# MEAN ABSOLUTE DEVIATION ABOUT MEDIAN AS A TOOL OF EXPLANATORY DATA ANALYSIS

**Elsayed A. E. Habib**
Department of Statistics and Mathematics, Benha University, Egypt & Management and Marketing Department, College of Business, University of Bahrain, P.O. Box 32038, Kingdom of Bahrain

**ABSTRACT**
The mean absolute deviation about median (MAD median) is often regarded as a robust measure of the scale of a distribution. In this paper it is shown that the MAD median is a very rich statistic and contains a lot of information not only about the scale but also about the shape of a distribution. MAD median is shown graphically using standardized empirical distribution function (SEDF chart). From this chart two concepts of skewness are introduced. One of them in terms of fat tail while the second in terms of tail length. More over the MAD median is used to compare and study the relationship between two variables through MAD median correlation coefficients and SEDF chart.

**Keyword:** *Correlation;empirical distribution function; skewness; tail weight.*

## 1 INTRODUCTION

Mean absolute deviation about median (MAD median) offers a direct measure of the scale of a random variable about its median and has many applications in different fields; see, for example, [1], [13] and [11].The MAD median is actually more efficient measure of scale than the standard deviationin life-like situations where small errors will occur in observation and measurement; see, [15], [9]. Exploratory data analysis(EDA; [16]; [3]; [8]) is a set of graphical techniques for finding interesting patterns in data or determining characteristics of a data. Most of these techniques work in part by hiding certain aspects of the data while making other aspects more clear. EDA was developed in the late 1970s when computer graphics first became widely available. It emphasizes robust and nonparametric methods, which make fewer assumptions about the shapes of curves and the distributions; see, [2].
In this paper it is shown that the MAD median contains a lot of information not only about the scale but also about the shape of a distribution and can be used as a tool of explanatory data analysis. MAD median is shown graphically using standardized empirical distribution function (SEDF chart) that by means of which the different parts of a distribution may be compared. From this chart two concepts of skewness are introduced. One of them in terms of fat tail (standardized heights below and above the median) while the second in terms of tail length. Moreover, the MAD median is used to compare and study the relationship between two variables through MAD median correlation coefficients and SEDF chart.
In Section 2 MAD median is expressed in terms of a covariance between the random variable and the indicator function for values above the median. In Section 3 new uses of MAD median are derived, MAD median correlation, MAD median skewness, SEDF chart and measures of skewness in terms of fat tail and tail length. Applications to univariate and bivariate data are given in Section 4. Section 5 is devoted to the conclusion.

## 2 MAD MEDIAN AS A COVARIANCE

Let $X_1, X_2, \dots, X_n$be a random sample from a distribution with, probability function $p(x)$, density function$f(x)$, quantile function $x(F) = F^{-1}(x) = Q(F)$where $0 < F < 1$, cumulative distribution function $F(x) = F_X = F$. Let$E(X) = \mu$,$Med(X) = v$. The MAD median is defined as

$$D_X(v) = E|X - v| \tag{1}$$

Habib [7]used the general dispersion function that defined by [12] as

$$D_X(a) = E|X - a| = a[2F_x(a) - 1] + E(X) - 2\int XI_{X<a}dP \tag{2}$$

To redefined MAD median as

$$D_X(v) = E(X) - 2E(XI_u) = E[(1 - 2I_u)X] \tag{3}$$

The indicator function for the values under the median is

$$I_u = \begin{cases} 1, & X < v \\ 0, & X \geq v \end{cases}$$

With $E(I_u) = 1/2$ and $Var(I_u) = 1/4$.
The above expression can be rewritten in terms of covariance as

$$D_X(v) = 2E(I)E(X) - 2E(XI) = -2Cov(X, I_u) \tag{4}$$

This is minus twice the covariance between the random variable $X$ and the indicator function for the values under the median$I_u$. Since $2Cov(X, I_o) = -2Cov(X, I_u)$, then

$$D_X(v) = 2Cov(X, I_o) \tag{5}$$

This is twice the covariance between the random variable $X$ and the indicator function for the values above the median$I_0$ and

$$I_o = \begin{cases} 1, & X > v \\ 0, & X \leq v \end{cases}$$

Moreover it can express$D_X(v)$as a weighted average as

$$D_X(v) = 2Cov(X, I_o) = 2E(XI_o) - 2E(X)E(I_o) = E[(2I_o - 1)X] \tag{6}$$

Note that the weight $W = 2I_0 - 1$with $E(W) = 0$.

### 2.1 Sample MAD median

Consider a random sample $X_1, X_2, \dots, X_n$ of size $n$ from a population with density function$f(x)$, and cumulative distribution function $F(x)$where its corresponding order statistics is $X_{1:n}, \dots, X_{n:n}$. The sample estimate using the absolute value is

$$d_{\tilde{x}} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \tilde{x}| \tag{7}$$

This is the most used method for estimating $D_X(v)$ and $\tilde{x}$ is the sample median. Let the indicator function for values over median from a sample is

$$\hat{I}_o = \begin{cases} 1, & x > \tilde{x} \\ 0, & x \leq \tilde{x} \end{cases}$$

The covariance estimate is

$$d_{\tilde{x}} = 2\widehat{Cov}\left(x, \hat{I}_o\right) = \frac{2}{n}\sum_{i=1}^{n} x_i\, \hat{I}_o - \bar{x} \tag{8}$$

The weighted average estimator can be obtained as

$$d_{\tilde{x}} = \frac{1}{n}\sum_{i=1}^{n} \left(2\hat{I}_o - 1\right) x_i = \sum_{i=1}^{n} w_i(I) x_i \tag{9}$$

The weights are

$$w_i(I, n) = \frac{\left(2\hat{I}_o - 1\right)}{n} \tag{10}$$

## 3 NEW USES OF MAD MEDIAN

### 3.1 MAD median correlation

Let $(X_i, Y_i), i = 1, \dots, n$denote independent and identically distributed data pairs drawn from a bivariate population with joint distribution function $F_{X,Y}(x, y)$. Rearrange the data pairs in ascending order with respect to the magnitudes of $X$, the new sequence of data is $\left(X_{(i)}, Y_{[i]}\right)$ where $X_{(1)} < \cdots < X_{(n)}$ are termed the order statistics of $X$ and $Y_{[1]}, \dots, Y_{[n]}$ the associated concomitants. Reversing the role of $X$ and $Y$ the data pairs are $\left(X_{[i]}, Y_{(i)}\right)$ where $X_{[1]}, \dots, X_{[n]}$ are termed concomitants of $X$ and $Y_{(1)} <, \dots, < Y_{(n)}$ the order statistics of $Y$; see, [4].

Habib [7] defined MAD median correlation coefficients using the covariance representation as

$$\Omega(X, Y) = \frac{Cov(X, I_o(Y))}{Cov(X, I_o(X))} \tag{11}$$

This is the ratio of the covariance between $X$ and the indicator function of $Y$ and $X$ and its indicator function. Also

$$\Omega(Y, X) = \frac{Cov(Y, I_o(X))}{Cov(Y, I_o(Y))} \tag{12}$$

This is the ratio of the covariance between $Y$ and the indicator function of $X$ and $Y$ and its indicator function. In general, these two measures are asymmetrical and they are necessarily not equal. Two symmetric measures of MAD median correlation are suggested as

$$\Omega_1 = \frac{\Omega(X, Y) + \Omega(Y, X)}{2} \tag{13}$$

This is the center of two asymmetric measures and

$$\Omega_2 = \frac{D_X\Omega(X, Y) + D_Y\Omega(Y, X)}{D_X + D_Y} \tag{14}$$

This is a weighted average of the asymmetric measures.

Using the data pairs $(x_i, y_i)$ the estimation of the MAD median correlation is

$$\widehat{\Omega}(X,Y) = \frac{\sum_{i=1}^{n}\left(2\hat{I}_o - 1\right)x_{[i]}}{\sum_{i=1}^{n}\left(2\hat{I}_o - 1\right)x_{i:n}} \tag{15}$$

and

$$\widehat{\Omega}(Y,X) = \frac{\sum_{i=1}^{n}\left(2\hat{I}_o - 1\right)y_{[i]}}{\sum_{i=1}^{n}\left(2\hat{I}_o - 1\right)y_{i:n}} \tag{16}$$

For properties of these measures and comparison with Pearson's correlation coefficient, see, [7].

### 3.2 MAD median skeweness

From Munoz-Perez and Sanchez-Gomez [12]for $X \in L^1$, let

$$D_X(a) = E|X - a| \tag{17}$$

be a function of $a \in R$, called dispersion function of $X$ and

$$D_X^+(a) = E(X-a)^+ = E[\max(X-a, 0)] \tag{18}$$

and

$$D_X^-(a) = E(X-a)^- = E[\max(a-X, 0)] \tag{19}$$

be the right and left dispersion functions, respectively. Therefore,

$$E(X-a) = D_X^+(a) - D_X^-(a) \tag{20}$$

By replacing $a = v$ it can establish a very simple relation among mean, median, left and right $D_X(v)$ functions as

$$E(X) - v = D_X^+(v) - D_X^-(v) \tag{21}$$

Therefore the difference between the mean and the median is equal to the difference between the right and left dispersion functions in terms of median. In other words, the difference between the mean and the median is represented by the expected value of the difference in heights between the values above and below the median.

A measure of skewness that depends completely on $D_X(v)$ is

$$SK_D = \frac{D_X^+(v) - D_X^-(v)}{D_X(v)} \tag{22}$$

This measure is equivalent to Groeneveld and Meeden [5] measure of skewness. One advantage of this measure it can be shown graphically in terms of standardized left and right dispersion functions as following.

### 3.3 Standardized empirical distribution function (SEDF) chart

It can show$SK_D$ graphically by plotting $F$ against $H_{i:N} = \frac{E(X_{i:N})-v}{D_X(v)}$or $\frac{Med\,(X_{i:N})-v}{D_X(v)}$for population and $h_{i:n} = \frac{x_{i:n}-\tilde{x}}{d_{\tilde{x}}}$for sample where the heights above x-axis represent standardized right dispersion function $(h_+)$,the heights below x-axis represent standardized left dispersion function $(h_-)$ and the points on x-axis represent median value. Therefore, SEDF is a chart by means of which the different parts of a distribution may be compared. Moreover SEDF chart reflects information about the shape of a distribution (skewness, fat tail and tail length).

Without loss of generality for standard normal distribution and standard exponential distribution Figure 1shows SEDF chart using expected values and$N = 25$.

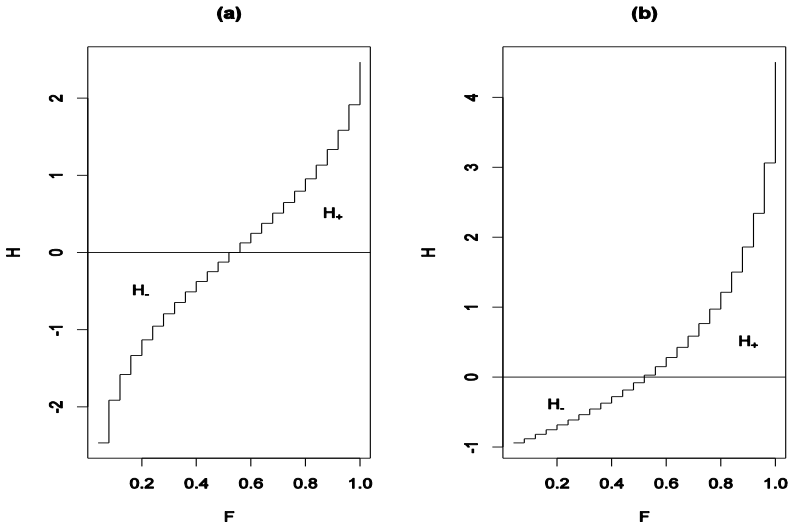


**Figure 1** SEDF graph (a) normal distribution $v = 0$and $D_X(v) = 0.798$ and (b) exponential distribution $v = D_X(v) = 0.693$for $N = 25$

It is clear that the SEDF chart reflects a lot of information about the shape of the distribution. Figure 1 (a) shows that the distribution

a. Is symmetric about its median.
b. Has equal tail lengths about 2.5 (symmetric in tail length).
c. Has equal fat tail (symmetric in fat tail).

Figure 1 (b) shows that the distribution is

a. Asymmetric about its median.
b. Long right tail (asymmetric in tail length about 4.5 times the left tail).
c. Much fat right tail (asymmetric in fat tail).

From SEDF chart it could suggest two concepts of skewness. One of them in terms of fat tail while the second in terms of tail length.

### 3.4 Measures of fat tail

We define measures of left and right fat tail that are applied to the half of probability mass lying to the left, respectively the right, side of the median of the distribution. MAD median may reflect information about the fat tail of a distribution by taking the expected values of the standardized heights below and right the median. A suggested measures of left ($L_{FT}$) and right ($R_{FT}$) fat tails are

$$L_{FT} = -E(H^-) = \frac{D_X^-(v)}{D_X(v)} \tag{23}$$

and

$$R_{FT} = E(H^+) = \frac{D_X^+(v)}{D_X(v)} \tag{24}$$

where

$$E(H_-) = E[\min(H, 0)] \tag{25}$$

and

$$E(H_+) = E[\max(H, 0)] \tag{26}$$

Properties of $L_{FT}$ and $R_{FT}$

1. $L_{FT}$ and $R_{FT}$ are location and scale invariant, i.e.
$$FT(aX + b) = TF(X)$$
for any $a > 0$ and $b \in R$.
2. The ranges are
$$0 \leq L_{FT}, R_{FT} \leq 1$$
3. In terms of fat tail, $L_{FT} = R_{FT} = 0.5$ for symmetric distribution, $R_{FT} > L_{FT}$ for right fat tail distribution, and $R_{FT} < L_{FT}$ for left fat tail distribution.
4. If a distribution is inverted, then
$$L_{FT}(-X) = R_{FT}(X)$$
5. For any distribution $L_{FT} + R_{FT} = 1$

These two measures could be estimated as

$$l_{FT} = \frac{-1}{n}\sum h_- \text{ and } r_{FT} = \frac{1}{n}\sum h_+ \tag{27}$$

where

$$h_- = min\left(0, \frac{x_{i:n} - \tilde{x}}{d_{\tilde{x}}}\right) \text{ and } h_+ = max\left(0, \frac{x_{i:n} - \tilde{x}}{d_{\tilde{x}}}\right) \tag{28}$$

Note that the measure of skewness in terms of fat tail is

$$SK_D = \frac{D_X^+(v) - D_X^-(v)}{D_X(v)} \tag{29}$$

This measure is equivalent to [5] and [6] measure of skewness $(\mu - v)/D_X(v)$.

### 3.5 Measures of tail length

MAD median could reflect information about the length of the tails of a distribution by finding the maximum and minimum values for heights above and under the median. The proposed measures of a left ($L_{TL}$) and a right ($R_{TL}$) tail lengths are

$$L_{TL} = -Min(H_{i:N}) \tag{30}$$

and

$$R_{TL} = Max(H_{i:N}) \tag{31}$$

Moreover a proposed measure of skewness in terms of tail lengths ($S_{TL}$)is

$$SK_{TL} = \frac{R_{TL} - L_{TL}}{R_{TL} + L_{TL}} = \frac{Max(H) + Min(H)}{Max(H) - Min(H)} \tag{32}$$

Properties of these measures are

1. $R_{TL}$,$L_{TL}$and $S_{TL}$ are location and scale invariant, i.e.
$$TL(aX + b) = TL(X)$$
for any $a > 0$and $b \in R$.
2. The range of $S_{TL}$ is
$$-1 \leq S_{TL} \leq 1$$
3. In terms of tail length $S_{TL} = 0$for symmetric distribution, $S_{TL} > 0$ for right tail length distribution, and $S_{TL} < 0$ for left tail length distribution.
4. If a distribution is inverted, then
$L_{TL}(-X) = R_{TL}(X)$ and $S_{TL}(-X) = -S_{TL}(X)$

Table 1 shows values of these measures from normal and exponential distributions for different values of $n$.

**Table 1** values of $L_{TL}$, $R_{TL}$ and $S_{TL}$ from normal (Nor) and exponential (Exp) distributions

| | $n$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 10 | 15 | 20 | 25 | 50 | 100 | 250 | 500 | 1000 |
| | | | | | Nor. | | | | | |
| $L_{TL}$ | 1.755 | 2.087 | 2.305 | 2.440 | 2.552 | 2.871 | 3.171 | 3.533 | 3.782 | 4.010 |
| $R_{TL}$ | 1.755 | 2.087 | 2.305 | 2.440 | 2.552 | 2.871 | 3.171 | 3.533 | 3.782 | 4.010 |
| $SK_{TL}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | | | | Exp. | | | | | |
| $R_{TL}$ | 2.571 | 3.381 | 3.936 | 4.305 | 4.613 | 5.556 | 6.523 | 7.821 | 8.811 | 9.805 |
| $L_{TL}$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $SK_{TL}$ | 0.44 | 0.543 | 0.595 | 0.623 | 0.644 | 0.695 | 0.734 | 0.773 | 0.796 | 0.815 |

Using the sample data $R_{TL}$, $L_{TL}$ and $SK_{TL}$could be estimated as

$$r_{TL} = max(h_{i:n}), l_{TL} = -\min(h_{i:n}) \tag{33}$$

and

$$sk_{TL} = \frac{max(h) + min(h)}{max(h) - min(h)} \tag{34}$$

where

$$h_{i:n} = \frac{x_{i:n} - \tilde{x}}{d_{\tilde{x}}} \tag{35}$$

## 4 APPLICATION

Consider consulting firm surveyed random samples of residents in two places. The firm is investigating the labor markets in these two communities for a client that is thinking of relocating its corporate offices. Educational level of the workforce in the two places is a key factor in making the relocation decision. The consulting firm surveyed 25 adults in each place and recorded the number of years of college attended in Tables 2 and 3.

**Table 2** number of years of college attended in place 1

| $x$ | 0 | 1 | 2 | 3 | |
|---|---|---|---|---|---|
| $f$ | 11 | 9 | 3 | 2 | |
| | | | proposed measures | | |
| $l_{TL}$ | $r_{TL}$ | $sk_{TL}$ | $l_{FT}$ | $r_{FT}$ | $sk_D$ |
| 1.389 | 2.778 | 0.333 | 0.611 | 0.389 | -0.222 |

and

**Table 3** number of years of college attended in place 2

| $x$ | 0 | 1 | 2 | | |
|---|---|---|---|---|---|
| $f$ | 12 | 10 | 3 | | |
| | | | proposed measure | | |
| $l_{TL}$ | $r_{TL}$ | $sk_{TL}$ | $l_{FT}$ | $r_{FT}$ | $sk_D$ |
| 1.667 | 1.667 | 0 | 0.8 | 0.2 | -0.6 |

Figure 2 shows the histogram and SEDF for data in Tables 2 and 3. Also the proposed measures are shown in Tables 2 and 3.For these data the histogram is showing strong right skewed while the measure $sk_d$ is a negative (-0.222 and

-0.6) that gives wrong conclusion if the histogram is used. If we look at SEDF we find the explanation directly. In terms of fat tail Figures 2 (b) and (c) show left fat tail, therefore $sk_D$ must be negative.By using numerical values, Tables 2 and 3 show that $l_{FT} = 0.611$and $r_{FT} = 0.389$ for place 1 and $l_{FT} = 0.8$ and $r_{FT} = 0.2$ for place 2. While in terms of tail length the SEDF chart shows that right tail length for place 1 data and equal tail length for place 2($sk_{TL} = 0.333$ and $sk_{TL} = 0$ respectively). By using the numerical values Tables 2 and 3 show that $l_{TL} = 1.389$and $r_{TL} = 2.778$ for place 1 while$l_{TL} = r_{tL} = 1.667$ and $sk_{TL} = 0$ for place 2.

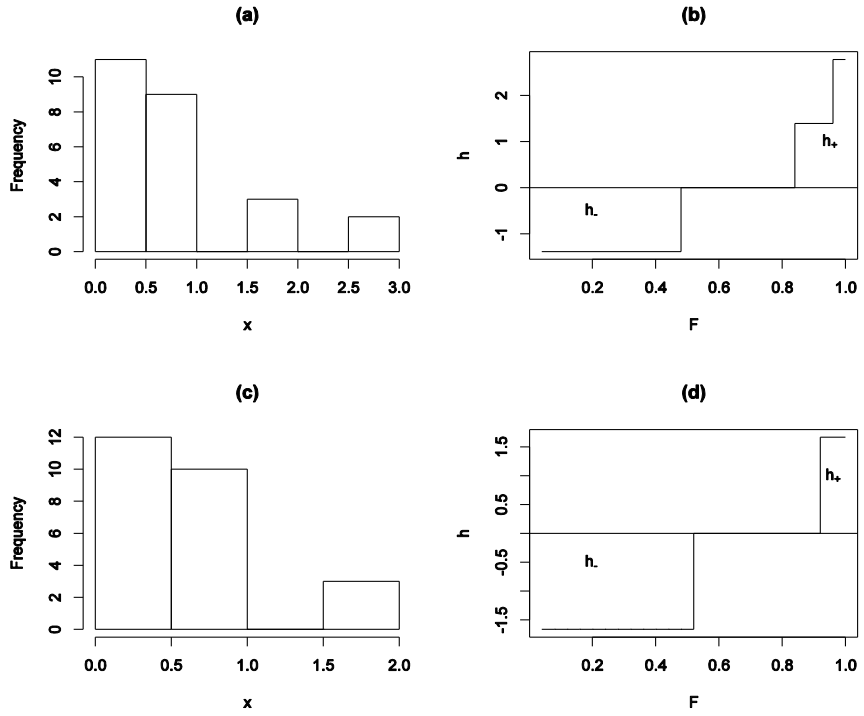


**Figure 2** Histogram and SEDF charts for the number of years of college attended.

### 4.1 Comparing two samples

Kacprzak and Chvojka [10] compared two methods of measuring mercury levels in fish. A new method, which they called "selective reduction", was compared to an established method, referred to as "the permanganate method". One advantage of selective reduction is that it allows simultaneous measurement of both inorganic mercury and methyl mercury. The mercury in each of 25 juvenile black marlin was measured by both techniques. The 25 measurements for each method are given in Table 4. Rice [14] used the paired $t$ test and sign rank test for looking for difference between the two methods. Both tests were insignificant using $\alpha = 0.05$. Table 4 gives the values of proposed measures. From this table it could see a difference in the right tail length.

**Table 4** measuring mercury levels using selective reduction and the permanganate methods.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Selective Reduction (S.R.) | | | | | | | | | | | | |
| 0.32 | 0.40 | 0.11 | 0.47 | 0.32 | 0.35 | 0.32 | 0.63 | 0.50 | 0.60 | 0.38 | 0.46 | 0.20 |
| 0.31 | 0.62 | 0.52 | 0.77 | 0.23 | 0.30 | 0.70 | 0.41 | 0.53 | 0.19 | 0.31 | 0.48 | |
| proposed measures | | | | | | | | | | | | |
| $l_{TL}$ | | $r_{TL}$ | | $sk_{TL}$ | | $l_{FT}$ | | $r_{FT}$ | | $sk_{FT}$ | | |
| 2.164 | | **2.761** | | **0.121** | | 0.435 | | **0.564** | | **0.128** | | |
| Permanganate (Per.) | | | | | | | | | | | | |
| 0.39 | 0.47 | 0.11 | 0.43 | 0.42 | 0.30 | 0.43 | 0.98 | 0.86 | 0.79 | 0.33 | 0.45 | 0.22 |
| 0.30 | 0.60 | 0.53 | 0.85 | 0.21 | 0.33 | 0.57 | 0.43 | 0.49 | 0.20 | 0.35 | 0.40 | |
| proposed measures | | | | | | | | | | | | |
| $l_{TL}$ | | $r_{TL}$ | | $sk_{TL}$ | | $l_{FT}$ | | $r_{FT}$ | | $sk_{FT}$ | | |
| 2.056 | | **3.535** | | **0.264** | | 0.411 | | **0.588** | | **0.177** | | |
| MAD med Correlation:$\widehat{\Omega}_1 = 0.928$, and Pearson's correlation $r = 0.854$ | | | | | | | | | | | | |

Figure 3 shows comparison between the two methods.It is clear that the difference in left tail is not so obvious while there may be a difference for upper tail of the distribution.

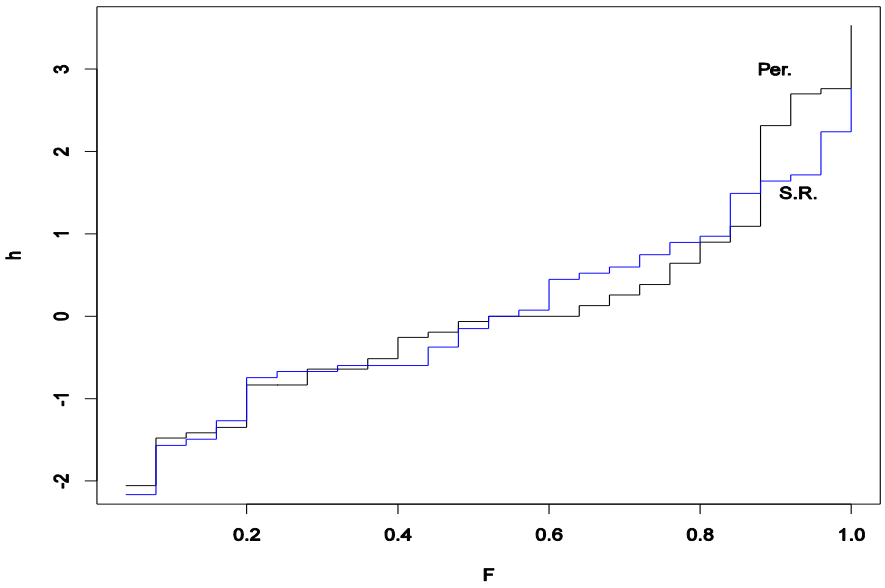


**Figure 3** SEDF chart for mercury levels using selective reduction (S.R.) and the permanganate (Per.) methods.

## 5 CONCLUSION

MAD med is used as a tool of explanatory data analysis where it contained several illuminating insights about the shape of a distribution and the correlation between two variables.Graphically MAD med is represented on SEDF chart that showed a lot of information about skewness, fat tail and tail length. The main advantage of this chart was that the different parts of a distribution may be compared. Moreover it was shown that the MDM med had been a powerful tool for comparing two samples and finding the correlation between them.

It may add line to SEDF chart at a specified value for some known distributions to detect outliers. As an example from Table 1 it may add lines at $\mp 2.75$for $n \leq 25$, $\mp 3.5$ for $25 < n \leq 100$and $\mp 4.25$ for $n > 100$for normal distribution to detect outliers. Moreover, it might recommend that the empirical and standard empirical distribution functions are shown together in one graph where the first chart shows minimum, maximum, median and percentiles of the data, the latter shows the shapes of the data.

## REFERENCES

[1]. Babu, C.J. and Roa, C.R., (1992) Expansions for statistics involving the mean absolute deviation. *Annals of the Institute of Statistical Mathematics*, **44**, 387-403.
[2]. Brys, G., Hubert, M. and Struyf, A. (2006) Robust measures of tail weight. *Computational Statistics & Data Analysis*, **50**, 733-759.
[3]. Cleveland, W. (1993)*Visualizing Data*.1st Ed., Hobart Press.
[4]. David, H. and Nagaraja, H., (2003)*Order Statistics.* 3rd Ed., Hoboken, NJ: Wiley-Interscience.
[5]. Groeneveld,R.A. and Meeden, G. (1984) Measuring skewness and kurtosis. *The Statistician*, **33**, 391-399.
[6]. Groeneveld and Meeden (2009) An improved skewness measure. users.stat.umn.edu/~gmeeden/papers/skew.pdf
[7]. Habib, E.A.E (2011) Correlation coefficients based on mean absolute deviation about median. *International Journal of Statistics and Systems*, **6**, 389-404.
[8]. Hoaglin, D.C., Mosteller, F. and Tukey, J.W. (2000) *Understanding Robust and Exploratory Data Analysis*. 1st Ed., John Wiley and Sons, New York.
[9]. HUBER, P., (1981) *Robust Statistics*.New York, John Wiley and Sons.
[10]. Kacprzak, J. AND Chvojka, R. (1976) Determination of methyl mercury in fish by flameless atomic absorption spectroscopy and comparison with an acid digestion method for total mercury. *Journal ofthe Association of Analytical Chemists*, **59**, 153-157.
[11]. Marona, R.A., Martin, D.R. and Yohai, V.J., (2006)*Robuststatistics: Theory and methods*.John Wiley & sons.
[12]. Munoz-Perez J. and Sanchez-Gomez(1990)A characterization of the distribution function: The dispersion function. *Statistics and Probability Letters*, **10**, 235-239.
[13]. Pham-Gia, T. and Hung, T.L.(2001)The mean and median absolute deviations.*Mathematical and Computer Modelling*, **34**, 921-936.
[14]. Rice, J.A. (2007) *Mathematical Staistics and Data Analysis*. 3rd Ed., Duxbury Press.
[15]. TUKEY, J. W. (1960) A survey of sampling from contaminated distributions. In *Contributions to Probability and Statistics*: *Essays in Honor of Harold Hotelling*(I. Olkin, S. G. Ghurye,W. Hoeffding, W. G. Madow and H. B. Mann, eds.) 448–485. Stanford Univ. Press.
[16]. Tukey, J. W. (1977) *Exploratory Data Analysis*. 1st Ed., Addison-Wesley, Reading, MA.