

General Parameter(s) Estimation using Sine Function-Based Estimators

Shobh Nath Tiwari

Banaras Hindu University <https://orcid.org/0009-0006-5930-7973>

Rajesh Singh

Banaras Hindu University

Madhulika Mishra

madhulika1707@gmail.com

Symbiosis International University

Research Article

Keywords: Auxiliary variable, Sine-type estimator, Bias, PRE, Simulation study, MSE

Posted Date: September 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11201277/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

General Parameter(s) Estimation using Sine Function-Based Estimators

Shobh Nath Tiwari¹, Rajesh Singh¹, Madhulika Mishra^{2,*}

¹Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005,
Uttar Pradesh, India

²Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune,
Maharashtra, India

Emails: shobh7285@gmail.com; rsinghstat@bhu.ac.in; madhulika1707@gmail.com

*Corresponding: madhulika1707@gmail.com

Abstract

In sampling, the estimation of population parameters constitutes a vital component of statistical analysis. Many researchers have developed various estimation methods for population parameters by incorporating auxiliary data. This study proposes a new sine-based difference estimator to estimate the general parameter utilizing supplementary data under simple random sampling without replacement (SRSWOR). A traditional estimator $t_{(p,q)}$ is employed for the assessment of population parameters such as the population mean, population mean square, coefficient of variation, and standard deviation. The mean squared error (MSE) and bias for the developed estimator are derived using first order approximation methods. We performed both empirical and simulation analyses to assess the efficiency of the proposed estimator and to validate the theoretical findings. Theoretical derivations and empirical analyses demonstrate that the suggested estimator exhibits improved accuracy over existing and competing sine estimators examined in this study.

Keywords: Auxiliary variable, Sine-type estimator, Bias, PRE, Simulation study, MSE.

1. Introduction

In sampling theory, the integration of auxiliary variables with study variables is widely employed to enhance the precision of population parameter estimation. Several techniques exist for utilizing auxiliary information to enhance the performance of estimators. These approaches rely on selecting a sample that adequately represents the population, and if the population is relatively homogeneous, unit selection can be carried out through simple random sampling. Various authors have carried out substantial work on population parameter estimation utilizing auxiliary information, for instance, see Srivastava [1], Das and Tripathi [2], Kadilar and Cingi [3], Singh et al. [4] and Gupta and Shabbir [5]. In their 1997 study, Rizvi and Singh [6] examined the problem of population mean estimation and developed several estimator classes, highlighting their properties. Solanki and Singh [7], Pal and Singh [8], and Kumar Adichwal et al. [9] made use of known auxiliary variable to construct estimator concerning general parameters. Additional related studies can be found in Singh et al. [10], Srivastava and Jhaji [11-12], Singh and Tiwari [13], Yadav et al. [14], Abu-Dayyeh et al. [15], Almulhim et al. [16], Alzahrani and Almohaimed [17] and Aloraini et al. [18-19].

Although extensive work has been conducted on the estimation of population parameters (e.g., mean and variance), the development of sine-type estimators remains comparatively underexplored, even though sine functions provide powerful tools for representing periodic behaviors such as sound and light waves, tidal cycles, seasonal weather patterns, electrical oscillations, and cyclical patterns in stock prices (Bhattacharyya et al. [20]). Some significant studies have focused on sine-type techniques, such as the work by Bhattacharyya et al. [20] on a sine-exponential hybrid imputation approach, Yunusa et al. [21], who developed a sine-based median estimator. By leveraging known auxiliary information, Singh et al. [22] proposed sine-based estimators to estimate mean. Notably, the estimation of general population parameters using sine-type functions has remained largely unexplored. To address this gap and inspired by previous research, we introduce new sine-type estimators for estimating the general parameters of the variable of interest, leveraging existing auxiliary information under simple random sampling.

2. Symbols and Notations

Consider a finite population $W = (w_1, w_2, \dots, w_N)$ of N identifiable elements, from which a subset of n units are randomly selected using the SRSWOR technique. Let Z_i and Y_i denote the auxiliary and study population variables, respectively, for the i^{th} unit ($i = 1, 2, \dots, N$). Correspondingly, let z_i and y_i represent the values of the i^{th} sampled unit ($i = 1, 2, \dots, n$). The symbols and notations used throughout the study are defined in Table 1.

Table 1: Symbols and notations.

Symbols	Terminology
$\bar{Z} = N^{-1} \sum_{i=1}^N z_i$	The population mean of the auxiliary variable z
$\bar{Y} = N^{-1} \sum_{i=1}^N y_i$	The population mean of the study variable y
$\bar{z} = n^{-1} \sum_{i=1}^n z_i$	The sample mean of z
$\bar{y} = n^{-1} \sum_{i=1}^n y_i$	The sample mean of y
$S_z^2 = \mu_{02}$	The Variance of the auxiliary variable z
$S_y^2 = \mu_{20}$	The Variance of the study variable y
$C_z = S_z/\bar{Z}$	The coefficient of variation of z
$C_y = S_y/\bar{Y}$	The coefficient of variation of y
$S_{yz} = \mu_{11}$	The population covariance between y and z
$\rho_{zy} = S_{yz}/(S_z S_y)$	Correlation coefficient between y and z

$$\lambda_{mn} = \frac{\mu_{mn}}{(\mu_{20}^{m/2} \mu_{02}^{n/2})}, \mu_{mn} = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{Y})^m (z_i - \bar{Z})^n$$

and (m, n) being non-negative integers.

Further, we define,

$$\bar{y} = \bar{Y}(1 + \eta_0), S_y^2 = s_y^2(1 + \eta_1), \bar{z} = \bar{Z}(1 + \eta_2), S_z^2 = s_z^2(1 + \eta_3)$$

such that

$$E(\eta_i) = 0, \forall i = 0, 1, 2, 3.$$

Considering the first-order approximation and ignoring fpc (finite population correction) term:

$$\begin{aligned} E(\eta_0^2) &= n^{-1}C_y^2, E(\eta_1^2) = n^{-1}(\lambda_{40} - 1), E(\eta_2^2) = n^{-1}C_z^2, E(\eta_3^2) = n^{-1}(\lambda_{04} - 1), \\ E(\eta_0\eta_1) &= n^{-1}\lambda_{30}C_y, E(\eta_0\eta_2) = n^{-1}\rho_{yz}C_zC_y, E(\eta_0\eta_3) = n^{-1}\lambda_{12}C_y, E(\eta_1\eta_2) = n^{-1}\lambda_{21}C_z, \\ E(\eta_1\eta_3) &= n^{-1}(\lambda_{22} - 1), E(\eta_2\eta_3) = n^{-1}\lambda_{03}C_z. \end{aligned}$$

3. Description of the general parameter

The general parameter is formulated as:

$$t_{(p,q)} = \bar{Y}^p S_y^q. \quad (3.1)$$

Where; $(p, q) = (-1, 1), (0, 1), (0, 2), (1, 0)$.

It should be noted that $\bar{y}$ and s_y^2 yield unbiased estimations of $\bar{Y}$ and S_y^2 , respectively.

When p and q take certain values, the general parameter $t_{(p,q)}$ can be written as:

- i. Taking $p = 1$ and $q = 0$, the parameter $t_{(p,q)} \rightarrow t_{(1,0)} = \bar{Y}$.
- ii. Taking $p = 0$ and $q = 2$, the parameter $t_{(p,q)} \rightarrow t_{(0,2)} = S_y^2$.
- iii. Taking $p = -1$ and $q = 1$, the parameter $t_{(p,q)} \rightarrow t_{(-1,1)} = C_y$.
- iv. Taking $p = 0$ and $q = 1$, the parameter $t_{(p,q)} \rightarrow t_{(0,1)} = S_y$.

The conventional estimator corresponding to the general parameter $t_{(p,q)}$ is denoted by;

$$\hat{t}_{(p,q)} = \bar{y}^p s_y^q \quad (3.2)$$

Where, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, and $s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$.

By expressing Eq. (3.2) with respect to η' s, we have:

$$\hat{t}_{(p,q)} = \bar{Y}^p S_y^q (1 + \eta_0)^p (1 + \eta_1)^{\frac{q}{2}} \quad (3.3)$$

Expanding Eq. (3.3), we obtain:

$$\hat{t}_{(p,q)} = t_{(p,q)} \left(1 + p\eta_0 + \frac{q}{2}\eta_1 + \frac{p(p-1)}{2}\eta_0^2 + \frac{pq}{2}\eta_0\eta_1 + \frac{q(q-2)}{8}\eta_1^2 \right)$$

Or

$$\hat{t}_{(p,q)} - t_{(p,q)} = t_{(p,q)} \left(p\eta_0 + \frac{q}{2}\eta_1 + \frac{p(p-1)}{2}\eta_0^2 + \frac{pq}{2}\eta_0\eta_1 + \frac{q(q-2)}{8}\eta_1^2 \right) \quad (3.4)$$

By taking the square of Eq. (3.4) and discarding higher-order terms in η' s, we get:

$$(\hat{t}_{(p,q)} - t_{(p,q)})^2 = t_{(p,q)}^2 \left(p^2\eta_0^2 + \frac{q^2}{4}\eta_1^2 + pq\eta_0\eta_1 \right) \quad (3.5)$$

Applying the expectation to Eq. (3.5) provides:

$$MSE(\hat{t}_{(p,q)}) = \frac{t_{(p,q)}^2}{n} \left(p^2 C_y^2 + \frac{q^2}{4} (\lambda_{40} - 1) + pq \lambda_{30} C_y \right). \quad (3.6)$$

4. Proposed estimator

Inspired by earlier studies of Singh et al. [22] and Rao [23], we introduce a new class of estimators for estimating population parameter $t_{(p,q)}$ as follows:

$$t^* = [\hat{t}_{(p,q)} + k(\bar{Z} - \bar{z})] \left[1 + \sin \left\{ \frac{w_1(\bar{Z} - \bar{z})}{\bar{z}} \right\} \right] \left[1 + \sin \left\{ \frac{w_2(S_z^2 - s_z^2)}{s_z^2} \right\} \right] \quad (4.1)$$

In this formulation, w_1 and w_2 are scalars taking values (0, 1), while k is selected in such a way that MSE of developed estimator (t^*) is minimized.

By expanding Eq. (4.1) in terms of the error components (η 's) up to the first-order, we obtain the following expression:

$$t^* = \left[t_{(p,q)}(1 + \eta_0)^p(1 + \eta_1)^{\frac{q}{2}} - k\bar{Z}\eta_2 \right] [(1 - w_1\eta_2)(1 - w_2\eta_3)] \quad (4.2)$$

By carrying out the expansion of Eq. (4.2), we get:

$$t^* = \left[\left\{ t_{(p,q)} \left(1 + p\eta_0 + \frac{q}{2}\eta_1 + \frac{p(p-1)}{2}\eta_0^2 + \frac{pq}{2}\eta_0\eta_1 + \frac{q(q-1)}{8}\eta_1^2 \right) - k\bar{Z}\eta_2 \right\} \right. \\ \left. \{ 1 - w_2\eta_3 - w_1\eta_2 + w_1w_2\eta_2\eta_3 \} \right] \quad (4.3)$$

On taking the difference with general parameter ($t_{(p,q)}$) in Eq. (4.3), we get:

$$t^* - t_{(p,q)} = \left[t_{(p,q)} \left\{ \begin{aligned} & p\eta_0 - w_1\eta_2 - w_2\eta_3 + w_1w_2\eta_2\eta_3 - pw_2\eta_0\eta_3 - pw_1\eta_0\eta_2 + \frac{q}{2}\eta_1 \\ & - \frac{q}{2}w_2\eta_1\eta_3 - \frac{q}{2}w_1\eta_1\eta_2 + \frac{p(p-1)}{2}\eta_0^2 + \frac{pq}{2}\eta_0\eta_1 + \frac{q(q-2)}{8}\eta_1^2 \\ & - \{ k\bar{Z}\eta_2 - kw_2\bar{Z}\eta_2\eta_3 - kw_1\bar{Z}\eta_2^2 \} \end{aligned} \right\} \right] \quad (4.4)$$

Evaluating the expected value of Eq. (4.4) gives:

$$Bias(t^*) = \left[t_{(p,q)} \left\{ \begin{aligned} & w_1w_2n^{-1}\lambda_{03}C_z - pw_2n^{-1}\lambda_{12}C_y - pw_1n^{-1}\rho C_zC_y - \frac{q}{2}w_2n^{-1}(\lambda_{22} - 1) \\ & - \frac{q}{2}w_1n^{-1}\lambda_{21}C_z + \frac{p(p-1)}{2}n^{-1}C_y^2 + \frac{pq}{2}n^{-1}\lambda_{30}C_y + \frac{q(q-2)}{8}n^{-1}(\lambda_{40} - 1) \\ & + \{ kw_2\bar{Z}n^{-1}\lambda_{03}C_z + kw_1\bar{Z}n^{-1}C_z^2 \} \end{aligned} \right\} \right] \quad (4.5)$$

By taking the square of Eq. (4.4) and dropping (η 's) terms of order higher than two, the following result is derived:

$$(t^* - t_{(p,q)})^2 = \left[t_{(p,q)} \left\{ p\eta_0 - w_1\eta_2 - w_2\eta_3 + \frac{q}{2}\eta_1 \right\} - k\bar{Z}\eta_2 \right]^2 \quad (4.6)$$

Applying the expectation to Eq. (4.6) results in:

$$MSE(t^*) = MSE(\hat{t}_{(p,q)}) + \frac{t_{(p,q)}^2}{n} \left[\begin{aligned} & w_1C_z^2 + w_2(\lambda_{04} - 1) + 2w_1w_2(\lambda_{03}C_z) - w_1(2ppC_zC_y + q\lambda_{21}C_z) \\ & - 2k\bar{Z}C_z^2t_{(p,q)}^{-1} - w_2(2p\lambda_{12}C_y + q(\lambda_{22} - 1) - 2k\bar{Z}\lambda_{03}C_zt_{(p,q)}^{-1}) \\ & + k^2\bar{Z}^2C_z^2t_{(p,q)}^{-2} - 2kp\bar{Z}\rho C_zC_yt_{(p,q)}^{-1} - qk\bar{Z}\lambda_{21}C_zt_{(p,q)}^{-1} \end{aligned} \right] \quad (4.7)$$

Or

$$MSE(t^*) = MSE(\hat{t}_{(p,q)}) + \frac{t_{(p,q)}^2}{n} [Aw_1^2 + Bw_2^2 + 2w_1w_2C - w_1D - w_2E + F]. \quad (4.8)$$

Where,

$$A = C_z^2.$$

$$B = \lambda_{04} - 1.$$

$$C = \lambda_{03} C_z.$$

$$D = 2p\rho C_z C_y + q\lambda_{21} C_z - 2k\bar{Z} C_z^2 t_{(p,q)}^{-1}.$$

$$E = 2p\lambda_{12} C_y + q(\lambda_{22} - 1) - 2k\bar{Z}\lambda_{03} C_z t_{(p,q)}^{-1}.$$

$$F = k^2 \bar{Z}^2 C_z^2 t_{(p,q)}^{-2} - 2kp\bar{Z}\rho C_z C_y t_{(p,q)}^{-1} - qk\bar{Z}\lambda_{21} C_z t_{(p,q)}^{-1}.$$

The optimal values of w_1 and w_2 that minimize the MSE of (t^*) are determined as,

$$w_{1(opt)} = \frac{(EC-BD)}{2(C^2-AB)} \text{ and } w_{2(opt)} = \frac{(DC-AE)}{2(C^2-AB)}. \quad (4.9)$$

Employing the optimal weights $w_{1(opt)}$ and $w_{2(opt)}$ from Eq. (4.9) in Eq. (4.8), the resulting expression is given by:

$$MSE(t^*)_{min} = MSE(\hat{t}_{(p,q)}) + \frac{t_{(p,q)}^2}{n} \left[F + \left\{ \frac{(AE^2 + BD^2 - 2CDE)}{4(C^2 - AB)} \right\} \right]. \quad (4.10)$$

Remark 4.1: Note that the expressions for the optimal values of (w_1, w_2) at Eq. (4.9) depend on unknown population parameters. These quantities may be estimated with reasonable accuracy from a pilot survey, available sample data, or prior knowledge acquired over time; see Das and Tripathi [2], Srivastava and Jhaji [11], Singh and Solanki [24] and Adichwal et al. [9].

When $\bar{Z}$ is given, the sine-based difference estimator for the population parameter $t_{(p,q)}$, is derived by substituting $w_2 = 0$ in Eq. (4.1), and is given by:

$$t_1^* = [\hat{t}_{(p,q)} + k(\bar{Z} - \bar{z})] \left[1 + \sin \left\{ \frac{w_1(\bar{Z} - \bar{z})}{\bar{Z}} \right\} \right] \quad (4.11)$$

Here, w_1 is a scalar assuming values $(0, 1)$, and k is selected for the purpose of minimizing the MSE of the estimator (t_1^*) .

The minimum MSE corresponding to estimator (t_1^*) is given by:

$$MSE(t_1^*)_{min} = MSE(\hat{t}_{(p,q)}) + \frac{t_{(p,q)}^2}{n} \left[F - \frac{D^2}{4A} \right]. \quad (4.12)$$

Table 2 provides various estimators formulated from Eq. (4.11), based on specific choices of the parameters p, q, k, w_1 and w_2 .

Table 2: Specific cases of the estimator t_1^* .

Estimators	p	q	k	w_1	w_2
$t_{1(0)}^* = \hat{t}_{(p,q)} \left[1 + \sin \left\{ \frac{(\bar{Z} - \bar{z})}{\bar{Z}} \right\} \right]$	p	q	0	1	0
$t_{1(1)}^* = \bar{y}$	1	0	0	0	0
$t_{1(2)}^* = \bar{y} \left[1 + \sin \left\{ \frac{(\bar{Z} - \bar{z})}{\bar{Z}} \right\} \right]$	1	0	0	1	0
$t_{1(3)}^* = [\bar{y} + k(\bar{Z} - \bar{z})] \left[1 + \sin \left\{ \frac{(\bar{Z} - \bar{z})}{\bar{Z}} \right\} \right]$	1	0	k	1	0

$t_{1(4)}^* = [\bar{y} + k(\bar{Z} - \bar{z})]$ (Difference Estimator)	1	0	k	0	0
--	---	---	-----	---	---

When S_Z^2 is known, the sine-type ratio estimator for the parameter $t_{(p,q)}$ is derived by assigning ($k = 0, w_1 = 0$) in Eq. (4.1) leads to:

$$t_2^* = \hat{t}_{(p,q)} \left[1 + \sin \left\{ \frac{w_2(S_Z^2 - s_z^2)}{S_Z^2} \right\} \right] \quad (4.13)$$

The scalar w_2 can take values 0 or 1, whereas k is determined such that the MSE of (t_2^*) reaches its minimum.

The expression for the least MSE of the estimator (t_2^*) is given by:

$$MSE(t_2^*)_{min} = MSE(\hat{t}_{(p,q)}) - \frac{t_{(p,q)}^2}{n} \left[\frac{E^2}{4B} \right]. \quad (4.14)$$

Table 3 provides various estimators formulated from Eq. (4.13), based on specific choices of the parameters p, q, k, w_1 and w_2 .

Table 3: Specific cases of the estimator t_2^* .

Estimators	p	q	k	w_1	w_2
$t_{2(1)}^* = s_y^2$	0	2	0	0	0
$t_{2(2)}^* = \hat{t}_{(p,q)} \left[1 + \sin \left\{ \frac{(S_Z^2 - s_z^2)}{S_Z^2} \right\} \right]$	p	q	0	0	1
$t_{2(3)}^* = s_y^2 \left[1 + \sin \left\{ \frac{(S_Z^2 - s_z^2)}{S_Z^2} \right\} \right]$	0	2	0	0	1

5. Efficiency comparison

$$MSE(\hat{t}_{(p,q)}) - MSE(t^*)_{min} = \frac{t_{(p,q)}^2}{n} \left[\left\{ \frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} \right\} - F \right] \geq 0. \quad (5.1)$$

$$MSE(t_1^*)_{min} - MSE(t^*)_{min} = \frac{t_{(p,q)}^2}{n} \left[\left\{ \frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} \right\} - \frac{D^2}{4A} \right] \geq 0. \quad (5.2)$$

$$MSE(t_2^*)_{min} - MSE(t^*)_{min} = \frac{t_{(p,q)}^2}{n} \left[\left\{ \frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} \right\} - \frac{E^2}{4B} - F \right] \geq 0. \quad (5.3)$$

It is clear from Eqs. (5.1) - (5.3) that suggested sine-based estimators achieve better performance than estimators $\hat{t}_{(p,q)}$, t_1^* and t_2^* .

6. Proposed estimator for the population mean

By assigning values $(p, q, k, w_1, w_2) = (1, 0, k, w_1, w_2)$ into Eq. (4.1), the estimator (t^*) formulated by:

$$t^@ = [\bar{y} + k(\bar{Z} - \bar{z})] \left[1 + \sin \left\{ \frac{w_1(\bar{Z} - \bar{z})}{\bar{Z}} \right\} \right] \left[1 + \sin \left\{ \frac{w_2(S_Z^2 - s_z^2)}{S_Z^2} \right\} \right] \quad (6.1)$$

MSE expression for estimator $t^@$, is presented by:

$$MSE(t^@) = MSE(\bar{y}) + \frac{\bar{y}^2}{n} [Aw_1^2 + Bw_2^2 + 2w_1w_2C - w_1D - w_2E + F]. \quad (6.2)$$

The minimum MSE of $(t^@)$ in Eq. (6.2) occurs at the values:

$$w_{1(opt)} = \frac{(EC-BD)}{2(C^2-AB)} \quad \text{and} \quad w_{2(opt)} = \frac{(DC-AE)}{2(C^2-AB)}. \quad (6.3)$$

Substituting the values of $w_{1(opt)}$ and $w_{2(opt)}$ into Eq. (6.2), we obtain:

$$MSE(t^@)_{min} = MSE(\bar{y}) + \frac{\bar{y}^2}{n} \left[F + \left\{ \frac{(AE^2 + BD^2 - 2CDE)}{4(C^2 - AB)} \right\} \right] \quad (6.4)$$

The minimum MSE values for the estimators $t_{1(1)}^*$, $t_{1(2)}^*$, $t_{1(3)}^*$ and $t_{1(4)}^*$ are expressed as:

$$MSE(t_{1(1)}^*) = Var(\bar{y}) = \frac{\bar{y}^2 C_y^2}{n}. \quad (6.5)$$

$$MSE(t_{1(2)}^*) = Var(\bar{y}) + \frac{\bar{y}^2}{n} [C_z^2 - 2\rho C_y C_z]. \quad (6.6)$$

$$MSE(t_{1(3)}^*) = Var(\bar{y}) - \frac{\bar{y}^2}{n} \rho^2 C_y^2. \quad (6.7)$$

$$MSE(t_{1(4)}^*) = Var(\bar{y}) - \frac{\bar{y}^2}{n} \rho^2 C_y^2. \quad (6.8)$$

7. Efficiency comparison

$$MSE(t_{1(1)}^*) - MSE(t^@)_{min} = \frac{\bar{y}^2}{n} \left[C_y^2 + \left\{ \frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} - F \right\} \right] \geq 0.$$

$$MSE(t_{1(2)}^*) - MSE(t^@)_{min} = \frac{\bar{y}^2}{n} \left[C_z^2 - 2\rho C_y C_z + \frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} - F \right] \geq 0.$$

$$MSE(t_{1(3)}^*) - MSE(t^@)_{min} = \frac{\bar{y}^2}{n} \left[\frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} - \frac{\bar{y}^2}{n} \rho^2 C_y^2 - F \right] \geq 0.$$

$$MSE(t_{1(4)}^*) - MSE(t^@)_{min} = \frac{\bar{y}^2}{n} \left[\frac{(2CDE - AE^2 - BD^2)}{4(C^2 - AB)} - \frac{\bar{y}^2}{n} \rho^2 C_y^2 - F \right] \geq 0.$$

8. Empirical illustration

The performance of the developed estimators is examined through two real-world datasets to establish their efficiency and practical relevance.

The following section provides a description of the population datasets:

Population 1. (Source: Suzuki Motor Corporation [25], (Feb 1 to Mar 25, 2025)).

Y : Daily high stock price of Suzuki Motor Corporation.

Z : Daily opening stock price of Suzuki Motor Corporation.

Descriptive statistics of the population 1 are shown as follows.

$$\bar{Z} = 49.449, \quad n = 14, \quad \bar{Y} = 51.002, \quad C_z = 0.031, \quad C_y = 0.026, \quad \rho_{zy} = 0.724, \quad \lambda_{12} = 0.3395, \\ \lambda_{03} = 0.3312, \quad \lambda_{04} = 2.6491.$$

Table 4: Comparison of MSEs and PREs for various estimators based on population 1.

Estimator	MSE	PRE
-----------	-----	-----

$t_{1(1)}^*$	0.122	100.000
$t_{1(2)}^*$	0.086	141.718
$t_{1(3)}^*$	0.058	210.098
$t_{1(4)}^*$	0.058	210.098
$t^@$	0.057	212.995

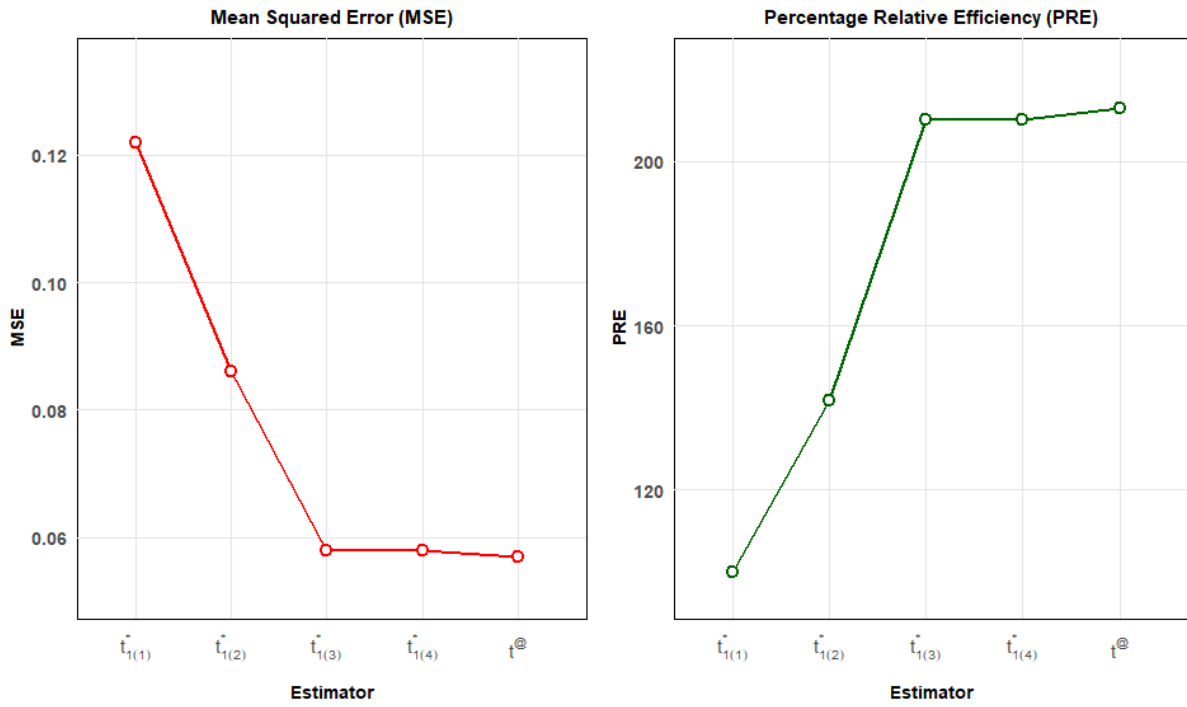


Figure 1: Graphical representation of the estimator's MSE and PRE results using population 1.

The PRE is computed through the following equation:

$$PRE = \frac{MSE(t_{1(1)}^*)}{MSE(\text{Sine based Estimator})} \times 100.$$

Population 2. (Source: Cochran [26], p.325)).

Y: Individuals in each block.

Z: Rooms in each block.

Descriptive statistics of the population 2 are shown as follows.

$\bar{Z} = 58.8$, $n = 10$, $\bar{Y} = 101.1$, $C_z = 0.1281$, $C_y = 0.1450$, $\rho_{zy} = 0.650$, $\lambda_{12} = 0.5714$, $\lambda_{03} = 0.4861$, $\lambda_{04} = 2.2387$.

Table 5: Comparison of MSEs and PREs for various estimators based on population 2.

Estimator	MSE	PRE
-----------	-----	-----

$t_{1(1)}^*$	21.490	100.000
$t_{1(2)}^*$	13.582	158.228
$t_{1(3)}^*$	12.411	173.160
$t_{1(4)}^*$	12.411	173.160
$t^@$	11.012	195.156

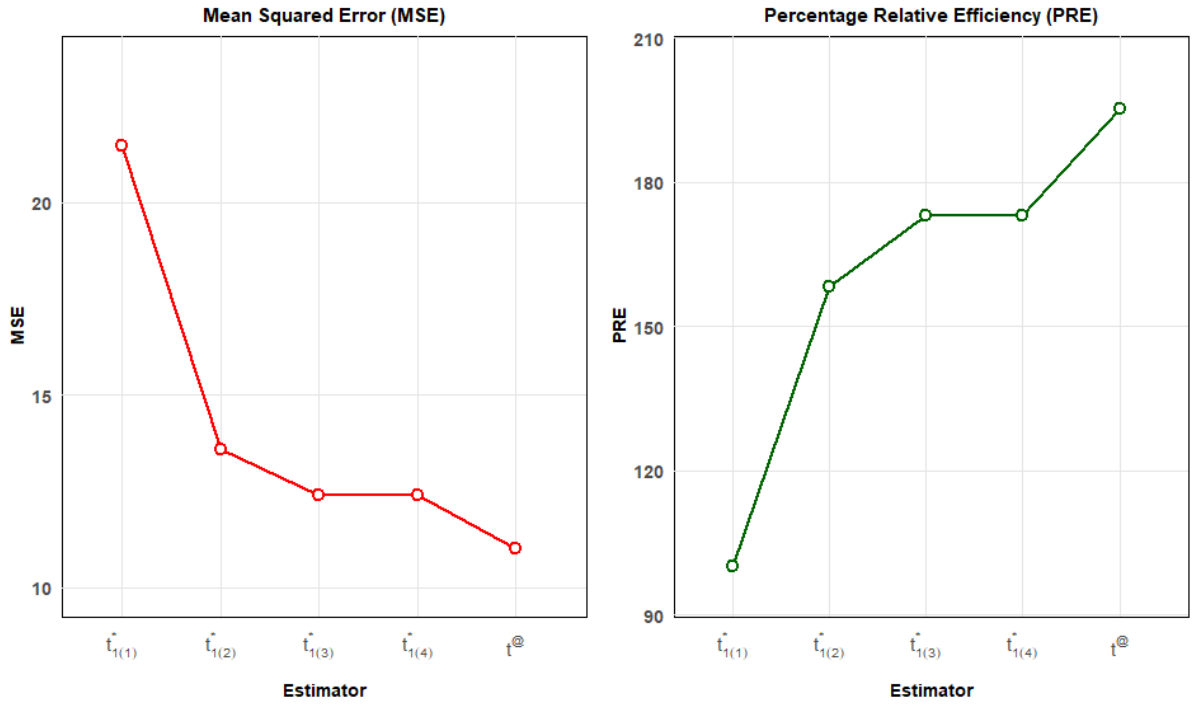


Figure 2: Graphical representation of the estimator's MSE and PRE results using population 2.

As shown in Table 4 and 5, the efficiency of proposed estimator $t^@$ surpasses that of conventional mean estimator $t_{1(1)}^*$ as well as the other comparable estimators $t_{1(2)}^*$, $t_{1(3)}^*$ and $t_{1(4)}^*$, given that its PRE is comparatively higher.

9. Simulation analysis

The following procedure computationally compares the estimators $t_{1(1)}^*$, $t_{1(2)}^*$, $t_{1(3)}^*$, and $t_{1(4)}^*$ to the estimator $t^@$, as implemented using the algorithm developed by Singh et al. [22]. The process of simulation is executed step by step as outlined below:

- The R software was utilized to generate simulated datasets of Z and Y , each consisting of $N=1000$ observations, drawn from two distributions - Bivariate Normal distribution, and Gamma distribution.
- Normal Distribution: defined by a mean vector $\mu = [24, 42]$ and covariance matrix Σ given as

$$\Sigma = \begin{bmatrix} \sigma_y^2 & \sigma_y \sigma_z \rho_{zy} \\ \sigma_y \sigma_z \rho_{zy} & \sigma_z^2 \end{bmatrix}$$

where, $\sigma_y^2 = 16$, $\sigma_z^2 = 9$ and $\rho_{zy} = 0.3, 0.5, 0.7, 0.9$;

Gamma Distribution: characterized by its shape parameter $k_z = 9, k_y = 6$ and rate parameter $\lambda_z = 17, \lambda_y = 12$ and $\rho_{zy} = 0.3, 0.5, 0.7, 0.9$.

Table 6: Comparison of MSE and PRE values of estimators using bivariate normal population.

N = 1000				
ρ_{zy}	n	Estimator	MSE	PRE
0.3	40	$t_{1(1)}^*$	0.400569	100.000
		$t_{1(2)}^*$	0.775805	51.632
		$t_{1(3)}^*$	0.355335	112.729
		$t_{1(4)}^*$	0.355335	112.729
		$t^@$	0.346097	115.739
	50	$t_{1(1)}^*$	0.319976	100.000
		$t_{1(2)}^*$	0.620551	51.563
		$t_{1(3)}^*$	0.285368	112.127
		$t_{1(4)}^*$	0.285368	112.127
		$t^@$	0.279330	114.551
	60	$t_{1(1)}^*$	0.265895	100.000
		$t_{1(2)}^*$	0.517879	51.343
		$t_{1(3)}^*$	0.238196	111.628
		$t_{1(4)}^*$	0.238196	111.628
		$t^@$	0.234120	113.572
0.5	40	$t_{1(1)}^*$	0.400594	100.000
		$t_{1(2)}^*$	0.565115	70.887
		$t_{1(3)}^*$	0.292826	136.802
		$t_{1(4)}^*$	0.292826	136.802
		$t^@$	0.285346	140.388
	50	$t_{1(1)}^*$	0.319999	100.000
		$t_{1(2)}^*$	0.451996	70.796
		$t_{1(3)}^*$	0.235276	136.010
		$t_{1(4)}^*$	0.235276	136.010
		$t^@$	0.230305	138.945
	60	$t_{1(1)}^*$	0.265926	100.000
		$t_{1(2)}^*$	0.377430	70.457
		$t_{1(3)}^*$	0.196665	135.217
		$t_{1(4)}^*$	0.196665	135.217
		$t^@$	0.193284	137.583
		$t_{1(1)}^*$	0.400559	100.000
		$t_{1(2)}^*$	0.354598	112.961

0.7	40	$t_{1(3)}^*$	0.199119	201.165
		$t_{1(4)}^*$	0.199119	201.165
		$t^@$	0.194019	206.453
	50	$t_{1(1)}^*$	0.319966	100.000
		$t_{1(2)}^*$	0.283577	112.832
		$t_{1(3)}^*$	0.160046	199.921
		$t_{1(4)}^*$	0.160046	199.921
		$t^@$	0.156667	204.232
	60	$t_{1(1)}^*$	0.265884	100.000
		$t_{1(2)}^*$	0.236851	112.257
		$t_{1(3)}^*$	0.133902	198.565
		$t_{1(4)}^*$	0.133902	198.565
		$t^@$	0.131603	202.034
0.9	40	$t_{1(1)}^*$	0.400469	100.000
		$t_{1(2)}^*$	0.144217	277.683
		$t_{1(3)}^*$	0.074179	539.864
		$t_{1(4)}^*$	0.074179	539.864
		$t^@$	0.072237	554.380
	50	$t_{1(1)}^*$	0.319885	100.000
		$t_{1(2)}^*$	0.115285	277.472
		$t_{1(3)}^*$	0.059640	536.355
		$t_{1(4)}^*$	0.059640	536.355
		$t^@$	0.058397	547.769
	60	$t_{1(1)}^*$	0.265798	100.000
		$t_{1(2)}^*$	0.096264	276.111
		$t_{1(3)}^*$	0.049928	532.361
		$t_{1(4)}^*$	0.049928	532.361
		$t^@$	0.049074	541.623

(c) Statistical parameters were calculated for the simulated dataset.

(d) SRSWOR was employed to select samples of sizes $n=40, 50, 60$ from the simulated population.

(e) The performance indicators, MSE and PRE, for both the newly suggested and conventional estimators were computed across various correlation levels ρ_{zy} based on the chosen samples.

Table 7: Comparison of MSE and PRE values of estimators using gamma distribution.

N = 1000				
ρ_{zy}	n	Estimator	MSE	PRE
		$t_{1(1)}^*$	0.001043	100.000

0.3	40	$t_{1(2)}^*$	0.001236	84.371
		$t_{1(3)}^*$	0.000926	112.537
		$t_{1(4)}^*$	0.000926	112.537
		$t^@$	0.000900	115.808
	50	$t_{1(1)}^*$	0.000834	100.000
		$t_{1(2)}^*$	0.000991	84.218
		$t_{1(3)}^*$	0.000746	111.859
		$t_{1(4)}^*$	0.000746	111.859
		$t^@$	0.000728	114.536
	60	$t_{1(1)}^*$	0.000693	100.000
		$t_{1(2)}^*$	0.000825	84.0401
		$t_{1(3)}^*$	0.000622	111.329
		$t_{1(4)}^*$	0.000622	111.329
		$t^@$	0.000610	113.529
0.5	40	$t_{1(1)}^*$	0.001042	100.000
		$t_{1(2)}^*$	0.000897	116.189
		$t_{1(3)}^*$	0.000767	135.939
		$t_{1(4)}^*$	0.000767	135.939
		$t^@$	0.000743	140.203
	50	$t_{1(1)}^*$	0.000834	100.000
		$t_{1(2)}^*$	0.000719	115.957
		$t_{1(3)}^*$	0.000617	135.034
		$t_{1(4)}^*$	0.000617	135.034
		$t^@$	0.000602	138.571
	60	$t_{1(1)}^*$	0.000693	100.000
		$t_{1(2)}^*$	0.000598	115.712
		$t_{1(3)}^*$	0.000515	134.339
		$t_{1(4)}^*$	0.000515	134.339
		$t^@$	0.000505	137.231
0.7	40	$t_{1(1)}^*$	0.001042	100.000
		$t_{1(2)}^*$	0.000555	187.689
		$t_{1(3)}^*$	0.000524	198.685
		$t_{1(4)}^*$	0.000524	198.685
		$t^@$	0.000507	205.351
	50	$t_{1(1)}^*$	0.000833	100.000
		$t_{1(2)}^*$	0.000445	187.327
		$t_{1(3)}^*$	0.000422	197.248
		$t_{1(4)}^*$	0.000422	197.248
		$t^@$	0.000411	202.690
		$t_{1(1)}^*$	0.000692	100.000

	60	$t_{1(2)}^*$	0.000370	186.937
		$t_{1(3)}^*$	0.000353	196.168
		$t_{1(4)}^*$	0.000353	196.168
		$t^@$	0.000345	200.653
0.9	40	$t_{1(1)}^*$	0.001042	100.000
		$t_{1(2)}^*$	0.000209	496.369
		$t_{1(3)}^*$	0.000197	528.026
		$t_{1(4)}^*$	0.000197	528.026
		$t^@$	0.000190	547.842
	50	$t_{1(1)}^*$	0.000833	100.000
		$t_{1(2)}^*$	0.000168	495.784
		$t_{1(3)}^*$	0.000159	524.027
		$t_{1(4)}^*$	0.000159	524.027
		$t^@$	0.000154	540.136
	60	$t_{1(1)}^*$	0.000692	100.000
		$t_{1(2)}^*$	0.000139	494.751
		$t_{1(3)}^*$	0.000132	520.978
		$t_{1(4)}^*$	0.000132	520.978
		$t^@$	0.000129	534.429

- (f) The complete procedure for calculating the MSEs under SRS was executed 12,000 iterations to ensure accuracy.

Simulation experiments were conducted to evaluate the efficiency of the suggested mean estimator. According to the data in Tables 6 and 7, the MSEs and PREs highlight the comparative performance of the estimators.

10. Findings and Discussion

The findings presented in Table 4 reveal that the difference-cum-sine-type estimator $t^@$ attains the minimum MSE (0.057) and a relatively higher PRE (212.99%), thereby demonstrating its advantage over the conventional mean and some sine-type estimators. Likewise, the outcomes reported in Table 5 indicate that the developed estimator $t^@$ achieves a lower MSE (11.012) along with a higher PRE (195.16%), further confirming its superiority over existing and competing sine estimators. In addition, Figs. 1 and 2 depict line diagram illustrating the MSE and PRE of different estimators corresponding to population 1 and population 2, respectively. The simulation results derived through the stepwise procedures are displayed in Tables 6 and 7. As indicated in Tables 6 and 7, for correlation coefficient values $\rho = 0.3, 0.5, 0.7, 0.9$ and sample sizes $n = 40, 50, 60$. The proposed estimator $t^@$ demonstrates higher efficiency compared to the traditional mean estimator $t_{1(1)}^*$ along with the competing estimators $t_{1(2)}^*$, $t_{1(3)}^*$ and $t_{1(4)}^*$.

11. Conclusion

This article presents an improved sine-type estimator for the general parameter related to the study variable. By specifying appropriate scalar values, the general parameter reduces to particular statistical measures such as the mean, variance, standard deviation, and coefficient of variation. A new class of

estimators is introduced to achieve better estimation of the general parameter by incorporating known auxiliary information. The properties of the proposed estimators have been established up to the first-order approximation. According to the findings, the traditional mean estimator, along with Singh et al. [22] and difference estimator, can be represented as members of the proposed class. An empirical study using two real datasets (Tables 4 and 5) and simulated datasets from bivariate normal and gamma distributions (Tables 6 and 7) demonstrates that the proposed estimator consistently outperforms the sample mean and competing sine estimators, achieving minimum MSE and higher PRE. Evidence from the literature indicates that the estimator with the smallest MSE exhibits the highest efficiency. Therefore, the findings from the empirical and simulation analyses indicate that the suggested estimator $t^@$ yields improved results over the conventional mean estimator $t_{1(1)}^*$ and competing sine estimators $t_{1(2)}^*$, $t_{1(3)}^*$ and $t_{1(4)}^*$. Hence, we recommend the proposed class of estimators for application in practical settings.

12. Future works

The findings of this study provide a foundation for future research aimed at exploring additional aspects and potential applications of the proposed methodology. Specifically, future work may focus on:

- (1) Utilizing multiple auxiliary variables to enhance estimation.
- (2) Incorporating considerations of both measurement errors and non-response effects.
- (3) Extending the methodology to alternative sampling designs such as stratified, ranked set, and double sampling.
- (4) Applying two-phase and stratified two-phase sampling techniques within the proposed methodology.

Data Availability

All the data used for this study can be found inside the manuscript.

Ethical approval

This research does not involve experimentation on Human Participants or Animals.

Conflict of interest

The authors declare that they have no conflicts of interest.

Funding

The authors declare that no funds or other grants were received for the preparation of this manuscript.

References

- [1] Srivastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of the American Statistical Association*, 66(334), 404–407. <https://doi.org/10.1080/01621459.1971.10482277>.
- [2] Das, A. K., & Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya: The Indian Journal of Statistics, Series C*, 40, 139–148.

- [3] Kadilar, C., & Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151(3), 893–902. [https://doi.org/10.1016/S0096-3003\(03\)00803-8](https://doi.org/10.1016/S0096-3003(03)00803-8).
- [4] Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). Improvement in estimating the population mean using exponential estimator in simple random sampling. In *Auxiliary information and a priori values in construction of improved estimators* (Vol. 33).
- [5] Gupta, S., & Shabbir, J. (2008). On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics*, 35(5), 559–566. <https://doi.org/10.1080/02664760701835839>.
- [6] Rizvi, S. A. H., & Singh, R. K. (1997). On estimation of population mean using known coefficient of variation. *Microelectronics Reliability*, 37(5), 841–843. [https://doi.org/10.1016/S0026-2714\(96\)00113-8](https://doi.org/10.1016/S0026-2714(96)00113-8).
- [7] Solanki, R. S., & Singh, H. P. (2015). An improved class of estimators for the general population parameters using auxiliary information. *Communications in Statistics - Theory and Methods*, 44(20), 4241–4262. <https://doi.org/10.1080/03610926.2013.784999>.
- [8] Pal, S. K., & Singh, H. P. (2022). An efficient new approach for estimating the general parameter using auxiliary variable in sample surveys. *African Matematika*, 33, 70. <https://doi.org/10.1007/s13370-022-00983-0>.
- [9] Adichwal, N. K., Ahmadini, A. A. H., Raghav, Y. S., Singh, R., & Ali, I. (2022). Estimation of general parameters using auxiliary information in simple random sampling without replacement. *Journal of King Saud University - Science*, 34(2), 101754. <https://doi.org/10.1016/j.jksus.2021.101754>.
- [10] Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). Auxiliary information and a priori values in construction of improved estimators. *arXiv*. <https://doi.org/10.48550/arXiv.0712.0096>.
- [11] Srivastava, S. K., & Jhaji, H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya: The Indian Journal of Statistics, Series C*, 42, 87–96.
- [12] Srivastava, S. K., & Jhaji, H. S. (1981). A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika*, 68, 341–343.
- [13] Singh, R., & Tiwari, S. N. (2025). Improved estimator for population mean utilizing known medians of two auxiliary variables under neutrosophic framework. *Neutrosophic Systems and Applications*, 25(1), 38–52. <https://doi.org/10.61356/j.nswa.2025.25443>.
- [14] Yadav, S. K., Singh, R., & Tiwari, S. N. (2025). Estimation of population mean utilizing two neutrosophic auxiliary variables with imprecise information. *Uncertain Discourse and Applications*, 2(1), 17–31. <https://doi.org/10.48313/uda.v2i1.51>.
- [15] Abu-Dayyeh, W. A., Ahmed, M. S., Ahmed, R. A., & Muttlak, H. A. (2003). Some estimators of a finite population mean using auxiliary information. *Applied Mathematics and Computation*, 139(2–3), 287–298. [https://doi.org/10.1016/S0096-3003\(02\)00180-7](https://doi.org/10.1016/S0096-3003(02)00180-7).
- [16] Almulhim, F. A., Aljohani, H. M., Aldallal, R., Mustafa, M. S., Alsolmi, M. M., Elshenawy, A., & Alrashidi, A. (2024). Estimation of finite population mean using dual auxiliary information under non-response with simple random sampling. *Alexandria Engineering Journal*, 100, 286–299. <https://doi.org/10.1016/j.aej.2024.04.058>.

- [17] Alzahrani, M. R., & Almohaimeed, M. (2025). A novel approach for estimation of population mean with dual use of in stratified random sampling. *Alexandria Engineering Journal*, 119, 64–72. <https://doi.org/10.1016/j.aej.2025.01.057>.
- [18] Aloraini, B., Diab, L. S., & Ahmad, S. (2025). Regression cum exponential type estimator for population proportion using auxiliary attributes: Application on radiation data and simulation study. *Journal of Radiation Research and Applied Sciences*, 18(2), 101479. <https://doi.org/10.1016/j.jrras.2025.101479>.
- [19] Aloraini, B. (2025). Improved estimation of population parameter in the existence of nonresponse using auxiliary information. *AIMS Mathematics*, 10(5), 12312–12342. <https://doi.org/10.3934/math.2025558>.
- [20] Bhattacharyya, D., Singh, G. N., Jawa, T. M., Sayed-Ahmed, N., & Pandey, A. K. (2021). An exponential-cum-sine type hybrid imputation technique for missing data. *Computational Intelligence and Neuroscience*, 2021, 4845569. <https://doi.org/10.1155/2021/4845569>.
- [21] Yunusa, M. A., Muili, J. O., Audu, A., & 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. <https://doi.org/10.13005/OJPS08.01.05>.
- [22] Singh, R., Kumari, A., Dubey, S., & Tiwari, S. N. (2025). Some novel sine-type estimators for finite population mean utilizing known auxiliary information. *Quality & Quantity*, 1–28. <https://doi.org/10.1007/s11135-025-02347-9>.
- [23] Rao, T. J. (1991). On certain methods of improving ratio and regression estimators. *Communications in Statistics – Theory and Methods*, 20(10), 3325–3340. <https://doi.org/10.1080/03610929108830705>.
- [24] Singh, H. P., & Solanki, R. S. (2013). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers*, 54(2), 479–497. <https://doi.org/10.1007/s00362-012-0445-2>.
- [25] Suzuki Motor Corporation. SZKMY stock price, news, Q. & H. F. Yahoo Finance (2025). <https://finance.yahoo.com/quote/SZKMY/history?p=SZKMY>.
- [26] Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.