Volume 69, Issue 3, 2025
Journal of Scientific Research
of
The Banaras Hindu University

Construction of Almost Unbiased Estimator for Coefficient of Variation using Exponential and Sine Estimators

Rajesh Singh[1], Shobh Nath Tiwari[2], Yashveer Singh[3], Sunil Kumar Yadav[*4]

[1]Professor, Dept. of Statistics, Institute of Science, Banaras Hindu University, Varanasi
Email ID: rsinghstat@gmail.com
[2]Research scholar, Dept. of Statistics, Institute of Science, Banaras Hindu University, Varanasi
Email ID: shobh7285@gmail.com
[3]Research scholar, Dept. of Statistics, Institute of Science, Banaras Hindu University, Varanasi
Email ID: yashveerbhu@gmail.com
[*4]Research scholar, Dept. of Statistics, Institute of Science, Banaras Hindu University, Varanasi
Email ID: ysunilkumar40@gmail.com

Abstract: **In survey sampling, the coefficient of variation (CV) is a vital measure of relative variability. This study proposes a novel class of sine-type estimators for estimating the finite population CV under simple random sampling without replacement (SRSWOR), utilizing known auxiliary information. There are number of estimators for estimating CV, we introduce new ratio, product, and difference-type estimator that incorporate sine and exponential transformations. Furthermore, an almost unbiased estimator is constructed by optimally combining three proposed estimators. The theoretical properties of the estimators, including bias and mean squared error (MSE), are derived and compared. Through empirical analysis using real datasets and simulation studies, we consistently observe that the proposed estimators—particularly the almost unbiased estimator—achieve lower MSE and higher percent relative efficiency (PRE) than existing estimators. These findings highlight the effectiveness of exponential-sine type estimators in enhancing estimator performance, especially when auxiliary information is strongly correlated with the study variable**

Index Terms: **Almost unbiased estimator, Auxiliary information, Coefficient of variation, Exponential-sine type estimator, Sine-type estimators**

I. Introduction

In the field of sampling theory, particularly within the context of survey sampling, it is widely recognized that the use of auxiliary information plays a crucial role in enhancing the precision of population parameter estimates. Auxiliary information refers to the data or attributes related to the population under study that are known prior to sampling and are often correlated with the variable of interest. Such information may include demographic data, historical records, or other relevant characteristics that are easier or cheaper to measure. The incorporation of this auxiliary information at the estimation stage allows statisticians and researchers to construct more efficient estimators.

As a result, estimators that utilize auxiliary information often require smaller sample sizes to achieve the same level of accuracy, which in turn reduces cost and effort in the data collection process. Moreover, the gain in efficiency from using auxiliary information is particularly significant when the auxiliary variable is highly correlated with the study variable. In such cases, methods such as ratio estimators, regression estimators, and product-type estimators are frequently used to make more accurate inferences about the population mean or total. These methods leverage the known relationship between the auxiliary and study variables to adjust the estimates accordingly, thereby increasing their reliability. In contrast, estimators that ignore auxiliary information may suffer from larger variances and reduced accuracy. They fail to take advantage of the potential correlation structure between the study and auxiliary variables, leading to less precise results. Therefore, in modern survey sampling practice, the use of auxiliary information is not only encouraged but often considered essential for producing high-quality statistical estimates.

Many researchers have worked on improving the estimation of population parameters like the mean, variance, and standard deviation by using auxiliary information. Notable contributions have been made by Srivastava, S.K. (1967), Singh and Solanki (2012), Shai and Ray (1980), Srivastava and Jhajj (1981), Singh and Kumar (2011), Sharma et al. (2013), Isaki, C.T. (1983), Jhajj et al. (2006), Naik and Gupta (1996), Das and Tripathi (1992) and their studies introduced methods that make better use of the relationship between the study variable and auxiliary variable, leading to more accurate and efficient estimates.

Although many researchers focused on estimating population parameters, the estimation of the CV received relatively less attention for a long time. However, some work has been done in this area. Das and Tripathi (1992) were among the first to propose an estimator for the CV under simple random sampling without replacement (SRSWOR). Later, Patel and Shah (2009) also contributed to this field. Singh et.al. (2018) worked on estimating the CV using auxiliary variable. Singh, R and Yadav (2024) has proposed as the estimators in case of CV. Archana and Rao (2014) proposed estimators for estimating finite population CV. Singh and Kumari (2022) suggested some novel estimators of population CV using SRSWOR.

Although many studies have focused on estimating the population mean, sine-type estimators have not been explored much. This study introduces improved sine-type estimators specifically for estimating the CV using known auxiliary information under simple random sampling. To the best of our knowledge, this is the first time sine-type estimators are being used for CV estimation. Previous works, such as Bhattacharyya et al. (2021) and Yunusa et al. (2023), explored sine-type approaches for estimating the population mean.

This study focuses on estimating the population CV by using information from a single auxiliary variable under simple random sampling without replacement (SRSWOR). For this, we constructed new sine-type estimators by modifying existing estimators. These include estimators suggested by Cochran W.G. (1940), Murthy M.N. (1964), Bahl and Tuteja (1991) exponential ratio, product-type, as well as Rao T.J. (1991)'s difference estimators, which are used to estimate the CV. After forming these sine-type estimators, we used a method by Singh and Singh (1991,1993) to make an almost unbiased estimator. This means the final estimator gives results that are very close to the true population value as it reduces bias. Our aim is to improve accuracy and reduce the error in estimating the CV using auxiliary information.

In this study, two mathematical functions are used: the exponential function and the sine function. These functions are chosen because they are commonly applied in real-life situations. The exponential function is often used to model natural growth and decay, such as the growth of bacteria, human population, the spread of diseases, and compound interest. On the other hand, the sine function is useful for modeling periodic phenomena like sound waves, light waves, ocean tides, seasonal temperature changes, electrical currents, and even GPS movements.

II. NOTATIONS AND TERMINOLOGY

Consider a finite population $U = U_1, U_2, U_3 \dots\dots.. U_N$ of size N consist of distinct and identifiable units. Suppose X and Y denotes the corresponding auxiliary and study variables and let Y_i and X_i is the i^{th} unit of the population $U_i(i = 1,2,3 \dots\dots..N)$. in this population case we define,

Notations	
$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$	Population mean of the study variable.
$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$	Population mean of the auxiliary variable.
$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$	Population mean square of the study variable.
$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$	Population mean square of the auxiliary variable.
$S_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})$	Population covariance.

Let us define sampling errors for both mean and variance of study and auxiliary variables as follows;

$$e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x}-\bar{X}}{\bar{X}}, e_2 = \frac{s_y^2-S_y^2}{S_y^2}, \quad e_3 = \frac{s_x^2-S_x^2}{S_x^2}, e_4 = \frac{s_{xy}-S_{xy}}{S_{xy}}$$

such that,

$$\bar{y} = \bar{Y}(1+e_0), \bar{x} = \bar{X}(1+e_1), s_y^2 = S_y^2(1+e_2),$$
$$s_x^2 = S_x^2(1+e_3), s_{xy} = S_{xy}(1+e_4)$$

and,

$$E(e_0) = E(e_1) = E(e_2) = E(e_3) = E(e_4) = 0.$$
$$E(e_0^2) = \gamma C_y^2, E(e_1^2) = \gamma C_x^2, E(e_2^2) = \gamma(\lambda_{40}-1), E(e_3^2) = \gamma(\lambda_{04}-1), E(e_4^2) = \gamma\left(\frac{\lambda_{22}}{\rho_{xy}^2}-1\right)$$
$$E(e_0e_1) = \gamma\rho_{xy}C_yC_x, E(e_0e_2) = \gamma C_y\lambda_{30},$$
$$E(e_0e_3) = \gamma C_y\lambda_{12}$$
$$E(e_1e_2) = \gamma C_x\lambda_{21}, E(e_1e_3) = \gamma C_x\lambda_{03}, E(e_2e_3) = \gamma(\lambda_{22}-1), E(e_0e_4) = \gamma C_y\left(\frac{\lambda_{21}}{\rho_{xy}}\right)$$
$$E(e_1e_4) = \gamma C_x\left(\frac{\lambda_{12}}{\rho_{xy}}\right), \qquad E(e_2e_4) = \gamma\left(\frac{\lambda_{31}}{\rho_{xy}}-1\right),$$
$$E(e_3e_4) = \gamma\left(\frac{\lambda_{13}}{\rho_{xy}}-1\right).$$

Here, $\gamma = \frac{1}{n}(1-f), f = \frac{n}{N}$, f is known as sampling fraction and C_y and C_x are the population coefficients of variations of study variable Y and auxiliary variable X respectively and defined as $C_y = \frac{S_y}{\bar{Y}}$ and $C_x = \frac{S_x}{\bar{X}}$. Also ρ_{yx} is the correlation coefficient between X and Y,

$$\rho_{yx} = \frac{S_{yx}}{S_yS_x}.$$

In general form, $\mu_{rs} = \frac{\sum_{i=1}^{N}(y_i-\bar{y})^r(x_i-\bar{x})^s}{N-1}$ and $\lambda_{rs} = \frac{\mu_{rs}}{\left(\mu_{20}^{\frac{r}{2}}\mu_{02}^{\frac{s}{2}}\right)}$.

III. EXISTING ESTIMATORS

Before proposing the estimator, it is important to review some existing estimators for the population CV. This section discusses a few of these estimators.

The usual estimator for the population CV is given by,

$$t_0 = \widehat{C_y} = \frac{s_y}{\bar{y}} = \frac{S_y(1+e_2)^{\frac{1}{2}}}{\bar{Y}(1+e_0)} \quad (3.1)$$

Where, $\widehat{C_y} = C_y(1 - e_0 + e_0^2 + \frac{1}{2}e_2 - \frac{1}{2}e_0e_1 - \frac{3}{8}e_2^2)$.

The mean square error (MSE) expression for the given usual estimator t_0 is given by

$$MSE(t_0) = C_y^2\gamma(C_y^2 + 0.25(\lambda_{40} - 1) - C_y\lambda_{30}) \quad (3.2)$$

Singh et al. (2018) proposed ratio-type, exponential ratio-type and difference–type estimators for CV of the study variable Y using mean of auxiliary variable as

$$t_{s1} = \widehat{C_y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha} \quad (3.3)$$

$$t_{s2} = \widehat{C_y}\exp\left(\beta\left(\frac{\bar{X}-\bar{x}}{\bar{X}-\bar{x}}\right)\right) \quad (3.4)$$

$$t_{s3} = \widehat{C_y} + d_1(\bar{X} - \bar{x}) \quad (3.5)$$

Where, α, β and d_1 are real constants.

MSE expressions for the estimators t_{s1}, t_{s2} and t_{s3} is given as,

$$MSE(t_{s1}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + \alpha^2C_x^2 - C_y\lambda_{30} + 2\rho C_yC_x - \alpha C_x\lambda_{21}\right) \quad (3.6)$$

$$MSE(t_{s2}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + \beta^2C_x^2 - C_y\lambda_{30} + \beta\rho C_yC_x - \frac{1}{2}\beta C_x\lambda_{21}\right) \quad (3.7)$$

$$MSE(t_{s3} = \left(\gamma\left(C_y^2\left(C_y^2 - C_y\lambda_{30} + \frac{1}{4}(\lambda_{40} - 1)\right) + d_1^2\bar{X}^2C_x^2 + 2d_1\bar{X}\rho C_yC_x - d_1\bar{X}C_yC_x\lambda_{21}\right)\right) \quad (3.8)$$

where,

$$\alpha = \frac{\lambda_{03}-2\rho C_y}{2C_x}, \beta = \frac{\lambda_{21}-2\rho_{yx}C_y}{C_x}, \quad d_1 = \frac{\lambda_{21}-2\rho_{yx}C_y}{2\bar{X}C_x}.$$

Archana and Rao (2014), proposed the following estimators for estimating the finite population CV given as:

$$t_{AR} = \widehat{C_y}\left(\frac{\bar{X}}{\bar{x}}\right) \quad (3.9)$$

$$t_{AR1} = \widehat{C_y}\left(\frac{\bar{x}}{\bar{X}}\right) \quad (3.10)$$

$$t_{AR2} = \widehat{C_y}\left(\frac{S_x^2}{s_x^2}\right) \quad (3.11)$$

$$t_{AR3} = \widehat{C_y}\left(\frac{s_x^2}{S_x^2}\right) \quad (3.12)$$

The MSE expressions for the estimators t_{AR}, t_{AR1}, t_{AR2} *and* t_{AR3}, are respectively given by:

$$MSE(t_{AR}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + C_x^2 - C_x\lambda_{21} - C_y\lambda_{30} + 2\rho C_yC_x\right) \quad (3.13)$$

$$MSE(t_{AR1}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + C_x^2 + C_x\lambda_{21} - C_y\lambda_{30} - 2\rho C_yC_x\right) \quad (3.14)$$

$$MSE(t_{AR2}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + (\lambda_{04} - 1) - (\lambda_{22} - 1) - C_y\lambda_{21} + 2C_y\lambda_{30}\right) \quad (3.15)$$

$$MSE(t_{AR2}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + (\lambda_{04} - 1) + (\lambda_{22} - 1) - C_y\lambda_{21} - 2C_y\lambda_{30}\right) \quad (3.16)$$

IV PROPOSED ESTIMATORS BASED ON SINE-TYPE ESTIMATOR

Let, $t_0 = \hat{C}_y = \frac{s_y}{\bar{y}}$ (4.1)

Bias and MSE of the estimator t_0 *is*

$$B(t_0) = C_y\gamma[C_y^2 + \frac{1}{2}C_y\lambda_{30} - \frac{1}{8}(\lambda_{40} - 1) \quad (4.2)$$

$$M.S.E(t_0) = C_y^2\gamma\left[C_y^2 - \frac{1}{4}(\lambda_{40} - 1) - C_y\lambda_{30}\right] \quad (4.3)$$

Motivated by Cochran W.G. (1940), we have proposed sine ratio type estimator in case of CV.

$$t_1 = \widehat{C_y}\left[1 + \sin\left(\frac{S_x^2-s_x^2}{S_x^2}\right)\right] \quad (4.4)$$

The bias of the estimator is given as follows:

$$Bias(t_1) = C_y\gamma\left[C_y^2 + C_y\lambda_{12} - \frac{1}{2}(\lambda_{22} - 1) - \frac{1}{2}C_y\lambda_{30} - \frac{1}{8}(\lambda_{40} - 1)\right] \quad (4.5)$$

The mean squared error (MSE) of the sine ratio-type estimator is given as follows:

$$MSE(t_1) = C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + (\lambda_{04} - 1) - C_y\lambda_{30} - (\lambda_{22} - 1) + 2C_y\lambda_{12}\right] \quad (4.6)$$

Product sine type estimator are constructed as motivated by Murthy M.N. (1964) to estimate CV is defined as follow:

$$t_2 = \widehat{C_y}\left[1 + \sin\left(\frac{s_x^2-S_x^2}{S_x^2}\right)\right] \quad (4.7)$$

The bias of the estimator for the product sine type estimator is given as

$$Bias(t_2) = C_y\gamma\left[C_y^2 - C_y\lambda_{12} + \frac{1}{2}(\lambda_{22} - 1) - \frac{1}{2}C_y\lambda_{30} - \frac{1}{8}(\lambda_{40} - 1)\right] \quad (4.8)$$

The mean squared error (MSE) of the product sine-type estimator is given as follows:

$$M.S.E.(t_2) = C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40} - 1) + (\lambda_{04} - 1) - C_y\lambda_{30} + (\lambda_{22} - 1) - 2C_y\lambda_{12}\right] \quad (4.9)$$

Following Bahl and Tuteja (1991) approach, we develop an innovative ratio-type estimator that combines the exponential and sine estimator to CV.

$$t_3 = C_y^2\exp\left[\mathrm{Sin}\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right)\right] \quad (4.10)$$

For the developed exponential-sine combined ratio estimator, the bias is expressed as follows:

$$Bias(t_3) = C_y\gamma\left[C_y^2 - \frac{1}{8}(\lambda_{40}-1) + \frac{3}{8}(\lambda_{04}-1) - \frac{1}{2}C_y\lambda_{30} - \frac{1}{4}(\lambda_{22}-1) + \frac{1}{2}C_y\lambda_{12}\right] \quad (4.11)$$

The mean squared error (MSE) of the exponential-sine ratio-type estimator is given by:

$$M.S.E.(t_3) = C_y^2\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + \frac{1}{4}(\lambda_{04}-1) - C_y\lambda_{30} - \frac{1}{2}(\lambda_{22}-1) + C_y\lambda_{12}\right] \quad (4.12)$$

Extending the work of Bahl and Tuteja (1991), we introduce a novel product-type estimator that combines the exponential and sine estimator, defined as follows:

$$t_4 = C_y^2 \exp\left[\text{Sin}\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right)\right] \quad (4.13)$$

The bias of the proposed estimator, based on the exponential-sine product form is as follows:

$$Bias(t_4) = C_y\gamma\left[C_y^2 - \frac{1}{8}(\lambda_{40}-1) - \frac{1}{8}(\lambda_{04}-1) - \frac{1}{2}C_y\lambda_{30} + \frac{1}{4}(\lambda_{22}-1) - \frac{1}{2}C_y\lambda_{12}\right] \quad (4.14)$$

The mean squared error (MSE) of proposed exponential-sine product estimator is as follows:

$$M.S.E(t_4) = C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + \frac{1}{4}(\lambda_{04}-1) + \frac{1}{2}(\lambda_{22}-1) - C_y\lambda_{12} - C_{y\lambda_{30}}\right] \quad (4.15)$$

Extending the difference estimator framework of Rao T.J. (1991), we propose a novel sine-type difference estimator for enhanced finite CV estimation. The estimator is defined as follows:

$$t_5 = \widehat{C_y}\left[1 + k\sin\left(\frac{S_x^2 - s_x^2}{S_x^2}\right)\right] \quad (4.16)$$

The bias and mean squared error (MSE) of the proposed sine-type estimator in difference estimator t_5 are derived as follows:

$$Bias(t_5) = C_y\gamma\left[C_y^2 + kC_y\lambda_{12} - \frac{1}{2}k(\lambda_{22}-1) - \frac{1}{2}C_y\lambda_{30} - \frac{1}{8}(\lambda_{40}-1)\right] \quad (4.17)$$

$$M.S.E(t_5) = C_y^2\gamma\left[C_y^2 + \frac{1}{2}(\lambda_{40}-1) + k^2(\lambda_{04}-1) - C_y\lambda_{30} - k(\lambda_{22}-1) + 2kC_y\lambda_{12}\right] \quad (4.18)$$

After minimizing this, the minimum mean squared error (MSE) for estimator t_5 is achieved when the tuning parameter K takes the value

$$k = \frac{1}{2}\left(\frac{(\lambda_{22}-1)}{\lambda_{04}-1}\right) - \left(\frac{C_y\lambda_{12}}{\lambda_{04}-1}\right) \quad (4.19)$$

Substituting the optimal parameter value k into (3.27) we obtain the minimum mean squared error for estimator t_5.

$$M.S.E(t_5) = C_y^2\gamma\left[C_y^2 + \frac{1}{2}(\lambda_{40}-1) + k^2(\lambda_{04}-1) - C_y\lambda_{30} - k(\lambda_{22}-1) + 2kC_y\lambda_{12}\right] \quad (4.20)$$

V PROPOSED ALMOST UNBIASED ESTIMATOR FOR COV COMBINING EXPONENTIAL AND SINE ESTIMATOR

Let,

$t_0 = \hat{C}_y = \frac{s_y}{\bar{y}}$, $t_1 = C_y^2 \exp\left[\text{Sin}\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right]$ and $t_2 = C_y^2 \exp\left[\text{Sin}\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right)\right]$ are the three estimators such that t_0 , t_1 and t_2 $\in$ L , where L denotes the set of all possible estimators for estimating the coefficient of variation C_y.

Adopting Singh, R. and Singh, S. (1991, 1993), we have proposed a linear combination of three initial estimators (t_0, t_1 *and* t_2) to construct an almost unbiased estimator t_h for the estimating unknown population CV. By imposing a linear constraint $t_h = \sum_{i=0}^{2}\alpha_i t_i \in$ L and solving for the weights α_i, we derive explicit formulas for α_1, α_2 *and* α_3 that depend on the correlation and relative variance of the auxiliary variables. This approach reduces bias to order $O(n^{-1})$ and ensures higher efficiency, particularly when the study and auxiliary variables are correlated.

Following Singh, R and Singh, S. (1991, 1993), we proposed an almost unbiased estimator t_h as follows-

$$t_h = \sum_{i=0}^{2}\alpha_i t_i \in L \quad (5.1)$$

The extended form of the estimator is as follow-

$$t_h = \alpha_0\widehat{C_y} + \alpha_1\widehat{C_y}\exp\left[sin\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right] + \alpha_2\widehat{C_y}\exp\left[sin\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right)\right] \quad (5.2)$$

$$\text{For } \sum_{i=0}^{2}\alpha_i = 1,\ \alpha_i \in R \quad (5.3)$$

where $\alpha_i (i = 0, 1, 2)$ denotes the statistical constants and R denotes the set of real numbers

Table 1: Member Estimators of the Proposed Exponential-sine type estimator.

α_0	α_1	α_2	Estimators
1	0	0	$\widehat{C_y}$
0	1	0	$\widehat{C_y}\exp\left[sin\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right]$
0	0	1	$\widehat{C_y}\exp\left[sin\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right)\right]$

Expressing Equation (5.2) in terms of error terms (e's) through first-order approximation, the expression is obtaining as

$$t_h = \alpha_0 C_y(1+e_2)^{\frac{1}{2}}(1+e_0)^{-1} + \alpha_1 C_y(1+e_2)^{\frac{1}{2}}(1+e_0)^{-1}\exp\left[sin\left(-\frac{e_3}{2+e_3}\right)\right] + \alpha_2 C_y(1+e_2)^{\frac{1}{2}}(1+e_0)^{-1}\exp\left[sin\left(\frac{e_3}{2+e_3}\right)\right] \quad (5.4)$$

After simplification, the expression reduces to the following form:

$$t_h = C_y\left[1 + e_0 + \frac{1}{2}e_2 + e_0^2 - e_0e_2 - \frac{1}{8}e_2^2 + (\alpha_2 - \alpha_1)\frac{1}{2}e_3 + \left(\frac{3}{8}\alpha_1 - \frac{1}{8}\alpha_2\right)e_3^2 - (\alpha_2 - \alpha_1)\frac{1}{4}e_0e_3 + (\alpha_2 - \alpha_1)\frac{1}{4}e_2e_3\right] \quad (5.5)$$

By subtracting the population CV from both sides of Equation (5.5) and taking expectations, we derive the bias expression as-

$$Bias(t_h) = \gamma C_y\left[C_y^2 - C_y\lambda_{30} - \frac{1}{8}(\lambda_{40}-1) - H_1C_y\lambda_{12} + \frac{1}{2}H_1(\lambda_{22}-1) + (3\alpha_1 - \alpha_2)\frac{1}{8}\gamma(\lambda_{04}-1)\right] \quad (5.6)$$

Where,

$$H_1 = \left(\frac{1}{2}\alpha_2 - \frac{1}{2}\alpha_1\right) \quad (5.7)$$

From the equation (5.5), we can write after ignoring higher order term as follows-

$$(t_h - C_y) \cong C_y \left[\frac{1}{2}e_2 - e_0 + H_1 e_3\right] \quad (5.8)$$

By squaring both sides of equation (5.8) and then taking the expectation, we obtain the mean squared error (MSE) of the estimator, up to the first-order approximation, as follows:

$$MSE(t_h) = C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + H_1^2(\lambda_{40}-1) - C_y\lambda_{30} + H_1(\lambda_{22}-1) - 2H_1C_y\lambda_{12}\right] \quad (5.9)$$

Equation (5.9) will be minimum when,

$$H_1 = \left(\frac{C_y\lambda_{12}}{\lambda_{04}-1}\right) - \frac{1}{2}\left(\frac{\lambda_{22}-1}{\lambda_{04}-1}\right) \quad (5.10)$$

Putting the value of H_1 in equation we get Min. MSE of the proposed estimator t_h.

$$MSE(t_h) = C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + H_1^2(\lambda_{40}-1) - C_y\lambda_{30} + H_1(\lambda_{22}-1) - 2H_1C_y\lambda_{12}\right] \quad (5.11)$$

From equation (5.7) and (5.10)

$$H_1 = \left(\frac{1}{2}\alpha_2 - \frac{1}{2}\alpha_1\right) = \frac{C_y\lambda_{12}}{\lambda_{04}-1} - \frac{1}{2}\frac{(\lambda_{22}-1)}{\lambda_{04}-1} \quad (5.12)$$

From equations (5.3) and (5.12), we have only two equations with three unknowns. It is not possible to find unique values for $\alpha_i\text{'}s\ (i = 0,1,2)$. To determine the values of the unknowns, we impose a linear restriction as follows:

$$\sum_{i=1}^{2} \alpha_i B(t_i) = 0 \quad (5.13)$$

Such that

$$\alpha_0 B(t_0) + \alpha_1 B(t_1) + \alpha_2 B(t_2) = 0 \quad (5.14)$$

Here $B(t_i)$ denotes the bias in the $i^{th} i = 0, 1, 2$ estimator.

Equation can be written in the matrix from as,

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ B(t_0) & B(t_1) & B(t_2) \end{bmatrix}\begin{bmatrix}\alpha_0 \\ \alpha_1 \\ \alpha_2\end{bmatrix} = \begin{bmatrix} 1 \\ H_1 \\ 0\end{bmatrix} \quad (5.15)$$

Where $B(t_0)$, $B(t_1)$ and $B(t_2)$ are defined in equation (4.2), (4.11) and (4.14).

From the system of equation , we get the unique values of $\alpha_i's$.

$$\alpha_0 = \frac{\left(-\frac{1}{2}B(t_2) - \frac{1}{2}B(t_1) - H_1B(t_2) + H_1B(t_1)\right)}{B(t_0) - \frac{1}{2}B(t_1) - \frac{1}{2}B(t_2)} \quad (5.16)$$

$$\alpha_1 = \frac{\left(H_1B(t_2) + \frac{1}{2}B(t_0) - H_1B(t_0)\right)}{B(t_0) - \frac{1}{2}B(t_1) - \frac{1}{2}B(t_2)} \quad (5.17)$$

$$\alpha_2 = \frac{\left(-H_1B(t_1) + H_1B(t_0) + \frac{1}{2}B(t_0)\right)}{B(t_0) - \frac{1}{2}B(t_1) - \frac{1}{2}B(t_2)} \quad (5.18)$$

Where,

$$\alpha_0 + \alpha_1 + \alpha_2 = 1 \quad (5.19)$$

The use of these $\alpha_i's(for\ i = 0,1\ and\ 2)$eliminates the bias of the proposed estimator up to the order of $O(n^{-1})$.

VI THEORETICAL COMPARISON

We have compared the proposed estimator t_h with all other existing estimator considered here,

1. By (5.11) and (4.3) estimator t_h will be more efficient than t_0 if,
   $Min.MSE(t_h) < Min.MSE(t_0)$ i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 - \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30}\right] \quad (6.1)$$
2. By (5.11) and (4.6) estimator t_h will be more efficient than $t_1$if,
   $Min.MSE(t_h) < Min.MSE(t_1)$ i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + (\lambda_{04}-1) - C_y\lambda_{30} - (\lambda_{22}-1) + 2C_y\lambda_{12}\right] \quad (6.2)$$
3. By (5.11) and (4.9) estimator t_h will be more efficient than t_2 if,
   $Min.MSE(t_h) < Min.MSE(t_2)$i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + (\lambda_{04}-1) - C_y\lambda_{30} + (\lambda_{22}-1) - 2C_y\lambda_{12}\right] \quad (6.3)$$
4. By (5.11) and (4.12) estimator t_h will be more efficient than $t_3$if,
   $Min.MSE(t_h) < Min.MSE(t_3)$i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + \frac{1}{4}(\lambda_{04}-1) - \frac{1}{2}(\lambda_{22}-1) - C_y\lambda_{12} + C_{y\lambda_{30}}\right] \quad (6.4)$$
5. By (5.11) and (4.15) estimator t_h will be more efficient than t_4 if ,
   $Min.MSE(t_h) < Min.MSE(t_4)$ i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 + \frac{1}{4}(\lambda_{40}-1) + \frac{1}{4}(\lambda_{04}-1) + \frac{1}{2}(\lambda_{22}-1) - C_y\lambda_{12} - C_{y\lambda_{30}}\right] \quad (6.5)$$
6. By (5.11) and (4.18) estimator t_h will be more efficient than $t_5$if,
   $Min.MSE(t_h) < Min.MSE(t_5)$i.e.
   $$Min.MSE(t_h) < C_y^2\gamma\left[C_y^2 + \frac{1}{2}(\lambda_{40}-1) + k^2(\lambda_{04}-1) - C_y\lambda_{30} - k(\lambda_{22}-1) + 2kC_y\lambda_{12}\right] \quad (6.6)$$

VII EMPIRICAL STUDY

We use the following data sets for empirical study:

Data 1: [Source: Singh S. (2003), p.1116]

X: Number of fish caught in year 1993, Y: Number of fish caught in year 1995.

The data statistics are as follows:

N=69, n=40 ,C_x= 1.38, C_y=1.35, ρ =0.96, λ_{21}=2.19, λ_{12}=2.3, λ_{40}=7.66, λ_{04}=9.84, λ_{30}=1.11,λ_{03}=2.52, λ_{22}=8.19, $\bar{Y}$=4514.89, $\bar{X}$=4591.07

Table 2. The Bias, MSE and PRE of the existing and proposed estimators in case of data set 1

Estimator	Bias	MSE	PRE
t_0	0.0034	0.0381	100
t_1	0.0068	0.1886	20.1949
t_2	0.0104	0.2261	16.8431
t_3	0.0470	0.0710	53.6263
t_4	-0.0088	0.0898	42.4184
t_5	0.00303	0.0375	101.4480
$t_h(Min.)$	0.0000	0.0371	102.6448

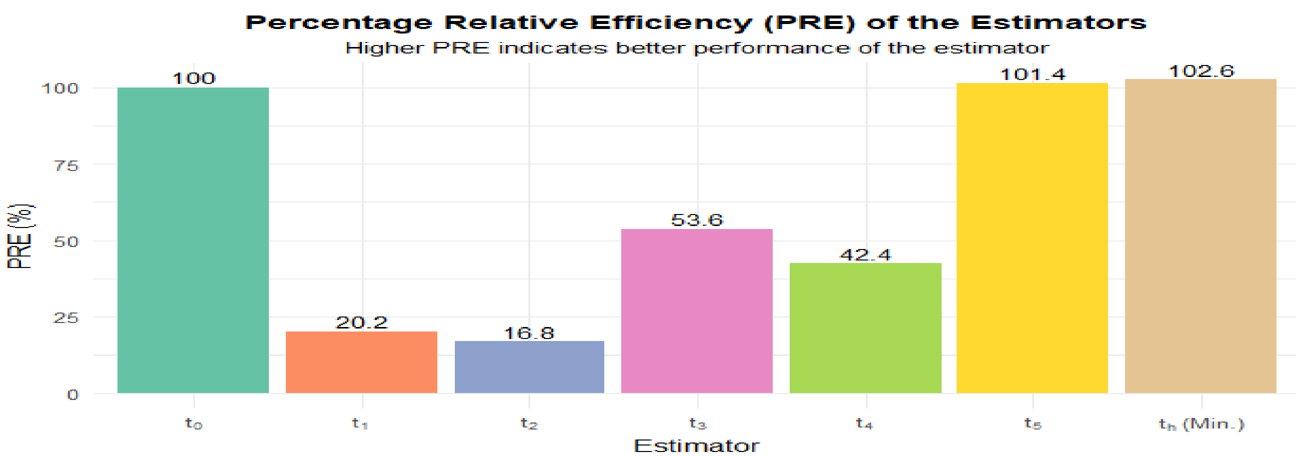

Fig.1 This fig. shows the bar graph of proposed novel sine type estimators with PRE values.

Data 2: [Source: Murthy M.N. (1967), p.399]
X: Area under wheat in 1963, Y: Area under wheat in 1964.
The data statistics are as follows:
N=34, n=15 ,C_x= 0.72, C_y=0.75, ρ=0.98, λ_{21}=1.0045, λ_{12}=0.9406, λ_{40}=3.6161, λ_{04}=2.8266, λ_{30}=1.1128, λ_{03}=0.9206,λ_{22}=3.01133, $\overline{Y}$ =199.44, $\overline{X}$ =208.88

Table 3. The Bias, MSE and PRE of the existing and proposed estimators in case of Data set 2.

Estimators	Bias	MSE	PRE
t_0	-0.0051	0.0080	100.00
t_1	-0.0135	0.0337	23.75
t_2	0.0033	0.0589	13.60
t_3	-0.0065	0.0113	70.94
t_4	-0.0073	0.0239	33.54
t_5	-0.00646	0.00697	114.8362
$t_{h(Min)}$	0.0000	0.00697	114.8362

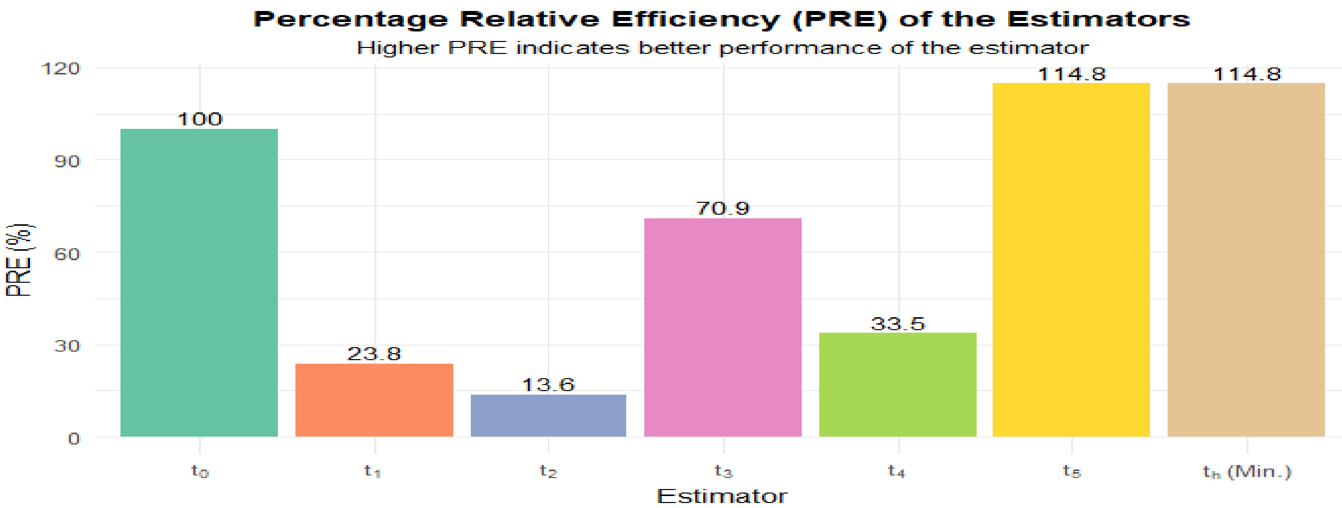

Fig 2. This fig. shows the bar graph of proposed novel sine type estimators with PRE values.

Table 4: Values of $\alpha_i (i = 0, 1, 2)$.

Scalars	Data 1	Data 2
α_0	1.4150	3.7370
α_1	-0.1521	-1.2041
α_2	-0.2629	-1.5329
$H_1 = \frac{1}{2}(\alpha_2 - \alpha_1)$	-0.0554	-0.0554

Using these values of $\alpha_i (i = 0, 1, 2)$ given in the table , one can reduce the bias up to the order $O(n^{-1})$ in estimator t_h.

VIII SIMULATION STUDY

In this section we have done simulation study to find the relative efficiency and there mean square errors of the proposed estimators i.e. almost unbiased estimators.

The following steps have been used for the simulation:

1. We have generated bivariate random observations of size N=1000 units from a bivariate normal distribution with parameters μ_x =3, σ_x=2, and μ_y =5, σ_y = 3 and ρ =0.80, 0.90.
2. A sample of size n = 200, 300 and 400 has been selected from this simulated population.
3. Sample statistics that is the sample mean, sample variance, and the values of the suggested and existing estimators of population mean are calculated for this sample.
4. Steps (2) and (3) are repeated m=10, 000 times.
5. The MSE of every estimator is calculated through the minimum MSE formula,
6. We use the following expression to obtain the percent relative efficiency (PRE)

$$PRE = \left(\frac{MSE(t_0)}{MSE(t_i) \text{ or } MSE(t_{estimators})} \right) * 100 \text{ for } i = 0, 1, 2, 3, 4, 5.$$

Table 5. The Bias, MSE and PRE of the proposed sine type estimators in case of Simulation study

N = 1000					
ρ	**n**	**Estimator**	**Bias**	**MSE**	**PRE**
	200	t_0	0.0003	0.0011	100
		t_1	-0.0009	0.0021	51.1056
		t_2	0.0015	0.0049	21.6457
		t_3	0.0012	0.0009	110.3397
		t_4	0.0003	0.0024	44.6857

		t_5	-0.0001	0.0009	124.0248
		$t_{h(Min)}$	0.000	0.0009	124.0248
0.8	300	t_0	0.00009	0.0007	100
		t_1	-0.0008	0.0012	59.4219
		t_2	0.00101	0.0033	21.2563
		t_3	0.0006	0.0006	126.1003
		t_4	0.0002	0.0016	43.4061
		t_5	-0.0002	0.0005	135.1579
		$t_{h(Min)}$	0.0000	0.0005	135.1579
	400	t_0	0.00006	0.0004	100
		t_1	-0.0005	0.0007	57.7696
		t_2	0.0007	0.0021	19.2557
		t_3	0.0004	0.0003	133.3383
		t_4	0.0002	0.0010	40.3047
		t_5	-0.0002	0.0003	143.7376
		$t_{h(Min)}$	0.0000	0.0003	143.7376

Table 6: The Bias, MSE and PRE of the proposed sine type estimators in case of Simulation study.

N = 1000					
ρ	**n**	**Estimator**	**Bias**	**MSE**	**PRE**
0.9	200	t_0	0.0002	0.0011	100
		t_1	-0.0014	0.0018	60.8545
		t_2	0.0018	0.0055	19.6737
		t_3	0.0011	0.0008	136.8257
		t_4	0.0005	0.0026	40.8048
		t_5	-0.0004	0.0007	145.5433
		$t_{h(Min)}$	0.0000	0.0007	145.5433
	300	t_0	0.0002	0.0008	100
		t_1	-0.0011	0.0012	64.9499
		t_2	0.0014	0.0043	18.2620
		t_3	0.0007	0.0005	155.5487
		t_4	0.0004	0.0021	38.2996
		t_5	-0.0003	0.0005	162.9053
		$t_{h(Min)}$	0.0000	0.0005	162.9053
	400	t_0	0.0001	0.0005	100
		t_1	-0.0008	0.0007	78.5802
		t_2	0.0010	0.0028	19.1422

		t_3	0.0004	0.0003	174.1520
		t_4	0.0003	0.0014	39.2164
		t_5	-0.0003	0.0003	176.6944
		$t_{h(Min)}$	0.0000	0.0003	176.6944

Table 7: Values of $\alpha_i(i = 0, 1, 2)$

n	Scalars	(N= 1000, ρ =0.80)	(N= 1000, ρ=0.90)
200	α_0	1.9673	1.8129
	α_1	-0.1927	-0.0426
	α_2	-0.7746	-0.7704
	H_1	-0.2910	-0.3639
	k	0.2910	0.3639
300	α_0	1.7025	1.7347
	α_1	-0.0069	0.0251
	α_2	-0.6956	-0.7598
	H_1	-0.3444	-0.3924
	k	0.3444	0.3924
400	α_0	1.7050	1.6940
	α_1	-0.0009	0.0924
	α_2	-0.7040	-0.7864
	H_1	-0.3515	-0.4394
	k	0.3515	0.4394

Using these values of $\alpha_i(i = 0, 1, 2)$ given in the table , one can reduce the bias up to the order $O(n^{-1})$ in estimator t_h .

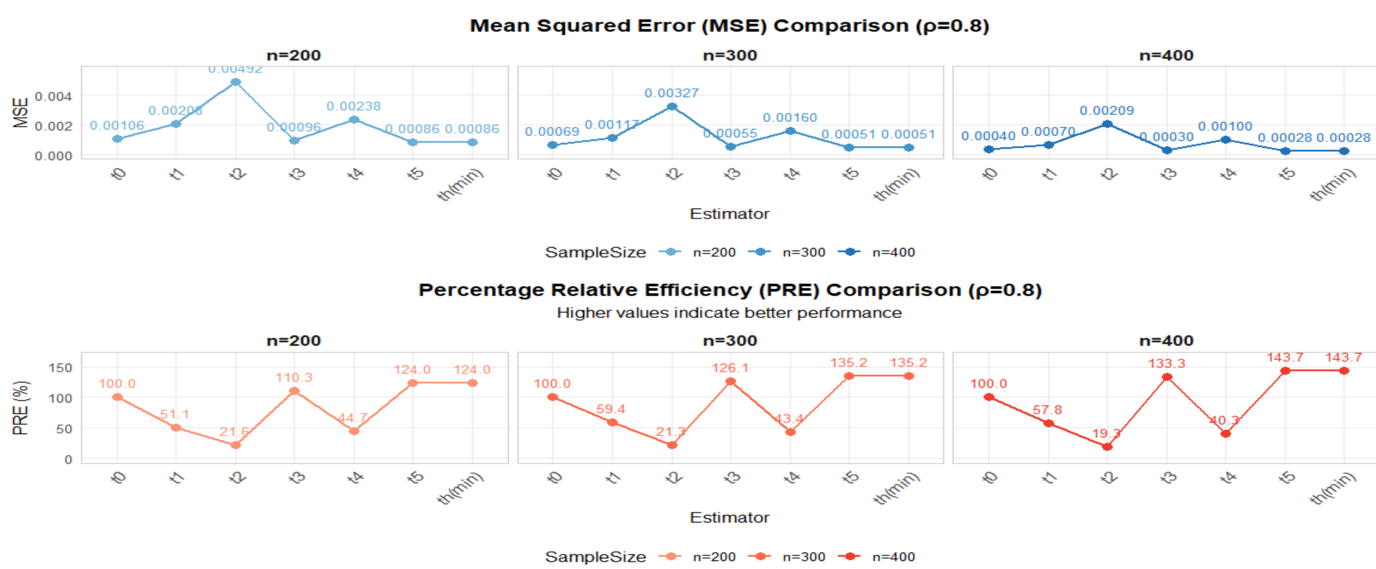

Fig 3. This fig. shows the line graph of novel proposed sine type estimators with their MSE and PRE values.

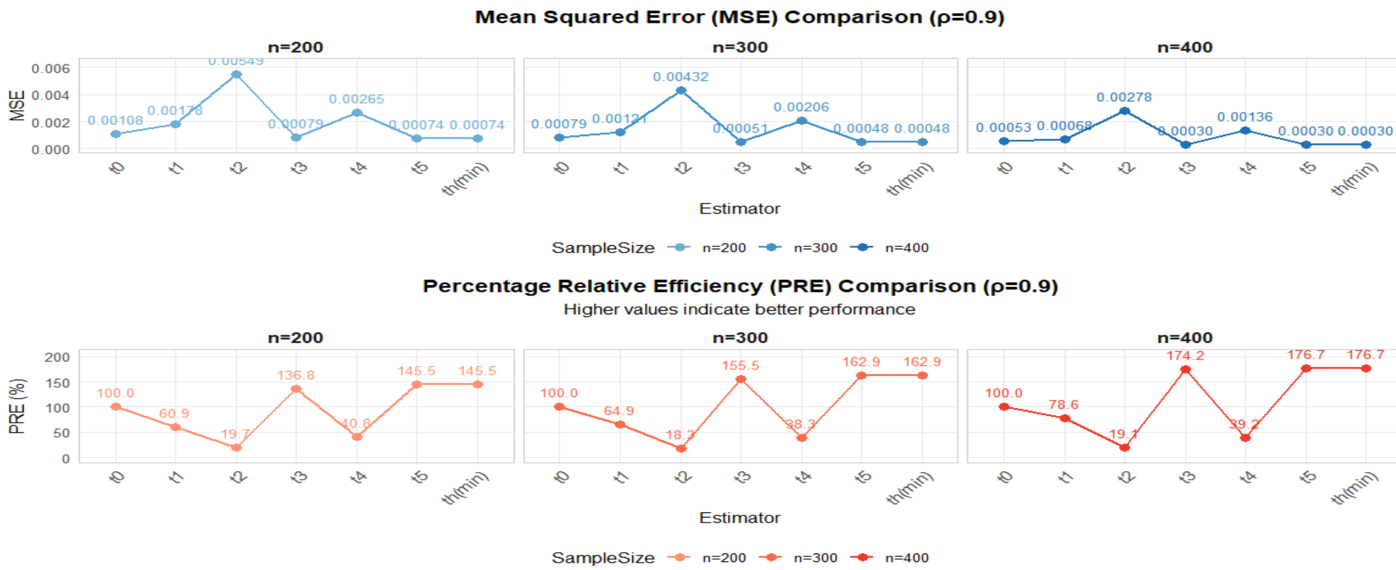

Fig 4. This fig. shows the line graph of novel proposed sine type estimators with their MSE and PRE values.

IX RESULTS AND DISCUSSION

This paper systematically evaluates the performance of the proposed estimator $t_{h(min.)}$ for the CV against existing estimators using both real datasets (fish catch and wheat area) and simulated bivariate normal data. Results consistently demonstrate that $t_{h(min.)}$ achieves the lowest mean square error (MSE) across all scenarios, with highest percentage relative efficiency (PRE). In the fish data, the estimator t_5 and $t_{h(min.)}$ proves the best estimators, as it shows the lowest mean square error (MSE) of 0.0371 and the highest per cent relative efficiency (PRE) of 102.64. Similarly, for the wheat data, estimators t_5 and $t_{h(min.)}$ shows the best performance, with a mean square error of 0.00697 and a percent relative efficiency of 114.83, respectively.

The study also includes a simulation using artificially generated data from a bivariate normal distribution. It explores sample sizes of 200, 300, and 400 under two levels of correlation ρ=0.80 and ρ=0.90. The findings show that estimators perform better when the correlation is higher and the sample size increases. In every scenario, $t_{h(min.)}$ and t_5 consistently perform the best, showing the smallest mean square error and the highest percent relative efficiency. At the largest sample size and highest correlation value (400, ρ=0.90), estimator $t_{h(min.)}$ performs best with a PRE of 176.69.

The graphs highlight that the proposed estimator $t_{h(min.)}$ and t_5 are the most efficient across diverse scenarios, making them preferable choices for population mean estimation. The line graph clearly illustrates that $t_{h(min.)}$ and t_5 are optimal, especially with larger samples and higher correlations.

X CONCLUSION

This paper introduces a new class of sine-type estimators for the efficient estimation of the CV in the context of SRSWOR, utilizing known auxiliary information. By incorporating sine and exponential transformations into traditional ratio, product, and difference-type estimator, the study enhances estimation precision significantly. Additionally, an almost unbiased estimator is developed by linearly combining selected exponential-sine estimators for minimizing bias up to the first order approximation.

Theoretical analysis of bias and MSE show that the new estimators are superior to existing estimators. These findings are confirmed by tests on real datasets like fish catch and wheat area, and by extensive simulations with different sample sizes and correlation values. The almost unbiased estimator t_h and the sine-type difference estimator t_5 have the lowest MSE and highest PRE, making them strong choices for real scenario. These results suggest that the proposed sine-type and exponential-sine estimators, especially the almost unbiased estimator, offer a valuable advancement in the estimation of population CV. These estimators are particularly effective when the auxiliary variable is correlated with the study variable, making them a preferred choice for practitioners in survey sampling.

REFERENCES

Srivastava, S. K. (1967). An estimator using auxiliary information in sample surveys. *Calcutta Statistical Association Bulletin, 16*(2–3), 121–132.

Singh, H. P., and Solanki, R. S. (2012). An efficient class of estimators for the population mean using auxiliary information in systematic sampling. *Journal of Statistical Theory and Practice, 6*, 274–285.

Sahai, A., and Ray, S. K. (1980). An efficient estimator using auxiliary information. *Metrika, 27*(1), 271–275.

Srivastava, S. K., and Jhajj, H. S. (1981). A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika, 68*(1), 341–343.

Singh, R., and Kumar, M. (2011). A note on transformations on auxiliary variable in survey sampling. *Model Assisted Statistics and Applications, 6*(1), 17–19.

Sharma, P., Verma, H., Sanaullah, A., and Singh, R. (2013). Some exponential ratio product type estimators using information on auxiliary attributes under second order approximation. *International Journal of Statistics and Economics, 12*(3), 58–66.

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

Jhajj, H. S., Sharma, M. K., and Grover, L. K. (2006). A family of estimators of population mean using information on auxiliary attribute. *Pakistan Journal of Statistics, 22*(1), 43–50.

Naik, V. D., and Gupta, P. C. (1996). A note on estimation of mean with known population proportion of an auxiliary character. *Journal of the Indian Society of Agricultural Statistics, 48*(2), 151–158.

Das, A. K., and Tripathi, T. P. (1992). Use of auxiliary information in estimating the coefficient of variation. *Aligarh Journal of Statistics, 12*, 51–58.

Patel, P. A., and Rina, S. (2009). A Monte Carlo comparison of some suggested estimators of coefficient of variation in finite population. *Journal of Statistics Sciences, 1*(2), 137–147.

Singh, R., Mishra, M., Singh, B. P., Singh, P., and Adichwal, N. K. (2018). Improved estimators for population coefficient of variation using auxiliary variable. *Journal of Statistics and Management Systems, 21*(7), 1335–1355.

Singh, R., and Yadav, S. K. (2024). Efficient estimators of population coefficient of variation under simple random sampling using single auxiliary variable. *Journal of Statistics Applications & Probability Letters, 11*(3), 109–121.

Archana, V., and Rao, A. (2014). Some improved estimators of coefficient of variation from bi-variate normal distribution: A Monte Carlo comparison. *Pakistan Journal of Statistics and Operations Research*, 87–105.

Singh, R., and Kumari, A. (2022). Improved estimators of population coefficient of variation under simple random sampling. *Asian Journal of Probability and Statistics, 19*(4), 22–36.

Bhattacharyya, D., Singh, G. N., Jawa, T. M., Sayed-Ahmed, N., and Pandey, A. K. (2021). An exponential-cum-sine-type hybrid imputation technique for missing data. *Computational Intelligence and Neuroscience, 2021*.

Yunusa, M. A., Muili, J. O., Audu, A., and Singh, R. V. K. (2023). A sine type median based estimator for the estimation of population mean. *Oriental Journal of Physical Sciences, 8*(1), 21–26.

Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *Journal of Agricultural Science, 30*(2), 262–275.

Murthy, M. N. (1964). Product method of estimation. *Sankhya: The Indian Journal of Statistics, Series A, 26*(1), 69–74.

Bahl, S., and Tuteja, R. K. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences, 12*(1), 159–164.

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

Singh, S., and Singh, R. (1993). Almost filtration of bias precipitates: A new approach. *Journal of the Indian Society of Agricultural Statistics, 45*(2), 214–218.

Singh, S., and Singh, R. (1991). Almost filtration of bias precipitates: A new technique. *Aligarh Journal of Statistics, 11*, 5–8.

Singh, R., Mishra, M., Singh, B. P., Singh, P., and Adichwal, N. K. (2018). Improved estimators for population coefficient of variation using auxiliary variable. *Journal of Statistics and Management Systems, 21*(7), 1335–1355.

Singh, S. (2003). *Advanced sampling theory with applications: How "Michael" selected "Amy"* (Vol. 2). Springer.

Murthy, M. N. (1967). *Sampling theory and methods*. Statistical Publishing Society.