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

S1110-256X(15)00027-9
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 $4 \leq n \leq 11$.

$F(x)$		$n=4$	$n=6$	$n=8$	$n=10$	$n=5$	$n=7$	$n=9$	$n=11$
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

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

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