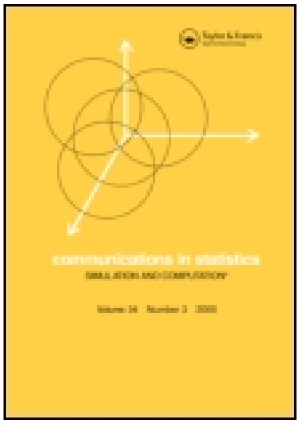
This article was downloaded by: [Memorial University of Newfoundland]
On: 11 September 2013, At: 05:56
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Communications in Statistics - Simulation and Computation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lssp20

Quantile Estimation Using Ranked Set Samples from a Population with Known Mean

M. Mahdizadeh [a] & N. R. Arghami [a]
[a] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran
Published online: 13 Jun 2012.

To cite this article: M. Mahdizadeh & N. R. Arghami (2012) Quantile Estimation Using Ranked Set Samples from a Population with Known Mean, Communications in Statistics - Simulation and Computation, 41:10, 1872-1881, DOI: 10.1080/03610918.2011.624236

To link to this article: http://dx.doi.org/10.1080/03610918.2011.624236

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Communications in Statistics—Simulation and Computation®, 41: 1872–1881, 2012
Copyright © Taylor & Francis Group, LLC
ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610918.2011.624236

Taylor & Francis
Taylor & Francis Group

Quantile Estimation Using Ranked Set Samples from a Population with Known Mean

M. MAHDIZADEH AND N. R. ARGHAMI

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Ranked set sampling (RSS) is a cost-efficient technique for data collection when the units in a population can be easily judgment ranked by any cheap method other than actual measurements. Using auxiliary information in developing statistical procedures for inference about different population characteristics is a well-known approach. In this work, we deal with quantile estimation from a population with known mean when data are obtained according to RSS scheme. Through the simple device of mean-correction (subtract off the sample mean and add on the known population mean), a modified estimator is constructed from the standard quantile estimator. Asymptotic normality of the new estimator and its asymptotic efficiency relative to the original estimator are derived. Simulation results for several underlying distributions show that the proposed estimator is more efficient than the traditional one.

Keywords Mean-correction; Quantile estimation; Ranked set sampling.

Mathematics Subject Classification 62G30; 62G99.

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

1. Introduction

The method of ranked set sampling (RSS) is applicable when ranking the sampling units can be done easily without actual measurement on them, which may be difficult or expensive. McIntyre (1952) proposed this sampling strategy when studying population mean of pasture yields in the context of agriculture. He claimed that in estimating the population mean, RSS based estimator is more efficient than that obtained by simple random sampling (SRS). The basic version of RSS introduced by McIntyre can be elucidated as follows. First, the experimenter draws k independent simple random samples, each of size k from the population. Then the units within the rth ($r = 1, \ldots, k$) sample are subjected to judgement ordering, with negligible cost, and the unit possessing ith lowest rank is identified. Finally, the identified units are measured. Proceeding in this way, we attain a ranked set sample of size k. The success of this sampling scheme highly depends on the accuracy of

Received January 27, 2010; Accepted September 13, 2011
Address correspondence to M. Mahdizadeh, Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, P.O. Box 91775-1159, Mashhad, Iran; E-mail: mahdizadeh.m@live.com

ranking the k units. For this reason, the set size k must be kept small to reduce errors in judgement ranking. If needed, this process can be replicated m times (cycles) to yield a sample of desired size $n = mk$. Each cycle involves identifying k^2 units which only k of them are selected for actual quantification. These n measured observations are said to constitute the ranked set sample denoted by $\{X_{[r]j} : r = 1, \ldots, k; j = 1 \ldots, m\}$, where $X_{[r]j}$ is the rth judgement order statistic from the jth cycle. In the absence of ranking errors, $X_{[r]j}$ has the same marginal distribution as the rth order statistic from a simple random sample of size k. It is worth noting that the n resulting measurements are mutually independent.

Ranked set sampling received little attention for years until Halls and Dell (1966) examined its effectiveness for estimating weight of herbage in a pine forest. This was perhaps due to lack of mathematical foundation for RSS. Theoretical support to this technique, without referring to McIntyre's work, was given by Takahasi and Wakimoto (1968) and, independently, by Dell (1969) while investigating McIntyre's proposal. Since then there have been a lot of contributions to statistical inference based on RSS. The recent book by Chen et al. (2004) presented a comprehensive survey and references on this topic. We now briefly mention some of the literature, especially parts related to our work.

Nonparametric approach to RSS has been developed in a large number of articles. Dell and Clutter (1972) formally showed that the sample mean using RSS is an unbiased estimator of the population mean regardless of ranking errors and it has a smaller variance than the sample mean using SRS when the number of measured units are the same. Stokes and Sager (1988) characterized a ranked set sample as a sample from a conditional distribution, conditioning on a multinomial random vector, and applied RSS to the estimation of the cumulative distribution function. Kvam and Samaniego (1994) discussed nonparametric maximum likelihood estimation based on ranked set samples. Chen (1999) studied the kernel method of density estimation in RSS. Chen (2000) considered quantile estimation from RSS data and found that RSS method can substantially improve the efficiency of the quantile estimators.

Being the most informative parameter of a distribution, mean is communicated in scientific media more than any other population parameter. For example, World Health Organization (WHO) has very reliable estimates of some health indices (such as milligram of iron per liter of blood of 12 years old children of various countries). The distribution of such indices are however not derived or at least not published. So there are cases in which we may be interested in the quantiles of a distribution with known mean. Now suppose that there is a simple and cheap method of measuring the variable of interest in addition to the relatively expensive method of measurement of the same quantity. The former can be used to order the units and the latter is then applied for actual quantification.

The use of auxiliary information to improve estimation accuracy is a key topic in sampling theory. Control variate is a well-established method for improving the efficiency of quantile estimation when auxiliary information is available. It works by adjusting the standard quantile estimator using an estimated relationship between the control variate and the variable of interest. In this work, we adopt this technique for quantile estimation using ranked set samples when the population mean is known.

This article is structured as follows. In Sec. 2, we propose the mean-corrected quantile estimator in RSS and present some theoretical results concerning its

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

asymptotic normality and asymptotic relative efficiency with respect to the standard quantile estimator in RSS. Section 3 contains results of Monte Carlo simulations carried out to reveal the properties of the proposed estimator. We end in Sec. 4 with a summary.

2. Mean-Corrected Quantile Estimator

To present the main result of this section, we first review some basic concepts regarding quantile estimation. For convenience, the following notations will be used. Let f and F denote the probability density function (pdf) and cumulative distribution fuction (cdf) of the population of interest, respectively. We also denote the pdf and cdf and mean of the rth order statistic in a set of size k as $f_{[r]}$, $F_{[r]}$, and $\mu_{[r]}$, respectively. Finally, the pdf of Beta random variable with parameters α and β is indicated by $b(\alpha, \beta; x)$, and its cdf by $B(\alpha, \beta; x)$. Throughout this article, we assume that there is no errors in ranking the sampling units.

Suppose we wish to estimate the pth quantile

$$\xi_p = \inf\{x : F(x) \geq p\}$$

using the ranked set sample $\{X_{[r]j} : r = 1, \ldots, k; j = 1 \ldots, m\}$ of size $n = mk$, from F. The corresponding empirical distribution function is given by

$$\widehat{F}^{\star}_n(x) = \frac{1}{mk} \sum_{r=1}^{k} \sum_{j=1}^{m} I\{X_{[r]j} \leq x\},$$

where $I\{A\}$ denotes the indicator function of the set A. The estimate of pth quantile based on the ranked set sample is given by

$$\hat{\xi}^{\star}_p = \inf\{x : \widehat{F}^{\star}_n(x) \geq p\}.$$

The following two theorems (Chen, 2000) which state, respectively, the asymptotic normality and Bahadur representation for $\hat{\xi}^{\star}_p$, are of main importance and will be used in the sequel.

Theorem 2.1. *Suppose that the density function f is positive in a neighborhood of ξ_p and is continuous at ξ_p and the judgement ranking is perfect. Then*

$$\sqrt{n}(\hat{\xi}^{\star}_p - \xi_p) \xrightarrow{\mathscr{L}} N\left(0, \frac{\sigma^2_{k,p}}{f^2(\xi_p)}\right),$$

where

$$\sigma^2_{k,p} = \frac{1}{k} \sum_{r=1}^{k} B(r, k-r+1; p)[1 - B(r, k-r+1; p)]$$

and $\xrightarrow{\mathscr{L}}$ denotes convergence in law.

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

Theorem 2.2. *Suppose that the density function f is positive in a neighborhood of ξ_p and is continuous at ξ_p. Then*

$$\hat{\xi}_p^\star = \xi_p + \frac{p - \widehat{F}_n^\star(\xi_p)}{f(\xi_p)} + R_n,$$

where with probability one,

$$R_n = O(n^{-3/4}(\log n)^{3/4})$$

as $n \to \infty$.

When the population of interest has known mean, it is possible to improve $\hat{\xi}_p^\star$ by an application of control variate method, i.e., mean-correction: subtract off the sample mean and add on the known population mean. Such a strategy was also employed by Breidt (2004) in the context of SRS. We now present a central limit theorem for the mean-corrected sample quantile, along with its asymptotic efficiency relative to the standard sample quantile in RSS.

Theorem 2.3. *Let $\{X_{[r]j} : r = 1, \ldots, k; j = 1 \ldots, m\}$ be a ranked sample of size $n = mk$ drawn from F, where F has known mean μ and finite unknown variance σ^2. Suppose that F has a continuous derivative in a neighborhood of ξ_p, with $F'(\xi_p) = f(\xi_p) > 0$. Define*

$$\beta_r = Cov(I\{X_{[r]1} \le \xi_p\}, X_{[r]1}).$$

Then for the mean-corrected sample quantile

$$\tilde{\xi}_p^\star = \hat{\xi}_p^\star - \frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{m} X_{[r]j} + \mu, \tag{1}$$

we have $E(\tilde{\xi}_p^\star) = E(\hat{\xi}_p^\star)$ and

$$\sqrt{n}(\tilde{\xi}_p^\star - \xi_p) \xrightarrow{\mathcal{L}} N\left(0, \frac{\sigma_{k,p}^2 + \frac{2}{k}\sum_r \beta_r f(\xi_p) + [\sigma^2 - \frac{1}{k}\sum_r(\mu_{[r]} - \mu)^2] f^2(\xi_p)}{f^2(\xi_p)}\right).$$

The asymptotic relative efficiency of $\tilde{\xi}_p^\star$ relative to $\hat{\xi}_p^\star$ is

$$ARE = \frac{\sigma_{k,p}^2}{\sigma_{k,p}^2 + \frac{2}{k}\sum_r \beta_r f(\xi_p) + [\sigma^2 - \frac{1}{k}\sum_r(\mu_{[r]} - \mu)^2] f^2(\xi_p)}.$$

Proof. From Theorem 2.2, we can write

$$\tilde{\xi}_p^\star = \xi_p + \frac{p}{f(\xi_p)} + \mu - \left(\frac{\widehat{F}_n^\star(\xi_p)}{f(\xi_p)} + \frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{m} X_{[r]j}\right) + R_n,$$

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

where $R_n \to 0$ almost surely as $n \to \infty$. Hence, the results follow from a standard central limit theorem and making the observation that

$$\begin{aligned} n\mathrm{Var}\left(\frac{\widehat{F}_n^{\star}(\xi_p)}{f(\xi_p)} + \frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{m} X_{[r]j}\right) &= \frac{\sigma_{k,p}^2}{f^2(\xi_p)} + \left[\sigma^2 - \frac{1}{k}\sum_{r=1}^{k}(\mu_{[r]} - \mu)^2\right] \\ &\quad + \frac{2}{nf(\xi_p)}\mathrm{Cov}\left(\sum_{s=1}^{k}\sum_{j=1}^{m} I\{X_{[s]j} \le \xi_p\}, \sum_{r=1}^{k}\sum_{i=1}^{m} X_{[r]i}\right) \\ &= \frac{\sigma_{k,p}^2}{f^2(\xi_p)} + \left[\sigma^2 - \frac{1}{k}\sum_{r=1}^{k}(\mu_{[r]} - \mu)^2\right] \\ &\quad + \frac{2m}{nf(\xi_p)}\sum_{r=1}^{k}\mathrm{Cov}(I\{X_{[r]1} \le \xi_p\}, X_{[r]1}), \end{aligned}$$

in which we have used the fact that

$$\mathrm{Var}\left(\frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{m} X_{[r]j}\right) = \frac{1}{n}\left[\sigma^2 - \frac{1}{k}\sum_{r=1}^{k}(\mu_{[r]} - \mu)^2\right]$$

and that $X_{[r]j}$ and $X_{[r']j}$ are independent for any $r \neq r'$.

To have better understanding of the behavior of $\tilde{\xi}_p^{\star}$, we look at its counterpart in SRS. Let $X_1, \ldots, X_n$ be a simple random sample of size n, from F, with empirical distribution function given by

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \le x\}.$$

The pth quantile is then estimated by

$$\hat{\xi}_p = \inf\{x : \widehat{F}_n(x) \ge p\},$$

and subsequently the mean-corrected sample quantile is

$$\tilde{\xi}_p = \hat{\xi}_p - \frac{1}{n}\sum_{i=1}^{n} X_i + \mu. \tag{2}$$

Breidt (2004) derived asymptotic relative efficiency of $\tilde{\xi}_p$ relative to $\hat{\xi}_p$ as

$$\mathrm{ARE} = \frac{p(1-p)}{p(1-p) + 2\beta_0 f(\xi_p) + \sigma^2 f^2(\xi_p)},$$

where

$$\beta_0 = \mathrm{Cov}(I\{X_i \le \xi_p\}, X_i).$$

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

We can write

$$\beta_0 = \mathrm{E}(I\{X_i \le \xi_p\}X_i) - \mathrm{E}(I\{X_i \le \xi_p\})\mathrm{E}(X_i) = \int_{-\infty}^{\xi_p} xf(x)dx - \mu \mathrm{P}(X_i \le \xi_p),$$

and by taking derivative of β_0 with respect to p we have

$$\frac{d\beta_0}{dp} = \frac{d}{dp}\left(\int_{-\infty}^{F^{-1}(p)} xf(x)dx - \mu p\right) = \frac{1}{F'(\xi_p)}\xi_p f(\xi_p) - \mu = \xi_p - \mu,$$

where the second equality is an application of the Leibnitz differentiation formula. He deduced that β_0 is a decreasing function of p for $\xi_p < \mu$, achieves its minimum at $\xi_p = \mu$ and is increasing function of p for $\xi_p > \mu$. Since $\beta_0 \to 0$ as $p \to 0$ or $p \to 1$, it follows that β_0 is negative for $0 < p < 1$. The negative covariance makes it possible for $\tilde{\xi}_p$ to outperform $\hat{\xi}_p$.

We are going to examine what happens in the case of RSS. Similarly,

$$\begin{aligned}\beta_r &= \mathrm{E}(I\{X_{[r]1} \le \xi_p\}X_{[r]1}) - \mathrm{E}(I\{X_{[r]1} \le \xi_p\})\mathrm{E}(X_{[r]1})\\ &= \int_{-\infty}^{\xi_p} xf_{[r]}(x)dx - \mu_{[r]}\mathrm{P}(X_{[r]1} \le \xi_p),\end{aligned}$$

and

$$\begin{aligned}\frac{d\beta_r}{dp} &= \frac{d}{dp}\left(\int_{-\infty}^{F_{[r]}^{-1}(B(r,k-r+1;p))} xf_{[r]}(x)dx - \mu_{[r]}B(r,k-r+1;p)\right)\\ &= \frac{1}{F'_{[r]}(\xi_p)}\xi_p f_{[r]}(\xi_p)b(r,k-r+1;p) - \mu_{[r]}b(r,k-r+1;p)\\ &= b(r,k-r+1;p)(\xi_p - \mu_{[r]}).\end{aligned}$$

But we have

$$\begin{aligned}\sum_{r=1}^{k} b(r,k-r+1;p) &= \sum_{r=1}^{k}\frac{k!}{(r-1)!(k-r)!}p^{r-1}(1-p)^{k-r}\\ &= k\sum_{s=0}^{k-1}\frac{(k-1)!}{s!(k-1-s)!}p^{s}(1-p)^{k-1-s} = k.\end{aligned}$$

Hence,

$$\frac{d}{dp}\left(\sum_r \beta_r\right) = k(\xi_p - \sum_r w_r\mu_{[r]}),$$

where $w_r = b(r,k-r+1;p)/k$ and $\sum_r w_r = 1$. It is observed that instead of μ in the expression of $\frac{d\beta_0}{dp}$, a weighted sum of the mean of the k order statistics appears, which makes it harder to justify negativity of the covariances explicitly. As a referee pointed out, the following elegant proof is available for determining the sign of β_r's. It is easy to show that for any random variable X, $\mathrm{E}(I\{X \le \xi_p\}X) = \mathrm{E}(X \mid X \le \xi_p)P(X \le \xi_p)$. Then, it follows that

$$Cov(I\{X \le \xi_p\}, X) = E(I\{X \le \xi_p\}X) - E(X)P(X \le \xi_p)$$

$$= E(X \mid X \le \xi_p)P(X \le \xi_p) - E(X)P(X \le \xi_p)$$
$$= [E(X \mid X \le \xi_p) - E(X)]P(X \le \xi_p) < 0.$$

As noted by Breidt (2004), a general form of the control variate estimator (2) is

$$\tilde{\xi}_p(\theta) = \hat{\xi}_p - \theta\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right),$$

which has asymptotically zero bias and attains minimum asymptotic variance (MAV) at

$$\theta_p = \frac{-\beta_0}{\sigma^2 f(\xi_p)}.$$

The above class of estimators can also be constructed in RSS, i.e.,

$$\tilde{\xi}^{\star}_p(\theta) = \hat{\xi}^{\star}_p - \theta\left(\frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{m} X_{[r]j} - \mu\right).$$

From Theorem 2.3, it follows that an asymptotic approximation for the variance of the MAV estimator is

$$\mathrm{Var}(\tilde{\xi}^{\star}_p(\theta)) \approx \frac{\sigma^2_{k,p}}{nf^2(\xi_p)} + \frac{2\theta}{nf(\xi_p)}\frac{1}{k}\sum_r \beta_r + \frac{\theta^2}{n}\left[\sigma^2 - \frac{1}{k}\sum_r(\mu_{[r]} - \mu)^2\right] = h(\theta).$$

By setting $h'(\theta)$ equal to zero, we have

$$\frac{dh}{d\theta} = \frac{2}{nf(\xi_p)}\frac{1}{k}\sum_r \beta_r + \frac{2\theta}{n}\left[\sigma^2 - \frac{1}{k}\sum_r(\mu_{[r]} - \mu)^2\right] = 0,$$

which is solved by

$$\theta^{\star}_p = \frac{-\frac{1}{k}\sum_r \beta_r}{[\sigma^2 - \frac{1}{k}\sum_r(\mu_{[r]} - \mu)^2]f(\xi_p)}, \tag{3}$$

and

$$\frac{d^2h}{d\theta^2} = 2\mathrm{Var}\left(\frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{n} X_{[r]j}\right) > 0,$$

ensures us that MAV occurs at $\theta^{\star}_p$. In practical situations, however, (3) is unknown and its estimation (which is quantile-specific) requires estimating covariances, variance and density function. To this end, the following estimators are available.

Each β_r is estimated by an expression such as numerator of the sample correlation coefficient with appropriate choice of random variables. Consider an ANOVA performed on the ranked set sample with rank of units as the factor. Let MSE and MST be the mean-square error and mean-square due to treatment, i.e.,

$$\mathrm{MSE} = \frac{1}{k(m-1)}\sum_r\sum_j (X_{[r]j} - \hat{\mu}_{[r]})^2$$

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

and

$$\mathrm{MST} = \frac{1}{k-1}\left\{ \sum_r \sum_j (X_{[r]j} - \hat{\mu})^2 - k(m-1)\mathrm{MSE} \right\},$$

where $\hat{\mu}_{[r]} = \sum_j X_{[r]j}/m$ and $\hat{\mu} = \sum_r \sum_j X_{[r]j}/(mk)$. MacEachern et al. (2002) developed an unbiased estimator of σ^2 in RSS. Their estimator is equivalently expressed as

$$\hat{\sigma}^2 = \frac{1}{mk}\{(k-1)\mathrm{MST} + (mk - k + 1)\mathrm{MSE}\}.$$

Also, the kernel estimator of f at a point $x \in R$ is given by

$$\hat{f}_{\mathrm{RSS}}(x) = \frac{1}{mkh}\sum_{r=1}^{k}\sum_{j=1}^{m} K\left(\frac{x - X_{[r]j}}{h}\right),$$

where K is usually a standard probability density function symmetric about zero, and h is a bandwidth to be determined. Chen (1999) showed that given a fixed h, $\hat{f}_{\mathrm{RSS}}$ has smaller variance than its SRS counterpart based on the same sample size. We use the standard normal density function as K with $h = 1.06\, sn^{-1/5}$, where s is the sample standard deviation. Finally, when estimating $f(\xi_p)$, $\hat{f}_{\mathrm{RSS}}(\hat{\xi}_p^\star)$ is used.

3. Simulation Results

In this section, the Monte Carlo approach is used to assess the performance of the proposed estimators through relative efficiencies e1, e2, and e3 defined as efficiency of the mean-corrected RSS estimator relative to the basic RSS estimator, efficiency of the MAV RSS estimator relative to the basic RSS estimator, and efficiency of the mean-corrected RSS estimator relative to its SRS counterpart, respectively. These values are estimated based on 10,000 replications at sample size $n = 50$. In generating ranked set samples, the set size $k = 2$ was used.

Table 1 contains simulated relative efficiencies for 0.05, 0.25, 0.5, 0.75, and 0.95 quantiles under the following models: (A) uniform(0,1), (B) normal(0,1), (C) t(10), (D) t(25), (E) beta(2,2), (F) exponential(1), (G) chi-square(8), (H) gamma(3,1), (I) inverse-gaussian(1,2), and (J) gumbel(0,1). This set of distributions includes a variety of functional shapes. It is to be noted that (A)–(E) are symmetric, (F)–(I) are right-skewed, and (J) is left-skewed.

It can be seen from Table 1 that for (A)–(E), the mean-corrected estimator is superior to the standard estimator in RSS. Moreover, the corresponding relative efficiency gains its maximum at $p = 0.5$ and damps away symmetrically towards $p = 0.05$ and $p = 0.95$. The results suggest that mean correction is deficient in a very light tail (such as the left tail in a right-skewed distribution, or a bounded tail). It is observed that for symmetric distributions, apart from a few cases, efficiency of the MAV estimator relative to the standard estimator is not higher than that of the mean-corrected estimator relative to the standard estimator. Indeed, the MAV estimator tends to be similar to $\tilde{\xi}_p^\star$ in these cases. Agreeing with this claim, the normal distribution has $\theta_p \equiv 1$, and $\tilde{\xi}_p(\theta_p)$ is the same as $\tilde{\xi}_p$. One can find that for asymmetric distributions, the MAV estimator outperforms the mean-corrected

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013

Table 1
Simulated relative efficiencies under several underlying distributions

		p							*p*				
Dist.	Eff.	0.05	0.25	0.5	0.75	0.95	Dist.	Eff.	0.05	0.25	0.5	0.75	0.95
A	e1	0.75	1.90	2.52	1.94	0.78	F	e1	0.08	0.40	1.29	2.17	1.39
	e2	1.14	1.87	2.94	1.88	1.04		e2	0.93	0.96	1.52	2.36	1.71
	e3	1.20	1.04	1.09	1.02	1.19		e3	1.26	1.10	1.00	1.05	1.06
B	e1	1.29	1.82	2.16	1.81	1.31	G	e1	0.78	1.15	1.93	2.15	1.40
	e2	1.26	1.77	2.08	1.76	1.27		e2	1.10	1.36	1.88	2.25	1.54
	e3	1.01	1.00	1.02	1.00	1.03		e3	1.11	1.03	1.02	1.04	1.03
C	e1	1.29	1.73	1.87	1.74	1.31	H	e1	0.66	1.03	1.84	2.16	1.39
	e2	1.28	1.71	1.85	1.73	1.30		e2	1.08	1.34	1.82	2.31	1.55
	e3	1.03	1.00	0.98	1.02	1.01		e3	1.12	1.02	1.02	1.03	1.05
D	e1	1.32	1.81	2.07	1.79	1.32	I	e1	0.54	0.82	1.54	2.15	1.40
	e2	1.29	1.78	1.99	1.75	1.28		e2	1.07	1.20	1.68	2.33	1.65
	e3	1.02	1.00	1.01	1.00	0.99		e3	1.13	1.05	1.02	1.03	1.03
E	e1	1.15	1.91	2.54	1.93	1.14	J	e1	0.86	1.08	1.74	2.10	1.38
	e2	1.18	1.86	2.57	1.85	1.17		e2	1.11	1.31	1.85	2.26	1.56
	e3	1.06	1.01	1.04	1.02	1.06		e3	1.10	1.02	1.04	1.02	1.03

estimator and this is more evident for 0.05 quantile. Finally, it is seen that except for the extreme tails in (A), and lower tails in (F)–(J), the mean-corrected estimator in SRS and RSS are more or less similar.

4. Conclusion

In this article, we employed the control variate method to improve quantile estimation using ranked set samples when the population that yielded the sample has a known mean. The proposed estimator is the standard quantile estimator which is mean corrected: the sample mean subtracted off and the known population mean added on. Asymptotic results for the mean-corrected estimator are provided and compared with those in SRS. Simulation results, for several underlying distributions, integrate the presentation to illustrate the asymptotic theory and show that the mean-corrected estimator surpasses the standard quantile estimator in RSS. We are working on a more systematic approach of utilizing the additional information (known mean) to propose quantile estimator which markedly dominates its competitor in SRS. Another interesting extension is towards distribution function estimation in the above setup. Results of research in these directions will be reported in future works.

Acknowledgement

We would like to thank the anonymous referees and the Associate Editor for providing helpful comments that greatly improved the initial version of this article.

References

Breidt, F. J. (2004). Simulation estimation of quantiles from a distribution with known mean. *J. Computat. Graph. Statist.* 13(2):487–498.

Chen, Z. (1999). Density estimation using ranked set sampling data. *Environmental Ecological Statistics* 6:135–146.

Chen, Z. (2000). On ranked-set sample quantiles and their application. *Journal of Statistical Planning and Inference* 83:125–135.

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

Dell, T. R. (1969). The Theory and some Applications of Ranked Set Sampling. Ph.D. Thesis, University of Georgia, Athens, GA.

Dell, T. R., Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics* 28:545–555.

Halls, L. K., Dell, T. R. (1966). Trial of ranked set sampling for forage yields. *Forest Sci.* 12(1):22–26.

Kvam, P. H., Samaniego, F. J. (1994). Nonparametric maximum likelihood estimation based on ranked set samples. *Journal of the American Statistical Association* 86:526–537.

MacEachern, S. N., Ozturk, O., Wolfe, D. A., Stark, G. V . (2002). A new ranked set sample estimator of variance. *Jornal of the Royal Statistical Society B* 64:177–188.

McIntyre, G. A. (1952). A method of unbiased selective sampling using ranked sets. *Australian Jornal of Agriculture Research* 3:385–390.

Stokes, S. L., Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution function. *Journal of the American Statistical Association* 83:374–381.

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 Statistics and Mathematics* 20:1–31.

Downloaded by [Memorial University of Newfoundland] at 05:56 11 September 2013