_{r^*,N}\theta^2 C_x^2 C_y)(\rho_{yx} \\ \quad - \rho_{xr_x}\rho_{yr_x}) + 4\,\theta C_x(-1+\rho_{xr_x}^2)[-1+\theta_{r^*,N}C_y^2(1-\rho_{y.xr_x}^2)]\end{array}\right]}{8\bar{X}C_x(-1+\rho_{xr_x}^2)[1+\theta_{r^*,N}C_y^2(1-\rho_{y.xr_x}^2)]}$$

and

$$\beta_{3(opt.)} = \frac{\bar{Y}(8-\theta_{r^*,N}\theta^2 C_x^2)C_y(\rho_{xr_x}\rho_{yx} - \rho_{yr_x})}{8\bar{R}C_r(-1+\rho_{xr_x}^2)[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]}.$$

where $\rho_{y.xr_x}^2 = \frac{\rho_{yx}^2+\rho_{yr_x}^2-2\rho_{yx}\rho_{yr_x}\rho_{xr_x}}{1-\rho_{xr_x}^2}$ is coefficient of multiple determination of Y on X and R.

The bias and minimum $MSE(\hat{\hat{Y}}_{Haq.})$ are given as:

$$Bias(\hat{\hat{Y}}^*_{Haq.}) \cong \frac{1}{8}\Big[-8\bar{Y}+4\theta_{r^*,N}\theta C_x(\bar{X}C_x\beta_1+\bar{U}C_r\beta_3\rho_{rx}) \\ +\bar{Y}\beta_1\Big\{8+\theta_{r^*,N}\theta C_x\big(3\theta C_x-4C_y\rho_{xy}\big)\Big\}\Big]. \quad (29)$$

and

$$MSE(\hat{\hat{Y}}^*_{Haq.})_{min.} \cong \frac{\theta_{r^*,N}\bar{Y}^2\begin{bmatrix}64C_y^2(1-\rho^2_{y.xr_x})-\theta_{r^*,N}\theta^4C_x^4 \\ -16\theta_{r^*,N}\theta^2C_x^2C_y^2(1-\rho^2_{y.xr_x})\end{bmatrix}}{64[1+\theta_{r^*,N}C_y^2(1-\rho^2_{y.xr_x})]}. \quad (30)$$

In Section 4, we propose new procedure for imputing the missing values by utilizing some extra auxiliary information like raw moments.

4. Proposed Method of Imputation

Correct use of auxiliary information about the study variable can enhance the performance of the estimation procedure. If the study and auxiliary variables are correlated with each other, then the second raw moment of the auxiliary variable is also correlated with the study variable. The utilization of the second raw moment is more effective than ranking, especially in those situations, when the ranking of the auxiliary information is done at high cost or is difficult. On the basis of this logic, we propose a new class of the estimators for imputing the missing values by utilizing the second raw moment of the auxiliary variable for the estimation of finite population mean. The suggested class of estimators can incorporate the supplementary information in the form of the second raw moment. Let $\rho_{xu}=S_{xu}/(S_xS_u)$ be the correlation coefficient between X and U.
The imputation procedure for the use of the second raw moment of the auxiliary information is described as follows:

$$\hat{Y}_j=\begin{cases}\hat{Y}_j & \text{if } j\varepsilon G \\ \frac{1}{(n-f_1)}\Big[\Big\{k_1\hat{\hat{Y}}_{r^*}+k_2(\bar{X}-\bar{x}_{r^*})+k_3(\bar{U}-\bar{u}_{r^*})\Big\} & \\ \exp\Big\{\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\Big\}-f_1\hat{\hat{Y}}_{r^*}\Big] & \text{if } j\varepsilon G^c,\end{cases} \quad (31)$$

The point estimator for the population mean for using the above mentioned imputation procedure in (31), is defined as:

$$\hat{\hat{Y}}^*_{Pr}=\Big\{k_1\bar{y}_{r^*}+k_2(\bar{X}-\bar{x}_{r^*})+k_3(\bar{U}-\bar{u}_{r^*})\Big\}\exp\Big\{\frac{a(\bar{X}-\bar{x}_{r^*})}{a(\bar{X}+\bar{x}_{r^*})+2b}\Big\}. \quad (32)$$

where k_1,k_2 and k_3 are suitably chosen constants, which can be determined by minimizing the mean square error. We can rewrite the proposed estimator for imputing the missing values in terms of error as:

$$\hat{\hat{Y}}^{*}_{Pr_1} = \left(k_1\bar{Y}(1+e_0) - k_2\bar{X}e_1 - k_3\bar{U}e_5\right)\left(1 - \frac{\theta}{2}e_1 + \frac{3}{8}\theta^2 e_1^2\right).$$

The bias of the proposed estimator is:

$$\begin{aligned} Bias(\hat{\hat{Y}}^{*}_{Pr}) \cong & \frac{1}{8}\Big[-8\bar{Y} + 4\theta_{r^*,N}\theta C_x\Big(\bar{X}C_x k_1 + \bar{U}C_u k_3\rho_{ux}\Big) \\ & +\bar{Y}k_1\Big\{8 + \theta_{r^*,N}\theta C_x\Big(3\theta C_x - 4C_y\rho_{xy}\Big)\Big\}\Big]. \end{aligned} \quad (33)$$

The mean squared error of the proposed imputation procedure is given as:

$$\begin{aligned} MSE(\hat{\hat{Y}}^{*}_{Pr}) \cong & \bar{Y}^2 + \theta_{r^*,N}\bar{X}C_x k_2\Big(-\bar{Y}\theta + \bar{X}k_1\Big) + \theta_{r^*,N}\bar{U}C_u^2 k_3^2 + \theta_{r^*,N}\bar{U}C_x C_u \\ & \Big(-\bar{Y}\theta + 2\bar{X}k_1\Big) + \bar{Y}^2 k_1^2\Big[1 + \theta_{r^*,N}\Big\{C_y^2 + \theta C_x\Big(\theta C_x - 2C_y\rho_{xy}\Big)\Big\}\Big] \\ & +\frac{1}{4}\bar{Y}k_1\Big[-8\bar{Y} + \theta_{r^*,N}C_x\Big\{\theta C_x\Big(-3\bar{Y}\theta + 8\bar{X}k_2\Big) + 8\bar{U}\theta C_u k_3\rho_{xu} \\ & +4C_y\Big(\bar{Y} - 2\bar{X}k_2\Big)\rho_{xy}\Big\} - 8\bar{U}C_u C_y\theta_{r^*,N}k_3\rho_{uy}\Big]. \end{aligned} \quad (34)$$

The optimum values of the unknown constants $[k_i \,\text{for}\, i = 1,2,3.]$ are determined by minimizing (34), which can be expressed as:

$$k_{1(opt.)} = \frac{8 - \theta_{r^*,N}\theta^2 C_x^2}{8[1 + \theta_{r^*,N}C_y^2(1 - \rho_{yx}^2)]},$$

$$k_{2(opt.)} = \frac{\bar{Y}\left[\begin{array}{r}\theta_{r^*,N}\theta^3 C_x^3(-1+\rho_{xu_x}^2) + (-8C_y + \theta_{r^*,N}\theta^2 C_x^2 C_y)(\rho_{yx} - \rho_{xu_x}\rho_{yu_x}) \\ +4\theta C_x(-1+\rho_{xu_x}^2)\{-1+\theta_{r^*,N}C_y^2(1-\rho_{y.xu_x}^2)\}\end{array}\right]}{8\bar{X}C_x(-1+\rho_{xu_x}^2)[1+\theta_{r^*,N}C_y^2(1-\rho_{y.xu_x}^2)]}$$

and

$$k_{3(opt.)} = \frac{\bar{Y}(8 - \theta_{r^*,N}\theta^2 C_x^2)C_y(\rho_{xu_x}\rho_{yx} - \rho_{yu_x})}{8\bar{U}C_u(-1+\rho_{xu_x}^2)[1+\theta_{r^*,N}C_y^2(1-\rho_{yx}^2)]}.$$

where $\rho_{y.xu_x}^2 = \frac{\rho_{yx}^2 + \rho_{yu_x}^2 - 2\rho_{yx}\rho_{yu_x}\rho_{xu_x}}{1-\rho_{xu_x}^2}$ is coefficient of multiple determination of Y on X and U in simple random sampling.

$$MSE(\hat{\hat{Y}}^{*}_{Pr})_{min.} \cong \frac{\theta_{r^*,N}\bar{Y}^2\left[\begin{array}{r}64C_y^2(1-\rho_{y.xu_x}^2) - \theta_{r^*,N}\theta^4 C_x^4 \\ -16\theta_{rN}\theta^2 C_x^2 C_y^2(1-\rho_{y.xu_x}^2)\end{array}\right]}{64[1+\theta_{r^*,N}C_y^2(1-\rho_{y.xu_x}^2)]}. \quad (35)$$

Table 1: Some special cases of existing and proposed imputation methods

a	b	$\hat{\hat{Y}}^{*}_{GK}$	$\hat{\hat{Y}}^{*}_{Haq.}$	$\hat{\hat{Y}}^{*}_{Pr}$
1	C_x	$\hat{\hat{Y}}^{1}_{GK}$	$\hat{\hat{Y}}^{1}_{Haq.}$	$\hat{\hat{Y}}^{1}_{Pr}$
1	$N\tilde{X}$	$\hat{\hat{Y}}^{2}_{GK}$	$\hat{\hat{Y}}^{2}_{Haq.}$	$\hat{\hat{Y}}^{2}_{Pr}$
$N\tilde{X}$	C_x	$\hat{\hat{Y}}^{3}_{GK}$	$\hat{\hat{Y}}^{3}_{Haq.}$	$\hat{\hat{Y}}^{3}_{Pr}$
C_x	$N\tilde{X}$	$\hat{\hat{Y}}^{4}_{GK}$	$\hat{\hat{Y}}^{4}_{Haq.}$	$\hat{\hat{Y}}^{4}_{Pr}$
1	ρ_{xy}	$\hat{\hat{Y}}^{5}_{GK}$	$\hat{\hat{Y}}^{5}_{Haq.}$	$\hat{\hat{Y}}^{5}_{Pr}$
C_x	ρ_{xy}	$\hat{\hat{Y}}^{6}_{GK}$	$\hat{\hat{Y}}^{6}_{Haq.}$	$\hat{\hat{Y}}^{6}_{Pr}$
ρ_{xy}	C_x	$\hat{\hat{Y}}^{7}_{GK}$	$\hat{\hat{Y}}^{7}_{Haq.}$	$\hat{\hat{Y}}^{7}_{Pr}$
$N\tilde{X}$	ρ_{xy}	$\hat{\hat{Y}}^{8}_{GK}$	$\hat{\hat{Y}}^{8}_{Haq.}$	$\hat{\hat{Y}}^{8}_{Pr}$
ρ_{xy}	$N\tilde{X}$	$\hat{\hat{Y}}^{9}_{GK}$	$\hat{\hat{Y}}^{9}_{Haq.}$	$\hat{\hat{Y}}^{9}_{Pr}$
1	$N\bar{X}$	$\hat{\hat{Y}}^{10}_{GK}$	$\hat{\hat{Y}}^{10}_{Haq.}$	$\hat{\hat{Y}}^{10}_{Pr}$

5. Efficiency Comparison

Here, we define the regulatory conditions under which the proposed estimators can perform better than their existing estimators, which are given by
(i) By (26) and (35), $MSE(\hat{\hat{Y}}_{GK}) - MSE(\hat{\hat{Y}}^{*}_{pr}) > 0$, if

$$\rho_{uy} > \rho_{xu}\rho_{xy} - \sqrt{\rho_{xy}\left(1-\rho_{xu}^{2}\right)\left(1-\rho_{xy}\right)}. \tag{36}$$

(ii) By (30) and (35), $MSE(\hat{\hat{Y}}_{Haq.}) - MSE(\hat{\hat{Y}}^{*}_{pr}) > 0$, if

$$\rho_{uy} > \frac{\sqrt{\left(1-\rho_{xu}^{2}\right)}\left(\rho_{wy}-\rho_{xw}\rho_{xy}\right)}{\sqrt{1-\rho_{xw}^{2}}} + \rho_{xy}\rho_{xu}. \tag{37}$$

Conditions (i) and (ii) are satisfied, then the proposed estimators for imputing the missing responses perform better than their counterparts.

6. Application

For the relative comparison of the proposed class of estimators with existing ones in terms of efficiency, we consider real life as well as simulated data, sets which are discussed in the following subsections.

6.1. Numerical Study

We consider the following four real life data sets for the practical application of the proposed class of estimator and obtained the percentage relative efficiencies of the existing and proposed estimators. The data description is given below as:
Population 1: [Source: Singh (2003)]
y= Estimated number of fish caught by marine recreational fishermen in year 1995 and x = estimated number of fish caught by marine recreational fishermen in year 1994.

$$\begin{aligned} N &= 69, n=40, \bar{Y}=14.0225, \bar{X}=147.0425, \bar{R}=100.5, \bar{U}=28955.59, \\ S_y^2 &= 27.22185, S_x^2=7370.95, S_w^2=3350\, S_u^2=653591180, S_{xy}=350.3902, \\ S_{uy} &= 98116.68\, S_{xu}=2123923, S_{ry}=234.8867, S_{wx}=4959.526, S_{wu}=1438183, \\ \rho_{xy} &= 0.7822, \rho_{uy}=0.7355817, \rho_{wy}=0.7778165, \rho_{xu}=0.967662, \rho_{uw}=0.97193, \\ \rho_{wx} &= 0.998058 \end{aligned}$$

Population 2: [Source: James et al. (2013)]
y= total sales and x = expenditure on TV advertisement

$$\begin{aligned} N &= 200, n=40, \bar{Y}=14.0225, \bar{X}=177.5965, \bar{R}=100.5, \bar{U}=73653530, \\ S_y^2 &= 27.22185, S_x^2=8057.097, S_u^2=4.4\,e^{+16}, S_{xy}=376.3316, S_{uy}=98116.68, \\ S_{xu} &= 1.4\,e^{+12}, S_{ry}=94080.28, S_{wx}=106830.7, S_{wu}=1.4\,e^{+12}, \rho_{xy}=0.9601, \\ \rho_{uy} &= 0.8554, \rho_{wy}=0.7689, \rho_{xu}=0.9283, \rho_{uw}=0.5208, \rho_{wx}=0.75434 \end{aligned}$$

Population 3: [Source: James et al. (2013)]
y= Income and x = education

$$\begin{aligned} N &= 30, n=15, \bar{Y}=16, \bar{X}=50.1455, \bar{R}=15.5, \bar{U}=2946.634, \\ S_y^2 &= 13.2712, S_x^2=446.9652, S_w^2=77.5\, S_u^2=4340687, S_{xy}=74.31184, \\ S_{uy} &= 7344.01, S_{xu}=43477.52, S_{ry}=30.7390, S_{wx}=106830.7, S_{wu}=18115.9, \\ \rho_{xy} &= 0.9648, \rho_{uy}=0.9676, \rho_{wy}=0.9584, \rho_{xu}=0.9283, \rho_{uw}=0.9870, \\ \rho_{wx} &= 0.9925 \end{aligned}$$

Population 4: [Source: James et al. (2013)]
y= Income and x = education + seniority

$$\begin{aligned} N &= 30, n=15, \bar{Y}=15.5, \bar{X}=110.2483, \bar{R}=15.5, \bar{U}=15249.32, \\ S_y^2 &= 729.7176, S_x^2=3201.347, S_w^2=77.5, S_u^2=179829664, S_{xy}=872.8027, \\ S_{uy} &= 186487.9, S_{xu}=741453.5, S_{ry}=130.5645, S_{wx}=491.1011, S_{wu}=1438183, \\ \rho_{xy} &= 0.5710, \rho_{uy}=0.5148, \rho_{wy}=0.5490, \rho_{xu}=0.97720, \rho_{uw}=0.9494, \\ \rho_{wx} &= 0.98594 \end{aligned}$$

For the relative efficiencies of the proposed and existing imputation procedures, we consider the following expression:

$$PRE(.) = \frac{Var(\hat{\hat{Y}}_M)}{MSE(\hat{\hat{Y}}_k)} \quad \text{for} \quad k = G.K, Haq., Pr. \tag{38}$$

To check the relative performance of the given procedures, we consider the response rate between 25% to 80% in all of the four populations. By the use of varying response rate, we are able to illustrate the relative performance of the imputation procedure in an significant way. Based on the results given in Table 2 and 3, we conclude that the estimator $\hat{\hat{Y}}_{GK}$, $\hat{\hat{Y}}_{Haq.}$ and $\hat{\hat{Y}}_{pr}$ remain better as compared to $\hat{\hat{Y}}_M$. At varying response rate, the inter-class efficiency of the available estimators is varying slightly over their entire range. After observing Table 2 and 3 in detail, we can say that there exists an inverse relationship between the response rate and PRE's. At low response rate, all the given estimators can perform better as compared to the mean estimator than a high response rate. For intra-class efficiency, we can observe that the proposed estimators can outperform the existing estimators. At the response rate (25%), PRE of the $\hat{\hat{Y}}_{GK}$, and $\hat{\hat{Y}}_{Haq.}$ is $1411.1340, 1502.4550$ and $261.4669, 262.9224$ for the first and second population, but at the same point PRE of $\hat{\hat{Y}}_{pr}$ is 1608.0930 and 266.3743 respectively. In population 3 and 4, PRE of the existing one is $1507.4520, 1508.4190$ and $154.8800, 156.4693$ respectively. The PRE value of the suggested estimator is 1741.5110 and 164.7871 respectively.
Overall, we can say that, the utilization of the second raw moment of the auxiliary variable has significant effect on the estimation of population parameters rather than utilizing the ranks of the supplementary information, even when the ranking of the auxiliary information is inexpensive.

6.2. Empirical Study

An empirical study of any strategy or procedure is helpful to draw the actual picture of the performance for the respective phenomena by assuming some known value of the population parameters. For empirical illustration of the existing and proposed methods of imputing non-response, we consider the following steps to generate the artificial data sets, which are defined as follows:

• We can generate first two artificial data sets by using the bivariate normal population with mean $\mathbf{A} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$ and variance $\mathbf{V} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$, and last two data sets are generated by using the gamma distribution with **Q**= $\begin{bmatrix} a \\ b \end{bmatrix}$ under following parametric values:

• **Artificial Data Set 1:**

$$\mathbf{A} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 6 & 3 \\ 3 & 8 \end{bmatrix}$$

Table 2: PRE(.) of the existing and proposed estimators by real life data sets

r^*	Estimators			Population 1			Population 2		
				$\hat{\hat{Y}}^*_{GK}$	$\hat{\hat{Y}}^*_{Haq.}$	$\hat{\hat{Y}}^*_{pr}$	$\hat{\hat{Y}}^*_{GK}$	$\hat{\hat{Y}}^*_{Haq.}$	$\hat{\hat{Y}}^*_{pr}$
10	$\hat{Y}^1_{GK}$	$\hat{Y}^1_{Haq.}$	$\hat{Y}^1_{pr}$	1411.1340	1502.5440	1608.0930	261.4669	262.9224	266.3743
	$\hat{Y}^2_{GK}$	$\hat{Y}^2_{Haq.}$	$\hat{Y}^2_{pr}$	1295.5650	1375.8530	1468.0350	258.8331	260.2699	263.6769
	$\hat{Y}^3_{GK}$	$\hat{Y}^3_{Haq.}$	$\hat{Y}^3_{pr}$	1411.2400	1502.6620	1608.2260	261.4941	262.9498	266.4023
	$\hat{Y}^4_{GK}$	$\hat{Y}^4_{Haq.}$	$\hat{Y}^4_{pr}$	1295.6230	1375.9150	1468.1010	258.8329	260.2696	263.6767
	$\hat{Y}^5_{GK}$	$\hat{Y}^5_{Haq.}$	$\hat{Y}^5_{pr}$	1411.1680	1502.5820	1608.1360	261.4577	262.9132	266.3649
	$\hat{Y}^6_{GK}$	$\hat{Y}^6_{Haq.}$	$\hat{Y}^6_{pr}$	1411.1900	1502.6060	1608.1630	261.4323	262.8875	266.3388
	$\hat{Y}^7_{GK}$	$\hat{Y}^7_{Haq.}$	$\hat{Y}^7_{pr}$	1411.1290	1502.5390	1608.0870	261.4594	262.9149	266.3666
	$\hat{Y}^8_{GK}$	$\hat{Y}^8_{Haq.}$	$\hat{Y}^8_{pr}$	1411.2400	1502.6620	1608.2260	261.4941	262.9498	266.4023
	$\hat{Y}^9_{GK}$	$\hat{Y}^9_{Haq.}$	$\hat{Y}^9_{pr}$	1295.5600	1375.8480	1468.0300	258.8330	260.2697	263.6768
	$\hat{Y}^{10}_{GK}$	$\hat{Y}^{10}_{Haq.}$	$\hat{Y}^{10}_{pr}$	1295.5170	1375.8020	1467.9810	258.8331	260.2699	263.6769
20	$\hat{Y}^1_{GK}$	$\hat{Y}^1_{Haq.}$	$\hat{Y}^1_{pr}$	1331.6930	1416.2790	1513.6120	259.2250	260.6695	264.0950
	$\hat{Y}^2_{GK}$	$\hat{Y}^2_{Haq.}$	$\hat{Y}^2_{pr}$	1286.4050	1366.6910	1458.8710	258.1407	259.5774	262.9845
	$\hat{Y}^3_{GK}$	$\hat{Y}^3_{Haq.}$	$\hat{Y}^3_{pr}$	1331.7320	1416.3220	1513.6610	259.2361	260.6807	264.1064
	$\hat{Y}^4_{GK}$	$\hat{Y}^4_{Haq.}$	$\hat{Y}^4_{pr}$	1286.4290	1366.7170	1458.8990	258.1406	259.5773	262.9844
	$\hat{Y}^5_{GK}$	$\hat{Y}^5_{Haq.}$	$\hat{Y}^5_{pr}$	1331.7060	1416.2930	1513.6280	259.2213	260.6657	264.0912
	$\hat{Y}^6_{GK}$	$\hat{Y}^6_{Haq.}$	$\hat{Y}^6_{pr}$	1331.7140	1416.3020	1513.6380	259.2108	260.6552	264.0805
	$\hat{Y}^7_{GK}$	$\hat{Y}^7_{Haq.}$	$\hat{Y}^7_{pr}$	1331.6910	1416.2770	1513.6100	259.2219	260.6664	264.0919
	$\hat{Y}^8_{GK}$	$\hat{Y}^8_{Haq.}$	$\hat{Y}^8_{pr}$	1331.7320	1416.3220	1513.6610	259.2361	260.6807	264.1064
	$\hat{Y}^9_{GK}$	$\hat{Y}^9_{Haq.}$	$\hat{Y}^9_{pr}$	1286.4030	1366.6890	1458.8690	258.1406	259.5774	262.9844
	$\hat{Y}^{10}_{GK}$	$\hat{Y}^{10}_{Haq.}$	$\hat{Y}^{10}_{pr}$	1286.3850	1366.6700	1458.8490	258.1407	259.5774	262.9845
30	$\hat{Y}^1_{GK}$	$\hat{Y}^1_{Haq.}$	$\hat{Y}^1_{pr}$	1306.9350	1389.4510	1484.3000	258.4836	259.9244	263.3413
	$\hat{Y}^2_{GK}$	$\hat{Y}^2_{Haq.}$	$\hat{Y}^2_{pr}$	1283.3520	1363.6380	1455.8170	257.9099	259.3466	262.7537
	$\hat{Y}^3_{GK}$	$\hat{Y}^3_{Haq.}$	$\hat{Y}^3_{pr}$	1306.9560	1389.4730	1484.3250	258.4895	259.9303	263.3473
	$\hat{Y}^4_{GK}$	$\hat{Y}^4_{Haq.}$	$\hat{Y}^4_{pr}$	1283.3650	1363.6510	1455.8310	257.9098	259.3466	262.7536
	$\hat{Y}^5_{GK}$	$\hat{Y}^5_{Haq.}$	$\hat{Y}^5_{pr}$	1306.9420	1389.4580	1484.3080	258.4816	259.9224	263.3392
	$\hat{Y}^6_{GK}$	$\hat{Y}^6_{Haq.}$	$\hat{Y}^6_{pr}$	1306.9460	1389.4620	1484.3130	258.4761	259.9169	263.3336
	$\hat{Y}^7_{GK}$	$\hat{Y}^7_{Haq.}$	$\hat{Y}^7_{pr}$	1306.9340	1389.4500	1484.2990	258.4820	259.9228	263.3396
	$\hat{Y}^8_{GK}$	$\hat{Y}^8_{Haq.}$	$\hat{Y}^8_{pr}$	1306.9560	1389.4730	1484.3250	258.4895	259.9303	263.3473
	$\hat{Y}^9_{GK}$	$\hat{Y}^9_{Haq.}$	$\hat{Y}^9_{pr}$	1283.3510	1363.6370	1455.8160	257.9099	259.3466	262.7536
	$\hat{Y}^{10}_{GK}$	$\hat{Y}^{10}_{Haq.}$	$\hat{Y}^{10}_{pr}$	1283.3420	1363.6270	1455.8050	257.9099	259.3466	262.7537

Table 3: PRE(.) of the existing and proposed estimators by real life data sets

r^*	Estimators			Population 3			Population 4		
				$\hat{\bar{Y}}^*_{GK}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$	$\hat{\bar{Y}}^*_{GK}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$
4	$\hat{\bar{Y}}^1_{GK}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	1507.4520	1508.4190	1741.5110	154.8800	156.4693	164.7871
	$\hat{\bar{Y}}^2_{GK}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	1449.6250	1450.5270	1667.1470	152.4071	153.9680	162.1362
	$\hat{\bar{Y}}^3_{GK}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	1509.2460	1510.2150	1743.8630	152.4055	153.9664	162.1345
	$\hat{\bar{Y}}^4_{GK}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	1449.6130	1450.5150	1667.1330	154.9062	156.4958	164.8154
	$\hat{\bar{Y}}^5_{GK}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	1505.2450	1506.2100	1738.6210	154.8770	156.4663	164.7840
	$\hat{\bar{Y}}^6_{GK}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	1500.3310	1501.2900	1732.1930	154.8499	156.4388	164.7547
	$\hat{\bar{Y}}^7_{GK}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	1507.3880	1508.3550	1741.4270	154.8606	156.4496	164.7662
	$\hat{\bar{Y}}^8_{GK}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	1509.2450	1510.2140	1743.8610	154.9062	156.4958	164.8154
	$\hat{\bar{Y}}^9_{GK}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	1449.6240	1450.5260	1667.1460	152.4057	153.9666	162.1346
	$\hat{\bar{Y}}^{10}_{GK}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	1449.6240	1450.5270	1667.1470	152.4071	153.9680	162.1362
8	$\hat{\bar{Y}}^1_{GK}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	1472.8720	1473.8010	1697.1490	151.1097	152.6824	160.9131
	$\hat{\bar{Y}}^2_{GK}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	1448.9680	1449.8710	1666.4890	150.0889	151.6498	159.8179
	$\hat{\bar{Y}}^3_{GK}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	1473.5970	1474.5270	1698.0940	150.0883	151.6492	159.8172
	$\hat{\bar{Y}}^4_{GK}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	1448.9630	1449.8660	1666.4830	151.1204	152.6933	160.9246
	$\hat{\bar{Y}}^5_{GK}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	1471.9800	1472.9080	1695.9870	151.1085	152.6812	160.9118
	$\hat{\bar{Y}}^6_{GK}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	1469.9870	1470.9130	1693.3920	151.0974	152.6700	160.8998
	$\hat{\bar{Y}}^7_{GK}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	1472.8460	1473.7750	1697.1160	151.1017	152.6744	160.9045
	$\hat{\bar{Y}}^8_{GK}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	1473.5960	1474.5260	1698.0930	151.1204	152.6933	160.9246
	$\hat{\bar{Y}}^9_{GK}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	1448.9680	1449.8700	1666.4890	150.0883	151.6492	159.8173
	$\hat{\bar{Y}}^{10}_{GK}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	1448.9680	1449.8710	1666.4890	150.0889	151.6498	159.8179
12	$\hat{\bar{Y}}^1_{GK}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	1461.6890	1462.6060	1682.8530	149.8684	151.4358	159.6378
	$\hat{\bar{Y}}^2_{GK}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	1448.7490	1449.6520	1666.2700	149.3162	150.8771	159.0451
	$\hat{\bar{Y}}^3_{GK}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	1462.0780	1462.9950	1683.3590	149.3159	150.8768	159.0448
	$\hat{\bar{Y}}^3_{GK}4$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	1448.7460	1449.6490	1666.2670	149.8742	151.4416	159.6440
	$\hat{\bar{Y}}^5_{GK}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	1461.2090	1462.1260	1682.2290	149.8678	151.4351	159.6371
	$\hat{\bar{Y}}^6_{GK}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	1460.1370	1461.0520	1680.8350	149.8618	151.4290	159.6306
	$\hat{\bar{Y}}^7_{GK}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	1461.6750	1462.5920	1682.8350	149.8641	151.4314	159.6332
	$\hat{\bar{Y}}^8_{GK}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	1462.0780	1462.9950	1683.3590	149.8742	151.4416	159.6440
	$\hat{\bar{Y}}^9_{GK}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	1448.7490	1449.6520	1666.2700	149.3159	150.8768	159.0448
	$\hat{\bar{Y}}^{10}_{GK}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	1448.7490	1449.6520	1666.2700	149.3162	150.8771	159.0451

- **Artificial Data Set 2:**

$$\mathbf{A} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 8 & 4 \\ 4 & 10 \end{bmatrix}$$

- **Artificial Data Set 3:**

$$\mathbf{Q} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

- **Artificial Data Set 4:**

$$\mathbf{Q} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

The main purpose of generating the two different data sets from the same distribution is to find the pattern of PRE with respect to their parametric values. In Data sets 3 and 4, the study variable is generated as $y = (r_{xy} \times x) + e$, where $e \sim N(0,1)$ and r_{yx} is the sample correlation coefficient between y and x.

• Here, we can select the sample of size n form N units, randomly, and select randomly r units out of n sample units and impute the dropped units by using the above mentioned imputation procedures, then compute the relevant statistics.

• Repeat the process 30000 (say H) times and obtain the value of $\hat{\bar{Y}}_k^*$. The mean squared error of the given estimator is obtained by using the following expression, as:

$$MSE(\hat{\bar{Y}}_k^*) = \frac{1}{H} \sum_{i=1}^{H} \left((\hat{\bar{Y}}_k^*)_i - \bar{Y} \right)^2 \tag{39}$$

At the specified values of parameters and $n = 50$, the behaviour of normal distribution, gamma distribution and self-generated study variable is shown in Appendix (Figure: 1). By utilizing the artificial data sets, mean squared errors of the given procedures are reported below. On the behalf of numerical findings, which are reported in Tables 4 and 5, we see that the relative performance of the existing and proposed imputation method is similar to the reported results in Table 2 and 3. By the use of simulated data sets (which are generated by bivariate normal and gamma distribution under certain regulatory conditions) the performances of the existing and proposed estimators are better than the mean estimator. As given by the reported results in Table 2 and 3, PRE of respective imputation procedure decreases as the response rate increases, but as a whole these are better than traditional estimators. After comprehensive examination of Tables 4 and 5, we can easily understand that our proposed class of estimators performs significantly better than existing and mean imputation procedures even in high response rate. As the parametric values of the population constants increase in normal population, the performance of all the estimator increase. But in the case of positively dispersed population, there is an inverse relationship between PRE's and parametric values.

Table 4: PRE(.) of existing and proposed estimators by using artificial data set (.).

r^*	Estimator			Artificial Data Sets 1			Artificial Data Sets 2		
				$\hat{\bar{Y}}^*_{G.K}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$	$\hat{\bar{Y}}^*_{G.K}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$
10	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	118.6620	105.9573	119.6891	124.0394	127.1311	132.0750
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	120.2631	106.3291	122.0707	135.0018	134.6481	138.5286
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	101.4103	117.0941	123.6927	115.5257	119.5066	126.8564
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	120.2941	106.3447	122.1189	136.0608	135.7178	141.6042
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	109.6636	116.9660	121.5692	125.0806	128.1636	132.6883
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	113.8250	114.6901	115.1804	126.1819	127.2321	131.1512
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	119.6756	100.1994	120.9550	127.8607	130.1780	132.4522
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	102.8872	117.2912	120.0964	115.6963	119.6452	126.6902
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	115.7024	105.5199	120.5186	136.5683	136.1019	139.8346
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	120.7377	106.3721	121.5206	137.3745	136.8175	138.5315
20	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	117.8110	105.2151	119.8508	120.7829	123.9798	129.3023
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	121.2069	105.0492	120.9690	134.9703	134.5377	137.2016
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	105.2075	119.1198	127.4633	118.2557	121.9277	128.7035
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	120.3429	105.1115	122.1641	134.4532	134.1529	139.9697
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	110.8191	116.4726	118.5771	125.1330	128.1416	132.7055
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	114.3393	114.3480	115.7193	128.5056	130.6469	132.4355
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	120.1568	100.1455	124.3512	128.6391	130.8064	132.7405
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	103.7974	105.2274	106.1870	116.6377	120.5083	127.8116
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	119.6394	105.0234	120.5388	134.5111	134.2217	135.1581
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	116.2761	105.1145	119.2251	134.6683	134.4694	137.5448
30	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	119.7622	105.1340	120.6659	122.5217	125.6854	130.4396
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	119.0610	104.6887	119.7440	134.7711	134.4483	135.1757
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	105.2325	118.6783	127.5033	117.8911	121.6284	128.4568
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	119.8297	104.5238	123.7315	133.8278	133.4614	136.3154
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	112.3861	116.9674	118.0800	124.5436	127.5873	132.0415
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	114.4807	114.5614	115.8958	131.6862	133.5503	135.1994
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	118.2454	100.0115	119.6349	129.0047	131.1589	133.2060
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	105.5192	106.8852	107.7843	117.4968	121.3278	128.5287
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	120.2420	104.7962	120.4320	133.3648	134.0260	135.7968
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	115.5720	104.6800	119.5149	132.7433	132.2578	136.8836

Table 5: PRE(.) of existing and proposed estimators by using artificial data set (.).

r^*	Estimator			Artificial Data Sets 3			Artificial Data Sets 4		
				$\hat{\bar{Y}}^*_{G.K}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$	$\hat{\bar{Y}}^*_{G.K}$	$\hat{\bar{Y}}^*_{Haq.}$	$\hat{\bar{Y}}^*_{pr}$
10	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	192.8977	192.5636	193.1662	136.4316	136.2749	136.7016
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	194.0581	194.0559	194.9967	140.5252	140.5304	140.6584
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	188.0287	187.8559	188.1561	136.3089	136.3474	136.8219
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	191.5353	191.5554	191.7439	140.5487	140.5449	141.2544
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	189.2338	189.3205	189.9834	138.3192	138.3297	139.3096
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	191.3141	191.5745	191.9685	139.2955	139.2150	139.4283
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	192.1016	192.8594	192.8723	139.1179	139.1532	139.4509
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	187.5624	187.3498	187.7200	138.2495	138.2914	140.1726
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	194.5011	194.4885	194.4346	140.8154	140.8160	141.8146
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	194.1044	194.0683	194.5535	138.5544	138.5559	138.7474
20	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	153.2724	152.6555	153.8245	118.6956	118.5047	119.0081
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	154.8089	154.7715	155.7848	140.5252	140.5304	140.7584
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	151.1041	150.8019	151.3284	122.4189	122.4683	122.7324
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	153.5339	153.4891	153.7509	125.0272	125.0842	125.8300
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	153.1577	153.3651	153.8868	123.5171	123.5145	123.5189
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	148.0252	149.0603	149.9335	123.0634	122.0864	123.1859
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	157.3814	157.4595	157.6159	124.3142	124.3697	124.6227
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	152.6729	152.4070	152.8817	122.5352	122.5805	123.9481
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	155.2700	155.2155	156.2721	124.9516	124.9523	125.0465
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	154.7365	154.6740	154.7385	124.1267	124.1129	124.1383
30	$\hat{\bar{Y}}^1_{G.K}$	$\hat{\bar{Y}}^1_{Haq.}$	$\hat{\bar{Y}}^1_{pr}$	137.6494	136.8790	138.3680	112.2768	112.0773	112.6044
	$\hat{\bar{Y}}^2_{G.K}$	$\hat{\bar{Y}}^2_{Haq.}$	$\hat{\bar{Y}}^2_{pr}$	141.3239	141.2753	142.2912	118.8472	118.8500	118.9965
	$\hat{\bar{Y}}^3_{G.K}$	$\hat{\bar{Y}}^3_{Haq.}$	$\hat{\bar{Y}}^3_{pr}$	139.0538	138.7795	139.2670	116.4861	116.5465	116.7900
	$\hat{\bar{Y}}^4_{G.K}$	$\hat{\bar{Y}}^4_{Haq.}$	$\hat{\bar{Y}}^4_{pr}$	141.5240	141.4602	141.5615	119.0015	119.0045	119.0104
	$\hat{\bar{Y}}^5_{G.K}$	$\hat{\bar{Y}}^5_{Haq.}$	$\hat{\bar{Y}}^5_{pr}$	139.7605	139.9683	140.4578	118.0333	118.0357	118.1293
	$\hat{\bar{Y}}^6_{G.K}$	$\hat{\bar{Y}}^6_{Haq.}$	$\hat{\bar{Y}}^6_{pr}$	130.9078	132.1338	132.6111	116.6496	116.5642	116.8055
	$\hat{\bar{Y}}^7_{G.K}$	$\hat{\bar{Y}}^7_{Haq.}$	$\hat{\bar{Y}}^7_{pr}$	139.3605	139.5683	139.8578	118.9265	118.9745	119.8550
	$\hat{\bar{Y}}^8_{G.K}$	$\hat{\bar{Y}}^8_{Haq.}$	$\hat{\bar{Y}}^8_{pr}$	138.2340	137.9276	138.4662	116.2475	116.3058	116.4495
	$\hat{\bar{Y}}^9_{G.K}$	$\hat{\bar{Y}}^9_{Haq.}$	$\hat{\bar{Y}}^9_{pr}$	140.9842	140.9393	141.9538	118.7914	118.7869	118.7950
	$\hat{\bar{Y}}^{10}_{G.K}$	$\hat{\bar{Y}}^{10}_{Haq.}$	$\hat{\bar{Y}}^{10}_{pr}$	142.1794	142.1593	142.2399	118.2523	118.2476	118.6504

7. Conclusions

In this study, we assume that the non-response which occurred in the study is MCAR. Our main objective is to introduce the idea of utilizing the second raw moment of the auxiliary variable for the imputation of missing values, especially for those situations when the ranking of the auxiliary information is difficult or expensive. The proposed imputation method provides better results in terms of efficiency than the existing procedures. From Tables 2, 3, 4 and 5, it can be easily understand that the proposed imputation procedure performs better than Grover and Kaur (2014) and Haq et al. (2017) estimators. Thus, we recommend the proposed estimator for the imputation of missing values and for a precise estimation of the population mean.
The current work can easily be extended to other domains of survey sampling such as the estimation population quartiles (Q_1 and Q_3) and population variance under the stratified and other sampling schemes. Another possible extension of the current work is to estimate the population parameter of the sensitive variable with the non-sensitive auxiliary variable, when the non-response occurs after the utilization of the randomized response model, as in Mohamed et al. (2016) and Sohail et al. (2017). This work is deferred to the later article, which is currently in progress for handling the non-response.

REFERENCES

AHMED, M., AL-TITI, O., AL-RAWI, Z., ABU-DAYYEH, W., (2006). Estimation of a population mean using different imputation methods. Statistics in Transition 7 (6), pp. 1247–1264.

BAHL, S., TUTEJA, R., (1991). Ratio and product type exponential estimators. Journal of Information and Optimization Sciences, 12 (1), pp. 159–164.

COCHRAN, W., (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. The Journal of Agricultural Science, 30 (2), pp. 262–275.

GROVER, L. K., KAUR, P., (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. Communications in Statistics-Simulation and Computation, 43 (7), pp. 1552–1574.

HAQ, A., KHAN, M., HUSSAIN, Z., (2017). A new estimator of finite population mean based on the dual use of the auxiliary information. Communications in Statistics-Theory and Methods, 46 (9), pp. 4425–4436.

HEITJAN, D. F., BASU, S., (1996). Distinguishing "missing at random" and "missing completely at random". The American Statistician, 50 (3), pp. 207–213.

HERRERA, C. N. B., AL-OMARI, A. I., (2011). Ranked set estimation with imputation of the missing observations: the median estimator. Revista Investigacion Operacional, 32 (1), pp. 30–37.

JAMES, G., WITTEN, D., HASTIE, T., AND TIBSHIRANI, R., (2013). An introduction to statistical learning, Vol. 6, Springer.

LITTLE, R. J., RUBIN, D. B., (2014). Statistical analysis with missing data. John WileySons.

MOHAMED, C., SEDORY, S. A., SINGH, S., (2016). Imputation using higher order moments of an auxiliary variable. Communications in Statistics-Simulation and Computation, (just-accepted), pp. 00–00.

RAO, T., (1991). On certain methods of improving ration and regression estimators. Communications in Statistics-Theory and Methods, 20 (10), pp. 3325–3340.

RUBIN, D. B., (1976). Inference and missing data. Biometrika, 63 (3), pp. 581–592.

SINGH, S., (2003). Advanced Sampling Theory With Applications: How Michael Selected Amy, Vol. 2, Springer ScienceBusiness Media.

SOHAIL, M. U., SHABBIR, J., AHMED, S., (2017). Modified class of ratio and regression type estimators for imputing scrambling response, Pakistan, Journal of Statistics, 33 (4), pp. 277–300.

APPENDIX

In Figure 1, we can show the shape of different distributions according to their respective parametric values. In Figure (a), the behaviour of normal distribution is shown according to their respective population parameters. The shape of gamma distribution is expressed in Figure (b) and standard normal distribution is shown in Figure (c). The trend of study variable is shown under the normal and gamma distribution in Figure (d) and (e) respectively. In both Figures, the study variable has an increasing trend.

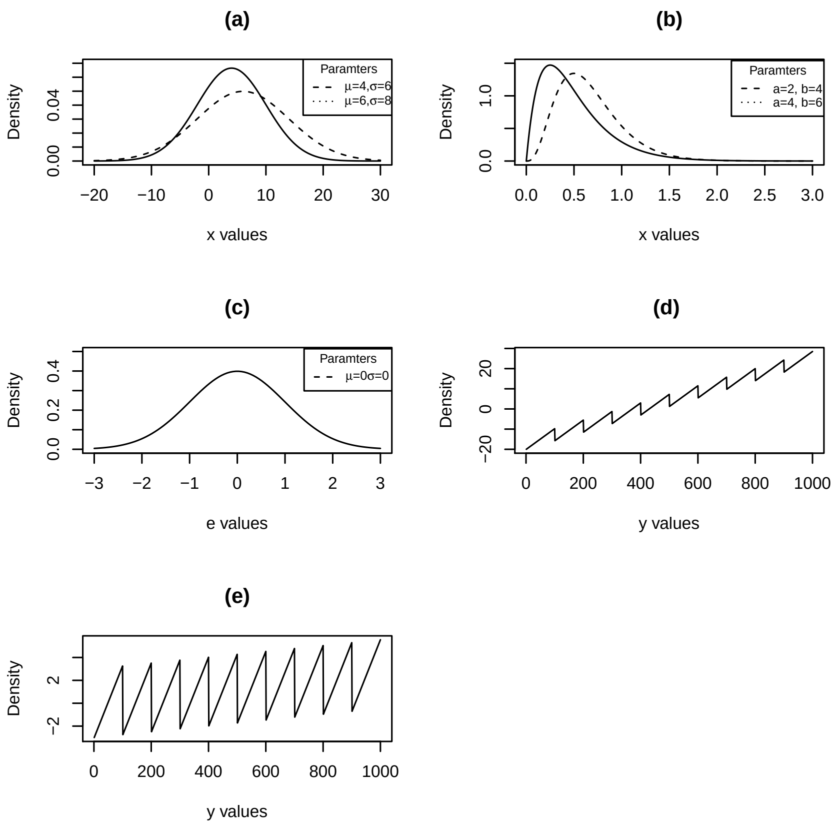

Figure 1: Shape of different distribtions according to their parametric values