AN EFFICIENT CLASS OF CHAIN DUAL TO RATIO TYPE IMPUTATION METHOD FOR ESTIMATING FINITE POPULATION MEAN UNDER DOUBLE SAMPLING IN CASE OF MISSING DATA

Deepjan Gohain and B. K. Singh
Department of Mathematics, North Eastern Regional Institute of science and Technology
Nirjuli, Itanagar-791109, India
Email: deepjangohain@gmail.com, bksinghstat@gmail.com

Abstract

In this present study we have suggested an efficient class of chain dual to ratio type imputation method and corresponding point estimator of finite population mean with imputation of missing data under double sampling. The Bias and Mean Square Error of the proposed estimator have been derived in terms of parameters. Theoretical and empirical studies are carried out to verify the superiorities of the proposed estimator over existing estimators.

Key words: Bias, MSE, Exponential estimators, Imputation, Simple random sampling, Double sampling.

1. INTRODUCTION

Generally population parameters are estimated by using a suitable corresponding statistics, known as estimator. Thus for estimating population mean, sample mean is the most appropriate estimator. Although it is unbiased, it has a large amount of variation. Therefore we seek such an estimator which may be biased but has smaller Mean Square Error (*MSE*) as compared to sample mean. This is achieved through the use of an auxiliary variable (X) that has highly positive or negative correlation with the study variable (Y). Cochran (1940) used the information of auxiliary variable that has highly positive correlation with the study variable to estimate the finite population mean of the study variable known as Classical ratio estimator. Further these techniques are extended by Srivenkataraman (1980). Bhal and Tuteja (1991) proposed a dual to ratio type estimator and exponential ratio type estimator of finite population mean under simple random sampling. The above estimators are used to estimate finite population mean under the assumption that the population mean $\bar{X}$ of the auxiliary variable is available. When $\bar{X}$ unknown then two-phase or double sampling method is used to overcome it, suggested by Neyman (1938). Kumar and Bhal (2006), Singh and Vishwakarma (2007), Kalita and Singh (2013), Sukla (2019) and many others are proposed ratio type estimator of finite population mean under two phrase sampling. Sometimes it is seen that auxiliary variable X is highly correlated with another variable Z is relatively less correlated to the study variable in comparison to the auxiliary variable X. In such situation the information of both X and Z are used to estimate the finite population mean of the study variable. Chand (1975) introduced a technique of chaining the information of two auxiliary variables with the study variable in two phrase sampling. Further similar work were extended by Singh and Choudhary (2012), Kalita et al. (2013), Singh and Majhi (2014), Singh and Ahmed (2015) and many others. The above methods are applicable in use when all the information of the variables X, Y& Z are available in the sample. In practical situations some unit information are not available due to non-response of the respondents or any other causes in the sample(s). Imputations technique is one of the most common scientific techniques to replace such missing information by using the information available in the sample. Mean method of imputation, ratio method of imputation, compromised method of imputation etc. are well known imputation techniques to estimate the finite population mean of the study variable. Other imputation methods developed by many scholars viz., Upadhyaya and Singh (1999), Ahmed *et al.* (2006), Shukla and Thakur(2008), Shukla et al. (2009), Thakur et al. (2013), Kadilar and Cingi (2008), Singh *et al.* (2014), Singh *et al.* (2015), Diana and Peri (2010), Gogoi and Singh (2017,2018, 2019), Nath and Singh (2018), Nath *et al*. (2020),

2. NOTATIONS

Let $\Omega = \{1, 2, 3, \ldots, N\}$ be a finite population of size N and Y is the study variable and X, Z are the auxiliary variables. $\bar{Y}$, $\bar{X}$ and $\bar{Z}$ are the population mean of the variable Y, X and Z respectively, and letting $\bar{X}$ be unknown.

Consider a first phase sample S_1 of size n_1 drawn from the population Ω by using SRSWOR method and a second sample S of size $n(n < n_1)$ drawn from S_1.

Let the second sample S contains $r(r < n)$ responding units forming a sub space A and $(n-r)$ non-responding units with sub space A^C, such that $S = A \cup A^C$. For every $i \in A$, y_i is observed available. For $i \in A^C$, y_i values are missing and imputed values are computed. The i^{th} value x_i and z_i of auxiliary variables are used as a source of imputation for missing data when $i \in A^C$, assuming that in S_1 the data $\{(x_i, z_i); i \in S\}$ *and* $\{(x_i, z_i);\ i \in S_1\}$ are known.

$$\bar{y}_n = \frac{1}{n}\sum_{i\in S} Y_i \qquad \bar{x}_n = \frac{1}{n}\sum_{i\in S} X_i \qquad \bar{x}_r = \frac{1}{r}\sum_{i\in A} X_i$$

$$\bar{y}_r = \frac{1}{r}\sum_{i\in A} Y_i \qquad \bar{z}_r = \frac{1}{r}\sum_{i\in A} Z_i \qquad \bar{x}_1 = \frac{1}{n_1}\sum_{i\in S_1} x_i$$

$$\bar{y}_1 = \frac{1}{n_1}\sum_{i\in S_1} y_i \qquad \bar{z}_1 = \frac{1}{n_1}\sum_{i\in S_1} z_i$$

$$\bar{X} = \frac{1}{N}\sum_{i\in \Omega} X_i \qquad \bar{Y} = \frac{1}{N}\sum_{i\in \Omega} Y_i \qquad \bar{Z} = \frac{1}{N}\sum_{i\in \Omega} Z_i$$

$$S_X^2 = \frac{1}{N-1}\sum_{i\in\Omega}\left(X_i - \bar{X}\right)^2 \qquad S_Y^2 = \frac{1}{N-1}\sum_{i\in\Omega}\left(Y_i - \bar{Y}\right)^2$$

$$S_Z^2 = \frac{1}{N-1}\sum_{i\in\Omega}\left(Z_i - \bar{Z}\right)^2 \qquad S_{XY} = \frac{1}{N-1}\sum_{i\in\Omega}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)$$

$$S_{YZ} = \frac{1}{N-1}\sum_{i\in\Omega}\left(Y_i - \bar{Y}\right)\left(Z_i - \bar{Z}\right) \qquad S_{ZX} = \frac{1}{N-1}\sum_{i\in\Omega}\left(Z_i - \bar{Z}\right)\left(X_i - \bar{X}\right)$$

$$\rho_{YX} = \frac{S_{XY}}{S_X S_Y} \qquad \rho_{YZ} = \frac{S_{YZ}}{S_Y S_Z} \qquad \rho_{ZX} = \frac{S_{ZX}}{S_Z S_X}$$

$$C_X = \frac{S_X}{\bar{X}} \qquad C_Y = \frac{S_Y}{\bar{Y}} \qquad C_Z = \frac{S_Z}{\bar{Z}}$$

$$\theta_{a,b} = \frac{1}{a} - \frac{1}{b}$$

3. PROPOSED METHOD OF IMPUTATION AND CORRESPONDING POINT ESTIMATOR:

Kumar and Bhal (2006) suggested a dual to ratio estimator of population mean of the study variable $\bar{Y}$ in double sampling as

$$\bar{y}_{d\,\mathrm{Re}}^{dc} = \bar{y}_n \left[\frac{n_1\bar{x}_1 - n\bar{x}_n}{\left(n_1 - n\right)\bar{x}_1} \right]$$

Motivated by Kumar and Bhal (2006), we have proposed the following chain dual to ratio type imputation method in double sampling

$$y_{\cdot i} = y_i \qquad if \qquad i \in A$$

$$= \frac{\bar{y}_r}{n-r}\left[n\psi^{\beta} - r \right] \qquad if \qquad i \in A^C$$

$$where \;\; \psi = \left[\frac{n_1\bar{x}_1\left(\frac{\bar{Z}+\alpha}{\bar{z}+\alpha}\right) - n\bar{x}_r}{(n_1-n)\bar{x}_1\left(\frac{\bar{Z}+\alpha}{\bar{z}+\alpha}\right)} \right]$$

Under the above imputation method, the point estimator of population mean $\bar{Y}$ is

$$\bar{y}_{SG}^{dc} = \frac{1}{n}\left\{ \sum_{i\in A} y_i\psi^{\beta} + \sum_{i\in A}\bar{y}_r\psi^{\beta} \right\}$$

$$= \frac{1}{n}\left[r\bar{y}_r\psi^{\beta} + (n-r)\,\bar{y}_r\psi^{\beta} \right]$$

$$= \bar{y}_r\left[\frac{n_1\bar{x}_1\left(\frac{\bar{Z}+\alpha}{\bar{z}+\alpha}\right) - n\bar{x}_r}{(n_1-n)\bar{x}_1\left(\frac{\bar{Z}+\alpha}{\bar{z}+\alpha}\right)} \right]^{\beta}$$

Where α and β are constants to be determined such that the estimator t_{SG}^{dc} has minimum MSE

Remarks :

a. If $\beta = 0$,the proposed estimator reduces to the estimator

   $t_{SG}^{dc} = \bar{y}_r$ which is the point estimator of mean method of imputation.

b. If $\beta = 1$ *and* $\alpha = \alpha_1$, the proposed estimator reduces to the estimator

$$\bar{y}_{SG}^{dc} = \bar{y}_r \left[\frac{n_1 \bar{x}_1 \left(\dfrac{\bar{Z} + \alpha_1}{\bar{z} + \alpha_1} \right) - n\bar{x}_r}{(n_1 - n)\bar{x}_1 \left(\dfrac{\bar{Z} + \alpha_1}{\bar{z} + \alpha_1} \right)} \right] = t_1 (say)$$

c. If $\alpha = 0$ *and* $\beta = \beta_1$, the proposed estimator reduces to the estimator

$$\bar{y}_{SG}^{dc} = \bar{y}_r \left[\frac{n_1 \bar{x}_1 \dfrac{\bar{Z}}{\bar{z}_1} - n\bar{x}_r}{(n_1 - n)\bar{x}_1 \dfrac{\bar{Z}}{\bar{z}_1}} \right]^{\beta_1} = t_2 (say)$$

d. If $\alpha = 0$ *and* $\beta = 1$, the proposed estimator reduces to the estimator

$$\bar{y}_{SG}^{dc} = \bar{y}_r \left[\frac{n_1 \bar{x}_1 \dfrac{\bar{Z}}{\bar{z}_1} - n\bar{x}_r}{(n_1 - n)\bar{x}_1 \dfrac{\bar{Z}}{\bar{z}_1}} \right] = t_3 (say)$$

4. PROPERTIES OF THE PROPOSED ESTIMATOR

To derive the properties of the proposed estimator we have consider the following transformations

$\bar{y}_r = \bar{Y}(1 + e_0)$ $\quad$ $\bar{x}_r = \bar{X}(1 + e_1)$ $\quad$ $\bar{x}_1 = \bar{X}(1 + e_2)$ $\quad$ $\bar{z}_1 = \bar{Z}(1 + e_3)$

such that $e_i \in (-1, 1) \quad \forall i = 0, 1, 2, 3$

$$E(e_0)=E(e_1)=E(e_2)=E(e_3)=0$$

$$E(e_0^2)=\theta_{r,N}C_Y^2 \qquad E(e_1^2)=\theta_{r,N}C_X^2 \qquad E(e_2^2)=E(\delta_1\delta_2)=\theta_{n_1,N}C_X^2$$

$$E(e_3^2)=\theta_{n_1,N}C_Z^2 \qquad E(e_0e_1)=\theta_{r,N}\rho_{XY}C_YC_X \qquad E(e_0e_2)=\theta_{n_1,N}\rho_{XY}C_YC_X$$

$$E(e_0e_3)=\theta_{n_1,N}\rho_{YZ}C_YC_Z \qquad E(e_1e_3)=E(e_2e_3)=\theta_{n_1,N}\rho_{ZX}C_ZC_X$$

Expressing t_{SG}^{dc} in terms of e's retaining the terms up to first order of approximation, we have

$$\bar{y}_{SG}^{dc}-\bar{Y}=\bar{Y}\left\{e_0+g\beta P+g^2\left(\frac{\beta^2-\beta}{2}\right)Q\right\} \qquad (i)$$

$$\text{where } P=e_1-e_2+\phi e_3+\phi e_1e_3-e_1e_2-\phi e_2e_3+e_2^2+e_0e_1-e_0e_2-\phi e_0e_3$$

$$Q=e_1^2+e_2^2+\phi^2e_3^2-2e_1e_2+2\phi e_1e_3-2\phi e_2e_3$$

$$g=\frac{n}{n_1-n} \qquad \phi=\left(1+\frac{\alpha}{\bar{Z}}\right)^{-1}$$

Theorem 4.1Bias and *MSE* of the proposed estimator up to first order approximation is

$$Bias(\bar{y}_{SG}^{dc})=\bar{Y}\left\{\left(\frac{\beta^2-\beta}{2}\right)g^2\left(\theta_{r,n_1}C_X^2+\theta_{n,N}\phi^2C_Z^2\right)-g\beta\left(\theta_{r,n_1}\rho_{XY}C_YC_X+\theta_{n,N}\phi\rho_{YZ}C_YC_Z\right)\right\} \qquad (ii)$$

$$MSE(\bar{y}_{GS}^{dc})=\bar{Y}^2\left\{\theta_{r,N}C_Y^2+g^2\beta^2\left(\theta_{r,n_1}C_X^2+\theta_{n_1,N}\phi^2C_Z^2\right)-2g\beta\left(\theta_{r,n_1}\rho_{XY}C_YC_X+\theta_{n_1,N}\phi\rho_{YZ}C_YC_Z\right)\right\} \qquad (iii)$$

Proof: Taking expectations on both sides of equation (i) and putting the corresponding expected values, we have

$$Bias(\bar{y}_{SG}^{dc})=\bar{Y}\left\{-g\beta E(P)+g^2\left(\frac{\beta^2-\beta}{2}\right)E(Q)\right\}$$

$$=\bar{Y}\left\{\left(\frac{\beta^2-\beta}{2}\right)g^2\left(\theta_{r,n_1}C_X^2+\theta_{n_1,N}\phi^2C_Z^2\right)-g\beta\left(\theta_{r,n_1}\rho_{XY}C_YC_X+\theta_{n_1,N}\phi\rho_{YZ}C_YC_Z\right)\right\}$$

Squaring both the sides of equation (i) and taking expectation retaining the terms up to first order approximation, we have

$$MSE(\bar{y}_{SG}^{dc}) = \bar{Y}^2 \left\{ E\left(e_0^2\right) + g^2\beta^2 E\left(P_1\right) + 2g\beta E\left(Q_1\right) \right\} \quad (iv)$$

Where

$$P_1 = e_2^2 + e_1^2 + \phi^2 e_3^2 - 2e_1e_2 - 2\phi e_2e_3 + 2\phi e_1e_3$$

$$Q_1 = e_0e_2 - e_0e_1 - \phi e_0e_3$$

Putting the corresponding expected values in equation (iv), we have

$$MSE(t_{SG}^{dc}) = \bar{Y}^2 \left\{ \theta_{r,N} C_Y^2 + g^2\beta^2 \left(\theta_{r,n_1} C_X^2 + \theta_{n_1,N} \phi^2 C_Z^2 \right) - 2g\beta \left(\theta_{r,n_1} \rho_{XY} C_X C_Y + \theta_{n_1,N} \phi \rho_{YZ} C_Y C_Z \right) \right\}$$

Theorem 4.2 Minimum MSE of the proposed estimator $\bar{y}_{GS}^{dc}$ is

$$MSE\left(\bar{y}_{SG}^{dc}\right)_{\min} = \bar{Y}^2 \left\{ \theta_{r,N} C_Y^2 - \left(\theta_{r,n_1} \rho_{XY}^2 + \theta_{n_1,N} \rho_{YZ}^2 \right) C_Y^2 \right\}$$

Proof: Differentiating equation (iii) with respect to ϕ and β and equating to zero, we have

$$\beta_{opt} = \frac{\rho_{XY}}{g} \frac{C_Y}{C_X} \quad and \quad \phi_{opt} = \frac{\rho_{YZ}}{\rho_{XY}} \frac{C_X}{C_Z}$$

Again

$$\left[\frac{\delta^2}{\delta\beta^2} MSE(\bar{y}_{SG}^{dc}) \right]_{\substack{\beta=\beta_{opt} \\ \phi=\phi_{opt}}} = 2\bar{Y}^2 g^2 \left(\theta_{r,n_1} C_X^2 + \theta_{n_1,N} \phi_{opt}^2 C_Z^2 \right) > 0$$

$$\left[\frac{\delta^2}{\delta\phi^2} MSE(\bar{y}_{SG}^{dc}) \right]_{\substack{\beta=\beta_{opt} \\ \phi=\phi_{opt}}} = 2\bar{Y}^2 g^2 \beta_{opt}^2 \left(\theta_{n_1,N} C_Z^2 \right) > 0$$

$$\left[\frac{\delta^2}{\delta\beta^2} MSE(\bar{y}_{SG}^{dc}) \frac{\delta^2}{\delta\phi^2} MSE(\bar{y}_{SG}^{dc}) - \left\{ \frac{\delta^2}{\delta\beta\delta\phi} MSE(\bar{y}_{SG}^{dc}) \right\}^2 \right]_{\substack{\beta=\beta_{opt} \\ \phi=\phi_{opt}}} = A\rho_{YX}^2 + B\rho_{YZ}^2 - B\rho_{YZ}^2 = A\rho_{YX}^2 > 0$$

Where

$$A=4\bar{Y}^4g^2\theta_{r,n_1}\theta_{n_1,N}C_Y^2C_Z^2>0 \ \ and \ \ B=4\bar{Y}^4g^2\theta_{n_1,N}^2C_Y^2C_Z^2>0$$

Thus putting the value of $\beta=\beta_{opt}$ and $\phi=\phi_{opt}$ in equation (iii) we have minimum Mean Square Error of $\bar{y}_{SG}^{dc}$

$$\begin{aligned} MSE(\bar{y}_{SG}^{dc})_{\min} &= \bar{Y}^2\left\{\theta_{r,N}C_Y^2+g^2\beta_{opt}^2\left(\theta_{r,n_1}C_X^2+\theta_{n_1,N}\phi_{opt}^2C_Z^2\right)-2g\beta_{opt}\left(\theta_{r,n_1}\rho_{XY}C_YC_X+\theta_{n_1,N}\phi_{opt}\rho_{YZ}C_YC_Z\right)\right\} \\ &= \bar{Y}^2\left\{\theta_{r,N}C_Y^2-\left(\theta_{r,n_1}\rho_{XY}^2+\theta_{n_1,N}\rho_{YZ}^2\right)C_Y^2\right\} \end{aligned}$$

Theorem 4.3 Minimum Bias and Mean Square Error of the estimator t_1 is

$$Bias(t_1)_{\min}=-g\bar{Y}\left\{\theta_{r,n_1}\rho_{YX}C_YC_X+\hat{\phi}_1\theta_{n_1,N}\rho_{YZ}C_YC_Z\right\}$$

$$MSE(t_1)_{\min}=\bar{Y}^2\left\{\theta_{r,N}C_Y^2-\theta_{n_1,N}\rho_{YZ}^2C_Y^2+\theta_{r,n_1}\left(g^2C_X^2-2g\rho_{YX}C_YC_X\right)\right\}$$

$$\hat{\phi}_1=\left(1+\frac{\hat{\alpha}_1}{\bar{Z}}\right)^{-1}=\frac{\rho_{YZ}C_Y}{gC_Z}$$

Theorem 4.4 Minimum Bias and Mean Square Error of the estimator t_2 is

$$Bias(t_2)_{\min}=-\bar{Y}\left\{g^2\left(\frac{\hat{\beta}_1^2-\hat{\beta}_1}{2}\right)\left(\theta_{r,n_1}C_X^2+\theta_{n_1,N}C_Z^2\right)-g\hat{\beta}_1\left(\theta_{r,n_1}\rho_{YX}C_YC_X+\theta_{n_1,N}\rho_{YZ}C_YC_Z\right)\right\}$$

$$MSE(t_2)_{\min}=\bar{Y}^2\left\{\theta_{r,N}C_Y^2-\frac{\left(\theta_{r,n_1}\rho_{XY}C_XC_Y+\theta_{n_1,N}\rho_{YZ}C_YC_Z\right)^2}{\theta_{r,n_1}C_X^2+\theta_{n_1,N}C_Z^2}\right\}$$

$$\hat{\beta}_1=\frac{\theta_{r,n_1}\rho_{XY}C_XC_Y+\theta_{n_1,N}\rho_{YZ}C_YC_Z}{g\left(\theta_{r,n_1}C_X^2+\theta_{n_1,N}C_Z^2\right)}$$

Theorem 4.5 Minimum Bias and Mean Square Error of the estimator t_3 is

$$Bias\left(t_3\right) = -g\bar{Y}\left(\theta_{r,n_1}\rho_{XY}C_X C_Y + \theta_{n_1,N}\rho_{YZ}C_Y C_Z\right)$$

$$MSE(\mathrm{t}_3) = \bar{Y}^2\left\{\theta_{\mathrm{r},N}C_Y^2 - \theta_{\mathrm{r,n_1}}\rho_{YX}^2 C_Y^2 - \theta_{\mathrm{n_1,N}}\rho_{YZ}^2 C_Y^2 + \theta_{\mathrm{r,n_1}}\left(gC_X - \rho_{YX}C_Y\right)^2 + \theta_{\mathrm{n_1},N}\left(gC_Z - \rho_{YZ}C_Y\right)^2\right\}$$

5. SOME EXISTING IMPUTATION METHODS

Here we consider some existing imputation methods for estimating the population mean of the study variable $\bar{Y}$ where the population mean of the auxiliary variable $\bar{X}$ is not used.

5.1 Mean method of imputation

Under this imputation method

$$\begin{aligned} y_{\cdot i} &= y_i & &\text{if} \quad i \in A \\ &= \bar{y}_r & &\text{if} \quad i \in A^C \end{aligned}$$

The corresponding point estimator is

$$\bar{y}_m = \bar{y}_r$$

Lemma 1. The Bias and Variance of the point estimator $\bar{y}_r$ is

$$Bias\left(\bar{y}_r\right) = 0$$
$$Var\left(\bar{y}_r\right) = \bar{Y}^2\theta_{r,N}C_Y^2$$

5.2 Ratio method of imputation [Lee and Sarndal (1994)]

Under this imputation method

$$y_{0i} = y_i \qquad if \qquad i \in A$$
$$= \hat{b}x_i \qquad if \qquad i \in A^C$$

Where

$$\hat{b} = \frac{\sum_{i \in A} y_i}{\sum_{i \in A} y_i}$$

The corresponding point estimator is

$$\overline{y}_R = \overline{y}_r \frac{\overline{x}_n}{\overline{x}_r}$$

Lemma 2. The Bias and Mean Square Error of the point estimator $\overline{y}_R$ is

$$Bias\left(\overline{y}_R\right) = \overline{Y}\theta_{r,N}\left(C_X^2 - \rho_{YX}C_YC_X\right)$$
$$MSE\left(\overline{y}_R\right) = \overline{Y}^2\left[\theta_{r,N}C_Y^2 + \theta_{r,\mathrm{n}}\left(C_X^2 - 2\rho_{YX}C_YC_X\right)\right]$$

5.3 Compromised method of imputation [Singh and Horn (2000)]

Under this imputation method

$$y_{0i} = \beta\frac{n}{r}y_i + \left(1-\beta\right)\hat{b}x_i \qquad if \qquad i \in A$$
$$= \left(1-\beta\right)\hat{b}x_i \qquad if \qquad i \in A^C$$

The corresponding point estimator is

$$\overline{y}_{comp} = \beta\overline{y}_r + \left(1-\beta\right)\overline{y}_r\frac{\overline{x}_n}{\overline{x}_r}$$

Where β is a suitable chosen constant such that the Mean Square Error of the point estimator $\overline{y}_{comp}$ has minimum.

Lemma 3. The Bias and Mean Square Error of the point estimator $\bar{y}_{comp}$ is

$$Bias\left(\bar{y}_{comp}\right)_{\min} = \bar{Y}\theta_{r,n}\left(1-\rho_{XY}\frac{C_Y}{C_X}\right)\rho_{XY}C_X C_Y$$

$$MSE\left(\bar{y}_{comp}\right)_{\min} = \bar{Y}^2\left(\theta_{r,\mathrm{N}}C_Y^2 - \theta_{r,n}\rho_{YX}^2 C_Y^2\right)$$

5.4 Dual to ratio method of imputation [Nath and Singh(2018)]

Under this imputation methods

$$y_{1i} = y_i \qquad if \qquad i \in A$$
$$= \frac{1}{n-r}\left[n\bar{y}_r\left\{\frac{n_1\bar{x}_1 - n\bar{x}_n}{(n_1-n)\bar{x}_1}\right\}^{\gamma_1} - r\bar{y}_r\right] \qquad if \qquad i \in A^c$$

$$y_{2i} = y_i \qquad if \qquad i \in A$$
$$= \frac{1}{n-r}\left[n\bar{y}_r\left\{\frac{n_1\bar{x}_n - n\bar{x}_r}{(n_1-n)\bar{x}_n}\right\}^{\gamma_1} - r\bar{y}_r\right] \qquad if \qquad i \in A^c$$

$$y_{3i} = y_i \qquad if \qquad i \in A$$
$$= \frac{1}{n-r}\left[n\bar{y}_r\left\{\frac{n_1\bar{x}_1 - n\bar{x}_r}{(n_1-n)\bar{x}_1}\right\}^{\gamma_3} - r\bar{y}_r\right] \qquad if \qquad i \in A^c$$

Where γ_1, γ_2 and γ_3 are suitably chosen constants such that the Mean Square Error of the resulting point estimator is minimum.

The corresponding point estimators are

$$\overline{y}_{iedr1}^{d}=\overline{y}_{r}\left[\frac{n_{1}\overline{x}_{1}-n\overline{x}_{n}}{\left(n_{1}-n\right)\overline{x}_{1}}\right]^{\gamma_{1}}$$

$$\overline{y}_{iedr2}^{d}=\overline{y}_{r}\left\{\frac{n_{1}\overline{x}_{n}-n\overline{x}_{r}}{\left(n_{1}-n\right)\overline{x}_{n}}\right\}^{\gamma_{2}}$$

$$\overline{y}_{iedr3}^{dc}=\overline{y}_{r}\left\{\frac{n_{1}\overline{x}_{1}-n\overline{x}_{r}}{\left(n_{1}-n\right)\overline{x}_{1}}\right\}^{\gamma_{3}}$$

Lemma 4. The minimum Bias and Mean Square Error of the point estimators $\overline{y}_{IRP1}^{dc},\overline{y}_{IRP2}^{dc}$ and $\overline{y}_{IRP3}^{dc}$ are

$$Bias\left(\overline{y}_{iedr1}^{d}\right)_{\min}=-\overline{Y}\left\{\theta_{2n,2n_{1}}\left(g\rho_{YX}C_{X}C_{Y}+\rho_{YX}^{2}C_{Y}^{2}\right)\right\}$$

$$Bias\left(\overline{y}_{iedr2}^{d}\right)_{\min}=-\overline{Y}\left\{\theta_{2r,2n}\left(g\rho_{YX}C_{X}C_{Y}+\rho_{YX}^{2}C_{Y}^{2}\right)\right\}$$

$$Bias\left(\overline{y}_{iedr3}^{d}\right)_{\min}=-\overline{Y}\left\{\theta_{2r,2n_{1}}\left(g\rho_{YX}C_{X}C_{Y}+\rho_{YX}^{2}C_{Y}^{2}\right)\right\}$$

$$MSE\left(\overline{y}_{iedr1}^{d}\right)_{\min}=\overline{Y}^{2}\left\{\theta_{r,\mathrm{N}}C_{Y}^{2}-\theta_{n,n_{1}}\rho_{YX}^{2}C_{Y}^{2}\right\}$$

$$MSE\left(\overline{y}_{iedr2}^{d}\right)_{\min}=\overline{Y}^{2}\left\{\theta_{r,\mathrm{N}}C_{Y}^{2}-\theta_{r,n}\rho_{YX}^{2}C_{Y}^{2}\right\}$$

$$MSE\left(\overline{y}_{iedr3}^{d}\right)_{\min}=\overline{Y}^{2}\left\{\theta_{r,\mathrm{N}}C_{Y}^{2}-\theta_{r,n_{1}}\rho_{YX}^{2}C_{Y}^{2}\right\}$$

6. COMPARISON WITH EXISTING ESTIMATORS.

In this section we compare the minimum Mean Square Error of the proposed estimator with existing estimators considered in this study.

6.1 Comparison with Mean Method of Imputation

$$V\left(\bar{y}_m\right)-MSE\left(\bar{y}_{SG}^{dc}\right)=\bar{Y}^2\left(\theta_{r,n_1}\rho_{YX}^2+\theta_{n_1,N}\rho_{YZ}^2\right)C_Y^2>0$$

6.2 Comparison with Ratio Method of Imputation

$$MSE\left(\bar{y}_R\right)-MSE\left(\bar{y}_{SG}^{dc}\right)=\bar{Y}^2\left\{\theta_{r,n}\left(C_X-\rho_{YX}\right)^2+\theta_{\mathrm{r},n_1}\rho_{XY}^2+\theta_{n_1,N}\rho_{YZ}^2\right\}C_Y^2>0$$

6.3 Comparison with Compromised Method of Imputation

$$MSE\left(\bar{y}_{com}\right)_{\min}-MSE\left(\bar{y}_{SG}^{dc}\right)_{\min}=\bar{Y}^2\left(\theta_{\mathrm{r},n_1}\rho_{XY}^2+\theta_{n_1,\mathrm{N}}\rho_{YZ}^2\right)C_Y^2>0$$

6.4 Comparison with dual to Ratio Method of Imputation

$$MSE\left(\bar{y}_{iedr1}^{d}\right)_{\min}-MSE\left(\bar{y}_{SG}^{dc}\right)_{\min}=\bar{Y}^2\left(\theta_{r,n}\rho_{YX}^2+\theta_{n_1,\mathrm{N}}\rho_{YZ}^2\right)C_Y^2>0$$

$$MSE\left(\bar{y}_{iedr2}^{d}\right)_{\min}-MSE\left(\bar{y}_{SG}^{dc}\right)_{\min}=\bar{Y}^2\left(\theta_{n,n_1}\rho_{YX}^2+\theta_{n_1,\mathrm{N}}\rho_{YZ}^2\right)C_Y^2>0$$

$$MSE\left(\bar{y}_{iedr3}^{d}\right)_{\min}-MSE\left(\bar{y}_{SG}^{dc}\right)_{\min}=\bar{Y}^2\left(\theta_{n_1,\mathrm{N}}\rho_{YZ}^2C_Y^2\right)>0$$

We have found that the Mean Square Error of the proposed estimator has less Mean Square Error of the existing estimators considered in this study. Therefore it is concluded that the proposed estimator is always better than the estimators $\bar{y}_m, \bar{y}_R, \bar{y}_{comp}, \bar{y}_{iedr1}^{d}, \bar{y}_{iedr2}^{d}$ *and* $\bar{y}_{iedr3}^{d}$

7. EMPIRICAL STUDY

To examine the performances among the proposed estimators and existing estimators of the finite population mean considered in this research work, we have considered the following two populations. The percent relative efficiencies of the proposed imputation methods with respect to the ratio and compromised methods of imputation are given as

$$PRE_I = \frac{MSE\left(\bar{y}_R\right)}{MSE(.)} \times 100 \quad \text{and} \quad PRE_{II} = \frac{MSE\bar{y}\left(\bar{y}_{comp}\right)}{MSE(.)} \times 100$$

Population P_1 **:** This data has been taken from Sukhatme and Chand (1977) which has been reproduced in Singh *et al.* (1940). The particulars of the data are as under:

Y: Apple trees of bearing age in 1964

X: Bushels of apples harvesting in 1964

Z: Bushels of apples harvesting in 1959

$N = 200$

$\bar{Y} = 1031.82$ $\bar{X} = 2934.58$ $\bar{Z} = 3651.49$

$\rho_{YX} = 0.93$ $\rho_{YZ} = 0.77$ $\rho_{XZ} = 0.84$

$C_Y^2 = 2.55280$ $C_X^2 = 4.02504$ $C_Z^2 = 2.09379$

The combination of r, n and n_1 are respectively taken as 20,25and 60.

Population P_2 **:** This data was artificially generated for three variables Y, X and Z by Shukla and Thakur (2008). Considering Y as study variable and X and Z respectively the main and additional auxiliary variables, we get the following population values:

$N = 200$

$\bar{Y} = 42.485$ $\bar{X} = 18.515$ $\bar{Z} = 20.52$

$\rho_{YX} = 0.8734$ $\rho_{YZ} = 0.8667$ $\rho_{XZ} = 0.9943$

$C_Y^2 = 0.10804$ $C_X^2 = 0.1410$ $C_Z^2 = 0.10864$

The combination of r, n and n_1 are respectively taken as 22,30and 80.

Table 1: *MSE* of the existing estimators

Estimator	*MSE* for Population I	*MSE* for Population II
$\bar{y}_r$	122303.02201	7.88928
$\bar{y}_R$	101679.01403	6.25712
$\bar{y}_{comp}$	98796.38118	6.08609
$\bar{y}_{iedr1}^d$	67454.97696	4.78988
$\bar{y}_{iedr2}^d$	98796.38118	6.08609
$\bar{y}_{iedr3}^d$	43947.63093	2.98675

Table 2: *MSE* of the Proposed Estimators

Estimator	*MSE* for Population I	*MSE* for Population II
t_1	32408.46074	2.11525
t_2	25412.78682	1.90063
t_3	32889.04043	2.21801
$\bar{y}_{SG}^{dc}$	**25147.83918**	**1.88817**

Table 3: *PRE* of the Proposed Estimators

Estimator	*Population I*		*Population II*	
	PRE_I	PRE_{II}	PRE_I	PRE_{II}
t_1	313.7422	304.8475	295.8099	287.7244
t_2	400.1097	388.7664	329.2129	320.2144
t_3	309.1577	300.3930	282.2105	274.3942
$\overline{y}_{SG}^{dc}$	**404.3251**	**392.8623**	**331.3854**	**322.3274**

8. CONCLUSION

The mean square error of the proposed estimator is derived in terms of population parameters. *MSE* of the proposed estimator $\overline{y}_{GS}^{dc}$ is compared with the *MSE* of existing estimators $\overline{y}_m$, $\overline{y}_R$, $\overline{y}_{comp}$, $\overline{y}_{iedr1}^{d}$, $\overline{y}_{iedr2}^{dc}$ and $\overline{y}_{iedr3}^{dc}$ .It has been found theoretically that the proposed estimator has lesser *MSE* than the estimators considered for comparisons. Considering two population data sets the *MSE* and *PRE* of the proposed and existing estimator are given in table 1, 2 and 3. It has been found that the proposed estimator performs better in both the populations. Thus, proposed estimator $\overline{y}_{GS}^{dc}$ is preferable in use over other estimators.

References

[1] Ahmed, M. S., Al-Titi, O., Al-Rawi, Z. & Abu-Dayyeh, W. (2006), "Estimation of a population mean using different imputation methods". Statistics in Transition, Vol.7, No.6, pp. 1247-1264.

[2] Bhal, S. and Tetuja, R.K.(1991), "Ratio and Product type exponential Eastimator". Information and Optimization Sciences, Vol. 12 pp. 159-163

[3] Chand, L. (1975), "Some ratio type estimator based on two or more auxiliary variables". Unpublished Ph.D. dissertation, Iowa State University, Ames. Iowa

[4] Cochran, W.G. (1940), "The Estimation of the Yields of the Cereal Experiments by Sampling for the ratio of Grain to Total Produce." Journal of Agricultural Science, Vol.30, No. 2, pp. 262-275. doi:10.1017/01621459.1962.10500551

[5] Diana, G., Francesco Perri, P. (2010), "Improved estimators of population mean for missing data." Communications in Statistics Theory and Methods, Vol. 39 No. 18 pp. 3245-3251

[6] Gogoi, U., Singh, B.K. (2017), "Estimation of Population Mean using Exponential dual to ratio type Compromised Imputation for missing data in Survey Sampling" Journal of Statistics Application & Probability., Vol. 6 No. 6 pp. 515-522

[7] Gogoi, U., Singh, B.K. (2018), "Estimation of Population Mean using Imputation Techniques in Sample Survey." International Journal of Statistics and Economics, Vol. 19 No. 3 pp.90-103

[8] Gogoi, U., Singh, B.K. (2019), "Study of a one parameter class of ratio cum product type compromised Imputation technique in survey

sampling" International Journal of Mathematics and Statistics, Vol. 20 No. 2 pp. 16-30

[9] Kadilar, C. And Chingi, H. (2008), "Estimators for the population mean in the case of missing data." Communication in Statistics, Theory and Methods, Vol.37, No. 14 pp. 2226-2236

[10] Kalita, D., Singh, B. K. and Choudhary, S. (2013), "Improved exponential chain ratio and product type estimator for finite population mean in double sampling". International Journal of Statistical Science. Vol. 13, pp. 21-37

[11] Kalita, D., Singh, B.K. (2013), "Exponential dual to ratio and dual to product-type estimators for finite population mean in double sampling", Elixir Statistics. Vol. 59 pp.15458-15470.

[12] Kumar, M. and Bhal, S. (2006), "Class of dual to ratio estimators for double sampling." Statistical Papers, Vol. 47 pp. 319-326

[13] Lee, H. and Sarndal, C. E. (1994), "Experiments with variance estimation from survey data with imputed values. Journal of official Statistical. Vol. 10, No. 3 pp. 231-243

[14] Nath, K. and Singh, B. K. (2018), "Population mean estimation using ratio cum product compromised method of imputation in two phase sampling scheme." Asian Journal of Mathematics and Statistics, DOI: 10.3923/ajms.2018 Vol. 11 No. 1 pp. 27-39

[15] Nath. K., Gohain, D., Gogoi, N. and Singh, B.K. (2020), " Study of imputation based generalized class of dual to product cum dual to ratio estimators for missing data in two phase sampling." https://doi.org/10.28919/jmcs/6671, Vol.11 No. 6 pp. 7921-7938

[16] Neyman, J. (1938), "Contribution to the theory of sampling human population", Journal of the American Statistical Association, Vol. 52 pp. 511-522.

[17] Shukla, A. K. (2019), "Efficient Classes of Modified Ratio and Product Estimators for Estimating Population Mean under Two-Phase Sampling", International Journal of Engineering Research and Technology(IJERT). Vol. 8 No.12 pp.659-663

[18] Shukla, A. K. (2019), "Efficient Classes of Modified Ratio and Product Estimators for Estimating Population Mean under Two-Phase Sampling", International Journal of Engineering Research and Technology(IJERT). Vol. 8 No.12 pp.659-663

[19] Shukla, D. And Thakur, N. S. (2008), "Estimation of mean imputation of missing data using factor type estimators." Statistics in Transition , Vol. 9 No. 1 pp. 33-48

[20] Shukla, D., Thakur, N. S. and Pathak, S. (2013), "Some new aspects on imputation in sampling." American Journal of Mathematics and Computer Science Research, Vol. 6 No. 1 pp.5-15

[21] Shukla. D., Thakur, N.S., Pathak, S. and Rajput, D.S. (2009), "Estimation of mean under imputation of missing data using factor type estimator in two phase sampling." Statistics in Transition, Vol. 10 No. 3, pp. 397-414

[22] Singh S. & Horn S. (2000), "Compromised imputation in survey sampling", Metrika, Vol.51, pp. 266-276.

[23] Singh, A. K., Singh, P. and Singh, V.K. (2014), " Imputation methods of missing data for estimating the population mean using simple random sampling." Global Journal of Advance Research (GJAR). Vol. 1 No. 2 pp. 253-263

[24] Singh, B. K. and Choudhary, S. (2012), "Exponential chain ratio and product type estimator for finite population mean under double sampling scheme". Global Journal of science Frontier Research Mathematics and Decision Science. Vol. 12 No. 6, pp. 13-24

[25] Singh, G.N., and Majhi, D. (2014), "Some chain type exponential estimators of population mean in two phase sampling", Statistics in Transition, Vol. 15 No.2 pp.221-230

[26] Singh, H.P. and Vishwakarma, G.K. (2007), Modified Exponential Ratio and Product Estimators for finite Population Mean in two phase sampling", Vol. 30, No. 3 pp. 217-225

[27] Singh, P., Singh, A. K. and Singh, V. K. (2015), " On the use of compromised imputation for missing data using factor type estimators." J. Sts. Applic. Prob. Lett. Vol. 24 pp. 105-113

[28] Srivenkataramana Tracy D.S. (1980), "An alternative to Ratio in Sample Survey", Ann Ins Stat Math. Vol. 32 pp. 111-120

[29] Thakur, N. S., Yadav, K. and Pathak, S. (2013), "On mean estimation with imputation in two-phase sampling design." Research Journal of Mathematical and Statistical Sciences, Vol. 1 No. 3 pp. 1-9

[30] Upadhaya, L. N. and Singh, H.P. (1999), "Use of Transformed Auxiliary variable in Estimating the finite Population Mean." Biometrical Journal, Vol. 5 No.5 pp.