



Thailand Statistician
October 2022; 20(4): 754-765
<http://statassoc.or.th>
Contributed paper

An Efficient Class of Estimators for Finite Population Mean Using Auxiliary Variable

Kuldeep Kumar Tiwari*[a] and Puneet Kumar [b]

[a] School of Mathematics, Shri Mata Vaishno Devi University, Katra, Jammu and Kashmir, India

[b] Independent Researcher, Jammu, Jammu and Kashmir, India

*Corresponding author; e-mail: kuldeep.smvd@gmail.com

Received: 3 July 2020
Revised: 22 August 2020
Accepted: 3 January 2021

Abstract

We have proposed a general class of estimators for population mean using auxiliary variable in simple random sampling. Our estimator contains various estimators from literature and many more can be developed from it. The expressions for bias and mean square error(MSE) are derived. It is found that the estimator in question is better than the other available alternatives. The Monte-Carlo simulation is carried out in various cases to show the validity of the estimator.

Keywords: Auxiliary variable, study variable, estimator, simulation, mean squared error, efficiency.

1. Introduction

In the present era, we deal with a huge amount of data where precision and accuracy are highly required. The accuracy can be increased by reducing bias and precision is obtained by minimizing variance. The mean square error (MSE) of an estimator $\hat{\theta}$ is defined as $MSE(\hat{\theta}) = Variance(\hat{\theta}) + [Bias(\hat{\theta})]^2$. In order to get a reliable result in a sampling survey, it is desirable to use an estimator which involves minimum MSE.

There have been consistent efforts to make estimators more efficient. For the first time, Cochran (1940) used auxiliary information to estimate the population mean by developing the ratio estimator. Since then, the use of auxiliary information in estimation becomes a good tool to enhance the efficiency of the estimators. Following Cochran (1940), Murthy (1964) proposes the product method of estimation. Murthy(1964) also shows that whereas the ratio estimators perform better when there is a positive correlation between study variable and auxiliary variable, the product estimators perform better when the correlation is negative. Again Cochran (1977) suggested the regression estimator and found it outshines both ratio and product estimators. When the regression line passes through the origin, the efficiency of the ratio, product and regression estimators coincides.

Various known information related to auxiliary variables such as mean, variance, coefficient of variation, coefficient of kurtosis, coefficient of skewness, correlation coefficient, regression coefficient, standard deviation, etc has been used in literature. Some of them are Sisodia and Dwivedi (1981), Upadhyaya and Singh (1984), Kothwala and Gupta (1989), Singh and Singh (1993), Kumar and Agarwal (1997), Singh and Tailor (2003), Kadilar and Cingi (2004), Kadilar and Cingi (2006),

Khoshnevisan et al. (2007), Gupta and Shabbir (2008), Yan and Tian (2010), Adewara et al. (2012), Yadav and Kadilar (2013), Yadav et al. (2017), Pal et al. (2017), Baghel and Yadav (2020) etc.

Let $\Pi = \{\pi_1, \pi_2, \dots, \pi_N\}$ be a finite populaion of size N . Let y_i and x_i be the i^{th} observation of study variable y and auxiliary variable x . The other notations are sample mean of the study variable $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, sample mean of the auxiliary variable $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, population mean of the study variable $\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i$, population mean of the auxiliary variable $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$, population variance of study variable $S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2$, population variance of auxiliary variable $S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$, covariance between variables

$$S_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y}), C_y = \frac{S_y}{\bar{Y}}, C_x = \frac{S_x}{\bar{X}}, \rho = \frac{S_{xy}}{S_x S_y}, \beta_{yx} = \frac{S_{xy}}{S_x^2},$$

$$C_{xy} = \rho C_x C_y, K_{yx} = \rho \frac{C_y}{C_x}, \lambda = \frac{1}{n} - \frac{1}{N}, m = \frac{\bar{X}}{\bar{Y}}.$$

2. Some Existing Estimators

We are now giving a brief review of some of the existing estimators related to the proposed general class of estimators.

The sample mean per unit estimator $\bar{y}$, combined ratio, product and regression estimators are classic estimators as discussed earlier. In 1991, Rao gave a difference type estimator in the search of better efficiency. Bahl and Tuteja (1991) used the exponential function to define a new ratio and product type estimator. Grover and Kaur (2011) proposed a difference-cum-exponential type estimator and discussed its properties. An exponential ratio type estimator suggested by Kadilar in 2016. She showed that the optimal MSE of her proposed estimator is the same as that of the regression estimator. The discussed estimators and their MSEs are given in Table 1.

Table 1 Some known estimators and their *MSEs*

Source	Estimator	<i>MSE</i>
Usual estimator	$T_1 = \bar{y}$	$\lambda \bar{Y}^2 C_y^2$
Ratio estimator	$T_2 = \bar{y} \left(\frac{\bar{X}}{\bar{x}} \right)$	$\lambda \bar{Y}^2 [C_y^2 + C_x^2 - 2C_{xy}]$
Regression estimator	$T_3 = \bar{y} + \beta_{yx}(\bar{X} - \bar{x})$	$\lambda \bar{Y}^2 C_y^2 (1 - \rho^2)$
Rao(1991)	$T_4 = k\bar{y} + t(\bar{X} - \bar{x})$	$\bar{Y}^2 \left[1 - \frac{1}{1 + \lambda C_y^2 (1 - \rho^2)} \right]$
Bahl and Tuteja(1991)	$T_5 = \bar{y} \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\lambda \bar{Y}^2 \left[C_y^2 + \frac{C_x^2}{4} - C_{xy} \right]$
Grover and Kaur(2011)	$T_6 = [k\bar{y} + t(\bar{X} - \bar{x})] \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\frac{\lambda \bar{Y}^2 [16(1 - \rho^2)(4 - \lambda C_x^2) C_y^2 - \lambda C_x^2]}{64[1 + \lambda C_y^2 (1 - \rho^2)]}$
Kadilar(2016)	$T_7 = \bar{y} \left(\frac{\bar{x}}{\bar{X}} \right)^\delta \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\lambda \bar{Y}^2 C_y^2 (1 - \rho^2)$

3. Proposed Class of Estimator

There is no estimator which perform well in every situation. So its better if we have a class of estimator that contains various estimator and also gives some efficient estimator. The focus of this paper is to develop a class of estimators which perform efficiently for larger domain.

Motivated by Grover & Kaur (2011) and Kadilar (2016), we propose a general class of estimators as

$$t_{(\delta, \alpha)} = [k\bar{y} + t(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right)^\delta \exp \alpha \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \quad (1)$$

where k, t, δ, α are suitable constants.

Generalized class of estimators $t_{(\delta, \alpha)}$ contains various estimators from literature for different values of k, t, δ, α . We are giving a list of member estimators and corresponding values of k, t, δ, α in Table 2.

Table 2 Known members of the proposed class of estimator

k	t	δ	α	Estimator	Source
1	0	0	0	$\bar{y}$	Usual estimator
1	0	1	0	$\bar{y} \left(\frac{\bar{X}}{\bar{x}} \right)$	Ratio estimator
1	0	-1	0	$\bar{y} \left(\frac{\bar{x}}{\bar{X}} \right)$	Product estimator
1	β_{yx}	0	0	$\bar{y} + \beta_{yx}(\bar{X} - \bar{x})$	Regression estimator
1	t	0	0	$\bar{y} + t(\bar{X} - \bar{x})$	Difference type
k	t	0	0	$k\bar{y} + t(\bar{X} - \bar{x})$	Rao (1991)
1	0	0	1	$\bar{y} \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	Bahl and Tuteja (1991)
1	0	0	-1	$\bar{y} \exp \left(\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}} \right)$	Bahl and Tuteja (1991)
k	t	0	1	$[k\bar{y} + t(\bar{X} - \bar{x})] \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	Grover and Kaur (2011)
1	0	$-\delta$	1	$\bar{y} \left(\frac{\bar{x}}{\bar{X}} \right)^\delta \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	Kadilar (2016)
1	0	2	0	$\bar{y} \left(\frac{\bar{X}}{\bar{x}} \right)^2$	Kadilar and Cingi (2003)
1	0	-2	0	$\bar{y} \left(\frac{\bar{x}}{\bar{X}} \right)^2$	Kadilar and Cingi (2003)
1	0	δ	0	$\bar{y} \left(\frac{\bar{X}}{\bar{x}} \right)^\delta$	Shrivastava (1967)
k	$k\beta_{yx}$	0	0	$k[\bar{y} + \beta_{yx}(\bar{X} - \bar{x})]$	Bedi and Hajela (1984)
k	β_{yx}	0	0	$k\bar{y} + \beta_{yx}(\bar{X} - \bar{x})$	Kaur (1985)
k	$(1 - k)$	0	0	$k\bar{y} + (1 - k)(\bar{X} - \bar{x})$	Jain (1987)
1	β_{yx}	1	0	$[\bar{y} + \beta_{yx}(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right)$	Kadilar and Cingi (2004)

4. Expressions for Bias and MSE

To obtain the expressions for bias and mean square error (MSE) of proposed class of estimator $t_{(\delta, \alpha)}$, we have to use the error terms $e_y = \frac{\bar{y} - \bar{Y}}{\bar{Y}}$ and $e_x = \frac{\bar{x} - \bar{X}}{\bar{X}}$ such that $E(e_y) = E(e_x) = 0$ and $E(e_y^2) = \lambda C_y^2$, $E(e_x^2) = \lambda C_x^2$, $E(e_y e_x) = \lambda C_{xy}$.

Express Equation (1) in terms of error, we have

$$t_{(\delta, \alpha)} = [k\bar{Y}(1 + e_y) + t(\bar{X} - \bar{X}(1 + e_x))] \left(\frac{\bar{X}}{\bar{X}(1 + e_x)} \right)^\delta \exp \alpha \left(\frac{\bar{X} - \bar{X}(1 + e_x)}{\bar{X} + \bar{X}(1 + e_x)} \right).$$

Expand the expression assuming $|e_i| < 1, i = x, y$ and ignoring the terms having $e_i, i = x, y$ degree greater than two, we have

$$t_{(\delta, \alpha)} = [k\bar{Y} + k\bar{Y}e_y - t\bar{X}e_x] \left[1 - \delta e_x + \frac{\delta(\delta + 1)}{2} e_x^2 \right] \left[1 - \frac{\alpha}{2} e_x + \left(1 + \frac{\alpha}{2} \right) \frac{\alpha}{4} e_x^2 \right]$$

or

$$t_{(\delta, \alpha)} - \bar{Y} = (k - 1)\bar{Y} + k\bar{Y}e_y - [t\bar{X} + k\bar{Y}(\delta + \frac{\alpha}{2})]e_x - k\bar{Y}(\delta + \frac{\alpha}{2})e_x e_y + \left[\left(t\bar{X} + \frac{k\bar{Y}}{2} \right) \left(\delta + \frac{\alpha}{2} \right) + \frac{k\bar{Y}}{2} \left(\delta + \frac{\alpha}{2} \right)^2 \right] e_x^2. \quad (2)$$

Taking expectation on both sides of Equation (2), we get the bias of $t_{(\delta, \alpha)}$ as

$$B(t_{(\delta, \alpha)}) = (k-1)\bar{Y} - k\bar{Y}(\delta + \frac{\alpha}{2})\lambda C_{xy} + \left[\left(t\bar{X} + \frac{k\bar{Y}}{2} \right) \left(\delta + \frac{\alpha}{2} \right) + \frac{k\bar{Y}}{2} \left(\delta + \frac{\alpha}{2} \right)^2 \right] \lambda C_x^2.$$

To get MSE, squaring Equation (2) on both sides and simplifying, we get

$$(t_{(\delta, \alpha)} - \bar{Y})^2 = \bar{Y}^2 [(k-1)^2 + k^2 e_y^2 - \{2k(2k-1)(\delta + \frac{\alpha}{2}) + 2ktm\} e_x e_y + \{t^2 m^2 + k(2k-1)(\delta + \frac{\alpha}{2})^2 + 2tm(2k-1)(\delta + \frac{\alpha}{2}) + k(k-1)(\delta + \frac{\alpha}{2})\} e_x^2] \quad (3)$$

where $m = \frac{\bar{X}}{\bar{Y}}$.

Taking expectation on both sides of Equation (3), we get the MSE of $t_{(\delta, \alpha)}$ as

$$MSE(t_{(\delta, \alpha)}) = \bar{Y}^2 [(k-1)^2 + k^2 \lambda C_y^2 - \{2k(2k-1)(\delta + \frac{\alpha}{2}) + 2ktm\} \lambda C_{xy} + \{t^2 m^2 + k(2k-1)(\delta + \frac{\alpha}{2})^2 + 2tm(2k-1)(\delta + \frac{\alpha}{2}) + k(k-1)(\delta + \frac{\alpha}{2})\} \lambda C_x^2]$$

or

$$MSE(t_{(\delta, \alpha)}) = \bar{Y}^2 [1 + \eta_1 k + \eta_2 k^2 + \eta_3 t + \eta_4 t^2 + \eta_5 kt] \quad (4)$$

where

$$\begin{aligned} \eta_1 &= -2 + 2(\delta + \frac{\alpha}{2})\lambda C_{xy} - (\delta + \frac{\alpha}{2})^2 \lambda C_x^2 - (\delta + \frac{\alpha}{2})\lambda C_x^2 \\ \eta_2 &= 1 + \lambda C_y^2 - 4(\delta + \frac{\alpha}{2})\lambda C_{xy} + 2(\delta + \frac{\alpha}{2})^2 \lambda C_x^2 + (\delta + \frac{\alpha}{2})\lambda C_x^2 \\ \eta_3 &= -2m(\delta + \frac{\alpha}{2})\lambda C_x^2 \\ \eta_4 &= \lambda C_x^2 m^2 \\ \eta_5 &= -2m\lambda C_{xy} - 4m(\delta + \frac{\alpha}{2})^2 \lambda C_x^2. \end{aligned}$$

We know that “If a function $f(x, y)$ having second order partial derivative f_{xx} , f_{yy} and f_{xy} continuous at critical point (x_o, y_o) , then $f(x, y)$ is minimum at (x_o, y_o) whenever $D = f_{xx}(x_o, y_o)f_{yy}(x_o, y_o) - f_{xy}^2(x_o, y_o) > 0$ and $f_{xx}(x_o, y_o) > 0$ ”.

To get minimum MSE , differentiate Equation (4) partially with respect to t and equate to zero,

$$\eta_5 k + 2\eta_4 t + \eta_3 = 0 \quad (5)$$

again, differentiate Equation (4) partially with respect to k and equate to zero,

$$2\eta_2 k + \eta_5 t + \eta_1 = 0. \quad (6)$$

On solving Equations (5) and (6), we get the optimum value of k and t as

$$k_o = \frac{\eta_3 \eta_5 - 2\eta_1 \eta_4}{4\eta_2 \eta_4 - \eta_5^2}, \quad t_o = \frac{\eta_1 \eta_5 - 2\eta_2 \eta_3}{4\eta_2 \eta_4 - \eta_5^2}. \quad (7)$$

Now, calculate second order partial derivatives of $MSE(t_{(\delta, \alpha)})$ at (k_o, t_o) , we have

$$MSE(t_{(\delta, \alpha)})_{tt} = 2\bar{Y}^2 \eta_4, \quad MSE(t_{(\delta, \alpha)})_{kk} = 2\bar{Y}^2 \eta_2, \quad MSE(t_{(\delta, \alpha)})_{tk} = \bar{Y}^2 \eta_5.$$

So,

$$D = \bar{Y}^4 [4\eta_2 \eta_4 - \eta_5^2] \text{ and } MSE(t_{(\delta, \alpha)})_{tt} = 2\bar{Y}^2 \eta_4.$$

Now, if $\bar{Y} = 0$ then $MSE(t_{(\delta, \alpha)})$ in Equation (4) becomes zero, so no need for minimization. If $\bar{Y} \neq 0$, then $D > 0$ whenever $4\eta_2\eta_4 - \eta_5^2 > 0$. Also, for $\bar{Y} \neq 0$, $MSE(t_{(\delta, \alpha)})_{tt} > 0$, since $\eta_4 = \lambda C_x^2 m^2 > 0$ as λ is always positive and other terms, coefficient of variation C_x and $m = \frac{\bar{X}}{\bar{Y}}$ are squared.

So, k_o and t_o in Equation (7) are points of minima if their denominator term is positive. That is, $MSE(t_{(\delta, \alpha)})$ is minimum at (k_o, t_o) if $4\eta_2\eta_4 - \eta_5^2 > 0$.

Substitute the optimum value of k and t in Equation (4), we get the minimum MSE of $t_{(\delta, \alpha)}$ as

$$MSE_{min}(t_{(\delta, \alpha)}) = \bar{Y}^2 \left[1 - \frac{\eta_1^2\eta_4 + \eta_2\eta_3^2 - \eta_1\eta_3\eta_5}{4\eta_2\eta_4 - \eta_5^2} \right].$$

5. Consistency of the Proposed Class of Estimator

Let $\hat{\theta}_n$ is an estimator of θ , where n in suffix represents sample size. $\hat{\theta}_n$ is said to be a consistent estimator of θ if

$$\lim_n MSE(\hat{\theta}_n) = 0.$$

That is, if n increases infinitely then $MSE(\hat{\theta}_n)$ approaches to zero.

Consistency is an asymptotic property and proposed class of estimator $t_{(\delta, \alpha)}$ is defined over finite population. So, assume population size N is infinitely large that we can take a sample of any size.

Since $\lambda = \frac{1}{n} - \frac{1}{N}$. So, if n and N increases infinitely then λ approaches to zero. Applying limit $\lambda \rightarrow 0$ in Equation (4), we have

$$\lim_{\lambda \rightarrow 0} MSE(t_{(\delta, \alpha)}) = \bar{Y}^2(k-1)^2. \quad (8)$$

So, the proposed class of estimators is not consistent except $\bar{Y} = 0$ or $k = 1$. We have no control over $\bar{Y}$, but k is a constant that can take any suitable value.

Use $k = 1$ in Equation (8), we have

$$\lim_{\lambda \rightarrow 0} MSE(t_{(\delta, \alpha)}) = 0. \quad (9)$$

Equation (9) concludes that, for $k = 1$, the proposed class of estimators $t_{(\delta, \alpha)}$ generates consistent estimator.

6. Efficiency Comparison

In this section, we have obtained the conditions under which the proposed class of estimators $t_{(\delta, \alpha)}$ is more efficient than other considered estimators. The required condition is obtained by using their $MSEs$.

To compare the efficiency of proposed class of estimators, we express the $MSEs$ given in Table 1 in terms of η 's. we have

$$MSE(T_1) = \bar{Y}^2(\eta_6 - 1)$$

$$MSE(T_2) = \bar{Y}^2(\eta_6 + \eta_7 - 2K_{yx}\eta_7 - 1)$$

$$MSE(T_3) = \bar{Y}^2(\eta_6 - 1)(1 - \rho^2)$$

$$MSE(T_4) = \bar{Y}^2 \left[1 - \frac{1}{1 + (\eta_6 - 1)(1 - \rho^2)} \right]$$

$$MSE(T_5) = \bar{Y}^2 \left[(\eta_6 - 1) + \frac{\eta_7}{4} (1 - 4K_{yx}) \right]$$

$$MSE(T_6) = \bar{Y}^2 \left[1 - \frac{64 + \eta_7 \{ 16(\eta_6 - 1)(1 - \rho^2) + \eta_7 \}}{64\{\rho^2 + \eta_6(1 - \rho^2)\}} \right]$$

$$MSE(T_7) = \bar{Y}^2(\eta_6 - 1)(1 - \rho^2)$$

where $\eta_6 = 1 + \lambda C_y^2$, $\eta_7 = \lambda C_x^2$.

Now we have to find the inequalities $MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_i) \leq 0$, $i = 1, 2, \dots, 7$. This will give the condition where the proposed class of estimators is better than other referred estimators.

$$\begin{aligned}
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_1) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 (\eta_6 - 1) \leq 0 \\
&\quad \eta_6 + \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq 2 \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_2) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 (\eta_6 + \eta_7 - 2K_{yx} \eta_7 - 1) \leq 0 \\
&\quad \eta_6 + \eta_7 - 2K_{yx} \eta_7 + \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq 2 \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_3) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 (\eta_6 - 1)(1 - \rho^2) \leq 0 \\
&\quad (\eta_6 - 1)(1 - \rho^2) + \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq 1 \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_4) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 \left[1 - \frac{1}{1 + (\eta_6 - 1)(1 - \rho^2)} \right] \leq 0 \\
&\quad \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq \frac{1}{1 + (\eta_6 - 1)(1 - \rho^2)} \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_5) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 \left[(\eta_6 - 1) + \frac{\eta_7}{4} (1 - 4K_{yx}) \right] \leq 0 \\
&\quad \eta_6 + \frac{\eta_7}{4} (1 - 4K_{yx}) + \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq 2 \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_6) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 \left[1 - \frac{64 + \eta_7 \{16(\eta_6 - 1)(1 - \rho^2) + \eta_7\}}{64\{\rho^2 + \eta_6(1 - \rho^2)\}} \right] \leq 0 \\
&\quad \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq \frac{64 + \eta_7 \{16(\eta_6 - 1)(1 - \rho^2) + \eta_7\}}{64\{\rho^2 + \eta_6(1 - \rho^2)\}} \\
MSE_{min}(t_{(\delta, \alpha)}) - MSE(T_7) &= \bar{Y}^2 \left[1 - \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \right] - \bar{Y}^2 (\eta_6 - 1)(1 - \rho^2) \leq 0 \\
&\quad (\eta_6 - 1)(1 - \rho^2) + \frac{\eta_1^2 \eta_4 + \eta_2 \eta_3^2 - \eta_1 \eta_3 \eta_5}{4\eta_2 \eta_4 - \eta_5^2} \geq 1.
\end{aligned}$$

To show the efficiency of the proposed class of estimators numerically, we are considering four particular cases of $t_{(\delta, \alpha)}$ for different choices of δ and α such as

Table 3 The value of $t_{(\delta, \alpha)}$

δ	α	$t_{(\delta, \alpha)}$
1	1	$t_{(1,1)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right) \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$
2	1	$t_{(2,1)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right)^2 \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$
1	2	$t_{(1,2)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right) \exp 2 \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$
-1	-1	$t_{(-1,-1)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{x}}{\bar{X}} \right) \exp \left(\frac{\bar{x} - \bar{X}}{\bar{X} + \bar{x}} \right)$

In Table 3, we have randomly chosen simple combination of δ and α . It is not necessary that the estimators in Table 3 will perform well for every population. But, for a particular population, we can find the value of δ and α for which $PRE(t_{(\delta, \alpha)}, T_1)$ is high. Required values of δ and α for a particular population can be found by using Hit and Trial Method or by graphing δ and α against $PRE(t_{(\delta, \alpha)}, T_1)$. To explain it numerically, we use $t_{(4,7)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right)^4 \exp 7 \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$ and $t_{(3,3.5)} = [k_o \bar{y} + t_o(\bar{X} - \bar{x})] \left(\frac{\bar{X}}{\bar{x}} \right)^3 \exp 3.5 \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$ in empirical study.

7. Empirical Study

For numerical illustration of proposed class of estimators, we use the populations given in Table 4.

Table 4 Populations

Population	Source	N	n	$\bar{Y}$	$\bar{X}$	C_y	C_x	ρ
1	Khoshnevisan et al. (2007)	20	8	19.55	18.8	0.3552	0.3943	-0.9199
2	Koyuncu & Kadilar (2009)	923	180	436.4345	11440.5	1.7183	1.864	0.9543
3	Gupta & Shabbir (2008)	200	50	500	25	15	2	0.9
4	Steel & Torrie (1960)	30	8	0.6860	4.6537	0.4803	0.2295	0.1794
5	Gupta & Kothwala (1990)	400	100	6.56	36.71	0.9634	0.9932	-0.40

In Table 5, we are giving a relative comparison of estimators with usual estimator $T_1 = \bar{y}$ that is computed in terms of percent relative efficiency (PRE) using Equation (10).

$$PRE(t, T_1) = \frac{MSE(T_1)}{MSE(t)} \times 100. \tag{10}$$

Table 5 Percent relative efficiency(PRE) of estimators with $T_1 = \bar{y}$

Estimator	$PRE(., T_1)$				
	Population-1	Population-2	Population-3	Population-4	Population-5
T_1	100	100	100	100	100
T_2	23.39	940.38	128.57	94.62	34.63
T_3	650.26	1119.67	526.31	103.33	119.05
T_4	651.20	1120.99	863.81	105.44	119.74
T_5	42.93	386.13	113.06	102.95	59.59
T_6	654.07	1128.99	877.04	105.57	119.98
T_7	650.26	1119.67	526.31	103.33	119.05
$t_{(1,1)}$	654.22	1141.58	889.92	105.10	119.24
$t_{(2,1)}$	704.38	5171.78	893.89	104.01	120.47
$t_{(1,2)}$	961.15	1386.70	905.43	104.59	119.10
$t_{(-1,-1)}$	938.68	4144.92	797.78	103.99	120.43

Table 5 shows that the estimators $t_{(1, 1)}$, $t_{(2, 1)}$, $t_{(1, 2)}$ and $t_{(-1, -1)}$ performs better than the other estimators for Population- 1, 2 and 3 with sufficient gain in efficiencies. For Population- 4 and 5, the performance of all the estimators in Table 5 is almost similar. Using Hit and Trial Method, we have found that for $\delta = 4$ and $\alpha = 7$, the proposed class of estimators performs well with $PRE(t_{(4, 7)}, T_1) = 1026.45$ in Population-4. While for Population-5, $\delta = 3$ and $\alpha = 3.5$ works good with $PRE(t_{(3, 3.5)}, T_1) = 929.59$.

8. Monte-Carlo Simulation Study

In this section, we have performed The Monte-Carlo simulation using R software to check the validity of the proposed estimator. We have generated six different populations of size $N = 1500$. The cases under which the populations are generated are given below.

Case-1a Model $Y = -2 - 3X + \varepsilon$ used where X is taken from normal distribution with mean 1.2 and variance 1.5. Sample size $n = 400$.

Case-1b Model $Y = 2 + 3X + \varepsilon$ used where X is taken from normal distribution with mean 1.2 and variance 1.5. Sample size $n = 400$.

Case-2a Model $Y = -2 - 3X + \varepsilon$ used where X is taken from exponential distribution with rate 0.7. Sample size $n = 350$.

Case-2b Model $Y = 2 + 3X + \varepsilon$ used where X is taken from exponential distribution with rate 0.7. Sample size $n = 350$.

Case-3a Model $Y = -2 - 3X + \varepsilon$ used where X is taken from uniform distribution with minimum -1 and maximum 3. Sample size $n = 300$.

Case-3b Model $Y = 2 + 3X + \varepsilon$ used where X is taken from uniform distribution with minimum -1 and maximum 3. Sample size $n = 300$.

In all the cases, The error term ε follows the normal distribution with mean 0 and variance 1. For all the six populations, PREs calculated using Equation (10). 100000 replications made to get better approximation of results.

Table 6 Percent relative efficiency (PRE) of estimators with $T_1 = \bar{y}$ using Monte-Carlo simulation

Estimator	$PRE(., T_1)$					
	Case-1a	Case-1b	Case-2a	Case-2b	Case-3a	Case-3b
T_1	100	100	100	100	100	100
T_2	292.83	292.89	386.06	386.00	205.17	205.19
T_3	2019.63	2019.72	1904.40	1904.12	1299.81	1299.68
T_4	2019.75	2019.84	1904.51	1904.23	1299.95	1299.83
T_5	1061.77	1061.63	834.46	834.48	975.15	975.01
T_6	2125.92	2126.01	1908.07	1907.79	1303.52	1303.39
T_7	2119.63	2119.72	1904.40	1904.12	1299.81	1299.68
$t_{(1,1)}$	2157.92	2158.01	1924.38	1924.10	1318.60	1318.46
$t_{(2,1)}$	4201.29	4201.29	2703.89	2703.37	2233.36	2232.86
$t_{(1,2)}$	2455.03	2455.12	2073.03	2072.70	1466.82	1466.64
$t_{(-1,-1)}$	4153.76	4153.77	2693.91	2693.39	2210.30	2209.82

Simulation results in Table 6 show that the PRE values of proposed estimators are higher than existing estimators in all cases irrespective of population distribution and sample size, which validate that proposed estimators are more efficient as well as consistent in comparison with existing estimators.

9. Conclusions

The proposed class of estimators $t_{(\delta, \alpha)}$ is comprehensive that it embraces many estimators suggested earlier as well as can produce numerous new and efficient estimators. From empirical study (Table 5) and simulation study (Table 6), we have shown that the proposed estimators are more efficient than usual estimator T_1 , classical ratio estimator T_2 , regression estimator T_3 , Rao (1991)'s estimator T_4 , Bahl and Tuteja (1991)'s estimator T_5 , Grover and Kaur (2011)'s estimator T_6 and Kadilar (2016)'s estimators T_7 . Referring to the results, we find the proposed class of estimators applicable in practice. Thus we recommend it for practical and future studies.

References

- Adewara AA, Singh R, Kumar M. Efficiency of some modified ratio and product estimators using known value of some population parameters. *Int J Appl Sci Technol*. 2012; 2(2): 76-79.
- Baghel S, Yadav SK. Restructured class of estimators for population mean using an auxiliary variable under simple random sampling scheme. *JAMSI*. 2020; 16(1): 91-75.
- Bahl S, Tuteja RK. Ratio and product type exponential estimator. *J Inform Optim Sci*. 1991; 12(1): 159-164.
- Bedi PK, Hajela D. An estimator for population mean utilizing known coefficient of variation and auxiliary variable. *J Statist Res*. 1984; 18: 29-34.
- Cochran WG. The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *J Agric Sci*. 1940; 30(2): 262-275.
- Cochran WG. Sampling techniques. New York: John Wiley and Sons; 3rd ed. 1977.
- Gupta PC, Kothwala NH. A study of second order approximation for some product type estimators. *J Ind Soc Agri Statist*. 1990; 42(2): 171-185.
- Gupta S, Shabbir J. On improvement in estimating the population mean in simple random sampling. *J Appl Stat*. 2008; 35(5): 559-566.
- Grover LK, Kaur P. An improved estimator of the finite population mean in simple random sampling. *Model Assist Stat Appl*. 2011; 6(1): 47-55.
- Jain RK. Properties of estimators in simple random sampling using auxiliary variable. *Metron*. 1987; 45(1-2): 265-271.
- Kadilar C, Cingi H. A study on the chain ratio type estimator, *Hacet J Math Stat*. 2003; 32: 105-108.
- Kadilar C, Cingi H. Ratio estimators in simple random sampling. *Appl Math Comput*. 2004; 151: 893-902.
- Kadilar C, Cingi H. An improvement in estimating the population mean by using the correlation coefficient. *Hacet J Math Stat*. 2006; 35(1): 103-109.
- Kadilar GO. A New Exponential Type Estimator for the Population Mean in Simple Random Sampling. *J Mod Appl Stat Methods*. 2016; 15(2): 207-214.
- Kaur P. An efficient regression type estimator in survey sampling. *Biometrical*. 1985; 27: 107-110.
- Khoshnevisan M, Singh R, Chauhan P, Sawan N, Smarandache F. A general family of estimators for estimating population mean using known value of some population parameters. *Far East J Theor Stat*. 2007; 22(2), 181-191.
- Kothwala NH, Gupta PC. Estimation of population mean with knowledge of coefficient of variation with p auxiliary variables. *Metron*. 1989; 47: 107-119.
- Koyuncu N, Kadilar C. Efficient estimators for the population mean. *Hacet J Math Stat*. 2009; 38(2): 217-225.
- Kumar P, Agarwal SK. Alternative estimators for the population totals in multiple characteristic surveys. *Commun Stat Theo Meth*. 1997; 26(10): 2527-2537.
- Murthy M. Product method of estimation. *Sankhya A*. 1964; 26(1): 69-74.
- Pal SK, Singh HP, Kumar S, Chatterjee K. A family of efficient estimators of the finite population mean in simple random sampling. *J Stat Comput Simul*. 2017; 88(5): 920-934.
- R Core Team. R: A language and environment for statistical computing. The R Foundation for Statistical Computing. Vienna: Austria; version 3.2.3 (2015-12-10), x86_64-pc-linux-gnu (64-bit). Available from: <https://www.R-project.org>.
- Rao TJ. On certain methods of improving ratio and regression estimators. *Commun Stat- Theory Methods*. 1991; 20(10): 3325-3340.
- Singh HP, Singh VP. A general class of unbiased estimators of a parameter. *Calcutta Stat Assoc Bull*. 1993; 43: 127-131.
- Singh HP, Tailor R. Use of known correlation coefficient in estimating the finite population mean. *Stat Transit*. 2003; 6(4): 555-560.

- Sisodia BVS, Dwivedi VK. A modified ratio estimator using coefficient of variation of auxiliary variable, *J Indian Soc Agric Stat.* 1981; 33: 13-18.
- Srivastava SK. An estimator using auxiliary information in sample surveys. *Calcutta Stat Assoc Bull.* 1967; 16, 121-132.
- Upadhyaya LN, Singh HP. On the estimation of the population mean with known coefficient of variation. *Biom J.* 1984; 26: 915-922.
- Steele RGD, Torrie JH. Principles and procedures of statistics. New York: McGraw Hill Book Co. 1960.
- Yadav DK, Kumar R, Misra S, Yadav SK. An improved estimator of population mean using information on median of the study Variable. *Int J Math Trends Technol.* 2017; 46(2).
- Yadav SK, Kadilar C. Efficient family of exponential estimators for the population mean. *Hacet J Math Stat.* 2013; 42(6): 671-677.
- Yan Z, Tian B. Ratio method to the mean estimation using coefficient of skewness of auxiliary variable. *ICICA.* 2010; Part II, CCIS 106, 103-110.

Appendix

R code for simulation:

```
#####
d = 1 # delta
a = 1 # alpha
#####
L = 100000 # No. of replication
N = 1500 # Population size
n = 400 # Sample size for sample 1a& 1b
# n = 350 # Sample size for sample 2a& 2b
# n = 300 # Sample size for sample 3a& 3b
#####
X = rnorm(N, 1.2, 1.5) # Population model-1a
Y = -2 - 3 * X + rnorm(N, 0, 1)
# X = rnorm(N, 1.2, 1.5) # Population model-1b
# Y = 2 + 3 * X + rnorm(N, 0, 1)
# X = rexp(N, 0.7) # Population model-2a
# Y = -2 - 3 * X + rnorm(N, 0, 1)
# X = rexp(N, 0.7) # Population model-2b
# Y = 2 + 3 * X + rnorm(N, 0, 1)
# X = runif(N, -1, 3) # Population model-3a
# Y = -2 - 3 * X + rnorm(N, 0, 1)
# X = runif(N, -1, 3) # Population model-3b
# Y = 2 + 3 * X + rnorm(N, 0, 1)
#####
y_bar_vector = array(0, c(L))
x_bar_vector = array(0, c(L))
Sx_vector = array(0, c(L))
Sy_vector = array(0, c(L))
r_vector = array(0, c(L))
#####
for(i in 1:L){
  set.seed(i)
  x = rnorm(n, 1.2, 1.5) # sample-1a
```

```

y = -2 - 3 * x + rnorm(n, 0, 1)
# x = rnorm(n, 1.2, 1.5) # sample-1b
# y = 2 + 3 * x + rnorm(n, 0, 1)
# x = rexp(n, 0.7) # sample-2a
# y = -2 - 3 * x + rnorm(n, 0, 1)
# x = rexp(n, 0.7) # sample-2b
# y = 2 + 3 * x + rnorm(n, 0, 1)
# x = runif(n, -1, 3) # sample-3a
# y = -2 - 3 * x + rnorm(n, 0, 1)
# x = runif(n, -1, 3) # sample-3b
# y = 2 + 3 * x + rnorm(n, 0, 1)
y_bar_vector[i] = mean(y)
x_bar_vector[i] = mean(x)
Sx_vector[i] = sd(x)
Sy_vector[i] = sd(y)
r_vector[i] = cor(x, y)
}

#####
y_bar = mean(y_bar_vector) # sample mean of y
x_bar = mean(x_bar_vector) # sample mean of x
Sx = mean(Sx_vector) # SD of x
Sy = mean(Sy_vector) # SD of y
r = mean(r_vector) # rho
Cx = Sx/x_bar # coefficient of variation for x
Cy = Sy/y_bar # coefficient of variation for y
Cxy = r * Cx * Cy
K = r * Cy/Cx
m = x_bar/y_bar
R = y_bar/x_bar
l = 1/n - 1/N # lambda

#####
B1 = -2 + 2 * (d + a/2) * l * Cxy - (d + a/2)2 * l * Cx2 - (d + a/2) * l * Cx2 # eta.1
B2 = 1 + l * Cy2 - 4 * (d + a/2) * l * Cxy + 2 * (d + a/2)2 * l * Cx2 + (d + a/2) * l * Cx2 # eta.2
B3 = -2 * m * (d + a/2) * l * Cx2 # eta.3
B4 = l * Cx2 * m2 # eta.4
B5 = -2 * m * l * Cxy + 4 * m * (d + a/2) * l * Cx2 # eta.5
B6 = 1 + l * Cy2 # eta.6
B7 = l * Cx2 # eta.7

#####
k = (B3 * B5 - 2 * B1 * B4) / (4 * B2 * B4 - B52) # k optimum
t = (B1 * B5 - 2 * B2 * B3) / (4 * B2 * B4 - B52) # t optimum

#####
V = y_bar2 * (B6 - 1) # variance of usual estimator
MSE_T2R = y_bar2 * (B6 + B7 - 2 * K * B7 - 1) # MSE of ratio estimator
MSE_T3RE = y_bar2 * (B6 - 1) * (1 - r2) # MSE of regression estimator
MSE_T4RAO = y_bar2 * (1 - (1 / (B6 * (1 - r2) + r2))) # MSE of Rao (1991) estimator
MSE_T5BTR = y_bar2 * (B6 - 1 + (B7/4) * (1 - 4 * K)) # MSE of Bahl and Tuteja (1991) estimator
MSE_T6GK = y_bar2 * (1 - ((64 + B7 * (16 * (B6 - 1) * (1 - r2) + B7)) / (64 * (B6 * (1 - r2) + r2)))) # MSE of Grover and Kaur (2011)
MSE_T7GOK = y_bar2 * (B6 - 1) * (1 - r2) # MSE of Kadilar (2016) estimator
p = B12 * B4 + B2 * B32 - B1 * B3 * B5

```

```

 $q = 4 * B2 * B4 - B5^2$ 
 $MSE = y\_bar^2 * (1 - p/q)$  # MSE of proposed class of estimator
#####
 $PRE1 = (V * 100) / V$ 
 $PRE2 = (V * 100) / MSE.T2R$ 
 $PRE3 = (V * 100) / MSE.T3RE$ 
 $PRE4 = (V * 100) / MSE.T4RAO$ 
 $PRE5 = (V * 100) / MSE.T5BTR$ 
 $PRE6 = (V * 100) / MSE.T6GK$ 
 $PRE7 = (V * 100) / MSE.T7GOK$ 
 $PRE8 = (V * 100) / MSE$ 
#####

```