L-Moments and calibration based variance estimators under double stratified random sampling scheme: an application of covid-19 pandemic

1,2 Usman Shahzad*, 1 Ishfaq Ahmad, 3,4 Ibrahim Mufrah Almanjahie, 2 Muhammad Hanif, 5 Nadia H. Al - Noor

1 Department of Mathematics and Statistics, International Islamic University, Islamabad 46000, Pakistan.
2 Department of Mathematics and Statistics - PMAS-Arid Agriculture University, Rawalpindi 46300, Pakistan.
3 Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia.
4 Statistical Research and Studies Support Unit, King Khalid University, Abha 62529, Saudi Arabia.
5 Department of Mathematics, College of Science, Mustansiriyah University, Baghdad 10011, Iraq.

Email Corresponding Author: usman.stat@yahoo.com (Shahzad, U.), Cell: 00923315995587
Email: ishfaq.ahmad@iiu.edu.pk (Ahmad, I.)
Email: imalmanjahi@kku.edu.sa (Almanjahie, I.)
Email: mhpuno@hotmail.com (Hanif, M.)
Email: nadialnoor@uomustansiriyah.edu.iq (Al - Noor, N.)

Abstract

The presence of extreme events gives rise to outrageous results regarding population parameters and their estimates using traditional moments. Traditional moments are usually influenced by extreme observations. In this paper, we propose some new calibration estimators under the L-Moments scheme for variance which is one of the most important population parameters. Some suitable calibration constraints under double stratified random sampling are also defined for these estimators. Our proposed estimators based on L-Moments are relatively more robust in presence of extreme values. The empirical efficiency of proposed estimators is also calculated through simulation. Covid-19 pandemic data from January 22, 2020, to August 23, 2020, is considered for simulation study.
Keywords: Extreme observations, Variance estimation, L-Moments, Calibration, Double stratified random sampling.

1 Introduction

Auxiliary information can be utilized at various stages to improve the efficiency of the variance estimator for a finite population case. There are numerous real life examples where a roughly linear relationship between the study variable $Y$ and the auxiliary variable $X$ exists. Consider height and weight: taller people tend to be heavier. Similarly, body mass index and total cholesterol: there is a direct positive relationship between body mass index and total cholesterol. Such sort of linear relationship allows researchers to use auxiliary variable $X$ for improved estimation of any parameter of study variable $Y$. For more discussion on auxiliary information, interested readers may refer to Koyuncu [1], Al-Omari [2], Zaman [3, 4], Naz et al. [5, 6], and Shahzad et al. [7, 8]. An alternative method for the situations in which an abundance of auxiliary information is available is ranked set sampling due to McIntyre [9]. It is more cost-efficient than simple random sampling method, for example, see Adel Rastkhiz et al. [10], Zamanzade and Vock [11], Zamanzade and Wang [12], Zamanzade and Mahdizadeh [13], Zamanzade and Wang [14], Mahdizadeh and Zamanzade [15] and Shahzad et al. [16].

Double sampling is a technique where information related to auxiliary variable $X$ is not available at first-phase, and the information related to study variable $Y$ is available on a smaller sub-sample chosen from the first-phase sampling. For instance, let us think about the situation, where a physiologist needs to assess variation (variance) of leaf zone (area) for another strain of wheat. It might, in some cases, not be alluring to pluck all the leaves in the total population of 120 plants and get weight $X$ for constructing a variance estimator of the leaf zone $Y$. It will, accordingly, be more proper to choose a substantially large first-stage sample of leaves and measure weight $X$ for the sample leaves. A sub-sample from this underlying sample of leaves could then be chosen to decide the leaf zone. An estimate for the variance of weight from all 120 plants could then be acquired from the perceptions made on the first-phase sample. This estimate of weight $X'$ would then be able to utilize instead of population weight $X$, in the variance estimation of the leaf zone $Y$.

Let $(Y, X)$ belong to the population $\Omega = \{v_{11}, v_{12}, \ldots, v_n\}$ of size $N$. $\Omega$ is stratified (grouped) into $H''$ strata. Further, $\phi_h = \frac{N_h}{N}$ is stratum weight where $N_h$ is denoting the size of $h^{th}$ stratum for $h = 1, 2, \ldots, H''$. The overall size of the population containing all the strata is denoted by $\sum_{h=1}^{H''} N_h = N$. 

Now the first phase simple random sample (srs) of size \( n'_h \) is drawn without replacement from the \( h^{th} \) stratum such that \( \sum_{h=1}^{H''} n'_h = n' \), and then a second stage sample \( n_h \) \( (n_h < n'_h) \) is selected. In light of this double stratified random sampling design, the traditional estimator of variance is as given below

\[
T_o = \sum_{h=1}^{H''} \phi_h s_{y'h}^2
\]

where \( s_{y'h}^2 \) is denoting the traditional variance of study variable in \( h^{th} \) stratum for \( h = 1, 2, \ldots, H'' \).

It is worth to note that \( T_o \) is the traditional unbiased variance estimator under double stratified random sampling. \( T_o \) is based on traditional moments and hence usually influenced by extreme values. In literature, much of the developments have been done to tackle this issue regarding mean estimation. For instance, Zaman and Bulut [17, 18] introduced robust regression techniques for controlling the influence of extreme values. Ali et al. [19] extended their idea for the mean estimation of a sensitive variable. Abid et al. [20] used some non-conventional descriptive measures of statistics for variance estimation. It is interesting to note that the estimates of mean and variance used in all these described studies are based on traditional moments. On contrary, in this study, we get Linear Moments (L-Moments) motivation to construct some new estimators of variance based upon L-Moments characteristics of auxiliary and study variables rather than traditional moments. L-Moments are highly robust in presence of extreme observations and can provide a suitable estimate of population variance under a double stratified sampling scheme.

Motivated by the above documented developments, in this article, we propose two new estimators for estimating population variance by employing more meticulous use of an auxiliary variable. The objective is met by utilizing the L-Moments characteristics such as L-scale, L-location, L-skewness, and L-kurtosis of auxiliary variable. The applicability of the proposition is further demonstrated in double stratified random sampling scheme by employing covid-19 data set taken from four continents of the world. Moreover, keenly persuaded comparative investigation with respect to traditional unbiased variance estimator is conducted by means of numerical simulations. The simulation evaluation reveals the superior performance of the proposed estimators.

The rest of the article is arranged in the following major parts. In section 2, we present preliminaries with reference to L-Moments along with proposed estimators. The simulation based performance evaluation is persuaded in section 3, whereas the general conclusion is documented in section 4.
2 L-Moments and proposed estimators

2.1 Extreme events and L-Moments

We observe extreme events in various fields of life. The ongoing covid-19 pandemic is also an example of an extreme event. These events influence badly on human culture. Therefore, it is important to control the effect of extreme values of these events for getting a better estimate of the population parameters such as mean, variance and quantiles. As we mentioned earlier, our study is based on variance estimation which is one of the most important population parameter. Up to our knowledge, different variance estimators are developed based on traditional moments which are usually influenced by extreme values. An elective procedure that has the critical ability to settle this issue is L-Moments. L-Moments are exceptionally affected by extreme values as compared to traditional moments (Hosking [21]).

L-Moments are based on linear combinations of order statistics. Note that "L" in L-Moments representing their linearity. Hence these moments are free from higher powers and known as linear moments. This is also one of the major differences between traditional (nonlinear) moments and L-Moments. For the auxiliary variable $X$, $h^{th}$ stratum, one may define the population L-Moments as follows (Hosking [21])

\[
\begin{align*}
L_{1x} &= E(X_{1:1}) \\
L_{2x} &= \frac{E(X_{2:2} - X_{1:2})}{2} \\
L_{3x} &= \frac{E(X_{3:3} - 2X_{2:3} + X_{1:3})}{3} \\
L_{4x} &= \frac{E(X_{4:4} - 3X_{3:4} + 3X_{2:4} + X_{1:4})}{4}.
\end{align*}
\]

The sample L-moments can be written as:

\[
\begin{align*}
\hat{L}_{1x} &= \left( \frac{n}{1} \right)^{-1} \sum_{k=1}^{n} x_{h(k)}, \\
\hat{L}_{2x} &= \frac{1}{2} \left( \frac{n}{2} \right)^{-1} \sum_{k=1}^{n} \left\{ \binom{k-1}{1} - \binom{n-k}{1} \right\} x_{h(k)}, \\
\hat{L}_{3x} &= \frac{1}{3} \left( \frac{n}{3} \right)^{-1} \sum_{k=1}^{n} \left\{ \binom{k-1}{2} - 2 \binom{k-1}{1} \right\} \left\{ \binom{n-k}{1} + \binom{n-k}{2} \right\} x_{h(k)}, \\
\hat{L}_{4x} &= \frac{1}{4} \left( \frac{n}{4} \right)^{-1} \sum_{k=1}^{n} \left\{ -3\binom{k-1}{2} \binom{n-k}{1} + 3\binom{k-1}{1} \binom{n-k}{2} - \binom{n-k}{3} \right\} x_{h(k)},
\end{align*}
\]
where $x_{h(k)}$ is representing $k^{th}$ order statistics with binomial coefficient ($\binom{x}{k}$). Further, we can write the mathematical expressions of L-Moments for the study variable by adapting the structure of the sample and population L-Moments related to auxiliary variable. For a detailed study about L-Moments see (Hosking and Wallis [22]).

Some notations for upcoming proposed work in light of L-Moments with respect to $h^{th}$ stratum are given below:

- $\bar{x}_{h\ell} = L_{1x_{h\ell}}, \hat{x}_{h\ell} = \hat{L}_{1x_{h\ell}}$ are population and sample means (L-location) of auxiliary variable based on L-Moments.
- $\bar{y}_{h\ell} = L_{1y_{h\ell}}, \hat{y}_{h\ell} = \hat{L}_{1y_{h\ell}}$ are population and sample means (L-location) of study variable based on L-Moments.
- $S^2_{hx\ell} = L^2_{2x_{h\ell}}, s^2_{hx\ell} = \hat{L}^2_{1x_{h\ell}}$ are population and sample variance (L-dispersion) of auxiliary variable based on L-Moments.
- $S^2_{hy\ell} = L^2_{2y_{h\ell}}, s^2_{hy\ell} = \hat{L}^2_{1y_{h\ell}}$ are population and sample variance (L-dispersion) of study variable based on L-Moments.

### 2.2 Calibration approach and proposed variance estimators

Calibration is one of the general methods for parameter estimation in which we improve the original weights $\phi_h$ by minimizing chi-square or any other suitable loss function. The improved weights are named calibrated weights. However, the minimization of the loss function is based on some suitable calibration constraints. These constraints belong to the auxiliary variable. Deville and Sarndal [23] initially developed the concept of calibration based estimation of parameters. Tracy et al. [24] introduced the idea of calibration based estimation in double stratified random sampling. Koyuncu [25] extended her idea by defining the new and unique constraint i.e. combination of original and calibrated weights. It is interesting to note that descriptive statistics (mean and variance etc) used in these studies were based upon traditional moments. However, no attempt is made to utilize the L-Moments characteristics, which are substantially robust in presence of extreme values. In this study, we propose L-Moments characteristics based calibration estimators of population variance under double stratified random sampling scheme as:

$$G_{st(j)} = \sum_{h=1}^{H''} \phi_{h} \bar{s}_{hy\ell}^2 \quad \text{for } j = 1, 2 \quad (2.1)$$
where \( \vartheta'_{h} \) are calibration weights, using the chi-square loss function

\[
L(\vartheta'_{h}, \phi_{h}) = \sum_{h=1}^{H'} \frac{(\vartheta'_{h} - \phi_{h})^2}{\phi_{h} \Delta_{h}} 
\]  
(2.2)

and subject to the following calibration constraints

\[
\sum_{h=1}^{H'} \vartheta'_{h} \bar{x}_{h, \ell} = \sum_{h=1}^{H'} \phi_{h} \bar{X}_{h, \ell} 
\]  
(2.3)

\[
\sum_{h=1}^{H'} \vartheta'_{h} s_{h, x, \ell}^2 = \sum_{h=1}^{H'} \phi_{h} S_{h, x, \ell}^2 
\]  
(2.4)

\[
\sum_{h=1}^{H'} \vartheta'_{h} \hat{\tau}_{h, x, \ell{(j)}} = \sum_{h=1}^{H'} \phi_{h} \tau_{h, x, \ell{(j)}} 
\]  
(2.5)

where \( \tau_{h, x, \ell{(j)}} \) (for \( j = 1, 2 \)) are the population L-skewness and L-kurtosis of the auxiliary variable \( X \), respectively. Similarly, \( \hat{\tau}_{h, x, \ell{(j)}} \) are the sample L-skewness and L-kurtosis of the auxiliary variable \( X \), respectively. \( \Delta_{h} \) is suitably chosen weights to decide different forms of estimators, see Koyuncu (2018).

The Lagrange function is given by

\[
\Omega = \sum_{h=1}^{H'} \frac{(\vartheta'_{h} - \phi_{h})^2}{\phi_{h} \Delta_{h}} - 2\lambda'_{1} \left( \sum_{h=1}^{H'} \vartheta'_{h} \bar{x}_{h, \ell} - \sum_{h=1}^{H'} \phi_{h} \bar{X}_{h, \ell} \right) 
\]  
- \[
2\lambda'_{2} \left( \sum_{h=1}^{H'} \vartheta'_{h} s_{h, x, \ell}^2 - \sum_{h=1}^{H'} \phi_{h} S_{h, x, \ell}^2 \right) 
\]  
- \[
2\lambda'_{3} \left( \sum_{h=1}^{H'} \vartheta'_{h} \hat{\tau}_{h, x, \ell{(j)}} - \sum_{h=1}^{H'} \phi_{h} \tau_{h, x, \ell{(j)}} \right). 
\]  
(2.6)

Minimizing the chi-square loss function (2.2) subject to the calibration constraints (2.3), (2.4) and (2.5) gives the calibration weights for stratified sampling as follows:

\[
\vartheta'_{h} = \phi_{h} + \phi_{h} \Delta_{h} (\lambda'_{1} \bar{x}_{h, x} + \lambda'_{2} s_{h, x, \ell}^2 + \lambda'_{3} \hat{\tau}_{h, x, \ell{(j)}}) 
\]  
(2.7)

Substituting (2.7) into (2.3), (2.4) and (2.5) respectively gives the following system of equations:

\[
\begin{bmatrix}
P_{11} & P_{12} & P_{13} \\
P_{12} & P_{22} & P_{23} \\
P_{13} & P_{23} & P_{33}
\end{bmatrix}
\begin{bmatrix}
\lambda'_{1} \\
\lambda'_{2} \\
\lambda'_{3}
\end{bmatrix} =
\begin{bmatrix}
P_{10} \\
P_{20} \\
P_{30}
\end{bmatrix} 
\]  
(2.8)

Solving the system of equations in (2.8) for \( \lambda'_{a} \) gives
\[
\lambda_1' = \frac{(P_{13}P_{23} - P_{12}P_{33})(P_{12}P_{20} - P_{22}P_{10}) - (P_{13}P_{22} - P_{12}P_{23})(P_{12}P_{30} - P_{23}P_{10})}{(P_{12}^2 - P_{11}P_{22})(P_{13}P_{23} - P_{12}P_{33}) - (P_{13}P_{22} - P_{12}P_{23})(P_{12}P_{13} - P_{11}P_{23})}
\]

\[
\lambda_2' = \frac{(P_{13}P_{23} - P_{12}P_{33})(P_{12}P_{10} - P_{11}P_{20}) - (P_{12}P_{22} - P_{23}P_{11})(P_{13}P_{20} - P_{12}P_{30})}{(P_{12}^2 - P_{11}P_{22})(P_{13}P_{23} - P_{12}P_{33}) - (P_{13}P_{22} - P_{12}P_{23})(P_{12}P_{13} - P_{11}P_{23})}
\]

\[
\lambda_3' = \frac{(P_{12} - P_{11}P_{22})(P_{13}P_{20} - P_{12}P_{30}) - (P_{13}P_{22} - P_{12}P_{23})(P_{12}P_{10} - P_{11}P_{20})}{(P_{12}^2 - P_{11}P_{22})(P_{13}P_{23} - P_{12}P_{33}) - (P_{13}P_{22} - P_{12}P_{23})(P_{12}P_{13} - P_{11}P_{23})}
\]

where

\[
P_{11} = \sum_{h=1}^{H''} \phi_h \Delta_h \bar{x}_h^2, \quad P_{22} = \sum_{h=1}^{H''} \phi_h \Delta_h \bar{s}_{h\ell}^2, \quad P_{33} = \sum_{h=1}^{H''} \phi_h \Delta_h \hat{\tau}^2_{h\ell \ell (j)}
\]

\[
P_{12} = \sum_{h=1}^{H''} \phi_h \Delta_h \bar{x}_h \bar{s}_{h\ell}, \quad P_{13} = \sum_{h=1}^{H''} \phi_h \Delta_h \bar{x}_h \hat{\tau}_{h\ell \ell (j)}, \quad P_{23} = \sum_{h=1}^{H''} \phi_h \Delta_h \hat{\tau}^2_{h\ell \ell (j)}
\]

\[
P_{10} = \sum_{h=1}^{H''} \phi_h (\bar{x}_h - \bar{x}_h), \quad P_{20} = \sum_{h=1}^{H''} \phi_h (\bar{s}_{h\ell}^2 - \bar{s}_{h\ell}^2), \quad P_{30} = \sum_{h=1}^{H''} \phi_h (\hat{\tau}_{h\ell \ell (j)} - \hat{\tau}_{h\ell \ell (j)})
\]

Putting \(\lambda_1'\) in (2.7) and the resulting equation in (2.1) while setting \(\Delta_h = 1\), gives the proposed estimator for population variance as follows:

\[
G_{stj} = \sum_{h=1}^{H''} \phi_h s_{h\ell}^2 + D_{1h(\alpha)} P_{10} + D_{2h(\alpha)} P_{20} + D_{3h(\alpha)} P_{30}
\]

(2.9)

where

\[
D_{1h(\alpha)} = \frac{T_{12}[T_{14}(T_{25}T_{33} - T_{23}^2) + T_{24}(T_{13}T_{23} - T_{12}T_{23}) + T_{34}(T_{12}T_{23} - T_{13}T_{25})]}{(T_{12}^2 - T_{11}T_{22})(T_{13}T_{23} - T_{12}^2) - (T_{13}T_{22} - T_{12}T_{23})(T_{12}T_{13} - T_{11}T_{23})}
\]

\[
D_{2h(\alpha)} = \frac{T_{12}[T_{14}(T_{13}T_{23} - T_{12}T_{33}) + T_{24}(T_{11}T_{33} - T_{13}^2) + T_{34}(T_{12}T_{13} - T_{11}T_{23})]}{(T_{12}^2 - T_{11}T_{22})(T_{13}T_{23} - T_{12}^2) - (T_{13}T_{22} - T_{12}T_{23})(T_{12}T_{13} - T_{11}T_{23})}
\]

\[
D_{3h(\alpha)} = \frac{T_{12}[T_{14}(T_{13}T_{22} - T_{12}T_{23}) + T_{24}(T_{12}T_{13} - T_{11}T_{23}) + T_{34}(T_{11}T_{22} - T_{12}^2)]}{(T_{12}^2 - T_{11}T_{22})(T_{13}T_{23} - T_{12}^2) - (T_{13}T_{22} - T_{12}T_{23})(T_{12}T_{13} - T_{11}T_{23})}
\]

\[
T_{11} = \sum_{h=1}^{H''} \phi_h \bar{x}_h^2, \quad T_{22} = \sum_{h=1}^{H''} \phi_h \bar{s}_{h\ell}^4, \quad T_{33} = \sum_{h=1}^{H''} \phi_h \hat{\tau}_{h\ell \ell (j)}^2
\]

\[
T_{12} = \sum_{h=1}^{H''} \phi_h \bar{x}_h \bar{s}_{h\ell}, \quad T_{13} = \sum_{h=1}^{H''} \phi_h \bar{x}_h \hat{\tau}_{h\ell \ell (j)}, \quad T_{14} = \sum_{h=1}^{H''} \phi_h \bar{x}_h \bar{s}_{h\ell}^2
\]
\[ T_{23} = \sum_{h=1}^{H''} \phi_h s_{hx}^2 \hat{\tau}_{hx_{(j)}}, \quad T_{24} = \sum_{h=1}^{H''} \phi_h s_{hx}^2 s_{hy}, \quad T_{34} = \sum_{h=1}^{H''} \phi_h \hat{\tau}_{hx_{(j)}^2} s_{hy}. \]

### 3 Simulation study

For simulation study, we use the covid-19 pandemic data concerning the total number of recoveries as study variable \( Y \) and the total number of cases as auxiliary variable \( X \) in 4 continents of the world (as 1: Africa 2: Asia 3: Europe 4: North America) from January 22, 2020, up to August 23, 2020 (Source: https://www.worldometers.info/coronavirus).

Each continent representing a stratum. There are 49 countries in Asia, 57 countries in Africa, 48 countries in Europe and 39 countries in North America. The number of countries representing the size of each stratum. The scatter plot \((X, Y)\) for each continent is provided in Figures 1-4. These figures clearly show that covid-19 pandemic data have the issue of extreme values. Hence, the proposed L-Moments based variance estimators are suitable choices in light of Sec. 2. The design of the sampling is formed by random selection of a large first-phase sample \( n'_h \) from each continent. Note that the size of \( n'_h \) is 60% for each \( h^{th} \) stratum. After the selection of preliminary sample \( n'_h \) from each continent, we select 1000 times for the second-phase sample whose sizes are denoted with \( n_h \). The size of \( n_h \) is 25% for each \( h^{th} \) stratum. The detailed characteristics of the data are provided in Table 1. It is worth noting that the traditional correlation coefficient between \((X, Y)\) is also provided in Table 1 where the strength of the relationship justifies that the total number of cases can be used as an auxiliary variable \( X \) for estimating study variable i.e. total number of recoveries \( Y \). The empirical mean square error and percentage relative efficiency results are calculated from the following formulae

\[
MSE(G_{st(j)}) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{N} (G_{st(j)} - \sigma_y^2)^2
\]
\[
PRE(G_{st(j)}) = \frac{MSE(T_o)}{MSE(G_{st(j)})} \times 100.
\]

From Table 2, it appears that the PRE values associated with the proposed estimators are greater than 100. This means that the proposed estimators \((G_{st(1)}, G_{st(2)})\) actually outperform the traditional estimator \(T_o\). This conclusion is also achieved through Figure 5, as it is clear that the highest values of the MSE are associated with the traditional estimator, and thus its performance is low compared to the proposed estimators when the data includes extreme values.

4 Conclusion

In the current study, we introduced a new way of estimating population variance using L-Moments and calibration approach under double stratified random sampling. A simulation study has been performed using covid-19 pandemic data set as real life application of proposed estimators. Simulation based percentage relative efficiency results are provided in Table 2 and mean square error results are provided graphically in Figure 5. These results clearly show that proposed estimators have high efficiencies with small mean square error as compared to traditional unbiased variance estimator under double stratified sampling scheme. Hence, it is recommended to use proposed estimators in the presence of extreme observations. We would like to mention that some other estimators can also be derived in the forthcoming studies by adding the suitable calibration constraints based on L-Moments characteristics of auxiliary information, such as L-Moments based coefficient of variation or skewness of the auxiliary variable, to the proposed estimators given here, as in the studies of Shahzad et al. [8].

Acknowledgments The authors wish to thank editor-in-chief, Prof. S.T.A. Niaki and three anonymous referees for constructive comments which helped to improve a previous version of this paper. The authors are grateful to the Deanship 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.1/189/41. Ibrahim Mufrah Almanjahie received the grant.
References

[1] Koyuncu, N. “Efficient estimators of population mean using auxiliary attributes”, Applied Mathematics and Computation, 218, pp. 10900–10905 (2012).

[2] 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).

[3] Zaman, T. “Improvement of modified ratio estimators using robust regression methods”, Applied Mathematics and Computation, 348, pp. 627-631 (2019).

[4] Zaman, T. “An efficient exponential estimator of the mean under stratified random sampling”, Mathematical Population Studies, In Press, (2020).

[5] Naz, F., Abid, M., Nawaz, T., and Pang, T. “Enhancing efficiency of ratio-type estimators of population variance by a combination of information on robust location measures”, Scientia Iranica, 27(4), pp. 2040–2056 (2020a).

[6] Naz, F., Nawaz, T., Pang, T., and Abid, M. “Use of nonconventional dispersion measures to improve the efficiency of ratio-type estimators of variance in the presence of outliers”, Symmetry, 12(16), pp. 1–26 (2020b).

[7] Shahzad, U., Hanif, M., Sajjad, I., Anas, M. M. “Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information”, Scientia Iranica, In Press, (2020a).

[8] Shahzad, U., Ahmad, I., Almanjahie, I., Al-Noor, N. H., Hanif, M. “A new class of L-Moments based calibration variance Estimators”, Computers Materials and Continua, 66(3), pp. 3013-3028 (2021).

[9] McIntyre, G. A. “A method of unbiased selective sampling using ranked sets”, Australian Journal of Agricultural Research, 3, pp. 358-390 (1952).

[10] Adel Rastkhiz, S. E., MobiniDehkordi, A., Yadollahi Farsi, J., and Azar, A. “A new approach to evaluating entrepreneurial opportunities”, Journal of Small Business and Enterprise Development, In Press, (2019).

[11] Zamanzade, E., Vock, M. “Variance estimation in ranked set sampling using a concomitant variable”, Statistics and Probability Letters, 105, pp. 1-5. 2015.
[12] Zamanzade, E. and Wang, X. “Estimation of population proportion for judgment post-stratification”, *Computational Statistics and Data Analysis*, **112**, pp. 257-269 (2017).

[13] Zamanzade, E. and Mahdizadeh, M. “A more efficient proportion estimator in ranked set sampling”, *Statistics and Probability Letters*, **129**, pp. 28-33 (2017).

[14] Zamanzade, E. and Wang, X. “Proportion estimation in ranked set sampling in the presence of tie information”, *Computational Statistics*, **33**(3), pp. 1349-1366 (2018).

[15] Mahdizadeh, M. and Zamanzade, E. “Efficient body fat estimation using multistage pair ranked set sampling”, *Statistical Methods in Medical Research*, **28**(1), pp. 223-234 (2019).

[16] Shahzad, U., Ahmad, I., Oral, E., Hanif, M., Almanjahie, I. “Estimation of the population mean by successive use of an auxiliary variable in median ranked set sampling”, *Mathematical Population Studies*, In Press, (2020c).

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

[18] Zaman, T. and Bulut, H. “Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling”, *Communications in Statistics - Theory and Methods*, **49**(14), pp. 3407-3420 (2020).

[19] Ali, N., Ahmad, I., Hanif, M., and Shahzad, U. “Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information”, *Communications in Statistics - Theory and Methods*, In Press, (2019).

[20] Abid, M., Ahmed, S., Tahir, M., Zafar Nazir, H., and Riaz, M. “Improved ratio estimators of variance based on robust measures”, *Scientia Iranica*, **26**(4), pp. 2848-2494 (2019).

[21] Hosking, J. R. “L-moments: Analysis and estimation of distributions using linear combinations of order statistics”, *Journal of the Royal Statistical Society: Series B (Methodological)*, **52**(1), pp. 105-124 (1990).

[22] Hosking, J. R. M., Wallis, J. R. “Regional frequency analysis: an approach based on L-moments”, Cambridge university press, (2005).

[23] Deville, J. C., and Srndal, C. E. “Calibration estimators in survey sampling”, *Journal of the American statistical Association*, **87**(418), pp. 376-382 (1992).
[24] Tracy, D.S., Singh, S., Arnab, R. “Note on calibration in stratified and double sampling”, Survey Methodology, 29, pp. 99–104 (2003).

[25] Koyuncu, N. “Calibration estimator of population mean under stratified ranked set sampling design”, Communications in Statistics-Theory and Methods, 47(23), pp. 5845–5853 (2018).
### Table 1: L-Moments characteristics of Covid-19 data set

| Stratum-I | Stratum-II | Stratum-III | Stratum-IV |
|-----------|------------|-------------|------------|
| $N_1 = 57$ | $N_2 = 49$ | $N_3 = 48$ | $N_4 = 39$ |
| $n_1' = 34$ | $n_2' = 29$ | $n_3' = 29$ | $n_4' = 23$ |
| $n_1 = 9$ | $n_2 = 7$ | $n_3 = 7$ | $n_4 = 6$ |
| $\rho_{x_1y_1} = 0.9985$ | $\rho_{x_2y_2} = 0.9987$ | $\rho_{x_3y_3} = 0.9352$ | $\rho_{x_4y_4} = 0.9995$ |
| $\bar{X}_1 = 20760.74$ | $\bar{X}_2 = 127006.30$ | $\bar{X}_3 = 69487.58$ | $\bar{X}_4 = 176996.7$ |
| $\bar{Y}_1 = 15840.37$ | $\bar{Y}_2 = 99818.78$ | $\bar{Y}_3 = 41214.02$ | $\bar{Y}_4 = 99015.46$ |
| $S_1x_1 = 17345.22$ | $S_2x_2 = 104999.5$ | $S_3x_3 = 55359.53$ | $S_4x_4 = 173204.70$ |
| $S_1y_1 = 13624.12$ | $S_2y_2 = 81958.28$ | $S_3y_3 = 33099.61$ | $S_4y_4 = 96747.00$ |
| $\tau_1x_{(1)} = 0.81740$ | $\tau_1x_{(2)} = 0.75474$ | $\tau_1x_{(3)} = 0.69328$ | $\tau_1x_{(4)} = 0.96329$ |
| $\tau_2x_{(1)} = 0.70339$ | $\tau_2x_{(2)} = 0.60642$ | $\tau_2x_{(3)} = 0.47242$ | $\tau_2x_{(4)} = 0.92148$ |

### Table 2: PREs of estimators

| $T_0$ | $G_{st(1)}$ | $G_{st(2)}$ |
|-------|-------------|-------------|
| 100   | 8471.74     | 8469.912    |
Figure 1: For h=1

Figure 2: For h=3

Figure 3: For h=2

Figure 4: For h=4
Figure 5: MSE of estimators

List of table captions
Table 1: L-Moments characteristics of Covid-19 data set
Table 2: PREs of estimators

List of figure captions
Figure 1: For h=1
Figure 2: For h=3
Figure 3: For h=2
Figure 4: For h=4
Figure 5: MSE of estimators
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 for his Ph.D. 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 35 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 B.Sc. Degree in Statistics, Physics and Mathematics from Bahudin Zakaria University, Multan Pakistan in 1999. After-ward, he received M.Sc. and M.Phil degrees in Statistics from Quaid-e-Azam University, Islamabad, Pakistan in 2003 and 2005 respectively. Finally, he received a Ph.D 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 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 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 Ph.D. 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 interest are Modeling and Analysis of Ion Channel Data, Hidden Markov Models, The M Algorithm, Finite Mixture Models MCMC, Computational Methods in Statistics, Time Series Modeling, Applied Statistics, Sampling and Functional Statistics.

Muhammad Hanif obtained his M.Sc. degree in statistics from the University of Agriculture, Faisal-
abad, 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.

Nadia H. Al-Noor completed her Ph.D. degree in Mathematical Statistics. Currently, she is serving as a Prof. of Mathematical Statistics in 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 thesis. Her main research interest is in Probability and Mathematical Statistics, Survey Sampling, Statistical Inference, Reliability Theory, Robust Regression, and Non-Parametric Estimation Methods.