

# The efficiency of L ranked set sampling in estimating the distribution function

Amer Ibrahim Al-Omari

Received: 20 February 2014 / Accepted: 8 October 2014 / Published online: 4 November 2014  
© African Mathematical Union and Springer-Verlag Berlin Heidelberg 2014

**Abstract** In this paper, L ranked set sampling (LRSS) technique (Al-Nasser in *Simul Comput*, 6:33–43, 2007) is considered for estimating the distribution function of a random variable. The suggested estimator of the distribution function is compared with its counterparts using simple random sampling (SRS) and ranked set sampling (RSS) schemes. It is found that the suggested LRSS estimator of the distribution function is biased and is more efficient than that of the SRS and RSS for a given  $x$  based on the number of measured units.

**Keywords** Distribution function · L ranked set sampling · Simple random sampling · Relative efficiency

**Mathematics Subject Classification** 62D05

## 1 Introduction

The RSS is suggested by [10] for estimating the population mean of pasture and forage yields as efficient method than the SRS. The RSS can be implemented in situations when the sampling units in a study can be more easily ranked than quantified.

The RSS as suggested [10] can be described as:

**Step 1:** Randomly select  $n^2$  units from the target population.

**Step 2:** Allocate the  $n^2$  selected units as randomly as possible into  $n$  sets, each of size  $n$ .

**Step 3:** Rank the units within each set with respect to a variable of interest. This may be based on personal professional judgment or done based on a concomitant variable correlated with the variable of interest.

---

A. I. Al-Omari (✉)  
Department of Mathematics, Faculty of Science, Al al-Bayt University,  
Mafraq, Jordan  
e-mail: alomari\_amer@yahoo.com

**Step 4:** Choose a sample for actual quantification by including the smallest ranked unit in the first set, the second smallest ranked unit in the second set, the process continues in this way until the largest ranked unit is selected from the last set.

**Step 5:** Repeat Steps 1 through 4 for  $m$  cycles to obtain a sample of size  $mn$  for actual quantification.

Let  $X_1, X_2, \dots, X_n$  be a SRS of size  $n$  with a continuous probability density function (pdf)  $\varphi(x)$  and cumulative distribution function (cdf)  $\Phi(x)$ , with a finite mean  $\mu$  and variance  $\sigma^2$ . Let  $X_{11h}, X_{12h}, \dots, X_{1nh}; X_{21h}, X_{22h}, \dots, X_{2nh}; \dots; X_{n1h}, X_{n2h}, \dots, X_{nmh}$  be  $n$  independent SRS each of size  $n$  in the  $h$ th cycle ( $h = 1, 2, \dots, m$ ).

Assume that  $X_{i(1:n)h}, X_{i(2:n)h}, \dots, X_{i(n:n)h}$  are the order statistics of the  $i$ th sample  $X_{i1h}, X_{i2h}, \dots, X_{inh}$ , ( $i = 1, 2, \dots, n$ ) in the  $h$ th cycle ( $h = 1, 2, \dots, m$ ). Then,  $X_{1(1:n)h}, X_{2(2:n)h}, \dots, X_{n(n:n)h}$  is a RSS of size  $n$ . The cdf  $\Phi_{(j:n)}(x)$  of the  $j$ th order statistics  $X_{(j:n)}$ , is given by

$$\Phi_{(j:n)}(x) = \sum_{i=j}^n \binom{n}{i} [\Phi(x)]^i [1 - \Phi(x)]^{n-i}, \quad -\infty < x < \infty, \tag{1}$$

and the pdf  $\varphi_{(j:n)}(x)$  is given by

$$\varphi_{(j:n)}(x) = \binom{n}{j} [\Phi(x)]^{j-1} [1 - \Phi(x)]^{n-j} \varphi(x). \tag{2}$$

Also,  $E(X_{(j:n)}) = \mu_{(j:n)} = \int_{-\infty}^{\infty} x\varphi_{(j:n)}(x)dx$  and  $\text{Var}(X_{(j:n)}) = \sigma_{(j:n)}^2 = \int_{-\infty}^{\infty} (x - \mu_{(j:n)})^2 \varphi_{(j:n)}(x)dx$ . See [8] for more details. [14] independently introduced the same RSS method and they showed that

$$\varphi(x) = \frac{1}{n} \sum_{j=1}^n \varphi_{(j:n)}(x), \quad \mu = \frac{1}{n} \sum_{j=1}^n \mu_{(j:n)} \quad \text{and} \quad \sigma^2 = \frac{1}{n} \sum_{j=1}^n \sigma_{(j:n)}^2 + \frac{1}{n} \sum_{j=1}^n [E(X_{(j:n)}) - \mu]^2.$$

As a modification of the usual RSS [1] suggested the LRSS as a robust method for estimating the population mean. The LRSS method consists of the following steps:

**Step 1:** Select  $n$  samples each of size  $n$  from the target population.

**Step 2:** Rank the units within each sample with respect to a variable of interest by visual inspection or any cost free method.

**Step 3:** Select the LRSS coefficient  $k = [\alpha n]$ , where  $0 \leq \alpha < 0.5$ , and  $[x]$  is largest integer value less than or equal to  $x$ .

**Step 4:** For each of the first  $k$  ranked samples, select the units with the rank  $k + 1$  for actual measurement.

**Step 5:** For each of the last  $k$  ranked samples, i.e., the  $(n - k)$ th to the  $n$ th ranked samples, select the units with the rank  $n - k$  for actual measurement.

**Step 6:** For  $j = k + 2, k + 3, \dots, n - k$ , the unit with rank  $j$  in the  $j$ th ranked sample is selected for actual measurement.

**Step 5:** The procedure can be repeated  $m$  cycles if needed to obtain a sample of size  $nm$  units.

Note that if  $k = 0$ , then the LRSS will be reduced to the usual RSS. The LRSS is used for estimating the linear regression model by [2].

Let  $X_{1(k+1:n)}, X_{2(k+1:n)}, \dots, X_{k(k+1:n)}, X_{k+1(k+1:m)}, X_{k+2(k+2:n)}, \dots, X_{n-k(n-k:n)}, X_{n-k+1(n-k:n)}, X_{n-k+2(n-k:n)}, \dots, X_{n(n-k:n)}$  be a LRSS of size  $n$ . The LRSS estimator of the population mean is defined as

$$\bar{X}_{LRSS(k,n)} = \frac{1}{n} \left( \sum_{i=1}^k X_{i(k+1:n)} + \sum_{i=k+1}^{n-k} X_{i(i:n)} + \sum_{i=n-k+1}^n X_{i(n-k:n)} \right).$$

Its variance is given by

$$Var(\bar{X}_{LRSS(k,n)}) = \frac{1}{n^2} \left( \sum_{i=1}^k Var(X_{i(k+1:n)}) + \sum_{i=k+1}^{n-k} Var(X_{i(i:n)}) + \sum_{i=n-k+1}^n Var(X_{i(n-k:n)}) \right)$$

For more information about RSS see for example [12] suggested extreme ranked set samples (ERSS) for estimating a population mean. [6] suggested multistage RSS for estimating the population mean. Estimation of mean based on modified robust ERSS is considered by [3]. [11] used ERSS and median RSS for estimating the distribution function. [9] considered the distribution function estimation using extreme median RSS. [4] considered ratio type estimators of population mean using auxiliary information based on median RSS.

The remaining part of this paper is organized as follows: In Sect. 2, estimation of the distribution function is considered using SRS, RSS, and the suggested LRSS. Numerical comparisons are given in Sect. 3. Finally, in Sect. 4 some conclusions and recommendations are presented.

### 2 Estimation of $\Phi(x)$ using SRS, RSS and LRSS

Let  $\Phi_{SRS}(x)$  be the empirical distribution function of a SRS of size  $nm$ ,  $X_1, X_2, \dots, X_{nm}$  from  $\Phi(x)$ . It is well known that  $\Phi_{SRS}(x)$  is an unbiased estimator of  $\Phi(x)$  for a given  $x$ , with variance  $Var[\Phi_{SRS}(x)] = \frac{1}{nm} \Phi(x) [1 - \Phi(x)]$ , and  $\Phi_{SRS}(x)$  is a consistent estimator of  $\Phi(x)$ , (See [7]).

Recently, a new estimator of  $\Phi(x)$  is suggested by [13] using RSS for fixed  $x$  as  $\Phi_{RSS}(x) = \frac{1}{nm} \sum_{h=1}^m \sum_{j=1}^n I(X_{j(j:n)h} \leq x)$ , where  $I(\cdot)$  is an indicator function. They proved that  $E[\Phi_{RSS}(x)] = \frac{1}{nm} \sum_{h=1}^m \sum_{j=1}^n \Phi_{j(j:n)h}(x) = \Phi(x)$ ,  $Var[\Phi_{RSS}(x)] = \frac{1}{n^2m} \sum_{j=1}^n \Phi_{(j:n)}(x) [1 - \Phi_{(j:n)}(x)]$ , and  $\frac{\Phi_{RSS}(x) - E[\Phi_{RSS}(x)]}{\sqrt{Var[\Phi_{RSS}(x)]}} \xrightarrow{D} N(0, 1)$  as  $m \rightarrow \infty$  when  $x$  and  $n$  are fixed.

The suggested LRSS estimator of the distribution function  $\Phi(x)$  is defined as

$$\begin{aligned} \Phi_{LRSS}(x) = \frac{1}{nm} & \left[ \sum_{h=1}^m \sum_{j=1}^k I(X_{j(k+1:n)h} \leq x) + \sum_{h=1}^m \sum_{j=k+1}^{n-k} I(X_{j(j:n)h} \leq x) \right. \\ & \left. + \sum_{h=1}^m \sum_{j=n-k+1}^n I(X_{j(n-k:n)h} \leq x) \right], \end{aligned} \tag{3}$$

with variance

$$\begin{aligned} \text{Var} [\Phi_{LRSS}(x)] &= \frac{k}{n^2} \{ \Phi_{(k+1:n)}(x) [1 - \Phi_{(k+1:n)}(x)] + \Phi_{(n-k:n)}(x) [1 - \Phi_{(n-k:n)}(x)] \} \\ &\quad + \frac{1}{n^2} \sum_{j=1}^{n-2k} \Phi_{(j:n)}(x) [1 - \Phi_{(j:n)}(x)], \end{aligned} \tag{4}$$

where

$$\Phi_{(r:n)}(x) = I_{\Phi(x)}(r, n - r + 1) = \frac{1}{B(a, b)} \int_0^{\Phi(x)} t^{a-1} (1 - t)^{b-1} dt, \tag{5}$$

$$B(a, b) = \int_0^1 t^{a-1} (1 - t)^{b-1} dt, \quad a > 0, \quad b > 0, \quad \text{and} \quad B(a, b) = \frac{(a - 1)!(b - 1)!}{(a + b - 1)!}.$$

It is well known that when we use the ranked set sampling method we select a small sample size in order to rank the units visually or by any cheap method. So to increase the sample size we repeat the process  $m$  cycles for fixed value of the sample size. Therefore, in the following two propositions we assumed that the number of cycles gets large.

**Proposition 1**

$$\begin{aligned} E [\Phi_{LRSS}(x)] &= \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] + \frac{1}{n} \sum_{j=k+1}^{n-k} \Phi_{(j:n)}(x), \\ &= \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] - \frac{1}{n} \left[ \sum_{j=1}^k \Phi_{(j:n)}(x) + \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \right] \\ &\quad + \Phi(x) \end{aligned} \tag{6}$$

with bias

$$\begin{aligned} \text{Bias} [\Phi_{LRSS}(x)] &= E [\Phi_{LRSS}(x)] - \Phi(x) \\ &= \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] \\ &\quad - \frac{1}{n} \left[ \sum_{j=1}^k \Phi_{(j:n)}(x) + \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \right]. \end{aligned} \tag{7}$$

*Proof* To prove (6), take the expectation of (3) as

$$\begin{aligned} E [\Phi_{LRSS}(x)] &= \frac{1}{nm} E \left[ \sum_{h=1}^m \sum_{j=1}^k I (X_{j(k+1:n)h} \leq x) + \sum_{h=1}^m \sum_{j=k+1}^{n-k} I (X_{j(j:n)h} \leq x) \right. \\ &\quad \left. + \sum_{h=1}^m \sum_{j=n-k+1}^n I (X_{j(n-k:n)h} \leq x) \right] \\ &= \frac{1}{n} \left\{ k [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] + E \left[ \sum_{h=1}^m \sum_{j=k+1}^{n-k} I (X_{j(j:n)h} \leq x) \right] \right\} \end{aligned}$$

**Table 1** The efficiency and bias values of  $\Phi_{LRSS}(x)$  with respect to  $\Phi_{SRS}(x)$  for  $n = 4, 5$

$\Phi(x)$	$n = 4$		$n = 5$		
	$k = 0$	$k = 1$	$k = 0$	$k = 1$	$k = 2$
0.01	1.0211	14.5442	1.0534	<b>12.5349</b>	18.6471
		-0.0097		-0.0097	-0.0100
0.10	1.2896	1.8754	1.3887	1.6776	1.7811
		-0.0712		-0.0656	-0.0915
0.20	1.5382	1.2975	1.6919	1.3120	1.0130
		-0.0961		-0.0795	-0.1418
0.30	1.7010	1.1671	1.8684	1.2944	0.9146
		-0.0829		-0.0681	-0.1365
0.40	1.8100	1.1406	1.9820	1.4093	0.9377
		-0.0464		-0.0354	-0.0825
0.50	<b>1.8322</b>	1.1674	<b>2.0062</b>	1.3992	1.0002
		-0.0009		0.0360	0.0002
0.60	1.8308	1.1679	2.0061	1.3992	0.9664
		0.0502		0.0360	0.0823
0.70	1.7045	1.1568	1.8680	1.3044	0.9185
		0.0841		0.0658	0.1368
0.80	1.5421	1.2866	1.6622	1.2734	1.0450
		0.0964		0.0816	0.1411
0.90	1.2944	1.8714	1.3745	1.7056	1.7909
		0.0718		0.0653	0.0909
0.99	1.0326	<b>15.4601</b>	1.0854	<b>11.9669</b>	<b>20.6824</b>
		0.0097		0.0096	0.0100

$$\begin{aligned}
 &= \frac{1}{n} \left\{ k [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] + \sum_{j=k+1}^{n-k} \Phi_{(j:n)}(x) \right\} \\
 &= \frac{1}{n} \left\{ k [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] + \sum_{j=1}^n \Phi_{(j:n)}(x) \right. \\
 &\quad \left. - \sum_{j=1}^k \Phi_{(j:n)}(x) - \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \right\} \\
 &= \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] + \frac{1}{n} \sum_{j=1}^n \Phi_{(j:n)}(x) - \frac{1}{n} \sum_{j=1}^k \Phi_{(j:n)}(x) \\
 &\quad - \frac{1}{n} \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \\
 &= \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] - \frac{1}{n} \left[ \sum_{j=1}^k \Phi_{(j:n)}(x) + \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \right] \\
 &\quad + \Phi(x)
 \end{aligned}$$

**Table 2** The efficiency and bias values of  $\Phi_{LRSS}(x)$  with respect to  $\Phi_{SRS}(x)$  for  $n = 6, 7$

$\Phi(x)$	$n = 6$			$n = 7$			
	$k = 0$	$k = 1$	$k = 2$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
0.01	1.0657	<b>9.5794</b>	<b>17.1181</b>	1.0245	<b>8.2039</b>	13.6128	14.0961
		-0.0095	-0.0100		-0.0094	-0.0100	-0.0100
0.10	1.4552	1.5523	1.5290	1.5550	1.5113	1.3597	1.3285
		-0.0581	-0.0915		-0.0525	-0.0890	-0.0972
0.20	1.8131	1.3572	0.9057	1.9769	1.4077	0.8957	0.7016
		-0.0655	-0.1425		-0.0537	-0.1302	-0.1672
0.30	2.0755	1.4886	0.8753	2.1832	1.6747	0.9323	0.6466
		-0.0489	-0.1372		-0.0349	-0.1197	-0.1728
0.40	2.1869	1.6872	0.9985	2.3471	1.9590	1.1684	0.8241
		-0.0247	-0.0834		-0.0172	-0.0693	-0.1100
0.50	<b>2.1955</b>	1.7772	1.1259	<b>2.3495</b>	2.1147	1.3362	1.0015
		0.0000	0.0000		0.0006	-0.0007	0.0011
0.60	2.1923	1.6615	1.0074	2.3442	1.9440	1.1890	0.8189
		0.0223	0.0815		0.0154	0.0687	0.1111
0.70	2.0350	1.4912	0.8793	2.2047	1.6898	0.9330	0.6473
		0.0479	0.1360		0.0339	0.1179	0.1734
0.80	1.8392	1.3512	0.9321	1.9479	1.4521	0.9075	0.6988
		0.0654	0.1422		0.0528	0.1292	0.1671
0.90	1.4676	1.5675	1.5284	1.5447	1.5144	1.3609	1.2894
		0.0586	0.0914		0.0532	0.0885	0.0971
0.99	1.0694	9.4590	16.1666	1.0375	7.9279	<b>13.7425</b>	<b>14.1188</b>
		0.0095	0.0100		0.0094	0.0100	0.0100

The proof of (7) is straightforward by assuming  $Bias [\Phi_{LRSS}(x)] = E [\Phi_{LRSS}(x)] - \Phi(x)$ .

Note that for fixed  $k$  as  $n \rightarrow \infty$ , the bias approaches zero while it does not depends on  $m$ . □

**Proposition 2** For fixed  $x$  and  $n$ , when  $m \rightarrow \infty$ , the random variable  $Z = \frac{\Phi_{LRSS}(x) - E[\Phi_{LRSS}(x)]}{\sqrt{Var[\Phi_{LRSS}(x)]}}$  converges in distribution to the standard normal distribution.

for

*Proof* Suppose that

$$\Psi_w = \frac{\sum_{j=1}^k I(X_{j(k+1:n)w} \leq x)}{n} + \frac{\sum_{j=k+1}^{n-k} I(X_{j(j:n)w} \leq x)}{n} + \frac{\sum_{j=n-k+1}^n I(X_{j(n-k:n)w} \leq x)}{n} \quad \text{for } w = 1, 2, \dots, m.$$

Then,

**Table 3** The efficiency and bias values of  $\Phi_{LRSS}(x)$  with respect to  $\Phi_{SRSS}(x)$  for  $n = 8$

$\Phi(x)$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
0.01	1.1051	6.8522	12.0100	12.2135
		-0.0093	-0.0100	-0.0100
0.10	1.6340	1.4694	1.2545	1.1497
		-0.0474	-0.0851	-0.0972
0.20	2.1307	1.5297	0.9210	0.6332
		-0.0425	-0.1146	-0.1663
0.30	2.3558	1.9222	1.0751	0.6037
		-0.0245	-0.0961	-0.1746
0.40	2.5092	2.2288	1.4134	0.8392
		-0.0099	-0.0520	-0.1092
0.50	<b>2.5451</b>	2.3672	1.6671	1.0573
		-0.0003	-0.0003	-0.0005
0.60	2.4667	2.2404	1.4287	0.8228
		0.0106	0.0526	0.1102
0.70	2.3147	1.8858	1.0556	0.6160
		0.0246	0.0956	0.1734
0.80	1.6651	1.4532	1.2480	1.1420
		0.0482	0.0851	0.0974
0.90	1.6651	1.4532	1.2480	1.1420
		0.0482	0.0851	0.0974
0.99	1.0319	<b>7.1302</b>	<b>12.2858</b>	<b>12.3700</b>
		0.0093	0.0100	0.0100

$$E(\Psi_w) = \frac{k}{n} [\Phi_{(k+1:n)}(x) + \Phi_{(n-k:n)}(x)] - \frac{1}{n} \left[ \sum_{j=1}^k \Phi_{(j:n)}(x) + \sum_{j=n-k+1}^n \Phi_{(j:n)}(x) \right] + \Phi(x) < \infty,$$

$$V(\Psi_w) = \text{Var}[\Phi_{LRSS}(x)] = \frac{k}{n^2} \{ \Phi_{(k+1:n)}(x) [1 - \Phi_{(k+1:n)}(x)] + \Phi_{(n-k:n)}(x) [1 - \Phi_{(n-k:n)}(x)] \} + \frac{1}{n^2} \sum_{j=1}^{n-2k} \Phi_{(j:n)}(x) [1 - \Phi_{(j:n)}(x)] < \infty,$$

where  $P(X_{(j:n)} \leq x) = \sum_{i=j}^n \binom{n}{i} F^i(x)(1 - F(x))^{n-i}$ .

Since the random variables  $\Psi_w$ 's are independent and identically (IID), then by the central limit theorem (CLT) we have  $\frac{\sqrt{m}[\Psi_w - E(\Psi_w)]}{\sqrt{\text{Var}(\Psi_w)}}$  converges in distribution to  $N(0, 1)$ .  $\square$

**Table 4** The efficiency and bias values of  $\Phi_{LRSS}(x)$  with respect to  $\Phi_{SRSS}(x)$  for  $n = 9, 10$

$\Phi(x)$	$n = 9$					$n = 10$				
	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
0.01	1.0840	6.0551	<b>10.8584</b>	11.0127	11.0339	1.0654	5.7599	9.4327	9.4800	9.8053
		-0.0092	-0.0100	-0.0100	-0.0100		-0.0091	-0.0100	-0.0100	-0.0100
0.10	1.7226	1.4238	1.1643	1.0379	1.0057	1.7869	1.4550	1.0988	0.9295	0.9048
		-0.0434	-0.0815	-0.0962	-0.0991		-0.0388	-0.0777	-0.0946	-0.0991
0.20	2.2090	1.6550	0.9576	0.6087	0.5209	2.2669	1.7416	1.0383	0.6068	0.4677
		-0.0327	-0.0998	-0.1584	-0.1800		-0.0264	-0.0872	-0.1471	-0.1804
0.30	2.5110	2.0728	1.2213	0.6327	0.4639	2.5823	2.2636	1.4302	0.7089	0.4282
		-0.0188	-0.0765	-0.1584	-0.2003		-0.0123	-0.0590	-0.1351	-0.2007
0.40	2.6689	2.4513	1.7531	0.9456	0.6836	2.7961	2.6065	2.0422	1.1503	0.6610
		-0.0073	-0.0370	-0.0962	-0.1305		-0.0046	-0.0257	-0.0771	-0.1331
0.50	<b>2.6884</b>	2.5668	2.0157	1.3054	1.0222	2.7913	2.7719	2.3278	1.5347	1.0572
		-0.0002	-0.0003	0.0000	0.0003		0.0014	-0.0005	0.0007	0.0004
0.60	2.6556	2.4335	1.7125	0.9168	0.6672	<b>2.8013</b>	2.6995	2.0470	1.1024	0.6495
		0.0064	0.0367	0.0963	0.1349		0.0037	0.0263	0.0008	0.1333
0.70	2.5213	2.0923	1.2472	0.6276	0.4624	2.5530	2.2852	1.4203	0.6989	0.4247
		0.0171	0.0747	0.1572	0.2017		0.0126	0.0585	0.1371	0.2008
0.80	2.1977	1.6446	0.9595	0.6067	0.5070	2.3049	1.7716	1.0350	0.6098	0.4668
		0.0332	0.1011	0.1587	0.1805		0.0276	0.0861	0.1469	0.1800
0.90	1.7576	1.4419	1.1700	1.0480	1.0100	1.7972	1.4476	1.1282	0.9411	0.8967
		0.0433	0.0817	0.0961	0.0990		0.0384	0.0775	0.0946	0.0990
0.99	1.1494	<b>6.4796</b>	10.4116	<b>11.0226</b>	<b>11.3417</b>	1.0982	<b>6.0413</b>	<b>9.5039</b>	<b>9.8651</b>	<b>9.9737</b>
		0.0092	0.0010	0.0100	0.0100		0.0092	0.0100	0.0100	0.0100

**Table 5** The efficiency of  $\Phi_{RSS}(x)$  with respect to  $\Phi_{SRSS}(x)$  for  $n = 4, 5, \dots, 10$

$\Phi(x)$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
0.01	1.0286	1.0451	1.0388	1.0644	1.1355	1.0712	1.1018
0.10	1.3071	1.3798	1.4832	1.5550	1.6539	1.7035	1.7929
0.20	1.5471	1.6973	1.8193	1.9610	2.0421	2.2084	2.2984
0.30	1.6880	1.8933	2.0402	2.2060	2.3306	2.4760	2.6122
0.40	1.7891	1.9969	2.2099	2.2982	2.4696	2.6347	2.7623
0.50	1.8887	2.0644	2.2449	2.3500	2.5601	2.6801	2.8271
0.60	1.7841	1.9545	2.2061	2.3337	2.4787	2.6349	2.7775
0.70	1.6816	1.9025	2.0456	2.1952	2.3528	2.4917	2.6052
0.80	1.5507	1.6731	1.8057	1.9757	2.0610	2.1988	2.3004
0.90	1.2972	1.3743	1.4716	1.5503	1.6241	1.7450	1.8179
0.99	1.1206	1.0146	1.0458	1.0781	1.0523	1.1562	1.0764



**Table 6** The best values of  $k$  for each sample size which give the large efficiency using LRSS with respect to SRS

	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$k$	1	2	2	3	3	4	4

**Table 7** The best values of the sample size for each  $\Phi(x)$  using LRSS

$\Phi(x)$	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
$n$	5	4	10	10	10	10	10	10	10	4	5

### 3 Numerical Comparisons

In this section, a simulation study is conducted to investigate the efficiency of LRSS method in estimating the distribution function  $\Phi(x)$ . The suggested estimator is compared with its competitors using RSS and SRS for samples of sizes,  $n = 4, 5, 6, \dots, 10$ . The efficiency of  $\Phi_{RSS}(x)$  and  $\Phi_{LRSS}(x)$  with respect to  $\Phi_{SRS}(x)$ , respectively are defined as

$$Eff [\Phi_{RSS}(x), \Phi_{SRS}(x)] = \frac{Var [\Phi_{SRS}(x)]}{Var [\Phi_{RSS}(x)]} = \frac{n\Phi(x) [1 - \Phi(x)]}{\sum_{j=1}^n \Phi_{(j:n)}(x) [1 - \Phi_{(j:n)}(x)]}, \tag{8}$$

and

$$Eff [\Phi_{LRSS}(x), \Phi_{SRS}(x)] = \frac{Var [\Phi_{SRS}(x)]}{MSE [\Phi_{LRSS}(x)]} = \frac{n\Phi(x) [1 - \Phi(x)]}{k \{ \Phi_{(k+1:n)}(x) [1 - \Phi_{(k+1:n)}(x)] + \Phi_{(n-k:n)}(x) [1 - \Phi_{(n-k:n)}(x)] \} + \sum_{j=1}^{n-2k} \Phi_{(j:n)}(x) [1 - \Phi_{(j:n)}(x)] + [Bias(\Phi_{LRSS}(x))]^2} \tag{9}$$

The bold fonts in the Tables 1, 2, 3 and 4 are the best values of the efficiency using LRSS relative to SRS with fixed  $k$  for estimating the distribution function  $\Phi(x)$ . It can be seen that as  $\Phi(x)$  goes to zero (0.01) or 1 (0.99), the efficiency values are the largest. As an example, for  $n = 10$  and  $k = 2$ , the efficiency values are 9.4327 and 9.5039 for  $\Phi(x) = 0.01$  and 0.99, respectively. But for other values when,  $0.10 \leq \Phi(x) \leq 0.90$ , the efficiency values are increases when  $\Phi(x)$  is approximately 0.5.

For the RSS method as it is shown in Table 5, [13] showed that the  $\Phi_{RSS}(x)$  is more efficient than  $\Phi_{SRS}(x)$ . But  $\Phi_{LRSS}(x)$  is more efficient than  $\Phi_{RSS}(x)$  for most of cases considered in this study. See as an example, when  $n = 8$  and  $\Phi(x) = 0.99$ , the efficiency values of RSS and LRSS when  $k = 2$ , are 1.0523 and 12.3858, respectively.

From Table 6, we can select the value of  $k$  with each sample size to get the best value of the efficiency.

Table 7 involved the suitable values of the sample size for each value of  $\Phi(x)$ . However, we can note that the best values of the sample size are  $n = 5$  for  $\Phi(x) = 0.01, 0.99$ ,  $n = 4$  for  $\Phi(x) = 0.10, 0.90$ , and  $n = 10$  for  $0.2 \leq \Phi(x) \leq 0.8$ .

## 4 Conclusion

A new estimator of the distribution function is suggested using L ranked set sampling method. The suggested estimator is compared with the SRS and RSS estimators of the distribution function. It is found that LRSS estimator is biased and it is more efficient than the SRS and RSS for most of cases considered in this study. It is recommended to use the LRSS for estimating the distribution of a continuous random variable.

**Acknowledgments** The author is thankful to the editor and the anonymous reviewers for their valuable comments and suggestions that significantly improved the original version of the article.

## References

1. Al-Nasser, A.D.: L-Ranked set sampling: a generalization procedure for robust visual sampling. *Communications in statistics. Simul. Comput.* **6**, 33–43 (2007)
2. Al-Nasser, A.D., Radaideh, A.: Estimation of simple linear regression model using L ranked set sampling. *Int. J. Open Probl. Comput. Sci. Math.* **1**(1), 18–33 (2008)
3. Al-Omari, A.I.: Estimation of mean based on modified robust extreme ranked set sampling. *J. Stat. Comput. Simul.* **81**(8), 1055–1066 (2011)
4. Al-Omari, A.I.: Ratio estimation of population mean using auxiliary information in simple random sampling and median ranked set sampling. *Stat. Probab. Lett.* **82**(11), 1883–1990 (2012)
5. Al-Omari, A.I., Al-Hadhrami, S.A.: On maximum likelihood estimators of the parameters of a modified Weibull distribution using extreme ranked set sampling. *J. Mod. Appl. Stat. Methods* **10**(2), 607–617 (2011)
6. Al-Saleh, M.F., Al-Omari, A.I.: Multistage ranked set sampling. *J. Stat. Plan. Inference* **102**(2), 273–286 (2002)
7. Bahadur, R.R.: A note on quantiles in large samples. *Ann Math Stat* **37**(3), 577–580 (1966)
8. David, H.A., Nagaraja, H.N.: *Order Stat.*, 3rd edn. Wiley, New Jersey (2003)
9. Kim, D.H., Kim, D.W., Kim, G.H.: On the estimation of the distribution function using extreme median ranked set sampling. *J. Korean Data Anal. Soc.* **7**(2), 429–439 (2005)
10. McIntyre, G.A.: A method for unbiased selective sampling using ranked sets. *Aust. J. Agric. Res.* **3**, 385–390 (1952)
11. Samawi, H., Al-Sageer, O.M.: On the estimation of the distribution function using extreme and median ranked set sampling. *Biom. J.* **43**(3), 357–373 (2001)
12. Samawi, H.M., Mohammad, S., Abu-Dayyeh, W.: Estimating the population mean using extreme ranked set sampling. *Biometrical J* **38**(5), 577–586 (1996)
13. Stokes, S.L., Sager, T.W.: Characterization of a ranked set sample with application to estimating distribution functions. *J Am Stat Assoc* **83**(402), 374–381 (1988)
14. Takahasi, K., Wakimoto, K.: On the unbiased estimates of the population mean based on the sample stratified by means of ordering. *Ann. Inst. Stat. Math.* **20**, 1–31 (1968)