Volume 66, Issue 1, 2022

Journal of Scientific Research

of

The Banaras Hindu University

Modified Compromized Type Method of Imputation for Estimating Population Mean

Yusuf H Aliyu[1], Amos A Adewara[2], Ahmed Audu[3], Omotayo A Abidoye[4], Issa Sulaiman[5], Muhammed B Aliyu[6]

[1]National Bureau of Statistics, Kwara State, Nigeria
[2,4,6]Department of Statistics, University of Ilorin, Kwara State, Nigeria
[3]Department of Mathematics, UsmanuDanfodiyo University, Sokoto, Nigeria,
[5]Department of Mathematical Sciences University of Maiduguri, Borno State, Nigeria
yusufaliko@gmail.com, aadewara@unilorin.edu.ng, ahmed.audu@udusok.edu.ng, abidoye@unilorin.edu.ng, isstatisticsman01@gmail.com, mbaliyu9@gmail.com

Abstract: **The study proposed an alternative imputation scheme for the schemes which converged to sample mean as the values of unknown parameters in their estimators converged to zero. The estimator of the population means for the proposed scheme as well as the bias and MSE were derived. The efficiency condition under which the modified estimator is more efficient than existing ones were also presented. Empirical study using four sets of populations was conducted and the results revealed that the proposed estimator was more efficient.**

Index Terms: **Estimator; Population mean; Imputation scheme; Bias; Mean Squared Error (MSE)**

I. Introduction

Data obtained from sampling surveys often face the problem of non-response or missing values. These missing values create difficulty in analysis, processing and handling of data. The problem of non-response has been considered by many authors including Singh and Horn (2000), Singh and Deo (2003), Wang and Wang (2006), Kadilar and Cingi (2008), Toutenburg et al. (2008), Singh (2009), Diana and Perri (2010), Al-Omari et al. (2013), Singh et al. (2014), Gira (2015), Singh et al. (2016), Singh, et al. (2010), Bhushan and Pandey (2016) and Prasad (2017), Audu et al. (2020a, b, c), Audu et al. (2021). Singh and Deo (2003) and Prasad (2017) estimators converged to sample mean as the values of unknown parameters in their estimators converged to zero while Singh and Horn (2000), Singh et al. (2014) estimators converged to sample mean as the values of unknown parameters converged to one. These converges lead to lose of information on the auxiliary variables which in turn reduces their efficiencies. These limitations identified above prompt the present study, some existing literature on imputation schemes and estimators are reviewed and their properties were presented.

II. Methodology

A. Some Existing Related Imputation Schemes

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

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,

$$y_{.i} = \begin{cases} y_i & \text{if i} \in R \\ \bar{y}_r & \text{if i} \in R^c \end{cases} \tag{1}$$

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

$$\bar{y}_r = \frac{1}{r}\sum_{i \in R} y_i \tag{2}$$

The bias and MSE of $\hat{\bar{\theta}}_{mean}$ are given by (3) and (4) respectively.

$$Bias(\hat{\theta}_{mean}) = 0 \quad (3)$$

$$MSE(\hat{\theta}_{mean}) = \varphi_{r,N}\bar{Y}^2 C_Y^2 \quad (4)$$

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 is 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

$$y_{.i} = \begin{cases} y_i & \mathrm{i} \in R \\ \hat{\beta}\mathrm{x_i} & \mathrm{i} \in R^c \end{cases} \quad (5)$$

The estimator of population mean denoted by $\hat{\bar{\theta}}_{ratio}$, its bias and MSE under the method of ratio imputation are given as

$$\hat{\theta}_{ratio} = \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r} \quad (6)$$

where

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

S

$$Bias(\hat{\theta}_{ratio}) = \varphi_{r,n}\bar{Y}\left(C_X^2 - C_{YX}\right) \quad (7)$$

$$MSE(\hat{\theta}_{ratio}) = \varphi_{r,N}\bar{Y}^2 C_Y^2 + \varphi_{r,n}\bar{Y}^2\left(C_X^2 - 2C_{YX}\right) \quad (8)$$

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 (9).

$$y_{.i} = \begin{cases} \lambda \frac{n}{r} y_i + (1-\lambda)\hat{\beta}x_i & \mathrm{i} \in R \\ (1-\lambda)\hat{\beta}x_i & \mathrm{i} \in R^c \end{cases} \quad (9)$$

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

$$\hat{\theta}_{cmp} = \bar{y}_r\left(\lambda + (1-\lambda)\frac{\bar{x}_n}{\bar{x}_r}\right) \quad (10)$$

The bias and MSE of $\hat{\bar{\theta}}_{cmp}$ up to first order approximation are given by (11) and (12) respectively as;

$$Bias\left(\hat{\theta}_{cmp}\right) = (1-\lambda)\varphi_{r,n}\bar{Y}\left(C_X^2 - C_{YX}\right) \quad (11)$$

$$MSE\left(\hat{\theta}_{cmp}\right) = \varphi_{r,N}\bar{Y}^2 C_Y^2 + \varphi_{r,n}\bar{Y}^2\left((1-\lambda)^2 C_X^2 - 2(1-\lambda)C_{YX}\right) \quad (12)$$

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

$$MSE\left(\hat{\theta}_{cmp}\right)_{\min} = \bar{Y}^2 C_Y^2\left(\varphi_{r,N} - \varphi_{r,n}\rho_{YX}^2\right) \quad (13)$$

Singh and Deo (2003) incorporated power transformation parameter to $\hat{\theta}_{ratio}$ and obtain $\hat{\theta}_{SH}$ as

$$\hat{\theta}_{SD} = \bar{y}_r\left(\frac{\bar{x}_n}{\bar{x}_r}\right)^{\alpha} \quad (14)$$

$\hat{\theta}_{SD}$ attained optimum when $\alpha = RS_{YX} / S_X^2$ and the $MSE\left(\hat{\theta}_{SH}\right)_{\min}$ is given by

$$MSE(\hat{\theta}_{SH}) = MSE\left(\hat{\theta}_{ratio}\right) - \psi_{r,n} S_X^2\left(\beta_{reg} - R\right)^2 \quad (15)$$

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

Kadilar and Cingi (2008) modified the work of Kadilar and Cingi (2004) in the case of missing observations and suggested the following estimators of population mean

$$\hat{\theta}_{KC1} = \left(\bar{y}_r + \hat{\beta}_{reg}\left(\bar{X} - \bar{x}_r\right)\right)\frac{\bar{X}}{\bar{x}_r} \quad (16)$$

$$\hat{\vartheta}_{KC2} = (\bar{y}_r + \hat{\beta}_{reg}(\bar{X} - \bar{x}_n))\frac{\bar{X}}{\bar{x}_n} \quad (17)$$

$$\hat{\vartheta}_{KC3} = (\bar{y}_r + \hat{\beta}_{reg}(\bar{X} - \bar{x}_n))\frac{\bar{x}_n}{\bar{x}_r} \quad (18)$$

The MSEs of $\hat{\theta}_{KC1}, \hat{\theta}_{KC2}$ and $\hat{\theta}_{KC3}$ are given in (19), (20) and (21) respectively

$$MSE(\hat{\theta}_{KC1}) = MSE\left(\hat{\theta}_{mean}\right) + \psi_{r,N} S_X^2 \left(R^2 - \beta_{reg}^2\right) \quad (19)$$

$$MSE(\hat{\theta}_{KC2}) = MSE\left(\hat{\theta}_{mean}\right) + \psi_{n,N} S_X^2 \left(R^2 - \beta_{reg}^2\right) \quad (20)$$

$$MSE(\hat{\theta}_{KC3}) = MSE\left(\hat{\theta}_{mean}\right) + \psi_{r,n}\left(S_X^2 \left(R + \beta_{reg}\right)^2 - 2\left(R + \beta_{reg}\right) S_{YX}\right) \quad (21)$$

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

$$y_{.i} = \begin{cases} \nu \dfrac{n}{r} y_i + (1-\nu)\bar{y}_r \exp\left(\dfrac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) & \text{if } i \in R \\ (1-\nu)\bar{y}_r \exp\left(\dfrac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) & \text{if } i \in R^c \end{cases} \quad (22)$$

The point estimator $\hat{\hat{\theta}}_{ExpCmp}$, bias and MSE are given as:

$$\hat{\theta}_{ExpCmp} = \nu \bar{y}_r + (1-\nu)\bar{y}_r \exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) \quad (24)$$

$$Bias\left(\hat{\theta}_{ExpCmp}\right) = (1-\nu)\varphi_{r,N} \bar{Y}\left(\frac{3}{8} C_X^2 - \frac{1}{2} C_{YX}\right) \quad (25)$$

$$MSE\left(\hat{\theta}_{ExpCmp}\right) = \varphi_{r,N} \bar{Y}^2 \left(C_Y^2 + \frac{(1-\nu)^2}{4} C_X^2 - (1-\nu) C_{YX}\right) \quad (26)$$

$$MSE\left(\hat{\theta}_{ExpCmp}\right)_{\min} = \varphi_{r,N} \bar{Y}^2 C_Y^2 \left(1 - \rho_{YX}^2\right) \quad (27)$$

Prasad (2017) proposed ratio exponential estimator given as

$$\hat{\theta}_{Prasad} = \eta \hat{t}_0 \exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r + 2\rho_{YX} / \beta_2(x)}\right) \quad (28)$$

where

$$\eta = \bar{Y}^2 / \left(\bar{Y}^2 + \psi_{r,N}\left(S_y^2 + 0.25\mu^2 R^2 S_X^2 - \mu R S_{YX}\right)\right),$$

$$MSE\left(\hat{\theta}_{Prasad}\right)_{\min} = \frac{\bar{Y}^2 \left(\psi_{r,N}\left(C_Y^2 + 0.25\mu^2 C_X^2 - \mu C_{YX}\right)\right)}{\left(1 + \psi_{r,N} \bar{Y}^2 \left(C_Y^2 + 0.25\mu^2 C_X^2 - \mu C_{YX}\right)\right)} \quad (29)$$

III. Proposed Modified Scheme

Motivated by the schemes proposed by Singh and Horn (2000), Singh and Deo (2003), Singh et al. (2014) and Prasad (2017), the following imputation scheme for estimating finite population mean is suggested;

$$y_{.i} = \begin{cases} y_i & \text{if } i \in R \\ \dfrac{1}{n-r}\left(n\left(\theta_2 \bar{y}_r + (1-\theta_2)\bar{y}_r \dfrac{\bar{x}_r}{\bar{X}}\right)\exp\left(\dfrac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) - r\bar{y}_r\right) & \text{if } i \in R^c \end{cases} \quad (30)$$

where θ_2 is unknown function of study and auxiliary variables which minimized the mean squared error of the proposed scheme.

A. *Equation of Proposed Estimator of the Modified Scheme*

The point estimator for the proposed scheme is obtained as

$$t^{new} = \frac{1}{n}\left(\sum_{i \in \Phi} y_i + \sum_{i \in \Phi^c} \frac{1}{n-r}\left(n\left(\theta_2 \bar{y}_r + (1-\theta_2)\bar{y}_r \frac{\bar{x}_r}{\bar{X}}\right)\exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) - r\bar{y}_r\right)\right) \quad (31)$$

Simplify (31) to obtain the estimator of the modified scheme

$$t^{new} = \left(\theta_2 \bar{y}_r + (1-\theta_2)\bar{y}_r \frac{\bar{x}_r}{\bar{X}}\right)\exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) \quad (32)$$

B. *Properties (Bias and MSE) for Estimator of the Modified Scheme*

In this subsection, the bias and MSE for estimator of the modified scheme are derived and discussed.

To derive the properties of the estimator obtained from the modified scheme, the following error terms are defined:

$e_0 = \dfrac{\bar{y}_r - \bar{Y}}{\bar{Y}}$, $e_1 = \dfrac{\bar{x}_r - \bar{X}}{\bar{X}}$ such that $\lim_{r \to n} |e_i| \approx 0, i = 0,1$.

This implies that $\bar{y}_r = \bar{Y}(1+e_0)$, $\bar{x}_r = \bar{X}(1+e_1)$.

The expectation in terms of e_0 and e_1 are given in (33);

$$\left.\begin{aligned} &E(e_0) = E(e_1) = 0,\ E\left(e_0^2\right) = \left(\frac{1}{r} - \frac{1}{N}\right) C_Y^2 \\ &E\left(e_1^2\right) = \left(\frac{1}{r} - \frac{1}{N}\right) C_X^2,\ E(e_0 e_1) = \left(\frac{1}{r} - \frac{1}{N}\right)\rho_{YX} C_Y C_X \end{aligned}\right\} \quad (33)$$

Similarly, express (32) in terms of error terms e_0 and e_1, t_2^{new} is of the form given in (34)

$$t^{new} = \left(\theta_2 \bar{Y}(1+e_0) + (1-\theta_2)\bar{Y}(1+e_0)\frac{\bar{X}(1+e_1)}{\bar{X}}\right)\exp\left(\frac{\bar{X} - \bar{X}(1+e_1)}{\bar{X} + \bar{X}(1+e_1)}\right) \quad (34)$$

Simplify (34) up to first order approximation,

$$t^{new} = \bar{Y}\left(1 + e_0 + \left(\frac{1}{2} - \theta_2\right)e_1 - \left(\frac{1}{8} - \theta_2\right)e_1^2 - \left(\frac{1}{2} + \theta_2\right)e_0 e_1\right) \quad (35)$$

Subtract $\bar{Y}$ from both sides of (35)

$$t^{new} - \bar{Y} = \bar{Y}\left(e_0 + \left(\frac{1}{2} - \theta_2\right)e_1 - \left(\frac{1}{8} - \theta_2\right)e_1^2 - \left(\frac{1}{2} + \theta_2\right)e_0 e_1\right) \quad (36)$$

Take expectation of (36) and apply the results of (33), the bias of t_2^{new} is obtained as:

$$Bias\left(t^{new}\right) = \bar{Y}\left(\frac{1}{r} - \frac{1}{N}\right)\left(-\left(\frac{1}{8} - \theta_2\right)C_X^2 - \left(\frac{1}{2} + \theta_2\right)\rho_{YX} C_Y C_X\right) \quad (37)$$

Similarly, square both sides of (36) up to second degree approximation, take expectation and apply the results of (33), the MSE of t^{new} is obtained as:

$$MSE\left(t^{new}\right) = \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) \quad (38)$$

C. Minimum MSEs of t^{new}

To obtain the minimum MSE of t^{new},(38) is partially differentiated with respect to θ_2 respectively and equate to zero as:

$$\frac{\partial MSE\left(t^{new}\right)}{\partial \theta_2} = \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(-2\left(\frac{1}{2} - \theta_2\right)C_X^2 - 2\rho_{YX} C_Y C_X\right) = 0 \quad (39)$$

Solve for θ_2 in (39), we obtained

$$\theta_2 = \frac{1}{2} + \rho_{YX}\frac{C_Y}{C_X} \quad (40)$$

Substitute (40) in (38) respectively, the minimum MSE of t^{new} is obtained as

$$MSE\left(t^{new}\right) = \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)C_Y^2\left(1 - \rho_{YX}^2\right) \quad (41)$$

D. Efficiency Comparison of Proposed Estimator t^{new} with some Related Estimators

i. $MSE\left(\hat{\theta}_{mean}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2 C_Y^2 - \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) > 0$$

$$\theta_2 < 2\frac{\rho_{YX} C_Y}{C_X} - \frac{1}{2} \quad (42)$$

ii. $MSE\left(\hat{\theta}_{ratio}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2 C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\bar{Y}^2\left(C_X^2 - 2C_{YX}\right) - \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) > 0$$

$$\theta_2 > \frac{1}{2} + \frac{\rho_{YX} C_Y}{C_X} - \frac{1}{C_X}\sqrt{\rho_{YX}^2 C_Y^2 - \left(\frac{1}{r} - \frac{1}{n}\right)\left(\frac{1}{r} - \frac{1}{N}\right)^{-1}\left(C_X^2 - 2C_{YX}\right)} \quad (43)$$

iii. $MSE\left(\hat{\theta}_{cmp}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2 C_Y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\bar{Y}^2\left(\left(1 - \lambda\right)^2 C_X^2 - 2\left(1 - \lambda\right)C_{YX}\right) - \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) > 0$$

$$\theta_2 < -\frac{1}{2} + \frac{\rho_{YX} C_Y}{C_X} + \frac{1}{C_X}\sqrt{\left(\frac{1}{r} - \frac{1}{n}\right)\left(\frac{1}{r} - \frac{1}{N}\right)^{-1}\left(\left(\lambda - 1\right)^2 C_X^2 - 2\left(\lambda - 1\right)C_{YX}\right) - \rho_{YX}^2 C_Y^2} \quad (44)$$

iv. $MSE\left(\hat{\theta}_{\exp cmp}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2\left(C_Y^2 + \left(\frac{1 - v}{2}\right)^2 C_X^2 - \left(1 - v\right)\rho_{YX} C_Y C_X\right) - \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) > 0$$

$$\theta_2 > -\frac{1}{2} + v + 4\frac{\rho_{YX} C_Y}{C_X} \quad (45)$$

v. $MSE\left(\hat{\theta}_{prasad}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2\left(\left(C_Y^2 + 0.25\mu^2 C_X^2 - \mu C_Y C_X\right)\right) / \left(1 + \left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2\left(C_Y^2 + 0.25\mu^2 C_X^2 - \mu C_Y C_X\right)\right) - \bar{Y}^2\left(\frac{1}{r} - \frac{1}{N}\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 + 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right) > 0$$

$$\left(\left(0.5\mu\right)^2 - \left(\frac{1}{2} - \theta_2\right)^2\right)C_X^2 - \left(\mu - 2\left(\frac{1}{2} - \theta_2\right)\right)\rho_{YX} C_Y C_X >$$
$$\left(\frac{1}{r} - \frac{1}{N}\right)\bar{Y}^2\left(C_Y^2 + 0.25\mu^2 C_X^2 - \mu\rho_{YX} C_Y C_X\right)\left(C_Y^2 + \left(\frac{1}{2} - \theta_2\right)^2 C_X^2 - 2\left(\frac{1}{2} - \theta_2\right)\rho_{YX} C_Y C_X\right)$$

(46)

vi. $MSE\left(\hat{\theta}_{kc1}\right) - MSE\left(t^{new}\right) > 0$

$$\left(\frac{1}{r}-\frac{1}{N}\right)\bar{y}^2 C_y^2 + \psi_{r,N} S_x^2\left(R^2 - B_{reg}^2\right) - \bar{y}^2\left(\frac{1}{r}-\frac{1}{N}\right)\bar{y}^2 C_y^2 - \left(\frac{1}{r}-\frac{1}{N}\right)\bar{y}^2\left(\frac{1}{2}-\theta_2\right)^2 C_x^2$$

$$-2\left(\frac{1}{r}-\frac{1}{N}\right)\bar{y}^2\left(\frac{1}{2}-\theta_2\right)^2 \rho_{yx} C_y C_x > 0$$

$$\theta_2 < \frac{1}{2} - \frac{\rho_{xy} C_y}{C_x} + \frac{1}{C_x}\sqrt{\psi_{r,N} S_x^2\left(R^2 - B_{reg}^2\right)\left(\frac{1}{2}-\frac{1}{N}\right)^{-1} - \rho_{xy}^2 C_y^2}$$

(47)

vii. $MSE\left(\hat{\theta}_{kc2}\right) - MSE\left(t^{new}\right) > 0$

$$MSE\left(\hat{\theta}_{mean}\right) - \psi_{r,N} S_x^2\left(R^2 - B_{reg}^2\right) - \bar{y}^2\left(\frac{1}{r}-\frac{1}{N}\right) C_y^2 + \left(\frac{1}{2}-\theta_2\right)^2 C_{x\backslash}^2 - 2\left(\frac{1}{2}-\theta_2\right)\rho_{xy} C_y > 0$$

$$\theta_2 < \frac{1}{2} - \frac{\rho_{xy} C_y}{C_x} + \frac{1}{C_x}\sqrt{\psi_{r,N} S_x^2\left(R^2 - B_{reg}^2\right)\left(\frac{1}{2}-\frac{1}{N}\right)^{-1} \bar{y}^2 - \rho_{xy}^2 C_y^2}$$

ix. $MSE\left(\hat{\theta}_{kc3}\right) - MSE\left(t^{new}\right) > 0$

$$MSE\left(\hat{\theta}_{mean}\right) + \psi_{r,n}\left(S_x^2\left(R + B_{reg}\right)^2 - 2\left(R + B_{reg}\right)S_{xy}\right) - MSE\left(t_2^{new}\right) > 0$$

$$\theta_2 < \frac{1}{2} - \frac{\rho_{xy} C_y}{C_x} + \frac{1}{C_x}\sqrt{\left[\psi_{r,n}\left(R + B_{reg}\right)^2 - 2\left(R + B_{reg}\right)S_{xy}\right]\left[\bar{y}^2\left(\frac{1}{r}-\frac{1}{N}\right)\right]^{-1} - \rho_{xy} C_y^2}$$

(48)

IV. Empirical Study

In this section, efficiency of the proposed modified estimator was compared with that of some existing estimators numerically.

Table 1: Data used for empirical study

Parameters	Pop. 1 (Murthy, 1967)	Pop. 2 (Murthy, 1967)	Pop. 3 Cochran, (1977)	Pop. 4 (Sarndal et al., 1992)
N	80	80	10	284
n	25	25	5	35
r (Assumed)	20	20	4	25
$\bar{Y}$	5182.638	5182.638	56.9	29.36
$\bar{X}$	285.125	1126.463	54.2961	245.088
C_Y	0.3542	0.3542	0.1840	1.76
C_X	0.9485	0.7507	0.1621	2.43
$\beta_1(x)$	1.2680	1.0237	0.4956	8.77
$\beta_2(x)$	3.5360	2.8306	2.5932	88.88
ρ_{YX}	0.9140	0.9140	0.9237	0.961

Table 2: MSE of Some Estimators and $\hat{t}^{new}$ using Populations 1 & 2

Estimators	Population 1			Population 2		
	BIAS	MSE	PRE	BIAS	MSE	PRE
$\hat{\theta}_{mean}$	0	126366	100	0	126366	100
$\hat{\theta}_{ratio}$	-96084.0	203055.8	62.23213	-28483.1	143279.8	88.19523
$\hat{\theta}_{cmp}$	89.97976	98215.14	128.6624	36.55499	96508.36	130.9378
$\hat{\theta}_{SH}$	-10.6729	98215.14	128.6624	-9.36634	96508.36	130.9378
$\hat{\theta}_{\exp cmp}$	13.53135	20800.34	607.5187	22.47284	14399.95	877.5447
$\hat{\theta}_{KC1}$	-310.565	926966.2	13.63221	-50.7525	582030.6	21.71122
$\hat{\theta}_{KC2}$	-227.747	713472.8	17.71139	-37.2185	460520	27.43984
$\hat{\theta}_{KC3}$	30.7674	339859.4	37.18184	16.24669	247876.5	50.9794
$\hat{\theta}_{Proposed}$	1759.951	43417.63	291.0476	2295.238	16152.71	782.3206
$\hat{t}^{new}$	-14.4796	20700.05	610.5000	-29.1756	14300.01	883.7000

Table 3: MSE of Some Estimators and $\hat{t}^{new}$ using Populations 3&4

Estimators	Population 3			Population 4		
	BIAS	MSE	PRE	BIAS	MSE	PRE
$\hat{\theta}_{mean}$	0	16.44188	100	0	97.40446	100
$\hat{\theta}_{ratio}$	-0.20684	11.77569	139.6256	142.2329	74.59707	130.5741
$\hat{\theta}_{cmp}$	-0.00346	11.76569	139.7443	0.865281	69.22218	140.7128
$\hat{\theta}_{SH}$	-0.08028	11.76569	139.7443	-1.16949	69.22218	140.7128
$\hat{\theta}_{\exp cmp}$	0.1132175	2.413311	681.2996	-0.09055	7.449396	1307.548
$\hat{\theta}_{KC1}$	-0.02323	15.17423	108.354	4.034864	193.1298	50.43471
$\hat{\theta}_{KC2}$	-0.01549	15.59678	105.4184	2.77077	163.1397	59.70616
$\hat{\theta}_{KC3}$	-0.00218	16.01933	102.6378	0.607891	127.3945	76.45891
$\hat{\theta}_{Prasad}$	59.67895	6.285896	261.5678	20.56411	14.34378	679.0712
$\hat{t}^{new}$	-0.39759	2.4001	685.0000	-5.09416	7.4004	1316.200

Tables 2 and 3 showed the results of the empirical study on the BIAS, MSE and PRE of some existing estimators and proposed modified estimators $\hat{t}^{new}$ using data sets of Populations 1, 2, 3 and 4. The results obtained showed that the proposed modified estimator $\hat{t}^{new}$ has the minimum MSEs and highest PREs among other estimators considered in the study. This implies that the proposed modified method $\hat{t}^{new}$ demonstrated high level of efficiency over others and can produce better estimate of population mean in the presence of non-response or missing observation on the average.

V. Conclusion

This study suggested new imputation scheme $\hat{t}^{new}$ as an alternative to Singh and Deo (2003), Prasad (2017), Singh and Horn (2000), Singh et al. (2014). The efficiency of the proposed modified estimator $\hat{t}^{new}$ over other estimators was demonstrated using four (4) sets of populations and the results revealed that the proposed modified estimator $\hat{t}^{new}$ has minimum MSEs and highest PREs. From the results of empirical study, it is concluded that the proposed modified estimator $\hat{t}^{new}$ is recommended for usage in Sample Survey.

VI. References

Al-Omari, A. I., Bouza, C. N. and Herrera, C. (2013). Imputation methods of missing data for estimating the population mean using simple random sampling with known correlation coefficient. Quality and Quantity, 47, 353-365.

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

Audu, A., Ishaq, O. O., Isah, U., Muhammed, S., Akintola, K. A. Rashida, A., Abubakar, A.(2020a). On the Class of ournal of Exponential-Type Imputation Estimators of Population Mean with KnownPopulation Mean of Auxiliary Variable. NIPES J Science and Technology Research 2(4), pp.1 – 11. https://doi.org/10.37933/nipes/2.4.2020.1

Audu, A, Ishaq, O. O., Muili, J. O., Zakari, Y., Ndatsu, A. M., Muhammed, S. (2020b): On the Efficiency of Imputation Estimators using Auxiliary Attribute. Continental J. Applied Sciences, 15 (1), 1-13. DOI: 10.5281/zenodo.3721046

Audu, A., Ishaq, O. O., Zakari, Y., Wisdom, D. D., Muili J. and Ndatsu, A. M. (2020c). Regression-cum-exponential ratio imputation class of estimators of population mean in the presence of non-response. Science Forum Journal of Pure and Applied Science, 20, 58-63. DOI: http://dx.doi.org/10.5455/sf.71109

Bhushan, S. and Pandey, A. P. (2016). Optimal imputation of missing data for estimation of population mean. Journal of Statistics and Management Systems, 19 (6), 755-769.

Cochran, W. G. (1977). Sampling Techniques, 3rd edn, Wiley and Sons, 1977.

Diana, G. and Perri, P. F., (2010). Improved estimators of the population mean for missing data, Communications in Statistics- Theory and Methods, 39, 3245-3251.

Gira, A. A. (2015). Estimation of population mean with a New Imputation Methods. Applied Mathematical Sciences, 9(34), 1663-1672.

Kadilar, C. and Cingi, H. (2004). Ratio Estimators in Simple Random Sampling. Appl. Math. Computat. 151, 893 – 902

Kadilar, C. and Cingi, H. (2008). Estimators for the population mean in the case of missing data, Communications in Statistics- Theory and Methods, 37, 2226-2236.

Murthy, M. N. (1967). Sampling theory and methods. Statistical Publishing Society, Calcutta, India.

Prasad, S. (2017). A study on new methods of ratio exponential type imputation in sample surveys. Hacettepe Journal of Mathematics and Statistics, DOI: 10.15672/HJMS.2016.392.

Sarndal, C. E., Swensson, B. and Wretman, J. H. (1992). Model

assisted survey sampling. New York: Springer-Verlag

Singh, P., Bouza, C. and Singh, R. (2019): Generalized exponential estimator for estimating the population mean using auxiliary variable. Jour. of Sci. Res., 63, 273 – 280

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

Singh, A. K., Singh, P. and Singh, V. K. (2014). Exponential-Type Compromised Imputation in Survey Sampling, J. Stat. Appl. 3 (2), 211-217.

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

Singh, S. and Deo, B. (2003). Imputation by power transformation. Statistical Papers, 44, 555-579. Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. Metrika, 51, 267-s276.

Singh, S., (2009). A new method of imputation in survey sampling. Statistics, 43, 499-511.

Toutenburg, H., Srivastava, V. K. and Shalabh, A. (2008). Imputation versus imputation of missing values through ratio method in sample surveys. Statistical Papers, 49, 237-247.

Wang, L. and Wang, Q., (2006). Empirical likelihood for parametric model under imputation for missing data. Journal of Statistics and Management Systems, 9 (1), 1-13.

***