Kurdish Studies
September 2024
Volume: 12, No: 5, pp 1060 -1066
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3406

Improved Estimation Of Mean Under Ranked Set Sampling Using Auxiliary Information

Muhammed Eisa[1], Muhammad Iqbal[1], Hameed Ali[1*], Khazan Sher[1], Imad Khan[2]

[1*]Department of Statistics University of Peshawar, Pakistan.
[2]Department of Statistics Abdul Wali Khan University Mardan, Pakistan.

***Corresponding Author:** Hameed Ali
*hameedali@aup.edu.pk,

Abstract
This paper aims to introduce a new approach to efficient mean estimators under ranked set sampling. In cases where the ranks of the auxiliary variables exhibit a positive correlation with the primary research variable, the proposed estimators are derived from a linear combination of conventional estimators such as product, ratio, and exponential. As performance metrics, the estimators' Mean Square Error (MSE) and Bias are utilized to assess their effectiveness. In addition to deriving the theoretical properties of the suggested estimators, the conditions under which they perform better than the current estimators are examined. Additionally, real data sets are used to empirically evaluate the proposed estimators, and the findings show that, under ranked set sampling, the suggested estimators consistently produce the best results under all circumstances. A data-driven simulation study demonstrates that the proposed estimators outperform the current estimators covered in the literature.

Keywords: Ranked set sampling, Auxiliary information, Transformation, Bias, MSE, Simulation study

1. Introduction

Sampling is crucial for the sampling survey statisticians to make conclusions about the population based on the data gathered from the sample in order to draw inferences. However, time and cost are two major aspects that are crucial to consider when creating a sampling technique in order to ensure a feasible sampling practice. These two elements should be taken into account at all times, especially when utilizing agricultural, biological, and ecological surveys as part of sample methodologies. When choosing a sample, traditional sample designs like simple random sampling (SRS) do not take these two criteria into account. The best sampling approach is ranked set sampling as an alternative to SRS. The ranks of auxiliary variables that exhibit a positive correlation with the primary study variable are considered in rank set sampling. In ranked set sampling, it might be difficult to estimate parameters like mean, variance, etc. in an efficient manner. It is commonly known from a number of research publications that using supplementary data progressively improves the estimates' efficiency. By using rankings of the auxiliary variables that have a positive correlation with the primary study variable, this work aims to construct efficient estimators for the estimation of the population mean. Various researchers have created their estimators to calculate the population mean by various techniques, such as Ali, Asim, Ijaz, et al., (2024); Ali, Asim, & Sher, (2024); Sher et al., (2024); Grover & Kaur, (2014); H. P. Singh et al., (2009) etc. in SRS scheme. However, the pioneering work of estimating the mean using *RSS* was due to McIntyre, (1952), who estimated the mean pasture, although he did not give his sampling method a specific name. The work of McIntyre, (1952) was extended by different researchers like Halls & Dell, (1966) using the weighted mean Lynne Stokes, (1977) and introduced the procedure of using an auxiliary variable instead of judgement. Under *RSS* Muttlak, (1998), Yu & Lam, (1997), Muttlak, (1998), (Muttlak, 2003), Koyuncu, (2015) etc., offered simple ratio, ratio and regression type estimators to estimate finite population mean. Takahasi & Wakimoto, (1968) provided mathematically based theory and developed the product type estimator with different strategies suggested by Bouza et al., (1982). Significant work is done by Dell & Clutter, (1972) in identifying factors in error ranks that lead to a reduction or loss of efficiency. Similarly, (Rather et al., 2022) recently developed an exponential ratio estimator type and modified ratio and regression estimators, followed by (Al-Omari and Bouza, 2014). To estimate the finite population mean efficiently, a generalized exponential ratio estimator developed by D.-Z. Khan & Muhammad, (2019). L. Khan & Shabbїr, (2016) studied Hartley-Rose type estimators and proposed a class of estimators for estimating the finite population mean of the main study variable under ranked set sample where the population parameter of the supplemental variables is known.

2. Methodology

In RSS, m random sets of size m are generated from a finite population. All of the set's members are ranked according to the auxiliary or study factors. It is assumed that there is inaccuracy in the primary variable Y's ranking since the auxiliary variable X's ranking is based on perfect ranking. To select the largest member from the final m-ranked set, the smallest member from the first ranked set with the lowest rank is chosen, followed by the smallest member from the second ranked set with the second lowest rank, and so on. The cycle for choosing an exact sample of size m is finished with this step. To choose a sample of size n, this cycle can alternatively be completed r times. Thus, units in a sample are chosen using this process. The i-th order statistics in the i-th set for the auxiliary variable and the scoring order of the study variable, respectively, can be used to indicate the units of the study variable (y) and the auxiliary variable (x) under RSS. To derive the estimators' biases and Mean Square Errors, the following terms are defined:

$$\gamma=\frac{1}{mr},K_{x(i)}=M_{x(i)}-\mu_{\bar{X}},\ K_{y(i)}=M_{y(i)}-\mu_{\bar{Y}},K_{yx(i)}=\left(M_{y(i)}-\mu_{\bar{Y}}\right)\left(M_{x(i)}-\mu_{\bar{X}}\right),\ W_y^2=\frac{1}{m^2r\mu_{\bar{Y}}^2}\sum_{i=1}^{m}K_{y(i)}^2,$$

$$W_x^2=\frac{1}{m^2r\mu_{\bar{X}}^2}\sum_{i=1}^{m}K_{x(i)}^2,W_{yx}=\frac{1}{m^2r\mu_{\bar{X}}\mu_{\bar{Y}}}\sum_{i=1}^{m}K_{yx(i)},\ \bar{y}_{(rss)}=\mu_{\bar{Y}}(1+e_0),\bar{x}_{(rss)}=\mu_{\bar{X}}(1+e_1)$$

Also the following terms can be defined such that

$E(e_i)=0, \quad i=0,1,2, \quad E(e_0^2)=\gamma C_y^2-W_y^2=H_y^2, \quad E(e_1^2)=\gamma C_x^2-W_x^2=G_x^2, E(e_0e_1)=\gamma C_{yx}-W_{yx}=J_{yx}.$

The coefficients of variation for main study variable and auxiliary variable are denoted by C_y and C_x

The mean of study variable is representing by $\mu_{\bar{Y}}$ while $\mu_{\bar{X}}$ representing the true mean of supplementary variable. Here $M_{y(i)}$ &$M_{x(i)}$ rely on order statistics from some appropriate distribution.

Using auxiliary information (Dell and Clutter, 1972) introduced estimators under RSS through which the population mean can be estimated. They defined means of each variable of auxiliary information X and main study variable Y using various samples.

$$\mu_{\bar{X}(RSS)}=\frac{1}{rm}\sum_{i=1}^{m}\sum_{j=1}^{r}X_{i(i:m)j} \qquad \mu_{\bar{Y}(RSS)}=\frac{1}{rm}\sum_{i=1}^{m}\sum_{j=1}^{r}Y_{i(i:m)j}$$

And their variances and covariance are

$$Var\left(\mu_{\bar{X}(RSS)}\right)=\frac{\delta^2}{rm}-\frac{1}{rm^2}\sum_{i=1}^{m}K_{x(i)}^2, \tag{1}$$

$$Var\left(\mu_{\bar{Y}(RSS)}\right)=\frac{\delta^2}{rm}-\frac{1}{rm^2}\sum_{i=1}^{m}K_{y(i)}^2, \tag{2}$$

$$Cov\left(\mu_{\bar{X}(RSS)},\mu_{\bar{Y}(RSS)}\right)=\frac{\delta^2}{rm}-\frac{1}{rm^2}\sum_{i=1}^{m}K_{xy(i)},$$

(Muttlak, 2003) and (Kadilar et al., 2009) suggested the finite population mean under RSS, which is defined as

$$\bar{y}_{rRSS}=\frac{\bar{y}_{RSS}}{\bar{x}_{RSS}}\mu_{\bar{X}} \tag{3}$$

With Bias & MSE is given by

$$Bias(\bar{y}_{rRSS})\cong\mu_{\bar{Y}}\left(G_x^2-J_{yx}\right), \tag{4}$$

$$MSE(\bar{y}_{rRSS})\cong\mu_{\bar{Y}(RSS)}^2\left(H_y^2+G_x^2-2J_{yx}\right). \tag{5}$$

When there is a negative correlation between the primary study variable (Y) and the auxiliary variable (X), the product estimator is also employed. Given is the product estimator.

$$\bar{y}_{pRSS}=\bar{y}_{rss}\frac{\bar{x}_{rss}}{\mu_{\bar{X}}}, \tag{6}$$

The Bias & MSE of the estimator is given as

$$Bias(\bar{y}_{pRSS}) \cong \mu_{\bar{Y}} J_{yx}, \tag{7}$$

$$MSE(\bar{y}_{pRSS}) \cong \mu^2_{\bar{Y}(RSS)} (H_y^2 + G_x^2 + 2J_{yx}), \tag{8}$$

(Rather et al., 2022), Kadilar et al., (2009) developed an exponential estimator utilizing the auxiliary information under ratio estimators. The proposed estimator is provided below

$$\hat{\bar{y}}_k = \bar{y}_{[n]} \exp \left[\frac{\mu_{\bar{X}}}{\mu_{\bar{X}} + k(\bar{x}_{(n)} - \mu_{\bar{X}})} - 1 \right] \tag{9}$$

The estimators developed by (Rather et al., 2022), (Kadilar et al., 2009), their properties are evaluated through the suggested estimator Bias & MSE which is given below

$$Bias\ (\hat{\bar{y}}_k) \cong \mu_{\bar{Y}} \left[\frac{3R^2k^2}{2}\left(\frac{1}{mr}\left(S_x^2 - \frac{1}{m}\sum_{i=1}^m K_{x[i]}^2\right)\right) - \frac{kR}{mr}\left(S_{yx} - \frac{1}{m}\sum_{i=1}^m K_{yx[i]}\right)\right], \tag{10}$$

$$MSE_{\min}(\hat{\bar{y}}_k) \cong \frac{1}{mr}\left(S_y^2 - \frac{1}{m}\sum_{i=1}^m K_{y[i]}^2\right) - \frac{\frac{1}{mr}\left(S_{yx} - \frac{1}{m}\sum_{i=1}^m K_{yx[i]}\right)^2}{\left(S_x^2 - \frac{1}{m}\sum_{i=1}^m K_{x[i]}^2\right)}, \tag{11}$$

This study's main objective is to use ranked set sampling to estimate the population's true mean of the main study variable. Rank set sampling estimation techniques and their application are necessary to determine the population's true mean. The following is a list of the suggested methods and estimators for determining the population mean.

3. Proposed Estimators

3.1 First Proposed Estimator under RSS

Several researchers suggested their efficient estimators under *RSS* like (Rather et al., 2022) Kadilar et al., (2009). Motivated by (Bahl and Tuteja, 1991),(Shabbir et al., 2021), Under RSS, to determine the finite population mean, the following estimators are recommended:

$$\bar{y}_{EI(I)} = \left[K_1 + K_2\bar{y}_{RSS} + \tau_3\left(\mu_{\bar{X}}/\bar{x}_{RSS}\right)\right] exp\left[\frac{\mu_{\bar{X}} - \bar{x}_{RSS}}{\mu_{\bar{X}} + \bar{x}_{RSS}}\right] \tag{12}$$

WhereK_1, K_2 and K_3 are optimization constants, that minimizes the MSE of the estimator.
Rewriting (12) in term of error due to sampling given in section 2, we proceed as following to find Bias and MSE:

$$\bar{y}_{EI(I)} - \mu_{\bar{Y}} \cong \left[K_1 + \mu_{\bar{Y}} K_2 - \mu_{\bar{Y}}\right] + K_2 \mu_{\bar{Y}} e_0 - \left(\frac{1}{2}K_1 + K_3 \mu_{\bar{X}} - \mu_{\bar{Y}} \frac{K_2}{2}\right) e_1 - \left(\frac{3}{8}K_1 + K_3 \frac{\mu_{\bar{X}}}{2} + \frac{3}{8}\mu_{\bar{Y}} K_2\right) e_1^2 - \frac{K_2 \mu_{\bar{Y}}}{2} e_0 e_1,$$

$$Bias(\bar{y}_{EI(I)}) \cong \left[K_1 + \mu_{\bar{Y}} K_2 - \mu_{\bar{Y}}\right] - \left(\frac{3}{8} K_1 + K_3 \frac{\bar{X}}{2} + \frac{3}{8} \mu_{\bar{Y}} K_2\right) G_x^2 - \frac{K_2 \mu_{\bar{Y}}}{2} J_{yx}, \tag{13}$$

$$\left(MSE\ (\bar{y}_{EI(I)}) = \left[K_1 + \mu_{\bar{Y}} K_2 - \mu_{\bar{Y}}\right]^2 + \left[\left(\frac{1}{2} K_1 + K_3 \mu_{\bar{X}} - \mu_{\bar{Y}} \frac{K_2}{2}\right)^2 + 2\left[K_1 + \mu_{\bar{Y}} K_2 - \mu_{\bar{Y}}\right]\left(\frac{3}{8}K_1 + K_3 \frac{\mu_{\bar{X}}}{2} + \frac{3}{8}\mu_{\bar{Y}} K_2\right)\right] G_x^2 + K_2^2 \mu_{\bar{Y}}^2 H_y^2 - 2\mu_{\bar{Y}} K_2 \left[\frac{1}{2}\left(K_1 + \mu_{\bar{Y}} K_2 - \mu_{\bar{Y}}\right) + \frac{1}{2}\left(K_1 + K_3 \mu_{\bar{X}} - \mu_{\bar{Y}} K_2\right)\right] J_{yx}\right) \tag{14}$$

K_1, K_2 &K_3, are optimum values of obtained from eq. (14) which is given as

$$K_1 = \frac{1}{4}\frac{\mu_{\bar{Y}}\left(14(G_x^2)^3 + 18(G_x^2)^2H_y^2 + 15(H_x^2)^2J_{yx} - 16G_x^2H_y^2 - 14G_x^2H_y^2 + 16J_{yx}^2\right)}{4(G_x^2)^3 + 5(G_x^2)^2J_{yx} - 4G_x^2H_y^2 - 4G_x^2J_{yx}^2 + 4J_{yx}^2}$$

$$K_2 = \frac{1}{4}\frac{(G_x^2)^2(2G_x^2 + J_{yx})}{4(G_x^2)^3 + 5(G_x^2)^2H_y^2 + 4(G_x^2)^2J_{yx} - 4G_x^2H_y^2 - 4G_x^2J_{yx}^2 + 4J_{yx}^2}$$

$$K_3 = -\frac{\mu_{\bar{Y}}}{4\,\mu_{\bar{X}}}\frac{\left(6(G_x^2)^3 + 7(G_x^2)^2 H_y^2 + 5(G_x^2)^2 J_{yx} - 8G_x^2 H_y^2 - 6G_x^2 J_{yx}^2 + 8J_{yx}^2\right)}{\left(4(G_x^2)^3 + 5(G_x^2)^2 H_y^2 + 4(G_x^2)^2 J_{yx} - 4G_x^2 H_y^2 - 4G_x^2 J_{yx}^2 + 4J_{yx}^2\right)}$$

By putting the optimum value of τ_1, τ_2 and τ_3 in (14). We get

$$MSE\left(\bar{y}_{EI(I)}\right)_{\min} \cong \frac{\mu_{\bar{Y}}^2 \left(G_x^2\right)^2 \left[G_x^2 H_y^2 - J_{yx}^2\right]}{\left(G_x^2\right)^2\left[16\left(4G_x^2 + 5H_y^2 + 4J_{yx}\right) - \frac{64\left(H_y^2 + J_{yx}^2\right)}{G_x^2}\right] + 64J_{yx}^2} \quad (15)$$

3.2 Second Proposed Estimator under RSS

Taking motivation from Shabbir et al., (2021), Mehta (Ranka) & Mandowara, (2016) R. Singh et al., (2008), the proposed estimator under ranked set sampling is given as

$$\bar{y}_{ES(II)} = \left[K_4 + K_5\bar{y}_{RSS} + K_6\left(\mu_{\bar{X}}/\bar{x}_{RSS}\right)\right]\exp\left[\frac{\bar{x}_{RSS} - \mu_{\bar{X}}}{\bar{x}_{RSS} + \mu_{\bar{X}}}\right] \quad (16)$$

To find Bias$\bar{y}_{EI(II)}$ and MSE $s\bar{y}_{EI(II)}$, rewriting eq. (14), as following

$$Bias\left(\bar{y}_{ES(II)}\right) \cong K_4 + (K_5 - 1)\,\mu_{\bar{Y}} + \frac{1}{2}\left(K_5\,\mu_{\bar{Y}}\,J_{yx} - \left(\frac{K_4}{4} + \frac{K_5\,\mu_{\bar{Y}}}{4} + K_6\,\mu_{\bar{X}}\right)G_x^2\right)$$

$$MSE\left(\bar{y}_{ES(II)}\right) \cong \left(K_4 + (K_5 - 1)\,\mu_{\bar{Y}}\right)^2 + K_5^2\,\mu_{\bar{Y}}^2\,H_y^2 + K_5\,\mu_{\bar{Y}}\left(3K_4 + 3K_5\,\mu_{\bar{Y}} - 2K_5\,\mu_{\bar{X}} - 2\,\mu_{\bar{Y}}\right)J_{yx}$$

$$+\left(\left(K_6\,\mu_{\bar{X}} - \frac{K_4}{2} - \frac{K_5\,\mu_{\bar{Y}}}{2}\right)^2 - \left(K_4 + (K_5 - 1)\,\mu_{\bar{Y}}\right)\left(\frac{K_4}{4} + \frac{K_5\,\mu_{\bar{Y}}}{4} + K_6\,\mu_{\bar{X}}\right)\right)G_x^2 \quad (17)$$

The optimum values of, K_4, K_5 and K_6 are obtained as

$$K_4 = \frac{\left(10(G_x^2)^2 H_y^2 + 9(G_x^2)^2 J_{yx} + 26G_x^2 J_{yx}^2\right)\bar{Y} + 8G_x^2 H_y^2 - 12G_x^2 J_{yx} - 8J_{yx}^2}{4G_x^2\left(4G_x^2 H_y^2 + 5J_{yx}^2\right)},$$

$$K_5 = -\frac{3J_{yx}\left(3G_x^2\,\mu_{\bar{Y}} - 4\right)}{4\,\mu_{\bar{Y}}\left(4G_x^2 H_y^2 + 5J_{yx}^2\right)},$$

$$K_6 = -\frac{\left(2\left(G_x^2\right)^2 H_y^2 + 7G_x^2 J_{yx}^2\right)\mu_{\bar{Y}} + 8G_x^2 H_y^2 + 4J_{yx}^2}{4\,\mu_{\bar{X}}\,G_x^2\left(4G_x^2 H_y^2 + 5J_{yx}^2\right)}.$$

Now Putting these optimum values of in $MSE\left(\bar{y}_{ES(II)}\right)_{\min}$, we get

$$MSE\left(\bar{y}_{ES(II)}\right)_{\min} \cong \frac{\left(G_x^2 H_y^2 - J_{yx}^2\right)\left(3G_x^2\,\mu_{\bar{Y}} - 4\right)^2}{64G_x^2\left(G_x^2 H_y^2 + \frac{5}{4}J_{yx}^2\right)}. \quad (18)$$

4. Numerical Analysis

The following two datasets are used to compare the performance of the competing and proposed estimators to the traditional mean estimators under ranked set sampling. Summary statistics of the data are given below

Summary Statistics of data-I Source: (Vallient and Royall, 2000)

y: Deaths due to Breast cancer during 1950-1969

x:population of female in adult age during 1960 $\mu_{\bar{Y}}$

$N = 301$, $n = 12$, $m = 3$, $r = 4$,

$\mu_{\bar{X}} = 11288.18100$, $\mu_{\bar{Y}} = 39.81500$, $\rho = 0.96171$, $\beta_2(x) = 10.719$,

$\mu_{\bar{X}_1} = 13780.8334$, $\mu_{\bar{X}_2} = 121852.410$, $\mu_{\bar{X}_3} = 13290.617$, $C_y = 1.279114$,

$C_x = 1.221107$, $C_{x1} = 1.221106$, $C_{x2} = 1.221107$, $C_{x3} = 1.21918$,

$C_{yx} = 1.51015,$ $C_{yx1} = 1.51014,$ $C_{yx2} = 1.51014,$ $C_{yx3} = 1.50913,$

Summary Statistics of data set-II Source: (Vallient and Royall, 2000)

y :Number of patients discharged, x :Number of beds,

$$N = 393, \quad n = 15, \quad m = 3, r = 5,$$

$$\mu_{\bar{X}} = 274.70, \quad \mu_{\bar{Y}} = 814.65,$$

$$\rho = 0.9105, \quad \beta_2(x) = 3.5670, \quad \mu_{\bar{X}_1} = 214.13, \mu_{\bar{X}_2} = 980.63,$$

$$\mu_{\bar{X}_3} = 216.78, C_y = 0.7239, \quad C_x = 0.7762, \quad C_{x1} = 0.7729,$$

$$C_{x2} = 0.7756, C_{x3} = 0.7634, \quad C_{yx} = 0.5116, \quad C_{yx1} = 0.5094,$$

$$C_{yx2} = 0.5112, C_{yx3} = 0.5031,$$

Data Set-I & Set-II are used to find Percent Relative Efficiency (*PRE*) of the suggested and existing estimators which are provided in the below Table.2.1 and Table 2.2. Where PRE Calculated by

$$PRE = \left(\frac{\bar{y}_{RSS}}{\bar{y}_{(j)}}\right) \times 100 \text{ where } j = rRSS, PRSS, k, EI(I)\&EI(II)$$

Table .1 *PRE* of proposed and existing estimators against the classical estimator $\bar{y}_{rss}$ under *RSS* using Data Set-I

z	r	n	$\bar{y}_{rss}$	$\bar{y}_{rRSS}$	$\bar{y}_{pRSS}$	$\hat{y}_k$	$\bar{y}_{EI(I)}$	$\bar{y}_{EI(II)}$
3	3	9	100	321.8028	29.4942	339.0200	10784.6377	7721.3051
	4	12	100	255.6787	30.1785	276.6335	13046.0481	7101.1332
	5	15	100	203.8473	31.8959	234.9460	33960.4117	6956.0493
	Condition		337.6241	163.8434	1138.022	148.94146	1	5.88313026
4	3	12	100	255.7760	30.1785	276.5992	13046.0481	7101.6619
	4	16	100	199.4052	32.1450	225.7158	39760.1648	6998.4611
	5	20	100	163.9005	33.1753	193.2744	63702.4904	7541.5669
	Condition		639.0049	393.0392	2041.3	344.03197	1	6.44770576
5	3	15	100	209.8325	29.8656	233.7291	33966.3552	6956.0493
	4	20	100	164.8944	30.1727	190.1539	63655.5784	7537.8063
	5	25	100	135.1092	31.5642	165.5030	91938.4690	9398.9859
	Condition		917.3946	691.7070	2826.329	561.31059	1	8.78393178

Table .2 *PRE* of proposed & competing estimators in comparison with the classical

m	r	n	$\bar{y}_{rss}$	$\bar{y}_{rRSS}$	$\bar{y}_{pRSS}$	$\hat{y}_k$	$\bar{y}_{EI(I)}$	$\bar{y}_{EI(II)}$
3	3	9	100	62.8914	38.2272	108.6782	67613.986	747.1152
	4	12	100	48.8627	46.1779	103.4068	97191.463	662.1831
	5	15	100	40.0135	58.4588	109.9455	125893.58	729.5428
	Condition		1259.9258	3228.898	2190.006	1156.555	1	173.80053
4	3	12	100	48.8993	46.5472	102.2886	99058.559	667.9221
	4	16	100	37.5702	64.9104	112.5237	137819.70	802.6948
	5	20	100	30.8308	108.1561	162.5537	175437.55	1435.4121
	Condition		1755.3655	5882.0543	1638.205	1086.933	1	123.30554
5	3	15	100	40.0362	59.0604	106.5339	128209.90	737.2655
	4	20	100	30.8899	107.400	155.8833	175442.87	1423.4249
	5	25	100	25.1705	669.538	1813.695	211709.10	11433.600
	Condition		2118.0810	8759.9458	317.6732	117.7919	1	19.517931

4.2. Simulation Study

The performance of the suggested estimators of finite population mean under ranked set sampling is evaluated through a simulation study that generates data from a bivariate normal population. In accordance with RSS technique, samples are drawn from a finite population, and estimates for each sample are computed based on simulation study circumstances. This process is then repeated 10,000 times. Each sample's sample estimations of the MSEs are computed, then averaged throughout the course of all 10,000 iterations. The following formulas are used to determine the RE, which are then reported in Table 3.

$$Var(\bar{y}_{rss}) = \frac{1}{\bar{Y}^2}\left(\frac{1}{10000}\sum_{i=1}^{10000}\left(\bar{y}_{Rss} - \mu_{\bar{Y}}\right)^2\right) \quad \text{and} \quad RE(\bar{y}_{(j)}) = \frac{Var(\bar{y}_{rss})}{MSE(\bar{y}_{(j)})}$$

Table .3 *REs* of the estimators through Simulation studies

m	r	n	$\bar{y}_{rss}$	$\bar{y}_{rRSS}$	$\bar{y}_{pRSS}$	$\hat{y}_k$	$\bar{y}_{EI(I)}$	$\bar{y}_{EI(II)}$
3	3	9	100	391.085	55.3701	6863.794	48872.125	6998.5405
	4	12	100	508.133	67.6857	6792.339	97725.657	7495.3273
	5	15	100	608.058	89.2903	7196.521	205855.073	8054.3072
4	3	12	100	73.9280	71.2912	6844.583	65683.864	7850.6979
	4	16	100	615.617	92.6902	6753.449	95879.696	9271.7919
	5	20	100	717.093	114.373	6948.431	278718.974	8192.8887
5	3	15	100	586.314	83.7240	6811.577	79388.7823	7079.5741
	4	20	100	835.021	120.8050	7223.613	141526.126	8398.2794
	5	25	100	1108.931	2710.56	7989.071	555904.261	8714.6455

5. Discussion

This study includes state-of-the-art work on the linear combination of classical estimators, such as product, ratio, and exponential estimators, to effectively estimate the mean among ranking samples. The suggested estimators performed exceptionally well if there is a positive correlation between the main variable of interest and the auxiliary variables when tested using real datasets and simulated experiments. The linear combination of the classical estimators utilized in the estimators' creation results in a significant reduction of the mean square error (MSE) of the suggested estimators. The theoretical characteristics of the generated estimators are derived and contrasted with the estimators that currently exist in Section 4.1. Additionally, every circumstance in which the suggested estimators perform better than rival estimators is covered in detail. According to a numerical analysis, each of these requirements is satisfied. An empirical analysis was conducted in Section 4.2. The performance of the proposed and competing estimators is displayed as a line graph for the provided populations when all requirements are met. Figures 1 and 2 demonstrate how the suggested estimator significantly lowers MSE and, as a result, increases PRE. The conditions under which the recommended estimators outperform are found to be satisfied and a noticeably high relative efficiency (REs) is achieved are also noted in section 4.3 of the simulation research. As demonstrated in Table 3., the suggested estimators consistently outperform the rival estimators.

6. Conclusion

In this Article, two novel estimators are introduced by employing linear combination of classical estimators like, product, ratio and exponential estimators under ranked set sampling methods for predicting the mean of finite population. Both estimators $\bar{y}_{EI(I)}$ and $\bar{y}_{EI(II)}$ are compared with the already existing estimators under RSS through two real data sets and simulation studies. The tables 1, 2, and 3, demonstrate that the suggested estimators of mean for finite population are more efficient than the existing estimators. It is also observed that efficiency is gain through linear combination and by the usage of proper optimization constants. Furthermore, the simulation studies demonstrated that the proposed estimators exhibit great efficiency when compared to other competing estimators. Therefore, based on the outcomes of simulation studies and real data sets, we conclude that $\bar{y}_{EI(I)}$ and $\bar{y}_{EI(II)}$are preferred. Therefore, due to the outperformance of $\bar{y}_{EI(I)}$ and $\bar{y}_{EI(II)}$ with respect to already existing estimators, for determining the finite population mean, these estimators can be recommended.

The broad benefits and new ideas for parameter estimation in this research will help many industries that use rank sampling. In addition to RSS, additional designs such as basic random sampling, random sampling with stratification, adaptive sampling, etc. can be effectively generalized using this work. In a similar vein, this study can be extended to estimate other parameters, such as variance, coefficients of variation, etc. By using the concepts of the suggested estimators, researchers and practitioners can provide population mean estimates that are more precise and trustworthy. All things considered, this research is extremely important because it clarifies estimation under ranked sampling schemes and advances statistical methods for trustworthy data interpretation.

Conflict of Interest Declaration

The work provided in this study was not impacted by any known financial or personal links, as the authors confirm.

Data availability

The study's findings are authenticated by the data, which is available within the article

Funding: No funding was received.

References

1. Ali, H., Asim, S. M., Ijaz, M., Zaman, T., & Iftikhar, S. (2024). Improvement in variance estimation using transformed auxiliary variable under simple random sampling. *Scientific Reports*, *14*(1), 8117. https://doi.org/10.1038/s41598-024-58841-x

2. Ali, H., Asim, S. M., & Sher, K. (2024). The impact of transformations on the performance of variance estimators of finite population under adaptive cluster sampling with application to ecological data. Journal of King Saud University-Science, 103287.
3. Sher, K., Ameeq, M., Hassan, M. M., Albalawi, O., & Afzal, A. (2024). Development of improved estimators of finite population mean in simple random sampling with dual auxiliaries and its application to real world problems. Heliyon, 10(10).
4. Bouza, C., Matemática, F., & Lázaro, S. (1982). *RANKED SET SAMPLING FOR THE PRODUCT*.
5. Dell, T. R., & Clutter, J. L. (1972). Ranked Set Sampling Theory with Order Statistics Background. *Biometrics*, *28*(2), 545–555. https://doi.org/10.2307/2556166
6. *Finite Population Sampling and Inference: A Prediction Approach | Wiley*. (n.d.). Wiley.Com. Retrieved August 17, 2023, from https://www.wiley.com/en-us/Finite+Population+Sampling+and+Inference%3A+A+Prediction+Approach-p-9780471293415
7. Grover, L., & 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*.
8. https://doi.org/10.1080/03610918.2012.736579
9. Halls, L. K., & Dell, T. R. (1966). Trial of Ranked-Set Sampling for Forage Yields. *Forest Science*, *12*(1), 22–26. https://doi.org/10.1093/forestscience/12.1.22
10. Kadilar, C., Unyazici, Y., & Cingi, H. (2009). Ratio estimator for the population mean using ranked set sampling. *Statistical Papers*, *50*, 301–309. https://doi.org/10.1007/s00362-007-0079-y
11. Khan, D.-Z., & Muhammad, I. (2019). *Ratio-type Estimator of Population Mean Based on Ranked Set Sampling*. *15*, 445–449.
12. Khan, L., & Shabbïr, J. (2016). A class of Hartley-Ross type unbiased estimators for population mean using ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, *45*(3), Article 3.
13. https://dergipark.org.tr/en/pub/hujms/issue/43486/524456
14. Koyuncu, N. (2015). New difference-cum-ratio and exponential type estimators in median ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, *45*(50), 1–1. https://doi.org/10.15672/HJMS.2015509378
15. Lynne Stokes, S. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics - Theory and Methods*, *6*(12), 1207–1211. https://doi.org/10.1080/03610927708827563
16. McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural*
17. *Research*, *3*(4), 385–390. https://doi.org/10.1071/ar9520385
18. Mehta (Ranka), N., & Mandowara, V. L. (2016). A modified ratio-cum-product estimator of finite population mean using ranked set sampling. *Communications in Statistics - Theory and Methods*, *45*(2), 267–276. https://doi.org/10.1080/03610926.2013.830748
19. Muttlak, H. A. (1998). Median ranked set sampling with concomitant variables and a comparison with ranked set sampling and regression estimators. *Environmetrics*, *9*(3), 255–267. https://doi.org/10.1002/(SICI)1099-095X(199805/06)9:3<255::AID-ENV302>3.0.CO;2-2
20. *New Exponential Ratio Estimator in Ranked Set Sampling | Pakistan Journal of Statistics and Operation Research*. (n.d.). Retrieved August 17, 2023, from https://pjsor.com/pjsor/article/view/3921
21. *(PDF) Review of ranked set sampling: Modifications and applications*. (n.d.). Retrieved August 17, 2023, from https://www.researchgate.net/publication/266383828_Review_of_ranked_set_sampling_Modifications_and_applications
22. Shabbir, J., Gupta, S., & Onyango, R. (2021). On Using the Conventional and Nonconventional Measures of the Auxiliary Variable for Mean Estimation. *Mathematical Problems in Engineering*, *2021*, e3741620. https://doi.org/10.1155/2021/3741620
23. Singh, H. P., Upadhyaya, L. N., & Tailor, R. (2009). Ratio-cum-product type exponential estimator. *Statistica*, *69*(4), Article 4. https://doi.org/10.6092/issn.1973-2201/3561
24. 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*, 105–115.
25. Takahasi, K., & 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), 1–31. https://doi.org/10.1007/BF02911622
26. Yu, P. L. H., & Lam, K. (1997). Regression Estimator in Ranked Set Sampling. *Biometrics*, *53*(3), 1070. https://doi.org/10.2307/2533564