Revista Colombiana de Estadística
Junio 2014, volumen 37, no. 1, pp. 199 a 211

A New Difference-Cum-Exponential Type Estimator of Finite Population Mean in Simple Random Sampling

Un nuevo estimador tipo diferencia-cum-exponencial de la media de una población finita en muestras aleatorias simple

Javid Shabbir[1,a], Abdul Haq[1,b], Sat Gupta[2,c]

[1]Department of Statistics, Quaid-I-Azam University, Islamabad, Pakistan

[2]Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, USA

Abstract

Auxiliary information is frequently used to improve the accuracy of the estimators when estimating the unknown population parameters. In this paper, we propose a new difference-cum-exponential type estimator for the finite population mean using auxiliary information in simple random sampling. The expressions for the bias and mean squared error of the proposed estimator are obtained under first order of approximation. It is shown theoretically, that the proposed estimator is always more efficient than the sample mean, ratio, product, regression and several other existing estimators considered here. An empirical study using 10 data sets is also conducted to validate the theoretical findings.

***Key words*:** Ratio estimator, Auxiliary Variable, Exponential type estimator, Bias, MSE, Efficiency.

Resumen

Información auxiliar se utiliza con frecuencia para mejorar la precisión de los estimadores al estimar los parámetros poblacionales desconocidos. En este trabajo, se propone un nuevo tipo de diferencia-cum-exponencial estimador de la población finita implicar el uso de información auxiliar en muestreo aleatorio simple. Las expresiones para el sesgo y el error cuadrático medio del estimador propuesto se obtienen en primer orden de aproximación. Se muestra teóricamente, que el estimador propuesto es siempre más eficiente que la media de la muestra, la relación de, producto, regresión y varios otros

[a]Professor. E-mail: javidshabbir@gmail.com
[b]Lecturer. E-mail: aaabdulhaq@gmail.com
[c]Professor. E-mail: sngupta@uncg.edu

estimadores existentes considerados aquí. Un estudio empírico utilizando 10 conjuntos de datos también se lleva a cabo para validar los resultados teóricos.

***Palabras clave*:** estimador de razón, variables auxiliares, estimador tipo exponencial, sesgo, error cuadrático medio.

1. Introduction

In sample surveys, auxiliary information can be used either at the design stage or at the estimation stage or at both stages to increase precision of the estimators of population parameters. The ratio, product and regression methods of estimation are commonly used in this context. Recently many research articles have appeared where authors have tried to modify existing estimators or construct new hybrid type estimators. Some contribution in this area are due to Bahl & Tuteja (1991), Singh, Chauhan & Sawan (2008), Singh, Chauhan, Sawan & Smarandache (2009), Yadav & Kadilar (2013), Haq & Shabbir (2013), Singh, Sharma & Tailor (2014) and Grover & Kaur, (2011, 2014).

Consider a finite population $U = \{U_1, U_2, \ldots, U_N\}$. We draw a sample of size n from this population using simple random sampling without replacement scheme. Let y and x respectively be the study and the auxiliary variables and y_i and x_i, respectively be the observations on the ith unit. Let $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample means and $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ and $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$, be the corresponding population means. We assume that the mean of the auxiliary variable ($\bar{X}$) is known. Let $s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ and $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ be the sample variances and $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$and $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$, be the corresponding population variances. Let ρ_{yx} be the correlation coefficient between y and x. Finally let $C_y = \frac{S_y}{\bar{Y}}$ and $C_x = \frac{S_x}{\bar{X}}$ respectively be the coefficients of variation for y and x.

In order to obtain the bias and mean squared error (MSE) for the proposed estimator and existing estimators considered here, we define the following relative error terms: Let $\delta_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $\delta_1 = \frac{\bar{x}-\bar{X}}{\bar{X}}$, such that $E(\delta_i) = 0$ for $(i = 0, 1)$, $E(\delta_0^2) = \lambda C_y^2$, $E(\delta_1^2) = \lambda C_x^2$ and $E(\delta_0\delta_1) = \lambda\rho_{yx}C_yC_x$, where $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$.

In this paper, our objective is to propose an improved estimator of the finite population mean using information on a single auxiliary variable in simple random sampling. Expressions for the bias and mean squared error (MSE) of the proposed estimator are derived under first order of approximation. Based on both theoretical and numerical comparisons, we show that the proposed estimator outperforms several existing estimators. The outline of the paper is as follows: in Section 2, we consider several estimators of the finite population mean that are available in literature. The proposed estimators are given in Section 3 along with the corresponding bias and MSE expressions. In Section 4, we provide theoretical comparisons to evaluate the performances of the proposed and existing estimators. An empirical study is conducted in Section 5, and some concluding remarks are given in Section 6.

2. Some Existing Estimators

In this section, we consider several estimators of finite population mean.

2.1. Sample Mean Estimator

The variance of the sample mean $\bar{y}$, the usual unbiased estimator, is given by

$$Var(\bar{y}) = \lambda \bar{Y}^2 C_y^2 \tag{1}$$

2.2. Traditional Ratio and Product Estimators

Using information on the auxiliary variable, Cochran (1940) suggested a ratio estimator $\hat{\bar{Y}}_R$ for estimating $\bar{Y}$. It is given by

$$\hat{\bar{Y}}_R = \bar{y}(\frac{\bar{X}}{\bar{x}}) \tag{2}$$

The MSE of $\hat{\bar{Y}}_R$, to first order of approximation, is given by

$$MSE(\hat{\bar{Y}}_R) \approx \lambda \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x) \tag{3}$$

On similar lines, Murthy (1964) suggested a product estimator ($\hat{\bar{Y}}_P$), given by

$$\hat{\bar{Y}}_P = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right) \tag{4}$$

The MSE of $\hat{\bar{Y}}_P$, to first order of approximation, is given by

$$MSE(\hat{\bar{Y}}_P) \approx \lambda \bar{Y}^2 (C_y^2 + C_x^2 + 2\rho_{yx} C_y C_x) \tag{5}$$

The ratio and product estimators are widely used when the correlation coefficient between the study and the auxiliary variable is positive and negative, respectively. Both of the estimators, $\hat{\bar{Y}}_R$ and $\hat{\bar{Y}}_P$, show better performances in comparison with $\bar{y}$ when $\rho_{yx} > \frac{C_x}{2C_y}$ and $\rho_{yx} < -\frac{C_x}{2C_y}$, respectively.

2.3. Regression Estimator

The usual regression estimator $\hat{\bar{Y}}_{Reg}$ of $\bar{Y}$, is given by

$$\hat{\bar{Y}}_{Reg} = \bar{y} + b(\bar{X} - \bar{x}) \tag{6}$$

where b is the usual slope estimator of the population regression coefficient β (Cochran 1977). The estimator $\hat{\bar{Y}}_{Reg}$ is biased, but the bias approaches zero as the sample size n increases.

Asymptotic variance of $\hat{\hat{Y}}_{Reg}$, is given by

$$Var(\hat{\hat{Y}}_{Reg}) = \lambda \bar{Y}^2 C_y^2 (1 - \rho_{yx}^2) \tag{7}$$

The regression estimator $\hat{\hat{Y}}_{Reg}$ performs better than the usual mean estimator $\bar{y}$, ratio estimator $\hat{\bar{Y}}_R$ and product estimator $\hat{\bar{Y}}_P$ when $\lambda \bar{Y}^2 \rho_{yx}^2 C_y^2 > 0$, $\lambda \bar{Y}^2 (C_x - \rho_{yx} C_y)^2 > 0$ and $\lambda \bar{Y}^2 (C_x + \rho_{yx} C_y)^2 > 0$, respectively.

2.4. Bahl & Tuteja (1991) Estimators

Bahl & Tuteja (1991) suggested ratio-and product type estimators of $\bar{Y}$, given respectively by

$$\hat{\hat{Y}}_{BT,R} = \bar{y} \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{8}$$

and

$$\hat{\hat{Y}}_{BT,P} = \bar{y} \exp\left(\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}}\right) \tag{9}$$

The MSEs of $\hat{\hat{Y}}_{BT,R}$ and $\hat{\hat{Y}}_{BT,P}$, to first order of approximation, are given by

$$MSE(\hat{\hat{Y}}_{BT,R}) \approx (1/4) \lambda \bar{Y}^2 (4C_y^2 + C_x^2 - 4\rho_{xy} C_y C_x) \tag{10}$$

and

$$MSE(\hat{\hat{Y}}_{BT,P}) \approx (1/4) \lambda \bar{Y}^2 (4C_y^2 + C_x^2 + 4\rho_{xy} C_y C_x) \tag{11}$$

2.5. Singh et al. (2008) Estimator

Following Bahl & Tuteja (1991), Singh et al. (2008) suggested a ratio-product exponential type estimator $\hat{\hat{Y}}_{S,RP}$ of $\bar{Y}$, given by

$$\hat{\hat{Y}}_{S,RP} = \bar{y}[\alpha \exp(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}) + (1 - \alpha) \exp(\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}})] \tag{12}$$

where α is an arbitrary constant.

The minimum MSE of $\bar{Y}_{S,RP}$, up to first order of approximation, at optimum value of α, i.e., $\alpha_{(opt)} = \frac{1}{2} + \frac{\rho_{yx} C_y}{C_x}$, is given by

$$MSE_{\min}(\hat{\hat{Y}}_{S,RP}) \approx \lambda \bar{Y}^2 (1 - \rho_{yx}^2) C_y^2 = Var(\hat{\hat{Y}}_{Reg}) \tag{13}$$

The minimum MSE of $\hat{\hat{Y}}_{S,RP}$ is exactly equal to variance of the linear regression estimator ($\hat{\hat{Y}}_{Reg}$).

2.6. Rao (1991) Estimator

Rao (1991) suggested a regression-type estimator of $\bar{Y}$, given by

$$\hat{\bar{Y}}_{R,Reg} = k_1\bar{y} + k_2(\bar{X} - \bar{x}) \quad (14)$$

where k_1 and k_2 are suitably chosen constants.

The minimum MSE of $\bar{Y}_{R,Reg}$, upto first order of approximation, at optimum values of k_1 and k_2, i.e., $k_{1(opt)} = \frac{1}{1+\lambda(1-\rho_{yx}^2)C_y^2}$ and $k_{2(opt)} = -\frac{\bar{Y}\rho_{yx}C_y}{\bar{X}C_x[-1+\lambda(-1+\rho_{yx}^2)C_y^2]}$, is given by

$$MSE_{\min}(\hat{\bar{Y}}_{R,Reg}) \approx \bar{Y}^2\left\{1 + \frac{1}{-1+\lambda(-1+\rho_{yx}^2)C_y^2}\right\} \quad (15)$$

2.7. Grover & Kaur (2011) Estimator

Following Rao (1991) and Bahl & Tuteja (1991), Grover & Kaur (2011) suggested an exponential type estimator of $\bar{Y}$, given by

$$\hat{\bar{Y}}_{GK} = [d_1\bar{y} + d_2(\bar{X} - \bar{x})]\exp(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}) \quad (16)$$

where d_1 and d_2 are suitably chosen constants.

The minimum MSE of $\hat{\bar{Y}}_{GK}$, up to first order of approximation, at optimum values of d_1 and d_2 i.e., $d_{1(opt)} = \frac{-8+\lambda C_x^2}{8\left\{-1+\lambda(-1+\rho_{yx}^2)C_y^2\right\}}$ and

$$d_{2(opt)} = \frac{\bar{Y}[-8\rho_{yx}C_y + C_x\left\{4 - \lambda C_x^2 + \lambda\rho_{yx}C_yC_x + 4\lambda(-1+\rho_{yx}^2)C_y^2\right\}}{8\bar{X}C_x\left\{-1+\lambda(-1+\rho_{yx}^2)C_y^2\right\}}$$

is given by

$$MSE_{\min}(\hat{\bar{Y}}_{GK}) \approx \frac{\lambda\bar{Y}^2[\lambda C_x^4 - 16(-1+\rho_{yx}^2)(-4+\lambda C_x^2)C_y^2]}{64[-1+\lambda(-1+\rho_{yx}^2)C_y^2]} \quad (17)$$

Grover & Kaur (2011) derived the result

$$MSE_{min}(\hat{\bar{Y}}_{GK}) \approx Var(\hat{\bar{Y}}_{Reg}) - \frac{\lambda^2\bar{Y}^2\left\{C_x^2 + 8(1-\rho_{yx}^2)C_y^2\right\}^2}{64\left\{1+\lambda(1-\rho_{yx}^2)C_y^2\right\}} \quad (18)$$

Equation (18) shows that $\hat{\bar{Y}}_{GK}$ is more efficient than the linear regression estimator $\hat{\bar{Y}}_{Reg}$.

Since regression estimator $\hat{\bar{Y}}_{Reg}$ is always better than $\bar{y}$, $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_P$, $\hat{\bar{Y}}_{BT,R}$, $\hat{\bar{Y}}_{BT,P}$, it can be argued that $\hat{\bar{Y}}_{GK}$ is also always better than these estimators.

3. Proposed Estimator

In this section, an improved difference-cum-exponential type estimator of the finite population mean $\bar{Y}$ using a single auxiliary variable is proposed. Expressions for the bias and MSE of the proposed estimator are obtained upto first order of approximation.

The conventional difference estimator ($\hat{\bar{Y}}_D$) of $\bar{Y}$, is given by

$$\hat{\bar{Y}}_D = \bar{y} + w_1(\bar{X} - \bar{x}) \tag{19}$$

where w_1 is a constant.

From (8), (12), and (14), a difference-cum-exponential type estimator ($\hat{\bar{Y}}^*_D$) of $\bar{Y}$may be given by

$$\hat{\bar{Y}}^*_D = \left[\hat{\bar{Y}}^*_{S,RP} + w_1(\bar{X} - \bar{x})\right] \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{20}$$

where $\hat{\bar{Y}}^*_{S,RP} = \frac{\bar{y}}{2}\left[\exp\left(\frac{\bar{X}-\bar{x}}{\bar{x}+\bar{X}}\right) + \exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right)\right]$ is the average of exponential ratio and exponential product estimators $\hat{\bar{Y}}_{BT,R}$ and $\hat{\bar{Y}}_{BT,P}$ respectively.

Following Searls (1964) and Bahl & Tuteja (1991), Yadav & Kadilar (2013) suggested the following estimator for $\bar{Y}$:

$$\hat{\bar{Y}}_{YK} = w_2\, \bar{y} \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{21}$$

where w_2 is a suitably chosen constant.

By combining the ideas in (20) and (21), a modified difference-cum-exponential type estimator of $\bar{Y}$, is given by

$$\hat{\bar{Y}}^*_P = [\hat{\bar{Y}}^*_{S,RP} + w_1(\bar{X} - \bar{x}) + w_2\bar{y}] \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{22}$$

where w_1 and w_2 are unknown constants to be determined later.

Rewriting $\hat{\bar{Y}}^*_P$ as

$$\hat{\bar{Y}}^*_P = \left[\frac{\bar{y}}{2}\left\{\exp(\frac{\bar{X} - \bar{x}}{\bar{x} + \bar{X}}) + \exp\left(\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}}\right)\right\} + w_1(\bar{X} - \bar{x}) + w_2\bar{y}\right] \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right)$$

Solving $\hat{\bar{Y}}^*_P$ in terms of $\delta_i (i = 0, 1)$, to first order of approximation, we can write

$$\begin{aligned}\hat{\bar{Y}}^*_P - \bar{Y} \approx\ & \bar{Y}w_2 + \bar{Y}\delta_o - \frac{1}{2}\bar{Y}\delta_1 - \bar{X}\delta_1 w_1 + \bar{Y}\delta_o w_2 - \frac{1}{2}\bar{Y}\delta_1 w_2 \\ & - \frac{1}{2}\bar{Y}\delta_o\delta_1 + \frac{1}{2}\bar{Y}w_1^2 + \frac{1}{2}\bar{X}\delta_1^2 w_1 - \frac{1}{2}\bar{Y}\delta_o\delta_1 w_2 + \frac{3}{8}\bar{Y}\delta_1^2 w_2\end{aligned} \tag{23}$$

Taking expectation on both sides of (23), we get the bias of $\hat{\bar{Y}}^*_P$, given by

$$Bias(\hat{\bar{Y}}^*_P) \approx \frac{1}{8}[8\bar{Y}w_2 + \lambda C_x^2\left\{4\bar{X}w_1 + \bar{Y}(4 + 3w_2)\right\} - 4\bar{Y}\lambda C_Y C_x(1 + w_2)\rho_{yx}] \tag{24}$$

Squaring both sides of (23) and using first order of approximation, we get

$$\begin{aligned}(\hat{\bar{Y}}_P^* - \bar{Y})^2 \approx\ & \bar{Y}^2 w_2^2 + \bar{Y}^2\delta_o^2 - \bar{Y}^2\delta_o\delta_1 \\ & + \frac{1}{4}\bar{Y}^2\delta_1^2 - 2\bar{X}\bar{Y}\delta_o\delta_1 w_1 + \bar{X}\bar{Y}\delta_1^2 w_1 + \bar{X}^2\delta_1^2 w_1^2 + 2\bar{Y}^2\delta_o^2 w_2 \\ & - 3\bar{Y}^2\delta_o\delta_1 w_2 + \frac{3}{2}\bar{Y}^2\delta_1^2 w_2 - 2\bar{X}\bar{Y}\delta_o\delta_1 w_1 w_2 + 2\bar{X}\bar{Y}\delta_1^2 w_1 w_2 \\ & + \bar{Y}^2\delta_o^2 w_2^2 - 2\bar{Y}^2\delta_o\delta_1 w_2^2 + \bar{Y}^2\delta_1^2 w_2^2\end{aligned} \tag{25}$$

Taking expectation on both sides of (25), the MSE of $\hat{\bar{Y}}_P^*$, to first order of approximation, is given by

$$\begin{aligned}MSE(\hat{\bar{Y}}_P^*) \approx\ & \frac{1}{4}\lambda C_x^2\left\{(\bar{Y} + 2\bar{X}w_1)^2 + 2\bar{Y}(3\bar{Y} + 4\bar{X}w_1)w_2 + 4\bar{Y}^2 w_2^2\right\} \\ & + \bar{Y}^2\left\{w_2^2 + \lambda C_Y^2(1 + w_2)^2\right\} \\ & - \bar{Y}\lambda\rho_{yx}C_y C_x(1 + w_2)(\bar{Y} + 2\bar{X}w_1 + 2\bar{Y}w_2)\end{aligned} \tag{26}$$

Partially differentiating (26) with respect to w_1 and w_2, we get

$$\frac{\partial MSE(\hat{\bar{Y}}_P^*)}{\partial w_1} = \bar{X}\lambda C_x\left\{-2\bar{Y}\rho_{yx}C_y(1 + w_2) + C_x(\bar{Y} + 2\bar{X}w_1 + 2\bar{Y}w_2)\right\}$$

$$\begin{aligned}\frac{\partial MSE(\hat{\bar{Y}}_P^*)}{\partial w_2} =\ & \frac{1}{2}\bar{Y}[4\bar{Y}\left\{w_2 + \lambda C_y^2(1 + w_2)\right\} - 2\lambda\rho_{yx}C_y C_x\left\{2\bar{X}w_1 + \bar{Y}(3 + 4w_2)\right\} \\ & + \lambda C_x^2\left\{4\bar{X}w_1 + \bar{Y}(3 + 4w_2)\right\}]\end{aligned}$$

Setting $\frac{\partial MSE(\hat{\bar{Y}}_P^*)}{\partial w_2}$=0 for $i = 0, 1$, the optimum values of w_1 and w_2 are given by

$$w_{1(opt)} = \frac{\bar{Y}[-4\rho_{yx}C_y + C_x\left\{2 - \lambda C_x^2 + \lambda\rho_{yx}C_y C_x + 2\lambda(-1 + \rho_{yx}^2)C_y^2\right\}]}{4\bar{X}C_x\left\{-1 + \lambda(-1 + \rho_{yx}^2)C_y^2\right\}}$$

and $w_{2(opt)} = \frac{\lambda(C_x^2 - 4(-1+\rho_{yx}^2)C_y^2)}{4\left\{-1+\lambda(-1+\rho_{yx}^2)C_y^2\right\}}$, respectively.

Substituting the optimum values of w_1 and w_2 in (26), we can obtain the minimum MSE of $\hat{\bar{Y}}_P^*$, as given by

$$MSE_{\min}(\hat{\bar{Y}}_P^*) \approx \frac{\lambda\bar{Y}^2\left\{\lambda C_x^4 - 8(-1 + \rho_{yx}^2)(-2 + \lambda C_x^2)C_y^2\right\}}{16\left\{-1 + \lambda(-1 + \rho_{yx}^2)C_y^2\right\}} \tag{27}$$

After some simplifications, (27) can be written as

$$MSE_{\min}(\hat{\bar{Y}}_P^*) \approx MSE(\hat{\bar{Y}}_{Reg}) - (T_1 + T_2) \tag{28}$$

where $T_1 = \frac{\lambda^2\bar{Y}^2\left\{C_x^2+8(1-\rho_{yx}^2)C_y^2\right\}^2}{64\left\{1+\lambda(1-\rho_{yx}^2)C_y^2\right\}}$ and $T_2 = \frac{\lambda^2\bar{Y}^2 C_x^2\left\{3C_x^2+16(1-\rho_{yx}^2)C_y^2\right\}}{64\left\{1+\lambda(1-\rho_{yx}^2)C_y^2\right\}}$

Note that both quantities, T_1 and T_2, are always positive.

4. Efficiency Comparisons

In this section, we compare the proposed estimator with the existing estimators considered in Section 2 and derive the following observations:

Observation (i): By (1) and (28)

$$Var(\bar{y}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \lambda \bar{Y}^2 \rho_{yx}^2 C_y^2 + T_1 + T_2 > 0$$

Observation (ii): By (3) and (28)

$$MSE(\hat{\hat{Y}}_R) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \lambda \bar{Y}^2 (C_x - \rho_{yx} C_y)^2 + T_1 + T_2 > 0$$

Observation (iii): By (5), and (28)

$$MSE(\hat{\hat{Y}}_P) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \lambda \bar{Y}^2 (C_x + \rho_{yx} C_y)^2 + T_1 + T_2 > 0$$

Observation (iV): By (7), (13) and (28)

$$MSE(\hat{\hat{Y}}_{Reg}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = MSE(\hat{\hat{Y}}_{S,RP}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = T_1 + T_2 > 0$$

Observation (V): By (10) and (28)

$$MSE(\hat{\hat{Y}}_{BT,R}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \frac{1}{4} \lambda \bar{Y}^2 (C_x - 2\rho_{yx} C_y)^2 + T_1 + T_2 > 0$$

Observation (Vi): By (11) and (28)

$$MSE(\hat{\hat{Y}}_{BT,P}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \frac{1}{4} \lambda \bar{Y}^2 (C_x + 2\rho_{yx} C_y)^2 + T_1 + T_2 > 0$$

Observation (Vii): By (15) and (28)

$$MSE(\hat{\hat{Y}}_{R,Reg}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = \frac{\lambda^2 \bar{Y}^2 C_x^2 \left\{ C_x^2 + 16(1-\rho_{yx}^2) C_y^2 \right\}}{64 \left\{ 1 + \lambda (1-\rho_{yx}^2) C_y^2 \right\}} + T_2 > 0$$

Observation (Viii): By (18) and (28)

$$MSE(\hat{\hat{Y}}_{GK}) - MSE_{\min}(\hat{\hat{Y}}_P^*) = T_2 > 0$$

In the light of the eight observations made above, we can argue that the proposed estimator performs better than all of the estimators considered here.

5. Empirical Study

In this section, we consider 10 real data sets to numerically evaluate the performances of the proposed and the existing estimators considered here.

Population 1: [Source: Cochran (1977), pp. 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:
$N = 256$, $n = 100$, $\bar{Y} = 56.47$, $\bar{X} = 44.45$, $C_y = 1.42$, $C_x = 1.40$, $\rho_{yx} = 0.887$.

Population 2: [Source: Murthy (1977), pp. 228] Let y be the output and x be the number of workers. The summary statistics for this data set are:
$N = 80$, $n = 10$, $\bar{Y} = 51.8264$, $\bar{X} = 2.8513$, $C_y = 0.3542$, $C_x = 0.9484$, $\rho_{yx} = 0.915$.

Population 3: [Source: Das (1988)] Let y be the number of agricultural laborers for 1971 and x be the number of agricultural laborers for 1961. The summary statistics for this data set are:
$N = 278$, $n = 25$, $\bar{Y} = 39.068$, $\bar{X} = 25.111$, $C_y = 1.4451$, $C_x = 1.6198$, $\rho_{yx} = 0.7213$.

Population 4: [Source: Steel, Torrie & Dickey (1960), pp. 282] Let y be the log of lef burn in sacs and x be the chlorine percentage. The summary statistics for this data set are:
$N = 30$, $n = 6$, $\bar{Y} = 0.6860$, $\bar{X} = 0.8077$, $C_y = 0.7001$, $C_x = 0.7493$, $\rho_{yx} = -0.4996$.

Population 5: [Source: Maddala (1977), pp. 282] Let y be the consumption per capita and x be the deflated prices of veal. The summary statistics for this data set are:
$N = 16$, $n = 4$, $\bar{Y} = 7.6375$, $\bar{X} = 75.4343$, $C_y = 0.2278$, $C_x = 0.0986$, $\rho_{yx} = -0.6823$.

Population 6: [Source: Kalidar & Cingi (2007)] Let y be the level of apple production (in 100 tones) and x be the number of apple trees in 104 villages in the East Anatolia Region in 1999. The summary statistics for this data set are:
$N = 104$, $n = 20$, $\bar{Y} = 6.254$, $\bar{X} = 13931.683$, $C_y = 1.866$, $C_x = 1.653$, $\rho_{yx} = 0.865$.

Population 7: [Source: Kalidar & Cingi (2005)] Let y be the apple production amount in 1999 and x be the number of apple trees in 1999 in Black sea region of Turkey. The summary statistics for this data set are:
$N = 204$, $n = 50$, $\bar{Y} = 966$, $\bar{X} = 26441$, $C_y = 2.4739$, $C_x = 1.7171$, $\rho_{yx} = 0.71$.

Population 8: [Source: Cochran (1977)] Let y be the number of 'placebo' children and x be the number of paralytic polio cases in the placebo group. The summary statistics for this data set are:
$N = 34$, $n = 10$, $\bar{Y} = 4.92$, $\bar{X} = 2.59$, $C_y = 1.01232$, $C_x = 1.07201$, $\rho_{yx} = 0.6837$.

Population 9: [Source: Srivnstava, Srivastava & Khare (1989)] Let y be the measurement of weight children and x be the mid-arm circumference of children. The summary statistics for this data set are:
$N = 55$, $n = 30$, $\bar{Y} = 17.08$, $\bar{X} = 16.92$, $C_y = 0.12688$, $C_x = 0.07$, $\rho_{yx} = 0.54$.

Population 10: [Source: Sukhatme & Chand (1977)] Let y be the apple trees of bearing age in 1964 and x be the bushels harvested in 1964. The summary statistics for this data set are:
$N = 200$, $n = 20$, $\bar{Y} = 1031.82$, $\bar{X} = 2934.58$, $C_y = 1.59775$, $C_x = 2.00625$, $\rho_{yx} = 0.93$.

In Table 1, the MSE values and percent relative efficiencies (PREs) of all the estimators considered here are reported based on Populations 1-10.

We observe from Table 1 that:

1. The ratio estimator ($\hat{\bar{Y}}_R$) performs better than $\bar{y}$ in Populations 1, 3, 6-10 because the condition $\rho_{yx} > \frac{C_x}{2C_y}$ is satisfied. In other Populations 2, 4 and 5, its performance is poor.

2. The product estimator ($\hat{\bar{Y}}_P$) performs better than $\bar{y}$ in Population 5 because the condition $\rho_{yx} < -\frac{C_x}{2C_y}$ is satisfied.

3. The exponential ratio estimator ($\hat{\bar{Y}}_{BT,R}$) performs better than $\bar{y}$ in Populations 1-3, 6-10 because the condition $\rho_{yx} > \frac{C_x}{4C_y}$ is satisfied.

4. The exponential product estimator ($\hat{\bar{Y}}_{BT,P}$) performs better than $\bar{y}$ in Populations 4 and 5 because the condition $\rho_{yx} < -\frac{C_x}{4C_y}$ is satisfied.

5. It is also observed that, regardless of positive or negative correlation between the study and the auxiliary variable, the estimators, $\hat{\bar{Y}}_{Reg}$, $\hat{\bar{Y}}_{R,Reg}$, $\hat{\bar{Y}}_{GK}$ and $\hat{\bar{Y}}_P^*$, always perform better than the unbiased sample mean, ratio and product estimators considered here in all populations. Among all competitive estimators, the proposed estimator ($\hat{\bar{Y}}_P^*$) is preferable.

6. Conclusion

In this paper, we have suggested an improved difference-cum-exponential type estimator of the finite population mean in simple random sampling using information on a single auxiliary variable. Expressions for the bias and MSE of the proposed estimator are obtained under first order of approximation. Based on both the theoretical and numerical comparisons, we showed that the proposed estimator always performs better than the sample mean estimator, traditional ratio and product estimators, linear regression estimator, Bahl & Tuteja (1991) estimators, Rao (1991) estimator, and Grover & Kaur (2011) estimator. Hence, we recommend the use of the proposed estimator for a more efficient estimation of the finite population mean in simple random sampling.

Acknowledgments

The authors are thankful to the Editor-in-Chief and two anonymous referees for their valuable comments and suggestions that led to an improved version of the article.

Table 1: MSE values and PREs of different estimators with respect to $\bar{y}$.

Population		Estimators								
		$\bar{y}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_P$	$\hat{\bar{Y}}_{Reg}, \hat{\bar{Y}}_{S,RP}$	$\hat{\bar{Y}}_{BT,R}$	$\hat{\bar{Y}}_{BT,P}$	$\hat{\bar{Y}}_{R,Reg}$	$\hat{\bar{Y}}_{GK}$	$\hat{\bar{Y}}_P^*$
1	MSE	39.1829	8.7384	145.8014	8.355	14.4389	82.9704	8.3332	8.3012	8.2551
	PRE	100	448.3998	26.8742	468.975	271.3702	47.2252	470.2037	472.0147	474.6535
2	MSE	29.4854	96.4018	385.3576	4.7995	10.0951	154.5729	4.7909	4.4372	3.5644
	PRE	100	30.586	7.6514	614.345	292.0779	19.0754	615.4427	664.5098	827.2156
3	MSE	116.031	74.1901	449.4337	55.6631	58.6653	246.2871	53.7046	52.2123	50.3002
	PRE	100	156.3967	25.8171	208.4522	197.7846	47.1121	216.0543	222.2292	230.6769
4	MSE	0.0308	0.0989	0.0331	0.0231	0.056	0.0231	0.022	0.0216	0.021
	PRE	100	31.1051	92.9307	133.2623	54.9124	133.0384	139.7975	142.7234	146.3193
5	MSE	0.5676	1.0091	0.3387	0.3033	0.7618	0.4265	0.3018	0.3016	0.3015
	PRE	100	56.2431	167.5887	187.1024	74.5067	133.0649	188.0754	188.163	188.2545
6	MSE	5.4999	1.3871	18.2446	1.3847	2.3645	10.7933	1.3374	1.2933	1.2349
	PRE	100	396.4953	30.1454	397.18	232.6006	50.9568	411.2418	425.2586	445.3895
7	MSE	86226.1674	42781.393	212750.8436	42759.5564	54118.7925	139103.5178	40886.0545	40403.4109	39865.5124
	PRE	100	201.5506	40.5292	201.6536	159.3276	61.9871	210.8938	213.4131	216.2926
8	MSE	1.7511	1.1791	6.2503	0.9325	0.9742	3.5097	0.8979	0.8773	0.8519
	PRE	100	148.5052	28.0157	187.7743	179.747	49.8911	195.0081	199.5885	205.5391
9	MSE	0.0712	0.0504	0.1352	0.0504	0.0554	0.0978	0.0504	0.0504	0.0504
	PRE	100	141.1359	52.6256	141.1632	128.5059	72.7795	141.1876	141.1903	141.1931
10	MSE	122303.2646	29494.9337	600785.7109	16523.171	27689.8347	313335.2233	16270.6541	14996.4826	12647.4936
	PRE	100	414.6585	20.3572	740.1925	441.6901	39.0327	751.6801	815.5463	967.0158

[Recibido: octubre de 2013 — Aceptado: marzo de 2014]

References

Bahl, S. & Tuteja, R. (1991), 'Ratio and product type exponential estimators', *Journal of Information and Optimization Sciences* **12**(1), 159–164.

Cochran, W. G. (1940), 'The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce', *The Journal of Agricultural Science* **30**(02), 262–275.

Cochran, W. G. (1977), *Sampling Techniques*, 3 edn, John Wiley and Sons, New York.

Das, A. (1988), Contribution to the Theory of Sampling Strategies Based on Auxiliary Information, Ph.D. thesis, Bidhan Chandra Krishi Viswavidyalay, Nadia West Bengal, India.

Grover, L. K. & Kaur, P. (2011), 'An improved estimator of the finite population mean in simple random sampling', *Model Assisted Statistics and Applications* **6**(1), 47–55.

Grover, L. K. & Kaur, P. (2014), 'A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable', *Communications in Statistics-Simulation and Computation* **43**(7), 1552–1574.

Haq, A. & Shabbir, J. (2013), 'Improved family of ratio estimators in simple and stratified random sampling', *Communications in Statistics-Theory and Methods* **42**(5), 782–799.

Kalidar, C. & Cingi, H. (2005), 'A new estimator using two auxiliary variables', *Applied Mathematics and Computation* **162**(2), 901–908.

Kalidar, C. & Cingi, H. (2007), 'Improvement in variance estimation in simple random sampling', *Communication in Statistics-Theory and Methods* **36**, 2075–2081.

Maddala, G. S. (1977), *Econometrics*, Economics handbook series, McGraw Hills Publication Company, New York.

Murthy, M. (1964), 'Product method of estimation', *Sankhya A* **26**(1), 69–74.

Murthy, M. N. (1977), *Sampling: Theory and Methods*, Statistical Pub. Society.

Rao, T. (1991), 'On certain methods of improving ratio and regression estimators', *Communications in Statistics-Theory and Methods* **20**(10), 3325–3340.

Searls, D. T. (1964), 'The utilization of a known coefficient of variation in the estimation procedure', *Journal of the American Statistical Association* **59**(308), 1225–1226.

Singh, H. P., Sharma, B. & Tailor, R. (2014), 'Hartley-Ross type estimators for population mean using known parameters of auxiliary variate', *Communications in Statistics-Theory and Methods* **43**(3), 547–565.

Singh, R., Chauhan, P. & Sawan, N. (2008), 'On linear combination of ratio and product type exponential estimator for estimating the finite population mean', *Statistics in Transition* **9**(1), 105–115.

Singh, R., Chauhan, P., Sawan, N. & Smarandache, F. (2009), 'Improvement in estimating the population mean using exponential estimator in simple random sampling', *Bulletin of Statistics and Economics* **3**(13), 13–18.

Srivnstava, R. S., Srivastava, S. & Khare, B. (1989), 'Chain ratio type estimator for ratio of two population means using auxiliary characters', *Communications in Statistics-Theory and Methods* **18**(10), 3917–3926.

Steel, R., Torrie, J. & Dickey, D. (1960), *Principles and Procedures of Statistics*, McGraw-Hill Companies, Michigan.

Sukhatme, B. & Chand, L. (1977), Multivariate ratio-type estimators, *in* 'Proceedings of the Social Statistics Section', American Statistical Association, Michigan, pp. 927–931.

Yadav, S. K. & Kadilar, C. (2013), 'Improved exponential type ratio estimator of population variance', *Revista Colombiana de Estadística* **36**(1), 145–152.