# Journal of Modern Applied Statistical Methods

### Volume 12 | Issue 1

Article 16

5-1-2013

# Improved Estimators in Finite Population Surveys: Theory and Applications

Sunil Kumar SOSU, Indian Statistical Institute - Kolkata

Part of the <u>Applied Statistics Commons</u>, <u>Social and Behavioral Sciences Commons</u>, and the <u>Statistical Theory Commons</u>

#### **Recommended** Citation

Kumar, Sunil (2013) "Improved Estimators in Finite Population Surveys: Theory and Applications," *Journal of Modern Applied Statistical Methods*: Vol. 12 : Iss. 1, Article 16. DOI: 10.22237/jmasm/1367381700

## Improved Estimators in Finite Population Surveys: Theory and Applications

Sunil Kumar SOSU, Indian Statistical Institute - Kolkata

Improved estimators are proposed for estimating the population mean  $\overline{Y}$  of the study variable y using auxiliary variable x in simple random sampling. Explicit expression for the bias and MSE of the proposed family are derived to the first order of approximation. The proposed estimators are compared with other estimators and theoretical findings are illustrated by two numerical examples.

Key words: Ratio estimator, product estimator, regression estimator, bias, mean square error.

#### Introduction

Consider a simple random sample of size ndrawn without replacement from a finite population  $U = (U_1, U_2, ..., U_N)$  of N units, let  $\overline{y}$  and  $\overline{x}$  be the sample mean estimates of  $\overline{Y}$ and  $\overline{X}$ , respectively, the population means of the study variable y and auxiliary variable x. To date in the literature of survey sampling, the efficiencies of estimators of unknown population means of a study variable y have been increased by the use of known information on an auxiliary variable x which is highly correlated with study variable y, the well-known ratio estimator is

$$t_R = \frac{-\overline{X}}{\overline{X}}.$$
 (1)

When the correlation between auxiliary variable x and study variable y is highly negative, then the conventional product estimator for  $\overline{Y}$  is defined as

$$t_P = \overline{y} \frac{x}{\overline{X}} \,. \tag{2}$$

Sunil Kumar is a Visiting Scientist in the SOSU, ISI Kolkata. Email him at: sunilbhougal06@gmail.com. Theoretically, it has been established that, in general, the linear regression estimator is more efficient than the ratio and product estimators except in the case where the regression line passes through the neighborhood of the origin, thus the classical regression estimator is

$$t_g = \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right). \tag{3}$$

Furthermore, to find more precise estimates several authors have used prior values of certain population parameter(s). Searls (1964) used known coefficient of variation (CV) of study variable at estimation stage. Sisodiya and Dwivedi (1981) extended the Searls (1964) work by using the known CV of the auxiliary variable for estimating population mean of study variable *v* in a ratio method of estimation. Thus, use of prior value of coefficient of kurtosis in estimating the population variance of study variable y was first conducted by Singh, et al. (1973). It was later, used by Sen (1978), Upadhyaya and Singh (1984) and Searls and Interpanich (1990) in the estimation of population mean of a study variable. Further, Singh and Tailor (2003) proposed a modified ratio estimator by using a known value of a correlation coefficient. This study suggests a new family of estimators to estimate a population mean  $\overline{Y}$  of a study variable y by using estimators from Khoshnevisan, et al.

(2007); the optimum cases of the suggested family of estimators are also obtained.

The Suggested Family of Estimators

A family of estimators proposed by Khoshnevisan, et al. (2007) for estimating the population mean is

$$t = \overline{y} \left\{ \frac{a\overline{X} + c}{\alpha(a\overline{x} + c) + (1 - \alpha)(a\overline{X} + c)} \right\}^{g}$$
(4)

where  $a(\neq 0)$ , *c* are either real numbers or functions of known parameters of an auxiliary variable *x* such as Standard Deviation  $(S_x)$ , Coefficient of Variation  $(C_x)$ , Skewness  $(\beta_1(x))$ , Kurtosis  $(\beta_2(x))$  and Correlation Coefficient  $(\rho_{yx})$  of the population, and  $\alpha$  and *g* are suitably chosen scalars.

The expressions of bias and MSE of the Khoshnevisan's estimator are respectively given as

$$B(t) = \left(\frac{1-f}{n}\right)\overline{Y}\left\{\frac{g(g+1)}{2}\alpha^{2}\upsilon^{2}C_{x}^{2} - \alpha\upsilon gC_{yx}\right\}$$
(5)

and

$$MSE(t) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} + \alpha^{2} v^{2} g^{2} R^{2} S_{x}^{2} - 2\alpha v g R S_{yx} \right\}$$

$$(6)$$

which is minimum, when  $\alpha = \frac{K}{\nu g}$ , then

$$\min .MSE(t) = \left(\frac{1-f}{n}\right) (1-\rho_{yx}^2) S_y^2 = MSE(t_g).$$
(7)

Following Khoshnevisan, et al. (2007), a class of estimators for estimating the population mean  $\overline{Y}$  of the study variable y can be defined by combining regression estimator with t, the

Khoshnevisan, et al. (2007) estimator for full information case is

T =

$$\left\{\overline{y} + b_{yx}\left(\overline{X} - \overline{x}\right)\right\} \left\{\frac{a\overline{X} + c}{\alpha(a\overline{x} + c) + (1 - \alpha)(a\overline{X} + c)}\right\}^{g},$$
(8)

where  $\{\overline{y} + b_{yx}(\overline{X} - \overline{x})\}$  is the classical regression estimator,  $b_{yx}$  is the sample estimate of  $\beta_{yx}, \overline{y} = (l/n) \sum_{i=1}^{n} y_i, \overline{x} = (l/n) \sum_{i=1}^{n} x_i,$  $\overline{X} = (l/N) \sum_{i=1}^{N} X_i, \overline{Y} = (l/N) \sum_{i=1}^{N} Y_i,$  $b_{yx} = (s_{yx}/s_x^2), \beta_{yx} = (S_{yx}/S_x^2),$  $s_{xy} = (l/(n-1)) \sum_{i=1}^{n} (y_i - \overline{y})(x_i - \overline{x}),$  $s_x^2 = (l/(n-1)) \sum_{i=1}^{n} (x_i - \overline{x})^2,$  $S_{xy} = (l/(N-1)) \sum_{i=1}^{N} (y_i - \overline{Y})(x_i - \overline{X}),$  and  $S_x^2 = (l/(N-1)) \sum_{i=1}^{N} (x_i - \overline{X})^2.$ 

To obtain the bias and MSE of the class of estimators T,

$$\overline{y} = \overline{Y}(1 + \varepsilon_0),$$
  

$$\overline{x} = \overline{X}(1 + \varepsilon_1),$$
  

$$s_{xy} = S_{xy}(1 + \varepsilon_2),$$

and

$$s_x^2 = S_x^2 (1 + \varepsilon_3),$$

such that 
$$E(\varepsilon_i) = 0, \forall i = 0 \text{ to } 3$$
,  
 $E(\varepsilon_0^2) = \left(\frac{1-f}{n}\right) \frac{S_y^2}{\overline{Y}^2}, \quad E(\varepsilon_1^2) = \left(\frac{1-f}{n}\right) \frac{S_x^2}{\overline{X}},$   
 $E(\varepsilon_0 \varepsilon_1) = \left(\frac{1-f}{n}\right) \frac{S_{yx}}{\overline{Y}\overline{X}}, \quad C_x = \frac{S_x}{\overline{X}}, C_y = \frac{S_y}{\overline{X}},$   
 $C_{yx} = \frac{S_{xy}}{\overline{Y}\overline{X}}, \quad \rho_{yx} = \frac{\mu_{11}}{\mu_{20}^{1/2}\mu_{02}^{1/2}},$   
 $E(\varepsilon_1 \varepsilon_2) = \left(\frac{1-f}{n}\right) \frac{\lambda_{12}S_x}{\overline{X}\rho_{yx}},$ 

$$E(\varepsilon_{1}\varepsilon_{3}) = \left(\frac{1-f}{n}\right) \frac{\lambda_{03}S_{x}}{\overline{X}}, \quad \lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}, \text{ and}$$
$$\mu_{rs} = \frac{1}{N} \sum_{i=1}^{N} (y_{i} - \overline{Y})^{r} (x_{i} - \overline{X})^{s}, \quad (r, s) \text{ being}$$

non-negative integers.

Expanding T in terms of  $\mathcal{E}'s$  results in

$$T = \overline{Y} \left\{ l + \varepsilon_0 - A\varepsilon_1 (l + \varepsilon_2) (l + \varepsilon_3)^{-l} \right\} (l + \alpha v \varepsilon_1)^{-g},$$
(9)

where  $A = (\beta_{yx}/R)$ ,  $R = (\overline{Y}/\overline{X})$ ,  $v = \frac{a\overline{X}}{a\overline{X} + c}$ . It is assumed that  $|\mathcal{E}_3| < 1$  and  $|\alpha v \mathcal{E}_1| < 1$  so that  $(1 + \mathcal{E}_3)^{-1}$  and  $(1 + \alpha v \mathcal{E}_1)^{-g}$ are expandable in terms of  $\mathcal{E}'s$ . Further, expanding the right hand side of (9), in terms of  $\mathcal{E}'s$  having power greater than two, results in

$$(T - \overline{Y}) = \overline{Y} \begin{cases} \varepsilon_0 - A\varepsilon_1 - g\alpha v\varepsilon_1 - g\alpha v\varepsilon_0 \varepsilon_1 \\ -A\varepsilon_1 \varepsilon_2 + A\varepsilon_1 \varepsilon_3 + g\alpha v\varepsilon_1^2 \\ + \frac{g(g+1)}{2}\alpha^2 v^2 \varepsilon_1^2 \end{cases} \end{cases}.$$
(10)

Taking expectations of both sides on (10), give the bias of T to the first degree of approximation, as

$$B(T) = \begin{bmatrix} \frac{1}{\overline{X}} \left(\frac{1-f}{n}\right) \left\{ \frac{g(g+1)}{2} \alpha^2 v^2 R S_x^2 + g \alpha v \left(A - \beta_{yx}\right) R S_x^2 \right\} - \left(\frac{\lambda_{12}}{\rho_{yx}} - \lambda_{03}\right) S_x \end{bmatrix}$$
(11)

Squaring both sides of (10) and neglecting terms of  $\mathcal{E}$ 's having power greater than two, results in

$$(T - \overline{Y})^{2} = \overline{Y}^{2} \left\{ \varepsilon_{0} - (A + g\alpha v) \varepsilon_{1} \right\}^{2}$$
$$= \overline{Y}^{2} \left\{ \varepsilon_{0}^{2} + A^{2} \varepsilon_{1}^{2} + g^{2} \alpha^{2} v^{2} \varepsilon_{1}^{2} - 2A \varepsilon_{0} \varepsilon_{1} \right\}$$
$$-2g\alpha v \varepsilon_{0} \varepsilon_{1} + 2Ag\alpha v \varepsilon_{1}^{2}$$
(12)

Taking expectations of both sides on (12), gives the MSE of T to the first degree of approximation as

$$MSE(T) = \left(\frac{1-f}{n}\right) \left[S_{y}^{2} + R(A+g\alpha\nu)\left\{R(A+g\alpha\nu) - 2\beta_{yx}\right\}S_{x}^{2}\right]$$
(13)

which is minimum when, for example,  $\alpha = 0 = \alpha_0$ . Thus, the resulting minimum MSE of *T*, is given by

$$min.MSE(T) = \left(\frac{1-f}{n}\right) (1-\rho_{yx}^2) S_y^2$$
$$= MSE(t_g)$$
$$= min.MSE(t)$$
(14)

The ratio-cum-regression estimators (see Table 1) are in the same family of (8) and the bias and mean squared error (MSE) in (13) for these estimators can be expressed as

$$B(T_{i}) = \begin{cases} \frac{1}{\overline{X}} \left(\frac{1-f}{n}\right) \begin{cases} \left(2-\beta_{yx}\right) RS_{x}^{2} \\ -\left(\frac{\lambda_{12}}{\rho_{yx}}-\lambda_{03}\right) S_{x} \end{cases}; \\ i = 1 \\ \frac{1}{\overline{X}} \left(\frac{1-f}{n}\right) \begin{cases} \left(\frac{v_{(i-1)}}{2} \\ +1-\beta_{yx}\end{array}\right) RS_{x}^{2} - \left(\frac{\lambda_{12}}{\rho_{yx}}-\lambda_{03}\right) S_{x} \end{cases}; \\ i = 3, 5, 7, \dots, 17 \end{cases}$$

$$(15)$$

and

$$MSE(T_{i}) = \begin{cases} \left(\frac{1-f}{n}\right) \left\{S_{y}^{2} - \left(\beta_{yx}^{2} - R^{2}\right)S_{x}^{2}\right\};\\ i = l\\ \left(\frac{1-f}{n}\right) \left\{S_{y}^{2} - \left(\beta_{yx}^{2} - v_{\underline{(i-l)}}R^{2}\right)S_{x}^{2}\right\},\\ i = 3, 5, 7, \dots, 17. \end{cases}$$
(16)

For the product-cum-regression estimators given in Table 1, the bias and MSE of the estimators are

$$B(T_{j}) = \begin{cases} \frac{1}{\overline{X}} \left(\frac{1-f}{n}\right) \left\{ \left(\beta_{yx} - 1\right) RS_{x}^{2} - \left(\frac{\lambda_{12}}{\rho_{yx}} - \lambda_{03}\right) S_{x} \right\}; \\ j = 2 \\ \frac{1}{\overline{X}} \left(\frac{1-f}{n}\right) \left\{ \left(\beta_{yx} - 1\right) v_{(j/2)-l} RS_{x}^{2} - \left(\frac{\lambda_{l2}}{\rho_{yx}} - \lambda_{03}\right) S_{x} \right\}, \\ j = 4, 6, 8, \dots, 18 \end{cases}$$

$$(17)$$

$$MSE(T_{j}) = \begin{cases} \left(\frac{1-f}{n}\right) \left\{S_{y}^{2} + \left(\beta_{yx}^{2} + R^{2}\right)S_{x}^{2}\right\}; \\ j = 2 \\ \left(\frac{1-f}{n}\right) \left\{S_{y}^{2} + \left(\beta_{yx}^{2} + v_{(j/2)-l}R^{2}\right)S_{x}^{2}\right\}, \\ j = 4, 6, 8, \dots, 18 \end{cases}$$
(18)

where 
$$v_1 = \frac{\overline{X}}{\overline{X} + C_x}$$
,  $v_2 = \frac{\beta_2(x)\overline{X}}{\beta_2(x)\overline{X} + C_x}$ ,  
 $v_3 = \frac{C_x\overline{X}}{C_x\overline{X} + \beta_2(x)}$ ,  $v_4 = \frac{\overline{X}}{\overline{X} + S_x}$ ,

$$\nu_{5} = \frac{\beta_{1}(x)\overline{X}}{\beta_{1}(x)\overline{X} + S_{x}}, \nu_{6} = \frac{\beta_{2}(x)\overline{X}}{\beta_{2}(x)\overline{X} + S_{x}},$$
$$\nu_{7} = \frac{\overline{X}}{\overline{X} + \rho_{yx}}, \text{ and } \nu_{8} = \frac{\overline{X}}{\overline{X} + \beta_{2}(x)}.$$

Many more estimators can be generated from the proposed estimator in (8) by substituting the different values of  $\alpha$ , g, a and c.

## Efficiency Comparisons

The expressions of MSE of various estimators to the first degree of approximation are

$$Var(\overline{y}) = \left(\frac{1-f}{n}\right)S_{y}^{2};$$
(19)

$$MSE(t_{R}) = \left(\frac{1-f}{n}\right) \left(S_{y}^{2} + R^{2}S_{x}^{2} - 2RS_{yx}\right)$$
(20)

$$MSE(T_{1}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - R^{2}\right) S_{x}^{2} \right\}$$
(21)

$$MSE(T_{3}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{1}R^{2}\right) S_{x}^{2} \right\};$$
(22)

$$MSE(T_{5}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{2}R^{2}\right) S_{x}^{2} \right\};$$
(23)

$$MSE(T_{7}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{3}R^{2}\right) S_{x}^{2} \right\};$$
(24)

$$MSE(T_{9}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{4}R^{2}\right) S_{x}^{2} \right\};$$
(25)

$$MSE(T_{11}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{5}R^{2}\right) S_{x}^{2} \right\};$$
(26)

## IMPROVED ESTIMATORS IN FINITE POPULATION SURVEYS

| Ratio-Cum-Regression Estimators $(g = 1)$  | Product-Cum-Regression Estimators $(g = -1)$  |   | а            | С              |
|--|---|---|--------------|----------------|
| $T_{1} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \frac{\overline{X}}{\overline{x}}$  | $T_{2} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \frac{\overline{x}}{\overline{X}}$   | 1 | 1            | 0              |
| $T_{3} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{X} + C_{x}}{\overline{x} + C_{x}} \right)$                           | $T_4 = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{x} + C_x}{\overline{X} + C_x} \right)$                        | 1 | 1            | $C_x$          |
| $T_5 = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_2(x)\overline{X} + C_x}{\beta_2(x)\overline{x} + C_x} \right)$             | $T_6 = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_2(x)\overline{x} + C_x}{\beta_2(x)\overline{X} + C_x} \right)$    | 1 | $\beta_2(x)$ | $C_x$          |
| $T_{7} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{C_{x} \overline{X} + \beta_{2}(x)}{C_{x} \overline{x} + \beta_{2}(x)} \right)$ | $T_8 = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{C_x \overline{x} + \beta_2(x)}{C_x \overline{X} + \beta_2(x)} \right)$  | 1 | $C_x$        | $\beta_2(x)$   |
| $T_9 = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{X} + S_x}{\overline{x} + S_x} \right)$                                 | $T_{10} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{x} + S_x}{\overline{X} + S_x} \right)$                     | 1 | 1            | S <sub>x</sub> |
| $T_{11} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_1(x)\overline{X} + S_x}{\beta_1(x)\overline{x} + S_x} \right)$          | $T_{12} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_1(x)\overline{x} + S_x}{\beta_1(x)\overline{X} + S_x} \right)$ | 1 | $\beta_1(x)$ | S <sub>x</sub> |
| $T_{13} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_2(x)\overline{X} + S_x}{\beta_2(x)\overline{x} + S_x} \right)$          | $T_{14} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\beta_2(x)\overline{x} + S_x}{\beta_2(x)\overline{X} + S_x} \right)$ | 1 | $\beta_2(x)$ | S <sub>x</sub> |
| $T_{15} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{X} + \rho_{yx}}{\overline{x} + \rho_{yx}} \right)$                  | $T_{16} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{x} + \rho_{yx}}{\overline{X} + \rho_{yx}} \right)$         | 1 | 1            | $ ho_{yx}$     |
| $T_{17} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{X} + \beta_2(x)}{\overline{x} + \beta_2(x)} \right)$                | $T_{18} = \left\{ \overline{y} + b_{yx} \left( \overline{X} - \overline{x} \right) \right\} \left( \frac{\overline{x} + \beta_2(x)}{\overline{X} + \beta_2(x)} \right)$       | 1 | 1            | $\beta_2(x)$   |

| Table 1: Some Members of the Family of Estimators of $T$ |  |
|--|--|

$$MSE(T_{13}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{6}R^{2}\right) S_{x}^{2} \right\};$$
(27)

$$MSE(T_{15}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{7}R^{2}\right) S_{x}^{2} \right\};$$
(28)

$$MSE(T_{17}) = \left(\frac{1-f}{n}\right) \left\{ S_{y}^{2} - \left(\beta_{yx}^{2} - \nu_{8}R^{2}\right) S_{x}^{2} \right\}.$$
(29)

The efficiency comparison of the proposed estimator T at its optimum with the usual unbiased estimator  $\overline{y}$ , ratio estimator  $t_r$  and  $T_i$ ; i = 1,3,5,7,...,17, the results are:

$$Var(\overline{y}) - \min .MSE(T) = \rho_{yx}^2 S_y^2 > 0$$

 $MSE(t_{R}) - \min .MSE(T) = \left\{ R^{2} + \beta_{yx} \left( \beta_{yx} - 2R \right) \right\} S_{x}^{2} > 0$   $MSE(T_{1}) - \min .MSE(T) = R^{2}S_{x}^{2} > 0$   $MSE(T_{3}) - \min .MSE(T) = v_{1}R^{2}S_{x}^{2} > 0$   $MSE(T_{5}) - \min .MSE(T) = v_{2}R^{2}S_{x}^{2} > 0$   $MSE(T_{7}) - \min .MSE(T) = v_{3}R^{2}S_{x}^{2} > 0$   $MSE(T_{9}) - \min .MSE(T) = v_{4}R^{2}S_{x}^{2} > 0$   $MSE(T_{11}) - \min .MSE(T) = v_{5}R^{2}S_{x}^{2} > 0$   $MSE(T_{13}) - \min .MSE(T) = v_{6}R^{2}S_{x}^{2} > 0$   $MSE(T_{15}) - \min .MSE(T) = v_{7}R^{2}S_{x}^{2} > 0$   $MSE(T_{17}) - \min .MSE(T) = v_{8}R^{2}S_{x}^{2} > 0$   $MSE(T_{17}) - \min .MSE(T) = v_{8}R^{2}S_{x}^{2} > 0$   $MSE(T_{17}) - \min .MSE(T) = v_{8}R^{2}S_{x}^{2} > 0$  $MSE(T_{17}) - \min .MSE(T) = v_{8}R^{2}S_{x}^{2} > 0$ 

$$MSE(T_{1}) - Var(\overline{y}) < 0; \text{ if } R^{2} < \beta_{yx}^{2}$$
$$MSE(T_{i}) - Var(\overline{y}) < 0; \text{ if } \nu_{j}R^{2} < \beta_{yx}^{2};$$
$$i = 3, 5, 7, 9, 11, 13, 15, 17 \text{ and } j = 1 \text{ to } 8.$$

If these defined conditions are true, then the proposed estimator is more efficient than the usual unbiased estimator  $\overline{y}$  and unbiased ratio estimator  $t_R$ .

**Empirical Study** 

Population I (Koyuncu & Kadilar, 2009)

Data concerning primary and secondary schools for 923 districts of Turkey in 2007 (Source: Ministry of Education, Republic of Turkey), taking the number of teachers as study variable and the number of students as auxiliary variable in both primary and secondary schools. A sample of size n = 180 was selected from the dataset and the correlations observed between auxiliary and study variables are positive. Therefore, ratio estimators were used for the estimation of the population mean. The summary statistics about the population are: N = 923, n = 180,  $S_v = 749.9395$ ,  $S_x = 21331.1315$ ,  $\overline{Y} = 436.4345$ ,  $\overline{X} = 11440.4984$  $\rho_{vx} = 0.9543$ ,  $\beta_2(x) = 18.7208 \text{ and } \beta_1(x) = 3.9365.$ 

Population II (Singh, 2001)

Consider the data used by Anderson (1958). The parameters of the population consist of 25 families as follows: y: head length of second son; x: head breadth of first son; N = 25, n = 7,  $C_y = 0.0546$ ,  $C_x = 0.0488$ ,  $\overline{Y} = 183.84$ ,  $\overline{X} = 151.12$ ,  $\rho_{yx} = 0.6932$ ,  $\beta_2(x) = 2.6519$ , and  $\beta_1(x) = 0.0002$ .

Here, the percent relative efficiencies (PRE) of different suggested estimators were computed with respect to the usual unbiased estimator  $\overline{y}$  for both populations. The outcomes are shown in Table 2.

| Population I             |                        | Population II            |                        |  |
|--------------------------|------------------------|--------------------------|------------------------|--|
| Ratio Type<br>estimators | $PRE(., \overline{y})$ | Ratio Type<br>estimators | $PRE(., \overline{y})$ |  |
| $\overline{\mathcal{Y}}$ | 100.00                 | $\overline{\mathcal{Y}}$ | 100.00                 |  |
| $t_r$                    | 939.7177               | $t_r$                    | 178.6661               |  |
| $T_1$                    | 78.94465               | $T_1$                    | 75.85505               |  |
| $T_3$                    | 78.95661               | $T_3$                    | 75.86713               |  |
| $T_5$                    | 78.94529               | $T_5$                    | 75.85961               |  |
| $T_7$                    | 79.00904               | $T_7$                    | 87.91403               |  |
| $T_9$                    | 199.8646               | $T_9$                    | 78.05581               |  |
| $T_{11}$                 | 112.5773               | $T_{11}$                 | 191.3018               |  |
| <i>T</i> <sub>13</sub>   | 86.20194               | <i>T</i> <sub>13</sub>   | 76.6948                |  |
| <i>T</i> <sub>15</sub>   | 78.95077               | <i>T</i> <sub>15</sub>   | 6.02636                |  |
| T <sub>17</sub>          | 79.06471               | <i>T</i> <sub>17</sub>   | 76.50771               |  |
| $T_{(opt.)}$             | 1119.6777              | $T_{(opt.)}$             | 192.5025               |  |

Table 2: Percent Relative Efficiencies of Different Suggested Ratio Type Estimators with Respect to the Usual Unbiased Estimator  $\overline{y}$ 

#### Conclusion

It is envisaged from Table 2 that the proposed estimator at its optimum is more efficient than among all discussed estimators. Also, the estimators  $T_1$ ,  $T_3$ ,  $T_5$ ,  $T_7$ ,  $T_{13}$ ,  $T_{15}$  and  $T_{17}$  are less efficient than the usual unbiased estimator because the condition  $v_j R^2 < \beta_{yx}^2$ ; i = 3,5,7,9,11,13,15,17 and j = 1 to 8 is not satisfied for both the populations.

#### References

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York, NY: John Wiley & Sons, Inc.

Khoshnevisan, M., et al. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s). *Far East Journal of Theoretical Statistics*, 22(2), 181-191.

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

Searls, D. T. (1964). The utilization of known coefficient of variation in the estimation procedure. *Journal of the American Statistics Association*, *59*, 1225-1226.

Searls, D. T., & Intarapanich, P. (1990). A note on an estimator for the variance that utilizes the kurtosis. *The American Statistician*, 44(4), 295-296.

Sen, A. R. (1978). Estimation of the population mean when the coefficient of variation is known. *Communication in Statistics, Theory and Methods*, 7, 657-672.

Singh, G. N. (2001). On the use of transformed auxiliary variable in estimation of population mean in two phase sampling. *Statistics in Transition*, 5(3), 405-416.

Singh, J., Pandey, B. N., & Hiran, K. (1973). On the utilization of a known coefficient of kurtosis in the estimation procedure of variance. *Annals of the Institute of Statistical Mathematics*, 25, 51-55.

Singh, H. P., & Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition*,6, 555-560.

Sisodia, B. V. S., & Dwivedi, V. K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of Indian Society Agricultural Statistics*, *33*, 13-18.

Upadhyaya, L. N., & Singh, H. P. (1984).On the estimator of the population mean with known coefficient of variation. *Biometrical Journal*, *26*(*8*), 915-922.