Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 6, Issue 1 (2013): 35-46

AN IMPROVEMENT OVER DIFFERENCE METHOD OF ESTIMATION OF POPULATION MEAN

Housila P. Singh, Anjana Rathour and Ramkrishna S. Solanki

School of Studies in Statistics, Vikram University, Ujjain-456010, India.
Email: ramkssolanki@gmail.com

(Received October 08, 2012)

Abstract

This paper suggests the difference method of estimation of the population mean of the study variable using information on an auxiliary variable with its properties. The optimum estimator in the suggested method has been identified alongwith its mean square error formula. It has been identified that the suggested method is more general and efficient than other existing methods. An empirical study is carried out to judge the merits of proposed method over other traditional methods by using three natural population data sets.

Key words: Study variable, Auxiliary variable, Mean square error, Bias, Simple random sampling.

AMS Classification: 62D05.

1. Introduction

The auxiliary information on the finite population under study is quite often available from previous experience, census or administrative data base in sample surveys. The sampling literature describes wide variety of techniques for using auxiliary information to improve the sampling design and/or obtain more efficient estimators. Ratio, product, difference and regression methods are good examples in this context [see Singh, S. (2003)]. In the present investigation we have suggested an alternative difference method of estimation for population mean which is more general (i.e. includes some traditional methods) and efficient than other existing methods of estimations by using information on an auxiliary variable.

Consider a finite population $U=(U_1,U_2,...,U_N)$ of N identifiable units. Let y and x denote the study variable and auxiliary variable taking values y_i and x_i respectively on the i^{th} unit U_i of the population $U,(i=1,2,...,N)$; and a sample of size $n(<N)$ is drawn by simple random sampling without replacement (SRSWOR) from the population U. Let $\bar{y}\ [=n^{-1}\sum_{i=1}^{n}y_i]$ and $\bar{x}[=n^{-1}\sum_{i=1}^{n}x_i]$ denote the sample means of the study variable y and auxiliary variable x respectively.

The remaining part of the paper is organized as follows. Sec.2 gives the brief review of some traditional methods of estimating population mean of the study variable. In Sec.3, a new difference method of estimation of population mean is described and the expressions for its asymptotic bias and mean square error (*MSE*) are obtained. Sec.4 addresses the problem of efficiency comparisons, while in Sec.5 an empirical study is carried out to evaluate the performance of different methods using three natural population data sets.Sec.6 concludes the paper with final remarks.

2. Reviewing methods of estimation of the population mean

It is very well known that sample mean $\bar{y}$ is an unbiased estimator of population mean $\bar{Y}$ and under SRSWOR its variance is given by

$$Var(\bar{y}) = \theta S_y^2 = \theta \bar{Y}^2 C_y^2 , \qquad (2.1)$$

where

$\theta = n^{-1}(1-f)$, $f = (n/N)$ (sample fraction), $S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ (population mean square of y) and $C_y^2 = (S_y^2/\bar{Y}^2)$ (population coefficient of variation of y).

For estimating population mean $\bar{Y}$, Cochran (1940) and Robson (1957), [Murthy (1964)] have been define usual ratio and product estimator respectively as

$$\bar{y}_R = \bar{y}(\frac{\bar{X}}{\bar{x}}) , \qquad (2.2)$$

$$\bar{y}_P = \bar{y}(\frac{\bar{x}}{\bar{X}}) , \qquad (2.3)$$

where $\bar{X}$, the population mean of auxiliary variable x is assumed to be known.

To the first degree of approximation the mean square errors (*MSE*s) of the ratio estimator $\bar{y}_R$ and the product estimator $\bar{y}_P$ are respectively given by

$$MSE(\bar{y}_R) = \theta \bar{Y}^2[C_y^2 + C_x^2 - 2\rho C_y C_x] , \qquad (2.4)$$

$$MSE(\bar{y}_P) = \theta \bar{Y}^2[C_y^2 + C_x^2 + 2\rho C_y C_x] , \qquad (2.5)$$

where $C_x^2 = (S_x^2/\bar{X}^2)$ (population coefficient of variation of x), $S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$ (population mean square of x), $\rho = (S_{xy}/S_x S_y)$ (population correlation coefficient between y and x) and $S_{xy} = (N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y})$ (population covariance between x and y).

Hansen et al. (1953) has defined the difference estimator for estimating population mean $\bar{Y}$ of the study variable y as

$$\bar{y}_d = [\bar{y} + \alpha_2(\bar{X} - \bar{x})] , \qquad (2.6)$$

where α_2 is suitable chosen constant.

To the first degree of approximation the *MSE* of the difference estimator $\bar{y}_d$ is given by

$$MSE(\bar{y}_d) = \theta[S_y^2 + \alpha_2^2 S_x^2 - 2\alpha_2 S_{xy}] , \qquad (2.7)$$

which is minimized for

$$\alpha_2 = (\frac{S_{xy}}{S_x^2}) = \beta \text{, (say).} \tag{2.8}$$

Substitution of α_2 at (2.8) in (2.6), yields the optimum estimator known as the usual liner regression estimator

$$\bar{y}_\beta = [\bar{y} + \beta(\bar{X} - \bar{x})], \tag{2.9}$$

where β (population regression efficient of y on *x*) is assumed to be known.

To the first degree of approximation the *MSE* (or variance) of $\bar{y}_\beta$ [or minimum *MSE* of $\bar{y}_d$] is given by

$$MSE(\bar{y}_\beta) = Var(\bar{y}_\beta) = MSE_{\min}(\bar{y}_d) = \theta \bar{Y}^2 (1-\rho^2) C_y^2 . \tag{2.10}$$

It follows from (2.10) that the difference estimator $\bar{y}_d$ at its optimum condition is equal efficient as that of the usual linear regression estimator $\bar{y}_\beta$.

Some modifications over the difference estimator $\bar{y}_d$ have been given Bedi and Hajela (1984), Jain (1987), Rao (1991) and Dubey and Singh (2001). These estimators are respectively designed as

$$\bar{y}_{BH} = \alpha_1 [\bar{y} + \beta(\bar{X} - \bar{x})], \quad \text{[Bedi and Hajela (1984)]} \tag{2.11}$$

$$\bar{y}_J = [\alpha_1 \bar{y} + \alpha_2 (\bar{X} - \bar{x})], (\alpha_1 + \alpha_2) = 1; \quad \text{[Jain (1987)]} \tag{2.12}$$

$$\bar{y}_{RG} = [\alpha_1 \bar{y} + \alpha_2 (\bar{X} - \bar{x})], \quad \text{[Rao (1991)]} \tag{2.13}$$

$$\bar{y}_{DS} = [\alpha_1 \bar{y} + \alpha_2 \bar{x} + (1 - \alpha_1 - \alpha_2) \bar{X}], \quad \text{[Dubey and Singh (2001)]} \tag{2.14}$$

where β is as same as defined earlier and α_1 and α_2 are suitable chosen constants .

To the first degree of approximation, the MSE_s of $\bar{y}_{BH}$, $\bar{y}_{RG}$, $\bar{y}_J$ and $\bar{y}_{DS}$ are respectively given by

$$MSE(\bar{y}_{BH}) = \bar{Y}^2 [1 + \alpha_1^2 \{\theta C_y^2 (1 - \rho^2) + 1\} - 2\alpha_1], \tag{2.15}$$

$$MSE(\bar{y}_J) = \bar{Y}^2 [1 + \theta^2 R^{-2} C_x^2 - 2\alpha_1 \{1 + \theta R^{-1} (R^{-1} C_x^2 + \rho C_y C_x)\}$$
$$+ \alpha_1^2 \{1 + \theta (C_y^2 + R^{-2} C_x^2 + 2R^{-1} \rho C_y C_x)\}], \tag{2.16}$$

$$MSE(\bar{y}_{RG}) = \bar{Y}^2 [1 + \alpha_1^2 \{1 + \theta C_y^2\} + \alpha_2^2 R^{-2} C_x^2 \theta - 2\alpha_1 \alpha_2 R^{-1} \rho C_y C_x \theta - 2\alpha_1], \tag{2.17}$$

$$MSE(\bar{y}_{DS}) = \bar{Y}^2 \theta [(\alpha_1^2 C_y^2 + \alpha_2^2 R^{-2} C_x^2 + 2\alpha_1 \alpha_2 R^{-1} \rho C_y C_x)] + (\alpha_1 - 1)^2 (\bar{Y} - \bar{X})^2, \tag{2.18}$$

where $R = (\bar{Y} / \bar{X})$.

The *MSE*s of $\bar{y}_{BH}$, $\bar{y}_{RG}$, $\bar{y}_J$ and $\bar{y}_{DS}$ are respectively minimized for

$$\alpha_1 = [1 + \theta C_y^2 (1 - \rho^2)]^{-1} = \alpha_{BH}^* \text{, (say)} \tag{2.19}$$

$$\alpha_1 = \frac{[1+\theta R^{-1}(R^{-1}C_x^2+\rho C_y C_x)]}{[1+\theta\{C_y^2+R^{-1}(R^{-1}C_x^2+2\rho C_y C_x)\}]} = \alpha_J^*, \text{ (say)} \tag{2.20}$$

$$\left.\begin{aligned} \alpha_1 &= [1+\theta C_y^2(1-\rho^2)]^{-1} = \alpha^*_{RG_1} \\ \alpha_2 &= [RK\{1+\theta C_y^2(1-\rho^2)\}^{-1}] = \alpha^*_{RG_2} \end{aligned}\right\}, \text{ (say)} \tag{2.21}$$

$$\left.\begin{aligned} \alpha_1 &= [1+MSE(\bar{y}_\beta)(\bar{Y}-\bar{X})^{-2}]^{-1} = \alpha^*_{DS_1} \\ \alpha_2 &= -\beta\alpha^*_{DS_1} = \alpha^*_{DS_2} \end{aligned}\right\}, \text{ (say)} \tag{2.22}$$

where $K = \rho(C_y / C_x)$ and β is same as defined earlier.

Thus the resulting minimum *MSE*s of $\bar{y}_{BH}$, $\bar{y}_{RG}$, $\bar{y}_J$ and $\bar{y}_{DS}$ are respectively given by

$$MSE_{\min}(\bar{y}_{BH}) = \frac{\bar{Y}^2\theta C_y^2(1-\rho^2)}{[1+\theta C_y^2(1-\rho^2)]} = \frac{MSE(\bar{y}_\beta)}{[1+\theta C_y^2(1-\rho^2)]}, \tag{2.23}$$

$$MSE_{\min}(\bar{y}_J) = \frac{\bar{Y}^2\theta C_y^2[1+\theta R^{-2}C_x^2(1-\rho^2)]}{[1+\theta\{C_y^2+R^{-1}(R^{-1}C_x^2+2\rho C_y C_x)\}]}, \tag{2.24}$$

$$MSE_{\min}(\bar{y}_{RG}) = \frac{\bar{y}^2\theta C_y^2(1-\rho^2)}{[1+\theta C_y^2(1-\rho^2)]} = \frac{MSE(\bar{y}_\beta)}{[1+\theta C_y^2(1-\rho^2)]} = MSE_{\min}(\bar{y}_{BH}), \tag{2.25}$$

$$MSE_{\min}(\bar{y}_{DS}) = \frac{MSE(\bar{y}_\beta)}{[1+MSE(\bar{y}_\beta)(\bar{Y}-\bar{X})^{-2}]}. \tag{2.26}$$

We would like to mention here that the estimators which we have discussed earlier such as usual ratio $(\bar{y}_R)$, usual product $(\bar{y}_P)$, usual difference $(\bar{y}_d)$, usual linear regression $(\bar{y}_\beta)$, Bedi and Hajela (1984) $(\bar{y}_{BH})$, Jain (1987) $(\bar{y}_J)$, Rao (1991) $(\bar{y}_{RG})$ and Dubey and Singh (2001) $(\bar{y}_{DS})$ used only population mean $\bar{X}$ as a auxiliary information on variable *x*. In the next section we have suggested the difference -type class of estimators which is more general (i.e. some existing estimators are members of suggested class of estimators) and efficient than other exiting estimators by using several parameter of auxiliary information such as mean $(\bar{X})$, mean square (S_x^2), coefficient of variation (C_x^2) etc.

3. Proposed class of estimators

We considered the following difference-type class of estimators for estimating population mean $\bar{Y}$ of study variable *y* as

$$T = [\alpha_1\bar{y} + \alpha_2\bar{x}^* + (1-\alpha_1-\alpha_2)\bar{X}^*](\frac{\bar{X}^*}{\bar{x}^*})^{\propto}, \tag{3.1}$$

where (α_1, α_2) are suitably chosen scalars such that *MSE* of proposed class of estimators T is minimum, $\bar{x}^*(=\eta\,\bar{x}+\lambda)$, $\bar{X}^*(=\eta\,\bar{X}+\lambda)$ with (η, λ) are either constants or function of some known population parameters such as mean $(\bar{X})$, mean square (S_x^2), coefficient of variation (C_x) and coefficient of kurtosis $(\beta_2(x))$ of the auxiliary variable *x* [see Upadhyaya and Singh (1999), Khoshnevisan et al. (2007), Grover and Kaur (2011), Singh and Solanki (2011) and Sanaullah et al. (2012)] and α being constant which take finite values for designing the different estimators. It is interesting to note that some existing estimators have been founded members of proposed class of estimators T for different values of $(\alpha_1, \alpha_2, \alpha, \eta, \lambda)$, which is summarized in Table 1.

Estimators	Values of constants				
	α_1	α_2	α	η	λ
$\bar{y}$ [usual unbiased]	1	0	0	-	-
$\bar{y}_R$ [usual ratio]	1	0	1	1	0
$\bar{y}_P$ [usual product]	1	0	-1	1	0
$\bar{y}_d$ [usual difference]	1	α_2	0	-1	$\bar{X}$
$\bar{y}_\beta$ [usual linear regression]	1	β	0	-1	$\bar{X}$
$\bar{y}_{BH}$ [Bedi and Hajela (1984)]	α_1	$\alpha_1\beta$	0	-1	$\bar{X}$
$\bar{y}_J$ [Jain (1987)]	α_1	$\alpha_2=(1-\alpha_1)$	0	-1	$\bar{X}$
$\bar{y}_{RG}$ [Rao (1991)]	α_1	α_2	0	-1	$\bar{X}$
$\bar{y}_{DS}$ [Dubey and Singh (2001)]	α_1	α_2	0	1	0

Table 1: Some known members of suggested class of estimators *T*.

To obtain the bias and *MSE* of proposed class of estimators *T*, we have define $\bar{y}=\bar{Y}(1+e_0)$ and $\bar{x}=\bar{X}(1+e_1)$ such that

$$E(e_0)=E(e_1)=0$$

and to the first degree of approximation

$$E(e_0^2)=\theta C_y^2,\ E(e_1^2)=\theta C_x^2,\ E(e_0 e_1)=\theta\,\rho C_y C_x.$$

Expressing (3.1) in terms of *e*’s, we have

$$T = [\alpha_1 \bar{Y}(1+e_0) + \alpha_2 \bar{X}^* \tau e_1 + (1-\alpha_1)\bar{X}^*](1+\tau e_1)^{-\alpha}, \tag{3.2}$$

where $\tau = \eta \bar{X}(\bar{X}^*)^{-1}$.

We assume $|\tau e_1| < 1$, so that the term $(1+\tau e_1)^{-\alpha}$ is expandable. Thus by expanding the right hand side of (3.2) and neglecting the terms of e's having power greater than two, we have

$$T = \bar{Y}[(\frac{1}{R^*} - \frac{\alpha\tau}{R^*} e_1 + \frac{\alpha(\alpha+1)}{2}\frac{\tau^2}{R^*} e_1^2) + \alpha_1\{\gamma + e_0 - \gamma\alpha\tau e_1 - \alpha\tau e_0 e_1$$
$$+ \gamma \frac{\alpha(\alpha+1)}{2}\tau^2 e_1^2\} + \alpha_2\{(\frac{\tau}{R^*})e_1 - \frac{\alpha\tau^2}{R^*} e_1^2\}]$$

or

$$(T-\bar{Y}) = \bar{Y}[\alpha_1\{\gamma + e_0 - \gamma\alpha\tau e_1 - \alpha\tau e_0 e_1 + \frac{\gamma\alpha(\alpha+1)}{2}\tau^2 e_1^2\}$$
$$+ \alpha_2(\frac{\tau}{R^*})(e_1 - \alpha\tau e_1^2) - \{\gamma + (\frac{\alpha\tau}{R^*})e_1 - \frac{\alpha(\alpha+1)}{2}\frac{\tau^2}{R^*} e_1^2\}], \tag{3.3}$$

where $\gamma = (1-(R^*)^{-1})$ and $R^* = (\bar{Y}/\bar{X}^*)$.

Taking expectation on both sides of (3.3) we get the bias of T to the first degree of approximation as

$$B(T) = \bar{Y}[\alpha_1\{\gamma + \theta\alpha\tau C_x^2((\gamma(\alpha+1)\tau/2) - K)\} - \alpha_2\theta\alpha\tau^2(1-\gamma)C_x^2$$
$$-\gamma + \theta()\alpha(\alpha+1)(1-\gamma)\tau^2/2C_x^2]. \tag{3.4}$$

Squaring both sides of (3.3) and neglecting terms of e's having power greater than two we have

$$(T-\bar{Y})^2 = \bar{Y}^2[\alpha_1^2\{\gamma^2 + e_0^2 + \gamma^2\alpha^2\tau^2 2e_1^2 + 2\gamma e_0 - 2\gamma^2\alpha\tau e_1$$
$$-2\gamma\alpha\tau e_0 e_1 - 2\gamma\alpha\tau e_0 e_1 + \gamma^2\alpha(\alpha+1)\tau^2 e_1^2\}$$
$$+\alpha_2^2(\tau^2/R^{*^2})e_1^2 + \{\gamma^2(\alpha^2\tau^2/R^{*^2})e_1^2 + (2\gamma\alpha\tau/R^*)e_1$$
$$-\gamma\alpha(\alpha+1)(\tau^2/R^*)e_1^2\} + 2\alpha_1\alpha_2(\tau/R^*)\{\gamma(e_1 - \alpha\tau e_1^2)$$
$$+e_0 e_1 - \gamma\alpha\tau e_1^2\} - 2\alpha_1\{\gamma^2 + \gamma e_0 - \gamma^2\alpha\tau e_1 - \gamma\alpha\tau e_0 e_1$$
$$+(\gamma^2\alpha(\alpha+1)/2)\tau^2 e_1^2 + (\alpha\tau\gamma/R^*)e_1 + (\alpha\tau/R^*)e_0 e_1$$
$$-(\gamma\alpha^2\tau^2/R^*)e_1^2 - (\gamma\alpha(\alpha+1)\tau^2/2R^*)e_1^2\}$$
$$-2\alpha_2(\tau/R^*)\{\gamma e_1 - \gamma\alpha\tau e_1^2 + (\alpha\tau/R^*)e_1^2\}]. \tag{3.5}$$

Taking expectation on both sides of (3.5), we get the *MSE* of proposed class of estimator T, to the first degree of approximation as

$$MSE(T) = \bar{Y}^2[A + \alpha_1^2 A_1 + \alpha_2^2 A_2 + 2\alpha_1\alpha_2 A_3 - 2\alpha_1 A_4 - 2\alpha_2 A_5], \tag{3.6}$$

where

$$A = [\gamma^2 + \theta\alpha\tau^2 C_x^2 (1-\gamma)\{\alpha(1-\gamma) - \gamma(\alpha+1)\}],$$
$$A_1 = [\gamma^2 + \theta\{C_y^2 + \gamma\alpha\tau\, C_x^2 (\gamma\tau(2\alpha+1) - 4K)\}],$$
$$A_2 = [\tau^2 (1-\gamma)^2 \theta C_x^2],$$
$$A_3 = [\theta\tau C_x^2 (1-\gamma)(K - 2\gamma\alpha\tau)],$$
$$A_4 = [\gamma^2 + \theta\alpha\tau C_x^2 \{K(1-2\gamma) - \alpha\tau\gamma(1-\gamma) - (\tau(\alpha+1)\gamma(1-2\gamma)/2)\}],$$
$$A_5 = [\alpha\tau^2 (1-\gamma)(1-2\gamma)\theta C_x^2].$$

Differentiating (3.6) with respect to (α_1, α_2) and equating them to zero, we get the following normal equations

$$\begin{bmatrix} A_1 & A_3 \\ A_3 & A_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} A_4 \\ A_5 \end{bmatrix}. \qquad (3.7)$$

Solving (3.7) we get the optimum values of α_1 and α_2 respectively as

$$\alpha_1 = [\frac{(A_2 A_4 - A_3 A_5)}{(A_1 A_2 - A_3^2)}] = \alpha_1^*, \text{(say)} \qquad (3.8)$$

$$\alpha_2 = [\frac{(A_1 A_5 - A_3 A_4)}{(A_1 A_2 - A_3^2)}] = \alpha_2^*, \text{(say)}. \qquad (3.9)$$

Substituting (3.8) and (3.9) in (3.4) and (3.6),we get the minimum bias and *MSE* of proposed class of estimators *T* respectively as

$$B_o(T) = \bar{Y}[-\gamma + (\alpha(\alpha+1)(1-\gamma)\tau^2/2)C_x^2 + \alpha_1^*\{\gamma + \theta\alpha\tau C_x^2((\gamma(\alpha+1)\tau/2) - K)\}$$
$$- \theta\alpha\tau^2(1-\gamma)\alpha_2^* C_x^2], \qquad (3.10)$$

$$MSE_{\min}(T) = \bar{Y}^2[A - \frac{(A_2 A_4^2 - 2A_3 A_4 A_5 + A_1 A_5^2)}{(A_1 A_2 - A_3^2)}]. \qquad (3.11)$$

Thus we established the following theorem.

Theorem 3.1: To the first degree of approximation,

$$MSE(T) \geq \bar{Y}^2[A - \frac{(A_2 A_4^2 - 2A_3 A_4 A_5 + A_1 A_5^2)}{(A_1 A_2 - A_3^2)}]$$

with equality holding if

$$\alpha_1 = \alpha_1^* \text{ and } \alpha_2 = \alpha_2^*.$$

4. Efficiency comparisons

In this section we have made theoretical comparisons between some estimators/classes of estimators, which we have discussed earlier.

From (2.1), (2.4), (2.5), and (2.10) we have

$$Var(\bar{y}) - MSE(\bar{y}_\beta)\ [MSE_{\min}(\bar{y}_d)] = \theta\bar{Y}^2 C_y^2 \rho^2 \geq 0, \qquad (4.1)$$

$$MSE(\bar{y}_R) - MSE(\bar{y}_\beta)\ [MSE_{\min}(\bar{y}_d)] = \theta\bar{Y}^2 C_y^2 (1-K)^2 \geq 0, \qquad (4.2)$$

$$MSE(\bar{y}_P) - MSE(\bar{y}_\beta)\ [MSE_{\min}(\bar{y}_d)\ ] = \theta \bar{Y}^2 C_y^2 (1+K)^2 \geq 0\ . \tag{4.3}$$

From (4.1)-(4.3), it is observed that the usual regression estimator $\bar{y}_\beta$ and usual difference estimator $\bar{y}_d$ (at their optimum condition) are more efficient than

(i) the usual unbiased estimator $\bar{y}$, provided $\rho \neq 0$. For $\rho = 0$, the estimators $\bar{y}$, $\bar{y}_\beta$ and $\bar{y}_d$ are equally efficient.

(ii) the usual ratio estimator $\bar{y}_R$, provided $K \neq 1$ (or $R \neq \beta$). In case $K = 1$ (or $R = \beta$), the estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_\beta$ and $\bar{y}_d$ are equally efficient.

(iii) the usual product estimator $\bar{y}_P$, provided $K \neq -1$ (or $R \neq -\beta$). In case $K = -1$ (or $R = -\beta$) the estimators $\bar{y}$, $\bar{y}_P$, $\bar{y}_\beta$ and $\bar{y}_d$ are equally efficient.

From (2.10), (2.23), (2.25), and (2.26), we have

$$MSE(\bar{y}_\beta)\ [\text{or } MSE_{\min}(\bar{y}_d)\ ] - MSE_{\min}(\bar{y}_{BH})\ [\text{or } MSE_{\min}(\bar{y}_{RG})\ ]$$
$$= \frac{MSE_{\min}(\bar{y}_{BH}) MSE(\bar{y}_\beta)}{\bar{Y}^2} \geq 0\ , \tag{4.4}$$

$$MSE(\bar{y}_\beta)\ [\text{or } MSE_{\min}(\bar{y}_d)\ ] - MSE_{\min}(\bar{y}_{DS}) = [MSE(\bar{y}_\beta)]^2 (\bar{Y} - \bar{X})^{-2} \geq 0\ . \tag{4.5}$$

It is observed from (4.4) and (4.5) that the estimators $\bar{y}_{BH}$ [Bedi and Hajela (1984)], $\bar{y}_{RG}$ [Rao (1991)] and $\bar{y}_{DS}$ [Dubey and Singh (2001)] are more efficient [at its optimum conditions] than the usual regression estimator $\bar{y}_\beta$ and hence more efficient than the estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_P$ and $\bar{y}_d$.

From (2.23), (2.24), (2.25) and (2.26), we have

$$MSE_{\min}(\bar{y}_{DS}) < MSE_{\min}(\bar{y}_{BH})\ [\text{or } MSE_{\min}(\bar{y}_{RG})\ ] \text{ if } 1 < 2R\ , \tag{4.6}$$

$$MSE_{\min}(\bar{y}_{DS}) < MSE_{\min}(\bar{y}_J) \text{ if } \rho^2 > \frac{2\theta K C_x^2 R^{-1}}{2\theta K C_x^2 R^{-1} + 1}\ . \tag{4.7}$$

Remark 4.1: The estimators $\bar{y}, \bar{y}_R, \bar{y}_P, \bar{y}_d, \bar{y}_\beta, \bar{y}_{BH}, \bar{y}_J, \bar{y}_{RG}$ and $\bar{y}_{DS}$ are members of suggested class of estimators *T* (see Table 1), therefore the *Var*/*MSE*s of estimators $\bar{y}, \bar{y}_R, \bar{y}_P, \bar{y}_d, \bar{y}_\beta, \bar{y}_{BH}, \bar{y}_J, \bar{y}_{RG}$ and $\bar{y}_{DS}$ are always greater than or equal to the *MSE* of class of estimators *T*.

In addition some new members of suggested difference-type class of estimators *T* [for different values of (α, η, λ)] have been summarized in Table 2.

Estimators	Values of constants		
	α	η	λ
$T_1 = [\alpha_1 \bar{y} + \alpha_2 \bar{x} + (1-\alpha_1-\alpha_2)\bar{X}](\frac{\bar{X}}{\bar{x}})$	1	1	0
$T_2 = [\alpha_1 \bar{y} + \alpha_2 (\bar{x}+\rho) + (1-\alpha_1-\alpha_2)(\bar{X}+\rho)](\frac{\bar{X}+\rho}{\bar{x}+\rho})$	1	1	ρ
$T_3 = [\alpha_1 \bar{y} + \alpha_2 (\bar{x}+C_x^2) + (1-\alpha_1-\alpha_2)(\bar{X}+C_x^2)](\frac{\bar{X}+C_x^2}{\bar{x}+C_x^2})$	1	1	C_x^2
$T_4 = [\alpha_1 \bar{y} + \alpha_2 (\rho\bar{x}+\bar{X}) + (1-\alpha_1-\alpha_2)(\rho\bar{X}+\bar{X})](\frac{\rho\bar{X}+\bar{X}}{\rho\bar{x}+\bar{X}})$	1	ρ	$\bar{X}$
$T_5 = [\alpha_1 \bar{y} + \alpha_2 (\bar{x}+\bar{X}) + 2(1-\alpha_1-\alpha_2)\bar{X}](\frac{\bar{x}+\bar{X}}{2\bar{X}})$	-1	1	$\bar{X}$
$T_6 = [\alpha_1 \bar{y} + \alpha_2 (1-\bar{x}C_x^2) + (1-\alpha_1-\alpha_2)(1-\bar{X}C_x^2)](\frac{1-\bar{x}C_x^2}{1-\bar{X}C_x^2})$	-1	$-C_x^2$	1
$T_7 = [\alpha_1 \bar{y} + \alpha_2 (C_x^2-\bar{x}) + (1-\alpha_1-\alpha_2)(C_x^2-\bar{X})](\frac{C_x^2-\bar{x}}{C_x^2-\bar{X}})$	-1	-1	C_x^2
$T_8 = [\alpha_1 \bar{y} + \alpha_2 (\rho\bar{x}+\bar{X}) + (1-\alpha_1-\alpha_2)(\rho\bar{X}+\bar{X})](\frac{\rho\bar{x}+\bar{X}}{\rho\bar{X}+\bar{X}})$	-1	ρ	$\bar{X}$

Table 2: Some new members of suggested class of estimators *T*.

5. Empirical study

To evaluate the performance of estimators T_i, (i = 1, 2, …, 8) which are members of the suggested class of estimators *T*, over other competitors, we have considered three population data sets. The descriptions of population data sets are as follows.

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

y: Number of inhabitants in 1930.

x: Number of inhabitants in 1920.

$N = 196$, $n = 49$, $\bar{Y} = 404.0955$, $\bar{X} = 103.1$, $C_y = 0.9634$, $C_x = 1.0126$, $\rho = 0.982$.

Population II: [Das (1988)]

y: Number of agricultural labourers in 1961.

x: Number of agricultural labourers in 1971.

$N = 278,\ n = 30,\ \overline{Y} = 39.0680,\ \overline{X} = 25.1104,\ C_y = 1.445,\ C_x = 1.6198,\ \rho = 0.7213.$

Population III: [Singh, S. (2003)]

y: Amount of real estate farm loans in different states during 1997.

x: Amount of non real estate farm loans in different states during 1997.

$N = 50,\ n = 8,\ \overline{Y} = 2573.4794,\ \overline{X} = 878.16,\ C_y = 1.0529,\ C_x = 1.2352,\ \rho = 0.8038$.

We have computed the percent relative efficiencies (*PRE*s) of different estimators t_o, with respect to the usual unbiased estimator $\bar{y}$ as

$$PRE(t_0, \bar{y}) = \frac{Var(\bar{y})}{MSE / MSE_{\min}(t_o)} * 100, \quad (5.1)$$

and the results are displayed in Table 3.

Estimator (t_o)	$PRE(t_o, \bar{y})$		
	Population – I	Population – II	Population – III
$\bar{y}$	100.00	100.00	100.00
$\bar{y}_R$	2731.37	156.38	203.95
$\bar{y}_P$	26.50	25.82	23.46
$\bar{y}_d$, $\bar{y}_\beta$	2803.00	208.45	282.56
$\bar{y}_{BH}$, $\bar{y}_{RG}$	2804.57	214.66	294.20
$\bar{y}_J$	102.41	114.12	120.20
$\bar{y}_{DS}$	2805.83	257.10	309.38
T_1	2805.87	261.21	314.50
T_2	2805.89	266.85	314.52
T_3	2805.89	287.27	314.54
T_4	2809.45	765.14	363.89
T_5	2832.59	343.97	333.04
T_6	**4439.95**	364.21	**1445.08**
T_7	4426.54	292.13	1173.17
T_8	2831.27	**1077.72**	319.79

Table 3: *PRE*s of different estimators with respect to $\bar{y}$.

*bold letters indicate the largest *PRE* in relevant population.

It is observed from Table 3 that

(i) The performance of the usual linear regression estimator $\bar{y}_\beta$ and usual difference estimators $\bar{y}_d$ (at their optimum condition) are equal as well as better than the usual unbiased estimator $\bar{y}$, usual ratio estimator $\bar{y}_R$ [Cochran (1940)] and usual product estimator $\bar{y}_P$ [Robson (1957) and Murthy (1967)].

(ii) The estimators T_i, (i = 1, 2, …, 8) [members of the suggested class of estimators T], $\bar{y}_{BH}$ [Bedi and Hajela (1984)], $\bar{y}_J$ [Jain (1987)], $\bar{y}_{RG}$ [Rao (1991)] and $\bar{y}_{DS}$ [Dubey and Singh (2001)] are more efficient than the usual linear regression estimator $\bar{y}_\beta$ and hence more efficient than the estimators $\bar{y}, \bar{y}_R, \bar{y}_P$ and $\bar{y}_d$.

(iii) The performance of the estimators $\bar{y}_{BH}, \bar{y}_{RG}, \bar{y}_J$ and $\bar{y}_{DS}$ are inferior [i.e. having smaller *PRE*s] to the estimators T_i, (i = 1, 2, …, 8) which are members of suggested class of estimators *T*.

(iv) The estimators T_6 [utilizes the information on $\bar{X}$ and C_x^2] is best in the sense of having largest *PRE* among all the estimators discussed here in the population data sets I and III , while the estimators T_8 [utilizes the information on $\bar{X}$ and ρ] performed better than all the estimators in the population data set II.

6. Conclusion

In the present study, we have suggested the difference-type class of estimators of the population mean of a study variable when information on an auxiliary variable is known in advance. The asymptotic bias and mean square error formulae of suggested class of estimators have been obtained. The asymptotic optimum estimator in the suggested class has been identified with its properties. Some traditional methods of estimation of population mean such as usual unbiased, ratio, product, difference, linear regression, and methods proposed by Bedi and Hajela (1984), Jain (1987), Rao (1991), and Dubey and Singh (2001) have been found members of suggested class of estimators. Thus the present study unifies several results. In addition, some new members of suggested class of estimators have been also generated. An empirical study is carried out to throw light on the performance of the suggested method over already exiting methods. Further empirical studies carried out in paper clearly reflect the usefulness of the suggested method in practice. However this conclusion cannot be extrapolated due to limited empirical study.

Acknowledgement

The authors are highly grateful to the referees for their encouraging suggestions which were very much useful in improving the quality of the paper.

References

1. Bedi, P.K. and Hajela, D. (1984). An estimator for population mean utilizing known coefficient of variation, J. Statist. Res., 18, p. 29-38.

2. Cochran, W.G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce, J. Agri. Soc., 30, p. 262-275.
3. Cochran, W.G. (1977). Sampling Techniques. 3rd ed., John Wiley and Sons, New York, USA.
4. Das, A.K. (1988). Contribution to the theory of sampling strategies based on auxiliary information, Ph.D. thesis submitted to B.C.K.V., Mohanpur, Nadia, West Bengal, India.
5. Dubey, V. and Singh, S.K. (2001). An improved regression estimator for estimating population mean, J. Ind. Soc. Agri. Statist., 54, p. 179-183.
6. Grover, L.K. and Kaur, P. (2011). An improved exponential estimator of finite population mean in simple random sampling using an auxiliary attribute, Appl. Math. Comp., 218(7), p. 3093-3099.
7. Hansen, M.N., Hurwitz, W.N. and Madow, W.G. (1953). Sample Survey Methods and Theory, John Wiley and Sons, Now York, USA.
8. Jain, R.K. (1987). Properties of estimators in simple random sampling using auxiliary variable, Metron, 45(1-2), p. 265-271.
9. Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s), Far East J. Theo. Statist. 22(2), p. 181-191.
10. Murthy, M.N. (1964). Product method of estimation, Sankhya, 26, p. 69-74.
11. Rao T.J. (1991). On certain methods of improving ratio and regression estimators, Commun. Statist. Theo. Meth., 20(10), p. 3325-3340.
12. Robson, D.S. (1957). Application of multivariate polykays to the theory of unbiased ratio type estimators, J. Amer. Statist. Assoc., 52, p. 511-522.
13. Sanaullah, A., Khan, H., Ali, H.A. and Singh, R. (2012). Improved exponential ratio-type estimators in survey sampling, J. Rel. Statist. Stud., 5(2), p. 119-132.
14. Singh, H.P. and Solanki, R.S. (2011). Generalized ratio and product methods of estimation in survey sampling, Pak. J. Statist. Oper. Res., 7(2), p. 245-264.
15. Singh, S. (2003). Advanced Sampling Theory with Application How Michael "selected" Amy. Kluwer Academic Publishers, The Netherlands.
16. Upadhyaya, L.N. and Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population mean, Biometrical J., 41, p. 627-636.