International Journal of Statistics and Probability; Vol. 13, No. 4; December 2024
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Moving Set Size Ranked Set Sampling

Mohammad Fraiwan Al-Saleh[1] & Hanan Youns Adam[1]

[1] Faculty of Science, Department of Statistics, Yarmouk University-Jordan

Correspondence: Mohammad Fraiwan Al-Saleh, Faculty of Science, Department of Statistics. Yarmouk University-Jordan.

Received: October 10, 2024 Accepted: November 21, 2024 Online Published: December 11, 2024

doi:10.5539/ijsp.v13n4p64 URL: https://doi.org/10.5539/ijsp.v13n4p64

Abstract

A new variation of ranked set sampling technique, entitled "Moving set size ranked set sampling", (MSSRSS), is suggested. The suggested variation allows for the increase of the number of units that can be chosen by judgement. The suitability of this modified technique for the estimation of parameters of the underlying distribution of interest is investigated. In addition, the estimation of the cumulative distribution function using this variation is considered. The results indicate that this modified technique can be useful in practice.

Keywords: ranked set sampling, simple random sampling, ranking error, concomitant variable, efficiency

1. Introduction

McIntyre (1952) introduced Ranked set sampling (RSS) technique; it is an alternative technique to simple random sampling (SRS). An annotated bibliography of RSS was given by Kaur, et al. (1995). A RSS of size m can be obtained using the following steps:

1. m random samples, each of size m, are chosen from the population of interest.
2. From the first sample, the minimum value (first order statistic) is identified by judgment, from the second sample, the second minimum value is identified,
3. Step 1,2 continue until the maximum unit is identified from the last sample.

The m chosen units are identified visually or by any inexpensive method(by judgment), with respect to the variable of interest. The RSS consists of the m chosen units from the m^2 selected units. To lower the chance of too much ranking errors, m should be small (2, 3 or 4). Increasing the size of the chosen sample can be done by repeating the above process r times to obtain a RSS of size $n = rm$.

In sampling from an infinite population, if there is no ranking error, then the RSS of size m consists of m independent order statistics. On the other hand, SRS consists of m dependent order statistics.

Let $Y^{j}_{(i:m)}$ is the i^{th} order statistic for the i^{th} sample of size m at the j^{th} cycle. Thus, the elements of RSS of size $n = rm$ can be denoted by:

$$\left\{Y^{j}_{(i:m)} : 1 \le i \le m, 1 \le j \le r\right\},$$

Note that for each i, $Y^{1}_{(i:m)}, Y^{2}_{(i:m)}, \ldots\ldots, Y^{r}_{(i:m)}$ are *iid* $f_{(i:m)}$, while for each j, $Y^{j}_{(1:m)}, Y^{j}_{(2:m)}, \ldots\ldots, Y^{j}_{(m:m)}$, are only independent; $f_{(i:m)}$ is the pdf of the i^{th} order statistic of a random sample of size m with common density f and cumulative distribution function cdf F:

$$f_{(i:m)}(x) = m\binom{m-1}{i-1}F^{i-1}(x)\left[1-F(x)\right]^{m-i} f(x).$$

In 1968, Takahasi & Wakimoto obtained the main properties of RSS. Assume that f is the underlying density function. Assume that the population mean is μ and the variance σ^2. Let $\mu_{(i:m)}$ =the mean of $(i:m)^{th}$ order statistic and $\sigma^2_{(i:m)}$ be its variance. The two authors highlighted the following important results:

$$f(x) = \sum_{i=1}^{m} f_{(i:m)}(x)/m.$$

$$\mu = \frac{1}{m}\sum_{i=1}^{m}\mu_{(i:m)}, \qquad \sigma^2 = \frac{1}{m}\sum_{i=1}^{m}\sigma^2_{(i:m)} + \frac{1}{m}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2 .$$

Using RSS, let

$$\hat{\mu}_{RSS} = \frac{1}{rm}\sum_{k=1}^{r}\sum_{i=1}^{m} X^{(k)}_{(i:m)}$$

be an estimator of μ and let:

$$\hat{\mu}_{SRS} = \frac{1}{rm}\sum_{i=1}^{rm} X_i$$

be the estimator of μ based SRS, $X_1, X_2, \ldots, X_n$. Clearly both estimators are unbiased. The relative efficiency(RE) of $\hat{\mu}_{RSS}$ with respect to $\hat{\mu}_{SRS}$ is:

$$RE\left(\hat{\mu}_{RSS};\hat{\mu}_{SRS}\right) = \frac{Var\left(\hat{\mu}_{SRS}\right)}{Var\left(\hat{\mu}_{RSS}\right)}.$$

They showed that

$$1 \le RE\left(\hat{\mu}_{RSS},\hat{\mu}_{SRS}\right) \le (m+1)/2 \quad .$$

The upper bound of RE is attained if the distribution of interest is uniform.

In 1972, Dell & Clutter investigated he effect of ranking error on the efficiency. They concluded that the basic result remains valid in the presence of ranking errors:

$$f(x) = \frac{1}{m}\sum_{i=1}^{m} f_{[i:m]}(x) .$$

$f_{[i:m]}(x)$ is the probability density function(pdf) of the realized order statistic produced by the ranking process (Judgement Ranking).

RSS for bivariate random variables was considered by Stokes (1977). let (X,Y) be the bivariate random variable. X is the variable of interest and Y is a variable that is not of direct interest but is relatively easy to measure and has strong relation with X. It is usually called the concomitant or subsidiary variable. The ranking of the units is done based on Y The resulting judgment order statistic is denoted by $X_{[i:m]}$. The estimation of the cdf $F(t)$ was considered by Stokes & Sager (1988):

Let $F(x) = P\left(X \le x\right)$, then it can be estimated, based on RSS, by

$$\hat{F}_{RSS}(x) = \frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} I_{\left(X^{j}_{(i:m)} \le x\right)}(x) .$$

$I_{\left(X^{j}_{(i:m)} \le x\right)}(x) = 1 \; if \; X^{j}_{(i:m)} \le x \quad and \; 0 \; o.w.$ (the indicator function). They showed that $\hat{F}_{RSS}(x)$ is an unbiased of $F(x)$ and is more efficient than $\hat{F}_{SRS}(x)$. Samawi, H., et al. (1996) investigated different variations of extreme ranked set sample (ERSS), which don't need a complete ranking, to estimate the μ.

Al-Saleh and Al Kadiri, introduced Double RSS (DRSS) variation of RSS in 2000, which was generalized to Multistage RSS by Al-Saleh and Al-Omary in (2002). Al Odat and Al-Saleh introduced another variation of RSS called moving extremes ranked set sampling in 2001, MERSS. Further investigation of This technique was done by Al-Saleh & Al

Hadhrami (2003). The estimation of Odds ratio, based on two types of MERSS, was considered by Samawi and Al-Saleh, M. in 2012. The estimation of the parameter of Downton's bivariate exponential distribution, using MERSS, was studied by Hanandeh and Al-Saleh (2013). Al-Saleh & Naamneh (2016) compared the properties of "5 Number Summery" obtained based on SRS, RSS and MERSS. Al-Saleh & Darabseh (2017) investigated the use of RSS techniques for the estimation of the parameter of the skew normal. For more works on RSS and its variations, see M. Al-Saleh and Abdul Wahab (2023), Maabreh and Al-Saleh (2023), Taconeli & Bonat (2020), Ozturk and wolfe (2000), (2001), Ozturk (2002), Al-Saleh and Ahmad (2018), Al-Saleh and Alshboul (2018), Al-Saleh and Zheng (2002), Zheng and Al-Saleh, M. (2002), Al-Saleh and Boreni (2018), Al-Saleh and Alshboul (2018), Al-Saleh & Darabseh (2017), Al-Saleh and Ababneh (2015), Samawi and Al-Saleh (2013), Samuh & Al-Saleh (2011), Al-Saleh and Mo'ath Mohammad (2023), Al-Saleh and Tani (2023), Al-Saleh and Abdul Wahab (2023).

In RSS, since ranking of units should be done by judgement and to avoid ranking error, the set size m should be small. Some distributions, especially skewed ones, may not benefit from RSS if m is small. MERSS allows for an improvement of efficiency because m can be taken larger.

In this paper, another variation of RSS technique is considered; "Moving set size ranked set sampling (MSSRSS)". It is the usual "ranked set sampling with varied set size". This will allow for the quantification of larger number of units even when m is small. The elements of the sample obtained using the MSSRSS for one cycle($r=1$) are denoted by:

$$X_{(i:j)},\ 1 \le i \le j, 1 \le j \le m.$$

Appropriateness of this modified technique for making inference about a distribution or its parameters will be investigated. The estimation of the population mean using MSSRSS under perfect ranking is considered in section 2, estimation under imperfect ranking is considered in section 3**,** section 4 deals with the estimation of distribution function based on MSSRSS. Conclusions and suggested future works are given in section 5.

2.1 Estimation the Population Mean Using MSSRSS With Perfect Ranking

For fixed m, MSSRSS consists of m ranked set samples of size $1,2,...,m$, respectively. The elements are:

$$\{X_{(i:j)} : X_{(1:1)}, X_{(1:2)}, X_{(2:2)}, X_{(1:3)}, X_{(2:3)}, X_{(3:3)}, \ldots, X_{(1:m)}, X_{(2:m)}, \ldots, X_{(m:m)}\}$$

$X_{(i:j)}$ is the i^{th} order statistic of a SRS of size j, $i=1,2,.....,j,\quad j=1,2,.....,m$. Clearly, $X_{(i:j)}$ are independent. The size of the obtained sample is:

$$n=\sum_{j=1}^{m} j = m(m+1)/2.$$

Assume that the underlying cdf F, is absolutely continuous with pdf, f, mean μ and variance σ^2. The pdf of $X_{(i:j)}$ is

$$f_{(i:j)}(x) = j\binom{j-1}{i-1}F^{i-1}(x)\left(1-F(x)\right)^{j-i} f(x).$$

Using the above identity of Takahasi and Wakimoto, we have:

$$\sum_{i=1}^{j} f_{(i:j)}(x) = jf(x).$$

Thus,

$$\sum_{j=1}^{m}\sum_{i=1}^{j} f_{(i:j)}(x) = \sum_{j=1}^{m} j f(x) = f(x)\sum_{j=1}^{m} j = \frac{m(m+1)}{2} f(x) = nf(x);$$

$$\mu = \int_{-\infty}^{\infty} xf(x)dx \quad = \frac{1}{n}\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\mu_{(i:j)}\right).$$

$$\mu_{(i:j)} = \int_{-\infty}^{\infty} x f_{(i:j)}(x)dx .$$

Let

$$\hat{\mu}_{MSSRSS} = \overline{X}_{MSSRSS} = \frac{2\sum_{j=1}^{m}\sum_{i=1}^{j} X_{(i:j)}}{m(m+1)}$$

be an estimator μ, which is the sample average

Now,

$$E\left(\overline{X}_{MSSRSS}\right) = \frac{2}{m(m+1)}\sum_{j=1}^{m}\left[\sum_{i=1}^{j} E\left(X_{(i:j)}\right)\right] = \frac{2}{m(m+1)}\sum_{j=1}^{m}\sum_{i=1}^{j}\mu_{(i:j)} = 2\sum_{j=1}^{m} j\mu_{(i:j)} / m(m+1) = \mu .$$

Therefore, $\hat{\mu}_{MSSRSS}$ is an unbiased estimator of μ.

$$Var\left(\overline{X}_{MSSRSS}\right) = \left(\frac{4\sum_{j=1}^{m}\left(\sum_{i=1}^{j} \operatorname{var}\left(X_{(i:j)}\right)\right)}{m^2(m+1)^2}\right) = \left(\frac{4\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\sigma_{(i:j)}^2\right)}{m^2(m+1)^2}\right).$$

By Takahasi & Wakimoto (1968), we have for each j:

$$\sigma^2 = \frac{1}{j}\left[\sum_{i=1}^{j}\sigma_{(i:j)}^2 + \sum_{i=1}^{j}\left(\mu_{(i:j)} - \mu\right)^2\right].$$

i.e.

$$\sum_{i=1}^{j}\sigma_{(i:j)}^2 = j\sigma^2 - \sum_{i=1}^{j}\left(\mu_{(i:j)} - \mu\right)^2 .$$

Thus,

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

$$= \frac{2}{m(m+1)}\sigma^2 - \left(\frac{2}{m(m+1)}\right)^2\sum_{j=1}^{m}\sum_{i=1}^{j}\left(\mu_{(i:j)} - \mu\right)^2 .$$

.

Let $X_1, X_2, \ldots, X_n$ be a SRS of size n from the same population. If

$$\hat{\mu}_{SRS} = \overline{X}_{SRS} = \frac{2\sum_{i=1}^{m(m+1)/2} X_i}{m(m+1)} .$$

Then,

$$E\left(\overline{X}_{SRS}\right) = \mu \quad \& \quad Var\left(\overline{X}_{SRS}\right) = \left(\frac{2}{m(m+1)}\right)\sigma^2 .$$

Clearly, $Var(\hat{\mu}_{MSSRSS})$ is smaller than $Var(\hat{\mu}_{SRS})$. Also,

$$RE\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)=\frac{Var\left(\hat{\mu}_{SRS}\right)}{Var\left(\hat{\mu}_{MSSRSS}\right)}=\left[1-2\frac{\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left(\mu_{(i:j)}-\mu\right)^2\right)}{m(m+1)\sigma^2}\right]^{-1}.$$

Thus, $RE\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)>1$. $\hat{\mu}_{MSSRSS}$ is more efficient than $\hat{\mu}_{SRS}$.

2.2. Efficiency of MSSRSS for Some of the Most Popular Distributions

(i) **Uniform distribution:** Assume that the sampling distribution is uniform with parameter θ, $U(0,\theta)$.

$f(x)=\frac{1}{\theta}$, $F(x)=\frac{x}{\theta}$, $0\le x\le\theta$. Thus, the pdf of $X_{(i:j)}$ is:

$$f_{(i:j)}(x)=j\binom{j-1}{i-1}F^{i-1}(x)\left(1-F(x)\right)^{j-i}f(x)=\frac{j!}{(i-1)!(j-i)!}\left(x/\theta\right)^{i-1}\left(1-x/\theta\right)^{j-i}\frac{1}{\theta}.$$

Clearly, the distribution of $\frac{1}{\theta}X_{(i:j)}$ Beta distribution with parameters $\alpha=i$ & $\beta=j-i+1$, $Beta(i,j-i+1)$.

Thus, $\mu_{(i:j)}=\frac{i}{j+1}\theta$, $Var\left(X_{(i:j)}\right)=\frac{i(j-i+1)}{(j+1)^2(j+2)}\theta^2$.

$$E\left(\hat{\mu}_{SRS}\right)=\frac{\theta}{2}=\mu \quad \& \quad Var\left(\hat{\mu}_{SRS}\right)=\frac{\theta^2}{6m(m+1)},$$

$$Var\left(\hat{\mu}_{MSSRSS}\right)=\frac{4}{m^2(m+1)^2}\left[\left(\frac{m(m+1)}{2}\right)\frac{\theta^2}{12}-\sum_{j=1}^{m}\sum_{i=1}^{j}\left(\frac{i}{j+1}\theta-\frac{\theta}{2}\right)^2\right].$$

Using the above results, the relative efficiency $\hat{\mu}_{MSSRSS}$ w.r.t. $\hat{\mu}_{SRS}$ is

$$RE\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)=\left[1-\frac{24\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left(\frac{i}{j+1}-\frac{1}{2}\right)^2\right)}{m(m+1)}\right]^{-1}.$$

(ii) Exponential Distribution: Suppose the sampling distribution is exponential with parameter θ. Then

$$f(x)=\frac{1}{\theta}e^{\frac{-x}{\theta}} \quad , \quad F(x)=1-e^{\frac{-x}{\theta}} \quad ,x\ge 0.$$

For this distribution, the mean is θ and the variance is θ^2. Thus,

$$f_{(i:j)}(x)=j\binom{j-1}{i-1}\frac{1}{\theta}\left(1-e^{\frac{-x}{\theta}}\right)^{i-1}\left(e^{\frac{-x}{\theta}}\right)^{j-i+1},$$

$$E\left(\hat{\mu}_{SRS}\right)=\theta\,,\ \ Var\left(\hat{\mu}_{SRS}\right)=2\theta^2/m(m+1)\,.$$

$$Var\left(\hat{\mu}_{MSSRSS}\right)=\left(\frac{2}{m(m+1)}\right)^2\left(\frac{m(m+1)}{2}\theta^2-\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left(\mu_{(i:j)}-\theta\right)^2\right)\right).$$

The relative efficiency of $\hat{\mu}_{MSSRSS}$ w.r.t $\hat{\mu}_{SRS}$ is

$$RE\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)=\frac{Var\left(\hat{\mu}_{SRS}\right)}{Var\left(\hat{\mu}_{MSSRSS}\right)}=\left[1-\frac{2}{m(m+1)}\sum_{j=1}^{m}\sum_{i=1}^{j}\left(\mu_{(i:j)}-1\right)^2\right]^{-1}.$$

(iii) Normal Distribution: Assume that X is $N(0,1)$ then the density function of $X_{(i:j)}$ is:

$$f_{(i:j)}(x)=j\binom{j-1}{i-1}\Phi^{i-1}(x)\left[1-\Phi(x)\right]^{j-i}\varphi(x)\,.$$

Thus,

$$E\left(X_{(i:j)}\right)=\frac{j!}{(i-1)!(j-i)!}\int_{-\infty}^{\infty}x(\Phi(x))^{i-1}\left[1-\Phi(x)\right]^{j-i}\phi(x)dx\ \ .$$

Using Scientific Work Place program(SWP), we obtained some values of $\mu_{(i:j)}$.
For general normal distribution, $N(\theta,\sigma^2)$,

$$E\left(\hat{\mu}_{SRS}\right)=\theta\ \ \&\ \ Var\left(\hat{\mu}_{SRS}\right)=2\sigma^2/m(m+1)\,,$$

$$Var\left(\hat{\mu}_{MSSRSS}\right)=\left(\frac{2}{m(m+1)}\right)^2\left(\frac{m(m+1)}{2}-\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left(\frac{\mu_{(i:j)}-\theta}{\sigma}\right)^2\right)\right)\sigma^2\,.$$

The value of the efficiency the estimator for normal distribution does not depend on the parameters, θ and σ^2 .

Thus, $$RE\left(\hat{\mu}_{MSSRSS},\hat{\mu}_{SRS}\right)=\frac{2/m(m+1)}{\left(\frac{4}{m^2(m+1)^2}\right)\left[\left(m(m+1)/2\right)-\sum_{j=1}^{m}\sum_{i=1}^{j}\mu_{(i:j)}^2\right]}=\left[1-\sum_{j=1}^{m}\sum_{i=1}^{j}\mu_{(i:j)}^2/2m(m+1)\right]^{-1}.$$

(iv) Half Normal Distribution: Assume that the sampling distribution is a half normal distribution; its pdf and cdf, using Al-Saleh & Darabseh (2017), is:

$$f(x)=\frac{2}{\sigma}\varphi\left(\left(x-\theta\right)/\sigma\right),\ x>\theta,\ \&\ 0\text{ o.w.},\ \ F(x)=2\Phi\left(\left(x-\theta\right)/\sigma\right)-1,\quad x>\theta,\ \&\text{ zero o.w.}$$

Also,

$$E(X)=\theta+\sigma\sqrt{\frac{2}{\pi}}\ \ \&\ \ Var(X)=\left(1-\frac{2}{\pi}\right)\sigma^2\,.$$

The pdf of $X_{(i:j)}$ is:

$$f_{(i:j)}(x)=\frac{j!}{(i-1)!(j-i)!}\left(2\Phi\left(\frac{x-\theta}{\sigma}\right)-1\right)^{i-1}\left[2-2\Phi\left(\frac{x-\theta}{\sigma}\right)\right]^{j-i}\frac{2}{\sigma}\phi\left(\frac{x-\theta}{\sigma}\right).$$

Clearly, the value of the relative efficiency for this distribution does not depend on parameters, thus, let $\mu=0$ & $\sigma^2=1$:

$$f_{(i:j)}(x)=\frac{j!}{(i-1)!(j-i)!}(2\Phi(x)-1)^{i-1}[2-2\Phi(x)]^{j-i}\,2\phi(x).$$

Therefore,

$$E\left(X_{(i:j)}\right)=\frac{j!}{(i-1)!(j-i)!}\int_0^\infty 2z\left(2\Phi(z)-1\right)^{i-1}\left[2-2\Phi(z)\right]^{j-i}\varphi(z)dz$$

Again, using SWP program we obtained the values of $\mu_{i:j}$.

For half-normal distribution:

$$E\left(\hat{\mu}_{SRS}\right)=\sqrt{\frac{2}{\pi}}+\theta \quad \& \quad (\text{when } \sigma^2=1)$$

$$Var\left(\hat{\mu}_{SRS}\right)=\frac{2\left(1-\frac{2}{\pi}\right)}{m(m+1)}$$

$$Var\left(\hat{\mu}_{MSSRSS}\right)=\left(\frac{2}{m(m+1)}\right)^2\left(\frac{m(m+1)}{2}\left(1-\frac{2}{\pi}\right)-\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left(\mu_{(i:j)}-\sqrt{2/\pi}\right)^2\right)\right)$$

Thus,

$$RE\left(\hat{\mu}_{MSSRSS},\hat{\mu}_{SRS}\right)=\frac{\frac{2}{m(m+1)}(1-2/\pi)}{\left(\frac{4}{m^2(m+1)^2}\right)^2\left[(0.5m(m+1))(1-2/\pi)-\sum_{j=1}^{m}\sum_{i=1}^{j}\left(\mu_{(i:j)}-\sqrt{2/\pi}\right)^2\right]}=\left[1-\frac{1}{2}\frac{\sum_{j=1}^{m}\sum_{i=1}^{j}\left(\mu_{(i:j)}-\sqrt{2/\pi}\right)^2}{m(m+1)(1-2/\pi)}\right]^{-1}.$$

The total number of elements need to be measured in MSSRSS is $n=\frac{1}{2}m(m+1)$; this is also the sample size of the obtained SRS. If m is odd, then we can obtain an equivalent RSS of size $n=\frac{1}{2}m(m+1)$, by taking $\frac{1}{2}(m+1)$ RSS samples with set size m each. Otherwise, the three methods may not be truly comparable. For example, for $m=3$, the three sampling techniques can be compared based on $n=6$: MSSRSS of size 6, SRS of size 6 and 2 RSS of size 3 each. One important thing that should be taken into account when comparing the methods is the **total number of elements that need to be chosen to obtain the required sample**.

The Relative efficiency for estimating the population mean using MSSRSS w.r.t. SRS and RSS, for $m=2,3,4,5$, is given in Table (2.1). The values of $RE\left(\hat{\mu}_{RSS},\hat{\mu}_{SRS}\right)$ for half-normal distribution is taken from Al-Saleh and Darabseh (2017); for the other three distributions, see Stokes (1995).

Table 2.1. $Eff\left(\hat{\mu}_{RSS};\hat{\mu}_{SRS}\right)$, $Eff\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)$, $Eff\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{RSS}\right)$

Distribution	m	$Eff\left(\hat{\mu}_{RSS};\hat{\mu}_{SRS}\right)$	$Eff\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{SRS}\right)$	$Eff\left(\hat{\mu}_{MSSRSS};\hat{\mu}_{RSS}\right)$
Uniform $(0,\theta)$	2	1.500	1.285	-
	3	2.000	1.565	0.78
	4	2.500	1.840	-
	5	3.000	2.113	0.70
Exp(θ)	2	1.330	1.200	-
	3	1.640	1.380	0.84
	4	1.920	1.560	-
	5	2.190	1.724	0.79
Normal (θ,σ^2)	2	1.470	1.269	-
	3	1.910	1.526	0.79
	4	2.350	1.775	-
	5	2.770	2.016	0.73
Half Normal (θ,σ^2)	2	1.429	1.250	-
	3	1.841	1.490	0.81
	4	2.239	1.720	-
	5	2.628	1.940	0.74

Based on Table (2.1), we can see that MSSRSS gives a more efficient unbiased estimator of μ than SRS. The efficiency is increasing in *m*. While when we compare the efficiency of using MSSRSS with respect to RSS, we see that the efficiency is smaller than one. However, the total number of elements needed to obtain MSSRSS is smaller than the needed elements to get RSS with the same size. For example: with $m=3$ we need 14 element to obtain a MSSRSS of size 6, while, we need 18 elements to obtain RSS of size 6. In general, we need $\sum_{j=1}^{m} j^2=\frac{1}{6}m(m+1)(2m+1)$ observations to obtain a MSSRSS of size $n=0.5m(m+1)$, but we need for m odd, $m^2(m+1)/2$ observations to get a RSS of the same size. Now,

$$\tfrac{1}{2}m^2(m+1)/\tfrac{1}{6}m(m+1)(2m+1)=3m/(2m+1),$$

which is larger than one and tends to 1.5 as m gets large. In addition, judgment ranking is easier when selecting MSSRSS than RSS, i.e. the new procedure, though less efficient that RSS, is **<u>easier to execute and may be less prone to error</u>** than RSS procedure.

3. Estimation the Mean Using Moving Set Size RSS Under Imperfect Ranking

In section 2, it is claimed that the obtained MSSRSS is error free; i.e. there is no ranking error. It is rarely happened that we can perfectly rank a set of elements by judgment, in particular when the set size is not small. To reduce difficulty of ranking and the effect of ranking error, another random variable that is related to the variable of interest is usually used. MSSRSS is obtained based on this concomitant variable. In this section, the relative efficiency of MSSRSS w.r.t. SRS when error in ranking may exist is investigated.

For fixed m, to obtain a MSSRSS based on concomitant variable, we assume that variable of interest Y is related to another variable X that is easy to rank without error. MSRSS is obtained based on this concomitant variable using the following steps:

1. For $j=1,2,\ldots,m$, obtain a MSSRSS of size *j* based on the concomitant variable X.
2. Measure the value of Y that correspond to each element obtained in (1).

The elements of the MSSRSS with concomitant variable are:

$$\left\{\left(X_{(i:j)},Y_{[i:j]}\right); i=1,\ldots,j,\ j=1,\ldots,m\right\}.$$

$X_{(i:j)}$ is the i^{th} order statistics of a random sample of size *j* and $Y_{[i:j]}$ is the corresponding y-value; it is the concomitant to the actual order statistic, so it is the concomitant order statistic. The MSSRSS consists of $\frac{1}{2}m(m+1)$ $\left(X_{(i:j)},Y_{[i:j]}\right)$.

It was shown by S. Yang (1977) that the conditional distribution of the concomitant order statistic given the order statistic is the same as the conditional distribution of the variable of interest given the concomitant variable:

$$\left(Y_{[i:j]} \mid X_{(i:j)}\right) \overset{D}{=} \left(Y \mid X\right) \; .$$

Thus, the relative efficiency of the estimator based on $Y_{[i:j]}$ w.r.t the corresponding one based on SRS can be obtained. We will calculate the efficiency when ranking is imperfect for two well-known bivariate distributions:

(i) Bivariate Normal Distribution

Assume that $(X,Y) \sim BN\left(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho\right)$, where ρ is the correlation coefficient of X & Y. Their pdf is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \frac{-1}{2\left(1-\rho^2\right)}\left[\frac{\left(x-\mu_x\right)^2}{\sigma^2{}_x} - 2\rho\left(\frac{\left(x-\mu_x\right)\left(y-\mu_y\right)}{\sigma_x\sigma_y}\right) + \frac{\left(y-\mu_y\right)^2}{\sigma^2{}_y}\right]$$

Also, $f_{X,Y}(x,y) = f_X(x)\, f_{Y|X}(y|x)$, where $f_X(x)$ is the marginal pdf of X which is $N\left(\mu_x, \sigma_x^2\right)$ & $f_{Y|X}(y|x)$ is the pdf of $(Y|X=x)$ which is $N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x), \sigma_y^2\left(1-\rho^2\right)\right)$.

Using Yang (1977)

$$f_{Y_{[i:j]}|X_{(i:j)}}\left(y_{[i:j]} \mid x_{(i:j)}\right) = f_{Y|X}\left(y_{[i:j]} \mid x_{(i:j)}\right).$$

Thus,

$$E\left(Y_{[i:j]} \middle| X_{(i;j)}\right) = \mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(X_{(i:j)} - \mu_x\right),$$

and

$$E\left(Y_{[i:j]}\right) = E\left(E\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right) = \mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(E(X_{(i:j)}) - \mu_x\right).$$

Also,

$$Var\left(Y_{[i:j]} \middle| X_{(i:j)}\right) = \sigma_y^2\left(1-\rho^2\right).$$

Now,

$$\text{“}Var\left(Y_{[i:j]}\right) = E\left(Var\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right) + Var\left(E\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right).$$

But,

$$E\left[Var\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right] = \sigma_y^2\left(1-\rho^2\right),$$

and

$$Var\left(E\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right) = Var\left(\mu_y + \rho\sigma_y\left(\left(X_{(i:j)} - \mu_x\right)/\sigma_x\right)\right) = \rho^2\sigma_y^2 Var\left(\left(X_{(i:j)} - \mu_x\right)/\sigma_x\right),$$

where

$$Var\left(\frac{X_{(i:j)} - \mu_x}{\sigma_x}\right) = \frac{1}{\sigma_x^2}\sigma^2_{x(i:j)} = \frac{1}{\sigma_x^2}\sigma^2_{(i:j)}\sigma_x^2 = \sigma^2_{(i:j)}\text{”},$$

$\sigma^2_{(i:j)}$ is the variance of the i^{th} standard normal order statistic of a SRS of size j and $\sigma^2_{x(i:j)}$ is the variance of the i^{th} order statistic of a SRS of size j from the distribution of X (M. Al-Saleh and AL-Ananbeh (2007)).

Therefore,

$$Var\left(Y_{[i:j]}\right) = \sigma_y^2\left(1-\rho^2\right) + \rho^2\sigma_y^2\sigma^2_{(i:j)}$$

Let $Y_1, Y_2, \ldots, Y_m$ be a SRS of size $m(m+1)/2$ from a population with mean μ_y and variance σ_y^2. Then

$$\hat{\mu}_{SRS} = \overline{Y}_{SRS} = \frac{2\sum_{i=1}^{m(m+1)/2} Y_i}{m(m+1)},$$

$$E\left(\overline{Y}_{SRS}\right) = \mu_y \ \& \ Var\left(\overline{Y}_{SRS}\right) = \frac{2\sigma_y^2}{m(m+1)}.$$

$$RE\left(\hat{\mu}_{yMSSRSS}, \hat{\mu}_{ySRS}\right) = \frac{Var\left(\hat{\mu}_{ySRS}\right)}{Var\left(\hat{\mu}_{yMSSRSS}\right)},$$

where

$$\hat{\mu}_{yMSSRSS} = \overline{Y}_{MSSRSS} = \frac{2\sum_{j=1}^{m}\sum_{i-1}^{j} Y_{[i:j]}}{m(m+1)},$$

with

$$E\left(\hat{\mu}_{yMSSRSS}\right) = \frac{2\sum_{j=1}^{m}\sum_{i-1}^{j} E\left(Y_{[i:j]}\right)}{m(m+1)}$$

$$= \frac{2\sum_{j=1}^{m}\left(\sum_{i=1}^{j}\left[\mu_y + \rho\frac{\sigma_y}{\sigma_x}E\left(X_{(i:j)} - \mu_x\right)\right]\right)}{m(m+1)}$$

$$= \mu_y + \frac{2\rho\frac{\sigma_y}{\sigma_x}\sum_{j=1}^{m}\left(\sum_{i=1}^{j}E\left(X_{(i:j)} - \mu_x\right)\right)}{m(m+1)},$$

Using $\sum_{i=1}^{j} E\left(X_{(i:j)}\right) = j\mu_x$, we have $E\left(\overline{Y}_{MSSRSS}\right) = \mu_y$. Thus, the estimator $\hat{\mu}_{yMSSRSS}$ is unbiased.

$$Var\left(\overline{Y}_{MSSRSS}\right) = \frac{4}{m^2(m+1)^2}\sum_{j=1}^{m}\sum_{i-1}^{j}\left(\sigma_y^2\left(1-\rho^2\right) + \rho^2\sigma_y^2\sigma^2_{(i:j)}\right)$$

$$= \frac{\frac{m(m+1)}{2}\sigma_y^2(1-\rho^2)+\rho^2\sigma_y^2\sum_{j=1}^{m}\sum_{i=1}^{j}\sigma_{(i:j)}^2}{\left(\frac{m(m+1)}{2}\right)^2} = \frac{\sigma_y^2(1-\rho^2)}{\left(\frac{m(m+1)}{2}\right)}+\rho^2\sigma_y^2\frac{\sum_{j=1}^{m}\sum_{i=1}^{j}\sigma_{(i:j)}^2}{\left(\frac{m(m+1)}{2}\right)^2}$$

Thus, $Var\left(\overline{Y}_{MSSRSS}\right)=\frac{2\sigma_y^2\left(1-\rho^2\right)}{m\left(m+1\right)}+\rho^2\sigma_y^2 Var\left(\overline{X}_{MSSRSS}\right)$.

The relative efficiency of $\hat{\mu}_{yMSSRSS}$ w.r.t $\hat{\mu}_{ySRS}$ is

$$RE\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)=\frac{1}{(1-\rho^2)+\rho^2\frac{m(m+1)}{2}Var\left(\overline{X}_{MSSRSS}\right)}.$$

If $\sigma_y^2=\sigma_x^2=1$, then

$$RE\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)=\frac{1}{1-\rho^2\left(1-\left[Eff\left(\hat{\mu}_{xMSSRSS},\hat{\mu}_{xSRS}\right)\right]^{-1}\right)}.$$

Clearly, for fixed m, $RE\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)$ is increasing in ρ and increasing in m for fixed ρ. $Eff\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)$for $m=2,3,4,5$ and for different values of ρ are obtained. Table (3.1) below contains these values. It can be seen that efficiency is larger than 1 and is increasing in m for fixed ρ and increasing in ρ for fixed m. For small ρ, the efficiency is very close to 1. Thus, for this method to be beneficial, the two variables must have a strong linear relation.

Table 3.1. $RE\left(\hat{\mu}_{yMSSRSS};\mu_{ySRS}\right)$ for Bivariate normal dist

$m\downarrow$ $\lvert\rho\rvert\rightarrow$	0	0.2	0.4	0.6	0.8	$\lvert\rho\rvert\rightarrow 1$
2	1	1.008	1.035	1.083	1.157	1.269
3	1	1.014	1.058	1.142	1.283	1.526
4	1	1.018	1.075	1.187	1.388	1.775
5	1	1.020	1.088	1.222	1.476	2.016

(ii) Downton's Bivariate Exponential Distribution

"*Downton's Bivariate Exponential Distribution*", (DBE), was introduced by Downton (1970); for more details, see M. Al-Saleh and Y. Diab (2009)). (X,Y) has $DBE(\theta_x,\theta_y,\rho)$ if its pdf is

$$f(x,y;\theta_x,\theta_y,\rho)=\frac{1}{\theta_x\theta_y(1-\rho)}\exp\left[-\frac{1}{(1-\rho)}\left(\frac{x}{\theta_x}+\frac{y}{\theta_y}\right)\right]\times I_0\left[\frac{2\sqrt{\rho xy}}{(1-\rho)\sqrt{\theta_x\theta_y}}\right]$$

x,y,θ_x,θ_y, are positive, $0\le\rho<1$.

$$I_0(z)=\sum_{k=0}^{\infty}\frac{(z/2)^{2k}}{k!^2}$$

$I_0(z)$ is the modified Bessel function of the first kind of order zero. The marginal distribution of X and Y are exponential with parameters θ_x & θ_y, respectively. It can be seen that (See M.F. Al-Saleh and Y. Diab, 2009)

" $E(Y \mid X = x) = (1-\rho)\theta_y + \rho \frac{\theta_y}{\theta_x} x$ & $Var(Y \mid X = x) = (1-\rho)^2 \theta_y^2 + 2\rho(1-\rho) \frac{\theta_y^2}{\theta_x} x$".

Since $I_0(0) = 1$, X and Y are indep. if $\rho = 0$. Also, $0 \le \rho < 1$. Using Yang (1977):

$$E\left(Y_{[i;j]} \middle| X_{(i:j)}\right) = (1-\rho)\ \theta_y + \rho \frac{\theta_y}{\theta_x} X_{(i:j)},$$

and

$$E\left(Y_{[i:j]}\right) = E\left(E\left(Y_{[i;j]} \middle| X_{(i:j)}\right)\right) = (1-\rho)\theta_y + \rho \frac{\theta_y}{\theta_x} E\left(X_{(i:j)}\right).$$

Also,

$$Var\left(Y_{[i:j]} \middle| X_{(i:j)}\right) = (1-\rho)^2 \theta_y^2 + 2\rho(1-\rho)\frac{\theta_y^2}{\theta_x} X_{(i:j)},$$

and

$$Var\left(Y_{[i:j]}\right) = E\left[Var\left(Y_{[i;j]} \middle| X_{(i:j)}\right)\right] + Var\left[E\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right],$$

where

$$E\left[Var\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right] = (1-\rho)^2 \theta_y^2 + 2\rho(1-\rho)\frac{\theta_y^2}{\theta_x} E\left(X_{(i:j)}\right),$$

and

$$Var\left[E\left(Y_{[i:j]} \middle| X_{(i:j)}\right)\right] = Var\left[(1-\rho)\ \theta_y + \rho \frac{\theta_y}{\theta_x} X_{(i:j)}\right]$$

$$= \rho^2 \frac{\theta_y^2}{\theta_x^2} Var\left(X_{(i:j)}\right).$$

Hence,

$$Var\left(Y_{[i:j]}\right) = (1-\rho)^2 \theta_y^2 + 2\rho(1-\rho)\frac{\theta_y^2}{\theta_x}\theta_x + \rho^2 \frac{\theta_y^2}{\theta_x^2} Var\left(X_{(i:j)}\right)$$

$$= \left(1-\rho^2\right) \theta_y^2 + \rho^2 \frac{\theta_y^2}{\theta_x^2} Var\left(X_{(i:j)}\right).$$

Now,

$$E\left(\bar{Y}_{MSSRSS}\right) = \frac{2}{m(m+1)} \sum_{j=1}^{m} \sum_{i-1}^{j} E\left(Y_{[i:j]}\right) = \frac{2}{m(m+1)} \sum_{j=1}^{m} \sum_{i=1}^{j} \left[(1-\rho)\ \theta_y + \rho \frac{\theta_y}{\theta_x} E\left(X_{(i:j)}\right)\right] = \theta_y .$$

Thus, the estimator $\hat{\mu}_{yMSSRSS}$ is unbiased. Also,

$$Var\left(\overline{Y}_{MSSRSS}\right)=\frac{4}{m^2(m+1)^2}\sum_{j=1}^{m}\left(\sum_{i-1}^{j}Var\left(Y_{[i:j]}\right)\right)=\frac{4}{m^2(m+1)^2}\sum_{j=1}^{m}\left(\sum_{i-1}^{j}\left(\left(1-\rho^2\right)\theta_y^2+\rho^2\frac{\theta_y^2}{\theta_x^2}Var\left(X_{(i:j)}\right)\right)\right)$$

$$=\frac{2\left(1-\rho^2\right)\theta_y^2}{m(m+1)}+4\rho^2\frac{\theta_y^2}{\theta_x^2}\frac{\sum_{j=1}^{m}\sum_{i=1}^{j}Var\left(X_{(i:j)}\right)}{m^2(m+1)^2}=\frac{2\left(1-\rho^2\right)\theta_y^2}{m(m+1)}+\rho^2\frac{\theta_y^2}{\theta_x^2}Var\left(\overline{X}_{MSSRSS}\right).$$

Thus,

$$RE\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)=\frac{Var\left(\overline{Y}_{SRS}\right)}{Var\left(\overline{Y}_{MSSRSS}\right)}=\frac{\frac{2\theta_y^2}{m(m+1)}}{\frac{2\theta_y^2}{m(m+1)}\left(1-\rho^2\right)+\rho^2\frac{\theta_y^2}{\theta_x^2}Var\left(\overline{X}_{MSSRSS}\right)}=\frac{1}{\left(1-\rho^2\right)+\rho^2\left[RE\left(\hat{\mu}_{xMSSRSS},\hat{\mu}_{xSRS}\right)\right]^{-1}}.$$

The values of $RE\left(\hat{\mu}_{yMSSRSS};\hat{\mu}_{ySRS}\right)$ for $m=2,3,4,5$ and different values of ρ are obtained. The values are in Table (3.2). Clearly, the relative efficiency is always larger than 1 and is increasing in m ;it is very close to 1 when ρ is small. As $\rho\to 1$, the efficiency converges to $Eff\left(\hat{\mu}_{xMSSRSS},\hat{\mu}_{xSRS}\right)$.

Table 3.2. $Eff\left(\hat{\mu}_{yMSSRSS},\hat{\mu}_{ySRS}\right)$ for Downton's bivariate exponential for m =2,3,4,5

$m\downarrow$ $\rho\to$	0	0.2	0.4	0.6	0.8	$\rho\to 1$
2	1	1.007	1.027	1.064	1.120	1.200
3	1	1.011	1.046	1.111	1.216	1.385
4	1	1.014	1.061	1.148	1.297	1.558
5	1	1.017	1.072	1.178	1.368	1.724

Thus, error in ranking reduces $Eff\left(\hat{\mu}_{yMSSRSS},\hat{\mu}_{ySRS}\right)$. The effect of ranking errors can be reduced if the supplementary variable X has a strong linear relation with the variable of interest Y. However, no matter how large is the ranking error, efficiency continues to be large than one.

4. Estimation of Distribution Function Using MSSRSS

The estimation of the cdf using MSSRSS is considered in this section. In particular, we will compare the estimator of the distribution function based on MSSRSS with the corresponding estimator based on SRS.

Let $X_1,X_2,\ldots,X_n$ be a SRS of size ;$m(m+1)/2$ with common cdf $F(t)$. Assume that $F(t)$ is absolutely continuous. For a given t, the well-known estimator of $F(t)$ is:

$$\hat{F}_{SRS}(t)=\frac{2\sum_{i=1}^{m(m+1)/2}I(X_i\le t)}{m(m+1)}.$$

$\hat{F}_{SRS}(t)$ is the empirical cdf.

Now, $E(I(X_i\le t))=P(X_i\le t)=F(t)$. Thus, $E(\hat{F}_{SRS}(t))=F(t),$ i.e. $\hat{F}_{SRS}(t)$ is an unbiased estimator of $F(t)$. Also,

$$Var\left(I\left(X_i\le t\right)\right)=F(t)\left(1-F(t)\right).$$

Thus,

$$Var\left(\hat{F}_{SRS}(t)\right)=\frac{4\sum_{i=1}^{m(m+1)/2}F(t)\left(1-F(t)\right)}{m^2\left(m+1\right)^2}=\frac{2F(t)(1-F(t))}{m(m+1)}.$$

For more detail, see Al-Saleh and Dana Ahmad (2018).

Let $X_{(i:j)}, i=1,2,....,j, \quad j=1,2,....,m$ be a MSSRSS . The empirical distribution function based on MSSRSS is:

$$\hat{F}_{MSSRSS}(t)=\frac{1}{\frac{m(m+1)}{2}}\sum_{j=1}^{m}\sum_{i=1}^{j}I\left(X_{(i;j)}\leq t\right).$$

Thus,

$$E\left(\hat{F}_{MSSRSS}(t)\right)=\frac{2\sum_{j=1}^{m}\sum_{i=1}^{j}F^{i}_{(i:j)}(t)}{m(m+1)}.$$

Using the identity of Takahasi and Wakimoto (1968, we have:

$$\sum_{i=1}^{j}F_{(i:j)}(t)=jF(t), \quad j=1,2,\dots m,$$

thus,

$$\sum_{j=1}^{m}\sum_{i=1}^{j}F_{(i:j)}(t)=\sum_{j=1}^{m}jF(t)=F(t)\sum_{j=1}^{m}j=\frac{m(m+1)}{2}F(t).$$

Therefore, $E\left(\hat{F}_{MSSRSS}(t)\right)=F(t)$, i.e. $\hat{F}_{MSSRSS}(t)$ is an unbiased estimator for $F(t)$. Also, $Var\left(I\left(X_{(i:j)}\leq t\right)\right)=F_{(i:j)}(t)\left(1-F_{(i;j)}(t)\right)$.

Thus,

$$Var\left(\hat{F}_{MSSRSS}(t)\right)=\frac{4}{m^2\left(m+1\right)^2}\sum_{j=1}^{m}\sum_{i=1}^{j}F_{(i:j)}(t)\left(1-F_{(i:j)}(t)\right)$$

$$=\frac{4}{m^2\left(m+1\right)^2}\sum_{j=1}^{m}\left(jF(t)-\sum_{i=1}^{j}F^2_{(i:j)}(t)\right)$$

$$=\frac{1}{\left(m(m+1)/2\right)^2}\sum_{j=1}^{m}\left(F(t)\frac{m(m+1)}{2}-\sum_{j=1}^{m}\sum_{i=1}^{j}F^2_{(i:j)}(t)\right).$$

The relative efficiency is:

$$RE\left(\hat{F}_{MSSRSS}(t);\hat{F}_{SRS}(t)\right)=\frac{Var\left(\hat{F}_{SRS}(t)\right)}{Var\left(\hat{F}_{MSSRSS}(t)\right)}$$

$$= \frac{F(t)\left(1-F(t)\right)}{F(t) - \dfrac{2\sum_{j=1}^{m}\sum_{i=1}^{j} F_{(i:j)}^{2}(t)}{m\left(m+1\right)}}.$$

$F_{(i:j)}(t)$ is the cdf of the i^{th} order statistic of a sample of size j:

$$F_{(i:j)}(t) = P(\text{at least } i \text{ of the } j \text{ values are } \le x) = \sum_{k=i}^{j}\binom{j}{k}F^{i}(t)\left(1-F(t)\right)^{j-k}.$$

Numerical values of $RE\left(\hat{F}_{MSSRSS}(t);\hat{F}_{SRS}\left(t\right)\right)$ are given in Table 4.1.

Table 4.1. $Eff(\hat{F}_{MSSRSS}(t);\hat{F}_{SRS}(t))$ for some values of m and $F(t)$

$F(t)$ / m	0.2	0.5	0.7	0.9
2	1.12	1.20	1.16	1.06
3	1.23	1.37	1.30	1.13
4	1.34	1.52	1.44	1.19
5	1.43	1.66	1.56	1.24

Clearly, $\ddot{F}_{MSSRSS}(t)$ is more efficient than $\ddot{F}_{SRS}(t)$. The relative efficiency is increasing in m for fixed $F(t)$. For fixed m, the efficiency is increasing in $F(t)$ for $F(t) \le 0.5$ and decreasing for $F(t) > 0.5$. The best values are at $F(t) = 0.5$.

5. Conclusions

In this paper, another variation of RSS, moving set size ranked set sampling (MSSRSS), is introduced and investigated. It turned out that this new technique is more efficient than SRS for estimating the population mean and more convenient than RSS; it needs less effort and it is less prone to ranking error. Error in ranking reduces the efficiency of MSSRSS w.r.t SRS, however, no matter of how large the ranking error, efficiency continues to be larger than one. For estimating the distribution function $F(t)$, it turned out that the estimator based on MSSRSS is more efficient than the estimator based on SRS. The method can be investigated parametrically, i.e. the use of some parametric methods of estimation such as the maximum likelihood, method of moments, etc. Also, the amount of information in MSSRSS can be compared to that in SRS and RSS.

Acknowledgment

The authors are very thankful to the referees for the comments and the editor of the journal.

Conflicts of interest/Competing interest/ Funding

We affirm that there is no conflict of interest regarding the publication of this article. The article is unfunded.

References

Al-Odat, M. T., & Al-Saleh, M. F. (2001). A variation of ranked set sampling. *Journal of Applied Statistical Science, 10*, 137-146.

Al-Saleh, M. F., & Mohammad, M. (2023). Estimation of the common mean of two normal distributions using ranked set sampling and some of its variations. *Communications in Statistics - Simulation and Computation*. Advance online publication. https://doi.org/10.1080/03610918.2023.xxxxxxxx

Al-Saleh, M. F., & Abdul Wahab, B. (2023). A close look at the estimation of population size. *International Journal of Statistics and Probability, 12*(4), 19-25. https://doi.org/10.5539/ijsp.v12n4p19

Al-Saleh, M. F., & Tani, L. (2023). Estimation of the population size using ranked set sampling and some of its variations. *Journal of Probability and Statistical Science, 21*(2), 1-15.

Al-Saleh, M. F., & Boreni, D. (2018). Estimation of the distribution function using moving extreme ranked set sampling (MERSS). In C. Bouza-Herrera & A. Al-Omari (Eds.), *Ranked set sampling: 65 years improving the accuracy in data gathering* (1st ed., pp. 43-58). ELSEVIER.

Al-Saleh, M. F., & Alshboul, H. (2018). Other approaches to bivariate ranked set sampling. *Communications for Statistical Applications and Methods, 25*, 283-296. https://doi.org/10.29220/csam.2018.25.4.283

Al-Saleh, M. F., & Darabseh, M. (2017). Inference on the skew normal distribution using ranked set sampling. *International Journal of Computational and Theoretical Statistics, 4*, 65-76.

Al-Saleh, M. F., & Naamneh, A. (2016). Properties of the elements of SRS, RSS, MERSS. *Journal of Applied Statistical Science, 22*, 75-85.

Al-Saleh, M. F., & Ababneh, A. (2015). Test for accuracy in ranking in moving extreme ranked set sampling. *International Journal of Computational and Theoretical Statistics, 2*, 67-77.

Al-Saleh, M. F., & Samawi, H. (2010). On estimating the odds using moving extreme ranked set sampling. *Statistical Methodology, 7*, 133-140. https://doi.org/10.1016/j.stamet.2009.09.002

Al-Saleh, M. F., & Diab, Y. A. (2009). Estimation of the parameters of Downton's bivariate exponential distribution using ranked set sampling scheme. *Journal of Statistical Planning and Inference, 139*, 277-286. https://doi.org/10.1016/j.jspi.2008.04.025

Al-Saleh, M. F., & Al-Ananbeh, A. M. (2007). Estimation of the means of the bivariate normal using moving extreme ranked set sampling with concomitant variable. *Statistical Papers, 48*, 179-195. https://doi.org/10.1007/s00362-006-0306-6

Al-Saleh, M. F., & Al-Hadrami, S. A. (2003). Estimation of the mean of the exponential distribution using moving extremes ranked set sampling. *Statistical Papers, 44*, 367-382. https://doi.org/10.1007/s00362-003-0164-z

Al-Saleh, M. F., & Al-Kadri, M. A. (2000). Double ranked set sampling. *Statistics & Probability Letters, 48*, 205-212. https://doi.org/10.1016/S0167-7152(00)00022-7

Al-Saleh, M. F., & Al-Omary, A. (2002). Multi-stage ranked set sampling. *Journal of Statistical Planning and Inference, 102*, 273-286. https://doi.org/10.1016/S0378-3758(01)00206-3

Al-Saleh, M. F., & Zheng, G. (2002). Estimation of bivariate characteristics using ranked set sampling. *The Australian & New Zealand Journal of Statistics, 44*, 221-232. https://doi.org/10.1111/1467-842X.00225

Dell, T. R., & Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics, 28*, 545-553. https://doi.org/10.2307/2528497

Downton, F. (1970). Bivariate exponential distribution in reliability theory. *Journal of the Royal Statistical Society. Series B (Methodological), 32*, 408-417. https://doi.org/10.1111/j.2517-6161.1970.tb00857.x

Hanandeh, A., & Al-Saleh, M. F. (2013). Inference on Downton's bivariate exponential distribution based on moving extreme ranked set sampling. *Austrian Journal of Statistics, 42*, 161-179. https://doi.org/10.17713/ajs.v42i3.56

Kaur, A., Patil, G. P., Sinha, A. K., & Taillie, C. (1995). Ranked set sampling: An annotated bibliography. *Environmental and Ecological Statistics*, *2*, 25-54. https://doi.org/10.1007/BF00569672

Maabreh, A., & Al-Saleh, M. F. (2023). Estimation of the location parameter of Cauchy distribution using some variations of the RSS technique. *Mathematics and Statistics, 10*(1), 263-276. https://doi.org/10.13189/ms.2022.100128

McIntyre, G. (1952). A method for unbiased selective sampling using ranked set sampling. *Australian Journal of Agricultural Research, 3*, 385–390. https://doi.org/10.1071/AR9520385

Ozturk, O. (2002). Ranked set sample inference under a symmetry restriction. *Journal of Statistical Planning and Inference, 102*, 317-336. https://doi.org/10.1016/S0378-3758(01)00192-6

Ozturk, O., & Wolfe, D. A. (2001). A new ranked set sampling protocol for the signed rank test. *Journal of Statistical Planning and Inference*, *96*, 351-370. https://doi.org/10.1016/S0378-3758(00)00352-9

Ozturk, O., & Wolfe, D. A. (2000). Alternative ranked set sampling protocol for the sign test. *Statistics & Probability Letters, 47*, 15-23. https://doi.org/10.1016/S0167-7152(99)00157-6

Samawi, H., & 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, 42*, 17-24. https://doi.org/10.1016/j.jkss.2012.03.004

Samawi, H. M., Ahmed, M. S., & Abu-Dayyeh, W. (1996). Estimating the population mean using extreme RSS. *Biometrical Journal, 38*, 577-586. https://doi.org/10.1002/bimj.4710380504

Samuh, M., & Al-Saleh, M. F. (2011). The effectiveness of multistage ranked set sampling in stratifying the population. *Communications in Statistics - Theory and Methods, 40*, 1063-1080. https://doi.org/10.1080/03610920903444011

Stokes, S. L. (1995). Parametric RSS. *Annals of the Institute of Statistical Mathematics, 47*(3), 465-482. https://doi.org/10.1007/BF00773466

Stokes, S. L., & Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. *Journal of the American Statistical Association, 83*, 374-381. https://doi.org/10.1080/01621459.1988.10478590

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics - Theory and Methods*, *A6*, 1207-1211. https://doi.org/10.1080/03610927708827563

Taconeli, C. A., & Bonat, W. H. (2020). On the performance of estimation methods under ranked set sampling. *Computational Statistics*. https://doi.org/10.1007/s00180-020-00953-9

Takahasi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics, 20*, 1-31. https://doi.org/10.1007/BF02911285

Yang, S. S. (1977). General distribution theory of the concomitants of order statistics. *Annals of the Institute of Statistical Mathematics, 5*, 996-1002. https://doi.org/10.1007/BF02481056

Zheng, G., & Al-Saleh, M. F. (2002). Modified maximum likelihood estimators based on RSS. *Annals of the Institute of Statistical Mathematics*, *54*, 641-650. https://doi.org/10.1023/A:1022421326631

Copyrights