TWMS J. App. and Eng. Math. V.10, N.4, 2020, pp. 1009-1022

SKEW-NORMAL REVISITED VIA SOME RANKED SET SAMPLING SCHEMES

H. ESFANDYARIFAR[1], M. SALEHI[1], §

Abstract. Ranked set sampling (RSS) was first introduced by McIntyre (1952) as a competitor of simple random sampling (SRS), the most common tool in the statistical methods. When the sample size is not large enough, it may be difficult to obtain a representative subset from the population based on SRS, but RSS and its generalizations overcome to this shortcoming. These sampling schemes usually work based on judgment ranking of the sample units. The present paper investigates the performance of the mentioned schemes when the underlying distribution is the well-known Azzalini's skew-normal (SN) distribution. It also answers to an important question, that is, which kind of rank-based sampling methods is appropriate when the parent distribution is SN? To this end, the maximum (penalized) likelihood estimation as well as the method of moments are applied as the estimation approaches of the skewness parameter of SN distribution. Comparison of the estimators is carried out via their mean squared error and the Pitman measure of closeness criteria through a simulation study. Results show that the suggested scheme is highly dependent on the sign of the skewness parameter.

Keywords: Maximum penalized estimation, Median ranked set sampling, Modified ranked set sampling, Ranked set sampling, Skew-Normal distribution.

AMS Subject Classification: 62C10, 62F07

1. Introduction

Different types of problem arise when we seek to collect data on agriculture, forestry, environment, reliability studies and so on. It has been observed that in most of these situations, actual measurement of the sample units is expensive or time-consuming but (judgment) ranking of them is cost-effective and not so much difficult. The RSS can provide an efficient basis for estimating parameters of these kind of variables. For example, in forestry, it is easy to judge approximately by visual inspection about which of the several trees contains the largest volume of wood, which one is the next largest, and so on, whereas it is much more expensive to actually measure the amount of wood in each of the trees. Similar situations may arise in environmental applications, where we want to assess the status of a hazardous waste site. In these cases, our knowledge on physical characteristics of sites or photos and records will enable us to rank the sites in terms of high to low

[1] Imam Ali University, Department of Mathematics, Iran.
e-mail: hassan_esfandyarifar@yahoo.com; ORCID: https://orcid.org/0000-0002-2481-0898.
e-mail: salehi.sms@neyshabur.ac.ir; ORCID: https://orcid.org/0000-0002-0620-0340.

§ Manuscript received: December 30, 2018; accepted: April 24, 2020.

levels of toxic pollution and thus would limit the number of expensive samples necessary to assess the status of the hazardous waste site [see Chen et al. (2003) for more examples]. Thus, in such situations, RSS has been utilized as an alternative to the commonly used SRS in statistical inference during the last few decades.

RSS was first introduced by McIntyre (1952) and then was discussed in detail by Takahasi and Wakimoto (1968). Dell and Clutter (1972) showed that RSS is more efficient than SRS even with an error in ranking (imperfect ranking). Since then, several authors have shown interest in RSS in different fields of study. For instance, they introduced some generalized/simplified versions of RSS, namely, Modified RSS [see, Stokes (1980)], Median RSS (MedRSS) [see, Muttlak (1997)], unified ranked sampling [see, Matthews and Wolfe (2016)] and etc. For more details on the RSS and the relevant works, one can refer to the review papers Wolfe (2012), Al-Omari and Bouza (2014) and Sevinc et al. (2019). Existence of a wide variety of rank-based sampling designs led some researches to the parametric and non-parametric comparitive studies. Salehi et al. (2015), Dey et al. (2017) and Esemen and Gurler (2018) are just few instances of this class. In this paper, we investigate SRS, RSS, Minimum RSS (MinRSS) and Maximum RSS (MaxRSS) as special cases of modified RSS, and MedRSS via maximum likelihood estimation (MLE) and method of moments estimation (MME) when the parent distribution is the well known Azzalini's skew-normal. A random variable Z is said to have a standard skew-normal distribution with the skewness parameter λ ($\in \mathbb{R}$), denoted by $Z \sim SN(\lambda)$, if its density is as follows [Azzalini (1985)]

$$\phi(z;\lambda) = 2\phi(z)\Phi(\lambda z), \qquad z \in \mathbb{R}, \tag{1}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ stands for the density and the distribution function of the standard normal distribution, respectively. When $\lambda = 0$, the density (1) reduces to that of standard normal distribution, while it is right-skewed for a positive λ and left-skewed for a negative λ. Thus, it will be interesting to know, whether it is customary to alternatively utilize the Modified RSS instead of the other mentioned sampling schemes when λ is moderately far from zero (i.e., skewness appears in the density)? Hence, the rest of the paper is organized as follows. In Section 2, we describe the rank-based sampling schemes used as competitors of SRS. Section 3 presents the estimation tools of the skewness parameter λ. A simulation study is carried out in Section 4. Finally, Section 5 concludes.

2. Models description

This section provides a concise description of the procedure of constructing the sampling plans considered in the sequel.

RSS. When the sample size is not enough large, it may be difficult to obtain a representative subset from the population via SRS. To overcome this deficiency, RSS was proposed in order to attempt to provide a sample that is more probably to extend the range of the population of interest than SRS. This is performed by employing the information the ranking of a (some) concomitant covariate(s). An RSS of size $n = rm$ is obtained as the following procedure. An SRS of size m is derived from the population. None of the units are measured at this step, but are instead ranked with respect to the variable of interest. An important point is that the ranking of the units is performed by judgment without taking an actual measurement. In this regard, visual inspections, expert opinions or high-correlated concomitant variables may be employed for judgment ranking. Then, the unit which seems to be the smallest one among the m units is measured. This is the first observation of the RSS, denoted by $X_{(1)1,1}$. Another SRS of size m is then obtained and also judgment ranked as already. At this step, the unit judgment ranked which sounds to be the second smallest one is measured, and this will be the second observation of the RSS, denoted by $X_{(2)2,1}$. This procedure continues until the largest (mth) judgement rank is measured and finalizes the last observation of the RSS, denoted by $X_{(m)m,1}$. Finally, the sample $\left(X_{(1)1,1}, X_{(2)2,1}, \cdots, X_{(m)m,1}\right)$ is called a one-cycle RSS of size m. Repeating the above process r times more, gives an r-cycle RSS containing $n = mr$ units, denoted by $\mathbf{X}_{RSS} = \left\{X_{(i)i,j},\ i = 1, \cdots, m,\ j = 1, \cdots, r\right\}$, where $X_{(i)i,j}$ is the ith judgement ranked unit from the ith SRS in the jth cycle.

Modified RSS. As mentioned earlier, if the needed conditions of running the RSS are provided, it is more efficient than SRS in most of scenarios. But, when the set size m is not small, the ranking error appears. On the other hand, inspecting the extreme values is not so difficult. Hence, Stokes (1980) proposed Modified RSS, denoted by MinRSS and MaxRSS in this paper. The procedure of obtaining an r-cycle MinRSS (MaxRSS) with set size m is almost similar to that of RSS, but in all steps the smallest (largest) judgment ranked unit are measured. Thus, the ranking error of this plan will be negligible. Let us denote a MinRSS and MaxRSS of size $n = rm$ with $\mathbf{X}_{MinRSS} = \left\{X_{(1)i,j},\ i = 1, \cdots, m,\ j = 1, \cdots, r\right\}$ and $\mathbf{X}_{MaxRSS} = \left\{X_{(m)i,j},\ i = 1, \cdots, m,\ j = 1, \cdots, r\right\}$, respectively.

Median RSS. The following algorithm may be applied in order to obtain an r-cycle MedRSS with set size m: if m is odd, then the sample is selected by measuring the judgment median of each SRS. For the even set sizes, suppose that we obtain $\frac{m}{2} + \frac{m}{2}$ SRS's of size m. Then, the largest judgment ranked units are measured from the first half of the SRS's, while the smallest jadgment ranked units are measured from the second portion of the SRS's. By repeating the above process r times, the desired sample size $n = mr$ will be achieved. More specifically, if $\mathbf{X}_{MedRSS}$ stands for the mentioned sample, then, it will be of the form

$$\left\{X_{(\frac{m+1}{2})i,j},\ i = 1, \cdots, m,\ j = 1, \cdots, r\right\},$$

when m is odd, and

$$\left\{X_{(\frac{m}{2})i,j},\ i = 1, \cdots, \frac{m}{2},\ j = 1, \cdots, r\right\} \cup \left\{X_{(\frac{m}{2}+1)i,j},\ i = \frac{m}{2} + 1, \cdots, m,\ j = 1, \cdots, r\right\},$$

when m is even.

3. Estimation approaches

Suppose that $\mathbf{x}_{SRS} = (x_1, \cdots, x_n)$ is an SRS of size n from the SN distribution with density given by (1). Then, the log-likelihood function of the parameter λ given $\mathbf{x}_{SRS}$ is readily obtained as

$$\ell_{SRS}(\lambda;\ \mathbf{x}_{SRS}) = \text{const.} + \sum_{i=1}^{n} \log \Phi(\lambda x_i), \tag{2}$$

where 'const.' is free of λ. Here, there is a problem in maximizing the above log-likelihood function when the observations have all the same sign, i.e. $x_i > 0$ or $x_i < 0$, for $i = 1, \cdots, n$. Since, in the former, the function $\ell_{SRS}(\cdot)$ will be an increasing function of λ while in the latter it will be decreasing. As a result, MLE of λ, say $\hat{\lambda}_{SRS}$, takes the extreme values $+\infty$ and $-\infty$, respectively, which may not be the actual values of λ in practice. The mentioned event has the non-zero probability [see, Azzalini and Arellano-Valle, 2013]

$$\pi_n(\lambda) = \left(\frac{1}{2} - \frac{\arctan \lambda}{\pi}\right)^n + \left(\frac{1}{2} + \frac{\arctan \lambda}{\pi}\right)^n, \tag{3}$$

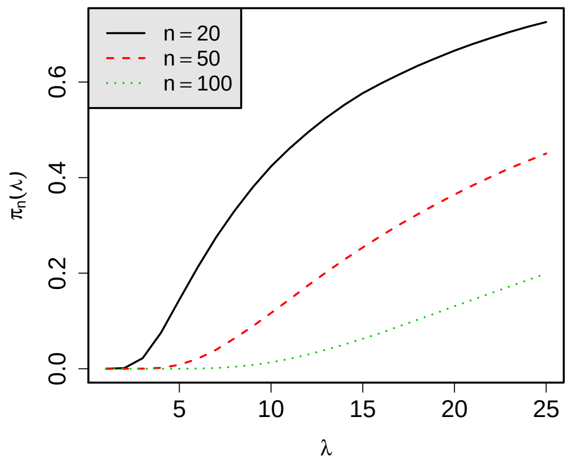


FIGURE 1. The probability of a divergent $\hat{\lambda}_{SRS}$ based on an SRS of size n from $SN(\lambda)$.

which increases for small n or large values of $|\lambda|$ (see Figure 1). A solution for this problem is to use the penalization of the log-likelihood function which is a tool to correct some undesirable behaviour of the common MLE. In this approach, the problematic log-likelihood function is subtracted by a non-negative function, say $Q(\cdot)$, which penalizes the divergent values of the parameter, as a result, there exists at least one finite solution in the optimization problem. In this regards, Azzalini and Arellano-Valle (2013) proposed a penalty function for the skew-normal setting as

$$Q(\lambda) = c_1 \log(1 + c_2\lambda^2), \tag{4}$$

where $c_1 = 0.875913$ and $c_2 = 0.856250$. So, one can maximize

$$\ell_{SRS}(\lambda;\ \mathbf{x}_{SRS}) - Q(\lambda), \tag{5}$$

instead of $\ell_{SRS}(\lambda;\ \mathbf{x}_{SRS})$ itself.

Now assume that $\mathbf{x}_{RSS}$, $\mathbf{x}_{MinRSS}$, $\mathbf{x}_{MaxRSS}$ and $\mathbf{x}_{MedRSS}$ are observations of $\mathbf{X}_{RSS}$, $\mathbf{X}_{MinRSS}$, $\mathbf{X}_{MaxRSS}$ and $\mathbf{X}_{MedRSS}$ explained in Section 2. The same problem mentioned above may also be happened when these observations are used, but the probability of being divergent is different for each of them. We have estimated the probability of a divergent MLE based on $\mathbf{x}_{RSS}$ via a Monte Carlo simulation for some selected values of the sample size and the skewness parameter. Figure 2 exhibits the results. As it is observed from the figure, the behavior of the mentioned probability is almost the same as the one based on an $\mathbf{x}_{SRS}$ shown in Figure 1. However, it seems that the problem of reaching to a divergent MLE is more intensive here. Hence, to overcome to this problem, one may follow the same procedure of Azzalini and Arellano-Valle (2013). By utilizing the penalty function (4), the penalized log-likelihood functions of the parameter λ given the observed samples $\mathbf{x}_{RSS}$,

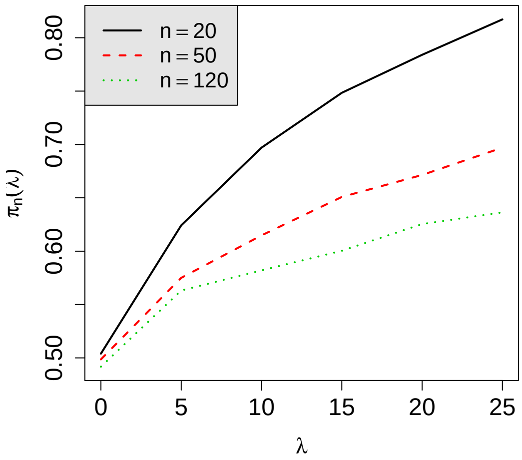


FIGURE 2. The probability of a divergent $\hat{\lambda}_{RSS}$ based on an RSS of size n from $SN(\lambda)$.

$\mathbf{x}_{MinRSS}$, $\mathbf{x}_{MaxRSS}$ and $\mathbf{x}_{MedRSS}$ are respectively derived as follows;

$$\begin{aligned}\ell_{RSS}(\lambda;\ \mathbf{x}_{RSS}) &= \sum_{j=1}^{r}\sum_{i=1}^{m}\log\phi_{i:m}(x_{(i)i,j};\lambda) - Q(\lambda)\\ &= \text{const.} + \sum_{j=1}^{r}\sum_{i=1}^{m}\big\{(i-1)\log\Phi(x_{(i)i,j};\lambda) + (m-i)\log\bar{\Phi}(x_{(i)i,j};\lambda)\\ &\quad + \log\Phi(\lambda x_{(i)i,j})\big\} - Q(\lambda), \qquad (6)\end{aligned}$$

where $\Phi(\cdot;\lambda)$ denotes the distribution function of $SN(\lambda)$, $\bar{\Phi}(\cdot;\lambda) \equiv 1-\Phi(\cdot;\lambda)$ and $\phi(\cdot;\lambda)$ is given by (1). Also, $\phi_{i:n}(z;\lambda)$ stands for the density of the ith order statistics from a random sample of size n arising from $SN(\lambda)$, i.e.

$$\phi_{i:n}(z;\lambda) = i\binom{n}{i}\phi(z;\lambda)\Phi^{i-1}(z;\lambda)\bar{\Phi}(z;\lambda)^{n-i},$$

where $\phi(z;\lambda)$ is given by (1).
The penalized log-likelihood function for the cases of MinRSS and MaxRSS are respectively obtained as

$$\begin{aligned}\ell_{MinRSS}(\lambda;\ \mathbf{x}_{MinRSS}) &= \sum_{j=1}^{r}\sum_{i=1}^{m}\log\phi_{1:m}(x_{(1)i,j};\lambda) - Q(\lambda)\\ &= \text{const.} + \sum_{j=1}^{r}\sum_{i=1}^{m}\big\{(m-1)\log\bar{\Phi}(x_{(1)i,j};\lambda) + \log\Phi(\lambda x_{(1)i,j})\big\} - Q(\lambda)\end{aligned} \qquad (7)$$

and

$$\begin{aligned}\ell_{MaxRSS}(\lambda;\ \mathbf{x}_{MaxRSS}) &= \sum_{j=1}^{r}\sum_{i=1}^{m}\log\phi_{m:m}(x_{(m)i,j};\lambda) - Q(\lambda)\\ &= \text{const.} + \sum_{j=1}^{r}\sum_{i=1}^{m}\left\{(m-1)\log\Phi(x_{(m)i,j};\lambda) + \log\Phi(\lambda x_{(m)i,j})\right\} - Q(\lambda),\end{aligned} \tag{8}$$

respectively. As mentioned in Section 2, the unites measured in the MedRSS scheme are different for odd and even set size m. The penalized log-likelihood of MedRSS for the former and the latter respectively are

$$\begin{aligned}\ell_{MedRSS}(\lambda;\ \mathbf{x}_{MedRSS}) &= \sum_{j=1}^{r}\sum_{i=1}^{m}\log\phi_{\frac{m+1}{2}:m}(x_{(\frac{m+1}{2})i,j};\lambda) - Q(\lambda)\\ &= \text{const.} + \sum_{j=1}^{r}\sum_{i=1}^{m}\left\{\frac{m-1}{2}\left(\log\Phi(x_{(\frac{m+1}{2})i,j};\lambda) + \log\bar{\Phi}(x_{(\frac{m+1}{2})i,j};\lambda)\right)\right.\\ &\quad \left. + \log\Phi(\lambda x_{(\frac{m+1}{2})i,j})\right\} - Q(\lambda)\end{aligned} \tag{9}$$

and

$$\begin{aligned}\ell_{MedRSS}(\lambda;\ \mathbf{x}_{MedRSS}) &= \sum_{j=1}^{r}\left\{\sum_{i=1}^{\frac{m}{2}}\log\phi_{\frac{m}{2}:m}(x_{(\frac{m}{2})i,j};\lambda) + \sum_{i=\frac{m}{2}+1}^{m}\log\phi_{\frac{m}{2}+1:m}(x_{(\frac{m}{2}+1)i,j};\lambda)\right\}\\ &\quad -Q(\lambda)\\ &= \text{const.} + \sum_{j=1}^{r}\sum_{i=1}^{\frac{m}{2}}\left\{\left(\frac{m}{2}-1\right)\log\Phi(x_{(\frac{m}{2})i,j};\lambda) + \frac{m}{2}\log\bar{\Phi}(x_{(\frac{m}{2})i,j};\lambda)\right.\\ &\quad \left. + \log\Phi(\lambda x_{(\frac{m}{2})i,j})\right\} + \sum_{j=1}^{r}\sum_{i=\frac{m}{2}+1}^{m}\left\{\frac{m}{2}\log\Phi(x_{(\frac{m}{2}+1)i,j};\lambda)\right.\\ &\quad \left. + \left(\frac{m}{2}-1\right)\log\bar{\Phi}(x_{(\frac{m}{2}+1)i,j};\lambda) + \log\Phi(\lambda x_{(\frac{m}{2}+1)i,j})\right\} - Q(\lambda).\end{aligned} \tag{10}$$

A numerical method must be employed in order to obtain the maximum (penalized) estimators by maximizing the penalized log-likelihoods (5)-(10).

Recently Salehi and Doostparast (2015) obtained an explicit expression for the moments of the order statistics arising from the skew-normal distribution in terms of the multivariate normal orthant probabilities. In fact, they found the expectation of the ith order statistics from a random sample of size r coming from the $SN(\lambda)$ distribution to be a linear combination as

$$\mu_{i:r}(\lambda) = \sum\left\{a_\lambda\Phi_{2k}\left(\mathbf{0};\tilde{\boldsymbol{\Omega}}_1\right) + b_\lambda\Phi_{2k}\left(\mathbf{0};\tilde{\boldsymbol{\Omega}}_2\right)\right\}, \tag{11}$$

where a_λ and b_λ are some real values, k is a positive integer number and $\tilde{\boldsymbol{\Omega}}_i$'s are some positive definite dispersion matrices (for more details see Salehi and Doostparast, 2015). Their results can be applied for obtaining the MME of the parameter λ when the rank-based sampling schemes are the only information about the $SN(\lambda)$ distribution. For instance, let $\mathbf{X}_{RSS} = \left\{X_{(i)i,j},\ i = 1, \cdots, m,\ j = 1, \cdots, r\right\}$ be an RSS of size $n = rm$ and

$$\bar{X}_{(i)} = \frac{1}{r}\sum_{j=1}^{r} X_{(i)i,j}$$

be the sample mean of the ith order statistics. Then, the MME of λ based on the $\mathbf{X}_{RSS}$ is derived as

$$\tilde{\lambda}_{RSS} = \frac{1}{r}\sum_{j=1}^{r} \tilde{\lambda}_i, \tag{12}$$

where $\tilde{\lambda}_i$ is the solution of the following equation

$$\mu_{i:r}(\tilde{\lambda}_i) = \bar{X}_{(i)}, \qquad i = 1, \ldots, r.$$

The same manner can be used in order to obtain the MME of λ on the basis of the other rank-based sampling plans.

4. Simulation study

This section provides a Monte Carlo simulation for comparing the sampling schemes as well as the estimating approaches which lead to get an answer to the opening question raised in the abstract. For simplicity of the comparison, all of the estimators obtained from the schemes are compared with the MLE of λ in the SRS plan, say $\hat{\lambda}_{SRS}$. More precisely, the following relative efficiency is utilized

$$\text{RE}(T, \hat{\lambda}_{SRS}) = \frac{\text{MSE}(T, \lambda)}{\text{MSE}(\hat{\lambda}_{SRS}, \lambda)},$$

where T is an estimator (MLE or MME) of the unknown parameter λ and MSE denotes the mean squared error. To be more closer in our comparison, the Pitman measure of closeness (PMC) is used as another criterion. It actually works on the basis of the probability instead of the moment. More specifically, if T_1 and T_2 are two common estimators of a real-valued parameter θ, with parameter space Θ, the PMC criterion is defined as the following probability

$$\text{PMC}(T_1, T_2|\theta) = \Pr(|T_1 - \theta| < |T_2 - \theta|).$$

Then, it is said that the estimator T_1 is Pitman-closer than T_2 if

$$\text{PMC}(T_1, T_2|\theta) \geq \text{PMC}(T_2, T_1|\theta),\ \forall\ \theta \in \Theta,$$

with strict inequality holding for at least one θ. Under the most commonly situation $\Pr(T_1 = T_2) = 0$, the above property holds if the simplified condition

$$\text{PMC}(T_1, T_2|\theta) \geq \frac{1}{2},\ \forall\ \theta \in \Theta, \tag{13}$$

is satisfied (still with strict inequality holding for at least one θ). PMC was originally introduced by Pitman (1937), as a criterion of comparison between two competing estimators. This concept is based on measuring the frequency with which one estimator is closer to an unknown parameter compared with a competing estimator. We refer the reader to the valuable publication by Keating et al. (1993) for further details on PMC.

Here, we consider perfect ranking, i.e. no error occurs in judgment ranking. The algorithm of the simulation is organized as follows.

Algorithm 1.

(i) Select λ from $\{-1, 1, 0, 3\}$, which respectively causes the data to be left-skewed, right-skewed, symmetric (about 0) and (almost) extremely right skewed, generate SRS, RSS, MinRSS, MaxRSS and MedRSS of sizes $n = mr$, $(m, r) \in \{(2, 5), (3, 10), (3, 20), (4, 20), (5, 20)\}$, as explained in Section 2. It should be noted that the set sizes are chosen small enough to avoid ranking error.

(ii) Compute the estimators mentioned in Sections 3 based on the observed samples generated in Step (i).

(iii) Repeat Steps (i) and (ii) for M =1000 times to get M observations of the estimators computed in Step (ii).

(iv) If $(t_i^{(1)}, ..., t_i^{(M)})$ are the observations of the estimator T_i, $i = 1, 2$, then the following criteria will be employed for the comparison purpose

$$\text{Bias}(T_1) = \frac{1}{M}\sum_{j=1}^{M} t_1^{(j)} - \lambda, \tag{14}$$

$$\text{RE}(T_1, \hat{\lambda}_{SRS}) = \frac{\sum_{j=1}^{M}(t_1^{(j)} - \lambda)^2}{\sum_{j=1}^{M}(\hat{\lambda}_{SRS}^{(j)} - \lambda)^2}, \tag{15}$$

$$\text{PMC}(T_1, T_2|\lambda) = \frac{1}{M}\sum_{j=1}^{M} I(|t_1^{(j)} - \lambda| < |t_2^{(j)} - \lambda|),$$

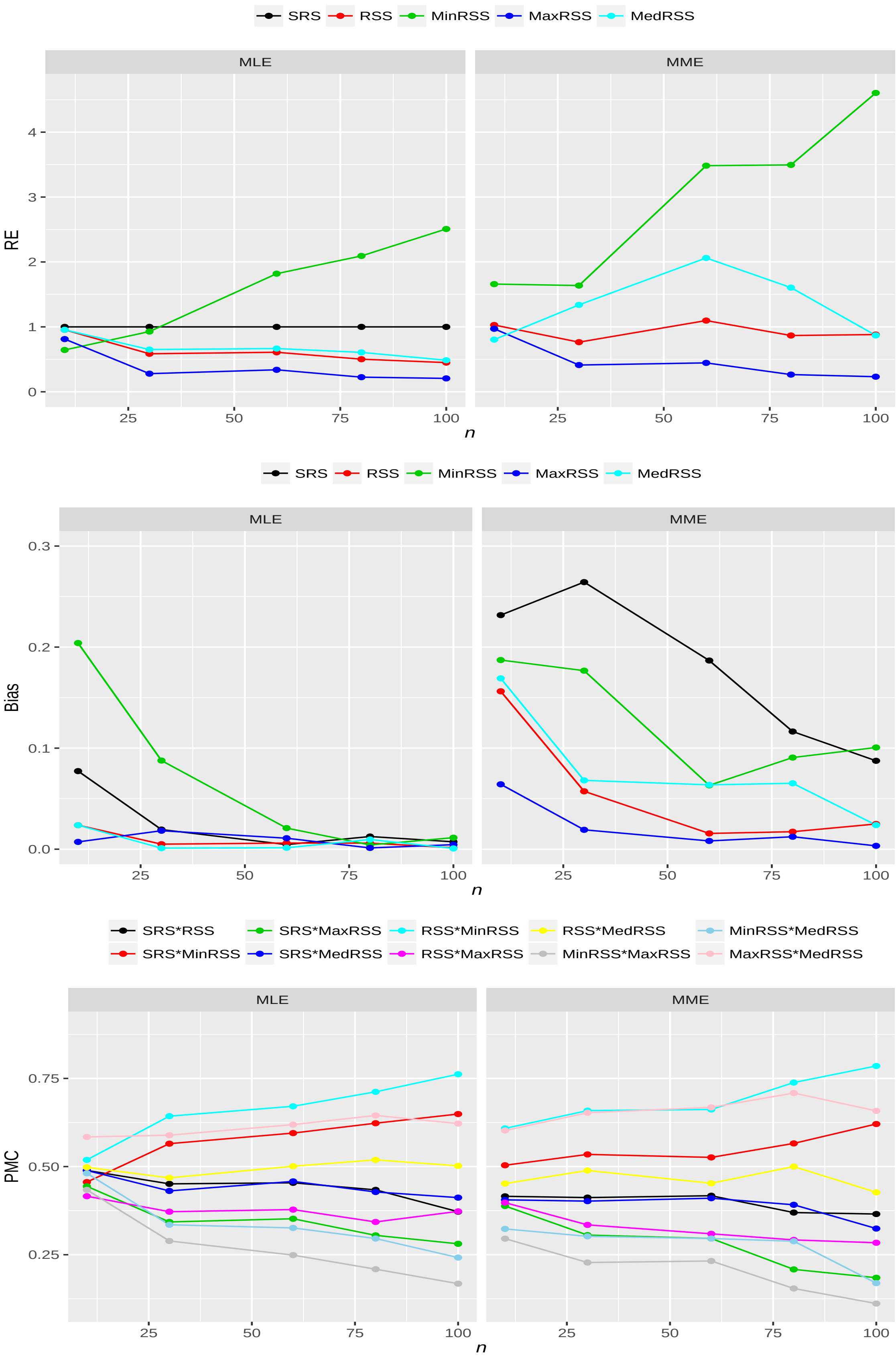
where $I(A)$ is the indicator function of the set A.

The all computations have been carried out with R software (R Core Team, 2018). It is also to be noted that we have used the routines `optimize` and `uniroot` where both are included in the basic `R` package `stats`. The former, which uses a combination of golden section search and successive parabolic interpolation (see Brent, 1973), has been employed for maximizing the log-likelihoods (5)-(10), while the latter has been applied for solving the equation (11).

The results of the simulation performed based on Algorithm 1 are displayed in Figures 3-6. In order to be convenient in comparison, both of estimators have the same Y-scale in each criterion.

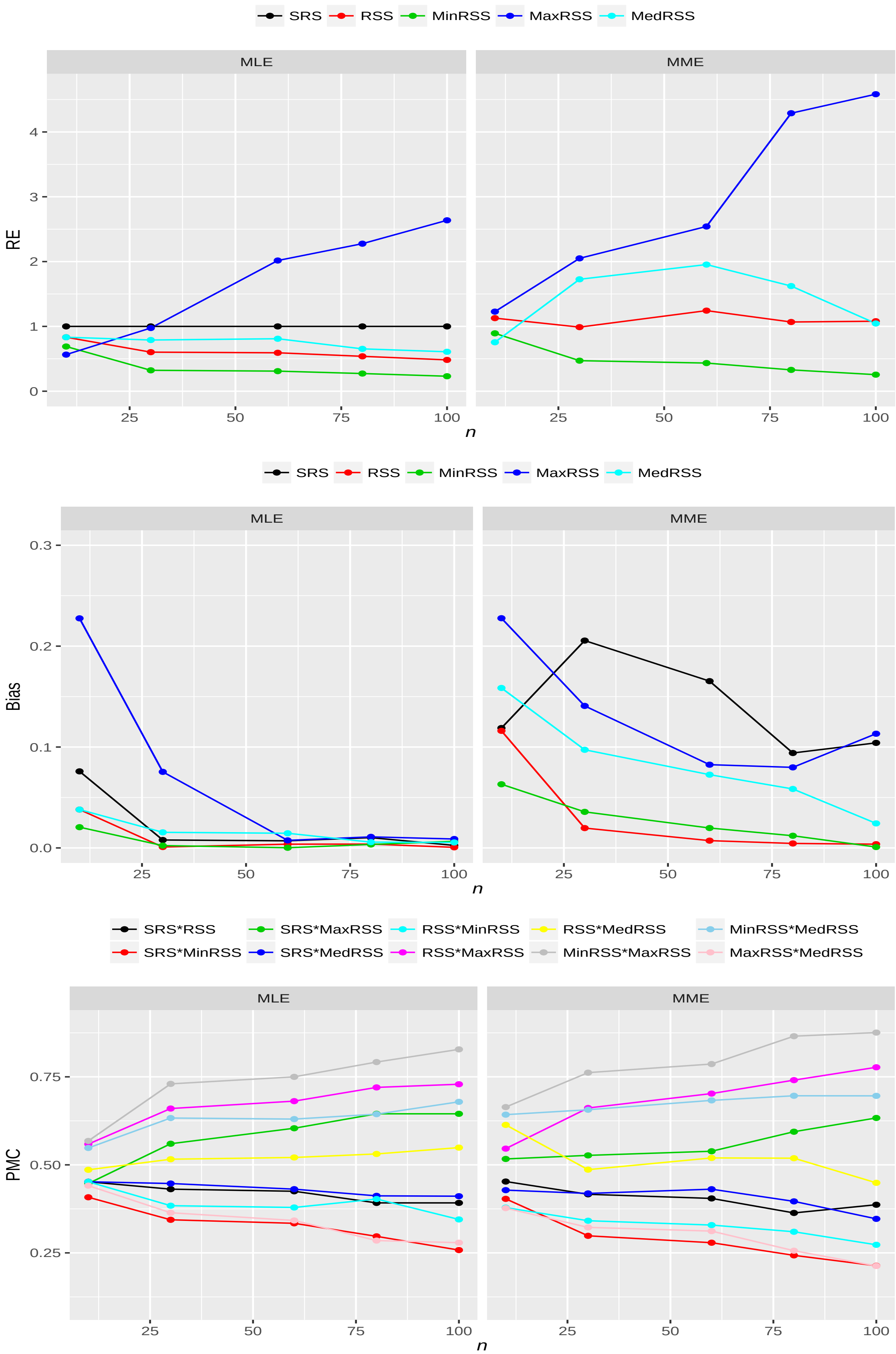
From Figures 3-6 (as well as some materials not shown here) it is observed that the results obtained based on the both of PMC and RE confirm each other. The following points may be extracted from these figures;

- The scheme which outperforms the other competitors based on the both estimation methods, is strongly dependent on the sign of the skewness parameter λ. More precisely, the MinRSS works well when the distribution is right-skewed ($\lambda > 0$), while the MaxRSS performance is the best when the distribution is left-skewed ($\lambda < 0$). Also, MedRSS and RSS have better performance than the other ones in the symmetric case $\lambda = 0$.
- As it was expected for the non-zero λ's, the efficiencies of the schemes MinRSS and MaxRSS almost everywhere play the role of lower-upper bounds and those of RSS and MedRSS move almost in the same line between these limits.
- Regardless of the values of λ, MedRSS and RSS are superior to SRS in case of MLE (however, this is not always true in the MME of MedRSS). While, MLE of SRS is superior to that of MinRSS (MaxRSS) for negative (positive) λ's, specially for the larger sample sizes.
- Generally, MLE has a better performance than MME.



* The black solid line is disappeared in the right panel of the RE figure duo to being out of scale.

FIGURE 3. Comparison of the various sampling schemes via Bias, RE and PMC criteria in SN(-1).



* The black solid line is disappeared in the right panel of the RE figure duo to being out of scale.

FIGURE 4. Comparison of the various sampling schemes via Bias, RE and PMC criteria in SN(1).

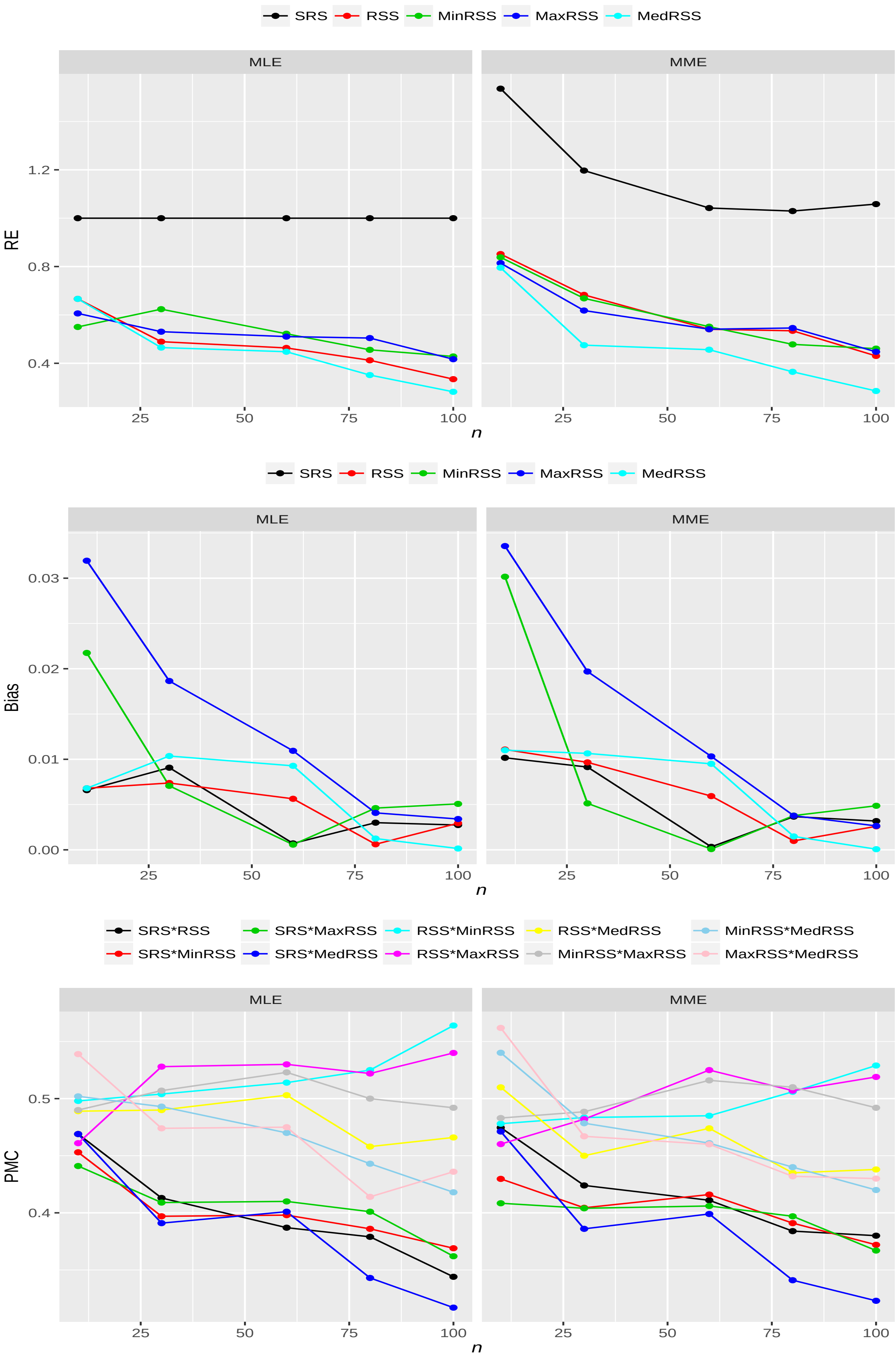
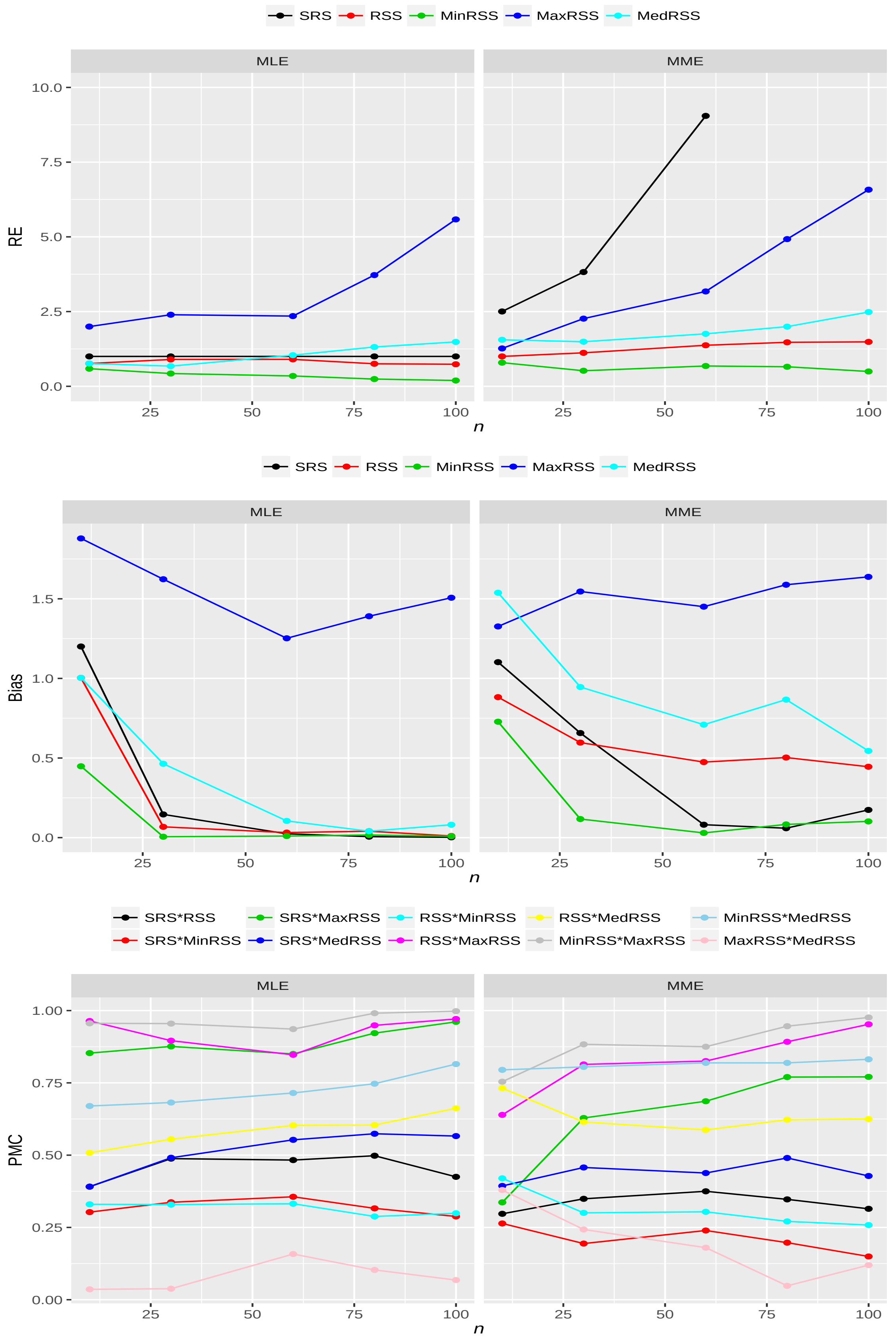


FIGURE 5. Comparison of the various sampling schemes via Bias, RE and PMC criteria in SN(0).



* Some segments of the black solid line (in the right panel of the RE figure) are vanished duo to being out of scale.

FIGURE 6. Comparison of the various sampling schemes via Bias, RE and PMC criteria in SN(3).

- RE and PMC criteria confirm that the difference between the estimators are much more significant for larger sample sizes, i.e. if an RE is greater than one, it is an increasing function of n, otherwise, it will be a decreasing function of n. The same fact holds for PMC as well, but with respect to the cut-point 0.5.
- The all estimators overestimates the parameter λ. However, for non-zero λ's, the estimators obtained based on the MedRSS, RSS as well as the best model for a given λ, are more closer to the unbiasedness than those of the other plans. Also, in the symmetric case, the all estimators are almost unbiased but still overestimate.

5. Conclusions

The paper was started off to answer to this question: among the well-known rank-based sampling schemes, namely, RSS, MedRSS and Modified RSS, as the competitors of SRS, the most frequently used method, which one is appropriate when the underlying distribution is SN(λ). To this end, the two estimation procedures MLE and MME were employed. MLE of λ was obtained based on maximizing the penalized likelihood function instead of the ordinary one. Simulation results showed that the scheme which outperforms the other competitors based on the both estimation approaches, is highly dependent on the sign of the skewness parameter λ. In more details, the MinRSS is suitable when the $\lambda > 0$, while the MaxRSS is the best when $\lambda < 0$. Also, MedRSS and RSS have better performance than the other ones when $\lambda = 0$.

References

[1] Al-Omari, I. and Bouza, C. N. (2014). Review of ranked set sampling: Modifications and applications, Revista Investigación Operacional, 3, 215-240.

[2] Azzalini, A. (1985). A class of distributions which includes the Normal ones, Scandinavian journal of statistics, 12, 171-178.

[3] Azzalini, A. and Arellano-Valle, R.B. (2013). Maximum penalized likelihood estimation for skew-normal and skew-t distributions, Journal of statistical planning and inference, 143, 419-433.

[4] Brent, R. (1973). Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, New Jersey.

[5] Sevinc, B., Cetintav, B., Esemen, M., and Gurler, S. (2019). RSSampling: A Pioneering Package for Ranked Set Sampling. R Journal, 11, 401.

[6] Chen, Z., Bai, Z. and Sinha, B. K. (2004). Ranked Set Sampling: Theory and Applications. *Springer-Vergla*, New York.

[7] Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics back ground, Biometrics, 28, 545-555.

[8] Esemen, M. and Gurler, S. (2018). Parameter estimation of generalized Rayleigh distribution based on ranked set sample. Journal of Statistical Computation and Simulation, 88, 615-628.

[9] Dey, S., Salehi, M. and Ahmadi, J. (2106). Rayleigh Distribution Revisited Via Ranked Set Sampling. METRON, 75, 69-85.

[10] Keating, J.P., Mason, R.L. and Sen, P. K. (1993). Pitman's Measure of Closeness : A Comparison of Statistical Estimators, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

[11] Matthews, M. and Wolfe, DA. (2016). Unified ranked set sampling, Statistics and probability letters, DOI: http://dx.doi.org/10.1016/j.spl.2016.11.015.

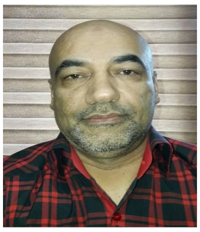
[12] McIntyre, G. (1952). A method for unbiased selective sampling using ranked set, Australian Journal of Agricultural Research, 3, 385-390.

[13] Muttlak, H. A. (1997). Median ranked set sampling, Journal of Applied Statistical Science, 6, 577-586.

[14] Pitman, E. (1937). The "closest" estimates of statistical parameters. Mathematical Proceedings of the Cambridge Philosophical Society, 33, 212-222.

[15] Salehi, M. and Doostparast, M. (2015). Expressions for moments of order statistics and records from the skew-normal distribution in terms of multivariate normal orthant probabilities. Journal of statistical methods and application, 24, 547-568.

[16] Salehi, M., Ahmadi, J. and Balakrishnan, N. (2015). Prediction of order statistics and record values based on ordered ranked set sampling. Journal of Statistical Computation and Simulation, 85, 77-88.
[17] Stokes, S. L. (1980). Inferences on the correlation coefficient in bivariate normal populations from ranked set samples, Journal of the American Statistical Association, 75, 989-995.
[18] 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.
[19] Wolfe DA. (2012). Ranked set sampling: its relevance and impact on statistical inference. ISRN Probability and Statistics, 1-32, DOI:10.5402/2012/568385.
[20] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.



H. Esfandyarifar was born in Ali Abad Katool, Iran, in 1971. He received his MSc degree in statistics-inference from Tarbiat Modares University, Iran, in 1997 and received PhD degree in statistics-inference from Payam Noor University, Iran, in 2017. He is an assistant professor in department of mathematics and statistics in Imam Ali University, Iran. His research interests are as follows: Ranked set sampling, Hierarchical Bayesian inference and sampling.

M. Salehi was born in Mashhad, Iran, in 1987. He received his MSc and PhD degrees in statistics-inference from Ferdowsi University of Mashhad, Iran, in 2011 and 2015. He has spent a six-month sabbatical leave under supervision of Prof. Adelchi Azzalini in Padova University of Italy (2014-2015). Now, he is an assistant professor in department of mathematics and statistics in University of Neyshabur, Iran. His area of research are as follows: Ordered data, Ranked set sampling and its variants, Flexible distributions and stress-strength reliability.