

A General Class of Dual to Ratio Estimators

Housila P. Singh¹ and Pragati Nigam^{2*}

* Corresponding Author



1. School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India, hpsujn@gmail.com
2. School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India, pragatinigam1@gmail.com

Abstract

In this paper we have considered the problem of estimating the population mean using auxiliary information in sample surveys. A class of dual to ratio estimators has been defined. Exact expressions for bias and mean squared error of the suggested class of dual to ratio estimators have been obtained. In particular, properties of some members of the proposed class of dual to ratio estimators have been discussed. It has been shown that the proposed class of estimators is more efficient than the sample mean, ratio estimator, dual to ratio estimator and some members of the suggested class of estimators in some realistic conditions. Some numerical illustrations are given in support of the present study.

Keywords: Study variable, Auxiliary variable, Population mean, Bias, Mean squared error.

Mathematical Subject Classification: 62D05

1. Introduction

The use of ratio method of estimation is quite effective if the correlation between the study variable y and the auxiliary variable x is positive (high). On the other hand, if this correlation is negative (high) product method of estimation is employed for estimating population mean \bar{Y} of the study variable y . It is to be mentioned that the ratio estimator suffers with a drawback that it does not provides exact bias and mean squared error while the product method of estimation provide the exact bias and mean squared error. But in practice positive correlation between the two variables (y, x) are generally encountered, while negative correlation situation is not much as compared to positive correlation.

Srivenkataramana (1980) and Bandyopadhyay (1980) have advocated the use of product method of estimation in case of positive correlation using simple transformation which induce the negative correlation even if the correlation between the two variables (y, x) is positive. Later many other authors have worked on the ratio and product methods of estimations such as Kadilar and Cingi (2004), Grover and Kaur (2011), Singh et al (2015), Singh and Yadav (2018), Pal et al (2018, 2019) etc.

Consider a finite population $U = (U_1, U_2, \dots, U_N)$ of N units on which the study variable y and the auxiliary variable x are defined. The units are identifiable in the sense that they can be uniquely labeled from 1 to N and the label of each unit is known. Let the values of the variables (y, x) for U_i be $(y_i, x_i, i = 1, 2, \dots, N)$. Let (\bar{Y}, \bar{X}) be the population means of the variables (y, x) respectively. Suppose that a simple random sample of size n is drawn without replacement from U for estimating the population mean \bar{Y} of the study variable y . Let (\bar{y}, \bar{x}) be the sample means of (y, x) respectively based on n observations drawn from the population U . Using the transformation:

$$x_i^* = (1 + g)\bar{X} - gx_i, i = 1, 2, \dots, N \quad (1.1)$$

Srivenkataramana (1980) and Bandyopadhyay (1980) proposed a dual to ratio estimator for \bar{Y} as

$$\bar{y}_{SB} = \bar{y} \frac{\bar{x}^*}{\bar{X}} = \bar{y} \frac{(\bar{X} - f\bar{x})}{(1-f)\bar{X}}, \quad (1.2)$$

where $f = \frac{n}{N}$, $g = \frac{n}{N-n}$, $\bar{x}^* = \frac{(N\bar{X} - n\bar{x})}{(N-n)}$ is the mean of the unobserved units in the population U and \bar{X} is the known population mean of the auxiliary variable x . The exact bias of the estimator \bar{y}_{SB} is given by

$$B(\bar{y}_{SB}) = -\left(\frac{1-f}{n}\right)g \frac{S_{yx}}{\bar{X}}, \quad (1.3)$$

where $S_{yx} = \frac{1}{(N-1)} \sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X})$.

For exact mean squared error of the estimator \bar{y}_{SB} the reader is referred to Srivenkataramana (1980).

To the first degree of approximation, the mean squared error (MSE) of the estimator \bar{y}_{SB} is given by

$$MSE(\bar{y}_{SB}) = \left(\frac{1-f}{n}\right)^2 \bar{Y}^2 [C_y^2 + gC_x^2(g-2k)], \quad (1.4)$$

where $k = \rho \frac{C_y}{C_x}$, $C_y = \frac{S_y}{\bar{Y}}$, $C_x = \frac{S_x}{\bar{X}}$, $\rho = \frac{S_{yx}}{(S_y S_x)}$,

$S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2$ and $S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$.

Under simple random sampling without replacement ($SRSWOR$) the variance/ MSE of the usual unbiased estimator \bar{y} of population mean \bar{Y} is given by

$$MSE(\bar{y}) = \left(\frac{1-f}{n}\right) \bar{Y}^2 C_y^2 = \left(\frac{1-f}{n}\right) S_y^2. \quad (1.5)$$

For estimating the population mean \bar{Y} , when the correlation between the two variables (y, x) is positive and population mean \bar{X} of the auxiliary variable x is known, the classical ratio estimator is defined by

$$\bar{y}_R = \bar{y} \frac{\bar{X}}{\bar{x}}. \quad (1.6)$$

To the first degree of approximation, the bias and MSE of the ratio estimator \bar{y}_R are respectively given by

$$B(\bar{y}_R) = \left(\frac{1-f}{n}\right) \bar{Y} C_x^2 (1-k), \quad (1.7)$$

$$MSE(\bar{y}_R) = \left(\frac{1-f}{n}\right) \bar{Y}^2 [C_y^2 + C_x^2(1-2k)], \quad (1.8)$$

In the present paper we have suggested a generalized version of the dual to ratio estimator \bar{y}_{SB} along with its properties. Numerical illustration is given in support of the present study.

2. Generalized Version of the Dual to Ratio Estimator \bar{y}_{SB}

Keeping in view the form of the dual to ratio estimator \bar{y}_{SB} at (1.2), we define its generalized version for population mean \bar{Y} as

$$\bar{y}_H = \bar{y} \frac{(\bar{X} - b\bar{x})}{(1-b)\bar{X}} \quad (2.1)$$

where 'b' is suitable chosen scalar such that $b \in [0, 1)$. We note that

(i) for $b=0$, $\bar{y}_H \rightarrow \bar{y}$

$$(ii) \quad \text{for } b = f = \frac{n}{N}, \bar{y}_H \rightarrow \bar{y}_{SB} = \bar{y} \frac{(\bar{X} - f\bar{x})}{(1-f)\bar{X}}$$

which is due to Srivenkataramana (1980) and Bandyopadhyay (1980).

$$(iii) \quad \text{for } b = \frac{f}{1+f},$$

$$\bar{y}_H \rightarrow \bar{y}_{H(1)} = \bar{y} \frac{\{(1+f)\bar{X} - f\bar{x}\}}{\bar{X}}, \quad (2.2)$$

$$(iv) \quad \text{for } b = \frac{(1-f)}{(1+f)},$$

$$\bar{y}_H \rightarrow \bar{y}_{H(2)} = \bar{y} \frac{\{(1+f)\bar{X} - (1-f)\bar{x}\}}{2f\bar{X}}, \quad (2.3)$$

$$(v) \quad \text{for } b = \frac{1}{2} \Rightarrow (1-b) = \frac{1}{2},$$

$$\bar{y}_H \rightarrow \bar{y}_{H(3)} = \bar{y} \frac{(2\bar{X} - \bar{x})}{\bar{X}}, \quad (2.4)$$

$$(vi) \quad \text{for } b = \frac{1}{(1+f)} \Rightarrow (1-b) = \frac{f}{(1+f)},$$

$$\bar{y}_H \rightarrow \bar{y}_{H(4)} = \bar{y} \frac{\{(1+f)\bar{X} - \bar{x}\}}{f\bar{X}}. \quad (2.5)$$

Many more acceptable estimators can be generated from the proposed estimator \bar{y}_H for different choices of b .

The exact bias and MSE of the suggested estimator \bar{y}_H are respectively given by

$$B(\bar{y}_H) = -\frac{(1-f)}{n} \frac{b}{(1-b)} \frac{S_{yx}}{\bar{X}} \quad (2.6)$$

$$MSE(\bar{y}_H) = \frac{(1-f)}{n} \bar{Y}^2 \left[C_y^2 + \frac{b}{(1-b)} C_x^2 \left\{ \frac{b}{(1-b)} - 2k \right\} \right] + \bar{Y}^2 \left[\left\{ 2 \frac{b^2}{(1-b)^2} V_{12} - 2 \frac{b}{(1-b)} V_{21} + \frac{b^2}{(1-b)^2} V_{22} \right\} \right], \quad (2.7)$$

where V_{ij} are the relative central moments defined by $V_{ij} = E \left\{ \left(\frac{\bar{y} - \bar{Y}}{\bar{Y}} \right)^i \left(\frac{\bar{x} - \bar{X}}{\bar{X}} \right)^j \right\}$, (i, j) being non-negative integers.

As remarked by Murthy (1967, pp.380-381) that the V_{ij} with $i + j > 2$ are generally small, so we neglect the terms with $V_{ij}, i + j > 2$. Thus to the first degree of approximation, the MSE of the proposed estimator \bar{y}_H is given by

$$MSE(\bar{y}_H) = \frac{(1-f)}{n} \bar{Y}^2 \left[C_y^2 + \frac{b}{(1-b)} C_x^2 \left\{ \frac{b}{(1-b)} - 2k \right\} \right] \quad (2.8)$$

The biases and mean squared errors of the estimators belonging to the suggested estimator \bar{y}_H can be easily obtained from (2.6), (2.7) and (2.8) just by putting different values of the scalar ' b '.

3. Efficiency Comparison

In this section we have obtained the regions of preferences in which the suggested estimator \bar{y}_H is better than the usual unbiased estimator \bar{y} , ratio estimator \bar{y}_R and dual to ratio estimator \bar{y}_{SB} .

From (1.5) and (2.8) we have

$$MSE(\bar{y}) - MSE(\bar{y}_H) = \frac{(1-f)}{n} \frac{b}{(1-b)} \bar{Y}^2 C_x^2 \left(2k - \frac{b}{1-b} \right)$$

which is positive if

$$2k - \frac{b}{1-b} > 0$$

i.e. if

$$0 \leq b < \frac{2k}{2k+1}.$$

From (1.8) and (2.8) we have

$$MSE(\bar{y}_R) - MSE(\bar{y}_H) = \frac{(1-f)}{n} \bar{Y}^2 C_x^2 \left(1 - \frac{b}{1-b}\right) \left(\frac{b}{1-b} - 2k + 1\right)$$

which is non-negative if

$$\left. \begin{array}{l} \text{either } \left(1 - \frac{1}{2k}\right) < b < \frac{1}{2} \\ \text{or } \frac{1}{2} < b < \left(1 - \frac{1}{2k}\right) \end{array} \right\} \quad (3.2)$$

or equivalently,

$$\min.\left\{\frac{1}{2}, \left(1 - \frac{1}{2k}\right)\right\} < b < \max.\left\{\frac{1}{2}, \left(1 - \frac{1}{2k}\right)\right\} \quad (3.3)$$

Further from (1.4) and (2.8) we have

$$MSE(\bar{y}_{SB}) - MSE(\bar{y}_H) = \left(\frac{1-f}{n}\right) \bar{Y}^2 C_x^2 \left(g - \frac{b}{1-b}\right) \left(g + \frac{b}{1-b} - 2k\right)$$

which is positive if

$$\left(g - \frac{b}{1-b}\right) \left(g + \frac{b}{1-b} - 2k\right) > 0$$

$$\left. \begin{array}{l} \text{either } \frac{2(k-g)}{(2k-g+1)} < b < \frac{1}{(1+g)} \\ \text{or } \frac{1}{(1+g)} < b < \frac{2(k-g)}{(2k-g+1)} \end{array} \right\} \quad (3.4)$$

i.e. if

or equivalently

$$\min.\left\{\frac{1}{(1+g)}, \frac{2(k-g)}{(2k-g+1)}\right\} < b < \max.\left\{\frac{1}{(1+g)}, \frac{2(k-g)}{(2k-g+1)}\right\} \quad (3.5)$$

Thus we state the following theorem.

Theorem 3.1: The proposed estimator \bar{y}_H is more efficient than:

(i) the usual unbiased estimator \bar{y} if

$$b \in \left[0, \frac{2k}{2k+1}\right),$$

(ii) the usual ratio estimator \bar{y}_R if

$$\min.\left\{\frac{1}{2}, \left(1 - \frac{1}{2k}\right)\right\} < b < \max.\left\{\frac{1}{2}, \left(1 - \frac{1}{2k}\right)\right\},$$

(iii) the dual to ratio estimator \bar{y}_{SB} if

$$\min.\left\{\frac{1}{(1+g)}, \frac{2(k-g)}{(2k-g+1)}\right\} < b < \max.\left\{\frac{1}{(1+g)}, \frac{2(k-g)}{(2k-g+1)}\right\}.$$

Putting $b = \frac{f}{(1+f)} = \frac{g}{(1+2g)}$, $b = \frac{(1-f)}{(1+f)} = \frac{1}{(1+2g)}$, $b = \frac{1}{2}$, and $b = \frac{(1+g)}{(1+2g)} = \frac{1}{(1+f)}$ in (2.8), we get the MSEs of the estimators $\bar{y}_{H(1)}$, $\bar{y}_{H(2)}$, $\bar{y}_{H(3)}$ and $\bar{y}_{H(4)}$ to the first degree of approximation respectively as

$$MSE(\bar{y}_{H(1)}) = \frac{(1-f)}{n} \bar{Y}^2 \left[C_y^2 + \frac{g}{(1+g)} C_x^2 \left\{ \frac{g}{(1+g)} - 2k \right\} \right], \quad (3.6)$$

$$MSE(\bar{y}_{H(2)}) = \frac{(1-f)}{n} \bar{Y}^2 \left[C_y^2 + \frac{1}{2g} C_x^2 \left\{ \frac{1}{2g} - 2k \right\} \right], \quad (3.7)$$

$$MSE(\bar{y}_{H(3)}) = \frac{(1-f)}{n} \bar{Y}^2 [C_y^2 + C_x^2(1-2k)] = MSE(\bar{y}_R), \quad (3.8)$$

$$MSE(\bar{y}_{H(4)}) = \frac{(1-f)}{n} \bar{Y}^2 \left[C_y^2 + \frac{(1+g)}{g} C_x^2 \left\{ \frac{(1+g)}{g} - 2k \right\} \right]. \quad (3.9)$$

Corollary 3.1: The proposed estimator $\bar{y}_{H(1)}$ is more efficient than:

- (i) the usual unbiased estimator \bar{y} if

$$k > \frac{g}{2(1+g)},$$

- (ii) the ordinary ratio estimator \bar{y}_R if

$$k < \frac{1}{2} \left\{ 1 + \frac{g}{1+g} \right\},$$

- (iii) the dual to ratio estimator \bar{y}_{SB} if

$$k < \frac{g}{2} \left\{ 1 + \frac{1}{1+g} \right\}.$$

Corollary 3.2: The proposed estimator $\bar{y}_{H(2)}$ is more efficient than:

- (i) the usual unbiased estimator \bar{y} if

$$k > \frac{1}{4g},$$

- (ii) the ordinary ratio estimator \bar{y}_R if

$$k < \frac{1}{2} \left\{ 1 + \frac{1}{2g} \right\},$$

- (iii) the dual to ratio estimator \bar{y}_{SB} if

$$\left. \begin{array}{l} \text{either } k < \frac{(4g^2+1)}{4g}, g > \frac{1}{2} \\ \text{or } k > \frac{(4g^2+1)}{4g}, g < \frac{1}{2} \end{array} \right\}$$

Corollary 3.3: The proposed estimator $\bar{y}_{H(3)}$:

- (i) is more efficient than, the usual unbiased estimator \bar{y} if

$$k > \frac{1}{2},$$

- (ii) and the ratio estimator \bar{y}_R are equally efficient.

- (iii) is more efficient than the dual to ratio estimator \bar{y}_{SB} if

$$\left. \begin{array}{l} \text{either } k < \frac{(1+g)}{2}, g < 1 \\ \text{or } k > \frac{(1+g)}{2}, g > 1 \end{array} \right\}$$

Corollary 3.4: The proposed estimator $\bar{y}_{H(4)}$ is more efficient than:

(i) the usual unbiased estimator \bar{y} if

$$k > \frac{(1+g)}{2g},$$

(ii) the ratio estimator \bar{y}_R if

$$k > \left\{ 1 + \frac{1}{2g} \right\},$$

(iii) the dual to ratio estimator \bar{y}_{SB} if

$$\left. \begin{array}{l} \text{either } k < \frac{(g^2 + g + 1)}{2g}, (g^2 + g + 1) > 0 \\ \text{or } k > \frac{(g^2 + g + 1)}{2g}, (g^2 + g + 1) < 0 \end{array} \right\}$$

4. Optimum Choice of Scalar ‘b’

Differentiating (2.8) with respect to b partially and equating to zero, we get the optimum value of b as

$$b = \frac{k}{(1+k)} = b_{opt} \text{ (say).} \quad (4.1)$$

Substitution of b_{opt} in place of b in (2.1) yields the optimum estimator

$$\bar{y}_{HO} = \bar{y} \frac{\{\bar{X} + k(\bar{X} - \bar{x})\}}{\bar{X}} \quad (4.2)$$

for the population mean \bar{Y} of the study variable y .

Putting (4.1) in (2.8) we get the minimum MSE of the proposed estimator \bar{y}_H (or the MSE of the optimum estimator \bar{y}_{HO})

$$MSE(\bar{y}_{HO}) = \frac{(1-f)}{n} S_y^2 (1 - \rho^2). \quad (4.3)$$

Thus we state the following theorem.

Theorem 4.1: To the first degree of approximation,

$$MSE(\bar{y}_H) \geq \frac{(1-f)}{n} S_y^2 (1 - \rho^2)$$

with equality holding if

$$b = \frac{k}{(1+k)}.$$

It is to be noted that the optimum estimator \bar{y}_{HO} at (4.2) can be used in practice only when the exact value of k is known. The exact value of k is rarely available. However in repeated surveys or studies based on multiphase sampling, where information regarding the same variables is collected on several occasions, it is possible to guess accurately the values of certain parameters. This problem has been discussed among others by Murthy (1967, pp.96-99) and Reddy (1978). Hence we assume that k can be guessed quite accurately. In turn a good approximation of b_0

for b_{opt} can be obtained. We judge below to what extent b_0 may deviate from b_{opt} and yields an estimator more efficient than \bar{y}_R or \bar{y} .

5. Allowable Departure from Optimum

The optimum value b_{opt} of b depends on the value of k which is a function of unknown population parameters such as ρ, C_y and C_x . However, the values of the parameters ρ, C_y and C_x can be guessed quite accurately from the past data or experience gathered in due course of time. Hence k can be guessed.

Let k_0 be the guessed value of k such that

$$k_0 = k(1 + \varepsilon), \quad (5.1)$$

where ε indicates the error in the guessed value k_0 .

Putting (5.1) in (4.2) we get the resulting estimator

$$\begin{aligned} \tilde{y}_{HO} &= \bar{y} \frac{\{\bar{X} + k(1 + \varepsilon)(\bar{X} - \bar{x})\}}{\bar{X}} \\ &= \bar{y} \left\{ 1 - k \frac{(\bar{x} - \bar{X})}{\bar{X}} (1 + \varepsilon) \right\} \end{aligned} \quad (5.2)$$

for population mean \bar{Y} .

Writing (5.2) in terms of e_0 and e_1 we have

$$\begin{aligned} \tilde{y}_{HO} &= \bar{Y}(1 + e_0)\{1 - ke_1(1 + \varepsilon)\} \\ &= \bar{Y}[1 + e_0 - ke_1(1 + \varepsilon) - ke_0e_1(1 + \varepsilon)] \end{aligned}$$

or

$$(\tilde{y}_{HO} - \bar{Y}) = \bar{Y}[e_0 - ke_1(1 + \varepsilon) - ke_0e_1(1 + \varepsilon)] \quad (5.3)$$

Squaring both sides of (5.3) and neglecting terms of e 's having power greater than two we have

$$(\tilde{y}_{HO} - \bar{Y})^2 = \bar{Y}^2[e_0^2 + k^2e_1^2(1 + \varepsilon)^2 - 2ke_0e_1(1 + \varepsilon)] \quad (5.4)$$

Taking expectation of both sides of (5.4) we get the MSE of \tilde{y}_{HO} to the first degree of approximation as

$$\begin{aligned} MSE(\tilde{y}_{HO}) &= \frac{(1-f)}{n} \bar{Y}^2 [C_y^2 + k^2(1 + \varepsilon)^2 C_x^2 - 2k(1 + \varepsilon) \rho C_x^2] \\ &= \frac{(1-f)}{n} \bar{Y}^2 [C_y^2 - k^2 C_x^2 + k^2 \varepsilon^2 C_x^2] \\ &= \frac{(1-f)}{n} S_y^2 (1 - \rho^2) \left[1 + \frac{\rho^2 \varepsilon^2}{(1 - \rho^2)} \right] \\ &= MSE(\bar{y}_{HO}) \left[1 + \frac{\rho^2 \varepsilon^2}{(1 - \rho^2)} \right] \\ &\Rightarrow \frac{[MSE(\tilde{y}_{HO}) - MSE(\bar{y}_{HO})]}{MSE(\bar{y}_{HO})} = \frac{\rho^2 \varepsilon^2}{(1 - \rho^2)} < \delta \end{aligned} \quad (5.5)$$

That is the proportional increase in MSE of \tilde{y}_{HO} over that of \bar{y}_{HO} is less than δ if

$$|\varepsilon| < \sqrt{\frac{\delta(1 - \rho^2)}{\rho^2}}, \quad (5.6)$$

where δ being a positive constant.

Thus, it follows from (5.6) that to ensure only a small relative increase in MSE, $|\varepsilon|$ must be close to zero if ρ is high but can depart substantially from zero if ρ is just moderate.

6. Estimator Based on Estimated Optimum

It is observed from (4.2) that the asymptotically optimum estimator (AOE) \bar{y}_{HO} presupposes the knowledge about k . If the value of k cannot be guessed quite accurately, then the only alternative left to the investigator is to replace k in (4.2) by its consistent estimator \hat{k} obtained from the data in hand. Thus the estimator of the population mean, \bar{Y} of y on the estimated optimum is

$$\hat{\bar{y}}_{HO} = \bar{y} \frac{\{\bar{X} + \hat{k}(\bar{X} - \bar{x})\}}{\bar{X}} \quad (6.1)$$

where $\hat{k} = \frac{\bar{X} s_{yx}}{\bar{y} s_x^2}$, (\bar{X} is known)

$$= \frac{\hat{\beta}}{\hat{R}} \quad (6.2)$$

is a consistent estimate of $\beta = \frac{S_{yx}}{S_x^2}$ the regression coefficient of y on x ; $s_{yx} = \frac{1}{(n-1)} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$,

$s_x^2 = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2$, and $\hat{R} = \left(\frac{\bar{y}}{\bar{x}}\right)$ is an estimate of the population ratio $R = \left(\frac{\bar{Y}}{\bar{X}}\right)$. Substitution of $\hat{k} = \frac{\hat{\beta}}{\hat{R}}$ in

(6.1) we get the estimator

$$\hat{\bar{y}}_{HO} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x}) \quad (6.3)$$

which is regression estimator of population mean \bar{Y} . Thus the procedure outlined here is an alternative way of obtaining regression estimator for the population mean \bar{Y} . To the first degree of approximation, the variance/MSE of the estimator $\hat{\bar{y}}_{HO}$ is given by

$$Var(\hat{\bar{y}}_{HO}) = \frac{(1-f)}{n} S_y^2 (1 - \rho^2). \quad (6.4)$$

Remark 6.1: A generalized version of dual to product estimator due to Srivenkataramana (1980) and Bandyopadhyay (1980) can be given by

$$\bar{y}_H^* = \bar{y} \frac{(1-b)\bar{X}}{(\bar{X} - b\bar{x})}, \quad (6.5)$$

where b is same as defined earlier.

For $b=0$, \bar{y}_H^* reduces to usual unbiased estimator \bar{y} while for $b=f$, it reduces to Srivenkataramana (1980) and Bandyopadhyaya (1980) dual to product estimator defined by

$$\bar{y}_{SB}^* = \bar{y} \frac{(1-f)\bar{X}}{(\bar{X} - f\bar{x})}.$$

7. Empirical Study

To judge the merits of the suggested estimator over usual unbiased estimator \bar{y} , ratio estimator \bar{y}_R , dual to ratio estimator \bar{y}_{SB} , we have considered three natural population data sets whose descriptions are given below.

Population I: Source: Sukhatme and Chand (1977): Let y be the apple tress of bearing age in 1964 and x be the bushels of apples harvested in 1964. The summary statistics for this data set are:

$$\bar{Y} = 1031.82, \bar{X} = 2934.58, C_y^2 = 2.55280, C_x^2 = 4.02504, \rho = 0.93, N = 200, n = 20.$$

Population II: Source: Kadilar and Cingi (2003): let y be the apple production amount and x be the number of apple tress in 854 villages of Turkey in 1999. The summary statistics for this data set are:

$$\bar{Y} = 1536, \bar{X} = 24357, C_y^2 = 17.49701, C_x^2 = 4.072365, \rho = 0.82, N = 106, n = 9.$$

Population III: Source: Cochran (1977,p.196): Let y be the peach production in bushels in an orchard and x be the number of peach trees in the orchard in North Carolina in June 1946. The summary statistics for this data set are:

$$\bar{Y} = 56.47, \bar{X} = 44.45, C_y^2 = 2.0164, C_x^2 = 1.96, \rho = 0.88, N = 256, n = 100.$$

We have computed the range of b in which the proposed estimator \bar{y}_H is better than \bar{y} , \bar{y}_R and \bar{y}_{SB} . Also we have calculated the percent relative efficiencies (*PREs*) of the proposed estimator \bar{y}_H with respect to usual unbiased estimator \bar{y} , \bar{y}_R and \bar{y}_{SB} for different values of b . Findings are shown in Table 7.1

Table 7.1: The ranges of b under which the suggested estimator \bar{y}_H is better than \bar{y} , \bar{y}_R and \bar{y}_{SB} .

Population	Range of b under which the suggested estimator \bar{y}_H is better than		
	\bar{y}	\bar{y}_R	\bar{y}_{SB}
I	[0,0.60)	(0.32,0.50)	(0.53,0.90)
II	[0,0.77)	(0.50,0.71)	(0.75,0.92)
III	[0,0.64)	(0.44,0.50)	(0.24,0.61)

Table 7.2: The percent relative efficiency (*PRE*) of \bar{y}_H with respect to \bar{y} , \bar{y}_R and \bar{y}_{SB} for different values of b .

b	Population								
	I			II			III		
	<i>PRE</i> of \bar{y}_H with respect to			<i>PRE</i> of \bar{y}_H with respect to			<i>PRE</i> of \bar{y}_H with respect to		
	\bar{y}	\bar{y}_R	\bar{y}_{SB}	\bar{y}	\bar{y}_R	\bar{y}_{SB}	\bar{y}	\bar{y}_R	\bar{y}_{SB}
0	100.00	*	*	100	*	*	100	*	*
0.085	125.53	*	*	107.70	*	100.00	118.22	*	*
0.1	131.59	*	100.00	109.29	*	101.49	122.30	*	*
0.15	156.00	*	119.31	115.26	*	107.03	138.58	*	*
0.2	194.30	*	147.66	122.44	*	113.70	160.39	*	*
0.25	252.10	*	191.58	131.21	*	121.84	190.48	*	*
0.3	346.44	*	263.28	142.11	*	131.97	233.12	*	*
0.32491	414.67	100.00	315.13	148.56	*	137.95	260.84	*	*
0.35	501.13	120.85	380.84	155.90	*	144.76	294.07	*	*
0.391	664.76	160.32	505.19	170.08	*	157.93	360.04	*	100.18
0.4	695.76	167.80	528.75	173.62	*	161.22	375.94	*	104.61
0.426 (ϕ)	740.17	178.50	562.50	184.84	*	171.64	421.32	*	117.24
0.445	709.23	171.04	538.99	194.08	*	180.22	449.35	100.21	125.04
0.45	691.66	166.80	525.63	196.67	*	182.63	455.20	101.52	126.67
0.474 (h)	569.10	137.25	432.49	210.07	*	195.07	468.97	104.59	130.50
0.5	414.66	100.00	315.12	226.48	100	210.30	448.40	100	124.77
0.53667	243.83	*	185.30	252.64	111.55	234.60	359.40	*	100.00
0.55	199.69	*	151.76	262.70	115.00	243.94	318.11	*	*
0.578	131.76	*	100.13	283.33	125.10	263.09	233.68	*	*
0.59698	100.00	*	*	295.24	130.36	274.16	184.50	*	*
0.6	*	*	*	296.84	131.07	275.64	177.45	*	*
0.63 (λ)	*	*	*	305.25	134.79	283.45	119.04	*	*

0.64277	*	*	*	303.11	133.84	281.47	100.00	*	*
0.65	*	*	*	299.97	132.45	278.55	*	*	*
0.7	*	*	*	237.50	104.87	220.55	*	*	*
0.70582	*	*	*	226.50	100.00	210.32	*	*	*
0.75	*	*	*	138.67	*	128.77	*	*	*
0.76779	*	*	*	107.70	*	100.00	*	*	*
3									
0.77269	*	*	*	100.00	*	*	*	*	*
0.8	*	*	*	*	*	*	*	*	*
0.85	*	*	*	*	*	*	*	*	*
0.9	*	*	*	*	*	*	*	*	*
0.95	*	*	*	*	*	*	*	*	*

*Stands for $PRE(\bar{y}_H, \cdot)$ less than 100, (φ) stands for optimum value of b for population I, $(\hat{\lambda})$ stands for optimum value of b for population II, (\hat{h}) stands for optimum value of b for population III.

It is observed from Table 7.1 that the proposed class of estimators \bar{y}_H is better than:

- (i) \bar{y} , \bar{y}_R and \bar{y}_{SB} respectively in the ranges:
 $[0, 0.60)$, $(0.32, 0.50)$ and $(0.53, 0.90)$ of b for population I.
- (ii) \bar{y} , \bar{y}_R and \bar{y}_{SB} respectively in the ranges:
 $[0, 0.77)$, $(0.50, 0.71)$ and $(0.75, 0.92)$ of b for population II.
- (iii) \bar{y} , \bar{y}_R and \bar{y}_{SB} respectively in the ranges:
 $[0, 0.64)$, $(0.44, 0.50)$ and $(0.24, 0.61)$ of b for population III.

Table 7.2 exhibits that the gain in efficiency by using the proposed class of estimators \bar{y}_H over \bar{y} (which does not utilize the auxiliary information) is the largest followed by \bar{y}_{SB} and then \bar{y}_R for all the population data sets I, II, and III. Largest gain in efficiency is observed at optimum value b_{opt} of b for all the populations I, II, III considered here. Finally, we conclude from Tables 7.1 and 7.2 that there is enough scope of choosing the value of scalar 'b' for obtaining estimators better than \bar{y} , \bar{y}_R and \bar{y}_{SB} . Thus we recommend the use of the proposed class of estimators \bar{y}_H in practice.

References

1. Bandyopadhyay, S. (1980). Improved ratio and product estimator. Sankhya. 42, 43-49.
2. Cochran, W. G. (1977). Sampling techniques. 3rd edition. New York: Wiley.
3. Grover, L. K. and Kaur, P. (2011). An improved estimator of the finite population mean in simple random sampling. Model Assisted Statistics and Applications, 6(1), 47-55.
4. Kadilar, C. and Cingi, H. (2003). Ratio estimators in Stratified random sampling. Biometrical Journal. 45(2), 218-225.
5. Kadilar, C. and Cingi, H. (2004). Ratio estimators in simple random sampling. Applied Mathematics and Computation, 151(2), 893-902.
6. Murthy, M. N. (1967). Sampling theory and Methods. Statistical Publishing Society, Calcutta, India.
7. Pal, S. K., Singh, H. P. and Solanki, R. S. (2019). A new efficient class of estimators of finite population mean in simple random sampling. Afrika Mathematica, 31, 595-607.
8. Pal, S. K., Singh, H. P., Kumar, S. and Chatterjee, K. (2018). A family of efficient estimators of finite population mean in simple random sampling. Journal of Statistical Computation and Simulation, 88(5), 920-934.
9. Reddy, V. N. (1978). A study on the use of prior knowledge on certain population parameters in estimation. Sankhya. C, 40, 29-37.

10. Singh, H. P. and Yadav, A. (2018). A two parameter ratio-product-ratio-type exponential estimator for finite population mean in sample surveys. *Pakistan Journal of Statistics and Operation Research*, 14(2), 215-232.
11. Singh, H. P., Solanki, R. S. and Singh, A. (2015). A generalized ratio-cum-product estimator for estimating the finite population mean in survey sampling. *Communications in Statistics-Theory and Methods*, 45(1), 158-172.
12. Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys. *Biometrika*. 67(1),pp. 199-204.
13. Sukhatme, B. V. and Chand, L. (1977). Multivariate ratio type estimators. *Proceedings of American Statistical Association, Social Statistical Section*, 927-931.