ISSN: 0974-8571 Vol.14 No. 1 June, 2022

International Journal of Computational Intelligence in Control

MODIFIED MEDIAN RANKED SET SAMPLING SCHEME FOR ESTIMATING POPULATION MEAN

Dr. Zahid Khan*, Dr. Muhammad Azeem*, Dr. Muhammad Haroon Aftab, Dr. Muhammad Ismail*****

*Department of Statistics, University of Malakand, Chakdara, Lower Dir, Khyber Pakhtunkhwa, Pakistan

**Department of Mathematics and Statistics, The University of Lahore, Lahore, Pakistan.

***Department of Statistics, COMSATS University Islamabad, Lahore Campus-Pakistan

Date of Submission: 11th January 2022 Revised: 27th March 2022 Accepted: 19th April 2022

Abstract

Ranked set sampling (RSS) is an efficient and cost-effective sampling design for estimation of population mean. RSS is modified into new sampling schemes to get more efficient estimator of population mean. Median ranked set sampling (MRSS) is one of them which estimate sample mean more efficiently than RSS. In order to make MRSS more applicable scheme, this study suggests a new sampling scheme named modified median ranked set sampling (MMRSS). Simulation study is conducted which shows that the suggested scheme is estimating population mean more efficiently than simple random sampling (SRS), ranked set sampling (RSS), and extreme ranked set sampling (ERSS). A real data set is considered for illustration of the proposed scheme.

Keywords: ranked set sampling, median ranked set sampling, extreme ranked set sampling, efficient, simple random sampling, simulation study, real data set.

1. Introduction

Ranked set sampling (RSS) is effectivein situation where visual ordering of set units can be done easily, but the exact quantification of the units is difficult and expensive. Mclntyre (1952) introduced RSS to estimate the average yield of pasture and forage. Takahasi and Wakimoto (1968) investigated that RSS provides unbiased estimator of population mean. They also proved that sample mean under RSS is more precise than that based on SRS.Dell and Clutter (1972) showed that RSS-based sample mean is not only unbiased estimator of population mean but also it is at least as efficient as the sample mean with SRS, regardless of whether the ranking is perfect or not. Stokes (1977) explained that concomitant variables which are easily available can be used to rank the variable of interest. Samawi et al., (1996) presented extreme ranked set sampling (ERSS) for estimation of population mean. Muttlak (1997) presented median ranked set sampling (MRSS) which is more efficient than its counterpart designs.Jemain et al. (2008) introduced balance group ranked set sampling (BGRSS) where the n units of size n is distributed into three groups, then, minimum units from the sets of first group, median units from second group, maximum units from the third group are taken for actual quantification. Similar to this method, Sevinc et al. (2018) suggested partial balance group ranked set sampling (PBGRSS) which present flexible sampling plans than BGRSS. Recently, Noor-ul-Amin et al., (2018) proposed new ranked set sampling (NRSS) scheme for estimation of population mean. For more modified schemes of RSS see (Al-Nasser & Mustafa, 2009; Samawi, 2011; Bani-Mustafa, 2011; Haq et al., 2013; Salehi & Ahmadi, 2014; Abbasi & Shahd, 2017; Majd and Saba, 2018; Al-Mawan et al., 2018; Noor-ul-Amin et al., 2018).

1.1 Ranked Set Sampling

The procedure of RSS for identifying $m$units is briefly explained as;randomlydraw$m$ simple random samples each of size m. Then, rank the units within each sample with respect to variable of interest byvisual inspection or byanyeconomical method. Then, select smallest ranked unit from the first sample, second smallest ranked unit from the second sample, the process is continue until largest ranked unit is selected from mth sample. The procedure is repeated r times to get $n = rm$ units.

Let X be the study variable with probability density function $f_{x(x)}$ and cumulative distribution function $F_{x(x)}$ with mean μ_x and variance σ_x^2. Let $X_{11}, X_{12} \dots \dots, X_{1m}, X_{21}, X_{22} \dots \dots, X_{2m}, \dots, X_{m1}, X_{m2} \dots \dots, X_{mm}$ be the m independent simple random samples each of size m drawn from $f_{x(x)}$. Applying the procedure of RSS, the mean and variance are follows,

$$\overline{X}_{RSS} = \frac{1}{mr}\sum_{h=1}^{r}\sum_{i=1}^{m} X_{i(i:m)h} \qquad (1)$$

$$Var(\overline{X}_{RSS}) = \frac{\sigma^2}{rm} - \frac{1}{rm^2}\sum_{i=1}^{m}\left(\mu_{i(i:m)} - \mu\right)^2 \qquad (2)$$

1.2 Extreme Ranked Set Sampling (ERSS)

The procedure of ERSS is briefly describe as, selectm^2 units from the target population using simple random sampling method, distribute them into m sets each of size m units.Rank the units within each set by visual inspection or any economical method, then, select the lowest ranked unit from the first$\frac{m}{2}$ sets, and identify largest ranked unit from the other $\frac{m}{2}$ sets select, for even sample size m. In case when$m$ is odd, the procedure is same except selecting the median unit from the mth set. The sample mean in case of even sample size$m$ is follows,

$$\overline{\mathrm{X}}_{\mathrm{ERSS(even)}} = \frac{1}{\mathrm{mr}}\sum_{\mathrm{h}=1}^{\mathrm{r}}\left(\sum_{\mathrm{i}=1}^{\frac{\mathrm{m}}{2}}\mathrm{X}_{\mathrm{i(1:m)h}} + \sum_{\mathrm{i}=\frac{\mathrm{m}+2}{2}}^{\mathrm{m}}\mathrm{X}_{\mathrm{i(m:m)h}}\right) \quad (3)$$

And its variance is,

$$\mathrm{Var}(\overline{\mathrm{X}}_{\mathrm{ERSS(even)}}) = \frac{1}{2\mathrm{mr}}\left(\sigma^2_{(1:\mathrm{m})} + \sigma^2_{(\mathrm{m:m})}\right) \quad (4)$$

Similarly, the mean of the ERSS for odd sample size is follows,

$$\overline{\mathrm{X}}_{\mathrm{ERSS(odd)}} = \frac{1}{\mathrm{mr}}\sum_{\mathrm{h}=1}^{\mathrm{r}}\left(\sum_{\mathrm{i}=1}^{\frac{\mathrm{m}-1}{2}}\mathrm{X}_{\mathrm{i(1:m)h}} + \sum_{\mathrm{i}=\frac{\mathrm{m}+1}{2}}^{\mathrm{m}-1}\mathrm{X}_{\mathrm{i(m:m)h}} + \mathrm{X}_{\mathrm{m}\left(\frac{\mathrm{m}+1}{2}:\mathrm{m}\right)\mathrm{h}}\right) \quad (5)$$

With variance:

$$\mathrm{Var}(\overline{\mathrm{X}}_{\mathrm{ERSS(odd)}}) = \frac{1}{\mathrm{m}^2\mathrm{r}}\left(\frac{\mathrm{m}-1}{2}\left(\sigma^2_{(1:\mathrm{m})} + \sigma^2_{(\mathrm{m:m})}\right) + \sigma^2_{\left(\frac{\mathrm{m}+1}{2}:\mathrm{m}\right)}\right) \quad (6)$$

1.3 Median Ranked Set Sampling

The procedure of MRSS is briefly describe as; select m^2 simple random samples from the target population. Distribute these units into $m$sets each of size $m$units, and rank them visually or by any inexpensive method. Then, if the sample size m is even, identify $\left(\frac{m}{2}\right)$th smallest ranked unit from the first $\left(\frac{m}{2}\right)$ sets, and select $\left(\frac{m+2}{2}\right)$ th smallest ranked unit from the other $\left(\frac{m}{2}\right)$ sets. In case when the sample size m is odd, identify $\left(\frac{m+1}{2}\right)$th smallest ranked unit form all sets for actual quantification. The estimator of population mean under MRSS is follows,

$$\bar{X}_{MRSS(even)}=\frac{1}{rm}\sum_{h=1}^{r}\left(\sum_{i=1}^{\frac{m}{2}}X_{i\left(\frac{m}{2}:m\right)j}+\sum_{i=\frac{m}{2}+1}^{m}X_{i\left(\frac{m+2}{2}:m\right)h}\right) \qquad (7)$$

$$\bar{X}_{MRSS(odd)}=\frac{1}{rm}\left(\sum_{h=1}^{r}\sum_{i=1}^{m}X_{i\left(\frac{m+1}{2}:m\right)h}\right) \qquad (8)$$

The variance of $\bar{X}_{MRSS(even)}$ and $\bar{X}_{MRSS(odd)}$ are,

$$Var\left(\bar{X}_{MRSS(evene)}\right)$$
$$=\frac{1}{rm}\sum_{h=1}^{r}\left(\sum_{i=1}^{\frac{m}{2}}Var\left(X_{i\left(\frac{m}{2}:m\right)h}\right)\right.$$
$$\left.+\sum_{i=\frac{m}{2}+1}^{m}Var\left(X_{i\left(\frac{m+2}{2}:n\right)h}\right)\right) \qquad (9)$$

$$Var\left(\bar{X}_{MRSS(odd)}\right)$$
$$=\frac{1}{rm}\left(\sum_{h=1}^{r}\sum_{i=1}^{m}Var\left(X_{i\left(\frac{m+1}{2}:m\right)h}\right)\right) \qquad (10)$$

Median ranked set sampling is more efficient method than SRS, RSS, ERSS and many other modified RSS schemes. This design is inapplicable when median ranked unit within all or some sets cannot be identified. We introduce a new scheme where MRSS is valid even median ranked unit cannot be recognized in some sets. The procedure of proposed design is in the next section.

2. Modified Median Ranked Set Sampling (MMRSS)

The suggested design MMRSS is mixture of MRSS and ERSS. In this scheme, median ranked units are selected in those sets where it can be identified. In the remaining sets, ERSS is applied to select the units for actual quantification.The procedure of proposed design is described as,

Step 1: Randomly select m^2units from the target population, and allocate them randomly into m sets, each of size m units.

Step 2: Rank the units within each set by visual inspection or by any economical method.

Step 3: Select c $(c \leq m)$ units from the$c$ sets using MRSS procedure, where c denote those sets where the median ranked unit can be identified.

Step 4: Select the remaining $(m-c)$ units from $(m-c)$ sets using the usual ERSS procedure.

Step 5: Repeat Steps 1 through 4 for r cycles to obtain a sample of size $mr$for actual measurement.

For $c = 0$,the proposed design is equivalent to ERSS, and for $c = m$the proposed design is same to MRSS. Thus the proposed design is special case of both MRSS and ERSS designs.

2.1 Example of MMRSS

For $m = 5, c = 2$, the MMRSS can be selected as follows,

$$\begin{bmatrix} X_{1(1:5)}, X_{2(1:5)}, \boxed{X_{3(1:5)}}, X_{5(1:5)}, X_{5(1:5)} \\ X_{1(2:5)}, X_{2(2:5)}, \boxed{X_{3(2:5)}}, X_{5(2:5)}, X_{5(2:5)} \\ \boxed{X_{1(3:5)}}, X_{2(3:5)}, X_{3(3:5)}, X_{5(3:5)}, X_{5(3:5)} \\ X_{1(4:5)}, X_{2(4:5)}, X_{3(4:5)}, X_{4(4:5)}, \quad \boxed{X_{5(4:5)}} \\ \boxed{X_{1(5:5)}}, X_{2(5:5)}, X_{3(5:5)}, X_{4(5:5)}, \boxed{X_{5(5:5)}} \end{bmatrix}$$

Thus, $X_{3(1:5)}, X_{3(2:5)}, X_{1(3:5)}, X_{5(4:5)}, \frac{1}{2}\left(X_{1(5:5)} + X_{5(5:5)}\right)$ are MMRSS samples.

2.3 Estimation of the population mean

To calculate sample mean and variance of MMRSS, fourcases are considered.

1. When m, c and $m-c$ are even,

$$\bar{X}_{MMRSS} = \frac{1}{mr}\sum_{h=1}^{r}\left(\sum_{i=1}^{\frac{c}{2}} X_{i\left(\frac{m}{2}:m\right)h} + \sum_{i=\frac{c}{2}+1}^{c} X_{i\left(\frac{m}{2}+1:m\right)h} + \sum_{i=c+1}^{\frac{m-c}{2}} X_{i(1:m)h} + \sum_{i=\frac{m-c}{2}+1}^{m} X_{i(m:m)h}\right) \quad (11)$$

$$var(\bar{X}_{MMRSS}) = \frac{1}{m^2 r}\left(\sum_{i=1}^{\frac{c}{2}} \sigma^2_{\left(\frac{m}{2}:m\right)} + \sum_{i=\frac{c}{2}+1}^{c} \sigma^2_{\left(\frac{m}{2}+1:m\right)} + \sum_{i=c+1}^{\frac{m-c}{2}} \sigma^2_{(1:m)} + \sum_{i=\frac{m-c}{2}+1}^{m} \sigma^2_{(m:m)}\right)$$

$$= \frac{1}{2m^2r}\left[c(\sigma^2_{\left(\frac{m}{2}:m\right)} + \sigma^2_{\left(\frac{m}{2}+1:m\right)}) + (m - c)\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right] \quad (12)$$

2. When m and c are odd and $m - c$ is even, the mean is follows,

$$\bar{X}_{MMRSS} = \frac{1}{mr}\sum_{h=1}^{r}\left(\sum_{i=1}^{c} X_{i\left(\frac{m+1}{2}:m\right)h} + \sum_{i=c+1}^{\frac{m-c}{2}} X_{i(1:m)h} + \sum_{i=\frac{m-c}{2}+1}^{m} X_{i(m:m)h}\right) \quad (13)$$

$$var(\bar{X}_{MMRSS}) = \frac{1}{m^2r}\left(\sum_{i=1}^{c}\sigma^2_{\left(\frac{m+1}{2}:m\right)} + \sum_{i=c+1}^{\frac{m-c}{2}}\sigma^2_{(1:m)} + \sum_{i=\frac{m-c}{2}+1}^{m}\sigma^2_{(m:m)}\right)$$

$$= \frac{1}{2m^2r}\left[2c\sigma^2_{\left(\frac{m+}{2}:m\right)} + (m - c)\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right] \quad (14)$$

3. When $m$is odd ,$c$is even, and $m - c$ is odd,

$$\bar{X}_{MMRSS} = \frac{1}{mr}\sum_{h=1}^{r}\left(\sum_{i=1}^{c} X_{i\left(\frac{m+1}{2}:m\right)h} + \sum_{i=c+1}^{\frac{m-c-1}{2}} X_{i(1:m)h} + \sum_{i=\frac{m-c}{2}+1}^{m-1} X_{i(m:m)h} + \frac{1}{2}\left(X_{1(m:m)h} + X_{m(m:m)h}\right)\right) \quad (15)$$

$$var(\bar{X}_{MMRSS}) = \frac{1}{m^2 r}\left(\sum_{i=1}^{c}\sigma^2_{\left(\frac{m+1}{2}:m\right)} + \sum_{i=c+1}^{\frac{m-c-1}{2}}\sigma^2_{(1:m)} + \sum_{i=\frac{m-c-1}{2}+1}^{m-1}\sigma^2_{(m:m)} + \frac{1}{2}\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right)$$

$$= \frac{1}{2m^2 r}\left[2c\sigma^2_{\left(\frac{m+1}{2}:m\right)} + (m - c + 1)\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right] \quad (16)$$

4. When $m$is even, c is odd, and $m-c$is odd,

$$\bar{X}_{MMRSS} = \frac{1}{mr}\sum_{h=1}^{r}\left(\sum_{i=1}^{\frac{c-1}{2}} X_{i\left(\frac{m}{2}:m\right)h} + \sum_{i=\frac{c-1}{2}+1}^{c-1} X_{i\left(\frac{m}{2}+1:m\right)h} + \frac{1}{2}\left(X_{c(1:m)h} + X_{c(m:m)h}\right) + \sum_{i=c+1}^{\frac{m-c-1}{2}} X_{i(1:m)h} + \sum_{i=\frac{m-c-1}{2}+1}^{m-1} X_{i(m:m)h} + \frac{1}{2}\left(X_{m(1:m)h} + X_{m(m:m)h}\right)\right) \quad (17)$$

$$var(\bar{X}_{MMRSS}) = \frac{1}{m^2 r}\left(\sum_{i=1}^{\frac{c-1}{2}}\sigma^2_{\left(\frac{m}{2}:m\right)} + \sum_{i=\frac{c}{2}+1}^{c-1}\sigma^2_{\left(\frac{m}{2}+1:m\right)} + \frac{1}{2}\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right) + \sum_{i=c+1}^{\frac{m-c-1}{2}}\sigma^2_{(1:m)} + \sum_{i=\frac{m-c-1}{2}+1}^{m-1}\sigma^2_{(m:m)} + \frac{1}{2}\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right)$$

$$= \frac{1}{2m^2 r}\left[(c-1)\left(\sigma^2_{\left(\frac{m}{2}:m\right)} + \sigma^2_{\left(\frac{m}{2}+1:m\right)}\right) + (m - c + 1)\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right)\right] \quad (18)$$

3. Simulation Study

In this section, efficiency of estimator of population mean under MMRSS method relative to SRS method is compare with REs of RSS, ERSS, and MRSS. Four symmetric distributions; normal(0,1), beta(3,3), logistic(0,1), and lognormal(0,1) are considered. The simulation results is presented in table 1, which shows that the estimator of proposed design is perform better than RSS, ERSS and SRS. However, the MRSS is performed well than the proposed distribution. But MRSS is not applicable when median ranked unit is not identified within all or some sets.

Table 1

Efficiency of MMRSS, ERSS, MRSS and RSS relative to SRS

Distribution	Design	m=5	m=6	m=7
N(0,1)	MMRSS(c=2)	3.1043	2.4843	3.3045
	MMRSS(c=4)	3.4401	3.3302	3.7752
	ERSS	2.4862	2.4060	2.7321
	RSS	2.774	3.161	3.684
	MRSS	3.5052	4.1898	4.5935
beta(3,3)	MMRSS(c=2)	2.8784	3.3013	3.7648
	MMRSS(c=4)	3.0094	3.2612	3.8744
	ERSS	2.8516	3.1987	3.7320
	RSS	2.9097	3.0392	3.8159
	MRSS	3.0094	3.5008	3.9535
Logistic(0,1)	MMRSS(c=2)	3.0926	2.2737	2.5469
	MMRSS(c=4)	4.0879	3.0335	3.6313
	ERSS	2.0638	1.7979	2.0185
	RSS	2.5368	2.906	3.305
	MRSS	4.0879	5.0333	5.7816
Lognormal(0,1)	MMRSS(c=2)	7.3967	1.1845	1.1998
	MMRSS(c=4)	10.4358	1.6469	2.0232
	ERSS	0.9256	0.6601	0.6192
	RSS	1.3748	1.6335	1.9430
	MRSS	10.1821	11.86	16.1720

4. Real Data Set

A real data set is taken for illustration and comparison of the proposed scheme MMRSS with SRS, RSS and ERSS. Following (Statistics online computation resource data, 1993) dataset of heights and weights of 18 years old children collected in Hong Kong by Growth Survey. The variable of interest is height of the 18 years old children. We are interested to select $m = 4$samples using suggested scheme MMRSS, RSS, ERSS and SRS. To do so, 16 units are taken randomly from the mentioned survey data. Then, procedure of MMRSS, ERSS, RSS and SRS is applied to identify $m = 4$ units.The mean and variance for the suggested scheme MMRSS, RSS, ERSS and SRS are presented in table 2. The variance of MMRSS is less than RSS, ERSS and SRS, proving the superiority on other schemes.

Table 2. Summery statistics of samples taken at different sampling schemes for m=4, r=1.

Estimator	MMRSS(c=2)	MMRSS(c=4)	RSS	ERSS	SRS
Mean	68.01	68.24	67.18	67.71	68.22
Variance	4.59	4.26	8.04	5.52	9.89

5. Conclusion

In the present paper, a new sampling schememodified median raked set sampling (MMRSS) is suggested. The proposed design merges the procedure of MRRS with ERSS to increase the efficiency of mean estimator. Also, the MMRSS ensure the use of MRSS even in some sets median ranked units cannot be identified. Numerical study showed that the MMRSS is performed well in estimating population mean.

References

[1] Mclntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research , 3(4), 385-90.

[2] 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, 1–3.

[3] Dell, T. R & Clutter. J. L. (1972). Ranked set sampling theory with order statistics background. Biometrics , 28, 545-555.

[4] Stoke, S. L. (1977). Ranked set sampling with concomitant variables. Communications in Statistics-Theory and Methods , 6, 1207-1211.

[5] Samawi, H. M., Ahmed, M. S., & Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Journal,* 38(5), 577-586.

[6] Muttlak, H. A. (1997). Median ranked set sampling. *Journal of Applied Statistical Sciences* , 6,245-255.

[7] Jemain, A. A., Al-Omari, A., & Ibrahim, K. (2008). Some variations of ranked set sampling. *Electronic Journal of Applied Statistical Analysis*, 1(1), 1-15.

[8] Sevinc, B., Gurler, S., &Cetintav, B. (2018). Partial groups ranked set sampling and mean estimation. *Journal of Statistical Computation and Simulation* , 88 ,1-12.

[9] Noor-ul-Amin, M., Tayyab, M., and Hanif, M. (2018). Mean estiamtaion using even orderstatistics. *Punjab University Journal of Mathematics* , 51(1), 91-99.

[10] Al-Nasser, A. D., and Mustafa, A. B. (2009.). Robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 79(7), 859-867.

[11] Samawi, H. M. (2011). Varied set size ranked set sampling with applications to mean and ratio estimation. *International Journal of Modelling and Simulation* , 31(1), 6-13.

[12] Bani-Mustafa, A., Al-Nasser, A. D., & Aslam, M. (2011). Folded ranked set sampling for asymmetric distributions. *Communications for Statistical Applications and Methods*,18(1), 147-153.

[13] Haq, A., Brown, J., Moltchanova, E., & Al Omari, A. I. (2013). Partial ranked set sampling design. *Environmetrics*, 24, 201-207.

[14] Salehi, M., & Ahmadi, J. (2014.). Record ranked set sampling scheme. *Metron*, 72(3), 351-365.

[15] Abbasi, A. M., and Shahd, M. Y. (2017). Manipulation-based ranked set sampling scheme. *Pakistan Journal of Statistics and Operation Research*, 13(4), 775-798.

[16] Al-Mawan, N.M., Rady.E.A., & Rashwan. N. I. (2018). Median double ranked set sampling. *Journal of Advances in Mathematics*, 14(1), 7503-7511.

[17] Majd, M. H. H., & Saba, R. A. (2018). Robust extreme double ranked set sampling. *Journal of Statistical Computation and Simulation*, 88(9), 1749-1758.

[18] Noor-ul-Amin, M., Arif, F., & Hanif, M. (2018). A new scheme using the ranked sets. *Journal of Statistics and Management Systems,* 21(8), 1565–1573.