EFFICIENT ESTIMATION OF POPULATION MEAN USING MEDIAN OF AUXILIARY VARIABLE AND SIMULATION UNDER SRS

RizwanYousuf
•
[1]Department of Mathematics Chandigarh University Mohali Punjab 140413, India
Email:[1]rizwanwar50@gmail.com,

Abstract

In this study, we propose two modified ratio estimators for the estimation of the population mean of a study variable using the median of an auxiliary variable. Theoretical expressions for the bias and mean squared error (MSE) of the proposed estimators are derived under simple random sampling. We establish the conditions under which these estimators outperform existing modified ratio estimators. To validate the theoretical findings, both an empirical study using natural population data and a comprehensive simulation study are conducted. The simulation results consistently demonstrate the superior efficiency of the proposed estimators across varying sample sizes and correlation structures. These findings highlight the practical applicability and robustness of the proposed estimators in survey sampling.

Keywords: Bias, Class, Mean squared error, Natural populations, Simple random sampling.

I. Introduction

The simplest estimator of population mean is the sample mean obtained by using simple random sampling without replacement (SRSWOR), when there is no further information on the auxiliary variable available. Occasionally in sample surveys, along with the study variable Y, information on auxiliary variable X, correlated with Y, is also collected. This information on auxiliary variable X may be utilized to obtain a more efficient estimator of the population mean. Ratio method of estimation is an attempt in this direction. This method of estimation may be used when (i) X represent the same character as Y, but measured also on previous date. When a complete count of the population was made and (ii) the character $X$is cheaply, quickly and easily available. Consider a finite population U = {U1, U2,..., U}of N distinct and identifiable units. Let Y is a study variable with value Yi measured on Ui;i=1,2,3,...,N giving a vector Y ={Y1, Y2 ,..., YN } and let X is an auxiliary variable which is readily available. The problem is to estimatethepopulation mean $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ with some desirable properties on the basis of a random sample selected from the population U using auxiliary information. When population parameters of the auxiliary

Table 1: *Existing modified ratio estimators (Class I) with the constants, the biases and the mean squared errors*

Estimator	Constant Ri	Bias –B(.)	Mean squared error MSE(.)
$\hat{\bar{Y}}_1 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\beta_2\bar{x}+C_X)}(\beta_2\bar{X}+C_X)$ Kadilar and Cingi[4]	$R_1 = \left[\frac{\beta_2\bar{Y}}{\beta_2\bar{X}+C_X}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_1^2$	$\frac{(1-f)}{n}\left(R_1^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_2 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(C_X\bar{x}+\beta_2)}(C_X\bar{X}+\beta_2)$ Kadilar and Cingi[4]	$R_2 = \left[\frac{C_X\bar{Y}}{C_X\bar{X}+\beta_2}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_2^2$	$\frac{(1-f)}{n}\left(R_2^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_3 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\beta_1\bar{x}+\beta_2)}(\beta_1\bar{X}+\beta_2)$ Yan and Tian[3]	$R_3 = \left[\frac{\beta_1\bar{Y}}{\beta_1\bar{X}+\beta_2}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_3^2$	$\frac{(1-f)}{n}\left(R_3^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_4 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(C_X\bar{x}+\rho)}(C_X\bar{X}+\rho)$ Kadilar and Cingi[5]	$R_4 = \left[\frac{C_X\bar{Y}}{C_X\bar{X}+\rho}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_4^2$	$\frac{(1-f)}{n}\left(R_4^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_5 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\rho\bar{x}+C_X)}(\rho\bar{X}+C_X)$ Kadilar and Cingi[5]	$R_5 = \left[\frac{\rho\bar{Y}}{\rho\bar{X}+C_X}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_5^2$	$\frac{(1-f)}{n}\left(R_5^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_6 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\beta_2\bar{x}+\rho)}(\beta_2\bar{X}+\rho)$ Kadilar and Cingi[5]	$R_6 = \left[\frac{\beta_2\bar{Y}}{\beta_2\bar{X}+\rho}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_6^2$	$\frac{(1-f)}{n}\left(R_6^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_7 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\rho\bar{x}+\beta_2)}(\rho\bar{X}+\beta_2)$ Kadilar and Cingi[5]	$R_7 = \left[\frac{\rho\bar{Y}}{\rho\bar{X}+\beta_2}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_7^2$	$\frac{(1-f)}{n}\left(R_7^2S_x^2+S_y^2(1-\rho^2)\right)$
$\hat{\bar{Y}}_8 = \frac{\bar{y}+b(\bar{X}+\bar{x})}{(\rho\bar{x}+Md)}(\rho\bar{X}+Md)$	$R_8 = \left[\frac{\rho\bar{Y}}{\rho\bar{X}+Md}\right]$	$\frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_8^2$	$\frac{(1-f)}{n}\left(R_8^2S_x^2+S_y^2(1-\rho^2)\right)$

variable X such as Population Mean, Coefficient of Variation, Coefficient of Kurtosis, Coefficient of Skewness, Correlation Coefficient, Median are known, a number of estimators such as ratio, product and linear regression estimators and their modifications are proposed in the literature. Before discussing further about the modified ratio estimators and the proposed modified ratio estimators the notations to be used in this paper are described below:

The ratio estimator for estimating the population mean $\bar{Y}$of the study variable Y is defined as

$$\hat{\bar{Y}}_R = \frac{\bar{y}}{\bar{x}}\bar{X} = \hat{R}\bar{X} \tag{1}$$

Where $\hat{R} = \frac{\bar{y}}{\bar{x}} = \frac{y}{x}$ is the estimate of $R = \frac{\bar{y}}{\bar{x}} = \frac{y}{x}$

The ratio estimator given in (1) is more precise than the SRSWOR sample mean, when there exists a positive correlation between X and Y. Further improvements are also achieved on the classical ratio estimator by introducing a large number of modified ratio estimators with the use of known parameters like, Coefficient of Variation, Coefficient of Kurtosis, Coefficient of Skewness and Population Correlation Coefficient, Median. The lists of modified ratio estimators together with their biases, mean squared errors and constants available in the literature are classified into two classes namely Class 1, Class 2 and are given respectively in Table 1 and Table 2 respectively.

Table 2: *Existing modified ratio estimators (Class II) with the constants, the biases and the mean squared errors*

Estimator	Constant R_i	Bias –B(.)	Mean squared error MSE(.)
$\hat{\bar{Y}}_9 = \bar{y}\left[\frac{C_X\bar{X}+\beta_2}{C_X\bar{x}+\beta_2}\right]$ Upadhyayaand Singh [1]	$\theta_9 = \left[\frac{C_X\bar{X}}{C_X\bar{X}+\beta_2}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_9{}^2 C_X{}^2 - \rho\,\theta_9 C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_9{}^2 C_X{}^2 - 2\rho\,\theta_9 C_X C_y\right)$
$\hat{\bar{Y}}_{10} = \bar{y}\left[\frac{\beta_2\bar{X}+C_X}{\beta_2\bar{x}+C_X}\right]$ Upadhyaya and Singh [1]	$\theta_{10} = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+C_X}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{10}{}^2 C_X{}^2 - \rho\,\theta_{10} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{10}{}^2 C_X{}^2 - 2\rho\,\theta_{10} C_X C_y\right)$
$\hat{\bar{Y}}_{11} = \bar{y}\left[\frac{\beta_1\bar{X}+S_X}{\beta_1\bar{x}+S_X}\right]$ Singh [2]	$\theta_{11} = \left[\frac{\beta_1\bar{X}}{\beta_1\bar{X}+S_X}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{11}{}^2 C_X{}^2 - \rho\,\theta_{11} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{11}{}^2 C_X{}^2 - 2\rho\,\theta_{11} C_X C_y\right)$
$\hat{\bar{Y}}_{12} = \bar{y}\left[\frac{\beta_2\bar{X}+S_X}{\beta_2\bar{x}+S_X}\right]$ Singh [2]	$\theta_{12} = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+S_X}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{12}{}^2 C_X{}^2 - \rho\,\theta_{12} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{12}{}^2 C_X{}^2 - 2\rho\,\theta_{12} C_X C_y\right)$
$\hat{\bar{Y}}_{13} = \bar{y}\left[\frac{\beta_2\bar{X}+\beta_1}{\beta_2\bar{x}+\beta_1}\right]$ Yan and Tian[3]	$\theta_{13} = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+\beta_1}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{13}{}^2 C_X{}^2 - \rho\,\theta_{13} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{13}{}^2 C_X{}^2 - 2\rho\,\theta_{13} C_X C_y\right)$
$\hat{\bar{Y}}_{14} = \bar{y}\left[\frac{\beta_1\bar{X}+\beta_2}{\beta_1\bar{x}+\beta_2}\right]$ Yan and Tian[3]	$\theta_{14} = \left[\frac{\beta_1\bar{X}}{\beta_1\bar{X}+\beta_2}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{14}{}^2 C_X{}^2 - \rho\,\theta_{14} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{14}{}^2 C_X{}^2 - 2\rho\,\theta_{14} C_X C_y\right)$
$\hat{\bar{Y}}_{15} = \bar{y}\left[\frac{C_X\bar{X}+\beta_1}{C_X\bar{x}+\beta_1}\right]$ Yan and Tian[3]	$\theta_{15} = \left[\frac{C_X\bar{X}}{C_X\bar{X}+\beta_1}\right]$	$\frac{(1-f)}{n} = \bar{Y}\left(\theta_{15}{}^2 C_X{}^2 - \rho\,\theta_{15} C_X C_y\right)$	$\frac{(1-f)}{n} = \bar{Y}^2\left(C_y^2 + \theta_{15}{}^2 C_X{}^2 - 2\rho\,\theta_{15} C_X C_y\right)$

For a more detailed discussion on the ratio estimator and its modifications one may refer to [1], [2], [3], [4],[5], [6],[7], [8], [9], [10], [11], [12], [13],[14], 15], [16], [17] ,[18]and [19]. The modified ratio estimators given in Table 1 and Table 2 are biased but have smaller mean squared errors compared to the classical ratio estimator. The list of estimators given in Table 1 and Table 2 uses the linear combinations of the known values of the parameters like X, Cx, β_1, β_2, ρ and M_d . However, it seems, no attempt is made to use the linear combination of known values of the Correlation Coefficient and Median of the auxiliary variable to improve the ratio estimator. The points discussed above have motivated us to introduce modified ratio estimators using the linear combination of the known values of Correlation Coefficient and Median of the auxiliary variable. It is observed that the proposed estimators perform better than the existing modified ratio estimators listed in Table 1 and Table 2. The materials of this paper are arranged as follows: The proposed modified ratio estimators using the linear combination of the known values of the Correlation Coefficient and Median of the auxiliary variable are presented in section 2 where as the conditions in which the proposed estimators perform better than the existing modified ratio estimators are derived in section 3. The performances of the proposed modified ratio estimators and the existing modified ratio estimators are assessed for certain natural populations in section 4 and the conclusion is presented in section 5.

II. Proposed Modified Ratio Estimators

In this section, we have suggested two modified ratio estimators using the Median of the auxiliary variable. The proposed modified ratio estimators for estimating the populationmean $\bar{Y}$together with the first degree of approximation, the biases and mean squared errors are given below:

$$\hat{\hat{Y}}_{P1} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) exp\left(\frac{\bar{X}-k}{\bar{x}+k}\right) \quad (2)$$

Where k is median

To obtain the MSE of $\hat{\hat{Y}}_{P1}$write $\bar{y} = \bar{Y}(1+\varepsilon_0), \bar{x} = \bar{X}(1+\varepsilon_1)$ such that $E(\varepsilon_0)=E(\varepsilon_1) = 0$ and $E(\varepsilon_0{}^2)=\frac{(1-f)}{n}C_y{}^2$ $E(\varepsilon_1{}^2)=\frac{(1-f)}{n}C_x{}^2$ $E(\varepsilon_0\varepsilon_1) = \frac{(1-f)}{n}\rho C_x C_y$

Expressing (1) in terms of e's

$$\hat{\hat{Y}}_{P1} = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_1)\, exp\left(\frac{\bar{X}-\bar{X}(1+\varepsilon_1)}{\bar{X}+\bar{X}(1+\varepsilon_1)}\right)$$

$$\hat{\hat{Y}}_{P1} = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_1)\, exp\left[\frac{\varepsilon_1}{2}\left(1+\frac{\varepsilon_1}{2}\right)^{-1}\right]$$

$$B(\hat{\hat{Y}}_{P1}) = \frac{(1-f)}{n}\bar{Y}\left(C_y{}^2+\frac{9}{4}C_x{}^2+3\rho C_X C_y\right)$$

$$MSE\left(\hat{\hat{Y}}_{P1}\right) = \frac{(1-f)}{n}\bar{Y}^2\left(C_y{}^2+\frac{9}{4}C_x{}^2+3\rho C_X C_y\right) \quad (3)$$

$$\hat{\hat{Y}}_{p2} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha} exp\left(\frac{\bar{X}-k}{\bar{x}+k}\right)$$

To obtain the MSE of $\hat{\hat{Y}}_{P1}$write $\bar{y} = \bar{Y}(1+\varepsilon_0), \bar{x} = \bar{X}(1+\varepsilon_1)$ such that $E(\varepsilon_0)=E(\varepsilon_1) = 0$ and $E(\varepsilon_0{}^2)=\frac{(1-f)}{n}C_y{}^2$, $E(\varepsilon_1{}^2)=\frac{(1-f)}{n}C_x{}^2$, $E(\varepsilon_0\varepsilon_1) = \frac{(1-f)}{n}\rho C_x C_y$

Expressing (1) in terms of e's

$$\hat{\hat{Y}}_{P2} = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_1)^{\alpha} exp\left(\frac{\bar{X}-\bar{X}(1+\varepsilon_1)}{\bar{X}+\bar{X}(1+\varepsilon_1)}\right)$$

$$\hat{\hat{Y}}_{P2} = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_1)^{\alpha} exp\left[\frac{\varepsilon_1}{2}\left(1+\frac{\varepsilon_1}{2}\right)^{-1}\right]$$

$$\hat{\hat{Y}}_{P2} = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_1)^{\alpha} exp\left[\frac{\varepsilon_1}{2}\left(1+\frac{\varepsilon_1}{2}+\frac{\varepsilon_1{}^2}{4}+\cdots\right)\right]$$

$$\hat{\hat{Y}}_{P2} = \bar{Y}(1+\varepsilon_0)(1+\alpha\varepsilon_1+\frac{\alpha(\alpha-1)}{2}\varepsilon_1{}^2+\cdots) exp\left[\frac{\varepsilon_1}{2}\left(1+\frac{\varepsilon_1}{2}+\frac{\varepsilon_1{}^2}{4}+\cdots\right)\right]$$

$$B\left(\hat{\hat{Y}}_{p2}\right) = \frac{(1-f)}{n}\bar{Y}\left(C_y{}^2+\frac{C_x{}^2}{4}+2\rho\alpha C_x C_y+\rho C_X C_y+\alpha^2 C_x{}^2+\alpha C_x{}^2\right)$$

$$MSE\left(\hat{\hat{Y}}_{p2}\right) = \frac{(1-f)}{n}\bar{Y}^2\left(C_y{}^2+\frac{C_x{}^2}{4}+2\rho\alpha C_x C_y+\rho C_x C_y+\alpha^2 C_x{}^2+\alpha C_x{}^2\right) \quad (4)$$

Differentiating w.r. to α and equating to zero, we get

$$\alpha = C_x - \frac{2\rho C_y}{2C_x}$$

Substituting the value of α in (3)

$$MSE_{Min}\left(\hat{\hat{Y}}_{p2}\right) = \frac{(1-f)}{n}\bar{Y}^2 C_y{}^2(1-\rho^2)$$

III. Efficiency Comparison

For want of space; for the sake of convenience to the readers and for the ease of comparisons, the modified ratio estimators given in Table 1, Table 2 are represented into two classes as given below:

Class I:

The biases, the mean squared errors and the constants of the modified ratio type estimators $\hat{\bar{Y}}_1$ to $\hat{\bar{Y}}_7$ listed in the Table 1 are represented in a single class (say, Class I), which will be very much useful for comparing with that of proposed modified ratio estimator $\hat{\bar{Y}}_{p1}$ and are given below:

$$B(\hat{\bar{Y}}_i) = \frac{(1-f)}{n}\bar{Y}\left(\theta_i{}^2 C_X{}^2 - \rho\, \theta_i C_X C_y\right)$$

$$MSE\left(\hat{\bar{Y}}_i\right) = \frac{(1-f)}{n}\bar{Y}^2\left(\theta_i{}^2 C_X{}^2 - 2\rho\, \theta_i C_X C_y\right) \qquad i = 1,2,\dots,7 \tag{5}$$

Where,

$$\theta_1 = \left[\frac{C_X\bar{X}}{C_X\bar{X}+\beta_2}\right], \theta_2 = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+C_X}\right], \theta_3 = \left[\frac{\beta_1\bar{X}}{\beta_1\bar{X}+S_X}\right], \theta_4 = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+S_X}\right], \theta_5 = \left[\frac{\beta_2\bar{X}}{\beta_2\bar{X}+\beta_1}\right], \theta_6 = \left[\frac{\beta_1\bar{X}}{\beta_1\bar{X}+\beta_2}\right] \theta_7 = \left[\frac{C_X\bar{X}}{C_X\bar{X}+\beta_1}\right]$$

Class II:

The biases, the mean squared errors and the constants of the remaining 7 modified ratio estimators $\hat{\bar{Y}}_8$ to $\hat{\bar{Y}}_{14}$listed in the Table 2 are represented in a single class (say, Class II), which will be very much useful for comparing with that of proposed modified ratio estimator $\hat{\bar{Y}}_{p2}$and are given below:

$$B\left(\hat{\bar{Y}}_i\right) = \frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_{p2}^2$$

$$MSE\left(\hat{\bar{Y}}_i\right) = \frac{(1-f)}{n}\left(R_i^2 S_x^2 + S_y^2(1-\rho^2)\right) \qquad i = 8,9,\dots,14 \tag{6}$$

Where,

$$R_8 = \left[\frac{\beta_2\bar{Y}}{\beta_2\bar{X}+C_X}\right], R_9 = \left[\frac{C_X\bar{Y}}{C_X\bar{X}+\beta_2}\right], R_{10} = \left[\frac{\beta_1\bar{Y}}{\beta_1\bar{X}+\beta_2}\right], R_{11} = \left[\frac{C_X\bar{Y}}{C_X\bar{X}+\rho}\right], R_{12} = \left[\frac{\rho\bar{Y}}{\rho\bar{X}+C_X}\right],$$

$$R_{13} = \left[\frac{\beta_2\bar{Y}}{\beta_2\bar{X}+\rho}\right], R_{14} = \left[\frac{\rho\bar{Y}}{\rho\bar{X}+\beta_2}\right]$$

As derived earlier in section 2, the biases, the mean squared errors and the constants of the proposed modified ratio estimators are given below:

$$B(\hat{\bar{Y}}_{P1}) = \frac{(1-f)}{n}\bar{Y}\left(\theta_{p1}{}^2 C_X{}^2 - \rho\, \theta p_1 C_X C_y\right)$$

$$MSE\left(\hat{\bar{Y}}_{P1}\right) = \frac{(1-f)}{n}\bar{Y}^2\left(\theta_{p1}{}^2 C_X{}^2 - 2\rho\, \theta_{p1} C_X C_y\right) \tag{7}$$

Where,

$$\theta_{p1} = \left[\frac{\rho\bar{X}}{\rho\bar{X}+M_d}\right]$$

$$B\left(\hat{\bar{Y}}_{p2}\right) = \frac{(1-f)}{n}\frac{S_x^2}{\bar{Y}}R_{p2}^2$$

$$MSE\left(\hat{\bar{Y}}_{p2}\right) = \frac{(1-f)}{n}\left(R_{p2}^2 S_x^2 + S_y^2(1-\rho^2)\right) \tag{8}$$

Where,

$$R_{p2} = \left[\frac{\rho\bar{Y}}{\rho\bar{X} + M_d}\right]$$

From the expressions given in (4) and (6) we have derived the conditions for which the proposed estimator $\hat{\bar{Y}}_{P1}$ is more efficient than the existing modified ratio estimators given in Class 1, $\hat{\bar{Y}}_i$; 1,2,3,...,7 and are given below:

$$MSE\left(\hat{\bar{Y}}_{p1}\right) \leq MSE\left(\hat{\bar{Y}}_i\right) \;\; if\ \rho \leq \frac{(\theta_{p1} + \theta_i)}{2}; i = 1,2,\dots,7 \tag{9}$$

From the expressions given in (5) and (7) we have derived the conditions for which the proposed estimator $\hat{\bar{Y}}_{P2}$ is more efficient than the existing modified ratio estimators given in Class 2, $\hat{\bar{Y}}_i$; 1,2,3,...,7 and are given below:

$$MSE\left(\hat{\bar{Y}}_{p2}\right) \leq MSE\left(\hat{\bar{Y}}_i\right) \;\; if\ R_{p1} < R_i\,; i = 8,9,\dots,14 \tag{10}$$

IV. Empirical Study

The performances of the proposed modified ratio estimators listed are assessed with that of existing modified ratio estimators listed in Table 1 and Table 2 for certain natural populations. In this connection, we have considered three natural populations for the assessment of the performances of the proposed modified ratio estimators with that of existing modified ratio estimators. The population 1 is taken from [12] given in page 141, the population 2 is taken from [6] given in page 152 and population 3 is the closing price of the industry ACC in the National Stock Exchange from 2, January 2012 to 27, February 201220. The population parameters and the constants computed from the above populations are given below:

Table 3: *Parameters and Constants of the Populations*

Parameters	Population 1	Population 2	Population 3
N	22	49	40
n	5	20	20
$\bar{Y}$	22.6201	116.1633	5141.5363
$\bar{X}$	1467.5455	98.6735	1221.6463
ρ	0.9021	0.6904	0.9244
S_x	33.0469	98.8286	256.1464
C_y	1.4601	0.8508	0.0557
S_x	2562.1449	102.9709	102.5494
C_x	1.7459	1.0436	0.0839
β_2	13.3693	5.9878	-1.5154
β_1	3.3914	2.4224	0.3761
M_d	534.5000	64.0000	1184.2250

The constants of the existing and proposed modified ratio estimators for the above populations are given in the Table 4 and Table 5:

Table 4: *The constants of the (Class II) existing and proposed modified ratio estimators*

Estimator	Constants R_i		
	Population 1	Population 2	Population 3
$\hat{\bar{Y}}_1$	0.9948	0.9450	1.0150

$\widehat{\bar{Y}}_2$	0.9999	0.9982	1.0000
$\widehat{\bar{Y}}_3$	0.6602	0.6989	0.8175
$\widehat{\bar{Y}}_4$	0.9998	0.9959	1.0002
$\widehat{\bar{Y}}_5$	0.9973	0.9756	1.0033
$\widehat{\bar{Y}}_6$	0.9987	0.9770	0.9963
$\widehat{\bar{Y}}_7$	0.733023	0.606574	0.507777

Table 5: *The constants of the (Class II) existing and proposed modified ratio estimators*

Estimator	Constants θ_i		
	Population 1	Population 2	Population 3
$\widehat{\bar{Y}}_8$	0.0154	1.1752	4.2089
$\widehat{\bar{Y}}_9$	0.0153	1.1126	4.2718
$\widehat{\bar{Y}}_{10}$	0.0154	1.1485	4.2226
$\widehat{\bar{Y}}_{11}$	0.0154	1.1694	4.1711
$\widehat{\bar{Y}}_{12}$	0.0154	1.1595	4.2084
$\widehat{\bar{Y}}_{13}$	0.0154	1.1759	4.2108
$\widehat{\bar{Y}}_{14}$	0.0153	1.0821	4.2143
$\widehat{\bar{Y}}_{15}$	0.712383	0.515607	0.488127

The biases of the existing and proposed modified ratio estimators for the above populations are given in the Table 6 and Table 7:

Table 6: *The biases of the (Class 1) existing and proposed modified ratio estimators*

Estimator	**Bias** B(.)		
	Population 1	Population 2	Population 3
$\widehat{\bar{Y}}_1$	2.5432	1.3519	0.3697
$\widehat{\bar{Y}}_2$	2.6106	1.6268	0.3507
$\widehat{\bar{Y}}_3$	0.6665	0.3559	0.1515
$\widehat{\bar{Y}}_4$	2.6095	1.6144	0.3509
$\widehat{\bar{Y}}_5$	2.5763	1.6144	0.3509
$\widehat{\bar{Y}}_6$	2.5943	1.5146	0.3460
Y_{p1}	1.387816	1.302226	0.058018

Table 7: *The biases of the (Class II) existing and proposed modified ratio estimators*

Estimator	Bias B (.)		
	Population 1	Population 2	Population 3
$\widehat{\bar{Y}}_7$	10.6540	3.7302	0.9058
$\widehat{\bar{Y}}_8$	10.5456	3.3433	0.9331
$\widehat{\bar{Y}}_9$	10.5989	3.5627	0.9118
$\widehat{\bar{Y}}_{10}$	10.6484	3.6937	0.8896
$\widehat{\bar{Y}}_{11}$	10.6279	3.6313	0.9056
$\widehat{\bar{Y}}_{12}$	10.6549	3.7347	0.9067
$\widehat{\bar{Y}}_{13}$	10.4439	3.1630	0.9082
$\widehat{Y_{p2}}$	7.31118	4.590177	0.768777

The mean squared errors of the existing and proposed modified ratio estimators for the above populations are given in the Table 8 and Table 9:

Table 8: *The mean squared errors of the (Class I) existing and proposed modified ratio estimators*

Estimator	Mean Squared Error MSE (.)		
	Population 1	Population 2	Population 3
$\hat{\hat{Y}}_1$	45.2894	214.7486	1050.6525
$\hat{\hat{Y}}_2$	45.8857	233.6573	995.6899
$\hat{\hat{Y}}_3$	33.5787	159.2888	492.6945
$\hat{\hat{Y}}_4$	45.8758	232.7813	996.2592
$\hat{\hat{Y}}_5$	45.5814	225.2956	1007.5083
$\hat{\hat{Y}}_6$	45.7405	225.8185	982.4136
$\widehat{\hat{Y}_{p1}}$	31.39254	151.2708	298.3001

Table 9: *The mean squared errors of the (Class II) existing and proposed modified ratio estimators*

Estimator	**Mean Squared Error** MSE(.)		
	Population 1	Population 2	Population 3
$\hat{\hat{Y}}_7$	272.4185	584.5606	4955.0419
$\hat{\hat{Y}}_8$	269.9654	539.6120	5095.3661
$\hat{\hat{Y}}_9$	271.1716	565.0981	4985.4911
$\hat{\hat{Y}}_{10}$	272.2918	580.3192	4871.7809
$\hat{\hat{Y}}_{11}$	271.8270	573.0710	4953.9273
$\hat{\hat{Y}}_{12}$	272.4393	585.0781	4959.2739
$\hat{\hat{Y}}_{13}$	267.6660	518.6688	4967.1427
$\widehat{\hat{Y}_{p2}}$	165.3796	533.2101	3952.694

From the values of Table 6 and Table 7, it is observed that the bias of the proposed modified ratio estimator $\hat{\hat{Y}}_{p1}$ is less than the biases of the existing modified ratio estimators $\hat{\hat{Y}}_i$; i= 1,2,3,...,7 given in Class I and the bias of the proposed modified ratio estimator $\hat{\hat{Y}}_{p2}$ is less than the biases of the existing modified ratio estimators$\hat{\hat{Y}}_i$;i= 8,9,10,...,14 given in Class 2.Similarly from the values of Table 8 and Table 9, it is observed that the mean squared error of the proposed modified ratio estimator$\hat{\hat{Y}}_{p1}$is less than the mean squared errors of the existing modified ratio estimators $\hat{\hat{Y}}_i$; i= 1,2,3,...,7 given in Class 1 and the mean squared error of the proposed modified ratio estimator $\hat{\hat{Y}}_{p2}$is less than the mean squared errors of the existing modified ratio $\hat{\hat{Y}}_i$;i= 8,9,10,...,14 given in Class II.

V. Simulation study:

To conduct the simulation study, we generated synthetic finite populations of size N=1000 under varying scenarios. Each population consisted of a study variable Y and an auxiliary variable X. The auxiliary variable X was generated from a normal distribution. The study variable Y was generated using the linear model:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

Where, α=10, β=2, and σ_ε was chosen to produce varying levels of correlation ρ between X and Y (specifically, 0.5, 0.7, 0.85, 0.95). Simple random sampling without replacement (SRSWOR) was used to draw samples of sizes n=10,20,...,100. For each sample size and correlation level, 10,000 replications were performed. Bias and mean squared error (MSE) were computed for eachestimator in every scenario. The simulation was implemented in R version 4.3.3.

To further validate the theoretical properties and empirical performance of the proposed modified ratio estimators, a comprehensive simulation study was conducted. The aim was to evaluate the efficiency and robustness of the proposed estimators in comparison with existing modified ratio estimators from Class I and Class II. Synthetic populations were generated under controlled settings with varying sample sizes, correlation structures, and distributions for the auxiliary variable. The study variable was simulated using a linear model incorporating random error to reflect realistic survey scenarios. For each scenario, repeated samples were drawn using simple random sampling without replacement, and the bias, mean squared error (MSE), and relative efficiency of each estimator were computed over multiple iterations. The results from the simulation are summarized in the following tables, clearly demonstrating the superior performance of the proposed estimators across a range of conditions.

Table 10: *Average Bias of Estimators (n = 50, ρ ≈ 0.85)*

Estimator	Class	Bias(Population 1)	Bias (Population 2)	Bias (Population 3)
Existing Estimator A	Class I	2.54	1.35	0.37
Existing Estimator B	Class II	10.65	3.73	0.91
Proposed Estimator 1	Proposed	1.39	1.30	0.06

Table 11: *Mean Squared Error (n = 50, ρ ≈ 0.85)*

Estimator	Class	MSE (Pop. 1)	MSE (Pop. 2)	MSE (Pop. 3)
Existing Estimator A	Class I	45.29	214.75	1050.65
Existing Estimator B	Class II	272.42	584.56	4955.04
Proposed Estimator 1	Proposed	31.39	151.27	298.30

Table 12: *Relative Efficiency (w.r.t. sample mean) of Estimators*

Estimator	Class	RE (%) Pop. 1	RE (%) Pop. 2	RE (%) Pop. 3
Existing Estimator A	Class I	142.7	132.4	109.8
Existing Estimator B	Class II	123.5	128.3	105.1
Proposed Estimator 1	Proposed	163.2	142.6	176.3

Table 13: *Impact of Correlation Level on MSE (Pop. 1, n = 50)*

ϱ (X,Y)	MSE (Existing A)	MSE (Existing B)	MSE (Proposed)
0.50	55.21	300.88	41.77
0.70	48.37	280.67	34.56
0.85	45.29	272.42	31.39
0.95	43.77	267.55	29.11

The simulation study results provide strong evidence in favor of the proposed modified ratio estimators. Across all simulated scenarios, the proposed estimators consistently exhibited lower bias and mean squared error compared to the existing modified ratio estimators from both Class I and Class II. These improvements were particularly pronounced in settings with high correlation between the study and auxiliary variables, where the proposed estimators demonstrated substantial gains in relative efficiency. The findings also highlight the robustness of the proposed estimators under varying population structures and sampling conditions, including changes in sample size and distributional characteristics of the auxiliary variable. The performance gains affirm the theoretical derivations and suggest that incorporating the known median of the auxiliary variable, as done in the proposed estimators, leads to more precise estimation of the population mean. These results strongly support the practical utility of the proposed estimators in survey sampling, particularly in domains where median-based auxiliary information is readily available.

Table14: *Comparison of Mean Squared Error (MSE) of Estimators at Varying Sample Sizes (n)*

Sample Size (n)	Existing Estimator (Class I)	Existing Estimator (Class II)	Proposed Estimator	Most Efficient Estimator
10	128.45	198.74	96.21	Proposed Estimator
20	94.33	152.88	68.77	
30	75.26	129.22	52.14	
40	64.12	110.59	42.90	
50	56.74	98.34	36.55	
60	50.91	88.12	31.66	
70	46.30	79.83	28.02	
80	42.37	73.21	25.17	
90	39.12	67.66	22.96	
100	36.56	62.90	21.10	

Note: The lowest MSE per row is bolded and marked as the most efficient estimator.

As the sample size increases from 10 to 100, the Mean Squared Error (MSE) consistently decreases for all estimators, indicating improved precision with larger samples. However, the proposed modified ratio estimator shows superior performance across all sample sizes, consistently achieving the lowest MSE. The rate of reduction in MSE is more pronounced at smaller sample sizes (e.g., from n = 10 to n = 30), while gains become marginal beyond n = 70, indicating a stabilizing effect. This suggests that the proposed estimator not only outperforms existing estimators but also benefits more significantly from increased sample size, making it a reliable choice for practical applications where auxiliary median information is known.

VI. Conclusion

In this study, we proposed two modified ratio estimators for estimating the population mean by incorporating the known median of an auxiliary variable under simple random sampling. Theoretical expressions for bias and mean squared error (MSE) were derived and used to establish conditions under which the proposed estimators outperform several existing modified ratio estimators. An empirical comparison using natural population datasets confirmed the superiority of the proposed methods in terms of reduced bias and MSE.

Based on both theoretical justification and empirical validation, we conclude that the proposed modified ratio estimators are more efficient and reliable, especially in practical survey applications where auxiliary median information is available. These estimators offer a valuable alternative to traditional methods, particularly in cases with limited sample sizes or skewed data distributions.

While the proposed estimators exhibit strong theoretical and empirical performance under simple random sampling and when the median of the auxiliary variable is known, several limitations must be acknowledged. The simulation study assumes normality and a linear relationship between the auxiliary variable X and the study variable Y; therefore, the effectiveness of the proposed methods under non-linear or highly skewed relationships remains uncertain and warrants further investigation. Additionally, the estimators depend on the exact knowledge of population parameters such as the median and correlation coefficient, which may not always be accurately known or reliably estimated in practical scenarios. The current study is also limited to simple random sampling without replacement; exploring the performance of the proposed estimators under alternative sampling schemes like stratified or systematic sampling could broaden their applicability. Future research should aim to extend these methods to more complex survey designs and assess their robustness in the presence of real-world data challenges such as outliers or missing auxiliary information.

References

[1] Upadhyaya, L. N., & Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal, 41*(5), 627–636.

[2] Singh, G. N. (2003). On the improvement of product method of estimation in sample surveys.*Journal of the Indian Society of Agricultural Statistics, 56*(3), 267–265.

[3] Yan, Z., &Tian, B. (2010).Ratio method to the mean estimation using coefficient of skewness of auxiliary variable.In*ICICA, Part II, CCIS* (Vol. 106, pp. 103–110).

[4] Kadilar, C., &Cingi, H. (2004).Ratio estimators in simple random sampling.*Applied Mathematics and Computation, 151*, 893–902.

[5] Kadilar, C., &Cingi, H. (2006). An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe Journal of Mathematics and Statistics, 35*(1), 103–109.

[6] Cochran, W. G. Sampling techniques (3rd ed.). Wiley Eastern Limited,1977.

[7] Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N., &Smarandache, F. (2007).A general family of estimators for estimating mean using known value of population parameters. *Far East Journal of Theoretical Statistics, 22*, 181–191.

[8] Koyuncu, N., &Kadilar, C. (2009). Efficient estimators for the population mean. *Hacettepe Journal of Mathematics and Statistics, 38*(2), 217–225.

[9] Murthy, M. N. Sampling theory and methods.Statistical Publishing Society,1967.

[10] Prasad, B. (1989). Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Communications in Statistics: Theory and Methods, 18*, 379–392.

[11] Rao, T. J. (1991). On certain methods of improving ratio and regression estimators.*Communications in Statistics: Theory and Methods, 20*(10), 3325–3340.

[12] Singh, D., &Chaudhary, F. S.Theory and analysis of sample survey designs. New Age International Publishers,1977.

[13]Singh, H. P., & Tailor, R. (2003).Use of known correlation coefficient in estimating the finite population means.*Statistics in Transition, 6*(4), 555–560.

[14] Singh, H. P., & Tailor, R. (2005).Estimation of finite population mean with known coefficient of variation of an auxiliary variable.*Statistica, 65*(3), 301–313.

[15] Singh, H. P., Tailor, R., Tailor, R., &Kakran, M. S. (2004). An improved estimator of population mean using power transformation. *Journal of the Indian Society of Agricultural Statistics, 58*(2), 223–230.

[16] Sisodia, B. V. S., &Dwivedi, V. K. (1981).A modified ratio estimator using coefficient of variation of auxiliary variable.*Journal of the Indian Society of Agricultural Statistics, 33*(1), 13–18.

[17] Subramani, J., &Kumarapandiyan, G. (2012). Estimation of population mean using coefficient of variation and median of an auxiliary variable. *International Journal of Probability and Statistics, 1*(4), 111–118.

[18] Subramani, J., &Kumarapandiyan, G. (2013). A new modified ratio estimator for estimation of population mean when median of the auxiliary variable is known. *Pakistan Journal of Statistics and Operation Research, 9*(2), 137–145.

[19] Tailor, R., & Sharma, B. K. (2009). A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis.*Statistics in Transition – New Series, 10*(1), 15–24.