Journal of the Nigerian Association of Mathematical Physics
***Volume 65, (October 2022 – August 2023 Issue*), pp95 – 102**

BOOTSTRAPPING CERTAIN MEASURES OF LOCATION

[1]Awariefe C. and [2]Ilo H.O.

[1]Department of Statistics, Delta State University of Science and Technology, Ozoro.
[2]Department of Computer Science, Delta State University of Science and Technology, Ozoro.

Abstract

The focus of this paper is on consistent estimates of the standard error of certain measures of location. The bootstrap approach was adopted to compute the standard error for assessing the relative efficiencies of some measures of location. The R statistical package was employed to obtain data from some distributions and real data for the analysis. Employing the re-sampling procedure inherent in bootstrapping, it was established analytically that bootstrap standard errors are smaller for the median estimator than their counterpart. The median was found to be the most robust since it produces the least bootstrap standard error and relative efficiency of less than one when compared to the other estimators under study. However, the mean was the most efficient compared to the other measures of the location under study when the distribution is normally distributed.

***Keywords*:** Bootstrap Standard Error, Trimmed Mean, Median, Winsorized Mean, Relative Efficiency.

1. Introduction

Bootstrapping was made popular by [1]; it is a database simulation method for statistical inference based on data re-sampling which can be used to study the variability of estimated characteristics of a probability distribution of a set of observations. Bootstrapping is a computer-intensive method used to estimate bias and standard errors and the construction of confidence intervals for parameters when the distribution is unknown. [2] and [3] investigated the performances of some measures of location by comparing them when the underlying distribution has heavier tails than normal. Normal distribution may provide a good approximation to most distributions that arise in practice. But unfortunately, empirical investigations indicate that departures from normality that have practical importance are common in applied works. In particular, distributions can be highly skewed, they can have heavy tails and random samples often have outliers. [4] article showed that the classical identically and independently distributed bootstrap is a valid procedure for estimating the sampling distribution of certain symmetric measures of location. [5] used the criteria for coverage accuracy and average interval length to assess the effectiveness of the sufficient, balanced, and conventional bootstrapping procedures for estimating three alternative bootstrap confidence interval methods. Both symmetric and skewed distributions were taken into account. [6] establishes empirically the utilization of bootstrapping for the major aim of producing efficient estimates of regression parameters. [7] demonstrates that M-regression parameters produce standard errors that are inconsistent and unstable, specifically when the sample under investigation is small. The alternative approach on offer in the paper was the bootstrap. [8] offers an early theoretical and practical analysis of the asymptotic theory of the bootstrap for robust measures of location. Specifically, the author described the method of estimating the bootstrap standard error of any robust measures of location; the literature also introduced some outlier detection methods and some graphical methods of summarizing data. When researchers are faced with making the best choice in choosing the appropriate measures of location estimators, the most robust estimators must be chosen to make valid conclusions. Consequently, the purpose of this paper is an attempt to assess the robustness of certain measures of central tendencies using the bootstrap approach by subjecting the estimators to real and simulated data. This paper reports yet another contribution to the kinds of research efforts described above; that is, research efforts directed towards the study of the performance of the bootstrap in certain traditional and robust measures of locations. Specifically, we demonstrate empirically that the bootstrap is a veritable instrument to enhance the efficiency of robust estimates. The rest of the paper is structured as follows: section 2 reviews some measures of location. The general

Corresponding Author**:** Awariefe C., Email: awariefe@gmail.com, Tel: +2348033471856

bootstrap procedure and bootstrap standard error are produced in section 3. The description and data analysis results are presented in section 4. The conclusions of this article are presented in section 5.

2. Materials and Methods

Review of Some Measures of Locations

2.1 The Mean

The arithmetic average, commonly known as the arithmetic mean or just the mean, is one of the most popular and practical measures of location. It is the most commonly used and also of the greatest importance of the three averages. The sum of a set of measurements divided by the total number of measurements is the mean of the set. Mathematically the mean can be expressed as a set of measurements $m_1, m_2, m_3, \ldots\ldots.., m_q$, and defined as:

$$\bar{M} = \frac{\sum_i^q x_i}{q} \quad (1)$$

where q is the sample size

If m_1,m_2,m_3,…., $m_q$occur with frequencies frequency f_1,f_2,f_3,…….., f_q respectively, then mean is defined as follows:

$$\bar{M} = \frac{\sum_i^q f_q x_q}{\sum_i^q f_q} \quad (2)$$

2.2 The median

The median (Med) is a measurement that sits in the middle of a set of data. For example, if data are arranged from smallest to largest or vice versa, the median (Med) is simply the value, say, y that falls in the middle. Calculating the median, the observation is ordered from lowest to highest value or vice versa. $y_1 \le y_2 \le y_3 \le \ldots\ldots.\le y_q$ [9]. The median is given by:

Med = $y_{\frac{(q+1)}{2}}$,

or $\frac{(q+1)^{th}}{2}\ term$, if q is an odd number (3)

and

Med =$0.5y\left(\frac{q}{2}\right)$+ $0.5y\left(\frac{q}{2}+1\right)$,

or $\frac{(\frac{q}{2})^{th} term}{2} + \frac{(\frac{q}{2}+1)^{th} term}{2}$, if n is an even number. (4)

Equivalently, the median minimizes the absolute value objective function

$$\sum_{i=1}^q |y_i - \hat{\mu}| = 0 \quad (5)$$

Taking the derivatives of the above equation gives the shape of the influence function

$$IF_{Med(y)} = \begin{cases} 1\ for\ y > 0 \\ 0\ for\ y = 0 \\ -1\ for\ y < 0 \end{cases} \quad (6)$$

as the bounded influence function indicates, the median is highly resistant to outliers, its robustness is also reflected in its breakdown point (BDP) =0.5. The disadvantage of the median is that it has relatively low efficiency compared to the mean when the distribution is normal [9] in this situation the sampling variance of the mean is:

$$\frac{\sigma^2}{q} \quad (7)$$

whereas the sampling variance of the median is:

$$\frac{\pi\sigma^2}{2q} \quad [10], \quad (8)$$

2.3 Estimation of Standard Error of the Sample Median

[11] derived a fairly simple method to estimate the standard error of the median. Trimmed means contain the usual sample median, Med when the maximum amount of trimming is used. To apply the McKean-Schrader method, compute:

$$K = \frac{q+1}{2} - z_{0.995}\sqrt{\frac{q}{4}} \quad (9)$$

Where k is rounded to the nearest integer and $z_{0.995}$ is the 0.995 quantile of a standard normal distribution. Let the observed values be in ascending order, $m_{1,} \le m_{2,} \le m_3 \ \ldots\ldots. \le m_q$.

Then the estimate of the standard error of the sample median is computed with

$$\left(\frac{M_{q-k+1} - M_{(k)}}{2z_{0.995}}\right)^2 \quad (10)$$

2.4 The Trimmed Mean

Consider a set of observations $[m_{1,} m_{2,} m_3 \ \ldots\ldots.. m_q]$ to be a random number that are ordered from smallest to largest, $m_{1,} \le m_{2,} \le m_3 \le \cdots \ldots.. \le m_q$, where m_i is called ith order statistics. The k% trimmed mean, m_t is the average of the values that remain after removing the bottom k% (i.e., smallest) and the top k% (i.e., largest) observations from the original set. The standard error of the sample mean can be relatively large when sampling from a heavy-tailed distribution, and the sample mean estimates a non-robust measure of location [8]. Let $g = [m_q]$, where $[m_q]$ is the value of $m_q$rounded down to the nearest integer. The sample trimmed mean is

$$\bar{M}_t = \frac{M_{(g+1)+\cdots+\cdots+} M_{(n-g)}}{q-2g} \quad (11)$$

The definition of sample trimmed mean given above is the one commonly used [8]. The variance of the trimmed mean is:

$$\frac{s_w^2}{q.(1-2\gamma)^2} \quad (12)$$

w is the winsorized variance (see below), q is the sample size (before trimming), and is the proportion of trimming.

$$s_w^2 = \frac{\sum_{i=0}^{q}(Wi-\bar{w})^2}{q-1} \quad (13)$$

The standard error of the trimmed mean is:

$$\frac{s_w}{\sqrt{q(1-2\gamma)}} \quad (14)$$

The best amount of trimming depends on particular circumstances; [12] recommends 20%. However, smaller sample sizes might result in the adoption of less trimming (e.g., 10%).

2.5 Estimation of Standard Error of the Trimmed Mean

To estimate the standard error of the trimmed mean based on a random sample of n observations, first winsorized the observation by transforming the ith observation, M_i using

$$W_i = \begin{cases} M_{(g+1),} & if\, M_I \le M_{(g+1)} \\ M_i, & if\, M_{(g+1)} < M_i < M_{(q-g)} \\ M_{(q-g),} & if\, M_i \ge M_{(q-g)} \end{cases} \text{[8]} \quad (15)$$

Where $M_{(1)} \le \ldots\ldots \le M_{(q)}$are the ordered statistics= (γq), q=sample size and Winsorization indicates that the g smallest values are pulled in and set equal to $M_{(g+1)}$and the g largest values are pulled in and set equal to $M_{(q-g)}$.Compute sample variance of the W_i values yield s_w^2, the Winsorized sample variance. The standard error of the trimmed mean is estimated to be:

$$\frac{s_w}{\sqrt{q}(1-2\gamma)} \quad (16)$$

2.6 The Winsorized Mean

The Winsorized mean is a robust measure of location. It involves the computation of the mean after replacing given parts of the probability distribution at the minimal and maximal end with the most remaining extreme values. It is a very useful measure of location because it is less sensitive to the presence of outliers in a dataset [13]. Let F be any distribution, and let c_γ and $c_{1-\gamma}$ be γ and 1- γ quantiles respectively. Then a γ-winsorized analogue of F is the distribution:

$$F_w(c) = \begin{cases} 0, & if\ c < c_\gamma \\ \gamma, & if\ c = c_\gamma \\ F(c), & if\ c_\gamma < c < c_{1-\gamma} \\ 1, & if\ c \ge c_{1-\gamma} \end{cases} \text{[14]} \quad (17)$$

The left tail is pulled in so that the probability of observing the value $c = c_\gamma$ is at γ, and the probability of observing any value less than c_γ after winsorization is 0, similarly the right tail is pulled in so that, after winsorization, the probability of observing a value greater than $c_{1-\gamma}$ is 0. The Winsorized mean of the distribution is:

$$\mu_w = \int_{c_\gamma}^{c_{1-\gamma}} CdF(c) + \gamma(c_\gamma + c_{1-\gamma}) \quad (18)$$

The Winsorized mean pays more attention to the central portion of a distribution by transforming the tails. The result is that the Winsorized mean is closer to the central portion of a distribution. [14]

2.7 Estimation of Standard Error of the Sample Winsorized Mean

The standard errors of the Winsorized mean estimate can be computed by:

$$\frac{q-1}{q-2g-1}\left(\frac{s_w}{\sqrt{q}}\right) \quad \text{(Dixon and Tukey, 1986)[15]} \quad (19)$$

where g= (γq) is the number of observations Winsorized in each tail, and q-2g is the number of observations that are not Winsorized.

3. General Bootstrap Procedure.

Bootstrapping is a highly computer-intensive method that is used to estimate standard errors and set confidential intervals for parameters when the form of the distribution is unknown [1], [2]. This discussion lead to the following general procedure to approximate the sampling distribution of a statistic $S = S\ (m_{1,}^{*}, m_{2}^{*}, \ldots,. m_{B}^{*}.)$ based on an observed simple random sample **M**= $(m_1, m_2, \ldots \ldots \ldots, m_q)$ of size q: where $m_{1,}^{*}, m_{2}^{*}, .\, m_{B}^{*}$ are the statistic of interest from the sampled data.

(I) Create many samples $m_1, m_2, \ldots \ldots \ldots, m_B$ where B is the bootstrapped sample size with replacement.

(II) Calculate the statistic of interest S= $m_{1,}^{*}, m_{2}^{*}, .\, m_{B}^{*}$.for each resample. where $m_{1,}^{*}, m_{2}^{*}, .\, m_{B}^{*}$ are the statistic of interest from the sampled data. The distribution of the resample statistics is called a *bootstrap distribution*.

(III) The bootstrap distribution gives information about the sampling distribution of the original statistic *S*. In particular, the bootstrap distribution gives us some idea about the center, spread, and shape of the sampling distribution of sampled data.

3.1 Bootstrap Estimates of Standard Error

[8] developed a strategy based on the general bootstrap procedure in section 3, for estimating the standard error of any estimator, the steps are stated below:

(I) Sample n from a population and compute $\bar{x}$

(II) Repeat (I) a moderate to a large number B (is the bootstrap, that is, number of times the statistics is sampled randomly), of the times to come up with estimates $m_{1,}^{*}, m_{2}^{*}, \ldots,. m_{B}^{*}$.

(III) Use the standard deviation of the B estimates in (II) to estimate the standard error.

Hence, the bootstrap standard error of a statistics is the standard deviation of the bootstrap distribution of that statistic. If the statistic of interest is the sample mean $\bar{x}$, the bootstrap standard error based on B samples is:

$$SE_{boot}, \bar{m} = \sqrt{\frac{1}{B-1}} \sum (\bar{m}^{*} - \frac{1}{B} \sum \bar{m}^{*})^2 \qquad (20)$$

In practical terms, a bootstrap sample is obtained by re-sampling with replacement n observations from, $\bar{m}_{1,} \bar{m}_{2},, \ldots,., \bar{m}_q$. This is can be done with the R program command: sample (m, size=length (m), replace=T) the R function bootse stated below:

bootse (m, nboot=1000, est=estimator) can be used to compute a bootstrap standard error of virtually any estimator. Where, x is any random variable containing the data. The argument nboot represents B, the number of bootstrap samples and defaults to 1000 if not specified. The argument est. indicates, the estimator in which the standard error is to be computed [8].

4. Application

4.1 Description of Data and Discussion of Results

The data in table 4.1 below are estimates of populations of 27 Scottish seabird (cormorant) colonies, from [16].

Table 4.1: estimates of populations of 27 Scottish seabird (cormorant) colonies.

S/N	Colony	Number of breeding pairs	S/N	Colony	Number of breeding pairs
1	Hascosay	36	15	Uyea	275
2	Whalsay	56	16	Stronsay	285
3	Mulkle	65	17	Hoy	310
4	St.Niniama	95	18	Papa Stour	348
5	S.W.Unit	136	19	Eday & Calif	354
6	Noss	141	20	Sumburgh	371
7	Bay of Bousay	146	21	Deerines	436
8	Rousay	150	22	Rapness	468
9	S.E.Yell	154	23	Fetlar	500
10	Wats Ness	156	24	S. Ronaldsay	521
11	Burra	191	25	Fair Isle	1530
12	Muckle Roe	232	26	N.W. Unst	1696
13	Noup	246	27	Foula	2000
14	P. Westray	260			

Source: [17]

Analysis of Real data of Table 4.1

Table 4.2: The bootstrap standard errors of the estimators under study from the secondary data in Table 4.1.

Measures of Locations	Bootstrap std error
Mean	93.15011
Median	50.92888
Winsorized mean	85.88737
Trimmed mean(0.1)	89.94682
Trimmed mean(0.2)	55.33853

Source: Author's computation

In this study, mean, median, winsorized mean, 10% and 20% trimmed mean, are compared in terms of their asymptotic relative efficiency values and mean squared error of $\hat{\theta}$ ($MSE(\hat{\theta})$). The asymptotic relative efficiency of estimator with $\hat{\theta}_{2,}$ respect to is $\hat{\theta}_{1,}$ defined as:

$$\text{EFF}\ (\hat{\theta}_{2,}\hat{\theta}_{1,})=\frac{Var(\hat{\theta}_{1,})}{Var(\hat{\theta}_{2),}} \qquad (21)$$

if this ratio of (21) is less than 1, it can be said that $\hat{\theta}_1$ is asymptotically more efficient than $\hat{\theta}_2$. Asymptotic relative efficiency is a useful concept that enable us to make comparisons of competing estimators.

Definition of Notations

Let Eff1, Eff2, Eff3 and EFF4, denote the asymptotic relative efficiencies of the median (med), Winsorized mean, trimmed mean (0.1), and trimmed mean (0.2) concerning mean.

Eff5, Eff6 and EFF7 denote the asymptotic relative efficiencies of winsorized mean, trimmed mean (0.1), and trimmed mean (0.2), concerning the median.

Eff8 and EFF9 denote the asymptotic relative efficiency of the trimmed mean (0.1), and trimmed mean (0.2) concerning winsorized mean.

EFF10 denote the asymptotic relative efficiency of the trimmed mean (0.1) concerning the trimmed mean (0.1). These values can be found in Table 4.3 for the real data. Table 4.2 showed that the standard error of the median is relatively smaller; this is followed by the 20% trimmed mean than the other estimators. However, the mean produced the highest standard error as expected, followed by the 10% trimmed mean; mean and 10 % trimmed mean.

Table 4.3: The relative efficiencies of each estimator for data in table 4.1

	Estimates	**Relative efficiency**
	EFF1-median wrt mean	0.5467
	EFF2- Winmeanwrt mean	0.9220
	EFF3-Trimmed mean(0.1) wrt mean	0.9656
	EFF4-Trimmed mean(0.2) wrt mean	0.5940
	EFF 5- Winmeanwrt median	1.6864
	EFF 6- Trimmean(0.1) wrt median	1.7661
	EFF 7- Trimmean(0.2) wrt median	1.0865
	EFF 8- Trimmean(0.1) wrtWinmean	1.0473
	EFF 9- Trimmean(0.2) wrtWinmean	0.6443
	EFF 10- Trimmean(0.1) wrtTrimmean(0.2)	1.6254

Source: Author's computation

In Table 4.3, when the efficiency of the measures of location is compared using their respective bootstrap standard errors all relative efficiencies concerning the mean are less than one. This means that the sample mean is less efficient than all other measures of location. When the relative efficiencies of the measures of location are compared to the median, it can be seen that all the relative efficiencies s greater than 1 which indicates that the median is the most efficient measure of location when the data are from a contaminated distribution. The trimmed mean (0.1) has relative efficiencies greater than 1 when compared to winsorized mean and trimmed mean (0.2) respectively, this indicates that it is less efficient. The relative efficiency of the trimmed mean (0.2) is less than 1 when compared to winsorized mean, which indicates that it is more efficient than winsorized mean when a dataset is from a contaminated distribution.

Table 4.4: Simulated standard normal distribution data, x=rnorm (27, 0, 1)

Measures of Locations	Bootstrap std error
Mean	0.1644
Median	0.1795
Winsorized mean	0.1853
Trimmed mean(0.1)	0.1769
Trimmed mean(0.2)	0.1696

Source: Author's computation

Table 4.5: The relative efficiencies of each estimator for data in table 4.4

	Estimates	**Relative efficiency**
	EFF1-median wrt mean	1.0918
	EFF2- Winmeanwrt mean	1.1271
	EFF3-Trimmed mean(0.1) wrt mean	1.0760
	EFF4-Trimmed mean(0.2) wrt mean	1.0316
	EFF 5- Winmeanwrt median	1.0323
	EFF 6- Trimmean(0.1) wrt median	0.9855
	EFF 7- Trimmean(0.2) wrt median	0.9448
	EFF 8- Trimmean(0.1) wrtWinmean	0.9546
	EFF 9- Trimmean(0.2) wrtWinmean	0.9152
	EFF 10- Trimmean(0.1) wrtTrimmean(0.2)	1.0430

Source: Author's computation

Table 4.6: Simulated exponential distribution data, y=rexp (27, rate=4)

Measures of Locations	Bootstrap std error
Mean	0.0628
Median	0.0475
Winsorized mean	0.0520
Trimmed mean(0.1)	0.0530
Trimmed mean(0.2)	0.0474

Source: Author's computation

Table 4.7: The relative efficiencies of each estimator for data in Table 4.6

	Estimates	**Relative efficiency**
	EFF1-median wrt mean	0.7564
	EFF2- Winmeanwrt mean	0.8280
	EFF3-Trimmed mean(0.1) wrt mean	0.8439
	EFF4-Trimmed mean(0.2) wrt mean	0.7548
	EFF 5- Winmeanwrt median	1.0947
	EFF 6- Trimmean(0.1) wrt median	1.1158
	EFF 7- Trimmean(0.2) wrt median	0.9979
	EFF 8- Trimmean(0.1) wrtWinmean	1.0192
	EFF 9- Trimmean(0.2) wrtWinmean	0.9115
	EFF 10- Trimmean(0.1) wrtTrimmean(0.2)	1.1181

Source: Author's computation

In Table 4.4 when samples are taken from standard normal distribution the mean bootstrap standard error value is the smallest, this is followed by the error values of the 20% and 10% trimmed mean. In table 4.5 the mean again has the least standard error when samples are drawn from standard normal distribution; it has relative efficiency of less than 1 when compared with other estimators. In Table 4.6 and 4.7 when samples are from skew distributions (exponential) the 20% trimmed mean and median has the smallest estimated standard error values, it is the same with their respective relative efficiencies less than 1 when compared with the other measures of location using their respective bootstrap standard errors. While the highest standard error is when the distribution is skewed and heavy tails.

5. Conclusion

In this article, we investigated the bootstrap approach to compute the standard errors and relative efficiencies of some measures of location. Based on the results of real data and those generated from different distributions, we concluded that the median is the most robust, since it has the least standard error and relative efficiency of less than one when compared to trimmed means, winsorized mean and mean estimators which are easily affected with the presence of outliers in a dataset, followed by 10% and 20% trimmed means, we recommend that they should be considered for statistical tests that utilize robust location estimators as an alternative approach when the underlying distribution has heavier tails than normal.

Appendix

The R functions used for this research are available in an R package and can be downloaded from www.R-project.org. To install the R package, created by Felix Shonbrodt. Wilcox (2012).Use the command
Install. Packages (‘”WRS,repos=’’http://R-Forge.R-project.org’’).
Access to the function is gained via the R command
Library (WRS)

The Bootstrap Algorithm for Computing Standard error of an Estimator

Step 1

Draw a bootstrap sample, $x, x_2^*, \dots \dots \dots, x_n^*$ from observed values $x_1, x_2, \dots \dots, x_n$.the values are randomly sample n observation with replacement from $x_1, x_2 \dots \dots ., x_n$.

Step 2

Using the bootstrap sample from step 1 compute the value of $\hat{\mu}_m$ and call the result $\hat{\theta}_m^*$, repeat the process B times yielding $\hat{\theta}_{m1}^*, \hat{\theta}_{m2}^*, \dots \dots, \hat{\theta}_{mb}^*$.

Step 3. Set

$\bar{\theta}^* = \frac{1}{B}\sum_{b=1}^{n} \hat{\theta}_{mb}^*$

Step 4: Set $\hat{s}_{mboot}^2 = \frac{1}{B-1}\sum_{b=1}^{n}\left(\hat{\theta}_{mb}^* - \bar{\theta}^*\right)^2$ the bootstrap estimate of the standard error.

```
> source(file.choose())
>x=c(36,56,65,95,136,141,146,150,154,156,191,232,246,260,275,285,310,348,354,371,436,468,500,521,1530,1696,2000)
> bootse(x,nboot=1000,est=mean)
> bootse(x,nboot=1000,est=winmean)
> bootse(x,nboot=1000,est=median)
> bootse(x,nboot=1000,est=mean,trim=0.1)
> bootse(x,nboot=1000,est=mean,trim=0.2)
> source(file.choose())
> x=rnorm(27,0,1)
> bootse(x,nboot=1000,est=mean)
> y=bootse(x,nboot=1000,est=winmean)
> bootse(x,nboot=1000,est=winmean)
> bootse(x,nboot=1000,est=median)
> bootse(x,nboot=1000,est=mean,trim=0.1)
> bootse(x,nboot=1000,est=mean,trim=0.2)
> bootse(x,nboot=1000,est=winmean)
> y=rexp(27,rate=4)
> bootse(y,nboot=1000,est=mean)
> bootse(y,nboot=1000,est=median)
> bootse(y,nboot=1000,est=winmean)
> bootse(y,nboot=1000,est=mean,trim=0.1)
> bootse(y,nboot=1000,est=mean,trim=0.2)
```

References

[1]. Efron, B. (1979). Bootstrap methods: *Annals of Statistics* 7, 1-26.

[2]. Efron, B. and Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman &Hall. U.S.A.

[3]. Özdemir, A.F. (2010). Comparing measures of location when the underlying distribution has heavier tails than normal. *statistikçilerDergisi* 3,8-16

[4]. Liu, R. Y. (1988). Bootstrap Procedures under Some Non-i.i.d Models. *Annals of Statistics*. Vol.18, No. 4, 1696-1708

[5]. Yildiztepe, E. (2020). Comparison of Conventional, Balanced and Sufficient Bootstrapping Approaches via Confidence Intervals and Efficiency. *Mugla Journal of Science and Technology*, 6(2), 111-120.

[6]. Breiman, P. (1996). On Robust Estimation. *Annals of Statistics*, 9. 1196 - 1217.

[7]. Hamadu, D. (2012). A Bootstrap Approach to Robust Regression. *International Journal of Applied Science and Technology.* Vol. 2 No. 9.

[8]. Wilcox, R. R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier Academic Press, Third Edition. New York.

[9]. Andersen, R. (2012). *Modern methods for robust regression*. University of Toronto. Sage University Paper Series. Canada.

[10]. Kenney, J.F and Keeping, E. S. (1962). Mathematics of Statistics. Part 2, Second edition, Van Nostrand. University of Michigan, U.S.A.

[11]. McKean, J.W and Schrader, R.M. (1984). Geometry of Robust Procedures in Linear Models. *Journal of Royal Statistical Society. Series B (Methodological) Wiley:* Vol. 42, No.3, 366-371

[12]. Wilcox, R. R. (2005). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier Academic Pres, Second Edition. New York.

[13]. Dodge, Y. (2003). *The Oxford Dictionary of Statistical Terms*. Oxford University Press.

[14]. Wilcox, R. R and Keselman, H. J. (2003). Modern Robust Data Analysis Methods: Measures of Central Tendency. Psychological Methods, 8(3), 254-274.

[15]. Dixon, W.J and Tukey, J.W. (1968). Approximate Behaviour of the Distribution of Winsorized (Trimming/Winsorization). *American Statistical Association and American Society for Quality. Vol. 10, No 1, 83-98.*

[16]. Furness, R.W., and Birkhead, T.R. (1984). Seabird colony distributions suggest for food supplies during the breeding season. *Nature, 311, 655-656.*

[17]. Hamilton, L.C. (1992**).** *Regression with Graphics- Second Course in Applied Statistics*. Brooks/Cole Publishing Company. Belmot, California.