Journal of King Saud University – Science 32 (2020) 2835–2844

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

Almost unbiased optimum estimators for population mean using dual auxiliary information

Muhammad Irfan *, Maria Javed, Sajjad Haider Bhatti, Muhammad Ali Raza, Tanvir Ahmad

Department of Statistics, Government College University, Faisalabad, Pakistan

ARTICLE INFO

Article history:
Received 9 October 2019
Revised 6 July 2020
Accepted 6 July 2020
Available online 11 July 2020

Keywords:
Auxiliary variable
Hartley-Ross type estimator
Ranked auxiliary variable
Unbiased
Variance

ABSTRACT

One eminent disadvantage of many existing optimal estimators/class of estimators is that they are typically biased. In this article, we proposed an optimum class of unbiased estimators for estimating the population mean under simple random sampling without replacement (SRSWOR) scheme. Proposed class is a blend of three concepts: 1) information on auxiliary variable, 2) the ranks of auxiliary variable and 3) Hartley-Ross type unbiased estimation procedure. Expressions for the bias and the minimum variance of the new class are derived up to first degree of approximation. To highlight the application of proposed class, five real data sets are used. Numerical findings confirm that the new class behaves efficiently as compared to traditional unbiased estimator and other almost unbiased estimators under study. In addition, Monte Carlo simulation study is conducted through two real populations to assess the performance of proposed class against competitors. On the basis of theoretical and numerical findings, it is concluded that new proposed class can generate optimum unbiased estimators under SRSWOR scheme. Therefore, use of proposed class is recommended for future applications.

1. Introduction

Utilizing the auxiliary information to boost the efficiency of estimators is a common practice in the theory of survey sampling. The auxiliary information such as standard deviation S_x, coefficient of variationC_x, coefficient of skewness $\beta_{1(x)}$, coefficient of kurtosis $\beta_{2(x)}$, coefficient of correlation ρ_{yx} etc. may play positive role in the selection of sample, strata, type of estimators or in estimation. If this auxiliary information is positively (high) correlated with study variable, ratio estimators are preferred and in case it is negatively (high) correlated, product estimators are used. In this context, some notable contributions were made by Upadhyaya and Singh (1999), Singh and Tailor (2003), Kadilar and Cingi (2006a, 2006b), Gupta and Shabbir (2008), Shabbir and Gupta (2011), Haq and Shabbir (2013), Singh and Solanki (2013), Irfan et al. (2019a, 2019b), Raza et al. (2020) and many others.

Consider a sample of n pair of observations $(y_i, x_i), i = 1, 2, 3, \cdots, n$ for the study and auxiliary variables, respectively are selected from a finite population $\Theta = \{\Theta_1, \Theta_2, \Theta_3, \cdots, \Theta_N\}$of size $''N''$ under simple random sampling without replacement (SRSWOR) subject to the constraint$n < N$. Let r denotes the ranks of auxiliary variable and r_i denotes the $i$thvalue of r in the population. Important measures related to study variable y, auxiliary variable x and the ranks of auxiliary variable r are described in Table 1.

Remark 1.1.

- $\bar{y}, \bar{x}$and$\bar{r}$ *are the unbiased estimators of* $\bar{Y}, \bar{X}$and$\bar{R}$, *respectively*.
- s_y^2, s_x^2ands_r^2 are also unbiased estimators of S_y^2, S_x^2 and S_r^2, respectively.
- Similarly,$s_{yx}, s_{yr}$and$s_{xr}$are the unbiased estimators of their population parameters S_{yx}, S_{yr}andS_{xr} respectively.

2. Unbiased/Almost unbiased estimators from literature

Usually, ratio and product type of estimators of population mean are biased and inconsistent and thus can lead to erroneous inferences. Several researchers have attempted to reduce the bias from these estimators as unbiasedness is one of the important properties of estimators. Unbiased ratio and product

* Corresponding author.
E-mail addresses: mirfan@gcuf.edu.pk (M. Irfan), mariajaved@gcuf.edu.pk (M. Javed), ali.raza@gcuf.edu.pk (M.A. Raza), dr_tanvir@gcuf.edu.pk (T. Ahmad).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksus.2020.07.006
1018-3647/

Table 1
Measures related to study variable, auxiliary variable and the ranks of auxiliary variable.

$\left.\begin{array}{l}\overline{Y}=N^{-1}\sum_{i=1}^{N}y_i\\ \overline{X}=N^{-1}\sum_{i=1}^{N}x_i\\ \overline{R}=N^{-1}\sum_{i=1}^{N}r_i\end{array}\right\}$Population means	$\left.\begin{array}{l}\overline{y}=n^{-1}\sum_{i=1}^{n}y_i\\ \overline{x}=n^{-1}\sum_{i=1}^{n}x_i\\ \overline{r}=n^{-1}\sum_{i=1}^{n}r_i\end{array}\right\}$Sample means
$\left.\begin{array}{l}S_y^2=\frac{\sum_{i=1}^{N}(y_i-\overline{Y})^2}{(N-1)}\\ S_x^2=\frac{\sum_{i=1}^{N}(x_i-\overline{X})^2}{(N-1)}\\ S_r^2=\frac{\sum_{i=1}^{N}(r_i-\overline{R})^2}{(N-1)}\end{array}\right\}$Population variances	$\left.\begin{array}{l}s_y^2=\frac{\sum_{i=1}^{n}(y_i-\overline{y})^2}{(n-1)}\\ s_x^2=\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{(n-1)}\\ s_r^2=\frac{\sum_{i=1}^{n}(r_i-\overline{r})^2}{(n-1)}\end{array}\right\}$Sample variances
$\left.\begin{array}{l}\rho_{yx}=(S_yS_x)^{-1}S_{yx}\\ \rho_{yr}=(S_yS_r)^{-1}S_{yr}\\ \rho_{xr}=(S_xS_r)^{-1}S_{xr}\end{array}\right\}$Correlation coefficients	$\left.\begin{array}{l}C_y=(\overline{Y})^{-1}S_y\\ C_x=(\overline{X})^{-1}S_x\\ C_r=(\overline{R})^{-1}S_{rx}\end{array}\right\}$Coefficients of variation
$\left.\begin{array}{l}S_{yx}=(N-1)^{-1}\sum_{i=1}^{N}(y_i-\overline{Y})(x_i-\overline{X})\\ S_{yr}=(N-1)^{-1}\sum_{i=1}^{N}(y_i-\overline{Y})(r_i-\overline{R})\\ S_{xr}=(N-1)^{-1}\sum_{i=1}^{N}(x_i-\overline{X})(r_i-\overline{R})\end{array}\right\}$Population covariances	
$\left.\begin{array}{l}s_{yx}=(n-1)^{-1}\sum_{i=1}^{n}(y_i-\overline{y})(x_i-\overline{x})\\ s_{yr}=(n-1)^{-1}\sum_{i=1}^{n}(y_i-\overline{y})(r_i-\overline{r})\\ s_{xr}=(n-1)^{-1}\sum_{i=1}^{n}(x_i-\overline{x})(r_i-\overline{r})\end{array}\right\}$Sample covariances	
$\varphi=\left(\frac{1}{n}-\frac{1}{N}\right)$Finite population correction factor	

type estimators have also been discussed by Hartely and Ross (1954), Robson (1957), Murthy and Nanjamma (1959), Biradar and Singh (1992a, 1992b, 1995), Sahoo et al. (1994) and Javed et al. (2019).

This section presents a comprehensive detail of unbiased/almost unbiased estimators of population mean under simple random sampling scheme from literature.

2.1. Traditional unbiased estimator

The traditional unbiased estimator of population mean along with its variance is

$$\overline{y}_0^{(u)}=\overline{y} \tag{1}$$

$$V\left(\overline{y}_0^{(u)}\right)=\varphi\overline{Y}^2C_y^2 \tag{2}$$

2.2. Hartley and Ross (1954) *estimator*

Hartley and Ross (1954) suggested an unbiased ratio type estimator for estimating population mean as below

$$\overline{y}_{HR}^{(u)}=\overline{p}^{(0)}\overline{X}+\frac{n(N-1)}{N(n-1)}\left(\overline{y}-\overline{p}^{(0)}\overline{x}\right) \tag{3}$$

$$\text{where}\,\overline{p}^{(0)}=n^{-1}\sum_{i=1}^{n}p_i^{(0)},p_i^{(0)}=\frac{y_i}{x_i}$$

The variance of this estimator, to the first order of approximation, is equal to the mean square error of the usual ratio estimator (see Singh and Mangat (1996)).

$$V\left(\overline{y}_{HR}^{(u)}\right)\cong\varphi\overline{Y}^2\left[C_y^2+C_x^2-2\rho_{yx}C_yC_x\right] \tag{4}$$

2.3. Singh et al. (2014) *estimators*

Singh et al. (2014) considered the estimators of Kadilar and Cingi (2006c) and Upadhyaya and Singh (1999) to propose the following Hartley-Ross type unbiased estimators for population mean.

$$\overline{y}_{S1}^{(u)}=\overline{p}^{(1)}\overline{X}^{(1)}+\frac{n(N-1)}{N(n-1)}\left(\overline{y}-\overline{p}^{(1)}\overline{x}^{(1)}\right) \tag{5}$$

$$\overline{y}_{S2}^{(u)}=\overline{p}^{(2)}\overline{X}^{(2)}+\frac{n(N-1)}{N(n-1)}\left(\overline{y}-\overline{p}^{(2)}\overline{x}^{(2)}\right) \tag{6}$$

$$\text{where}\,\overline{p}^{(1)}=n^{-1}\sum_{i=1}^{n}p_i^{(1)},p_i^{(1)}=\frac{y_i}{C_xx_i+\rho_{yx}}=\frac{y_i}{x_i^{(1)}},\overline{X}^{(1)}=C_x\overline{X}+\rho_{yx}$$

$$\overline{p}^{(2)}=n^{-1}\sum_{i=1}^{n}p_i^{(2)},p_i^{(2)}=\frac{y_i}{C_xx_i+\beta_{2(x)}}=\frac{y_i}{x_i^{(2)}},\overline{X}^{(2)}=C_x\overline{X}+\beta_{2(x)}$$

here ρ_{yx} is the coefficient of correlation between study variable y and auxiliary variable x and $\beta_{2(x)}$ is the coefficient of kurtosis of auxiliary variable x.

Variance of $\overline{y}_{S1}^{(u)}$ and $\overline{y}_{S2}^{(u)}$ are respectively given below.

$$V\left(\overline{y}_{S1}^{(u)}\right)\cong\varphi\left(S_y^2+\left(\overline{P}^{(1)}S_{x^{(1)}}\right)^2-2\overline{P}^{(1)}S_{yx^{(1)}}\right) \tag{7}$$

$$V\left(\overline{y}_{S2}^{(u)}\right)\cong\varphi\left(S_y^2+\left(\overline{P}^{(2)}S_{x^{(2)}}\right)^2-2\overline{P}^{(2)}S_{yx^{(2)}}\right) \tag{8}$$

$$\text{where}\,\overline{P}^{(\star)}=N^{-1}\sum_{i=1}^{N}p_i^{(\star)},S_{x^{(\star)}}=\sqrt{(N-1)^{-1}\sum_{i=1}^{N}\left(x_i^{(\star)}-\overline{X}^{(\star)}\right)^2},$$

$$S_{yx^{(\star)}}=(N-1)^{-1}\sum_{i=1}^{N}\left(y_i-\overline{Y}\right)\left(x_i^{(\star)}-\overline{X}^{(\star)}\right),\text{for}\star=1,2$$

2.4. Cekim and Kadilar (2016) *estimators*

A general class of Hartley-Ross type unbiased estimators was developed by Cekim and Kadilar (2016) from special version of estimators of Khoshnevisan et al. (2007) as given below

$$\overline{y}_{CK1}^{(u)}=\overline{p}^{(3)}\overline{X}^{(3)}+\frac{n(N-1)}{N(n-1)}\left(\overline{y}-\overline{p}^{(3)}\overline{x}^{(3)}\right) \tag{9}$$

$$\text{where}\,\overline{p}^{(3)}=n^{-1}\sum_{i=1}^{n}p_i^{(3)},p_i^{(3)}=\frac{y_i}{\alpha x_i+\beta}=\frac{y_i}{x_i^{(3)}},\overline{X}^{(3)}=\alpha\overline{X}+\beta$$

$\alpha(\neq0)$ and β are either known constants or functions of any known population parameters of auxiliary variable including coefficient of skewness, coefficient of kurtosis, coefficient of variation and coefficient of correlation etc.

Variance of $\overline{y}_{CK1}^{(u)}$ is given as under

$$V\left(\bar{y}^{(u)}_{CK1}\right) \cong \varphi\left(S_y^2 + \left(\bar{P}^{(3)}S_{x^{(3)}}\right)^2 - 2\bar{P}^{(3)}S_{yx^{(3)}}\right) \quad (10)$$

where $\bar{P}^{(3)} = N^{-1}\sum_{i=1}^{N} p_i^{(3)}, S_{x^{(3)}} = \sqrt{(N-1)^{-1}\sum_{i=1}^{N}\left(x_i^{(3)} - \bar{X}^{(3)}\right)^2}$,

$$S_{yx^{(3)}} = (N-1)^{-1}\sum_{i=1}^{N}(y_i - \bar{Y})\left(x_i^{(3)} - \bar{X}^{(3)}\right)$$

Remark 2.1. *It is worth pointing out that if we have*

1) $\alpha = C_x$ and $\beta = \rho_{yx}$ in $p_i^{(3)}$, then
$p_i^{(3)} = p_i^{(1)}$ and $\bar{y}^{(u)}_{CK1} = \bar{y}^{(u)}_{S1}$
2) $\alpha = C_x$ and $\beta = \beta_{2(x)}$ in $p_i^{(3)}$, then
$p_i^{(3)} = p_i^{(2)}$ and $\bar{y}^{(u)}_{CK1} = \bar{y}^{(u)}_{S2}$

Another class proposed by Cekim and Kadilar (2016) using the special version of Koyuncu and Kadilar (2009) is defined below

$$\bar{y}^{(u)}_{CK2} = k_1\bar{y}\left[\frac{\alpha\bar{X}+\beta}{\gamma(\alpha\bar{x}+\beta)+(1-\gamma)(\alpha\bar{X}+\beta)}\right]^t - k_1\bar{y}\varphi\left[\frac{t(t+1)}{2}\gamma^2\lambda^2C_x^2 - t\gamma\lambda\frac{s_{yx}}{\bar{y}\bar{X}}\right] - (k_1-1)\bar{y} \quad (11)$$

here $\lambda = \frac{\alpha\bar{X}}{\alpha\bar{X}+\beta}, t = 1, \gamma = 1,\ k_1$ is the weight to be determined such that the variance becomes minimum and α and β are the same as defined earlier.

$$V\left(\bar{y}^{(u)}_{CK2}\right) \cong \varphi\bar{Y}^2\left[\left\{k_1^2t^2\gamma^2\lambda^2C_x^2 - 2k_1t\gamma\lambda C_{yx} + C_y^2\right\} - \varphi\left\{k_1t\gamma\lambda\left(\frac{t+1}{2}\gamma\lambda C_x^2 - C_{yx}\right)\right\}^2 - \varphi k_1t\gamma\lambda\left\{2\frac{C_{yx}}{\rho_{yx}}(k_1t\gamma\lambda C_x\theta_{12x} - C_y\theta_{21x}) - (t+1)\gamma\lambda C_x^2\left(k_1t\gamma\lambda C_{yx} - C_y^2\right)\right\}\right] \quad (12)$$

Differentiating Eq. (12) with respect to k_1 and equating to zero, we get the optimal value of k_1 as follows

$$k_{1(Opt)} = \frac{A}{B}$$

where

$$A = t\gamma\lambda\left[C_{yx}\left(1 - \frac{\varphi C_y\theta_{21x}}{\rho_{yx}}\right) + \frac{(t+1)}{2}\varphi\gamma\lambda C_x^2C_y^2\right]$$

$$B = t^2\gamma^2\lambda^2\left[C_x^2 + \varphi\left\{C_{yx}\left((t+1)\gamma\lambda C_x^2 - 2\frac{C_x\theta_{12x}}{\rho_{yx}}\right) - \left(\frac{(t+1)}{2}\gamma\lambda C_x^2 - C_{yx}\right)^2\right\}\right]$$

Putting the optimal value of k_1 in Eq. (12), we get the minimum variance as

$$V_{min}\left(\bar{y}^{(u)}_{CK2}\right) \cong \varphi\bar{Y}^2\left[C_y^2 - \frac{A^2}{B}\right] \quad (13)$$

3. Methodology

All contributions for efficient estimation of population mean under simple random sampling scheme and alike published work are based on only the utilization of original auxiliary information. None of them tried the dual use of auxiliary information to explore the unbiased estimators for population mean under simple random sampling.

Recently, Irfan et al. (2020) and Javed and Irfan (2020) used an additional information of the auxiliary variable called ranked auxiliary variable to develop efficient estimators under simple and stratified random sampling.

First time, we initiated a blend of three concepts to explore an optimum class of almost unbiased estimators for estimating the population mean:

i) information on auxiliary variable
ii) the ranks of auxiliary variable
iii) Hartley-Ross type unbiased estimation

A class of biased estimators proposed by Haq et al. (2017) is as follows:

$$\bar{y}_H = \left[k_2\bar{y} + k_3(\bar{X} - \bar{x}) + k_4(\bar{R} - \bar{r})\right]\exp\left(\frac{\alpha(\bar{X} - \bar{x})}{\alpha(\bar{X} + \bar{x}) + 2\beta}\right) \quad (14)$$

Bias of the class given in Eq. (14) is derived, up to first order of approximation as

$$Bias(\bar{y}_H) \cong -\bar{Y} + \frac{1}{2}\varphi\lambda C_x\left\{k_3\bar{X}C_x + k_4\bar{R}C_r\rho_{xr}\right\} + k_2\bar{Y}\left\{1 + \varphi\lambda C_x\left(\frac{3}{8}\lambda C_x - \frac{1}{2}\frac{S_{yx}}{\bar{Y}S_x}\right)\right\} \quad (15)$$

Subtracting Eq. (15) from Eq. (14), we obtained the expression given below

$$\left[k_2\bar{y} + k_3(\bar{X} - \bar{x}) + k_4(\bar{R} - \bar{r})\right]\exp\left(\frac{\alpha(\bar{X} - \bar{x})}{\alpha(\bar{X} + \bar{x}) + 2\beta}\right) + \bar{Y} - \frac{1}{2}\varphi\lambda C_x\left\{k_3\bar{X}C_x + k_4\bar{R}C_r\rho_{xr}\right\} - k_2\bar{Y}\left\{1 + \varphi\lambda C_x\left(\frac{3}{8}\lambda C_x - \frac{1}{2}\frac{S_{yx}}{\bar{Y}S_x}\right)\right\} \quad (16)$$

After some simplification and replacing the parameters $\bar{Y}$ and S_{yx} by their unbiased estimators $\bar{y}$ and s_{yx} in Eq. (16), we have

$$\bar{y}_H + \bar{y} - \frac{1}{2}k_3\varphi\lambda\bar{X}C_x^2 - \frac{1}{2}k_4\varphi\lambda C_{xr}\bar{R} - k_2\bar{y} - \frac{3}{8}k_2\varphi\lambda^2\bar{y}C_x^2 + \frac{1}{2}k_2\varphi\lambda\frac{s_{yx}}{\bar{X}}$$

So, the proposed class of almost unbiased estimators is as follows:

$$\bar{y}^{(u)}_P = \bar{y}_H + \bar{y}(1 - k_2) - \frac{1}{2}\varphi\lambda\left(k_3\bar{X}C_x^2 + k_4\bar{R}C_{xr} + \frac{3}{4}k_2\lambda\bar{y}C_x^2 - k_2\frac{s_{yx}}{\bar{X}}\right) \quad (17)$$

where $\lambda = \frac{\alpha\bar{X}}{\alpha\bar{X}+\beta}$, ($k_2, k_3$ and k_4) are the suitable weights to be chosen. $\alpha(\neq 0)$ and β are either known constants or functions of any known population parameters of auxiliary variable including coefficient of skewness $\beta_{1(x)}$, coefficient of kurtosis $\beta_{2(x)}$, coefficient of variation C_x and coefficient of correlation ρ_{yx}etc.

Following are the relative error terms along with their expectations, used to derive the expressions for the bias, variance and minimum variance of the proposed estimators.

$$\omega_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}, \omega_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}, \omega_2 = \frac{\bar{r} - \bar{R}}{\bar{R}}, \omega_3 = \frac{s_{yx} - S_{yx}}{S_{yx}}$$

such that

$E(\omega_i) = 0$ for $i = 0, 1, 2, 3$

$$E(\omega_0^2) = \varphi C_y^2, E(\omega_1^2) = \varphi C_x^2, E(\omega_2^2) = \varphi C_r^2, E(\omega_3^2) = \varphi\left(\frac{\theta_{22x}}{\rho_{yx}} - 1\right),$$

$$E(\omega_0\omega_1) = \varphi\rho_{yx}C_yC_x = \varphi C_{yx}, E(\omega_0\omega_2) = \varphi\rho_{yr}C_yC_r = \varphi C_{yr},$$

$$E(\omega_0\omega_3) = \varphi\left(\frac{C_y\theta_{21x}}{\rho_{yx}}\right), E(\omega_1\omega_2) = \varphi\rho_{xr}C_xC_r = \varphi C_{xr},$$

$$E(\omega_1\omega_3) = \varphi\left(\frac{C_x\theta_{12x}}{\rho_{yx}}\right), E(\omega_2\omega_3) = \varphi\left(\frac{C_r\theta_{12r}}{\rho_{yr}}\right).$$

In order to obtain the values of $\theta_{21x}, \theta_{12x}, \theta_{22x}$ and θ_{12r}, following expressions are helpful.

$$\left.\begin{array}{l}\theta_{abx} = \frac{\mu_{abx}}{\left(\mu_{20x}^{\frac{a}{2}}\right)\left(\mu_{02x}^{\frac{b}{2}}\right)}, \mu_{abx} = \frac{\sum_{i=1}^{N}\left(y_i-\overline{Y}\right)^a\left(x_i-\overline{X}\right)^b}{N}\\ \theta_{abr} = \frac{\mu_{abr}}{\left(\mu_{20r}^{\frac{a}{2}}\right)\left(\mu_{02r}^{\frac{b}{2}}\right)}, \mu_{abr} = \frac{\sum_{i=1}^{N}\left(y_i-\overline{Y}\right)^a\left(r_i-\overline{R}\right)^b}{N}\end{array}\right\} \text{ for } a = b = 0,1,2$$

After rewriting $\overline{y}_P^{(u)}$ in terms of relative errors and expanding up to first order of approximation, we get

$$\begin{aligned}\overline{y}_P^{(u)} \cong & \left[k_2\overline{Y}(1+\omega_0) - k_3\overline{X}\omega_1 - k_4\overline{R}\omega_2\right]\left[1 - \frac{1}{2}\lambda\omega_1 + \frac{3}{8}\lambda^2{\omega_1}^2\right]\\ & + \overline{Y}(1+\omega_0) - \frac{1}{2}k_3\varphi\lambda\overline{X}C_x^2 - \frac{1}{2}k_4\varphi\lambda C_{xr}\overline{R} - k_2\overline{Y}(1+\omega_0)\\ & - \frac{3}{8}k_2\varphi\lambda^2C_x^2\overline{Y}(1+\omega_0) + \frac{1}{2}k_2\varphi\lambda\frac{S_{yx}(1+\omega_3)}{\overline{X}}\end{aligned} \tag{18}$$

Subtracting $\overline{Y}$ from both sides of Eq. (18), we have

$$\begin{aligned}\left(\overline{y}_P^{(u)} - \overline{Y}\right) \cong & \overline{Y}\left(1 - \frac{3}{8}k_2\varphi\lambda^2C_x^2\right)\omega_0 - \left(\frac{1}{2}k_2\overline{Y}\lambda + k_3\overline{X}\right)\omega_1\\ & - k_4\overline{R}\omega_2 + \frac{1}{2}k_2\varphi\lambda\frac{S_{yx}}{\overline{X}}\omega_3 - \frac{1}{2}k_2\overline{Y}\lambda\omega_0\omega_1\\ & + \frac{1}{2}k_4\overline{R}\lambda\omega_1\omega_2 + \left(\frac{3}{8}k_2\overline{Y}\lambda^2 + \frac{1}{2}k_3\overline{X}\lambda\right)\omega_1^2\\ & - \frac{1}{2}\varphi\lambda\left(k_3\overline{X}C_x^2 + k_4C_{xr}\overline{R} + \frac{3}{4}k_2\overline{Y}\lambda C_x^2 - k_2\overline{Y}C_{yx}\right)\end{aligned} \tag{19}$$

Taking expectation on both sides of Eq. (19) to get the $Bias\left(\overline{y}_P^{(u)}\right)$

$$\begin{aligned}Bias\left(\overline{y}_P^{(u)}\right) &= E\left(\overline{y}_P^{(u)} - \overline{Y}\right)\\ &= -\frac{1}{2}k_2\overline{Y}\lambda\varphi C_{yx} + \frac{1}{2}k_4\overline{R}\lambda\varphi C_{xr} + \left(\frac{3}{8}k_2\overline{Y}\lambda^2 + \frac{1}{2}k_3\overline{X}\lambda\right)\varphi C_x^2\\ &\quad - \frac{1}{2}\varphi\lambda\left(k_3\overline{X}C_x^2 + k_4C_{xr}\overline{R} + \frac{3}{4}k_2\overline{Y}\lambda C_x^2 - k_2\overline{Y}C_{yx}\right)\end{aligned}$$

$$\begin{aligned}E\left(\overline{y}_P^{(u)} - \overline{Y}\right) = & -\frac{1}{2}k_2\overline{Y}\lambda\varphi C_{yx} + \frac{1}{2}k_4\overline{R}\lambda\varphi C_{xr} + \frac{3}{8}k_2\overline{Y}\lambda^2\varphi C_x^2\\ & + \frac{1}{2}k_3\overline{X}\lambda\varphi C_x^2 + \frac{1}{2}k_2\overline{Y}\lambda\varphi C_{yx} - \frac{1}{2}k_4\overline{R}\lambda\varphi C_{xr}\\ & - \frac{3}{8}k_2\overline{Y}\lambda^2\varphi C_x^2 - \frac{1}{2}k_3\overline{X}\lambda\varphi C_x^2\end{aligned}$$

During simplification, all the terms cancel out and we get zero bias which shows that the proposed class generates *almost unbiased* estimators. As the first order approximation is used in deriving the expression therefore the term "*almost*" is added here. So,

$$Bias\left(\overline{y}_P^{(u)}\right) = E\left(\overline{y}_P^{(u)} - \overline{Y}\right) \cong 0$$

Squaring both sides of Eq. (19) and taking the expectation, we get the variance of proposed estimators up to first order of approximation as:

$$\begin{aligned}V\left(\overline{y}_P^{(u)}\right) \cong & \overline{Y}^2\varphi C_y^2 - \overline{Y}^2\varphi\lambda A_1k_2 - 2\overline{XY}\varphi C_{yx}k_3 - 2\overline{YR}\varphi C_{yr}k_4\\ & + \overline{Y}^2\lambda^2\varphi A_2k_2^2 + \overline{X}^2\varphi C_x^2A_3k_3^2 + \overline{R}^2\varphi C_r^2A_4k_4^2\\ & + \overline{XY}\varphi\lambda C_xA_5k_2k_3 + \overline{RY}\varphi\lambda A_6k_2k_4 + 2\overline{RX}\varphi C_{xr}A_3k_3k_4\end{aligned} \tag{20}$$

where

$$A_1 = \frac{3}{4}\varphi\lambda C_y^2C_x^2 + C_{yx} - \frac{\varphi C_{yx}C_y\theta_{21x}}{\rho_{yx}}$$

$$A_2 = \frac{1}{4}C_x^2 - \frac{9}{64}\varphi\lambda^2C_x^4 - \frac{1}{4}\varphi C_{yx}^2 - \frac{\varphi C_{yx}C_x\theta_{12x}}{2\rho_{yx}} + \frac{3}{4}\varphi\lambda C_x^2C_{yx}$$

$$A_3 = 1 - \frac{1}{4}\varphi\lambda^2C_x^2$$

$$A_4 = 1 - \frac{1}{4}\varphi\lambda^2C_x^2\rho_{xr}^2$$

$$A_5 = C_x + \frac{5}{4}\varphi\lambda C_xC_{yx} - \frac{\varphi C_{yx}\theta_{12x}}{\rho_{yx}} - \frac{3}{8}\varphi\lambda^2C_x^3$$

$$A_6 = \frac{3}{4}\varphi\lambda C_x^2C_{yr} + C_{xr} - \frac{\varphi C_{yx}C_r\theta_{12r}}{\rho_{yr}} + \frac{1}{2}\varphi\lambda C_{xr}C_{yx} - \frac{3}{8}\varphi\lambda^2C_x^2C_{xr}$$

Partially differentiating Eq. (20) with respect to k_2, k_3 and k_4 and equating them to zero, we get the optimal values of k_2, k_3 and k_4 as follows.

$$k_{2(opt)} = \frac{2C_1}{\lambda C_2}$$

$$k_{3(opt)} = \frac{\overline{Y}[A_1B_2C_2 - 4A_2B_2C_1 - A_6(B_3C_2 - B_6C_1)]}{\overline{X}C_xA_5B_2C_2}$$

$$k_{4(opt)} = \frac{\overline{Y}[B_3C_2 - B_6C_1]}{\overline{R}B_2C_2}$$

Placing these optimal values in Eq. (20), we obtained the minimum variance as given by

$$\begin{aligned}V_{min}\left(\overline{y}_P^{(u)}\right) \cong & \frac{1}{C_xA_5^2B_2^2C_2^2}\overline{Y}^2\varphi\Big[C_xA_5^2B_2^2\left\{C_y^2C_2^2 - 2A_1C_1C_2 + 4A_2C_1^2\right\}\\ & + D_1\left\{C_xA_3D_1 - 2C_{yx}A_5B_2C_2 + 2C_xA_5^2B_2C_1\right\}\\ & + C_xA_5^2D_2\left\{C_r^2A_4D_2 - 2C_{yr}B_2C_2 + 2A_6B_2C_1\right\} + 2C_{xr}A_3A_5D_1D_2\Big]\end{aligned} \tag{21}$$

where

$$B_1 = C_xA_1A_3 - C_{yx}A_5, B_2 = C_{xr}A_3A_6 - C_r^2C_xA_4A_5, B_3 = C_{xr}A_1A_3 - C_{yr}C_xA_5$$

$$B_4 = C_xA_3A_6 - C_{xr}A_3A_5, B_5 = C_x\left(4A_2A_3 - A_5^2\right), B_6 = 4C_{xr}A_2A_3 - C_xA_5A_6$$

$$C_1 = B_1B_2 - B_3B_4, C_2 = B_2B_5 - B_4B_6$$

$$D_1 = A_1B_2C_2 - 4A_2B_2C_1 - A_6(B_3C_2 - B_6C_1), D_2 = (B_3C_2 - B_6C_1)$$

4. Results and discussion

In this section, we evaluated the performance of proposed class of estimators as compared to other unbiased/almost unbiased estimators. For this purpose, we selected five real life data sets with different correlation coefficients (first three with positive and last two with negative) between study variable and auxiliary variable. The descriptions of the populations are given below.

Population 1: [Source: Singh and Mangat (1996), p. 369]

$y =$ Number of tube wells

$x =$ Net irrigated area(in hectares) for 69 villages of Doraha development block of

Punjab, India

$N = 69, n = 10, \overline{Y} = 135.2609, \overline{X} = 345.7536, \overline{R} = 35,$

$C_y = 0.8422, C_x = 0.8422, C_r = 0.5732, \rho_{yx} = 0.9224, \rho_{yr} = 0.7136,$

$\rho_{xr} = 0.8185, \beta_{2(x)} = 7.2159, \beta_{1(x)} = 2.3808$

Population 2: [Source: Cochran (1977), p.152]

$y =$ Population size in 1930

$x =$ Population size in 1920

$N = 49, n = 12, \overline{Y} = 127.7959, \overline{X} = 103.1429, \overline{R} = 25,$

$C_y = 0.9634, C_x = 1.0122, C_r = 0.5715, \rho_{yx} = 0.9817, \rho_{yr} = 0.7207,$

$\rho_{xr} = 0.7915, \beta_{2(x)} = 5.1412, \beta_{1(x)} = 2.2553$

Population 3: [Source: Singh and Mangat (1996), p. 369]

$y =$ Number of tube wells

$x =$ Number of tractors for 69 villages of Doraha development block of Punjab, India

$N = 69, n = 10, \overline{Y} = 135.2609, \overline{X} = 21.2319, \overline{R} = 35,$

$C_y = 0.8422, C_x = 0.7969, C_r = 0.5726, \rho_{yx} = 0.9119, \rho_{yr} = 0.7364,$

$\rho_{xr} = 0.8616, \beta_{2(x)} = 3.7653, \beta_{1(x)} = 1.8551$

Population 4: [Source: Gujarati (2004), p. 433]

$y =$ Average miles per gallons

$x =$ Top Speed(miles per hour) of 81 cars

$N = 81, n = 16, \overline{Y} = 33.8346, \overline{X} = 112.4568, \overline{R} = 41,$

$C_y = 0.2972, C_x = 0.1256, C_r = 0.5728, \rho_{yx} = -0.6908,$
$\rho_{yr} = -0.7298,$

$\rho_{xr} = 0.8456, \beta_{2(x)} = 4.1454, \beta_{1(x)} = 1.9016$

Population 5: [Source: Gujarati (2004), p. 433]

$y =$ Average miles per gallons

$x =$ Cubic feet of cab space of 81 cars

$N = 81, n = 18, \overline{Y} = 33.8346, \overline{X} = 98.7654, \overline{R} = 41,$

$C_y = 0.2972, C_x = 0.2258, C_r = 0.5727, \rho_{yx} = -0.3683,$
$\rho_{yr} = -0.4732,$

$\rho_{xr} = 0.9245, \beta_{2(x)} = 0.9202, \beta_{1(x)} = -0.5902$

We calculated the variances of all the estimators i.e. $\overline{y}_0^{(u)}, \overline{y}_{HR}^{(u)}, \overline{y}_{S1}^{(u)}, \overline{y}_{S2}^{(u)}, \overline{y}_{CK1}^{(u)}, \overline{y}_{CK2}^{(u)}$ and $\overline{y}_P^{(u)}$for the populations 1–5. Expressions for the variances of all the existing and proposed estimators are given in section 1 & section 3 in detail. All empirical results are summarized in Tables 2-6.

In case of positive correlation between study variable and auxiliary variable (populations 1–3), some important observations are made from Tables 2-4 as follows:

- $\overline{y}_{HR}^{(u)}$ performs better than $\overline{y}_0^{(u)}$.
- It is worth pointing out that $\overline{y}_{HR}^{(u)}$ has less variance than $\overline{y}_{S1}^{(u)}, \overline{y}_{S2}^{(u)}$ and $\overline{y}_{CK1}^{(u)}$.
- All the proposed estimators have minimum variance as compared to $\overline{y}_0^{(u)}, \overline{y}_{HR}^{(u)}, \overline{y}_{S1}^{(u)}, \overline{y}_{S2}^{(u)}, \overline{y}_{CK1}^{(u)}$ and $\overline{y}_{CK2}^{(u)}$.
- A deep insight of columns of $\overline{y}_P^{(u)}$ reveals that the value of $(\alpha, \beta) = \left(\beta_{2(x)}, C_x\right)$ provides the least variance among all proposed estimators.

In case of negative correlation between study variable and auxiliary variable (populations 4–5), following important considerations are made from Tables 5-6:

- It is perceived that $\overline{y}_0^{(u)}$ performs better than $\overline{y}_{HR}^{(u)}$.

Table 2
Minimum variance of different estimators for population 1.

Estimator	Variance	Classes of Estimators				
		α	β	$\overline{y}_{CK1}^{(u)}$	$\overline{y}_{CK2}^{(u)}$	$\overline{y}_P^{(u)}$
$\overline{y}_0^{(u)}$	1109.534	1	C_x	196.0392	152.9700	**146.4461**
$\overline{y}_{HR}^{(u)}$	173.4709	1	$\beta_{2(x)}$	184.9776	153.2623	**146.9481**
$\overline{y}_{S1}^{(u)}$	195.5583	$\beta_{2(x)}$	C_x	197.5358	152.9325	**146.3866**
$\overline{y}_{S2}^{(u)}$	183.1241	ρ_{yx}	C_x	195.8957	152.9736	**146.4518**
		$\beta_{2(x)}$	ρ_{yx}	197.5143	152.9330	**146.3874**
		ρ_{yx}	$\beta_{2(x)}$	184.0929	153.2871	**146.9945**
		$\beta_{2(x)}$	S_x	165.6300	153.9599	**149.1641**
		1	ρ_{yx}	195.8893	152.9738	**146.4521**
		C_x	$\beta_{2(x)}$	183.1241	153.3145	**147.0466**
		C_x	ρ_{yx}	195.5583	152.9821	**146.4655**
		1	$\beta_{1(x)}$	193.0592	153.0459	**146.5696**

*Bold values indicate minimum variances

- It is important to mention that $\overline{y}_0^{(u)}$ has less variance than $\overline{y}_{S1}^{(u)}, \overline{y}_{S2}^{(u)}$ and $\overline{y}_{CK1}^{(u)}$.
- All proposed estimators have minimum variance as compared to existing estimators.
- $(\alpha, \beta) = \left(\rho_{yx}, \beta_{2(x)}\right)$ is an appropriate choice in order to get the minimum variance among all the proposed estimators.

4.1. A simulation study

It is clearly observed from numerical findings that the proposed class provides almost unbiased and efficient estimators for estimating population mean in case of SRSWOR. In addition, this superiority is assessed through a Monte Carlo simulation study using R

Table 3
Minimum variance of different estimators for population 2.

Estimator	Variance	Classes of Estimators				
		α	β	$\overline{y}_{CK1}^{(u)}$	$\overline{y}_{CK2}^{(u)}$	$\overline{y}_{P}^{(u)}$
$\overline{y}_0^{(u)}$	953.8721	1	C_x	191.0767	37.6156	**16.6971**
$\overline{y}_{HR}^{(u)}$	39.0457	1	$\beta_{2(x)}$	60.3292	37.5857	**17.4239**
$\overline{y}_{S1}^{(u)}$	194.9164	$\beta_{2(x)}$	C_x	300.6648	37.5954	**16.5412**
$\overline{y}_{S2}^{(u)}$	60.9889	ρ_{yx}	C_x	189.4141	37.6160	**16.7006**
		$\beta_{2(x)}$	ρ_{yx}	301.8640	37.5952	**16.5423**
		ρ_{yx}	$\beta_{2(x)}$	59.3432	37.5826	**17.4398**
		$\beta_{2(x)}$	S_x	45.8849	36.1171	**19.4647**
		1	ρ_{yx}	193.8267	37.6150	**16.6914**
		C_x	$\beta_{2(x)}$	60.9889	37.5876	**17.4137**
		C_x	ρ_{yx}	194.9164	37.6148	**16.6891**
		1	$\beta_{1(x)}$	118.8250	37.6290	**16.9256**

*Bold values indicate minimum variances

Table 4
Minimum variance of different estimators for population 3.

Estimator	Variance	Classes of Estimators				
		α	β	$\overline{y}_{CK1}^{(u)}$	$\overline{y}_{CK2}^{(u)}$	$\overline{y}_{P}^{(u)}$
$\overline{y}_0^{(u)}$	1109.5340	1	C_x	192.3828	207.8877	**169.8932**
$\overline{y}_{HR}^{(u)}$	188.2657	1	$\beta_{2(x)}$	202.3949	208.5434	**171.9661**
$\overline{y}_{S1}^{(u)}$	188.7435	$\beta_{2(x)}$	C_x	205.5316	207.5081	**169.3616**
$\overline{y}_{S2}^{(u)}$	215.0774	ρ_{yx}	C_x	191.3648	207.9301	**169.9596**
		$\beta_{2(x)}$	ρ_{yx}	204.5481	207.5305	**169.3906**
		ρ_{yx}	$\beta_{2(x)}$	206.9720	208.5160	**172.1645**
		$\beta_{2(x)}$	S_x	211.8509	208.4709	**172.3533**
		1	ρ_{yx}	190.9106	207.9504	**169.9920**
		C_x	$\beta_{2(x)}$	215.0774	208.4339	**172.4686**
		C_x	ρ_{yx}	188.7435	208.0659	**170.1863**
		1	$\beta_{1(x)}$	187.1201	208.3344	**170.7391**

*Bold values indicate minimum variances

Table 5
Minimum variances of different estimators for population 4.

Estimator	Variance	Classes of Estimators				
		α	β	$\overline{y}_{CK1}^{(u)}$	$\overline{y}_{CK2}^{(u)}$	$\overline{y}_{P}^{(u)}$
$\overline{y}_0^{(u)}$	5.0712	1	C_x	9.1068	2.6919	**2.2790**
$\overline{y}_{HR}^{(u)}$	8.9364	1	$\beta_{2(x)}$	8.9279	2.6920	**2.2791**
$\overline{y}_{S1}^{(u)}$	9.3860	$\beta_{2(x)}$	C_x	9.1112	2.6919	**2.2790**
$\overline{y}_{S2}^{(u)}$	7.9898	ρ_{yx}	C_x	9.1211	2.6919	**2.2789**
		$\beta_{2(x)}$	ρ_{yx}	9.1204	2.6919	**2.2789**
		ρ_{yx}	$\beta_{2(x)}$	9.4125	2.6917	**2.2788**
		$\beta_{2(x)}$	S_x	8.9596	2.6920	**2.2790**
		1	ρ_{yx}	9.1451	2.6919	**2.2789**
		C_x	$\beta_{2(x)}$	7.9898	2.6925	**2.2794**
		C_x	ρ_{yx}	9.3859	2.6918	**2.2789**
		1	$\beta_{1(x)}$	9.0257	2.6919	**2.2790**

*Bold values indicate minimum variances

software. For this purpose, two real populations are used. Different sample sizes i.e. ($n = 180$ and 220) and ($n = 18$ and 20) are used for both real populations.

Following steps are performed to carry out the simulation study:

Step 1. Select a SRSWOR of size n from the population of size N.

Step 2. Use sample data from step 1 to find the variance/minimum variance of all the

existing and proposed estimators.

Step 3. Step 1 and step 2 are repeated 10,000 times.

Step 4. Obtain 10,000 values for variance of each estimator.

Step 5. Average of 10,000 values, obtained in step 4 is the variance of each estimator.

Remark 4.1. *The following expression is used for calculation of variance/minimum variance for all estimators considered in this study*:

$$Var(\bar{y}^*) = \frac{\sum_{i=1}^{10000}(\bar{y}^* - \bar{Y})^2}{10000} \text{where} \bar{y}^* = \bar{y}_0^{(u)}, \bar{y}_{HR}^{(u)}, \bar{y}_{S1}^{(u)}, \bar{y}_{S2}^{(u)}, \bar{y}_{CK1}^{(u)}, \bar{y}_{CK2}^{(u)} \text{ and } \bar{y}_P^{(u)}.$$

4.1.1. Real population 1

We used a real data of primary and secondary schools for 923 districts of Turkey in 2007, taking number of teachers as study variable and number of students as auxiliary variable (Source: Koyuncu and Kadilar, 2009). Some important parameters of the data set are:

$$\bar{Y} = 436.4345, \bar{X} = 11440.5, \bar{R} = 462,$$

$$C_y = 1.7183, C_x = 1.8645, C_r = 0.5770,$$

Table 6
Minimum variances of different estimators for population 5.

Estimator	Variance	Classes of Estimators				
		α	β	$\bar{y}_{CK1}^{(u)}$	$\bar{y}_{CK2}^{(u)}$	$\bar{y}_P^{(u)}$
$\bar{y}_0^{(u)}$	4.3690	1	C_x	10.0399	3.7985	**3.3712**
$\bar{y}_{HR}^{(u)}$	9.3363	1	$\beta_{2(x)}$	9.9681	3.7984	**3.3715**
$\bar{y}_{S1}^{(u)}$	10.2410	$\beta_{2(x)}$	C_x	10.0379	3.7985	**3.3713**
$\bar{y}_{S2}^{(u)}$	9.6624	ρ_{yx}	C_x	10.1291	3.7985	**3.3709**
		$\beta_{2(x)}$	ρ_{yx}	10.1062	3.7986	**3.3710**
		ρ_{yx}	$\beta_{2(x)}$	10.3397	3.7986	**3.3703**
		$\beta_{2(x)}$	S_x	8.2728	3.7976	**3.3778**
		1	ρ_{yx}	10.1028	3.7985	**3.3710**
		C_x	$\beta_{2(x)}$	9.6624	3.7984	**3.3725**
		C_x	ρ_{yx}	10.2410	3.7986	**3.3706**
		1	$\beta_{1(x)}$	10.1266	3.7985	**3.3709**

*Bold values indicate minimum variances

Table 7
Minimum variance of estimators based on simulation through real population 1.

Estimator	Variance	Classes of Estimators				
		α	β	$\bar{y}_{CK1}^{(u)}$	$\bar{y}_{CK2}^{(u)}$	$\bar{y}_P^{(u)}$
$n = 180$						
$\bar{y}_0^{(u)}$	2510.8720	1	C_x	505.7589	206.0152	**178.9120**
$\bar{y}_{HR}^{(u)}$	271.3697	1	$\beta_{2(x)}$	478.7737	208.0823	**180.6443**
$\bar{y}_{S1}^{(u)}$	506.8906	$\beta_{2(x)}$	C_x	505.8466	204.0754	**176.6485**
$\bar{y}_{S2}^{(u)}$	490.9651	ρ_{yx}	C_x	508.3117	207.6117	**180.2311**
		$\beta_{2(x)}$	ρ_{yx}	505.5180	207.5761	**180.9583**
		ρ_{yx}	$\beta_{2(x)}$	483.6694	202.0577	**175.2135**
		$\beta_{2(x)}$	S_x	229.4068	206.3638	**179.6351**
		1	ρ_{yx}	504.0403	208.9553	**182.3116**
		C_x	$\beta_{2(x)}$	490.9651	200.0236	**173.7984**
		C_x	ρ_{yx}	506.8906	203.3060	**176.9591**
		1	$\beta_{1(x)}$	498.5304	203.1903	**176.5236**
$n = 220$						
$\bar{y}_0^{(u)}$	1938.0569	1	C_x	390.0107	161.6810	**144.0454**
$\bar{y}_{HR}^{(u)}$	209.1543	1	$\beta_{2(x)}$	372.9320	162.4285	**144.7053**
$\bar{y}_{S1}^{(u)}$	393.2290	$\beta_{2(x)}$	C_x	393.0053	158.8819	**141.2770**
$\bar{y}_{S2}^{(u)}$	385.2977	ρ_{yx}	C_x	390.6282	160.2888	**142.3775**
		$\beta_{2(x)}$	ρ_{yx}	393.9777	162.1509	**144.2795**
		ρ_{yx}	$\beta_{2(x)}$	373.2435	163.9430	**146.4945**
		$\beta_{2(x)}$	S_x	178.2684	159.8451	**142.1195**
		1	ρ_{yx}	392.5001	163.9137	**146.3558**
		C_x	$\beta_{2(x)}$	385.2977	158.3126	**140.8176**
		C_x	ρ_{yx}	393.2290	160.8272	**143.2344**
		1	$\beta_{1(x)}$	388.8904	158.9783	**141.3371**

*Bold values indicate minimum variances

Table 8
Minimum variance of estimators based on simulation through real population 2.

Estimator	Variance	Classes of Estimators α	β	$\bar{y}^{(u)}_{CK1}$	$\bar{y}^{(u)}_{CK2}$	$\bar{y}^{(u)}_{P}$
$n = 18$						
$\bar{y}^{(u)}_0$	4.3528	1	C_x	7.8021	2.0194	**1.5434**
$\bar{y}^{(u)}_{HR}$	7.7071	1	$\beta_{2(x)}$	7.6910	2.0369	**1.5463**
$\bar{y}^{(u)}_{S1}$	8.0659	$\beta_{2(x)}$	C_x	7.8296	2.0300	**1.5328**
$\bar{y}^{(u)}_{S2}$	6.8708	ρ_{yx}	C_x	7.8169	2.0165	**1.5264**
		$\beta_{2(x)}$	ρ_{yx}	7.8105	2.0501	**1.5596**
		ρ_{yx}	$\beta_{2(x)}$	8.1228	2.0344	**1.5473**
		$\beta_{2(x)}$	S_x	7.6585	2.0146	**1.5305**
		1	ρ_{yx}	7.8603	2.0412	**1.5490**
		C_x	$\beta_{2(x)}$	6.8708	2.0101	**1.5387**
		C_x	ρ_{yx}	8.0659	2.0610	**1.5616**
		1	$\beta_{1(x)}$	7.7222	2.0231	**1.5402**
$n = 20$						
$\bar{y}^{(u)}_0$	3.8208	1	C_x	6.8307	1.8090	**1.3954**
$\bar{y}^{(u)}_{HR}$	6.7615	1	$\beta_{2(x)}$	6.6893	1.7862	**1.3836**
$\bar{y}^{(u)}_{S1}$	7.0436	$\beta_{2(x)}$	C_x	6.8060	1.8016	**1.4004**
$\bar{y}^{(u)}_{S2}$	5.9986	ρ_{yx}	C_x	6.8150	1.8009	**1.3975**
		$\beta_{2(x)}$	ρ_{yx}	6.8163	1.7871	**1.3887**
		ρ_{yx}	$\beta_{2(x)}$	7.0883	1.8057	**1.3974**
		$\beta_{2(x)}$	S_x	6.6915	1.7788	**1.3871**
		1	ρ_{yx}	6.8542	1.7990	**1.3909**
		C_x	$\beta_{2(x)}$	5.9986	1.8138	**1.4053**
		C_x	ρ_{yx}	7.0436	1.8051	**1.3953**
		1	$\beta_{1(x)}$	6.7303	1.8095	**1.3998**

*Bold values indicate minimum variances

$$\rho_{yx} = 0.9543, \rho_{yr} = 0.6444, \rho_{xr} = 0.6307,$$

$$\beta_{1(x)} = 3.9365, \beta_{2(x)} = 18.7208$$

4.1.2. Real population 2

This real data relates to 81 cars in which average miles per gallons (MPG) is taken as a study variable and top speed, miles per hour (SP) as an auxiliary variable. (Source: Gujarati (2004), p. 433). Some important parameters of the data set are:

$$\bar{Y} = 33.8346, \bar{X} = 112.4568, \bar{R} = 41,$$

$$C_y = 0.2972, C_x = 0.1256, C_r = 0.5728,$$

$$\rho_{yx} = -0.6908, \rho_{yr} = -0.7298, \rho_{xr} = 0.8456,$$

$$\beta_{1(x)} = 1.9016, \beta_{2(x)} = 4.1454$$

Variances calculated for different sample sizes through real populations 1–2 are reported in Tables 7-8. Simulation study, alike in applications to real data reveals that

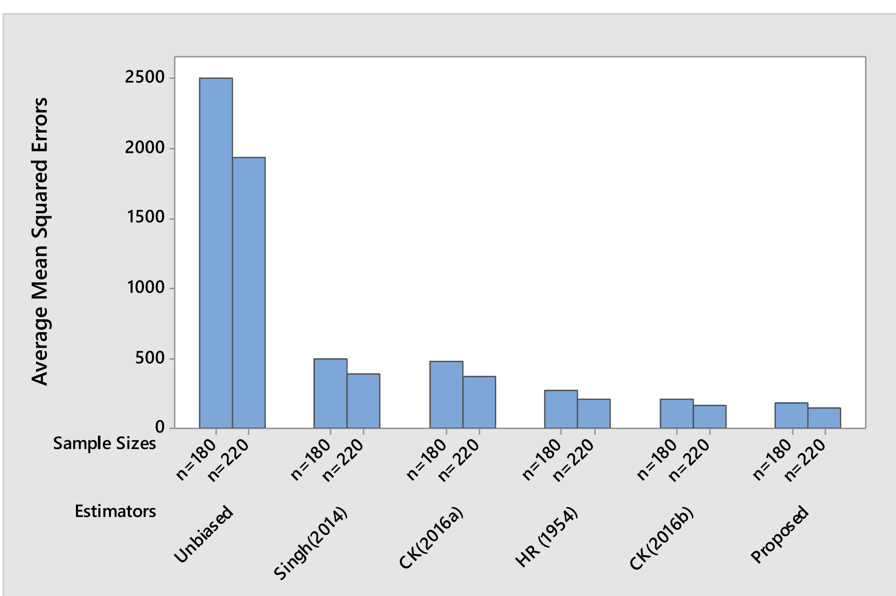

Fig. 1. Minimum variance of estimators based on simulation through real population 1.

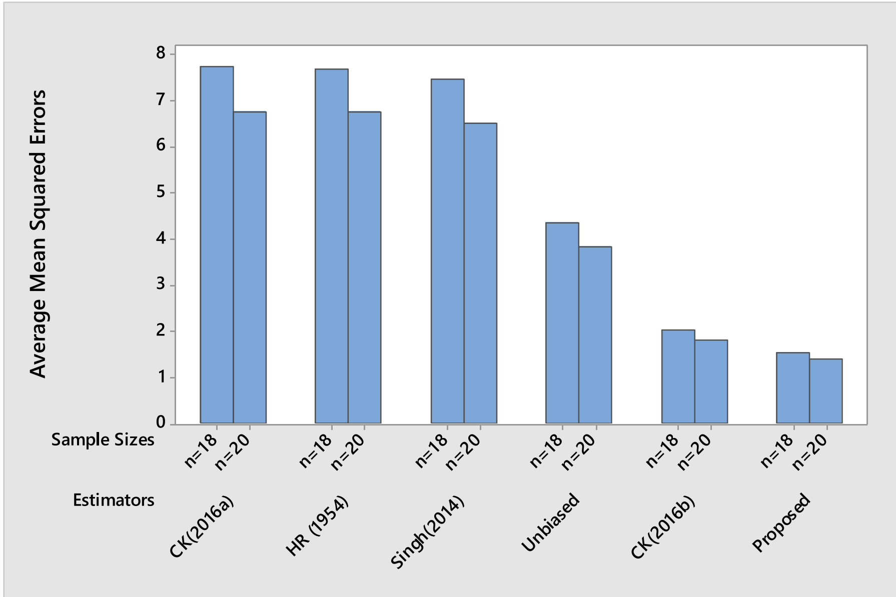

Fig. 2. Minimum variance of estimators based on simulation through real population 2.

- $\bar{y}_{HR}^{(u)}$ is more efficient than $\bar{y}_0^{(u)}$ in case of positive correlation between study variable and auxiliary variable (see Table 7) but less efficient in case of negative correlation (see Table 8).
- By increasing the sample size, variance of all the estimators reduces.
- Proposed estimators $\bar{y}_P^{(u)}$ have minimum variance as compared to all other estimators.

The performance of the proposed estimators $\bar{y}_P^{(u)}$ as compared to $\bar{y}_0^{(u)}, \bar{y}_{HR}^{(u)}, \bar{y}_{S1}^{(u)}, \bar{y}_{S2}^{(u)}, \bar{y}_{CK1}^{(u)}$ and $\bar{y}_{CK2}^{(u)}$ are also shown graphically for both populations considered in simulation study. Figs. 1-2 comprise the average of mean squared errors of the estimators based on different sample sizes. From Figs. 1 & 2, it can be seen that: 1) By increasing the sample size, variance of all the estimators reduces. 2) Proposed estimators $\bar{y}_P^{(u)}$ have minimum variance as compared to all other estimators under study.

5. Conclusion

We proposed a new class of almost unbiased estimators for estimating population mean under SRSWOR. This class is developed through the Hartley-Ross type estimation using the information of auxiliary variable and the ranks of auxiliary variable. Minimum variance of proposed class is derived up to first degree of approximation. Five real life data sets are used to check the numerical performance of new estimators. A comparison of new class is made with existing unbiased/almost unbiased estimators. A simulation study through two real data sets is also conducted to assess the potential of suggested class. On the basis of numerical findings, it is concluded that new class can generate optimum almost unbiased estimators. Therefore, use of proposed class is recommended for future applications.

The possible extensions of this work are to estimate the: 1) finite population mean under other sampling designs like stratified random sampling, double sampling, rank set sampling etc. 2) other unknown finite population parameters including median, variance and proportions etc. 3) population mean in the presence of non-sampling errors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Biradar, R.S., Singh, H.P., 1992a. A note on almost unbiased ratio-cum-product estimator. Metron 40 (1–2), 249–255.
Biradar, R.S., Singh, H.P., 1992b. On a class of almost unbiased ratio estimators. Biomed. J. 34 (8), 937–944.
Biradar, R.S., Singh, H.P., 1995. A class of unbiased ratio estimators. J. Indian Soc. Agric. Stat. 47 (3), 230–239.
Cekim, H.O., Kadilar, C., 2016. New unbiased estimators with the help of Hartley-Ross type estimators. Pakistan J. Stat. 32 (4), 247–260.
Cochran, W.G., 1977. Sampling Techniques. John Wiley and Sons, New York.
Gujarati, D.N., 2004. Basic Econometrics. The McGraw-Hill Companies, New York.
Gupta, S., Shabbir, J., 2008. On improvement in estimating the population mean in simple random sampling. J. Appl. Stat. 35 (5), 559–566.
Haq, A., Khan, M., Hussain, Z., 2017. A new estimator of finite population mean based on the dual use of the auxiliary in formation. Communications in Statistics- Theory and Methods 46 (9), 4425–4436.
Haq, A., Shabbir, J., 2013. Improved family of ratio estimators in simple and stratified random sampling. Communications in Statistics- Theory and Methods 42 (5), 782–799.
Hartley, H.O., Ross, A., 1954. Unbiased ratio estimators. Nature 174, 270–272.
Irfan, M., Javed, M., Lin, Z., 2019a. Improved estimation of population mean through known conventional and non-conventional measures of auxiliary variable. Iran. J. Sci. Technol., Trans. A: Sci. 43 (4), 1851–1862.
Irfan, M., Javed, M., Lin, Z., 2019b. Enhanced estimation of population mean in the presence of auxiliary information. J. King Saud Univ.- Sci. 31 (4), 1373–1378.
Irfan, M., Javed, M., & Bhatti, S., H. (2020). Difference-type-exponential estimators based on dual auxiliary information under simple random sampling. Scientia Iranica: Transactions on Industrial Engineering (E), Accepted.
Javed, M., Irfan, M., 2020. A simulation study: new optimal estimators for population mean by using dual auxiliary information in stratified random sampling. J. Taibah Univ. Sci. 14 (1), 557–568.
Javed, M., Irfan, M., Pang, T., 2019. Hartley-Ross type unbiased estimators of population mean using two auxiliary variables. Scientia Iranica: Trans. Ind. Eng. (E) 26 (6), 3835–3845.
Kadilar, C., Cingi, H., 2006a. An improvement in estimating the population mean by using the correlation coefficient. Hacettepe J. Math. Stat. 35 (1), 103–109.

Kadilar, C., Cingi, H., 2006b. Improvement in estimating the population mean in simple random sampling. Appl. Math. Lett. 19, 75–79.

Kadilar, C., Cingi, H., 2006c. A new ratio estimator using correlation coefficient. Inter-Stat, 1–11.

Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N., Smarandache, F., 2007. A general family of estimators for estimating population mean using known value of some population parameter(s). Far East Journal of Theoretical Statistics 22, 181–191.

Koyuncu, N., Kadilar, C., 2009. Efficient estimators for the population mean. Hacettepe Journal of Mathematics and Statistics 38 (2), 217–225.

Murthy, M.N., Nanjamma, N.S., 1959. Almost unbiased estimator based on interpenetrating sub-sample estimates. Sankhya 21, 381–392.

Raza, M. A., Nawaz, T., & Aslam, M. (2020). On designing CUSUM charts using ratio-type estimators for monitoring the location of normal processes. Scientia Iranica: Transactions on Industrial Engineering (E), Accepted.

Robson, D.S., 1957. Application of multivariate polykays to the theory of unbiased ratio type estimation. J. Am. Stat. Assoc. 50, 1225–1226.

Sahoo, J., Sahoo, L.N., Mohanty, S., 1994. An alternative approach to estimation in two phase sampling using two auxiliary variables. Biometrical Journal 36, 293–298.

Shabbir, J., Gupta, S., 2011. On estimating finite population mean in simple and stratified random sampling. Communications in Statistics- Theory and Methods 40 (2), 199–212.

Singh, R., Mangat, N.S., 1996. Elements of survey sampling. Kluwer Academic Publishers, Norwell, MA.

Singh, H.P., Sharma, B., Tailor, R., 2014. Hartley-Ross type estimators for population mean using known parameters of auxiliary variate. Communications in Statistics- Theory and Methods 43, 547–565.

Singh, H.P., Tailor, R., 2003. Use of known correlation coefficient in estimating the finite population means. Stat. Transition 6 (4), 555–560.

Singh, H.P., Solanki, R.S., 2013. An efficient class of estimators for the population mean using auxiliary information. Comm. Stat.- Theory Meth. 42, 145–163.

Upadhyaya, L.N., Singh, H.P., 1999. Use of transformed auxiliary variable in estimating the finite population mean. Biometrical J. 41 (5), 627–636.