IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Optimized Logarithmic Ratio Type Estimator For Improved Estimation Of Population Mean Using Auxiliary Information

Mukesh Kumar[1] and Anchal Yadav[1*]

Department of Statistics, University of Lucknow, Lucknow-226007, Uttar Pradesh, India

Abstract

Estimation of parameters say mean, variance, median and coefficient of variation etc. is necessary under various studies. The bias and mean square error (MSE) of the estimator reflect the appropriateness and usefulness of the estimator. The improvement in the efficiency of the estimator using auxiliary information motivates the researches to suggest new estimators. Logarithmic ratio type estimator is one of the new classes of estimators where improvement has been suggested by only few authors in the literature and hence contribution to such estimators attempted for possible improvement in the efficiency of the estimators. In this study, we introduced an improved logarithmic ratio type estimator for estimating the finite population means using auxiliary information under simple random sampling scheme. The bias and mean square error (MSE) expressions for the proposed estimator have been derived up to the first order of approximation. The theoretical efficiency comparison has been carried out with the existing estimators. Also, the MSE and percentage relative efficiency (PRE), results has been calculated with real data set and artificial data set generated by using R software with bivariate normal distribution by simulation approach. The result shows that the proposed estimator performs well as compared to the other existing estimators of this class, also some new application areas might be explored in respect of such estimators.

Keywords: Auxiliary information, logarithmic ratio-type estimators, Mean squared error, Bias, simple random sampling, Simulation Study

1. Introduction

In survey sampling, auxiliary information is frequently utilized to enhance the accuracy of population parameter estimation. Ratio estimators are typically employed when a significant positive correlation exists in the relationship between the study variable and the auxiliary variable is considered, especially in cases where the regression line intersects the origin. Similarly, product-type estimators are suitable when there exist a strong negative linear relationship and the regression line passes through the origin. In contrast, when the regression line fails to pass through the origin, regression-type estimators become more appropriate. Various scholars have contributed substantial and commendable work to this domain. Cochran (1940) first proposed the conventional ratio estimator, marking the beginning of the use of auxiliary information in estimation. Later, Murthy (1964) developed the classical product-type estimator. Srivastava (1967) made a notable contribution by introducing a power transformation estimator that generalized both ratio and product estimators. Further progress was achieved by Bahl and Tuteja (1991), who proposed exponential ratio and product-type estimators for estimating the population mean of the study variable, based on auxiliary information. Further many researchers have made significant contributions to this area, several researchers, such as Kadilar and Cingi (2004), Khoshnevisan et al. (2007), Onyeka (2012), Chauhan and Singh (2014), Subramani and Ajith (2016), and Madhulika et al. (2017), have contributed to this field. The arithmetic mean is seen as the basic measure of a population and is more effective in yielding precise results when the population is homogeneous. Numerous efficient estimators for the finite population mean have been developed over time by various researchers under simple random sampling techniques. As an illustration, Singh and Tailor (2005) introduced efficient estimators that make uses of the recognized values of the correlation coefficient. Kadilar and Cingi (2006) presented methods integrating the coefficient of kurtosis to refine mean estimates, and Singh and Vishwakarma (2009) formulated a general framework for estimating the population mean through successive sampling. Haq and Shabbir (2013) introduced an improved family of ratio-type estimators, whereas Singh and Solanki suggested efficient class of estimators that utilizes auxiliary information. Subramani and Kumarpandiyan (2012) proposed a modified ratio-type estimator that utilizes the known median of an auxiliary variable. Muneer et al. (2017) developed estimators for the population mean under both simple and stratified random sampling frameworks, incorporating two auxiliary variables. Additionally, Izunobi and Onyeka (2019) introduced efficient logarithmic ratio and product type estimators in the context of simple random sampling. Building on this foundation, the present study aims to propose novel logarithmic ratio and product-type estimators that leverage auxiliary information to improve the estimation of the finite population mean. These log-based estimators extend beyond traditional and exponential forms, demonstrating superior performance when the logarithm of the auxiliary mean exceeds a particular value, as supported by theoretical and empirical evidence. Cekim et al. (2020) introduced a new family of ln-type estimators for population variance estimation under SRS, highlighting their higher efficiency under specific theoretical conditions. Yunusa et al. (2021) proposed a logarithmic ratio-type estimator for estimating the population coefficient of variation, using log transformations on both population and sample variances of the auxiliary variable. Further, Zaman (2023) and Adejumobi (2023) developed efficient logarithm-type estimators of the population mean under SRSWOR, outperforming previous approaches. Most recently, Qureshi et al. (2024) presented two ln-type estimators for estimating the mean of sensitive study variables using auxiliary information under basic probability sampling designs.

2. Notations and terminology

Consider a finite population $\psi = \{\psi_1, \psi_2, \ldots, \psi_N\}$consisting of Nunits, from which a sample of size has been drawn using simple random sampling without replacement (SRSWOR). Let $y_i$and $x_i$denote the values of the study and auxiliary variables, respectively, with population means $\bar{Y}$and $\bar{X}$. The coefficients of variation for $X$and $Y$are defined as $C_x^2 = \frac{S_x^2}{\bar{X}^2}$and $C_y^2 = S_y^2/\bar{Y}^2$, while the regression coefficient is $b = S_{xy}/S_x^2$, and the correlation coefficient between $X$and $Y$is $\rho = S_{xy}/(S_x S_y)$. The kurtosis coefficients of the study and auxiliary variables are $\phi_{2(y)}$and $\phi_{2(x)}$. Sample variances are given by $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$and $s_x^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ population variances by $S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ and $S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$

To derive the bias and mean squared error (MSE) for the proposed and existing estimators, we define the sampling error terms

$e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_1 = \frac{\bar{x}-\bar{X}}{\bar{X}}$, such that $E(e_0) = E(e_1) = 0$, $E(e_0^2) = (1-f)/n\, C_y^2$, $E(e_1^2) = (1-f)/n\, C_x^2$, and $E(e_0 e_1) = (1-f)/n\, \rho C_x C_y$.

3. Existing Estimators

The following discussion focuses on selected estimators from the existing literature.

1. The finite population mean $\bar{y}$may be estimated using the classical ratio estimator, whose variance is given by

$$\text{Var}\ (\bar{y}_R)= \bar{y}(\bar{\psi}/\bar{x}) \quad (1)$$

$$\text{Var}\ (\bar{y}_R)= \lambda\bar{\Psi}^2[\ C_x^2 + C_y^2 - 2\rho C_x C_y] \quad (2)$$

Where, $\lambda=\frac{1-f}{n}$

2. Robson (1957) introduced the product estimator, which was subsequently modified by

Murthy (1964) and is presented below:

$$\bar{y}_P= \bar{y}(\bar{x}/\bar{\psi}) \quad (3)$$

The expression for the mean squared error is formulated as:

$$(\bar{y}_P)=\lambda\bar{\Psi}^2[\ C_x^2 + C_y^2 + 2\rho C_x C_y]. \quad (4)$$

3. The regression estimator was developed by Watson (1937) is as follows:

$$\bar{y}_{RE}=\bar{y}+\text{b}\ (\bar{\psi} - \bar{x}) \quad (5)$$

Its corresponding MSE is defined as follows:

$$\text{MSE}\ (\bar{y}_{RE}) = \lambda\bar{\Psi}^2 C_y^2(1\text{-}\rho^2). \quad (6)$$

4. The exponential ratio-type estimator, proposed by Bahl and Tuteja (1991), is given as follows:

$$t_{exp}=\bar{y}.\exp\left(\frac{\bar{\psi}-\bar{x}}{\bar{\psi}+\bar{x}}\right) \quad (7)$$

The mean squared error (MSE) for this estimator is given by:

$$\text{MSE}\ (t_{exp})= \lambda\bar{\Psi}^2\left[C_y^2 + \frac{C_x^2}{4} - \rho C_x C_y\right]. \quad (8)$$

5. The following represents the estimator developed by Updhyaya and Singh:

$$\bar{y}_{US}=\left[\bar{y}\frac{(\bar{\psi}\ \beta_{2x}+C_x)}{(\beta_{2x}+C_x)}\right]\frac{(\bar{\psi}\beta_{2x})}{(\beta_{2x}+C_x)} \quad (9)$$

The expression for its MSE is as follows.

$$\text{MSE}\ (\bar{y}_{US}) =\left[(C_y^2 + \left(\frac{(\bar{\psi}\ \beta_{2x})}{(\beta_{2x}+C_x)}\right)^2 C_x^2 - 2\left(\frac{(\bar{\psi}\beta_{2x})}{(\beta_{2x}+C_x)}\right)\rho C_x C_y\right]. \quad (10)$$

6. Kadilar and Cingi (2004) developed new classes of ratio-type estimators employing different transformations of the auxiliary variable. An example of these estimators is:

$$\bar{Y}_{KCi}= [\, \bar{y}+\text{b}\,(\bar{\psi} - \bar{x})]\alpha_i,\ \text{i=1…, 5} \tag{11}$$

$$\alpha_1=\frac{\bar{\psi}}{\bar{x}},\quad \alpha_2=\frac{\bar{\psi}+C_x}{\bar{x}+C_x},\ \alpha_3=\frac{\bar{\psi}+\beta_{2(x)}}{\bar{x}+\beta_{2(x)}},\ \alpha_4=\frac{\bar{\psi}\beta_{2(x)}+C_x}{\bar{x}\beta_{2(x)}+C_x},\ \alpha_5=\frac{\bar{\psi}C_x+\beta_{2(x)}}{\bar{x}\,C_x+\beta_{2(x)}}$$

is the sample regression coefficient, and the values of $\alpha_i$are defined as follows:

The bias and MSE of $\bar{Y}_{KCi}$ (i=1…, 5), to first order of approximation, are given by

$$\text{Bias}\ (\bar{Y}_{KCi}) \approx \lambda\bar{Y}C_x^2\alpha_i^{*2}, \tag{12}$$

$$\text{MSE}\ (\bar{Y}_{KCi}) \approx \lambda\bar{Y}^2\left[\alpha_i^{*2}C_x^2 + C_y^2\ (1-\rho^2)\right]. \tag{13}$$

Where α_1^*=1, $\alpha_2^*=\frac{\bar{X}}{(\bar{X}+C_x)}$, $\alpha_3^*=\frac{\bar{X}}{(\bar{X}+\beta_{2(x)})}$, $\alpha_4^*=\frac{\bar{X}\beta_{2(x)}}{(\bar{X}\beta_{2(x)}+C_x)}$ and $\alpha_5^*=\frac{\bar{X}}{(\bar{X}C_x+\beta_{2(x)})}$

7. The logarithmic ratio-type estimator, proposed by Izunobi and Onyeka (2019), is given as follows:

$$\bar{y}_{IO}= \bar{y}\left[\frac{\ln(\bar{\psi})}{\ln(\bar{x})}\right] \tag{14}$$

Its corresponding MSE can be expressed as follows:

$$\text{MSE}\ (\bar{y}_{IO})= \lambda\bar{\Psi}^2\left[C_x^2\left(\frac{1}{\ln(\bar{\psi})}\right)^2 + C_y^2 - 2\left(\frac{1}{\ln(\bar{\psi})}\right)\rho C_x C_y\right]. \tag{15}$$

8. A novel logarithmic ratio-type estimator, introduced by Zaman and Iftikhar (2023), can be expressed as:

$$\bar{y}_{AS}=\bar{y}\left[\frac{\ln(\beta_{2x}\bar{\psi}+C_x)}{\ln(\beta_{2x}\bar{x}+C_x)}\right]. \tag{16}$$

The terms $\beta_{2x}$and $C_x$denote the kurtosis coefficient and the coefficient of variation of the auxiliary variable.

Its corresponding MSE can be expressed as follows:

$$\text{MSE}\ (\bar{y}_{AS}) = \lambda\bar{\Psi}^2\left[C_y^2 + C_x^2\frac{\theta^2}{k^2} - 2\rho C_x C_y\frac{\theta}{k}\right]. \tag{17}$$

4. **The Proposed logarithmic Estimator**:

Motivated by the work of Zaman et al. (2023), we introduce an improved logarithmic-type estimator, described as follows:

$$t_{AM}^{v}=[k_1\ \bar{y}+k_2(\bar{X}-\bar{x})]\left[\frac{\ln(a\bar{X}+b)}{\ln(a\bar{x}+b)}\right] \tag{18}$$

The values of $k_1$and $k_2$are chosen such that the mean squared error of $t_{AM}^{v}$is minimized.

The estimator t_{AM}^{v} can be represented in terms of sampling errors using a first-order approximation, as follows:

$$t_{AM}^{v}=[k_1\,\bar{Y}\,(1+e_0)+k_2(\bar{X}-\bar{X}(1+e_1))]\left[\frac{\ln(a\bar{X}+b)}{\ln(a(\bar{X}(1+e_1)+b)}\right] \quad (19)$$

After subtracting $\bar{Y}$ we get from both the sides, we get,

$$t_{AM}^{v}-\bar{Y}=\Big[(k_1-1)\,\bar{Y}-\frac{k_1}{k}\,\theta e_1\bar{Y}-\frac{\theta^2}{2k}k_1{e_1}^2\bar{Y}+k_1e_0\bar{Y}-\frac{k_1}{k}\,\theta\bar{Y}e_0e_1-k_2\bar{X}e_1+\frac{k_2}{k}\,\theta e_1^2\bar{X}+\frac{\theta^2}{2k}\,\bar{X}k_2e_1{e_1}^2\Big] \quad (20)$$

Where, $\theta=\frac{a\bar{X}}{a\bar{X}+b}$ and $k=ln(a\bar{X}+b)$

By taking expectation on both the sides and taking the terms of maximum higher order up to degree 2, we get expression for the bias:

$$E(t_{AM}^{v}-\bar{Y})=\mathrm{E}[(k_1\text{-}1)\,\bar{Y}-\frac{k_1}{k}\,\theta e_1\bar{Y}\text{-}\frac{\theta^2}{2k}k_1{e_1}^2\bar{Y}+k_1e_0\bar{Y}-\frac{k_1}{k}\,\theta\bar{Y}e_0e_1-k_2\bar{X}e_1+\frac{k_2}{k}\,\theta e_1^2\bar{X}+\frac{\theta^2}{2k}\bar{X}k_2e_1{e_1}^2] \quad (21)$$

$$\mathrm{Bias}(t_{AM}^{v})=[(k_1-1)\,\bar{Y}-\frac{\theta^2}{2k}k_1\lambda\bar{Y}C_x^2-\frac{k_1}{k}\,\theta\lambda\rho C_xC_y\bar{Y}+\frac{k_2}{k}\,\theta\bar{X}\,\lambda C_x^2] \quad (22)$$

Now, squaring both the sides of $(t_{AM}^{v}-\bar{Y})$ and taking expectation, we get the expression for the MSE:

$$MSE(t_{AM}^{v})=[(k_1-1)^2\bar{Y}^2+\frac{k_1^2}{k^2}\theta^2\bar{Y}^2\lambda C_x^2+k_2^2\bar{X}^2\lambda C_x^2-\frac{2\theta}{k}\bar{X}\bar{Y}k_1k_2\lambda C_x^2+k_1^2\bar{Y}^2\lambda C_y^2-\frac{2k_1^2}{k}\theta\bar{Y}^2\lambda\rho C_xC_y+2k_1k_2\bar{X}\bar{Y}\,\lambda\rho C_xC_y] \quad (23)$$

Where, $\boldsymbol{\lambda}=\left(\frac{1}{n}-\frac{1}{N}\right)$

$$MSE(t_{AM}^{v})=\;[(k_1-1)^2A+k_1^2B+k_1^2F+k_2^2C-k_1k_2D+k_1k_2G] \quad (24)$$

Where,

A=$\bar{Y}^2$, B=$\frac{\theta^2}{k^2}\bar{Y}^2\lambda C_x^2$, C=$\bar{X}^2\lambda C_x^2$, D=$\frac{2\theta}{k}\bar{X}\bar{Y}\lambda C_x^2$, E=A λC_y^2, F=$\frac{2}{k}\theta A\lambda\rho C_xC_y$, G=$2\bar{Y}\lambda\rho C_xC_y$

Now,

By differentiating the above equation with respect to k_1 and k_2 and setting the derivatives to zero, the optimum values of k_1 and k_2 are obtained as follows:

$$k_1^*=\frac{2A}{2A+2(C+E-F)-\frac{(D-G)^2}{2B}}$$ and

$$k_2^*=\frac{\frac{2A}{2A+2(C+E-F)-\frac{(D-G)^2}{2B}}(D-G)}{2B}$$

Using the values k_1^* and k_2^* that minimize MSE, the minimum mean squared error of the proposed logarithmic estimator can be expressed as:

$$Min.\text{MSE}\,(t_{AM}^{v}) = \frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2} \qquad (25)$$

Where, $\boldsymbol{\tau = 2A + 2(C + E - F) - \frac{(D-G)^2}{2B}}$

$\Delta = \boldsymbol{D - G}$

5. Empirical Study

This section utilizes two data sets to compare the performance of the proposed estimator with that of existing ones. The sources and descriptive statistics are presented as follows:

Data set I: Singh and Chaudhary (1986)

x=Area under wheat in 1971

y=Area under wheat in 1974

N=34, n=20, $\bar{\Psi}$=856.4118, $\bar{\psi}$=208.8824, ρ=0.4491, S_y=733.1407, C_y=0.8561, S_x=150.5060, C_x=0.7205, β_{2x}=0.0974

Data set II: Murthy (1967)

y=output for 80 factories in a region

x=fixed capital

N=80, n=20, $\bar{\Psi}$=51.8264, $\bar{\psi}$=11.2646, ρ=0.7941, S_y=18.3569, C_y=0.3542, S_x=8.4062, C_x=0.9484, β_{2x}=2.8664

The MSE values and corresponding PRE of all estimators relative to the traditional ratio estimator are shown in Table 1. The calculation is given by:

PRE=$\frac{MSE(\bar{y}_R)}{MSE(\bar{y}_i)} \times 100$ where i=$\bar{y}_{US}$,$\bar{y}_{kc1}$, $\bar{y}_{kc2}$, $\bar{y}_{kc3}$, $\bar{y}_{kc4}$, $\bar{y}_{kc5}$, $\bar{y}_{kIO}$, $\bar{y}_{AS}$, t_{AM}^{v}

Table 1. MSE values of the estimators along with PRE values.

Estimators	Data set 1		Data set 2	
	MSE	PRE	MSE	PRE
$\bar{y}_R$	27271.90247	100	26.44515	100.000
$\bar{y}_{US}$	10299.06594	264.799	42.081696	373.016
$\bar{y}_{kc1}$	16673.80929	163.561	95.265576	164.772
$\bar{y}_{kc2}$	16620.01056	164.090	81.741208	192.035
$\bar{y}_{kc3}$	16666.50403	163.633	77.482230	202.590
$\bar{y}_{kc4}$	16146.76108	168.900	84.583337	185.582
$\bar{y}_{kc5}$	16663.67288	163.660	76.669039	204.739
$\bar{y}_{IO}$	9775.684135	278.976	43.872832	357.788
$\bar{y}_{AS}$	9203.137592	296.332	4.701784	562.446
t_{AM}^{v}	`8806.5678	309.6771	3.135423	844.892

Table1 shows for that the proposed estimator has minimum the MSE value and higher PRE value as compared to $\bar{y}_{US}$,$\bar{y}_{kc1}$, $\bar{y}_{kc2}$, $\bar{y}_{kc3}$, $\bar{y}_{kc4}$, $\bar{y}_{kc5}$, $\bar{y}_{kIO}$ and $\bar{y}_{AS}$

6. Efficiency Comparisons

The proposed logarithmic-type estimators have been compared with existing competing estimators, and the conditions for their efficiency have also been outlined.

- Comparative analysis of the proposed logarithmic estimator t_{AM}^{ν} with the sample mean $\bar{y}_R$:

MSE ($\bar{y}_R$)- *min*.MSE(t_{AM}^{ν})>0 or,

$$\lambda\bar{\Psi}^2[\, C_x^2 + C_y^2 - 2\rho C_x C_y] - \left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right] > 0. \tag{26}$$

Which shows our proposed logarithmic ratio type estimator performs better than sample mean ($\bar{y}_R$)

- Evaluating the performance of the proposed logarithmic estimator t_{AM}^{ν} in relation to the Updhyaya and Singh estimator $\bar{y}_{US}$:

$$\text{MSE}(\bar{y}_{US}) - \text{min.MSE}(t_{AM}^{\nu}) > 0$$

$$\left[(C_y^2 + \left(\frac{(\bar{\psi}\,\beta_{2x})}{(\beta_{2x}+C_x)}\right)^2 C_x^2 - 2(\frac{(\bar{\psi}\beta_{2x})}{(\beta_{2x}+C_x)})\rho C_x C_y\right] - \left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right] > 0. \tag{27}$$

Which shows our proposed logarithmic ratio type estimator performs better than the Updhyaya and Singh estimator $\bar{y}_{US}$.

- Comparison of the proposed logarithmic estimator t_{AM}^{ν} and the Kadilar and Cingi ($kc1$) estimator:

$$\text{MSE}(\bar{y}_{kc1}) - \text{min.MSE}(t_{AM}^{\nu}) > 0$$

$$\lambda\,\bar{\Psi}^2\,[C_x^2 + C_y^2\,(1-\rho^2)] - \left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right] > 0. \tag{28}$$

Which shows our proposed logarithmic ratio type estimator performs better than the Kadilar and Cingi estimator (kc_1).

- Assessment of the proposed logarithmic estimator t_{AM}^{ν} against the Kadilar and Cingi estimator ($kc2$):

$$\text{MSE}(\bar{y}_{kc2}) - \text{min.MSE}(t_{AM}^{\nu}) > 0$$

$$\lambda\,\bar{\Psi}^2\left[\left(\frac{\bar{\psi}}{\bar{\psi}+C_x}\right)^2 C_x^2 + C_y^2\,(1-\rho^2)\right] - \left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right] > 0. \tag{29}$$

Which shows our proposed logarithmic ratio type estimator performs better than the Kadilar and Cingi estimator (kc2).

- The performance of the proposed logarithmic estimator t_{AM}^{v} is analyzed relative to the Kadilar and Cingi estimator $(kc3)$.

MSE $(\bar{y}_{kc3})$- min.MSE(t_{AM}^{v})>0

$$\lambda\,\bar{\Psi}^2\left[\left(\frac{\bar{\psi}}{\bar{\psi}+\beta_{2x}}\right)^2 C_x^2 \;+\; C_y^2\,(1-\rho^2)\right]-\left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right]>0. \tag{30}$$

Which shows our proposed logarithmic ratio type estimator performs better than the Kadilar and Cingi estimator $(kc3)$.

- An evaluation is made to contrast the proposed logarithmic estimator t_{AM}^{v} with Kadilar and Cingi's $(kc4)$ estimator.

MSE $(\bar{y}_{kc4})$- min.MSE(t_{AM}^{v})>0

$$\lambda\,\bar{\Psi}^2\left[\left(\frac{\bar{\psi}\beta_{2x}}{\bar{\psi}\beta_{2x}+C_x}\right)^2 C_x^2 +\; C_y^2\,(1-\rho^2)\right]-\left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right]>0. \tag{31}$$

Which shows our proposed logarithmic ratio type estimator performs better than the Kadilar and Cingi estimator $(kc4)$.

- The effectiveness of the proposed logarithmic estimator t_{AM}^{v} is analyzed relative to Kadilar and Cingi's $(kc5)$ estimator.

MSE $(\bar{y}_{Kc5})$- min.MSE(t_{AM}^{v})>0

$$\lambda\,\bar{\Psi}^2\left[\left(\frac{\bar{\psi}C_x}{\bar{\psi}\rho C_x+\beta_{2x}}\right)^2 C_x^2 +\; C_y^2\,(1-\rho^2)\right]-\left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right]>0. \tag{32}$$

- The efficiency and accuracy of the proposed logarithmic estimator t_{AM}^{v} are compared with the Izunobi and Onyeka (2019) estimator $\bar{y}_{IO}$.

MSE $(\bar{y}_{IO})$- *min.*MSE(t_{AM}^{v})>0

$$\lambda\bar{\Psi}^2\left[C_x^2\left(\frac{1}{\ln(\bar{\psi})}\right)^2 + C_y^2 - 2\left(\frac{1}{\ln(\bar{\psi})}\right)\rho C_x C_y\right]-\left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right]>0. \tag{33}$$

Which shows our proposed logarithmic ratio type estimator performs better than the

Izunobi and Onyeka (2019) $\bar{y}_{IO}$.

- Comparison of the performance of the proposed logarithmic estimator t_{AM}^{v} and the Zaman and Iftikhar (2023) estimator $\bar{y}_{AS}$:

MSE $(\bar{y}_{AS})$- *min.*MSE(t_{AM}^{v})>0

$$\lambda\bar{\Psi}^2\left[C_y^2 + C_x^2\frac{\theta^2}{k^2} - 2\rho C_x C_y\frac{\theta}{k}\right] - \left[\frac{A(\tau-2A)^2+4A^2(B+F)+\frac{A^2\Delta^2(C-2B)}{B^2}}{\tau^2}\right] > 0. \quad (34)$$

which shows our proposed logarithmic ratio type estimator performs superior than the Zaman and Iftikahar $\bar{y}_{AS}$.

MSE Analysis of the Estimators with the help of Line Chart:

MSE Analysis of the estimators (Data se-1)

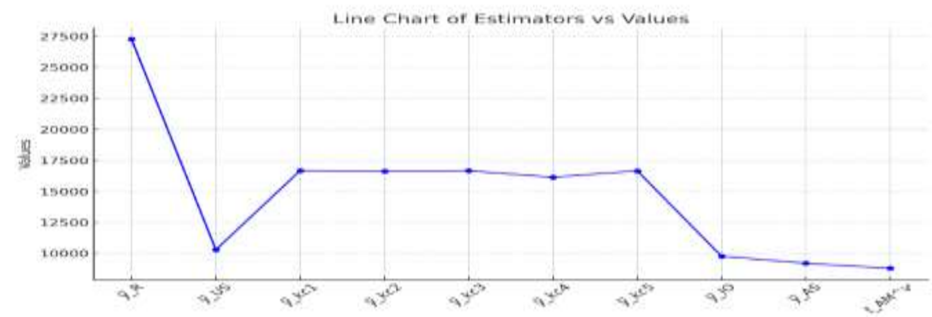


Fig: 1

FInterpretation of the Line Chart:

The chart illustrates how different estimators perform by comparing their values. $\bar{y}_R$ has the largest value, indicating it could be less consistent or accurate. In contrast, t_{AM}^{v}, having the lowest value, likely represents the most accurate or dependable estimator. The other estimators fall somewhere in between, reflecting a range of estimation qualities.

MSE Analysis of the estimators (Data se-2)

Fig:2

Interpretation of the Line Chart:

The chart compares the values across different estimators. $\bar{y}_R$ Stands out with the highest value, suggesting it might be less reliable or more prone to error. Conversely, $\bar{y}_{AS}$ and t_{AM}^{v}, having the lowest values, likely reflect higher accuracy or stability. The other estimators show intermediate results, representing moderate efficiency.

7. Simulation Study

The aim of this simulation study is to compare how the proposed estimator performs in comparison with the established estimator under different distributional conditions. Simulation is an essential tool for evaluating estimator efficiency in a controlled environment, allowing researchers to experiment with different conditions and assess the estimator's adaptability. It also sheds light on how the estimators might perform when real data is either not available or inappropriate. Simulated data drawn from a normal population is used to examine the efficiency of the proposed as well as the existing logarithmic ratio estimators.

For the simulation study, we used the R program to generate a bivariate normal population of size 1000, which was then divided into five sample sizes: $n = 100,200,300,400,500$. Samples were drawn using simple random sampling without replacement, with population parameters $\mu = (20,30)$ and the specified covariance matrix.

$$\sum = \begin{bmatrix} \sigma_y^2 & \sigma_y \sigma_x \rho_{xy} \\ \sigma_y \sigma_x \rho_{xy} & \sigma_x^2 \end{bmatrix}$$

The table below presents the MSE values of the estimator for sample sizes $n = 100,200,300,400,500$, based on 5000 simulations. The MSE was calculated using the following formula:

$$\text{MSE}\,(t^*) = \frac{1}{1000} \sum_i^{1000} (t^* - \bar{Y}_N)^2$$

After utilizing the necessary parameters and the value of correlation coefficient ρ the MSE results are obtained and compiled in the table-2

Table 2: MSE of the Estimator Across Different Sample Sizes from artificially generated data ($n = 100, 200, 300, 400, 500$)

n	$\bar{y}_R$	$\bar{y}_{US}$	$\bar{y}_{kc1}$	$\bar{y}_{kc2}$	$\bar{y}_{kc3}$	$\bar{y}_{kc4}$	$\bar{y}_{kc5}$	$\bar{y}_{IO}$	$\bar{y}_{AS}$	t_{AM}^{v}(prop.)
100	0.108516	0.092691	0.170895	0.148104	0.167198	0.099499	0.165916	0.092534	0.090228	**0.067026**
200	0.047457	0.040786	0074837	0.065011	0.073222	0.043942	0.072641	0.040725	0.039707	**0.023750**
300	0.027663	0.023781	0.043665	0.037948	0.042723	0.025668	0.042382	0.023755	0.023156	**0.022133**
400	0.017739	0.015246	0.028025	0.024357	0.027420	0.016473	0.027200	0.015236	0.014847	**0.001372**
500	0.011815	0.010165	0.018668	0.016231	0.018266	0.010986	0.018119	0.010155	0.009899	**0.000737**

Simulation results from the bivariate normal population is shown in the above table shows that the proposed estimation t_{AM}^{v} performs better as compared to $\bar{y}_{US}$, $\bar{y}_{kc1}$, $\bar{y}_{kc2}$, $\bar{y}_{kc3}$, $\bar{y}_{kc4}$, $\bar{y}_{kc5}$, $\bar{y}_{kIO}$ and $\bar{y}_{AS}$.

8. Results and Discussion

- The analysis shows that t_{AM}^{v} is the most effective estimator in both data sets, achieving the lowest Mean Square Error (MSE) with values of **8806.57** for Data set 1 and **3.13** for Data set 2. Conversely, $\bar{y}_R$ have the highest MSE values of **27271.90** and **26.45**, indicating it is the least improved. The Percent Relative Efficiency (PRE) values further support this conclusion, with t_{AM}^{v} being around **310%** more efficient in Data set 1 and exceeding **845%** efficiency in Data set 2 relative to $\bar{y}_R$. Other estimators such as $\bar{y}_{US}$, $\bar{y}_{IO}$ and the $\bar{y}_{KC}$ series show some gains but still lag behind t_{AM}^{v}. These results highlight the robustness and accuracy of t_{AM}^{v}, especially when the data variability is lower.
- We have made the line chart for showing the MSE values graphically which presents a comparison of different estimators based on their values. The estimator $\bar{y}_R$ shows the highest value, suggesting it may be less accurate or consistent. On the other hand, t_{AM}^{v} and $\bar{y}_{AS}$ which have the lowest values, are likely the most reliable and stable estimators. The remaining estimators fall between these extremes, indicating varying levels of accuracy and efficiency.
- The simulation study indicates that all estimators become more accurate as the sample size increases from 100 to 500. The (t_{AM}^{v} prop.) estimator consistently achieves the smallest error values, suggesting it is the most dependable method. Meanwhile, estimators like $\bar{y}_{kc1}$ and $\bar{y}_{kc3}$ have higher values, implying lower accuracy. The other estimators show performance levels between these extremes and improve as more data is used.

9. Conclusions

In this work, we develop a new logarithmic ratio estimator to estimate the population mean. by incorporating auxiliary information under simple random sampling scheme. Theoretical findings, supported by bias and MSE expressions, along with numerical analysis, showed the proposed estimator demonstrates superior performance compared to existing estimators in terms of both reliability and accuracy. Tests using real data sets revealed that the estimator delivers lower MSE and greater efficiency, making it a strong candidate for population mean estimation and we have performed the simulation study to validate theoretical and empirical results of the suggested logarithmic ratio type estimator from which we also get the result as our proposed estimator have lower MSE than the existing estimators.

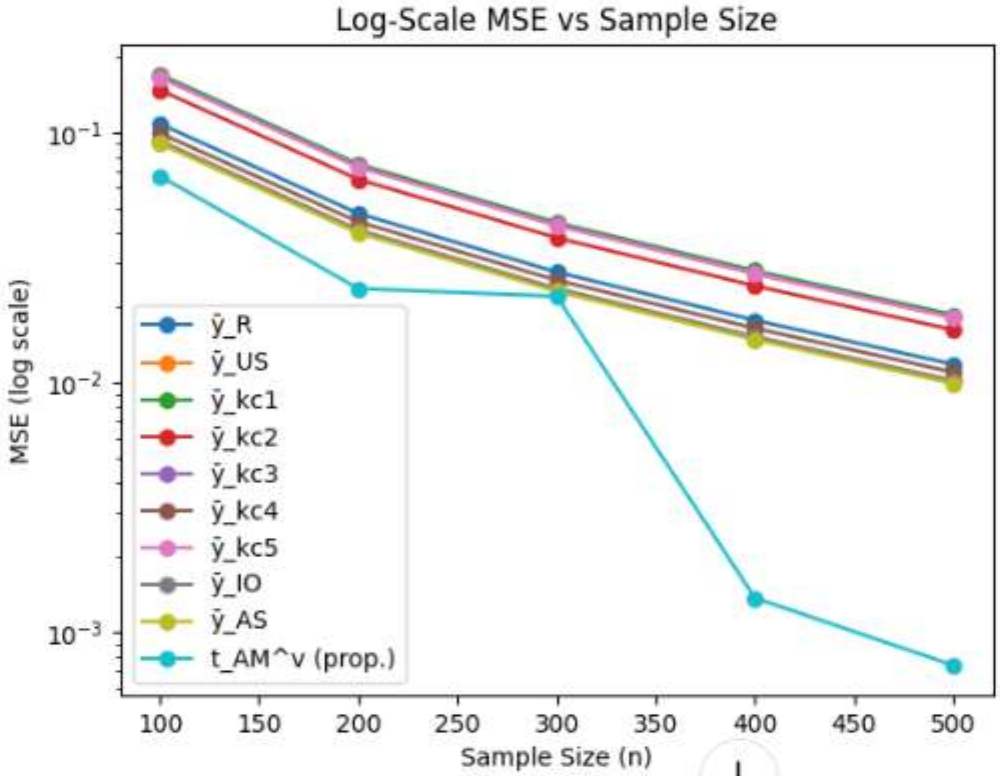
Acknowledgement

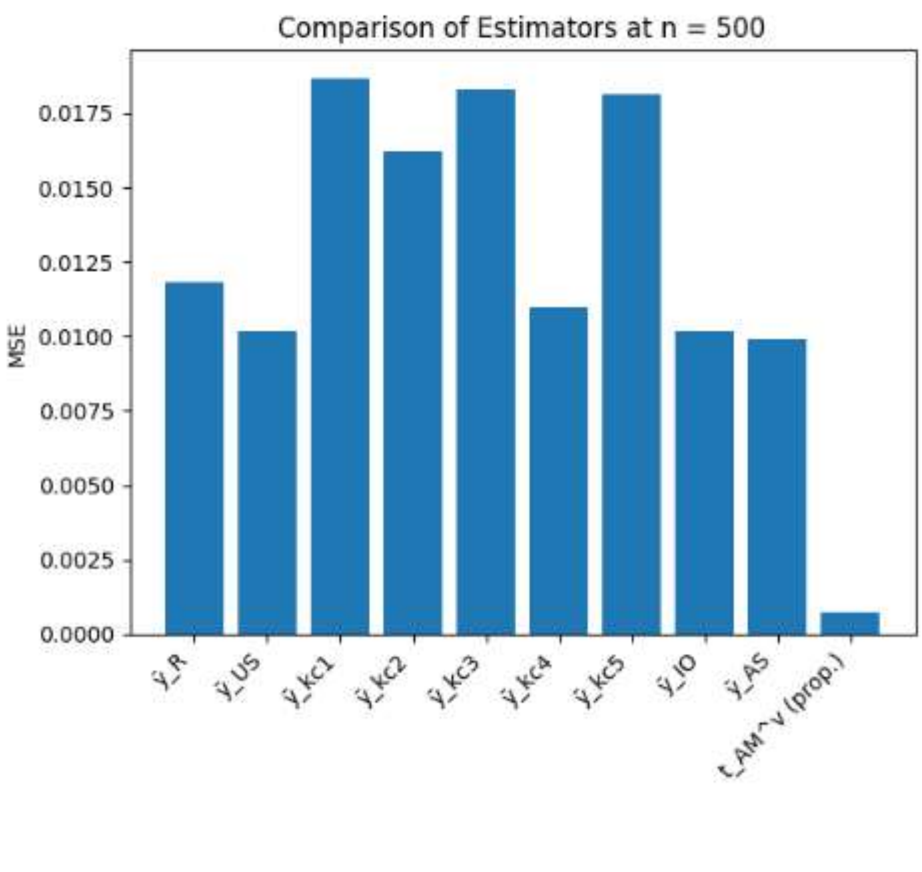
The author gratefully acknowledges the contributions of the editors, reviewers, and the corresponding author, whose advice, evaluations, and guidance were invaluable during the study.

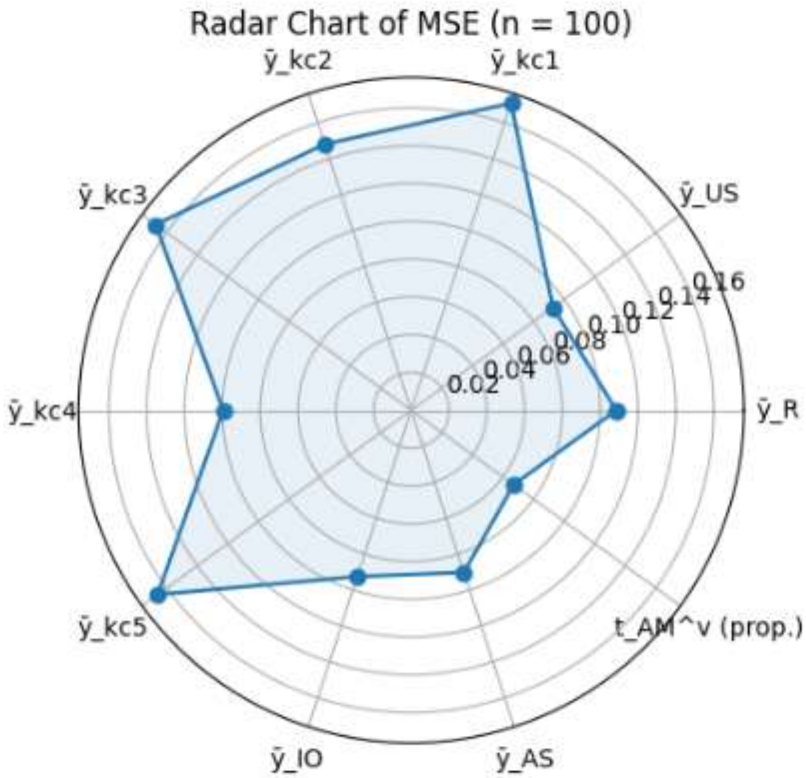
10. References

1. Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. The journal of agricultural science, 30(2), 262-275.
2. Murthy, M. N. (1967). Product method of estimation. Sankhyā: The Indian journal of statistics, series A, 69-74.
3. Srivastava, S. K. (1967). An estimator using auxiliary information in sample surveys. Calcutta Statistical Association Bulletin, 16(2-3), 121-132.
4. Tuteja, R. K., & Bahl, S. (1991). Multivariate product estimators. Calcutta Statistical Association Bulletin, 41(1-4), 109-116.
5. Kadilar, C., & Cingi, H. (2004). Ratio estimators in simple random sampling. Applied mathematics and computation, 151(3), 893-902.
6. Kadilar, C. (2006). New ratio estimators using correlation coefficient.
7. Singh, H. P., & Vishwakarma, G. K. (2009). A general procedure for estimating population mean in successive sampling. Communications in Statistics—Theory and Methods, 38(2), 293-308.

8. Haq, A., & Shabbir, J. (2013). Improved family of ratio estimators in simple and stratified random sampling. Communications in Statistics-Theory and Methods, 42(5), 782-799.
9. Muneer, S., Shabbir, J., & Khalil, A. (2017). Estimation of finite population mean in simple random sampling and stratified random sampling using two auxiliary variables. Communications in Statistics-Theory and Methods, 46(5), 2181-2192.
10. Singh, R., & Solanki, R. S. (2012). Estimation of population mean using coefficient of variation and median of an auxiliary variable. *Statistical Papers*, *53*(2), 263–277.
11. Adejumobi, A., Yunusa, M. A., Erinola, Y. A., & Abubakar, K. (2023). An efficient logarithmic ratio type estimator of finite population mean under simple random sampling. International Journal of Engineering and Applied Physics, 3(2), 700-705.
12. Yunusa, M. A., Audu, A., Musa, N., Beki, D. O., Rashida, A., Bello, A. B., & Hairullahi, M. U. (2021). Logarithmic ratio-type estimator of population coefficient of variation. Asian Journal of Probability and Statistics, 14(2), 13-22.
13. Izunobi, C. H., & Onyeka, A. C. (2019). Logarithmic ratio and product-type estimators of population mean in simple random sampling. International Journal of Advanced Statistics and Probability, 7(2), 47-55.
14. Onyeka, A. C., Ogbumuo, D. T., & Izunobi, C. (2019). Estimation of population mean in stratified random sampling when using auxiliary information in the presence of non-response. Far East Journal of Theoretical Statistics, 55(2), 151-167.
15. Singh, H. P., & Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of an auxiliary character. Statistica, 65(3), 301-313.
16. Zaman, T., & Iftikhar, S. (2023). A New Logarithmic ratio type estimator of population mean for simple random sampling: a simulation study. Journal of Science and Arts, 23(4), 839-848.
17. Cekim, H. O., & Kadilar, C. (2020). Ln-type variance estimators in simple random sampling. Pakistan Journal of Statistics and Operation Research, 16(4), 689-696.
18. Qureshi, M. N., Faizan, Y., Shetty, A., Ahelali, M. H., Hanif, M., & Alamri, O. A. (2024). Ln-type estimators for the estimation of the population mean of a sensitive study variable using auxiliary information. Heliyon, 10(1).
19. Singh, D., & Chaudhary, F. S. (1986). Theory and analysis of sample survey designs.


Log-Scale MSE vs Sample Size
MSE (log scale)
Sample Size (n)
ȳ_R
ȳ_US
ȳ_kc1
ȳ_kc2
ȳ_kc3
ȳ_kc4
ȳ_kc5
ȳ_IO
ȳ_AS
t_AM^v (prop.)


Comparison of Estimators at n = 500
MSE
ȳ_R
ȳ_US
ȳ_kc1
ȳ_kc2
ȳ_kc3
ȳ_kc4
ȳ_kc5
ȳ_IO
ȳ_AS
t_AM^v (prop.)



Log-Scale Line Chart

- Very important when MSE values differ greatly in magnitude
- Makes small MSE differences (especially for the proposed estimator) visible
- Highly recommended for journal-quality figures