Kurdish Studies
September 2024
Volume: 12, No: 5, pp 1320 -1327
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3476

Family Of Enhanced Estimators For Population Mean Using Auxiliary Information Under Simple Random Sampling

Shumaila Wazir[1], Muhammad Atif[1*], Neelam[1], Hassan Yar Ud Din[1], Soofia Iftikhar[2], Muhammad Farooq[1], Qamruz Zaman[1], Mujeeb Hussain[3]

[1]Department of Statistics, University of Peshawar, Pakistan
[2]Department of Statistics, Shaheed Benazir Bhutto Women University, Peshawar, Pakistan
[3]Departemt of Statistics, Government College Peshawar, Peshawar, Pakistan

***Corresponding author:** Muhammad Atif
*dratif96@gmail.com

Abstract
In this paper, we present an enhanced class of exponential-ratio-type estimators for estimating the finite population mean using simple random sampling without replacement. The proposed estimators depend on the careful and appropriate selection of the parameters of auxiliary information. The inclusion of auxiliary variable decreases the mean square error (MSE) and boosts the accuracy and efficiency. The expressions for bias and mean square error up to order one, along with the theoretical conditions are derived. The results show latest estimators performed better than the prior ones. In order to assess the estimators performance numerically, we use real data set to calculate their mean square error and percentage relative efficiency (PRE). Using numerical results, a comparative study show that the proposed estimators yielded more accurate results than existing ones. Therefore, the proposed family of enhanced estimators can be utilized to achieve better results as compared to the existing mean estimators for the population.

Key Words: Auxiliary variable, study variable, bias (B), MSE, Percentage relative efficiency.

Introduction:
Incorporating additional auxiliary information and relevant mathematical functions along with the research variable will enhance the accuracy and efficiency of estimating population parameters. Some statistical techniques, namely ratio, regression, and exponential estimators, are used to estimate the finite population mean. When there is a significant positive relationship between the research variable and auxiliary variables, then the ratio estimator is favorable. If it has an adverse correlation, then the product estimator is implemented. To strengthen the estimates of population parameters, we frequently make use of auxiliary variables. Cochran [1] is often regarded as the pioneer of the ratio estimation formula in survey sampling. He played a significant role in this field.

Numerous researchers have presented a variety of ratio-type estimators involving one or more auxiliary variables for estimating finite population means, for instance, Grover and Kaur [2] have contributed ratio type exponential estimators for estimating population means involving auxiliary variables in SRS. Gupta and Shabbir [3] focused on improving the estimation of the population mean in SRS. Haq and Shabbir [4], enhanced the family of ratio estimators in simple and stratified random sampling, Kadilar et al. [5], showed improvement in estimating the population means by using the correlation coefficient as auxiliary information. Koyuncu [6] proposed a new family of exponential estimators with the use of auxiliary attributes.

Various researchers have used product estimators when the correlation between research and auxiliary variables is negative. Several articles have been published in the case of an exponential ratio estimator using auxiliary variables by multiple researchers such as, Pandey and Dubey [7], modified product estimator using the coefficient of variation as an auxiliary variable. Ray and Sahai [8], proposed efficient families of ratio and product-type estimators, Singh et al. [9], used an improved estimator of the population mean using power transformation and so on. Similarly, exponential ratio cum product estimator using auxiliary attributes was proposed by Ahmad et al.[10], Kumar and Sunil [11] developed a generalized exponential-type estimator for the population mean using auxiliary attributes, focusing on improving the exponential ratio product estimator in the context of non-response. Haq and Shabbir [12] added a stepped-forward class of estimators for the finite population suggesting in SRS the use of an auxiliary attribute. This study presents a new kind of estimator applying exponential ratio-cum-product approaches, relying on the foundation laid by prior work [13].

The primary objective of this paper is to generate a new family of exponential ratio estimators that utilizes known population parameters such as mean, standard deviation, coefficient of correlation, covariance, coefficient of skewness, and coefficient of variation as an auxiliary variable in the context of SRSWOR.

Techniques and Terminologies:

A finite population having N units, expressed as $G = \{g_1, g_2, g_3, ...g_N\}$. We select a random sample of size n through SRSWOR, where each unit has an equal chance to appear. We estimate the mean of the research variable and auxiliary variable by "Y" and "X". For the i^{th} (i=1, 2,..., N) unit from G, let y_i stand for RG and x_i stand for AG. Let the population mean and sample mean of RG be expressed as $\overline{Y} = \frac{\sum_{i=1}^{N} y_i}{N}$, $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n}$ and their standard deviation is expressed as $S_y = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \overline{Y})^2}{N-1}}$, $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \overline{y})^2}{n-1}}$ respectively. The population and sample mean of AG is represented as $\overline{X} = \frac{\sum_{i=1}^{N} x_i}{N}$ and $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ while their standard deviations are $S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{X})}{N-1}}$ and $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$. $C_y = \frac{S_y}{\overline{Y}}$ and $C_x = \frac{S_x}{\overline{X}}$ be the co-efficient of variation of RG and AG.

$\beta_{1(x)} = \frac{N(\sum_{i=1}^{N} Xi - \overline{X})^3}{(N-1)(N-2)S^3}$, $\beta_{2(x)} = \frac{N(N+1)\sum_{i=1}^{N}(X_i - \overline{X})^4}{(N-1)(N-2)(N-3)S^4} - \frac{3(N-1)^2}{(N-2)(N-3)}$ be Coefficient of skewness and kurtosis of AG.

$T = C_{yx} = \rho_{yx} C_y C_x$

Error terms can be represented as $e_0 = \frac{(\overline{y} - \overline{Y})}{\overline{Y}}$ and $e_1 = \frac{(\overline{x} - \overline{X})}{\overline{X}}$, Expected values of error term can be expressed as

$E(e_0) = RC_y^2$ and $E(e_1) = RC_x^2$ where $R = \frac{(1-f)}{n} = \frac{(1-\frac{n}{N})}{n} = \frac{(N-n)}{Nn}$,

$\alpha = (\frac{u\overline{X}}{u\overline{X} + v})$, $K = \frac{a\overline{X}}{2(a\overline{X} + b)}$, $\omega = \frac{\overline{X}}{\overline{X} + \beta_{2(x)}}$ *and* $\omega_1 = \frac{\overline{X}\beta_{2(x)}}{\overline{X}\beta_{2(x)} + C_x}$

The inclusion of various types of auxiliary variables with variable of research not only decreases the bias but also increases the efficiency of the new and modified estimators.

Existing Estimators:

Various mean estimators and its MSE are given below:

The usual mean unbiased estimator is $\overline{U}_1 = \overline{y}$ (1)

$$\mathrm{Var}(\overline{U}_1) = R\overline{Y}^2 C_y^2 \quad (2)$$

Cochran [14] endorsed the conventional ratio estimator to estimate the population means, when there is a positive relationship between research and auxiliary variables. The usual ratio estimator and its MSE are as follow;

$$\overline{U}_2 = \frac{\overline{y}}{\overline{x}}\overline{X} \quad (3)$$

$$MSE(\overline{U}_2) = R\overline{Y}^2 (C_y^2 + C_x^2 - 2T) \quad (4)$$

The exponential ratio estimator idea was put forward by Bahl and Tuteja [15]

$$\overline{U}_3 = \overline{y}\exp(\frac{\overline{X} - \overline{x}}{\overline{X} + \overline{x}}) \quad (5)$$

$$MSE(\bar{U}_3) = R\bar{Y}^2 (C_y^2 + (\frac{1}{4})C_x^2 - T) \quad (6)$$

Singh et al. [16] proposed an exponential ratio type estimator:

$$\bar{U}_4 = \bar{y} \exp(\frac{u(\bar{X} - \bar{x})}{u(\bar{X} - \bar{x}) + 2\,v}) \quad (7)$$

$$MSE(\bar{U}_4) = R\bar{Y}^2 (C_y^2 + (\frac{1}{4})\alpha^2 C_x^2 - \alpha T) \quad (8)$$

Singh et al. [17] further introduced an exponential ratio estimator with the inclusion of auxiliary variable, given as:

$$\bar{U}_5 = \bar{y} \exp[\frac{(a\bar{X} + b) - (a\,\bar{x} + b)}{(a\bar{X} + b) + (a\bar{x} + b)}] \quad (9)$$

Where a is not identical to zero and b either be the real numbers or known values of population parameters.

$$MSE(\bar{U}_5) = R\bar{Y}^2 (C_y^2 + K^2 C_x^2 - 2KT) \quad (10)$$

Singh and Kakran [18] introduced a ratio estimator that increases the efficiency of the ratio estimator by adding the coefficient of kurtosis and the research variable.

$$\bar{U}_6 = \bar{y} (\frac{\bar{X} + \beta_{2(x)}}{\bar{x} + \beta_{2(x)}}) \quad (11)$$

$$MSE(\bar{U}_6) = R\bar{Y}^2 (C_y^2 + C_x^2 \omega(\omega - 2T)) \quad (12)$$

Upadhyaya and Singh [19] suggested an estimator by including the coefficient of variation and the coefficient of kurtosis as auxiliary variables.

$$\bar{U}_7 = \bar{y} (\frac{\bar{X}\beta_{2(x)} + C_x}{\bar{x}\beta_{2(x)} + C_x}) \quad (13)$$

$$MSE(\bar{U}_7) = R\bar{Y}^2 (C_y^2 + C_x^2\, \omega_1(\omega_1 - 2T)) \quad (14)$$

Proposed Estimator:

We present a new family of modified exponential ratio estimators for estimating the population mean with the help of known auxiliary variables. The recommended estimator is more efficient than all prior estimators, which decreases the bias and increases the PRE of the estimator which is defined as

$$\bar{U}_p = \left[\theta_1 \bar{y} + \theta_2 \left(\bar{X} - \bar{x}\right) - \theta_3 \left(\frac{\bar{X}}{\bar{x}}\right)\right] \exp\left[\frac{a\left(\bar{X} - \bar{x}\right)}{a\left(\bar{X} - \bar{x}\right) + 2b}\right] \quad (15)$$

In this context, θ_1, θ_2 *and* θ_3 are constants that will optimize the resulting MSE. The constants a and b are generalizing constants that could take value from the known parameters of the auxiliary variables. By adding various auxiliary information, we obtained different estimators.

The above estimator, in term of error can be expressed as

$$\bar{U}_p = \left[\theta_1 \bar{Y}\left(1 + \varepsilon_0\right) + \theta_2 \left\{\bar{X} - \bar{X}\left(1 + \varepsilon_1\right)\right\} - \theta_3 \left(\frac{\bar{X}}{\bar{X}(1 + \varepsilon_1)}\right)\right] \exp\left[\frac{a(\bar{X} - \bar{X}(1 + \varepsilon_1)}{a(\bar{X} + \bar{X}(1 + \varepsilon_1) + 2b}\right] \quad (16)$$

Simplified form of estimator using error term

$$\bar{U}_p = \left[\theta_1 \bar{Y}\left(1 + \varepsilon_0\right) + \theta_2 \left(\bar{X} - \bar{X}\right)\left(1 + \varepsilon_1\right)\right] - \theta_3 \left(\frac{\bar{X}}{\bar{X}(1 + \varepsilon_1)}\right) \exp\left[\frac{-\phi\varepsilon_1}{2}\left(1 + \frac{\phi\varepsilon_1}{2}\right)\right] \quad (17)$$

Where $\phi = \frac{a\,\bar{X}}{a\,\bar{X} + b}$

The difference of $\bar{Y}$ from both sides has shown in Eq (20)

$$\begin{aligned} \bar{U}_p - \bar{Y} = {} & \theta_1 \bar{Y}[1 - (\frac{1}{2})\phi\varepsilon_1 + (\frac{3}{8})\phi^2\varepsilon_1^2 + \varepsilon_0 - (\frac{1}{2})\phi\varepsilon_0\varepsilon_1] - \theta_2\, \bar{X}[\varepsilon_1 - (\frac{1}{2})\phi\varepsilon_1^2] \\ & -\theta_3[1 - (\frac{1}{2})\phi\varepsilon_1 + (\frac{3}{8})\phi^2\varepsilon_1^2 - \varepsilon_1 + (\frac{1}{2})\phi\varepsilon_1^2 + \varepsilon_1^2] - \bar{Y} \end{aligned} \quad (18)$$

Bias of proposed estimator

$$Bias(\bar{U}_P) = \theta_1 \bar{Y}[1+(\frac{3}{8})\phi^2 \lambda C_x^2 - (\frac{1}{2})\phi\lambda T] + \theta_2[(\frac{1}{2})\phi \bar{X} \lambda C_x^2] - \theta_3[1+(\frac{3}{8})\phi^2 \lambda C_x^2 + (\frac{1}{2})\phi\lambda C_x^2 + \lambda C_x^2] - \bar{Y} \quad (19)$$

MSE of proposed estimator

$$MSE(\bar{U}_p) = \theta_1^2[\bar{Y}^2 + \bar{Y}^2\phi^2\lambda C_x^2 + \bar{Y}^2\lambda C_y^2 - 2\bar{Y}^2\phi\lambda Cyx] + \theta_2^2[\bar{X}^2 \lambda C_x^2] + \theta_3^2[1+\phi^2\lambda C_x^2 + 3\lambda C_x^2 + 2\phi\lambda C_x^2] + \bar{Y}^2 + \theta_1\theta_2[2\phi\bar{Y}\,\bar{X}\lambda C_x^2 - 2\bar{Y}\,\bar{X}\lambda C_{yx}] + \theta_1\theta_3[-2\bar{Y} - 2\bar{Y}\phi^2\lambda C_x^2 - 2\bar{Y}\phi\lambda C_x^2 - 2\bar{Y}\lambda C_x^2 + 2\bar{Y}\lambda C_{yx} + 2\bar{Y}\phi\lambda C_{yx}] - \theta_2\theta_3[2\phi\bar{X}\lambda C_x^2 + 2\bar{X}\lambda C_x^2] - \theta_1[2\bar{Y}^2 + (\frac{3}{4})\phi^2\bar{Y}^2\lambda C_x^2 - \phi\bar{Y}^2\lambda T] - \theta_2[(\frac{1}{2})\phi\bar{X}\bar{Y}\lambda C_x^2] + \theta_3[\bar{Y} + (\frac{3}{8})\bar{Y}\phi^2\lambda C_x^2 + (\frac{1}{2})\bar{Y}\phi\lambda C_x^2 + \bar{Y}\lambda C_x^2] \quad (20)$$

Simplified form of MSE

$$MSE(\bar{U}_P) = \theta_1^2 A + \theta_2^2 B + \theta_3^2 C + \bar{Y}^2 + \theta_1\theta_2 D + \theta_1\theta_3 E - \theta_2\theta_3 F - \theta_1 G - \theta_2 H + \theta_3 I \quad (21)$$

Where

$A = \bar{Y}^2[1+\phi^2\lambda C_x^2 + \lambda C_y^2 - 2\phi\lambda Cyx]$, $B = [\bar{X}^2 \lambda C_x^2]$, $C = [1+\phi^2\lambda C_x^2 + 3\lambda C_x^2 + 2\phi\lambda C_x^2]$,

$D = 2\bar{Y}\bar{X}\lambda[\phi C_x^2 - \lambda C_{yx}]$, $E = 2\bar{Y}[-1-\phi^2\lambda C_x^2 - \phi\lambda C_x^2 - \lambda C_x^2 + \lambda C_{yx} + \phi\lambda C_{yx}]$,

$F = 2\bar{X}\lambda C_x^2[\phi+1]$, $G = \bar{Y}^2[2+(\frac{3}{4})\phi^2\lambda C_x^2 - \phi\lambda C_{yx}]$, $H = (\frac{1}{2})\phi\bar{X}\bar{Y}\lambda C_x^2$,

$I = \bar{Y}[1+(\frac{3}{8})\phi^2\lambda C_x^2 + (\frac{1}{2})\phi\lambda C_x^2 + \lambda C_x^2]$

To get the ideal values of $\theta_1,\ \theta_2\ and\ \theta_3$, we differentiate the eq (21) and equate to zero.

$$\frac{\partial MSE(\bar{U}_p)}{\partial\theta_1} = 0 \quad (22)$$

$$2A\theta_1 + D\theta_2 + E\theta_3 - G = 0 \quad (23)$$

Similarly

$$\frac{\partial MSE(\bar{U}_p)}{\partial\theta_2} = 0 \quad (24)$$

$$2B\theta_2 + D\theta_1 - F\theta_3 - H = 0 \quad (25)$$

Also

$$\frac{\partial MSE(\bar{U}_p)}{\partial\theta_3} = 0 \quad (26)$$

$$2\theta_3 C + \theta_1 E - \theta_2 F + I = 0 \quad (27)$$

By solving the equations (23, 25 and 27) simultaneously, we obtain $\theta_1,\ \theta_2$ and θ_3 values

$$\theta_1 = \frac{\frac{1}{2}(4BCG + 2IBE - 2CDH + IDF - EFH - F^2G)}{(4ABC - AF^2 - BE^2 - CD^2 - DEF)} \quad (28)$$

$$\theta_2 = \frac{\frac{1}{2}(4ACH + 2IAF - 2CDG - IDE - E^2H - EFG)}{(4ABC - AF^2 - BE^2 - CD^2 - DEF)} \tag{29}$$

$$\theta_3 = \frac{-\frac{1}{2}(4IAB - 2AFH + 2BEG - ID^2 - DEH - DFG)}{(4ABC - AF^2 - BE^2 - CD^2 - DEF)} \tag{30}$$

By putting θ_1, θ_2 and θ_3 values, we get

$$MSE(\bar{U}_p)_{(\min)} = \left[\frac{1}{4\left(4ABC - AF^2 - BE^2 - CD^2 + DEF\right)}\right]\begin{bmatrix} 16\,ABCY^2 - 4AF^2Y^2 - 4BE^2Y^2 \\ +4DEFY^2 + 4AB - 4ACH^2 - \\ 2IDFG - 4BCG^2 + 4IBEG + \\ 4CDGH - D^2 - 2IDEH + \\ 4IAFH + E^2H^2 - 2EFGH \\ +F^2G^2 \end{bmatrix} \tag{31}$$

Simplified form of MSE

$$MSE(\bar{U}_P) = \frac{1}{4}[\psi_1][\psi_2] \tag{32}$$

$$\psi_1 = \frac{1}{4ABC - AF^2 - BE^2 - CD^2 + DEF} \tag{33}$$

$$\psi_2 = 16\,ABCY^2 - 4AF^2Y^2 - 4BE^2Y^2 + 4DEFY^2 + 4AB \tag{34}$$

Family of Proposed Estimators

Using different choices of the parameters of the auxiliary variables in eq (15), different family members of the proposed estimators is given in Table 1.

Table 1: Different family members of the proposed estimators

Estimators	A	B
$\bar{U}_{p(\min)} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{\beta_{2(x)}(\bar{X}-\bar{x})}{\beta_{2(x)}(\bar{X}-\bar{x})+2S_{yx}}]$	$\beta_{2(x)}$	S_{yx}
$\bar{U}_{p1} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{\beta_{2(x)}(\bar{X}-\bar{x})}{\beta_{2(x)}(\bar{X}-\bar{x})+2S_x}]$	$\beta_{2(x)}$	S_x
$\bar{U}_{p2} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{S_x(\bar{X}-\bar{x})}{S_x(\bar{X}-\bar{x})+2S_{yx}}]$	S_x	S_{yx}
$\bar{U}_{p3} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{C_x(\bar{X}-\bar{x})}{C_x(\bar{X}-\bar{x})+2S_y}]$	C_x	S_y
$\bar{U}_{p4} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{\rho_{yx}(\bar{X}-\bar{x})}{\rho_{yx}(\bar{X}-\bar{x})+2S_y}]$	ρ_{yx}	S_y
$\bar{U}_{p5} = [\theta_1\bar{y} + \theta_2(\bar{X}-\bar{x}) - \theta_3(\frac{\bar{X}}{\bar{x}})]\exp[\frac{C_x(\bar{X}-\bar{x})}{C_x(\bar{X}-\bar{x})+2\beta_{2(x)}}]$	C_x	$\beta_{2(x)}$

Theoretical Comparisons:

By using the SRS technique, we have performed a theoretical contrast between a proposed estimator and the prior estimators. Existing estimators are not as efficient as the proposed estimator. List of estimators is given below.

(i) By comparing eq (32) with eq (2): $MSE(\bar{U}_p) \le MSE(\bar{U}_1)$ iff

$$\frac{1}{4}[\psi_1][\psi_2] - R\bar{Y}^2 C_y^2 \le 0$$

(ii) Comparing eq (32) and eq (4): $MSE(\bar{U}_P) \le MSE(\bar{U}_2)$ iff

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + C_x^2 - 2T)] \le 0$$

(iii) Comparing eq (32) and eq (6): $MSE(\bar{U}_P) \le MSE(\bar{U}_3)$

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + (\frac{1}{4})C_x^2 - T)] \le 0$$

(iv) Comparing eq (32) and eq (8): $MSE(\bar{U}_P) \le MSE(\bar{U}_4)$

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + (\frac{1}{4})\alpha^2 C_x^2 - \alpha T)] \le 0$$

(vi) Comparing eq(32) and eq(10): $MSE(\bar{U}_P) \le MSE(\bar{U}_5)$

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + K^2 C_x^2 - 2KT)] \le 0$$

(vii) Comparing eq (32) and eq (12): $MSE(\bar{U}_p) \le MSE(\bar{U}_6)$

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + C_x^2 \omega(\omega - 2T))] \le 0$$

(viii) Comparing eq (32) and eq (14): $MSE(\bar{U}_P) \le MSE(\bar{U}_7)$

$$\frac{1}{4}[\psi_1][\psi_2] - [R\bar{Y}^2 (C_y^2 + C_x^2 \omega_1(\omega_1 - 2T))] \le 0$$

Numerical Evaluation:

For numerical illustration, we use two data sets to evaluate the efficiency of current and proposed estimators. Data 1 is taken from Kadilar and Cingi [20] shown below. The information set represents Turkey's Aegean area accumulated using a simple random sampling scheme. The data set represents the quantity of apple production (as the study variable "y") and the number of apple trees (as the auxiliary variable "x") in 106 Aegean Place villages in 1990. Data 2 is taken from Kadilar and Cingi [21] shown below. MSE and PRE of existing and proposed estimators for data 1 and data 2 are provided in Table 2. The results proved that the proposed estimator shows an optimum result compared to other prior estimators. Different values of the data sets are described below:

Data 1:

N=106, n=20, $\bar{Y}$ =2212.59, $\bar{X}$ =27421.7, C_y =5.22, C_x =2.10, ρ_{yx} =0.86,

S_x =57460.61, S_y =11551.53, S_{yx} =568176176.1, $\beta_{2(x)}$ =34.57

Data 2:

N=200, n=50, $\bar{Y}$ =500, $\bar{X}$ =25, C_y =15, C_x =2, ρ_{yx} =0.9543,

S_x =50, S_y =7500, S_{yx} =35496.89, $\beta_{2(x)}$ =50

Table 2: MSE and PRE values for suggested and existing estimators using two data sets

S.No	Estimators	MSE/PRE	Data1	Data2
1	$\bar{U}_1$	MSE	5411348	843750
		PRE	100	100
2	$\bar{U}_2$	MSE	2542740	644032.5
		PRE	212.8156	131.0105
3	$\bar{U}_3$	MSE	3758095	740141.2
		PRE	143.9918	113.9985
4	$\bar{U}_4$	MSE	3758098	740310.3

		PRE	143.9917	113.9725
5	$\bar{U}_5$	MSE	3758956	791008.1
		PRE	143.9588	106.6677
6	$\bar{U}_6$	MSE	2545251	773844.2
		PRE	212.6057	109.0336
7	$\bar{U}_7$	MSE	2542745	644327.6
		PRE	212.8152	130.9505
8	$\bar{U}_{p(\min)}$	MSE	**225094.6**	**48513.08**
		PRE	**2404.033**	**1739.222**
9	$\bar{U}_{p_1}$	MSE	1120056	81919.39
		PRE	483.132	1029.976
10	$\bar{U}_{p_2}$	MSE	1091100	48513.08
		PRE	495.9533	1739.222
11	$\bar{U}_{p_3}$	MSE	1115730	48827.59
		PRE	485.0052	1728.019
12	$\bar{U}_{p_4}$	MSE	1065047	48481.94
		PRE	508.0854	1740.339
13	$\bar{U}_{p_5}$	MSE	1111852	77781.15
		PRE	486.6967	1084.774

Table 3: Numerical verification of conditions (i)-(vii) derived in efficiency comparisons

Conditions	Existing Estimators	Data	Proposed Estimators					
			$\bar{U}_{p(\min)}$	$\bar{U}_{p_1}$	$\bar{U}_{p_2}$	$\bar{U}_{p_3}$	$\bar{U}_{p_4}$	$\bar{U}_{p_5}$
i	$\bar{U}_1$	1	5186253	4291292	4320248	4295618	4346301	4299496
		2	795236.9	761830.6	795236.9	794922.4	795268.1	765968.9
ii	$\bar{U}_2$	1	2317645	1422684	1451640	1427010	1477693	1430888
		2	595519.4	562113.1	595519.4	595204.9	595550.6	566251.4
iii	$\bar{U}_3$	1	3533000	2638039	2666995	2642365	2693048	2646243
		2	691628.1	658221.8	691628.1	691313.6	691659.3	662360.1
iv	$\bar{U}_4$	1	3533003	2638042	2666998	2642368	2693051	2646246
		2	691797.2	658390.9	691797.2	691482.7	691828.4	662529.2
v	$\bar{U}_5$	1	3533861	2638900	2667856	2643226	2693909	2647104
		2	742495	709088.7	742495	742180.5	742526.2	713227
vi	$\bar{U}_6$	1	2320156	1425195	1454151	1429521	1480204	1433399
		2	725331.1	691924.8	725331.1	725016.6	725362.3	696063.1
vii	$\bar{U}_7$	1	2317650	1422689	1451645	1427015	1477698	1430893
		2	595814.5	562408.2	595814.5	595500	595845.7	566546.5

Results and Discussions:

This work proposed an exponential family of estimators under simple random sampling for finite population mean by incorporating auxiliary information. We used two real data sets to assess the performance of existing and proposed family of estimators through MSE and PRE criterion. Proposed family of estimators yields varying outcomes by using different known parameters such as $C_x, C_y, S_x, S_y, S_{yx}$ *and* $\beta_{2(x)}$ *etc.*

The subsequent main statements are presented as

i. Table 2 shows that all the proposed estimators $\bar{U}_{p_{(q)}}$ (q = min, 1, 2… 5) having smallest MSEs and largest PREs comparatively existing estimators. This reveals that proposed estimators outperform than existing ones.

ii. The combo of various known parameters $C_x, C_y, S_x, S_y, S_{yx}, \rho_{yx}$ *and* $\beta_{2(x)}$ yields different outcomes for proposed estimators.

iii. The proposed estimator $\bar{U}_{p(\min)}$ have minimum MSE and higher PRE for both the data sets i.e. (225094.6, 2404.033) and (48513.08, 1739.222).

iv. The theoretical efficiency criteria are verified as all of the numerical discrepancies are greater than zero.

v. Table 3 reveals that all the numerical differences are greater than zero which verifies the efficiency conditions described theoretically.

Conclusion:

A refined family of estimators for the finite population mean utilizing auxiliary information, has been proposed. The bias and mean squared error (MSE) are theoretically derived up to the first-order approximation. The proposed estimators are compared with existing ones both theoretically and numerically, using MSE and PRE as benchmarks. Numerical comparisons are conducted using two real-world datasets. The results clearly show that the proposed estimators consistently achieve lower MSEs as compared to traditional estimators while estimating the finite population mean through simple random sampling.

References:

1. W.G. Cochran, Sampling Techniques, John Wiley & Sons, 1977.
2. L.K. Grover, P. Kaur, A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable, Commun. Stat. Simulat. Comput. 43 (2014) 1552–1574.
3. S. Gupta, J. Shabbir, on improvement in estimating the population mean in simple random sampling, J. Appl. Stat. 35 (2008) 559–566.
4. A. Haq, J. Shabbir, Improved family of ratio estimators in simple and stratified random sampling, Commun. Stat.-Theory Methods 42 (2013) 782–799.
5. C. Kadilar, H. Ci˙ngi˙, H. Cingi, AN improvement in estimating the population mean by using the correlation coefficient, Hacet. J. Math. Stat. 35 (2006) 103–109.
6. N. Koyuncu, Efficient estimators of population mean using auxiliary attributes, Appl. Math. Comput. 218 (2012) 10900–10905.
7. B.N. Pandey, V. Dubey, Modified product estimator using coefficient of variation of auxiliary variate, Assam Stat. Rev. 2 (1988) 64–66.
8. S.K. Ray, A. Sahai, Efficient families of ratio and product-type estimators, Biometrika 67 (1980) 211–215.
9. H.P. Singh, Rajesh Tailor, Ritesh Tailor, M.S. Kakran, An improved estimator of population mean using power transformation, J. Indian Soc. Agric. Stat. 58 (2004) 223–230.
10. S. Ahmad, M. Arslan, A. Khan, J. Shabbir, A generalized exponential-type estimator for population mean using auxiliary attributes, PLoS One 16 (2021) e0246947.
11. S. Kumar, S. Bhougal, Estimation of the population mean in Presence of non-response, Commun. Stat. Appl. Methods 18 (2011) 537–54
12. A. Haq, J. Shabbir, An improved class of estimators of finite population mean in simple random sampling using an auxiliary attribute, J. Stat. Theory Pract. 12 (2018) 282–289, https://doi.org/10.1080/15598608.2017.1349012
13. J. Shabbir, S. Gupta, R. Onyango, On using the conventional and nonconventional measures of the auxiliary variable for mean estimation, Math. Probl Eng. 2021 (2021) 1–13.
14. W. G. Cochran, "(e estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce," The Journal of Agricultural Science, vol. 30, no. 2, pp. 262–275, 1940.
15. S. Bahl and R.K Tuteja, Ratio and product type exponential estimators. Information and Optimization Sciences 12 (1), 159-163 (1991).
16. R. Singh, P.Chauhan, N. Swan, and F. Smarandache, "Improvement in estimating the population mean using exponential estimator in simple random sampling," Bulletin of StatisticsandEconomics,vol.3,pp.13–18,2009.
17. Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). Improvement in estimating the population mean using exponential estimator in simple random sampling. *Auxiliary Information and a priori Values in Construction of Improved Estimators*, *33*.
18. H. P. Singh and M. S. Kakran, A modified ratio estimator using known coefficient of kurtosis of an auxiliary character, Revised version submitted to Journal of Indian Society of Agricultural Statistics, New Delhi, India, 1993.
19. L. N. Upadhyaya and H. P. Singh, Use of transformed auxiliary variable in estimating the finite population mean, Biom. J. 41(5) (1999), 627-636.
20. C. Kadilar and H. Cingi, Ratio estimator in simple random sampling, J. Appl. Math. Comput. 151 (2004), 893-902.
21. Kadilar , C. and Cingi , H. 2006.New ratio estimators using correlation coefficient . *Interstat* , 4 : 1 – 11 .