Regression-type Imputation Class of Estimators using Auxiliary Attributes

A. Audu¹, O. O. Ishaq²*, A. Abubakar¹, K. A. Akintola³, U. Isah¹, A. Rashida⁴ and S. Muhammed¹

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

Authors’ contributions:
This work was carried out in collaboration among all authors. Authors AA, OO, AA, KA, 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.

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 \( \Psi \) as the set of \( r \) units response and \( \Psi^c \) 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,

\[
y_j = \begin{cases} 
y_i & \text{if } i \in \Psi \\
\bar{y} & \text{if } i \in \Psi^c
\end{cases}
\]

(2.1)

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

\[
\hat{\tau}_0 = \frac{1}{r} \sum_{i \in R} y_i = \bar{y}
\]

(2.2)

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

\[
Var(\hat{\tau}_0) = \left( \frac{1}{r} - \frac{1}{N} \right) S^2_y
\]

(2.3)

where \( S^2_y = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2 \), \( \bar{Y} = \frac{1}{N} \sum_{i=1}^{N} y_i \)

Ratio imputation estimator \( \tau_1 \) and imputation estimators proposed by Singh and Horn [10] \( \tau_2 \), Singh and Deo [11] \( \tau_3 \), Ahmed et al. [25] \( \tau_4 \), Singh [15] \( \tau_5 \), Singh et al. [18] \( \tau_6 \) and Singh and Gogoi [26] \( \tau_7 \) when auxiliary character is qualitative and their MSEs are given below;

\[
\hat{\tau}_i = \frac{\bar{y} P_n}{P_r}
\]

(2.4)
\[
MSE(\hat{\tau}_1) = \left(\frac{1}{r} - \frac{1}{N}\right)S_y^2 + \left(\frac{1}{r} - \frac{1}{n}\right)\left(R^2S_a^2 - 2\rho_{rs}R'S_rS_a\right)
\]  
(2.5)

\[
\hat{\tau}_2 = \overline{Y}_r \left(\lambda + (1 - \lambda)\frac{P_{ur}}{P_r}\right)
\]  
(2.6)

\[
MSE(\hat{\tau}_2)_{\text{min}} = S_y^2\left(\frac{1}{r} - \frac{1}{N}\right) - \left(\frac{1}{r} - \frac{1}{n}\right)\rho_{Y\tau}^2
\]  
(2.7)

\[
\hat{\tau}_3 = \overline{Y}_r \left(\frac{P_{ur}}{P_r}\right)^\beta
\]  
(2.8)

\[
MSE(\hat{\tau}_3)_{\text{min}} = S_y^2\left(\frac{1}{r} - \frac{1}{N}\right) - \left(\frac{1}{r} - \frac{1}{n}\right)\rho_{Y\tau}^2
\]  
(2.9)

\[
\hat{\tau}_4 = \overline{Y}_r \left(\frac{P}{P_r}\right)^\beta
\]  
(2.10)

\[
MSE(\hat{\tau}_4) = S_y^2\left(\frac{1}{r} - \frac{1}{N}\right) - \left(\frac{1}{r} - \frac{1}{n}\right)\rho_{Y\tau}^2
\]  
(2.11)

\[
\hat{\tau}_5 = \frac{\overline{Y}_rP_{ur}}{\alpha P_r + (1 - \alpha)P_{ur}}
\]  
(2.12)

\[
MSE(\hat{\tau}_5)_{\text{min}} = S_y^2\left(\frac{1}{r} - \frac{1}{N}\right) - \left(\frac{1}{r} - \frac{1}{n}\right)\rho_{Y\tau}^2
\]  
(2.13)

\[
\hat{\tau}_6 = \kappa \overline{Y}_r + (1 - \kappa)\overline{Y}_r \exp\left(\frac{P - P_r}{P + P_r}\right)
\]  
(2.14)

\[
MSE(\hat{\tau}_6)_{\text{min}} = \left(\frac{1}{r} - \frac{1}{N}\right)S_y^2\left(1 - \rho_{Y\tau}^2\right)
\]  
(2.15)

\[
\hat{\tau}_7 = \overline{w}\overline{Y}_r + (1 - \overline{w})\overline{Y}_r \exp\left(\frac{p^* - P}{p^* + P}\right)
\]  
(2.16)

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

\[
MSE(\hat{\tau}_7)_{\text{min}} = \left(\frac{1}{r} - \frac{1}{N}\right)S_y^2\left(1 - \rho_{Y\tau}^2\right)
\]  
(2.17)
where, \( p_r = \frac{1}{r} \sum_{i \in r} \pi_i, p_n = \frac{1}{n} \sum_{i \in s} \pi_i, \rho_{iy} = \frac{S_{iy}}{S_{y}^{2}}, R' = \frac{\bar{y}}{P}, S_{\pi} = \frac{1}{N-1} \sum_{i=1}^{N} (\pi_i - P)^2, \)

\[
P = \frac{1}{N} \sum_{i=1}^{N} \pi_i, S_{y\pi} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})(\pi_i - P), \lambda = 1 - R' \beta_{x}, \phi = R' \beta_{x}, \kappa = 1 - 2R' \beta_{x},
\]

\[
\beta_{x} = \frac{S_{y\pi}}{S_{\pi}^{2}}
\]

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

\[
y_{i} = \begin{cases} y_{i} & \text{if } i \in \Psi \\
\frac{\bar{y}_{r}}{n-r} \left( n \left( \Lambda_1 \frac{P}{p_r} + \Lambda_2 \frac{p_r}{P} \right) \exp \left( \frac{p_r - P}{p_r + P} \right) - r \right) & \text{if } i \in \Psi^{c}
\end{cases}
\]

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

\[
\hat{t}_{8} = \bar{y}_{r} \left( \Lambda_1 \frac{P}{p_r} + \Lambda_2 \frac{p_r}{P} \right) \exp \left( \frac{p_r - P}{p_r + P} \right)
\]

\[
Bias(\hat{t}_{8}) = \bar{y} \left( \Lambda_1 \left( 1 + \frac{\theta}{8} \left( 3C_{p}^{2} - 4\rho C_{y}C_{p} \right) \right) + \Lambda_2 \left( 1 + \frac{\theta}{8} \left( 3C_{p}^{2} - 12\rho C_{y}C_{p} \right) \right) \right) - 1
\]

\[
MSE(\hat{t}_{8}) = \bar{y}^{2} \left( 1 + \Lambda_1 \Psi_{1} + \Lambda_2 \Psi_{2} - 2\Lambda_1 \Psi_{3} - 2\Lambda_2 \Psi_{4} + 2\Lambda_2 \Lambda_2 \Psi_{5} \right)
\]

where

\[
\Psi_{1} = 1 + \theta \left( C_{y}^{2} + \frac{1}{2} C_{p}^{2} - 2\rho C_{y}C_{p} \right), \Psi_{2} = 1 + \theta \left( C_{y}^{2} + 3C_{p}^{2} + 6\rho C_{y}C_{p} \right), \Psi_{3} = 1 + \theta \left( 3C_{p}^{2} - 4\rho C_{y}C_{p} \right)
\]

\[
\Psi_{4} = 1 + \theta \left( 3C_{p}^{2} - 12\rho C_{y}C_{p} \right), \Psi_{5} = 1 + \theta \left( C_{y}^{2} + 2\rho C_{y}C_{p} \right)
\]

\[
MSE(\hat{t}_{8})_{\text{min}} = \bar{y}^{2} \left( 1 - \frac{\Psi_{2} \Psi_{5}^{2} + \Psi_{1} \Psi_{4}^{2} - 2\Psi_{3} \Psi_{4} \Psi_{5}}{\Psi_{1} \Psi_{2} - \Psi_{5}^{2}} \right)
\]

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

\[
y_i = \begin{cases} y_i, & \text{if } i \in \Psi \\ \frac{y_i + \hat{\beta}_1 (P-p_r)}{\sigma_1 P_r + \sigma_2} \exp \left( \frac{P}{P + p_r} \right), & \text{if } i \in \Psi^c \end{cases}
\]

\[
y_i = \begin{cases} y_i, & \text{if } i \in \Psi \\ \frac{y_i + \hat{\beta}_2 (P^* - P)}{\sigma_1 P_r + \sigma_2} \exp \left( \frac{P^*}{P^* + P} \right), & \text{if } i \in \Psi^c \end{cases}
\]

where \( \sigma_1 \) and \( \sigma_2 \) 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, \( \hat{\beta}_1 = s_{y\pi} / s_{x\pi} \), \( s_{x\pi} = \frac{1}{r-1} \sum_{j=1}^{r-1} (y_j - \bar{y}_r)(\pi_j - p_r) \), \( s_{y\pi} = \frac{1}{r-1} \sum_{j=1}^{r-1} (y_j - \bar{y}_r)(\pi_j - p_r) \).

3.1 Remark

Note that \( \sigma_1 \neq \sigma_2 \) and \( \sigma_1 \neq 0 \).

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

\[
T_i = \frac{r}{n} \bar{y}_r + \left( 1 - \frac{r}{n} \right) \frac{y_i + \hat{\beta}_1 (P-p_r)}{\sigma_1 P_r + \sigma_2} \exp \left( \frac{P}{P + p_r} \right)
\]

\[
T_i^{(\ast)} = \frac{r}{n} \bar{y}_r + \left( 1 - \frac{r}{n} \right) \frac{y_i + \hat{\beta}_2 (P^* - P)}{\sigma_1 P_r + \sigma_2} \exp \left( \frac{P^*}{P^* + P} \right)
\]

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 \( \sigma_1 \) and \( \sigma_2 \)

| \( i \) | Estimators | Values of Constants |
|-------|------------|--------------------|
| \( i \) | \( \sigma_1 \) | \( \sigma_2 \) |
| 1     | \( T_i = \frac{r}{n} \bar{y}_r + \left( 1 - \frac{r}{n} \right) \frac{y_i + \hat{\beta}_1 (P-p_r)}{\sigma_1 P_r + \sigma_2} \exp \left( \frac{P}{P + p_r} \right) \) | \( C_X \) |
| $i$ | Estimators                                                                                                                                                                                                 | Values of Constants |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| 2   | $T_2 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{P - p_r}{P + p_r}\right) \left(\frac{P - p_r}{P + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 3   | $T_3 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{P + \beta_1 (\pi)}{P + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 4   | $T_4 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{P + S_\pi}{P + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 5   | $T_5 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{C_\pi + \beta_1 (\pi)}{C_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 6   | $T_6 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{C_\pi + \beta_2 (\pi)}{C_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 7   | $T_7 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{C_\pi + S_\pi}{C_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 8   | $T_8 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + C_\pi}{\beta_1 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 9   | $T_9 = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + S_\pi}{\beta_2 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 10  | $T_{10} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + S_\pi}{\beta_1 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 11  | $T_{11} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + S_\pi}{\beta_2 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 12  | $T_{12} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + S_\pi}{\beta_2 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 13  | $T_{13} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{\beta_1 (\pi) + S_\pi}{\beta_2 (\pi) + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 14  | $T_{14} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{S_\pi + \beta_1 (\pi)}{S_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 15  | $T_{15} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{S_\pi + \beta_1 (\pi)}{S_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
| 16  | $T_{16} = \frac{r}{n} \bar{y}_r + \left(1 - \frac{r}{n}\right) \bar{y}_r + \hat{\beta}_2 (P - p_r) \left(\frac{S_\pi + \beta_2 (\pi)}{S_\pi + p_r}\right)$ | $\omega_1$ $\omega_2$ |
4 Properties of the Estimators Suggested

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

\[
\text{MSE}(T_i) = \theta_{r,N} \left( S_r^2 + \left(1 - \frac{r}{n}\right)^2 \beta_x + \left(1 + \gamma_i\right) R \right)^2\left( S_r^2 - 2 \left(1 - \frac{r}{n}\right) \left( \beta_x + \left(1 + \gamma_i\right) R \right) S_{r^2} \right)
\]

(4.1)

where

\[
R = \frac{Y}{P}, \quad \gamma_1 = P/(P + C_x), \quad \gamma_2 = P/(P + \beta_1(x)), \quad \gamma_3 = P/(P + \beta_2(x)), \quad \gamma_4 = P/(P + S_x),
\]

\[
\gamma_5 = C_xP/(C_xP + \beta_1(x)), \quad \gamma_6 = C_xP/(C_xP + \beta_2(x)), \quad \gamma_7 = C_xP/(C_xP + S_x),
\]

\[
\gamma_8 = \beta_1(x)P/(\beta_1(x)P + C_x), \quad \gamma_9 = \beta_1(x)P/(\beta_1(x)P + \beta_2(x)), \quad \gamma_{10} = \beta_1(x)P/(\beta_1(x)P + S_x),
\]

\[
\gamma_{11} = \beta_2(x)P/(\beta_2(x)P + C_x), \quad \gamma_{12} = \beta_2(x)P/(\beta_2(x)P + \beta_1(x)), \quad \gamma_{13} = \beta_2(x)P/(\beta_2(x)P + S_x),
\]

\[
\gamma_{14} = S_xP/(S_xP + C_x), \quad \gamma_{15} = S_xP/(S_xP + \beta_1(x)), \quad \gamma_{16} = S_xP/(S_xP + \beta_2(x))
\]

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

\[
\text{MSE}(T_i) = \Delta \Sigma \Delta'
\]

(4.2)

where \( \Delta \) is a \( 1 \times 2 \) matrix, \( \Sigma \) is a \( 2 \times 2 \) variance-covariance matrix,

\[
\Delta = \begin{pmatrix}
\frac{\partial T_i}{\partial \bar{y}} & \frac{\partial T_i}{\partial X} & \frac{\partial T_i}{\partial \beta_z} \\
\frac{\partial T_i}{\partial \bar{X}} & \frac{\partial T_i}{\partial \beta_x} & \frac{\partial T_i}{\partial \beta_x} \\
\end{pmatrix}, \quad \Sigma = \begin{pmatrix}
\text{var}(\bar{y},r) & \text{cov}(\bar{y},\bar{x}) \\
\text{cov}(\bar{x},\bar{y}) & \text{var}(p_r) \\
\end{pmatrix}
\]

\[
\frac{\partial T_i}{\partial \bar{y}} = \frac{r}{n} + \left(1 - \frac{r}{n}\right) \frac{\alpha_iP + \alpha_2}{\alpha_ip_r + \alpha_2} \exp\left(\frac{P - p_r}{P + p_r}\right)
\]

(4.3)

\[
\frac{\partial T_i}{\partial \bar{y}} = \bar{y}, \quad \frac{\partial T_i}{\partial P} = P, \quad \frac{\partial T_i}{\partial \beta_z} = \beta_z = 1
\]

(4.4)

\[
\frac{\partial T_i}{\partial p_r} = -\left(1 - \frac{r}{n}\right) \left(\frac{\alpha_iP + \alpha_2}{\alpha_ip_r + \alpha_2}\right) \exp\left(\frac{P - p_r}{P + p_r}\right) \left(2P\left(\bar{y} + \hat{\beta}_{\bar{y}}(P - p_r)\right)\right) \left(\frac{P + p_r}{P + p_r}\right)
\]

\[
+ \hat{\beta}_{\bar{y}} + \alpha_i \left(\bar{y} + \hat{\beta}_{\bar{y}}(P - p_r)\right)\left(\alpha_ip_r + \alpha_2\right)^2
\]

(4.5)

\[
\frac{\partial T_i}{\partial \bar{x}} = -\left(1 - \frac{r}{n}\right) \left(\beta_x + \frac{\bar{y}}{2P} + \gamma_i\bar{y}\right)
\]

(4.6)

where \( \gamma_i = \alpha_iP / (\alpha_iP + \alpha_2) \)
So, from the definition of $\Delta$, we have

$$
\Delta = \begin{pmatrix}
1 & -\left(1 - \frac{r}{n}\right) \left(\beta_x + \frac{\bar{Y}}{2P} + \gamma_i \frac{\bar{Y}}{P}\right)
\end{pmatrix}
$$

(4.7)

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

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

$$
MSE\left(T_i^*\right) = \theta_{r,N} \left(S_i^2 + \left(1 - \frac{r}{n}\right) f^* \left(\beta_x + \frac{1}{2} \gamma_i \right) \bar{R} \right)^2 S_i^2 - 2 \left(1 - \frac{r}{n}\right) f^* \left(\beta_x + \frac{1}{2} + \gamma_i \right) \bar{R} S_i^2
$$

(4.8)

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

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

$$
MSE\left(T_i^*\right) = \Delta^* \Sigma \Delta^*
$$

(4.9)

where $\Delta$ is a $1 \times 2$ matrix, $\Sigma$ is a $2 \times 2$ variance-covariance matrix,

$$
\Delta^* = \begin{pmatrix}
\frac{\partial T_i^*}{\partial \bar{y}_r} & \frac{\partial T_i^*}{\partial \bar{x}_r} \\
\frac{\partial T_i^*}{\partial \bar{y}_r} & \frac{\partial T_i^*}{\partial \bar{x}_r}
\end{pmatrix}, \quad \Sigma = \begin{pmatrix}
\sigma_x, \sigma_x & \sigma_x, \sigma_y \\
\sigma_x, \sigma_y & \sigma_y, \sigma_y
\end{pmatrix}
$$

(4.10)

$$
\frac{\partial T_i^*}{\partial \bar{y}_r} = \frac{r}{n} + \left(1 - \frac{r}{n}\right) \left(\sigma_x P + \sigma_y^2\right) \exp\left(\frac{p^* - P}{p^* + P}\right)
$$

(4.11)

$$
\frac{\partial T_i^*}{\partial \bar{y}_r} = -\left(1 - \frac{r}{n}\right) \frac{N - n}{N - n} \exp\left(\frac{p^* - P}{p^* + P}\right) \left[2P\left(\bar{y}_r + \hat{\beta}_x (p^* - P)\right)\right] / \left(P + p^*\right)^2
$$

(4.12)

$$
\frac{\partial T_i^*}{\partial \bar{y}_r} = -\left(1 - \frac{r}{n}\right) \frac{N - n}{n} \left(\beta_x + \frac{\bar{Y}}{2P} + \gamma_i \frac{\bar{Y}}{P}\right)
$$

(4.13)

So, from the definition of $\Delta$, we have
\[
\Delta^* = 1 - \left(1 - \frac{r}{n}\right) \left(\frac{n}{N-n}\right) \left(\beta_x + \frac{\bar{Y}}{2P} + \frac{\gamma}{P}\right)
\]

(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]

\[ Y = \text{The number of villages in the circle,} \]

\[ \pi = \begin{cases} 
1, & \text{if } Y > 5 \\
0, & \text{if } Y \leq 5 
\end{cases} \]

Population II: Source [27]

\[ Y = \text{Area (in Acres) under wheat crop in the circles,} \]

\[ \pi = \begin{cases} 
1, & \text{if } Y > 5 \\
0, & \text{if } Y \leq 5 
\end{cases} \]

Table 2. Descriptive statistics of the populations

| Population | \( N = 89, n = 23, P = 0.124, \bar{Y} = 3.3596, \rho_{xY} = 0.766, C_Y = 0.6008, C_\rho = 2.6779, \beta_2 = 6.162 \) |
|------------|---------------------------------------------------------------------------------------------------------------|
| Population | \( N = 89, n = 23, P = 0.124, \bar{Y} = 1102, \rho_{xY} = 0.624, C_Y = 0.65, C_\rho = 2.6779, \beta_2 = 6.162 \)  |

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 | \( \hat{\tau}_{j}, j = 2, 3, \ldots, 7 \) | 0.1536047 | 147.1544 |
| \( \hat{\tau}_1 \) | 1.458 | 5.50317 | \( \hat{\tau}_8 \) | 0.3224198 | 70.10618 |
| 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 |
| Estimators | MSE   | PRE   | Estimators | MSE   | PRE   |
|------------|-------|-------|------------|-------|-------|
| \(\hat{\tau}_j\) | 21323.15 | 170.0692 | \(\hat{\tau}_i\) | 19541.98 | 145.5365 |
| \(\hat{\tau}_j\) | 23434.9 | 133.3795 | \(\hat{\tau}_i\) | 19081.74 | 149.0467 |
| \(\hat{\tau}_j\) | 31270.33 | 121.3604 | \(\hat{\tau}_i\) | 18121.18 | 156.9473 |
| \(\hat{\tau}_j\) | 29598.11 | 90.95109 | \(\hat{\tau}_i\) | 18267.79 | 155.6877 |
| \(\hat{\tau}_j\) | 47310.5 | 96.08959 | \(\hat{\tau}_i\) | 17434.78 | 163.1263 |
| \(\hat{\tau}_j\) | 21377.7 | 60.115 | \(\hat{\tau}_i\) | 19527.95 | 145.641 |
| \(\hat{\tau}_j\) | 31746.83 | 133.0391 | \(\hat{\tau}_i\) | 18083.31 | 157.276 |
| \(\hat{\tau}_j\) | 20487.58 | 89.58598 | \(\hat{\tau}_i\) | 19776.69 | 143.8092 |
| \(\hat{\tau}_j\) | 21135.36 | 138.8193 | \(\hat{\tau}_i\) | 19591.37 | 145.1696 |

Table 4. MSE and PRE of proposed and other estimators using population II
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 \( T^*_i, i = 1, 2, \ldots, 16 \) and \( T_i, T_2, T_3, T_4, T_5, T_6, T_9, T_{10}, T_{12}, T_{13}, T_{15}, T_{16} \) 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 \( T^*_i, i = 1, 2, \ldots, 16 \) and \( T_i, i = 1, 2, \ldots, 16 \) 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.
Audu et al.; ARJOM, 17(5): 1-13, 2021; Article no.ARJOM.63925

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;12(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 sampling. 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.

© 2021 Audu et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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