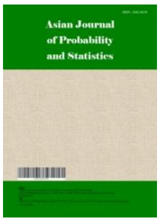

Asian Journal of Probability and Statistics

15(2): 59-74, 2021; Article no.AJPAS.75530
ISSN: 2582-0230

Exponential-Ratio-Type Imputation Class of Estimators using Nonconventional Robust Measures of Dispersions

Ahmed Audu[1*], Mojeed Abiodun Yunusa[1], Aminu Bello Zoramawa[1], Samaila Buda[2] and Ran Vijay Kumar Singh[3]

[1]Department of Mathematics, Usmanu Danfidiyo University, Sokoto, Nigeria.
[2]Department of Physics, Usmanu Danfidiyo University, Sokoto, Nigeria.
[3]Department of Mathematics, Kebbi State University of Science and Technology, Aliero, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2021/v15i230351
Editor(s):
(1) Dr. Manuel Alberto M. Ferreira, Lisbon University, Portugal.
Reviewers:
(1) Adji Achmad Rinaldo Fernandes, Brawijaya University, Indonesia.
(2) Jean Vaillant, University of Antilles, France.
Complete Peer review History: https://www.sdiarticle4.com/review-history/75530

Original Research Article

Received 09 August 2021
Accepted 18 October 2021
Published 28 October 2021

Abstract

Human-assisted surveys, such as medical and social science surveys, are frequently plagued by non-response or missing observations. Several authors have devised different imputation algorithms to account for missing observations during analyses. Nonetheless, several of these imputation schemes' estimators are based on known auxiliary variable parameters that can be influenced by outliers. In this paper, we suggested new classes of exponential-ratio-type imputation method that uses parameters that are robust against outliers. Using the Taylor series expansion technique, the MSE of the class of estimators presented was derived up to first order approximation. Conditions were also specified for which the new estimators were more efficient than the other estimators studied in the study. The results of numerical examples through simulations revealed that the suggested class of estimators is more efficient.

Keywords: Imputation; non-response; estimator; population mean; Mean Squared Error (MSE).

**Corresponding author: Email: ahmed.audu@udusok.edu.ng;*

1 Introduction

Surveys like medical and social science surveys do face the problem of non-response due to involvement of human in data collection. These missing values, due to non-response, in turn create complications in data handling and analysis. Over time, many methods have been developed to address the problem of estimating unknown parameters in the presence of missing values. Imputation is a common technique used to handle situations where data is missing. Missing values can be completed with specific substitutes and data can be analyzed using standard methods. Information about unit of characteristic of interest observed and auxiliary variable help improve the accuracy of demographic parameter estimates. Hansen and Hurtwitz [1] were the first to consider the problem of non-response. Several authors also proposed imputation methods to deal with non-response or missing values. Recent among them include Singh and Deo [2], Toutenburg et al. [3] Kadilar and Cingi [4], Singh [5], Singh and Horn [6], Gira [7], Audu and Singh [8] Kadilar and Cingi [9] Bhushan and Bandey [10], Singh et al. [11], Diana and Perri [12] Al-Omari et al. [13], Audu et al. [14-18], Singh et al. [19]. However, some of the estimators in aforementioned literatures depend on known parameters of the auxiliary variable which are influenced by outliers. In this study, new classes of ratio-type imputation method which utilized parameters that are free from outliers have been presented.

The remaining sections of this article were organized as follows: In section 2, a class of mean imputation schemes to obtain estimators of population mean that are not sensitive to outliers were proposed. They are based on non-conventional robust measures of the auxiliary variable. Distributional properties of the suggested estimators are given in section 3. Conditions for the efficiency of the new estimators with respect to some existing estimators were established in section 4. Simulation studies were presented in section 5 to assess the performance of the proposed scheme estimators with respect to Audu and Singh [8] estimators.

1.1 Notations

The following notations have been used

Y: Study variable.
X: Auxiliary variable.
$\bar{X}, \bar{Y}$: Population mean of the variables X and Y respectively
N: Population Size.
n: Size of the sample
r: Number of respondents.
R: Ratio of the population mean of study variable to population mean of auxiliary variable.
$\bar{x}_n$: The sample mean for the sample of size n.
$\bar{x}_r$: The sample mean of the variable X for set Φ
$\bar{y}_r$: The sample mean of the variable Y for set Φ
S_Y^2, S_X^2 : Population variance of the variables X and Y
S_Y, S_X : Population standard deviation of Y and X.
β_1 : Population coefficient of skewness of X.
β_2 : Population coefficient of kurtosis of X.
ρ_{YX} : Population coefficient of correlation between Y and X.
β_{rg} : Population regression coefficient.
C_Y, C_X : Population coefficient of variation of Y and X.
$G = \frac{4}{N-1}\sum_{i=1}^{N}\left(\frac{2i-N-1}{2N}\right)X_{(i)}$: Gini's mean difference for X.

$D = \frac{2\sqrt{\pi}}{N-1}\sum_{i=1}^{N}\left(i - \frac{N+1}{2N}\right)X_{(i)}$: Downtown's method for X.

$S_{pw} = \frac{\sqrt{\pi}}{N^2}\sum_{i=1}^{N}(2i - N - 1)X_{(i)}$: Probability weighted moments for X.

$\sum$: Population variance covariance matrix.

Let Φ denotes the set of r units response and Φ^c denotes the se of $n-r$ units non-response or missing out of n units sampled without replacement from the N units population. For each $i \in \Phi$, the value of y_i is observed. However, for unit $i \in \Phi^c$, y_i is missing but calculated using different methods of imputation.

Using mean method of imputation, values found missing is to be replaced by the mean of the rest of observed values. The study variable thereafter, takes the form given as,

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

Under the mean method of imputation, sample mean denoted by $\hat{\mu}_0$ can be derived as

$$\hat{\mu}_0 = r^{-1}\sum_{i \in R} y_i \tag{1.2}$$

The variance of $\hat{\mu}_0$ is given by (1.3).

$$Var(\hat{\mu}_0) = \psi_{r,N} S_Y^2 \tag{1.3}$$

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

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

$$y_{.i} = \begin{cases} y_i & \mathrm{i} \in \Phi \\ \hat{\beta}\mathrm{x_i} & \mathrm{i} \in \Phi^c \end{cases} \tag{1.4}$$

Under the ratio method of imputation, estimator of population mean denoted by $\hat{\mu}_1$ can be derived as

$$\hat{\mu}_1 = \hat{\mu}_0 \bar{x}_n \bar{x}_r^{-1} \tag{1.5}$$

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

The MSE of $\hat{\mu}_1$ up $O\left(n^{-1}\right)$ is given as:

$$MSE(\hat{\mu}_1) = MSE\left(\hat{\mu}_0\right) + \psi_{r,n}\left(R^2 S_X^2 - 2RS_{YX}\right) \tag{1.6}$$

where $S_{YX} = \rho_{YX} S_Y S_X$, , $S_X^2 = \left(N-1\right)^{-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2, \bar{X} = N^{-1}\sum_{i=1}^{N} x_i, \psi_{r,n} = r^{-1} - n^{-1}, R = \bar{Y} / \bar{X}$

Singh and Horn [6] utilized information from imputed values for responding and non-responding units as well, thereafter giving study variable the form given by (1.7). The scheme defined in (1.7) is called compromised imputation scheme as the responses of the respondents were computed by using linear combination of information on the study and auxiliary variables.

$$y_{.i} = \begin{cases} \lambda \dfrac{n}{r} y_i + (1-\lambda)\hat{\beta} x_i & \text{i} \in \Phi \\ (1-\lambda)\hat{\beta} x_i & \text{i} \in \Phi^c \end{cases} \tag{1.7}$$

Under this method of imputation, estimator of population mean denoted by $\hat{\mu}_2$ can be derived as

$$\hat{\mu}_2 = \hat{\mu}_0\left(\lambda + \left(1-\lambda\right)\bar{x}_n \bar{x}_r^{-1}\right) \tag{1.8}$$

The performance of $\hat{\mu}_2$ attained optimal when $\lambda = 1 - RS_{YX} / S_X^2$ and the $MSE\left(\hat{\mu}_2\right)_{\min}$ is given by

$$MSE\left(\hat{\mu}_2\right)_{\min} = MSE\left(\hat{\mu}_0\right) - \psi_{r,n}\beta_{rg}\rho_{YX} S_Y S_X \tag{1.9}$$

Nevertheless, the precision of $\hat{\mu}_2$ depends on and λ is function of $\bar{Y}, S_{YX}, \rho_{YX}$, unknown parameters of study variable y which makes $\hat{\mu}_2$ impracticable in real life application

Singh and Deo [2] incorporated power transformation parameter to $\hat{\mu}_2$ and obtain $\hat{\mu}_3$ as

$$\hat{\mu}_3 = \hat{\mu}_0\left(\bar{x}_n / \bar{x}_r\right)^{\alpha} \tag{1.10}$$

The performance of $\hat{\mu}_3$ attained optimal when $\alpha = RS_{YX} / S_X^2$ and the $MSE\left(\hat{\mu}_3\right)_{\min}$ is given by

$$MSE(\hat{\mu}_3) = MSE\left(\hat{\mu}_1\right) - \psi_{r,n} S_X^2\left(\beta_{rg} - R\right)^2 \tag{1.11}$$

where $\beta_{rg} = S_{YX} / S_X^2$

Nevertheless, the precision of $\hat{\mu}_3$ depends on α and α is a function of $\bar{Y}, S_{YX}, \rho_{YX}$, unknown parameters of study variable y which makes $\hat{\mu}_3$ impracticable in real life application.

Kadilar and Cingi [9] modified the work of Kadilar and Cingi [4] in the case of missing observations and suggested the following estimators of population mean

$$\hat{\mu}_4 = \left(\hat{\mu}_0 + \widehat{\beta}_{rg}\left(\bar{X} - \bar{x}_r\right)\right)\bar{X}\,\bar{x}_r^{-1} \tag{1.12}$$

$$\hat{\mu}_5 = \left(\hat{\mu}_0 + \widehat{\beta}_{rg}\left(\bar{X} - \bar{x}_n\right)\right)\bar{X}\,\bar{x}_n^{-1} \tag{1.13}$$

$$\hat{\mu}_6 = \left(\hat{\mu}_0 + \hat{\beta}_{rg}\left(\bar{x}_n - \bar{x}_r\right)\right)\bar{x}_n\,\bar{x}_r^{-1} \tag{1.14}$$

$$MSE(\hat{\mu}_4) = MSE\left(\hat{\mu}_0\right) + \psi_{r,N} S_X^2\left(R^2 - \beta_{rg}^2\right) \tag{1.15}$$

$$MSE(\hat{\mu}_5) = MSE\left(\hat{\mu}_0\right) + \psi_{n,N} S_X^2\left(R^2 - \beta_{rg}^2\right) \tag{1.16}$$

$$MSE(\hat{\mu}_6) = MSE\left(\hat{\mu}_0\right) + \psi_{r,n}\left(S_X^2\left(R + \beta_{rg}\right)^2 - 2\left(R + \beta_{rg}\right)S_{YX}\right) \tag{1.17}$$

Audu and Singh [8] studied $\hat{\mu}_4, \hat{\mu}_5$ and $\hat{\mu}_6$ suggested by Kadilar and Cingi [9], and proposed the following generalized class of imputation schemes;

$$y_{.i} = \begin{cases} y_i & \text{i} \in \Phi \\ \dfrac{\hat{\mu}_0 + \hat{\beta}_{rg}\left(\bar{X} - \bar{x}_r\right)}{\pi_1 \bar{x}_r + \pi_2}\left(\pi_1 \bar{X} + \pi_2\right)\exp\left(\dfrac{\varpi_1\left(\bar{X} - \bar{x}_r\right)}{\varpi_1\left(\bar{X} + \bar{x}_r\right) + 2\varpi_2}\right) & \text{i} \in \Phi^{\text{c}} \end{cases} \tag{1.18}$$

where π_1 and π_2 are known functions of auxiliary variables like coefficients of skewness $\beta_1(x)$, kurtosis $\beta_2(x)$, variation C_X, standard deviation S_X etc, and $\pi_1 \neq \pi_2$ and $\pi_1 \neq 0$.

The point estimators of finite population mean under these methods of imputation are given by

$$\hat{\mu}_i^{(*)} = \frac{r}{n}\hat{\mu}_0 + \left(1 - \frac{r}{n}\right)\frac{\hat{\mu}_0 + \hat{\beta}_{rg}\left(\bar{X} - \bar{x}_r\right)}{\pi_1 \bar{x}_r + \pi_2}\left(\pi_1 \bar{X} + \pi_2\right)\exp\left(\frac{\varpi_1\left(\bar{X} - \bar{x}_r\right)}{\varpi_1\left(\bar{X} + \bar{x}_r\right) + 2\varpi_2}\right) \tag{1.19}$$

Their proposed class of imputation estimators is independent of unknown parameters; hence it is practically applicable.

$$MSE(\hat{\mu}_i^{(*)}) = \psi_{r,N}(S_Y^2 + \Upsilon^2 S_X^2 - 2\Upsilon S_{YX}) \tag{1.20}$$

where, $\eta_1 = \dfrac{\pi_1 \bar{X}}{\pi_1 \bar{X} + \pi_2}$, $\eta_2 = \dfrac{\varpi_1 \bar{X}}{2(\varpi_1 \bar{X} + \varpi_2)}$, and $\Upsilon = \left(1 - \dfrac{r}{n}\right)(R(\eta_1 + \eta_2) + \beta_{rg})$

2 Proposed Estimator under Imputation

Having studied the scheme and estimators of Audu and Singh [8], which utilized known functions of the auxiliary variable, that are sensitive to outliers, we proposed the following schemes to obtain estimators that are not sensitive to outliers by using nonconventional robust measures of the auxiliary variable defined as

$$y.i = \begin{cases} y_i & i \in \Phi \\ \dfrac{\bar{y}_r + \beta_{rg}(\bar{X} - \bar{x}_r)}{\phi_j \bar{x}_r + \phi_k}(\phi_j \bar{X} + \phi_k)\exp\left(\dfrac{(\phi_j \bar{X} + \phi_k) - (\phi_j \bar{x}_r + \phi_k)}{(\phi_j \bar{X} + \phi_k) + (\phi_j \bar{x}_r + \phi_k)}\right) & i \in \Phi^c \end{cases} \tag{2.1}$$

The point estimator of the scheme above is given by

$$\hat{t}_a = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right) \quad (2.2)$$

where, $\Gamma = \{(G\times n),(D\times n),(S_{pw}\times n)\}$, $\phi_j,\phi_k \in \Gamma$, $a = 1,2,3,4,5,6$ G, D and S_{pw} are the Gini's mean difference, Downtown's method and Probability weighted moments, which are the nonconventional robust measures free from the influence of outliers.

Remark 1: Note that $\phi_j \neq \phi_k$, six different estimators for scheme (2.2) were obtained in the Table 1.

Table 1. Some member of $\hat{t}_a$ for different values of ϕ_j and ϕ_k

a	Estimators	Values of Constants ϕ_j	ϕ_k
1	$\hat{t}_1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(G\times n)$	$(D\times n)$
2	$\hat{t}_2 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(G\times n)$	$(S_{pw}\times n)$
3	$\hat{t}_3 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(D\times n)$	$(G\times n)$
4	$\hat{t}_4 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(D\times n)$	$(S_{pw}\times n)$
5	$\hat{t}_5 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(S_{pw}\times n)$	$(G\times n)$
6	$\hat{t}_6 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k)\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\bar{x}_r+\phi_k)}\right)$	$(S_{pw}\times n)$	$(D\times n)$

3 Properties of the Suggested Estimators

In this section, to obtain the MSE of the estimators suggested, let us define:
Theorem 1:

$$MSE(\hat{t}_a) = \psi_{r,N}\left(S_Y^2 - 2M_a\rho S_Y S_X + M_a^2 S_X^2\right) \quad (3.1)$$

Proof: The MSE of the proposed estimators can be expressed as in (3.2)

$$MSE(\hat{t}_a) = \Delta_1 \Sigma \Delta_1^T \quad for\ a = 1,2,3,4,5,6 \tag{3.2}$$

where $\Delta_1 = \left(\frac{\partial \hat{t}_a}{\partial \bar{y}_r} \Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} \quad \frac{\partial \hat{t}_a}{\partial \bar{x}_r} \Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} \right)$

$\Sigma = \begin{pmatrix} \psi_{r,N} S_Y^2 & \psi_{r,N} \rho S_Y S_X \\ \psi_{r,N} \rho S_Y S_X & \psi_{r,N} S_X^2 \end{pmatrix}$, $\psi_{r,N} = r^{-1} - N^{-1}$ *and* Δ_1 is a row matrix of partial derivatives of the estimator $\hat{t}_a$ with respect to $\bar{y}_r$ *and* $\bar{x}_r$ and the transpose of Δ is Δ_1^T

On differentiating the estimator $\hat{t}_a$ with respect to $\bar{y}_r$ partially, we have,

$$\frac{\partial \hat{t}_a}{\partial \bar{y}_r} = \frac{r}{n} + \left(1 - \frac{r}{n}\right) \frac{(\phi_j \bar{X} + \phi_k)}{(\phi_j \bar{x}_r + \phi_k)} \exp\left(\frac{(\phi_j \bar{X} + \phi_k) - (\phi_j \bar{x}_r + \phi_k)}{(\phi_j \bar{X} + \phi_k) + (\phi_j \bar{x}_r + \phi_k)} \right) \tag{3.3}$$

On setting $\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}$, we have,

$$\frac{\partial \hat{t}_a}{\partial \bar{y}_r} \Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = 1 \tag{3.4}$$

On differentiating the estimator $\hat{t}_a$ with respect to $\bar{x}_r$ partially, considering product rule, we obtain,

$$\frac{\partial \hat{t}_a}{\partial \bar{x}_r} = \left(1 - \frac{r}{n}\right)(\phi_j \bar{X} + \phi_k) \left(U \frac{\partial V}{\partial \bar{x}_r} + V \frac{\partial U}{\partial \bar{x}_r} \right) \tag{3.5}$$

For $U = \frac{\left(\bar{y}_r + \beta_{rg}(\bar{X} - \bar{x}_r)\right)}{\left(\phi_j \bar{x}_r + \phi_k\right)} \left(\phi_j \bar{X} + \phi_k\right)$ and $V = \exp\left(\frac{\left(\phi_j \bar{X} + \phi_k\right) - \left(\phi_j \bar{x}_r + \phi_k\right)}{\left(\phi_j \bar{X} + \phi_k\right) + \left(\phi_j \bar{x}_r + \phi_k\right)} \right)$ (3.6)

$$\frac{\partial V}{\partial \bar{x}_r} = \frac{\left(-\phi_j \left[(\phi_j \bar{X} + \phi_j \bar{x}_r + 2\phi_k) + (\phi_j \bar{X} - \phi_j \bar{x}_r) \right]\right)}{\left[\left(\phi_j \bar{X} + \phi_k\right) + \left(\phi_j \bar{x}_r + \phi_k\right) \right]^2} V \tag{3.7}$$

$$\frac{\partial U}{\partial \bar{x}_r} = \frac{\left[\left(\phi_j \bar{X} + \phi_k\right) \hat{\beta}_{rg} + \left(\bar{y}_r + \hat{\beta}_{rg}(\bar{X} - \bar{x}_r)\right) \phi_j \right]}{\left(\phi_j \bar{X} + \phi_k\right)^2} \tag{3.8}$$

On setting $\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}$, in equations (3.6), (3.7), (3.8), we obtain,

$$U \Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = \frac{\bar{Y}}{(\phi_j \bar{X} + \phi_k)} \tag{3.9}$$

$$V\Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = 1 \tag{3.10}$$

$$\frac{\partial V}{\partial \bar{x}_r}\Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -\frac{\phi_j}{2(\phi_j \bar{X} + \phi_k)} \tag{3.11}$$

$$\frac{\partial U}{\partial \bar{x}_r}\Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -\left[\frac{\beta_{rg}}{(\phi_j \bar{X} + \phi_k)} + \frac{\bar{Y}\phi_j}{(\phi_j \bar{X} + \phi_k)^2}\right] \tag{3.12}$$

Substituting equations (3.9), (3.10), (3.11), (3.12) into equation (3.5), we obtain,

$$\frac{\partial \hat{t}_i^2}{\partial \bar{x}_r}\Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -\left(1 - \frac{r}{n}\right)\left(\beta_{rg} + \frac{3\bar{Y}\phi_j}{2(\phi_j \bar{X} + \phi_k)}\right) \tag{3.13}$$

$$\frac{\partial \hat{t}_i^2}{\partial \bar{x}_r}\Big|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -M_a \tag{3.14}$$

where, $M_a = \left(1 - \frac{r}{n}\right)\left(\beta_{rg} + \frac{3\phi_j \bar{Y}}{2(\phi_j \bar{X} + \phi_k)}\right)$

Therefore,

$\Delta_1 = \begin{pmatrix} 1 & -M_a \end{pmatrix}$ and its transpose $\Delta_1^T = \begin{pmatrix} 1 & -M_a \end{pmatrix}^T$ (3.15)

Substitute (3.15) in (3.2), we obtained the mean square error of the estimator as

$$MSE(\hat{t}_a) = \begin{pmatrix} 1 & -M_a \end{pmatrix} \begin{pmatrix} \psi_{r,N} S_Y^2 & \psi_{r,N} \rho S_Y S_X \\ \psi_{r,N} \rho S_Y S_X & \psi_{r,N} S_X^2 \end{pmatrix} \begin{pmatrix} 1 \\ -M_a \end{pmatrix} \tag{3.16}$$

$$MSE(\hat{t}_i^2) = \psi_{r,N}\left(S_Y^2 - 2M_a \rho S_Y S_X + M_a^2 S_X^2\right) \tag{3.17}$$

Theorem 2: the estimators $\hat{t}_a$ $(a = 1,2,3,4,5,6)$ are consistent.

Proof: Let f(x) and g(x) be continuous function, then

$$\lim_{x \to p}\left(f(x) \pm g(x)\right) = \lim_{x \to p} f(x) \pm \lim_{x \to p} g(x), \quad p \neq \infty \tag{3.18}$$

$$\lim_{x \to p}\left(f(x) \times g(x)\right) = \lim_{x \to p} f(x) \times \lim_{x \to p} g(x), \quad p \neq \infty \tag{3.19}$$

$$\lim_{x\to p}\frac{f(x)}{g(x)}=\frac{\lim_{x\to p}f(x)}{\lim_{x\to p}g(x)},\quad p\neq\infty,\ \lim_{x\to p}g(x)\neq 0 \tag{3.20}$$

As $r\to N,\ \ n=N$. Using the results of (3.18), (3.19) and (3.20), we have

$$\lim_{r\to N}\left(\hat{t}_a\right)=\lim_{r\to N}\frac{r}{n}\lim_{r\to N}\left(\bar{y}_r\right)+\left(1-\lim_{r\to N}\frac{r}{n}\right)\frac{(\lim_{r\to N}\bar{y}_r+\lim_{r\to N}\hat{\beta}_{rg}(\bar{X}-\lim_{r\to N}\bar{x}_r))}{\phi_j\lim_{r\to N}\bar{x}_r+\phi_k}(\phi_j\bar{X}+\phi_k)$$
$$\exp\left(\frac{(\phi_j\bar{X}+\phi_k)-(\phi_j\lim_{r\to N}\bar{x}_r+\phi_k)}{(\phi_j\bar{X}+\phi_k)+(\phi_j\lim_{r\to N}\bar{x}_r+\phi_k)}\right) \tag{3.21}$$

Since $n=N$ if $r\to N$, then $\lim_{r\to N}\bar{y}_r=\bar{Y},\ \lim_{r\to N}\bar{x}_r=\bar{X},\ \lim_{r\to N}\frac{r}{n}=1$ and $\lim_{r\to N}\hat{\beta}_{rg}=\beta_{rg}$. Therefore,

$$\lim_{r\to N}\left(\hat{t}_a\right)=\bar{Y},\quad \forall\quad a=1,2,3,4,5,6 \tag{3.22}$$

4 Efficiency Comparisons

In this section, conditions for the efficiency of the new estimators over some existing related estimators were established.

Theorem 3: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_0$ if (4.1) is satisfied.

$$M_a<2\beta_{rg} \tag{4.1}$$

Proof: Minus (3.1) from (1.3), theorem 3 is proved.
Theorem 4: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_1$ if (4.2) is satisfied.

$$\psi_{r,N}\left(M_a^2-2\beta_{rg}M_a\right)-\psi_{r,n}\left(R^2-2\beta_{rg}R\right)<0 \tag{4.2}$$

Proof: Subtract (3.11) from (1.6), theorem 4 is proved.
Theorem 5: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_2$ and $\hat{\mu}_3$ if (4.3) is satisfied.

$$\psi_{r,n}\beta_{rg}^2+\psi_{r,N}\left(M_a^2-2\beta_{rg}\Upsilon\right)<0 \tag{4.3}$$

Proof: Subtract (3.1) from each of (1.9) and (1.11), theorem 5 is proved.
Theorem 6: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_4$ if (4.4) is satisfied.

$$\left(\beta_{rg}-M_a\right)^2-R^2<0 \tag{4.4}$$

Proof: Subtract (3.1) from (1.15), theorem 6 is proved.
Theorem 47: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_5$ if (4.5) is satisfied.

$$\psi_{r,N}\left(M_a^2-2\beta_{rg}M_a\right)-\psi_{n,N}\left(R^2-\beta_{reg}^2\right)<0 \tag{4.5}$$

Proof: Subtract (3.1) from (1.16), theorem 7 is proved.

Theorem 8: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_6$ if (4.6) is satisfied.

$$\psi_{r,N}\left(M_a^2-2\beta_{rg}M_a\right)-\psi_{r,n}\left(R^2-\beta_{reg}^2\right)<0 \tag{4.6}$$

Proof: Subtract (3.1) from (1.17), theorem 8 is proved.

Theorem 9: Estimator $\hat{t}_a$ is more efficient than $\hat{\mu}_i^{(*)}$ if (4.7) is satisfied.

$$\left(\Upsilon+M_a\right)\beta_{rg}+\left(M_a^2-\Upsilon^2\right)<0 \tag{4.7}$$

Proof: Subtract (3.1) from (1.20), theorem 9 is proved.

5 Numerical Examples

In this section, simulation studies were conducted to assess the performance of the estimators of the proposed scheme with respect to Audu and Singh [8] estimators. Data of size 1000 units were generated for study populations using function defined in Table 1. Samples of size 100 units from which 60 units were selected as respondents were randomly chosen 10,000 times by method of simple random sampling without replacement (SRSWOR). The Biases, MSEs and PREs of the considered estimators were computed using (4.39), (4.40), (4.41).

$$Bias\left(\hat{\theta}_d\right)=\frac{1}{10000}\sum_{d=1}^{10000}\left(\hat{\theta}_d-\bar{Y}\right),\hat{\theta}_d=\hat{\mu}_0,\hat{\mu}_i^{(*)},i=1,2,...,17,\hat{t}_a\left(\ i=1,2,3,4,5,6\right) \tag{4.39}$$

$$MSE\left(\hat{\theta}_d\right)=\frac{1}{10000}\sum_{d=1}^{10000}\left(\hat{\theta}_d-\bar{Y}\right)^2,\hat{\theta}_d=\hat{\mu}_0,\mu_i^{(*)},i=1,2,...,17,\hat{t}_a\left(\ i=1,2,3,4,5,6\right) \tag{4.40}$$

$$PRE\left(\hat{\theta}_d\right)=\left(\frac{MSE\left(\hat{\theta}_d\right)}{Var\left(\hat{\mu}_0\right)}\right)\times 100,\hat{\theta}_d=\hat{\mu}_0,\hat{\mu}_i^{(*)},i=1,2,...,17,\hat{t}_a\left(\ i=1,2,3,4,5,6\right) \tag{4.41}$$

Table 2. Populations used for Simulation Study

Populations	Auxiliary variable (x)	Study variable (y)
I	$X\sim beta(1.1,2.0)$	$Y=50+10X+20X^2+e,$
II	$X\sim gamma(10,25)$	$where,\ e\sim(0,4)$
III	$X\sim pois(0.5)$	
IV	$X\sim unif\ (0,0.4)$	

Tables 3, 4, 5 and 6 show the results of the biases, MSEs and PREs of the sample mean, Audu and Singh [8]estimators and estimators of the proposed scheme using the simulated data for populations I, II, III and IV defined in Table 2 respectively. The results revealed that the estimators $\hat{t}_a$ of the proposed scheme, have minimum biases, MSEs and higher PREs than the Sample mean and Audu and Singh [8] estimators with the exception of few cases where few members of Audu and Singh [8] estimators outperformed some members of the proposed estimators.

Table 3. Biases, MSEs and PREs of Proposed and Some Estimators using Pop. I

Estimators	Biases	MSEs	PREs
Sample mean $\hat{\mu}_0$	**0.005212425**	**0.08212077**	**100**
Audu and Singh [8]			
$\hat{\mu}_1^{(*)}$	0.01881164	0.5988081	13.71404
$\hat{\mu}_2^{(*)}$	0.006281752	0.1573342	52.1951
$\hat{\mu}_3^{(*)}$	0.00771724	0.2157503	38.06288
$\hat{\mu}_4^{(*)}$	0.01400318	0.120201	68.31951
$\hat{\mu}_5^{(*)}$	0.009613986	0.2862123	28.69225
$\hat{\mu}_6^{(*)}$	0.006427576	0.1636258	50.18815
$\hat{\mu}_7^{(*)}$	0.007157499	0.03267426	251.3317
$\hat{\mu}_8^{(*)}$	0.007912665	0.2232647	36.78179
$\hat{\mu}_9^{(*)}$	0.004970697	0.09101034	90.23235
$\hat{\mu}_{10}^{(*)}$	0.005403985	0.02719941	301.9212
$\hat{\mu}_{11}^{(*)}$	0.006415286	0.1630998	50.35002
$\hat{\mu}_{12}^{(*)}$	0.01825179	0.1904569	43.11778
$\hat{\mu}_{13}^{(*)}$	0.3306025	21.33659	0.3848824
$\hat{\mu}_{14}^{(*)}$	0.8956953	1218.282	0.006740701
$\hat{\mu}_{15}^{(*)}$	0.004744487	0.07359994	111.5772
$\hat{\mu}_{16}^{(*)}$	0.00497412	0.09123646	90.00871
$\hat{\mu}_{17}^{(*)}$	0.004917712	0.03289367	249.6552
Estimators of Proposed Scheme			
$\hat{t}_1$	0.04056205	0.05404431	151.9508
$\hat{t}_2$	0.04089361	0.05434909	151.0987
$\hat{t}_3$	-0.03405898	0.04264035	192.5893
$\hat{t}_4$	0.00220966	0.03396731	241.7641
$\hat{t}_5$	-0.03434565	0.04281589	191.7997
$\hat{t}_6$	0.001591584	0.03389004	242.3153

Table 4. Biases, MSEs and PREs of Proposed and Some Estimators using Pop. II

Estimators	Biases	MSEs	PREs
Sample mean $\hat{\mu}_0$	**-0.00238708**	**0.02055153**	**100**
Audu and Singh [8]			
$\hat{\mu}_1^{(*)}$	0.01224332	0.1290456	15.92579
$\hat{\mu}_2^{(*)}$	0.006784724	0.05279314	38.92841
$\hat{\mu}_3^{(*)}$	0.005456455	0.03706221	55.45144
$\hat{\mu}_4^{(*)}$	0.008066989	0.06924065	29.6813
$\hat{\mu}_5^{(*)}$	0.00908595	0.08303842	24.74942
$\hat{\mu}_6^{(*)}$	0.003340483	0.01579654	130.1015
$\hat{\mu}_7^{(*)}$	0.005124688	0.03338117	61.56624
$\hat{\mu}_8^{(*)}$	0.006160419	0.04521465	45.45325
$\hat{\mu}_9^{(*)}$	0.005198207	0.03418736	60.11442
$\hat{\mu}_{10}^{(*)}$	0.006411459	0.04822499	42.61593
$\hat{\mu}_{11}^{(*)}$	0.007531888	0.06224325	33.01808
$\hat{\mu}_{12}^{(*)}$	0.003381086	0.01614697	127.2779
$\hat{\mu}_{13}^{(*)}$	0.002833185	0.01166292	176.2125
$\hat{\mu}_{14}^{(*)}$	0.004979215	0.03180247	64.62243
$\hat{\mu}_{15}^{(*)}$	0.002918198	0.01232231	166.7831
$\hat{\mu}_{16}^{(*)}$	0.002534123	0.009460828	217.2276
$\hat{\mu}_{17}^{(*)}$	0.003465998	0.01688858	121.6889
Estimators of Proposed Scheme			
$\hat{t}_1$	0.04219346	0.02208199	93.06919
$\hat{t}_2$	0.04254799	0.0223906	91.78643
$\hat{t}_3$	-0.03787117	0.01712049	120.0405
$\hat{t}_4$	0.001111189	0.003437694	597.8289
$\hat{t}_5$	-0.03817979	0.01735068	118.4479
$\hat{t}_6$	0.000447942	0.003415049	601.7931

Table 5. Biases, MSEs and PREs of Proposed and Some Estimators using Pop. III

Estimators	Biases	MSEs	PREs
Sample mean $\hat{\mu}_0$	**-0.009713722**	**2.390206**	**100**
Audu and Singh [8]			
$\hat{\mu}_1^{(*)}$	0.1744741	4.818217	49.60768
$\hat{\mu}_2^{(*)}$	0.03957149	0.5903091	404.9075
$\hat{\mu}_3^{(*)}$	0.03908925	0.5832271	409.8242
$\hat{\mu}_4^{(*)}$	0.03088749	0.4790073	498.9915
$\hat{\mu}_5^{(*)}$	0.05693135	0.898523	266.015
$\hat{\mu}_6^{(*)}$	0.04675464	0.7061113	338.5027
$\hat{\mu}_7^{(*)}$	0.03580307	0.5375918	444.6135
$\hat{\mu}_8^{(*)}$	0.06873159	1.156279	206.7154
$\hat{\mu}_9^{(*)}$	0.04801837	0.7283173	328.1819
$\hat{\mu}_{10}^{(*)}$	0.0362162	0.5430662	440.1316
$\hat{\mu}_{11}^{(*)}$	0.06964129	1.177546	202.982
$\hat{\mu}_{12}^{(*)}$	0.06419486	1.053127	226.9626
$\hat{\mu}_{13}^{(*)}$	0.06333724	1.034183	231.1201
$\hat{\mu}_{14}^{(*)}$	0.09051003	1.715716	139.3124
$\hat{\mu}_{15}^{(*)}$	0.0338194	0.5124397	466.4365
$\hat{\mu}_{16}^{(*)}$	0.03345816	0.50807	470.4481
$\hat{\mu}_{17}^{(*)}$	0.02742584	0.4462391	535.6334
Estimators of Proposed Scheme			
$\hat{t}_1$	0.06620763	0.4403929	542.7439
$\hat{t}_2$	0.06673713	0.4411481	541.8148
$\hat{t}_3$	-0.05414129	0.4264173	560.532
$\hat{t}_4$	0.004647217	0.3941274	606.4551
$\hat{t}_5$	-0.05460819	0.4269713	559.8048
$\hat{t}_6$	0.003650053	0.3940539	606.5682

Table 6. Biases, MSEs and PREs of Proposed and Some Estimators using Pop. IV

Estimators	Biases	MSEs	PREs
Sample mean $\hat{\mu}_0$	**0.0002986744**	**0.03345174**	**100**
Audu and Singh [8]			
$\hat{t}_1^{(*)}$	0.03263777	0.5715367	5.852947
$\hat{t}_2^{(*)}$	0.01207208	0.1363139	24.54023
$\hat{t}_3^{(*)}$	0.03494594	0.6251446	5.35104
$\hat{t}_4^{(*)}$	0.006156071	0.03519193	95.05513
$\hat{t}_5^{(*)}$	0.02070915	0.3089493	10.82758
$\hat{t}_6^{(*)}$	0.03691146	0.6714911	4.98171
$\hat{t}_7^{(*)}$	0.006960825	0.0467871	71.49777
$\hat{t}_8^{(*)}$	0.01720137	0.2367268	14.13095
$\hat{t}_9^{(*)}$	0.008072876	0.06462869	51.75989
$\hat{t}_{10}^{(*)}$	0.008144402	0.06582323	50.82056
$\hat{t}_{11}^{(*)}$	0.007881181	0.06145104	54.4364
$\hat{t}_{12}^{(*)}$	0.005492981	0.05355526	62.46209
$\hat{t}_{13}^{(*)}$	0.03097614	0.5334939	6.270313
$\hat{t}_{14}^{(*)}$	0.08076056	1.88362	1.775928
$\hat{t}_{15}^{(*)}$	0.008613089	0.07375923	45.35261
$\hat{t}_{16}^{(*)}$	0.08302571	1.956675	1.709621
$\hat{t}_{17}^{(*)}$	0.007885434	0.06152115	54.37437
Estimators of Proposed Scheme			
$\hat{t}_1^2$	0.02550096	0.03836241	87.19925
$\hat{t}_2^2$	0.02569806	0.0384789	86.93527
$\hat{t}_3^2$	-0.01804812	0.03207996	104.2761
$\hat{t}_4^2$	0.002918864	0.03027092	110.5078
$\hat{t}_5^2$	-0.01821237	0.03212954	104.1152
$\hat{t}_6^2$	0.002558428	0.03022621	110.6713

6 Conclusions

From the results of the empirical study, it was obtained that some members of the proposed class of estimators especially $\hat{t}_4$ and $\hat{t}_6$ are more efficient than Audu and Singh [8] estimators and, therefore, they are recommended to estimate the population average when certain values of the variables of the study are missing in the study. In conclusion, the proposed class of imputation schemes is recommended for use when the characteristics of the population under study are characterized by outliers or extreme values.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Hansen MN, Hurwitz WN. The problem of non-response in sample surveys. Journal of the American Statistical Association. 1946; 41:517–529.

[2] Singh S, Deo B. Imputation by power transformation. Statistical Papers. 2008; 44:555-579.

[3] Toutenburg H, Srivastava VK, Shalabh A. Imputation versus imputation of missing values through ratio method in sample surveys. Statistical Papers. 2008; 49:237-247.

[4] Kadilar, C., Cingi, H. Ratio estimators in simple random sampling. Applied Mathematics and Computation. 2004, 151:893–902. doi:10.1016/S0096-3003(03)00803-8

[5] Singh S. A new method of imputation in survey sampling. Statistics. 2009; 43:499-511.

[6] Singh S, Horn S. Compromised imputation in survey sampling. Metrika. 2000; 51:267-276.

[7] Gira AA. Estimation of population mean with a New Imputation Methods. Applied Mathematical Sciences. 2015;9(34):1663-1672.

[8] Audu A, Singh RVK. Exponential-type regression compromised imputation class of estimators, Journal of Statistics and Management System. 2021;24(6):1253-1266.
DOI:10.1080/09720510.2020.1814501

[9] Kadilar C, Cingi H. Estimators for population mean in the case of missing data. Communication in Statistics-Theory and Methods. 2008; 37(14) :2226-2236. doi: 10.1080/03610920701855020.

[10] Bhushan S, Pandey AP. Optimal imputation of missing data for estimation of population mean. Journal of Statistics and Management Systems. 2016;19(6):755-769.

[11] Singh GN, Maurya S, Khetan M, Kadilar C. Some imputation methods for missing data in sample surveys. Hacettepe Journal of Mathematics and Statistics. 2016;45 (6):1865-1880.

[12] Diana G, Perri PF. Improved estimators of the population mean for missing data, Communications in Statistics- Theory and Methods. 2010; 39:3245-3251.

[13] Al-Omari AI, Bouza CN, Herrera C. Imputation methods of missing data for estimating the population mean using simple random sampling with known correlation coefficient. Quality and Quantity. 2013;47: 353-365.

[14] Audu A, Ishaq OO, Isah U, Muhammad S, Akintola K, A, Rashida A, Abubakar, A. On the Class of Exponential-Type Imputation Estimators of Population Mean with Known Population Mean of Auxiliary Variabl, NIPES Journal of Science and Technology. 2020a;2(4):1-11.
Available: https://doi.org/10.37933/nipes/2.4.2021.1

[15] Audu A, Ishaq OO, Zakari Y, Wisdom DD, Muili J, Ndatsu AM. Regression-cum-exponential ratio imputation class of estimators of population mean in the presence of non-response. Science forum (Journal of Pure and Applied Sciences). 2020b;58-63.
Available: http://dx.doi.org/10.5455/st.71109

[16] Audu A, Singh R, Khare S. New regression-type compromised imputation class of estimators with known parameters of auxiliary variable. Communications in statistics – Simulation and Computation. 2021a;1-13.
DOI: 10.1080/03610918.2021.1970182

[17] Audu A, Ishaq OO, Abubakar A, Akintola KA, Isah U, Rashida A, Muhammad S. Regression-type Imputation Class of Estimators using Auxiliary Attribute. Asian Research Journal of Mathematics. 2021b;17(5):1-13.
DOI: 10.9734/ARJOM/2021/v17i530296

[18] Audu A, Ishaq OO, Muili J, Zakari Y, Ndatsu AM, Muhammed S. On the Efficiency of Imputation Estimators using Auxiliary Attribute. Continental J. Applied Sciences. 2020c;15(1):1-13.
DOI: 10.5281/zenodo.3721046

[19] Singh AK, Singh P, Singh VK. Exponential-Type Compromised Imputation in Survey Sampling, J. Stat. Appl. 2014;3(2):211-217.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle4.com/review-history/75530