STATISTICS IN TRANSITION new series, December 2026
Vol. 27, No. 4, pp. xx–XX, https://doi.org/10.59139/stattrans-2026-xxx
Received – xx.xx.xxxx; accepted – xx.xx.xxxx

Quantile estimation for stratified random sampling using multiplicative semi-parametric bias reduction density

Dioggban Jakperik[1], **Michael Jackson Adjabui**[2], **Queensley C. Chukwudum**[3]

Abstract

In this study, a quantile estimator based on a multiplicative semi-parametric bias reduction density function is proposed for finite populations. Its theoretical properties have been studied and simulation studies performed to verify its practical potential. The proposed estimator performed better than the estimators that were compared to it. The simulation further showed that the estimates based on the proposed estimator improved as the sample size was increasing. In other words, the proposed estimator closely approximates the true values as sample size increases. For the sample of size 1,000, the bias reduction density estimators performed far better than the other estimators, with the proposed estimator turning out as the best. It is therefore desirable to use the proposed estimator when the sample size is relatively very large.

Key words: quantile, multiplicative semi-parametric, bias beduction, stratified random sampling.

1. Introduction

One of the most common sampling designs that fit so effectively in practice for many real-life applications without sacrificing precision is the stratified random sampling. This has many advantages as it allows for consideration of relevant population characteristics to obtain precise estimates of parameters. In the implementation of stratified random sampling, a heterogeneous population is subdivided into homogeneous non-overlapping units called strata and a simple random sampling scheme independently applied in each stratum to draw a sample of size n (Jakperik, Otieno and Okungu, 2021). The eventual estimate of the population parameter is a weighted quantity emanating from the subunits making up the whole population. The stratification variable(s) can take various forms and may be defined by the researcher or naturally occurring. Consequently, the stratification variable may be qualitative or quantitative.

In either case, stratification must be done in a way that enhances homogeneity within the strata and optimizes the precision of parameter estimates. Precision of parameter estimates is enhanced if the within strata variation is lower than the between strata variations,

[1]Biometry, University of Technology and Applied Sciences, Navrongo, Ghana.
E-mail: jdioggban@utas.edu.gh ORCID: https://orcid.org/0000-0003-3369-6123.
[2]Department of Mathematics, University of Technology and Applied Sciences, Navrongo, Ghana.
E-mail: madjabui@utas.edu.gh. ORCID: https://orcid.org/0000-0002-3078-9378.
[3]Department of Insurance and Risk Management, University of Uyo, Uyo, Nigeria.
E-mail: queensleyv@yahoo.com. ORCID: https://orcid.org/0000-0002-8998-1700.

© D. Jakperik, M. J. Adjabui, Q. C. Chukudum. Article available under the CC BY-SA 4.0 licence

otherwise the gain in precision for the stratified sample may not be superior to the simple random sample (Jakperik, Otieno and Okungu, 2021, Makumi, Otieno, Orwa and Habineza, 2022). This explains largely why most researchers prefer simple random sampling. Simple random sampling is suitable for homogeneous populations and generally applicable to large populations. Application of simple random sampling does not require sampling weights but frame construction. Homogeneous population means the elements of the population are generated by the same distribution function (Jakperik, Otieno and Okungu, 2021). Therefore, one needs a good distribution function to enhance the estimation of parameters. The most commonly used estimators are the kernel-based estimators (Watson, 1964, Nadaraya, 1964, Cochran, 1977). A kernel is a non-parametric smoothing technique commonly used in linearizing nonlinear estimators for improved estimation. Properties of the kernel estimator have been discussed in Section 2. Because the kernel estimator is a second order estimator and does not automatically correct for boundary bias, its precision is usually affected. Many researchers have proposed different estimators to improve the estimation precision, for instance, non-parametric biased corrected estimators (Makumi, Otieno, Orwa and Habineza, 2022, Abbas and Hani, 2020, Samawi and Al-Sagheer, 2001), and semiparametric estimators (Jakperik, Otieno, and Orwa, 2018). The semiparametric biased corrected estimators have been shown to be more efficient in estimating complex indices (Jakperik, Otieno, and Orwa, 2018). The issue of getting estimators with improved precision for finite population parameters has led researchers into exploring robust and parametric based mean and quantile estimators in slightly greater depth to enhance estimation precision. The quantile estimator, being robust to outliers and violations of assumptions of parametric distributions have been preferred by many researchers in proposing new estimators. Nevertheless, where the deviations are pervasive, the performance of those estimators are usually below expectations. Consequently, non-parametric and semi-parametric estimators have been proposed to improve the estimation precision (Eftekharian and Sawami, 2021, Makumi, Otieno, Orwa, and Habineza, 2022, Shahzad, Ahmad, Almanjahie, A-Noor, and Hanif, 2022). Although the alternative approaches improved estimation precision and bolstered confidence in the estimators based on the conventional methods, they are generally computationally intensive and time demanding (Shahzad, Ahmad, Almanjahie, A-Noor, and Hanif, 2022, Shahzad, Ahmad, Al-Noor, Hanif and Almanjahie, 2023). All of these exhibit significant limitations.

The rest of the paper is presented as follows: Section 2 contains notations and review of basic concepts on quantile estimation, Section 3 presents the Multiplicative Semi-parametric Bias Reduction density function used to develop the proposed estimator, while Sections 4 and 5 contain the simulation results and conclusions respectively.

2. Notations

Suppose U_N is a finite population subdivided into L distinctly unique strata with sizes N_h each such that $N = \sum_{h=1}^{L} N_h$. Each stratum contains unique observations of the variable of interest, Y with distinct indexes in the population. Let Y_{hi} be the measure of unit i in stratum h for $h = 1, \cdots, L$ and $i = 1, \cdots, N_h$. The distribution function, $F_N(t)$ of Y in the population

is given by

$$F_N(t) = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} I(Y_{hi} \leq t), \tag{1}$$

where

$$I = \begin{cases} 1, & \text{if } Y_{hi} \leq t, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

For each stratum, $F_N(t)$ is given by

$$F_{N_h}(t) = \frac{1}{N_h}\sum_{i=1}^{N_h} I(Y_{hi} \leq t), \tag{3}$$

given a population cdf as

$$F_N(t) = \frac{1}{N}\sum_{h=1}^{L} N_h F_{N_h}(t). \tag{4}$$

The population estimate of the p^{th} quantile is given by

$$Q(p) = \inf(t : F_N(t) \geq p), \tag{5}$$

with p taking on values between 0 and 1. In stratified random sampling, estimates are based on simple random samples independently drawn from the L strata. Thus, the sample estimate of the quantile function for the variable of interest is

$$F_n(t) = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_h I(y_{hi} \leq t), \tag{6}$$

where n_h is the size of the sample drawn from stratum h, y_{hi} is the realization of Y_{hi} based on the sample drawn from stratum h and $n = \sum_{h=1}^{L} n_h$. The weight for stratum h is given by $w_h = \frac{N_h}{n_h N} = \frac{W_h}{n_h}, W_h = \frac{N_h}{N}$, and

$$I = \begin{cases} 1, & \text{if } y_{hi} \leq t, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

The weighted empirical distribution function is thus given by

$$F_n(t) = \frac{1}{N}\sum_{h=1}^{L} N_h F_{n_h}(t), \tag{8}$$

where

$$F_{n_h} = \frac{1}{n_h}\sum_{h=1}^{n_h} I(y_{hi} \leq t). \tag{9}$$

Consequently, the sample estimate of the p^{th} quantile is given by

$$Q(p) = \inf(t : F_n(t) \geq p) \tag{10}$$

(Samawi, Chatterjee, Yin and Rochani, 2018).

In this paper, the following assumptions are used:

B_1. The underlying population of the h^{th} stratum has the cdf, $F_{N_h}(.)$ which is Holder continuous with a square-integrable second derivative for any $h = 1, \cdots, L$.

B_2. $K_h(.)$ is an absolutely continuous function with $\lim_{\to -a} K_h(x) = 0$ and $\lim_{\to a} K_h(x) = 1$, for $h = 1, \cdots, L$.

B_3. The Kernel function, $K_h(x)$, satisfies the following conditions for any x $K_h(x) = K_h(-x)$, $\int_{-a}^{a} K_h(x)\,dx = 1$, $\int_{-a}^{a} x^2 K_h(x)\,dx \neq 0$ (Abbas,and Hani, 2020, Jakperik, Otieno, and Orwa, 2018, Silvermann, 1986).

3. Multiplicative Semi-parametric Bias Reduction Density Function

The Multiplicative Semi-parametric Bias Reduction $(MSBR)$ density function proposed by Jakperik, Otieno, and Orwa (2018) is used to develop a quantile estimator. The distribution function is given by

$$\hat{f}(x)_{MSBR} = \frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)\frac{f\left(x,\hat{\theta}\right)}{f\left(X_i,\hat{\theta}\right)}, \tag{11}$$

where $\sum_{i=1}^{n} K_h(X_i - x), n, f\left(x,\hat{\theta}\right)$, and $f\left(X_i,\hat{\theta}\right)$ are the kernel density, sample size, non-parametric and parametric density estimators respectively (Jakperik, Otieno, and Orwa, 2018). This estimator automatically corrects for boundary bias, has a faster convergence rate, and attains optimum performance for large samples. The bias of the density estimator is given by

$$bias\left(\hat{f}(x)\right)_{MSBR} = \frac{1}{24}h^4\mu_4 f_0 r^{(iv)}(x) + o\left(h^4\right), \tag{12}$$

and the variance is given by

$$var\left(\left(\hat{f}(x)_{MSBR}\right)\right) = \frac{R(K)}{nh}f(x) - \frac{f(x)^2}{n}, \tag{13}$$

where $R(K) = \int_{-\infty}^{\infty} K(z)^2\,dz$ is the roughness of the kernel and $z = \frac{y-x}{h} \implies x = y + hz$. The theoretical properties and convergence of the density estimator have been studied (Jakperik, Otieno, and Orwa, 2018, Jakperik, Otieno, and Orwa, 2020). Suppose that Y_{hi} is the study variable; a variable that is of interest in the study, associated with the auxiliary variable, X_{hi}; a variable that is not the focus of the study but is related to the study variable and can enhance estimation precision of the study variable in the h^{th} stratum, associated with the auxiliary variable, X_{hi} in the h^{th} stratum. If a simple random sample of size n_h is taken from the h^{th} stratum of size N_h without replacement, the sampling fraction, $f_h = \frac{n_h}{N_h} \to 0$ as $n_h \to N_h$ and $N_h \to \infty$.

The empirical distribution function has been shown to converge in probability to the underlying distribution function in the stratum, thus $\hat{F}(x) \xrightarrow{p_r} F(x)$ and indeed is a strongly

consistent estimator for $F(x)$ (Bahadur, 1966). Because samples must be drawn from all strata, $F(x)$ is a weighted function of the empirical distribution functions from all strata, $F(x) = \sum_{h=1}^{L} W_h F_h(x)$, where W_h is a predetermined weight of stratum h (Silvermann, 1986, Bahadur, 1966). Let s be the sample with size n_h taken from stratum h through simple random sampling without replacement, and $j \in r \in h = N_h - s$ are non-sampled data from the h^{th} stratum. Assuming the units in the strata are generated by the super-population model

$$y_{hi} = m(x_{hi}) + \sigma(h_{hi})\,\varepsilon_{hi}, \tag{14}$$

where the error term is normal, independently and identically distributed with mean 0 and variance $\sigma_{x_{hi}}$, $\varepsilon_{hi} \sim NIID\left(0, \sigma_{x_{hi}}\right)$ with

$$Cov\left(y_{hi}, y_{hj}\right) = \begin{cases} \sigma^2(x_{hi}), & \text{if } i = 1, \cdots, N, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

(Fang and Yang, 2021, Makumi, Otieno, Orwa and Habineza, 2022).

The empirical distribution function for the stratified population is thus given by

$$F_{N_y}^{s}(t) = \frac{1}{N}\sum_{h=1}^{L} N_h \left(\frac{1}{N_h} \left[\sum_{j \in s \subset h} I(y_{hi} \le t) + \sum_{j \in r \subset h} I(y_{hi} \le t) \right] \right). \tag{16}$$

The first term in equation (16) is obtained from the sample while the second part is unknown. This must be estimated by a density function with good convergence to improve the estimation precision (Jakperik, Otieno, and Okungu, 2021). Given $0 < p < 1$, the p^{th} quantile for the underlying distribution is

$$\eta_{N_y}^{s}(p) = \inf\{t : F_N(t) \ge p\}, \tag{17}$$

and commonly defined as the inverse of the cumulative distribution function, $F_N^{-1}(p)$. The proposed quantile estimator based on the *MSBR* density is obtained as

$$\hat{\eta}_{MSBR}^{s} = \inf\{t \in U : \hat{F}_{MSBR}^{s-1}(t) > p\} = \hat{F}_{MSBR}^{s-1}(p), \tag{18}$$

where $\frac{1}{N} \le p \le \frac{N-1}{N}$. It is easy to see that $\hat{\eta}_{MSBR}^{s}$ satisfies the three assumptions listed in Section 2 above.

Substituting equation (18) into equation (16) and expanding around $\hat{\eta}_{MSBR}^{s}$ gives

$$\hat{F}_{MSBR}^{s}\left(\eta_{N_y}^{s}(p)\right) = p + \left[\hat{\eta}_{MSBR}^{s}(p)\,\eta_{N_y}^{s}(p)\right] f_{MSBR}^{s}\left(\eta_{N_y}^{s}(p)\right) + R_n. \tag{19}$$

where $F_{MSBR}^{s} = f_{MSBR}^{s} > 0$ and $R_n = O\left(n^{-\frac{1}{n}}\right)$ approaches 0 as $n \to \infty$ (Kiefer, 1967, Bahadur, 1966). Bahadur's representation (Bahadur, 1966) of the estimator is

$$\hat{\eta}_{MSBR}^{s}(p) = \eta_{N_y}^{s}(p) + \frac{p - \hat{F}_{MSBR}^{s}\left(\eta_{N_y}^{s}(p)\right)}{f_{MSBR}^{s}\left(\eta_{N_y}^{s}\right)} + R_n. \tag{20}$$

The expectation and variance of the proposed estimator are respectively

$$E\left[\hat{F}^{s}_{MSBR}\left(\eta^{s}_{N_y}(p)\right)\right] = Z_N = \frac{\sum_{i=1}^{N} Z_i}{N} = F^{s}_{N_y}\left(\eta^{s}_{N_y}\right) = p, \tag{21}$$

and

$$var\left[\hat{F}^{s}_{MSBR}\left(\hat{\eta}^{s}_{N_y}(p)\right)\right] = \frac{(1-f)\,p\,(1-p)}{n-1}, \tag{22}$$

following the findings of Francisco (1987). The theoretical properties such as unbiasedness, variance and mean-square error of this family of estimators are well studied and documented, see for example, (Abbas and Hani, 2020, Makumi,Otieno, Orwa and Habineza, 2022). The proposed estimator has a bias given by

$$bias\left(\hat{\eta}^{s}_{MSBR}(p)\right) = O\left(n^{-\frac{1}{2}}\right), \tag{23}$$

for $n \to \infty$, $O\left(n^{-\frac{1}{2}}\right)$ tails off to zero indicating that the proposed estimator is asymptotically unbiased. The mean square error for the estimator is given by

$$MSE = \left(\hat{\eta}^{s}_{MSBR}(p)\right) = \frac{1-f}{n-1}\frac{p\,(1-p)}{\left[f^{s}_{MSBR}\left(\eta^{s}_{N_y}\right)(p)\right]^2} + O\left(\frac{1}{n}\right). \tag{24}$$

The second part of equation (24) approaches zero as $n \to \infty$ and thus, the estimator is a consistent estimator; readers interested in the estimation procedure should read, for instance, (Makumi, Otieno, Orwa, and Habineza, 2022, Abbas and Hani, 2020, Chambers and Clarke, 2012, Serfling, 1980). The estimator for quantile estimation was developed using the multiplicative semi-parametric bias reduction density function, which has the ability to automatically correct boundary bias, produce estimates with optimum precision for large sample sizes and also perform satisfactorily for small and medium size samples. This density is then used to develop the quantile estimator thereby enhancing the usual characteristics of quantile estimators as being consistent and robust, especially where there are extreme values (Ayinzoya and Jakperik).

4. Simulation

A simulation study was conducted to assess the performance of the proposed estimator and compare it with its competitors. A population of 2000 units was simulated using a linear model given by

$$Y_i = 1 + 2\,(x_i + 0.5) + e_i, \tag{25}$$

and a non-linear population using the cosine function stated as

$$Y_i = \cos\left(1 + 2\,(x_i - 0.5)^2\right) + e_i. \tag{26}$$

The auxiliary variable was generated from a uniform distribution over the interval $[0,1]$. The error term, e_j was simulated from a standard normal distribution as $e_i \sim N(0,1)$. For each of the simulated populations, $Y_i's$ were divided into five homogeneous subgroups that are non-overlapping and of size 400 each. A sample of size $n = 200$ units was drawn from each of the populations for the five strata using proportional allocation and applying simple random sampling independently in all strata, resulted in the $n_h = 40$ units per stratum and replicated 500 times. Kernel smoothing was achieved through Epanecknikov kernel given by

$$K(v) = \frac{3}{4}\left(1-u^2\right) I_{\{|v|\leq 1\}}. \tag{27}$$

For each of the samples the proposed estimator was evaluated and each of the required model properties computed. The quantile function, which is the inverse of the $F_N(t)$, was evaluated using numerical methods since it is difficult to obtain an analytical solution for such a complex function. The proposed estimator, the Multiplicative Semi-parametric Bias Reduction Quantile Estimator $(MSBRQE)$, was compared to the Multiplicative Bias Corrected Quantile Estimator $(MBCQE)$ (Makumi, Otieno, Orwa, and Habineza, 2022) , the kernel-based estimator for stratified sampling with optimal allocation $(SCQE)$ by Sawami, Chatterjee, Yin and Rochani (2019), and the and kernel-based quantile estimator for different sampling schemes with optimum allocation $(ESQE)$ by Eftekharian and Sawami (2021). The results of the simulation are contained in Table 1.

Table 1. Unconditional biases, relative mean errors, and relative root-mean-square errors

		Linear			Cosine		
p	**Estimator**	**Bias**	**RME**	**RRMSE**	**Bias**	**RME**	**RRMSE**
0.25	ESQE	1.5021	1.5135	0.3203	0.3959	1.3901	0.5340
	SCQE	0.5271	1.2412	0.2567	1.5002	2.0140	0.4216
	MBCQE	0.3340	0.8201	0.1307	0.2132	0.9021	0.3407
	MSBRQE	0.3231	0.8102	0.1253	0.1210	0.7172	0.1388
0.50	ESQE	0.8910	0.8624	0.3462	0.5782	1.5481	0.3165
	SCQE	0.9942	1.0091	0.1986	0.9143	1.0112	0.2516
	MBCQE	0.6906	0.7576	0.1091	0.5256	0.5971	0.1634
	MSBRQE	0.6008	0.6515	0.0128	0.4320	0.5306	0.1574
0.75	ESQE	2.0162	1.9632	0.2698	2.0021	1.5924	0.4131
	SCQE	1.8379	1.3423	0.1907	1.6819	1.4866	0.3003
	MBCQE	1.7820	1.3106	0.1524	0.8342	1.0969	0.2134
	MSBRQE	1.6891	1.2120	0.1248	0.7855	0.8003	0.2014

The simulation results for the conditional biases, relative mean error, and relative root mean square error were implemented using a sample size, $n = 200$. The $MSBRQE$ is consistently providing smaller biases among the estimators for all levels, $p = 0.25, 0.5$, and 0.75. Both the $MBCQE$ and the $MSBRQE$ were far better in terms of bias, relative mean error, and relative root mean square error than both the $ESQE$ and $SCQE$.

The four estimators were then applied on the two independently simulated populations in turns to estimate the quantiles for $p = 0.25, 0.05,$ and 0.75 respectively for sample sizes $n = 200, 400, 1000$.

Table 2. Quantile Estimates of Estimators for $n = 200$

	Quantile Estimates			
Mean Function	Estimators	0.25	0.50	0.75
	Q(p)	0.2635	1.0393	1.9352
linear	ESQE	0.1980	0.9431	1.7869
	SCQE	0.3241	1.2460	1.8762
	MBCQE	0.2607	1.1740	1.8321
	MSBRQE	0.2623	1.0752	1.9214
Cosine	Q(p)	0.2839	0.4444	0.5199
	ESQE	0.1682	0.2258	0.4163
	SCQE	0.3142	0.5231	0.6723
	MBCQE	0.2319	0.2960	0.4185
	MSBRQE	0.2471	0.3802	0.4783

Generally, the *MSBRQE* closely estimated the true quantiles for $n = 200$ for both the linear and non-linear populations as seen in Table 2. The estimates for the *MSBRQE* were best followed by the *MBCQE*, and the *SCQE* for sample size $n = 200$. This is expected as the *MSBRQE* is a large sample estimator.

Table 3. Quantile Estimates of Estimators for $n = 400$

	Quantile Estimates			
Mean Function	Estimators	0.25	0.50	0.75
linear	Q(p)	0.3008	1.0246	1.7865
	ESQE	0.2461	0.9984	1.3730
	SCQE	0.4190	1.0508	1.3209
	MBCQE	0.2821	1.0147	1.6801
	MSBRQE	0.2973	1.0198	1.7253
Cosine	Q(p)	0.2789	0.4336	0.5148
	ESQE	0.17031	0.9573	0.4781
	SCQE	0.2650	0.3345	0.4092
	MBCQE	0.2657	0.4103	0.4721
	MSBRQE	0.2708	0.4289	0.4941

The improvement in the performance of the *MBCQE* and *MSBQE* was enhanced with increasing sample sizes. As the sample size increased to $n = 400$, the *MSBRQE* performed better, thus enhancing its estimates better than the *MSBCQE*, especially for $p = 0.5, 0.75$ as contained in Table 3. The other two estimators, however, performed poorly compared to the bias reduction estimators.

Table 4. Quantile Estimates of Estimators for $n = 1000$

		Quantile Estimates		
Mean Function	Estimators	0.25	0.50	0.75
linear	Q(p)	0.2123	1.0144	1.8447
	ESQE	0.1681	0.5932	1.6812
	SCQE	0.3093	1.1530	1.3705
	MBCQE	0.2281	0.9521	1.8110
	MSBRQE	0.2159	1.0091	1.8367
Cosine	Q(p)	0.2746	0.4293	0.5147
	ESQE	0.1306	0.2720	0.3209
	SCQE	0.3310	0.3604	0.2928
	MBCQE	0.2599	0.4108	0.4951
	MSBRQE	0.2678	0.4289	0.5059

For sample size $n = 1000$, the bias reduction estimators, *MBCQE* and *MSBRQE*, improved in precision greatly but were not exactly approximating the true quantiles. The differences in estimates made them preferred over the kernel-based estimators as they yielded estimates with small disparities from the true quantiles. The detailed results can be found in Table 4.

The estimators were then applied to the Net Income variable of the Ghana Living Standards Survey Round 7 data to assess its practical potential. The proposed estimator, *MSBRQE*, closely estimated the true quantiles with good precision. The *MBCQE* was the second best estimator on the application data followed by the kernel-based estimators as demonstrated in the simulation studies. The proposed bias reduction estimator, *MSBRQE*, provides estimates which are very close to the true quantile estimates.

Table 5. Quantile Estimates of Estimators for the GLSS7 Data

		Quantile Estimates		
Mean Function	Estimators	0.25	0.50	0.75
linear	Q(p)	932.6714	3329.9727	9335.6736
	ESQE	847.9752	2870.8052	8865.4071
	SCQE	876.5041	3029.5413	8965.6056
	MBCQE	910.5219	3107.8114	8992.4429
	MSBRQE	925.6984	3298.1332	9209.9870

5. Conclusions

A new quantile estimator for stratified random sampling is proposed. The simulation studies revealed that the proposed estimator has desirable properties comparable to the best performing quantile estimators. The optimum properties of the estimator are obtained when the sample size is significantly large. The proposed estimator automatically corrects bound-

ary bias and thus commands high precision relative to its competitors. For large samples, the computational time can considerably increase.

Further studies can be performed by constructing confidence intervals for the proposed estimator, estimating quantiles using the proposed estimator for different types of complex sampling schemes such as multistage sampling, multiphase sampling, among others. Also, the multiplicative semi-parametric bias reduction density used for constructing the proposed estimator can be extended to its bivariate form and used to estimate copulas for applications in survival analysis.

Acknowledgements

The authors thank the Ghana Statistical Service for granting permission to use the Ghana Living Standards Survey Round 7 data for the applications.

References

Abbas, E., Hani S., (2020). On Kernel-Based Quantile Estimation Using Different Stratified Sampling Scheme with Optimal Allocation. *Journal of Statistical Computation and Simulation*. https://doi.org/10.1080/00949655.2020.1839900.

Ayinzoya, R., Jakperik, D., (2023). Estimation of Finite Population Mean under Probability-Proportional-to-Size Sampling in the Presence of Extreme Values. *International Journal of Mathematics and Mathematical Sciences*. https://doi.org/10.1155/2023/3064736.

Bahadur, R. R., (1966). A note on quantiles in large samples. *Annals of Mathematical Statistics*, 37, pp. 577–580.

Chambers, R., Clark, R., (2012). *An Introduction to Model-Based Survey Sampling with Applications*, OUP, Oxford, UK.

Cochran, William, G., (1977). *Sampling Techniques*, John Wiley & Sons.

Eftekharian, A., Sawami H., (2021). On kernel-based quantile estimation using different stratified sampling schemes with optimal allocation. *Journal of Statistical Computation and Simulation*, 91(5), pp. 1040–1056.

Fang, Z., Qi, L. and Yang, K, X., (2021). A Simple Non-parametric Approach for Estimation and Inference of Conditional Quantile Functions. *Econometric Theory*. https://doi.org/10.1017/S0266466621000499.

Francisco C. A., (1987). *Estimation of Quantiles and the Interquartile Range in Complex Surveys*, PhD Thesis, Iowa State University, Ames, IA, USA.

Jakperik, D., Otieno, R. Odhiambo and Okungu, J. O., (2020). A Note on Nonparametric Regression Modeling Using a Density Function. *African Journal of Applied Statistics*, 7(2), pp. 993–1000.

Jakperik, D., Otieno, R. Odhiambo and Okungu, J. O., (2021). Estimating a Finite Population Total using a Density Function. *Eurasian Bulletin of Mathematics*, 4(1), pp. 32–37.

Jakperik, D., Otieno, R. Odhiambo and Orwa, G. O., (2018). A semi-parametric multiplicative bias reduction density with a parametric start. *Advances and Applications in Statistics*, 53(6), pp. 715–729.

Jakperik, D., Otieno, R. Odhiambo and Orwa, G. O., (2018). Variance Estimation for Poverty Indicators via Linearization Techniques. *Journal of Computations and Modeling*, 9(1), pp. 1–9.

Kiefer, J., (1967). On Bahadur's representation of sample quantiles. *The Annals of Mathematical Statistics*, 38(5), pp. 1323–1342.

Makumi, N., Otieno, R. Odhiambo, Orwa, G. O. and Habineza, A., (2022). On Quantile Estimation Based on Stratified Sampling Using Multiplicative Bias Correlation Approach. *Journal of Mathematics*.

Nadaraya E. A., (1964). Some New Estimates for Distribution Functions. *Theory of Probability and its Applications*, 9(3), pp. 497–500.

Samawi H. M., Al-Sagheer O. A., (2001). On the Estimation of the Distribution Function using Extreme and Median Ranked Set Sampling. *Biometrical Journal*, 43(3), pp. 357–373.

Samawi, H., Chatterjee,A., Yin, J. and Rochani, H., (2018). On quantiles estimation based on different stratified sampling with optimal allocation, *Communications in Statistics - Theory and Methods*. https://doi.org/10.1080/03610926.2018.1433856.

Sawami, H., Chatterjee, A., Yin, J. and Rochani H., (2019). On quantiles estimation based on different stratified sampling with optimal allocation. *Communications in Statistics - Theory and Methods*, 48(6), pp. 1529–1544.

Serfling, R., (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, NY, USA. Shahzad, U. Shazadi, S., Afshan, N.,Al-Noor, N. H., Alilah, D.,Hanif, M. and Anas M. M., (2021).Poisson regression based mean estimator. *Mathematical Problems in Engineering*, 2021(2), pp. 1–6.

Shahzad, U., Ahmad, I., Almanjahie, I.,A-Noor, N. H., and Hanif, M. (2022). Mean estimation using robust quantile estimation regression with two auxilliary variables. *Scientia Iranica*. https://doi.org/10.24200/SCI.2022.57170.5098.

Shahzad, U., Ahmad, I., Al-Noor, N. H., Hanif, M. and Almanjahie, I. M.,(2023). Robust estimation of the population mean using quantile regression under systematic sampling. *Mathematical Population Studies*, 30(3), pp. 195–207.

Shahzad, U., Ahmad, I., Al-Noor, N. H., Iftikhar, S., Abd Ellah, A. H. Benedict, T. J. and others, (2022). Särndal Approach and Separate Type Quantile Robust Regression Type Mean Estimators for Nonsensitive and Sensitive Variables in Stratified Random Sampling. *Journal of Mathematics*

Shahzad, U., Al-Noor, N. H., Afshan, N., Alilah, D., Hanif, M. and Anas, M. M., (2021). Minimum covariance determinant based quantile robust regression-type estimators for mean parameter. *Mathematical Problems in Engineering*, 2021, pp. 1–8. https://doi.org/10.1155/2021/5255839.

Silverman, B. W., (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Watson, G., (1964). Smooth regression analysis. *Shankya A.* 26, pp. 359–372.