IMPROVED ESTIMATORS OF POPULATION MEAN USING AUXILIARY VARIABLES IN RANKED SET SAMPLING

Rajesh Singh and Anamika Kumari*,
Department of StatisticsBanaras Hindu University, Varanasi (India)

ABSTRACT
This paper presents some improved estimators of population mean using auxiliary variables in Ranked Set Sampling. We have derived the expressions for bias and mean square errors up to the first order of approximation and shown that the proposed estimators under optimum conditions are more efficient than other estimators taken in this paper. In an attempt to verify the efficiencies of proposed estimators, theoretical results are supported by empirical study and simulation study for which we have considered two populations.

KEYWORDS: Study variable, auxiliary variable, bias, mean square error, and ranked set sampling.

MSC: 62D05

RESUMEN
Este artículo presenta algunos estimadores mejorados de la media poblacional utilizando variables auxiliares en el muestreo de conjuntos ordenados. Hemos derivado las expresiones para el sesgo y los errores cuadráticos medios hasta el primer orden de aproximación y hemos demostrado que los estimadores propuestos en condiciones óptimas son más eficientes que otros estimadores tomados en este artículo. En un intento de verificar las eficiencias de los estimadores propuestos, los resultados teóricos están respaldados por un estudio empírico y un estudio de simulación para los cuales hemos considerado dos poblaciones.

PALABRAS CLAVE: variable de estudio, variable auxiliar, sesgo, error cuadrático medio y muestreo de conjuntos ordenados

1. INTRODUCTION

Ranked Set Sampling (RSS) is an improved sampling method over Simple Random Set Sampling (SRS).McIntyre (1952) was the first to explain RSS for estimating the population means. McIntyre showed that the RSS estimator of a population means is an unbiased estimator of the population mean. He also showed that the RSS estimator of the population mean has more efficiency than the SRS estimator based on the same sample size. Takahasi and Wakimoto (1968) gave the necessary mathematical theory.
Samawi and Muttlak (1996) suggested ratio estimators of population mean in RSS and showed that the RSS estimators gave improved results over their SRS counterparts. Shiva (2006) compared RSS with SRS for estimation of the mean and the ratio. He concluded that RSS gives a better estimate for both the mean and the ratio. Singh et. al. (2012) suggested a general procedure for estimating the population mean using RSS. Bouza (2014) and Bouza et. al. (2018) provided a review of RSS, its modification, and its application.
In Ranked Set Sampling (RSS), we rank randomly selected units from the population merely by observation or prior experience after which only a few of these sampled units are measured. RSS is more cost-friendly than SRS because fewer samples need to be collected and measured.
RSS takes the following steps.
1. Select sampling units from the target population.
2. Randomly partitioned sampling units into disjoint subsets each having a pre-assigned size (usually taken to be ≤4, as it is convenient and minimizes ranking error)
3. Rank each sub-set.
4. Measure one suitable selected unit from each ranked sub-set.
Coming to the mathematical formulation of RSS, if in the RSS scheme, we want to select a sample of size k. We select k random sets each of size k from the target population. Each set is then ranked by observation/inspection/prior information or a convenient/cheap method.

Original observation After Ranking

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kk} \end{bmatrix} \rightarrow \begin{bmatrix} x_{1(1)} & x_{1(2)} & \cdots & x_{1(k)} \\ x_{2(1)} & x_{2(2)} & \cdots & x_{2(k)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{k(1)} & x_{k(2)} & \cdots & x_{k(k)} \end{bmatrix}$$

Here x_{ij} represents the jth observation in the ith set and $x_{i(j)}$ is the jth ordered statistic in the ith set. After Ranking, select the diagonal units and them. We have now $x_{1(1)}$, $x_{2(2)}$, ……, $x_{k(k)}$ by selecting the smallest ranked unit from the first row, the second smallest ranked unit from the second row, and so on until the largest unit from kth row selected. This will be the Ranked Set Sample (RSS): We can repeat the whole steps r times to obtain an RSS of size n=rk.

$$\bar{X}_{RSS} = \frac{1}{rk}\sum_{l=1}^{r}\sum_{i=1}^{k} x_{i(i)l} \tag{1.1}$$

$$var\ (\bar{X}_{RSS}) = \frac{\sigma^2}{n} - \frac{1}{rm^2}\sum_{i=1}^{k}(\mu_{(i)} - \mu)^2 \tag{1.2}$$

Where $\mu_{(i)}$the mean of the ith is ranked set and is given by

$$\mu_{(i)} = \frac{1}{r}\sum_{l=1}^{r} x_{i(i)l} \quad (1.3)$$

If we want to estimate the contamination level in an area, which is a costly process. We may rank the extent of defoliation i.e. black spot or deprivation of leaves of trees. Then select sampling units based on the ranking of the extent of defoliation and then measure the contamination level of only selected units after ranking.

2. EXISTING ESTIMATORS

In this paper, we take the situation where we use ranking on an auxiliary variable. Let $(y_{[i]}, x_{[i]})$ denote an ith judgment ordering in the ith set for the study variable Y corresponding to the ith order statistic of the ith set for the auxiliary variable X.

•The conventional estimator of the population mean $\bar{Y}$ based on RSS is given by

$$t_{RSS}=\bar{y}_{[n]} \quad (2.1)$$

The variance of the estimator t_{RSS} is given by

$$\text{Var}\ (t_{RSS})= \bar{Y}^2(\eta C_y^2 - D_{y[i]}^2) \quad (2.2)$$

Kadilar et al (2009) proposed the ratio estimator of the population mean $\bar{Y}$ as

$$t_{r,RSS} = \frac{\bar{y}_{[n]}}{\bar{x}_{[n]}}\bar{X} \quad (2.3)$$

The Mean Squared Error (MSE) of the estimator $t_{r,RSS}$ is given by

$$\text{MSE}\ (t_{r,RSS})= \bar{Y}^2[\eta(C_y^2+C_x^2 - 2\rho C_y C_x) - D_{y[i]}^2 - D_{x[i]}^2 + 2D_{yx[i]}] \quad (2.4)$$

•Philip and Lam (1997) proposed the regression type estimator for the population mean $\bar{Y}$ as

$$t_{reg,RSS} = \bar{y}_{[n]} + \hat{\beta}(\bar{X} - \bar{X}_{[n]}) \quad (2.5)$$

The MSE of the estimator $t_{reg,RSS}$ is given by

$$\text{MSE}\ (t_{reg,RSS})=\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx}-D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right] \quad (2.6)$$

Where

$\bar{y}_{[n]} = \frac{1}{n}\sum_{i=1}^{n} y_{[i]}$, $\bar{x}_{[n]} = \frac{1}{n}\sum_{i=1}^{n} x_{[i]}$, $\hat{\beta} = \frac{\hat{R}(\eta\hat{C}_{yx}-\hat{D}_{yx[i]})}{(\eta\hat{C}_x^2-\hat{D}_{x[i]}^2)}$, $\hat{C}_y = \frac{1}{\bar{y}_{[n]}}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_{[i]}-\bar{y}_{[n]})^2}$, $\hat{C}_x = \frac{1}{\bar{x}_{[n]}}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_{[i]}-\bar{x}_{[n]})^2}$

$\hat{C}_{yx} = \frac{1}{(n-1)\bar{y}_{[n]}\bar{x}_{[n]}}\sum_{i=1}^{n}(y_{[i]}-\bar{y}_{[n]})(x_{[i]}-\bar{x}_{[n]})$, $D_{y[i]}^2 = \frac{1}{k^2 r\bar{y}_{[n]}^2}\sum_{i=1}^{k}(\mu_{(iy)}-\bar{y}_{[n]})^2$, $D_{x[i]}^2 = \frac{1}{k^2 r\bar{x}_{[n]}^2}\sum_{i=1}^{k}(\mu_{(ix)}-\bar{x}_{[n]})^2$

$D_{yx[i]} = \frac{1}{k^2 r\bar{y}_{[n]}\bar{x}_{[n]}}\sum_{i=1}^{k}(\mu_{[iy]}-\bar{y}_{[n]})(\mu_{[ix]}-\bar{x}_{[n]})$, $\eta = \frac{1}{kr}$, $\hat{R} = \frac{\bar{y}_{[n]}}{\bar{x}_{[n]}}$

Where $\mu_{[iy]}$ and $\mu_{[ix]}$ are the means of the ith ranked set of Y and X respectively

To obtain bias and MSE of the estimators we need

$\bar{y}_{[n]} = \bar{Y}(1+\epsilon_0)$, $\bar{x}_{[n]} = \bar{X}(1+\epsilon_1)$, E (ϵ_0)=E (ϵ_1)=0, E (ϵ_0^2)=$\eta C_y^2 - D_{y[i]}^2$, E (ϵ_1^2)=$\eta C_x^2 - D_{x[i]}^2$, E $(\epsilon_0\epsilon_1) = \eta C_{yx} - D_{yx[i]}$

3. PROPOSED ESTIMATORS

Motivated by Mishra and Singh (2017) and Bhushan et. al. (2020), we suggest some estimators of the population mean $\bar{Y}$ using RSS as

$$t_{1,RSS} = \bar{y}_{[n]}\exp\left[\alpha\left(\frac{\bar{x}_{[n]}}{\bar{X}}-1\right)\right] \quad \text{…………….} \quad (3.1)$$

$$t_{2,RSS} = \bar{y}_{[n]}\left(1+\beta\, log\frac{\bar{x}_{[n]}}{\bar{X}}\right) \quad (3.2)$$

$$t_{3,RSS} = \bar{y}_{[n]}exp\left(\gamma log\frac{\bar{x}_{[n]}}{\bar{X}}\right) \quad (3.3)$$

Expressing $t_{1,RSS}$ in terms of ϵ's we get

$$t_{1,RSS} = \bar{Y}(1+\epsilon_0)\ exp\left[\alpha\left(\frac{\bar{X}(1+\epsilon_1)}{\bar{X}}-1\right)\right] \quad (3.4)$$

We know that

$$e^x = 1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots, -\infty < x < \infty$$

The bias of the estimator $t_{1,RSS}$ is given by

$$\text{Bias}\ (t_{1,RSS}) = \bar{Y}\left[\frac{\alpha^2}{2}\left(\eta C_x^2 - D_{x[i]}^2\right) + \alpha(\eta C_{yx} - D_{yx[i]})\right] \quad (3.5)$$

The MSE of the estimator $t_{1,RSS}$ is given by

$$\text{MSE}\ (t_{1,RSS}) = \bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 + \alpha^2\left(\eta C_x^2 - D_{x[i]}^2\right) + 2\alpha(\eta C_{yx} - D_{yx[i]})\right] \quad (3.6)$$

To find out the minimum MSE for, we partially differentiate equation wrt α and equating to zero we get

$$\alpha^* = -\frac{\left(\eta C_{yx} - D_{yx[i]}\right)}{\left(\eta C_x^2 - D_{x[i]}^2\right)} \quad (3.7)$$

Putting the optimum value of α in the equation we get a minimum MSE of $t_{1,RSS}$ as

Min MSE= $\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx} - D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right]$ (3.8)

Expressing $t_{2,RSS}$ in terms of ϵ's we get

$$t_{2,RSS} = \bar{Y}(1+\epsilon_0)\left(1+\beta\, log\frac{\bar{X}(1+\epsilon_1)}{\bar{X}}\right) \quad (3.9)$$

We know that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} \dots\dots\dots$$

The bias of the estimator $t_{2,RSS}$ is given by

Bias $(t_{2,RSS}) = \bar{Y}\left[\beta(\eta C_{yx} - D_{yx[i]}) - \frac{\beta^2}{2}(\eta C_x^2 - D_{x[i]}^2)\right]$ (3.10)

The MSE of the estimator $t_{2,RSS}$ is given by

MSE $(t_{2,RSS}) = \bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 + \beta^2(\eta C_x^2 - D_{x[i]}^2) + 2\beta(\eta C_{yx} - D_{yx[i]})\right]$ (3.11)

To find out the minimum MSE for, we partially differentiate equation wrt β and equating to zero we get

$$\beta^* = -\frac{(\eta C_{yx} - D_{yx[i]})}{(\eta C_x^2 - D_{x[i]}^2)} \quad (3.12)$$

Putting the optimum value of β in the equation we get a minimum MSE of $t_{2,RSS}$ as

Min MSE= $\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx} - D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right]$ (3.13)

Expressing $t_{3,RSS}$ in terms of ϵ's we get

$$t_{3,RSS} = \bar{Y}(1+\epsilon_0)\, exp\left(\gamma\, log\frac{\bar{X}(1+\epsilon_1)}{\bar{X}}\right) \quad (3.14)$$

The bias of the estimator $t_{3,RSS}$ is given by

Bias $(t_{3,RSS}) = \bar{Y}\left[\frac{(\gamma^2-\gamma)}{2}(\eta C_x^2 - D_{x[i]}^2) + \gamma(\eta C_{yx} - D_{yx[i]})\right]$ (3.15)

The MSE of the estimator $t_{3,RSS}$ is given by

MSE $(t_{3,RSS}) = \bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 + \gamma^2(\eta C_x^2 - D_{x[i]}^2) + 2\gamma(\eta C_{yx} - D_{yx[i]})\right]$ (3.16)

To find out the minimum MSE for $t_{3,RSS}$ we partially differentiate equation wrt γ and equating to zero we get

$$\gamma^* = -\frac{(\eta C_{yx} - D_{yx[i]})}{(\eta C_x^2 - D_{x[i]}^2)} \quad (3.17)$$

Putting the optimum value of β in the equation we get a minimum MSE of $t_{2,RSS}$ as

Min MSE= $\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx} - D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right]$ (3.18)

4. SOME OTHER PROPOSED ESTIMATORS

We propose modified estimators of population mean by $\bar{Y}$ using RSS as

$$t_{4,RSS} = w_1\bar{y} + w_2\left(\frac{\bar{X}}{\bar{x}}\right) exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \quad (4.1)$$

$$t_{5,RSS} = w_3\bar{y} + w_4 exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\left(1 + log\left(\frac{\bar{X}}{\bar{x}}\right)\right) \quad (4.2)$$

Expressing $t_{4,RSS}$ in terms of ϵ's we get

$$t_{4,RSS} = w_1\bar{Y}(1+\epsilon_0) + w_2(1+\epsilon_1)^{-1}\, exp\left(\frac{-\epsilon_1}{2+\epsilon_1}\right) \quad (4.3)$$

$$t_{4,RSS} - \bar{Y} = (w_1 - 1)\bar{Y} + w_1\bar{Y}\epsilon_0 + w_2\left(1 - \frac{3\epsilon_1}{2} + \frac{15{\epsilon_1}^2}{8}\right) \quad (4.4)$$

Bias $(t_{4,RSS})$=$\bar{Y}(w_1 - 1) + w_2\left[1 + \frac{15}{8}(\eta C_x^2 - D_{x[i]}^2)\right]$ (4.5)

Expressing $t_{5,RSS}$ in terms of ϵ's we get

$$t_{5,RSS} = w_3\bar{Y}(1+\epsilon_0) + w_4\, exp\left(\frac{-\epsilon_1}{2+\epsilon_1}\right)(1 + \log(1+\epsilon_1)) \quad (4.6)$$

$$t_{5,RSS} - \bar{Y} = (w_3 - 1)\bar{Y} + w_3\bar{Y}\epsilon_0 + w_4\left(1 + \frac{\epsilon_1}{2} - \frac{5{\epsilon_1}^2}{8}\right) \quad (4.7)$$

Bias $(t_{5,RSS})$=$\bar{Y}(w_3 - 1) + w_4\left[1 - \frac{5}{8}(\eta C_x^2 - D_{x[i]}^2)\right]$ (4.8)

We have to check these estimators for two conditions

CASE 1: SUM OF WEIGHTS IS UNITY ($\boldsymbol{w_1 + w_2 = 1}$ and $\boldsymbol{w_3 + w_4 = 1}$)

The MSE of the estimator $t_{4,RSS}$ is given by

MSE $(t_{4,RSS}) = \bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 + {w_2}^2\left(\eta C_x^2 - D_{x[i]}^2\right) - 2w_2(\eta C_{yx} - D_{yx[i]})\right]$ (4.9)

To find out the minimum MSE for, we partially differentiate equation wrt and equating to zero we get

$$ {w_2}^* = \frac{(\eta C_{yx} - D_{yx[i]})}{\left(\eta C_x^2 - D_{x[i]}^2\right)} \quad (4.10)$$

Putting the optimum value of $w_2$in the equation we get a minimum MSE of $t_{4,RSS}$ as

Min MSE= $\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx} - D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right]$ (4.11)

The MSE of the estimator $t_{5,RSS}$ is given by

MSE $(t_{5,RSS}) = \bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 + {w_4}^2\left(\eta C_x^2 - D_{x[i]}^2\right) - 2w_4(\eta C_{yx} - D_{yx[i]})\right]$ (4.12)

To find out the minimum MSE for $t_{5,RSS}$, we partially differentiate equation wrt w_2, and equating to zero we get

$$ {w_4}^* = \frac{(\eta C_{yx} - D_{yx[i]})}{\left(\eta C_x^2 - D_{x[i]}^2\right)} \quad (4.13)$$

Putting the optimum value of $w_4$in the equation we get a minimum MSE of $t_{5,RSS}$ as

Min MSE= $\bar{Y}^2\left[\eta C_y^2 - D_{y[i]}^2 - \frac{(\eta C_{yx} - D_{yx[i]})^2}{(\eta C_x^2 - D_x^2)}\right]$ (4.14)

CASE 2: THE SUM OF WEIGHTS IS FLEXIBLE ($\boldsymbol{w_1 + w_2 \neq 1}$ and $\boldsymbol{w_3 + w_4 \neq 1}$)

$$ t_{4,RSS} - \bar{Y} = (w_1 - 1)\bar{Y} + w_1\bar{Y}\epsilon_0 + w_2\left(1 - \frac{3\epsilon_1}{2} + \frac{15{\epsilon_1}^2}{8}\right) \quad (4.15)$$

Squaring on both sides we get

$(t_{4,RSS} - \bar{Y})^2 = \bar{Y}^2 + \bar{Y}^2{w_1}^2(1+\epsilon_0^2) + {w_2}^2(1 + 6\epsilon_1^2) - 2w_1\bar{Y}^2 - 2w_2\bar{Y}\left(1 + \frac{15{\epsilon_1}^2}{8}\right) + 2w_1w_2\bar{Y}(1 + \frac{15{\epsilon_1}^2}{8} - \frac{3\epsilon_0\epsilon_1}{2})$ (4.16)

Taking expectations on both sides we get

MSE$(t_{4,RSS}) = \bar{Y}^2 + \bar{Y}^2{w_1}^2(1+\eta C_y^2 - D_{y[i]}^2) + {w_2}^2\left(1 + 6\left(\eta C_x^2 - D_{x[i]}^2\right)\right) - 2w_1\bar{Y}^2 - 2w_2\bar{Y}\left(1 + \frac{15}{8}\left(\eta C_x^2 - D_{x[i]}^2\right)\right) + 2w_1w_2\bar{Y}(1 + \frac{15}{8}\left(\eta C_x^2 - D_{x[i]}^2\right) - \frac{3}{2}(\eta C_{yx} - D_{yx[i]}))$ (4.17)

MSE $(t_{4,RSS}) = C_1 + {w_1}^2A_1 + {w_2}^2B_1 - 2w_1C_1 - 2w_2D_1 + 2w_1w_2E_1$ (4.18)

Where $A_1 = \bar{Y}^2(1+\eta C_y^2 - D_{y[i]}^2)$, $B_1=\left(1 + 6\left(\eta C_x^2 - D_{x[i]}^2\right)\right)$, $C_1= \bar{Y}^2$, $D_1=\bar{Y}\left(1 + \frac{15}{8}\left(\eta C_x^2 - D_{x[i]}^2\right)\right)$,$E_1=\bar{Y}(1 + \frac{15}{8}\left(\eta C_x^2 - D_{x[i]}^2\right) - \frac{3}{2}(\eta C_{yx} - D_{yx[i]}))$

To find out the minimum MSE for, we partially differentiate equation wrt $w_1$and and equating to zero we get

$$ {w_1}^* = \frac{B_1C_1 - D_1E_1}{A_1B_1 - E_1^2} \quad (4.19)$$

$$ {w_2}^* = \frac{A_1D_1 - C_1E_1}{A_1B_1 - E_1^2} \quad (4.20)$$

Putting the optimum value of w_1 $and w_2$in the equation we get a minimum MSE of $t_{4,RSS}$ as

$$ \min MSE = C_1 + \frac{B_1C_1^2 + A_1D_1^2 - 2C_1D_1E_1}{E_1^2 - A_1B_1} \quad (4.21)$$

$$ t_{5,RSS} - \bar{Y} = (w_3 - 1)\bar{Y} + w_3\bar{Y}\epsilon_0 + w_4\left(1 + \frac{\epsilon_1}{2} - \frac{5{\epsilon_1}^2}{8}\right) \quad (4.22)$$

Squaring on both sides we get

$(t_{5,RSS} - \bar{Y})^2 = \bar{Y}^2 + \bar{Y}^2{w_3}^2(1+\epsilon_0^2) + {w_4}^2(1 - \epsilon_1^2) - 2w_3\bar{Y}^2 - 2w_4\bar{Y}\left(1 - \frac{5\epsilon_1^2}{8}\right) + 2w_3w_4\bar{Y}(1 - \frac{5{\epsilon_1}^2}{8} + \frac{\epsilon_0\epsilon_1}{2})$ (4.23)

Taking expectations on both sides we get

MSE $(t_{5,RSS}) = \bar{Y}^2 + \bar{Y}^2{w_3}^2(1+\eta C_y^2 - D_{y[i]}^2) + {w_4}^2\left(1 - \left(\eta C_x^2 - D_{x[i]}^2\right)\right) - 2w_3\bar{Y}^2 - 2w_4\bar{Y}\left(1 - \frac{5}{8}\left(\eta C_x^2 - D_{x[i]}^2\right)\right) + 2w_3w_4\bar{Y}(1 - \frac{5}{8}\left(\eta C_x^2 - D_{x[i]}^2\right) + \frac{1}{2}(\eta C_{yx} - D_{yx[i]}))$ (4.24)

MSE $(t_{4,RSS}) = C_2 + {w_1}^2A_2 + {w_2}^2B_2 - 2w_1C_2 - 2w_2D_2 + 2w_1w_2E_2$ (4.25)

Where $A_2 = \bar{Y}^2(1+\eta C_y^2 - D_{y[i]}^2)$, $B_2=\left(1 - \left(\eta C_x^2 - D_{x[i]}^2\right)\right)$, $C_2= \bar{Y}^2$, $D_2=\bar{Y}\left(1 - \frac{5}{8}\left(\eta C_x^2 - D_{x[i]}^2\right)\right)$, $E_2=\bar{Y}(1 - \frac{5}{8}\left(\eta C_x^2 - D_{x[i]}^2\right) + \frac{1}{2}(\eta C_{yx} - D_{yx[i]}))$

To find out the minimum MSE for, we partially differentiate equation wrt w_3 and equating to zero we get

$$ {w_3}^* = \frac{B_2C_2 - D_2E_2}{A_2B_2 - E_2^2} \quad (4.26)$$

$$ {w_4}^* = \frac{A_2D_2 - C_2E_2}{A_2B_2 - E_2^2} \quad (4.27)$$

Putting the optimum value of w_3 $and w_4$in the equation we get a minimum MSE of $t_{5,RSS}$ as

$$\min MSE = C_2 + \frac{B_2C_2^2 + A_2D_2^2 - 2C_2D_2E_2}{E_2^2 - A_2B_2} \tag{4.28}$$

5. NUMERICAL ILLUSTRATIONS

In this section, we compare the performance of the proposed estimators with the other estimators considered in this paper. For comparison, we have taken a population from Singh (2003) (page no. 1111(Appendix)).
From the above population, we took ranked set samples with size k=3 and the number of cycles r=4, 5,6,10. For these samples, we have calculated MSE and PRE for different estimators.

Table 1: The Mean Square Errors of the Estimators

Estimators	K=3, r=4 , n=rk=12	K=3, r=5, n=rk=15	K=3, r=6, n=rk=18	K=3, r=10 N=rk=30
t_{RSS}	11552.3791	19752.407	11065.05554	5748.64008
$t_{r,RSS}$	7301.3221	7170.007	6329.035581	4647.76857
$t_{reg,RSS}$	5992.717	5631.803	6323.455597	3157.62779
$t_{1,RSS}$	5992.717	5631.803	6323.455597	3157.62779
$t_{2,RSS}$	5992.717	5631.803	6323.455597	3157.62779
$t_{3,RSS}$	5992.717	5631.803	6323.455597	3157.62779
$t_{4,RSS}$	2104.430694	1532.508392	1795.968342	1323.04998
$t_{5,RSS}$	2724.46435	3974.48669	1056.137	1918.17813

From Table 1, it is observed that:
1. The estimators $t_{1,RSS}$, $t_{2,RSS}$ and $t_{3,RSS}$ are almost equally efficient estimators as linear regression estimators as these estimators show the MSE almost equal to the MSE of the linear regression estimator ($t_{reg,RSS}$).
These three estimators $t_{1,RSS}$, $t_{2,RSS}$ and $t_{3,RSS}$ are more efficient estimators than that of the conventional estimator and ratio estimator.
2. $t_{4,RSS}$ and $t_{5,RSS}$ are more efficient than other estimators used in this paper as our proposed estimators have greater PRE. It is observed that $t_{4,RSS}$, and $t_{5,RSS}$ are more efficient than convention, ratio estimator, and linear regression estimator.
The formula for Percent Relative Efficiency (PRE) is

$$\text{PRE (estimators)} = \frac{MSE(t_{RSS})}{MSE(estimator)} \times 100$$

Table 2 also shows that our proposed estimators perform better than the existing estimators. The MSE of the estimators decreases when the correlation and sample size increases for the population 1 and 2.

Table 2: The Mean Square Errors and Percentage Relative Efficiencies of the Estimators for n=12

Estimators	MSE	PRE
t_{RSS}	11552.3791	100.00
$t_{r,RSS}$	7301.3221	158.2231
$t_{reg,RSS}$	5992.717	192.7736

$t_{1,RSS}$	5992.717	192.7736
$t_{2,RSS}$	5992.717	192.7736
$t_{3,RSS}$	5992.717	192.7736
$t_{4,RSS}$	2104.430694	548.9550
$t_{5,RSS}$	2724.46435	424.0238

6. SIMULATION STUDY

We perform some simulation experiments to check the proposed estimator's Percent relative efficiency (PRE) with the conventional, ratio, and regression estimator. Also, we investigate the PRE of the proposed estimators when the sum of weights is unity wrt the same estimators and the when sum of weights is flexible.

This is done via the following steps

1. Generate two independent random variables X~ N (μ, σ^2) and $X_1 \sim N(\mu_1, \sigma_1^2)$ using the box Muller method (Johnson, 1987).
2. Set Y= ρX+$\sqrt{1-\rho^2}X_1$ where ρ=0.6, 0.7, 0.8, 0.9<1

3: Return the pair (Y, X).

4: Consider the population-I with the parameters μ_y =5, σ_y =3, and $\mu_x = 5, \sigma_x$=3 in step-1 and repeat steps 1 to 3 for 10000 times. This population will contain the same variance for the variables Y and X.

5: Similarly, generate the population-II with the parameters μ_y =3, σ_y =2, μ_x=5 and σ_y=3 in step-1 and repeat steps 1 to 3 for 10000 times. This population will have different variances for the variables Y and X

6. Three sets of size 3 are taken randomly from N=1000. These sets are ranked using the x. From the first set select the lowest value, from the second set the median value is selected and from the third set the largest value is selected.
7. Step 6 is repeated r times so we have RSS of size n=kr (k=3) r=4, 5,6,10.
8. Steps 6 and Step 7 are repeated 10000 times. The mean and the Variance of these values are computed.

Table 3: Comparison of MSE of the estimators

r_{yx}	n	estimators	Population 1 MSE	Population 2 MSE
0.9	12	t_{RSS}	0.152154	0.171324
		$t_{r,RSS}$	0.077241	0.015460
		$t_{reg,RSS}$	0.050724	0.004284
		$t_{1,RSS}$	0.047787	0.004976
		$t_{2,RSS}$	0.049700	0.004647
		$t_{3,RSS}$	0.047225	0.004434
		$t_{4,RSS}$	0.056221	0.004229
		$t_{5,RSS}$	0.055273	0.004335
	15	t_{RSS}	0.122575	0.129890
		$t_{r,RSS}$	0.061580	0.012281
		$t_{reg,RSS}$	0.040402	0.003538
		$t_{1,RSS}$	0.041187	0.003792
		$t_{2,RSS}$	0.040413	0.003671
		$t_{3,RSS}$	0.039520	0.003311
		$t_{4,RSS}$	0.039940	0.003172
		$t_{5,RSS}$	0.042219	0.003885
	18	t_{RSS}	0.104998	0.107005
		$t_{r,RSS}$	0.045434	0.009145
		$t_{reg,RSS}$	0.033634	0.002747

		$t_{1,RSS}$	0.032081	0.003183
		$t_{2,RSS}$	0.031449	0.002927
		$t_{3,RSS}$	0.031992	0.002621
		$t_{4,RSS}$	0.034619	0.002421
		$t_{5,RSS}$	0.035400	0.002720
	30	t_{RSS}	0.063604	0.064378
		$t_{r,RSS}$	0.031763	0.005540
		$t_{reg,RSS}$	0.018650	0.001707
		$t_{1,RSS}$	0.017680	0.001850
		$t_{2,RSS}$	0.019670	0.001828
		$t_{3,RSS}$	0.018549	0.001697
		$t_{4,RSS}$	0.019172	0.001419
		$t_{5,RSS}$	0.019821	0.001404
0.8	12	t_{RSS}	0.180392	0.185993
		$t_{r,RSS}$	0.134896	0.087700
		$t_{reg,RSS}$	0.092868	0.074471
		$t_{1,RSS}$	0.098835	0.076173
		$t_{2,RSS}$	0.094387	0.075587
		$t_{3,RSS}$	0.094563	0.070803
		$t_{4,RSS}$	0.092018	0.081049
		$t_{5,RSS}$	0.095503	0.740483
	15	t_{RSS}	0.131067	0.161216
		$t_{r,RSS}$	0.110187	0.071194
		$t_{reg,RSS}$	0.073311	0.055151
		$t_{1,RSS}$	0.074002	0.057728
		$t_{2,RSS}$	0.072264	0.056220
		$t_{3,RSS}$	0.071470	0.056787
		$t_{4,RSS}$	0.069733	0.057450
		$t_{5,RSS}$	0.071291	0.043243
	18	t_{RSS}	0.114785	0.128750
		$t_{r,RSS}$	0.090967	0.092254
		$t_{reg,RSS}$	0.062433	0.049902
		$t_{1,RSS}$	0.061425	0.045817
		$t_{2,RSS}$	0.063433	0.047118
		$t_{3,RSS}$	0.065415	0.045559
		$t_{4,RSS}$	0.058531	0.048685
		$t_{5,RSS}$	0.057302	0.045106
	30	t_{RSS}	0.070402	0.076343
		$t_{r,RSS}$	0.053824	0.034527
		$t_{reg,RSS}$	0.035038	0.027835
		$t_{1,RSS}$	0.034768	0.025014
		$t_{2,RSS}$	0.036705	0.026239
		$t_{3,RSS}$	0.036418	0.027316
		$t_{4,RSS}$	0.033901	0.027854

		$t_{5,RSS}$	0.035514	0.027549
0.7	12	t_{RSS}	0.1944850	0.202534
		$t_{r,RSS}$	0.173737	0.145953
		$t_{reg,RSS}$	0.120574	0.118416
		$t_{1,RSS}$	0.134713	0.121742
		$t_{2,RSS}$	0.119949	0.112254
		$t_{3,RSS}$	0.120265	0.113914
		$t_{4,RSS}$	0.113229	0.104075
		$t_{5,RSS}$	0.140403	0.109392
	15	t_{RSS}	0.151521	0.166570
		$t_{r,RSS}$	0.134287	0.112821
		$t_{reg,RSS}$	0.098746	0.090097
		$t_{1,RSS}$	0.106997	0.087018
		$t_{2,RSS}$	0.096792	0.082566
		$t_{3,RSS}$	0.101644	0.084592
		$t_{4,RSS}$	0.095310	0.082298
		$t_{5,RSS}$	0.096305	0.090351
	18	t_{RSS}	0.120667	0.144174
		$t_{r,RSS}$	0.111706	0.088692
		$t_{reg,RSS}$	0.081425	0.072458
		$t_{1,RSS}$	0.084035	0.068005
		$t_{2,RSS}$	0.086801	0.076613
		$t_{3,RSS}$	0.085555	0.074263
		$t_{4,RSS}$	0.080167	0.072055
		$t_{5,RSS}$	0.085309	0.077269
	30	t_{RSS}	0.078230	0.079345
		$t_{r,RSS}$	0.069852	0.058810
		$t_{reg,RSS}$	0.050767	0.044316
		$t_{1,RSS}$	0.048712	0.041729
		$t_{2,RSS}$	0.051286	0.045329
		$t_{3,RSS}$	0.048395	0.045208
		$t_{4,RSS}$	0.043242	0.043282
		$t_{5,RSS}$	0.045122	0.046804
0.6	12	t_{RSS}	0.220789	0.238781
		$t_{r,RSS}$	0.239418	0.222704
		$t_{reg,RSS}$	0.181706	0.203448
		$t_{1,RSS}$	0.182990	0.208017
		$t_{2,RSS}$	0.172726	0.195914
		$t_{3,RSS}$	0.182735	0.187646
		$t_{4,RSS}$	0.155211	0.182710
		$t_{5,RSS}$	0.175934	0.183729
	15	t_{RSS}	0.169128	0.196310
		$t_{r,RSS}$	0.188486	0.194962
		$t_{reg,RSS}$	0.150099	0.154603
		$t_{1,RSS}$	0.145505	0.140324
		$t_{2,RSS}$	0.150503	0.146587
		$t_{3,RSS}$	0.146181	0.148311
		$t_{4,RSS}$	0.134538	0.137897

		$t_{5,RSS}$	0.130544	0.140155
	18	t_{RSS}	0.142389	0.155581
		$t_{r,RSS}$	0.159748	0.150182
		$t_{reg,RSS}$	0.114844	0.132240
		$t_{1,RSS}$	0.107508	0.127191
		$t_{2,RSS}$	0.113107	0.129402
		$t_{3,RSS}$	0.110554	0.128706
		$t_{4,RSS}$	0.110810	0.109907
		$t_{5,RSS}$	0.139292	0.108098
	30	t_{RSS}	0.085515	0.102551
		$t_{r,RSS}$	0.091276	0.099424
		$t_{reg,RSS}$	0.073868	0.072887
		$t_{1,RSS}$	0.068966	0.075434
		$t_{2,RSS}$	0.073337	0.073721
		$t_{3,RSS}$	0.068292	0.074040
		$t_{4,RSS}$	0.062993	0.072472
		$t_{5,RSS}$	0.063988	0.072478

7. CONCLUSION

• In this article we have proposed estimators for the population mean in ranked set sampling using the information of auxiliary variables. The expressions for Bias and MSE of the suggested estimators have been derived up to the first order of approximation.

•Numerical Illustrations and simulation studies for comparing the efficiency of the proposed estimators with other estimators have been used. The results have been shown the Tables 1 & 2. The Tables show that the proposed estimators turn out to be more efficient as compared to the other estimators for both populations.

•The proposed estimators are found to be rather improved in terms of lesser MSE and greater PRE as compared to the existing estimators in both real and simulated data sets. We have taken ranked set samples of size 12, 15, 18, and 30 it is observed from the simulation that the MSE of the proposed estimators decreases as the sample size increase whereas the PRE of the suggested estimators increases as the sample size increase. For ρ=0.6, 0.7, 0.8, 0.9 it is observed from the simulation that the MSE of the proposed estimators decreases as the values of the correlation coefficient increase whereas the PRE of the suggested estimators increases as the values of the correlation coefficients increase.

• Based on our Numerical Illustrations and simulation study, we can conclude that our proposed estimators can be preferred over the other estimators taken in this paper in several real situations like agriculture sciences, mathematical sciences, biological sciences, poultry, business, economics, commerce, social sciences, etc.

RECEIVED: MAY, 2022
REVISED:OCTOBER, 2022.

REFERENCES

[1] BHUSHAN, S., GUPTA, R., SINGH, S., and KUMAR, A. (2020): A New Efficient Log-Type Class of Estimators Using Auxiliary Variable. **International Journal of Statistics and Systems**, 15, 19-28.

[2] BOUZA, C.N. and AL-OMARI, A.B. (2014): Review of Ranked Set Sampling: Modifications and Applications**. Revista Investigacion Operacional**, 35(3), 215-240.

[3] BOUZA, C.N., SINGH, P., AND SINGH, R. (2018): Ranked set sampling and optional scrambling randomized response modeling. **Revista Investigacion Operacional**. 39, 100-107.

[4] GANESH, S., and GANESLINGAM, S. (2006): Ranked set sampling vs simple random sampling in the estimation of the mean and the ratio. **Journal of Statistics and Management Systems**, 9, 459-472.

[5] JOHNSON, M. E. (1987): **Multivariate Statistical Simulation: A Guide to Selecting and Generating Continuous Multivariate Distributions**, 192, John Wiley and Sons, N.York.

[6] KADILAR, C., UNYAZICI, Y., and CINGI, H. (2009): Ratio estimator for the population means using ranked set sampling. **Stat. Papers**, 50, 301–309.

[7] MCINTYRE, G.A. (1952): A method of unbiased selective sampling using ranked sets. **Australian Journal of Agricultural Research**, 3, 385–390.

[8] MISHRA, P. and SINGH, R. (2017): A new log-product type estimator using auxiliary information. **Jour. Sci. Res**., 61(1and2), 179-183.

[9] SAMAWI, H.M., and MUTTLAK, H.A. (1996): Estimation of ratio using rank set sampling. **The Biometrical Journal**, 38 , 753–764.

[10] SINGH, H.P., TAILOR, R. and SINGH, S. (2014): General procedure for estimating the population means using ranked set sampling. **Journal of Statistical Computation and Simulation**, 84, 931-945.
[11] SINGH, S. (2003): **Advanced Sampling Theory with Applications.** Kluwer Academic Publishers, Dordrecht.
[12] TAKAHASI, K. and WAKIMOTO, K. (1968): On unbiased estimates of the population mean based on the sample stratified using ordering**. Annals of Institute of Statistical Mathematics**, 20, 1-31.