Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(7): 11992–12010.
DOI: 10.3934/math.2022668
Received: 01 February 2022
Revised: 26 March 2022
Accepted: 13 April 2022
Published: 21 April 2022

Research article

Logarithmic type predictive estimators under simple random sampling

Shashi Bhushan [1], Anoop Kumar[2], Md Tanwir Akhtar[3] and Showkat Ahmad Lone[4,*]

[1] University of Lucknow, Lucknow, U.P., India, 226007
[2] Department of Mathematics & Statistics, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, U.P., India, 226017
[3] Department of Public Health, College of Health Sciences, Saudi Electronic University, KSA
[4] Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, 11673, KSA

* **Correspondence:** Email: s.lone@seu.edu.sa.

Abstract: This study introduces a novel predictive estimation approach of the population mean based on logarithmic type estimators as predictor under simple random sampling. The bias and mean square error of the proffered predictive estimators are examined to the approximation of order one. The efficiency conditions are obtained and the performance of the proffered predictive estimators is examined regarding the contemporary predictive estimators existing till date. Further, a broad computational study is also administered utilizing few real and artificially rendered symmetric and asymmetric populations to exemplify the theoretical results.

Keywords: bias; efficiency; mean square error; predictive estimation
Mathematics Subject Classification: 62D05

1. Introduction

The basic objective of the survey practitioners in sample surveys is to obtain an efficient estimate of an unknown population parameter. Therefore, in sequence of improving the efficiency of estimators of parameters, the survey practitioners usually consider the additional information on an auxiliary variable X that is correlated with the study variable Y. [1] suggested the traditional ratio estimator of population mean under simple random sampling (SRS) provided the variable Y is positively correlated with the variable X. [2] investigated the traditional product estimator of population mean provided the variable Y is negatively correlated with the variable X. [3] mooted the exponential ratio and product estimators of population mean based on SRS. [4] introduced an improved mean estimation procedure under SRS. [5] proposed Kernel-based estimation of $P(X > Y)$ in ranked set sampling (RSS) whereas [6] developed an interval estimation of $P(X < Y)$ in RSS. [7] introduced entropy estimation from ranked set samples with application to test of fit. [8] suggested reliability estimation in multistage ranked set sampling ($MRSS$) whereas [9] investigated the estimation procedure of a symmetric distribution function in $MRSS$. Recently, [10–12] suggested various improved classes of estimators under RSS.

In real life scenarios, situations may also arise when the survey practitioners may be interested in evaluating the mean value of the variable being quantified for the non-sampled units with the help of available sample data. This approach is popularly established as predictive method of estimation which is based on superpopulation models and thus it is also established as model-based approach. This approach presumes that the parent population is a realization of random variables concerning to a superpopulation model. Under this superpopulation, the prior information about the population parameters namely variance, standard deviation, mean, coefficient of variation, etc is utilized to predict the non-sampled values of the study variable.

[13] developed some predictive estimators of population mean based on conventional mean, ratio and regression estimators as predictors for the mean of unobserved units in the population. Later on, [14] constructed predictive estimator of population mean using classical product estimator as a predictor for the mean of an unobserved units in the population and compared it with the conventional product estimator. Further, [15] introduced predictive estimators based on [3] exponential ratio and product estimators as predictors for the mean of an unobserved units of the population. Readers may also refer to few recent related studies like, [16–18] for more detailed study of predictive estimation approach.

The objective of the present manuscript is to proffer few novel logarithmic type predictive estimators under SRS for the mean of unobserved units of the population. The paper is organized in few sections. The Section 2 considers a thorough review of the existing predictive estimators and their properties. In Section 3, the proffered predictive estimators are given with their properties. The efficiency conditions are presented in Section 4 followed by a broad computational study given in Section 5. Lastly, the manuscript is ended in Section 6 with the conclusion.

2. Conventional predictive estimators

Consider a finite population $\kappa = (\kappa_1, \kappa_2, ..., \kappa_N)$ consist of N identifiable units labeled as 1,2,...,N. Let (x_i, y_i) be the observations on i^{th} population unit of the variables (X, Y). Let $\bar{x}$, $\bar{y}$ and $\bar{X}$, $\bar{Y}$ respectively be the sample means and population means of variables X and Y. It is presumed that the population mean $\bar{X}$ of variable X is known and the population mean $\bar{Y}$ of variable Y is computed by measuring a random sample of size n from the population κ utilizing simple random sampling with replacement ($SRSWR$). Let S be the aggregate of all possible samples from population κ such that for any given $s \in S$, let $\vartheta(s)$ be the number of specified units in s and $\bar{s}$ be the set of all those units of κ that are not in s.

The usual mean estimator of population mean $\bar{Y}$ consist of sampled units is given by

$$\bar{y}_s = \frac{1}{\vartheta(s)} \sum_{i \in s} y_i. \tag{2.1}$$

The usual mean estimator of population mean $\bar{Y}$ consist of non-sampled units is given by

$$\bar{Y}_{\bar{s}} = \frac{1}{(N - \vartheta(s))} \sum_{i \in \bar{s}} y_i. \tag{2.2}$$

[13] mooted a model based predictive approach in which a model is defined to predict the non-sampled values. Thus, under SRS for any given $s \in S$, we have the following model:

$$\bar{Y} = \frac{\vartheta(s)}{N} \bar{y}_s + \frac{N - \vartheta(s)}{N} \bar{Y}_{\bar{s}}. \tag{2.3}$$

Under SRS with size $\vartheta(s) = n$, the predictor for overall population mean is stated as

$$\bar{Y} = \frac{n}{N} \bar{y}_s + \frac{(N - n)}{N} \bar{Y}_{\bar{s}}. \tag{2.4}$$

Thus, the estimator for estimating the population mean $\bar{Y}$ is stated as

$$t = \frac{n}{N} \bar{y}_s + \frac{(N - n)}{N} T, \tag{2.5}$$

where T is the predictor of the mean $\bar{Y}_{\bar{s}}$ of unobserved values which is given as

$$T_1 = \bar{y}_s, \qquad \text{Usual mean estimator} \tag{2.6}$$

$$T_2 = \bar{y}_s \left(\frac{\bar{X}_{\bar{s}}}{\bar{x}_s} \right), \quad \text{Classical ratio estimator} \tag{2.7}$$

$$T_3 = \bar{y}_s + b(\bar{X}_{\bar{s}} - \bar{x}_s), \quad \text{Classical regression estimator} \tag{2.8}$$

$$T_4 = \bar{y}_s \left(\frac{\bar{x}_s}{\bar{X}_{\bar{s}}} \right), \quad \text{Classical product estimator} \tag{2.9}$$

$$T_5 = \bar{y}_s \exp\left(\frac{\bar{X}_{\bar{s}} - \bar{x}_s}{\bar{X}_{\bar{s}} + \bar{x}_s} \right), \quad \text{[3] exponential ratio estimator} \tag{2.10}$$

$$T_6 = \bar{y}_s \exp\left(\frac{\bar{x}_s - \bar{X}_{\bar{s}}}{\bar{x}_s + \bar{X}_{\bar{s}}} \right), \quad \text{[3] exponential product estimator} \tag{2.11}$$

$$T_7 = \bar{y}_s \left\{ 1 + \log\left(\frac{\bar{x}_s}{\bar{X}_{\bar{s}}} \right) \right\}^{\beta_1}, \quad \text{[19] estimator} \tag{2.12}$$

$$T_8 = \bar{y}_s \left\{ 1 + \beta_2 \log\left(\frac{\bar{x}_s}{\bar{X}_{\bar{s}}} \right) \right\}, \quad \text{[19] estimator} \tag{2.13}$$

where $\bar{x}_s = n^{-1} \sum_{i \in s} x_i$ and $\bar{X}_{\bar{s}} = (N - n)^{-1} \sum_{i \in \bar{s}} x_i = (N\bar{X} - n\bar{x}_s)/(N - n)$. Also, b is the regression coefficient of Y on X, β_1 and β_2 are duly opted scalars.

Now, corresponding to every predictors $T_i,\ i = 1, 2, ..., 8$, we obtain the predictive estimators $t_i,\ i = 1, 2, ..., 8$ using (2.5) as

$$t_1 = \bar{y}_s, \tag{2.14}$$

$$t_2 = \bar{y}_s \left(\frac{\bar{X}_{\bar{s}}}{\bar{x}_s} \right), \tag{2.15}$$

$$t_3 = \bar{y}_s + b(\bar{X}_{\bar{s}} - \bar{x}_s), \tag{2.16}$$

$$t_4 = \bar{y}_s \left\{ \frac{n\bar{X} + (N-2n)\bar{x}_s}{N\bar{X} - n\bar{x}_s} \right\}, \tag{2.17}$$

$$t_5 = f\bar{y}_s + (1-f)\bar{y}_s \exp\left\{ \frac{N(\bar{X} - \bar{x}_s)}{N(\bar{X} + \bar{x}_s) - 2n\bar{x}_s} \right\}, \tag{2.18}$$

$$t_6 = f\bar{y}_s + (1-f)\bar{y}_s \exp\left\{ \frac{N(\bar{x}_s - \bar{X})}{N(\bar{x}_s + \bar{X}) - 2n\bar{x}_s} \right\}, \tag{2.19}$$

$$t_7 = f\bar{y}_s + (1-f)\bar{y}_s \left\{ 1 + \log\left(\frac{\bar{x}_s}{\bar{X}_{\bar{s}}} \right) \right\}^{\beta_1}, \tag{2.20}$$

$$t_8 = f\bar{y}_s + (1-f)\bar{y}_s \left\{ 1 + \beta_2 \log\left(\frac{\bar{x}_s}{\bar{X}_{\bar{s}}} \right) \right\}, \tag{2.21}$$

where $f = n/N$.

[13] demonstrated that while using the usual mean estimator, ratio estimator and regression estimator as predictor $T_i,\ i = 1,2,3$ respectively, the predictive estimator $t_i,\ i = 1,2,3$ becomes the corresponding usual mean estimator T_1, ratio estimator T_2 and regression estimator T_3 respectively. Further, [14] demonstrated that when product estimator T_4 is used as predictor, the predictive estimator t_4 is rather different from the usual product estimator T_4. Later on, [15] demonstrated that when [3] exponential ratio and product estimators are used as predictor, the corresponding predictive estimators are rather different from the natural estimators $T_i,\ i = 5,6$ respectively. It is also observed that when the log type estimators envisaged by [15] are used as predictor, the corresponding predictive estimators are found to be rather different from the customary estimators $T_i,\ i = 7,8$.

To enhance the efficiency of the conventional estimators, [20] investigated a technique by multiplying a regulating constant ϕ (*say*) whose optimum value depend on the coefficient of variation which is a fairly stable quantity. Using [20] procedure, [16] defined the following improved estimators corresponding to the predictive estimators $t_i,\ i = 1,2,4$ as

$$t_9 = \phi_1 t_1 = \phi_1 \bar{y}_s, \tag{2.22}$$

$$t_{10} = \phi_2 t_2 = \phi_2 \bar{y}_s \left(\frac{\bar{X}_{\bar{s}}}{\bar{x}_s} \right), \tag{2.23}$$

$$t_{11} = \phi_3 t_4 = \phi_3 \bar{y}_s \left\{ \frac{n\bar{X} + (N-2n)\bar{x}_s}{N\bar{X} - n\bar{x}_s} \right\}, \tag{2.24}$$

where $\phi_i,\ i = 1,2,3$ are duly opted scalars to be determined.

Further, [16] developed the [20] based predictive estimators corresponding to the predictive estimators $t_i,\ i = 5,6$ as

$$t_{12} = \phi_4 t_5 = \phi_4 \left[f\bar{y}_s + (1-f)\bar{y}_s \exp\left\{ \frac{N(\bar{X} - \bar{x}_s)}{N(\bar{X} + \bar{x}_s) - 2n\bar{x}_s} \right\} \right], \tag{2.25}$$

$$t_{13} = \phi_5 t_6 = \phi_5 \left[f\bar{y}_s + (1-f)\bar{y}_s \exp\left\{ \frac{N(\bar{x}_s - \bar{X})}{N(\bar{x}_s + \bar{X}) - 2n\bar{x}_s} \right\} \right], \tag{2.26}$$

where ϕ_4 and ϕ_5 are duly opted scalars to be determined.

[17] suggested regression type predictive estimator corresponding to the predictive estimator t_3 as

$$t_{14} = \phi_6 f \bar{y}_s + (1-f)\{\phi_6 \bar{y}_s + b(\bar{X}_{\bar{s}} - \bar{x}_s)\}, \tag{2.27}$$

where ϕ_6 is a duly opted scalar to be determined.

The readers may refer to appendix A for the properties like, bias and mean square error (MSE) of the above predictive estimators.

3. Proposed predictive estimators

The motivation of this study is to examine an efficient alternative to survey practitioners under SRS. These predictive estimators provide a better alternative to the existing predictive estimators discussed in the previous section. In our proposal, motivated by [21], we suggest few novel logarithmic predictive estimators corresponding to the predictive estimators $t_i,\ i = 1,2$ for the population mean $\bar{Y}$ as

$$t_{sb_1} = \phi_7 f \bar{y}_s + (1-f)\phi_7 \bar{y}_s \left\{1 + \log\left(\frac{\bar{x}}{\bar{X}_s}\right)\right\}^{\beta_1}, \tag{3.1}$$

$$t_{sb_2} = \phi_8 f \bar{y}_s + (1-f)\phi_8 \bar{y}_s \left\{1 + \beta_2 \log\left(\frac{\bar{x}}{\bar{X}_s}\right)\right\}, \tag{3.2}$$

where ϕ_7, ϕ_8 and $\beta_i,\ i = 1,2$ are duly opted scalars.

Theorem 3.1. *The bias and minimum MSE of the proffered predictive estimators $t_{sb_i},\ i = 1,2$ are given by*

$$Bias(t_{sb_i}) = \bar{Y}(\phi_j Q_i - 1),\ j = 7,8, \tag{3.3}$$

$$minMSE(t_{sb_i}) = \bar{Y}^2\left(1 - \frac{Q_i^2}{P_i}\right), \tag{3.4}$$

where $\phi_{j(opt)} = \frac{Q_i}{P_i}$, $P_1 = 1 + f_1 C_y^2 + \left\{\beta_1(\beta_1 - 1) + \beta_1 f + \frac{\beta_1 f^2}{(1-f)} + \frac{\beta_1(\beta_1-1)}{(1-f)}\right\} f_1 C_x^2 + 4\beta_1 f_1 \rho_{xy} C_x C_y$, $Q_1 = 1 + \beta_1 f \rho_{xy} C_x C_y - \frac{\beta_1}{2}\left\{\frac{(1-2f)}{(1-f)} - \frac{(\beta_1-1)}{(1-f)}\right\} f_1 C_x^2$, $P_2 = 1 + f_1 C_y^2 + \beta_2\left\{\beta_2 - \frac{(1-2f)}{(1-f)}\right\} f_1 C_x^2 + 4\beta_2 f_1 \rho_{xy} C_x C_y$ *and* $Q_2 = 1 + \beta_2 f_1 \rho_{xy} C_x C_y - \frac{\beta_2(1-2f)}{2(1-f)} f_1 C_x^2$.

Proof. To derive the expressions of bias and MSE of various predictive estimators, let us assume that $\bar{y} = \bar{Y}(1+\epsilon_0)$, $\bar{x} = \bar{X}(1+\epsilon_1)$, such that $E(\epsilon_0) = E(\epsilon_1) = 0$, $E({\epsilon_0}^2) = f_1 C_y^2$, $E({\epsilon_1}^2) = f_1 C_x^2$ and $E(\epsilon_0, \epsilon_1) = f_1 \rho_{xy} C_x C_y$.

where $f_1 = (n^{-1} - N^{-1}) \cong 1/n$. Also, C_x and C_y are respectively the population coefficient of variations of variables X and Y and ρ_{xy} is the population coefficient of correlation between variables X and Y.

Using the above notations, we convert t_{sb_1} in $\epsilon' s$ as

$$t_{sb_1} - \bar{Y} = \bar{Y}\left(\phi_7 \left[\begin{array}{l} 1 + \epsilon_0 + \beta_1\epsilon_1 + \beta_1\left\{\frac{f^2}{2(1-f)} + \frac{(\beta_1-1)}{2(1-f)} - \frac{(1-f)}{2}\right\}\epsilon_1^2 \\ +\beta_1\epsilon_0\epsilon_1 \end{array}\right] - 1\right). \tag{3.5}$$

Taking expectation both the sides of (3.5), we get

$$Bias(t_{sb_1}) = \bar{Y}\left(\phi_7\left[\ 1 + \beta_1 f_1 \rho_{xy} C_x C_y - \frac{\beta_1}{2}\left\{\frac{(1-2f)}{(1-f)} - \frac{(\beta_1-1)}{(1-f)}\right\} f_1 C_x^2\ \right] - 1\right) \tag{3.6}$$

$$= \bar{Y}(\phi_7 Q_1 - 1). \tag{3.7}$$

Similarly, we can obtain bias of predictive estimator t_{sb_2}.

Now, squaring and applying expectation both the sides of (3.5), we get

$$MSE(t_{sb_1}) = \bar{Y}^2 \left(\begin{array}{l} 1 + \phi_7^2 \left[\begin{array}{l} 1 + f_1 C_y^2 + \left\{ \beta_1(\beta_1 - 1) + \beta_1 f + \frac{\beta_1 f^2}{(1-f)} + \frac{\beta_1(\beta_1-1)}{(1-f)} \right\} f_1 C_x^2 \\ +4\beta_1 f_1 \rho_{xy} C_x C_y \end{array} \right] \\ -2\phi_7 \left[1 + \beta_1 f_1 \rho_{xy} C_x C_y - \frac{\beta_1}{2} \left\{ \frac{(1-2f)}{(1-f)} - \frac{(\beta_1-1)}{(1-f)} \right\} f_1 C_x^2 \right] \end{array} \right), \tag{3.8}$$

which can be written as

$$MSE(t_{sb_1}) = \bar{Y}^2 \left(1 + \phi_7^2 P_1 - 2\phi_7 Q_1 \right). \tag{3.9}$$

On differentiating the above MSE expression regarding ϕ_7 and equating to zero, we get

$$\phi_{7(opt)} = \frac{Q_1}{P_1}. \tag{3.10}$$

Putting the value of $\phi_{7(opt)}$ in the $MSE(t_{sb_1})$, we get

$$minMSE(t_{sb_1}) = \bar{Y}^2 \left(1 - \frac{Q_1^2}{P_1} \right). \tag{3.11}$$

Similarly, the derivations of MSE of the estimator t_{sb_2} can be obtained. In general, we can write

$$MSE(t_{sb_i}) = \bar{Y}^2 \left(1 + \phi_j^2 P_i - 2\phi_j Q_i \right), \ i = 1, 2 \ \text{ and } \ j = 7, 8. \tag{3.12}$$

We note that the simultaneous optimization of ϕ_j and β_i of the MSE equation is not possible. So, we get the optimum values of β_i =$\beta_{i(opt)}$ given ϕ_j=1 and put it inside ϕ_j=$\phi_{j(opt)}$ to get (3.4). The optimum values of scalars ϕ_j are given by

$$\phi_{j(opt)} = \frac{Q_i}{P_i}, \tag{3.13}$$

where

$$P_1 = 1 + f_1 C_y^2 + \left\{ \beta_1(\beta_1 - 1) + \beta_1 f + \frac{\beta_1 f^2}{(1-f)} + \frac{\beta_1(\beta_1 - 1)}{(1-f)} \right\} f_1 C_x^2 + 4\beta_1 f_1 \rho_{xy} C_x C_y,$$

$$Q_1 = 1 + \beta_1 f \rho_{xy} C_x C_y - \frac{\beta_1}{2} \left\{ \frac{(1-2f)}{(1-f)} - \frac{(\beta_1 - 1)}{(1-f)} \right\} f_1 C_x^2,$$

$$P_2 = 1 + f_1 C_y^2 + \beta_2 \left\{ \beta_2 - \frac{(1-2f)}{(1-f)} \right\} f_1 C_x^2 + 4\beta_2 f_1 \rho_{xy} C_x C_y,$$

$$Q_2 = 1 + \beta_2 f_1 \rho_{xy} C_x C_y - \frac{\beta_2(1-2f)}{2(1-f)} f_1 C_x^2.$$

The optimum values of β_i, $i = 1, 2$ are given by

$$\beta_{i(opt)} = -\rho_{xy} \frac{C_y}{C_x}. \tag{3.14}$$

□

We would like to note that the MSE expression stated in (3.4) is important in order to determine the efficiency conditions of next sections.

Corollary 3.1. *The proposed predictive estimator t_{sb_1} dominate the proposed predictive estimator t_{sb_2}, iff*

$$\frac{Q_2^2}{P_2} < \frac{Q_1^2}{P_1}, \tag{3.15}$$

and contrariwise. Otherwise, both are equally efficient when equality holds in (3.15).

Proof. On comparing the minimum MSE of both the proffered estimators, we get (3.15). □

We can merely obtain (3.15) whether it retains in practice is through a computational study carried out in Section 5.

4. Efficiency conditions

In the present section, the efficiency conditions are derived by comparing the minimum MSE of the proffered predictive estimators $t_{sb_i},\ i = 1, 2$ from (3.4):

(1) with the MSE of the predictive estimator t_1 from (A.1) and get,

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2. \tag{4.1}$$

(2) with the MSE of the predictive estimator t_2 from (A.3) and get,

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 - f_1 C_x^2 + f_1 \rho_{xy} C_x C_y. \tag{4.2}$$

(3) with the minimum MSE of the predictive estimator t_3 from (A.4) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 + f_1 \rho_{xy}^2 C_y^2. \tag{4.3}$$

(4) with the MSE of the predictive estimator t_4 from (A.8) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 - f_1 C_x^2 - f_1 \rho_{xy} C_x C_y. \tag{4.4}$$

(5) with the minimum MSE of the predictive estimator t_5 from (A.10) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 - \frac{1}{4} f_1 C_x^2 + f_1 \rho_{xy} C_x C_y. \tag{4.5}$$

(6) with the minimum MSE of the predictive estimator t_6 from (A.12) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 - \frac{1}{4} f_1 C_x^2 - f_1 \rho_{xy} C_x C_y. \tag{4.6}$$

(7) with the minimum MSE of the predictive estimator t_7 from (A.15) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 + f_1 \rho_{xy}^2 C_y^2. \tag{4.7}$$

(8) with the minimum MSE of the predictive estimator t_8 from (A.18) and get

$$\frac{Q_i^2}{P_i} > 1 - f_1 C_y^2 + f_1 \rho_{xy}^2 C_y^2. \tag{4.8}$$

(9) with the minimum MSE of the predictive estimator t_9 from (A.19) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{MSE(t_1)}{(\bar{Y}^2 + MSE(t_1))}. \tag{4.9}$$

(10) with the minimum MSE of the predictive estimator t_{10} from (A.20) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{(MSE(t_2) - \{Bias(t_2)\}^2)}{(\bar{Y}^2 + MSE(t_2) + 2\bar{Y}Bias(t_2))}. \tag{4.10}$$

(11) with the minimum MSE of the predictive estimator t_{14} from (A.28) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{MSE(t_3)}{(\bar{Y}^2 + MSE(t_3))}. \tag{4.11}$$

(12) with the minimum MSE of the predictive estimator t_{11} from (A.21) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{(MSE(t_4) - \{Bias(t_4)\}^2)}{(\bar{Y}^2 + MSE(t_4) + 2\bar{Y}Bias(t_4))}. \tag{4.12}$$

(13) with the minimum MSE of the predictive estimator t_{12} from (A.24) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{(MSE(t_5) - \{Bias(t_5)\}^2)}{(\bar{Y}^2 + MSE(t_5) + 2\bar{Y}Bias(t_5))} \tag{4.13}$$

(14) with the minimum MSE of the predictive estimator t_{13} from (A.27) and get

$$\frac{Q_i^2}{P_i} > 1 - \frac{(MSE(t_6) - \{Bias(t_6)\}^2)}{(\bar{Y}^2 + MSE(t_6) + 2\bar{Y}Bias(t_6))}. \tag{4.14}$$

Under the above conditions, the proffered predictive estimators dominate the reviewed predictive estimators in SRS. Further, these conditions hold in practice is verified through a broad computational study using various real and artificially generated symmetric and asymmetric populations. Also, it is worth mentioning that the population coefficient of variations and coefficient of correlation are stable quantities and therefore, the optimum values of both proposed and existing estimators can be estimated using sample data.

5. Computational study

In tandem of the theoretical results, a broad computational study is carried out under the four heads namely, numerical study using real populations, simulation study using real populations, simulation study using artificially generated symmetric and asymmetric populations and discussion of computational results.

5.1. Numerical study using real populations

We consider six natural populations to perform the numerical study. The source of the populations, the nature of the variables Y and X and the values of different parameters are described below.

Population 1: Source: ([22], pp. 1115), Y=season average price per pound during 1996, X=season average price per pound during 1995, N=36, n=12, $\bar{Y}$=0.2033, $\bar{X}$=0.1856, S_y^2=0.006458, S_x^2=0.005654 and ρ_{xy}=0.8775.

Population 2: Source: ([22], pp. 1113), Y=duration of sleep (in minutes), X=age of old persons ($\geq$ 50 years), N=30, n=8, $\bar{Y}$=384.2, $\bar{X}$=67.267, S_y^2=3582.58, S_x^2=85.237 and ρ_{xy}=-0.8552.

Population 3: Source: ([23], pp. 228), Y=output for 80 factories in a region, X=number of workers for 80 factories in a region, N=80, n=35, $\bar{Y}$=5182.637, $\bar{X}$=285, S_y^2=3369642, S_x^2=73188.3 and ρ_{xy}=0.9150.

Population 4: Source: ([24], pp. 653-659), Y=real estate values according to 1984 assessment (in millions of kroner), X=number of municipal employees in 1984, N=284, n=75 $\bar{Y}$=3077.525, $\bar{X}$=1779.063, S_y^2=22520027, S_x^2=18089178 and ρ_{xy}=0.94.

Population 5: Source: ([22], pp. 1116), Y=number of fish caught by marine recreational fisherman in 1995, X=number of fish caught by marine recreational fisherman in 1993, N=69, n=28 $\bar{Y}$=4514.89, $\bar{X}$=4591.07, S_y^2=37199578, S_x^2=39881874 and ρ_{xy}=0.9564.

Population 6: The data is chosen from [25] based on apple production and number of apple trees in 7 regions of Turkey during 1999. However, we take only the data of South Anatolia region consist of 69 villages. (Origin: Institute of Statistics, Republic of Turkey). The essential statistics are presented as, Y=amount of apple yield in South Anatolia region, X=quantity of apple trees in South Anatolia region, N=69, n=22 $\bar{Y}$=71.347, $\bar{X}$=3165.029, S_y^2=12289.72, S_x^2=15723128 and ρ_{xy}=0.9177.

For the above populations, we have calculated the percent relative efficiency (PRE) of different predictive estimators T with respect to ($w.r.t.$) the usual mean estimator t_1 as follows.

$$PRE = \frac{V(t_1)}{MSE(T)} \times 100. \tag{5.1}$$

The results of the numerical study calculated for the above discussed populations are displayed in Table 1 by MSE and PRE.

Table 1. Results of simulation study using real populations.

	Population 1		Population 2		Population 3		Population 4		Population 5		Population 6	
Estimators	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
t_1	0.000519	100.0000	434.6420	100.0000	95002.81	100.0000	302269.90	100.0000	1306989.0	100.0000	551.7252	100.0000
t_2	0.000166	312.5867	1385.5170	31.3703	320553.70	29.6371	184797.60	163.5680	174803.5	747.6902	117.1800	470.8355
t_3	0.000119	434.9081	116.7285	372.3529	15449.03	614.9437	35146.1	860.0365	111433.1	1172.8910	87.0246	633.9874
t_4	0.001964	26.4355	153.5288	283.1012	1246500.00	7.6215	1487555.00	20.3199	5160970.0	25.3244	1772.1600	31.1329
t_5	0.000206	251.8145	826.3593	52.5972	35647.24	266.5081	110057.10	274.6482	400671.8	326.1994	236.2164	233.5677
t_6	0.001105	46.9801	210.3652	206.6131	498620.40	19.0531	761436.00	39.6973	2893755.0	45.1658	1063.7070	51.8681
$t_i,\ i = 7, 8$	0.000119	434.9081	116.7285	372.3529	15449.03	614.9437	35146.17	860.0365	111433.1	1172.8910	87.0246	633.9874
t_9	0.000511	101.4308	433.0463	100.3685	94653.17	100.3694	289418.30	104.4405	1227636.0	106.4639	496.6125	111.0977
t_{10}	0.000163	317.2549	1355.3220	32.0692	297609.70	31.9219	153695.70	196.6677	165981.4	787.4310	114.3368	482.5438
t_{11}	0.001760	29.4973	153.2289	283.6554	1086787.00	8.7416	987991.50	30.5943	3236364.0	40.3844	1019.0590	54.1406
t_{12}	0.000205	252.3068	817.8529	53.1442	35590.03	266.9366	108684.00	278.1181	399312.5	327.3099	231.8114	238.0061
t_{13}	0.001060	48.9902	210.2373	206.7387	481698.60	19.7224	664626.50	45.4796	2373909.0	55.0564	822.3618	67.0903
t_{14}	0.000119	436.341	116.6131	372.7215	15439.75	615.3131	34964.35	864.5088	110819.8	1179.3830	85.5087	645.2266
t_{sb_1}	0.000113	456.0985	116.0060	374.6722	13336.01	712.3780	25317.32	1193.9250	52077.9	2509.6780	63.4117	870.0673
t_{sb_2}	0.000118	438.1760	116.0764	374.4446	15389.20	617.3344	31571.27	957.4207	66962.0	1951.8350	68.1361	809.7390

5.2. *Simulation study using real populations*

In order to generalize the findings of numerical study, a simulation study is carried out using some real populations. The steps involved in the simulation study are as follows:

Step 1. Consider the real populations discussed in subsection 5.1.
Step 2. Draw a simple random sample of size given in the respective populations using $SRSWR$ scheme.
Step 3. Compute the necessary statistics.
Step 4. Iterate the above steps 10,000 times and compute the MSE and PRE of various estimators.

The simulated PRE is computed as

$$PRE = \frac{\sum_{i=1}^{10000}(t_1 - \bar{Y})^2}{\sum_{i=1}^{10000}(T_i - \bar{Y})^2} \times 100. \tag{5.2}$$

The outcomes of the simulation study consist of the real populations are reported in Table 2 by MSE and PRE.

Table 2. Results of numerical study using real populations.

	Population 1		Population 2		Population 3		Population 4		Population 5		Population 6	
Estimators	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
t_1	0.000537	100.0000	448.3643	100.0000	96275.54	100.0000	300267.00	100.0000	1328554.0	100.0000	558.5513	100.0000
t_2	0.000135	396.9244	1469.7820	30.5055	366273.30	26.2851	146815.80	204.5196	118405.9	1122.0340	95.0817	587.7120
t_3	0.000123	434.7944	120.4454	372.2551	15671.25	614.3450	34951.08	859.1065	113324.4	1172.3460	88.1543	633.9874
t_4	0.002069	25.9611	120.7119	371.4332	1361119.00	7.0732	1897189	15.8269	5293654.0	25.0971	1748.331	31.9476
t_5	0.000195	275.4914	872.3524	51.3971	39419.24	244.2349	43107.56	696.5530	379110.9	350.4394	236.0278	236.6627
t_6	0.001162	46.2326	197.8174	226.6556	536842.20	17.9336	918294.20	32.6983	2966735.0	44.7816	1062.6520	52.5617
$t_i,\ i = 7, 8$	0.000123	434.7944	120.4454	372.2551	15671.25	614.3450	34951.08	859.1065	113324.4	1172.3460	88.1543	633.9874
t_9	0.000530	101.3	447.0065	100.3038	95931.69	100.3584	291040.10	103.1703	1247263.0	106.5176	503.3234	110.9738
t_{10}	0.000134	400.2691	1438.944	31.1592	338562.80	28.4365	128582.90	233.5201	116430.5	1141.0710	94.6848	590.1685
t_{11}	0.001855	28.9509	120.597	371.7873	1174318.00	8.1984	1214622.00	24.7210	3248543.0	40.8969	988.6951	56.4963
t_{12}	0.000194	275.6640	863.8603	51.9024	39373.75	244.5170	43107.55	696.5532	379063.5	350.4832	233.7415	238.9769
t_{13}	0.001114	48.1968	197.8105	226.6636	517500.20	18.6039	797916.30	37.6313	2409785.0	55.1316	813.3817	68.6718
t_{14}	0.000123	436.0944	120.3472	372.5589	15662.11	614.7034	34822.58	862.2769	112697.8	1178.8640	86.6536	644.9612
t_{sb_1}	0.000117	457.0226	119.7430	374.4387	13286.20	724.6280	20246.70	1483.0420	s48347.9	2747.9040	63.5411	879.7056
t_{sb_2}	0.000122	437.2748	119.7707	374.3522	15619.28	616.3890	34074.35	881.2114	69799.1	1903.3950	71.5180	781.7386

5.3. *Simulation study using artificially generated populations*

Following [26], we accomplish a simulation study using some artificially rendered populations. The simulation steps are are given as follows:

Step 1. Generate two families of symmetric populations such as Normal and Logistic and two families of asymmetric populations such as Gamma and Weibull each of size N=500. The data on variables X and Y are generated through the models $Y = 8.4 + \sqrt{(1-\rho_{xy}^2)}\, Y^* + \rho_{xy}\left(S_y/S_x\right)X^*$ and $X = 4.4 + X^*$ with particular values of parameters given in Tables 3 and 4.
Step 2. Draw a bivariate simple random sample of size n=50 using $SRSWR$ scheme from each population.
Step 3. Compute the required statistics.
Step 4. Iterate the above steps 10,000 times.

Table 3. Results of simulation study using artificially generated symmetric populations.

ρ_{xy}	0.3		0.5		0.7		0.9	
Estimators	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
$X^* \sim N(25, 45)$								
$Y^* \sim N(30, 50)$								
t_1	50.3968	100	51.1183	100	51.4405	100	50.8947	100
t_2	142.5336	35.3578	121.8881	41.9388	90.1208	57.0795	46.0322	110.5632
t_3	45.8610	109.8901	38.3387	133.3333	26.2347	196.0784	9.6700	526.3158
t_4	236.8842	21.2748	288.6117	17.7118	326.3872	15.7606	326.4349	15.5910
t_5	61.6371	81.7636	47.9703	106.5625	31.5773	162.9035	14.6288	347.9076
t_6	108.8125	46.3152	131.3322	38.9229	149.7106	34.3600	154.8301	32.8713
$t_i,\ i = 7, 8$	45.8610	109.8901	38.3387	133.3333	26.2347	196.0784	9.6700	526.3158
t_9	49.1092	102.6217	49.9274	102.3854	50.2591	102.3507	49.5476	102.7189
t_{10}	109.5109	46.0199	94.6505	54.0074	70.1504	73.3289	35.0234	145.3163
t_{11}	170.5642	29.5471	203.4361	25.1274	225.0427	22.8581	218.5347	23.2891
t_{12}	56.8564	88.6386	44.8309	114.0248	29.8436	172.3669	13.9891	363.8153
t_{13}	103.0177	48.9205	122.9221	41.5860	138.1792	37.2274	139.8083	36.4032
t_{14}	44.7919	112.5130	37.6642	135.7212	25.9233	198.4338	9.6201	529.0459
t_{sb_1}	43.6835	115.3680	35.9884	142.0410	23.9434	214.8419	7.7975	652.7058
t_{sb_2}	44.5380	113.1544	37.4848	136.3710	25.9061	198.5650	9.6033	529.9678
$X^* \sim Logis(1, 5)$								
$Y^* \sim Logis(2, 6)$								
t_1	2.8501	100	2.8076	100	2.7875	100	2.8168	100
t_2	7.1526	39.8471	5.6476	49.7137	4.0338	69.1047	2.2797	123.5622
t_3	2.5936	109.8901	2.1057	133.3333	1.4216	196.0784	0.5352	526.3158
t_4	12.0603	23.6323	13.9165	20.1748	15.6729	17.7858	17.1602	16.4151
t_5	3.3123	86.0467	2.4840	113.0280	1.6442	169.5347	0.8225	342.4645
t_6	5.7661	49.4288	6.6184	42.4212	7.4638	37.3478	8.2627	34.09123
$t_i,\ i = 7, 8$	2.5936	109.8901	2.1057	133.3333	1.4216	196.0784	0.5352	526.3158
t_9	2.7620	103.1910	2.7267	102.9653	2.7108	102.8291	2.7377	102.8914
t_{10}	5.5292	51.5465	4.4241	63.4618	3.1609	88.1867	1.7133	164.4075
t_{11}	8.6904	32.7964	9.7905	28.6769	10.7669	25.8902	11.4823	24.5323
t_{12}	3.0518	93.3901	2.3232	120.8504	1.5565	179.0859	0.7821	360.1578
t_{13}	5.4251	52.5356	6.1453	45.6874	6.8335	40.7925	7.4366	37.8784
t_{14}	2.5203	113.0837	2.0598	136.3038	1.4013	198.0152	0.5322	529.2192
t_{sb_1}	2.4553	116.0774	1.9658	142.8188	1.2952	215.2172	0.4348	647.8555
t_{sb_2}	2.5067	113.7006	2.0529	136.7633	1.4040	198.5382	0.5313	530.1453

Table 4. Results of simulation study using artificially generated asymmetric populations.

ρ_{xy}	0.3		0.5		0.7		0.9	
Estimators	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
$X^* \sim Gamma(0.8, 0.1)$								
$Y^* \sim Gamma(0.7, 0.5)$								
t_1	0.0554	100	0.0540	100	0.0534	100	0.0545	100
t_2	1.0098	5.4953	0.9505	5.6886	0.8719	6.1283	0.7581	7.1889
t_3	0.0504	109.8901	0.0405	133.3333	0.0272	196.0784	0.0103	526.3158
t_4	1.2956	4.2830	1.4319	3.7763	1.5512	3.4447	1.6323	3.3391
t_5	0.2583	21.4801	0.2180	24.8016	0.1731	30.8601	0.1211	44.9883
t_6	0.4012	13.8298	0.4587	11.7883	0.5128	10.4205	0.5582	9.7639
$t_i,\ i = 7, 8$	0.0504	109.8901	0.0405	133.3333	0.0272	196.0784	0.0103	526.3158
t_9	0.0554	100.0647	0.0540	100.0584	0.0534	100.0544	0.0544	100.0569
t_{10}	0.9673	5.7369	0.9118	5.9299	0.8376	6.3794	0.7293	7.4731
t_{11}	1.2346	4.4948	1.3621	3.9698	1.4730	3.6275	1.5470	3.5232
t_{12}	0.2548	21.7760	0.2151	25.1302	0.1709	31.2603	0.1195	45.5821
t_{13}	0.3999	13.8743	0.4568	11.8359	0.5103	10.4716	0.5549	9.8218
t_{14}	0.0504	109.9549	0.0405	133.3918	0.0272	196.1329	0.0103	526.3727
t_{sb_1}	0.0503	110.2239	0.0403	134.1225	0.0269	198.4795	0.0099	548.9376
t_{sb_2}	0.0504	110.0322	0.0404	133.5476	0.0271	196.4839	0.0103	528.3340
$X^* \sim Weibull(10, 9)$								
$Y^* \sim Weibull(10, 7)$								
t_1	8.0686	100	7.9500	100	7.8739	100	7.9031	100
t_2	40.5827	19.8818	35.3331	22.5002	28.0046	28.1167	18.0719	43.7314
t_3	7.3424	109.8901	5.9625	133.3333	4.0157	196.0784	1.5015	526.3158
t_4	61.4817	13.1235	71.2310	11.1609	78.2847	10.0581	78.7113	10.0406
t_5	13.5847	59.3945	10.3085	77.1207	6.6216	118.9127	2.8654	275.8113
t_6	24.0342	33.5712	28.2575	28.1342	31.7616	24.7908	33.1850	23.8152
$t_i,\ i = 7, 8$	7.3424	109.8901	5.9625	133.3333	4.0157	196.0784	1.5015	526.3158
t_9	7.8197	103.1830	7.7348	102.7818	7.6731	102.6170	7.6656	103.0977
t_{10}	25.0187	32.2501	22.2511	35.7286	17.9325	43.9089	11.6411	67.8896
t_{11}	34.6409	23.2921	38.9807	20.3948	41.5833	18.9354	40.2655	19.6275
t_{12}	11.6348	69.3483	8.9894	88.4374	5.8548	134.4865	2.5384	311.3420
t_{13}	22.2728	36.2261	25.7313	30.8963	28.3435	27.7805	28.7229	27.5150
t_{14}	7.1356	113.0748	5.8405	136.1181	3.9627	198.6987	1.4927	529.4229
t_{sb_1}	6.8280	118.1688	5.3627	148.2464	3.3633	s234.1102	0.8619	916.9281
t_{sb_2}	7.0541	114.3802	5.7602	138.0170	3.9119	201.2821	1.4796	534.1258

We have taken different values of correlation coefficient $\rho_{xy} = 0.3, 0.5, 0.7, 0.9$ to observe the deportment of the proffered predictive estimators. The MSE and simulated PRE of different predictive estimators T regarding the usual mean estimator t_1 are computed using the expression given in (5.2).

The simulation results for both the populations are displayed in Tables 3 and 4 by MSE and PRE for various values of correlation coefficient ρ_{xy}.

5.4. Discussion of computational results

The following discussion is drawn from the computational results displayed from Tables 1 to 4.

(i) From Table 1 consists of the results of numerical study of six real populations, the proposed predictive estimators $t_{sb_i},\ i = 1, 2$ show their ascendancy over the existing predictive estimators $t_i,\ i = 1, 2, ..., 14$ by minimum MSE and maximum PRE. The dominance of the proposed predictive estimators can also be observed from the histogram drawn from Figures 1 to 6 for

MSE and PRE.

(ii) The similar inclination can be observed from the findings of simulation study of Table 2 consist of the six real populations.

(iii) From Table 3 based on the simulation results for symmetric populations such as Normal and Logistic with different values of ρ_{xy} also exhibit the ascendancy of the proposed predictive estimators $t_{sb_i},\ i = 1,2$ over the existing predictive estimators $t_i,\ i = 1,2,...,14$ by minimum MSE and maximum PRE.

(iv) The similar conclusion can be drawn from Table 4 based on the asymmetric populations such as Gamma and Weibull.

(iv) From Tables 3 and 4 consist of the simulation results using artificially generated populations, it can be observed that the MSE of the proffered predictive estimators gradually declines as the value of correlation coefficient ρ_{xy} increases and contrariwise in sense of PRE in each population.

(v) Furthermore, from Tables 1 to 4 the proffered predictive estimator t_{sb_1} is found to be superior than the proposed predictive estimator t_{sb_2}.

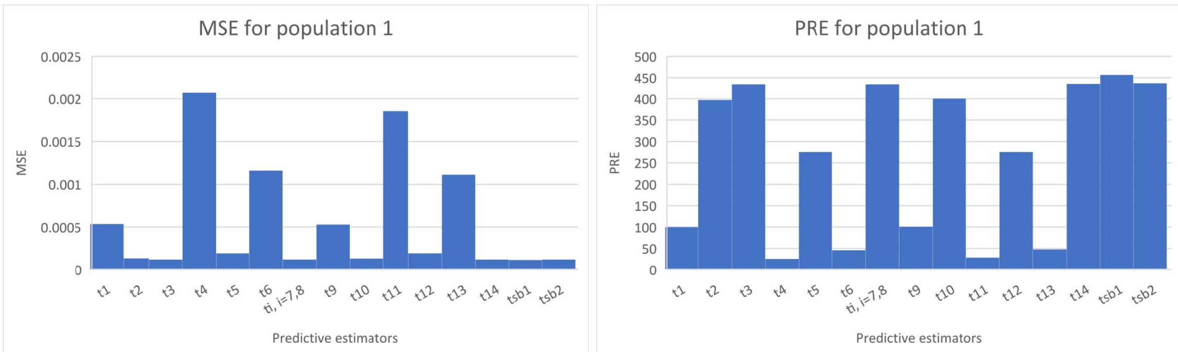

Figure 1. MSE and PRE for population 1.

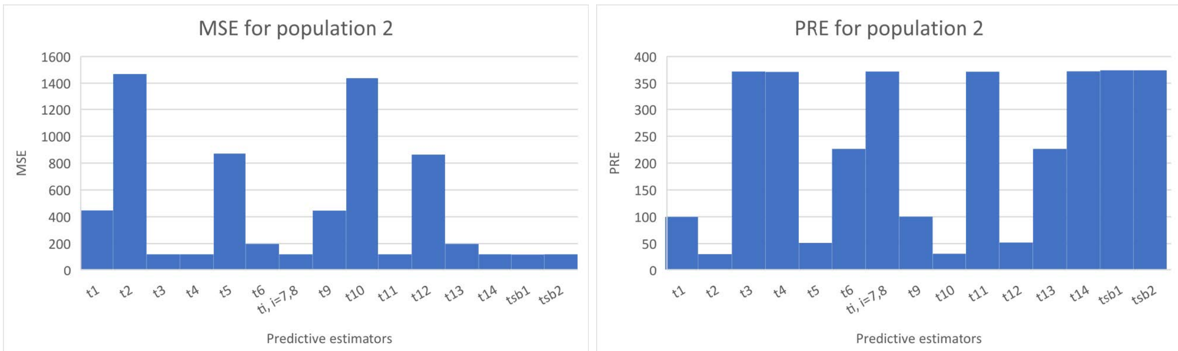

Figure 2. MSE and PRE for population 2.

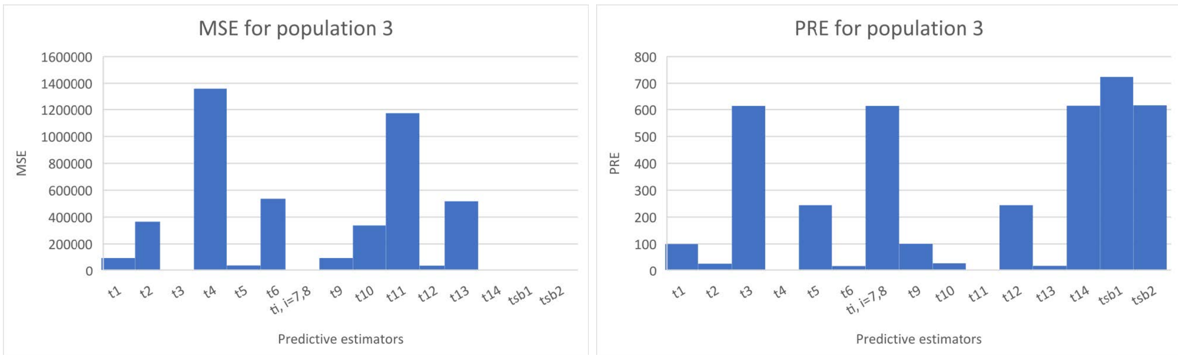

Figure 3. MSE and PRE for population 3.

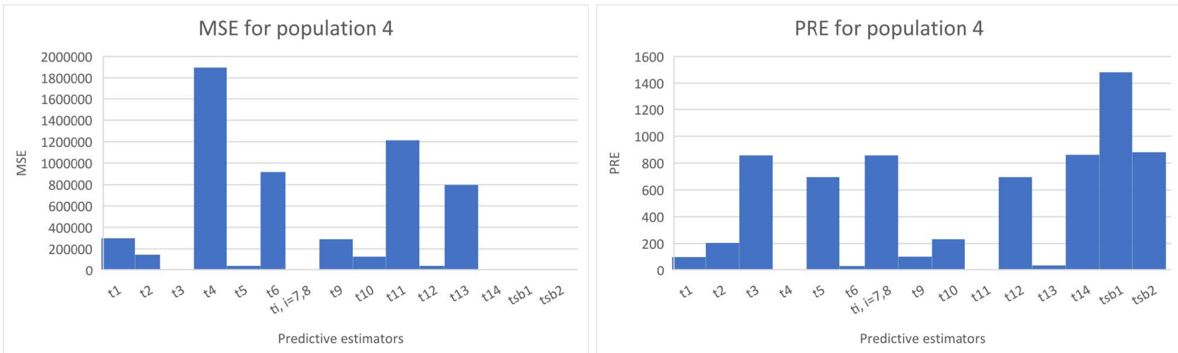

Figure 4. MSE and PRE for population 4.

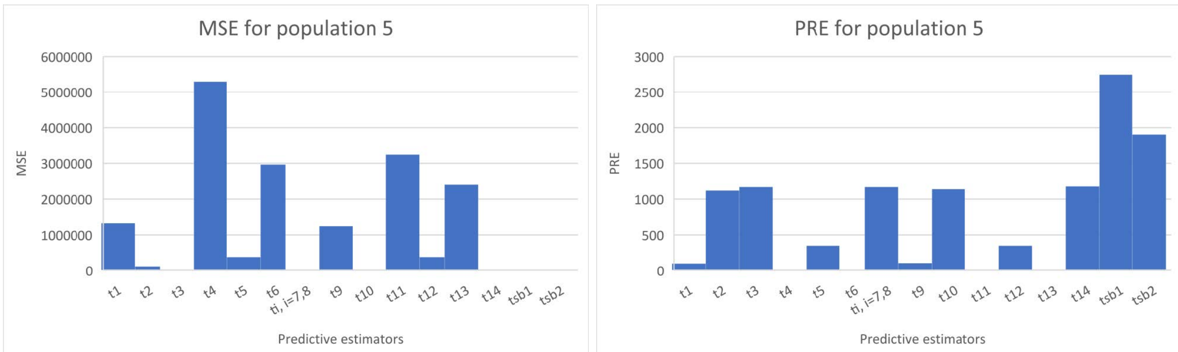

Figure 5. MSE and PRE for population 5.

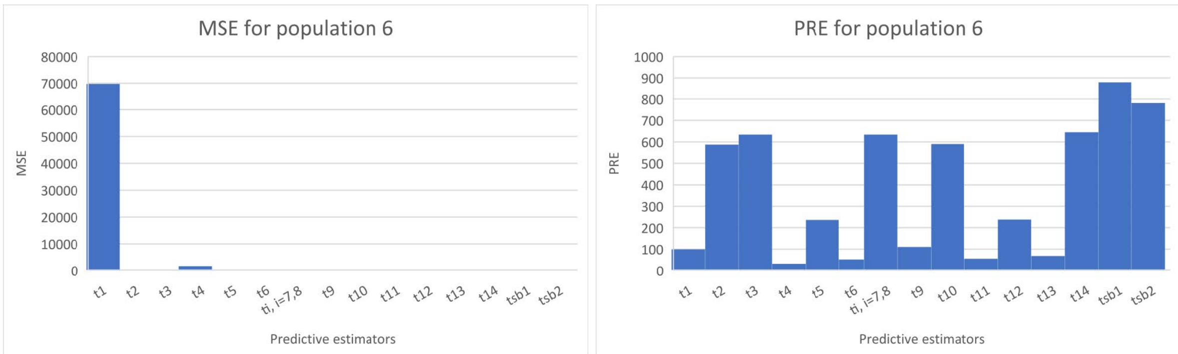

Figure 6. MSE and PRE for population 6.

6. Conclusions

In this manuscript, we have developed few novel logarithmic predictive estimators of population mean in SRS. The properties like bias and MSE of the proffered logarithmic predictive estimators are determined to the first order of approximation. The efficiency conditions have been obtained which are successively enhanced by a broad computational study using various real and artificially generated symmetric and asymmetric populations. From the computational results listed from Tables 1 to 4, we observe that:

(i) The proffered predictive estimators $t_{sb_i},\ i = 1, 2$ are found to be most efficient than the usual unbiased, ratio and regression predictive estimators due to Basu (1971), product predictive estimator due to Srivastava (1983), Bahl and Tuteja (1991) exponential ratio and product type predictive estimators, logarithmic type predictive estimators, Searls (1964) based predictive estimators defined and proposed by Singh et al. (2019) and Bhushan et al. (2020) predictive estimator.

(ii) The correlation coefficient ρ_{xy} demonstrate adverse effect over the MSE and favorable effect over the PRE of the proffered predictive estimators $t_{sb_i},\ i = 1, 2$ which can be seen from the simulation results of Tables 3 and 4.

(iii) The proffered predictive estimator t_{sb_1} performs better than the proposed predictive estimator t_{sb_2} in each real and simulated populations.

Thus, we enthusiastically recommend the utilization of the proffered predictive estimators to the survey professionals in real life. Moreover, in forthcoming studies, we are intended to develop the proposed predictive estimators using ranked set sampling.

Conflict of interest

The authors have no conflict of interest.

References

1. W. G. Cochran, The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce, *J. Agr. Sci.*, **30** (1940), 262–275. http://dx.doi.org/10.1017/S0021859600048012

2. D. S. Robson, Applications of multivariate polykays to the theory of unbiased ratio-type estimation, *J. Am. Stat. Assoc.*, **52** (1957), 511–522. http://dx.doi.org/10.1080/01621459.1957.10501407
3. S. Bahl, R. K. Tuteja, Ratio and product type exponential estimators, *J. Inform. Optim. Sci.*, **12** (1991), 159–164. https://doi.org/10.1080/02522667.1991.10699058
4. S. Bhushan, A. Kumar, S. Singh, S. Kumar, An improved class of estimators of population mean under simple random sampling, *Philippine Statist.*, **70** (2021), 33–47.
5. M. Mahdizadeh, E. Zamanzade, Kernel-based estimation of $P(X > Y)$ in ranked set sampling, *SORT Stat. Oper. Res. T.*, **1** (2016), 243–266.
6. M. Mahdizadeh, E. Zamanzade, Interval estimation of $P(X < Y)$ in ranked set sampling, *Comput. Stat.*, **33** (2018), 1325–1348. https://doi.org/10.1007/s00180-018-0795-x
7. E. Zamanzade, M. Mahdizadeh, Entropy estimation from ranked set samples with application to test of fit, *Rev. Colomb. Estad.*, **40** (2017), 223–241. https://doi.org/10.15446/RCE.V40N2.58944
8. M. Mahdizadeh, E. Zamanzade, Reliability estimation in multistage ranked set sampling, *REVSTAT Stat. J.*, **15** (2017), 565–581.
9. M. Mahdizadeh, E. Zamanzade, Estimation of a symmetric distribution function in multistage ranked set sampling, *Stat. Papers*, **61** (2020), 851–867. https://doi.org/10.1007/s00362-017-0965-x
10. S. Bhushan, A. Kumar, On optimal classes of estimators under ranked set sampling, *Commun. Stat. Theor. M.*, **51** (2020), 2610–2639. https://doi.org/10.1080/03610926.2020.1777431
11. S. Bhushan, A. Kumar, Log type estimators of population mean under ranked set sampling, *Predictive Analyt. Statist. Big Data: Concepts Model.*, **28** (2020), 47–74. https://doi.org/10.2174/9789811490491120010007
12. S. Bhushan, A. Kumar, An efficient class of estimators based on ranked set sampling, *Life Cycle Reliab. Saf. Eng.*, **11** (2022), 39–48. https://doi.org/10.1007/s41872-021-00183-y
13. D. Basu, *An essay on the logical foundation of survey sampling, Part I*, 1971.
14. S. K. Srivastava, Predictive estimation of finite population mean using product estimator, *Metrika*, **30** (1983), 93–99. https://doi.org/10.1007/BF02056907
15. H. P. Singh, R. S. Solanki, A. K. Singh, Predictive estimation of finite population mean using exponential estimators, *Statistika*, **94** (2014), 41–53.
16. A. Singh, G. K. Vishwakarma, R. K. Gangele, Improved predictive estimators for finite population mean using Searls technique, *J. Stat. Manage. Syst.*, **22** (2019), 1555–1571. https://doi.org/10.1080/09720510.2019.1630939
17. S. Bhushan, P. Jaiswal, S. Pandey, An improved predictive approach for estimation of population mean, *IJRASET*, **8** (2020), 251–256. https://doi.org/10.22214/ijraset.2020.31389
18. S. Bhushan, A. Kumar, Predictive estimation approach using difference and ratio type estimators in ranked set sampling, *J. Comput. Appl. Math.*, **410** (2022), 114214. https://doi.org/10.1016/j.cam.2022.114214
19. S. Bhushan, R. Gupta, S. K. Pandey, Some log-type classes of estimators using auxiliary information, *Int. J. Agric. Stat. Sci.*, **11** (2015), 487–491.

20. D. T. Searls, The utilization of a known coefficient of variation in the estimation procedure, *J. Am. Stat. Assoc.*, **59** (1964), 1225–1226. https://doi.org/10.1080/01621459.1964.10480765

21. S. Bhushan, A. Kumar, Novel log type class of estimators under ranked set sampling, *Sankhya B*, 2021, 1–27. https://doi.org/10.1007/s13571-021-00265-y

22. S. Singh, *Advanced sampling theory with applications: How Michael selected Amy*, Kluwer Academic Publishers, 2003.

23. M. N. Murthy, *Sampling: Theory and methods*, Statistical Publishing Society, 1967.

24. C. E. Sarndal, B. Swensson, J. Wretman, *Model assisted survey sampling*, Springer Science & Business Media, 2003.

25. C. Kadilar, H. Cingi, Ratio estimators in stratified random sampling, *Biomed. J.*, **45** (2003), 218–225. https://doi.org/10.1002/bimj.200390007

26. S. Singh, S. Horn, An alternative estimator for multi-character surveys, *Metrika*, **48** (1998), 99–107. https://doi.org/10.1007/PL00020899

Appendix A

The variance of predictive estimator t_1 is given by

$$V(t_1) = f_1\bar{Y}^2C_y^2. \tag{A.1}$$

The bias and MSE of predictive estimator t_2 are given by

$$Bias(t_2) = f_1\bar{Y}^2(C_x^2 - \rho_{xy}C_xC_y), \tag{A.2}$$
$$MSE(t_2) = f_1\bar{Y}^2(C_y^2 + C_x^2 - 2\rho_{xy}C_xC_y). \tag{A.3}$$

The MSE of predictive estimator t_3 is given by

$$MSE(t_3) = \bar{Y}^2f_1C_y^2 + \bar{X}^2b^2f_1C_x^2 - 2b\bar{X}\bar{Y}f_1\rho_{xy}C_xC_y. \tag{A.4}$$

The optimum value of b is obtained by minimizing $(A.4)$ w.r.t. b as

$$b_{(opt)} = \rho_{xy}\frac{S_y}{S_x}. \tag{A.5}$$

The minimum MSE at optimum value of b is given by

$$MSE(t_3) = \bar{Y}^2f_1C_y^2(1 - \rho_{xy}^2). \tag{A.6}$$

The bias and MSE of predictive estimator t_4 are given by

$$Bias(t_4) = f_1\bar{Y}\left(\frac{f}{(1-f)}C_x^2 + \rho_{xy}C_xC_y\right), \tag{A.7}$$
$$MSE(t_4) = f_1\bar{Y}^2\Big(C_y^2 + C_x^2 + 2\rho_{xy}C_xC_y\Big), \tag{A.8}$$

where $f = n/N$.

The bias and MSE of predictive estimator t_5 are given by

$$Bias(t_5) = \frac{\bar{Y}}{8} f_1 (3C_x^2 - 4f_1C_x^2 - 4\rho_{xy}C_xC_y), \tag{A.9}$$

$$MSE(t_5) = \bar{Y}^2 f_1 \left(C_y^2 + \frac{C_x^2}{4} - \rho_{xy}C_xC_y \right). \tag{A.10}$$

The bias and MSE of predictive estimator t_6 are given by

$$Bias(t_6) = \frac{\bar{Y}}{8} f_1 (4f_1C_x^2 + 4\rho_{xy}C_xC_y - 3C_x^2), \tag{A.11}$$

$$MSE(t_6) = \bar{Y}^2 f_1 \left(C_y^2 + \frac{C_x^2}{4} + \rho_{xy}C_xC_y \right). \tag{A.12}$$

The MSE of predictive estimator t_7 is given by

$$MSE(t_7) = \bar{Y}^2 \left[f_1C_y^2 + \beta_1^2 f_1 C_x^2 + 2\beta_1 f_1 \rho_{xy} C_x C_y \right]. \tag{A.13}$$

The optimum value of β_1 is obtained by minimizing $(A.13)$ w.r.t. β_1 as

$$\beta_{1(opt)} = -\rho_{xy}\frac{C_y}{C_x}. \tag{A.14}$$

The minimum MSE at optimum value of β_1 is

$$MSE(t_7) = \bar{Y}^2 f_1 C_y^2 (1 - \rho_{xy}^2). \tag{A.15}$$

The MSE of predictive estimator t_8 is given by

$$MSE(t_8) = \bar{Y}^2 \left[f_1C_y^2 + \beta_2^2 f_1 C_x^2 + 2\beta_2 f_1 \rho_{xy} C_x C_y \right]. \tag{A.16}$$

The optimum value of β_2 is obtained by minimizing $(A.16)$ w.r.t. β_2 as

$$\beta_{2(opt)} = -\rho_{xy}\frac{C_y}{C_x}. \tag{A.17}$$

The minimum MSE at optimum value of β_2 is

$$MSE(t_8) = \bar{Y}^2 f_1 C_y^2 (1 - \rho_{xy}^2). \tag{A.18}$$

The minimum MSE of predictive estimator t_9 under SRS is given by

$$minMSE(t_9) = \frac{\bar{Y}^2 MSE(t_1)}{\bar{Y}^2 + MSE(t_1)}. \tag{A.19}$$

The minimum MSE of predictive estimator t_{10} under SRS is given by

$$minMSE(t_{10}) = \bar{Y}^2 \left[\frac{MSE(t_2) - \{Bias(t_2)\}^2}{\bar{Y}^2 + MSE(t_2) + 2\bar{Y}Bias(t_2)} \right], \tag{A.20}$$

where $\phi_{2(opt)} = (\bar{Y}^2 + \bar{Y}Bias(t_2))/(\bar{Y}^2 + MSE(t_2) + 2\bar{Y}Bias(t_2))$.

The minimum MSE of predictive estimator t_{11} is given by

$$minMSE(t_{11}) = \bar{Y}^2\left[\frac{MSE(t_4) - \{Bias(t_4)\}^2}{\bar{Y}^2 + MSE(t_4) + 2\bar{Y}Bias(t_4)}\right], \tag{A.21}$$

where $\phi_{3(opt)} = (\bar{Y}^2 + \bar{Y}Bias(t_4))/(\bar{Y}^2 + MSE(t_4) + 2\bar{Y}Bias(t_4))$.

The MSE of predictive estimator t_{12} is given by

$$MSE(t_{12}) = (\phi_4 - 1)^2\bar{Y}^2 + \bar{Y}^2\phi_4^2 MSE(t_5) + 2\phi_4(\phi_4 - 1)\bar{Y}Bias(t_5). \tag{A.22}$$

The optimum value of ϕ_4 is obtained by minimizing $(A.22)$ w.r.t. ϕ_4 as

$$\phi_{4(opt)} = \frac{(\bar{Y}^2 + \bar{Y}Bias(t_5))}{(\bar{Y}^2 + MSE(t_5) + 2\bar{Y}Bias(t_5))}. \tag{A.23}$$

The minimum MSE at the optimum value of ϕ_4 is given by

$$minMSE(t_{12}) = \frac{\bar{Y}^2(MSE(t_5) - \{Bias(t_5)\}^2)}{(\bar{Y}^2 + MSE(t_5) + 2\bar{Y}Bias(t_5))}. \tag{A.24}$$

The MSE of predictive estimator t_{13} is given by

$$MSE(t_{13}) = (\phi_5 - 1)^2\bar{Y}^2 + \bar{Y}^2\phi_5^2 MSE(t_6) + 2\phi_5(\phi_5 - 1)\bar{Y}Bias(t_6). \tag{A.25}$$

The optimum value of ϕ_5 is obtained by minimizing $(A.25)$ w.r.t. ϕ_5 as

$$\phi_{5(opt)} = \frac{(\bar{Y}^2 + \bar{Y}Bias(t_6))}{(\bar{Y}^2 + MSE(t_6) + 2\bar{Y}Bias(t_6))}. \tag{A.26}$$

The minimum MSE at the optimum value of ϕ_5 is given by

$$minMSE(t_{13}) = \frac{\bar{Y}^2(MSE(t_6) - \{Bias(t_6)\}^2)}{(\bar{Y}^2 + MSE(t_6) + 2\bar{Y}Bias(t_6))}. \tag{A.27}$$

The minimum MSE of predictive estimator t_{14} under SRS is given by

$$MSE(t_{14}) = \frac{\bar{Y}^2 MSE(t_3)}{\bar{Y}^2 + MSE(t_3)}, \tag{A.28}$$

where $\phi_{6(opt)} = (\bar{Y}^2 + \bar{Y}Bias(t_3))/(\bar{Y}^2 + MSE(t_3) + 2\bar{Y}Bias(t_3))$.