Tanzania Journal of Science 47(5): 1753-1765, 2021
ISSN 0856-1761, e-ISSN 2507-7961

Improved Modified Ratio Estimation of Population Mean Using Information on Size of the Sample

Sikiru A Suleiman[1*] **and Amos A Adewara**[2]
[1]*Department of Statistics, Kwara State Polytechnic, Ilorin, Nigeria*
[2]*Department of Statistics, University of Ilorin, Ilorin, Nigeria*
**Corresponding author email: sikirusuleiman097@gmail.com*
Received 25 Jul 2021, Revised 2 Dec 2021, Accepted 6 Dec 2021, Published Dec 2021
DOI: https://dx.doi.org/10.4314/tjs.v47i5.22

Abstract

In sample surveys, auxiliary information is used for estimation to improve the efficiency of estimators. Increased precision can be obtained when the variable under study is highly correlated with auxiliary information. In this study, the sample size has been used as information for improved estimation of population mean of the main variable under study. A new modified generalized ratio type estimator of population mean has been proposed and the efficiency was examined using Murthy (1967) and Mukhopadhyay (2009) dataset. The large sample properties, the bias and the mean squared error of the newly proposed modified ratio estimator were obtained up to first order of approximation. The optimum value of the characterizing scalar which minimizes the mean squared error was obtained and the minimum value of the mean squared error of the proposed modified ratio estimator for this optimum value was also obtained. A theoretical comparison of the proposed modified ratio estimators was made with the other existing related estimators of population mean using auxiliary information. The conditions under which the proposed modified ratio estimators perform better than the other existing estimators of population mean are given. A numerical study was also carried out to see the performances of the proposed modified ratio estimators and some existing related ratio estimators of population mean and verify the conditions under which the proposed modified ratio estimators are better than some other existing related ratio estimators considered. It was shown that the proposed modified ratio estimators perform better than some existing related ratio estimators as they are having lower mean squared errors.

Keywords: Ratio Estimator, Sample size, Bias, Mean Squared Error, Efficiency.

Introduction

In sampling theory, estimation of the population parameters is necessary when the size of the population is very large (Gupta and Yadav 2018) and we wish to get the results in very shortest time and with minimum costs, fewer labor, etc. In order to estimate any parameter, the best estimator is the corresponding statistic. Thus, the sample mean is the most suitable estimator for estimating population mean, but it has a reasonably large sampling variance (Gupta and Yadav 2017). Our purpose is to search for the estimator with higher efficiency that has minimum variance or mean squared error. This aim is achieved through the use of auxiliary information provided by the auxiliary variables or auxiliary attributes. It is a well-established phenomenon that supplementary information provided by auxiliary variables often improves the accuracy of estimators of unknown population parameters. Ratio, product, and regression-type estimators are three such methods. Auxiliary information is obtained from auxiliary variable which is highly positively or negatively correlated with the main variable under study (Gupta and Yadav 2017).

Let the finite population under consideration consist of N distinct and identifiable units, and let (xi, yi), $i = 1,2, \ldots, n$ be a two variable sample of size n taken from bivariate variables (X, Y) through simple random sampling without sampling scheme. Let $\bar{X}$ and $\bar{Y}$ be the population means of the auxiliary and the study variables, respectively, and let $\bar{x}$ and $\bar{y}$ be the respective sample means and both are unbiased estimators of $\bar{X}$ and $\bar{Y}$, respectively. Let the correlation coefficient between the variables X and Y be denoted by ρ.

In this study, we have confined our work to positively correlated populations only and proposed seven ratio type estimators for improved estimation of the population mean with higher efficiencies. In addition, its large sample properties have been studied up to the first order of approximation. In sampling literature, many estimators have been proposed when a single auxiliary variable is involved, and they are found to be more efficient than the sample mean, the ratio and product estimators under some realistic conditions, as well as efficient as the regression estimator in the optimum case but the problem of the best estimator in terms of both efficiency and biasedness has not been fully exhausted. This work was another attempt in solving this problem. The aim of this research work was to improve the efficiency of some modified existing ratio type estimators of population mean using suitably chosen scalar such that the mean squared error of the proposed estimator is minimum.

Literature Review

Let U denote a finite population consisting of N units $\{U_1, U_2, \ldots, U_N\}$. Also let Y be study variable taking values $\{Y_1, Y_2, \ldots, Y_N\}$ and X be auxiliary variable taking values $\{X_1, X_2, \ldots, X_N\}$ on ith unit U_i of the population U.

$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ -population mean of the study variable Y.

$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ -population mean of the auxiliary variable X.

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ -sample mean of the study variable Y.

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ -sample mean of the auxiliary variable X.

$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2$ -finite population variance of study variable Y.

$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2$ -finite population variance of auxiliary variable X.

$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)\left(X_i - \bar{X}\right)$ -finite population covariance of X and Y.

$\rho_{yx} = \frac{S_{yx}}{S_x S_y}$ -Pearson's moment correlation coefficient of X and Y.

$C_y = \frac{S_y}{\bar{Y}}$ -coefficient of variation of Y.

$C_x = \frac{S_x}{\bar{X}}$ -coefficient of variation of X.

M_d -Median of the auxiliary variable X.

$\beta_1 = \frac{N\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^3}{(N-1)(N-2)S_x^3}$ -coefficient of skewness of auxiliary variable X.

$\beta_2 = \frac{N(N+1)\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^4}{(N-1)(N-2)(N-3)S_x^4} - \frac{3(N-1)^2}{(N-2)(N-3)}$ - coefficient of kurtosis of auxiliary variable X.

Review of existing estimators

As mentioned earlier, the most suitable estimator for estimating population mean $\bar{Y}$ is the sample mean $\bar{y}$ given by,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (1)$$

It is unbiased for population mean and its variance up to the first order of approximation is given by,

$$V(\bar{y}) = \frac{1-f}{n} S_y^2 = \frac{1-f}{n} \bar{Y}^2 C_y^2 = \gamma \bar{Y}^2 C_y^2 \qquad C_y = \frac{S_y}{\bar{Y}^2}, S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2, \gamma = \frac{1-f}{n}, f = \frac{n}{N}$$

(2)
where,

Table 1: Biases and mean squared errors (MSE) of some existing modified ratio estimators

S/No	Estimator	Constant	Bias	MSE
1.	$\hat{\bar{Y}}_1 = \bar{y}\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right)$ Sisodia and Dwivedi (1981)	$\delta_1 = \left(\frac{\bar{X}}{\bar{X}+C_x}\right)$	$\gamma\bar{Y}\left(\delta_1^2 C_x^2 - 2\delta_1\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_1^2 C_x^2 - 2\delta_1\rho C_x C_y\right)$
2.	$\hat{\bar{Y}}_2 = \bar{y}\left(\frac{\bar{X}C_x+\beta_2}{\bar{x}C_x+\beta_2}\right)$ Upadhyaya and Singh (1999)	$\delta_2 = \left(\frac{\bar{X}C_x}{\bar{X}C_x+\beta_2}\right)$	$\gamma\bar{Y}\left(\delta_2^2 C_x^2 - 2\delta_2\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_2^2 C_x^2 - 2\delta_2\rho C_x C_y\right)$
3.	$\hat{\bar{Y}}_3 = \bar{y}\left(\frac{\bar{X}+\rho}{\bar{x}+\rho}\right)$ Singh and Tailor (2003)	$\delta_3 = \left(\frac{\bar{X}}{\bar{X}+\rho}\right)$	$\gamma\bar{Y}\left(\delta_3^2 C_x^2 - 2\delta_3\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_3^2 C_x^2 - 2\delta_3\rho C_x C_y\right)$
4.	$\hat{\bar{Y}}_4 = \bar{y}\left(\frac{\bar{X}+\beta_2}{\bar{x}+\beta_2}\right)$ Singh et al. (2004)	$\delta_4 = \left(\frac{\bar{X}}{\bar{X}+\beta_2}\right)$	$\gamma\bar{Y}\left(\delta_4^2 C_x^2 - 2\delta_4\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_4^2 C_x^2 - 2\delta_4\rho C_x C_y\right)$
5.	$\hat{\bar{Y}}_5 = \bar{y}\left(\frac{\bar{X}+\beta_1}{\bar{x}+\beta_1}\right)$ Yan and Tian (2010)	$\delta_5 = \left(\frac{\bar{X}}{\bar{X}+\beta_1}\right)$	$\gamma\bar{Y}\left(\delta_5^2 C_x^2 - 2\delta_5\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_5^2 C_x^2 - 2\delta_5\rho C_x C_y\right)$
6.	$\hat{\bar{Y}}_6 = \bar{y}\left(\frac{\bar{X}+M_d}{\bar{x}+M_d}\right)$ Subramani and Kumarpandiyan (2013)	$\delta_6 = \left(\frac{\bar{X}}{\bar{X}+M_d}\right)$	$\gamma\bar{Y}\left(\delta_6^2 C_x^2 - 2\delta_6\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_6^2 C_x^2 - 2\delta_6\rho C_x C_y\right)$
7	$\hat{\bar{Y}}_7 = \bar{y}\left(\frac{\bar{X}+n}{\bar{x}+n}\right)$ Jerajuddin and Kishun (2016)	$\delta_7 = \left(\frac{\bar{X}}{\bar{X}+n}\right)$	$\gamma\bar{Y}\left(\delta_7^2 C_x^2 - 2\delta_7\rho C_x C_y\right)$	$\gamma\bar{Y}^2\left(C_y^2 + \delta_7^2 C_x^2 - 2\delta_7\rho C_x C_y\right)$

Cochran (1940) used the positively correlated auxiliary variable with the study variable and proposed the following usual ratio estimator of population mean as,

$$\hat{\bar{Y}}_r = \bar{y}\frac{\bar{X}}{\bar{x}} \qquad (3)$$

The above estimator is a biased estimator of population mean and its bias and mean squared error, up to the first order of approximation, respectively are,

$$B(\hat{\bar{Y}}_r) = \frac{1-f}{n}\bar{Y}[C_y^2 - \rho_{yx} C_y C_x] \qquad (4)$$

$$MSE(\hat{\bar{Y}}_r) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x]$$

Where,

$$C_x = \frac{S_x}{\bar{X}}, S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2, \rho_{yx} = \frac{Cov(x,y)}{S_x S_y},$$

$$Cov(x,y) = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X}),$$

In literature, various modified estimators of population mean of the study variable using auxiliary variables have been given by various authors. For detailed study of the modified ratio type estimators, latest references can be made of Kadilar and Cingi (2004, 2006(a, b), 2009), Singh (2003), Singh and Tailor (2003, 2005), Singh and Chaudhary (1986), Gupta and Misra (2006), Gupta and Yadav (2017 and 2018), Koyuncu and Kadilar (2009), Misra and Gupta (2008), Subramani (2013a), Subramani and Kumarapandiyan (2012(a,b,c), 2013, Subramani (2013b)), Tailor and Sharma (2009), Yan and Tian (2010), Yadav and Pandey (2011), Yadav and Adewara (2013), Yadav et al. (2014, 2015), Yadav et al. (2016(a, b, c, d)), Abid et al. (2016), Misra et al. (2012), Jerajuddin and Kishun (2016), Cochran (1977) and Tailor et al. (2011).

Thus, biases and mean squared errors of the above estimators may be written as,

$$B(\hat{\bar{Y}}_i) = \gamma\bar{Y}\left(\delta_i^2 C_x^2 - 2\delta_i \rho C_x C_y\right), \quad (5)$$
$$MSE(\hat{\bar{Y}}_i) = \gamma\bar{Y}^2\left(C_y^2 + \delta_i^2 C_x^2 - 2\delta_i \rho C_x C_y\right), \; i = 1,2,\ldots,7.$$

Materials and Methods

The proposed estimators

Motivated by Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Singh and Tailor (2003), Singh et al. (2004), Yan and Tian (2010), Subramani and Kumarpandiyan (2013), Jerajuddin and Kishun (2016) and Gupta and Yadav (2018) estimator of population mean, the following generalized estimators of the population mean using information on size of the sample were proposed as,

$$\xi_{p1} = \bar{y}\left[\alpha_1 + (1-\alpha_1)\left(\frac{\bar{X} + C_x n}{\bar{x} + C_x n}\right)\right] \quad (6)$$

$$\xi_{p2} = \bar{y}\left[\alpha_2 + (1-\alpha_2)\left(\frac{\bar{X}C_x + \beta_2 n}{\bar{x}C_x + \beta_2 n}\right)\right] \quad (7)$$

$$\xi_{p3} = \bar{y}\left[\alpha_3 + (1-\alpha_3)\left(\frac{\bar{X} + \rho n}{\bar{x} + \rho n}\right)\right] \quad (8)$$

$$\xi_{p4} = \bar{y}\left[\alpha_4 + (1-\alpha_4)\left(\frac{\bar{X} + \beta_2 n}{\bar{x} + \beta_2 n}\right)\right] \quad (9)$$

$$\xi_{p5} = \bar{y}\left[\alpha_5 + (1-\alpha_5)\left(\frac{\bar{X} + \beta_1 n}{\bar{x} + \beta_1 n}\right)\right] \quad (10)$$

$$\xi_{p6} = \bar{y}\left[\alpha_6 + (1-\alpha_6)\left(\frac{\bar{X} + M_d n}{\bar{x} + M_d n}\right)\right] \quad (11)$$

$$\xi_{p7} = \bar{y}\left[\alpha_7 + (1-\alpha_7)\left(\frac{\bar{X}M_d + n}{\bar{x}M_d + n}\right)\right] \quad (12)$$

Where, α_i *(i =1, 2, …, 7)* is a suitably chosen constant to be defined such that the mean squared error of the proposed estimator is minimum.

To study the large sample properties of the proposed modified ratio estimators, we have used the following approximations as:

$\bar{y} = \bar{Y}(1+e_0)$; $\bar{x} = \bar{X}(1+e_1)$ such that $E(e_i) = 0, (i = 0,1)$ and $E(e_0^2) = \frac{1-f}{n}C_y^2$

and $E(e_1^2) = \frac{1-f}{n}C_x^2$, $E(e_0 e_1) = \frac{1-f}{n}\rho C_x C_y$

Bias and MSE of ξ_{p1}

Expressing Equation (6) in terms of e_i's, we get

$$\xi_{p1} = \bar{Y}(1+e_0)\left[\alpha_1 + (1-\alpha_1)\left(\frac{\bar{X} + C_x n}{\bar{X}(1+e_1) + C_x n}\right)\right]$$

$$\xi_{p1} = \bar{Y}(1+e_0)\left[\alpha_1 + (1-\alpha_1)\left(\frac{1}{1 + \frac{\bar{X}}{\bar{X} + C_x n}e_1}\right)\right]$$

$$\xi_1 = \bar{Y}(1+e_0)\left[\alpha_1 + (1-\alpha_1)\left(1+\delta_{p1}e_1\right)^{-1}\right] \quad (13)$$

Where, $\delta_{p1} = \frac{\bar{X}}{\bar{X} + C_x n}$

We assume that $|e_1| < 1$, so that $(1+\delta_{p1}e_1)^{-1}$ may be expanded. Now expanding the right-hand side of Equation (13), we have,

$$\xi_{p1} = \bar{Y}(1+e_0)\left[\alpha_1 + (1-\alpha_1)\left(1+\delta_{p1}e_1 + \delta_{p1}^2 e_1^2\right)\right] \quad (14)$$

$$\xi_{p1} = \bar{Y}(1+e_0)\left[\alpha_1 + 1 - \delta_{p1}e_1 + \delta_{p1}^2 e_1^2 - \alpha_1 + \alpha_1\delta_{p1}e_1 - \alpha_1\delta_{p1}^2 e_1^2\right]$$

$$\xi_{p1} = \bar{Y}(1+e_0)\left[1-\delta_{p1}e_1+\delta_{p1}^2e_1^2+\alpha_1\delta_{p1}e_1-\alpha_1\delta_{p1}^2e_1^2\right] \quad (15)$$

Retaining the terms up to the first order of approximation, we have

$$\xi_{p1} = \bar{Y}\left[1+e_0-\delta_{p1}e_1-\delta_{p1}e_0e_1+\delta_{p1}^2e_1^2+\alpha_1\delta_{p1}e_1+\alpha_1\delta_{p1}e_0e_1-\alpha_1\delta_{p1}^2e_1^2\right] \quad (16)$$

Subtracting $\bar{Y}$ from both sides of Equation (16), we get

$$\xi_{p1}-\bar{Y} = \bar{Y}\left[e_0-\delta_{p1}e_1-\delta_{p1}e_0e_1+\delta_{p1}^2e_1^2+\alpha_1\delta_{p1}e_1+\alpha_1\delta_{p1}e_0e_1-\alpha_1\delta_{p1}^2e_1^2\right] \quad (17)$$

Taking expectations on both sides of Equation (17) and putting the values of different expectations, we get the bias of ξ_{p1} as

$$E(\xi_{p1}-\bar{Y}) = B(\xi_{p1}) = \bar{Y}E\left[e_0-\delta_{p1}e_1-\delta_{p1}e_0e_1+\delta_{p1}^2e_1^2+\alpha_1\delta_{p1}e_1+\alpha_1\delta_{p1}e_0e_1-\alpha_1\delta_{p1}^2e_1^2\right]$$

$$B(\xi_{p1}) = \frac{1-f}{n}\bar{Y}\left[-\delta_{p1}\rho C_xC_y+\delta_{p1}^2C_x^2+\alpha_1\delta_{p1}\rho C_xC_y-\alpha_1\delta_{p1}^2C_x^2\right] \quad (18)$$

Squaring both sides of Equation (17) and retaining the terms up to the first order of approximation, we have,

$$(\xi_{p1}-\bar{Y})^2 = \bar{Y}^2\left[e_0^2+\delta_{p1}^2e_1^2-2\delta_{p1}e_0e_1+\alpha_1^2\delta_{p1}^2e_1^2+2\alpha_1\delta_{p1}e_0e_1-2\alpha_1\delta_{p1}^2e_1^2\right] \quad (19)$$

Taking expectation on both sides of Equation (19) and putting the values of different expectations, we get the mean square error of ξ_{p1}, up to the first order of approximation, as

$$E(\xi_{p1}-\bar{Y})^2 = MSE(\xi_{p1}) = \bar{Y}^2E\left[e_0^2+\delta_{p1}^2e_1^2-2\delta_{p1}e_0e_1+\alpha_1^2\delta_{p1}^2e_1^2+2\alpha_1\delta_{p1}e_0e_1-2\alpha_1\delta_{p1}^2e_1^2\right]$$

$$MSE(\xi_{p1}) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2+\delta_{p1}^2C_x^2-2\delta_{p1}\rho C_xC_y+\alpha_1^2\delta_{p1}^2C_x^2+2\alpha_1\delta_{p1}\rho C_xC_y-2\alpha_1\delta_{p1}^2C_x^2\right] \quad 20)$$

which is minimum for α_1 when Equation (20) is partially differentiated with respect to α_1 and equate to zero, we have

$$\frac{\partial MSE(\xi_{p1})}{\alpha_1} = \frac{1-f}{n}\bar{Y}^2\left[2\alpha_1\delta_{p1}^2C_x^2+2\delta_{p1}\rho C_xC_y-2\delta_{p1}^2C_x^2\right]=0$$

$$\alpha_1\delta_{p1}^2C_x^2 = \delta_{p1}^2C_x^2-\delta_{p1}\rho C_xC_y$$

$$\alpha_{p1} = \frac{\delta_{p1}^2C_x^2-\delta_{p1}\rho C_xC_y}{\delta_{p1}^2C_x^2} = \frac{A_1}{B_1}$$

Thus, the minimum MSE of ξ_{p1} is,

$$MSE_{\min}(\xi_{p1}) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2+\delta_{p1}^2C_x^2-2\delta_{p1}\rho C_xC_y-\frac{A_1^2}{B_1}\right] \quad (21)$$

Similarly, the biases and mean squared errors (MSE) of others proposed estimators can be obtained in the same way. Thus, the following are generalized biases and mean squared errors (MSE) of the proposed estimators given by

$$B(\xi_{pi}) = \frac{1-f}{n}\bar{Y}\left[-\delta_{pi}\rho C_xC_y+\delta_{pi}^2C_x^2+\alpha_i\delta_{pi}\rho C_xC_y-\alpha_i\delta_{pi}^2C_x^2\right]$$

$$MSE_{\min}(\xi_{pi}) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2+\delta_{pi}^2C_x^2-2\delta_{pi}\rho C_xC_y-\frac{A_i^2}{B_i}\right] \quad (22)$$

Where,

$$\alpha_i = \frac{\delta_{pi}^2 C_x^2 - \delta_{pi}\rho C_x C_y}{\delta_{pi}^2 C_x^2} = \frac{A_i}{B_i},\ i = 1,2,...,7.$$

Thus, biases and mean squared errors (MSE) of the proposed estimators are given as:

$$B(\xi_{pi}) = \gamma\bar{Y}\left[-\delta_{pi}\rho C_x C_y + \delta_{pi}^2 C_x^2 + \alpha_i\delta_{pi}\rho C_x C_y - \alpha_i\delta_{pi}^2 C_x^2\right] \quad (22)$$

$$MSE_{\min}(\xi_{pi}) = \gamma\bar{Y}^2\left[C_y^2 + \delta_{pi}^2 C_x^2 - 2\delta_{pi}\rho C_x C_y - \frac{A_i^2}{B_i}\right] \quad (23)$$

Where, $\alpha_i = \frac{\delta_{pi}^2 C_x^2 - \delta_{pi}\rho C_x C_y}{\delta_{pi}^2 C_x^2} = \frac{A_i}{B_i},\ i = 1,2,...,7.$

Theoretical efficiency comparison

In this section, the proposed modified ratio estimators were compared theoretically with the other existing related ratio estimators of population mean in terms of their variances and mean squared errors (MSE) under simple random sampling without replacement scheme and thereby establishing their efficiency conditions.

Efficiency condition of $\xi_{pi}(i = 1,2,\ldots,7)$ over some related existing ratio estimators

From the MSE of proposed modified ratio estimator ξ_{pi} and Equation (2), proposed modified ratio estimator ξ_{pi} is better than the mean per unit estimator if,

$$V(\bar{y}) - MSE_{\min}(\xi_{pi}) = \gamma\bar{Y}^2\left[\delta_{pi}^2 C_x^2 - 2\delta_{pi}\rho C_x C_y - \frac{A_i^2}{B_i}\right] > 0$$

Or, $\delta_{pi}^2 C_x^2 - 2\delta_{pi}\rho C_x C_y > \frac{A_i^2}{B_i}$ (*i=1, 2,…, 7*) (24)

When Equation (24) is satisfied, ξ_{pi} is more efficient than $\bar{y}$.

From the MSE of proposed modified ratio estimator ξ_{pi} and Equation (2), proposed modified ratio estimator ξ_{pi} is better than the usual ratio estimator $\bar{y}_r$ by Cochran (1940) if,

$$MSE(\hat{\bar{Y}}_r) - MSE_{\min}(\xi_{pi}) = \gamma\bar{Y}^2\left[(R^2 - \delta_{pi}^2)C_x^2 - 2(R - \delta_{pi})\rho C_x C_y - \frac{A_i^2}{B_i}\right] > 0$$

Or, $(R^2 - \delta_{pi}^2)C_x^2 - 2(R - \delta_{pi})\rho C_x C_y > \frac{A_i^2}{B_i}$ (*i=1, 2, …, 7*) (25)

When Equation (25) is satisfied, ξ_{pi} is more efficient than $\bar{y}_r$.

From the MSE of proposed modified ratio estimator ξ_{pi} and MSE in Tables 1 and 2, proposed modified ratio estimator ξ_{pi} is better than the modified existing ratio type estimator by Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Singh and Tailor (2003), Singh et al. (2004), Yan and Tian (2010), Subramani and Kumarpandiyan (2013) and Jerajuddin and Kishun (2016) if,

$$MSE(\hat{\bar{Y}}_i) - MSE_{\min}(\xi_{pi}) = \gamma\bar{Y}^2\left[(R^2 - \delta_{pi}^2)C_x^2 - 2(R - \delta_{pi})\rho C_x C_y - \frac{A_i^2}{B_i}\right] >$$

Or, $(R^2-\delta_{pi}^2)C_x^2-2(R-\delta_{pi})\rho C_x C_y > \frac{A_i^2}{B_i},\quad i=1,2,...,7$ (26)

Table 2: Biases and mean squared errors (MSEs) of the proposed modified ratio estimators

S/No	Estimator	Constant	Bias	MSE
1.	ξ_{p1}	$\delta_{p1}=\frac{\bar{X}}{\bar{X}+C_x n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p1}\rho C_x C_y+\delta_{p1}^2C_x^2\\+\alpha_1\delta_{p1}\rho C_x C_y-\alpha_1\delta_{p1}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p1}^2C_x^2\\-2\delta_{p1}\rho C_x C_y-\frac{A_1^2}{B_1}\end{bmatrix}$
2.	ξ_{p2}	$\delta_{p2}=\frac{\bar{X}C_x}{\bar{X}C_x+\beta_2 n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p2}\rho C_x C_y+\delta_{p2}^2C_x^2\\+\alpha_2\delta_{p2}\rho C_x C_y-\alpha_2\delta_{p2}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p2}^2C_x^2\\-2\delta_{p2}\rho C_x C_y-\frac{A_2^2}{B_2}\end{bmatrix}$
3.	ξ_{p3}	$\delta_{p3}=\frac{\bar{X}}{\bar{X}+\rho n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p3}\rho C_x C_y+\delta_{p3}^2C_x^2\\+\alpha_3\delta_{p3}\rho C_x C_y-\alpha_3\delta_{p3}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p3}^2C_x^2\\-2\delta_{p3}\rho C_x C_y-\frac{A_3^2}{B_3}\end{bmatrix}$
4.	ξ_{p4}	$\delta_{p4}=\frac{\bar{X}}{\bar{X}+\beta_2 n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p4}\rho C_x C_y+\delta_{p4}^2C_x^2\\+\alpha_4\delta_{p4}\rho C_x C_y-\alpha_4\delta_{p4}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p4}^2C_x^2\\-2\delta_{p4}\rho C_x C_y-\frac{A_4^2}{B_4}\end{bmatrix}$
5.	ξ_{p5}	$\delta_{p5}=\frac{\bar{X}}{\bar{X}+\beta_1 n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p5}\rho C_x C_y+\delta_{p5}^2C_x^2\\+\alpha_5\delta_{p5}\rho C_x C_y-\alpha_5\delta_{p5}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p5}^2C_x^2\\-2\delta_{p5}\rho C_x C_y-\frac{A_5^2}{B_5}\end{bmatrix}$
6.	ξ_{p6}	$\delta_{p6}=\frac{\bar{X}}{\bar{X}+M_d n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p6}\rho C_x C_y+\delta_{p6}^2C_x^2\\+\alpha_6\delta_{p6}\rho C_x C_y-\alpha_6\delta_{p6}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p6}^2C_x^2\\-2\delta_{p6}\rho C_x C_y-\frac{A_6^2}{B_6}\end{bmatrix}$
7	ξ_{p7}	$\delta_{p7}=\frac{\bar{X}}{\bar{X}M_d+n}$	$\gamma\bar{Y}\begin{bmatrix}-\delta_{p7}\rho C_x C_y+\delta_{p7}^2C_x^2\\+\alpha_7\delta_{p7}\rho C_x C_y-\alpha_7\delta_{p7}^2C_x^2\end{bmatrix}$	$\gamma\bar{Y}^2\begin{bmatrix}C_y^2+\delta_{p7}^2C_x^2\\-2\delta_{p7}\rho C_x C_y-\frac{A_7^2}{B_7}\end{bmatrix}$

Dataset for empirical study

To judge the performance of the proposed modified ratio estimators and the existing related ratio estimators of population mean using auxiliary variable, we considered four natural populations from two sources. First two populations, populations 1 and 2 are from Murthy (1967), while populations 3 and 4 are from Mukhopadhyay (2009).

Murthy (1967)

Population 1: *Y* = Output for 80 factories in a region and *X*= Number of workers

$N=80, n=20, \bar{Y}=51.8264, \bar{X}=11.2646,$

$\rho=0.9413, C_y=0.3542, C_x=0.7507$

$\beta_1=1.0500, \beta_2=-0.0634, M_d=7.5750.$

Population 2: *Y* = Output for 80 factories in a region and *X* = Fixed Capital

$N=80, n=20, \bar{Y}=51.8264,$

$\bar{X}=11.2646, \rho=0.9413,$

$C_y=0.3542, C_x=0.9485$

$\beta_1=1.3006, \beta_2=0.6977, M_d=1.4800$

Mukhopadhyay (2009)

Population 3: *Y* = Output for 40 factories in a region and *X* = Number of workers

$N=40, n=8, \bar{Y}=50.7858, \bar{X}=2.3033,$

$\rho=0.8006, C_y=0.3295, C_x=0.8406$

$\beta_1 = 0.8799, \beta_2 = -0.4622, M_d = 1.2500$.

Population 4: Y = Output for 40 factories in a region and X = Fixed capital

$N = 40, n = 8, \overline{Y} = 50.7858$, $\overline{X} = 9.4543$,

$\rho = 0.8349$, $C_y = 0.3295, C_x = 0.6756$

$\beta_1 = 0.8799, \beta_2 = -0.4622, M_d = 7.0700$.

Results and Discussion

In this section, the performances of the proposed ratio estimators are evaluated and compared with the mentioned ratio estimators in Table 1 by using the population data of Murthy (1967) and Mukhopadhyay (2009). We apply the proposed and existing estimators to this data set, and the efficiency of the proposed modified ratio estimators over some existing related ratio estimators were investigated using real life data to support the theoretical comparisons in the previous section of this paper.

The numerical values of biases and the mean squared errors as well as percentage relative efficiency (PRE) of the newly proposed modified ratio estimators over other existing related ratio estimators of population mean using auxiliary variable for the four populations are as shown in Tables 3-6.

From Tables 3–6, it can be observed that some proposed modified ratio estimators are having lower biases when compared with other existing related ratio estimators, while the mean squared errors of the newly proposed modified ratio estimators were also lower as compared to other existing related ratio estimators.

Table 3: Biases, mean squared errors and percentage relative efficiency of the existing and newly proposed modified ratio estimators using population 1

Estimator	Constant	Bias	MSE	PRE
$\bar{y}$	0.000000	0.000000	12.63661	NA
$\hat{\bar{Y}}_r$	0.000000	0.608819	18.97931	66.581
$\hat{\bar{Y}}_1$	0.937521	0.050583	15.25812	82.819
$\hat{\bar{Y}}_2$	1.007554	0.131644	19.45925	64.939
$\hat{\bar{Y}}_3$	0.922882	0.034995	14.450269	87.448
$\hat{\bar{Y}}_4$	1.005660	0.129311	19.33831	65.345
$\hat{\bar{Y}}_5$	0.914735	0.026525	14.01128	90.189
$\hat{\bar{Y}}_6$	0.597921	-0.190136	2.782544	454.139
$\hat{\bar{Y}}_7$	0.360299	-0.208344	1.838908	687.180
ξ_{p1}	0.428661	-0.007525	1.439881	877.615
ξ_{p2}	1.176397	0.3562010	1.439996	877.545
ξ_{p3}	0.374356	-0.033941	1.439885	877.612
ξ_{p4}	1.126843	0.332096	1.439992	877.547
ξ_{p5}	0.349132	-0.046211	1.439985	877.552
ξ_{p6}	0.069208	-0.182376	1.439981	877.554
ξ_{p7}	0.810119	0.1780302	1.439988	877.550

Table 4: Biases, mean squared errors and percentage relative efficiency of the existing and newly proposed modified ratio estimators using population 2

Estimator	Constant	Bias	MSE	PRE
$\bar{y}$	0.000000	0.000000	12.63661	NA
$\hat{\bar{Y}}_r$	0.000000	1.151032	41.32765	30.577
$\hat{\bar{Y}}_1$	0.937521	0.087907	17.19251	73.501
$\hat{\bar{Y}}_2$	1.007554	0.155035	20.671513	61.131
$\hat{\bar{Y}}_3$	0.922882	0.097524	17.690914	71.430
$\hat{\bar{Y}}_4$	1.005660	0.168609	21.375015	59.119
$\hat{\bar{Y}}_5$	0.914735	0.004041	12.84606	98.370
$\hat{\bar{Y}}_6$	0.597921	-0.028866	11.140575	113.429
$\hat{\bar{Y}}_7$	0.360299	-0.121869	6.320581	199.928
ξ_{p1}	0.130666	-0.126073	2.056924	614.345
ξ_{p2}	0.162347	-0.107145	2.056889	614.355
ξ_{p3}	0.134805	-0.123600	2.056895	614.354
ξ_{p4}	0.169667	-0.1027724	2.056881	614.358
ξ_{p5}	0.098786	-0.1451187	2.056933	614.342
ξ_{p6}	0.087864	-0.1516441	2.056947	614.338
ξ_{p7}	0.174234	-0.100044	2.056875	614.359

Table 5: Biases, mean squared errors and percentage relative efficiency of the existing and newly proposed modified ratio estimators using population 3

Estimator	Constant	Bias	MSE	PRE
$\bar{y}$	0.000000	0.000000	28.0024	NA
$\hat{\bar{Y}}_r$	0.000000	2.46240	95.86411	29.211
$\hat{\bar{Y}}_1$	0.937521	0.276010	42.01979	66.641
$\hat{\bar{Y}}_2$	1.007554	3.735316	217.7034	12.863
$\hat{\bar{Y}}_3$	0.922882	0.304709	43.47728	64.407
$\hat{\bar{Y}}_4$	1.005660	3.151586	188.05822	14.890
$\hat{\bar{Y}}_5$	0.914735	0.189565	37.62961	74.416
$\hat{\bar{Y}}_6$	0.597921	0.0478559	30.43281	92.014
$\hat{\bar{Y}}_7$	0.360299	-0.324172	11.53909	242.674
ξ_{p1}	0.255126	-0.0661004	10.053897	278.523
ξ_{p2}	-0.827754	-1.285603	10.053953	278.521
ξ_{p3}	0.264501	-0.0555424	10.053889	278.523
ξ_{p4}	-1.168061	-1.668846	10.053966	278.521
ξ_{p5}	0.2281557	-0.096473	10.053941	278.522
ξ_{p6}	0.1872099	-0.142585	10.053976	278.521
ξ_{p7}	0.2646467	-0.0553781	10.053876	278.523

Table 6: Biases, mean squared errors and percentage relative efficiency of the existing and newly proposed modified ratio estimators using population 4

Estimator	Constant	Bias	MSE	PRE
$\bar{y}$	0.000000	0.00000	28.0024	NA
$\hat{\bar{Y}}_r$	0.000000	1.37415	49.85359	56.169
$\hat{\bar{Y}}_1$	0.937521	0.257278	41.06847	68.185
$\hat{\bar{Y}}_2$	1.007554	0.658753	61.45774	45.564
$\hat{\bar{Y}}_3$	0.922882	0.222518	69.30314	40.406
$\hat{\bar{Y}}_4$	1.005660	0.577651	57.33886	48.837
$\hat{\bar{Y}}_5$	0.914735	0.213064	38.82301	72.128
$\hat{\bar{Y}}_6$	0.597921	-0.321274	11.68628	239.618
$\hat{\bar{Y}}_7$	0.360299	-0.342432	10.61172	263.882
ξ_{p1}	0.636263	0.216217	8.483064	330.098
ξ_{p2}	2.374713	1.857124	8.483106	330.096
ξ_{p3}	0.586004	0.168778	8.483052	330.098
ξ_{p4}	1.642312	1.165818	8.483099	330.096
ξ_{p5}	0.573214	0.1567055	8.483031	330.099
ξ_{p6}	0.143216	-0.249166	8.483006	330.100
ξ_{p7}	0.893108	0.4586518	8.483088	330.097

Conclusion

From the results of empirical study using four natural population datasets, it can be concluded that the newly proposed modified ratio estimators in this study demonstrated high relative efficiency over existing related ratio estimators. From Table 3, all the newly proposed modified ratio estimators has PRE of about 877.6 which is higher than the PRE of all the existing related ratio estimations. This is also the case in Tables 4–6, where all the newly proposed modified ratio estimators have PRE of about 614.3 (Table 4), 278.5 (Table 5) and 330.1 (Table 6), respectively, which are higher than the PRE of all the existing related ratio estimations. In population 1, the newly proposed modified ratio estimator ξ_{p1} is the most efficient estimator with PRE of 877.615, followed by $\xi_{p3}, \xi_{p6}, \xi_{p5}, \xi_{p7}, \xi_{p4}$, and ξ_{p2} in that order. Also, in population 2, the newly proposed modified ratio estimator ξ_{p7} is the most efficient estimator with PRE of 614.359, followed by ξ_{p4}, ξ_{p2}, ξ_{p3}, ξ_{p1}, ξ_{p5} and ξ_{p6} in that order. Moreover, in population 3, the newly proposed modified ratio estimators ξ_{p1}, ξ_{p3} and ξ_{p7} are the most efficient estimators with PRE of 278.523, followed by ξ_{p5}, then, ξ_{p2}, ξ_{p4}, and ξ_{p6}. Finally, in population 4, the newly proposed modified ratio estimator ξ_{p6} is the most efficient estimator with PRE of 330.100, followed by ξ_{p5}, then ξ_{p1} and ξ_{p3}, ξ_{p7}, ξ_{p2} and ξ_{p4}, and in that order. In conclusion, the newly proposed modified ratio estimators are improved versions of Gupta and Yadav (2018) generalized estimator of population mean using information on size of the sample. Based on the empirical findings, the newly proposed modified ratio estimators are recommended for estimating finite population mean of any variable of interest.

References

Abid M, Abbas N, Sherwani RAK and Nazir HZ 2016 Improved ratio estimators for the population mean using non-conventional measure of dispersion. *Pak. J. Stat. Oper. Res.* XII (2): 353–367.

Cochran WG 1940 The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *J. Agric. Sci.* 30: 262–275.

Cochran WG 1977 Sampling Techniques. John Wiley and Sons, New York.

Gupta RK and Misra S 2006 Estimation of population variance using ratio type estimator in *Indian J. Math. Mathemat. Sci.* 2(2): 169–176.

Gupta RK and Yadav SK 2017 New efficient estimators of population mean using non-traditional measures of dispersion. *Open J. Stat.* 7(3): 394–404.

Gupta RK and Yadav SK 2018 Improved estimation of population mean using information on size of the sample. *Am. J. Math. Stat.* 8(2): 27–35.

Jerajuddin M and Kishun J 2016 Modified ratio estimators for population mean using size of the sample selected from population. *IJSRSET* 2(2): 10–16.

Kadilar C and Cingi H 2004 Ratio estimators in simple random sampling. *Appl. Math. Comput.* 151: 893–902.

Kadilar C and Cingi H 2006a An improvement in estimating the population mean by using the correlation co-efficient. *Hacettepe J. Math. Stat.* 35(1): 103–109.

Kadilar C and Cingi, H 2006b Improvement in estimating the population mean in simple random sampling. *Appl. Math. Lett.* 19: 75–79.

Kadilar C and Cingi H 2009 Advances in sampling theory-ratio method of estimation, Bentham Science Publishers.

Koyuncu N and Kadilar C 2009 Efficient estimators for the population mean. *Hacettepe J. Math. Stat.* 38(2): 217–225.

Misra S and Gupta RK 2008 Almost unbiased jacknifed ratio type estimator of population variance. *Int. J. Agric. Stat. Sci.* 4: 345–350.

Misra S, Gupta RK and Shukla AK 2012 Generalized class of estimators for estimation of finite population variance. *Int. J. Agric. Stat. Sci.* 8(2): 447-458.

Mukhopadhyay P 2009 Theory and methods of survey sampling. PHI Learning, 2nd edition, New Delhi.

Murthy MN 1967 Sampling Theory and Methods. Statistical Publishing Society, Calcutta.

Singh GN 2003 On the improvement of product method of estimation in sample surveys. *J. Indian Soc. Agric. Stat.* 56(3): 267–265.

Singh HP, Upadhyaya, LN and Chandra P 2004 A general family of estimators for estimating population mean using two auxiliary variables in two-phase sampling. *Stat. Transit.* 6(7): 1055–1077.

Singh D and Chaudhary FS 1986 Theory and analysis of sample survey designs. New Age International Publisher, New Delhi.

Singh HP and Tailor R 2003 Use of known correlation coefficient in estimating the finite population means. *Stat. Transit.* 6(4): 555–560.

Singh HP and Tailor R 2005 Estimation of finite population mean with known co-efficient of variation of an auxiliary variable. *Statistica. anno LXV.* 3: 301–313.

Sisodia BVS and Dwivedi VK 1981 A Modified ratio estimator using coefficient of variation of auxiliary variable. *J. Indian. Soc. Agri. Stat.* 33(1): 13–18.

Subramani J 2013a Generalized modified ratio estimator of finite population mean.

J. Modern Appl. Stat. Meth. 12(2): 121–155.

Subramani J and Kumarapandiyan G 2012a Estimation of population mean using coefficient of variation and median of an auxiliary variable. *Int. J. Prob. Stat.* 1(4): 111–118.

Subramani J, and Kumarapandiyan G 2012b Modified ratio estimators using known median and coefficient of kurtosis. *Am. J. Math. Stat.* 2(4): 95–100.

Subramani J and Kumarapandiyan G 2012c Estimation of population mean using known median and coefficient of skewness. *Am. J. Math. Stat.* 2(5): 101–107.

Subramani J and Kumarapandiyan G 2013a Estimation of population mean using deciles of an auxiliary variable. *Statistics in Transition-New Series* 14(1): 75–88.

Subramani J 2013b A new modified ratio estimator of population mean when median of the auxiliary variable is known. *Pak. J. Stat. Oper. Res.* 9(2): 137–145.

Tailor R, Parmar R, Kim JM and Tailor R 2011 Ratio-cum-product estimators of population mean using known population parameters of auxiliary variates. *Commun. Stat. Appl. Methods* 18(2): 155-164.

Tailor R and Sharma B 2009 A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis. *Stat. Transit.-New Ser.* 10(1): 15–24.

Upadhyaya LN and Singh HP 1999 Use of transformed auxiliary variable in estimating the finite population means. *Biom. J.* 41(5): 627–636.

Yadav SK and Pandey H 2011 Improved exponential estimators of population mean using qualitative auxiliary information under two phase sampling. *Investig. Math. Sci.* 1: 85-94.

Yadav SK and Adewara AA 2013 On improved estimation of population mean using qualitative auxiliary information. *Math. Theor. Model.* 3(11): 42-50.

Yadav SK, Mishra SS and Shukla AK 2014 Improved ratio estimators for population mean based on median using linear combination of population mean and median of an auxiliary variable. *Am. J. Oper. Res.* 4(2): 21-27.

Yadav SK, Mishra, SS and Shukla, AK 2015 Estimation approach to ratio of two inventory population means in stratified random sampling. *Am. J. Oper. Res.* 5(4): 96-101.

Yadav SK, Mishra SS, and Shukla AK, Kumar S and Singh RS 2016a Use of non-conventional measures of dispersion for improved estimation of population mean. *Am. J. Oper. Res.* 6(3): 69-75.

Yadav SK, Gupta SAT, Mishra SS and Shukla AK 2016b Modified ratio and product estimators for estimating population mean in two-phase sampling. *Am. J. Oper. Res.* 6(3): 61-68.

Yadav SK, Subramani J, Mishra SS and Shukla AK 2016c Improved ratio-cum-product estimators of population mean using known population parameters of auxiliary variables. *Am. J. Oper. Res.* 6(2): 48-54.

Yadav SK, Misra S, Mishra SS, and Chutiman N 2016d Improved ratio estimators of population mean in adaptive cluster sampling. *J. Stat. Appl. Prob. Lett.* 3: 1-6.

Yan Z and Tian B 2010 Ratio method to the mean estimation using coefficient of skewness of auxiliary variable. *Int. Conf. Inf. Comput. Appl.* 103-110.