Open Journal of Statistics, 2016, 6, 426-435
Published Online June 2016 in SciRes. http://www.scirp.org/journal/ojs
http://dx.doi.org/10.4236/ojs.2016.63038

Scientific Research Publishing

An Efficient Class of Estimators for the Finite Population Mean in Ranked Set Sampling

Lakhkar Khan[1,2], Javid Shabbir[2]

[1]Department of Statistics Government College Toru, Khyber Pukhtunkhwa, Pakistan
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
Email: lakhkarkhan.stat@gmail.com, javidshabbir@gmail.com

Received 14 February 2016; accepted 11 June 2016; published 14 June 2016

Abstract

In this paper, we propose a class of estimators for estimating the finite population mean of the study variable under Ranked Set Sampling (RSS) when population mean of the auxiliary variable is known. The bias and Mean Squared Error (MSE) of the proposed class of estimators are obtained to first degree of approximation. It is identified that the proposed class of estimators is more efficient as compared to [1] estimator and several other estimators. A simulation study is carried out to judge the performances of the estimators.

Keywords

Ranked Set Sampling, Auxiliary Variable, Bias, Mean Squared Error, Relative Efficiency

1. Introduction

The problem of estimation in the finite population mean has been widely considered by many authors in different sampling designs. In application, there may be a situation when the variable of interest cannot be measured easily or is very expensive to do so, but it can be ranked easily at no cost or at very little cost. In view of this situation, [2] introduced the Ranked Set Sampling (RSS) procedure. [3] proved the mathematical theory that the sample mean under RSS was an unbiased estimator of the finite population mean and more precise than the sample mean estimator under simple random sampling (SRS).

The auxiliary information plays an important role in increasing efficiency of the estimator. [4] suggested an estimator for population ratio in RSS and showed that it had less variance as compared to usual ratio estimator in simple random sampling (SRS).

In RSS, perfect ranking of elements was considered by [2] and [3] for estimation of population mean. In some situations, ranking may not be perfect. According to [5], the sample mean in RSS is an unbiased estimator of the

How to cite this paper: Khan, L. and Shabbir, J. (2016) An Efficient Class of Estimators for the Finite Population Mean in Ranked Set Sampling. *Open Journal of Statistics*, **6**, 426-435. http://dx.doi.org/10.4236/ojs.2016.63038

population mean regardless of errors in ranking of the elements. In [6], the ranking of elements was done on basis of the auxiliary variable instead of judgment. [1] suggested an estimator for population mean and ranking of the elements was observed on basis of the auxiliary variable. [7] had suggested a class of Hartley-Ross type unbiased estimators in RSS. [8] had also proposed unbiased estimators in RSS and stratified ranked set sampling.

In this paper, we suggest a class of estimators for the population mean, using known population mean of the auxiliary variable in RSS. It is shown that the proposed class of estimators outperforms as compared to the [9], [1] and several other estimators. Also some special cases of the proposed class are considered in **Table A1** (**Appendix**).

2. Ranked Set Sampling Procedure

In ranked set sampling (RSS), we select *m* random samples, each of size *m* units from the population, and rank the units within each sample with respect to a variable of interest. In order to facilitate the ranking, the design parameter *m*, is chosen to be small. From the first sample the unit having the lowest rank is selected, from the second sample the unit having second lowest rank is selected and the process is continued until from the last sample the unit having the highest rank is selected. In this way, we obtain *m* measured units, one from each sample. The cycle may be repeated *r* times until mr units have been measured. These $n = mr$ units form the RSS data.

Suppose that the variable of interest *Y* is difficult to measure and to rank, but there is the auxiliary variable *X*, which is correlated with *Y*. The variable *X* may be used to obtain the rank of *Y*. To perform the sampling procedure, *m* bivariate random samples, each of size *m* units are drawn from the population then each sample is ranked with respect to one of the variables *Y* or *X*. Here, we assume that the perfect ranking is done on basis of the auxiliary variable *X* while the ranking of *Y* is with error. An actual measurement from the first sample is then taken of the unit with the smallest rank of *X*, together with the variable *Y* associated with the smallest rank of *X*. From the second sample of size *m* the *Y* associated with the second smallest rank of *X* is measured. The process is continued until from the *m*th sample, the *Y* associated with the highest rank of *X* is measured. The cycle is repeated *r* times until $n = mr$ bivariate units have been measured out of the total $m^2 r$ selected units.

3. Some Existing Estimators and Notations

We consider a situation when rank the elements on the auxiliary variable. Let $\left(y_{[i]j}, x_{(i)j}\right)$ be the *i*th judgment ordering in the *i*th set for the study variable *Y* based on the *i*th order statistics of the *i*th set of the auxiliary variable *X* at the *j*th cycle. Based on RSS, the sample mean estimator $\left(\bar{y}_{RSS}\right)$ of the population mean $\left(\bar{Y}\right)$, is given by

$$\bar{y}_{RSS} = \bar{y}_{[rss]}, \tag{1}$$

where $\bar{y}_{[rss]} = \left(1/mr\right)\sum_{j=1}^{r}\sum_{i=1}^{m} y_{[i]j}$.

To obtain the bias and MSE of estimators, we define:

$$\bar{y}_{[rss]} = \bar{Y}\left(1+e_0\right), \bar{x}_{(rss)} = \bar{X}\left(1+e_1\right),$$

such that

$$E\left(e_0\right) = E\left(e_1\right) = 0\,,$$

and

$$E\left(e_0^2\right) = \gamma C_y^2 - W_y^2\,,\quad E\left(e_1^2\right) = \gamma C_x^2 - W_x^2\,,\quad E\left(e_0 e_1\right) = \gamma\rho C_y C_x - W_{yx}\,,$$

where

$$W_{yx} = \frac{1}{m^2 r\bar{X}\bar{Y}}\sum_{i=1}^{m}\tau_{yx(i)}, W_x^2 = \frac{1}{m^2 r\bar{X}^2}\sum_{i=1}^{m}\tau_{x(i)}^2, W_y^2 = \frac{1}{m^2 r\bar{Y}^2}\sum_{i=1}^{m}\tau_{y[i]}^2,$$

$\tau_{x(i)} = \left(\mu_{x(i)} - \bar{X}\right)$, $\tau_{y[i]} = \left(\mu_{y[i]} - \bar{Y}\right)$, $\tau_{yx(i)} = \left(\mu_{y[i]} - \bar{Y}\right)\left(\mu_{x(i)} - \bar{X}\right)$, $C_{yx} = \rho C_y C_x$. C_y and C_x are the coefficients of variation of Y and X respectively. It also be noted that the values of $\mu_{y[i]}$ and $\mu_{x(i)}$ are the means of ith order statistics from some specific distributions (see [10]).

The variance of $\bar{y}_{RSS}$ under RSS scheme, is given by

$$Var\left(\bar{y}_{RSS}\right)=\bar{Y}^2\left(\gamma C_y^2-W_y^2\right). \tag{2}$$

[4] proposed an estimator of the population ratio $R=\frac{\bar{Y}}{\bar{X}}$ under RSS as:

$$\hat{R}_{RSS}=\frac{\bar{y}_{[rss]}}{\bar{x}_{(rss)}}. \tag{3}$$

When population mean ($\bar{X}$) of the auxiliary variable (X) is known, and the variables Y and X are positively correlated, [9] proposed the ratio estimator for population mean ($\bar{Y}$) based on RSS as

$$\bar{y}_{rRSS}=\frac{\bar{y}_{[rss]}}{\bar{x}_{(rss)}}\bar{X}. \tag{4}$$

The bias and MSE of $\bar{y}_{rRSS}$, up to the first degree of approximation, are given by

$$Bias\left(\bar{y}_{rRSS}\right)=\bar{Y}\left[\gamma\left(C_x^2-\rho C_y C_x\right)-\left(W_x^2-W_{yx}\right)\right] \tag{5}$$

and

$$MSE\left(\bar{y}_{rRSS}\right)\cong\bar{Y}^2\left[\gamma\left(C_y^2+C_x^2-2\rho C_y C_x\right)-\left(W_y^2+W_x^2-2W_{yx}\right)\right]. \tag{6}$$

When population mean ($\bar{X}$) of the auxiliary variable (X) is known, and the variables Y and X are negatively correlated, then the product estimator based on RSS is defined as:

$$\bar{y}_{pRSS}=\bar{y}_{[rss]}\frac{\bar{x}_{(rss)}}{\bar{X}}. \tag{7}$$

The bias and MSE of $\bar{y}_{pRSS}$, up to the first degree of approximation, are given by

$$Bias\left(\bar{y}_{pRSS}\right)=\bar{Y}\left(\gamma\rho C_y C_x-W_{yx}\right) \tag{8}$$

and

$$MSE\left(\bar{y}_{pRSS}\right)\cong\bar{Y}^2\left\{\gamma\left(C_y^2+C_x^2+2\rho C_y C_x\right)-\left(W_y^2+W_x^2+2W_{yx}\right)\right\}. \tag{9}$$

[11] suggested an estimator under RSS and is defined as:

$$\bar{y}_{sRSS}=\lambda\bar{y}_{[rss]}, \tag{10}$$

where λ is suitably chosen constant.

The minimum bias and MSE of $\bar{y}_{sRSS}$ at optimum value of λ *i.e.*

$$\lambda_{(opt)}=\frac{1}{\left(1+\gamma C_y^2-W_y^2\right)}$$

are given by

$$Bias\left(\bar{y}_{sRSS}\right)_{min}=-\frac{\bar{Y}\left(\gamma C_y^2-W_y^2\right)}{\left(1+\gamma C_y^2-W_y^2\right)} \tag{11}$$

and

$$MSE\left(\bar{y}_{sRSS}\right)_{min}\cong\frac{\bar{Y}^2\left(\gamma C_y^2-W_y^2\right)}{\left(1+\gamma C_y^2-W_y^2\right)}. \tag{12}$$

The difference-type estimator for population mean ($\bar{Y}$) based on RSS, is given by

$$\bar{y}_{d(RSS)} = \bar{y}_{[rss]} + d\left(\bar{X} - \bar{x}_{(rss)}\right), \tag{13}$$

where d is a constant.

The minimum variance of $\bar{y}_{d(RSS)}$ at optimum value of d *i.e.*

$$d_{(opt)} = \frac{R\left(\gamma C_{yx} - W_{yx}\right)}{\left(\gamma C_x^2 - W_x^2\right)}$$

is given by

$$Var\left(\bar{y}_{dRSS}\right)_{min} \cong \bar{Y}^2\left[\gamma C_y^2 - W_y^2 - \frac{\left(\gamma C_{yx} - W_{yx}\right)^2}{\left(\gamma C_x^2 - W_x^2\right)}\right]. \tag{14}$$

Following [12], [1] suggested a class of estimators of the population mean ($\bar{Y}$), based on RSS as:

$$\bar{y}_{S(RSS)} = \lambda_1\bar{y}_{[rss]} + \lambda_2\bar{y}_{[rss]}\left[\frac{a\bar{X} + b}{\alpha\left(a\bar{x}_{(rss)} + b\right) + (1-\alpha)\left(a\bar{X} + b\right)}\right]^g, \tag{15}$$

where α is a suitably chosen constant, a and b are either real numbers or functions of known parameters of the auxiliary variable X, g is a scalar which takes value of 1 (for generating ratio-type estimators) and -1 (for generating product-type estimators) and $\left(\lambda_1, \lambda_2\right)$ are constants whose sum need not be unity.

The bias of $\bar{y}_{S(RSS)}$, is given by

$$Bias\left(\bar{y}_{S(RSS)}\right) = \bar{Y}\left[(\lambda_1 + \lambda_2 - 1) + \lambda_2 g\alpha^2\theta^2\left(\frac{g+1}{2}\right)\left(\gamma C_x^2 - W_x^2\right) - \lambda_2 g\alpha\theta\left(\gamma\rho C_Y C_x - W_{yx}\right)\right]. \tag{16}$$

The MSE of $\bar{y}_{S(RSS)}$, to first degree of approximation, is given by

$$MSE\left(\bar{y}_{S(RSS)}\right) \cong \bar{Y}^2\left[1 + \lambda_1^2\left(A_s - A_w\right) + \lambda_2^2\left(B_s - B_w\right) + 2\lambda_1\lambda_2\left(C_s - C_w\right) - 2\lambda_1 - 2\lambda_2\left(D_s - D_w\right)\right], \tag{17}$$

where

$$A_s = \left(1 + \gamma C_y^2\right),$$

$$A_w = W_y^2,$$

$$B_s = 1 + \gamma\left\{\left(C_y^2 + g\left(2g+1\right)\theta^2\alpha^2C_x^2 - 4g\alpha\theta C_y x\right)\right\},$$

$$B_w = W_y^2 + g\left(2g+1\right)\theta^2\alpha^2W_x^2 - 4g\alpha\theta W_{yx},$$

$$C_s = 1 + \gamma\left(C_y^2 - 2g\theta\alpha C_{yx} + \frac{g\left(g+1\right)}{2}\theta^2\alpha^2C_x^2\right),$$

$$C_w = W_y^2 - 2g\theta\alpha C_{yx} + \frac{g\left(g+1\right)}{2}\theta^2\alpha^2W_x^2,$$

$$D_s = 1 + \gamma\left\{\frac{g\left(g+1\right)}{2}\theta^2\alpha^2C_x^2 - g\theta\alpha C_{yx}\right\},$$

$$D_w = \frac{g\left(g+1\right)}{2}\theta^2\alpha^2W_x^2 - g\theta\alpha W_{yx}.$$

We discuss two cases.

Case 1: Sum of weights is unity (*i.e.* $\lambda_1 + \lambda_2 = 1$).

Solving (17), the optimum value of λ_1, is given by

$$\lambda_{1(opt)} = \frac{\left[1+\left(B_s - B_w\right)-\left(C_s - C_w\right)-\left(D_s - D_w\right)\right]}{\left[\left(A_s - A_w\right)+\left(B_s - B_w\right)-2\left(C_s - C_w\right)\right]}.$$

Substituting $\lambda_{1(opt)}$ in (17), we get the minimum MSE of $\bar{y}_{S(RSS)1}$, given by

$$MSE\left(\bar{y}_{S(RSS)1}\right)_{min} \cong \bar{Y}^2\left[1+\left(B_s - B_w\right)-2\left(D_s - D_w\right)-\frac{\left(1+B_s - B_w - C_s + C_w - D_s + D_w\right)^2}{\left(A_s - A_w\right)+\left(B_s - B_w\right)-2\left(C_s - C_w\right)}\right]. \tag{18}$$

Case 2: Sum of weights is flexible (*i.e.* $\lambda_1 + \lambda_2 \neq 1$).

Solving (17), the optimum values of λ_1 and λ_2 are given by

$$\lambda_{1(opt)} = \frac{\left[\left(B_s - B_w\right)-\left(C_s - C_w\right)\left(D_s - D_w\right)\right]}{\left[\left(A_s - A_w\right)\left(B_s - B_w\right)-\left(C_s - C_w\right)^2\right]}$$

and

$$\lambda_{2(opt)} = \frac{\left[\left(A_s - A_w\right)\left(D_s - D_w\right)-\left(C_s - C_w\right)\right]}{\left[\left(A_s - A_w\right)\left(B_s - B_w\right)-\left(C_s - C_w\right)^2\right]}.$$

Substituting the optimum values of λ_1 and λ_2 in (17), we get

$$MSE\left(\bar{y}_{S(RSS)2}\right)_{min} \cong \bar{Y}^2\left[1-\frac{\left\{\left(B_s - B_w\right)-2\left(C_s - C_w\right)\left(D_s - D_w\right)+\left(A_s - A_w\right)\left(D_s - D_w\right)^2\right\}}{\left\{\left(A_s - A_w\right)\left(B_s - B_w\right)-\left(C_s - C_w\right)^2\right\}}\right]. \tag{19}$$

4. Proposed Class of Estimators

Following [1] and [12], we propose a class of estimators of the population mean ($\bar{Y}$), under RSS as

$$\bar{y}_{L(RSS)} = \left[k_1\bar{y}_{[rss]} + k_2\left(\bar{X} - \bar{x}_{(rss)}\right)\right]\left[\alpha\left\{exp\frac{\left(a\bar{X}+b\right)-\left(a\bar{x}_{(rss)}+b\right)}{\left(a\bar{X}+b\right)+\left(a\bar{x}_{(rss)}+b\right)}\right\}+\left(1-\alpha\right)\frac{\left(a\bar{X}+b\right)}{\left(a\bar{x}_{(rss)}+b\right)}\right], \tag{20}$$

where α is a suitably chosen constant, *a* and *b* are either real numbers or the functions of known parameters of the auxiliary variable *X* and $\left(k_1, k_2\right)$ are constants whose sum need not be unity. From (20) we can generate a large number of estimators for the different values of the constants (**Table A1** in **Appendix**). The proposed estimator $\bar{y}_{L(RSS)}$ can be written in terms of e_0 and e_1 as

$$\bar{y}_{L(RSS)} = \left[k_1\bar{Y}\left(1+e_0\right)-k_2\bar{X}e_1\right]\left[\alpha\left(1-\frac{\theta e_1}{2}+\frac{3\theta^2 e_1^2}{8}\right)+\left(1-\alpha\right)\left(1+\theta e_1\right)^{-1}\right], \tag{21}$$

where $\theta = \dfrac{a\bar{X}}{\left(a\bar{X}+b\right)}$.

Solving (21), we have

$$\begin{aligned}\bar{y}_{L(RSS)} - \bar{Y} = \Bigg[&\bar{Y}\left(k_1 - 1\right) - k_1\bar{Y}\left(1-\frac{\alpha}{2}\right)\theta e_1 + \left(1-\frac{5\alpha}{8}\right)k_1\bar{Y}\theta^2 e_1^2 + k_1\bar{Y}e_0 \\ &- k_1\bar{Y}\left(1-\frac{\alpha}{2}\right)\theta e_0 e_1 - k_2\bar{X}e_1 + k_2\bar{X}\left(1-\frac{\alpha}{2}\right)\theta e_1^2\Bigg].\end{aligned} \tag{22}$$

Taking expectation of both sides of above equation, we get bias of $\bar{y}_{L(RSS)}$, given by

$$\begin{aligned}Bias\left(\bar{y}_{L(RSS)}\right) = {}&\bar{Y}\left(k_1 - 1\right) + k_1\bar{Y}\theta^2\left(1-\frac{5\alpha}{8}\right)\left(\gamma C_x^2 - W_x^2\right) - k_1\bar{Y}\theta\left(1-\frac{\alpha}{2}\right)\left(\gamma C_{yx} - W_{yx}\right) \\ &+ k_2\bar{X}\theta\left(1-\frac{\alpha}{2}\right)\left(\gamma C_x^2 - W_x^2\right).\end{aligned} \tag{23}$$

Squaring both sides of Equation (22) and ignoring higher order terms of e's, we have

$$\left(\bar{y}_{L(RSS)}-\bar{Y}\right)^2=\bar{Y}^2\left(k_1-1\right)^2+k_1^2\bar{Y}^2\left\{e_0^2+\left(1-\frac{\alpha}{2}\right)^2\theta^2e_1^2-2\left(1-\frac{\alpha}{2}\right)\theta e_0e_1\right\}$$
$$+k_2^2\bar{X}^2e_1^2+2k_1\left(k_1-1\right)\bar{Y}^2\left\{\left(1-\frac{5\alpha}{8}\right)\theta^2e_1^2-\left(1-\frac{\alpha}{2}\right)\theta e_1\right\}$$
$$+2k_2\left(k_1-1\right)\overline{YX}\left\{\left(1-\frac{\alpha}{2}\right)\theta e_1^2-e_1\right\}+2k_1k_2\overline{YX}\left\{\left(1-\frac{\alpha}{2}\right)\theta e_1^2-e_0e_1\right\}.$$

Taking expectation of both sides of above equation, we obtain the MSE of $\bar{y}_{L(RSS)}$ as given by

$$MSE\left(\bar{y}_{L(RSS)}\right)=\bar{Y}^2\left(k_1-1\right)^2+k_1^2\left(E_s-E_w\right)+k_2^2\left(F_s-F_w\right)+2k_1\left(k_1-1\right)\left(G_s-G_w\right)$$
$$+2k_2\left(k_1-1\right)\left(H_s-H_w\right)+2k_1k_2\left(I_s-I_w\right), \tag{24}$$

where

$$E_s=\bar{Y}^2\gamma\left[C_y^2+\left(1-\frac{\alpha}{2}\right)^2\theta^2C_x^2-2\left(1-\frac{\alpha}{2}\right)\theta C_{yx}\right],$$

$$E_w=\bar{Y}^2\left[W_y^2+\left(1-\frac{\alpha}{2}\right)^2\theta^2W_x^2-2\left(1-\frac{\alpha}{2}\right)\theta W_{yx}\right],$$

$$F_s=\bar{X}^2\gamma C_x^2,$$

$$F_w=\bar{X}^2W_x^2,$$

$$G_s=\bar{Y}^2\gamma\left[\left(1-\frac{5\alpha}{8}\right)\theta^2C_x^2-\left(1-\frac{\alpha}{2}\right)\theta C_{yx}\right],$$

$$G_w=\bar{Y}^2\left[\left(1-\frac{5\alpha}{8}\right)\theta^2W_x^2-\left(1-\frac{\alpha}{2}\right)\theta W_{yx}\right],$$

$$H_s=\bar{X}\bar{Y}\gamma\left(1-\frac{\alpha}{2}\right)\theta C_x^2,$$

$$H_w=\bar{X}\bar{Y}\left(1-\frac{\alpha}{2}\right)\theta W_x^2,$$

$$I_s=\bar{X}\bar{Y}\gamma\left[\left(1-\frac{\alpha}{2}\right)\theta C_x^2-C_{yx}\right],$$

$$I_w=\bar{X}\bar{Y}\left[\left(1-\frac{\alpha}{2}\right)\theta W_{x(i)}^2-W_{yx}\right].$$

We discuss two cases.

Case 1: Sum of weights is unity (*i.e.* $k_1+k_2=1$).

The optimum value of k_1, is given by

$$k_{1(opt)}=\frac{\left[\bar{Y}^2+\left(F_s-F_w\right)+\left(G_s-G_w\right)-2\left(H_s-H_w\right)-\left(I_s-I_w\right)\right]}{\left[\bar{Y}^2+\left(E_s-E_w\right)+\left(F_s-F_w\right)+2\left(G_s-G_w\right)-2\left(H_s-H_w\right)-2\left(I_s-I_w\right)\right]}$$

Thus, the minimum MSE of $\bar{y}_{L(RSS)}$, is given by

$$MSE\left(\bar{y}_{L(RSS)1}\right)_{min}\cong\frac{\left(E_s-E_w\right)\left\{\bar{Y}^2-2\left(H_s-H_w\right)+\left(F_s-F_w\right)\right\}-\left\{\left(I_s-I_w\right)-\left(G_s-G_w\right)\right\}^2}{\left[\bar{Y}^2+\left(E_s-E_w\right)+\left(F_s-F_w\right)+2\left(G_s-G_w\right)-2\left(H_s-H_w\right)-2\left(I_s-I_w\right)\right]} \tag{25}$$

Case 2: Sum of weights is flexible (*i.e.* $k_1 + k_2 \neq 1$).

For $(k_1 + k_2 \neq 1)$, the *MSE* of $\bar{y}_{L(RSS)}$ in Equation (24) is minimized for

$$k_{1(opt)} = \frac{(F_s - F_w)\{\bar{Y}^2 + (G_s - G_w)\} - (H_s - H_w)\{(H_s - H_w) + (I_s - I_w)\}}{(F_s - F_w)\{\bar{Y}^2 + 2(G_s - G_w) + (E_s - E_w)\} - \{(H_s - H_w) + (I_s - I_w)\}^2}$$

and

$$k_{2(opt)} = \frac{(H_s - H_w)\{(E_s - E_w) + (G_s - G_w)\} - (I_s - I_w)\{\bar{Y}^2 + (G_s - G_w)\}}{(F_s - F_w)\{\bar{Y}^2 + 2(G_s - G_w) + (E_s - E_w)\} - \{(H_s - H_w) + (I_s - I_w)\}^2}.$$

Substituting the optimum values of k_1 and k_2 in (24), we get

$$\begin{aligned} MSE\left(\bar{y}_{L(RSS)2}\right)_{min} &= \bar{Y}^2\left(k_{1(opt)} - 1\right)^2 + k_{1(opt)}^2(E_s - E_w) + k_{2(opt)}^2(F_s - F_w) \\ &\quad + 2k_{1(opt)}\left(k_{1(opt)} - 1\right)(G_s - G_w) + 2k_{2(opt)}\left(k_{1(opt)} - 1\right)(H_s - H_w) \\ &\quad + 2k_{1(opt)}k_{2(opt)}(I_s - I_w). \end{aligned} \tag{26}$$

Note: It is difficult to make the theoretical comparison due to complexity, therefore we adopt the numerical study.

5. Simulation Study

We use the same data set as earlier used by [1], and perform some simulation study to investigate the performances of the estimators.

Population (source: [13]).

Y = Number of acres devoted to farms during 1992 (ACRES92).

X = Number of large farms during 1992 (LARGEF92).

$$N = 3059 \qquad \rho_{yx} = 0.677428$$
$$\bar{Y} = 308582.4 \qquad \bar{X} = 56.5$$
$$S_y = 425312.8 \qquad S_x = 72.3$$

We set $r = 10$ and $m = 5$ to select a sample of $n = mr = 50$ units from the population of size $N = 3059$. To compute the values of W_y^2, W_x^2 and W_{yx} by simulation, we explain our simulation methodology as follow.

Here W_y^2, W_x^2 and W_{yx} can be written as

$$W_y^2 = \frac{1}{m^2 r}\sum_{i=1}^{m}\left(RDY[i] - 1\right)^2,$$

$$W_x^2 = \frac{1}{m^2 r}\sum_{i=1}^{m}\left(RDX(i) - 1\right)^2,$$

and

$$W_{yx} = \frac{1}{m^2 r}\sum_{i=1}^{m}\left(RDX(i) - 1\right)\left(RDY[i] - 1\right),$$

where

$$RDY[i] = \frac{\mu_{y[i]}}{\bar{Y}} \quad \text{and} \quad RDX(i) = \frac{\mu_{x(i)}}{\bar{X}}, \; i = 1, 2, \cdots, m.$$

To find the possible values of the ratio $RDY[i]$ for $m = 5$, we generate $e_i \sim N(0,1)$ and calculate $RDY[1] = 0.25 + 0.08e_1$, $RDY[2] = 0.50 + 0.08e_2$, $RDY[3] = 1.00 + 0.08e_3$, $RDY[4] = 1.25 + 0.08e_4$, and $RDY[5] = 1.75 + 0.08e_5$. It means that when the first smallest value is selected from the ranked set sample, the expected ratio of that value to the population mean could be close to 0.25, and when the second smallest value is selected the ratio of that value to the population mean could be close to 0.50, and when the third smallest value is selected the expected ratio of that value to the population mean will close to 1. Similarly, the expected ratio of the fourth and fifth values could be close to 1.25 and 1.75 respectively. In each case we weighted error term e_i with a small number 0.08 to make sure that the ratio $RDY[i]$ remains positive. In other words, it means that we are generating $e_i \sim N(0, 0.08)$. Thus, the possible values of the ratio $RDY[i]$ are expected to remain close to those we are considering here. Similarly, for the possible values of the ratio $RDX(i)$, we consider $RDX(1) = 0.25 + 0.05e_1$, $RDX(2) = 0.50 + 0.05e_2$, $RDX(3) = 1.00 + 0.05e_3$, $RDX(4) = 1.25 + 0.05e_4$, and $RDX(5) = 1.75 + 0.05e_5$, where $e_i \sim N(0,1)$. Here we weighted e_i with a small number 0.05 because it may be less risky to rank the auxiliary variable *X* than the study variable *Y*. Thus the values of $W^2_{y[i]}$, $W^2_{x(i)}$, and $W_{yx(i)}$ are obtained through this simulation and are represented in the last three columns of **Table 1**.

Table 1. PREs of proposed class of estimators through simulation.

a	b	g	α	$R(0,1)$	$R(0,2)$	$R(0,3)$	$R(0,4)$	$R(0,5)$	$R(0,6)$	$W^2_{x(i)}$	$W^2_{y(i)}$	$W_{yx(i)}$
−1.5	−1.5	−1	0.1	140.6	103.2	160.9	161.4	153.2	**164.5**	0.00573	0.00574	0.00573
−1.5	−1.5	−1	0.5	139.3	103.2	159.9	160.5	163.8	**164.4**	0.00590	0.00604	0.00596
−1.5	−1.5	−1	0.9	148.1	103.4	167.1	167.5	165.5	**171.8**	0.00462	0.00404	0.00431
−1.5	−1.5	1	0.1	144.5	103.3	164.1	164.8	157.3	**167.8**	0.00516	0.00485	0.00499
−1.5	−1.5	1	0.5	132.4	103.0	154.5	156.8	157.5	**158.7**	0.00689	0.00764	0.00725
−1.5	−1.5	1	0.9	144.6	103.3	164.2	168.6	163.4	**168.8**	0.00514	0.00482	0.00497
−1.5	0	−1	0.1	136.9	103.1	158.0	158.6	148.0	**161.5**	0.00625	0.00658	0.00641
−1.5	0	−1	0.5	137.6	103.1	158.6	159.2	162.0	**163.0**	0.00615	0.00642	0.00628
−1.5	0	−1	0.9	142.5	103.3	162.4	162.9	162.7	**167.0**	0.00546	0.00530	0.00538
−1.5	0	1	0.1	130.1	103.0	152.8	153.7	141.0	**156.1**	0.00520	0.00816	0.00766
−1.5	0	1	0.5	140.9	103.2	161.2	163.5	164.9	**165.7**	0.00568	0.00567	0.00567
−1.5	0	1	0.9	137.2	103.1	158.3	162.7	159.5	**162.9**	0.00620	0.00651	0.00635
−1.5	1.5	−1	0.1	140.8	103.2	161.1	161.6	150.5	**164.8**	0.00569	0.00570	0.00569
−1.5	1.5	−1	0.5	140.3	103.2	160.7	161.2	164.1	**165.3**	0.00576	0.00582	0.00578
−1.5	1.5	−1	0.9	135.2	103.1	156.7	157.3	158.7	**161.1**	0.00649	0.00697	0.00673
−1.5	1.5	1	0.1	138.2	103.2	159.1	159.9	147.8	**162.7**	0.00605	0.00629	0.00616
−1.5	1.5	1	0.5	139.2	103.2	159.8	162.2	163.0	**164.3**	0.00592	0.00602	0.00598
−1.5	1.5	1	0.9	143.3	103.3	163.1	168.0	163.9	**168.2**	0.00533	0.00513	0.00522
1.5	−1.5	−1	0.1	133.4	103.1	155.4	156.0	142.9	**158.8**	0.00672	0.00743	0.00706
1.5	−1.5	−1	0.5	140.8	103.2	161.1	161.6	164.5	**165.7**	0.00569	0.00578	0.00576
1.5	−1.5	−1	0.9	140.3	103.2	160.8	161.3	162.0	**165.4**	0.00575	0.00578	0.00576
1.5	−1.5	1	0.1	142.3	103.2	162.4	163.1	152.1	**166.1**	0.00546	0.00540	0.00541
1.5	−1.5	1	0.5	145.3	103.3	164.7	167.1	168.7	**169.5**	0.00504	0.00467	0.00484
1.5	−1.5	1	0.9	139.1	103.2	159.9	164.3	161.1	**164.6**	0.00592	0.00605	0.00598
1.5	0	−1	0.1	133.4	103.0	155.4	156.0	144.4	**158.8**	0.00672	0.00743	0.00658
1.5	0	−1	0.5	140.8	103.2	161.1	161.6	164.8	**165.6**	0.00569	0.00568	0.00566
1.5	0	−1	0.9	145.9	103.3	165.2	165.6	164.8	**169.6**	0.00496	0.00453	0.00473
1.5	0	1	0.1	142.3	103.3	162.4	163.1	153.6	**166.1**	0.00545	0.00540	0.00540
1.5	0	1	0.5	141.6	103.2	161.8	164.1	165.6	**166.3**	0.00557	0.00551	0.00553
1.5	0	1	0.9	140.3	103.2	160.7	165.1	161.4	**165.3**	0.00576	0.00582	0.00578
1.5	1.5	−1	0.1	139.2	103.2	159.9	160.4	151.8	**163.5**	0.00591	0.00605	0.00597
1.5	1.5	−1	0.5	133.0	103.0	155.1	155.7	158.1	**159.2**	0.00679	0.00749	0.00713
1.5	1.5	−1	0.9	137.3	103.1	158.4	158.9	159.0	**162.7**	0.00619	0.00650	0.00634
1.5	1.5	1	0.1	141.7	103.2	161.9	162.4	154.4	**165.5**	0.00555	0.00551	0.00520
1.5	1.5	1	0.5	142.3	103.3	162.3	164.6	166.4	**166.8**	0.00548	0.00534	0.00540
1.5	1.5	1	0.9	135.2	103.1	156.8	161.0	157.8	**161.2**	0.00648	0.00701	0.00672

We investigate the percentage relative efficiency (*PRE*) of ratio estimator $\bar{y}_{rRSS} = \hat{\theta}_1$ (say), the Searls estimator $\bar{y}_{sRSS} = \hat{\theta}_2$, the difference estimator $\bar{y}_{dRSS} = \hat{\theta}_3$, [1] estimator $\bar{y}_{S(RSS)} = \hat{\theta}_4$ when $\lambda_1 + \lambda_2 \neq 1$ with respect to conventional estimator $\bar{y}_{RSS} = \hat{\theta}_0$ (say). We also calculate PRE of the proposed class of estimators, say, $\bar{y}_{L(RSS)1} = \hat{\theta}_5$ when $(k_1 + k_2 = 1)$ and when $(k_1 + k_2 \neq 1)$, say, $\bar{y}_{L(RSS)2} = \hat{\theta}_6$, with respect to $\bar{y}_{RSS} = \hat{\theta}_0$. The PRE of our proposed estimator and other existing estimators $\hat{\theta}_j$, $j = 1, 2, \cdots, 6$, with respect to conventional estimator $\bar{y}_{RSS} = \hat{\theta}_0$, is defined as

$$PRE\left(\hat{\theta}_0, \hat{\theta}_j\right) = \frac{MSE\left(\hat{\theta}_0\right)}{MSE\left(\hat{\theta}_j\right)} \times 100, \quad j = 1, 2, \cdots, 6. \tag{27}$$

The *PREs* of our proposed estimator and other existing estimators with respect to conventional estimator are given in **Table 1**.

6. Conclusions

Since a, b, g and α are the fixed constants in [1] estimator and in the proposed class of estimators. There can be a large number of combinations for different values of these constants. Here, only limited number of results are reported in **Table 1**. Obviously, it can be observed through the simulation study in **Table 1**, that the proposed class of estimators is more efficient than all considered estimators. Its PRE increases from 164.5 to 171.8 when α changes from 0.1 to 0.9 but decreases slightly when α is close to 0.5. Generally, we can say *PRE* of proposed class increases as value of α increases for fixed values of constants *a*, *b* and *g* [1]. Class of estimators has maximum PRE 167.5, but it is less efficient as compared to the proposed class of estimators for all the choices of constants reported in **Table 1**. Also from the **Table 1**, we can see that other competitor estimators are also less efficient than the proposed class of estimators. If we make comparison between the two proposed cases then the class of estimators in Case 2 $(k_1 + k_2 \neq 1)$ is more precise than the Case 1 $(k_1 + k_2 = 1)$. We can see from **Table 1** that by fixing the values of *a* and *b* at -1.5, the proposed classes of estimators give more precise results when the value of α is away form 0.5, either close to 0 or 1. While by fixing positive values of the constants *a* and *b*, we get more precise results for α close to 0.5.

Therefore, the proposed class of estimators can be preferred over its competitive estimators in application under *RSS*.

Acknowledgements

The authors wish to thank the editor and the anonymous referees for their suggestions which led to improvement in the earlier version of the manuscript.

References

[1] Singh, H.P., Tailor, R. and Singh, S. (2014) General Procedure for Estimating the Population Mean Using Ranked Set Sampling. *Journal of Statistical Computation and Simulation*, **84**, 931-945. http://dx.doi.org/10.1080/00949655.2012.733395

[2] Mclntyre, G. (1952) A Method for Unbiased Selective Sampling, Using Ranked Sets. *Crop and Pasture Science*, **3**, 385-390. http://dx.doi.org/10.1071/AR9520385

[3] Takahasi, K. and Wakimoto, K. (1968) On Unbiased Estimates of the Population Mean Based on the Sample Stratified by Means of Ordering. *Annals of the Institute of Statistical Mathematics*, **20**, 1-31. http://dx.doi.org/10.1007/BF02911622

[4] Samawi, H.M. and Muttlak, M.A. (1996) Estimation of Ratio Using Ranked Set Sampling. *Biometrical Journal*, **38**, 753-764. http://dx.doi.org/10.1002/bimj.4710380616

[5] Dell, T. and Clutter, J. (1972) Ranked Set Sampling Theory with Order Statistics Background. *Biometrics*, 545-555. http://dx.doi.org/10.2307/2556166

[6] Stokes, S.L. (1977) Ranked Set Sampling with Concomitant Variables. *Communication in Statistics*: *Theory and Methods*, **6**, 1207-1211. http://dx.doi.org/10.1080/03610927708827563

[7] Khan, L. and Shabbir, J. (2015) A Class of Hartley-Ross Type Unbiased Estimators for Population Mean Using Ranked Set Sampling. *Hacettepe Journal of Mathematics and Statistics*. http://dx.doi.org/10.15672/HJMS.20156210579

[8] Khan, L. and Shabbir, J. (2016) Hartley-Ross Type Unbiased Estimators Using Ranked Set Sampling and Stratified Ranked Set Sampling. *North Carolina Journal of Mathematics and Statistics*, **2**, 10-22.

[9] Kadilar, C., Unyazici, Y. and Cingi, H. (2009) Ratio Estimator for the Population Mean Using Ranked Set Sampling. *Statistical Papers*, **50**, 301-309. http://dx.doi.org/10.1007/s00362-007-0079-y

[10] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (2012) A First Course in Order Statistics. Vol. 54, Siam.

[11] Searls, D.T. (1964) The Utilization of a Known Coefficient of Variation in the Estimation Procedure. *Journal of the American Statistical Association*, **59**, 1225-1226. http://dx.doi.org/10.1080/01621459.1964.10480765

[12] 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 Journal of Theoretical Statistics*, **22**, 181-191.

[13] Lohr, S. (1999) Sampling: Design and Analysis. Duxbury Press, Boston.

Appendix

Table A1. Some special cases of the proposed class of estimators.

k_1	k_2	α	a	b	Estimator	Remarks
1	0	0	0	1	$\bar{y}_{s(RSS)} = \bar{y}_{[rss]}$	Usual RSS mean estimator
1	0	0	1	0	$\bar{y}_{r(RSS)} = \bar{y}_{[rss]}\left(\dfrac{\bar{X}}{\bar{x}_{(rss)}}\right)$	Usual RSS ratio estimaotr
λ	0	0	1	0	$\bar{y}_{sr(RSS)} = \lambda\bar{y}_{[rss]}\left(\dfrac{\bar{X}}{\bar{x}_{(rss)}}\right)$	Kadilar *et al*. (2009) ratio type estimator
1	β	0	0	1	$\bar{y}_{reg(RSS)} = \left[\bar{y}_{[rss]} + \beta\left(\bar{X} - \bar{x}_{(rss)}\right)\right]$	Regression type estimator
1	k_2	0	0	1	$\bar{y}_{d(RSS)} = \left[\bar{y}_{[rss]} + k_2\left(\bar{X} - \bar{x}_{(rss)}\right)\right]$	Difference type estimator
1	k_2	0	1	0	$\bar{y}_{dr(RSS)} = \left[\bar{y}_{[rss]} + k_2\left(\bar{X} - \bar{x}_{(rss)}\right)\right]\left(\dfrac{\bar{X}}{\bar{x}_{(rss)}}\right)$	Difference-ratio estimator
k_1	k_2	0	1	0	$\bar{y}_{gdr(RSS)} = \left[k_1\bar{y}_{[rss]} + k_2\left(\bar{X} - \bar{x}_{(rss)}\right)\right]\left(\dfrac{\bar{X}}{\bar{x}_{(rss)}}\right)$	Generalied difference-ratio estimator
1	β	0	1	0	$\bar{y}_{regr(RSS)} = \left[\bar{y}_{[rss]} + \beta\left(\bar{X} - \bar{x}_{(rss)}\right)\right]\left(\dfrac{\bar{X}}{\bar{x}_{(rss)}}\right)$	Regression-ratio estimator
1	0	1	1	0	$\bar{y}_{e(RSS)} = \bar{y}_{[rss]}\exp\left(\dfrac{\bar{X} - \bar{x}_{(rss)}}{\bar{X} + \bar{x}_{(rss)}}\right)$	Exponential type estimator
1	β	1	1	0	$\bar{y}_{rege(RSS)} = \left[\bar{y}_{[rss]} + \beta\left(\bar{X} - \bar{x}_{(rss)}\right)\right]\exp\left(\dfrac{\bar{X} - \bar{x}_{(rss)}}{\bar{X} + \bar{x}_{(rss)}}\right)$	Regression-exponential type estimator