**Journal of Management (JOM)**
Volume 7, Issue 2, March – April 2020, pp. 14-18, Article ID: JOM_07_02_003
Available online at http://www.iaeme.com/jom/issues.asp?JType=JOM&VType=7&IType=2
Journal Impact Factor (2020): 6.1633 (Calculated by GISI) www.jifactor.com
ISSN Print: 2347-3940 and ISSN Online: 2347-3959
DOI: 10.34218/JOM.7.2.2020.003


# ESTIMATION OF LOCATION PARAMETER USING ADAPTIVE AND SOME OTHER METHODS

**Chikhla Jun Gogoi**
Department of Statistics, Dibrugarh University, Dibrugarh, Assam, India.

**Bipin Gogoi**
Professor, Department of Statistics, Dibrugarh University, Dibrugarh, Assam, India

**ABSTRACT**

*The purpose of this article to propose a comparative study in estimation methods for location parameter between adaptive and some other methods. Underlying distributions for this study are Normal, exponential and logistic. Adaptive estimation methods are found to give more accurate results for asymmetric distributions compared to symmetric distributions. Monte Carlo simulation method is used here to fulfill the objectives. Results are displayed here in tables.*

**Keywords:** Adaptive, Simulation, Gastwirth, Winsorized mean, trimmed mean, Hogg-Lehmann Estimator

**Cite this Article:** Chikhla Jun Gogoi and Bipin Gogoi, Estimation of Location Parameter Using Adaptive and Some Other Methods, *Journal of Management (JOM)*, 7 (2), 2020, pp. 14-18.
http://www.iaeme.com/JOM/issues.asp?JType=JOM&VType=7&IType=2

## I. INTRODUCTION

A generally found important problem in statistics is the estimation of location parameter for any population. For example, the sample mean $\bar{x}$ can be thought of as an estimate of the distribution mean $\mu$. But here arise some questions like how good is this estimate? What makes an estimate good? Can we say anything about the closeness of an estimate to an unknown parameter?

There exist various methods of estimation of location parameters. In history we found that many robust estimators of location parameters have been proposed over the last two decades. Significant work in the area has been done among others by Andrews et al. (1972), Huber (1964), Hogg (1972,1974), and Gastwirth(1966). Hogg (1974) developed adaptive robust procedures which use the tail length of a distribution as a basis for determining the proportion of observations to be trimmed from each end of an ordered sample. Wegman and Carroll (1977) made a comparative study using Monte Carlo simulation method. Frost and Ali (1981) also made a Monte Carlo studies of some adaptive robust procedure for location. Hogg and Lenth (1984) used a selection statistic to identify the situation and trimming is used here.

In this article a study is made to compare the adaptive procedure and some other procedures for estimation of location parameter for different symmetric and asymmetric distributions.

## 2. TEST PROCEDURES

Let us consider n order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ corresponding to a random sample $Y_1, Y_2, \ldots, Y_n$ from a distribution of continuous type possessing the density $f(y,\theta)$ which is symmetric about $\theta$. In accordance to history it is seen that many authors suggested different estimator to estimate the population location parameter. Some robust estimators which will be compared with adaptive WLS are given below.

Gastwirth(1966) formed a very simple estimator defined as

$$T = 0.3X_{[n/3+1]} + 0.4\,\text{median} + 0.3Y_{n-[n/3]}$$

The Winsorized mean is defined as

$$W = \frac{1}{n}\left\{(r_1+1)X_{r_1+1} + \sum_{i=r_1+2}^{n-r_2-1} X_i + (r_2+1)X_{n-r_2}\right\}$$

The $100\alpha\%$ trimmed mean is defined as

$$T(\alpha) = \frac{pX_{1+[n\alpha]} + \sum_{i-2+[n\alpha]}^{n-[n\alpha]-1} X_i + pX_{n-[n\alpha]}}{n(1-2\alpha)}$$

where $p = 1+[n\alpha] - n\alpha$, [t] refers to the greatest integer function.

The Hogg-Lehmann Estimator is given by

$$\text{H-L} = \text{median}\left\{\frac{X_i+X_j}{2}\right\}, \quad 1 \le i \le j \le$$

## 3. ADAPTIVE PROCEDURES

It is known that for the normal distribution , the .05 trimmed mean ( or simply mean) is the best estimator. Rothenberg, Fisher and Tilanus(1964) argue that the 3/8 trimmed mean is very robust in Cauchy case. However, the best amount of trimming is dependent on the distribution from which the sample arises. In order to obtain an adaptive robust procedure for trimming , a statistic devised by Hogg(1974) and based on tail length was considered. The tail length, which is used as a measure of a distribution's tail thickness, is defined as

$TL = \frac{\bar{U}(\beta) - \bar{L}(\beta)}{\bar{U}(\alpha) - \bar{L}(\alpha)}$ where $\bar{U}(t)$ *and* $\bar{L}(t)$ are the means of the upper and lower t.n order statistics, respectively. Hogg(1979) suggested that $\beta = 0.2$ and $\alpha = 0.5$ is a good choice for this purpose.Thus , the above tail length formula for n=30 can be written as

$$TL = \frac{\frac{1}{6}\sum_{25}^{30} Y_i - \frac{1}{6}\sum_{1}^{6} Y_i}{\frac{1}{15}\sum_{16}^{30} Y_i - \frac{1}{15}\sum_{1}^{15} y_i}$$

Hogg and Length(1984) had provided some distinct criterion for the selection of estimators, such as the mean is selected as the estimator , if the empirical distribution has characteristic of tail length TL which short with TL<1.81. Again under the condition of moderate tail for the empirical distribution with 1.81<TL<1.87 then the 10% trimmed mean is selected as the estimator. Similarly median is selected as the estimator if empirical distribution has tails that are so long with TL>1.87. This estimation procedure, avoids unnecessary trimming of extreme values. Also the estimator is reasonably robust, even to the presence of a high percentage of outliers.

## 4. ADAPTIVE WLS ESTIMATOR

Let $y_i$ denote the ith observation and $\tilde{y}$ denote the median of the observations {$y_1,...,y_n$). In order to find the appropriate weights, we begin by estimating the cdf of a symmetric distribution. Then we obtain the set $Y_D$={$y_1 - \tilde{y},...,y_n - \tilde{y}, -(y_1 - \tilde{y}), \dots, -(y_n - \tilde{y})$}. Next we compute a robust estimate of the standard deviation as $\hat{\sigma}$ =$IQR_D/1.349$, where $IQR_D$ is the inter quartile range of $Y_D$.

The cantered and standardised observations are defined as

$z_i = \frac{y_{i-\tilde{y}}}{\hat{\sigma}}$ for i=1,2,...,n. We then double these centred and standardised observations to obtain the set $Z_D$={$z_1,...,z_n,-z_1,...,-z_n$} and denote the ith element in $Z_D$ as $z_{D,i}$. At a point z we define the smoothed cdf as

$$\widehat{F_h}(z; Z_D) = \frac{1}{2n}\sum_{i=1}^{2n} \emptyset(\frac{z - z_{D,i}}{h})$$

Where $\emptyset(.)$ is the cdf of standard normal distribution and $h=1.26n^{-1/3}$ is the smoothing constant. In order to weight the observations, we need to compare $z_i$ to the corresponding normal score, which is given by $\emptyset^{-1}[\widehat{F_h}(z_i; Z_D)]$. The weights are determined by $w_i = \frac{\emptyset^{-1}[\widehat{F_h}(z_i; Z_D)]}{z_i}$ for i= 1,2,...,n. If the differences tend to normal distribution, then $\emptyset^{-1}[\widehat{F_h}(z_i; Z_D)]$ should approximate $z_i$ for all the observations, which will produce weights near one. If an outlier that produces an extremely large or small value of $z_i$ then it will be given a small weight.

After calculation of weights, they are applied to the observations to obtain the adaptive estimate. The WLS estimate of β in the model $y_i = \beta + \epsilon_i$ is $\hat{\beta} = \frac{\sum_{i=1}^{n} w_i^2 y_i}{\sum_{i=1}^{n} w_i^2}$ . Consequently $\hat{\beta}$ is our adaptive estimator of the mean of the distribution.

## 5. SIMULATION STUDY

A simulation study is performed in order to compare the adaptive WLS estimator to median, the 10% trimmed mean and the mean. For this study, we generate data from the Normal, Logistic, Exponential distribution. We used sample sizes of n = 10, 20,30 and for each distribution and sample size we generated 5000 data sets.

**Table 1** The Estimated Values of the Mean=1 for Various Estimators under Normal Distribution:

| **Sample Size** → | **n=10** | **n=20** | **n=30** |
|---|---|---|---|
| Mean estimator | 1.0019 | 1.0006 | 0.9999 |
| Median estimator | 1.0028 | 1.0012 | 1.0009 |
| 10% trimmed mean estimator | 0.8884 | 0.9106 | 0.9531 |
| Adaptive WLS estimator | 2.8723 | 2.5720 | 2.0098 |
| Gastwirth estimator | 1.5893 | 1.4590 | 1.3175 |
| Winsorized mean | 1.5412 | 1.3016 | 1.2891 |
| Hogg-Lehmann Estimator | 0.8881 | 0.9032 | 0.9511 |

**Table 2** The Estimated m.s.e's for Various Estimators of the Mean under Normal Distribution:

| **Sample size** → | **10** | **20** | **30** |
|---|---|---|---|
| Mean | 0.0000036 | 0.00000012 | 0.00000002 |
| Median | 0.000093 | 0.000087 | 0.0000078 |

| **Sample size →** | **10** | **20** | **30** |
|---|---|---|---|
| 10% Trim Mean | 0.001245 | 0.00107 | 0.000941 |
| Adaptive WLS | 0.35055 | 0.29981 | 0.13940 |
| Gastwirth Estimator | 0.13473 | 0.115239 | 0.102455 |
| Winsorized Mean Estimator | 0.108864 | 0.10233 | 0.09452 |
| Hogg-Lehmann Estimator | 0.00125 | 0.00109 | 0.0009 |

**Table 3** The Estimated Values of the Mean=1 for Various Estimators under Logistic Distribution:

| **Sample Size →** | n=10 | n=20 | n=30 |
|---|---|---|---|
| Mean estimator | 1.0009 | 1.0005 | 1.00009 |
| Median estimator | 2.00978 | 2.00891 | 2.00800 |
| 10% trimmed mean estimator | 1.00192 | 1.00282 | 1.00011 |
| Adaptive WLS estimator | 0.5548 | 0.66111 | 0.86840 |
| Gastwirth estimator | 1.50908 | 1.00827 | 1.00710 |
| Winsorized mean | 1.00022 | 1.0067 | 1.00902 |
| Hogg-Lehmann Estimator | 1.00018 | 1.0027 | 1.00012 |

**Table 4** The Estimated m.s.e.'s for the Different Estimators of the Mean of Logistic Distribution:

| **Sample size →** | **10** | **20** | **30** |
|---|---|---|---|
| Mean | 0.00907 | 0.00638 | 0.00285 |
| Median | 0.4618 | 0.3288 | 0.27933 |
| 10% Trim Mean | 0.00992 | 0.00841 | 0.00375 |
| Adaptive WLS | 0.6489 | 0.6082 | 0.5819 |
| Gastwirth Estimator | 0.28911 | 0.2487 | 0.1903 |
| Winsorized Mean Estimator | 0.00612 | 0.00448 | 0.00160 |
| Hogg-Lehmann Estimator | 0.00054 | 0.00011 | 0.00009 |

**Table 5** The Estimated Values of the Mean=1 for Various Estimators under Exponential Distribution:

| **Sample size →** | **n=10** | **n=20** | **n=30** |
|---|---|---|---|
| Mean estimator | 0.6590 | 0.7799 | 0.8810 |
| Median estimator | 1.3378 | 1.2891 | 1.1199 |
| 10% trimmed mean estimator | 0.9976 | 0.9991 | 0.9995 |
| Adaptive WLS estimator | 1.0001 | 0.9999 | 1.00002 |
| Gastwirth estimator | 0.7621 | 0.8769 | 0.9890 |
| Winsorized mean | 0.9749 | 0.9828 | 0.9920 |
| Hogg-Lehmann Estimator | 0.9855 | 0.9934 | 1.00023 |

**Table 6** The Estimated m.s.e.'s for the Different Estimators of the Mean of Exponential Distribution:

| **Sample size →** | **10** | **20** | **30** |
|---|---|---|---|
| Mean | 0.01161 | 0.01096 | 0.000931 |
| Median | 0.0128 | 0.01063 | 0.01008 |
| 10% Trim Mean | 5.781*10^(-5) | 3.63*10^(-5) | 1.46*10^(-5) |
| Adaptive WLS | 3.11*10^(-6) | 2.629*10^(-6) | 1.848*10^(-6) |

| **Sample size →** | **10** | **20** | **30** |
|---|---|---|---|
| Gastwirth Estimator | 0.005896 | 0.00317 | 0.00199 |
| Winsorized Mean Estimator | 6.89*10^(-4) | 4.90*10^(-4) | 3.94*10^(-4) |
| Hogg-Lehmann Estimator | 2.143*10^(-4) | 1.904*10^(-4) | 1.57*10^(-4) |

## 6. CONCLUSION

It is seen that in case of Normal distribution the estimated value for mean is very close to population mean=1. Hence accordingly the m.s.e. of mean is less than the other estimators. So if the parent distribution of the data set is normal then mean estimator is to be preferred than the other estimators. Here the Hogg-Lehmann and trimmed mean estimator also give a good performance although their m.s.e. slightly more than mean estimator..

The Logistic distribution has a moderate tail. So according to the criterion the m.s.e. of Winsorized mean, Hogg-Lehmann Estimator and 10% trimmed mean estimator for Logistic distribution is less than the other estimators. So, these estimators are preferable under Logistic distribution.

The Exponential distribution being a skewed distribution having long tail, the adaptive WLS estimator shows a better result followed by 10% trimmed mean estimator.

## REFERENCES


[1] Bean, R.(1974): Asymptotically Efficient Adaptive Rank Estimates in Location Models. *Annals of Statistics, Vol. 2, 63-74*

[2] Beran, R.(1978): An Efficient and Robust Adaptive Estimator of Location . A*nnals of Statistics, Vol.6, No 2,292-313.*

[3] Bickel,P.J.(1982): On Adaptive Estimation. *Annals of Statistics, Vol. 1, No 3, 647-671*

[4] Croux, C. and Dehon, C.(2014), Robust estimation of location and scale. *Wiley States Ref Statistic Reference online received from https://doi.org/10.1002/9781118445112.stat07416*

[5] D'Agostino, R. B. and Cureton, E. E. (1973): A Class of Simple Linear Estimators of the Standard Deviation of the Normal Distribution. *Journal of the American Statistical Association Vol. 68, 207–210.*

[6] Faraway, J.J.(1992): Smoothing in Adaptive Estimation. *The Annals of Statistics.1992 Vol.20, No 1,414-427*

[7] Hill, N. J. and Padmanabhan, A. R., (1991): Some Adaptive Robust Estimators Which Work with Real Data. *Biometrical Journal, Vol.*33, 81-101.

[8] Hogg, R.V.; Uthoff, V.A., Randles, R.H., and Devenport, A.S.(1972): 'On the Selection of the Underlying Distribution and Adaptive Estimation' , *Journal of the American Statistical Association.,Vol.67,597-600.*

[9] Hogg, R.(1967): Some Observations on Robust Estimation. *Journal of the American Statistical Association, Vol. 62, 1179-1186.*

[10] Jaeckel Louis A.(1971): Robust Estimates of Location: Symmetry and Asymmetry Contamination, *Annals of Mathematical Statistics., Vol.42,1020-34*