Journal of Modern Applied Statistical Methods

Volume 23 | Issue 1 Article 3

Modified Exponential Ratio-Type Estimator of Population Mean in Stratified Sampling using Calibration Approach

Neha Garg
School of Sciences, Indira Gandhi National Open University, India,
neha1garg@gmail.com

Housila P. Singh
School of Studies in Statistics, Vikram University, India

Menakshi Pachori
School of Sciences, Indira Gandhi National Open University, India

Recommended Citation

Neha Garg, Housila P. Singh, Menakshi Pachori (2024). Modified Exponential Ratio-Type Estimator of Population Mean in Stratified Sampling using Calibration Approach. Journal of Modern Applied Statistical Methods, 23(1), https://doi.org/10.56801/Jmasm.V23.i1.3

Journal of Modern Applied Statistical Methods
2024, Vol. 23, No. 1,
Doi: 10.56801/Jmasm.V23.i1.3

ISSN 1538 – 9472

Modified Exponential Ratio-Type Estimator of Population Mean in Stratified Sampling using Calibration Approach

Neha Garg
School of Sciences, Indira Gandhi National Open University, India

Housila P. Singh
School of Studies in Statistics, Vikram University, India

Menakshi Pachori
School of Sciences, Indira Gandhi National Open University, India

In this paper, the problem of estimation of finite population mean in stratified random sampling is considered. Two improved exponential logarithmic type calibration estimators for finite population mean have been proposed for stratified random sampling when auxiliary information related to variable under study is available for each stratum. To judge the performance of the proposed estimators, a simulation study has been carried out in R-software using two datasets, one real and another one artificial generated population. The proposed estimators have also been compared with the estimators developed by Bahl and Tuteja [1] and Singh [17] in case of stratified random sampling.

Keywords: Auxiliary information, Calibration estimation, Stratified sampling, Exponential, Logarithmic.

1. Introduction

The auxiliary information can be used both at designing and estimation stages. The use of auxiliary information for improving the precision of the estimators is well known when the auxiliary variable is highly correlated with the study variable. The ratio estimator is used for the estimation of population parameters when the study variable and the auxiliary variables are highly positively correlated to each other. Cochran [4] initiated the use of auxiliary information at estimation stage by suggesting ratio estimator for population mean.

Recently calibration is commonly used method in survey sampling to increase the precision of the estimators of population parameters by making use of available auxiliary information. Calibration can be defined as a method of adjusting weights in sampling by utilizing the available auxiliary information in order to estimate population mean, total, etc. of the variable under study. Deville and Sarndal [5] introduced the calibration approach as a procedure of minimizing a distance function subject to some calibration constraints. Later on, researchers Kim et al. [8], Singh and Arnab [16], Koyuncu and Kadilar [9], Mouhamed et al. [13], Koyuncu and Kadilar [10], Clement and Enang [3], Singh et al. [15], Basak et al. [2], Guha et al. [7], Kumari et al. [11], Garg and Pachori [6], etc., have also suggested some calibrated estimators of different population parameters under various sampling schemes.

In this paper, two exponential ratio type calibration estimators for population mean using logarithmic conditions have been proposed under stratified random sampling to obtain better estimator of population mean. The exponential ratio type calibrated estimator has been suggested under stratified random sampling when auxiliary information related to variable under study is available for each stratum. A simulation study has also been carried out to check the performance of the proposed estimators with the estimators suggested by Bahl and Tuteja [1] and Singh [17] in case of stratified random sampling on two datasets.

1.1 Notations used in Calibration Approach

Let us consider a finite population U of size N. Let y_i and x_i (i = 1, 2, …, N) be the values associated with the i^{th} unit of the study and auxiliary variables, respectively. A sample s = {1, 2, …, n} $\subset$ U of fixed size n is drawn using a probability sampling design P. The population total of the auxiliary variable, $X = \sum_{i \in U} x_i$ is assumed to be known. Deville and Sarndal [5] proposed the calibrated estimator as:

$$\hat{Y}_{ds} = \sum_{i \in s} w_i y_i \tag{1}$$

For the Horvitz and Thompson (1952) estimator given as:

$$\hat{Y}_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i} = \sum_{i \in s} d_i y_i \tag{2}$$

where $d_i = 1/\pi_i$ & $\pi_i = \Pr(i \in s)$ and $\pi_{ij} = \Pr(i, j \in s)$ are the inclusion probabilities of order one and two, respectively.

Minimization of Chi-square distance function $\sum_{i \in s} \frac{(w_i - d_i)^2}{d_i q_i}$ subject to its calibration constraints $\sum_{i \in s} w_i x_i = X$, Deville and Sarndal (1992) obtained the generalized regression (GREG) estimator of the population total Y as:

$$\hat{Y}_{GREG} = \sum_{i \in s} d_i y_i + \hat{\beta}_{ds} \left(X - \sum_{i \in s} d_i x_i \right) \tag{3}$$

where $\hat{\beta}_{ds} = \dfrac{\left(\sum_{i\in s} d_i q_i x_i y_i\right)}{\left(\sum_{i\in s} d_i q_i x_i^2\right)}$

The calibration estimator for population mean $\bar{Y}$ using two calibration constraints under stratified random sampling given by **Singh [17]** is

$$\bar{y}_s = \sum_{h=1}^{L} \Omega_h \bar{y}_h \tag{4}$$

where Ω_h are the new calibrated weights obtained by minimizing the Chi-square distance measure $\sum_{h=1}^{L} \frac{(\Omega_h - W_h)^2}{Q_h W_h}$, subject to two calibration constraints:

$$\sum_{h=1}^{L} \Omega_h = 1 \tag{5}$$

$$\sum_{h=1}^{L} \Omega_h \bar{x}_h = \sum_{h=1}^{L} W_h \bar{X}_h \tag{6}$$

The calibrated weight is given as:

$$\Omega_h = W_h + \left[\frac{(W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h) - (W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)}{(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h)^2} \right] (\bar{X} - \sum_{h=1}^{L} W_h \bar{x}_h)$$

Thus, the estimator given by Singh (2003) is given as

$$\bar{y}_s = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}_s (\bar{X} - \sum_{h=1}^{L} W_h \bar{x}_h) \tag{7}$$

where,

$$\hat{\beta}_s = \left[\frac{(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h \bar{y}_h)}{(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h)^2} \right]$$

2. PROPOSED CALIBRATION ESTIMATOR

Bahl and Tuteja [1] suggested an exponential ratio type estimator as:

$$\bar{y}_{bt} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \tag{8}$$

Let us consider a heterogeneous finite population U of size N which is divided into L homogeneous strata of sizes $N_1, N_2, \ldots, N_L$ such that $N = \sum_{h=1}^{L} N_h$. A sample of size n_h is drawn using simple random sampling without replacement (SRSWOR) from the h^{th} stratum such that $\sum_{h=1}^{L} n_h = n$, where n is the required sample size. Let study variable (Y) and auxiliary variable (X) are positively correlated with each other. Suppose y_{hi} and x_{hi} are the i^{th} units of Y and X, respectively, in the h^{th} stratum for i= 1, 2, ..., n_h and h = 1, 2, …, L.

$W_h = \frac{N_h}{N}$ and $f_h = \frac{n_h}{N}$ are the h^{th} stratum weight and sample fraction, respectively.

The traditional stratified estimator of population mean in stratified random sampling is given as:

$$\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h \tag{9}$$

Following Bahl and Tuteja [1], the exponential ratio type estimators given by Malik et al. [12] in stratified random sampling is defined as:

$$\bar{y}_{st.bt} = \sum_{h=1}^{L} W_h \bar{y}_h \exp\left(\frac{\bar{X}_h-\bar{x}_h}{\bar{X}_h+\bar{x}_h}\right) \tag{10}$$

This paper proposes two new exponential ratio type calibration estimators of population mean in stratified random sampling using two different sets of calibration constraints and calibration weights Ω_{h1} and Ω_{h2} for Case I and Case II, respectively, which are chosen in order to minimize the Chi-square type distance measure given as:

$$\sum_{h=1}^{L} \frac{(\Omega_{hi}-W_h)^2}{W_h Q_h};\ i=1,2 \tag{11}$$

Case I: The first proposed calibration estimator of population mean in stratified random sampling using one calibration constraint is given as:

$$\bar{y}_{bt.L1} = \sum_{h=1}^{L} \Omega_{h1} \bar{y}_h \exp\left(\frac{\bar{X}_h-\bar{x}_h}{\bar{X}_h+\bar{x}_h}\right) \tag{12}$$

subject to the following calibration constraints:

$$\sum_{h=1}^{L} \Omega_{h1} \log \bar{x}_h = \sum_{h=1}^{L} W_h \log \bar{X}_h \tag{12}$$

The Lagrange function is given as:

$$L_1 = \sum_{h=1}^{L} \frac{(\Omega_{h1} - W_h)^2}{W_h Q_{h1}} - 2\lambda(\sum_{h=1}^{L} \Omega_{h1} \log \bar{x}_h - \sum_{h=1}^{L} W_h \log \bar{X}_h) \tag{13}$$

Case II: The second proposed calibration estimator using two calibration constraint is given as:

$$\bar{y}_{bt.L2} = \sum_{h=1}^{L} \Omega_{h2} \bar{y}_h \exp\left(\frac{\bar{X}_h - \bar{x}_h}{\bar{X}_h + \bar{x}_h} \right) \tag{14}$$

subject to the following calibration constraints:

$$\sum_{h=1}^{L} \Omega_{h2} = \sum_{h=1}^{L} W_h \tag{15}$$

and

$$\sum_{h=1}^{L} \Omega_{h2} \log \bar{x}_h = \sum_{h=1}^{L} W_h \log \bar{X}_h \tag{16}$$

The Lagrange function for both constraints are given as:

$$L_2 = \sum_{h=1}^{L} \frac{(\Omega_{h2} - W_h)^2}{Q_h W_h} - 2\lambda_1(\sum_{h=1}^{L} \Omega_{h2} - \sum_{h=1}^{L} W_h) - 2\lambda_2(\sum_{h=1}^{L} \Omega_{h2} \log \bar{x}_h - \sum_{h=1}^{L} W_h \log \bar{X}_h) \tag{17}$$

where λ, λ_1 and λ_2 are the Lagrange's multiplier. To find the optimum value of Ω_{h1} and Ω_{h2}, we differentiate the Lagrange function given in equation (13) and (17) with respect to Ω_{h1} and Ω_{h2}, respectively, and equate it to zero. Thus, the calibration weights are obtained as:

$$\Omega_{h1} = W_h + \lambda(W_h Q_h \log \bar{x}_h) \tag{18}$$

and

$$\Omega_{h2} = W_h + W_h Q_h (\lambda_1 + \lambda_2 \log \bar{x}_h) \tag{19}$$

Here λ is determined by substituting the value of Ω_{h1} from equation (18) to equation (12) and λ_1, λ_2 are determined by substituting the value of Ω_{h2} from equation (19) to equations (15) and (16), so this leads to calibrated weight as:

$$\Omega_{h1} = W_h + (W_h Q_h \log \bar{x}_h)\left[\frac{\sum_{h=1}^{L} W_h(\log \bar{X}_h - \log \bar{x}_h)}{\sum_{h=1}^{L} W_h Q_h (\log \bar{x}_h)^2}\right] \tag{20}$$

and

$$\Omega_{h2} = W_h + W_h Q_h \left[\frac{-(\sum_{h=1}^{L} W_h(\log \bar{X}_h - \log \bar{x}_h))(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h)}{(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h) - (\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h)^2}\right] + W_h Q_h \log \bar{x}_h \left[\frac{(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h(\log \bar{X}_h - \log \bar{x}_h))}{(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h) - (\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h)^2}\right] \tag{21}$$

Thus, on substituting the value of Ω_{h1} and Ω_{h2} from equation (20) to equation (12) and from (21) to equation (15), we obtain the proposed calibrated estimators as:

$$\bar{y}_{bt.L1} = \sum_{h=1}^{L} W_h \bar{y}_{h.bt} + \hat{\beta}_{bt.L1}\left[\sum_{h=1}^{L} W_h(\log \bar{X}_h - \log \bar{x}_h)\right] \tag{22}$$

where,

$$\hat{\beta}_{bt.L1} = \left[\frac{\sum_{h=1}^{L} W_h Q_h (\log \bar{x}_h)\bar{y}_{h.bt}}{\sum_{h=1}^{L} W_h Q_h (\log \bar{x}_h)^2}\right]$$

and

$$\bar{y}_{bt.L2} = \sum_{h=1}^{L} W_h \bar{y}_{h.bt} + \hat{\beta}_{bt.L2}\left[\sum_{h=1}^{L} W_h(\log \bar{X}_h - \log \bar{x}_h)\right] \tag{23}$$

where,

$$\hat{\beta}_{bt.L2} = \left[\frac{(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{y}_{h.bt} \log \bar{x}_h) - (\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h \bar{y}_{h.bt})}{(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h) - (\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h)^2}\right]$$

and

$$\bar{y}_{h.bt} = \bar{y}_h \exp\left(\frac{\bar{X}_h - \bar{x}_h}{\bar{X}_h + \bar{x}_h}\right)$$

2.1 Expression for Mean Squared Error (MSE)

The calibration estimator for first suggested estimator can be rewritten by first order Taylor expansion as:

$$\bar{y}_{bt.L1} = \bar{Y} + \sum_{h=1}^{L} W_h \left(\bar{y}_h - \bar{Y}\right)^2 - \left\{\frac{\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h\right) \bar{Y}_{bt.h}\right)}{\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h\right)^2\right)}\right\} \times \left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h - \log \bar{x}_h\right)\right)$$

The mean squared error of the estimator up to second order of approximation is given as:

$$\mathrm{MSE}\left(\bar{y}_{bt.L1}\right) = E\left[\bar{y}_{L1} - \bar{Y}\right]^2$$

$$= \sum_{h=1}^{L} W_h^2 f_h' \left[\bar{Y}_h^2 \left(C_{yh} - \frac{C_{xh}}{2}\right)^2 + \beta_{bt.L1}^2 C_{xh}^2 - 2\beta_{bt.L1} \bar{Y}_h \left(\rho_h C_{xh} C_{yh} - C_{xh}^2\right)\right]$$

where $\beta_{bt.L1} = \dfrac{\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h\right) \bar{Y}_{bt.h}\right)}{\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h\right)^2\right)}$

The calibration estimator in case of second proposed estimator can be rephrased by first order Taylor expansion as:

$$\bar{y}_{bt.L2} = \bar{Y} + \sum_{h=1}^{L} W_h \left(\bar{y}_h - \bar{Y}\right)^2 - \left\{\frac{\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h\right) \bar{Y}_h\right) - \left(\sum_{h=1}^{L} W_h \bar{Y}_h\right)\left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)}{\left(\sum_{h=1}^{L} W_h \log \bar{X}_h^2\right) - \left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)^2}\right\} \times \left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h - \log \bar{x}_h\right)\right)$$

The mean squared error of the estimator up to second order of approximation is given as:

$$\mathrm{MSE}\left(\bar{y}_{bt.L2}\right)=E\left[\bar{y}_{L2}-\bar{Y}\right]^2$$

$$=\sum_{h=1}^{L} W_h^2 f_h' \left[\bar{Y}_h^2\left(C_{yh}-\frac{C_{xh}}{2}\right)^2+\beta_{bt.L2}^2 C_{xh}^2-2\beta_{bt.L2}\bar{Y}_h\left(\rho_h\, C_{xh}C_{yh}-C_{xh}^2\right)\right]$$

where $$\beta_{bt.L2}=\frac{\left(\sum_{h=1}^{L} W_h\left(\log\bar{X}_h\right)\bar{Y}_{bt.h}\right)-\left(\sum_{h=1}^{L} W_h\bar{Y}_{bt.h}\right)\left(\sum_{h=1}^{L} W_h\log\bar{X}_h\right)}{\left(\sum_{h=1}^{L} W_h\log\bar{X}_h^2\right)-\left(\sum_{h=1}^{L} W_h\log\bar{X}_h\right)^2}$$

Now we consider the different values of Q_h to obtain the different forms of the suggested calibration estimator as follows:

1. The calibration estimators when $Q_h=1$

Case-I: $$\bar{y}_{bt.L1}=\sum_{h=1}^{L} W_h\bar{y}_{h.bt}+\left[\hat{\beta}_{bt.L1}\sum_{h=1}^{L} W_h(\log\bar{X}_h-\log\bar{x}_h)\right]$$

where, $$\hat{\beta}_{bt.L1}=\left[\frac{(\sum_{h=1}^{L} W_h\bar{y}_{h.bt}\log\bar{x}_h)}{(\sum_{h=1}^{L} W_h(\log\bar{x}_h)^2}\right]$$

Case-II: $$\bar{y}_{bt.L2}=\sum_{h=1}^{L} W_h\bar{y}_{h.bt}+\left[\hat{\beta}_{bt.L2}\sum_{h=1}^{L} W_h(\log\bar{X}_h-\log\bar{x}_h)\right]$$

where, $$\hat{\beta}_{bt.L2}=\left[\frac{(\sum_{h=1}^{L} W_h\bar{y}_{h.bt}\log\bar{x}_h)-(\sum_{h=1}^{L} W_h\bar{y}_h.bt)(\sum_{h=1}^{L} W_h\log\bar{x}_h)}{(\sum_{h=1}^{L} W_h\log\bar{x}_h^2)-(\sum_{h=1}^{L} W_h\log\bar{x}_h)^2}\right]$$

2. The calibration estimators when $Q_h=\dfrac{1}{\bar{x}_h}$

Case-I: $$\bar{y}_{bt.L1}=\sum_{h=1}^{L} W_h\bar{y}_{h.bt}+\left[\hat{\beta}_{bt.L1}\sum_{h=1}^{L} W_h(\log\bar{X}_h-\log\bar{x}_h)\right]$$

where, $$\hat{\beta}_{bt.L1}=\left[\frac{\left(\sum_{h=1}^{L} W_h\bar{y}_h\frac{\log\bar{x}_h}{\bar{x}_h}\right)}{\left(\sum_{h=1}^{L} W_h\frac{\log\bar{x}_h^2}{\bar{x}_h}\right)}\right]$$

Case-II: $$\bar{y}_{bt.L2} = \sum_{h=1}^{L} W_h \bar{y}_h + \left[\hat{\beta}_{bt.L2} \sum_{h=1}^{L} W_h (\log \bar{X}_h - \log \bar{x}_h) \right]$$

where, $$\hat{\beta}_{bt.L2} = \left[\frac{\left(\sum_{h=1}^{L} \frac{W_h}{\bar{x}_h}\right)\left(\sum_{h=1}^{L} W_h \bar{y}_h \frac{\log \bar{x}_h}{\bar{x}_h}\right) - \left(\sum_{h=1}^{L} W_h \frac{\log \bar{x}_h}{\bar{x}_h}\right)\left(\sum_{h=1}^{L} W_h \frac{\bar{y}_h}{\bar{x}_h}\right)}{\left(\sum_{h=1}^{L} W_h \frac{\log \bar{x}_h^2}{\bar{x}_h}\right)\left(\sum_{h=1}^{L} \frac{W_h}{\bar{x}_h}\right) - \left(\sum_{h=1}^{L} W_h \frac{\log \bar{x}_h}{\bar{x}_h}\right)^2} \right]$$

3. The calibration estimators when $Q_h = \dfrac{1}{\log \bar{x}_h}$

Case-I: $$\bar{y}_{bt.L1} = \sum_{h=1}^{L} W_h \bar{y}_h + \left[\hat{\beta}_{bt.L1} \sum_{h=1}^{L} W_h (\log \bar{X}_h - \log \bar{x}_h) \right]$$

where, $$\hat{\beta}_{bt.L1} = \left[\frac{\left(\sum_{h=1}^{L} W_h \bar{y}_h\right)}{\left(\sum_{h=1}^{L} W_h \log \bar{x}_h\right)} \right]$$

Case-II: $$\bar{y}_{bt.L2} = \sum_{h=1}^{L} W_h \bar{y}_h + \left[\hat{\beta}_{bt.L2} \sum_{h=1}^{L} W_h (\log \bar{X}_h - \log \bar{x}_h) \right]$$

where, $$\hat{\beta}_{bt.L2} = \left[\frac{\left(\sum_{h=1}^{L} \frac{W_h}{\log \bar{x}_h}\right)\left(\sum_{h=1}^{L} W_h \bar{y}_h\right) - \left(\sum_{h=1}^{L} W_h \frac{\bar{y}_h}{\log \bar{x}_h}\right)}{\left(\sum_{h=1}^{L} W_h \log \bar{x}_h\right)\left(\sum_{h=1}^{L} \frac{W_h}{\log \bar{x}_h}\right) - 1} \right]$$

3. Simulation Study

A simulation study is carried out in order to study the performance of the proposed exponential type calibrated ratio estimators on two datasets: real and artificial populations. The stratified random samples are drawn using Proportional allocation and simple random sampling without replacement (SRSWOR) from each stratum. A simulated study generating 25,000 samples is performed using R-software.

The empirical percentage relative root mean squares error (%RRMSE) and percentage relative efficiency (%RE) of the estimators are computed using the following formulae:

$$\%RRMSE(\bar{y}_{\alpha}) = \sqrt{\frac{1}{25000}\sum_{i=1}^{25000}\left(\frac{(\bar{y}_{\alpha} - \bar{Y})}{\bar{Y}}\right)^2} \times 100 \ ; \ \alpha = s, st.bt, bt.L1, bt.L2$$

$$\%RE(\bar{y}_{\alpha}) = \frac{RRMSE(\bar{y}_s)}{RRMSE(\bar{y}_{\alpha})} \times 100; \quad \alpha = st.bt, bt.L1, bt.L2$$

3.1 Real Population

The mango production population (*https://data.gov.in/catalog/all-india-and-state-wise-area-and-production-fruits*) is considered for the simulation study. It comprises of 78 units, divided into two strata (North and South zone) of 30 and 48 sizes. The study variable Y is production of mangoes (in tonnes) and the auxiliary variable X is the area (in hectares) from the year 2009-2015. The mean of the study and the auxiliary variable are $\bar{Y} = 917.52$ and $\bar{X} = 106.67$, respectively, and the correlation coefficient between X and Y is $\rho_{xy} = 0.901$. The empirical percentage relative root mean squares error (%RRMSE) and percentage relative efficiency (%RE) of the estimators $\bar{y}_s$, $\bar{y}_{st.bt}$, $\bar{y}_{bt.L1}$ and $\bar{y}_{bt.L2}$ computed for mango population are given in Table 1 and 3, respectively.

3.2 Artificial Population

A finite population of size N=3000 is generated for 3 strata considering 1000 units in each stratum. The values for the auxiliary variable X are generated considering the Exponential distribution with varying values of the parameter for each stratum and the variable of interest Y is generated using the following models:

1st strata: $X_1 = Exp(1000, 2)$ and $Y_1 = 100 + (\beta_1 * X_1) + \varepsilon_1$

where $\beta_1 = 0.25$ and $\varepsilon_1 \sim N(0, 2)$

2nd strata: $X_2 = Exp(1000, 3)$ and $Y_2 = 200 + (\beta_2 * X_2) + \varepsilon_2$

where $\beta_2 = 0.50$ and $\varepsilon_2 \sim N(0, 4)$

3rd strata: $X_3 = Exp(1000, 8)$ and $Y_3 = 300 + (\beta_3 * X_3) + \varepsilon_3$

where $\beta_3 = 0.75$ and $\varepsilon_3 \sim N(0, 8)$

The empirical percentage relative root mean squares error (%RRMSE) and percentage relative efficiency (%RE) of the estimators $\bar{y}_s$, $\bar{y}_{st.bt}$, $\bar{y}_{bt.L1}$ and $\bar{y}_{bt.L2}$ computed for artificial population are given in Table 2 and 4, respectively.

Table 1. Percentage Relative Mean Square Error (%RRMSE) for Mango Population

Q_h	**Sample Size**	**%RRMSE** ($\bar{y}_s$)	**%RRMSE** ($\bar{y}_{st.bt}$)	**%RRMSE** ($\bar{y}_{bt.L1}$)	**%RRMSE** ($\bar{y}_{bt.L2}$)
1	**18**	24.47	16.03	14.10	11.63
	20	18.85	17.20	13.15	9.82
	22	16.70	18.60	12.34	9.33
	24	15.29	19.77	11.45	8.53
$\frac{1}{\bar{x}_h}$	**18**	24.47	16.03	15.86	11.63
	20	18.85	17.20	14.83	9.82
	22	16.70	18.60	13.72	9.33
	24	15.29	19.77	12.72	8.53
$\frac{1}{\log \bar{x}_h}$	**18**	24.47	16.03	14.53	11.63
	20	18.85	17.20	13.56	9.82
	22	16.70	18.60	12.68	9.33
	24	15.29	19.77	11.76	8.53

Table 2. Percentage Relative Root Mean Square Error (%RRMSE) for Artificial Population

Q_h	**Sample Size**	**%RRMSE** ($\bar{y}_s$)	**%RRMSE** ($\bar{y}_{st.bt}$)	**%RRMSE** ($\bar{y}_{bt.L1}$)	**%RRMSE** ($\bar{y}_{bt.L2}$)
1	**100**	9.40	5.56	3.17	2.77
	200	6.33	3.75	2.19	1.85
	300	5.04	3.00	1.77	1.46
	400	4.26	2.54	1.50	1.24
$\frac{1}{\bar{x}_h}$	**100**	9.41	5.56	3.03	2.56
	200	6.27	3.75	2.09	1.69
	300	4.98	3.00	1.68	1.33
	400	4.21	2.54	1.42	1.13
$\frac{1}{\log \bar{x}_h}$	**100**	9.51	5.56	3.35	3.03
	200	6.45	3.75	2.30	2.05
	300	5.14	3.00	1.85	1.63
	400	4.35	2.54	1.57	1.39

Table 3. Percentage Relative Efficiency (%RE) for Mango Population

Q_h	**Sample Size**	**%RE** ($\bar{y}_s$)	**%RE** ($\bar{y}_{st.bt}$)	**%RE** ($\bar{y}_{bt.L1}$)	**%RE** ($\bar{y}_{bt.L2}$)
1	**18**	100.00	152.64	173.59	210.48
	20	100.00	109.57	143.27	191.83
	22	100.00	89.79	135.29	179.09
	24	100.00	77.33	133.56	179.28
$\frac{1}{\bar{x}_h}$	**18**	100.00	152.64	154.28	210.48
	20	100.00	109.57	127.10	191.83
	22	100.00	89.79	121.72	179.09
	24	100.00	77.33	120.14	179.28
$\frac{1}{\log \bar{x}_h}$	**18**	100.00	152.64	168.44	210.48
	20	100.00	109.57	138.94	191.83
	22	100.00	89.79	131.72	179.09
	24	100.00	77.33	130.01	179.28

Table 4. Percentage Relative Efficiency (%RE) for Artificial Population

Q_h	**Sample Size**	**%RE** ($\bar{y}_s$)	**%RE** ($\bar{y}_{st.bt}$)	**%RE** ($\bar{y}_{bt.L1}$)	**%RE** ($\bar{y}_{bt.L2}$)
1	**18**	100.00	169.15	297.00	338.95
	20	100.00	168.74	288.53	342.41
	22	100.00	168.01	285.47	345.16
	24	100.00	167.88	284.59	344.68
$\frac{1}{\bar{x}_h}$	**18**	100.00	169.26	310.12	367.26
	20	100.00	167.12	299.65	370.74
	22	100.00	165.92	296.41	373.49
	24	100.00	165.67	295.31	373.05
$\frac{1}{\log \bar{x}_h}$	**18**	100.00	171.10	284.05	313.99
	20	100.00	171.97	280.02	314.33
	22	100.00	171.45	277.68	315.20
	24	100.00	171.20	277.29	313.92

4. Conclusion

In this paper, two exponential type ratio calibration estimators have been suggested using logarithmic mean of the auxiliary variable in defining the calibration constrains. A simulation study has been conducted on real and artificial datasets, in order to compare the efficiency of the proposed estimator with the estimators given by Bahl and Tuteja [1] and Singh [17]. It is observed from Table 1 and 4 that the proposed estimators are having less %RRMSE than the estimators given by Bahl and Tuteja [1] and Singh [17] for all values of Qh. It can be concluded that the proposed exponential type calibrated ratio estimators are more efficient than the estimators of Bahl and Tuteja [1] and Singh [17] under stratified random sampling as they are having less percentage relative efficiency for these datasets.

References

Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimator. Journal of Information and Optimization Sciences, 12(1), 159-163.

Basak, P., Sud, U. C. & Chandra, H. (2018). Calibration estimation of regression coefficient for two-stage sampling design using single auxiliary variable. Journal of the Indian Society of Agricultural Statistics, 72 (1), 1-6.

Clement, E. P. & Enang, E. I. (2017). On the efficiency of ratio estimator over the regression estimator. Communications in Statistics-Theory and Methods, 46(11), 357-5367.

Cochran, W.G. (1940). The estimation of yield of cereal experiments by sampling for the ratio of gain to total produce. Journal of Agricultural Science, 30, 262-275.

Deville, J. C. & Sarndal, C. E. (1992). Calibration estimators in survey sampling. Journal of the American Statistical Association, 87(418), 376-382.

Garg, N. & Pachori, M. (2020). Use of coefficient of variation in calibration estimation of population mean in stratified sampling. Communications in Statistics-Theory and Methods, 49(23), 5842−5852.

Guha, S., Sud, U. C. & Sisodia, B. V. S. (2018). Calibration approach-based chain ratio-product type estimator involving two auxiliary variables in two phase sampling. Journal of the Indian Society of Agricultural Statistics, 72 (3), 179-186.

Kim, J-M, Sungur, E. A. & Heo T-Y. (2007). Calibration approach estimators in stratified sampling. Statistics and Probability Letter, 77, 99-103.

Koyuncu, N., & Kadilar, C. (2013). Calibration estimator using different distance measures in stratified random sampling. International Journal of Modern Engineering Research, 3(1):415-419.

Koyuncu, N. & Kadilar, C. (2016). Calibration weighting in stratified random sampling. Communications in Statistics-Simulation and Computation, 45(7), 2267-2275.

Kumari, V., Chandra, H. & Bhar, L. M. (2018). Calibration estimator of regression coefficient using two auxiliary variables. Journal of the Indian Society of Agricultural Statistics, 72 (3), 193-199.

Malik S., Singh, V. K. & Singh, R. (2014). An Improved Estimator for Population Mean using Auxiliary Information in Stratified Random Sampling. Statistics in Transitions, 15(1), 59–66.

Mouhamed, A. M., Ei-Sheikh, A. A. & Mohamed, H. A. (2015). A new calibration estimator of Stratified random sampling. Applied Mathematical Sciences, 9(35), 1735-1744.

Sarndal, C. E., Swensson, B. & Wretman, J. (1992). Model assisted survey sampling. Springer Verlag, New York.

Singh, D., Sisodia, B. V. S., Rai, V. N. & Kumar, S. (2017). A calibration approach-based regression and ratio type estimators of finite population mean in two-stage stratified random sampling. Journal of the Indian Society of Agricultural Statistics, 71(3), 217-224.

Singh, S. & Arnab, R. (2011). On Calibration of design weights. Metron-International Journal of Statistics, LXIX (2),185-205.

Singh, S. (2003). Golden jubilee year 2003 of the linear regression estimator. Working paper at St. Cloud State University. St. Cloud, MN, USA.

Tracy, D. S., Singh, S. & Arnab, R. (2003). Note on Calibration in Stratified and Double Sampling. Survey Methodology, 29(1), 99-104.