An Improved Class of Estimators of Population Mean under Simple Random Sampling

Shashi Bhushan, Anoop Kumar,* and Saurabh Singh
Department of Mathematics & Statistics,
Dr. Shakuntala Misra National Rehabilitation University,
Lucknow, India

Sumit Kumar
School of Applied Science, K.K. University,
Biharsharif, Bihar, India

In this article, we consider an improved class of estimators of population mean using additional information under simple random sampling (SRS). The expressions of bias and mean square error of the proposed class of estimators are obtained up to first order of approximation. In addition, some well-known estimators have been identified as particular member of the proposed class of estimators. The theoretical results are established and empirical study has been carried out using real and simulated data sets. The findings appear to be rather satisfactory showing better improvement over the existing estimators.

Keywords: simple random sampling, mean square error, efficiency

1. Introduction

In sample surveys, usually the information on auxiliary variable such as standard deviation S_x, correlation coefficient ρ_{xy}, coefficient of skewness $\beta_1(x)$, coefficient of kurtosis $\beta_2(x)$ are known in advance. The use of these information helps to improve the efficiency of the suggested estimators. However, over recent years, a large number of ratio, product, regression and exponential type estimators based on different auxiliary information have been proposed by various authors like, Cochran (1977), Murthy (1967), Srivastava (1967), Walsh (1970), Sisodia and Dwivedi (1981), Upadhyaya et al. (1985), Pandey and Dubey (1988), Prasad (1989), Upadhyaya and Singh (1999), Singh and Tailor (2003), Singh (2003), Khoshnevisan et al. (2007), Koyuncu and Kadilar (2009) and more recently Bhushan et al. (2020 a, b, c) and Bhushan and Kumar (2020a, b). In this article,

* *corresponding author email: anoop.asy@gmail.com*

we suggest an improved class of estimators of population mean that contains a wide range of estimators existing till date.

Consider a finite population $U = (U_1, U_2, \ldots, U_N)$ consisting of N identifiable units. Let a sample of size n be selected from a finite population of size N using simple random sampling without replacement (SRSWOR). Let y_i and x_i denote the values of the study and auxiliary variables for the i^{th} ($i = 1, 2, \ldots, N$) unit of the population. Let $\bar{y} = 1/n\sum_{i=1}^{n} y_i$ and $\bar{Y} = 1/N\sum_{i=1}^{N} y_i$ be respectively the sample and population means of the study variable $y; \bar{x} = 1/n\sum_{i=1}^{n} x_i$ and $\bar{X} = 1/N\sum_{i=1}^{N} x_i$ be respectively the sample and population means of auxiliary variable $x; s_y = \sqrt{(n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}$ and $S_y = \sqrt{(N-1)^{-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2}$ be respectively the sample and population standard deviations of study variable $y; s_x = \sqrt{(n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ and $s_x = \sqrt{(N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})^2}$ be respectively the sample and population standard deviations of auxiliary variable $x; C_y = S_y / \bar{Y}$ and $C_x = S_x / \bar{X}$ be respectively the population coefficient of variations of study variable and auxiliary variable. Also, let ρ_{xy} be the population coefficient of correlation between auxiliary variable and study variable.

To derive the properties of the proposed class of estimators, let us assume that $\bar{y} = \bar{Y}(1+e_0)$, $\bar{x} = \bar{X}(1+e_1)$, such that $E(e_0) = E(e_1) = 0$ and

$$E(e_0^2) = fC_y^2$$
$$E(e_1^2) = fC_x^2$$
$$E(e_0 e_1) = f\rho_{xy}C_y C_x$$

where $f = (N-n)/Nn$.

The classical mean, ratio, product and regression estimators under SRS for the estimation of population mean $\bar{Y}$ of study variable y is given as

$$t_m = \bar{y} \tag{1.1}$$

$$t_{rat} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{1.2}$$

$$t_{pro} = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right) \tag{1.3}$$

$$t_{lr} = \bar{y} + \beta(\bar{X} - \bar{x}) \tag{1.4}$$

where β is the regression coefficient of y on x.

The mean square error (MSE) of the above estimators are given as

$$MSE(t_m) = f\bar{Y}^2 C_y^2 \tag{1.5}$$

$$MSE(t_{rat}) = f\bar{Y}^2 (C_y^2 + C_x^2 - 2\rho_{xy} C_x C_y) \tag{1.6}$$

$$MSE(t_{pro}) = f\bar{Y}^2 (C_y^2 + C_x^2 + 2\rho_{xy} C_x C_y) \tag{1.7}$$

$$MSE(t_{lr}) = f\bar{Y}^2 (C_y^2 + \beta^2 C_x^2 - 2\beta\rho_{xy} C_x C_y) \tag{1.8}$$

The minimum MSE of t_{lr} at optimum value of $\beta = \rho_{xy}\, S_y / S_x$ is given as

$$minMSE(t_{lr}) = f\bar{Y}^2 C_y^2 (1 - \rho_{xy}^2) \tag{1.9}$$

In this paper, we suggest a class of estimators for the estimation of population mean $\bar{Y}$ of study variable y with its properties which include a wide range of estimators namely, the usual mean, ratio, product, regression, Srivastava (1967), Walsh (1970), Sisodia and Dwivedi (1981), Upadhyaya et al. (1985), Pandey and Dubey (1988), Prasad (1989), Singh and Kakran (1993), Upadhyaya and Singh (1999), Singh (2003), Khoshnevisan et al. (2007), Koyuncu and Kadilar (2009) and Bhushan and Gupta (2014) estimators. The rest of the paper is organized as follows: Section 2 considers the proposed class of estimators and its known members with their properties. In Section 3, a numerical study over three real populations and a simulation study over two hypothetically generated populations are performed to support the theoretical results. The conclusion is drawn in Section 4.

2. Proposed Class of Estimators

The logarithmic function has some useful properties and play prominent role in different fields of science and non-science. We suggest an improved class of estimators by developing the logarithmic relationship between study variable y and auxiliary variable x as

$$t_b = \left[w_1\bar{y} + w_2\bar{y}\left(\frac{\bar{X}^*}{\theta\bar{x}^* + (1-\theta)\bar{X}^*} \right)^g \right]\left[1 + log\left(\frac{\bar{x}^*}{\bar{X}^*} \right) \right]^\eta \tag{2.1}$$

where θ and g are scalars that assume real values to design different estimators whereas w_1, w_2, and η are suitably chosen scalars. Also, $\bar{x}^* = a\bar{x} + b$ and $\bar{X}^* = a\bar{X} + b$ such that a and b are either real values or function of known parameters of auxiliary variable x namely, standard deviation S_x, coefficient of variation C_x, coefficient of skewness $\beta_1(x)$, coefficient of kurtosis $\beta_2(x)$ etc.

Using the notations defined in earlier section, we express the proposed class of estimators t_b in terms of e's as

$$t_b = \left[w_1\bar{Y}(1+e_0) + w_2\bar{Y}(1+e_0)\left(\frac{a\bar{X}+b}{\theta\{a\bar{X}(1+e_1)+b\}+(1-\theta)(a\bar{X}+b)} \right)^g \right]\left[1+log\left\{ \frac{a\bar{X}(1+e_1)+b}{a\bar{X}+b} \right\} \right]^\eta$$

$$t_b = [w_1\bar{Y}(1+e_0) + w_2\bar{Y}(1+e_0)(1+\theta v e_1)^{-g}][1+\log(1+ve_1)]^\eta$$

where $v = a\bar{X}/(a\bar{X}+b)$.

$$t_b = \left[w_1\bar{Y}(1+e_0) + w_2\bar{Y}(1+e_0)(1+\theta v e_1)^{-g} \right]\left[1+ve_1 - \frac{v^2e_1^2}{2} \right]^\eta$$

$$t_b = \left[w_1\bar{Y}(1+e_0) + w_2\bar{Y}(1+e_0)(1+\theta v e_1)^{-g} \right]\left[1+\eta ve_1 - \eta v^2 e_1^2 + \frac{\eta^2 v^2}{2} e_1^2 \right]$$

$$t_b - \bar{Y} = \bar{Y}\left[\begin{array}{c} w_1 + w_2 - 1 + w_1 e_0 + w_2 e_0 + w_1\eta v e_1 + w_2 \eta v e_1 - w_2 g\theta v e_1 \\ -w_2 g\theta v e_0 e_1 + w_2 \frac{g(g+1)}{2}\theta^2 v^2 e_1^2 + w_1 \eta v e_0 e_1 + w_2 \eta v e_0 e_1 \\ -w_2\eta g\theta v^2 e_1^2 - w_1\eta v^2 e_1^2 - w_2 \eta v^2 e_1^2 + w_1\frac{\eta^2}{2}v^2e_1^2 + w_2\frac{\eta^2}{2}v^2e_1^2 \end{array} \right] \qquad (2.2)$$

It is to be noted that the above expression is up to first order of approximation.

Taking expectation both the side of (2.2), we get the bias of the proposed class of estimators up to first order of approximation as

$$E(t_b - \bar{Y}) = \bar{Y}\left[\begin{array}{c} w_1 + w_2 - 1 + w_1 E(e_0) + w_2 E(e_0) + w_1\eta v E(e_1) + w_2 \eta v E(e_1) - w_2 g\theta v E(e_1) \\ -w_2 g\theta v E(e_0 e_1) + w_2 \frac{g(g+1)}{2}\theta^2 v^2 E(e_1^2) + w_1 \eta v E(e_0 e_1) + w_2 \eta v E(e_0 e_1) \\ -w_2\eta g\theta v^2 E(e_1^2) - w_1\eta v^2 E(e_1^2) - w_2 \eta v^2 E(e_1^2) + w_1\frac{\eta^2}{2}v^2E(e_1^2) + w_2\frac{\eta^2}{2}v^2E(e_1^2) \end{array} \right]$$

Putting the results discussed in earlier section, we get

$$Bias(t_b) = \bar{Y}[w_1 E_4 + w_2 E_5 - 1] \qquad (2.3)$$

Again, squaring both side of (2.2) and taking expectation, we obtain the *MSE* of the proposed class of estimators up to first order of approximation as

$$E(t_b-\bar{Y})^2=\bar{Y}^2\Big[1+w_1^2\{1+\gamma C_y^2+(2\eta^2v^2-2\eta v^2)\gamma C_x^2+4\eta v\gamma\rho_{xy}C_yC_x\}$$
$$+w_2^2\{1+\gamma C_y^2+\{2\eta^2v^2-2\eta v^2+g^2\theta^2v^2-4\eta g\theta v^2+g(g+1)\theta^2v^2\}\gamma C_x^2$$
$$+4v(\eta-g\theta)\gamma\rho_{xy}C_yC_x+2w_1w_2\{1+\gamma C_y^2+\left\{2\eta^2v^2-2\eta g\theta v^2-2\eta v^2+\frac{g(g+1)}{2}\theta^2v^2\right\}\gamma C_x^2$$
$$+2v(2\eta-g\theta)\gamma\rho_{xy}C_yC_x-2w_1\left\{1+\left\{\frac{\eta^2}{2}v^2-\eta v^2\right\}\gamma C_x^2+\eta v\gamma\rho_{xy}C_yC_x\right\}$$
$$-2w_2\left\{1+\left\{\frac{\eta^2}{2}v^2-\eta v^2-\eta g\theta v^2+\frac{g(g+1)}{2}\theta^2v^2\right\}\gamma C_x^2+(\eta v-g\theta v)\gamma\rho_{xy}C_y\,C_x\right\}\Big]$$

$$MSE(t_b)=\bar{Y}^2\left[1+w_1^2E_1+w_2^2E_2+2w_1w_2E_3-2w_1E_4-2w_2E_5\right] \quad (2.4)$$

where

$$E_1=\left[1+\gamma C_y^2+(2\eta^2v^2-2\eta v^2)\gamma C_x^2+4\eta v\gamma\rho_{xy}C_yC_x\right]$$

$$E_2=\Big[1+\gamma C_y^2+\{2\eta^2v^2-2\eta v^2+g^2\theta^2v^2-4\eta g\theta v^2+g(g+1)\theta^2v^2\}\gamma C_x^2$$
$$+4v(\eta-g\theta)\gamma\rho_{xy}C_yC_x$$

$$E_3=\Big[1+\gamma C_y^2+\left\{2\eta^2v^2-2\eta g\theta v^2-2\eta v^2+\frac{g(g+1)}{2}\theta^2v^2\right\}\gamma C_x^2$$
$$+2v(2\eta-g\theta)\gamma\rho_{xy}C_yC_x\Big]$$

$$E_4=\left[1+\left\{\frac{\eta^2}{2}v^2-nv^2\right\}\gamma C_x^2+\eta v\gamma\rho_{xy}C_yC_x\right]$$

$$E_5=\Big[1+\left\{\frac{\eta^2}{2}v^2-nv^2-\eta g\theta v^2+\frac{g(g+1)}{2}\theta^2v^2\right\}\gamma C_x^2$$
$$+(nv-g\theta v)v\gamma\rho_{xy}C_yC_x\Big]$$

The optimum values of w_1 and w_2 can be obtained by minimizing (2.4) whereas the optimum values of η can be obtained from the log type estimators envisaged by Bhushan and Gupta (2014) as

$$w_{1(opt)}=\frac{(E_2E_4-E_3E_5)}{(E_1E_2-E_3^2)} \quad (2.5)$$

$$w_{2(opt)}=\frac{(E_1E_5-E_3E_4)}{(E_1E_2-E_3^2)} \tag{2.6}$$

$$\eta_{(opt)}=-\rho_{xy}\frac{C_y}{C_x} \tag{2.7}$$

Now, putting the optimum values of w_1 and w_2 in (2.4), we get

$$MSE(t_b)=\bar{Y}^2\left[1+\left\{\frac{(E_2E_4-E_3E_5)}{(E_1E_2-E_3^2)}\right\}^2E_1+\left\{\frac{(E_1E_5-E_3E_4)}{(E_1E_2-E_3^2)}\right\}^2E_2\right.$$

$$\left.+2\frac{(E_2E_4-E_3E_5)}{(E_1E_2-E_3^2)}\frac{(E_1E_5-E_3E_4)}{(E_1E_2-E_3^2)}E_3-2\frac{(E_2E_4-E_3E_5)}{(E_1E_2-E_3^2)}E_4-2\frac{(E_1E_5-E_3E_4)}{(E_1E_2-E_3^2)}E_5\right]$$

After simplifying the above expression, we get the minimum MSE up to first order of approximation as

$$minMSE=\bar{Y}^2\left[\frac{(E_1E_5^2+E_2E_4^2-2E_3E_4E_5)}{(E_1E_2-E_3^2)}\right] \tag{2.8}$$

Theorem 2.1. *To the first order of approximation*

$$MSE(t_b)>\bar{Y}^2\left[1-\frac{(E_1E_5^2+E_2E_4^2-2E_3E_4E_5)}{(E_1E_2-E_3^2)}\right] \tag{2.9}$$

with equality holding if

$$w_1=w_{1\,(opt)}$$

$$w_2=w_{2\,(opt)}$$

It is worth mentioning that for different values of (w_1, w_2, η, θ, g,a, b), the proposed class of estimators t_b reduces to some well-known estimators of population mean $\bar{Y}$ which are reported below in Table 1. In addition to these, many other known estimators can also be generated by just putting different values of (w_1, w_2, η, θ, g, a, b). Further, we would like to note that the properties of the estimators belonging to proposed class of estimators t_b can be easily obtained from the expressions (2.3), (2.4) and (2.8) by suitably chosen values of scalars. Moreover, Theorem 2.1 shows the dominance of the suggested class of estimators over the existing estimators.

Some unknown members of the proposed class of estimators t_b are also produced which are disclosed in Table 2 for ready reference.

Table 1. Some Known Members of the Proposed Class of Estimators t_b

Estimator	w_1	w_2	η	θ	g	a	b
$t_m = \bar{y}$ Usual mean estimator	1	0	0	-	-	-	-
$t_r = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)$ Usual ratio estimator	0	1	0	1	1	1	0
$t_p = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)$ Usual product estimator	0	1	0	1	-1	1	0
$t_s = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\beta}$ Srivastava (1967) estimator	0	1	0	1	β	1	0
$t_w = \bar{y}\left(\frac{\bar{X}}{\bar{X}+\theta(\bar{x}-\bar{X})}\right)$ Walsh (1970) estimator	0	1	0		1	1	0
$t_w = \bar{y}\left[\frac{\bar{X}+C_x}{(\bar{x}-C_x)}\right]$ Sisodia and Dwivedi (1981) estimator	0	1	0	1	1	1	C_x
$t_u = w_1\bar{y} + w_2\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{g}$ Upadhyaya et al. (1985) estimator	w_1	w_2	0	1	G	1	0
$t_{pd} = \bar{y}\left[\frac{\bar{x}+Cx}{\bar{X}+Cx}\right]$ Pandey and Dubey (1988) estimator	0	1	0	1	-1	1	C_x
$t_p = k\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)$ Prasad (1989) estimator	0		0	1	1	1	0

$t_{sk} = \bar{y}\left[\dfrac{\bar{X}+\beta_2(x)}{\bar{x}+\beta_2(x)}\right]$ Singh and Kakran (1993) estimator	0	1	0	1	1	1	$\beta_{2(x)}$
$t_{up_1} = \bar{y}\left[\dfrac{\bar{X}C_x+\beta_{2(x)}}{\bar{x}C_x+\beta_{2(x)}}\right]$ Upadhyaya and Singh (1999) estimator	0	1	0	1	1	C_x	$\beta_{2(x)}$
$t_{up_2} = \bar{y}\left[\dfrac{\bar{X}\beta_{2(x)}+C_x}{\bar{x}\beta_{2(x)}+C_x}\right]$ Upadhyaya and Singh (1999) estimator	0	1	0	1	1	$\beta_{2(x)}$	
$t_{s_1} = \bar{y}\left[\dfrac{\bar{x}+\sigma_x}{\bar{X}+\sigma_x}\right]$ Singh (2003) estimator	0	1	0	1	-1	1	σ_x
$t_{s_2} = \bar{y}\left[\dfrac{\bar{x}\beta_{1(x)}+\sigma_x}{\bar{X}\beta_{1(x)}+\sigma_x}\right]$ Singh (2003) estimator	0	1	0	1	-1	$\beta_{(x)}$	σ_x
$t_{s_3} = \bar{y}\left[\dfrac{\bar{x}\beta_{2(x)}+\sigma_x}{\bar{X}\beta_{2(x)}+\sigma_x}\right]$ Singh (2003) estimator	0	1	0	1	-1	$\beta_{2(x)}$	σ_x
$t_{s_4} = \bar{y}\left[\dfrac{\bar{X}+\rho xy}{\bar{x}+\rho xy}\right]$ Singh (2003) estimator	0	1	0	1	1	1	ρxy
$t_{s_5} = \bar{y}\left[\dfrac{\bar{X}+\rho xy}{\bar{x}+\rho xy}\right]$ Singh (2003) estimator	0	1	0	1	1	1	ρxy
$t_k = \bar{y}\left(\dfrac{\bar{x}^*}{\theta\bar{x}^*+(1-\theta)\bar{X}^*}\right)^g$ Khoshnevisan et al. (2007) estimator	0	1	0		g	a	*b*
$t_{kk} = \lambda t_k$ Koyuncu and Kadilar (2009) estimator	0	λ	0		g	a	*b*
$t_{l_1} = \alpha_1\bar{y}\left[1+log\left(\dfrac{\bar{x}}{\bar{X}}\right)\right]^{\beta_1}$ Bhushan and Gupta (2014) estimator	α_1	0	β_1	-	-	1	0
$t_{l_2} = \alpha_2\bar{y}\left[1+log\left(\dfrac{\bar{x}^*}{\bar{X}^*}\right)\right]^{\beta_2}$ Bhushan and Gupta (2014) estimator	α_2	0	β_2	-	-	a	*b*

Table 2. Some Unknown Members of the Proposed Class of Estimators t_b

Estimators	*a*	*b*
$t_{b_1}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\bar{X}}{\theta\bar{x}+(1-\theta)\bar{X}}\right)^g\right]\left[1+log\left(\frac{\bar{x}}{\bar{X}}\right)\right]^\eta$	1	0
$t_{b_2}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\bar{X}+C_x}{\theta(\bar{x}+C_x)+(1-\theta)(\bar{X}+C_x)}\right)^g\right]\left[1+log\left(\frac{\bar{x}+C_x}{\bar{X}+C_x}\right)\right]^\eta$	1	C_x
$t_{b_3}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\beta_{2(x)}\bar{X}+C_x}{\theta(\beta_{2(x)}\bar{x}+C_x)+(1-\theta)(\beta_{2(x)}\bar{X}+C_x)}\right)^g\right]\left[1+log\left(\frac{\beta_{2(x)}\bar{x}+C_x}{\beta_{2(x)}\bar{X}+C_x}\right)\right]^\eta$	$\beta_{2(x)}$	C_x
$t_{b_4}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{C_x\bar{X}+\beta_{2(x)}}{\theta(C_x\bar{x}+\beta_{2(x)})+(1-\theta)(C_x\bar{X}+\beta_{2(x)})}\right)^g\right]\left[1+log\left(\frac{C_x\bar{x}+\beta_{2(x)}}{C_x\bar{X}+\beta_{2(x)}}\right)\right]^\eta$	C_x	$\beta_{2(x)}$
$t_{b_5}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\bar{X}+S_x}{\theta(\bar{x}+S_x)+(1-\theta)(\bar{X}+S_x)}\right)^g\right]\left[1+log\left(\frac{\bar{x}+S_x}{\bar{X}+S_x}\right)\right]^\eta$	1	S_x
$t_{b_6}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\beta_{1(x)}\bar{X}+S_x}{\theta(\beta_{1(x)}\bar{x}+S_x)+(1-\theta)(\beta_{1(x)}\bar{X}+S_x)}\right)^g\right]\left[1+log\left(\frac{\beta_{1(x)}\bar{x}+S_x}{\beta_{1(x)}\bar{X}+S_x}\right)\right]^\eta$	$B_{1(x)}$	S_x
$t_{b_7}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\beta_{2(x)}\bar{X}+S_x}{\theta(\beta_{2(x)}\bar{x}+S_x)+(1-\theta)(\beta_{2(x)}\bar{X}+S_x)}\right)^g\right]\left[1+log\left(\frac{\beta_{2(x)}\bar{x}+S_x}{\beta_{2(x)}\bar{X}+S_x}\right)\right]^\eta$	$\beta_{2(x)}$	S_x
$t_{b_8}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\bar{X}+\rho_{xy}}{\theta(\bar{x}+\rho_{xy})+(1-\theta)(\bar{X}+\rho_{xy})}\right)^g\right]\left[1+log\left(\frac{\bar{x}+\rho_{xy}}{\bar{X}+\rho_{xy}}\right)\right]^\eta$	*1*	ρ_{xy}
$t_{b_9}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\bar{X}+\beta_{2(x)}}{\theta(\bar{x}+\beta_{2(x)})+(1-\theta)(\bar{X}+\beta_{2(x)})}\right)^g\right]\left[1+log\left(\frac{\bar{x}+\beta_{2(x)}}{\bar{X}+\beta_{2(x)}}\right)\right]^\eta$	*1*	$\beta_{2(x)}$
$t_{b_{10}}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{C_x\bar{X}+\rho_{xy}}{\theta(C_x\bar{x}+\rho_{xy})+(1-\theta)(C_x\bar{X}+\rho_{xy})}\right)^g\right]\left[1+log\left(\frac{C_x\bar{x}+\rho_{xy}}{C_x\bar{X}+\rho_{xy}}\right)\right]^\eta$	C_x	ρ_{xy}
$t_{b_{11}}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\rho_{xy}\bar{X}+C_x}{\theta(\rho_{xy}\bar{x}+C_x)+(1-\theta)(\rho_{xy}\bar{X}+C_x)}\right)^g\right]\left[1+log\left(\frac{\rho_{xy}\bar{x}+C_x}{\rho_{xy}\bar{X}+C_x}\right)\right]^\eta$	ρ_{xy}	C_x
$t_{b_{12}}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\beta_{2(x)}\bar{X}+\rho_{xy}}{\theta(\beta_{2(x)}\bar{x}+\rho_{xy})+(1-\theta)(\beta_{2(x)}\bar{X}+\rho_{xy})}\right)^g\right]\left[1+log\left(\frac{\beta_{2(x)}\bar{x}+\rho_{xy}}{\beta_{2(x)}\bar{X}+\rho_{xy}}\right)\right]^\eta$	$\beta_{2(x)}$	ρ_{xy}
$t_{b_{13}}=\left[w_1\bar{y}+w_2\bar{y}\left(\frac{\rho_{xy}\bar{X}+\beta_{2(x)}}{\theta(\rho_{xy}\bar{x}+\beta_{2(x)})+(1-\theta)(\rho_{xy}\bar{X}+\beta_{2(x)})}\right)^g\right]\left[1+log\left(\frac{\rho_{xy}\bar{x}+\beta_{2(x)}}{\rho_{xy}\bar{X}+\beta_{2(x)}}\right)\right]^\eta$	ρ_{xy}	$\beta_{2(x)}$

3. Empirical Study

In order to enhance the theoretical results, we conducted an empirical study over three real populations and a simulation study over artificially generated two normal populations. It is important to note that we took $(\theta, g, a, b) = (1, 1, 1, 0)$ for empirical comparison of the proposed class of estimators with the existing estimator.

3.1. Real populations

The description of the populations is given below.

Populations 1: Pandey and Dubey (1988)

$N = 20$, $n = 8$, $\bar{Y} = 19.55$, $\bar{X} = 18.8$, $C_x^2 = 0.1555$, $C_y^2 = 0.1262$, $\rho_{xy} = 0.9199$, $\beta_{1(x)} = 0.5473$, $\beta_{2(x)} = 3.0613$

Populations 2: Cochran (1977) page 325

X = Number of rooms, Y = Number of persons

$N = 10$, $n = 4$, $\bar{Y} = 101.1$, $\bar{X} = 58.8$, $C_x = 0.1281$, $C_y = 0.1449$, $\rho_{xy} = 0.6515$,

$\beta_{1(x)} = 0.5764$, $\beta_{2(x)} = 0.3814$

Populations 3: Murthy (1967) page 228

X = Data on the number of workers, Y = Output for 80 factories in a region

$N = 80$, $n = 20$, $\bar{Y} = 51.8264$, $\bar{X} = 2.8513$, $C_x = 0.9484$, $C_y = 0.3542$, $\rho_{xy} = 0.915,0$

$\beta_{1(x)} = 0.6978$, $\beta_{2(x)} = 1.3005$

Using the above data sets, we have calculated the percent relative efficiency (PRE) of the various class of estimators T with respect to mean per unit estimator t_m with the help of following expression given below.

$$PRE = \frac{MSE(t_m)}{MSE(T)} \times 100 \tag{3.1}$$

The results are reported in Table 3 for each population that show the dominance of the proposed class of estimators over the existing estimators.

Table 3. ***PRE*** **of Different Estimators for Real Populations**

Estimators	Population 1	Population 2	Population 3
t_m	100	100	100
t_r	23.4261	124.1296	30.5859
t_p	521.9572	26.6426	7.6514
t^*	643.3392	173.7475	614.345
t_{sd}	23.9408	124.5408	73.5194
t_u	662.9184	173.7645	650.1155
t_{pd}	545.0977	26.7131	11.4756
t_p	23.6375	124.1558	30.9536
t_{sk}	27.2994	123.0393	98.3962
t_{up1}	32.7005	116.3779	104.1327
t_{up2}	23.5948	123.0436	60.8110
t_{s1}	586.118	30.1946	18.5068
t_{s2}	553.1894	26.9226	11.8894
t_{s3}	638.3463	26.4566	12.9089
t_{s4}	22.2064	125.9649	71.4541
t_{s5}	461.8686	26.9589	11.3396
t_{kk}	644.2710	173.9937	616.016
t_{li}, i = 1,2	717.3913	174.6481	1268.2310
$\boldsymbol{t_b}$	**718.3899**	**174.6510**	**1495.6000**

where $t^* = t_{lr}, t_s, t_w, t_k$

3.2 Simulation study

To evaluate the effectiveness of theoretical results, we perform a simulation study over artificially generated symmetric and asymmetric populations which are described below.

We generated a bivariate normal population of size N = 100 from bivariate normal distribution using R software with $\bar{X}$ = 5, $\bar{Y}$ = 10, σ_x = 10, σ_y = 15 and different values of correlation coefficient between x and y, i.e., ρ_{xy} = 0.6, 0.7, 0.8, 0.9.

We generated another bivariate normal population of size N = 500 from bivariate normal distribution using R software with $\bar{X}$ = 10, $\bar{Y}$ = 15, σ_x = 15 σ_y =

20 and different values of correlation coefficient between x and y, i.e., ρ_{xy} = 0.6, 0.7, 0.8, 0.9.

It is to be noted that we took various values of correlation coefficient to observe the behavior of the proposed class of estimators. Now, we have drawn a simple random sample without replacement of size n = 10 from population 1 and n = 50 from population 2. With 20000 iterations, PRE has been calculated using (3.1) and the outcomes of the simulation study for population 1 and 2 are summarized respectively in Table 4 and Table 5 which show the superiority of the proposed class of estimators t_b over the existing estimators. It has been also observed from simulation results that the PRE increases as the value of correlation coefficient increases.

Table 4. *PRE* of Different Estimators for First Artificial Population

ρ_{xy} **Estimators**	**0.6**	**0.7**	**0.8**	**0.9**
t_m	100	100	100	100
t_r	101.0235	128.4072	178.7238	298.819
t_p	29.4845	26.5863	23.8850	21.2049
t^*	143.5131	177.8600	249.2790	468.4963
t_{sd}	131.7374	170.9669	247.4537	465.7095
t_u	152.5079	179.3764	258.7186	314.2353
t_{pd}	39.6577	36.5709	33.6941	30.9200
t_p	104.3882	130.5718	179.2732	300.4003
t_{sk}	134.8123	174.2042	249.2236	454.2902
t_{up1}	123.8762	161.0592	234.4851	451.4049
t_{up2}	118.4886	153.4324	222.4433	422.6949
t_{s1}	63.4771	59.8022	56.2870	52.8294
t_{s2}	33.8213	30.7214	27.7901	24.8791
t_{s3}	41.2940	38.2766	35.3876	32.5044
t_{s4}	113.0772	149.0306	219.9263	427.0446
t_{s5}	32.7760	30.3503	28.0401	25.7003
t_{kk}	167.4342	201.9362	273.2664	462.0933
t_{li}, i = 1,2	216.8347	284.5069	450.2003	694.1310
t_b	**251.8103**	**344.7675**	**605.8373**	**875.1920**

where $t^* = t_{lr}, t_s, t_w, t_k$

Table 5. ***PRE*** **of Different Estimators for Second Artificial Population**

ρ_{xy} Estimators	0.6	0.7	0.8	0.9
t_m	100	100	100	100
t_r	104.8471	138.8311	205.6501	398.1569
t_p	29.3583	27.3620	25.5519	23.8702
t^*	146.7487	182.8806	257.6252	486.3080
t_{sd}	121.6439	161.5704	240.9921	477.5131
t_u	147.7398	182.9840	258.8232	518.2825
t_{pd}	33.9792	31.6971	29.6346	27.7292
t_p	105.3971	139.2860	205.9366	398.1694
t_{sk}	131.7491	173.9820	255.6094	478.8917
t_{up1}	124.7454	165.8426	247.2665	485.4327
t_{up2}	111.5293	147.9374	220.0929	433.1026
t_{s1}	59.5670	56.3277	53.3660	50.6220
t_{s2}	28.5793	26.9147	25.4524	24.1540
t_{s3}	37.5642	35.2375	33.2039	31.3960
t_{s4}	111.7965	150.2005	226.6745	455.1018
t_{s5}	31.1449	29.3482	27.7098	26.1743
t_{kk}	150.174	186.3007	261.0335	489.6952
t_{li}, $i = 1,2$	156.3934	195.8246	278.1517	539.7598
$\boldsymbol{t_b}$	**156.8058**	**196.3830**	**279.0534**	**542.2117**

where $t^* = t_{lr}, t_s, t_w, t_k$

4. Conclusion

In this paper, we considered an improved class of estimators for population mean and explored from the view point of superiority. The estimator was proposed and studied in order to obtain the best possible method of estimation over various entrants. In addition, the usual mean estimator, usual ratio estimator, usual regression estimator, Srivastava (1967) estimator, Walsh (1970) estimator, Sisodia and Dwivedi (1981) estimator, Upadhyaya et al. (1985) estimator, Pandey and Dubey (1988) estimator, Prasad (1989) estimator, Singh and Kakran (1993) estimator, Upadhyaya and Singh (1999) estimator, Singh (2003) estimators, Khoshnevisan et al. (2007), Koyuncu and Kadilar (2009) estimator, Bhushan and Gupta (2014) estimators are identified as the members of the suggested class of estimators for suitably chosen values of scalars. The properties of these estimators can also be obtained from the properties of proposed estimator by giving the suitable values of scalars. Thus, the proposed estimator unifies the properties of several other estimators. Further, to enhance the theoretical results, a numerical study was carried out using three different real data sets and a simulation study was carried out using two artificially generated data sets with different amount of correlation coefficient. The empirical results are found to be quite satisfactory showing the dominance of the proposed class of estimators over its counterparts when the correlation between study variable y and auxiliary variable x is positive. Thus, this study justifies its worthiness.

Acknowledgement

The authors are extremely thankful to the anonymous referees for their valuable comments on the paper and to the editor-in-chief Professor Jose Ramon G. Albert.

References

BHUSHAN, S. and GUPTA, R. 2014. "Improved Seals Type Logarithmic Estimators Using Auxiliary Information." Proceeding of the National Conference on Recent Advances in Statistical and Mathematical Sciences and their Applications (RASMSA), Kumaun University, Nainital.

BHUSHAN, S. and GUPTA, R., SINGH, S. and KUMAR, A. 2020a. "A Modified Class of Log Type Estimators for Population Mean using Auxiliary Information on Variable." *International Journal of Applied Engineering Research* 15(6): 612-627.

———. 2020b. "Some improved classes of estimators using auxiliary information." *International Journal for Research in Applied Science & Engineering Technology* 8(VI): 1088-1098.

———. 2020c. "A New Efficient Log-type Class of Estimators Using Auxiliary Variable." *International Journal of Statistics and Systems* 15(1): 19-28.

BHUSHAN, S. and KUMAR, A. 2020a. "On Optimal Classes of Estimators Under Ranked Set Sampling." *Communications in Statistics - Theory and Methods*, DOI:10.1080/03610926.2020.1777431.

———. 2020b. "Log Type Estimators of Population Mean Under Ranked Set Sampling. *Predictive Analytics using Statistics and Big Data: Concepts and Modelling* 28: 47-74. DOI:10.2174/97898114904911200010007.

COCHRAN, W.G. 1977. *Sampling Techniques*. John Wiley and Sons, New York, U.S.A.

KHOSHNEVISAN, M., SINGH, R., CHAUHAN, P. and SMARANDACHE, F. 2007. "A General Family of Estimators for Estimating Population Mean Using Known Value of Some Population Parameter(s). *East J Math* 22: 181-191.

KOYUNCU, N. and KADILAR, C. 2009. "Ratio and Product Estimators in Stratified Random Sampling." *J Stat Plan Inference* 139: 2552-2558.

MURTHY, M.N. 1967. *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.

PANDEY, B.N. and DUBEY, V. 1988. "Modified Product Estimator Using Coefficient of Variation of Auxiliary Variate, *Assam Statistical Rev.* 2(2): 64-66.

PRASAD, B. 1989. "Some Improved Ratio Type Estimators of Population Mean and Ratio in Finite Population Sample Surveys. *Commun Stat Theory Methods*18: 379-392.

SINGH, H.P. and KAKRAN, M.S. 1993. "A Modified Ratio Estimator Using Known Coefficient of Kurtosis of an Auxiliary Character." Unpublished.

SINGH, G.N. 2003. "On the Improvement of Product Method of Estimation in Sample Surveys." *Jour. Ind. Soc. Agri. Statistics* 56(3): 267-275.

SINGH H.P. and TAILOR, R. 2003. "Use of Known Correlation Coefficient in Estimating the Finite Population Mean." *Statistics in Transition* 6(4): 555-560.

SISODIA, B.V.S. and DWIVEDI, V.K. (1981). "A Modified Ratio Estimator Using Coefficient of Variation of Auxiliary Variable. *J. Ind. Soc. Agril. Statist*. 33(2): 13-18.

SRIVASTAVA, S.K. 1967. "An Estimator Using Auxiliary Information." *Calcutta Statist. ASSOC. BULL.* 16: 121-132.

UPADHYAYA, L.N., SINGH, H.P., VOS, J.W.E. 1985. "On the Estimation of Population Means and Ratios Using Supplementary Information. *Stat Neerl* 39(3): 309318.

UPADHYAYA, L.N., SINGH, H.P. 1999. "Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean." *Biometrical Journal* 41(5): 627-636.

WALSH, J.E. 1970. "Generalization of Ratio Estimator for Population Total." *Sankhya*, *A* 32: 99-106.