

Maximum Ranked Set Sampling for Estimating Population Maximum

Jaya L. Limbore

Research Scholar, JIT University, Rajasthan

Dr. Rajendra G. Gurao

Brihan Maharashtra College, Pune

Abstract - Ranked set sampling (RSS) was initially proposed by McIntyre (1952) for estimating the population mean when measurement is far more costly than judgmentally comparing and ranking sampling units. It has been established in the statistical literature that ranked set sampling provides a more efficient estimator of the population mean than simple random sampling (SRS). When the purpose of sampling is to estimate the population maximum, a variation of ranked set sampling, called maximum ranked set sampling (MRSS) has been proposed. This paper investigates the sampling properties of the sample maximum under MRSS. The expected value and sampling variance of the sample maximum are derived under MRSS and are compared with the corresponding quantities for simple random sampling and ranked set sampling.

Keywords: Ranked Set Sampling, Simple Random Sampling, Maximum Ranked Set Sampling, Sample Maximum, Expected Value, Sampling Variance, Relative Efficiency.

I. INTRODUCTION

Statistical inference from a sample usually attempts to estimate the population mean under some assumptions on the parent population or the under-lying probability distribution. Estimators obtained for this purpose inherit their statistical properties partly from the population and partly from the sampling design. It is therefore common to develop optimal sampling designs for specifies populations or probability distributions. Simple random sampling is known to be optimal for finite populations under the assumption of discrete uniform distribution of the variable of interest. However, the sample size requirement of simple random sampling can be prohibitively high if the population is not very homogeneous. Researchers have proposed various sampling designs that perform better than simple random sampling for heterogeneous populations. One of the more recent among these sampling designs is ranked set sampling. McIntyre (1952) proposed the method of Ranked Set Sampling (RSS) for estimating the population mean when measurements are prohibitively costly in comparison to making comparisons between sampling units without making measurement. He proposed a sampling design that would select an initial simple random sample of size $n = r \times m^2$ and organize it in small sets of size m each. Then sampling units are ranked within every set from highest to lowest. The final sample for making measurement is obtained by selecting rank one unit from set one, rank two unit from set two, and so on, thus making up the final sample of a much smaller size ($r \times m$) than the initial simple random sample ($n = r \times m^2$). This sample is known as a ranked set sample, because it is based on ranked sampling units within sets and hence are expected to be more representative of the population than a simple random sample of the same size. It may be easier to understand the method of ranked set sampling by explaining it symbolically. The initial simple random sample of size $n = r \times m^2$ is arranged in r cycles of size m^2 each, where the m^2 sampling units in every cycle arranged in m sets of size m each, as shown in Table (1).

Table 1: Set up of Ranked Set Sampling with Set Size m and r cycles.

Cycle	Set				
1	1	$X_{1:1}^{(1)}$	$X_{1:2}^{(1)}$...	$X_{1:m}^{(1)}$
	2	$X_{2:1}^{(1)}$	$X_{2:2}^{(1)}$...	$X_{2:m}^{(1)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(1)}$	$X_{m:2}^{(1)}$...	$X_{m:m}^{(1)}$
2	1	$X_{1:1}^{(2)}$	$X_{1:2}^{(2)}$...	$X_{1:m}^{(2)}$
	2	$X_{2:1}^{(2)}$	$X_{2:2}^{(2)}$...	$X_{2:m}^{(2)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(2)}$	$X_{m:2}^{(2)}$...	$X_{m:m}^{(2)}$
⋮	⋮	⋮	⋮	⋮	⋮
r	1	$X_{1:1}^{(r)}$	$X_{1:2}^{(r)}$...	$X_{1:m}^{(r)}$
	2	$X_{2:1}^{(r)}$	$X_{2:2}^{(r)}$...	$X_{2:m}^{(r)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(r)}$	$X_{m:2}^{(r)}$...	$X_{m:m}^{(r)}$

In every cycle, the sampling unit ranked 1 is selected from set number 1, sampling unit ranked 2 is selected from set number 2, and so on up to the set number m. In this way, a total of $r \times m$ sampling units are selected for inclusion in the final sample for making measurement. The final selection is indicated in Table (2) as boldface characters.

Table 2: Ranked Set Sample indicated by boldfaced sampling units.

Cycle	Set				
1	1	$X_{1:1}^{(1)}$	$X_{1:2}^{(1)}$...	$X_{1:m}^{(1)}$
	2	$X_{2:1}^{(1)}$	$X_{2:2}^{(1)}$...	$X_{2:m}^{(1)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(1)}$	$X_{m:2}^{(1)}$...	$X_{m:m}^{(1)}$
2	1	$X_{1:1}^{(2)}$	$X_{1:2}^{(2)}$...	$X_{1:m}^{(2)}$
	2	$X_{2:1}^{(2)}$	$X_{2:2}^{(2)}$...	$X_{2:m}^{(2)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(2)}$	$X_{m:2}^{(2)}$...	$X_{m:m}^{(2)}$
⋮	⋮	⋮	⋮	⋮	⋮
r	1	$X_{1:1}^{(r)}$	$X_{1:2}^{(r)}$...	$X_{1:m}^{(r)}$
	2	$X_{2:1}^{(r)}$	$X_{2:2}^{(r)}$...	$X_{2:m}^{(r)}$
	⋮	⋮	⋮	⋮	⋮
	4	$X_{m:1}^{(r)}$	$X_{m:2}^{(r)}$...	$X_{m:m}^{(r)}$

The ranked set sample to be used for making measurements is then obtained as follows.

$$RSS = \{X_{1:1}^{(1)}, X_{2:2}^{(1)}, \dots, X_{m:m}^{(1)}, X_{1:1}^{(2)}, X_{2:2}^{(2)}, \dots, X_{m:m}^{(2)}, X_{1:1}^{(r)}, X_{2:2}^{(r)}, \dots, X_{m:m}^{(r)}\}. \quad (1)$$

Ranked set sampling was initially proposed to improve precision over simple random sampling and it can be shown that the sampling variance of the ranked sample mean is smaller than that of simple random sample mean when ranking is perfect. Even when there are ranking errors, it has been shown that the sampling variance of ranked set sample mean cannot exceed that of simple random sample mean, the worst case leading to equality of the two. Variations in the original design of ranked set sampling have been suggested to overcome some shortcomings of ranked set sampling, especially when there may be ranking errors because sampling units are ranked judgmentally without any quantitative measurements.

There have been other considerations regarding the possible applications of ranked set sampling like any other classical sampling design. In particular, some variations have been introduced in ranked set sampling. Notable examples include ranked set sampling with unequal allocations, where all ranks are not equally represented in the final sample, but replications of ranks are decided according to the density function of the population values of the

variable of interest. Another variation involves selecting only the extremely ranked sampling units to form an extreme ranked set sample, the justification being that extreme values are easier to identify than central ones, and hence ranking errors would be minimized if extreme ranks are selected for forming the sample. One particular case of extreme ranked set sampling is maximum ranked set sampling (see Biradar and Santosha, 2014), where only the highest ranked sampling unit is selected from every set. It is interesting to note that this variation was initially proposed when the objective was to estimate the population mean. Our interest is in estimating the population maximum and therefore it is proposed to use the maximum ranked set sampling and investigate its statistical properties so that its efficiency can be compared with that of any other sampling design, especially simple random sampling and ranked set sampling as proposed by McIntyre (1952).

II. MAXIMUM RANKED SET SAMPLING

Maximum ranked set sampling is developed from ranked set sampling as follows. The initial steps of ranked set sampling are followed up to forming

$r \times m$ sets of size m each from the initial simple random sample of size $r \times m^2$. After forming the sets and ranking sampling units within every set, only the highest ranked sampling unit is selected from every set to form the maximum ranked set sample. Thus, the maximum ranked set sample is obtained by selecting the bold-faced sampling units shown in Table (3).

Table 3: Maximum Ranked Set Sample with Set Size m and r cycles.

Cycle	Set				
1	1	$X_{1:1}^{(1)}$	$X_{1:2}^{(1)}$...	$X_{1:m}^{(1)}$
	2	$X_{2:1}^{(1)}$	$X_{2:2}^{(1)}$...	$X_{2:m}^{(1)}$
	⋮	⋮	⋮	⋮	⋮
	m	$X_{m:1}^{(1)}$	$X_{m:2}^{(1)}$...	$X_{m:m}^{(1)}$
2	1	$X_{1:1}^{(2)}$	$X_{1:2}^{(2)}$...	$X_{1:m}^{(2)}$
	2	$X_{2:1}^{(2)}$	$X_{2:2}^{(2)}$...	$X_{2:m}^{(2)}$
	⋮	⋮	⋮	⋮	⋮
	m	$X_{m:1}^{(2)}$	$X_{m:2}^{(2)}$...	$X_{m:m}^{(2)}$
r	1	$X_{1:1}^{(r)}$	$X_{1:2}^{(r)}$...	$X_{1:m}^{(r)}$
	2	$X_{2:1}^{(r)}$	$X_{2:2}^{(r)}$...	$X_{2:m}^{(r)}$
	⋮	⋮	⋮	⋮	⋮
	m	$X_{m:1}^{(r)}$	$X_{m:2}^{(r)}$...	$X_{m:m}^{(r)}$

III. STATISTICAL PROPERTIES OF SAMPLE MAXIMUM USING MAXIMUM RANKED SET SAMPLING

Let the probability density function of the variable X of interest be denoted by $f(x)$ and the cumulative distribution function by $F(x)$. In Maximum Ranked Set Sampling, every set forms a random sample of size m . As a result, the cumulative distribution function of the set maximum $X_{i:1}^j$ for set $i = 1, 2, \dots, m$ and cycle $j = 1, 2, \dots, r$, is given by

$$F_m(x) = P(X_{(1)} \leq x) = [F(x)]^m; x \in R; \tag{2}$$

If the population mean and variance are respectively denoted by μ and σ^2 , the mean and variance of the sample maximum in a random sample of size m are respectively denoted by $\mu_{(m)}$ and $\sigma_{(m)}^2$. The set maximum is obtained from row-maxima and is defined as follows.

$$X_{\max}^j = \max(X_{1:1}, X_{2:1}, \dots, X_{m:1}), j = 1, 2, \dots, r; \tag{3}$$

with the corresponding mean and variance given by μ_{\max}^j and σ_{\max}^2 respectively. The maximum of the Maximum Ranked Set Sample is then defined by

$$X_{\max}(\text{MRSS}) = \max(X_{\max}^1, X_{\max}^2, \dots, X_{\max}^r), \tag{4}$$

with the mean and variance given respectively by $\mu_m 2_r$ and $\sigma_m 2_r$.

What these results show is that the largest observation in a Maximum Ranked Set Sample with set size m and r cycles is statistically equivalent to the largest observation in a simple random sample of size $m^2 r$ with only mr measurements. In this way, the relative cost of Maximum Ranked Set Sampling is only $1/m$ of the cost of an equivalent simple random sampling effort. The only important assumption in this result is that judgmental ranking is perfect without any ranking error. The case of possible ranking errors will be taken up in a subsequent study

IV. ILLUSTRATIVE EXAMPLE

Let us consider the case where the set size is given by $m = 4$ and the initial sample size is $n = 48$, so that there are $r = 3$ cycles in ranked set sampling. The initial sample of size $n = 48$ is arranged as follows in 12 set of size 4 each. The first 4 make up cycle 1, the next 4 cycle 2 and the last 4 make up cycle 3. This arrangement is shown in Table (4).

Table 4: Ranked Set Sampling Set-up for Set Size $m = 4$ and $r = 3$ cycles.

Cycle	Set				
1	1	$X_{1:1}^{(1)}$	$X_{1:2}^{(1)}$	$X_{1:3}^{(1)}$	$X_{1:4}^{(1)}$
	2	$X_{2:1}^{(1)}$	$X_{2:2}^{(1)}$	$X_{2:3}^{(1)}$	$X_{2:4}^{(1)}$
	3	$X_{3:1}^{(1)}$	$X_{3:2}^{(1)}$	$X_{3:3}^{(1)}$	$X_{3:4}^{(1)}$
	4	$X_{4:1}^{(1)}$	$X_{4:2}^{(1)}$	$X_{4:3}^{(1)}$	$X_{4:4}^{(1)}$
2	1	$X_{1:1}^{(2)}$	$X_{1:2}^{(2)}$	$X_{1:3}^{(2)}$	$X_{1:4}^{(2)}$
	2	$X_{2:1}^{(2)}$	$X_{2:2}^{(2)}$	$X_{2:3}^{(2)}$	$X_{2:4}^{(2)}$
	3	$X_{3:1}^{(2)}$	$X_{3:2}^{(2)}$	$X_{3:3}^{(2)}$	$X_{3:4}^{(2)}$
	4	$X_{4:1}^{(2)}$	$X_{4:2}^{(2)}$	$X_{4:3}^{(2)}$	$X_{4:4}^{(2)}$
3	1	$X_{1:1}^{(3)}$	$X_{1:2}^{(3)}$	$X_{1:3}^{(3)}$	$X_{1:4}^{(3)}$
	2	$X_{2:1}^{(3)}$	$X_{2:2}^{(3)}$	$X_{2:3}^{(3)}$	$X_{2:4}^{(3)}$
	3	$X_{3:1}^{(3)}$	$X_{3:2}^{(3)}$	$X_{3:3}^{(3)}$	$X_{3:4}^{(3)}$
	4	$X_{4:1}^{(3)}$	$X_{4:2}^{(3)}$	$X_{4:3}^{(3)}$	$X_{4:4}^{(3)}$

The maximum ranked set sample of size 12 is obtained by selecting the sampling units indicated by bold face in Table (5).

Table 5: Ranked Set Sample with Set Size $m = 4$ and $r = 3$ cycles.

Cycle	Set				
1	1	$X_{1:1}^{(1)}$	$X_{1:2}^{(1)}$	$X_{1:3}^{(1)}$	$X_{1:4}^{(1)}$
	2	$X_{2:1}^{(1)}$	$X_{2:2}^{(1)}$	$X_{2:3}^{(1)}$	$X_{2:4}^{(1)}$
	3	$X_{3:1}^{(1)}$	$X_{3:2}^{(1)}$	$X_{3:3}^{(1)}$	$X_{3:4}^{(1)}$
	4	$X_{4:1}^{(1)}$	$X_{4:2}^{(1)}$	$X_{4:3}^{(1)}$	$X_{4:4}^{(1)}$
2	1	$X_{1:1}^{(2)}$	$X_{1:2}^{(2)}$	$X_{1:3}^{(2)}$	$X_{1:4}^{(2)}$
	2	$X_{2:1}^{(2)}$	$X_{2:2}^{(2)}$	$X_{2:3}^{(2)}$	$X_{2:4}^{(2)}$
	3	$X_{3:1}^{(2)}$	$X_{3:2}^{(2)}$	$X_{3:3}^{(2)}$	$X_{3:4}^{(2)}$
	4	$X_{4:1}^{(2)}$	$X_{4:2}^{(2)}$	$X_{4:3}^{(2)}$	$X_{4:4}^{(2)}$
3	1	$X_{1:1}^{(3)}$	$X_{1:2}^{(3)}$	$X_{1:3}^{(3)}$	$X_{1:4}^{(3)}$
	2	$X_{2:1}^{(3)}$	$X_{2:2}^{(3)}$	$X_{2:3}^{(3)}$	$X_{2:4}^{(3)}$
	3	$X_{3:1}^{(3)}$	$X_{3:2}^{(3)}$	$X_{3:3}^{(3)}$	$X_{3:4}^{(3)}$
	4	$X_{4:1}^{(3)}$	$X_{4:2}^{(3)}$	$X_{4:3}^{(3)}$	$X_{4:4}^{(3)}$

The ranked set sample, in this way, will be as follows.

$$RSS = \{X_{1:1}^{(1)}, X_{2:1}^{(1)}, X_{3:1}^{(1)}, X_{4:1}^{(1)}, X_{1:1}^{(2)}, X_{2:1}^{(2)}, X_{3:1}^{(2)}, X_{4:1}^{(2)}, X_{1:1}^{(3)}, X_{2:1}^{(3)}, X_{3:1}^{(3)}, X_{4:1}^{(3)}\}. \quad (5)$$

As an illustrative example, we take a population of size $N = 1000$, and assume that $y_i = i$ for $i = 1, 2, \dots, 1000$. A random sample of size $n = 48$ is selected and it is given by
 180, 668, 960, 356, 206, 956, 84, 342, 951, 883, 122, 456, 555, 172, 861, 384, 299, 389, 717, 110, 263, 641, 407, 239, 237, 216, 424, 965, 746, 165, 793, 966, 364, 57, 231, 178, 811, 739, 75, 812, 63, 820, 829, 503, 90, 1000, 977, 473.

Arranged in the set up for ranked set sampling, this sample appears as follows (Table 6).

Table 6: Simple Random Sample of size $n = 48$ from a population of size $N = 1000$ arranged in Sets of Size $m = 4$ and $r = 3$ cycles.

Cycle	Set				
1	1	180	668	960	356
	2	206	956	84	342
	3	951	883	122	456
	4	555	172	861	384
2	1	299	389	717	110
	2	263	641	407	239
	3	237	216	424	965
	4	746	165	793	966
3	1	364	57	231	178
	2	811	739	75	812
	3	63	820	829	503
	4	90	1000	977	473

After arranging the sampling units within each set from highest to lowest, the following table is obtained as the set-up for Maximum Ranked Set Sampling (see Table 7).

Table 7: Simple Random Sample of size $n = 48$ from a population of size $N = 1000$ arranged in Sets of Size $m = 4$ and $r = 3$ cycles.

Cycle	Set				
1	1	960	668	356	180
	2	956	342	206	84
	3	951	883	456	122
	4	861	555	384	172
2	1	717	389	299	110
	2	641	407	263	239
	3	965	424	237	216
	4	966	793	746	165
3	1	364	231	178	57
	2	812	811	739	75
	3	829	820	503	63
	4	1000	977	473	90

Finally, the maximum ranked set sample is shown by the boldface number in Table 8.

Table 8: Simple Random Sample of size $n = 48$ from a population of size $N = 1000$ arranged in Sets of Size $m = 4$ and $r = 3$ cycles.

Cycle	Set				
1	1	960	668	356	180
	2	956	342	206	84
	3	951	883	456	122
	4	861	555	384	172
2	1	717	389	299	110
	2	641	407	263	239
	3	965	424	237	216
	4	966	793	746	165
3	1	364	231	178	57
	2	812	811	739	75
	3	829	820	503	63
	4	1000	977	473	90

The largest element from cycle 1 is $\max(960, 956, 951, 861) = 960$. The largest elements from cycle 2 is $\max(717, 641, 965, 966) = 966$. The largest element from cycle 3 is $\max(364, 811, 829, 1000) = 1000$.

Finally, the largest element in the entire maximum ranked set sample is 1000. This also happens to be the largest element in the initial simple random sample of size $n = 48$.

The results of this paper are obtained under the assumption that ranking is perfect and there are no ranking errors. There is always a possibility of error in ranking because ranking is based on judgment without making any measurement. The situation where ranking errors may be present is beyond the scope of this paper and is therefore not discussed in this paper. A subsequent paper will deal with the situation where ranking may be imperfect and may affect the performance of maximum ranked set sampling.

REFERENCES

- [1] Al-Nasser, A. D. and Mustafa, A. B. (2009). Robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, Vol. 79, No. 7, pp. 859 - 867.
- [2] Al-Omari, A. I. (2011). Estimation of mean based on modified robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, Vol. 81, No. 8, pp. 1055-1066.
- [3] Al-Omari, A.I. and Al-Nasser, A.D. (2012). On the population median estimation using robust extreme ranked set sampling. *Monte Carlo Methods and Applications*, Vol. 18, pp. 109-118.
- [4] Al-Saleh, M. F. and Samawi, H. M. (2010). On estimating the odds using Moving extreme ranked set sampling. *Statistical Methodology*, Vol. 7, No. 2, pp. 133 - 140.
- [5] Biradar, B. S. and Santosha, C. D. (2014). Estimation of the Mean of the Exponential Distribution Using Maximum Ranked Set Sampling with Unequal Samples. *Open Journal of Statistics*, Vol. 4, pp. 641 - 649.
- [6] Ahmed Ali Hanandeh, and Mohammad Fraiwan Al-Saleh (2013). Inference on Downtons Bivariate Exponential Distribution Based on Moving Extreme Ranked Set Sampling. *AUSTRIAN JOURNAL OF STATISTICS*, Vol. 42, No. 3, pp. 161-179.
- [7] McIntyre, G. A. (1952). A Method for Unbiased Selective Sampling Using Ranked Sets. *Australian Journal of Agricultural Research*, Vol. 3, pp. 385 - 390.
- [8] Muttalak, H. A. (1999). On extreme ranked set sampling with size biased probability of selection. *Far East Journal of Theoretical Statistics*, Vol. 3, pp. 319-329.
- [9] Muttalak, H. A. (2001). Regression estimators in extreme and median ranked set samples. *Journal of Applied Statistics*, Vol. 28, pp. 1003-1017.
- [10] Hani M. Samawi, Mohammad S. Ahmed and Walid Abu-Dayyeh (1996). Estimating the Population Mean Using Extreme Ranked Set Sampling. *Biometrical Journal*, Vol. 5, pp. 577 - 586.
- [11] Samawi, H. M. and Saeid, L. J. (2004). Stratified extreme ranked set sample with application to ratio estimators. *Journal of Modern Applied Statistical Methods*, Vol. 3, No. 1, pp. 117 -133.
- [12] Samawi, H. M. and Al-Sagheer, O. A. M. (2001). On the estimation of the distribution function using extreme and median ranked set sampling. *Biometrical Journal*, VOL. 43, PP. 357 - 373.
- [13] Samawi, H.M. and Al-Saleh, M.F. (2013). Valid estimation of odds ratio using two types of moving extreme ranked set sampling. *Journal of the Korean Statistical Society*, Vol. 42, pp. 17 - 24.
- [14] Samawi, H. M., Vogel, R. and Senkowski, C. K. (2010). On distribution free runs test for symmetry using extreme ranked set sampling with an application involving base de cit score. *Journal of Statistical Theory and Practice*, Vol. 4, pp. 289 - 301.