Asian Research Journal of Mathematics

17(5): 1-13, 2021; Article no.ARJOM.63925
ISSN: 2456-477X

Regression-type Imputation Class of Estimators using Auxiliary Attributes

A. Audu[1], O. O. Ishaq[2*], A. Abubakar[1], K. A. Akintola[3], U. Isah[1], A. Rashida[4] and S. Muhammed[1]

[1]*Department of Mathematics, Usmanu Danfidiyo University, Sokoto, Nigeria.*
[2]*Department of Statistics, Kano University of Science and Technology, Wudil, Nigeria.*
[3]*Department of Statistics, Oyo State College of Agriculture and Technology, Igboora, Nigeria.*
[4]*State College of Basic and Remedial Studies, Sokoto, Nigeria.*

Authors' contributions:

This work was carried out in collaboration among all authors. Authors AA, OOI, AA, KAA, UI, AR and SM designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors AA and OOI managed the analyses of the study. Authors AA and SM managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ARJOM/2021/v17i530296
Editor(s):
(1) Dr. Ruben Dario Ortiz Ortiz, Universidad Michoacana de San Nicolas de Hidalgo, Mexico.
Reviewers:
(1) Totohasina André, University of Antsiranana, Madagascar.
(2) Atiya R. Kazi, Mumbai University, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/63925

Received: 22 October 2020
Accepted: 28 December 2020
Published: 08 July 2021

Original Research Article

Abstract

Several imputation schemes and estimators have been proposed by different authors in sample survey. However, these estimators utilized quantitative information of auxiliary characters. In this study, some imputation methods were studied using qualitative information of auxiliary characters and two new imputation schemes using auxiliary attribute have been suggested. The mean squared errors of the proposed estimators were derived up to first order approximation using Taylor series approach. Numerical illustrations with two populations were conducted and the results revealed that the proposed estimator is more efficient.

Keywords: Imputation; non-response; estimator; population mean; attribute.

**Corresponding author: E-mail: ishaq.o.o@kustwudil.edu.ng, babington4u@gmail.com;*

1 Introduction

Different researches in sample survey have shown that auxiliary characters play important role in the enhancement of ratio, product and regression estimators of population characteristics especially when the study and auxiliary variables are strongly correlated. Several authors have employed the concept of auxiliary variables in the development and improvement estimators. Authors like Audu et al [1,2], Muili et al. [3], Audu and Adewara [4], Audu and Ishaq [5], Ishaq and Audu [6], Audu et al. [6], Singh and Audu [7], Singh and Audu [8], Ahmed et al. [9]. However, when the study variables are characterized by non-response, the aforementioned estimators are not applicable. Authors like Singh and Horn [10], Singh and Deo [11], Wang and Wang [12], Toutenburg et al. [13], Kadilar and Cingi [14], Singh [15], Diana and Perri [16], Al-Omari et al. [17], Singh et al. [18], Gira (2015), Singh et al. [19], Bhushan and Pandey [20], Prasad [21], Audu et al. [22-24], have studied different schemes and estimators in the presence of non-response. Situations arise when the auxiliary characters are qualitative in nature e.g. gender, marital status, family history on a disease, patient status with respect to disease, and of which the imputation schemes proposed by aforementioned authors will be impracticable. In the present study, we consider generalized imputation schemes when the auxiliary character is qualitative.

2 Existing Imputation Schemes using Auxiliary Attribute

Consider Ψ as the set of r units response and Ψ^c be the set of (N-n) units non-response sampled without replacement from the N units population.

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,

$$y_{.i} = \begin{cases} y_i & \text{if} \quad i \in \Psi \\ \bar{y}_r & \text{if} \quad i \in \Psi^c \end{cases} \tag{2.1}$$

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

$$\hat{\tau}_0 = \frac{1}{r}\sum_{i \in R} y_i = \bar{y}_r \tag{2.2}$$

The variance of $\hat{\tau}_0$ is given by (2.3).

$$Var(\hat{\tau}_0) = \left(\frac{1}{r} - \frac{1}{N}\right) S_Y^2 \tag{2.3}$$

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

Ratio imputation estimator τ_1 and imputation estimators proposed by Singh and Horn [10] τ_2, Singh and Deo [11] τ_3, Ahmed et al. [25] τ_4, Singh [15] τ_5, Singh et al. [18] τ_6 and Singh and Gogoi [26] τ_7 when auxiliary character is qualitative and their MSEs are given below;

$$\hat{\tau}_1 = \bar{y}_r \frac{p_n}{p_r} \tag{2.4}$$

$$MSE(\hat{\tau}_1) = \left(\left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2 + \left(\frac{1}{r}-\frac{1}{n}\right)\left(R^{*2}S_\pi^2 - 2\rho_{y\pi}R^{*}S_Y S_\pi\right)\right) \tag{2.5}$$

$$\hat{\tau}_2 = \bar{y}_r\left(\lambda + (1-\lambda)\frac{p_n}{p_r}\right) \tag{2.6}$$

$$MSE(\hat{\tau}_2)_{\min} = S_Y^2\left(\left(\frac{1}{r}-\frac{1}{N}\right)-\left(\frac{1}{r}-\frac{1}{n}\right)\rho_{Y\pi}^2\right) \tag{2.7}$$

$$\hat{\tau}_3 = \bar{y}_r\left(\frac{p_n}{p_r}\right)^{\phi} \tag{2.8}$$

$$MSE(\hat{\tau}_3)_{\min} = S_Y^2\left(\left(\frac{1}{r}-\frac{1}{N}\right)-\left(\frac{1}{r}-\frac{1}{n}\right)\rho_{Y\pi}^2\right) \tag{2.9}$$

$$\hat{\tau}_4 = \bar{y}_r\left(\frac{P}{p_r}\right)^{\beta} \tag{2.10}$$

$$MSE(\hat{\tau}_4) = S_Y^2\left(\left(\frac{1}{r}-\frac{1}{N}\right)-\left(\frac{1}{r}-\frac{1}{n}\right)\rho_{Y\pi}^2\right) \tag{2.11}$$

$$\hat{\tau}_5 = \frac{\bar{y}_r p_n}{\alpha p_r + (1-\alpha)p_n} \tag{2.12}$$

$$MSE(\hat{\tau}_5)_{\min} = S_Y^2\left(\left(\frac{1}{r}-\frac{1}{N}\right)-\left(\frac{1}{r}-\frac{1}{n}\right)\rho_{Y\pi}^2\right) \tag{2.13}$$

$$\hat{\tau}_6 = \kappa\bar{y}_r + (1-\kappa)\bar{y}_r \exp\left(\frac{P-p_r}{P+p_r}\right) \tag{2.14}$$

$$MSE(\hat{\tau}_6)_{\min} = \left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2\left(1-\rho_{Y\pi}^2\right) \tag{2.15}$$

$$\hat{\tau}_7 = w\bar{y}_r + (1-w)\bar{y}_r \exp\left(\frac{p^*-P}{p^*+P}\right) \tag{2.16}$$

where $p^* = (NP - np_r)/(N-n)$.

$$MSE(\hat{\tau}_7)_{\min} = \left(\frac{1}{r}-\frac{1}{N}\right)S_Y^2\left(1-\rho_{Y\pi}^2\right) \tag{2.17}$$

where, $p_r = \frac{1}{r}\sum_{i\in R}\pi_i, p_n = \frac{1}{n}\sum_{i\in S}\pi_i, \rho_{Y\pi} = \frac{S_{Y\pi}}{S_Y S_\pi}, R^* = \bar{Y}/P, S_\pi^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\pi_i - P)^2,$

$P = \frac{1}{N}\sum_{i=1}^{N}\pi_i, S_{Y\pi} = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(\pi_i - P\right), \lambda = 1 - R^*\beta_\pi, \phi = \beta = R^*\beta_\pi, \kappa = 1 - 2R^*\beta_\pi,$

$\beta_\pi = S_{Y\pi} / S_\pi^2$

Audu et al. [23] studied and proposed imputation schemes using auxiliary attribute given as:

$$y_{.i} = \begin{cases} y_i & \text{if} \quad i \in \Psi \\ \frac{\bar{y}_r}{n-r}\left(n\left(\Lambda_1 \frac{P}{p_r} + \Lambda_2 \frac{p_r}{P}\right)\exp\left(\frac{p_r - P}{p_r + P}\right) - r\right) & \text{if} \quad i \in \Psi^c \end{cases} \tag{2.18}$$

where $\Lambda_1 \neq 0$ and $\Lambda_2 \neq 0$ are unknown functions of study variable and auxiliary attribute.

The point estimators of finite population mean under this scheme denoted by t_8 is given by

$$\hat{\tau}_8 = \bar{y}_r\left(\Lambda_1 \frac{P}{p_r} + \Lambda_2 \frac{p_r}{P}\right)\exp\left(\frac{p_r - P}{p_r + P}\right) \tag{2.19}$$

$$Bias\left(\hat{\tau}_8\right) \approx \bar{Y}\left(\Lambda_1\left(1 + \frac{\theta_r}{8}\left(3C_p^2 - 4\rho C_Y C_p\right)\right) + \Lambda_2\left(1 + \frac{\theta_r}{8}\left(3C_p^2 - 12\rho C_Y C_p\right)\right) - 1\right) \tag{2.20}$$

$$MSE\left(\hat{\tau}_8\right) = \bar{Y}^2\left(1 + \Lambda_1^2\Psi_1 + \Lambda_2^2\Psi_2 - 2\Lambda_1\Psi_3 - 2\Lambda_2\Psi_4 + 2\Lambda_2\Lambda_2\Psi_5\right) \tag{2.21}$$

where

$\Psi_1 = 1 + \theta_r\left(C_Y^2 + \frac{1}{2}C_P^2 - 2\rho C_Y C_P\right), \Psi_2 = 1 + \theta_r\left(C_Y^2 + 3C_X^2 + 6\rho C_Y C_P\right), \Psi_3 = 1 + \frac{\theta_r}{8}\left(3C_P^2 - 4\rho C_Y C_P\right)$

$\Psi_4 = 1 + \frac{\theta_r}{8}\left(3C_P^2 - 12\rho C_Y C_P\right), \ \Psi_5 = 1 + \theta_r\left(C_Y^2 + 2\rho C_Y C_P\right)$

$$MSE\left(t_1^*\right)_{\min} = \bar{Y}^2\left(1 - \frac{\Psi_2\Psi_3^2 + \Psi_1\Psi_4^2 - 2\Psi_3\Psi_4\Psi_5}{\Psi_1\Psi_2 - \Psi_5^2}\right) \tag{2.22}$$

However, the existing estimators mentioned above are functions of unknown parameters which need to be estimated form sample observation before the estimators can be applicable in real life situations. To overcome the shortcoming identified above, two new class of imputation schemed are proposed to obtain new imputation estimators which are independent of unknown parameters.

3 Proposed Estimator under Imputation

Inspired by Audu and Singh [3], we proposed the following generalized class of imputation schemes;

$$y_{.i} = \begin{cases} y_i & if \quad \mathrm{i} \in \Psi \\ \dfrac{\bar{y}_r + \hat{\beta}_\pi (P - p_r)}{\varpi_1 p_r + \varpi_2} (\varpi_1 P + \varpi_2) \exp\left(\dfrac{P - p_r}{P + p_r} \right) & if \quad \mathrm{i} \in \Psi^{\mathrm{c}} \end{cases} \tag{3.1}$$

$$y_{.i} = \begin{cases} y_i & if \quad \mathrm{i} \in \Psi \\ \dfrac{\bar{y}_r + \hat{\beta}_\pi (p^* - P)}{\varpi_1 p^* + \varpi_2} (\varpi_1 P + \varpi_2) \exp\left(\dfrac{p^* - P}{p^* + P} \right) & if \quad \mathrm{i} \in \Psi^{\mathrm{c}} \end{cases} \tag{3.2}$$

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, $\hat{\beta}_\pi = s_{y\pi} / s_\pi^2$, $s_\pi^2 = \dfrac{1}{r-1} \sum_{j=1}^{r} (\pi_j - p_r)$, $s_{y\pi} = \dfrac{1}{r-1} \sum_{j=1}^{r} (y_j - \bar{y}_r)(\pi_j - p_r)$.

3.1 Remark

Note that $\varpi_1 \neq \varpi_2$ and $\varpi_1 \neq 0$.

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

$$T_i = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \frac{\bar{y}_r + \hat{\beta}_\phi (P - p_r)}{\varpi_1 p_r + \varpi_2} (\varpi_1 P + \varpi_2) \exp\left(\frac{P - p_r}{P + p_r} \right) \tag{3.3}$$

$$T_i^{(*)} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \frac{\bar{y}_r + \hat{\beta}_\phi (p^* - P)}{\varpi_1 p^* + \varpi_2} (\varpi_1 P + \varpi_2) \exp\left(\frac{p^* - P}{p^* + P} \right) \tag{3.4}$$

3.2 Remark

The proposed class of imputation estimators is independent of unknown parameters, hence it is practically applicable.

Table 1. Some member of T_i for different values of ϖ_1 and ϖ_2

i	**Estimators**	**Values of Constants**	
		ϖ_1	ϖ_2
1	$T_1 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \frac{\bar{y}_r + \hat{\beta}_\pi (P - p_r)}{p_r + C_\pi} (P + C_\pi) \exp\left(\frac{P - p_r}{P + p_r} \right)$	1	C_X

i	**Estimators**	**Values of Constants**	
		ϖ_1	ϖ_2
2	$T_2=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{p_r+\beta_1(\pi)}\left(P+\beta_1(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	1	$\beta_1(\pi)$
3	$T_3=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{p_r+\beta_2(\pi)}\left(P+\beta_2(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	1	$\beta_2(\pi)$
4	$T_4=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{p_r+S_\pi}\left(P+S_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	1	S_π
5	$T_5=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{C_\pi p_r+\beta_1(\pi)}\left(C_\pi P+\beta_1(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	C_π	$\beta_1(\pi)$
6	$T_6=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{C_\pi p_r+\beta_2(\pi)}\left(C_\pi P+\beta_2(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	C_π	$\beta_2(\pi)$
7	$T_7=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{C_\pi p_r+S_\pi}\left(C_\pi P+S_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	C_π	S_π
8	$T_8=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_1(\pi)p_r+C_\pi}\left(\beta_1(\pi)P+C_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_1(\pi)$	C_π
9	$T_9=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_1(\pi)p_r+\beta_2(\pi)}\left(\beta_1(\pi)P+\beta_2(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_1(\pi)$	$\beta_2(\pi)$
10	$T_{10}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_1(\pi)p_r+S_\pi}\left(\beta_1(\pi)P+S_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_1(\pi)$	S_π
11	$T_{11}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_2(\pi)p_r+C_\pi}\left(\beta_2(\pi)P+C_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_2(\pi)$	C_π
12	$T_{12}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_2(\pi)p_r+\beta_1(\pi)}\left(\beta_2(\pi)P+\beta_1(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_2(\pi)$	$\beta_1(\pi)$
13	$T_{13}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{\beta_2(\pi)p_r+S_\pi}\left(\beta_2(\pi)P+S_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	$\beta_2(\pi)$	S_π
14	$T_{14}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{S_\pi p_r+C_\pi}\left(S_\pi P+C_\pi\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	S_π	C_π
15	$T_{15}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{S_\pi p_r+\beta_1(\pi)}\left(S_\pi P+\beta_1(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	S_π	$\beta_1(\pi)$
16	$T_{16}=\frac{r}{n}\bar{y}_r+\left(1-\frac{r}{n}\right)\frac{\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)}{S_\pi p_r+\beta_2(\pi)}\left(S_\pi P+\beta_2(\pi)\right)\exp\left(\frac{P-p_r}{P+p_r}\right)$	S_π	$\beta_2(\pi)$

4 Properties of the Estimators Suggested

Theorem 4.1: The MSE of the suggested estimator $T_i,\ i=1,2,3,...,16$ to $\mathrm{O}\left(n^{-1}\right)$ is:

$$MSE\left(T_i\right)=\theta_{r,N}\left(S_Y^2+\left(1-\frac{r}{n}\right)^2\left(\beta_\pi+\left(\frac{1}{2}+\gamma_i\right)R^*\right)^2 S_\pi^2-2\left(1-\frac{r}{n}\right)\left(\beta_\pi+\left(\frac{1}{2}+\gamma_i\right)R^*\right)S_{Y\pi}\right) \tag{4.1}$$

where

$R^*=\bar{Y}/P,\ \gamma_1=P/\left(P+C_X\right),\gamma_2=P/\left(P+\beta_1(x)\right),\gamma_3=P/\left(P+\beta_2(x)\right),\gamma_4=P/\left(P+S_X\right)$,
$\gamma_5=C_X P/\left(C_X P+\beta_1(x)\right),\gamma_6=C_X P/\left(C_X P+\beta_2(x)\right),\gamma_7=C_X P/\left(C_X P+S_X\right),$
$\gamma_8=\beta_1(x)P/\left(\beta_1(x)P+C_X\right),\gamma_9=\beta_1(x)P/\left(\beta_1(x)P+\beta_2(x)\right),\gamma_{10}=\beta_1(x)P/\left(\beta_1(x)P+S_X\right),$
$\gamma_{11}=\beta_2(x)P/\left(\beta_2(x)P+C_X\right),\gamma_{12}=\beta_2(x)P/\left(\beta_2(x)P+\beta_1(x)\right),\gamma_{13}=\beta_2(x)P/\left(\beta_2(x)P+S_X\right),$
$\gamma_{14}=S_X P/\left(S_X P+C_X\right),\gamma_{15}=S_X P/\left(S_X P+\beta_1(x)\right),\ \gamma_{16}=S_X P/\left(S_X P+\beta_2(x)\right)$

Proof: $MSE\left(T_i\right)$ can be derived using up to $\mathrm{O}\left(n^{-1}\right)$ using Taylor series approach given as:

$$MSE\left(T_i\right)=\Delta\Sigma\Delta^{'} \tag{4.2}$$

where Δ is a 1×2 matrix, Σ is a 2×2 variance-covariance matrix, $\Delta=\left(\left.\frac{\partial T_i}{\partial \bar{y}_r}\right|\bar{Y},\bar{X},\beta_\pi \quad \left.\frac{\partial T_i}{\partial \bar{x}_r}\right|\bar{Y},\bar{X},\beta_\pi\right)$, $\Sigma=\begin{pmatrix}\mathrm{var}\left(\bar{y}_r\right) & \mathrm{cov}\left(\bar{y}_r\bar{x}_r\right)\\ \mathrm{cov}\left(\bar{x}_r\bar{y}_r\right) & \mathrm{var}\left(p_r\right)\end{pmatrix}$

$$\frac{\partial T_i}{\partial \bar{y}_r}=\frac{r}{n}+\left(1-\frac{r}{n}\right)\frac{\left(\varpi_1 P+\varpi_2\right)}{\varpi_1 p_r+\varpi_2}\exp\left(\frac{P-p_r}{P+p_r}\right) \tag{4.3}$$

$$\left.\frac{\partial T_i}{\partial \bar{y}_r}\right|\bar{y}_r=\bar{Y},\,p_r=P,\,\hat{\beta}_\pi=\beta_\pi \quad =1 \tag{4.4}$$

$$\begin{aligned}\frac{\partial T_i}{\partial p_r}=-\left(1-\frac{r}{n}\right)\left(\varpi_1 P+\varpi_2\right)\exp\left(\frac{P-p_r}{P+p_r}\right)&\left(2P\left(\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)\right)/\left(P+p_r\right)^2\right.\\ &\left.+\hat{\beta}_\pi+\varpi_1\left(\bar{y}_r+\hat{\beta}_\pi\left(P-p_r\right)\right)\right)/\left(\varpi_1 p_r+\varpi_2\right)^2\end{aligned} \tag{4.5}$$

$$\left.\frac{\partial T_i}{\partial p_r}\right|\bar{Y},P,\beta_\pi \quad =-\left(1-\frac{r}{n}\right)\left(\beta_\pi+\frac{\bar{Y}}{2P}+\gamma_i\frac{\bar{Y}}{P}\right) \tag{4.6}$$

where $\gamma_i=\varpi_1 P/\left(\varpi_1 P+\varpi_2\right)$

So, from the definition of Δ, we have

$$\Delta=\left(1 \qquad -\left(1-\frac{r}{n}\right)\left(\beta_{\pi}+\frac{\bar{Y}}{2P}+\gamma_i\frac{\bar{Y}}{P}\right)\right) \tag{4.7}$$

Put (4.7) in (4.2), we obtained (4.1).

Theorem 4.2: The MSE of the suggested estimator $T_i^*,\ i=1,2,3,\dots,16$ to $O\left(n^{-1}\right)$ is:

$$MSE\left(T_i^*\right)=\theta_{r,N}\left(S_Y^2+\left(1-\frac{r}{n}\right)^2 f^{*2}\left(\beta_{\pi}+\left(\frac{1}{2}+\gamma_i\right)R^*\right)^2 S_{\pi}^2-2\left(1-\frac{r}{n}\right)f^*\left(\beta_{\pi}+\left(\frac{1}{2}+\gamma_i\right)R^*\right)S_{Y\pi}\right) \tag{4.8}$$

where $f^*=n/\left(N-n\right)$

Proof: $MSE\left(T_i^*\right)$ can be derived using up to $O\left(n^{-1}\right)$ using Taylor series approach given as:

$$MSE\left(T_i^*\right)=\Delta^*\Sigma\Delta^{*'} \tag{4.9}$$

where Δ is a 1×2 matrix, Σ is a 2×2 variance-covariance matrix,

$$\Delta^*=\left(\left.\frac{\partial T_i^*}{\partial \bar{y}_r}\right|\bar{Y},\bar{X},\beta_{\phi} \qquad \left.\frac{\partial T_i^*}{\partial \bar{x}_r}\right|\bar{Y},\bar{X},\beta_{\pi}\right)$$

$$\frac{\partial T_i^*}{\partial \bar{y}_r}=\frac{r}{n}+\left(1-\frac{r}{n}\right)\frac{\left(\varpi_1 P+\varpi_2\right)}{\varpi_1 p^*+\varpi_2}\exp\left(\frac{p^*-P}{p^*+P}\right) \tag{4.10}$$

$$\left.\frac{\partial T_i^*}{\partial \bar{y}_r}\right|\bar{Y},P,\beta_{\pi}=1 \tag{4.11}$$

$$\begin{aligned}\frac{\partial T_i^*}{\partial p_r}=&-\left(1-\frac{r}{n}\right)\frac{n\left(\varpi_1 P+\varpi_2\right)}{N-n}\exp\left(\frac{p^*-P}{p^*+P}\right)\Big(2P\left(\bar{y}_r+\hat{\beta}_{\pi}\left(p^*-P\right)\right)/\left(P+p^*\right)^2\\&+\hat{\beta}_{\pi}+\varpi_1\left(\bar{y}_r+\hat{\beta}_{\pi}\left(p^*-P\right)\right)\Big)/\left(\varpi_1 p^*+\varpi_2\right)^2\end{aligned} \tag{4.12}$$

$$\left.\frac{\partial T_i^*}{\partial p_r}\right|\bar{Y},P,\beta_{\pi}=-\left(1-\frac{r}{n}\right)\left(\frac{n}{N-n}\right)\left(\beta_{\pi}+\frac{\bar{Y}}{2P}+\gamma_i\frac{\bar{Y}}{P}\right) \tag{4.13}$$

So, from the definition of Δ, we have

$$\Delta^* = \left(1 \qquad -\left(1-\frac{r}{n}\right)\left(\frac{n}{N-n}\right)\left(\beta_\pi + \frac{\bar{Y}}{2P} + \gamma_i \frac{\bar{Y}}{P}\right)\right) \tag{4.14}$$

Put (4.14) in (4.9), we obtained (4.8).

5 Numerical Illustration

For the empirical justification of the results, we consider five sets of real data. The performance of the proposed estimator is justified by comparing its MSE to those of some existing estimators considered in the study.

Population I: Source [25]

$Y =$ The number of villages in the circle,

$$\pi = \begin{cases} 1, & if\ Y > 5 \\ 0, & if\ Y \le 5 \end{cases}$$

Population II: Source [27]

$Y =$ Area (in Acres) under wheat crop in the circles,

$$\pi = \begin{cases} 1, & if\ Y > 5 \\ 0, & if\ Y \le 5 \end{cases}$$

Table 2. Descriptive statistics of the populations

Population I: $N=89, n=23, P=0.124, \bar{Y}=3.3596, \rho_{Y\phi}=0.766, C_Y=0.6008, C_\phi=2.6779, \beta_2=6.162$
Population II: $N=89, n=23, P=0.124, \bar{Y}=1102, \rho_{Y\phi}=0.624, C_Y=0.65, C_\phi=2.6779, \beta_2=6.162$

Table 3. MSE and PRE of proposed and other estimators using population I

Estimators	MSE	PRE	Estimators	MSE	PRE
$r=15$ (Assumed)					
$\hat{\tau}_0$	0.2260362	100	$\hat{\tau}_j, j=2,3,...,7$	0.1536047	147.1544
$\hat{\tau}_1$	1.458	5.50317	$\hat{\tau}_8$	0.3224198	70.10618
Proposed T_i			**Proposed** T_i^*		
T_1	0.1201507	188.1273	T_1^*	0.1259095	179.5228
T_3	0.1377838	164.0514	T_3^*	0.120258	187.9593
T_4	0.2031452	111.2683	T_4^*	0.1082503	208.8088
T_6	0.1905247	118.6388	T_6^*	0.1099639	205.5548
T_7	0.3427627	65.94538	T_7^*	0.09740895	232.0487

T_{11}	0.1203748	187.777	T_{11}^{*}	0.1258236	179.6453
T_{13}	0.2052654	110.119	T_{13}^{*}	0.1079811	209.3294
T_{14}	0.113595	198.9842	T_{14}^{*}	0.1286607	175.684
T_{16}	0.1188363	190.2081	T_{16}^{*}	0.1264228	178.7938
$r=20$ (Assumed)					
$\hat{\tau}_0$	0.1580726	100	$\tau_j, j=2,3,...,7$	0.1377013	114.7939
$\hat{\tau}_1$	0.5045625	31.32865	$\hat{\tau}_8$	0.2664749	59.31988
Proposed $\hat{\tau}_i$			**Proposed** $\hat{\tau}_i^{*}$		
T_1	0.0846603	186.714	T_1^{*}	0.1262716	125.1846
T_3	0.08073266	195.7976	T_3^{*}	0.1237309	127.7551
T_4	0.07279778	217.1393	T_4^{*}	0.1175568	134.4649
T_6	0.0738488	214.049	T_6^{*}	0.1184973	133.3976
T_7	0.06646145	237.841	T_7^{*}	0.1094977	144.3616
T_{11}	0.0846186	186.806	T_{11}^{*}	0.1262458	125.2102
T_{13}	0.07267728	217.4993	T_{13}^{*}	0.1174454	134.5925
T_{14}	0.08657643	182.5816	T_{14}^{*}	0.1274346	124.0422
T_{16}	0.08500116	185.9652	T_{16}^{*}	0.1264817	124.9766

Table 4. MSE and PRE of proposed and other estimators using population II

Estimators	MSE	PRE	Estimators	MSE	PRE
$r=15$ (Assumed)					
$\hat{\tau}_0$	28440.7	100	$\tau_j, j=2,3,...,7$	21016.58	135.3251
$\hat{\tau}_1$	169221.2	16.80682	$\hat{\tau}_8$	40929.16	69.48763
Proposed $\hat{\tau}_i$			**Proposed** $\hat{\tau}_i^{*}$		
T_1	21323.15	170.0692	T_1^{*}	19541.98	145.5365
T_3	23434.9	133.3795	T_3^{*}	19081.74	149.0467
T_4	31270.33	121.3604	T_4^{*}	18121.18	156.9473
T_6	29598.11	90.95109	T_6^{*}	18267.79	155.6877
T_7	47310.5	96.08959	T_7^{*}	17434.78	163.1263
T_{11}	21377.7	60.115	T_{11}^{*}	19527.95	145.641
T_{13}	31746.83	133.0391	T_{13}^{*}	18083.31	157.276
T_{14}	20487.58	89.58598	T_{14}^{*}	19776.69	143.8092
T_{16}	21135.36	138.8193	T_{16}^{*}	19591.37	145.1696
$r=20$ (Assumed)					
$\hat{\tau}_0$	19889.28	100	$\hat{\tau}_j, j=2,3,...,7$	17801.24	111.7297
$\hat{\tau}_1$	59483.79	33.43647	$\hat{\tau}_8$	33527.19	59.32283

Proposed T_i			**Proposed** T_i^*		
T_1	13389.26	148.5465	T_1^*	16992.92	117.0445
T_3	13074.11	152.1271	T_3^*	16763.14	118.6488
T_4	12469.74	159.5004	T_4^*	16191.86	122.8351
T_6	12550.8	158.4702	T_6^*	16286.56	122.1208
T_7	12145.41	163.7596	T_7^*	15461.62	128.6364
T_{11}	13382.57	148.6207	T_{11}^*	16988.32	117.0762
T_{13}	12454.04	159.7014	T_{13}^*	16172.33	122.9834
T_{14}	13549.78	146.7867	T_{14}^*	17100.58	116.3076
T_{16}	13420.46	148.2012	T_{16}^*	17014.27	116.8976

Tables 3 and 4 show the numerical results of MSE and PRE (percentage relative efficiency) of the proposed and other estimators considered in the study using population sets I and II respectively. Of all the estimators considered in the study, the proposed estimators have minimum MSEs and higher PREs for the two population sets except for data set I when $r = 15$ where proposed estimator T_7 performed poorly and for data set II when $r = 15$ where proposed estimators T_6, T_7, T_{11}, T_{14} performed poorly. This implies that the proposed estimators T_i^*, $i = 1, 2, ..., 16$ and $T_1, T_2, T_3, T_4, T_5, T_8, T_9, T_{10}, T_{12}, T_{13}, T_{15}, T_{16}$ demonstrated high level of efficiency over others and can produce better estimate of population mean in the presence of non-response on the average.

6 Conclusion

From the results of the empirical study in section 4, it was observed that the proposed estimators T_i^*, $i = 1, 2, ..., 16$ and T_i, $i = 1, 2, ..., 16$ with exception of T_6, T_7, T_{11}, T_{14}, are more efficient than other estimators considered in the study and therefore, it is recommended for use for estimating population mean when the study variable is associated with an attribute in the presence of non-response.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Audu A, Ishaq OO, Isah U, Muhammad S, Akintola KA, Rashida A, Abubakar A. On the class of exponential-type imputation estimators of population mean with known population mean of auxiliary variable. Journal of Science and Technology Research, 2020a;2(4):1-11.
Available:https://doi.org/10.37933/nipes/2.4.2020.1

[2] Audu A, Ishaq OO, Muili JO, Abubakar A, Rashida A, Akintola KA, Isah U. Modified estimators of population mean using robust multiple regression methods. Journal of Science and Technology Research. 2020a;2(4):12-20.
Available:https://doi.org/10.37933/nipes/2.4.2020.2

[3] Muili JO, Agwamba EN, Erinola YA, Yunusa MA, Audu A, Hamzat MA. A family of ratio-type estimators of population mean using two auxiliary variables. Asian Journal of Research in Computer Science; 2020.

DOI: 10.9734/AJRCOS/2020/v6i130152

[4] Audu A, Adewara AA. Modified factor-type estimators with two auxiliary variables under two-phase sampling. Anale. Seria Informatica. 2017;15(1):63-76.

[5] Audu A, Ishaq OO. Isaki, Olufadi and Kadilar (2014) Variance Estimators under Transformation of Sample Means. Journal of Nigerian Association of Mathematical Physics. 2016;38:167-172.

[6] Ishaq OO, Audu A. Alternative to mohanty (1967) and samiuddin and hanif (2006) regression-cum-ratio finite population mean estimators. Journal of Nigerian Association of Mathematical Physics. 2016;38:173-178.

[7] Singh RVK, Audu A. Improved exponential ratio–product type estimator for finite population mean. International Journal of Engineering Science and Innovative Technology. 2015;4(3):317-322.

[8] Singh RVK, Audu A. Efficiency of ratio estimators in stratified random sampling using information on auxiliary attribute. International Journal of Engineering Science and Innovative Technology. 2013;l2(1):166-172.

[9] Ahmed A, Adewara AA, Singh RVK. Class of ratio estimators with known functions of auxiliary variable for estimating finite population variance. Asian Journal of Mathematics and Computer Research. 2016;12(1):63-70.

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

[11] Singh S, Deo B. Imputation by power transformation. Statistical Papers. 2003;44:555-579.

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

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

[14] Kadilar C, Cingi H. Estimators for the population mean in the case of missing data, Communications in Statistics- Theory and Methods. 2008;37:2226-2236.

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

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

[17] 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.

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

[19] 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.

[20] 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.

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

[22] Audu A, Ishaq OO, Zakari Y, Wisdom DD, Muili JO, 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 Science. 2020c;20:58-63.
DOI: http://dx.doi.org/10.5455/sf.71109

[23] Audu A, Ishaq OO, Muili JO, Zakari Y, Ndatsu AM, Muhammed S. On the efficiency of imputation estimators using auxiliary attribute. Continental J. Applied Sciences. 2020d;15(1):1-13.
DOI: 10.5281/zenodo.3721046

[24] Audu A, Singh RVK. Exponential-type regression compromised imputation class of estimators. Journal of Statistics & Management Systems. 2020;1-14.
DOI: 10.1080/09720510.2020.1814501

[25] Ahmed MS, Al-Titi Z, Al-Rawi Z, Abu-Dayyeh W. Estimation of a population mean using different imputation methods. Trans. 2006;7:1247-1264.

[26] Singh, Gogoi. Estimation of population mean using ratio-cum-product imputation technique in sample survey under two-phase samping. Int. J. Math. Stat. 2017;19:26-46.

[27] Sukhatme PV, Suhatme BV. Sampling theory of surveys with applications. Ames, IA: Iowa State University Press; 1970.

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)
http://www.sdiarticle4.com/review-history/63925