Journal of the Egyptian Mathematical Society (2016) **24**, 303–308

Egyptian Mathematical Society

Journal of the Egyptian Mathematical Society

www.etms-eg.org
www.elsevier.com/locate/joems

Original Article

Quartile ranked set sampling for estimating the distribution function

Amer Ibrahim Al-Omari

Al al-Bayt University, Faculty of Science, Department of Mathematics, Mafraq, Jordan

Received 24 January 2014; revised 5 July 2014; accepted 15 April 2015
Available online 21 May 2015

Keywords

Distribution function;
Quartile ranked set sampling;
Simple random sampling;
Ranked set sampling;
Relative efficiency

Abstract Quartile ranked set sampling (QRSS) method is suggested by Muttlak (2003) for estimating the population mean. In this article, the QRSS procedure is considered to estimate the distribution function of a random variable. The proposed QRSS estimator is compared with its counterparts based on simple random sampling (SRS) and ranked set sampling (RSS) schemes. It is found that the suggested estimator of the distribution function of a random variable X for a given x is biased and more efficient than its competitors using SRS and RSS.

2010 Mathematics Subject Classification: 62G30; 62D05; 62G05; 62G07

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n having probability density function (pdf) $f(x)$ and cumulative distribution function (cdf) $F(x)$, with a finite mean μ and variance σ^2. Let $X_{11i}, X_{12i}, \ldots, X_{1ni}; X_{21i}, X_{22i}, \ldots, X_{2ni}; \ldots; X_{n1i}, X_{n2i}, \ldots, X_{nni}$ be n independent simple random samples each of size n in the ith cycle $(i = 1, 2, \ldots, m)$.

Let $F_{SRS}(x)$ denote the empirical distribution function of a simple random sample $X_1, X_2, \ldots, X_{nm}$ from $F(x)$. Bahadur [2] showed that $F_{SRS}(x)$ has the following properties:

1. $F_{SRS}(x)$ is an unbiased estimator of $F(x)$ for a given x.
2. $Var[F_{SRS}(x)] = \frac{1}{mn} F(x)[1 - F(x)]$.
3. $F_{SRS}(x)$ is a consistent estimator of $F(x)$.

The RSS was first suggested by McIntyre [3] as a method for estimating the mean of pasture and forage yields. The RSS is a useful method when the sampling units can be easily ranked than quantified. McIntyre proposed the ranked set sample mean as an estimator of the population mean and showed that the RSS mean estimator is unbiased and is more efficient than the SRS counterpart.

The RSS can be described as follows: randomly select n sets each of size n from the target population. Then, visually rank the units within each sample with respect to the variable of interest. From the first set of n units the smallest ranked unit is selected. From the second set of n units the second smallest ranked unit is selected. The process is continued until the largest ranked unit is measured from the nth set. To increase the

E-mail address: alomari_amer@yahoo.com
Peer review under responsibility of Egyptian Mathematical Society.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.joems.2015.01.006

Table 1 The relative precision of $F_{QRSS}(x)$ with respect to $F_{SRS}(x)$ and bias values of $F_{QRSS}(x)$ for $4 \le n \le 11$.

$F(x)$		$n=4$	$n=6$	$n=8$	$n=10$	$n=5$	$n=7$	$n=9$	$n=11$
0.01	RP	0.5270	7.9487	5.3923	**9.1690**	**11.0248**	7.0763	4.5962	8.2674
	Bias	0.0090	−0.0093	−0.0088	−0.0099	−0.0096	−0.0092	−0.0084	−0.0099
0.10	RP	0.6852	**1.4566**	1.1587	1.2156	**1.7024**	1.4095	1.1780	1.1780
	Bias	0.0718	−0.0429	−0.0066	−0.0650	−0.0655	−0.0356	0.0006	−0.0592
0.20	RP	0.9939	**1.3260**	1.0919	1.2896	1.3138	**1.4388**	1.1553	1.3873
	Bias	0.0973	−0.0254	0.0482	−0.0396	−0.0802	−0.0142	0.0527	−0.0257
0.30	RP	1.6994	1.6332	1.5037	**1.7397**	1.2961	1.9075	1.7570	**1.9325**
	Bias	0.0849	−0.0044	0.0727	0.0098	−0.0673	0.0062	0.0679	0.0201
0.40	RP	2.9902	2.2115	**3.3200**	2.9599	1.3908	2.5254	**3.3480**	3.1859
	Bias	0.0475	0.0041	0.0505	0.0224	−0.0367	0.0103	0.0447	0.0265
0.50	RP	4.3694	2.5640	**7.3573**	4.8399	1.4534	2.8914	**5.5273**	4.8417
	Bias	0.0005	0.0008	0.0000	−0.0005	0.0003	0.0005	0.0004	−0.0003
0.60	RP	3.0824	2.2121	**3.2736**	2.9261	1.3843	2.5467	**3.3917**	3.2440
	Bias	−0.0481	−0.0040	−0.0508	−0.0216	0.0354	−0.0105	−0.0443	−0.0256
0.70	RP	1.7065	1.6410	1.5109	**1.7521**	1.2757	1.8497	1.7220	**1.9539**
	Bias	−0.0848	0.0051	−0.0733	−0.0088	0.0662	−0.0067	−0.0686	−0.0194
0.80	RP	1.0080	**1.3748**	1.1220	1.2898	1.2997	**1.4229**	1.1603	1.3658
	Bias	−0.0960	0.0277	−0.0483	0.0391	0.0812	0.0136	−0.0536	0.0256
0.90	RP	0.6628	**1.4455**	1.1659	1.2054	**1.6756**	1.4006	1.1633	1.1870
	Bias	−0.0736	0.0428	0.0063	0.0650	0.0659	0.0352	0.0004	0.0592
0.99	RP	0.5126	7.3353	4.9833	**9.4109**	**11.8336**	7.3257	4.4370	8.6399
	Bias	−0.0098	0.0092	0.0086	0.0099	0.0096	0.0093	0.0084	0.0099

sample size, the whole process can be repeated m times to obtain a set of size nm units.

Let $X_{j(1:n)i}, X_{j(2:n)i}, \ldots, X_{j(n:n)i}$ be the order statistics of the jth sample $X_{j1i}, X_{j2i}, \ldots, X_{jni}$ $(j = 1, 2, \ldots, n)$ in the ith cycle $(i = 1, 2, \ldots, m)$. Then, the measured units $X_{1(1:n)i}, X_{2(2:n)i}, \ldots, X_{n(n:n)i}$ are denoted to the RSS. David and Nagaraja [4] showed that the cdf and the pdf of the jth order statistic $X_{(j:n)}$ are given by

$$F_{(j:n)}(x) = \sum_{i=j}^{n} \binom{n}{i} [F(x)]^{i}[1 - F(x)]^{n-i}, -\infty < x < \infty,$$

and

$$f_{(j:n)}(x) = \frac{n!}{(j-1)!(n-j)!}[F(x)]^{j-1}[1 - F(x)]^{n-j} f(x).$$

The mean and the variance of $X_{(j:n)}$ are given by $\mu_{(j:n)} = \int_{-\infty}^{\infty} x f_{(j:n)}(x)dx$ and $\sigma^2_{(j:n)} = \int_{-\infty}^{\infty} (x - \mu_{(j:n)})^2 f_{(j:n)}(x)dx$, respectively. Takahasi and Wakimoto [5] independently introduced the same method of RSS with mathematical theory and showed that

$$f(x) = \frac{1}{n}\sum_{j=1}^{n} f_{(j:n)}(x), \quad \mu = \frac{1}{n}\sum_{j=1}^{n}\mu_{(j:n)}, \quad \text{and}$$

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

For a fixed x, Stokes and Sager [6] suggested an estimator for $F(x)$ using RSS as

$$F_{RSS}(x) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} I\left(X_{j(j:n)i} \le x\right),$$

where $I(\cdot)$ is an indicator function. They proved the following:

1. $F_{RSS}(x)$ is an unbiased estimator for $F(x)$.
2. $Var[F_{RSS}(x)] = \frac{1}{mn^2}\sum_{j=1}^{n} F_{(j:n)}(x)[1 - F_{(j:n)}(x)]$.
3. $\frac{F_{RSS}(x) - E[F_{RSS}(x)]}{\sqrt{Var[F_{RSS}(x)]}}$ converges in distribution to the standard normal as $m \to \infty$ when x and n are fixed.

For more about estimation of the distribution function in ranked set sampling methods see Stokes and Sager [6], Samawi and Al-Sagheer [7], Kim and Kim [8], Al-Saleh and Samuh [9], and Ghosh and Tiwari [10].

The rest of this paper is organized as follows. In Section 2, we introduced the suggested estimation of the distribution function using QRSS method. The performance of the new estimator against its SRS and RSS counterparts is given in Section 3. Section 4, is devoted for some inferences about $F(x)$. In Section 5, some concluding remarks are provided.

2. Estimation of $F(x)$ using QRSS

The quartile ranked set sampling procedure as suggested by Muttlak [1] can be summarized as follows. Randomly select n samples each of size n units from the target population and rank the units within each sample with respect to the variable of interest. If the sample size n is even, select and measure from the first $n/2$ samples the $Q_1(n+1)$th smallest ranked unit of each sample, i.e., the first quartile, and from the second $n/2$ samples the $Q_3(n+1)$th smallest ranked unit of each sample, i.e., the third quartile. Note that, we always take the nearest integer of $Q_1(n+1)$th and $Q_3(n+1)$th where $Q_1 = 25\%$, and $Q_3 = 75\%$. If the sample size n is odd, select and measure from the first $(n-1)/2$ samples the $Q_1(n+1)$th smallest ranked unit of each sample and from the other $(n-1)/2$ samples the $Q_3(n+1)$th smallest ranked unit of each sample, and from one sample the median for that sample. The cycle can be repeated m times if needed to get a sample of size nm units.

If the sample size n is even, in the ith cycle $(i = 1, 2, \ldots, m)$, let $X_{j(Q_1(n+1):n)i}$ be the $(Q_1(n+1))$th smallest ranked unit of the jth sample $(j = 1, 2, \ldots, \frac{n}{2})$, and $X_{j(Q_3(n+1):n)i}$ be the $(Q_3(n+1))$th smallest ranked unit of the jth sample

$(j = \frac{n+2}{2}, \frac{n+4}{2}, \dots, n)$. Note that, the measured units $X_{1(Q_1(n+1):n)i}, X_{2(Q_1(n+1):n)i}, \dots, X_{\frac{n}{2}(Q_1(n+1):n)i}$ are independent and identically distributed (iid) random variables, and also $X_{\frac{n+2}{2}(Q_3(n+1):n)i}, X_{\frac{n+4}{2}(Q_3(n+1):n)i}, \dots, X_{n(Q_3(n+1):n)i}$ are iid. However, all units $X_{1(Q_1(n+1):n)i}, X_{2(Q_1(n+1):n)i}, \dots, X_{\frac{n}{2}(Q_1(n+1):n)i}, X_{\frac{n+2}{2}(Q_3(n+1):n)i}, X_{\frac{n+4}{2}(Q_3(n+1):n)i}, \dots, X_{n(Q_3(n+1):n)i}$ are mutually independent but not identically distributed. These measured units are denoted by QRSSE.

For odd sample size, let $X_{j(Q_1(n+1):n)i}$ be the $Q_1(n+1)$th smallest ranked unit of the jth sample $(j = 1, 2, \dots, \frac{n-1}{2})$, $X_{j(\frac{n+1}{2}:n)i}$ be the median of the jth sample of the rank $j = \frac{n+1}{2}$, and $X_{j(Q_3(n+1):n)i}$ be the $Q_3(n+1)$th smallest ranked unit of the jth sample $(j = \frac{n+3}{2}, \frac{n+5}{2}, \dots, n)$. Note that, the only measured units $X_{1(Q_1(m+1):m)i}, X_{2(Q_1(n+1):n)i}, \dots, X_{\frac{n-1}{2}(Q_1(n+1):n)i}$ are iid, and also $X_{\frac{n+3}{2}(Q_3(n+1):n)i}, X_{\frac{n+5}{2}(Q_3(n+1):n)i}, \dots, X_{n(q_3(n+1):n)i}$ are iid. However, all units $X_{1(Q_1(m+1):m)i}, X_{2(Q_1(n+1):n)i}, \dots, X_{\frac{n-1}{2}(Q_1(n+1):n)i}, X_{\frac{n+1}{2}(\frac{n+1}{2}:n)i}, X_{\frac{n+3}{2}(Q_3(n+1):n)i}, X_{\frac{n+5}{2}(Q_3(n+1):n)i}, \dots, X_{n(q_3(n+1):n)i}$ are mutually independent but not identically distributed. These measured units are denoted by QRSSO.

Following Stokes and Sager [6], let $V' = (V_1, V_2, \dots, V_n)$ be a multinomial random vector with nm trials and $P = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ be a probability vector, and assume that the nm random variables were found by first observing V and then selecting Vj units randomly from a population with pdf $f_{(j:n)}(x)$ $(j = 1, 2, \dots, n)$. Let $Y_1, Y_2, \dots, Y_{nm}$ denote to the obtained mn units.

Also, following Stokes and Sager [6] we have the following theorem based on QRSS method.

Theorem 1. *Based on the same conditions of Theorem 1 in Stokes and Sager [6] we have the following:*

(1) *For even sample size,* $\{Y_1, Y_2, \dots, Y_{nm} | V = (\frac{nm}{2}, 0, 0, \dots, 0, \frac{nm}{2})\}$ *has the same probability structure as* $\{X_{j(Q_1(n+1):n)i}, X_{l(Q_3(n+1):n)i};\ j = 1, 2, \dots, \frac{n}{2};\ l = \frac{n+2}{2}, \frac{n+4}{2}, \dots, n;\ i = 1, 2, \dots, m\}$, *a QRSSE from the same population.*
(2) *For odd sample size,* $\{Y_1, Y_2, \dots, Y_{nm} | V = v_1 = \frac{(n-1)m}{2}, 0, 0, \dots, 0; v_{(\frac{n+1}{2})} = m; 0, 0, \dots, 0, v_n = \frac{(n-1)m}{2}\}$,

has the same probability structure as

$$\left\{X_{j(Q_1(n+1):n)i}, X_{t(t:n)i} X_{l(Q_3(n+1):n)i};\ j = 1, 2, \dots, \frac{n-1}{2};\ t \frac{n+1}{2} :\right.$$
$$\left. l = \frac{n+3}{2}, \frac{n+5}{2}, \dots, n;\ i = 1, 2, \dots, m\right\},$$

a QRSSO from the same population.

Proof. The proof is directly and similar to the proof of Theorem 1 in Stokes and Sager [6]. □

The suggested QRSSE estimator of $F(x)$ is given by

$$F_{QRSSE}(x) = \frac{1}{nm}\sum_{i=1}^{m}\sum_{j=1}^{\frac{n}{2}} I\left(X_{j(Q_1(n+1):n)i} \le x\right) \frac{1}{nm}\sum_{i=1}^{m}\sum_{j=\frac{n+2}{2}}^{n} I\left(X_{j(Q_3(n+1):n)i} \le x\right), \tag{1}$$

with variance

$$Var\left[F_{QRSSE}(x)\right] = \frac{1}{2nm}\left\{F_{Q_1}(x)\left[1 - F_{Q_1}(x)\right] + F_{Q_3}(x)\left[1 - F_{Q_3}(x)\right]\right\}. \tag{2}$$

And the suggested QRSSO estimator of $F(x)$ is given by

$$F_{QRSSO}(x) = \frac{1}{nm}\sum_{i=1}^{m}\sum_{j=1}^{\frac{n-1}{2}} I\left(X_{j(Q_1(n+1):n)i} \le x\right) + \frac{1}{nm} I\left(X_{\frac{n+1}{2}(\frac{n+1}{2}:n)i} \le x\right)$$
$$+ \frac{1}{nm}\sum_{i=1}^{m}\sum_{j=\frac{n+3}{2}}^{n} I\left(X_{j(Q_3(n+1):n)i} \le x\right), \tag{3}$$

with variance

$$Var\left[F_{QRSSO}(x)\right] = \frac{1}{n^2 m}\left\{\frac{n-1}{2}\left[F_{Q_1}(x)\left(1 - F_{Q_1}(x)\right) + F_{Q_3}(x)\left(1 - F_{Q_3}(x)\right)\right] + F_{Q_2}(x)\left(1 - F_{Q_2}(x)\right)\right\}, \tag{4}$$

where

$$F_{Q_1}(x) = B\left(F(x); \frac{n-1}{4}, \frac{3n-3}{4}\right),$$
$$F_{Q_2}(x) = B\left(F(x); \frac{n+1}{2}, \frac{n+1}{2}\right),$$
$$F_{Q_3}(x) = B\left(F(x); \frac{3n-3}{4}, \frac{n-1}{4}\right),$$

and $B(w; a, b)$ is the incomplete beta function defined as $B(w; a, b) = \int_0^w u^{a-1}(1-u)^{b-1} du$.

Assume that n is odd, and $\frac{n+1}{2}$ is odd, and let $Q_1 = X_{(\frac{n+3}{4})}$, $Q_2 = X_{(\frac{n+1}{2})}$, and $Q_3 = X_{(\frac{3n+1}{4})}$, then

$$f_{Q_1}(x) = \frac{1}{B\left(\frac{n-1}{4}, \frac{3n-3}{4}\right)}\left[F(x)\right]^{\frac{n-1}{4}}\left[1 - F(x)\right]^{\frac{3n-3}{4}} f(x), \tag{5}$$

$$f_{Q_2}(x) = \frac{1}{B\left(\frac{n+1}{2}, \frac{n+1}{2}\right)}\left[F(x)\left(1 - F(x)\right)\right]^{\frac{n-1}{2}} f(x), \tag{6}$$

and

$$f_{Q_3}(x) = \frac{1}{B\left(\frac{3n-3}{4}, \frac{n-1}{4}\right)}\left[F(x)\right]^{\frac{3n-3}{4}}\left[1 - F(x)\right]^{\frac{n-1}{4}} f(x), \tag{7}$$

where $B(\delta, \vartheta)$ is the Beta function defined as

$$B(\delta, \vartheta) = \frac{\Gamma(\delta)\Gamma(\vartheta)}{\Gamma(\delta + \vartheta)} = \frac{(\delta - 1)!(\vartheta - 1)!}{(\delta + \vartheta - 1)!}. \tag{8}$$

Some properties of the QRSSE and QRSSO estimators of the distribution function are provided in the following propositions.

Proposition 1.

(1) $E[F_{QRSSE}(x)] = \frac{1}{2}[F_{Q_1}(x) + F_{Q_3}(x)]$.
(2) $E[F_{QRSSO}(x)] = \frac{n-1}{2n}\{F_{Q_1}(x)[1 - F_{Q_1}(x)] + F_{Q_3}(x)[1 - F_{Q_3}(x)]\} + \frac{1}{n}F_{Q_2}(x)$, *where* $F_{Q_1}(x), F_{Q_2}(x), F_{Q_3}(x)$, *are defined above.*

Proposition 2. *$F_{QRSSO}(x)$ and $F_{QRSSE}(x)$ are biased estimators of $F(x)$, with bias given by*

$$Bias\big[F_{QRSSE}(x)\big] = \frac{1}{2}\big[F_{Q_1}(x) + F_{Q_3}(x)\big] - F_{QRSSE}(x),$$

and

$$Bias\big[F_{QRSSO}(x)\big] = \frac{n-1}{2n}\{F_{Q_1}(x)[1 - F_{Q_1}(x)] + F_{Q_3}(x)[1 - F_{Q_3}(x)]\} + \frac{1}{n}F_{Q_2}(x) - F_{QRSSO}(x).$$

It is easy to prove Propositions 1 and 2 based on some algebra. From this proposition we can see that the suggested estimators are biased, but based on the calculations given in Section 3, the bias is close to zero.

Proposition 3. *For fixed n and $m \to \infty$, we have the following:*

1. $\frac{F_{QRSSE}(x) - E[F_{QRSSE}(x)]}{\sqrt{Var[F_{QRSSE}(x)]}}$ *converges in distribution to* $N(0, 1)$.
2. $\frac{F_{QRSSO}(x) - E[F_{QRSSO}(x)]}{\sqrt{Var[F_{QRSSO}(x)]}}$ *converges in distribution to* $N(0, 1)$.

Proof. Following Samawi and Al-Sagheer [7] and Kim and Kim [8], the proof of the first part can be done by assuming

$$Z_i = \frac{1}{n}\sum_{j=1}^{\frac{n}{2}} I\big(X_{j(Q_1(n+1):n)i} \le x\big) + \frac{1}{n}\sum_{j=\frac{n+2}{2}}^{n} I\big(X_{j(Q_3(n+1):n)i} \le x\big),$$

$i = 1, 2, \ldots, m.$

As Z_i's are independent and identically with finite mean and variance random variables, then based on the central limit theorem we can conclude $\frac{\sqrt{m}[\bar{Z}_i - E(Z_i)]}{\sqrt{Var(Z_i)}}$ converges in distribution to the standard normal distribution.

The second part can be proved similarly by assuming

$$Z_i = \frac{1}{n}\sum_{j=1}^{\frac{n-1}{2}} I\big(X_{j(Q_1(n+1):n)i} \le x\big) + \frac{1}{n} I\left(X_{\frac{n+1}{2}\left(\frac{n+1}{2}:n\right)i} \le x\right) + \frac{1}{n}\sum_{j=\frac{n+3}{2}}^{n} I\big(X_{j(Q_3(n+1):n)i} \le x\big). \qquad \square$$

3. Numerical comparisons

Numerical comparisons are considered in this section to investigate the performance of the suggested QRSS estimator of the distribution function with respect to SRS and RSS estimators based on the same number of measured units. We considered different sample sizes $4 \le n \le 11$. The relative precisions (RP) of $F_{RSS}(x)$ and $F_{QRSS}(x)$ with respect to $F_{SRS}(x)$ are defined as:

$$RP[F_{RSS}(x), F_{SRS}(x)] = \frac{Var[F_{SRS}(x)]}{Var[F_{RSS}(x)]}, \text{ and}$$

$$RP\big[F_{QRSS}(x), F_{SRS}(x)\big] = \frac{Var[F_{SRS}(x)]}{MSE\big[F_{QRSS}(x)\big]}, \tag{9}$$

where MSE is the mean squared error of an estimator.

The results are presented in Tables 1 and 2 using QRSS and RSS, respectively. The results in bold fonts in both tables are the best RP values of the distribution function estimators QRSS and RSS with respect to SRS for different values of $F(x)$. Based on QRSS from Table 1, the largest RP values are observed as $F(x)$ goes to zero or 1. For example, with $n = 5$ and $F(x) = 0.1, 0.99$, the RP values are 11.0248 and 11.8336, respectively. Otherwise, the RP values increase when $F(x)$ is close to 0.5. In all cases the bias is negligible and close to zero.

From Table 2 we observe that $F_{RSS}(x)$ is more efficient than $F_{SRS}(x)$. However, $F_{QRSS}(x)$ is better than $F_{RSS}(x)$ for most of cases considered in this study. As an example, when $n = 9$ and $F(x) = 0.4$, the relative precisions of RSS and QRSSO are 2.6347 and 3.3480, respectively.

The optimal values of the sample size using QRSSE and QRSSO methods for estimating $F(x)$ are summarized in Table 3.

From Table 3 it can be noted that QRSSO is more efficient than QRSSE for estimating $F(x)$. Also, the sample size $n = 5$ has most occurrence among other sample sizes considered in this study.

4. Inferences about the distribution function

In this section, a pointwise estimate of $F(x)$ is considered and some inferences about the population distribution are presented using QRSS. As m gets large, then based on Proposition 3, an approximate $100(1-\alpha)\%$ confidence interval for $E[F_\Sigma(x)]$, $\Sigma = QRSSE, QRSSO, RSS, SRS$, is given by

$$F_\Sigma(x) - Z_{\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]} < E[F_\Sigma(x)] < F_\Sigma(x) + Z_{\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]}, \tag{10}$$

where $Z_{\alpha/2}$ is the upper $100(\alpha/2)\%$ quantile of the $N(0, 1)$, and

Table 2 The relative precision of $F_{RSS}(x)$ with respect to $F_{SRS}(x)$ for $4 \le n \le 11$.

$F(x)$	$n = 4$	$n = 6$	$n = 8$	$n = 10$	$n = 5$	$n = 7$	$n = 9$	$n = 11$
0.01	0.9886	1.0388	**1.1355**	1.1018	1.0851	1.0644	1.0712	1.0749
0.10	1.3071	1.4832	1.6539	1.7929	1.3798	1.5550	1.7035	**1.9029**
0.20	1.5471	1.8193	2.0421	2.2984	1.6973	1.9610	2.2084	**2.3711**
0.30	1.6880	2.0402	2.3306	2.6122	1.8933	2.2060	2.4760	**2.7136**
0.40	1.7891	2.2099	2.4696	2.7623	1.9969	2.2982	2.6347	**2.8966**
0.50	1.8487	2.2449	2.5601	2.8271	2.0644	2.3500	2.6801	**2.9600**
0.60	1.7841	2.2061	2.4787	2.7775	1.9545	2.3337	2.6349	**2.9724**
0.70	1.6816	2.0456	2.3528	2.6052	1.9025	2.1952	2.4917	**2.7661**
0.80	1.5507	1.8057	2.0610	2.3004	1.6731	1.9757	2.1988	**2.4008**
0.90	1.2972	1.4716	1.6241	1.8179	1.3743	1.5503	1.7450	**1.8920**
0.99	1.1206	1.0458	1.0523	1.0764	1.0146	1.0781	**1.1562**	1.1301

Table 3 The best values of the sample size using QRSS for estimating $F(x)$.

$F(x)$	0.01	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.99
n	**5**	**5**	7	11	9	8	9	11	7	**5**	**5**

$$\widehat{Var}\left[F_{QRSSE}(x)\right] = \frac{1}{(m-1)n^2}\left\{\sum_{j=1}^{\frac{n}{2}} \hat{F}_{j(Q_1(n+1):n)}(x)\left[1-\hat{F}_{j(Q_1(n+1):n)}(x)\right] + \sum_{j=\frac{n+2}{2}}^{n} \hat{F}_{j(Q_3(n+1):n)}(x)\left[1-\hat{F}_{j(Q_3(n+1):n)}(x)\right]\right\} \quad (11)$$

$$\widehat{Var}\left[F_{QRSSO}(x)\right] = \frac{1}{(m-1)n^2}\left\{\begin{array}{l}\sum_{j=1}^{\frac{n-1}{2}} \hat{F}_{j(Q_1(n+1):n)}(x)\left[1-\hat{F}_{j(Q_1(n+1):n)}(x)\right] + \hat{F}_{Q_2}(x)\left[1-\hat{F}_{Q_2}(x)\right] \\ + \sum_{j=\frac{n+3}{2}}^{n} \hat{F}_{j(Q_3(n+1):n)}(x)\left[1-\hat{F}_{j(Q_3(n+1):n)}(x)\right]\end{array}\right\}, \quad (12)$$

$$Var[F_{RSS}(x)] = \frac{1}{mn^2}\sum_{j=1}^{n} F_{(j:n)}(x)\left[1-F_{(j:n)}(x)\right],$$

and

$$\widehat{Var}[F_{SRS}(x)] = \frac{1}{nm-1}\left[\hat{F}(x)\left(1-\hat{F}(x)\right)\right],$$

where based on QRSS, RSS, and SRS, respectively, we have

$$\hat{F}_{i(t:n)}(x) = \frac{1}{m}\sum_{i=1}^{m} I\left(X_{j(t:n)i} \le x\right),\ t = Q_1(n+1),\ Q_3(n+1),\ \frac{n+1}{2};\ j = 1, 2, \ldots, n, \quad (13)$$

$$\hat{F}_{i(j:n)}(x) = \frac{1}{m}\sum_{i=1}^{m} I\left(X_{j(j:n)i} \le x\right),\ j = 1, 2, \ldots, n, \quad (14)$$

and

$$\hat{F}_{SRS}(x) = \frac{1}{nm}\sum_{i=1}^{nm} I(Y_i \le x). \quad (15)$$

Based on the $100(1-\alpha)\%$ confidence interval of $E[F_\Sigma(x)]$, $\Sigma = QRSSE, QRSSO, RSS, SRS$, a $100(1-\alpha)\%$ confidence interval for $F(x)$ can be found as

$$P\left(Z_{\alpha/2} \le \frac{F_\Sigma(x) - E[F_\Sigma(x)]}{\sqrt{\widehat{Var}[F_\Sigma(x)]}} \le Z_{1-\alpha/2}\right) = 1-\alpha. \quad (16)$$

Solving (16) for $E[F_\Sigma(x)]$, $\Sigma = QRSSE, QRSSO, RSS, SRS$ to get

$$P\left(F_\Sigma(x) - Z_{1-\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]} \le E[F_\Sigma(x)] \le F_\Sigma(x) + Z_{\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]}\right) = 1-\alpha,$$

and the limits will be

$$Lower\ Bound(LB) = F_\Sigma(x) - Z_{1-\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]},$$

and

$$Upper\ Bound(UB) = F_\Sigma(x) + Z_{\alpha/2}\sqrt{\widehat{Var}[F_\Sigma(x)]}.$$

Finally, an approximate $100(1-\alpha)\%$ confidence for $F(x)$ can be obtained by solving the equations $2LB = h(F)$ and $2UB = h(F)$, numerically, or by any suitable method, where based on QRSSE we have

$$h(F) = F_{Q_1}(x) + F_{Q_3}(x),$$

and using QRSSO it will be

$$h(F) = \frac{1}{n}\left((n-1)\left\{F_{Q_1}(x)\left[1-F_{Q_1}(x)\right] + F_{Q_3}(x)\left[1-F_{Q_3}(x)\right]\right\} + 2F_{Q_2}(x)\right).$$

It is of interest to note here that any possible solution for last two equations should be singular since $h(F)$ is increasing function in $F(x)$.

5. Conclusion

In this study, QRSS procedure is considered to estimate the distribution function of a random variable. The QRSS is compared with RSS and SRS in estimating the distribution function. It is found that QRSS is more efficient than SRS and also it is more efficient than RSS in most cases considered in this study. The RP values are preferred as $F(x)$ goes to zero or 1 and increase when $F(x)$ is close to 0.5. The bias of the QRSS estimator of the distribution function is negligible.

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] H.A. Muttlak, Investigating the use of quartile ranked set samples for estimating the population mean, Appl. Math. Comput. 146 (2003) 437–443.
[2] R.R. Bahadur, A note on quantiles in large samples, Ann. Math. Stat. 37 (1996) 577–580.
[3] G.A. McIntyre, A method for unbiased selective sampling using ranked sets, Aust. J. Agric. Res. 3 (1952) 385–390.
[4] H.A. David, H.N. Nagaraja, Order Statistics, 3rd ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.
[5] K. Takahasi, K. Wakimoto, On the unbiased estimates of the population mean based on the sample stratified by means of ordering, Ann. Inst. Stat. Math. 20 (1968) 1–31.
[6] S.L. Stocks, T. Sager, Characterization of a ranked set sample with application to estimating distribution functions, J. Am. Stat. Assoc. 83 (1988) 374–381.

[7] H. Samawi, O.M. Al-Sageer, On the estimation of the distribution function using extreme and median ranked set sampling, Biometrical J. 43 (3) (2001) 357–373.
[8] D.H. Kim, D.W. Kim, G.H. Kim, On the estimation of the distribution function using extreme median ranked set sampling, JKDAS 7 (2) (2005) 429–439.
[9] M.F. Al-Saleh, M.H. Samuh, On multistage ranked set sampling for distribution and median estimation, Comput. Stat. Data Anal. 52 (4) (2008) 2066–2078.
[10] K. Ghosh, R.C. Tiwari, Estimating the distribution function using k-tuple ranked set samples, J. Stat. Plann. Infer. 138 (4) (2008) 929–949.