Statistics and Applications {**ISSN 2452-7395 (online)**}
Volume 17, No. 1, 2019 (New Series), pp 121–139

Ranked Set Sampling: A Review with New Initiative on Extreme Ranked Set Sampling

Neeraj Tiwari[1], Girish Chandra[2] and Raman Nautiyal[2]
[1]*Department of Statistics, Kumaun University, S.S.J. Campus, Almora, India-263601.*
[2]*Division of Forestry Statistics, Indian Council of Forestry Research and Education, Dehradun, India-248006*

Received: September 10, 2018; Revised: December 03, 2018; Accepted: December 10, 2018

Abstract

Ranked Set Sampling (RSS) is a sampling scheme, which is mainly used under the condition that the measurement of units are difficult or time consuming but the ranking of the units by some methods other than actual measurements (e.g. visual inspection) is relatively easy. This paper provides a critical review of the major developments and issues in RSS since its inception by McIntyre (1952) and suggests a new initiative on extreme RSS Samawi et al. (1996).

Key words: Ranked set sampling; Population ratio; Location and scale parameters.

1 Introduction

Understanding environmental issues and conducting related research is beset with certain peculiar problems, especially those that require data and their subsequent analysis. One of the basic objectives of researchers engaged in environmental and ecological studies is to obtain optimum precision with a reasonable cost and time associated with field and laboratory experiments. In many studies, it is observed that the measurement of the units depends upon a host of extraneous factors like difficulty of reaching the sampling unit, costs involved, destruction of the sampling unit (especially when repeated measurements need to be taken), etc. For example, in research related to entomology (infestation of bark eating caterpillar *Inderbela quadrinotato* in *Populus deltoides* commonly known as poplar), assessment of infestation of plant species by insect pests is regularly done, the measurement of each sampled tree is not only difficult but costly as well because larva of this caterpillar makes tunnel into the trunk of the tree and feeds bark tissues covering them with fecal ribbons. It is compulsory to remove the ribbon made by the caterpillar to confirm the presence of active insects inside which requires more resources in terms of effort, cost and time. The conventional sampling techniques like simple random sampling (SRS) fails here because the actual measurement of active insects present in each tree is not feasible. However, the ranking of the trees based upon the ribbon made by this insect is relatively easy. The number of ribbons and their sizes can easily be seen in the tree and hence the ranking based upon visual inspection can be done for all the sampled trees [Figure 1]. For such situations, an environmental sampling scheme introduced by McIntyre (1952), known as Ranked Set Sampling (RSS), is most appropriate and can be utilized to potentially increase the precision and reduce costs when actual measurements of the variable of interest

Corresponding Author: Neeraj Tiwari
Email: neerajtiwari.amo@gmail.com

is costly and / or time-consuming but the ranking of the set of items according to the variable can be done without actual measurements.

Figure 1: (A) larva of *I. quadrinotata* (B) Ribbon symptom made by *I. quadrinotata* (C) Heavy attach of *I. quadrinotata* in the poplar plantation field (D) Healthy poplar plantation in the field.

Patil (2002), in a note in Encyclopedia of Environmetrics, described the method of RSS in details. In order to obtain a ranked set sample of size $n=km$, where k is the number of sample units selected in each cycle (set size) and m is the total number of cycles, the following steps are carried out:

1. The k^2 units are randomly selected from the population
2. The k^2 units are then allocated as randomly as possible into k sets, each of size k.
3. The units within each set are then ranked based on a perception of relative values for the variable under interest. This may be done based on personal judgment, expert judgment or measurement of a covariate correlated with the variable under interest (concomitant variable) whose actual measurement is relatively easier and inexpensive.
4. Actual analysis is carried out on the sample obtained by including the smallest ranked unit in the first set, second smallest unit in the second set and so on up to the largest unit in the k^{th} set.
5. Repeat steps 1 to 4 for m cycles until the desired sample size $n=km$ [Figure 2], is obtained for analysis.

$$\begin{array}{l} Y_{(1:1)j}, Y_{(1:2)j}, \ldots Y_{(1:k)j} \\ Y_{(2:1)j}, Y_{(2:2)j}, \ldots Y_{(2:k)j} \\ . \\ . \\ Y_{(k:1)j}, Y_{(k:2)j}, \ldots Y_{(k:k)j} \end{array}$$

Figure 2: Elucidation of k^2 units in k sets each of size k for cycle j

This is the case of balanced RSS (*i.e*. the number of measurements made on each ranked units is same for all the ranks), which maintains the unbiasedness of SRS and, in addition, increases its representativeness of the true underlying population. An unbiased estimator of population mean under RSS (McIntyre, 1952) is given by,

$$\bar{Y}_{(k)RSS} = \frac{1}{km}\sum_{i=1}^{k}\sum_{j=1}^{m} Y_{(i:k)j} \ , \tag{1}$$

where, $Y_{(i:k)j}$ represents the measured value of the i^{th} rank order statistic under j^{th} cycle. Takahasi and Wakimoto (1968) proved that under the perfect ranking, the relative precision (RP) lies between 1 and $(k+1)/2$. It was established by Dell and Clutter (1972) that the RSS performs better than SRS even in presence of the ranking errors, which shows that the balanced RSS is never less precise than SRS.

An unbalanced RSS or RSS with unequal allocation is one in which the ranked order statistics are not measured the same number of times. Kaur et al. (1994, 1997) proposed the near optimum allocation models for skewed distributions to overcome the certain difficulties found in Neyman's optimum allocation procedure in which the knowledge of standard deviations of the order statistics were unknown. Their allocation procedure does not provide the integer allocation values. To overcome this difficulty, Tiwari and Chandra (2011) suggested new allocation procedure which ensures the gain of RSS quite closer to the above models. RP of these models for two skewed distributions, LN (0, 1) and Pareto (3), are shown in Table 1 and Table 2, respectively (Tiwari and Chandra, 2011).

Table 1: RP of different models of LN (0, 1) for k = 2(1)8.

Set size (k)	2	3	4	5	6	7	8
Balanced RSS	1.1872	1.3393	1.4711	1.5891	1.6971	1.7974	1.8914
Kaur et al (t-model)	1.482	2.039	2.324	2.714	3.098	3.266	3.656
Kaur et al ((s, t)-model)		2.039	2.595	3.030	3.283	3.815	4.198
Neyman's Model	1.482	2.039	2.595	3.088	3.560	4.067	4.532
Tiwari and Chandra model	1.482	1.859	2.174	2.449	2.695	2.920	3.128

Table 2: RP of different models of Pareto (3) for k = 2(1)8.

Set size (k)	2	3	4	5	6	7	8
Balanced RSS	1.136	1.242	1.330	1.407	1.475	1.537	1.594
Kaur et al (t-model)	1.449	2.055	2.466	2.632	3.156	3.463	3.629
Kaur et al ((s, t)-model)		2.055	2.591	3.099	3.487	3.800	4.272
Neyman's Model	1.449	2.055	2.591	3.099	3.631	4.114	4.619
Tiwari and Chandra model	1.449	1.772	2.026	2.236	2.416	2.575	2.718

For symmetric distributions, Kaur *et al*. (2000) suggested an optimal allocation model and compared it with the equal and Neyman allocation models in terms of the RP of the estimator of population mean. This model had a disadvantage that it does not contain all the order statistics and depends only on extreme or median order statistics. Chandra et al. (2015) proposed a systematic model for symmetric distributions to overcome the drawbacks in Kaur *et al.* (2000) model in the sense that measurements are made upon each rank order. RP of this method was found to be very close to the optimum allocation model and better than the balanced RSS and Neyman's model (See, Table 3 and Table 4 for two symmetric distributions). Yanagawa and Chen (1980) suggested a minimum variance linear unbiased median-mean estimator of population mean for a family of symmetric distributions.

Shirahata (1982) examined more general procedures that are unbiased for symmetric distributions.

Table 3: RP of different models for uniform distribution for *k* = 2(1)8

Set size (k)	2	3	4	5	6	7	8
Balanced RSS	1.500	2.000	2.500	3.000	3.500	4.000	4.500
Neyman's Model	1.500	1.846	2.381	2.983	3.500	4.081	4.563
Kaur et al model	1.500	2.222	3.125	4.200	5.445	6.857	8.437
Chandra et al. model	1.500	2.111	2.778	3.458	4.265	4.982	5.909

Table 4: RP of different models for normal distribution for *k* = 2(1)8

Set size (k)	2	3	4	5	6	7	8
Balanced RSS	1.467	1.914	2.347	2.770	3.186	3.595	3.999
Neyman's model	1.467	1.747	2.199	2.656	3.186	3.633	3.932
Kaur et al. model	1.467	2.229	2.774	3.486	4.062	4.752	5.342
Chandra et al. model	1.467	2.008	2.527	3.085	3.624	4.187	4.735

For the unbalanced RSS, Takahasi and Wakimoto (1968) showed that the RP of RSS relative to SRS lies between 0 and *k*.

2 RSS Versus Other Classical Designs

Since the inception of RSS by McIntyre (1952), there had been a large number of developments in the area of RSS. It has been compared with the other classical methods viz. SRS, stratified sampling, systematic sampling, multistage sampling etc. theoretically as well as on the basis of real applications in forestry, environmental sciences, epidemiology, agriculture etc. In this section, we shall critically discuss the differences between RSS and the other classical designs being frequently used in sample surveys. Relative advantages and disadvantages of these designs shall be revisited.

With the classical designs, the entire selection of sample units in the sample is made by the actual measurement of each of the selected units. For example, in SRS, the selected units require measurement cost and time towards each and every unit of the population to get a SRS selected. This does not guarantee to have an optimum sample due to the equal probabilities of selection of each unit of the population. Similar situations also prevail in other classical designs. With RSS, however, the researcher has a greater advantage of saving time and money on the selection of units. Despite this advantage, the objective is the same, that is, to select the unit (based on actual measurement or ranking method) and then estimate some function $Z(y)$ (or equivalently the population total, $Z(y)=\sum_{i=1}^{N} y_i$, the population mean, $Z(y)=\frac{1}{N}\sum_{i=1}^{N} y_i = \mu$ or finite population variance, $Z(y)=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\mu)^2$).

For many classical sampling plans, the measurement of all the sampled units is compulsory to estimate the parameters under study. On the other hand, for RSS, the measurement of all sampled units is not compulsory and there is drastic reduction in the number of units for actual measurement. RSS gains the advantages of the stratified random sampling, in which, the sampling unit from various order statistics are for measurement. Another important advantage of RSS is that classical estimators used in conjunction with

RSS are unbiased. There are numerous other advantages that RSS may have over the classical designs. In many instances RSS can be a more efficient design, i.e. the variances of estimators will be smaller for an equivalent amount of sampling effort (Takahasi and Wakimoto, 1968).

Furthermore, beyond the cost effectiveness, RSS has other advantages. It may be implemented in any of the situations where classical designs are used for drawing the valid inferences but the procedure, formulae and data computations may differ while using RSS. For example, if RSS data are used in non-parametric inferences, the Wilcoxon Rank Sum test will be used to test for differences in the medians of two populations, and as such the computations for the Wilcoxon Rank Sum test described in Bohn and Wolfe (1992, 1994) should be used rather than the standard computations [Gibbons and Chakraborty (2010)] that would have been used if the data had been obtained using SRS.

3 RSS in Estimating Various Population Parameters

In this section, we shall discuss the use of RSS in estimating various population parameters. For the sake of convenience and clarity, this section has been divided into five sub-sections to categorize the various population parameters under consideration.

3.1 Population Mean, Variance and Distribution Functions

A substantial work has been carried out by many researchers in estimation of mean for finite and infinite populations after McIntyre (1952). This was based upon the original RSS and its modifications. Takahasi and Wakimoto (1968) provided the complete theory of estimation of mean using RSS. The problem of estimating mean of lognormal distribution with known coefficient of variation has been studied by Shen (1994). It was shown that the use of RSS and its modifications results in improved estimators compared to the SRS. Samawi et al. (1996) suggested a variety of extreme RSS for mean estimation. Muttlak (1997) introduced a median RSS to estimate population mean. Similarly other researchers discussed the mean estimation using other versions of RSS. For example, using random selection in RSS (Li et al., 1999), double extreme RSS Samawi (2002), double RSS procedure Al-Saleh and Al-Kadiri (2000), moving extremes RSS Al-Saleh and Al-Hadhrami (2003), weighted modified RSS Muttlak and Abu-Dayyeh (2004), missing observation in RSS Bouza (2008), and folded RSS Bani- Mustafa et al. (2011). Singh et al. (2014) introduced a general procedure for estimation of mean using RSS.

Dell (1969) and Dell and Clutter (1972) provided various expressions of the variance of RSS. Estimating the population variance based on judgment ordered ranked set samples was considered by Stokes (1980). A nonparametric study was considered by Perron and Sinha (2004) for estimation of population variance under RSS. Biswas et al. (2013) proposed variance estimation in RSS based on finite population framework using the Jackknife technique. On estimation of population variance, one may refer to MacEachern et al. (2002), Al-Hadhrami and Al-Omari (2006) and Chen and Lim (2011).

The kernel estimators of probability density functions using RSS were suggested by Barabesi and Fattorini (2002) and Lim et al. (2014). Lam et al. (2002) suggested nonparametric estimators for the distribution function and the mean using the auxiliary information and concomitant variable in RSS process. Frey (2014) considered bootstrap confidence bands for the CDF using RSS. Other researchers who studied on the estimation of distribution function in details are Abu-Dayyeh et al. (2002), Wolfe (2004), Baraneso and Fattorini (2002), and Huang (1972). Sengupta and Mukhuti (2006) studied the unbiased estimation of an exponential distribution variance using RSS.

3.2 RSS in Estimating Population Ratio

The case of estimating population proportion from an unbalanced RSS under perfect ranking was considered by Chen et al. (2006). The probabilities of success for order statistics were the functions of the underlying population proportion. They have shown that the Neyman's allocation is optimal as it leads to minimum variance within the class of RSS estimators that are simple averages of the means of order statistics. Other versions of RSS, such as Median RSS (Samawi and Muttlak, 2001) and Double median RSS (Samawi and Tawalbeh, 2002), were also used to estimate the population ratio. For details about the ratio estimation in RSS, one may refer to Samawi and Saeid (2004), Al-Omari et al. (2009), Kadilar et al. (2009), Al-Omari (2012) and Mandowara and Mehta (2014).

3.3 RSS in Estimating Quantiles

The problem of quantile estimation for any distribution function using unbalanced RSS has already been discussed by Chen (2001) and Zhu and Wang (2005). If m_i observations are taken for measurements for i^{th} rank order statistic, $i = 1,2,...k$, $\sum_{i=1}^{k} m_i = n$ and let $X_{(i:k)j}$, $i = 1,2,...k$; $j = 1,2,...m_i$ denotes the j^{th} observation of i^{th} order statistic. Note that the observations corresponding to the i^{th} rank order statistic are *i.i.d.* with mean $\mu_{(i:k)}$ and variance $\sigma^2_{(i:k)}$. Let $F_{(i)}$ and $f_{(i)}$ be the CDF and pdf of i^{th} rank order statistic. In estimating the p^{th} quantile, ξ_p,

$$\xi_p = \inf \{x : F(x) \geq p\} \quad ,$$

Chen (2001) proved that the method of unbalanced RSS outperforms the methods of balanced RSS and SRS in terms of asymptotic relative efficiency (ARE). He used a probability vector $\mathbf{q} = (q_1, q_2, ...q_k)$ with $\sum_{i=1}^{k} q_i = 1$ and $0 \leqslant q_i \leqslant 1$, in constructing a new CDF $F_{\mathbf{q}}(x) = \sum_{i=1}^{k} q_i F_{(i)}(x)$ and corresponding $f_{\mathbf{q}}(x) = \sum_{i=1}^{k} q_i f_{(i)}(x)$ which is useful to find the asymptotical unbiased estimator of ξ_p, where, $f_{(i)}(x) = \frac{k!}{(i-1)!(k-i)!} F^{i-1}(x)(1-F(x))^{k-i} f(x) = b(F(x); i, k+1-i)$, $b(y;a,b) = \frac{y^{i-1}(1-y)^{b-1}}{\beta(a,b)}$, the pdf of the random variable which follows beta distribution with parameters a and b and $F_{(i)}(x) = B(F(x); i, k+1-i)$.

Zhu and Wang (2005) suggested a new weighted estimator of ξ_p by overcoming the certain shortcomings in Chen (2001). They suggested that as each $\hat{\xi}_i$, $i = 1,2,...k$ was a consistent estimator of ξ_p, therefore a weighted estimator could be constructed by combining them and assigning some weights to each $\hat{\xi}_i$, where, $\hat{\xi}_i(x) = \inf\{x : \hat{F}_{(i)}(x) \geq p\}$ and $\hat{F}_{(i)}(x) = \frac{1}{m_i}\sum_{j=1}^{m_i} I\left(X_{(i)j} \leq x\right)$ is the empirical CDF of $F_{(i)}(x)$. The estimator proposed by Zhu and Wang (2005) is

$$\hat{\xi}_W = \frac{\sum_{i=1}^{k_i} w_i \hat{\xi}_i}{\sum_{i=1}^{k_i} w_i}, \text{ where, } w_i = \frac{1}{AVAR\left(\xi_i\right)}$$

3.4 RSS in Non-Parametric Inference

An alternative to non-parametric Mann-Whitney-Wilcoxon (MWW) estimation and testing procedures using ranked set empirical distribution function was discussed by Bohn and Wolfe (1992). They gave the tables of the null distribution for perfect ranking and presented asymptotic relative efficiency comparisons between SRS MWW procedures and their ranked set analogues. Bohn (1996) further examined other non-parametric tests like sign test and signed rank test for ranked set samples while discussing the similarities and differences in the properties of the RSS procedures. Ozturk and Wolfe (2000) have investigated the effects of different RSS protocols on the sign test statistic under different sampling protocols like sequential, mid-range and fixed sampling designs. They have shown that the introduction of any correlation structure in quantified observation leads to a reduction in the Pitman efficiency of the design. Koti and Babu (1996) computed the exact null distribution of the RSS sign test. They compared power of this test with the SRS sign test for some continuous symmetric distributions and demonstrated the superiority of RSS over SRS. Kvan and Samaniego (1994) studied the nonparametric MLE of CDF F and demonstrated its existence and uniqueness.

3.5 Location and Scale Parameters of Specified Distributions

Many environmental data such as species abundance, length of abundant periods of infectious diseases, distribution of mineral resources in the Earth's crust, average concentration of an air pollutant such as sulfur dioxide, carbon monoxide, nitrous oxide etc. often approximately follows highly skewed distributions like lognormal distribution (Aitchison and Brown, 1957; Crow and Shimizu, 1988; Lee, 1992; Johnson et al., 1994; Sachs, 1997; Larsen, 1969). With the intention of adding content to the above literature, Chandra and Tiwari (2012) estimated the location and scale parameters of lognormal distribution using RSS. It was shown that with the increase of sample size, the use of RSS results in much improved estimator compared to the use of a SRS (Table 5).

Table 5: RP of RSS against SRS for the estimation in location and scale parameters of lognormal distribution for sample size *n* = 2(1)10.

n	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RP (location parameter)	0.633	0.682	0.772	0.870	0.968	1.064	1.157	1.249	1.339
RP (scale parameter)	0.813	0.995	1.231	1.486	1.749	2.016	2.286	2.557	2.829

Fei et al. (1994) estimated the parameters of Weibull distribution by the use of RSS. Lam et al. (1994) used RSS to estimate the location and scale parameters of the exponential distribution. With a generalized geometric distribution, Bhoj and Ahsanullah (1996) showed that the RP of the estimates of the parameters of the generalized geometric distribution obtained by using the RSS procedure are higher than those of the ordered least square estimates. Adatia (2000) generalized the RSS and used it for estimating the location and scale parameters of half-logistic distribution. Abu-Dayyeh et al. (2004) studied the estimation of the logistic distribution parameters using SRS and RSS with some of its modifications as extreme RSS and median RSS. Tiwari et al. (2015) estimated the location and scale parameters of the normal distribution. The utility of RSS and Partial RSS (PRSS) over the SRS had been demonstrated with the help of numerical illustration and their generalized variances were also calculated. Chandra et al. (2016) also used PRSS in estimation of location and scale parameters of lognormal distribution.

4 Comparisons of RSS With Other Classical Designs

Yu and Lam (1997) suggested the regression RSS estimators in the case of double sampling. It was shown that the RSS regression estimator was more efficient than ordinary RSS and SRS naïve estimators unless the correlation between study and concomitant variable is low (less than 0.4). The Bayes risk of the Bayes estimator using RSS method was found less than the Bayes risk of the Bayes estimator using SRS Al-Saleh et al. (2000). The performance of RSS using maximum likelihood estimator for estimation of the correlation coefficients in a bivariate normal distribution in comparison with other classical designs was investigated by Stokes (1980a). The asymptotic variance of the maximum likelihood estimator of correlation coefficient based only on the extreme study variable and their concomitant variables was less than that from random samples.

While estimating rare plant or animal species, Adaptive Cluster Sampling (ACS) was found appropriate (Thompson, 1990). Chandra et al. (2011) used RSS in the first phase of ACS and found that the proposed design appears to perform better than the existing procedures of ACS. This design was appropriate for the situation in which the value of the characteristic under study on the sampled places is low or negligible but the neighborhoods of these places may have a few scattered pockets of the same. The various estimators like those based on only initial sample, based on initial intersection probabilities and improvement of the estimators using the Rao-Blackwell theorem were attempted under this design.

5 Important Case Studies Using Rss

Under the application part of RSS, following studies have been carried out by various researchers:

Evans (1967) applied the RSS to regeneration surveys in areas direct-seeded to longleaf pine. He noted that despite of means based on both of RSS and SRS methods being not significantly different, the variances of the mean based upon RSS were significantly different than based upon SRS. Martin et al. (1980) applied the RSS procedure for estimating shrub phytomass in Appalachian Oak forests. Cobby et al. (1985) conducted four experiments at Hurley (UK) during 1983 to investigate the performance of RSS relative to SRS for estimation of herbage mass in pure grass swards, and of herbage mass and clover content in mixed grass-clover swards. Johnson et al. (1993) applied RSS method to estimate the mean of forest, grassland and other vegetation resources. Nussbaum and Sinha (1997) successfully used RSS in estimating mean Reid vapor pressure. Mode et al. (1999) investigated the conditions under which RSS becomes a cost-effective sampling method for ecological and environmental field studies where the rough but cheap measurement has a cost. They found that RSS estimates of the mean pool area for 20 of 21 streams were more precise than estimates of the pool area that would be obtained by physically measuring pool areas selected using SRS. Al-Saleh and Al-Shrafat (2001) studied the performance of RSS in estimation milk yield based on 402 sheep. Al-Saleh and Al-Omari (2002) used the multistage RSS to estimate the average of Olive yields in a field in West of Jordan. Husby et al. (2005) investigated the use of RSS in estimating the mean and median of a population using the crop production dataset from the United State Department of Agriculture. They found that the gain in efficiency for mean estimation using RSS is better for symmetric distribution than for asymmetric distribution, and vice versa in the case of median estimation. Kowalczyk (2005) applied the RSS procedure in market and consumer surveys. Ganeslingam and Ganesh (2006) applied the RSS method to estimate the population mean and the ratio using a real data set on body measurement. The authors used the data of the weight and height of 507 individuals. Halls and Dell (1966) coined McIntyre's method as RSS and applied it for estimating the weights of browse and herbage in a pine-hardwood forest of east Texas, USA. Wang et al. (2009) used the RSS in fisheries research. Tiwari and Pandey (2013) applied RSS in environmental investigations for real data set. Chandra et al. (2018) attempted a study in response estimation of the developmental programs implemented by the government and non-government organizations in successive phases by the use of RSS. For a detailed study about the applications of RSS in real life situations, the readers may refer to Dong et al. (2012) and Chen et al. (2004).

6 Perfect Versus Imperfect Rankings.

Clearly, RSS procedures perform best in the absence of any errors in the ranking process (perfect rankings). In many practical applications, however, we are faced with errors in ranking due to the use of rough and inexpensive measures. This is known as imperfect ranking. Minimal uncertainty in the rankings may not cause an excessive increase in the variance, but if the ranking process is not very reliable, the precision of RSS estimators (particularly unbalanced ones) may be reduced. Al-Omari and Bouza (2014) discussed the impact of perfect and imperfect ranking in details. Bohn and Wolf (1992) also discussed MWW procedures under imperfect ranking proposing a model for the probabilities of imperfect judgment rankings based on the concept of expected spacing and have used this model to study the properties of tests based on the ranked set analogue of the MWW statistic. Greater precision in the RSS estimator requires more accurate rankings in each set David and Levine (1972), Barnett and Moore (1997), Stark and Wolfe (2002).

Much of the literature detailing the improvements from RSS estimation is based on the assumption of perfect rankings. This is the natural starting point for developing theory, but it is not realistic in practice. With imperfect rankings, the precision of the unbalanced RSS estimator may actually be worse than that for the SRS estimator. This suggests that unbalanced RSS should not be used without first considering the settings of the study. If the amount of error in the ranking procedure is expected to be only minor, then it is safe to use RSS, balanced or unbalanced, to improve estimation. When considerable uncertainty exists regarding the exactness of rankings in the procedure, balanced RSS may be the safest approach to take.

7 A New Initiative on Extreme Ranked Set Sampling (Erss)

ERSS was introduced by Samawi et al. (1996) to avoid the imperfect ranking as the ranking of extreme rank order statistics (only first and last order statistics in case of even set size and additional of ranking of mid order statistics for the case of odd set size) was found rather easy. This method provides unbiased estimator of population mean if the nature of distribution of the population is symmetric and is found to be more efficient than SRS. However, the estimator of this method is not unbiased when the distribution is skewed. In the method of Samawi et al. (1996) first and last order statistics are allocated equal times, *i.e.* $n/2$ each, when n is even and for the case when n is odd, they proposed a method that allocates $(n-1)/2$ times each to first and last order statistics and one time to the middle rank order statistics.

In what follows, we discuss a new initiative on ERSS by proposing an Unbalanced ERSS (UERSS) procedure and propose an improved estimator for the case of skewed distributions when exact distribution is unknown. The proposed method results in increased efficiency in comparison to that obtained through the method of Samawi et al. (1996) for skewed distribution. The proposed UERSS procedure is explained as under:

In UERSS, lowest and highest order statistics are measured m_1 and $m_2 = am_1 (a > 1)$ times, respectively, such that the sample size $n = (1+a)m_1$. Let n sets of units, each of size m_1, are drawn from the unknown infinite skewed population $F(x)$ with mean μ and variance σ^2. From the first set of m_1 units, lowest ranked unit is selected for measurement. From the second set of m_1 units, largest ranked unit is selected for measurement. From the third set of m_1 units, the smallest ranked unit is selected for measurement. This process is continued until the m_1 smallest and m_1 largest ranked units are selected for measurements. From the remaining $n-2m_1$ sets, $n-2m_1$ largest ranked unit, one from each set, is selected for measurements. The unordered nm_1 units can be written as

$$y_{11}, y_{12}, ..., y_{1m_1}$$
$$y_{21}, y_{22}, ..., y_{2m_1}$$
$$...$$
$$y_{m_1 1}, y_{m_1 2}, ..., y_{m_1 m_1}$$
$$...$$
$$y_{n1}, y_{n2}, ..., y_{nm_1}$$

where, the rows are showing n independent random sets each of size m_1 and y_{ij} denotes the value of j^{th} unit of the i^{th} set, $i = 1, 2, ..., n, j = 1, 2, ..., m_1$. Let $y_{i(1)}, y_{i(2)}, ..., y_{i(m_1)}$ denote the values of ordered statistics of i^{th} sample $(i = 1, 2, ..., n)$. Then the RSS shall be $y_{1(1)}, y_{2(2)}, ..., y_{m_1(m_1)}, y_{m_1+1(1)}, y_{m_1+2(2)} ...$; ERSS (Samawi et al., 1996) shall be (a) $y_{1(1)}, y_{2(m_1)}, y_{3(1)} ..., y_{n-1(1)}, y_{n(m_1)}$ (if n is even) and (b) $y_{1(1)}, y_{2(m_1)}, y_{3(1)}, ..., y_{n-1(m_1)}, \frac{1}{2}\left(y_{n(1)} + y_{n(m_1)}\right)$(if n is odd); UERSS shall be $y_{1(1)}, y_{2(m_1)}, y_{3(1)}, ..., y_{2m_1-1(1)}, y_{2m_1(m_1)}, y_{2m_1+1(m_1)}, ..., y_{n-1(m_1)}, y_{n(m_1)}$, irrespective of n is even or odd.

As usual, based upon SRS in which first observation is measured from each set (say), unbiased estimator of μ is

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_{i1} \text{ with } Var(\bar{Y}) = \frac{\sigma^2}{n}$$

The estimator of μ based upon UERSS is

$$\bar{Y}_{UERSS} = \frac{(T_1 + T_2)}{2},$$

where, $T_1 = \frac{1}{m_1}\sum_{i=1}^{m_1} y_{2i-1(1)}$ and $T_2 = \frac{1}{n-m_1}\left[\sum_{i=1}^{m_1} y_{2i(1)} + \sum_{i=2m_1+1}^{n} y_{i(m_1)}\right]$

Here it is to be noted that, $y_{1(1)}, y_{3(1)}, ..., y_{2m_1-1(1)}$ are independently and identically distributed with mean and variance $\mu_{(1)}$ and $\sigma^2_{(1)}$, respectively. Similarly the units $y_{2(m_1)}, y_{4(m_1)}, ..., y_{2m_1(m_1)}, y_{2m_1+1(m_1)}, ..., y_{n(m_1)}$ are independently and identically distributed with mean and variance $\mu_{(m_1)}$ and $\sigma^2_{(m_1)}$, respectively. Also the units in T_1 and T_2 are independent.

Therefore, we may easily have

$$E(\bar{Y}_{UERSS}) = \frac{(\mu_{(1)} + \mu_{(m_1)})}{2} \text{ with}$$

$$Var(\bar{Y}_{UERSS}) = \frac{1}{4}\left[\frac{\sigma^2_{(1)}}{m_1} + \frac{\sigma^2_{(m_1)}}{n-m_1}\right]$$

For symmetric distributions, we have $\mu_{(i)} = \mu_{(m_1-i+1)}$ and $\sigma^2_{(i)} = \sigma^2_{(m_1-i+1)}$, where, $\mu_{(i)}$ and $\sigma^2_{(i)}$ denotes the mean and variance of ith order statistics. Therefore, for the symmetric distribution,

$$E(\bar{Y}_{UERSS}) = 0 \text{ and } Var(\bar{Y}_{UERSS}) = n\frac{\sigma^2_{(1)}}{4m_1(n-m_1)}.$$

This means that $\bar{Y}_{UERSS}$ is an unbiased estimator of μ for standard symmetric distribution. For the non-standard symmetric distributions, $\bar{Y}_{UERSS}$ is biased with bias =

$$\frac{2}{m_1}\sum_{i=1}^{m_1/2}\mu_{(i)} \text{ and } \frac{2}{m_1}\sum_{i=1}^{(m_1-1)/2}\mu_{(i)} + \frac{\mu_{\left(\frac{m_1+1}{2}\right)}}{m_1}.$$

It is to be noted that if $m_1 = n/2$ (for n is even), then $\bar{Y}_{UERSS}$ becomes $\bar{Y}_{ERSS}$ (estimator based upon the ordinary ERSS).

The unbiased estimator of μ based upon RSS is given by

$$\bar{Y}_{RSS} = \frac{y_{1(1)} + y_{2(2)} + ... + y_{m_1(m_1)} + y_{m_1+1(1)} + y_{m_1+2(2)} + ...upto\ n\ terms}{n},$$

with $Var\left(\bar{Y}_{RSS}\right)=\dfrac{1}{n^2}\sum_{l=1}^{m_1} a_l \sigma_{(l)}^2$, where, a_l is number of times l^{th} order statistic taken for measurement $\left(l=1,2,...,m_1\right)$ and $\sum_{l=1}^{m_1} a_l = n$.

Now, RP of UERSS with respect to ERSS is given by (for the sake of convenience of calculation we took n as even)

$$RP_{UERSS}=\frac{MSE\left(\bar{Y}_{ERSS}\right)}{MSE\left(\bar{Y}_{UERSS}\right)}=2\frac{\dfrac{\sigma_{(1)}^2+\sigma_{(m_1)}^2}{n}-\left(\mu-\dfrac{\left(\mu_{(1)}+\mu_{(m_1)}\right)}{2}\right)^2}{\dfrac{\sigma_{(1)}^2}{m_1}+\dfrac{\sigma_{(m_1)}^2}{am_1}-\left(\mu-\dfrac{\left(\mu_{(1)}+\mu_{(m_1)}\right)}{2}\right)^2}$$

Our aim is to find the optimum value of a such that RP_{UERSS} is maximum.

To maximize RP_{UERSS} means equivalently maximize $\dfrac{\dfrac{\sigma_{(1)}^2+\sigma_{(m_1)}^2}{n}}{\dfrac{\sigma_{(1)}^2}{m_1}+\dfrac{\sigma_{(m_1)}^2}{am_1}}=\dfrac{\sigma_{(1)}^2+\sigma_{(m_1)}^2}{(1+a)\left(\sigma_{(1)}^2+\dfrac{\sigma_{(m_1)}^2}{a}\right)}$.

Or, in other words, one needs to minimize $S=(1+a)\left(\sigma_{(1)}^2+\dfrac{\sigma_{(m_1)}^2}{a}\right)$.

This can be obtained by $\dfrac{\partial S}{\partial a}=0$. This results in,

$$a=\frac{\sigma_{(m_1)}}{\sigma_{(1)}}.$$

Similarly for odd n, the similar results may be obtained.

Again, $\dfrac{\partial^2 S}{\partial a^2}=2\dfrac{\sigma_{(m_1)}^2}{a^3}$, which means $\dfrac{\partial^2 S}{\partial a^2}$ is positive. Therefore, at $a=\dfrac{\sigma_{(m_1)}}{\sigma_{(1)}}$, S is minimum with the value $2\dfrac{\sigma_{(1)}^3}{\sigma_{(m_1)}}$.

The value of RP_{UERSS} at $a=\dfrac{\sigma_{(m_1)}}{\sigma_{(1)}}$ shall be

$$2\frac{\dfrac{\sigma_{(1)}\left(\sigma_{(1)}^2+\sigma_{(m_1)}^2\right)}{\left(\sigma_{(1)}+\sigma_{(m_1)}\right)m_1}-\left(\mu-\dfrac{\left(\mu_{(1)}+\mu_{(m_1)}\right)}{2}\right)^2}{\dfrac{\sigma_{(1)}\left(\sigma_{(1)}+\sigma_{(m_1)}\right)}{m_1}-\left(\mu-\dfrac{\left(\mu_{(1)}+\mu_{(m_1)}\right)}{2}\right)^2}.$$

From above one may show that $RP_{UERSS}>1$, if $2\dfrac{\sigma_{(1)}\left(\sigma_{(1)}^2+\sigma_{(m_1)}^2\right)}{\left(\sigma_{(1)}+\sigma_{(m_1)}\right)m_1}>\dfrac{\sigma_{(1)}\left(\sigma_{(1)}+\sigma_{(m_1)}\right)}{m_1}$ holds,

or, $\dfrac{\sigma_{(m_1)}}{\sigma_{(1)}}+\dfrac{\sigma_{(1)}}{\sigma_{(m_1)}}>2$.

The values of RPs for four skewed distributions (Pareto (2.5), Lognormal (0, 1), Weibull (0.5) and Gamma (0.5)) are shown in Table 6 to Table 9 respectively.

Table 6: Bias and RPs of UERSS and ERSS for Pareto (2.5) distribution with k=2(1)8

Set size $(k = m_1)$	2	3	4	5	6	7	8
Bias(UERSS)	0	-0.1121	-0.2243	-0.3284	-0.4243	-0.5132	-0.5960
Bias(ERSS)	0	0.1042	0.2243	0.3246	0.4243	0.5110	0.5960
MSE(UERSS)	0.0937	0.0507	0.0719	0.1221	0.1903	0.2711	0.3614
MSE(ERSS)	0.1425	0.0756	0.0898	0.1320	0.1997	0.2761	0.3673
RP(UERSS)	1.6499	1.0334	0.3445	0.1148	0.0465	0.0220	0.0120
RP(ERSS)	1.0847	0.6927	0.2758	0.1062	0.0443	0.0216	0.0118

Table 7: Bias and RPs of UERSS and ERSS for lognormal (0, 1) distribution with k=2(1)8

Set size $(k = m_1)$	2	3	4	5	6	7	8
Bias(UERSS)	0	-0.1984	-0.3969	-0.5790	-0.7450	-0.8968	-1.0367
Bias(ERSS)	0	0.1844	0.3969	0.5723	0.7450	0.8930	1.0367
MSE(UERSS)	0.2152	0.1152	0.1964	0.3591	0.5712	0.8159	1.0837
MSE(ERSS)	0.2736	0.1445	0.2198	0.3671	0.5827	0.8176	1.0906
RP(UERSS)	1.5092	0.9550	0.2651	0.0821	0.0326	0.0154	0.0084
RP(ERSS)	1.1872	0.7614	0.2369	0.0803	0.0319	0.0154	0.0083

Table 8: Bias and RPs of UERSS and ERSS for Weibull (0.5) distribution with k=2(1)8

Set size $(k = m_1)$	2	3	4	5	6	7	8
Bias(UERSS)	0	-0.4722	-0.9444	-1.3786	-1.7747	-2.1378	-2.4727
Bias(ERSS)	0	0.4388	0.9444	1.3625	1.7747	2.1286	2.4727
MSE(UERSS)	0.8478	0.5319	1.0569	2.0047	3.2218	4.6227	6.1551
MSE(ERSS)	1.2343	0.7202	1.1982	2.0539	3.2893	4.6334	6.1943
RP(UERSS)	1.6403	0.8858	0.2109	0.0629	0.0247	0.0116	0.0063
RP(ERSS)	1.1268	0.6541	0.1861	0.0614	0.0242	0.0116	0.0063

Table 9: Bias and RPs of UERSS and ERSS for Gamma (0.5) distribution with k=2(1)8

Set size $(k = m_1)$	2	3	4	5	6	7	8
Bias(UERSS)	0.0000	-0.0741	-0.1477	-0.2142	-0.2733	-0.3261	-0.3736
Bias(ERSS)	0.0000	0.0689	0.1477	0.2117	0.2733	0.3246	0.3736
MSE(UERSS)	0.0253	0.0129	0.0252	0.0478	0.0759	0.1071	0.1401
MSE(ERSS)	0.0279	0.0148	0.0270	0.0479	0.0768	0.1068	0.1406
RP(UERSS)	1.3757	0.9108	0.2214	0.0661	0.0263	0.0125	0.0069
RP(ERSS)	1.2448	0.7971	0.2064	0.0659	0.0260	0.0126	0.0069

It is evident from tables 6-9 that for all the four skewed distributions considered by us, the proposed UERSS method performs better than the method proposed by Samawi et al. (1996) in terms of RPs for small set sizes.

8 Conclusion and Discussion

Since the introduction of RSS by McIntyre (1952), significant theoretical work has been carried out in this area. Many suitable modifications of RSS are also available in the literature. Important areas from parametric setup to non parametric setup, estimation of mean, proportion, location, variances, scale and shape parameters of the distributions, correlation coefficients, distribution functions, density function etc. have been studied using RSS. Now, there is a time to implement RSS in the surveys in agriculture, forestry, environmental, ecological, medical, paramedical studies etc. and achieve the advantages of RSS over the other conventional sampling methods. This paper provides major advancement of theoretical work on RSS along with the important case studies. As a new initiative, UERSS has been proposed in Section 7 and the efficiency of the proposed estimator has been compared with the ERSS for four skewed distributions. The estimators obtained under ERSS for skewed distributions (Samawi et al., 1996) were biased but at the same time the possibility of ranking error was reduced. While using the UERSS, the small set size is recommended to retain the increased efficiency. The application of UERSS may be helpful in many real life situations with the advantage of avoiding improper ranking and increase in efficiency.

Acknowledgement

The authors are thankful to the referee for his constructive comments on the paper, which has led to considerable improvement in presentation of this manuscript.

References

Abu-Dayyeh, W., Al-Subuh, S.A. and Muttlak, H.A. (2004). Logistic parameters estimation using simple random sampling and ranked set sampling data. *Applied Mathematics and Computation*, **150**, 543-554.

Abu-Dayyeh, W., Samawi, H.M. and Bani-Hani, L.A. (2002). On distribution function estimation using double ranked set samples with application. *Journal of Modern Applied Statistical Method*, **1(2)**, 443-451.

Adatia, A. (2000). Estimation of parameters of the half-logistic distribution using generalized ranked set sampling. *Computational Statistics and Data Analysis*, **33**, 1-13.

Aitchison J. and Brown J.A.C. (1957). *The Log-normal Distribution*. Cambridge (UK): Cambridge University Press.

Al-Hadhrami, S.A. and Al-Omari, A.I. (2006). Bayesian inference on the variance of normal distribution using moving extremes ranked set sampling. *Journal of Modern Applied Statistical Methods*, **8(1)**, 273-281.

Al-Omari, A.I. (2012). Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling. *Statistics and Probability Letters*, **82**, 1883-1890.

Al-Omari, A.I. and Bouza, C.N. (2014). Review of ranked set sampling: modifications and applications. *Revista Investigación Operacional*, **35(3)**, 215-240.

Al-Omari, A. I., Jemain, A. A. and Ibrahim, K. (2009). New ratio estimators of the mean using simple random sampling and ranked set sampling methods. *Revista Investigación Operacional*, **3.0(2)**, 97-108.

Al-Saleh, F. M., Al-Shrafat, K. and Muttlak, H. A. (2000). Bayesian estimation using Ranked Set Sampling. *Biometrical Journal*, **42**, 3, 1-12.

Al-Saleh, M.F. and Al-Hadhrami, S. (2003). Parametric estimation for the location parameter for symmetric distributions using moving extremes ranked set sampling with application to tree data. *Environmetrics*, **14**, 651-664.

Al-Saleh, M.F. and Al-Kadiri, M. (2000). Double ranked set sampling. *Statistics and Probability Letters*, **48**, 205–212.

Al-Saleh, M.F. and Al-Omari, A.I. (2002). Multistage ranked set sampling. *Journal of Statistical Planning and Inference*, **102**, 273-286.

Al-Saleh, M.F. and Al-Shrafat, K. (2001). Estimation of milk yield using ranked set sampling. *Envirometrics*, **12**, 395-399.

Bani-Mustafa, A., Al-Nasser, A.D. and Aslam, M. (2011). Folded ranked set sampling for asymmetric distributions. *Communications of the Korean Statistical Society*, **18(1)**, 147-153.

Barabesi, L. and Fattorini, L. (2002). Kernel estimators of probability density functions by ranked set sampling. *Computational Statistics Part A, Theory & Methods*, **31(4)**, 597-610.

Baraneso, L. and Fattorini, L. (2002). Kernel estimators for probability density functions by ranked-set sampling. *Communication in Statistical Theory and Methods*, **31(4)**, 597–610.

Barnett, V. and Moore, K. (1997). Best linear unbiased estimates in ranked-set sampling with particular reference to imperfect ordering. *Journal of Applied Statistic*, **24**, 697-710.

Bhoj, D.S. and Ahsanullah, M. (1996). Estimation of parameters of the generalized geometric distribution using ranked set sampling. *Biometrics*, **52(2)**, 685-694.

Biswas, A., Ahmad, T. and Rai, A. (2013). Variance estimation using Jackknife method in ranked set sampling under finite population framework. *Journal of the Indian Society Agricultural Statistics*, **67(3)**, 345-353.

Bohn, L.L. (1996). *A Ranked-Set Sample Signed-Rank Statistic*. Department of Statistics, Technical Report Number 426. The University of Florida. Gainevesville, FL.

Bohn, L.L. and Wolfe, D.A. (1992). Nonparametric two-sample procedures for ranked set samples data. *Journal of the American Statistical Association*, **87(418)**, 552-561.

Bohn, L.L. and Wolfe, D.A. (1994). The effect of imperfect judgment rankings on properties of procedures based on the ranked set samples analog of the Mann-Whitney-Wilcoxon statistics. *Journal of the American Statistical Association*, **89(425)**, 168-176.

Bouza, C.N. (2008). Estimation of the population mean with Missing observations using product type estimators. *Revista Investigacion Operacional*, **29(3)**, 207-223.

Chandra, G., Pandey, R., Bhoj, D. S., Nautiyal, R., Ashraf, J. and Verma, M. R. (2018). Ranked set sampling approach for response estimation of developmental programs with linear impacts under successive phases, *Pakistan Journal of Statistics*, **34(2)**, 163-176.

Chandra, G. and Tiwari, N. (2012). Estimation of location and scale parameters of lognormal distribution using ranked set sampling. *Journal of Statistics and Applications*, **7(3-4)**, 139-152.

Chandra, G., Tiwari, N. and Chandra, H. (2011). Adaptive cluster sampling based on ranked sets, *Metodoloski Zvezki*, **8(1)**, 39-55.

Chandra, G., Tiwari, N. and Nautiyal, R. (2015): Near optimal allocation models for symmetric distributions in ranked set sampling. In Chandra, G., Nautiyal, R., Chandra, H., Roychoudury, N. and Mohammad, N. (Eds.), *Statistics in Forestry: Methods and Applications*. Coimbatore: *Bonfring Publication*, 85-90.

Chandra, G., Tiwari, N., Nautiyal, R. and Gupta, D. S. (2016). On partial ranked set sampling in parameter estimation of lognormal distribution. *International Journal of Statistical and Agricultural Sciences*, **12(2)**, 321-326.

Chen Z. (2001). The optimal ranked set sampling scheme for inference on population quantiles. *Statistica Sinica*, **11**, 23-37.

Chen, Z., Bai, Z. and Sinha, B. (2004). *Ranked Set sampling: Theory and Applications*. Springer Verlag. New York.

Chen, M. and Lim, J. (2011). Estimating variances of strata in ranked set sampling. *Journal of Statistical Planning and Inference*, **141,** 2513-2518.

Chen H., Stasny, E.A. and Wolfe D.A. (2006). Unbalanced ranked set sampling for estimating a population proportion. *Biometrics*, **62**, 150-158

Cobby, J.M., Ridout, M.S., Bassett, P.J. and Large, R.V. (1985). An investigation into the use of ranked set sampling on grass and grass-clover swards. *Grass and Forage Science*, **40**, 257-63.

Crow, E.L, Shimizu K. (Eds.) (1988). *Log-normal Distributions: Theory and Application*. New York, Dekker.

David, H.A. and Levine, D.N. (1972). Ranked set sampling in the presence of judgment error. *Biometrics*, **28**, 553–555.

Dell, T.R. (1969). *The Theory of Some Applications of Ranked Set Sampling*. Ph.D. thesis, University of Georgia, Athens, GA.

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

Dong, X.F., Cui, L.R. and Liu, F.Y. (2012). A further study on reliable life estimation under ranked set sampling. *Communications in Statistics – Theory and Methods*, **21**, 3888-3902.

Evans, M.J. (1967). *Application of Ranked Set Sampling to Regeneration, Surveys in Areas Direct-Seeded to Long Leaf Pine*. Master Thesis, School for Forestry and Wild-life Management, Louisiana State University, Baton Rouge, Louisiana.

Fei, H., Sinha, B. and Wu, Z. (1994). Estimation of parameters in two-parameter Weibull and extreme-value distributions using ranked set sampling. *Journal of Statistical Research*, **28**, 149-161.

Frey, J. (2014). Bootstrap confidence bands for the CDF using ranked-set sampling. *Journal of the Korean Statistical Society*, **43(3)**, 453-461.

Ganeslingam, S. and Ganesh, S. (2006). Ranked set sampling versus simple random sampling in the estimation of the mean and the ratio. *Journal of Statistics and Management Systems*, **2**, 459-472.

Gibbons, J. D. and Chakraborty, S. (2010). *Non-Parametric Statistical Inference*. 5th Edition, CRC Press.

Halls, L. S. and Dell, T. R. (1966). Trial of ranked set sampling for forage yields. *Forest Science*, **12(1)**, 22-26.

Huang, J. (1972). Asymptotic properties of the NPMLE of a distribution function based on ranked set samples. *The Annals of Statistics*, **25(3)**, 1036-1049.

Husby, C.E., Stansy, E.A. and Wolfe, D.A. (2005). An application of ranked set sampling for mean and median estimation using USDA crop production data. *Journal of Agricultural, Biological, and Environmental Statistics*, **10(3)**, 354-373.

Johnson, G.D., Paul, G.P. and Sinha, A.K. (1993). Ranked set sampling for vegetation research. *Abstracta Botanica*, **17**, 87-102.

Johnson, N.L., Kotz, S. and Balakrishnan N. (1994). *Continuous Univariate Distributions*. New York, Wiley.

Kadilar, C., Unyazici, Y. and Cingi, H. (2009). Ratio estimator for the population mean using ranked set sampling. *Statistical Papers*, **50**, 301-309.

Kaur, A., Patil, G.P. and Taillie, C. (1994). *Unequal Allocation Models for Ranked Set Sampling With Skew Distributions*. Technical Report 94-0930, Center for Statistical Ecology and Environmental Statistics, Pennsylvania State University, University Park.

Kaur, A., Patil, G.P. and Taillie, C. (1997). Unequal allocation models for ranked set sampling with skew distributions. *Biometrics*, **53**, 123-130.

Kaur, A., Patil, G.P. and Taillie, C. (2000). Optimal allocation for symmetric distributions in ranked sampling. *Annals of the Institute of Statistical Mathematics*, **52(2)**, 239-254.

Koti, K.M. and Babu, G.J. (1996). Sign test for ranked-set sampling. *Communications in Statistics-Theory and Methods*, **25**, 1617-1630.

Kowalczyk, B. (2005). Alternative sampling designs some applications of qualitative data in survey sampling. *Statistics in Transition*, **7(2)**, 427-443.

Kvam, P.H. and Samaniego, F.J. (1994). Nonparametric maximum likelihood estimation based on ranked set samples. *Journal of the American Statistical Association*, **89**, 526-537.

Lam, K., Sinha, B., and Wu, Z. (1994). Estimation of parameters in a two-parameter exponential distribution using ranked set sampling. *Annals of the Institute of Statistical Mathematics*, **46(4)**, 723-736.

Lam, K.F., Philip, L.H. and Lee, C.F. (2002). Kernel method for the estimation of the distribution function and the mean with auxiliary information in ranked set sampling. *Environmetrics*, **13**, 397-406.

Larsen, R.L. (1969). A New Mathematical model of air pollution concentration averaging time and frequency. *Journal of the Air Pollution Control Association*, **19**, 24-30.

Lee, E.T. (1992). *Statistical Methods for Survival Data Analysis*. New York, Wiley.

Li, D., Sinha, B.K. and Perron, F. (1999). Random selection in ranked set sampling and its applications. *Journal of Statistical Planning and Inference*, **76**, 185-201.

Lim, J., Chen, M., Park, S., Wang, X. and Stokes, L. (2014). Kernel density estimator from ranked set samples. *Communications in Statistics - Theory and Methods*, **43**, 2156–2168.

MacEachern, S.N., Ozturk, O., Wolfe, D.A. and Stark, G.V. (2002). A new ranked set sample estimator of variance. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **64**, 177-188.

Mandowara, V.L. and Mehta, N. (2014). Modified ratio estimators using stratified ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, **43(3)**, 461-471.

Martin, W., Sharik, T., Oderwald, R. and Smith, D. (1980). *Evaluation of Ranked Set Sampling for Estimating Shrub Phytomass in Appalachian Oak forests*. Publication Number FWS-4-80, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research*, **3**, 385-390.

Mode, N.A., Conquest, L.L. and Marker, D.M. (1999). Ranked set sampling for ecological research: accounting for the total costs of sampling. *Environmetrics*, **10**, 179-194.

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

Muttlak, H.A. and Abu-Dayyeh, W.A. (2004). Weighted modified ranked set sampling. *Applied Mathematics and Computation*, **151**, 645-657.

Nussbaum, B.D. and Sinha, B.K. (1997). Cost effective gasoline sampling using ranked set sampling. Proceedings of the Section on Statistics and the Environment. *American Statistical Association*, Alexandria, VA, 83-87.

Ozturk, O. and Wolfe, D.A. (2000). Alternative ranked set sampling protocols for the sign test. *Statistics and Probability Letters*, **47**, 15-23.

Patil, G. P. (2002). *Ranked Set Sampling*. In: El-Shaarawi, A.H., Piergorsch, W.W. (Eds), *Encyclopedia of Environmetrics*. John Wiley and Sons, New York.

Perron, F. and Sinha, B.K. (2004). Estimation of variance based on a ranked set sample. *Journal of Statistical Planning and Inference*, **120**, 21-28.

Sachs L. (1997). *Angewandte Statistik. Anwendung Statistischer Methoden*. Heidelberg (Germany), Springer.

Samawi, H.M. (2002). On double extreme ranked set sample with application to regression estimator. *Metron*, **2**, 53-66.

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

Samawi, H.M. and Muttlak, H.A. (2001). On the ratio estimation using median ranked set sampling. *Journal of Applied Statistical Science*, **10(2)**, 89-98.

Samawi, H.M. and Saeid, L.J. (2004). Stratified extreme ranked set sample with application to ratio estimators. *Journal of Modern Applied Statistical Methods*, **3(1)**, 117-133.

Samawi, H.M. and Tawalbeh, E.M. (2002). Double median ranked set sampling: Comparison to other double ranked set samples for mean and ratio estimators. *Journal of Modern Applied Statistical Methods*, **1(2)**, 428-442.

Sengupta, S. and Mukhuti, S. (2006). Unbiased variance estimation in a simple exponential population using ranked set samples. *Journal of Statistical Planning and Inference*, **139**, 1526-1553.

Shen, W.H. (1994). On the estimation of lognormal mean using a ranked set sample. *Sankhya: The Indian Journal of Statistics*, **56(3)**, 323-333.

Shirahata, S. (1982). An extension of the ranked set sampling theory. *Journal of Statistical Planning and Inference*, **6**, 65-72

Singh, H.P., Tailor, R. and Singh, S. (2014). General procedure for estimating the population mean using ranked set sampling. *Journal of Statistical Computation and Simulation*, **84(5)**, 931-945.

Stark, G.V. and Wolfe, D.A. (2002). Evaluating ranked-set sampling estimators with imperfect rankings. *Journal of Statistical Studies*, 77-103.

Stokes, S.L. (1980). Estimation of variance using judgment ordered ranked-set samples. *Biometrics*, **36**, 35-42.

Stokes, S.L. (1980a). Inferences on the correlation coefficient in bivariate normal populations from ranked set samples. *Journal of the American Statistical Association*, **75**, 989-95.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics* **20**, 1-31.

Thompson, S.K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association* **85**, 1050-1059.

Tiwari, N. and Chandra, G. (2011). A systematic procedure for unequal allocation for skewed Distributions in Ranked Set Sampling. *Journal of the Indian Society of Agricultural Statistics*, **65(3)**, 331-338.

Tiwari, N. and Pandey, G.S. (2013). Application of ranked set sampling design in environmental investigations for real data set. *Thailand Statistician*, **11(2)**, 173-184.

Tiwari, N., Pandey, G. S. and Chandra, G. (2015): *Estimation of Location and Scale Parameters of Normal Distribution Using Ranked Set Sampling*. In Chandra, G., Nautiyal, R., Chandra, H., Roychoudury, N. and Mohammad, N. (Eds.), Statistics in Forestry: Methods and Applications. Coimbatore: *Bonfringe Publication*, 78-84.

Wang, Y.G., Ye, Y. and Milton, D.A. (2009). Efficient designs for sampling and subsampling in fisheries research based on ranked sets. *Journal of Marine Science*, **66**, 928-934.

Wolfe, D.A. (2004). Ranked set sampling: an approach to more efficient data collection. *Statistical Science*, **19(4)**, 636-643.

Yanagawa, T. and Chen, S. (1980). The MG-procedure in ranked set sampling. *Journal of Statistical Planning and Inference*, **4**, 33-44.

Yu, P.L.H. and Lam, K. (1997). Regression estimator in ranked set sampling. *Biometrics*, **53**, 1070-1080.

Zhu, M. and Wang, Y. (2005). Quantile estimation from ranked set sampling data. *Sankhya, The Indian Journal of Statistics*, **67(2)**, 295-304.