Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol. 51(3)(2019) pp. 133-145

Modified Maximum Likelihood Integrated Robust Ratio Estimator in Simple Random Sampling

Azaz Ahmad
Department of Statistics,
National College of Business Administration and Economics Lahore, Pakistan.
Email:azazpu17@gmail.com

Aamir Sanaullah (corresponding author)
Department of Statistics,
COMSATS Univeristy Islambad, Lahore Campus, Pakistan.
Email:chaamirsanaullah@yahoo.com

Muhammad Hanif
National College of Business Administration and Economics Lahore, Pakistan.
Email:drhanif@ncbae.edu.pk

Received: 28 March, 2018 / Accepted: 10 September, 2018 / Published online: 28 December, 2018

Abstract. A new ratio estimator using proposed modified maximum likelihood estimator (MMLE) is suggested. The properties of the new ratio estimator with respect to robustness and efficiency are studied followed by various population models. The employment of proposed MMLE into ratio estimator yields robust estimates, which is very productive for non-normal data or the data contain outliers.

AMS (MOS) Subject Classification Codes: 62DO5; 60EO5; 62H12
Key Words: Ratio estimator, modified maximum likelihood estimation, long tail symmetric family distribution, mean square error.

1. Introduction

Practically, to estimate the population parameters usually have been a common issue for all fields particular in management, engineering, actuarial, medicine and social sciences. The mean of sample values of the study variable y for size n i.e. $t_1 = \sum_{j=1}^{n} y_j/n$ is an efficient estimator of the mean of population values of size N i.e. $\bar{Y} = \sum_{j=1}^{N} y_j / N$. Although it fulfills the properties of good estimator yet it is very responsive to outliers and even a single outlier can mislead its value. The auxiliary information is used by the survey statisticians to achieve a precise estimator of $\bar{Y}$. For instance, to improve the estimators plot size can be taken as additional information to estimate production of fruits in a specific field. To utilize auxiliary information the classical estimators such as ratio, product and regression estimators are used to estimate the population mean, see (Audu and Adewara [2], Noor-ul-Amin et al. [1]). If there is positive relation between an auxiliary variable z and the study variable y and line passes through the origin then the ratio estimator is most suitable. The ratio estimator is defined as,

$$t_2 = \frac{t_1}{\bar{z}}\bar{Z}, \tag{1.1}$$

with the mean square error (MSE)

$$MSE(t_2) = var(t_1) - 2Rcov(t_1, \bar{z}) + R^2 var(\bar{z}), \tag{1.2}$$

where $\bar{z} = \sum_{1=j}^{n} z_j/n$, $\bar{Z} = \sum_{j=1}^{N} z_j/N$, $var(\bar{z}) = [\frac{(1-f)}{n}]S_z^2$, $var(t_1) = [\frac{(1-f)}{n}]S_y^2$, $f = \frac{n}{N}$ and $R = \bar{Y}/\bar{Z}$.

Moreover, t_2 is more efficient than t_1 with condition $\rho > V_z/2V_y$, where $V_z = S_z/\bar{Z}$ and $V_y = S_y/\bar{Y}$;S_z and S_y are the standard deviations of z and y over the values of entire population, respectively, and ρ is population coefficient of correlation of z and y.

The above mentioned studies are based on the postulation that the y follows normal distribution. But in some situations if the outliers appear then the data do not follow the normal distribution, see (Jenkins et al. [7], Jabbari and Nasiri [6], Haddad and Alsmadi [5]). It is noticed that the efficiency of t_1 decreases under non-normal distribution. The MMLEs are proved very helpful to enhance the efficiency (Tiku and Bhasln [12]). The MMLEs are implemented in the situations where maximum likelihood estimators (MLEs) are not in closed form. It is also discussed at large in literature e.g. see (Tiku and Suresh [15]; Tiku and Vellaisamy [16],Tiku,Islam, and Selcuk[13] and Oral [11]). The MMLEs have same large sample properties as MLEs and for small sample size they are identified efficient as MLEs (Vaughan and Tiku[17]).

Following [15] and [16], suppose a linear regression model $y_j = \beta z_j + e_j$,$j = 1, 2, ..., n$, in which y follows a long tailed symmetric (LTS) family as,

$$f(y) : LTS(\mu_y, b, \lambda) = \frac{\Gamma(b)}{\lambda\sqrt{n}\Gamma(1/2)\Gamma(b-1/2)}[1 + \frac{1}{k}\{(y-\mu_y)/\lambda\}^2]^{-b}, -\infty < y < +\infty \tag{1.3}$$

where shape parameter is b and $k = 2b - 3$ and taking b is great than 2 with $E(y) = \mu_y$; $var(y) = \lambda^2$ and kurtosis of (1.3) is $\beta_2 = \frac{3}{(1-2/k)}$ for $k \to \infty$ it reduce to normal distribution.

Let the sample of order statistics from (1.3) as $y_{(1)} \leq y_{(2)} \leq ... \leq y_{(n)}$ .The MMLEs of mean (μ_y) and standard deviation (λ) as,

$$\hat{\mu_y} = \sum_{j=1}^{n} \gamma_j y_j / w, \tag{1.4}$$

and

$$\hat{\lambda} = (D + \sqrt{D^2 + 4nQ})/\sqrt{4n(n-1)}, \quad (1.5)$$

where, $w = \sum_{j=1}^{n} \gamma_j$; $D = (2b/k)\sum_{j=1}^{n} \alpha_j y_{(j)}$ and $Q = (2b/k)\sum_{j=1}^{n} \gamma_j (y_{(j)} - \hat{\mu_y})^2$ the coefficients (α_j, γ_j) are given by

$$\alpha_j = \frac{(2/k)t_{(j)}^3}{[1 + t_{(j)}^2/k]^2} \quad \text{and} \quad \gamma_j = \frac{[1 - t_{(j)}^2/k]}{[1 + t_{(j)}^2/k]^2}. \quad (1.6)$$

Since $E(\hat{\mu_y} - \bar{Y}) = 0$ by [16]

$$var(\hat{\mu_y}) = \gamma^t \Omega \gamma \frac{\lambda^2}{w} \quad \text{and} \quad cov(\hat{\mu_y}, \bar{y}) = \gamma^t \Omega w(\lambda^2/w), \quad (1.7)$$

where w is $n \times 1$ column vector through elements $1/n$ and γ is the vector consist of the elements of γ_j. Let $t_{(j)} = E(\frac{y_{(j)} - \mu_y}{\lambda})$ and Ω be the expected vector and variance covariance matrix of order statistics of LTS family which are tabularized by Tiku and Kumra [14] for $b \le 10$. Using $t_{(j)}$ and Ω, one may get solution for each of Eq.(1.4)-(1.7).

Oral and Oral [10] advised a ratio estimator when study variable y is not from the normal distribution following the MMLE by,

$$t_3 = (\hat{\mu_y}/\bar{z})\bar{Z}, \quad (1.8)$$

where$\hat{\mu_y}$ is the MMLE of μ_y. The MSE of (1.8) is given by,

$$MSE(t_3) = var(\hat{\mu_y}) - 2Rcov(\hat{\mu_y}, \bar{z}) + R^2 var(\bar{z}), \quad (1.9)$$

where $cov(\hat{\mu_y}, \bar{z}) = 1/\beta[cov(\hat{\mu_y}, t_1) - cov(\bar{e}, \beta\bar{z}_{[.]} + \bar{e}_{[.]}]$; $\bar{z}_{[.]} = \sum_{j=1}^{n} \gamma_j z_{[j]}/w$; $e_{[.]} = y_{(j)} - \beta z_{[j]}$ and $z_{[j]}$ is the concomitant of $y_{[j]}, 1 < j < n$. Hence $cov(\hat{\mu_y}, \bar{z}) = [cov(\hat{\mu_y}, t_1)/\beta - cov(\bar{e}, \bar{e}_{[.]})/\beta]$, where β is the slope of the model and $cov(\bar{e}, \bar{e}_{[.]}) = \gamma^t \Omega w(\lambda_e^2/w)$.

Note that if $f = n/N > 0.05$ then finite population correction (f.p.c) $(1 - f)$ can be used as $cov(\hat{\mu_y}, \bar{z}) = (1 - f)/\beta[cov(\hat{\mu_y}, t_1) - cov(\bar{e}, \bar{e}_{[.]})]$.

They showed that t_3 is robust in presence of wild observations and has always less MSE as compare to the MSE of t_2, except the under study distribution is normal. Following the MMLE, Kumar and Chhaparwal [5] suggested product estimator which is robust under existence of outliers.

In this paper, the objectives are twofold. The first objective is to propose the MMLEs to improve the robustness of any estimator, and the second objective is to propose an improved ratio estimator following the proposed MMLEs. The proposed MMLEs are presented in Section 2. A new ratio estimator following the proposed MMLEs is discussed in Section 3. Efficiency comparison and simulation study for presenting application and simulation study are presented respectively in Section 4 and 5.

2. Propsed Modified MLEs

The log likelihood equation from (1.3) can be given by,

$$L = -nlog\lambda - \frac{n}{2}logk - nlog\{\frac{\Gamma(b)}{\Gamma(1/2)\Gamma(b-1/2)}\} - b\sum_{j=1}^{n}[1+\frac{1}{k}\{(y-\mu_y)/\lambda\}^2]. \tag{2.1}$$

Differentiating (2.1) with respect to parameters (μ_y and λ) respectively, we have

$$\frac{\partial L}{\partial \mu_y} = \frac{2b}{k\lambda}\sum_{j=1}^{n} h(v_j), \tag{2.2}$$

and

$$\frac{\partial L}{\partial \lambda} = -\frac{n}{\lambda} + \frac{2b}{k\lambda}\sum_{j=1}^{n} v_j h(v_j), \tag{2.3}$$

where $h(v_j) = \frac{v_j}{1+v_j^2/k}$ and $v_j = \frac{y_j-\mu_y}{\lambda}$. The MLEs of μ_y and λ are the solution of (2.2) and (2.3) respectively, which does not have explicit solutions. In such situation replacement of particular portions of (2.1) by some appropriate approximation occasionally results into simpler and efficient estimators of the parameters. Thus, the estimators in literature are known as MMLEs e.g. see Mehrota and Nanda [9] and Tiku and Suresh [15].

Following[12], the variate $v_j = \frac{y_j-\mu_y}{\lambda}$ is replaced with ordered variate $v_{(j)} = \frac{y_{(j)}-\mu_y}{\lambda}$ in (2.2) and (2.3) respectively as,

$$\frac{\partial L}{\partial \mu_y} = \frac{2b}{k\lambda}\sum_{j=1}^{n} h(v_{(j)}), \tag{2.4}$$

and

$$\frac{\partial L}{\partial \lambda} = -\frac{n}{\lambda} + \frac{2b}{k\lambda}\sum_{j=1}^{n} v_{(j)} h(v_{(j)}), \tag{2.5}$$

where $h(v_{(j)}) = \frac{v_{(j)}}{1+v_{(j)}^2/k}$. The function $h(v_{(j)})$ is linearized taking up to first order approximation by the Taylor series expansion around $t_{(j)}$ as,

$$h(v_{(j)}) \cong h(t_{(j)}) + (v_{(j)} - t_{(j)})|h'(v_{(j)})|_{v_{(j)}=t_{(j)}}, \tag{2.6}$$

where $t_{(j)} = E(v_{(j)})$, Tiku and Kumra [14] constructed the table of $t_{(j)}$ taking $b = 2(0.5)10$, $n \leq 20$ for LTS family. Further, simplification of (2.6) can be yield as, $h(v_{(j)}) \cong \frac{(2/k)t_{(j)}^3}{[1+t_{(j)}^2/k]^2} + v_{(j)}\frac{[1-t_{(j)}^2/k]}{[1+t_{(j)}^2/k]^2}$ or alternatively can be given by,

$$h(v_{(j)}) \cong \alpha_j + v_{(j)}\gamma_j, \tag{2.7}$$

where $\alpha_j = \frac{(2/k)t_{(j)}^3}{[1+t_{(j)}^2/k]^2}$ and $\gamma_j = \frac{[1-t_{(j)}^2/k]}{[1+t_{(j)}^2/k]^2}$.

In Eq. (1.4) $y_{(j)}$ is being weighted by γ_j. The γ_j gives higher weights to the central observations and lower to the extreme observations. Consequently, the extremes receive minute weights that reduce the effect of non-normality and existence of outlier (Oral and Oral [10]), (see figure 1). Figure1 shows that the small sample size, γ_j gives lower weight to extreme observations but it is not so lower that can minimize the effect of extremes. If the sample size is increased, although it is useful for the protection from extreme values yet middle observations attain balance weights relatively instead higher weights. To deal with it, the weight function is modified as under to improve the robustness and to provide appropriate weight to middle observations particularly in small and large sample size, respectively,

$$\gamma_j^* = \frac{exp(-t_{(j)}/c)/[1+exp(-t_{(j)}/c)]^2}{\sum_{j=1}^{n} \frac{exp(-t_{(j)}/c)}{[1+exp(-t_{(j)}/c)]^2}}, \tag{2.8}$$

where c is a tuning constant and assumed to be known or may be chosen so as to make estimator good in sense of robustness and efficiency. This modification may be chosen for convenience in order to improve the efficiency of the linear estimates (Downton [4]). From the figure 2, it may be observed that γ_j^* overcomes the problems that one may face while using γ_j and attains the features, we have claimed. following (2.8) ,we consider (2.7) as,

$$h(v_{(j)}) \cong \alpha_j + v_{(j)}\gamma_j^*,$$

we can write (2.4) and (2.5) respectively as,

$$\frac{\partial L}{\partial \mu_y} \cong \frac{2b}{k\lambda}\sum_{j=1}^{n}(\alpha_j + v_{(j)}\gamma_j^*) = 0, \tag{2.9}$$

and

$$\frac{\partial L}{\partial \lambda} \cong -\frac{n}{\lambda} + \frac{2b}{k\lambda}\sum_{j=1}^{n}(\alpha_j + v_{(j)}\gamma_j^*) = 0. \tag{2.10}$$

Now proposed MMLEs for μ_y and λ can be obtained from (2.9) and (2.10) respectively as,

$$\hat{\mu_y^*} = \sum_{j=1}^{n} \gamma_j^* y_j, \tag{2.11}$$

and

$$\hat{\lambda^*} = (E + \sqrt{E^2 + 4nG})/\sqrt{4n(n-1)}, \tag{2.12}$$

where, $E = (2b/k)\sum_{j=1}^{n}\alpha_j y_{(j)}$ and $G = (2b/k)\sum_{j=1}^{n}\gamma_j^*(y_{(j)} - \hat{\mu_y^*})^2$ iff $\sum_{j=1}^{n}\gamma_j^* = 1$; $\sum_{j=1}^{n}\alpha_j = 0$.

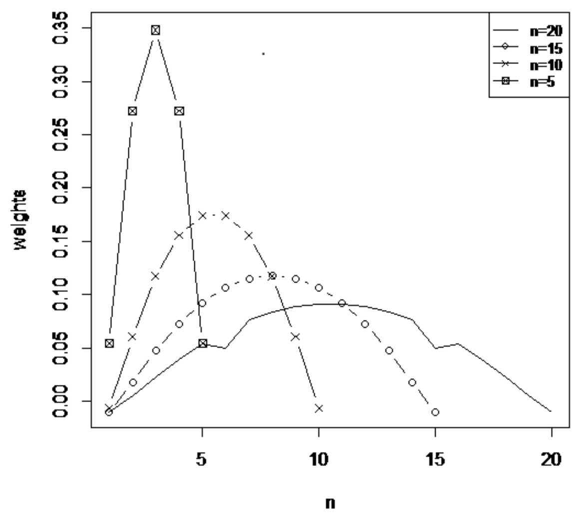

FIGURE 1. Weights function for sample size $n = 5, 10, 15$ and 20

μ_y^* is also the linear estimate in ordered observations and is unbiased mean estimator of population. The variance of μ_y^* and covariance between μ_y^* and t_1 can be obtained as,

$$var(\hat{\mu_y^*}) = \gamma^{*t}\Omega\gamma^*\lambda^2, \tag{2.13}$$

and

$$cov(\hat{\mu_y^*}, t_1) = \gamma^{*t}\Omega w\lambda^2, \tag{2.14}$$

where w is the $n \times 1$ column vector through elements $1/n$ and γ^{*t} is the transpose vector of γ^*.

3. PROPOSED ROBUST RATIO ESTIMATOR

Due to attractive features of the proposed weight function, we propose a robust ratio estimator following the proposed MMLEs of previous section by

$$t_p = \frac{\hat{\mu_y^*}}{\bar{z}}\bar{Z}. \tag{3.1}$$

For the derivation of approximate MSE of (3.1) we may follow proceed as,

$$t_p - \bar{Y} = \frac{\hat{\mu_y^*}}{\bar{z}}\bar{Z} - \bar{Y} \Rightarrow \bar{Z}(\hat{R} - R),$$

where $\hat{R} = \frac{\hat{\mu_y}^*}{\bar{z}}$ and $R = \bar{Y}/\bar{Z}$.

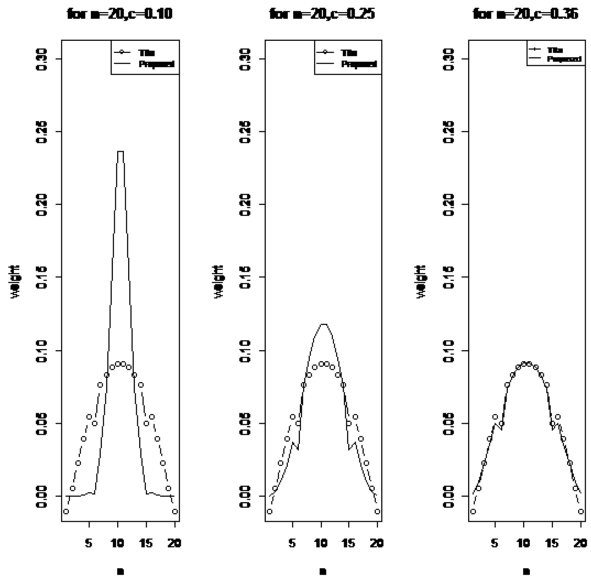

FIGURE 2. γ_j and γ_j^* at $c = 0.10, 0.25$, and 0.36 for $n = 20$

Following the Taylor series approximation of $\hat{R} - R$ about $(\bar{Z}, \bar{Y})$

$$g(\bar{z}, \hat{\mu_y^*}) \cong g(\bar{Z}, \bar{Y}) + (\bar{z} - \bar{Z})|\frac{\partial g(\bar{z}, \hat{\mu_y^*})}{\partial \bar{z}}|_{(\bar{z}=\bar{Z}),(\hat{\mu_y^*}=\bar{Y})} + (\hat{\mu_y^*} - \bar{Y})|\frac{\partial g(\bar{z}, \hat{\mu_y^*})}{\partial \hat{\mu_y^*}}|_{(\bar{z}=\bar{Z}),(\hat{\mu_y^*}=\bar{Y})}, \tag{3.2}$$

where $g(\bar{z}, \hat{\mu_y^*}) = \hat{R}$ and $g(\bar{Z}, \bar{Y}) = R$. Using (3.2) to the proposed estimator in order to obtain the MSE as,

$$\hat{R} - R \cong (\hat{\mu_y^*} - \bar{Y})\frac{1}{\bar{Z}} - (\bar{z} - \bar{Z})\frac{\bar{Y}}{\bar{Z}^2},$$

or

$$\bar{Z}^2 E(\hat{R} - R)^2 + E(\bar{z} - \bar{Z})^2 R^2 - 2R[E(\hat{\mu_y^*} - \bar{Y})(\bar{z} - \bar{Z})],$$

or the $MSE(t_p)$ is obtained as,

$$MSE(t_p) \cong var(\hat{\mu_y^*}) + R^2 var(\bar{z}) - 2R cov(\hat{\mu_y^*}, \bar{z}), \tag{3.3}$$

where $cov(\hat{\mu_y^*}, \bar{z}) = 1/\beta\{cov(\hat{\mu_y^*}, t_1) - cov(\beta\bar{z}_{[.]} + \bar{e}_{[.]}, \bar{e})\}$ or alternatively may be written as, $cov(\hat{\mu_y^*}, \bar{z}) = 1/\beta\{cov(\hat{\mu_y^*}, t_1) - cov(\bar{e}_{[.]}, \bar{e})\}$; $\bar{z}_{[.]} = \sum_{j=1}^{n} \gamma_j^* z_{[.]}$; $\bar{e}_{[.]} = \sum_{j=1}^{n} \gamma_j^* e_{[.]}$; $e_{[.]} = y_{(j)} - \beta z_{(j)}$ and $z_{(j)}$ is the concomitant of $y_(j)$, where $cov(\bar{e}_{[.]}, \bar{e}) = \gamma^{*t}\Omega w \lambda_e^2$.

4. Efficiencies comparison

To derive the conditions for which the suggested estimator is perform better than their competing estimator are given as,

$$MSE(t_p) < MSE(t_2),$$

$$cov(t_1, \bar{z}) < \frac{1}{2R}[E(t_1 - \bar{Y})^2 - E(\hat{\mu_y^*} - \bar{Y})^2] + cov(\hat{\mu_y^*}, \bar{z}),$$

$$cov(t_1, \bar{z}) < B_1, \quad (4.1)$$

where $B_1 = \frac{1}{2R}[E(t_1 - \bar{Y})^2 - E(\hat{\mu_y^*} - \bar{Y})^2] + cov(\hat{\mu_y^*}, \bar{z})$. If the condition (4.1) is satisfied then the proposed ratio estimator is more efficient than the classical ratio estimator.

$$MSE(t_p) < MSE(t_3),$$

$$cov(\hat{\mu_y}, \bar{z}) < \frac{1}{2R}[var(\hat{\mu_y}) - var(\hat{\mu_y^*})] + cov(\hat{\mu_y^*}, \bar{z}),$$

$$cov(\hat{\mu_y}, \bar{z}) < B_2, \quad (4.2)$$

where $B_2 = \frac{1}{2R}[var(\hat{\mu_y}) - var(\hat{\mu_y^*})] + cov(\hat{\mu_y^*}, \bar{z})$. Since the proposed ratio estimator is more efficient than the proposed Oral and Oral ratio estimator if the condition (4.2) is satisfied.

5. Numerical Illustration of a Practical Application

The data by Dobson and Barnett [3] in table 1 show the percentages of total calories obtained from complex carbohydrates, for twenty male insulin-dependent diabetics who had been on a high-carbohydrate diet for six months. Compliance with the regime was thought to be related to the percentage of calories as protein.

TABLE 1. Carboydrate and protein for twenty male insulin-dependent diabetics

Carb. (y)	Pro.(z)	Carb.(y)	Pro.(z)
33	14	50	17
40	15	51	19
37	18	30	19
27	12	36	20
30	15	41	15
43	15	42	16
34	14	46	18
48	17	24	13
30	15	35	18
38	14	37	14

Let the amount of carbohydrate and protein be the study variable (y) and an auxiliary variable (z) respectively. The Q-Q plot in figure 3 shows that the data of carboydrate is following the $LTS(b = 2.5)$ family. Considering $c = 0.35$, and using the equations (1.2),

TABLE 2. Computation of Example 1

$N = 20$	$var(\hat{\mu_y^*}) = 0.3956$
$n = 5$	$cov(t_1, \bar{z}) = 1.17$
$R = 2.3648$	$cov(\hat{\mu_y}, \bar{z}) = -0.4597$
$var(t_1) = 8.6274$	$cov(\hat{\mu_y^*}, \bar{z}) = -0.4851$
$var(\bar{z}) = 0.7405$	$B_1 = 1.2553$
$var(\hat{\mu_y}) = 0.0361$	$B_2 = -0.5611$

TABLE 3. MSE values of ratio estimators

$MSE(t_2)$	$MSE(t_3)$	$MSE(t_p)$
7.2349	6.3517	6.8312

(1.11) and (3.3) the values of the MSEs are evaluated and results are presented in Table 2 and 3. From Table 3, it may be concluded that non-normality is highly influenced on usual

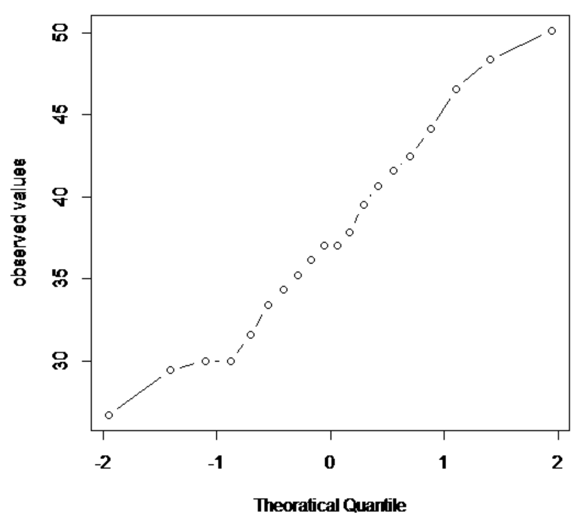

FIGURE 3. LTS family Q-Q plot of carbohydrate data for $b = 2.5$

ratio estimator. From table 3 it is also observed that the proposed ratio estimator t_p has less MSE value than the MSE of t_2 and reason is that the required condition in (4.1) is met for this data, see table 2, $cov(t_1, \bar{z}) = 1.17$ and $B_1 = 1.255354$ so $cov(t_1, \bar{z}) < B_1$.

From table 3, it is observed that required condition in (4.2) is not fulfilled for this data, and this is why the proposed ratio estimator t_p is attaining slightly larger value of MSE as compared to the MSE of t_3, see table 2,$cov(\hat{\mu_y}, \bar{z}) = -0.4597$ and $B_2 = -0.5611$ so $cov(\hat{\mu_y}, \bar{z}) > B_2$.

6. Monte- Carlo Simulation

Followed by R programming in simulations study, we apply the model $y_j = \beta z_j + e_j$ where e_j and z_j are generated independently, and compute y_j . Let e_j be random observation of error of the LTS population with zero mean and variance $\sigma_e^2 = \lambda_e^2$,$1 < j < N$, let $\prod_N$ represent the related population consisting of $(z_1, y_1), (z_2, y_2), ..., (z_N, y_N)$. In order to determine the MSEs of the t_p , we have to compute t_p for each sample. The coefficients (γ_j, γ_j^*) are calculated with $(b = 2.5, c = 0.10, 0.25, 0.38, 0.50)$. There would be T=$\binom{N}{n}$ possible samples of size n that could be drawn theoretically from $\prod_N$ using simple random sample (SRS). Obviously T is much large, therefore, a Monte Carlo stimulation study is carried out. Taking $N = 500$ in every simulation and z is from uniform distribution with parametric values 0 and 1, and take β with no loss of generality. To select the parametric value of $\sigma_e^2 = 1/12[1/\rho^2 - 1]$ such that $\rho = 0.65$. Generating $\prod_{500}$ pairs and then picked at random $S = 15000$ of all the possible $\binom{500}{n}$ SRS of size $n = 5, 10, 15$ and 20 from assumed populations, which provide 15000 values of t_p. For the comparison of the efficiencies of the t_p for a given n, compute the values of the MSEs, $MSE(t_2) = (1/S)\sum_{j=1}^{S}(t_2 - \bar{Y})^2$,$MSE(t_3) = (1/S)\sum_{j=1}^{S}(t_3 - \bar{Y})^2$ and $MSE(t_p) = (1/S)\sum_{j=1}^{S}(t_p - \bar{Y})^2$ following the models as,

1) True model $LTS(2.5, 0, 1)$
2) Outlier model of Dixon; $n - n_o$ observations from $LTS(2.5, 0, 1)$ and n_o (we do not know which) from $LTS(2.5, 0, 4)$, where n_o is calculated from the formula $[|\frac{n}{10} + \frac{1}{2}|]$.

To recognize that model (1), may be consider as the true population model for comparisons purpose and the model (2) is elected as its probable substitute. In Dixons outlier model (2), we adopt the procedure to inject the outliers into the each sample rather than the generated populations in order that all the samples drawn from $\prod_{100}$ contain outliers. With the attention that all models have the same variance as that of y , standardized the $e_j (j = 1, 2, ..., N)$ in all models. Replicated values of the MSEs of t_p and their corresponding relative efficiencies E are given in Table 4, where$E = \frac{MSE(t_1)}{MSE(t_g)}$ where $g = 1, 2, 3$ and p.

See table 4, for all values of c proposed estimator t_p has less MSE than the sample mean per unit t_1 in this study. The MSE of all estimators decreases as the sample size increases. The MSE of t_p for any specified sample size decreases as the value of c increases. As the value of c increases the relative efficiency of proposed estimator regarding t_2 and t_3 also increases in all choices of sample sizes. Proposed estimator t_p is better perform than the t_3 for $c \geq 0.38$.

It would be perceived that t_p has dual advantages with respect to robustness and efficiency, depending on the choice of survey statisticians. If they want to select efficient estimator with compromising on robustness or robust estimates with compromising on efficiency then both can be achieved with the choice of c.

TABLE 4. Mean square errors and relative efficiencies

	$n=5$					$n=10$					$n=15$					$n=20$				
True Model(1):$LTS(2.5,0,1)$: $Z\sim U(0,1)$																				
		Efficiency					Efficiency					Efficiency					Efficiency			
		c					c					c					c			
Estimators	MSE	0.10	0.25	0.38	0.50	MSE	0.10	0.25	0.38	0.50	MSE	0.10	0.25	0.38	0.50	MSE	0.10	0.25	0.38	0.50
t_1	0.0397	1.12	1.43	1.50	1.501	0.0189	1.25	1.6	1.75	1.81	0.0129	1.27	1.55	1.63	1.64	0.0101	1.64	1.98	2.08	2.09
t_2	0.0289	0.81	1.04	1.09	1.09	0.0116	0.77	0.98	1.08	1.11	0.0093	0.92	1.11	1.17	1.18	0.0065	1.05	1.27	1.33	1.33
t_3	0.0267	0.75	0.96	1.01	1.01	0.0110	0.73	0.94	1.02	1.06	0.0080	0.79	0.96	1.01	1.02	0.0048	0.79	0.96	1.01	1.01
$t_p(c=0.10)$	0.0355	1	1.28	1.34	1.34	0.0151	1	1.28	1.4	1.44	0.0102	1	1.22	1.28	1.29	0.0062	1	1.21	1.27	1.27
$t_p(c=0.25)$	0.0278	0.78	1	1.05	1.05	0.0118	0.78	1	1.09	1.13	0.0083	0.82	1	1.05	1.06	0.0051	0.83	1	1.05	1.05
$t_p(c=0.38)$	0.0265	0.75	0.95	1	1	0.0108	0.72	0.92	1	1.03	0.0079	0.78	0.95	1	1.01	0.0048	0.79	0.95	1	1
$t_p(c=0.50)$	0.0264	0.74	0.95	0.99	1	0.0104	0.69	0.89	0.97	1	0.0078	0.77	0.94	0.99	1	0.0048	0.79	0.95	1	1
Dixon outlier Model(2):$Z\sim U(0,1)$																				
t_1	0.0530	1.04	1.28	1.32	1.32	0.0260	1.25	1.57	1.68	1.71	0.0163	1.37	1.73	1.87	1.91	0.0105	1.16	1.42	1.56	1.62
t_2	0.0441	0.87	1.07	1.1	1.1	0.0189	0.91	1.14	1.23	1.24	0.0111	0.94	1.18	1.27	1.3	0.0074	0.82	1	1.1	1.14
t_3	0.0401	0.79	0.97	1	1	0.0155	0.75	0.94	1.01	1.02	0.0089	0.75	0.94	1.01	1.04	0.0069	0.76	0.93	1.02	1.06
$t_p(c=0.10)$	0.0509	1	1.23	1.27	1.27	0.0208	1	1.25	1.35	1.36	0.0119	1	1.26	1.36	1.39	0.0091	1	1.23	1.35	1.4
$t_p(c=0.25)$	0.0414	0.81	1	1.03	1.03	0.0166	0.8	1	1.07	1.09	0.0095	0.8	1	1.08	1.1	0.0074	0.82	1	1.1	1.14
$t_p(c=0.38)$	0.04001	0.79	0.97	1	1	0.0154	0.74	0.93	1	1.01	0.0088	0.74	0.92	1	1.02	0.0068	0.74	0.91	1	1.04
$t_p(c=0.50)$	0.0402	0.79	0.97	1	1	0.0152	0.73	0.92	0.99	1	0.0086	0.72	0.91	0.98	1	0.0065	0.72	0.88	0.96	1

7. Conclusion

This paper proposes a new ratio estimator utilizing the information provided by using an auxiliary variable that is measured on the whole population. The observed random variable is usually assumed to be normally distributed. However, in many real-life problems non-normal distributions are also encountered quite often. In such cases the classical estimators usually develop biases and their variances are also inflated. If a non-normal distribution exhibit long-tailed (kurtosis greater than 3) behavior, observing some large values in the sample are very much expected. In reference to the normal distribution such observations can be thought as outliers, however, they are not. In any case, the maximum likelihood method can be used in order to obtain maximum likelihood estimators of the parameters. However, in non-normal cases the likelihood equations are in intricate nonlinear form and do not yield estimators that are expressed analytically. Tiku and Suresh proposed modified maximum likelihood method through which they were able to get closed form estimators under non-normality. In this paper the authors borrow these estimators and with some modifications plugged them into their ratio estimator. It is shown, through simulation, that the ratio estimators so obtained are more efficient and comparatively robust in comparison with the other available estimators.

8. Acknowledgments

We would like to thank anonymous referees and editor for their valuable comments to improve the manuscript.

References

[1] M. N.-ul-Amin, M. Tayyab and M. Hanif, *Mean Estimation Using Even Order Ranked Set Sampling*,Punjab Univ. j. math. **51**, No. 1 (2019) 91-99.
[2] A. Audu and A. A. Adewara, , *A Modified Factor-type Estimator under Two-phase Sampling*,Punjab Univ. j. math. **49**, No. 2 (2017) 59-73.
[3] A. J. Doson and A. Barnet, *Introduction to Generalized Linear models*, Champan and Hall, New York, (2008), p96.
[4] F. Downton, *Linear estimates of parameters in the extreme value distribution*, Technometrics **8**, (1996) 3-17.
[5] F. Haddad and M. K Alsmadi, *Improvement of The Hotellings T Charts Using Robust Location Winsorized One Step M-Estimator (WMOM)*, Punjab Univ. j. math. **50**, No. 1 (2018) 97-112.
[6] N. M. Jabbari and P. Nasiri, *Bayesian Shrinkage Estimator of Parameter of Exponential Distribution with Outliers*, Punjab Univ. j. math. **50**, No. 2 (2018) 11-19.
[7] O. C. Jenkins, L. J. Ringer and H. O. Hartley, *Root estimators*, Journal of the American Statistical Association **68**, (1973) 414-419.
[8] S. Kumar and P. Chhaparwal, *A robust unbiased dual to product estimator for population mean through modified maximum likelihood in simple random sampling*, Cogent Mathematics **3**, (2016) 1-13.
[9] K. G. Mehrotra and P. Nanda, *Unbiased estimation of parameters by order statistics in the case of censored samples,* Biometrika **61**, (1974) 601-606.
[10] E. Oral and E. Oral, *A robust alternative to the ratio estimator under non-normality*, Statistics and Probability Letters **81**, (2011) 930-936.
[11] E. Oral, *Binary regression with stochastic covariates*, Communications in Statistics Theory and Methods **35**, (2006) 1429-1447.
[12] M. L. Tiku and P. Bhasln, *Usefulness of robust estimators in sample survey*, Communications in Statistics Theory and Methods **11**, (1982) 2597-2610.
[13] M. L. Tiku, M. Q. Islam and S. Selcuk, *Nonnormal Regression II. Symmetric Distributions,* Communications in Statistics Theory and Methods **30**, (2001) 1021-1045.

[14] M. L. Tiku and S. Kumra, *Expected values and variances and covariances of order statistics for a family of symmetric distributions,* In selected tables in mathematical statistics **8**, (1981) 141-270, American Mathematical Society, Providence RI, USA.

[15] M. L. Tiku and R. P. Suresh, *A new method of estimation for location and scale parameters*, Journal of Statistical Planning and Inference, **30**, (1992) 281-292.

[16] M. L. Tiku and P. Vellaisamy, *Improving efficiency of survey sample procedures through order statistics,* Journal of the Indian Society of Agricultural Statistics **49**, (1996) 363-385.

[17] D. C. Vaughan and M. L. Tiku, *Estimation and hypothesis testing for a nonnormal bivariate distribution with applications,* Mathematical and Computer Modelling **32**, (2000) 53-67.