Physical Education, Health and Social Sciences

https://journal-of-social-education.org E-ISSN: 2958-5996
P-ISSN: 2958-5988

An Enhanced Robust Type of Variance Estimator of Finite Population Variance

Taj Farin Khan[1], Abdul Samad [2], Fazal Hassan [3]*, Zanib Shabir [4]

[1] Lecturer at Iqra National University Swat Campus Khyber Pakhtunkhwa Pakistan,
[2] Mphil Scholar, Department of Statistics, Abdul Wali Khan University Mardan, Khyber Pakhtunkhwa Pakistan.
[3] Assistant Director ORIC at Iqra National University Swat Campus Khyber Pakhtunkhwa Pakistan,
[4] Mphil Scholar, Department of Mathematics and Statistics, The University of Haripur, Khyber Pakhtunkhwa,
Correspondence: Fazal Hassan, Assistant Director ORIC at Iqra National University Swat Campus, Khyber Pakhtunkhwa, Pakistan. E-mail address: Fazalhassan0349@gmail.com

DOI: https://doi.org/10.63163/jpehss.v3i3.561

Abstract
Background: Sampling is the only approach suitable for obtaining the most accurate estimate of the population parameter under consideration if it is significant and it is time- and money-consuming to conduct observations on each population unit. For a more effective estimation of population variance, several authors have provided a variety of estimators.
Objective: The research aims to find an estimate of the population variance of the study variable that is more effective than the competing estimators.
Materials and Methods: The estimator has been created using data on the tri-mean, population correlation, interquartile range, First quartile of auxiliary variable, Third quartile, Quartile deviation, Population mid-range of auxiliary variable, Downton's Method, Gini's Mean Difference, Percentile of auxiliary variable etc. Up to the first level of approximation, the equations for the mean squared error (MSE) of the proposed estimator have been developed. The suggested estimator has been theoretically compared to the competing population variance estimators.
Results: The MSEs and PREs of the proposed and current estimators are shown in Table 2 using the aforementioned real-world dataset. Given that it has the lowest mean squared error of the competing population variance estimators, it has been determined that the suggested estimator is the best one.
Conclusion: The proposed estimator must be used for the improved estimate of population variance, as it is superior to competing estimators of population variance.

Keywords: Auxiliary variable; Mean squared error; Numerical method; Percentage relative efficiency; Ratio estimator; Simulation

Competing Interests
The authors declare no conflicts of interest.

1 Introduction

In survey sampling, the requirement for additional data has long been recognised as creating accurate estimates of population characteristics such as the mean, median, quartiles, interquartile range, percentile, Population mid-range of auxiliary variables, Downton's Method, Gini's Mean Difference, coefficient of variation, coefficient of correlation and percentage[1] The literature on survey sampling presents a broad range of approaches for keeping additional information. The ratio, product, and regression type estimators are used to utilise the relationship between the study variable and the auxiliary variable[2]. These estimators frequently function better when there is a correlation between the research variable and the auxiliary variable, which increases their precision[3]. When there is a relationship between the study variable and the supplemental data, the rank of the supplementary data is likewise connected with the study variable, and this correlation makes it possible to employ the rank of the supplementary data as a crucial element in improving an estimator's accuracy[4].

Researchers have created ratio and product-type estimators for variance estimation where the population variance of the supplementary data is known in advance. A variance estimate for finite population parameters is considered when dealing with populations likely to be skewed in fields like agriculture, medicine, biology, and industry[5]. Variations may happen in various sectors, including economic, genetic, and environmental research. For instance, to determine where, how,

and when to plant his crop, an agriculturist must know how meteorological variables vary over time and from place to location[6].
Many writers have modified ratio estimators for finite population variance when an auxiliary variable is included in the literature. Authors like Isaki et al.[7] suggested a typical ratio estimator to estimate the population variance. Some essential references regarding variance estimation include Kadilar and Cingi[8]; Subramani and Kumarapandiyan[9]; Subramani and Kumarapandiyan[10]; Khan and Shabbir[11]; Bhat et al.[12]; Maqbool and Javaid[13]; Khalil et al.[14]; D.K. Sharma et al.[15].

2 Notations and Symbols

Y is the study variable associated with the auxiliary variable X, where and $X = \left\{X_1, X_2, ..., X_N\right\}$ are the n sample values. The sample means for the study and the auxiliary variables are $\bar{y}$ and $\bar{x}$, respectively. S_x^2 and S_y^2 are the population mean squares of X and Y, respectively. S_x^2 And S_y^2 represent the mean sample square for the size n randomly selected, without replacement. Y: Study variable, n: sample size, N: population size, X: Auxiliary variable, $\bar{x}$, $\bar{y}$: Auxiliary and study variables sample means, $\bar{x}$, $\bar{y}$: Auxiliary and study variables population means, f : Sampling fraction, p : Correlation coefficient, Cy , Cx : Coefficient of variation of study and auxiliary variable, β1(x): Skewness, Q1: First quartile of auxiliary variable, Q2: Second quartile, Q3: Third quartile, QD: Quartile deviation, TM: Tri-mean, Md: Median of auxiliary variable, MR: Population mid-range of auxiliary variable, G: Gini's Mean Difference

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i,\ \bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i,\ \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,\ S_y^2 = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2,\ S_x^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2,$$

$$MR = \frac{X_{(1)} + X_{(2)}}{2},\ TM = \frac{Q_1 + 2Q_2 + Q_3}{4},\ G_X = \frac{4}{N-1}\sum_{i=1}^{N}\left(\frac{2i - N - 1}{2N}\right)X_i,$$

$$D = \frac{2\sqrt{\pi}}{N(N-1)}\sum_{i=1}^{N}\left(i - \frac{N+1}{2}\right)X_i$$

3 Literature review

The unbiased estimator of the population variance is

$$t_0 = s_y^2$$

$$MSE\left(t_0\right) = \psi S_y^4\left(\lambda_{40} - 1\right) \tag{1}$$

Isaki et al.[7] suggested a typical ratio estimator estimate the population variance S_y^2 of the study variable using known auxiliary information on the population variance S_x^2 of the auxiliary variable, which is denoted as:

$$t_r = s_y^2\left(\frac{S_x^2}{s_x^2}\right)$$

$$MSE\left(t_r\right) = \psi S_y^4\left(\left(\lambda_{40} - 1\right) + \left(\lambda_{04} - 1\right) - 2\left(\lambda_{22} - 1\right)\right) \tag{2}$$

Kadilar and Cingi[8] developed a class of ratio estimators taking into account the coefficient of variation and kurtosis of the auxiliary variable as auxiliary information for improvement in the estimation of population variance, and it is defined as

$$t_{KCi} = s_y^2\left(\frac{S_x^2 + l_i}{s_x^2 + l_i}\right), i=1, 2, 3, 4$$

The estimator's t_{KCi} mean square error is,

$$MSE\left(t_{KCi}\right) = \psi S_y^4\left(\left(\lambda_{40} - 1\right) + L_i^2\left(\lambda_{04} - 1\right) - 2L_i\left(\lambda_{22} - 1\right)\right) \tag{3}$$

Where $l_1 = C_X, l_2 = \beta_{2x}, l_3 = \frac{\beta_{2x}}{C_x}, l_4 = \frac{C_x}{\beta_{2x}}$

Subramani and Kumarapandiyan[9] developed numerous ratio estimators based on quartiles and their use as an auxiliary variable.

$$t_{Z1}=s_y^2\left(\frac{S_x^2+Q_1}{s_x^2+Q_1}\right),\ t_{Z2}=s_y^2\left(\frac{S_x^2+Q_3}{s_x^2+Q_3}\right),\ t_{Z3}=s_y^2\left(\frac{S_x^2+Q_d}{s_x^2+Q_d}\right),$$

Where the first quartile, third quartile, interquartile range, semi-quartile range, and semi-quartile average are denoted by Q1, Q3, and Qd

Subramani and Kumarapandiyan[10] presented the following estimators to consider the information on the deciles of an auxiliary variable.

$$t_{Z4}=s_y^2\left(\frac{S_x^2+D_1}{s_x^2+D_1}\right),\ t_{Z5}=s_y^2\left(\frac{S_x^2+D_2}{s_x^2+D_2}\right),\ t_{Z6}=s_y^2\left(\frac{S_x^2+D_3}{s_x^2+D_3}\right),\ t_{Z7}=s_y^2\left(\frac{S_x^2+D_4}{s_x^2+D_4}\right),$$

$$t_{Z8}=s_y^2\left(\frac{S_x^2+D_5}{s_x^2+D_5}\right),\ t_{Z9}=s_y^2\left(\frac{S_x^2+D_6}{s_x^2+D_6}\right),\ t_{Z10}=s_y^2\left(\frac{S_x^2+D_7}{s_x^2+D_7}\right),\ t_{Z11}=s_y^2\left(\frac{S_x^2+D_8}{s_x^2+D_8}\right),$$

$$t_{Z12}=s_y^2\left(\frac{S_x^2+D_9}{s_x^2+D_9}\right),\ t_{Z13}=s_y^2\left(\frac{S_x^2+D_{10}}{s_x^2+D_{10}}\right),$$

An estimate based on the coefficient of correlation (ρ) and Q3 has been proposed by Khan and Shabbir[11] and is provided below:

$$t_{Z14}=s_y^2\left(\frac{\rho S_x^2+Q_3}{\rho s_x^2+Q_3}\right),$$

The estimator's t_{Zi} mean square error is

$$MSE\left(t_{Zi}\right)=\psi S_y^4\left(\left(\lambda_{40}-1\right)+P_i^2\left(\lambda_{04}-1\right)-2P_i\left(\lambda_{22}-1\right)\right) \tag{4}$$

Where

$$P_1=\left(\frac{S_x^2}{S_x^2+Q_1}\right), P_2=\left(\frac{S_x^2}{S_x^2+Q_3}\right), P_3=\left(\frac{S_x^2}{S_x^2+Q_d}\right), P_4=\left(\frac{S_x^2}{S_x^2+D_1}\right), P_5=\left(\frac{S_x^2}{S_x^2+D_2}\right),$$

$$P_6=\left(\frac{S_x^2}{S_x^2+D_3}\right), P_7=\left(\frac{S_x^2}{S_x^2+D_4}\right), P_8=\left(\frac{S_x^2}{S_x^2+D_5}\right), P_9=\left(\frac{S_x^2}{S_x^2+D_6}\right), P_{10}=\left(\frac{S_x^2}{S_x^2+D_7}\right),$$

$$P_{11}=\left(\frac{S_x^2}{S_x^2+D_8}\right), P_{12}=\left(\frac{S_x^2}{S_x^2+D_9}\right), P_{13}=\left(\frac{S_x^2}{S_x^2+D_{10}}\right),$$

Bhat et al.[12] updated robust estimators for estimating population variance using a linear combination of Downton's technique and deciles of the auxiliary variable to increase the estimate's accuracy and estimate the study's limited population variance.

$$t_{Bi}=s_y^2\left(\frac{S_x^2+b_i}{s_x^2+b_i}\right), i=1, 2, \ldots, 10$$

The mean square error of the estimator t_{Bi} is

$$MSE\left(t_{Bi}\right)=\psi S_y^4\left(\left(\lambda_{40}-1\right)+B_i^2\left(\lambda_{04}-1\right)-2B_i\left(\lambda_{22}-1\right)\right) \tag{5}$$

$$B_i=\left(\frac{\rho S_x^2}{\rho S_x^2+b_i}\right), i=1, 2, \ldots, 10$$

$$b_1=\left(D+D_1\right),\ b_2=\left(D+D_2\right),\ b_3=\left(D+D_3\right),\ b_4=\left(D+D_4\right),\ b_5=\left(D+D_5\right),\ b_6=\left(D+D_6\right)$$

$$,\ b_7=\left(D+D_7\right),\ b_8=\left(D+D_8\right),\ b_9=\left(D+D_9\right),\ b_{10}=\left(D+D_{10}\right).$$

Maqbool and Javaid's[13] estimator utilizing information on the tri-mean and inter-quartile range of auxiliary variables has been developed.

$$t_m=s_y^2\left(\frac{S_x^2+C_xS_x}{s_x^2+C_xS_x}\right),$$

$$MSE(t_m) = \psi S_y^4\left((\lambda_{40}-1)+M_1^2(\lambda_{04}-1)-2M_1(\lambda_{22}-1)\right) \quad (6)$$

Where $M_1 = \left(\frac{S_x^2}{S_x^2 + C_x S_x}\right)$

Khalil et al.[14] updated estimators for estimating population variance using a coefficient of variation (CV), standard deviation, mean, and median of auxiliary variables to increase the estimate's accuracy and estimate the study's limited population variance.

$$t_{K1} = s_y^2\left(\frac{S_x^2 + C_x\bar{X}}{s_x^2 + C_x\bar{X}}\right),\ t_{K2} = s_y^2\left(\frac{S_x^2 + (TM + Q_3)}{s_x^2 + (TM + Q_3)}\right),$$

The general mean square error of the estimator t_{K1} t_{K2} is

$$MSE(t_{Ki}) = \psi S_y^4\left((\lambda_{40}-1)+R_i^2(\lambda_{04}-1)-2R_i(\lambda_{22}-1)\right)\ i=1,2 \quad (7)$$

Where

$$R_1 = \left(\frac{S_x^2}{S_x^2 + C_x\bar{X}}\right), R_2 = \left(\frac{S_x^2}{S_x^2 + (TM + Q_3)}\right)$$

D.K. Sharma et al.[15] updated a Searls ratio type estimator for the primary variable using the available information on the tri-mean and the third quartile of the auxiliary variable for an enhanced population variance estimation.

$$t_{DK} = Ks_y^2\left(\frac{S_x^2 + (TM + Q_3)}{s_x^2 + (TM + Q_3)}\right),$$

The MSE of the estimator t_{DK} is

$$MSE(t_{DK}) = S_y^4\left(K^2\gamma(\lambda_{40}-1)+(3K^2-2K)R_2^2\gamma(\lambda_{04}-1)-2(2K^2-K)R_2\gamma(\lambda_{22}-1)+(K-1)^2\right) \quad (8)$$

The MSE of the suggested estimator is obtained for the optimum value of κ as,

$$K = \frac{A}{B}$$

Where

$$A = 1 + R_2^2\gamma(\lambda_{04}-1) - R_2\gamma(\lambda_{22}-1)$$

And

$$B = 1 + \gamma(\lambda_{40}-1) + 3R_2^2\gamma(\lambda_{04}-1) - 4R_2\gamma(\lambda_{22}-1)$$

$$MSE_{min}(t_{DK}) = S_y^4\left(1 - \frac{A^2}{B}\right) \quad (9)$$

4 Proposed Estimators

This study is motivated by Searls[16], who demonstrated that an improved estimator could be obtained by using some constant multiple of the sample mean as an estimator of Y, and this constant is obtained by minimizing the MSE of the suggested estimator and Yunusa et al.[17], and to improve estimation, we propose the following estimator as,

$$T_{S1} = \zeta s_y^2\left(\frac{S_x^2 + (MR + Q_3)}{s_x^2 + (MR + Q_3)}\right)$$

$$T_{S2} = \zeta s_y^2\left(\frac{S_x^2 + (MR + D_4D_1)}{s_x^2 + (MR + D_4D_1)}\right)$$

$$T_{M1} = \beta s_y^2\left(\frac{C_x S_x^2 + (MR + D\rho)}{C_x s_x^2 + (MR + D\rho)}\right)$$

$$T_{M2} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_1 \rho\right)}{C_x s_x^2 + \left(MR + D_1 \rho\right)} \right)$$

$$T_{M3} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_2 \rho\right)}{C_x s_x^2 + \left(MR + D_2 \rho\right)} \right)$$

$$T_{M4} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_3 \rho\right)}{C_x s_x^2 + \left(MR + D_3 \rho\right)} \right)$$

$$T_{M5} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_4 \rho\right)}{C_x s_x^2 + \left(MR + D_4 \rho\right)} \right)$$

$$T_{M6} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_5 \rho\right)}{C_x s_x^2 + \left(MR + D_5 \rho\right)} \right)$$

$$T_{M7} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_6 \rho\right)}{C_x s_x^2 + \left(MR + D_6 \rho\right)} \right)$$

$$T_{M8} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + D_7 \rho\right)}{C_x s_x^2 + \left(MR + D_7 \rho\right)} \right)$$

$$T_{M9} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_5 \rho\right)}{C_x s_x^2 + \left(MR + P_5 \rho\right)} \right)$$

$$T_{M10} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{10} \rho\right)}{C_x s_x^2 + \left(MR + P_{10} \rho\right)} \right)$$

$$T_{M11} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{15} \rho\right)}{C_x s_x^2 + \left(MR + P_{15} \rho\right)} \right)$$

$$T_{M12} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{20} \rho\right)}{C_x s_x^2 + \left(MR + P_{20} \rho\right)} \right)$$

$$T_{M13} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{25} \rho\right)}{C_x s_x^2 + \left(MR + P_{25} \rho\right)} \right)$$

$$T_{M14} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{30} \rho\right)}{C_x s_x^2 + \left(MR + P_{30} \rho\right)} \right)$$

$$T_{M15} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{35} \rho\right)}{C_x s_x^2 + \left(MR + P_{35} \rho\right)} \right)$$

$$T_{M16} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + P_{40} \rho\right)}{C_x s_x^2 + \left(MR + P_{40} \rho\right)} \right)$$

$$T_{M17} = \beta s_y^2 \left(\frac{C_x S_x^2 + \left(MR + Q_1 \rho\right)}{C_x s_x^2 + \left(MR + Q_1 \rho\right)} \right)$$

$$T_{M18} = \beta s_y^2 \left(\frac{C_x S_x^2 + (MR + Q_2\rho)}{C_x s_x^2 + (MR + Q_2\rho)} \right)$$

The suggested estimators can be expressed generally as follows,

$$T_{Si} = \zeta_i s_y^2 \left(\frac{S_x^2 + z_i}{s_x^2 + z_i} \right), \text{ i=1, 2.}$$

$$T_{Mh} = \beta_h s_y^2 \left(\frac{C_x S_x^2 + F_h}{C_x s_x^2 + F_h} \right), \text{ h=1, 2, ...,18}$$

Properties (MSE) of the proposed estimators

The following definitions are used to calculate the bias and MSE of the suggested estimators. e's sampling errors as

$e_0 = \frac{s_y^2 - S_y^2}{S_y^2}$ and $e_1 = \frac{s_x^2 - S_x^2}{S_x^2}$ such that $s_y^2 = S_y^2(1+e_0)$ and $s_x^2 = S_x^2(1+e_1)$, also

$E(e_0^2) = \Psi(\lambda_{40} - 1)$, $E(e_1^2) = \Psi(\lambda_{04} - 1)$ and $E(e_0 e_1) = \Psi(\lambda_{22} - 1)$,

The first generalized proposed estimator

$$T_{Si} = \zeta_i s_y^2 \left(\frac{S_x^2 + Z_i}{s_x^2 + Z_i} \right) \tag{10}$$

Where ζ_i is constant

Expressing (10) in error terms, we have

$$T_{Si} = \zeta_i S_y^2 (1+e_0) \left(\frac{S_x^2 + Z_i}{S_x^2 (1+e_1) + Z_i} \right)$$

$$T_{Si} = \zeta_i S_y^2 (1+e_0) \left(\frac{S_x^2 + Z_i}{S_x^2 + Z_i + S_x^2 e_1} \right)$$

By extending the term in the equation mentioned above, simplifying it, and bringing the terms up to approximately order one, we obtain

$$T_{Si} = \zeta_i S_y^2 (1+e_0)(1+\varphi_i e_1)^{-1}$$

suppose

$$\varphi_i = \frac{S_x^2}{S_x^2 + Z_i}, \text{ i=1, 2.}$$

Where

$$Z_1 = (MR + Q_3),\ Z_2 = (MR + D_4 D_1)$$

$$T_{Si} = \zeta_i S_y^2 (1+e_0)(1 - \varphi_i e_1 + \varphi_i^2 e_1^2)$$

S_y^2 is subtracted from both sides of the equation above, giving us,

$$T_{Si} - S_y^2 = \zeta_i S_y^2 (1+e_0)(1 - \varphi_i e_1 + \varphi_i^2 e_1^2) - S_y^2 \tag{11}$$

By simplifying equation (11) and solving for the MSE of T_{Si}, we may obtain the MSE by selecting an expectation and entering values for different expectations in the form of,

$$MSE(T_{Si}) = S_y^4 \left(\begin{array}{l} \zeta_i^2 \psi(\lambda_{40} - 1) + (3\zeta_i^2 - 2\zeta_i)\varphi_i^2 \psi(\lambda_{04} - 1) - 2(2\zeta_i^2 - \zeta_i)\varphi_i \psi(\lambda_{22} - 1) + \\ (\zeta_i - 1)^2 \end{array} \right) \tag{12}$$

The optimal value of ζ_i the proposed estimator's MSE is as follows:

$$\zeta_i = \left(\frac{c_i}{d_i}\right) \text{ i=1, 2.} \tag{13}$$

where

$c_i = 1 + \varphi_i^2\left(\lambda_{04} - 1\right) - \varphi_i \psi\left(\lambda_{22} - 1\right)$ i=1, 2.

and

$d_i = 1 + \psi\left(\lambda_{40} - 1\right) + 3\varphi_i^2 \psi\left(\lambda_{04} - 1\right) - 4\varphi_i \psi\left(\lambda_{22} - 1\right)$ i=1, 2.

The least value of the MSE of T_{Si}, for the optimal value of ζ_i in (13), is:

$$MSE_{min}\left(T_{Si}\right) = S_y^4\left(1 - \frac{c_i^2}{d_i}\right) \text{ i=1, 2.} \tag{14}$$

Second generalized estimator

$$T_{Mh} = \beta_h s_y^2\left(\frac{C_x S_x^2 + F_h}{C_x s_x^2 + F_h}\right) \tag{15}$$

where β_h is constant

Expressing (15) in error terms, we have

$$T_{Mh} = \beta_h S_y^2\left(1 + e_0\right)\left(\frac{C_x S_x^2 + F_h}{C_x S_x^2\left(1 + e_1\right) + F_h}\right)$$

By extending the term in the equation mentioned above, simplifying it, and bringing the terms up to approximately order one, we obtain,

$$T_{Mh} = \beta_h S_y^2\left(1 + e_0\right)\left(1 + \tau_h e_1\right)^{-1}$$

suppose $\tau_h = \dfrac{C_x S_x^2}{C_x S_x^2 + F_h}$, h=1, 2, …,18

Where

$F_1 = MR + D\rho$, $F_2 = MR + D_1\rho$, $F_3 = MR + D_2\rho$, $F_4 = MR + D_3\rho$, $F_5 = MR + D_4\rho$, $F_6 = MR + D_5\rho$, $F_7 = MR + D_6\rho$, $F_8 = MR + D_7\rho$, $F_9 = MR + P_5\rho$, $F_{10} = MR + P_{10}\rho$, $F_{11} = MR + P_{15}\rho$, $F_{12} = MR + P_{20}\rho$, $F_{13} = MR + P_{25}\rho$, $F_{14} = MR + P_{30}\rho$, $F_{15} = MR + P_{35}\rho$, $F_{16} = MR + P_{40}\rho$, $F_{17} = MR + Q_1\rho$, $F_{18} = MR + Q_2\rho$.

$$T_{Mh} = \beta_h S_y^2\left(1 + e_0\right)\left(1 - \tau_h e_1 + \tau_h^2 e_1^2\right)$$

S_y^2 is subtracted from both sides of the equation above, giving us,

$$T_{Mh} - S_y^2 = \beta_h S_y^2\left(1 + e_0 - \tau_h e_1 - \tau_h e_1 e_0 + \tau_h^2 e_1^2\right) - S_y^2 \tag{16}$$

By simplifying equation (16) and solving for the MSE of T_{Mh}, we may obtain the MSE by selecting an expectation and entering values for different expectations in the form of,

$$MSE\left(T_{Mh}\right) = S_y^4\left(\begin{array}{l}\beta_h^2\psi\left(\lambda_{40} - 1\right) + \left(3\beta_h^2 - 2\beta_h\right)\tau_h^2\psi\left(\lambda_{04} - 1\right) - 2\left(2\beta_h^2 - \beta_h\right)\tau_h\psi\left(\lambda_{22} - 1\right) + \\ \left(\beta_h - 1\right)^2\end{array}\right) \tag{17}$$

For the optimal value of β_h the proposed estimator's MSE is as follows:

$$\beta_h = \left(\frac{y_h}{z_h}\right) \text{ h=1, 2, …,18} \tag{18}$$

where

$y_h = 1 + \tau_h^2\left(\lambda_{04} - 1\right) - \tau_h \psi\left(\lambda_{22} - 1\right)$ h=1, 2, …,18

and

$$z_h = 1+\psi\left(\lambda_{40}-1\right)+3\tau_h^2\psi\left(\lambda_{04}-1\right)-4\tau_h\psi\left(\lambda_{22}-1\right) \quad h=1, 2, \ldots,18$$

The least value of the MSE of T_{Mh}, for the optimal value of β_h in (18), is:

$$MSE_{min}\left(T_{Mh}\right)=S_y^4\left(1-\frac{y_h^2}{z_h}\right) \quad h=1, 2, \ldots,18$$

5 Efficiency Comparisons

The proposed estimators T_{Si} and T_{Mh} are more efficient than the existing estimators if the following conditions are satisfied.

The proposed estimator outperforms the sample variance if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_0\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\lambda_{40}-1\right)$$

The proposed estimator performs better than Isaki et al.[7] estimator if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_r\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+\left(\lambda_{04}-1\right)-2\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms Kadilar and Cingi[8] estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_{KCi}\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+L_i^2\left(\lambda_{04}-1\right)-2L_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms Subramani and Kumarapandiyan[9], Subramani and Kumarapandiyan[10], and Khan and Shabbir[11] estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_{Zi}\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+P_i^2\left(\lambda_{04}-1\right)-2P_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Bhat et al. (2018) estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_{Bi}\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+B_i^2\left(\lambda_{04}-1\right)-2B_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Maqbool and Javaid (2017) estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_m\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+M_1^2\left(\lambda_{04}-1\right)-2M_1\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Khalil et al. (2018) estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE\left(t_{Ki}\right)$$

$$1-\frac{c_i^2}{d_i}<\psi\left(\left(\lambda_{40}-1\right)+R_i^2\left(\lambda_{04}-1\right)-2R_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than D.K. Sharma et al. (2022) estimators if,

$$MSE_{min}\left(T_{Si}\right)<MSE_{min}\left(t_{DK}\right)$$

$$1-\frac{c_i^2}{d_i}<1-\frac{A^2}{B}$$

The proposed estimator outperforms the sample variance if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_0\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\lambda_{40}-1\right)$$

The proposed estimator outperforms than Isaki et al.[7] estimator if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_r\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+\left(\lambda_{04}-1\right)-2\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Kadilar and Cingi[8] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_{KCi}\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+L_i^2\left(\lambda_{04}-1\right)-2L_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Subramani and Kumarapandiyan[9], Subramani and Kumarapandiyan[10], Khan and Shabbir[11] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_{Zi}\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+P_i^2\left(\lambda_{04}-1\right)-2P_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Bhat et al.[12] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_{Bi}\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+B_i^2\left(\lambda_{04}-1\right)-2B_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Maqbool and Javaid[13] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_m\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+M_1^2\left(\lambda_{04}-1\right)-2M_1\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than Khalil et al.[14] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE\left(t_{Ki}\right)$$

$$1-\frac{y_h^2}{z_h} < \psi\left(\left(\lambda_{40}-1\right)+R_i^2\left(\lambda_{04}-1\right)-2R_i\left(\lambda_{22}-1\right)\right)$$

The suggested estimator outperforms than D.K. Sharma et al.[15] estimators if,

$$MSE_{min}\left(T_{Mh}\right) < MSE_{min}\left(t_{DK}\right)$$

$$1-\frac{y_h^2}{z_h} < 1-\frac{A^2}{B}$$

6 Empirical Study

The efficiency criteria of proposed estimators over other competing estimators are verified in this section. As a result, we looked at the population provided by Yunusa et al.[17]. Yunusa et al.[17] used this data set from Murthy[18] on page 228 with fixed capital as X and output from 80 factories as Y. The MSE and PREs of the introduced and competing estimators have been quantitatively estimated. This population's parameters are shown in Table 1.

Table 1: Parameters of the population in Yunusa et al.[17]

N=80	n=20	S_x=8.4542	S_y=18.3569	C_y=0.3542	C_x=0.7507
$\bar{Y}$ = 51.8264	$\bar{X}$ =11.2624	λ_{40}=2.2667	λ_{04}=2.8664	ρ =0.9413	λ_{22}=2.2209
Q1=9.318	Q2=7.575	Q3=16.975	D=8.0138	tm=9.318	MR=17.955
D1=3.6	D2=4.6	D3=5.9	D4=6.7	D5=7.5	D6=8.5
D7=14.8	D8=18.1	D9=25	D10=34.8	P5=4.35	P10=5.9

P15=6.63	P20=7.45	P25=7.8	P30=8.7	P35=11.6	P40=15.3
P50=17.2	P55=19.3	P60=21.7	P65=23.55	P70=24.98	P75=25
P80=26.95	P85=27.8	P90=29.7	P95=30	P99=34.85	

Table 2: MSE and PRE of Existing and Proposed Estimators

Estimators	MSE	PRE
t_0	5393.89	100
t_r	2943.71	183.23
t_{KC1}	2887.45	186.80
t_{Z1}	2415.58	223.29
t_{Z2}	2181.58	247.24
t_{Z3}	2813.93	191.68
t_{Z4}	2698.47	199.88
t_{Z5}	2640.48	204.27
t_{Z6}	2570.87	209.80
t_{Z7}	2531.09	213.10
t_{Z8}	2493.48	216.31
t_{Z9}	2449.38	220.21
t_{Z10}	2234.73	241.36
t_{Z11}	2157.56	249.99
t_{Z12}	2052.96	262.73
t_{Z13}	1995.79	270.26
t_{Z14}	2158.91	249.91
t_{B1}	2330.77	231.42
t_{B2}	2298.09	234.71
t_{B3}	2259.14	238.75
t_{B4}	2237.03	241.11
t_{B5}	2216.26	243.37
t_{B6}	2192.08	246.06
t_{B7}	2079.00	259.44
t_{B8}	2041.86	264.66
t_{B9}	2000.22	269.66
t_{B10}	1999.59	269.74
t_m	2548.31	211.66
t_{K1}	2396.48	225.07
t_{K2}	2040.20	264.37
t_{DK}	1987.08	271.44
T_{S1}	1957.806	275.506
T_{S2}	1967.446	274.157
T_{M1}	1958.535	275.404
T_{M2}	1974.542	273.171

T_{M3}	1968.986	273.942
T_{M4}	1963.530	274.942
T_{M5}	1961.101	275.044
T_{M6}	1959.335	275.292
T_{M7}	1958.011	275.478
T_{M8}	1968.662	273.987
T_{M9}	1970.261	273.987
T_{M10}	1963.530	274.703
T_{M11}	1961.286	275.018
T_{M12}	1959.426	275.279
T_{M13}	1958.837	275.361
T_{M14}	1959.858	275.499
T_{M15}	1959.514	275.266
T_{M16}	1970.689	273.705
T_{M17}	1957.616	275.533
T_{M18}	1959.202	275.310

According to Table 2, the proposed estimator's mean square error is the lowest of the estimators considered in this study. Given that the recommended estimator has a more significant percentage relative efficiency, this suggests that the suggested estimator has been improved.

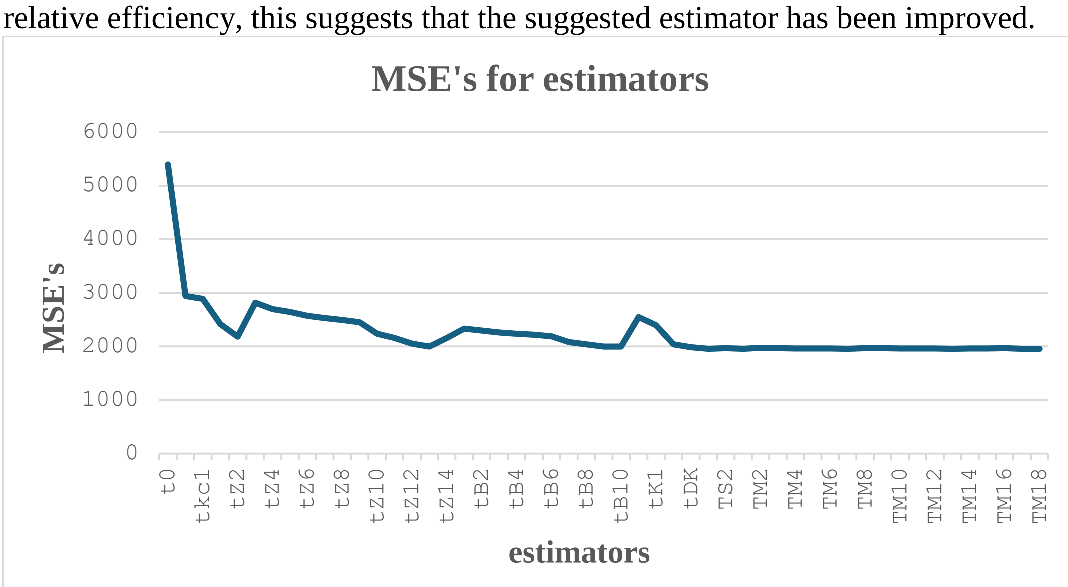

Figure1. MSE for the existing vs proposed estimators

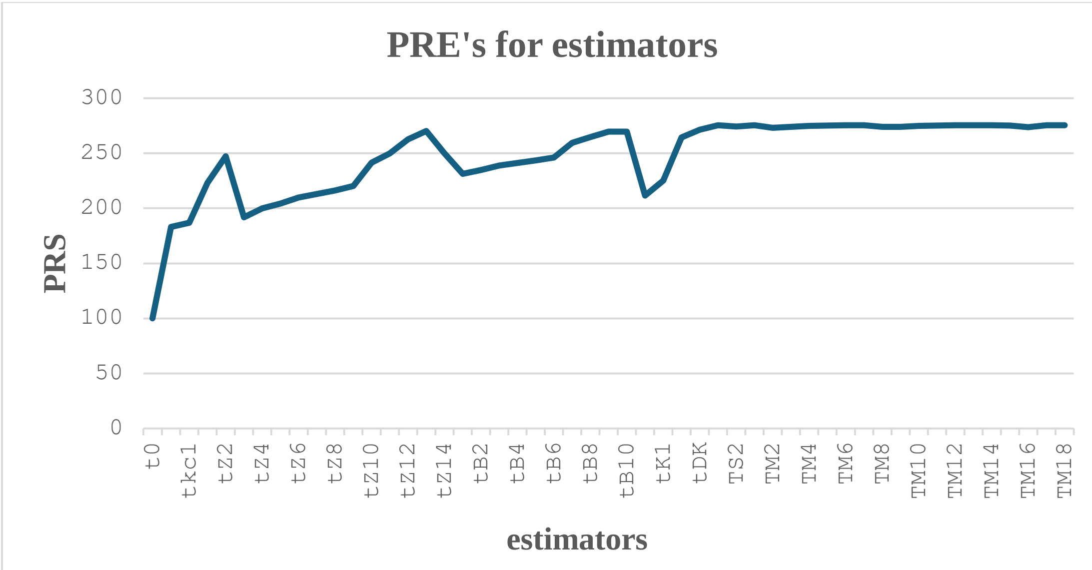

Figure2. PRE for the existing vs proposed estimators

Mean Squared Error (MSE) is the mean or Average Square of the difference between actual or parametric and estimated values. MSE is widely used to check the efficiency of am estimators in statistics, particularly in a sample survey. This study also used the MSE measure to validate proposed estimators. Figure 1 is the plot of MSE for the existing and proposed estimators. The values of the MSE are plotted on the y-axis, where the x-axis represents the various estimators. The plot clearly shows that the MSE for the proposed estimators are lower than the existing estimators. Therefore, in terms of MSE measures, the proposed estimators are better than the existing estimators.

In Figure 2, we have plotted the existing and proposed estimators' relative percentage efficiency. The relative percentage efficiency is used to check the performance of the estimators in percentage figures. In the plot, the y-axis shows the values of relative MSE in percentage while the x-axis shows the estimators. The figure shows that the percentage relative efficiency for the proposed estimators is higher than the existing estimators. Based on relative efficiency, we conclude that the proposed estimators are more excellent than the existing ones.

7 Conclusions

In a sample survey, parameter value estimate is essential for modelling various real-world issues when obtaining actual data is challenging. Numerous scholars specify adaptable improvements to the current estimators in this situation to get a better estimate for unknown population parameters. Recently, research has focused on altering the current estimators by utilizing auxiliary data to achieve such improvements. Ratio, product, and regression-type estimators are well-covered in the literature. This research study's goal is to create flexible variance estimators under simple random sampling by employing an auxiliary variable, intending to improve their efficacy compared to the current estimators.

We could construct different ratio-type variance estimators by employing the auxiliary information and the simple random sampling method used in this work. The auxiliary variable includes population Mean, Quartiles, tri-mean Correlation, etc. The suggested estimator is more effective than the other existing estimators taken into consideration in the study, according to the findings of the numerical analysis. Therefore, adopting the new proposed family of estimators for improved outcomes is advised when calculating population variance under a simple random sampling method.

References

Solanki, R.S., Singh, H.P. and Rathour, A., 2012. An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. International Scholarly Research Notices, 2012.

Singh, R., Chauhan, P., Smarandache, F. and Sawan, N., 2007. Auxiliary information and a priori values in construction of improved estimators. Infinite Study.

Dawud Adebayo, A. and Peter I, O., 2013. Effect of correlation level on the use of auxiliary variable in double sampling for regression estimation. Open Journal of Statistics, 2013.

Onyango, R., Oduor, B. and Odundo, F., 2022. Mean Estimation of a Sensitive Variable under Nonresponse Using Three-Stage RRT Model in Stratified Two-Phase Sampling. Journal of Probability and Statistics, 2022.

Ahmad, S., Hussain, S., Shabbir, J., Zahid, E., Aamir, M. and Onyango, R., 2022. Improved Estimation of Finite Population Variance Using Dual Supplementary Information under Stratified Random Sampling. Mathematical Problems in Engineering, 2022.

Raza, A., Razzaq, A., Mehmood, S.S., Zou, X., Zhang, X., Lv, Y. and Xu, J., 2019. Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. Plants, 8(2), p.34.

Isaki, C.T., 1983. Variance estimation using auxiliary information. Journal of the American Statistical Association, 78(381), pp.117-123.

KADILAR, C. and CİNGİ, H., 2006. Improvement in variance estimation using auxiliary information. Hacettepe Journal of mathematics and Statistics, 35(1), pp.111-115.

Subramani, J. and Kumarapandiyan, G., 2012. Estimation of population mean using co-efficient of variation and median of an auxiliary variable. International Journal of Probability and Statistics, 1(4), pp.111-118.

Subramani, J. and Kumarapandiyan, G., 2012. Estimation of variance using deciles of an auxiliary variable. In Proceedings of International Conference on Frontiers of Statistics and Its Applications (Vol. 33, pp. 143-149).

Khan, M. and Shabbir, J., 2013. A RatioType Estimator for the Estimation of Population Variance using Quartiles of an Auxiliary Variable. Journal of Statistics Applications & Probability, 2(3), p.319.

Bhat, M., Raja, T., Maqbool, S., Sofi, N., Rauf, A., Baba, S. and Shah, I., 2018. Robust Estimators for Estimation of Population Variance Using Linear Combination of Downton's Method and Deciles as Auxiliary Information. Advances in Research, 15(2), pp.1-7.

Maqbool, S. and Javaid, S., 2017. Variance estimation using linear combination of tri-mean and quartile average. American Journal of Biological and Environmental Statistics, 3(1), pp.5-9.

Khalil, M., Ali, M., Shahzad, U., Hanif, M. and Jamal, N., 2018. Improved Estimators of Population Variance Using Measure of Dispersion of Auxiliary Variable. Oriental Journal of Physical Sciences, 3(1), pp.33-39.

Sharma, D.K., Yadav, S.K. and Sharma, H., 2022. Improvement in estimation of population variance utilising known auxiliary parameters for a decision-making model. International Journal of Mathematical Modelling and Numerical Optimisation, 12(1), pp.15-28.

Searls, D.T., 1964. The utilization of a known coefficient of variation in the estimation procedure. Journal of the American Statistical Association, 59(308), pp.1225-1226.

YUNUSA, M.A., AUDU, A. and ADEJUMOBI, A., 2022. Oriental Journal of Physical Sciences.

Murthy, M.N., 1967. Sampling theory and methods. Sampling theory and methods.