Scientia Iranica E (2023) 30(3), 1245–1254

SCIENTIA IRANICA

Sharif University of Technology
**Scientia Iranica**
*Transactions E: Industrial Engineering*
http://scientiairanica.sharif.edu

# Mean estimation using robust quantile regression with two auxiliary variables

**U. Shahzad[a,b,*], I. Ahmad[a], I.M. Almanjahie[c,d], N.H. Al-Noor[e], M. Hanif[b]**

a. *Department of Mathematics and Statistics, International Islamic University, Islamabad 44000, Pakistan.*
b. *Department of Mathematics and Statistics - PMAS-Arid Agriculture University, Rawalpindi 46300, Pakistan.*
c. *Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia.*
d. *Statistical Research and Studies Support Unit, King Khalid University, Abha 62529, Saudi Arabia.*
e. *Department of Mathematics, College of Science, Mustansiriyah University, Baghdad 10011, Iraq.*

Received 14 November 2020; received in revised form 5 September 2020; accepted 5 September 2022

**KEYWORDS**
Quantile regression;
Robust measures;
Mean square error;
Simple random sampling;
Auxiliary information.

**Abstract.** In the presence of outliers in the data set, the utilization of robust regression tools for mean estimation is a widely established practice in survey sampling with single auxiliary variable. Recently, with the aid of some non-conventional location measures and traditional Ordinary Least Square (OLS), proposed a class of mean estimators using information on two supplementary variates under a simple random sampling framework. The utilization of non-traditional measures of location, especially in the presence of outliers, performed better than existing conventional estimators. In this study, a new class of estimators of mean utilizing quantile regression is proposed. The general forms of Mean Square Error (MSE) and Minimum Mean Square Error (MMSE) are also derived. The theoretical findings are being reinforced by different real-life data sets and simulation study.

© 2023 Sharif University of Technology. All rights reserved.

## 1. Introduction

The success of all kinds of development plans and projects in any country depends on efforts made in preparing, following up, and evaluating these plans. Consequently, data and information play an essential and pivotal role for decision-makers. The data and information can be provided in two ways: sampling surveys and comprehensive surveys. The sampling surveys gained popular acceptance being less expensive than comprehensive surveys of the finite population (see, [1,2]). The challenge faced by sampling survey is to find the best possible estimates based on existing data and any appropriate auxiliary or supplementary information that can help to improve the estimation and enhance the information about population. One of the primary concerns of survey sampling is mean estimation that can be improved by utilizing auxiliary information (see, among others, [3–16]).

On a variety of topics, numerous surveys have been conducted for the sake of data collection. Such surveys have become an accepted part of modern life. However, survey results have been increasingly influenced by growing trends in non-responses, with loss of accuracy. The presence of an unwanted number of irregular response outliers in the data may lead to false results. Outliers can change the regression parameters and degree of accuracy compared to those parameters

*. *Corresponding author. Tel.: 00923315995587*
*E-mail addresses: usman.stat@yahoo.com (U. Shahzad); ishfaq.ahmad@iiu.edu.pk (I. Ahmad); imalmanjahi@kku.edu.sa (I.M. Almanjahie); nadialnoor@uomustansiriyah.edu.iq (N.H. Al-Noor); mhpuno@hotmail.com (M. Hanif)*

doi: 10.24200/sci.2022.57170.5098

evaluated without such outliers (see, [17,18]). This indicates that a reasonable measure is required when the data is contaminated by the presence of outliers. Quantile Regression (QR) has emerged as a useful supplement to standard regression technique. Further, it is also robust to outliers. Boscovich laid down QR in the eighteenth century even before the establishment of Ordinary Least Square (OLS) regression estimators (see, [19,20]). For many decades, the use of QR was restricted due to financial issues or natural research, but now, it is used virtually in all fields of social and economic sciences. In light of the preceding lines, we are introducing the QR coefficient rather than OLS in mean estimation.

In this article, we propose a new family of estimators for estimation of population mean via a more scrupulous use of auxiliary variables. This objective is achieved by considering QR. The applicability of the scheme is demonstrated under simple random sampling framework by employing different data sets from various fields of inquiries. Moreover, a keen comparative investigation was carried out on the estimators of Abid et al. [1] and the proposed family by means of numerical evaluations.

The rest of this article is structured as follows: A brief review of the assigned to Abid et al. [1] is given in Section 2. A new class of quantile-regression-type estimators is proposed in Section 3. Moreover, expressions of large sample properties, namely Mean Square Error (MSE) and Minimum Mean Square Error (MMSE), are shown in Section 3. In Section 4, various numerical comparisons are made between real-life data sets and simulation study in order to shed light on the performance of the proposed estimators with respect to competitive estimators. The concluding remarks are given in Section 5.

## 2. Existing family of estimators

Let $\{\Omega = 1, \cdots, N\}$ be a finite population of $N$ identifiable units, and ($y_{>}0$, $x_1 > 0$, $x_2 > 0$) represent the subject variable ($Y$) and the supplementary variables ($X_1$, $X_2$), respectively. Suppose a sample of size $n$ be drawn from the population under Simple Random Sampling Without Replacement (SRSWOR) and let $\lambda = (\frac{1}{n} - \frac{1}{N})$. Abid et al. [1] introduced a class of estimators for estimating population mean as follows:

$$t_{a(i)} = Q_1 \bar{y}\left(\frac{\bar{X}_1\psi_{1(i)}+\beta_{1(i)}}{\bar{x}_1\psi_{1(i)}+\beta_{1(i)}}\right) + Q_2\bar{y}\left(\frac{\bar{X}_2\psi_{2(i)}+\beta_{2(i)}}{\bar{x}_2\psi_{2(i)}+\beta_{2(i)}}\right),$$
$$\text{for} \quad i = 1, 2, \cdots, 16, \tag{1}$$

where $\psi_{1(i)}$, $\beta_{1(i)}$, $\psi_{2(i)}$, and $\beta_{2(i)}$ represent known non-conventional and conventional measures of $X$ location such as mid-range ($MR_{x1}$, $MR_{x2}$), Hodges-Lehmann ($HL_{x1}$, $HL_{x2}$), tri-mean ($TM_{x1}$, $TM_{x2}$) and decile-mean ($DM_{x1}$, $DM_{x2}$), coefficient of variation ($C_{x1}$, $C_{x2}$), and coefficient of kurtosis ($\beta_2(x_1)$, $\beta_2(x_2)$) of the first and second supplementary variables. The correlation coefficients for ($Y$, $X_1$, $X_2$) are denoted as ($\rho_{yx1}$, $\rho_{yx2}$, $\rho_{x1x2}$). The sample means of ($Y$, $X_1$, $X_2$) are $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$. Further, ($\bar{X}_1$, $\bar{X}_2$) represent the population means of the first and second supplementary variables, respectively. Two tuning parameters, $Q_1$ and $Q_2$, are attached for minimizing the MSE of $t_{a(i)}$. All the family members determined by Abid et al. [1] are listed in Table 1. The MSE of $t_{a(i)}$ is given by:

**Table 1.** Family members of mean estimators of Abid et al. [1].

| $t_{a(i)}$ | $\psi_{1(i)}$ | $\beta_{1(i)}$ | $\psi_{2(i)}$ | $\beta_{2(i)}$ |
|---|---|---|---|---|
| $t_{a(1)}$ | 1 | $MR_{x1}$ | 1 | $MR_{x2}$ |
| $t_{a(2)}$ | 1 | $TM_{x1}$ | 1 | $TM_{x2}$ |
| $t_{a(3)}$ | 1 | $HL_{x1}$ | 1 | $HL_{x2}$ |
| $t_{a(4)}$ | 1 | $DM_{x1}$ | 1 | $DM_{x2}$ |
| $t_{a(5)}$ | $\beta_2(x_1)$ | $MR_{x1}$ | $\beta_2(x_2)$ | $MR_{x2}$ |
| $t_{a(6)}$ | $\beta_2(x_1)$ | $TM_{x1}$ | $\beta_2(x_2)$ | $TM_{x2}$ |
| $t_{a(7)}$ | $\beta_2(x_1)$ | $HL_{x1}$ | $\beta_2(x_2)$ | $HL_{x2}$ |
| $t_{a(8)}$ | $\beta_2(x_1)$ | $DM_{x1}$ | $\beta_2(x_2)$ | $DM_{x2}$ |
| $t_{a(9)}$ | $C_{x_1}$ | $MR_{x1}$ | $C_{x_2}$ | $MR_{x2}$ |
| $t_{a(10)}$ | $C_{x_1}$ | $TM_{x1}$ | $C_{x_2}$ | $TM_{x2}$ |
| $t_{a(11)}$ | $C_{x_1}$ | $HL_{x1}$ | $C_{x_2}$ | $HL_{x2}$ |
| $t_{a(12)}$ | $C_{x_1}$ | $DM_{x1}$ | $C_{x_2}$ | $DM_{x2}$ |
| $t_{a(13)}$ | $\rho_{yx_1}$ | $MR_{x1}$ | $\rho_{yx_2}$ | $MR_{x2}$ |
| $t_{a(14)}$ | $\rho_{yx_1}$ | $TM_{x1}$ | $\rho_{yx_2}$ | $TM_{x2}$ |
| $t_{a(15)}$ | $\rho_{yx_1}$ | $HL_{x1}$ | $\rho_{yx_2}$ | $HL_{x2}$ |
| $t_{a(16)}$ | $\rho_{yx_1}$ | $DM_{x1}$ | $\rho_{yx_2}$ | $DM_{x2}$ |

$$\begin{aligned} MSE(t_{a(i)}) = & \lambda\bar{Y}^2(C_y^2 + Q_1^2\vartheta_1^2C_{x_1}^2 + Q_2^2\vartheta_2^2C_{x_2}^2 \\ & -2Q_1\vartheta_1\rho_{yx_1}C_yC_{x_1} - 2Q_2\vartheta_2\rho_{yx_2}C_yC_{x_2} \\ & + 2Q_1Q_2\vartheta_1\vartheta_2\rho_{x_1x_2}C_{x_1}C_{x_2}), \end{aligned} \tag{2}$$

where the optimum value of $Q_1$ obtained by Eq. (3) as shown in Box I. Using the unity condition of weights $Q_1^* + Q_2^* = 1$, they found $Q_2^* = 1 - Q_1^*$.

## 3. QR-type estimators

Outliers are the observations in a data set that appear to be inconsistent with the rest of that data set. The presence of outliers significantly affects the mean estimation, which is one of the most important measures of central tendency. Mean estimators using OLS regression coefficient are the most ideal choices for the estimation of population mean, i.e., $\bar{Y}$. However, outliers may have a significant impact

$$Q_1^* = \frac{\vartheta_2^2 C_{x_2}^2 + \vartheta_1 \rho_{yx_1} C_y C_{x_1} - \vartheta_1 \vartheta_2 \rho_{x_1 x_2} C_{x_1} C_{x_2} - \vartheta_2 \rho_{yx_2} C_y C_{x_2}}{\vartheta_1^2 C_{x_1}^2 + \vartheta_2^2 C_{x_2}^2 - 2\vartheta_1 \vartheta_2 \rho_{x_1 x_2} C_{x_1} C_{x_2}}. \quad (3)$$

Box I

on the traditional regression coefficient calculated from OLS tool. Hence, the estimate of the population mean, i.e., $\bar{y}$, based upon OLS may indicate poor performance. One of the natural solutions is to adopt quantile regression. It can be used as a robust approach in circumstances when data is non-normal and contaminated with outliers. Further, it is also robust to outliers [19,21]. QR is similar to the customary least squares regression (OLS-R) in the sense that both of them explore connections among endogenous and exogenous variables. The primary distinction between the two is that OLS-R selects the parameter estimates with the least squared deviation from the regression line, while QR selects parameter estimates that have the least absolute deviation from the regression line. Therefore, in this study, we are proposing a group of quantile regression-type estimators by extending the idea of Abid et al. [1], as given below:

$$t_{p1} = Q_1\left[\bar{y} + b_{y.x1(0.10)}\left(\bar{X}_1 - \bar{x}_1\right)\right] + Q_2\left[\bar{y} + b_{y.x2(0.10)}\left(\bar{X}_2 - \bar{x}_2\right)\right], \quad (4)$$

$$t_{p2} = Q_1\left[\bar{y} + b_{y.x1(0.15)}\left(\bar{X}_1 - \bar{x}_1\right)\right] + Q_2\left[\bar{y} + b_{y.x2(0.15)}\left(\bar{X}_2 - \bar{x}_2\right)\right], \quad (5)$$

$$t_{p3} = Q_1\left[\bar{y} + b_{y.x1(0.25)}\left(\bar{X}_1 - \bar{x}_1\right)\right] + Q_2\left[\bar{y} + b_{y.x2(0.25)}\left(\bar{X}_2 - \bar{x}_2\right)\right], \quad (6)$$

$$t_{p4} = Q_1\left[\bar{y} + b_{y.x1(0.35)}\left(\bar{X}_1 - \bar{x}_1\right)\right] + Q_2\left[\bar{y} + b_{y.x2(0.35)}\left(\bar{X}_2 - \bar{x}_2\right)\right]. \quad (7)$$

In general, we can write the proposed family of estimators as follows:

$$t_{pi} = Q_1\left[\bar{y} + b_{y.x1(q)}\left(\bar{X}_1 - \bar{x}_1\right)\right] + Q_2\left[\bar{y} + b_{y.x2(q)}\left(\bar{X}_2 - \bar{x}_2\right)\right] \quad \text{for} \quad i = 1, 2, \cdots, 4, \quad (8)$$

with:

$$b_{y.x1(q)} = \text{argmin}_{\beta \epsilon R^p} \rho_q(v) \sum_{i=1}^{n} (y_i - \langle x_{1i}, \beta \rangle),$$

$$b_{y.x2(q)} = \text{argmin}_{\beta \epsilon R^p} \rho_q(v) \sum_{i=1}^{n} (y_i - \langle x_{2i}, \beta \rangle),$$

where $\rho_q(v)$ is a continuous piecewise linear function (or asymmetric absolute loss function) for quantile $q\epsilon(0,1)$, but nondifferentiable at $v = 0$. Note that all the notations of $\bar{y}_{p_i}$ have usual meanings as discussed in the previous section. However, $(b_{y.x1(q)},\ b_{y.x2(q)})$ are the QR coefficients. For a deep study of QR, interested readers may refer to Koenker and Hallock [22].

It is worth mentioning that we are using $q^{10\text{th}} = 0.10$, $q^{15\text{th}} = 0.15$, $q^{25\text{th}} = 0.25$, and $q^{35\text{th}} = 0.35$ quantiles for the purposes of the current article. We see from the consequences of the numerical study conducted in Section 4 that utilizing the QR coefficients, based on these referenced quantiles, will incredibly enhance the efficiencies of the proposed estimators. Note that the proposed class comprises four members based on these four referenced quantiles.

To obtain MSE, let us define $\bar{y} = (1 + \eta_y)\bar{Y}$, $\bar{x}_1 = (1 + \eta_{x1})\bar{X}_1$, and $\bar{x}_2 = (1 + \eta_{x2})\bar{X}_2$. Utilizing these notations $\eta_i$ $(i = y, x_1, x_2)$, we can write $E(\eta_y) = E(\eta_{x1}) = E(\eta_{x2}) = 0$, $E(\eta_y^2) = \lambda C_y^2$, $E(\eta_{x1}^2) = \lambda C_{x1}^2$, $E(\eta_{x2}^2) = \lambda C_{x2}^2$, $E(\eta_y \eta_{x1}) = \lambda C_{yx1}$, $E(\eta_y \eta_{x2}) = \lambda C_{yx2}$ and $E(\eta_{x1}\eta_{x2}) = \lambda C_{x1x2}$. Now, expanding $t_{pi}$ in terms of $\eta_y$, $\eta_{x1}$, and $\eta_{x2}$ as:

$$t_{pi} = \left[Q_1 \bar{Y}(1 + \eta_y) - b_{y.x1(q)}\bar{X}_1 \eta_{x1}\right] + \left[Q_2 \bar{Y}\left(1 + \eta_y\right) - b_{y.x2(q)}\bar{X}_2 \eta_{x2}\right]. \quad (9)$$

Eq. (9), applying expectation, we get a theoretical MSE of the estimator $t_{pi}$ up to the order $n^{-1}$ as follows:

$$MSE(t_{pi}) = \bar{Y}^2 + Q_1^2 \tau_A + Q_2^2 \tau_B + 2Q_1 Q_2 \tau_C - 2Q_1 \tau_D - 2Q_2 \tau_E, \quad (10)$$

where:

$$\tau_A = \left[\bar{Y}^2 + \lambda\left\{S_y^2 + B_{y.x1(q)}(B_{y.x1(q)} S_{x1} - 2\rho S_y) S_{x1}\right\}\right],$$

$$\tau_B = \left[\bar{Y}^2 + \lambda\{S_y^2 + B_{y.x2(q)}(B_{y.x2(q)} S_{x2} - 2\rho S_y) S_{x2}\}\right],$$

$$\tau_C = \left[\bar{Y}^2 + \lambda\{S_y^2 - B_{y.x2(q)}\rho_{y.x2} S_y S_{x2} - B_{y.x1(q)}\rho_{y.x1} S_y S_{x1} + B_{y.x1(q)} B_{y.x2(q)} \rho_{x1.x2} S_{x1} S_{x2}\}\right],$$

$$\tau_D = \tau_E = \bar{Y}^2.$$

By partially differentiating Eq. (10) with respect to $Q_1$ and $Q_2$, we obtained the optimum values as given by:

$$Q_1^{opt} = \left[\frac{\tau_B \tau_D - \tau_C \tau_E}{\tau_A \tau_B - \tau_C^2}\right],$$

and:

$$Q_2^{opt} = \left[\frac{\tau_A\tau_E - \tau_C\tau_D}{\tau_A\tau_B - \tau_C^2}\right].$$

Substitution of $Q_1^{opt}$ and $Q_2^{opt}$ into Eq. (10) provides the minimum MSE of $t_{pi}$ as follows:

$$MSE_{\min}(t_{pi}) = \left[\bar{Y}^2 - \frac{\tau_B\tau_D^2 - 2\tau_C\tau_D\tau_E + \tau_A\tau_E^2}{\tau_A\tau_B - \tau_C^2}\right]. \quad (11)$$

## 4. Numerical illustration

The current section is based on the performance evaluation of various estimators. The accuracy of the recommended estimators is assessed carefully in comparison with previous estimators. In Subsections 4.1 and 4.2, two data sets are considered: first, real-life data set free from outliers and the other real-life data set with outliers.

Generally, in the evaluation of new proposals, it is typical to adopt analytical derivations and find some conditions. These conditions help declare the superiority of one estimator over the others. However, the difficulty to verify these conditions is one of the most difficult problems associated with such analytical derivations. Their use in practice remains doubtful unless universal supremacy of a specific proposal has been defined with certainty. In addition, when parameters of interest for a population are not known, conditions that hold theoretically may not be fulfilled in real analyses. In light of such circumstances, yet, we have deliberately avoided creating comparisons on a theoretical basis with other considered estimators based on the MSE. Therefore, we conduct a variety of numerical-simulations in Subsection 4.3. Finally, we discuss the result of numerical illustrations in Subsection 4.4.

### *4.1. Practical study (Pop-1)*

In this subsection, we consider the Iris Data Set (IDS). The Iris "flower" data set, also named Fisher's IDS or sometimes Anderson's IDS, introduced by Fisher [23] and collected by Anderson [24] to quantify the variation of Iris flowers of different related species. IDS is one of the most well-known multivariate datasets used in data mining. It contains measurements "in centimeters" for the four variables including sepal length and width and petal length and width for 150 flowers from three species of Iris "Setosa, Versicolor, and Virginica". The study variable $Y$ is taken as "Sepal length", and the auxiliary variables $X_1$ and $X_2$ are taken as "Petal width" and "Petal length", respectively.

### *4.2. Robustness study of the proposed estimators (Pop-2)*

As in the earlier sections, it is mentioned that quantile regression is robust against outliers. Thus, in case outliers exist in the data, quantile regression coefficient performs efficiently as compared to other measures of locations. Thus, in the current sub-section, our recommended estimators are evaluated in case of outliers. For this purpose, the data set of Sukhatme and Sukhatme is considered [25]. Herein, $Y$ is taken as "area (acres) under wheat in 1937", $X_1$ is taken as "area (acres) under wheat in 1936", and $X_2$ is taken as "total cultivated area (acres) in 1931". Figure 1 shows non-normality. Box-plots and Scatter-plots in Figures 2, 3 and 4 point to the presence of outliers, individually and in combination, in $Y$, $X_1$, and $X_2$, respectively.

Hence, Pop-2 is suitable for the utilization of non-traditional measures as Abid et al. [1] and for the proposed class containing quantile regression. The remaining characteristics of all the two populations are provided in Table 2. The findings of Percentage Relative Efficiency (PRE) of the new class with respect to the estimators of Abid et al. [1] are presented numerically in Tables 3 and 4.

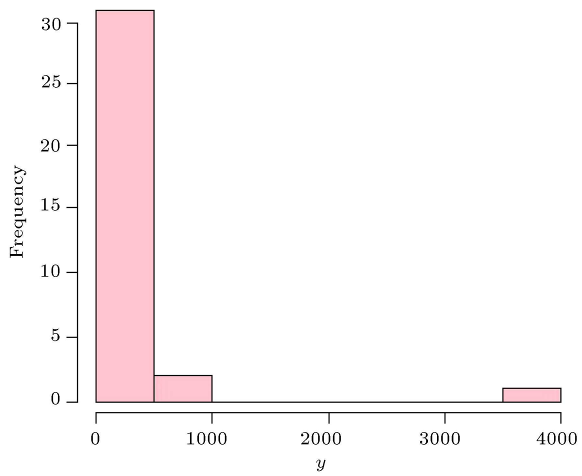


**Figure 1.** Histogram of study variable $Y$ of Population-2.

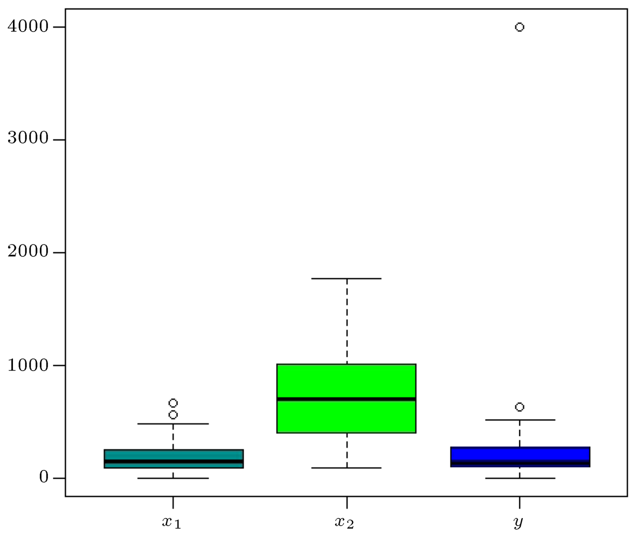


**Figure 2.** Box-plot of Population-2.

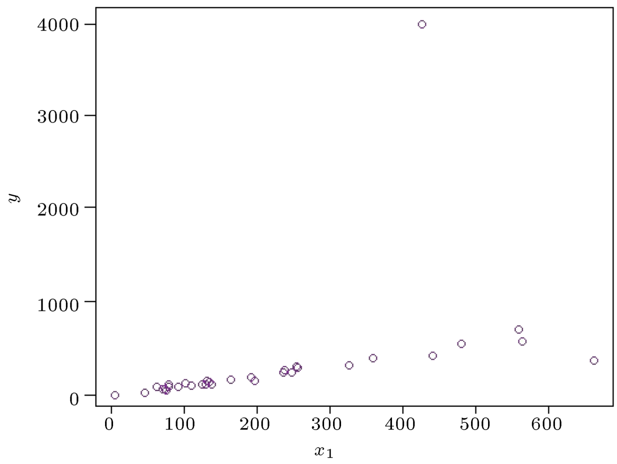


**Figure 3.** Scatter Plot ($x_1$, $y$) of Population-2.

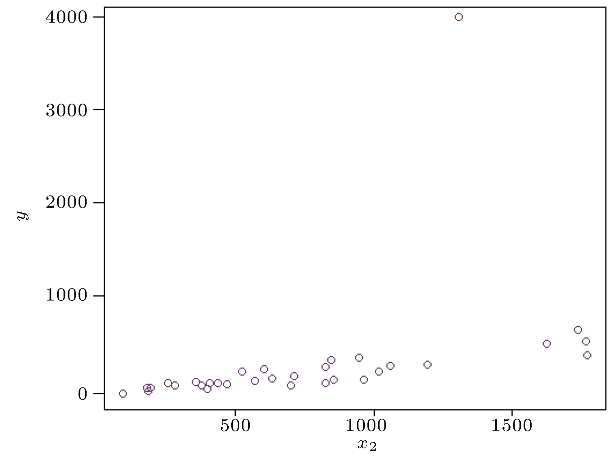


**Figure 4.** Scatter Plot ($x_2$, $y$) of Population-2.

### *4.3. Simulation study (Pop-3 & Pop-4)*

This sub-section is developed to assess the efficiency of the estimators $t_{p1} - t_{p4}$ with respect to the estimators $t_{a(1)} - t_{a(16)}$ based on simulation study. To make comparisons, two multi-variate Normal Distributions (ND) for ($Y$,$X_1$,$X_2$) with means $(\bar{Y}, \bar{X}_1, \bar{X}_2) = (4.9, 4.9, 4.9)$ and covariance matrices are given respectively as follows:

- Population 3:

$$\Sigma = \begin{bmatrix} 9.9 & 2.9 & 2.8 \\ 2.9 & 1.9 & 1.0 \\ 2.8 & 1.0 & 1.9 \end{bmatrix}, \qquad \rho_{yx_1} = 0.651,$$

$\rho_{yx_2} = 0.669$.

- Population 4:

$$\Sigma = \begin{bmatrix} 13.9 & 2.9 & 2.8 \\ 2.9 & 1.9 & 1.0 \\ 2.8 & 1.0 & 1.9 \end{bmatrix}, \qquad \rho_{yx_1} = 0.598,$$

$\rho_{yx_2} = 0.574$.

For the application of robust tools, we add noise in $Y$ [11]. From these populations, $K = 8000$ and SRSWOR with size $n = (250, 300)$ are selected for the $k$th sample and the estimators ($t_{pi}$, $t_{a(i)}$) are evaluated. In this way, for each ($t_{pi}$, $t_{a(i)}$)), the MSE is determined as:

$$MSE\left(\hat{\theta}\right) = \sum_{k=1}^{K} \left(\hat{\theta}^{(k)} - \bar{Y}\right)^2 / K,$$

where $\hat{\theta}^{(k)}$ denotes ($t_{pi}$, $t_{a(i)}$) estimators. The PRE is computed for comparison purposes:

$$PRE\left(\hat{\theta}\right) = \frac{MSE(t_{a(i)})}{MSE(t_{pi})} \times 100.$$

The PRE results of the simulation study are presented numerically in Tables 5 and 6.

**Table 2.** Characteristics of populations.

| | **Population-1** | **Population-2** |
|---|---|---|
| $N$ | 150 | 34 |
| $n$ | 25 | 10 |
| $\bar{Y}$ | 5.843333 | 307.2941 |
| $\bar{X}_1$ | 1.199333 | 218.4118 |
| $\bar{X}_2$ | 3.758 | 765.3529 |
| $C_y$ | 0.1417113 | 2.176777 |
| $C_{x1}$ | 0.6355511 | 0.7678148 |
| $C_{x2}$ | 0.4697441 | 0.6169129 |
| $\rho_{yx1}$ | 0.8179411 | 0.4143947 |
| $\rho_{yx2}$ | 0.8717538 | 0.3906281 |
| $\rho_{x1x2}$ | 0.9628654 | 0.8307546 |
| $\beta_2(x_1)$ | 1.340604 | 0.5274551 |
| $\beta_2(x_2)$ | 1.402103 | 0.1002026 |
| $MR_{x1}$ | 1.3 | 334 |
| $MR_{x2}$ | 3.95 | 933 |
| $TM_{x1}$ | 1.175 | 162.25 |
| $TM_{x2}$ | 3.85 | 705.25 |
| $HL_{x1}$ | 1.2 | 190 |
| $HL_{x2}$ | 3.65 | 718.5 |
| $DM_{x1}$ | 1.183333 | 206.4222 |
| $DM_{x2}$ | 3.738889 | 749.3333 |
| $B_{y.x1(0.10)}$ | 0.8235294 | 0.4727891 |
| $B_{y.x2(0.10)}$ | 0.3684211 | 0.1639929 |
| $B_{y.x1(0.15)}$ | 0.7727273 | 0.7121212 |
| $B_{y.x2(0.15)}$ | 0.375 | 0.196468 |
| $B_{y.x1(0.25)}$ | 0.8 | 0.8972332 |
| $B_{y.x2(0.25)}$ | 0.4 | 0.2234848 |
| $B_{y.x1(0.35)}$ | 0.8333333 | 0.9202733 |
| $B_{y.x2(tky)}$ | 0.3913043 | 0.2570379 |

**Table 3.** PRE of $t_{pi}$ w.r.t $t_{a(i)}$ in Pop-1.

| Estimators | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ |
|---|---|---|---|---|
| $t_{a(1)}$ | 219.5014 | 222.7373 | 222.2552 | 223.1740 |
| $t_{a(2)}$ | 211.8727 | 214.9961 | 214.5307 | 215.4176 |
| $t_{a(3)}$ | 244.2169 | 247.8172 | 247.2808 | 248.3030 |
| $t_{a(4)}$ | 227.3940 | 230.7463 | 230.2468 | 231.1987 |
| $t_{a(5)}$ | 20033.98 | 20329.33 | 20285.32 | 20369.18 |
| $t_{a(6)}$ | 32312.25 | 32788.60 | 32717.62 | 32852.88 |
| $t_{a(7)}$ | 15065.29 | 15287.39 | 15254.29 | 15317.35 |
| $t_{a(8)}$ | 21391.23 | 21706.59 | 21659.60 | 21749.14 |
| $t_{a(9)}$ | 100.7776 | 100.2338 | 100.0168 | 100.4303 |
| $t_{a(10)}$ | 100.90578 | 100.36388 | 100.14662 | 100.56062 |
| $t_{a(11)}$ | 100.2512 | 101.7292 | 101.5089 | 101.9286 |
| $t_{a(12)}$ | 100.50171 | 100.96859 | 100.75002 | 101.16652 |
| $t_{a(13)}$ | 194.1464 | 197.0086 | 196.5821 | 197.3948 |
| $t_{a(14)}$ | 187.2427 | 190.0031 | 189.5918 | 190.3755 |
| $t_{a(15)}$ | 215.9996 | 219.1839 | 218.7094 | 219.6136 |
| $t_{a(16)}$ | 201.3573 | 204.3257 | 203.8834 | 204.7263 |

**Table 4.** PRE of $t_{pi}$ w.r.t $t_{a(i)}$ in Pop-2.

| Estimators | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ |
|---|---|---|---|---|
| $t_{a(1)}$ | 128.7190 | 133.0835 | 135.3778 | 135.1200 |
| $t_{a(2)}$ | 123.4841 | 127.6711 | 129.8722 | 129.6248 |
| $t_{a(3)}$ | 124.7226 | 128.9516 | 131.1747 | 130.9249 |
| $t_{a(4)}$ | 125.0594 | 129.2998 | 131.5289 | 131.2784 |
| $t_{a(5)}$ | 117.4394 | 121.4214 | 123.5147 | 123.2795 |
| $t_{a(6)}$ | 117.5975 | 121.5849 | 123.6810 | 123.4454 |
| $t_{a(7)}$ | 117.5464 | 121.5321 | 123.6273 | 123.3918 |
| $t_{a(8)}$ | 117.5331 | 121.5182 | 123.6132 | 123.3778 |
| $t_{a(9)}$ | 129.0606 | 133.4367 | 135.7371 | 135.4786 |
| $t_{a(10)}$ | 123.3203 | 127.5017 | 129.6999 | 129.4529 |
| $t_{a(11)}$ | 124.4892 | 128.7103 | 130.9293 | 130.6799 |
| $t_{a(12)}$ | 125.0015 | 129.2399 | 131.4680 | 131.2176 |
| $t_{a(13)}$ | 133.9057 | 138.4460 | 140.8329 | 140.5646 |
| $t_{a(14)}$ | 126.0089 | 130.2815 | 132.5276 | 132.2752 |
| $t_{a(15)}$ | 127.9244 | 132.2619 | 134.5421 | 134.2859 |
| $t_{a(16)}$ | 128.6995 | 133.0633 | 135.3574 | 135.0996 |

### *4.4. Discussion*

As an overview, from the PRE results of the proposed estimators with respect to those of Abid et al. [1] presented in Tables 3–6, we note that the relative efficiency values of all estimators surpass 100, which clearly indicates that the proposed estimators are performing better than the one attributed to Abid et al. [1]. Furthermore, particularly, we observe that:

- From Table 3, containing results of Pop-1, the proposed estimators $t_{p_i}$; $i = 1, \cdots, 4$ record high efficiency and $t_{p_4}$, with all $t_{a(i)}$ except $t_{a(9)}$ and $t_{a(10)}$, appears to be the best. Also, from the values of PRE with respect to $t_{a(i)}$, it can be seen that:

$$PRE(t_{p_i}) = \begin{cases} PRE(t_{p_1}) > PRE(t_{p_4}) > PRE(t_{p_2}) \\ > PRE(t_{p_3}); \quad \text{w.r.t} \ \ t_{a(9)}, t_{a(10)} \\ \\ PRE(t_{p_4}) > PRE(t_{p_2}) > PRE(t_{p_3}) \\ > PRE(t_{p_1}); \qquad \text{otherwise} \end{cases}$$

  Regarding the existing estimators, the four highest values of efficiency were always associated with $t_{a(6)}$, $t_{a(8)}$, $t_{a(5)}$, and $t_{a(7)}$ and $\psi_{1(i)}$ and $\psi_{2(i)}$ are considered as the coefficients of kurtosis (recall Table 1).

- From Table 4, results of Pop-2, the proposed estimators $t_{p_i}$; $i = 1, \cdots, 4$ record high efficiency and $t_{p_3}$ appears to be the best. Also, from the values of PRE with respect to $t_{a(i)}$, it is seen that:

$$PRE(t_{p_3}) > PRE(t_{p_4}) > PRE(t_{p_2}) > PRE(t_{p_1}).$$

  Regarding the existing estimators, the four highest values of efficiency were always associated with $t_{a(13)}$, $t_{a(9)}$, $t_{a(1)}$, and $t_{a(16)}$, respectively. In addition, with different coefficients of $\psi_{1(i)}$ and $\psi_{2(i)}$, i.e., comparing each estimator with its counterpart, such that comparing $t_{a(1)}$ with $t_{a(5)}$, $t_{a(9)}$, and $t_{a(13)}$; comparing $t_{a(2)}$ with $t_{a(6)}$, $t_{a(10)}$, and $t_{a(14)}$; comparing $t_{a(3)}$ with $t_{a(7)}$, $t_{a(11)}$, and $t_{a(15)}$; and comparing $t_{a(4)}$ with $t_{a(8)}$, $t_{a(12)}$, and $t_{a(16)}$, we find that the results of the highest efficiency are associated with $t_{a(13)}$, $t_{a(14)}$, $t_{a(15)}$, and $t_{a(16)}$ corresponding to $\psi_{1(i)}$ and $\psi_{2(i)}$ as correlation coefficients.

- From Table 5, containing results of Pop-3, the proposed estimators $t_{p_i}$; $i = 1, \cdots, 4$ record high efficiency and $t_{p_4}$ and $t_{p_2}$ appear to be the best with $n = 250$ and $n = 300$, respectively. From the values of PRE with respect to all others, it is noted that:

$$PRE(t_{p_i}) = \begin{cases} PRE(t_{p_4}) > PRE(t_{p_3}) > PRE(t_{p_1}) \\ > PRE(t_{p_2}); \quad \text{with} \ \ n = 250 \\ \\ PRE(t_{p_2}) > PRE(t_{p_1}) > PRE(t_{p_4}) \\ > PRE(t_{p_3}); \quad \text{with} \ \ n = 300 \end{cases}$$

  Regarding the existing estimators, the four highest values of efficiency were always associated with $t_{a(9)}$, $t_{a(12)}$, $t_{a(11)}$, and $t_{a(10)}$. In addition, with different

**Table 5.** PRE of $t_{pi}$ w.r.t $t_{a(i)}$ in Pop-3.

| Estimators | $n = 250$ | | | | $n = 300$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ |
| $t_{a(1)}$ | 158.0752 | 157.0214 | 161.0813 | 162.0652 | 159.8053 | 159.9673 | 157.4025 | 158.0704 |
| $t_{a(2)}$ | 157.4370 | 156.3875 | 160.4310 | 161.4109 | 158.5655 | 158.7263 | 156.1813 | 156.8441 |
| $t_{a(3)}$ | 157.4808 | 156.4310 | 160.4757 | 161.4558 | 158.6164 | 158.7772 | 156.2314 | 156.8944 |
| $t_{a(4)}$ | 157.5117 | 156.4617 | 160.5071 | 161.4875 | 158.6498 | 158.8106 | 156.2643 | 156.9274 |
| $t_{a(5)}$ | 117.2443 | 116.4627 | 119.4740 | 120.2037 | 104.3757 | 104.4815 | 102.8063 | 103.2425 |
| $t_{a(6)}$ | 117.6107 | 116.8267 | 119.8473 | 120.5793 | 104.5747 | 104.6807 | 103.0023 | 103.4394 |
| $t_{a(7)}$ | 117.6351 | 116.8510 | 119.8723 | 120.6044 | 104.6463 | 104.7523 | 103.0728 | 103.5102 |
| $t_{a(8)}$ | 117.6407 | 116.8565 | 119.8780 | 120.6102 | 104.6893 | 104.7954 | 103.1152 | 103.5528 |
| $t_{a(9)}$ | 184.3911 | 183.1619 | 187.8977 | 189.0454 | 186.9523 | 187.1418 | 184.1412 | 184.9226 |
| $t_{a(10)}$ | 184.2794 | 183.0509 | 187.7839 | 188.9309 | 186.1350 | 186.3237 | 183.3363 | 184.1143 |
| $t_{a(11)}$ | 184.3388 | 183.1099 | 187.8444 | 188.9917 | 186.2145 | 186.4033 | 183.4146 | 184.1929 |
| $t_{a(12)}$ | 184.3697 | 183.1406 | 187.8759 | 189.0234 | 186.2636 | 186.4524 | 183.4629 | 184.2414 |
| $t_{a(13)}$ | 167.7246 | 166.6065 | 170.9143 | 171.9582 | 169.7884 | 169.9605 | 167.2354 | 167.9451 |
| $t_{a(14)}$ | 167.1990 | 166.0844 | 170.3786 | 171.4193 | 168.6136 | 168.7845 | 166.0783 | 166.7830 |
| $t_{a(15)}$ | 167.2511 | 166.1362 | 170.4318 | 171.4728 | 168.6823 | 168.8533 | 166.1459 | 166.8510 |
| $t_{a(16)}$ | 167.2845 | 166.1694 | 170.4658 | 171.5070 | 168.7262 | 168.8972 | 166.1892 | 166.8944 |

**Table 6.** PRE of $t_{pi}$ w.r.t $t_{a(i)}$ in Pop-4.

| Estimators | $n = 250$ | | | | $n = 300$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ | $t_{p1}$ | $t_{p2}$ | $t_{p3}$ | $t_{p4}$ |
| $t_{a(1)}$ | 121.1568 | 123.7984 | 124.0890 | 124.6420 | 148.5768 | 148.5108 | 148.1262 | 148.1145 |
| $t_{a(2)}$ | 120.8914 | 123.5272 | 123.8171 | 124.3689 | 149.1373 | 149.0710 | 148.6850 | 148.6732 |
| $t_{a(3)}$ | 120.8901 | 123.5259 | 123.8159 | 124.3676 | 149.1388 | 149.0726 | 148.6865 | 148.6748 |
| $t_{a(4)}$ | 120.8885 | 123.5243 | 123.8142 | 124.3660 | 149.1569 | 149.0906 | 148.7045 | 148.6928 |
| $t_{a(5)}$ | 101.8737 | 104.0949 | 104.3392 | 104.8042 | 120.9501 | 120.8964 | 120.5833 | 120.5738 |
| $t_{a(6)}$ | 101.9040 | 104.1258 | 104.3703 | 104.8354 | 120.4182 | 120.3647 | 120.0530 | 120.0435 |
| $t_{a(7)}$ | 101.8967 | 104.1184 | 104.3628 | 104.8279 | 120.3943 | 120.3409 | 120.0292 | 120.0197 |
| $t_{a(8)}$ | 101.8930 | 104.1146 | 104.3590 | 104.8240 | 120.3911 | 120.3376 | 120.0259 | 120.0165 |
| $t_{a(9)}$ | 135.0795 | 138.0247 | 138.3487 | 138.9652 | 165.5874 | 165.5139 | 165.0852 | 165.0722 |
| $t_{a(10)}$ | 134.8794 | 137.8203 | 138.1438 | 138.7594 | 166.0360 | 165.9622 | 165.5324 | 165.5193 |
| $t_{a(11)}$ | 134.8729 | 137.8135 | 138.1370 | 138.7526 | 166.0305 | 165.9568 | 165.5270 | 165.5139 |
| $t_{a(12)}$ | 134.8686 | 137.8091 | 138.1326 | 138.7482 | 166.0447 | 165.9709 | 165.5411 | 165.5280 |
| $t_{a(13)}$ | 127.8766 | 130.6648 | 130.9715 | 131.5551 | 156.9219 | 156.8522 | 156.4460 | 156.4336 |
| $t_{a(14)}$ | 127.6339 | 130.4168 | 130.7229 | 131.3054 | 157.4141 | 157.3442 | 156.9367 | 156.9243 |
| $t_{a(15)}$ | 127.6266 | 130.4092 | 130.7154 | 131.2978 | 157.4098 | 157.3399 | 156.9324 | 156.9200 |
| $t_{a(16)}$ | 127.6217 | 130.4043 | 130.7104 | 131.2929 | 157.4274 | 157.3575 | 156.9499 | 156.9375 |

coefficients of $\psi_{1(i)}$ and $\psi_{2(i)}$, we find that the results of the highest efficiency are associated with $t_{a(9)}$, $t_{a(12)}$, $t_{a(11)}$, and $t_{a(10)}$ corresponding to $\psi_{1(i)}$ and $\psi_{2(i)}$ as coefficients of variation.

- From Table 6, containing results of Pop-4, the proposed estimators $t_{p_i}$; $i = 1, \ldots, 4$ record high efficiency and $t_{p_4}$ and $t_{p_1}$ appear to be the best with $n = 250$ and $n = 300$, respectively. From the values of PRE with respect to all others, we note that:

$$PRE(t_{p_i}) = \begin{cases} PRE(t_{p_4}) > PRE(t_{p_3}) > PRE(t_{p_2}) \\ > PRE(t_{p_1}); \quad \text{with} \quad n = 250 \\ \\ PRE(t_{p_1}) > PRE(t_{p_2}) > PRE(t_{p_3}) \\ > PRE(t_{p_4}); \quad \text{with} \quad n = 300 \end{cases}$$

Regarding the existing estimators, the four highest values of efficiency were always associated with $t_{a(9)}$, $t_{a(10)}$, $t_{a(11)}$, and $t_{a(12)}$ when $n = 250$ and with $t_{a(12)}$, $t_{a(10)}$, $t_{a(11)}$, and $t_{a(9)}$ when $n = 300$. In addition, with different coefficients of $\psi_{1(i)}$ and $\psi_{2(i)}$, we find that the results of the highest efficiency for $n = 250$ and $n = 300$ are associated with $t_{a(9)}$, $t_{a(10)}$, $t_{a(11)}$, $t_{a(12)}$ and $t_{a(12)}$, $t_{a(10)}$, $t_{a(11)}$, $t_{a(9)}$, respectively, corresponding to $\psi_{1(i)}$ and $\psi_{2(i)}$ as coefficients of variation.

Overall, the results of the numerical illustration support utilizing the proposed quantile regression-type estimators for mean estimation based on the information of two auxiliary variables.

## 5. Conclusion

One of the main concerns of survey sampling is to enhance the mean estimation based on auxiliary information. Abid et al. [1] found that the utilization of non-conventional location measures of mean estimation, especially for a data set with outliers, was much better than the traditional and existing location estimators. In this paper, as an extension to the work of Abid et al. [1], a new family of estimators based on the information of two auxiliary variables was proposed to estimate the population mean through the use of quantile regression. Two real-life data sets besides simulation study were considered in the numerical illustration. The first real-life data set called Iris data set was free from outliers, as credited by Fisher [23], and the second real-life data set reflected the existence of outliers, credited by Sukhatme [25]. Simulation study was also performed for the purpose of evaluation. With all the cases considered in numerical illustration, the percentage relative efficiency clearly indicates the higher efficiency of the proposed estimators than the existing estimators attributed to Abid et al. [1]. Moreover, an additional feature of the proposed estimators is robustness to outliers. Last but not least, we opt for the use of the proposed estimators in the presence of outliers over the existing estimators attributed to Abid et al. [1].


## Acknowledgments

The authors are grateful to the Dean of Scientific Research at King Khalid University, Kingdom of Saudi Arabia for funding this study through the research groups program under project number R.G.P.1/64/42. Ibrahim Mufrah Almanjahie and Ishfaq Ahmad received the grant.


## References


1. Abid, M., Nazir, H.Z., Riaz, M., et al. "Improved ratio estimators using some robust measures", *Hacettepe Journal of Mathematics and Statistics*, **47**(5), pp. 1375–1393 (2018).

2. Sarndal, C.-E. and Lundstrom, S., *Estimation in Surveys with Nonresponse*, John Wiley and Sons Ltd (2005).

3. Al-Omari, A.I. "Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling", *Statistics and Probability Letters*, **82**(11), pp. 1883–1890 (2012).

4. Abid, M., Abbas, N., and Riaz, M. "Improved modified ratio estimators of population mean based on deciles", *Chiang Mai Journal of Science*, **43**(1), pp. 1311–1323 (2016).

5. Abid, M., Abbas, N., Zafar Nazir, H., et al. "Enhancing the mean ratio estimators for estimating population mean using non-conventional location parameters", *Revista Colombiana de Estadistica*, **39**(1), pp. 63–79 (2016).

6. Shahzad, U., Perri, P.F., and Hanif, M. "A new class of ratio-type estimators for improving mean estimation of nonsensitive and sensitive variables by using supplementary information", *Communications in Statistics-Simulation and Computation*, **48**(9), pp. 2566–2585 (2018).

7. Bulut, H. and Zaman, T. "An improved class of robust ratio estimators by using the minimum covariance determinant estimation", *Communications in Statistics-Simulation and Computation*, **51**(5), pp. 2457–2463 (2022).

8. Zaman, T. and Bulut, H. "Modified ratio estimators using robust regression methods", *Communications in Statistics-Theory and Methods*, **48**(8), pp. 2039–2048 (2019).

9. Zaman, T. and Bulut, H. "Modified regression estimators using robust regression methods and covariance matrices in stratied random sampling", *Communications in Statistics - Theory and Methods*, **49**(14), pp. 3407–3420 (2020).

10. Shahzad, U. and Hanif, M. "Some imputation based new estimators of population mean under non-response", *Journal of Statistics and Management Systems*, **22**(8), pp. 1381–1399 (2019).

11. Ali, N., Ahmad, I., Hanif, M., et al. "Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information", *Communications in Statistics-Theory and Methods*, **50**(4), pp. 979–992 (2021).

12. Hanif, M. and Shahzad, U. "Estimation of population variance using kernel matrix", *Journal of Statistics and Management Systems*, **22**(3), pp. 563–586 (2019).

13. Zaman, T. "Improvement of modified ratio estimators using robust regression methods", *Applied Mathematics and Computation*, **348**, pp. 627–631 (2019).

14. Shahzad, U., Al-Noor, N.H., Hanif, M., et al. "Imputation based mean estimators in case of missing data utilizing robust regression and variance-covariance matrices", *Communications in Statistics- Simulation and Computation*, **51**(8), pp. 4276–4295 (2022).

15. Shahzad, U., Hanif, M., Sajjad, I., et al. "Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information", *Scientia Iranica*, **29**(3), pp. 1705–1715 (2022).

16. Shahzad, U., Ahmad, I., Almanjahie, I., et al. "A new class of L-moments based calibration variance Estimators", *Computers Materials and Continua*, **66**(3), pp. 3013–3028 (2021).

17. Moore, D.S. and Kirkland, S., *The Basic Practice of Statistics*, **2**, New York, WH Freeman (2007).

18. Pedhazur, E.J., *Multiple Regression in Behavioral Research: Explanation and Prediction*, Thompson Learning Inc: New York (1997).

19. Koenker, R., *Quantile Regression*, Cambridge University Press, New York (2005).

20. Koenker, R. and Bassett Jr, G. "Regression quantiles", *Econometrica: Journal of the Econometric Society*, pp. 33–50 (1978).

21. Hao, L., Naiman, D.Q., and Naiman, D.Q., *Quantile Regression*, Sage (2007).

22. Koenker, R. and Hallock, K.F. "Quantile regression", *Journal of Economic Perspectives*, **15**(4), pp. 143–156 (2001).

23. Fisher, R.A. "The use of multiple measurements in taxonomic problems", *Annals of Augenics*, **7**(2), pp. 179–188 (1936).

24. Anderson, E. "The irises of the Gasp Peninsula", *Bulletin of the American Iris Society*, **59**, pp. 2–5 (1935).

25. Sukhatme, B.V. "Some ratio-type estimators in two-phase sampling", *Journal of the American Statistical Association*, **57**(299), pp. 628–632 (1962).


## Biographies

**Usman Shahzad** obtained his MSc degree in statistics from International Islamic University, Islamabad, Pakistan. He completed his MPhil degree in statistics from Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. Currently, he is studying for his PhD in statistics from International Islamic University, Islamabad, Pakistan. He served as a lecturer at Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. He has published more than 40 research papers in research journals. His research interests include survey sampling, extreme value theory, stochastic process, probability and non-parametric statistics.

**Ishfaq Ahmad** was born in Pakistan on March 26th, 1981. He received his BSc degree in Statistics, Physics, and Mathematics from Bahudin Zakaria University, Multan Pakistan in 1999. Afterward, he received MSc and MPhil degrees in Statistics from Quaid-e-Azam University, Islamabad, Pakistan in 2003 and 2005, respectively. Finally, he received a PhD degree in Probability and Mathematical Statistics from the Institute of Applied Mathematics, University of Chinese Academy of Science (UCAS) Beijing, China in 2010. Currently, he is working as an Associate Professor of Statistics at the Department of Mathematics and Statistics, Faculty of Basic and Applied Sciences, International Islamic University, Islamabad, Pakistan. Before this, he was working as an Assistant Professor at King Khalid University, Saudi Arabia. His main areas of research interest are extreme value theory, statistical inference, survey sampling, optimization, Bayesian analysis, and functional analysis. He has published more than 47 articles in journals of international repute, such as the International Journal of Climatology, Scientific reports, and many others.

**Ibrahim Mufrah Almanjahie** was born in Saudi Arabia in 1979. He received his BSc degree in Mathematics from King Khalid University, Saudi Arabia in 2002. He received the MSc degree in Mathematical and Statistical Sciences and PhD degree in Probability and Statistical Modeling from University of Western Australia, Australia in 2008 and 2015, respectively. Currently, he is an Associate Professor at the Department of Mathematics, College of Sciences, King Khalid University. He is also the President of the Saudi Association for Statistical Sciences. His main areas of research are modeling and analysis of ion channel data, hidden Markov models, M Algorithm, finite mixture models MCMC, computational methods in statistics, time series modeling, applied statistics, sampling, and functional statistics.

**Nadia H. Al-Noor** completed her PhD degree in

Mathematical Statistics. Currently, she is serving as a Professor of Mathematical Statistics at the Department of Mathematics, College of Science, Mustansiriyah University, Baghdad, Iraq. She has published more than 40 research papers in research journals. She supervised more than 10 postgraduate students' dissertations and theses. Her main research interests include probability and mathematical statistics, survey sampling, statistical inference, reliability theory, robust regression, and non-parametric estimation methods.

**Muhammad Hanif** obtained his MSc degree in Statistics from the University of Agriculture, Faisalabad, Pakistan. He completed his MPhil in statistics from Government College University, Lahore, Pakistan. He completed his PhD degree in Statistics from Zhejiang University, Hangzhou, China. Currently, he is an Associate Professor and serving as the Chairman of the Department of Mathematics and Statistics at Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. He has published more than 50 research papers in research journals. His main research interests include stochastic process, probability and mathematical statistics, survey sampling, and non-parametric estimation.