Enhanced Population Variance Estimation Using Auxiliary Information in Simple Random Sampling

Nirupama Sahoo[1] and Saradaprasan Jena[2]

[1]Assistant Professor, School of Statistics, Gangadhar Meher University, AmrutaVihar, Sambalpur, Odisha, India

[2]Ph.D. Research Scholar, School of Statistics, Gangadhar Meher University, AmrutaVihar, Sambalpur, Odisha, India

[1]ORCID: https://orcid.org/0000-0002-3244-3557

Abstract: In this study, we propose an estimator for the population variance of a study variable under simple random sampling without replacement (SRSWOR) scheme. The bias and mean squared error (MSE) expressions of the proposed estimator are derived up to the first order of approximation**.** The characterizing constant is optimized to achieve the minimum mean squared error, ensuring the estimator's efficiency. A theoretical comparison has also been conducted to show the proposed estimators in comparison to the other existing estimators. These efficiencies are validated on a number of real natural populations and simulated population.

Key words: Study variable, Auxiliary variable, Estimator, Bias, Mean Squared Error, Simulation, Percentage relative efficiency

1. Introduction

Sampling theory concisely prioritizes the computation of the relevant population parameter, but gathering data on each unit of the population is expensive and time intensive when population is large. Therefore, sampling from the targeted population is an alternative to it. The sample data is used to estimate the parameter in question using an estimator. The sample variance serves as the most suitable estimator for the population variance of the study variable, as it provides the most relevant statistics for the parameter in discussion. A significant limitation of the sample variance estimator is its very high sampling variance. We propose a biased estimator of the population variance that achieves a smaller mean squared error (MSE) compared to the sampling variance of the conventional sample variance estimator**.** This objective is accomplished through the appropriate use of an auxiliary variable that shows a strong positive or negative correlation with the study variable. Estimators incorporating auxiliary information are generally more efficient than those based solely on the sample variance, due to their effective use of additional information. Ratio-type estimators based on the ratio method of estimation are employed to enhance the estimation of the population variance σ_y^2 of the primary variable Y**,** particularly when the auxiliary variable $\boldsymbol{X}$ and the study variable Y exhibit a strong positive correlation, and the regression line of Y on $\boldsymbol{X}$ passes through the origin**.**When the variables Y and $\boldsymbol{X}$ exhibit a strong negative correlation, and the regression line of Y on $\boldsymbol{X}$ passes through the origin, alternative estimators such as product-type estimators are generally more suitable for estimating the population parameters.Conversely, Regression estimators within the regression estimation technique are utilized for an enhanced estimate of σ_y^2. If the regression line of Y on $\boldsymbol{X}$ does not pass through the origin or a point near it, alternative estimation methods should be considered instead of the ratio approach.

The literature has demonstrated that the efficiency of the estimator under consideration is enhanced when auxiliary information is appropriately applied during the estimating stage. For a more accurate calculation of the population variance, Singh and Singh [1] suggested a ratio estimator using auxiliary data. Singh and Singh [2] later devised an improved regression technique for assessing population variation in a two-phase sampling design.A useful family of chain estimators for the increased estimation of population variance using the sub-sampling method was presented by Jhajj et al. [3]. Furthermore, Shabbir and Gupta's [4] study focused mostly on developing auxiliary parameter-based variance estimation. Kadilar and Cingi[5] have been proposed to improve the estimation of population variance under the simple random sampling framework. Singh et al. [6] proposed ratio and product type estimators that are asymptotically unbiased for estimating the finite population variance, making use of the kurtosis information of an auxiliary variable in sample surveys.Grover [7] released an update on the enhanced estimate of σ_y^2 utilizing auxiliary parameters. Additionally, Singh and Solanki [8] proposed an estimator which also utilized auxiliary data.

However, utilizing supplementary factors, Yadav and Kadilar[9] suggested a two-parameter improved σ_y^2 estimator. Singh and Pal [10] suggested an enhanced family of estimators to estimate σ_y^2 utilizing auxiliary variable quartiles. Yadav et al. [11] used the inter-quartile range (IQR) and tri-mean of the auxiliary variable, which are known values, to create a superior σ_y^2estimate. Yadav et al. [12] have suggested a more accurate estimate of σ_y^2 by using the third quartile and the renowned tri-mean of the auxiliary variable. When outliers were present, Naz et al. [13] offered ratio-type estimators of σ_y^2 using non-traditional measures of dispersion X of that showed a strong association with the study variable Y. Olayiwola et al. [14] developed a unique exponential ratio estimate for population variance, which has shown benefits over many other current estimators of σ_y^2 . Bhushan et al. [15] introduced several modified classes of estimators for population variance by incorporating well-known auxiliary parameters. Ahmed and Hussein [16] developed a few ratio estimators of σ_y^2 using the ranked set sampling scheme's known auxiliary parameters. Sharma et al. [17] proposed a better variance estimator for decision-making that makes use of known auxiliary factors.

2. Review of existing estimators

This section presents various estimators of the population variance σ_y^2 of Y with their mean squared errors (MSEs) for a first order of approximation which is displayed in the table given below.

Table 1: Existing estimators and MSE of difference estimators

Sl. N	Estimator and Authors	Variance/MSE
1	$t_0 = s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ Sample variance Estimator	$V(t_0) = \gamma S_y^4(\lambda_{40} - 1)$
2	$t_r = s_y^2\left[\frac{S_x^2}{s_x^2}\right]$	$MSE(t_r) = \gamma S_y^4[(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)]$

	Isaki [18] Estimator	
3	$t_1 = s_y^2\left[\frac{S_x^2 - \beta_2}{s_x^2 - \beta_2}\right]$ Kadilar and Cingi [19]	$MSE(t_1) = \gamma S_y^4[(\lambda_{40} - 1) + R_1^2(\lambda_{04} - 1) - 2R_1(\lambda_{22} - 1)]$
4	$t_2 = s_y^2\left[\frac{S_x^2 C_x - \beta_2}{s_x^2 C_x - \beta_2}\right]$ Kadilar and Cingi [19]	$MSE(t_2) = \gamma S_y^4[(\lambda_{40} - 1) + R_2^2(\lambda_{04} - 1) - 2R_2(\lambda_{22} - 1)]$
5	$t_3 = s_y^2\left[\frac{S_x^2 - C_x}{s_x^2 - C_x}\right]$ Kadilar and Cingi [19]	$MSE(t_3) = \gamma S_y^4[(\lambda_{40} - 1) + R_3^2(\lambda_{04} - 1) - 2R_3(\lambda_{22} - 1)]$
6	$t_4 = s_y^2\left[\frac{S_x^2 \beta_2 - C_x}{s_x^2 \beta_2 - C_x}\right]$ Kadilar and Cingi[19]	$MSE(t_4) = \gamma S_y^4[(\lambda_{40} - 1) + R_4^2(\lambda_{04} - 1) - 2R_4(\lambda_{22} - 1)]$
7	$t_5 = s_y^2\left[\frac{S_x^2 + Q_1}{s_x^2 + Q_1}\right]$ Subramani and Kumarapandiyan [20]	$MSE(t_5) = \gamma S_y^4[(\lambda_{40} - 1) + R_5^2(\lambda_{04} - 1) - 2R_5(\lambda_{22} - 1)]$
8	$t_6 = s_y^2\left[\frac{S_x^2 + Q_3}{s_x^2 + Q_3}\right]$ Subramani and Kumarapandiyan[20]	$MSE(t_6) = \gamma S_y^4[(\lambda_{40} - 1) + R_6^2(\lambda_{04} - 1) - 2R_6(\lambda_{22} - 1)]$
9	$t_7 = s_y^2\left[\frac{S_x^2 + Q_R}{s_x^2 + Q_R}\right]$ Subramani and Kumarapandiyan [20]	$MSE(t_7) = \gamma S_y^4[(\lambda_{40} - 1) + R_7^2(\lambda_{04} - 1) - 2R_7(\lambda_{22} - 1)]$
10	$t_8 = s_y^2\left[\frac{S_x^2 + Q_D}{s_x^2 + Q_D}\right]$ Subramani and Kumarapandiyan[20]	$MSE(t_8) = \gamma S_y^4[(\lambda_{40} - 1) + R_8^2(\lambda_{04} - 1) - 2R_8(\lambda_{22} - 1)]$
11	$t_9 = s_y^2\left[\frac{S_x^2 + Q_A}{s_x^2 + Q_A}\right]$ Subramani and Kumarapandiyan[20]	$MSE(t_9) = \gamma S_y^4[(\lambda_{40} - 1) + R_9^2(\lambda_{04} - 1) - 2R_9(\lambda_{22} - 1)]$

12	$t_{10} = \left[W_1 s_y^2 + W_2 s_x^2\right] exp\left[\frac{(S_x^2 - s_x^2)}{(S_x^2 + s_x^2) + 2S_x}\right]$ Panda and Sahoo [21]	$MSE(t_{10}) = S_y^2\left[1 + \frac{(2A_3A_4A_5 - A_2A_3^2 - A_1A_4^2)}{(A_1A_2 - A_5^2)}\right]$
13	$t_{11} = s_y^2\left[\frac{S_x^2 + S_x\beta_2}{s_x^2 + S_x\beta_2}\right]$ Khan et al. [22]	$MSE(t_{11}) = \gamma S_y^4[(\lambda_{40} - 1) + R_{11}^2(\lambda_{04} - 1) - 2R_{11}(\lambda_{22} - 1)]$
14	$t_{12} = s_y^2\left[\frac{S_x^2}{s_x^2}\right] exp\left[\frac{(S_x^2 - s_x^2)}{S_x^2 + (a-1)s_x^2}\right]$ Sahoo and Jhankar[23]	$MSE(t_{11}) = \gamma S_y^4\left[(\lambda_{40} - 1) + (\lambda_{04} - 1) + \left(\frac{((\lambda_{22}-1)-(\lambda_{04}-1))}{(\lambda_{04}-1)}\right)^2(\lambda_{04} - 1) + 2\left(\frac{((\lambda_{22}-1)-(\lambda_{04}-1))}{(\lambda_{04}-1)}\right)\left((\lambda_{04} - 1) - (\lambda_{22} - 1)\right) - 2(\lambda_{22} - 1)\right]$
15	$t_{13} = s_y^2\left[\frac{S_x^2 + S_x\beta_2}{s_x^2 + S_x\beta_2}\right]^\delta$ Soni and Pandey[24]	$MSE(t_{13}) = \gamma S_y^4\left[(\lambda_{40} - 1) - \frac{(\lambda_{22} - 1)^2}{(\lambda_{04} - 1)}\right]$

Where,

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

$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i , \lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}},$

$\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^r(X_i - \bar{X})^s, S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2 , R_1 = \frac{S_x^2}{S_x^2 - \beta_2} , R_2 = \frac{S_x^2 C_x}{S_x^2 C_x - \beta_2} ,$

$R_3 = \frac{S_x^2}{S_x^2 - C_x} , R_4 = \frac{S_x^2\beta_2}{S_x^2\beta_2 - C_x} , R_5 = \frac{S_x^2}{S_x^2 + Q_1} , R_6 = \frac{S_x^2}{S_x^2 + Q_3} ,$

$R_7 = \frac{S_x^2}{S_x^2 + Q_R} , R_8 = \frac{S_x^2}{S_x^2 + Q_D} , R_9 = \frac{S_x^2}{S_x^2 + Q_A} ,$

$\theta = \frac{S_x^2}{S_x^2 + S_x},$

$A_1 = 1 + \frac{1}{n}(\lambda_{40} - 1) + \frac{\theta^2}{n}(\lambda_{04} - 1) - \frac{2\theta}{n}(\lambda_{22} - 1) ,$

$A_2 = 1 + \frac{(1-\theta)^2}{n}(\lambda_{04} - 1) - \frac{\theta}{n}(\lambda_{22} - 1),$

$$A_3 = -1 + \frac{\theta}{2n}(\lambda_{22} - 1) - \frac{3\theta^2}{8n}(\lambda_{04} - 1),$$

$$A_4 = -1 - \frac{\theta}{2n}(\lambda_{22} - 1) + \left(\frac{\theta}{2} - \frac{3\theta^2}{8}\right)(\lambda_{04} - 1)$$

$$A_5 = 1 + \frac{(\theta^2 - \theta)}{n}(\lambda_{04} - 1) - \left(1 - \frac{3\theta}{2}\right)\frac{1}{n}(\lambda_{22} - 1),$$

(W_1, W_2) are suitably choosen scalar.

$R_{11} = \frac{S_x^2}{S_x^2 + S_x\beta_2}$ and a is suitably choosen constant.

3. Proposed estimator

Inspired by the works discussed in the last section, we propose a unique estimator for S_y^2 in SRSWOR as follows:

$$T_R^* = k_1 s_y^2 \left(\frac{S_x^2}{s_x^2}\right)^{K_2} exp\left(\frac{C_x S_x^2 + \beta_2}{C_x s_x^2 + \beta_2}\right),$$

where k_j $(j = 1,2)$ are constants chosen to minimize the mean squared error.

To examine the bias and mean squared error (MSE) of the proposed estimator, the following standard approximations are applied:

$e_0 = (s_y^2 - S_y^2)/S_y^2$, $e_1 = (s_x^2 - S_x^2)/S_x^2$, $E(e_0) = 0, E(e_1) = 0$ and $E(e_0^2) = \gamma(\lambda_{40} - 1)$, $E(e_1^2) = \gamma(\lambda_{04} - 1)$, $E(e_0 e_1) = \gamma(\lambda_{22} - 1)$.

Theorem 1. The bias of the proposed estimator T_R^* for the population variance S_y^2 is expressed as

$$Bias(T_R^*) = S_y^2\left[k_1\{2 + (\delta^2 + k_2\delta + k_2(k_2 + 1))\gamma(\lambda_{04} - 1) - (\delta + 2k_2)\gamma(\lambda_{22} - 1)\} - 1\right],$$

where $\delta = \frac{C_x S_x^2}{C_x S_x^2 + \beta_2}$.

Proof.

$$T_R^* = k_1 s_y^2 \left(\frac{S_x^2}{s_x^2}\right)^{K_2} exp\left(\frac{C_x S_x^2 + \beta_2}{C_x s_x^2 + \beta_2}\right)$$

$$= k_1 S_y^2 (1 + e_0)(1 + e_1)^{-k_2} exp(1 - \delta e_1 + \delta^2 e_1^2)$$

$$= k_1 S_y^2 (1 + e_0)\left\{1 - k_2 e_1 + \frac{k_2(k_2 + 1)}{2!} e_1^2 - \cdots\right\}$$

$$\times \quad \{1 + (1 - \delta e_1 + \delta^2 e_1^2)\}$$

$$= k_1 S_y^2 (1 + e_0)\left\{1 - k_2 e_1 + \frac{k_2(k_2 + 1)}{2!} e_1^2 - \cdots\right\}$$

$$\times \ \{2 - \delta e_1 + \delta^2 e_1^2\}$$

Upon simplifying the above expression, we obtain

$$T_R^* = k_1 S_y^2 \begin{pmatrix} 2-\delta e_1+\delta^2 e_1^2-2k_2 e_1-k_2\delta e_1^2 \\ +k_2(k_2+1)e_1^2-2e_0-2\delta e_1 e_0-2k_2 e_1 e_0 \end{pmatrix}$$

$$= k_1 S_y^2 \begin{Bmatrix} 2+2e_0-(\delta+2k_2)e_1 \\ +(\delta^2-k_2\delta+k_2(k_2+1))e_1^2-(\delta+2k_2)e_1 e_0 \end{Bmatrix}$$

$$T_R^* - S_y^2 = S_y^2\left[k_1\left\{2+2e_0-(\delta+2k_2)e_1\begin{pmatrix}\delta^2-k_2\delta+ \\ k_2(k_2+1)\end{pmatrix}e_1^2-(\delta+2k_2)e_1e_0\right\}-1\right] \qquad (1)$$

The first-degree bias of the estimator T_R^* is derived by taking the expectation on both sides of equation (1) as follows:

$$E(T_R^*-S_y^2) = S_y^2\left[k_1\begin{Bmatrix}2+(\delta^2-k_2\delta+k_2(k_2+1))\gamma(\lambda_{04}-1) \\ -(\delta+2k_2)\gamma(\lambda_{22}-1)\end{Bmatrix}-1\right]$$

$$Bias(T_R^*) = S_y^2\left[k_1\{2+(\delta^2-k_2\delta+k_2(k_2+1))\gamma(\lambda_{04}-1)-(\delta+2k_2)\gamma(\lambda_{22}-1)\}-1\right]$$

(Proved)

Theorem 2. The mean squared error of the proposed estimators T_R^* , based on the first-order approximation, is given by

$$\text{MSE}(T_R^*) = S_y^4(k_1^2 Z_1 - 4k_1 + 1),$$

where $Z_1 = \{4+4\,\gamma(\lambda_{40}-1)+(\delta+2K_2)^2\gamma(\lambda_{04}-1)-4(\delta+2K_2)\gamma(\lambda_{22}-1)\}$.

Proof: From the equation (1)

$$T_R^*-S_y^2 = S_y^2\left[k_1\{2+2e_0-(\delta+2k_2)e_1+(\delta^2-k_2\delta+k_2(k_2+1))e_1^2-(\delta+2k_2)e_1e_0\}-1\right] \qquad (2)$$

The first-order mean squared error of the estimator T_R^* can be derived by squaring and taking expectations on both sides of equation (2) as follows:

$$MSE(T_R^*) = S_y^4[K_1^2\{4+4\gamma(\lambda_{40}-1)+(\delta+2K_2)^2\gamma(\lambda_{04}-1)-4(\delta+2K_2)\gamma(\lambda_{22}-1)+1-4K_1\}]$$

$$\Rightarrow MSE(T_R^*) = S_y^4(k_1^2 Z_1 - 4k_1 + 1)$$

Where $Z_1 = \begin{Bmatrix}4+4\,\gamma(\lambda_{40}-1)+(\delta+2K_2)^2\gamma(\lambda_{04}-1)- \\ 4(\delta+2K_2)\gamma(\lambda_{22}-1)\end{Bmatrix}$.

Note:

(a) Minimizing the MSE(T_R^*) against k_1 provides $k_{1(opt)}$ as

$$k_{1(opt)} = \frac{2}{Z_1}$$

Putting the values of $k_{1(opt)}$ in the MSE(T_R^*) , we obtain

$$\text{MSE}(T_R^*)_{min} = S_y^4\left(1-\frac{4}{Z_1}\right).$$

(b) Simultaneous optimization of k_1 and k_2 is quite complex. Hence, by setting $k_1 = 1$ in the estimator (T_R^*) and minimizing MSE(T_R^*) with respect to k_2 , the optimal value of k_2 can be determined as

$$k_2 = -\frac{\delta}{2} + \frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}.$$

4. Efficiency comparison

This section presents a comparison between the proposed estimator and the existing estimators of population variance. Table 2 outlines the conditions under which the proposed estimator outperforms the competing ones.The proposed estimator T_R^* exhibits greater efficiency than the competing estimators t_i $(i = 0, r, \ldots, 13)$, when the condition $MSE(t_i) - \text{MSE}(T_R^*)_{min} > 0$.

.

Table 2. Efficiency criteria of the proposed estimator compared to competing estimators

Estimator	Efficiency condition
t_0	$\frac{4}{Z_1} > \{1 - \gamma(\lambda_{40} - 1)\}$
t_r	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)\}]$
t_1	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_1^2(\lambda_{04} - 1) - 2R_1(\lambda_{22} - 1)\}]$
t_2	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_2^2(\lambda_{04} - 1) - 2R_2(\lambda_{22} - 1)\}]$
t_3	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_3^2(\lambda_{04} - 1) - 2R_3(\lambda_{22} - 1)\}]$
t_4	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_4^2(\lambda_{04} - 1) - 2R_4(\lambda_{22} - 1)\}]$
t_5	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_5^2(\lambda_{04} - 1) - 2R_5(\lambda_{22} - 1)\}]$
t_6	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_6^2(\lambda_{04} - 1) - 2R_6(\lambda_{22} - 1)\}]$
t_7	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_7^2(\lambda_{04} - 1) - 2R_7(\lambda_{22} - 1)\}]$
t_8	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_8^2(\lambda_{04} - 1) - 2R_8(\lambda_{22} - 1)\}]$
t_9	$\frac{4}{Z_1} > [1 - \gamma\{(\lambda_{40} - 1) + R_9^2(\lambda_{04} - 1) - 2R_9(\lambda_{22} - 1)\}]$
t_{10}	$\frac{4}{Z_1} > -\left[\frac{(2A_3A_4A_5 - A_2A_3^2 - A_1A_4^2)}{(A_1A_2 - A_5^2)}\right].$

t_{11}	$\frac{4}{Z_1} > [1-\gamma\{(\lambda_{40}-1)+R_{11}^2(\lambda_{04}-1)-2R_{11}(\lambda_{22}-1)\}]$
t_{12}	$\frac{4}{Z_1} > \left[1-\gamma\left[(\lambda_{40}-1)+(\lambda_{04}-1)+\left(\frac{((\lambda_{22}-1)-(\lambda_{04}-1))}{(\lambda_{04}-1)}\right)^2(\lambda_{04}-1)+2\left(\frac{((\lambda_{22}-1)-(\lambda_{04}-1))}{(\lambda_{04}-1)}\right)\left((\lambda_{04}-1)-(\lambda_{22}-1)\right)-2(\lambda_{22}-1)\right]\right]$
t_{13}	$\frac{4}{Z_1} > \left[1-\gamma\left\{(\lambda_{40}-1)-\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right\}\right]$

5. Numerical study

To assess the efficacy of the proposed estimator relative to competing estimators, we have examined two natural populations in this section. The study and auxiliary variables from two natural populations are as follows:

Population-1:(Source: Gupta and Shabbir [25])

Y: Number of agricultural laborers in 1971

X: Number of agricultural labourers in 1961

Population-2: (Source: Kadilar and Cingi [19])

Y: Apple production per 100 tons

X: Number of trees (each unit represents 100 apple trees)

Table 3 provides the parameters of both actual populations.

Table 3: The parameter of both natural population 1 and 2

Population-1	Population-2
$N = 278, n = 30, \bar{Y} = 39.068, \bar{X} = 25.111,$ $S_y = 56.457, S_x = 40.674, \lambda_{40} = 25.896,$ $\lambda_{04} = 38.889, \lambda_{22} = 26.814, \rho = 0.721$	$N = 173, n = 20, \bar{Y} = 4.04, \bar{X} = 98.44,$ $S_y = 9.46, S_x = 187.94, \lambda_{40}, = 27.96,$ $\lambda_{04} = 28.10, \lambda_{22} = 23.08, \rho = 0.891$

Table 4 shows the Mean Squared Errors (MSEs) of the proposed and competing estimators, along with the Percent Relative Efficiency (PRE) of each estimator relative to t_0 .

Table 4: MSE of the proposed and competing estimators and PRE with respect to t_0

Populaion-1			Population-2		
Estimator	MSE	PRE	Estimator	MSE	PRE
t_0	7521187.116	100	t_0	9547.72201	100
t_r	3778793.03	199.037	t_r	3506.02551	272.323
t_1	3983111.977	188.827	t_1	3511.14624	271.926
t_2	4862885.344	154.665	t_2	4862.88534	196.339

t_3	4862205.422	154.687	t_3	4862.20542	196.366
t_4	4862885.467	154.665	t_4	4862.88546	196.339
t_5	3468925.171	216.816	t_5	3504.57264	272.436
t_6	3022246.855	248.861	t_6	3502.82491	272.572
t_7	3220379.202	233.551	t_7	3504.16751	272.468
t_8	3468925.171	216.816	t_8	3506.54872	272.283
t_9	3220379.202	233.551	t_9	3503.61287	272.511
t_{10}	2545050.407	295.522	t_{10}	3876.9875	246.266
t_{11}	2846652.986	264.212	t_{11}	3497.55713	272.983
t_{12}	2208010.086	340.632	t_{12}	3176.70558	300.554
t_{13}	2208010.086	340.632	t_{13}	3176.70558	300.554
T_R^*	**1828705.171**	**411.284**	T_R^*	**2115.629**	**451.26**

6. Simulation Study

In addition, we have also performed a simulation study to validate the superiority of our proposed estimator.We generate $N = 300$ values (x_i, y_i) from a Bi-variate normal distribution with means (70,60) and standard deviation (15,10) with correlation coefficient value 0.90. A sample of sizes $n = 40$ and 30 has been considered respectively for this study. By simulating the data 600 times we have generated a finite population of size 300 and we calculate the following simulated mean, standard deviation and other parameters for the fixed correlation coefficient value.

Table 5: The parameter of both simulated population 3 and 4

Population-3	Population-4
$N = 300, n = 40, \bar{Y} = 62.632, \bar{X} = 72.017,$ $S_y = 5.366, S_x = 5.366, \lambda_{40} = 17.946,$ $\lambda_{04} = 15.303, \lambda_{22} = 14.089$	$N = 300, n = 30, \bar{Y} = 60.982, \bar{X} = 72.41,$ $S_y = 4.874, S_x = 4.874, \lambda_{40}, = 39.847,$ $\lambda_{04} = 26.663, \lambda_{22} = 30.108$

Table 6 shows the Mean Squared Errors (MSEs) of the proposed and competing estimators, along with the Percent Relative Efficiency (PRE) of each estimator relative to t_0.

Table 6: MSE of the proposed and competing estimators and PRE with respect to t_0

Populaion-3			Population-4		
Estimator	MSE	PRE	Estimator	MSE	PRE
t_0	93.003	100	t_0	173.435	100
t_r	27.831	334.170	t_r	28.105	617.102
t_1	29.311	317.299	t_1	26.047	665.854
t_2	56.541	164.487	t_2	130.311	133.095
t_3	27.931	332.971	t_3	27.832	623.149
t_4	27.879	333.597	t_4	28.001	619.386

t_5	54.806	169.696	t_5	112.065	154.763
t_6	61.877	150.309	t_6	122.063	142.086
t_7	38.865	239.298	t_7	65.871	263.294
t_8	31.124	298.816	t_8	47.131	367.982
t_9	58.682	158.487	t_9	117.499	147.605
t_{10}	58.21	159.771	t_{10}	125.3494	138.361
t_{11}	30.138	308.595	t_{11}	53.393	324.826
t_{12}	27.267	341.101	t_{12}	26.041	666.01
t_{13}	27.267	341.101	t_{13}	26.041	666.01
T_R^*	**24.616**	**377.818**	T_R^*	**22.163**	**782.552**

7. Results and Conclusion

In this paper, we offer a generalized ratio type estimator of population variance based on known auxiliary variables. The sample properties, bias, and MSE of the suggested estimator are established for the first order of approximation. The ideal value of the characterizing constant for the proposed estimator has been determined. The MSE of the minimal value of the suggested estimator has also been established for this optimal value of the characterizing constant. Conceptually, empirically and simulated, the suggested estimator is compared to the existing competing estimators. Table 4 and Table 6 show that, out of all the competing estimators, the suggested estimator T_R^* has the lowest MSE for both populations. Consequently, it is recommended that the suggested estimator be applied practically in a variety of application domains.

Acknowledgements

The author thanks to the Chief Editor for his constructive feedback and reviewer's thoughtful suggestions that helped enhance this study.

REFERENCES

[1] H.P. Singh, R. Singh. Improved ratio-type esimator forvariance using auxiliary information, Journal of the Indian Society of Agricultural Statistics, Vol. 54, No.3, 276-287,2001.

[2] H.P. Singh, R. Singh. Improved regression estimator for population variance in two-phase sampling, Journal of theIndian, Society of Agricultural Statistics, Vol. 56, No. 3, 267-275, 2003.

[3] S.S. Jhajj, R. Sharma, H.P. Singh. A family of chain-type estimators for finite population variance using two-phase sampling scheme, Journal of the Indian Society of AgriculturalStatistics, Vol.59,No. 2, 190–196, 2005.

[4] J. Shabbir, S. Gupta. Auxiliary information based estimationof population variance, Journal of the Indian Society ofAgricultural Statistics,Vol. 61, No. 3, 319–326, 2007.

[5] C. Kadilar, H. Cingi. Improvement in variance estimationusing auxiliary information, Hacettepe Journal of Mathematicsand Statistics, Vol.36, No. 1, 103–109, 2007.

[6] H.P. Singh, N. Malik, P. Kumar. A general class of ratio-product type estimators of finite population variance using information on kurtosis, Journal of the Indian Society of Agricultural Statistics,Vol. 62, No. 2, 160–170, 2008.

[7] P. Grover. Improved estimation of population variance using known parameters of auxiliary variable, Journal of the Indian Society of Agricultural Statistics, Vol. 64, No. 3, 367–372, 2010.

[8] H.P. Singh, R.S. Solanki. Improved estimation of population variance using auxiliary information in sample surveys, Journal of the Indian Society of Agricultural Statistics, Vol. 66, No. 1, 55–63, 2012.

[9] S. Yadav, C. Kadilar. An improved two-parameter ratio-product type estimator using auxiliary information, Hacettepe Journal of Mathematics and Statistics, Vol.43, No.6, 967–975, 2014.

[10] H.P. Singh, S. Pal. Improved estimation of finite population variance using known values of population quartiles, Journal of the Indian Society of Agricultural Statistics,Vol.70, No.2, 129–136, 2016.

[11] S. Yadav, P. Chandel, S. Sharma. Improved estimation of population variance using tri-mean and inter-quartile range of an auxiliary variable, Journal of the Indian Society of Agricultural Statistics, Vol. 71, No. 2, 141–148, 2017.

[12] S. Yadav, S. Sharma, P. Chandel. An improved estimator of population variance using tri-mean and third quartile of an auxiliary variable, Journal of the Indian Society of Agricultural Statistics, Vol.73, No. 2, 145–152, 2019.

[13] A. Naz, A. Hanif, M. Shabbir. Ratio estimators for population variance using some non-conventional measures of dispersion under SRS, Pakistan Journal of Statistics and Operation Research, Vol.16, No.1, 27–40, 2020.

[14] O.A. Olayiwola, T.S. Adewara, A.J. Adeleke. A novel exponential ratio-type estimator for estimating population variance using auxiliary information, Hacettepe Journal of Mathematics and Statistics, Vol.50, No. 2, 367–376, 2021.

[15] S. Bhushan, P. Chandel, R. Sharma. Modified classes of estimators for population variance using known auxiliary information, Journal of Statistical Theory and Practice, Vol.16, No.2, 2022.

[16] S.E. Ahmed, M.I. Hussein. Ratio-type estimators for population variance using known auxiliary parameters under ranked set sampling, Communications in Statistics - Theory and Methods, Vol.51, No.24, 8483–8496, 2022.

[17] S. Sharma, P. Chandel, S. Yadav. Improved estimator of population variance using known parameters of auxiliary variable for better decision making, Journal of the Indian Society of Agricultural Statistics, Vol.76, No.1, 59–66, 2022.

[18] C.T. Isaki. Variance estimation using auxiliary information, Journal of American Statistical Association, Vol.78, 117–123, 1983.

[19] C. Kadilar, H. Cingi. Improvement in variance estimation using auxiliary information, Journal of Mathematics and Statistics, Vol.35, 111–115, 2005.

[20] J. Subramani, G. Kumarapandiyan. Variance estimation using quartiles and their functions of an auxiliary variable, International Journal of Statistics and Applications, Vol.2, No.5, 67–72, 2012.

[21] K.B. Panda, N. Sahoo. Improved ratio-type exponential estimator for estimating population variance using auxiliary information, International Journal of Statistics and Systems, Vol.11, 187–195, 2016.

[22] S.A. Khan, M. Nawaz, K.M. Din. Estimation of population variance in simple random sampling using auxiliary information, International Journal of Computer Applications, Vol.175, No.33, 6–10, 2020.

[23] N. Sahoo, S.K. Jhankar. An exponential ratio-type estimator for estimating population variance with its efficiency, International Journal of Agricultural and Statistical Science, Vol.18, 1085–1089, 2022.

[24] S.S. Soni, H. Pandey. Generalized estimator of population variance utilizing auxiliary information in simple random sampling scheme, Journal of Probability and Statistical Science, Vol. 21, No.2, 23-33, 2023

[25] S. Gupta, J. Shabbir. Variance estimation in simple random sampling using auxiliary information, Journal of Mathematics and Statistics, Vol.37, 57–67, 2008.