Pakistan Journal of Statistics and Operation Research

A Multiplicative Bias Correction Technique for Estimating Quantile Function with an Application

Nicholas Makumi[1*], Romanus Odhiambo Otieno[2], George Otieno Orwa[3]
Alexis Habineza[4]

*Corresponding author

1. Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya, nicholas.makumi@jkuat.ac.ke
2. Department of Statistics and Actuarial Sciences, JKUAT, Nairobi, Kenya, rodhiambo@must.ac.ke
3. Department of Statistics and Actuarial Sciences, JKUAT, Nairobi, Kenya, gorwa@buc.ac.ke
4. Pan African University, Institute for Basic Sciences, Technology and Innovation (PAUSTI), Nairobi, Kenya, alexhabk87@gmail.com

Abstract

Smooth non-parametric quantile function estimators on basis of symmetric kernels exhibit boundary bias due to spill-over near the edges. An improved non-parametric estimator of a quantile function under simple random sampling without replacement is proposed, based on a multiplicative bias corrected distribution function. There is no spill-over around the edges with our new quantile estimator. The proposed quantile estimator's asymptotic properties are investigated. The suggested method is compared to existing estimators using real data set findings, demonstrating the improved performance.

Key Words: Quantiles estimation; Multiplicative Bias Correction; Finite population quantiles.

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

Simple random sampling (SRS) is widely utilized only when variables' values don't really change significantly and the population is homogeneous. SRS is among the most basic sampling procedures in many ways, and no further information is required. Furthermore, when using SRS to create a sample, sample weights aren't really required for evaluating data from a survey using, for example, regression or multivariate analysis. A downside of SRS is the complexity in managing accuracy and the inefficiencies of not using supplemental data, which could result in enormous samples that are unneeded. Furthermore, because no supplementary information is used, there is always the potential of a skewed sample.

In the sampling survey, we are time and again interested in studying the distribution of a certain interest variable, say Y. The efficient technique to illustrate the distribution function is by assessing the quantiles of the distribution. By definition, the distribution p th quantile, is the value Q, satisfying $P(Y \leq Q) = p$. In the literature, much emphasis has been placed on the p th quantile estimation problem. Majority of studies employ simple random sampling (SRS) to estimate the quantile utilizing kernel density function, for further information we direct the reader see for example, Nadaraya(1964), Lio and Padgett(1991), and Jones(1992).

It is commonly acknowledged that the sample quantile has a significant inefficiency. To address this issue, many researchers have suggested kernel-type estimators as smooth alternatives to the sample quantile. Nadaraya(1964) and Parzen(1979) are two early works on kernel estimation techniques of the quantile function. Reiss(1980) demonstrated that as the sample size grows, the sample quantile's asymptotic relative deficit with regard to a linear combination of finitely numerous order statistics diverges to infinite. Falk(1984) also looked at the sample quantile's asymptotic relative inadequacy in comparison to kernel-type quantile estimators. Yang(1985) investigated kernel-type quantile estimators' asymptotic properties. Padgett(1986) investigated on right-censored data to the earlier works. These conclusions are all based on symmetric kernel functions. The use of symmetric kernel functions causes boundary bias or spill-over consequences because the quantile function's domain is a bounded interval $(0,1)$. When fixed symmetric kernels are utilized, boundary bias is caused by incorrect weights of kernel functions around the quantile function's boundaries.

To eliminate boundary bias, Chen(1999) and Chen(2000) advocated using beta kernel estimators for density functions and regression curves. It is permissible to integrate a beta probability density function into smooth nonparametric estimators of quantile function because the intervention of a beta probability density function fits the range of the quantile function. We offer a novel quantile estimator predicated on the multiplicative bias corrected distribution function that is devoid of spill-over effects for simple random sampling without replacement in this study.

2. Multiplicative Bias Corrected Distribution Function Estimator

The usual practice of quantile estimation is to construct an estimator of the cdf of Y first, then to deduce an estimator of the $\alpha-$quantile of Y. In this section, a brief description of the multiplicative bias corrected distribution function estimator developed by Onsongo et al.(2018) is presented. Then used to derive the proposed quantile estimator as well as its asymptotic properties. Let Y be the survey variable associated with auxiliary variable X which are assumed to follow superpopulation model

$$y_i = \mu(x_i) + \sigma(x_i)e_i,\ i = 1, 2, ..., N \tag{1}$$

where $\sigma(x_i)$ is a function of x_i that takes account of heteroscedasticity and $e_i \sim iidN(0, \sigma^2)$, $E(y_i) = \mu(x_i)$ and

$$Cov(y_i, y_j) = \begin{cases} \sigma^2(x_i) & \text{if i=1,2,...,N} \\ 0 & \text{otherwise.} \end{cases}$$

As before, $y_1, \ldots, y_N$ represent Y values in U. Similarly, let $x_1, \ldots, x_N$ signify the values of X, respectively, in U. Assume that the value of Y is only recorded for elements of s_n, while the values of X are available for every element in U. Under model-based approach, Onsongo et al.(2018) proposed

$$\hat{F}_{MBC}(t) = \frac{1}{N}\left\{ \sum_{i\in s} I\left(y_i \le t\right) + \sum_{j\in r} \hat{H}\left(t - \hat{\mu}(x_j)\right) \right\} \tag{2}$$

The estimator in equation (2) can be viewed as a weighted sum of two estimators of one from observed values and one from the auxiliary information of the non-sample elements. That is,

$$\hat{F}_{MBC}(t) = N^{-1}\left[NF_{N_y}(t) + (N-n)F_r^*(t)\right],$$

where $F_{N_y}(\cdot)$ is the usual empirical distribution function and

$$F_r^*(t) = \sum_{i\in s}(N-n)^{-1}\sum_{j\in r}\hat{H}\left(t - \hat{\mu}(x_j)\right).$$

Based on some regularity conditions, Onsongo et al.(2018) showed that
$\hat{F}_{MBC}(t)$ is unbiased estimator of $F_N(t)$ that is

$$\mathrm{E}\left(\hat{F}_{MBC}(t)\right) = F_N(t) \tag{3}$$

and the analytic expression of the variance $\hat{F}_{MBC}(t) - F_{N_y}(t)$ is

$$\begin{aligned}\operatorname{Var}\left[\hat{F}_{MBC}(t) - F_N(t)\right] = \frac{1}{N^2}\sum_{i\in s}\left\{\sum_{j=1}^{N-n}\sum_{k=1}^{N-n} w_{ij}^* w_{ik}^* \left[H_i\left(t - \max\left(\hat{\mu}_j, \hat{\mu}_k\right)\right) - H_i\left(t - \hat{\mu}_j\right) H_i\left(t - \hat{\mu}_k\right)\right]\right\} \\ + \frac{1}{N^2}\left\{(N-n)P\left(y_j \le t\right)\left[1 - P\left(y_j \le t\right)\right]\right\}\end{aligned} \tag{4}$$

3. Proposed Quantile Estimator

A common use of distribution function estimators is to provide a quantile estimators. Now, we can use $\hat{F}_{MBC}(t)$ to estimate αth quantile $Q_{N_y}(\alpha)$. To this end, suppose that a sample of size n based on SRSWOR selected from the underlying population with an interested variable Y. Then, $Q_{N_y}(\alpha)$ can be estimated by

$$\hat{Q}_{MBC}(\alpha) = \inf\left\{t \in U : \hat{F}_{MBC}(t) \ge \alpha\right\} = \hat{F}_{MBC}^{-1}(\alpha) \tag{5}$$

this is a common way to define a quantile estimator based on a distribution function estimator. From equation (2), $\hat{Q}_{MBC}(\alpha)$ can be computed by numerically solving the equation

$$\hat{F}_{MBC}\left(\hat{Q}_{MBC}(\alpha)\right) = \alpha.$$

The following two theorems which state, respectively, the asymptotic normality and Bahadur representation for α, are of main importance and will be used in the sequel.

Theorem 3.1. *Serfling(1980) Assume that the density function f is positive in the vicinity of ξ_p and that it is continuous at ξ_p, and that the judgment ordering is perfect.*

$$\sqrt{n}\left(\hat{\xi}_p^* - \xi_p\right) \xrightarrow{D} N\left(0, \frac{\sigma_{k,p}^2}{f^2(\xi_p)}\right)$$

where $\xrightarrow{D}$ denotes convergence in law.

Theorem 3.2. *Francisco and Fuller(1991) Suppose x be a point in the interval A_1 that contains the interior point $q(\tau_1^0)$. Then, sample quantile can be expressed as,*

$$\hat{q}_{rn}(\tau) = q(\tau) - \left[f(q(\tau))\right]^{-1}\left[F_{rn}(q(\tau)) - F(q(\tau))\right] + R_{rn}^*(\tau)$$

with $R_{rn}^*(\tau) = o_p\left(n_r^{-1/2}\right)$ uniformly in τ for τ in W_1, where $W_1 = \left\{\tau : F(x) = \tau \;\; and \;\; x \in A_1\right\}$

Proof: Proof see (Francisco and Fuller, 1991).

With the conclusions from Theorem (3.2) above, the estimator $\hat{Q}_{MBC}(\alpha)$ may be written asymptotically as a linear function of the estimated distribution function assessed at the quantile $Q_{N_y}(\alpha)$ by the Bahadur representation, see (Chambers and Dunstan, 1986). Let $\mathrm{F_{MBC}}$ be multiplicative bias corrected distribution function of the density $\mathrm{f_{MBC}}$. Then, utilizing Taylor series expansion of a function $F_{MBC}\left(\hat{Q}_{MBC}(\alpha)\right)$ about $Q_{N_y}(\alpha)$, the following is obtained:

$$F_{MBC}\left(\hat{Q}_{MBC}(\alpha)\right) = F_{MBC}(Q_{N_y}(\alpha)) + f_{MBC}(Q_{N_y}(\alpha))\left[\hat{Q}_{MBC}(\alpha) - Q_{N_y}(\alpha)\right] + O(n^{-\frac{1}{2}}) \tag{6}$$

Where $F'_{MBC}(Q_{N_y}(\alpha)) = f_{MBC}(Q_{N_y}(\alpha))$, according to (1) since F_{MBC} contains two derivatives in a $Q_{N_y}(\alpha)$

neighborhood, this neighborhood is bound by the second derivative and $F'_{MBC}(Q_{N_y}(\alpha))$ is positive. From equation (6) the Bahadur's representation, Bahadur(1966) of $\hat{Q}_{MBC}(\alpha)$ is

$$\hat{Q}_{MBC}(\alpha) = Q_{N_y}(\alpha) + \frac{\left(\alpha - \hat{F}_{MBC}(Q_{N_y}(\alpha))\right)}{f_{MBC}(Q_{N_y}(\alpha))} + O(n^{-\frac{1}{2}}). \tag{7}$$

where $f_{MBC}(.)$ denotes the derivative of the limiting value of $F_{MBC}(.)$ as $N \longrightarrow \infty$, $\hat{F}_{MBC}(\hat{Q}_{MBC}(\alpha)) = \alpha$ and $O(n^{-1/2})$, according to Kiefer(1967) becomes negligible as $n \longrightarrow \infty$. The linear approximation previously used by Kuk and Mak(1989) and Chen and Wu(2002) helps to study the asymptotic properties of the estimator. According to equations (3), (4) and (7) it is easy to see that

$$E\left[\hat{Q}_{MBC}(\alpha)\right] = Q_{N_y}(\alpha) + O(n^{-\frac{1}{2}}), \tag{8}$$

and

$$\text{Var}\left(\hat{Q}_{MBC}(\alpha)\right) = \frac{1}{f^2_{MBC}(Q_{N_y}(\alpha))}\left[\frac{1-f}{N}\alpha\,(1-\alpha)\right]. \tag{9}$$

According to Kiefer(1967), $O(n^{-1/2})$ becomes negligible as $N \longrightarrow \infty$, the right-hand side of equation (8) tends to 0 and so $\hat{Q}_{MBC}(\alpha)$ is asymptotically unbiased. Furthermore, from equation (11) and (12) it can be seen that $\hat{Q}_{MBC}(\alpha)$ is asymptotically consistent estimator of $Q_{N_y}(\alpha)$. Moreover, $\hat{Q}_{MBC}(\alpha)$ has an asymptotic normal distribution as in (Serfling, 1980)

$$\text{N}\left(Q_{N_y}(\alpha), \frac{1-f}{N-1}\alpha\,(1-\alpha)\left[f_{MBC}(Q_{N_y}(\alpha))\right]^{-2}\right)$$

4. Estimators Included for Comparison in the Study

Although one of our aims is to develop estimators with respectable qualities in terms of bias, variance, and asymptotic mean squared error, we compare the new estimator given by equation (5) to some of the prominent quantile estimators proposed in the literature.

Firstly, in study estimator by Chambers and Dunstan(1986) is included, which is driven by the linear superpopulation model $y_k = \beta_0 + \boldsymbol{\beta}'\mathbf{x}_k + \epsilon_k, k \in U$, where ϵ_k forms a sequence identically and independently distributed random variables with mean zero and variance which is finite. Their estimator is described as follows:

$$\hat{Q}_{y,\text{CD},\alpha} = \inf\left\{t \mid \hat{F}_{y,\text{CD}}(t) \geq \alpha\right\} \tag{10}$$

where $\hat{F}_{y,\text{CD}}(t) = N^{-1}\left\{\Sigma_s H\left(t - y_k\right) + \sum_{U/s}\hat{G}\left(t - \hat{y}_k\right)\right\}$ represents a model-based distribution function estimator,

$$\hat{G}(u) = n^{-1}\sum_s H\left(u - \hat{\epsilon}_k\right)$$

Since this estimator (10) essentially assigns the unknown y_k for $k \in U/s$, it is important to highlight that it requires a comprehensive understanding of $\mathbf{x}_k$ for $k \in U$.

Also, empirical study includes the model-based estimator by Rao et al.(1990)

$$\widehat{F}^{\bullet}_{rkm}(t) = \frac{1}{N}\left\{\sum_{i\in s}\pi_i^{-1}\delta\left(y_i \leq t\right) + \left(\sum_{i\in U}\widehat{G}_i(t) - \sum_{i\in s}\pi_i^{-1}\widehat{G}_{ic}(t)\right)\right\}$$

with

$$\widehat{G}_i(t)=\frac{1}{\widehat{N}}\sum_{j\in s}\frac{1}{\pi_j}\delta\left(\widehat{u}_j\leq\frac{t-\widehat{R}x_i}{x_i^{1/2}}\right)$$
$$\widehat{G}_{ic}(t)=\left(\sum_{j\in s}\frac{\pi_i}{\pi_{ij}}\right)^{-1}\left[\sum_{j\in s}\frac{\pi_i}{\pi_{ij}}\delta\left(\widehat{u}_j\leq\frac{t-\widehat{R}x_i}{x_i^{1/2}}\right)\right],$$
$$\widehat{u}_j=\frac{y_j-\widehat{R}x_j}{x_j^{1/2}},\widehat{R}=\left[\sum_{i\in s}\frac{x_i}{\pi_i}\right]^{-1}\sum_{i\in s}\frac{y_i}{\pi_i}$$

where π_{ij} denotes the joint inclusion probability for the units i and j. Since the estimator $\widehat{F}^{\bullet}_{rkm}(t)$ is not always a monotone nondecreasing function, Rao et al.(1990) proposed to use the following estimator

$$\widehat{F}_{rkm}(t)=\max\left\{\widetilde{F}_{rkm}\left(y_{(i)}\right):y_{(i)}\leq t\right\}$$

where the $y_{(i)}$'s are the order statistics of the sample $\{y_i, i\in s\}$ and $\widetilde{F}_{rkm}\left(y_{(i)}\right)$ is defined by the following recursive formula

$$\widetilde{F}_{rkm}\left(y_{(i)}\right)=\max\left\{\widetilde{F}_{rkm}\left(y_{(i-1)}\right),\widehat{F}^{\bullet}_{rkm}\left(y_{(i)}\right)\right\}$$

with $\widetilde{F}_{rkm}\left(y_{(1)}\right)=\widehat{F}^{\bullet}_{rkm}\left(y_{(1)}\right)$. The Rao et al.(1990) estimator of quantile $Q_{RKM,\alpha}$ is given by

$$\widehat{Q}_{RKM;\alpha}=\widehat{F}^{-1}_{rkm}(\alpha). \tag{11}$$

A common kernel quantile estimator is also employed, which is based on a Nadaraya(1964) type kernel distribution function estimator

$$\hat{F}_n(x)=\frac{1}{n}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right)$$

whereby, K is determined from a kernel k as $K(x)=\int_{-\infty}^{x}k(t)dt$, h is a smoothing parameter. The quantile function's associated estimator is then defined by

$$\tilde{Q}_{NW}(p)=\inf\{x:\hat{F}_n(x)\geq p\},\ 0<p<1 \tag{12}$$

Finally, in our empirical study we include Dorfman and Hall estimator studied in Dorfman et al.(1993)

$$\widehat{F}_{DH}(t)=\frac{1}{N}\left[\sum_{i\in s}I\left(y_i\leq t\right)+\sum_{j\in r}\hat{G}\left(t-\hat{\mu}\left(x_j\right)\right)\right]$$

where $\hat{\mu}$ is the linear estimator of the mean function. The corresponding estimator of the quantile function is then defined by

$$\hat{Q}_{DH}(p)=\inf\{t:\widehat{F}_{DH}(t)\geq p\},\ 0<p<1 \tag{13}$$

Equations (8), (13), (14), (15) and (16) are respectively denoted by MBCQE, CDQE, RKMQE, NWQE and FAQE.

5. Application to real data set

This section illustrates the applications of our estimation approach on a dataset that involves a population, U of size $N=189$ from the United Nations Development Programme 2017. The United Nations looked into development in 189 countries around the world. The United Nations classified countries' development as either very high human development, high human development, medium human development, or low human development. According to UN figures from 2017, Kenya is among the countries with a medium level of development, ranking 143rd out of 189 countries evaluated. To rank human development index in 189 countries, the UN study employed the Human Development Index (HDI), Life expectancy at birth, Expected years of schooling, Mean years of schooling, Gross National Income (GNI) per capita, and GNI per capita rank minus HDI.

Knowing a country's GNI per capita is a solid starting point for determining the country's economic strengths and requirements, as well as the average citizen's level of living. The HDI and GNI figures were derived from the UN Development Programme 2017 data set to demonstrate the recommended methodologies. HDI serves as an auxiliary variable, while GNI serves as the research variable $y_i(i = 1, 2, 3, \ldots\ldots, 189)$. A scatter plot of HDI vs GNI is shown in Figure 1, as well as a line of best fit between GNI and HDI.

Our aim is to estimate the quantiles at 25%, 50% and 75% of GNI values for the population, assuming UN development Programme 2017 data set are representative of the entire population. From the scatter plot in Figure 1 we observe a quadratic relationship between HDI and GNI.

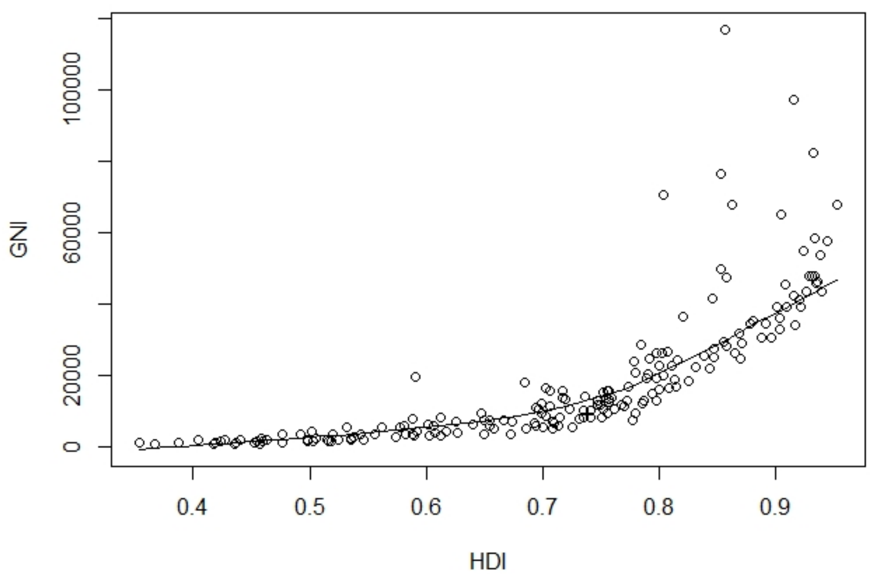

Figure 1: Scatter diagram for the population

Fitting a regression model to the data, the obtained model equation is of the form

$$\text{GNI} = 96512 \star \text{HDI} - 50408$$

The correlation coefficient between GNI and HDI is approximately 0.7515. This indicates a strong positive nonlinear relation between GNI and HDI.

We draw from the population samples of size $n = 50$, and 100 and compute quantile estimators listed below for probability levels $\alpha \in \{0.25,\ \ 0.5,\ \ 0.75\}$. Table 1 presents a comparison of empirical quantile estimator with proposed estimator and other estimators in the literature.

Table 1: Comparison of Empirical Quantile Estimator with other Estimators

		Quantile estimates		
Sample size	Estimators	0.25	0.50	0.75
	Q(p)	3843	11100	25393
$n = 50$	RKMQE	6720.611	9303.594	8463.536
	CDQE	7497.621	12089.5	10377.66
	FAQE	8351.48	7935.459	13461.61
	NWQE	1085.417	3055.923	3573.846
	MBCQE	4851.519	10274.23	15273.08
$n = 100$	RKMQE	1461.375	7676.824	9634.885
	CDQE	2683.90	7530.389	8827.565
	FAQE	1354.588	7753.122	7961.417
	NWQE	1367.84	7174.192	9692.346
	MBCQE	3792.2	7856.255	9999.85

The results exhibited in Table 1 shows that the MBCQE is closer to the empirical quantile. Therefore, MBCQE provides an almost flawless estimate of the empirical quantile function.

To study the performance of the proposed estimator in practice, 150 samples of various sizes were taken from population (United Nations development Programme 2017) according to simple random sampling without replacement. For each sample s we computed several estimators of the population quantiles $(\alpha = 0.25,\ 0.5,\ 0.75)$. We computed the bias and the mean squared error over the 150 samples.

Table 2 shows the estimated relative mean error, RME, and the relative root mean squared error, RRMSE, for each estimator considered, for the population quantiles (sample size $n = 50$ and 100).

Table 2: RME and RRMSE of the Quantile Estimators

	$\alpha = 0.25$		$\alpha = 0.5$		$\alpha = 0.75$	
Estimator	RME	RRMSE	RME	RRMSE	RME	RRMSE
			$n = 50$			
RKMQE	5966.513	1.552566	6957.031	0.6267596	17852.89	0.7030635
CDQE	4735.123	1.232142	6271.702	0.5650182	16795.58	0.6614255
FAQE	5530.96	1.43923	8556.715	0.7708752	13198.98	0.5197879
NWQE	8077.199	2.101795	8069.366	0.7269699	21820.89	0.859327
MBCQE	2779.99	0.7233906	5765.705	0.5194329	12887.84	0.5075353
			$n = 100$			
RKMQE	2453.382	0.6384027	4240.424	0.3820202	15828.63	0.6233464
CDQE	1370.384	0.3565923	4575.144	0.4121751	16763.01	0.6601428
FAQE	2589.154	0.6737325	4614.388	0.4157107	17480.43	0.6883955
NWQE	2507.451	0.6524723	4818.027	0.4340565	15764.34	0.6208144
MBCQE	812.6091	0.2114518	3929.739	0.3540305	15459.01	0.60879

From the values in Tables 2 it is clear that the overall performance of the Multiplicative bias corrected quantile estimate, MBCQE, is far superior to the usual one since it has minimum Relative Mean Error and Relative Root Mean squared Error at all levels of the $\alpha-$quantile.

The estimator's conditional performance was assessed and compared to that of other previous finite population quantile estimators. To accomplish this, 500 random samples of size 100 were chosen, and the mean of the auxiliary values x_i was calculated for each sample to get 200 $\bar{X}$ values. These sample means were then sorted in increasing order and then divided into 20-group clusters, yielding a total of 25 groups. To determine $\bar{\bar{X}}$, the group means of the means of the auxiliary variables were calculated. The RKMQE, CDQE, FAQE, NWQE and MBCQE estimators' means and biases were then computed. To get a better understanding of the pattern formed, the conditional biases were plotted against $\bar{\bar{X}}$. Figures $2-4$ exhibits plots of Conditional Bias (CB), Conditional RAB (CRAB) and Conditional MSE (CMSE) versus group means of means of the HDI for various values of α quantile. These figures shows that MBCQE have better performance than the other estimators.

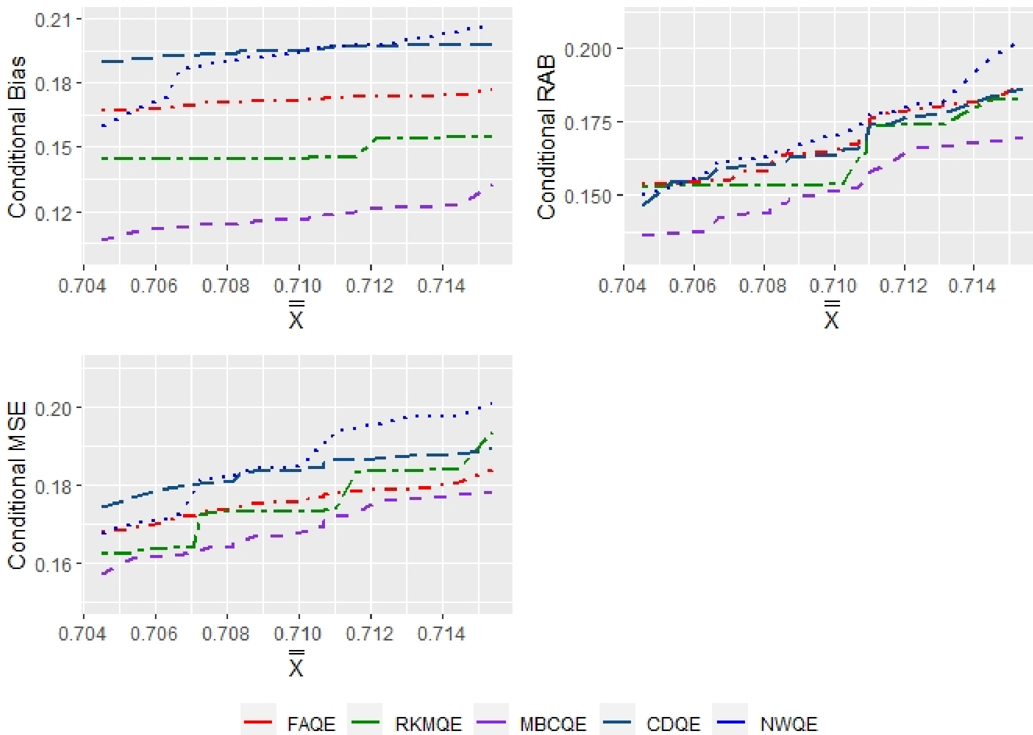

Figure 2: Plots of the behaviour of the CB, CRAB and CMSE for different estimators: $\alpha = 0.25$.

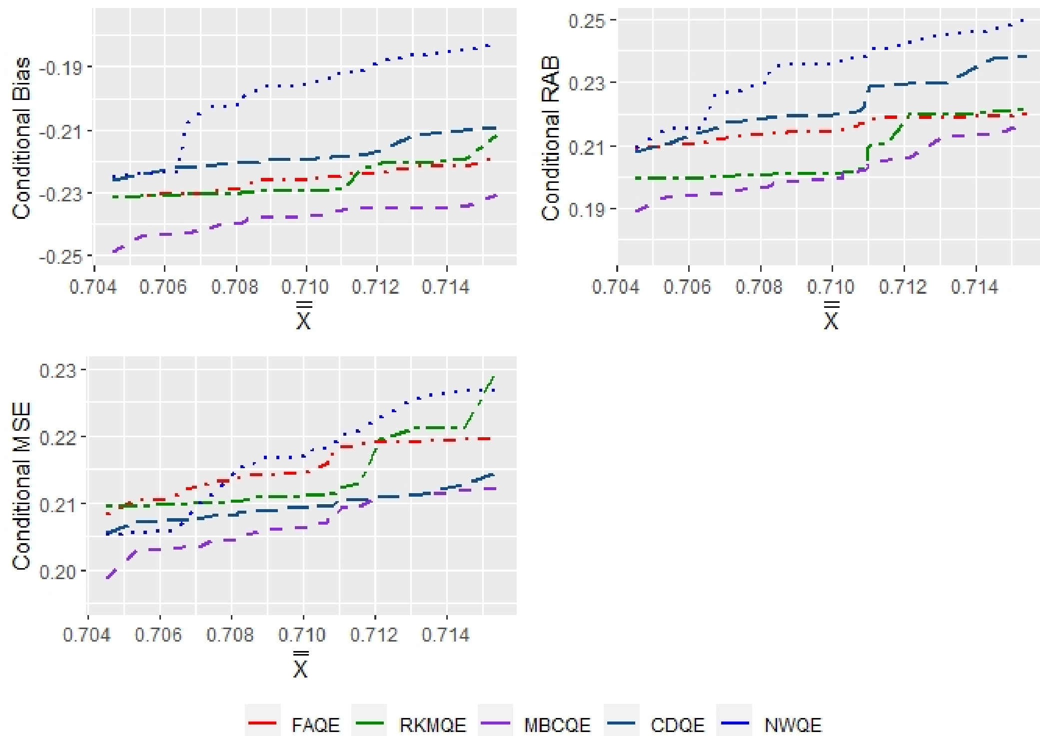

Figure 3: Plots of the behaviour of the CB, CRAB and CMSE for different estimators: $\alpha = 0.05$.

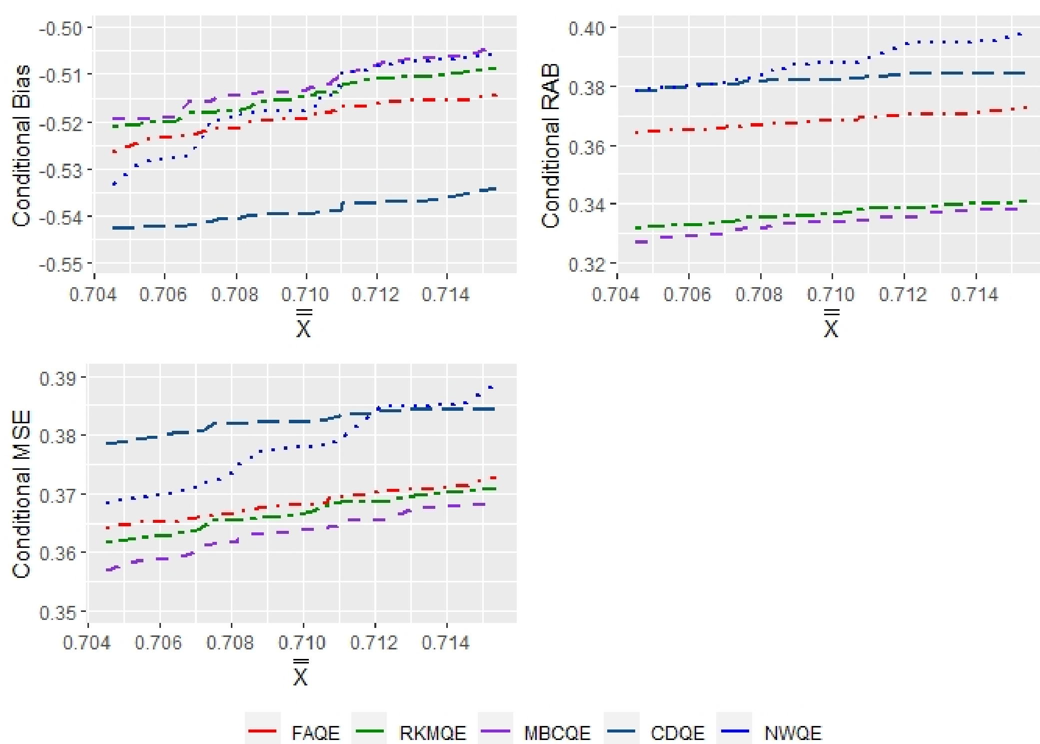

Figure 4: Plots of the behaviour of the CB, CRAB and CMSE for different estimators: $\alpha = 0.75$.

6. Conclusion

The quantile estimator based on simple random sampling without replacement has been developed. Investigation of the developed estimator's properties was done and discovered that it possesses asymptotic normal distributions. Under SRSWOR, its asymptotically unbiased estimator and asymptotically consistent estimator of population quantiles. It is clear from results that the quantile estimator based on SRSWOR results in a larger decrease of Bias than the one achieved by other estimators in literature used for comparison. In terms of performance, MBCQE has consistently produced results that are more precise than existing quantile estimators. We can therefore conclude that MBCQE can be used in estimating finite population quantiles for simple random sampling without replacement populations in various sectors since it yields very good results.

Other bias correction processes in quantile estimation, such as Adaptive Boosting and Bootstrap bias reduction methodologies, can be explored, as well as further research on the construction of confidence intervals for the recommended estimator.

References

1. Bahadur, R. R. (1966). A note on quantiles in large samples. *The Annals of Mathematical Statistics*, 37(3):577–580.
2. Chambers, R. L. and Dunstan, R. (1986). Estimating distribution functions from survey data. *Biometrika*, 73(3):597–604.
3. Chen, J. and Wu, C. (2002). Estimation of distribution function and quantiles using the model-calibrated pseudo empirical likelihood method. *Statistica Sinica*, pages 1223–1239.
4. Chen, S. X. (1999). Beta kernel estimators for density functions. *Computational Statistics & Data Analysis*, 31(2):131–145.
5. Chen, S. X. (2000). Beta kernel smoothers for regression curves. *Statistica Sinica*, pages 73–91.
6. Dorfman, A. H., Hall, P., et al. (1993). Estimators of the finite population distribution function using nonparametric regression. *The Annals of Statistics*, 21(3):1452–1475.
7. Falk, M. (1984). Relative deficiency of kernel type estimators of quantiles. *The Annals of Statistics*, pages 261–268.
8. Francisco, C. A. and Fuller, W. A. (1991). Quantile estimation with a complex survey design. *The Annals of statistics*, pages 454–469.
9. Jones, M. C. (1992). Estimating densities, quantiles, quantile densities and density quantiles. *Annals of the Institute of Statistical Mathematics*, 44(4):721–727.
10. Kiefer, J. (1967). On bahadur's representation of sample quantiles. *The Annals of Mathematical Statistics*, 38(5):1323–1342.
11. Kuk, A. Y. and Mak, T. (1989). Median estimation in the presence of auxiliary information. *Journal of the Royal Statistical Society: Series B (Methodological)*, 51(2):261–269.
12. Lio, Y. and Padgett, W. (1991). A note on the asymptotically optimal bandwidth for nadaraya's quantile estimator. *Statistics & Probability Letters*, 11(3):243–249.
13. Nadaraya, E. A. (1964). Some new estimates for distribution functions. *Theory of Probability & Its Applications*, 9(3):497–500.
14. Onsongo, W. M., Otieno, R. O., and Orwa, G. O. (2018). Bias reduction technique for estimating finite population distribution function under simple random sampling without replacement.
15. Padgett, W. J. (1986). A kernel-type estimator of a quantile function from right-censored data. *Journal of the American Statistical Association*, 81(393):215–222.
16. Parzen, E. (1979). Nonparametric statistical data modeling. *Journal of the American statistical association*, 74(365):105–121.
17. Rao, J., Kovar, J., and Mantel, H. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika*, pages 365–375.
18. Reiss, R.-D. (1980). Estimation of quantiles in certain nonparametric models. *The Annals of Statistics*, pages 87–105.
19. Serfling, R. (1980). Approximation theorems of mathematical statistics, john wiley and sons, new york.
20. Yang, S.-S. (1985). A smooth nonparametric estimator of a quantile function. *Journal of the American Statistical Association*, 80(392):1004–1011.