Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information

Usman Shahzad, Muhammad Hanif, Irsa Sajjad, Malik Muhammad Anas

Department of Mathematics and Statistics, International Islamic University, Islamabad, 46000, Pakistan.
Department of Mathematics and Statistics - PMAS-Arid Agriculture University, Rawalpindi, 46300, Pakistan.
Department of Lahore Business School - University of Lahore, Islamabad, 46000, Pakistan.
School of Science, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, P.R. China.

Email Corresponding Author: usman.stat@yahoo.com (Shahzad, U.), Cell: 00923315995587
Email: mhpuno@hotmail.com (Hanif, M.)
Email: irsasajjad@yahoo.com (Sajjad, I.)
Email: anas.uaar@gmail.com (Anas, M.)

Abstract

Traditional ordinary least square (OLS) regression is commonly utilized to develop regression-ratio-type estimators with traditional measures of location. Abid et al. [1] extended this idea and developed regression-ratio-type estimators with traditional and non-traditional measures of location. In this article, the quantile regression with traditional and non-traditional measures of location is utilized and a class of ratio type mean estimators are proposed. The theoretical mean square error (MSE) expressions are also derived. The work is also extended for two phase sampling (partial information). The pertinence of the proposed and existing group of estimators is shown by considering real data collections originating from different sources. The discoveries are empowering and prevalent execution of the proposed group of estimators is witnessed and documented throughout the article.

Keywords: Quantile regression; OLS regression; Ratio-type estimators; Simple random sampling;
Two stage sampling.

1 Introduction

It is a commonly known fact that this is the era of information. This fact not only highlights the volume and pace of information but also underlines the necessity of accurate flow of it. This is all linked with the true urge of collecting data. It also enables ourselves to profile the surroundings precisely which in turn helps in making an optimal decision. Till date, the sampling theory and methods are still considered as a core of literature of multidisciplinary research. Utility of surveys and sampling methods in applied research can be clearly seen in the surveys for instance, Eurostat on European Union Labor Force giving periodical statistics of labor participation. Another example depicting the importance of sampling method is market profiling-segmentation survey providing significant evidence to encourage market share investigation and National Co-morbidity survey evaluating the substance disorder and anxiety levels.

The most philosophical goal of sampling methods is the valuation of the occurrence of patterns or attributes present in inhabitants or population of research. Auxiliary information is widely used to bring more precision in sample estimates. Laplace in 1814, gave the justification and significance of practicing auxiliary information in easy but succinct manner. For instance, he revealed that the births register providing the conditions of the residents, can help in estimating the population of a large country without holding a census of its residents. But this is only possible if the ratio of masses to annual birth is known. In addition, McClellan et al. [2] practiced distinction of separations of heart assault injured individual’s home to closest cardiovascular catheterization medical clinic and the separation to the closest emergency clinic of any kind, as auxiliary information to help the estimation of impact of catheterization on patients health.

An alternative method for situations in which abundance of auxiliary information is available is ranked set sampling (RSS) method due to McIntyre [3], which is shown is far more cost-efficient than simple random sampling method. See Adel Rastkhiz et al. [4], Zamanzade and Vock [5], Zamanzade and Wang [6], Zamanzade and Mahdizadeh [7], Zamanzade and Wang [8] and Mahdizadeh and Zamanzade [9] for more information about this.

A vast amount of literature is available on ratio-type estimators for mean estimation. Such as Koyuncu [10], Shahzad el al. [11], Hanif and Shahzad [12] have developed some classes of estimators utilizing
supplementary information under simple random sampling scheme. However, for positive correlation, traditional regression-ratio-type estimator is better for the estimation of population parameters (see, e.g., Abid et al. [13]; Irfan et al. [14]; Naz et al. [15], Abid et al. [16]). Note that traditional regression-type-ratio estimators based on conventional regression coefficient i.e. known as OLS regression coefficient. However, the OLS estimate become inappropriate when data are contaminated by outliers. For solving this issue, there are some modifications available in literature (see, e.g., Zaman et al. [17]; Zaman and Bulut [18]; Zaman and Bulut [19]; Zaman [20]; Ali et al. [21]). Abid et al. [1] first time introduced the idea of utilizing non-traditional measures of location, and enhance the estimates of population mean. They found enhanced estimates of population mean utilizing non-traditional measures. The current study is based on an extension of Abid et al. [1] work, by incorporating quantile regression.

Outliers have a negative effect on accuracy of the estimators obtained by OLS technique. The presence of outliers disregard OLS model assumption (Hao et al., [22]). The outliers can change the regression parameters to be smaller or bigger than the parameters evaluated when the outliers were excluded in the data (Moore, [23]; Pedhazur, [24]). It was recommended that the outliers can be barred from the examination if they were, from exhaustive examinations, demonstrated to be non-legitimate perceptions. Yet, when the outliers are substantial, they can give new bits of knowledge about the nature of the data (Pedhazur, [24]). It implies that a measurable procedure is required that will capture the outliers in the investigation but is less impacted by their quality. An elective procedure that has greater ability to settle a few issues referenced before is called quantile regression. It was created by Boscovich in the eighteenth century, even before the possibility of least squares regression estimators developed (Koenker, [25]; Koenker and Bassett, [26]). In olden times, utilisations of quantile regression were restricted to financial matters or natural investigations, however now it is used in almost all the fields of social sciences as a data analysis device. In light of preceding lines, we are introducing quantile regression coefficient rather than OLS, in ratio type mean estimators, by extending the idea of Abid et al. [1].

Inspired by the above documented developments, in this article, we propose a new family of estimators for estimating population mean using more scrupulous use of auxiliary variable. The objective is met by utilizing the quantile regression, traditional and non-traditional measures of location of auxiliary variable. The applicability of the scheme is further demonstrated in simple and two stage random sampling frameworks by employing on three diverse data sets coming from various fields of inquiries. Moreover, keenly persuaded comparative investigation between Abid et al. [1] and proposed family, by means of numerical evaluations. The rest of the article is arranged in seven major parts. In section 2, we present
review of estimators constructed in the study conducted by Abid et al. [1]. The section 3 proposes a class of quantile regression-ratio-type estimators. Whereas, in section 4, Two stage sampling scheme version of reviewed and proposed estimators are provided. Numerical illustration and general findings are documented in section 5. Some final remarks are also given in section 6.

2  A review of Abid et al. estimators

Abid et al. [1] incorporated non-traditional measures of location with traditional measures of location for mean estimation. They used mid-range of $X$, denoted by $MR$, as a first non-traditional measure of location. Hodges-Lehmann estimator, denoted by $HL$, as a second non-traditional measure of location. They also used the weighted average of the population median and two extreme quartiles, known as Tri-mean, denoted by $TM$, as a third non-traditional measure of location. According to Abid et al. [1], these measures are highly sensitive in absence of normality and in presence of outliers. They introduced the following class of OLS regression-ratio-type estimators utilizing traditional and non-traditional measures of location as given by

$$\bar{y}_{ab_i} = \bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})(A_1 \bar{X} + A_2) \quad \text{for} \quad i = 1, 2, ..., 9$$

where $(\bar{X}, \bar{Y})$ be the population means and $(\bar{x}, \bar{y})$ with be the sample means when a simple random sample of size $n$ is drawn from the population. Further, $A_1$ and $A_2$ are either $(0,1)$ or some known population measures namely, $HL$, the Hodges Lemon, $MR$, the Mid Range, $TM$, the Tri Mean, $C_x$, the coefficient of variation, $\rho$, the coefficient of correlation, and $\hat{\beta}_{(ols)}$ is the OLS regression coefficient. The family members of $\bar{y}_{ab_i}$ are provided in Table 1. The MSE of $\bar{y}_{ab_i}$ is given below

$$MSE(\bar{y}_{ab_i}) = \theta[S_y^2 + u_i^2S_x^2 + 2\beta_{(ols)}u_iS_x\hat{\beta}_{(ols)}S_x - 2u_iS_{xy} - 2\beta_{(ols)}S_y] \quad \text{for} \quad i = 1, 2, ..., 9 \quad (2.1)$$

where $u_i = \frac{A_1 \bar{Y}}{A_1 \bar{X} + A_2}$ and $\theta = \left(1 - \frac{f}{n}\right)$ for $i=1,2,...,9$. Further, $S_y^2$ and $S_x^2$ are the unbiased variances, and $S_{xy}$ be the co-variance of $Y$ and $X$.
3 Proposed class of quantile regression-ratio-type estimators

Outliers are the observations in a data set which appear to be inconsistent with the rest of that data set. Presence of outliers significantly effect mean estimation which is one of the most important measure 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 significant impact on the traditional regression coefficient calculated from OLS tool. Hence the estimate of population mean i.e. \( (\bar{Y}) \), based upon OLS may indicate poor performance. One of the solution is to utilize quantile regression. It can be used as a robust approach in such circumstances i.e. whenever data is non-normal and contaminated with outliers. Because it is not sensitive to outliers (Hao et al., [22]; Koenker, [25]). Quantile regression is like customary OLS regression in a sense that both of them explore connections among endogenous and exogenous variables. The primary distinction is that OLS regression picks parameter esteems that have the least squared deviation from the regression line as the parameter estimates, while quantile regression picks parameter esteems that have the least absolute deviation from the regression line as the parameter estimates. So we are proposing a group of quantile regression-ratio-type estimators by extending the idea of abid et al. [1], as given below:

\[
\bar{y}_{pi} = \bar{y} + \hat{\beta}(q)(\bar{X} - \bar{x}) \frac{(A_1 \bar{x} + A_2)}{(A_1 \bar{x} + A_2)} \quad \text{for} \quad i = 1, 2, \ldots, 9,
\]

with

\[
\hat{\beta}(q) = \arg\min_{\beta \in \mathbb{R}^p} \rho_q(v) \sum_{i=1}^{n} (y_i - \langle x_i, \beta \rangle),
\]

where \( \rho_q(v) \) is a continuous piecewise linear function (or asymmetric absolute loss function), for quantile \( q \in (0, 1) \), but nondifferentiable at \( v = 0 \). Note that all the notations of \( \bar{y}_{pi} \) have usual meanings as discuss in previous section. However, \( \hat{\beta}(q) \) is the quantile regression coefficient for \( p = 2 \) variables. For deep study of quantile regression, interested readers may refer to Koenker and Hallock [27]. The family members of \( \bar{y}_{pi} \) are provided in Table 2. The MSE of proposed family of estimators is given below

\[
MSE(\bar{y}_{pi}) = \theta \left[ S_y^2 + u_i S_x^2 + 2\beta(q)u_i S_x^2 + \beta^2(q) S_x^2 - 2u_i S_{xy} - 2\beta(q)S_{xy} \right] \quad \text{for} \quad i = 1, 2, \ldots, 9. \quad (3.1)
\]

It is worth mentioning that we are using \( q^{15th} = 0.15 \), \( q^{25th} = 0.25 \) and \( q^{35th} = 0.35 \) quantiles for the purposes of current article. We see from the consequences of the numerical study conducted in Sec. 5 that utilizing the quantile regression coefficients, based on these referenced quantiles, incredibly enhance the efficiencies of proposed estimators. Note that utilizing these three referenced quantiles, proposed class
contain twenty seven members. For sack of readability, let us provide twenty seven members of proposed
class with their MSE in compact form, as follows

\[
y_{p_i} = \begin{cases} 
\frac{\bar{y} + \hat{\beta}_{(0.15)}(\bar{X} - \bar{x})}{(A_1\bar{x} + A_2)} (A_1\bar{X} + A_2) & \text{for } (q) = (0.15) = 1, 2, \ldots, 9 \\
\frac{\bar{y} + \hat{\beta}_{(0.25)}(\bar{X} - \bar{x})}{(A_1\bar{x} + A_2)} (A_1\bar{X} + A_2) & \text{for } (q) = (0.25) = 1, 2, \ldots, 9 \\
\frac{\bar{y} + \hat{\beta}_{(0.35)}(\bar{X} - \bar{x})}{(A_1\bar{x} + A_2)} (A_1\bar{X} + A_2) & \text{for } (q) = (0.35) = 1, 2, \ldots, 9,
\end{cases}
\]

\[
MSE(y_{p_i}) = \begin{cases} 
\theta[S_y^2 + u_i^2S_x^2 + 2\beta_{(0.15)}u_iS_x^2 + \beta_{(0.15)}^2S_y^2 - 2u_iS_{xy} - 2\beta_{(0.15)}S_{xy}] \\
\theta[S_y^2 + u_i^2S_x^2 + 2\beta_{(0.25)}u_iS_x^2 + \beta_{(0.25)}^2S_y^2 - 2u_iS_{xy} - 2\beta_{(0.25)}S_{xy}] \\
\theta[S_y^2 + u_i^2S_x^2 + 2\beta_{(0.35)}u_iS_x^2 + \beta_{(0.35)}^2S_y^2 - 2u_iS_{xy} - 2\beta_{(0.35)}S_{xy}]
\end{cases}
\]

[Table 2 Here]
4 Two stage sampling scheme (partial information)

One can use the two-stage sampling, exactly when the information on population mean of supplementary variable isn’t available. Neyman [28] was the pioneer who gave the possibility of this sampling scheme in assessing the population parameters. The two-stage sampling is monetarily insightful and more straightforward as well. This sampling plan is used to get the data about supplementary variable productively by choosing a more noteworthy sample from first stage and moderate size sample at the second stage. Sukhatme [29] used two-stage sampling plan to propose a general ratio type estimator. Cochran [30] is another reference for more comprehensive study about two-stage sampling.

Under two-stage sampling plan, we select a first stage sample of size \( n_a \) units from the population of size \( N \) with the assistance of SRSWOR plan. After that we select a second stage sample i.e. \( n_b \), a sub-sample of the first stage sample i.e. \( n_a \).

4.1 Reviewed and proposed estimators in two stage sampling scheme

In this sub-section, we are presenting Abid et al. [1] and proposed family of estimators under two-stage sampling scheme as given by

\[
\bar{y}_{ab} = \frac{\bar{y}_b + \hat{\beta}_{(ols)}(\bar{x}_a - \bar{x}_b)}{(A_1\bar{x}_b + A_2)}(A_1\bar{x}_a + A_2) \quad \text{for} \quad i = 1, 2, ..., 9
\]

\[
\bar{y}_{pi} = \frac{\bar{y}_b + \hat{\beta}_{(q)}(\bar{x}_a - \bar{x}_b)}{(A_1\bar{x}_b + A_2)}(A_1\bar{x}_a + A_2) \quad \text{for} \quad i = 1, 2, ..., 9
\]

where \((\bar{x}_b, \bar{y}_b)\) representing sample means at second stage and \(\bar{x}_a\) be the sample mean at first stage. Further, \(A_1\) and \(A_2\) have the same meanings as described in previous section. The family members of \(\bar{y}_{ab}\) and \(\bar{y}_{pi}\) are available in Table 1 and 2, respectively.

Utilizing Taylor series method we are obtaining MSE for \(\bar{y}_{ab}\) as follows

\[
MSE(\bar{y}_{ab}) = d_1 \sum d_1', \quad (4.1)
\]

where

\[
d_1 = \left[ \begin{array}{c}
\frac{\delta h(\bar{y}_b, \bar{x}_a, \bar{x}_b)}{\delta \bar{y}_b} |\bar{y}, \bar{x} \begin{array}{c}
\frac{\delta h(\bar{y}_b, \bar{x}_a, \bar{x}_b)}{\delta \bar{x}_a} |\bar{y}, \bar{x} \begin{array}{c}
\frac{\delta h(\bar{y}_b, \bar{x}_a, \bar{x}_b)}{\delta \bar{x}_b} |\bar{y}, \bar{x} \end{array}
\end{array}
\end{array}
\right],
\]

\[
d_1 = \left[ 1 \begin{array}{c}
(u_i + \beta_{(ols)}) \end{array} - (u_i + \beta_{(ols)}) \right].
\]
\[
\begin{bmatrix}
V(\bar{y}_b) & Cov(\bar{y}_b, \bar{x}_a) & Cov(\bar{y}_b, \bar{x}_b) \\
Cov(\bar{x}_a, \bar{y}_b) & V(\bar{x}_a) & Cov(\bar{x}_a, \bar{x}_b) \\
Cov(\bar{x}_b, \bar{y}_b) & Cov(\bar{x}_b, \bar{x}_a) & V(\bar{x}_b)
\end{bmatrix}
\]

where

\[
V(\bar{y}_b) = \lambda_2 S_y^2,
V(\bar{x}_a) = \lambda_1 S_x^2,
V(\bar{x}_b) = \lambda_2 S_y^2,
Cov(\bar{y}_b, \bar{x}_a) = Cov(\bar{x}_a, \bar{y}_b) = \lambda_1 S_{yx},
Cov(\bar{y}_b, \bar{x}_b) = Cov(\bar{x}_b, \bar{y}_b) = \lambda_2 S_{yx},
Cov(\bar{x}_a, \bar{x}_b) = Cov(\bar{x}_b, \bar{x}_a) = \lambda_1 S_x^2.
\]

Utilizing these defined notations of variances and co-variances, and hence substituting the values of \(d_1\) and \(\Sigma\) in equation (4.1), MSE expressions of \(\bar{y}_{ab}^{'}\) as follows

\[
MSE(\bar{y}_{ab}^{'}) = \lambda_2 S_y^2 + (\lambda_2 - \lambda_1)[(u_i + \beta_{(ols)})^2 S_x^2 - 2(u_i + \beta_{(ols)}) S_{yx}], \text{ for } i = 1, 2, ..., 9.
\]

By replacing \(\beta_{(ols)}\) with \(\beta_{(q)}\), we can easily find MSE of proposed class of estimators as follows

\[
MSE(\bar{y}_{p_i}^{'}) = \lambda_2 S_y^2 + (\lambda_2 - \lambda_1)[(u_i + \beta_{(q)})^2 S_x^2 - 2(u_i + \beta_{(q)}) S_{yx}],
\]

where \(\lambda_1 = \left(\frac{1}{n_a} - \frac{1}{N}\right)\) and \(\lambda_2 = \left(\frac{1}{n_b} - \frac{1}{N}\right)\).

The twenty seven family members of proposed class with their MSE in compact form, under partial information, as follows

\[
\bar{y}_{p_i}^{'}, \quad \begin{cases}
\frac{\bar{y}_b + \hat{\beta}_{(0.15)} (\bar{x}_a - \bar{x}_b)}{(A_1 \bar{x}_a + A_2)} (A_1 \bar{x}_a + A_2) & \text{for } (q) = (0.15); i = 1, 2, ..., 9 \\
\frac{\bar{y}_b + \hat{\beta}_{(0.25)} (\bar{x}_a - \bar{x}_b)}{(A_1 \bar{x}_a + A_2)} (A_1 \bar{x}_a + A_2) & \text{for } (q) = (0.25); i = 1, 2, ..., 9 \\
\frac{\bar{y}_b + \hat{\beta}_{(0.35)} (\bar{x}_a - \bar{x}_b)}{(A_1 \bar{x}_a + A_2)} (A_1 \bar{x}_a + A_2) & \text{for } (q) = (0.35); i = 1, 2, ..., 9.
\end{cases}
\]

(4.2)
\[ \text{MSE}(\hat{y}_{pi}) = \begin{cases} 
\lambda_2 S_y^2 + (\lambda_2 - \lambda_1)[(u_i + \beta_{(0.15)})^2 S_x^2 - 2(u_i + \beta_{(0.15)})S_{yx}] \\
\lambda_2 S_y^2 + (\lambda_2 - \lambda_1)[(u_i + \beta_{(0.25)})^2 S_x^2 - 2(u_i + \beta_{(0.25)})S_{yx}] \\
\lambda_2 S_y^2 + (\lambda_2 - \lambda_1)[(u_i + \beta_{(0.35)})^2 S_x^2 - 2(u_i + \beta_{(0.35)})S_{yx}] 
\end{cases} \] (4.3)

5 Numerical illustration

In this section, we are assessing the performance of proposed and existing estimators, using three real life data sets as given below:

**Population-1 (Pop-1):** We use the data set of Singh [31]. In this data set, “amount of non-real estate farm loans during 1977” is taken as auxiliary variable (X), while “amount of real estate farm loans during 1977” is taken as study variable (Y). Further \( N = 50 \), \( n = n_a = 20 \) and \( n_b = 16 \).

**Population-2 (Pop-2):** We use “UScereals” data set. Which describes 65 commonly available breakfast cereals in the USA, based on the information available on the mandatory food label on the packet. The measurements are normalized here to a serving size of one American cup. The data come from ASA Statistical Graphics Exposition and used by Venables and Ripley [32]. As the data set contains a number of variables. So, “grams of fibre in one portion” is taken as auxiliary variable (X), while “grams of potassium” is taken as study variable (Y). Further \( N = 65 \), \( n = n_a = 20 \) and \( n_b = 15 \).

**Population-3 (Pop-3):** Third population is also considered from “UScereals” data set. Where, “grams of sodium in one portion” is taken as auxiliary variable (X), while “Number of calories” is taken as study variable (Y). Further \( N = 65 \), \( n = n_a = 20 \) and \( n_b = 15 \).

The remaining characteristics of all the three populations are provided in Table 3. Further, we draw Histogram, Box-plot and Scatter-plot for all the three referenced populations, see figures 1-9. These figures show non-normality and presence of outliers, hence suitable for utilization of non-conventional measures of location as Abid et al. [1], and for quantile regression too.

[Table 3 Here]

[Figures 1 – 9 Here]
As each existing estimator based on $\beta_{(ols)}$ while their corresponding proposed estimators based on $\beta_{(0.15)}, \beta_{(0.25)}$, and $\beta_{(0.35)}$. So results of each existing estimator with their corresponding proposed estimators (say) $\hat{\theta}$, are provided in a same row, such as $\hat{\theta} = (\beta_{(ols)}, \beta_{(0.15)}, \beta_{(0.25)}, \beta_{(0.35)})$, in Tables 4 - 6, under complete information. Similarly, PRE of existing and proposed estimators (say) $\hat{\theta}'$ are also available in Tables 4 - 6, under partial information. Further, MSE of $(\hat{\theta}, \hat{\theta}')$ presented graphically in figures 10-12.

5.1 Results

The consequences of the numerical study are given in Tables 4-6 under complete and partial information setting. From Tables 4-6, we see that every one of the estimator $\bar{y}_{pi}$, with $i = 1, \ldots, 9$, considered in the proposed class, beat their corresponding estimator of existing class i.e. $\bar{y}_{abi}$. Also, we see that the proposed estimators are more productive than the existing estimators in case of partial information. From figures 10-12, we see that, proposed estimators have small MSE as compare to their corresponding ones. Hence, in both cases, the proficiency increase yielded by the new estimators is striking in populations 1-3.

6 Final remarks

In this paper, beginning from some ongoing contribution of Abid et. al [1] for mean estimation under design/configuration based sampling, we have constructed a class of quantile regression-ratio-type estimators for the population mean whenever the data is non-normal and contaminated with outliers. The proposed class beats existing ones. In case of complete information, where population mean of auxiliary variable is known, we have determined the MSE expressions, theoretically. Further, we have also introduced the existing and proposed classes to the case of partial information. In the article, three real life populations have been considered. The numerical computations well underscore the predominance of the proposed class in both the complete and partial information setting, at least for the considered experimental circumstances in the article. Hence, survey practitioners should use proposed class since it might expand the odds of getting progressively proficient evaluations of obscure objective parameters.
References


List of table captions
Table 1: Family members of Abid et al. (2016b)
Table 2: Family members of proposed class
Table 3: Characteristics of Populations
Table 4: PRE of proposed and existing estimators for Pop-1
Table 5: PRE of proposed and existing estimators for Pop-2
Table 6: PRE of proposed and existing estimators for Pop-3

List of figure captions
Figure 1: Pop-1
Figure 2: Boxplot Pop-1
Figure 3: Scatter-plot Pop-1
Figure 4: Pop-2
Figure 5: Boxplot Pop-2
Figure 6: Scatter-plot Pop-2
Figure 7: Pop-3
Figure 8: Boxplot Pop-3
Figure 9: Scatter-plot Pop-3
Figure 10: MSE of Population-1
Figure 11: MSE of Population-2
Figure 12: MSE of Population-3
Authors’ Biographies

Usman Shahzad obtained his M.Sc. degree in statistics from International Islamic University, Islamabad, Pakistan. He completed his M.Phil. in statistics from Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. Currently he is studying his Ph.D. in statistics from International Islamic University, Islamabad, Pakistan. He served as a lecturer in Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. He has published more than 25 research papers in research journals. His research interests include Survey Sampling, Extreme value theory, Stochastic Process, Probability and Non-parametric statistics.

Muhammad Hanif obtained his M.Sc. degree in statistics from University of Agriculture, Faisalabad, Pakistan. He completed his M.Phil. in statistics from Government College University, Lahore, Pakistan. He completed his Ph.D. degree in Statistics from Zhejiang University, Hangzhou, China. Currently he is an Associate Prof. and serving as a chairman of the Department of Mathematics and Statistics in Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. He has published more than 50 research papers in research journals. His main research interest is in Stochastic Process, Probability and Mathematical Statistics, Survey Sampling and Non Parametric Estimation.

Irsa Sajjad has completed her M.Sc. and M.Phil. in statistics from Pir Mehr Ali Shah Arid agriculture university Rawalpindi, Pakistan. She is serving as a lecturer in Department of Lahore Business School-University of Lahore, Islamabad, Pakistan. Her research interests include Sampling, Stochastic Process, Methods and Probability.

Malik Muhammad Anas has completed his M.Sc. and M.Phil. in statistics from Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. Currently he is studying his Ph.D. in statistics from Nanjing University of Science and Technology, Nanjing, China. He served as a lecturer in Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Pakistan. His research interests include Non-parametric regression, Stochastic Process, Kernel Smoothing and Sampling Methods.
### Table 1: Family members of Abid et al. (2016b)

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Under complete information</th>
<th>Under partial information</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}_{ab1}$</td>
<td>$(1, TM)$</td>
<td>$\tilde{y}_{ab1}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab2}$</td>
<td>$(C_x, TM)$</td>
<td>$\tilde{y}_{ab2}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab3}$</td>
<td>$(\rho, TM)$</td>
<td>$\tilde{y}_{ab3}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab4}$</td>
<td>$(1, MR)$</td>
<td>$\tilde{y}_{ab4}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab5}$</td>
<td>$(C_x, MR)$</td>
<td>$\tilde{y}_{ab5}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab6}$</td>
<td>$(\rho, MR)$</td>
<td>$\tilde{y}_{ab6}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab7}$</td>
<td>$(1, HL)$</td>
<td>$\tilde{y}_{ab7}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab8}$</td>
<td>$(C_x, HL)$</td>
<td>$\tilde{y}_{ab8}$</td>
</tr>
<tr>
<td>$\bar{y}_{ab9}$</td>
<td>$(\rho, HL)$</td>
<td>$\tilde{y}_{ab9}$</td>
</tr>
</tbody>
</table>

### Table 2: Family members of proposed class

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Under complete information</th>
<th>Under partial information</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}_{p1}$</td>
<td>$(1, TM)$</td>
<td>$\tilde{y}_{p1}$</td>
</tr>
<tr>
<td>$\bar{y}_{p2}$</td>
<td>$(C_x, TM)$</td>
<td>$\tilde{y}_{p2}$</td>
</tr>
<tr>
<td>$\bar{y}_{p3}$</td>
<td>$(\rho, TM)$</td>
<td>$\tilde{y}_{p3}$</td>
</tr>
<tr>
<td>$\bar{y}_{p4}$</td>
<td>$(1, MR)$</td>
<td>$\tilde{y}_{p4}$</td>
</tr>
<tr>
<td>$\bar{y}_{p5}$</td>
<td>$(C_x, MR)$</td>
<td>$\tilde{y}_{p5}$</td>
</tr>
<tr>
<td>$\bar{y}_{p6}$</td>
<td>$(\rho, MR)$</td>
<td>$\tilde{y}_{p6}$</td>
</tr>
<tr>
<td>$\bar{y}_{p7}$</td>
<td>$(1, HL)$</td>
<td>$\tilde{y}_{p7}$</td>
</tr>
<tr>
<td>$\bar{y}_{p8}$</td>
<td>$(C_x, HL)$</td>
<td>$\tilde{y}_{p8}$</td>
</tr>
<tr>
<td>$\bar{y}_{p9}$</td>
<td>$(\rho, HL)$</td>
<td>$\tilde{y}_{p9}$</td>
</tr>
</tbody>
</table>
Table 3: Characteristics of Populations

<table>
<thead>
<tr>
<th></th>
<th>(Pop-1)</th>
<th></th>
<th>(Pop-2)</th>
<th></th>
<th>(Pop-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{Y}$</td>
<td>555.4345</td>
<td>$X$</td>
<td>878.1624</td>
<td>$HL$</td>
<td>628.0575</td>
</tr>
<tr>
<td>$S_y$</td>
<td>584.826</td>
<td>$S_x$</td>
<td>1084.678</td>
<td>$MR$</td>
<td>1964.483</td>
</tr>
<tr>
<td>$HL$</td>
<td>628.0575</td>
<td>$TM$</td>
<td>536.4088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Sy$</td>
<td>584.826</td>
<td>$Sx$</td>
<td>878.1624</td>
<td>$MR$</td>
<td>15.1515</td>
</tr>
<tr>
<td>$Sx$</td>
<td>1084.678</td>
<td>$MR$</td>
<td>393.9394</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.8038</td>
<td>$\beta_{(ols)}$</td>
<td>0.4334034</td>
<td>$\beta_{(0.25)}$</td>
<td>0.3542911</td>
</tr>
<tr>
<td>$C_x$</td>
<td>1.235168</td>
<td>$\beta_{(0.15)}$</td>
<td>0.3447828</td>
<td>$\beta_{(0.35)}$</td>
<td>0.3570439</td>
</tr>
<tr>
<td>$C_y$</td>
<td>1.052916</td>
<td>$\beta_{(0.15)}$</td>
<td>0.3447828</td>
<td>$\beta_{(0.35)}$</td>
<td>0.3570439</td>
</tr>
</tbody>
</table>

- Table 4: PRE of proposed and existing estimators for Pop-1

<table>
<thead>
<tr>
<th></th>
<th>Under complete information</th>
<th>Under partial information</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}$</td>
<td>$\beta_{(ols)}$</td>
<td>$\beta_{(0.15)}$</td>
</tr>
<tr>
<td>$\hat{\theta}_1$</td>
<td>3769.87</td>
<td>4961.64</td>
</tr>
<tr>
<td>$\hat{\theta}_2$</td>
<td>3436.58</td>
<td>4510.95</td>
</tr>
<tr>
<td>$\hat{\theta}_3$</td>
<td>4176.24</td>
<td>5500.01</td>
</tr>
<tr>
<td>$\hat{\theta}_4$</td>
<td>6870.27</td>
<td>8480.31</td>
</tr>
<tr>
<td>$\hat{\theta}_5$</td>
<td>6313.21</td>
<td>7977.74</td>
</tr>
<tr>
<td>$\hat{\theta}_6$</td>
<td>7407.26</td>
<td>8866.77</td>
</tr>
<tr>
<td>$\hat{\theta}_7$</td>
<td>4056.99</td>
<td>5343.50</td>
</tr>
<tr>
<td>$\hat{\theta}_8$</td>
<td>3680.05</td>
<td>4840.89</td>
</tr>
<tr>
<td>$\hat{\theta}_9$</td>
<td>4509.32</td>
<td>5929.52</td>
</tr>
<tr>
<td>( \hat{\theta} )</td>
<td>( \beta_{(ols)} )</td>
<td>( \beta_{(0.15)} )</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>( \hat{\theta}_1 )</td>
<td>3254.74</td>
<td>4965.69</td>
</tr>
<tr>
<td>( \hat{\theta}_2 )</td>
<td>2509.54</td>
<td>3638.07</td>
</tr>
<tr>
<td>( \hat{\theta}_3 )</td>
<td>3334.54</td>
<td>5113.07</td>
</tr>
<tr>
<td>( \hat{\theta}_4 )</td>
<td>19012.92</td>
<td>36065.05</td>
</tr>
<tr>
<td>( \hat{\theta}_5 )</td>
<td>12379.90</td>
<td>24741.64</td>
</tr>
<tr>
<td>( \hat{\theta}_6 )</td>
<td>19610.21</td>
<td>36757.77</td>
</tr>
<tr>
<td>( \hat{\theta}_7 )</td>
<td>3643.28</td>
<td>5692.21</td>
</tr>
<tr>
<td>( \hat{\theta}_8 )</td>
<td>2730.46</td>
<td>4022.14</td>
</tr>
<tr>
<td>( \hat{\theta}_9 )</td>
<td>3741.34</td>
<td>5879.02</td>
</tr>
</tbody>
</table>

Table 6: PRE of proposed and existing estimators for Pop-3

<table>
<thead>
<tr>
<th>( \hat{\theta} )</th>
<th>( \beta_{(ols)} )</th>
<th>( \beta_{(0.15)} )</th>
<th>( \beta_{(0.25)} )</th>
<th>( \beta_{(0.35)} )</th>
<th>( \hat{\beta}' )</th>
<th>( \beta_{(ols)} )</th>
<th>( \beta_{(0.15)} )</th>
<th>( \beta_{(0.25)} )</th>
<th>( \beta_{(0.35)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\theta}_1 )</td>
<td>2488.90</td>
<td>2824.16</td>
<td>2981.56</td>
<td>3083.08</td>
<td>( \hat{\beta}'_1 )</td>
<td>1853.20</td>
<td>1935.58</td>
<td>1969.90</td>
<td>1990.75</td>
</tr>
<tr>
<td>( \hat{\theta}_2 )</td>
<td>3063.47</td>
<td>3403.58</td>
<td>3546.67</td>
<td>3632.31</td>
<td>( \hat{\beta}'_2 )</td>
<td>1986.79</td>
<td>2050.79</td>
<td>2075.07</td>
<td>2088.95</td>
</tr>
<tr>
<td>( \hat{\theta}_3 )</td>
<td>3098.65</td>
<td>3436.36</td>
<td>3577.02</td>
<td>3660.63</td>
<td>( \hat{\beta}'_3 )</td>
<td>1993.86</td>
<td>2056.48</td>
<td>2080.05</td>
<td>2093.43</td>
</tr>
<tr>
<td>( \hat{\theta}_4 )</td>
<td>2992.03</td>
<td>3335.88</td>
<td>3483.27</td>
<td>3572.64</td>
<td>( \hat{\beta}'_4 )</td>
<td>1972.09</td>
<td>2038.78</td>
<td>2064.49</td>
<td>2079.33</td>
</tr>
<tr>
<td>( \hat{\theta}_5 )</td>
<td>3490.14</td>
<td>3770.11</td>
<td>3867.94</td>
<td>3918.58</td>
<td>( \hat{\beta}'_5 )</td>
<td>2065.65</td>
<td>2110.31</td>
<td>2124.79</td>
<td>2132.08</td>
</tr>
<tr>
<td>( \hat{\theta}_6 )</td>
<td>3515.64</td>
<td>3789.35</td>
<td>3883.19</td>
<td>3930.94</td>
<td>( \hat{\beta}'_6 )</td>
<td>2069.93</td>
<td>2113.20</td>
<td>2127.01</td>
<td>2133.83</td>
</tr>
<tr>
<td>( \hat{\theta}_7 )</td>
<td>2495.39</td>
<td>2831.08</td>
<td>2988.55</td>
<td>3090.05</td>
<td>( \hat{\beta}'_7 )</td>
<td>1854.93</td>
<td>1937.14</td>
<td>1971.36</td>
<td>1992.15</td>
</tr>
<tr>
<td>( \hat{\theta}_8 )</td>
<td>3069.89</td>
<td>3409.59</td>
<td>3552.25</td>
<td>3637.53</td>
<td>( \hat{\beta}'_8 )</td>
<td>1988.09</td>
<td>2051.83</td>
<td>2075.99</td>
<td>2089.78</td>
</tr>
<tr>
<td>( \hat{\theta}_9 )</td>
<td>3104.98</td>
<td>3442.22</td>
<td>3582.43</td>
<td>3665.65</td>
<td>( \hat{\beta}'_9 )</td>
<td>1995.12</td>
<td>2057.49</td>
<td>2080.93</td>
<td>2094.22</td>
</tr>
</tbody>
</table>
Figure 1: Pop-1
Figure 2: Boxplot Pop-1
Figure 3: Scatter-plot Pop-1
Figure 4: Pop-2
Figure 5: Boxplot Pop-2
Figure 6: Scatter-plot Pop-2
Figure 7: Pop-3
Figure 8: Boxplot Pop-3
Figure 9: Scatter-plot Pop-3
Figure 10: MSE of Population-1

Figure 11: MSE of Population-2

Figure 12: MSE of Population-3