

# Minima-Maxima Plot

Elsayed A. H. Elamir<sup>1,2</sup>

<sup>1</sup>Department of Statistics and Mathematics, Benha University, Egypt

<sup>2</sup>Egypt & Management & Marketing Department, College of Business, University of Bahrain, Kingdom of Bahrain

---

**Abstract** A new graphical technique is proposed as a tool for studying the distributional behaviour of a set of observations since it is more intuitive and easily interpretable. This graph is based on minimum and maximum order statistics of the distribution. It has advantage of being stable at the extreme tails of the distribution unlike quantile quantile plot. Moreover, Minima-Maxima plot captures all information not only about the tails of the distribution but also about the whole distribution of the data and characterizes the distribution of the data completely. Therefore, the idea behind this plot is focusing on valuable information at the extreme tails while few numbers can describe the middle of the distribution. Accurate estimates of the location, scale and shape parameters may be obtained from the graph.

**Keywords** Extreme values, Histogram, Order statistics, Pareto distribution, Q-Q plot

---

## 1. Introduction

Graphical presentation of data is a vital tool in sciences. Good graph conveys a great deal of information and can be used to extract new conclusions while bad graph can be misleading and confusing. Given a random sample of univariate data points, a pertinent question is whether this sample comes from some specified distribution  $F$ . Decision techniques are based on how close the empirical distribution of the sample and the distribution  $F$  are for some sample size  $n$ .

A histogram is a graphical representation of the distribution of numerical data and was first introduced by Pearson (1895). The most common form of the histogram is obtained by splitting the range of the data into equal-sized bins. Then for each bin, the numbers of points from the data set that fall into each bin are counted. The classes can either be defined arbitrarily by the user or via some systematic rules. A number of theoretically derived rules have been proposed by Scott (1992). The purpose of a histogram is to graphically summarize the distribution of a univariate data set. The histogram graphically shows the location, spread, skewness and presence of multiple modes in the data.

Quantile-quantile (Q-Q) plot is commonly used device to graphically and informally test the goodness-of-fit of a sample in an exploratory way. It is used to plot the sample quantiles against the theoretical quantiles or other sample quantiles and then a visual check is made to see whether or not the points are close to a straight line; see, Chambers et al (1983), Cleveland (1994) and Cleveland and McGill (1988).

The pattern of points in the plot is used to compare the shapes of distributions, providing a graphical view of how properties such as location, scale and skewness. The use of Q-Q plots to compare two samples of data can be viewed as a non parametric approach to comparing their underlying distributions. A Q-Q plot is generally a more powerful approach to do this than the common technique of comparing histogram of the two samples, but requires more skill to interpret; see, Makkonen (2008), Wilk and Gnanadesikan (1968), Wicklin (2011).

The minima-maxima plot (M-M plot) is proposed based on the minimum and maximum of order statistics that can be theoretically computed from a specific distribution and can be estimated from a sample data. The M-M plot depends on giving more weights to the data at the extreme tails than the data at the centre of the distribution; therefore, its philosophy is “most important information for a distribution at the tails”. This plot captures all information not only about the tails of the distribution but also about the whole distribution of the data. The pattern of points in the M-M plot is used to compare the shapes of distributions and providing a graphical view of how properties such as location, scale, skewness and kurtosis. Like Q-Q plot, M-M plot is used to plot the data against theoretical extreme order statistics or sample extreme order statistics and then a visual check is made to see whether or not the points are close to a straight line but the M-M plot has more stability at the tails of the distribution than Q-Q plot. Note that all graphs and programming are done by R-software and the program is given in Appendix A.

The minimum and maximum order statistics and their characteristics to probability distributions are presented in Section 2. The M-M plots are proposed in Section 3 for one variable and two variables. An application to blood plasma levels one hour following chocolate consumption data is

---

\* Corresponding author:

shabib40@gmail.com (Elsayed A. H. Elamir)

Published online at <http://journal.sapub.org/statistics>

Copyright © 2016 Scientific & Academic Publishing. All Rights Reserved

studied in Section 4. Section 5 is devoted for conclusion.

## 2. Minimum and Maximum Order Statistics

Let  $X_1, \dots, X_n$  be a sample from a distribution function  $F$ , probability function  $f(x)$  and quantile function  $x(F)$ . When the  $X_i$ 's are arranged in ascending order of magnitude and then written as

$$X_{1:n} \leq \dots \leq X_{n:n}$$

$X_{r:n}$  is the  $r$ th order statistic. Since the event  $(X_{r:n} \leq x)$  occurs if and only if at least  $r$  of the  $X_i$ 's are less than or equal to  $x$ ,  $F_{r:n}$  is expressible in terms of  $F$  as the binomial tail probability

$$F_{r:n}(x) = Pr(X_{r:n} \leq x) = \sum_{j=r}^n \binom{n}{j} F^j(x) [1 - F(x)]^{n-j}$$

The expected value of order statistics is

$$E(X_{r:n}) = r \binom{n}{r} \int_0^1 x(F) F^{r-1} (1 - F)^{n-r} dF, \quad r \leq n$$

This can be re-written as

$$E(X_{r:n}) = r \binom{n}{r} E[x(F) F^{r-1} (1 - F)^{n-r}]$$

see; David (1981).

Let

$$M_{n:n} = \max\{X_1, \dots, X_n\}$$

Denote the maximum of the first  $n$  random variables. Its distribution function is given by

$$F_{n:n} = [F(x)]^n, \quad -\infty < x < \infty$$

As pointed out by Arnold et al. (2008) and Alpay (2016), clearly knowledge of the distribution of  $X_{n:n}$  determines  $F(x)$  completely. This is true since

$$F(x) = [F_{n:n}]^{1/n}, \quad -\infty < x < \infty$$

Moreover, Chan (1967) has shown that if  $E|X| < \infty$  then  $F(x)$  is uniquely determined by the sequence

$$\{E(X_{n:n}): n = 1, 2, 3, \dots\}$$

Let

$$M_{1:n} = \min\{X_1, \dots, X_n\}$$

Denote the minimum of the first  $n$  random variables. The distribution function is given by

$$F_{1:n} = 1 - [1 - F(x)]^n, \quad -\infty < x < \infty$$

Clearly knowledge of the distribution of  $X_{1:n}$  determines  $F$  completely. This is true since

$$F(x) = 1 - [1 - F_{1:n}]^{1/n}, \quad -\infty < x < \infty$$

Also, Chan (1967) has shown that if  $E|X| < \infty$  then  $F$  is uniquely determined by the sequence

$$\{E(X_{1:n}): n = 1, 2, 3, \dots\}$$

For example,

$$E(X_{1:n}) = 1/n, \quad n = 1, 2, 3, \dots$$

if and only if  $F$  is unit exponential

$$(F(x) = 1 - \exp(-x), \quad x > 0),$$

$$E(X_{n:n}) = 2n/(2n + 1), \quad n = 1, 2, 3, \dots$$

if and only if  $F$  is triangular ( $F(x) = x^2$ ,  $0 < x < 1$ ) and

$$E(X_{1:n}) = 1/(2^n - 1), \quad n = 1, 2, 3, \dots$$

if and only if  $F$  is geometric ( $P(X = x) = 2^{-x-1}$ ,  $x = 0, 1, 2, \dots$ ); see, for example, Huang (1989).

## 3. Minima-Maxima Plot

M-M plot is proposed through two functions or curves plotted on the same graph as following.

### Single variable

For a given data of size  $n$ ,  $x_1, x_2, \dots, x_n$ , the theoretical minima curve based on the expected value of order statistics is defined as

$$E(X_{1:k}) = kE[x(F)(1 - F)^{k-1}], \quad k = 1, 2, \dots, n$$

From Downton (1966) and Elamir and Seheult (2003) this can be estimated as

$$\hat{E}(X_{1:k}) = \frac{1}{\binom{n}{k}} \sum_{i=1}^k \binom{n-i}{k-1} x_{i:n}, \quad k = 1, 2, \dots, n$$

The theoretical maxima curve based on the expected value of order statistics is defined as

$$E(X_{k:k}) = kE[x(F)F^{k-1}], \quad k = 1, 2, \dots, n$$

From Downton (1966) and Elamir and Seheult (2004) this can be estimated as

$$\hat{E}(X_{k:k}) = \frac{1}{\binom{n}{k}} \sum_{i=1}^k \binom{i-1}{k-1} x_{i:n}, \quad k = 1, 2, \dots, n$$

The M-M plot consists of two curves, the minima curve is plotted as

$$\text{Minima curve} := (x_{i:n} \text{ versus } \hat{E}(X_{1:k})), \quad i = 1, \dots, n, \\ k = 1, \dots, n$$

This curve starts from the average  $\bar{x} = \hat{E}(X_{1:1})$  to the minimum value  $x_{1:n} = \hat{E}(X_{1:n})$ .

Also the maxima curve is plotted as

$$\text{Maxima curve} := (x_{i:n} \text{ versus } \hat{E}(X_{k:k})), \\ i = 1, \dots, n, k = 1, \dots, n$$

This curve starts from the average  $\bar{x} = \hat{E}(X_{1:1})$  to the maximum value  $x_{n:n} = \hat{E}(X_{n:n})$ .

Both curves should tell us the whole picture about the distribution function of random variable  $X$  for a given data. Also each curve in its own should reflect all the information about the whole distribution for the random variable  $X$  for a given data.

Figure 1 is created for standard normal distribution at  $n = 30$  based on the quantile function using plotting position  $(i - 0.5)/n$  on the x-axis and M-M values on y-axis using exact extreme order statistics from package *EnvStats* in R software as

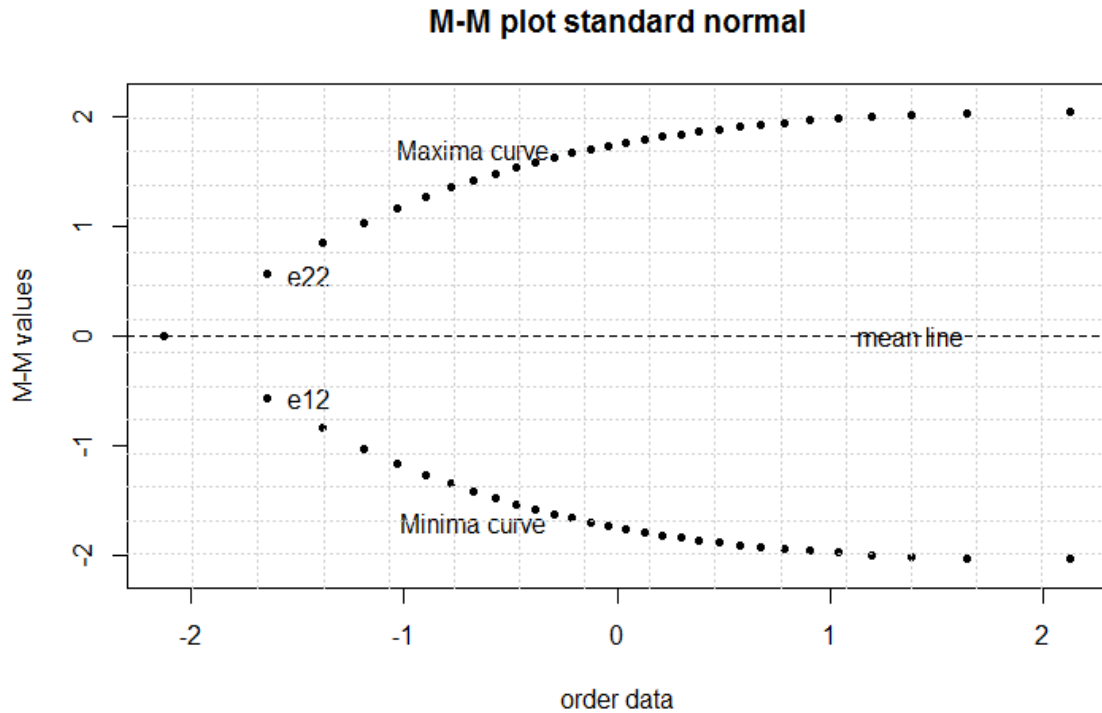


Figure 1. M-M plot for standard normal distribution at  $n = 30$

Minima curve

$$:= \left( \text{qnorm}\left(\frac{(i-0.5)}{30}\right) \text{ versus } \text{evNormOrdStatsScalar}(1,k), \right. \\ \left. i = 1, \dots, 30, k = 1, \dots, 30 \right)$$

and

Maxima curve

$$:= \left( \text{qnorm}\left(\frac{(i-0.5)}{30}\right) \text{ versus } \text{evNormOrdStatsScalar}(k,k), \right. \\ \left. i = 1, \dots, 30, k = 1, \dots, 30 \right)$$

There are a lot of information can be captured from M-M plot about the whole distribution. Some of these features are

1. the joint point between the two curves is the average. Figure 1 shows the average is 0.
2. the gini's measure of the scale of the distribution can always be obtained from the second value in each curve,  $E(X_{2:2}) - E(X_{1:2})$ . Figure 1 shows that  $e_{22} \approx 0.56$  and  $e_{12} \approx -0.56$ , therefore,  $G \approx 0.56 + 0.56 = 1.12$ .
3. if  $\mu \neq 0$ , the Gini index of inequality can be estimated as  $G/2\bar{x}$ ; see, Cowell (2011) and Elamir (2013).
4. the skewness can be detected easily. For symmetric distributions the two curves will be at equal distances from the mean line. Figure 1 shows the two curves at the same distances from the mean line. For right skewed distribution the maxima curve will be wider than minima curve from mean line. Also, the minima

curve will be wider than maxima curve for the left skewed distribution.

5. the starting and ending of the two curves should give a clear picture about peak and tail of the distribution.

As the order statistics are available for uniform, logistic, Laplace and exponential distributions, Figure 2 shows the behaviour of M-M plot for these distributions at  $n = 30$ . The differences among them are clear and some features are

1. the joint point for two curves at averages 0.5, 0,0 and 1.
2. the first three distributions are symmetric about average and fourth one is asymmetric to the right.
3. the sharpness is clear for the second and third distributions while others are more flat. The long tails are clear for third and fourth distributions while short tail is clear for first distribution.

Also the M-M plot can be done for discrete distributions. From Arnold et al. (2008) the extreme order statistics for binomial distribution can be obtained as

$$E(X_{1:n}) = \sum_{x=0}^{N-1} [1 - F(x)]^n$$

and

$$E(X_{n:n}) = \sum_{x=0}^{N-1} [1 - (F(x))^n]$$

From Arnold et al. (2008) the extreme order statistics for Poisson distribution are

$$E(X_{1:n}) = \sum_{x=0}^{\infty} [1 - F(x)]^n$$

and

$$E(X_{n:n}) = \sum_{x=0}^{\infty} [1 - (F(x))^n]$$

Figure 3 shows the M-M plot for binomial and Poisson distributions for selected parameters. Some features may be concluded as

1. the discreteness of the distributions is obvious.
2. the joint points or averages are 5, 1, 9.5 and 0.7, respectively.
3. the right skewness is clear for second and fourth distributions while left skewness for third distribution and symmetric for the first distribution are obvious.

The Pareto distribution is a heavy tailed distribution which is used a lot in economics, has a probability density function that can be written as

$$f(x; x_0, \alpha) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}}, \quad 0 < x_0 \leq x$$

Where  $\alpha$  is the shape parameter which measures the heaviness of the right tail and  $x_0$  is a scale parameter; see, Arnold (1983) and Michael (2010). The corresponding cumulative distribution function is

$$F(x; x_0, \alpha) = 1 - \left(\frac{x}{x_0}\right)^{-\alpha}$$

The expected value is

$$E(X) = \frac{\alpha x_0}{\alpha - 1}, \quad \alpha > 1$$

Also from Malik (1966) the order statistics are defined as

$$E(X_{i:n}) = x_0 \frac{\Gamma(n+1)\Gamma\left(n-i-\frac{1}{\alpha}+1\right)}{\Gamma(n-i+1)\Gamma\left(n-\frac{1}{\alpha}+1\right)}$$

For extreme order statistic,

$$E(X_{1:n}) = x_0 \frac{\Gamma(n+1)\Gamma\left(n-\frac{1}{\alpha}\right)}{\Gamma(n)\Gamma\left(n-\frac{1}{\alpha}+1\right)}$$

and

$$E(X_{n:n}) = x_0 \frac{\Gamma(n+1)\Gamma\left(-\frac{1}{\alpha}+1\right)}{\Gamma\left(n-\frac{1}{\alpha}+1\right)}, \quad \alpha > 1$$

Figure 4 shows the M-M plot for Pareto distribution with parameters (10,8), (10,5), (10,2) and (10,1.25). Note the changes for the curves with changes of  $\alpha$  in terms of skewness and tails especially at the beginning and ending of curves.

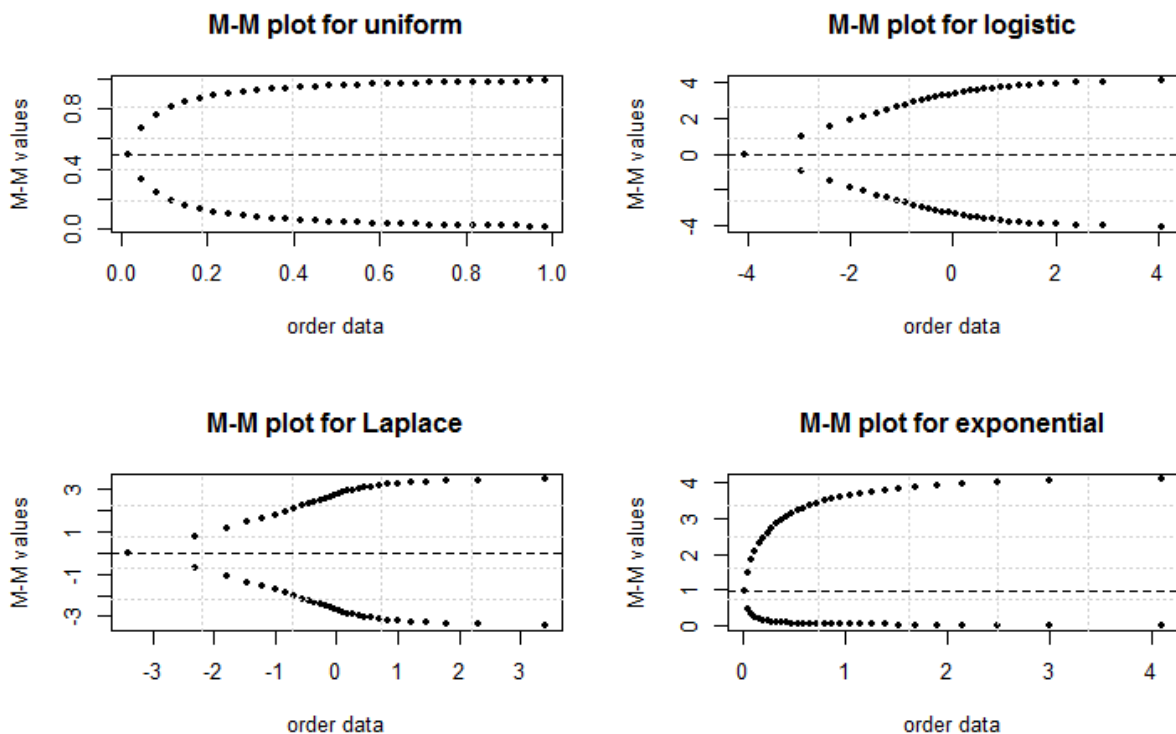


Figure 2. M-M plot for uniform, logistic, Laplace and exponential distributions at  $n = 30$

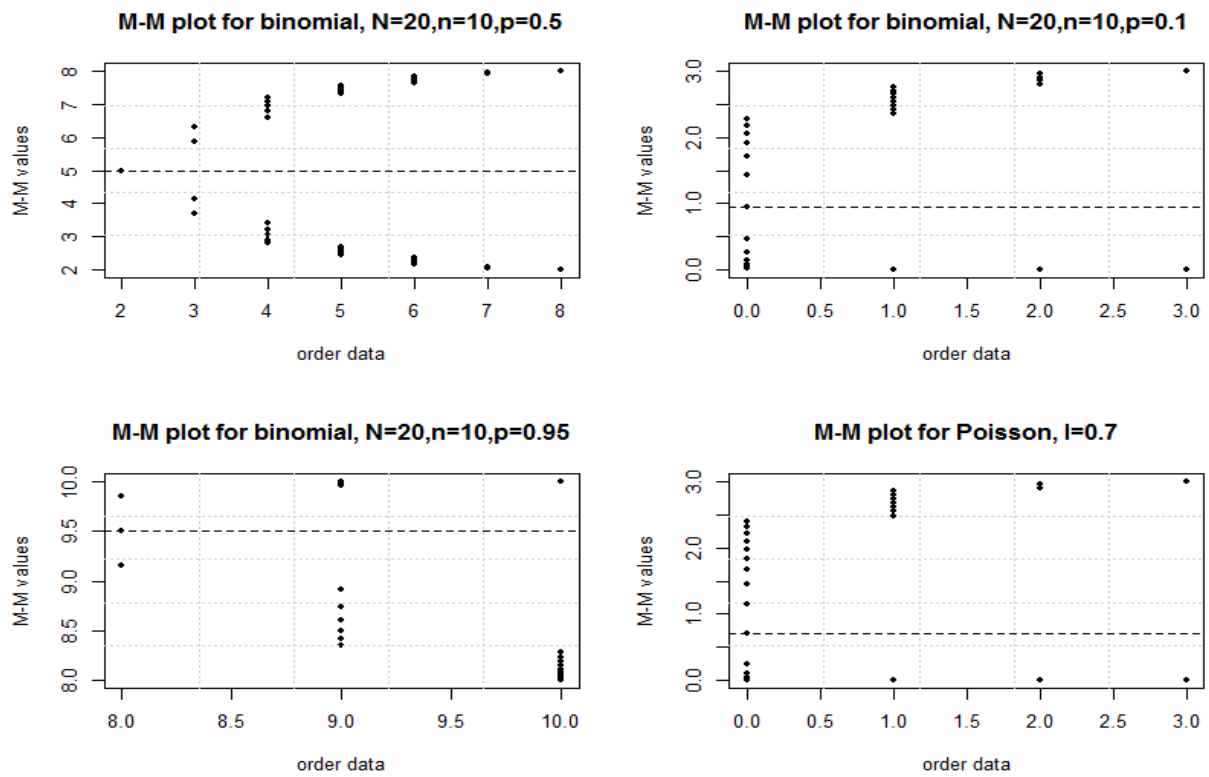


Figure 3. M-M plot for binomial and Poisson distributions for selected parameter. Note that  $\lambda = 0.7$  for Poisson

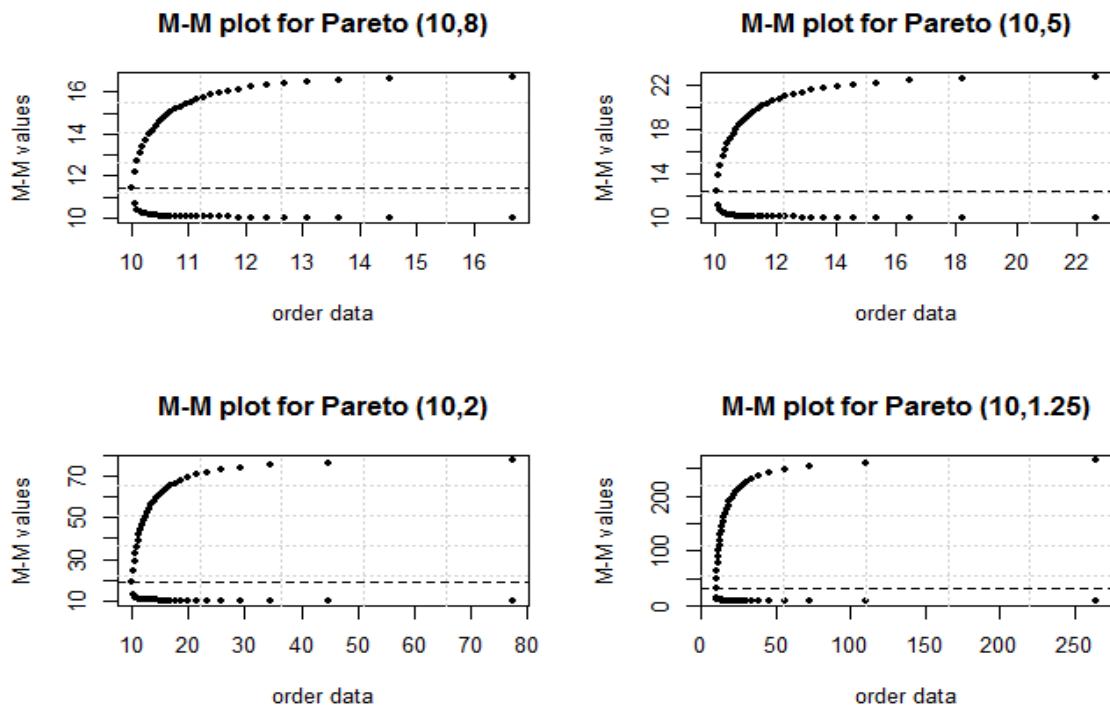


Figure 4. M-M plot for Pareto distribution with different values for shape parameter  $\alpha$  and  $n = 30$

## Two variables and M-M line plot

In the sense of Q-Q plot, M-M plot can be used to compare any data with some specified distribution or with any other data. M-M line plot considers the sample as a whole and plots the sample minima and maxima against the theoretical minima and maxima of the specified target distribution  $F$ . If a correct target distribution is given, the M-M plot hugs a straight line. Therefore, it can assess if a set of data plausibly came from some theoretical distribution by constructing M-M line plot. If there is another distribution for a random variable  $Y$ , therefore, the theoretical versus theoretical is

$$\text{M-M line} := \left( (E(X_{1:k}), E(X_{k:k})), (E(Y_{1:k}), E(Y_{k:k})), \right. \\ \left. k = 1, \dots, n \right)$$

and theoretical versus estimated is

$$\text{M-M line} := \left( (E(X_{1:k}), E(X_{k:k})), (\hat{E}(X_{1:k}), \hat{E}(X_{k:k})), \right. \\ \left. k = 1, \dots, n \right)$$

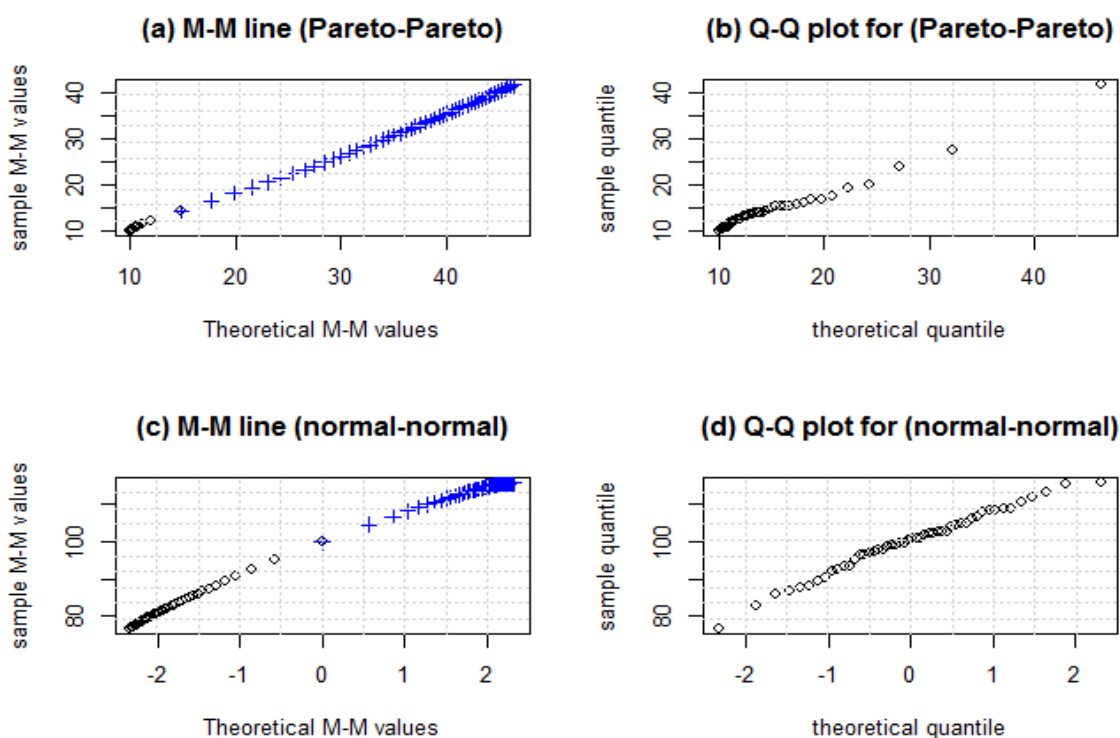
If this closes to straight line, the data follows the theoretical distribution. Like Q-Q plot, M-M line plot can be plotted without knowing the location and scale parameters where  $(E(X_{1:k}), E(X_{k:k}))$  can be computed from standard distributions. Moreover, M-M line plot can be used to check if two data have the same distribution using

$$\text{M-M line} := \left( (\hat{E}(X_{1:k}), \hat{E}(X_{k:k})), (\hat{E}(Y_{1:k}), \hat{E}(Y_{k:k})), \right. \\ \left. k = 1, \dots, n \right)$$

The points plotted in a M-M line plot are non-decreasing when viewed from left to right. If the two distributions being compared are identical, the M-M line plot follows the  $45^\circ$  line. If the differences between points  $e_{22}$  and  $e_{12}$  in M-M line plot is more for horizontal variable than vertical variable, the distribution plotted on the horizontal axis is more dispersed than the distribution plotted on the vertical axis and vice versa. If the difference between  $e_{22}$ -mean and mean  $- e_{12}$  is nearly zero the variable is symmetric, negative (left skewed) and (positive) right skewed. When M-M line plot is often S shaped this is indicating that one of the distributions has more than one mode and when it is often J shaped this is indicating that one of the distributions has much heavier tails than the other. This can be shown in the following cases.

### Case 1: theoretical versus estimated

Figure 5 shows M-M line and Q-Q plots for theoretical and simulated data from Pareto and normal distributions. It is clear that the M-M line plot is more stable than Q-Q plot for heavy tail distributions. For the normal distribution most of the data for M-M plot are concentrated at the extreme tails while few data at the middle. On the other hand most of the data for Q-Q plot are concentrated on the middle and few of them at the extreme tails.



**Figure 5.** M-M and Q-Q plots for theoretical Pareto (10,3) versus simulated Pareto (10,3) in (a) and (b), and theoretical normal (0,1) versus simulated normal (100,8) in (c) and (d) and  $n = 50$

**Case 2: estimated versus estimated**

Figure 6 shows two shapes for M-M line when one distribution has more than one mode like Beta distribution (0.5, 0.5), the shape is near from 'S' and when one distribution is very heavy, it shows 'J' shape.

Figure 7 shows the M-M line and Q-Q plots for simulated data from lognormal distribution (5, 0.5) and it is clear the difference between two curves at the extreme tails; see, Elamir (2016).

Also the Min and Max normal plots will complete the picture of QQ-norm plot especially at the extreme tails of the distribution. The Min-norm plot is proposed by plotting the exact minimum order statistics of size  $k$  from standard normal distribution that can be obtained from package *EnvStats* in R software versus estimated minimum order statistics from a data as

Min normal

plot:= (evNormOrdStatsScalar(1,  $k$ ) versus  $\hat{E}(X_{1:k})$ ,  
 $k = 1, 2, \dots, n$ )

Also the maximum plot is proposed as

Max normal plot

:= (evNormOrdStatsScalar( $k, k$ ) versus  $\hat{E}(X_{k:k})$ ,  $k = 1, 2, \dots, n$ )

The pattern of points in the Min-normal and Max-normal plots must show straight line or close to straight line; see for more details, Elamir (2016).

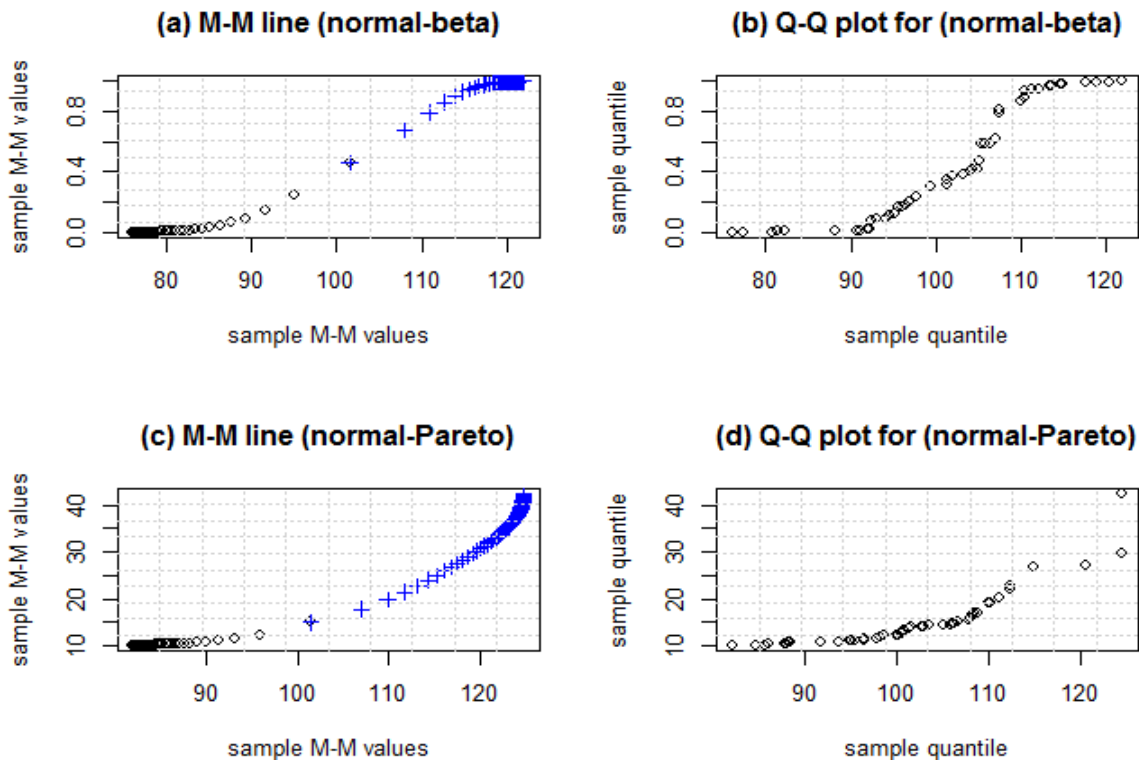
**Standard M-M plot**

Away from M-M line, M-M plot still has a great value in comparisons between data where it is possible to do more than one M-M plot on the graph paper to compare between actual data and theoretical data but one has to take care of location and scale. In this aspect, standard M-M plot can be proposed as exactly M-M plot except is done for standardised values

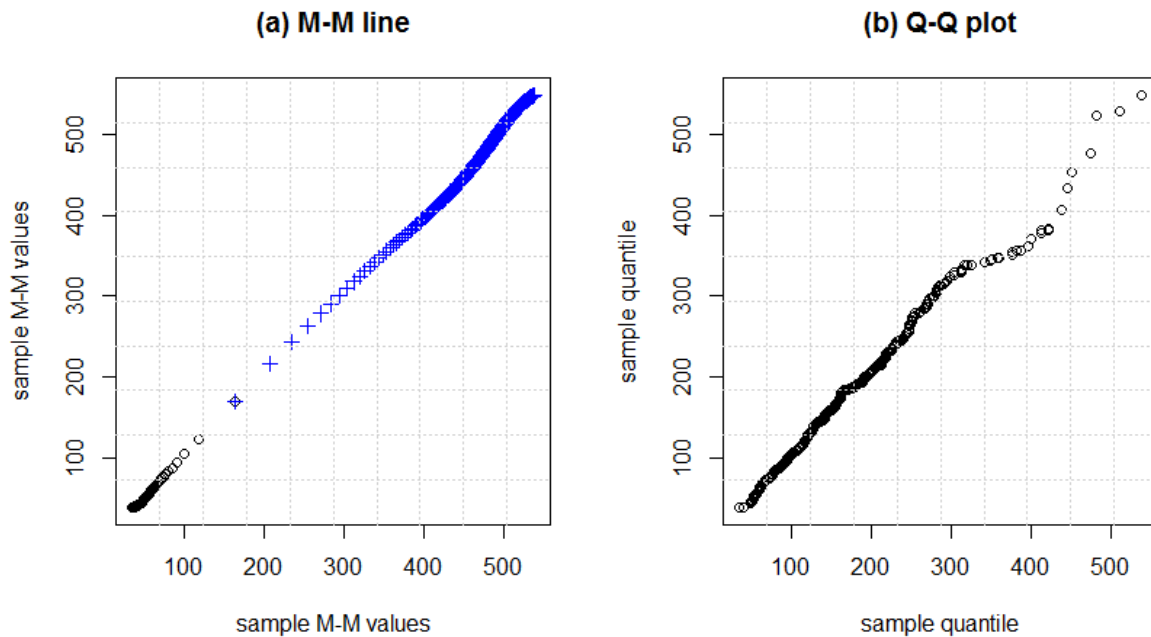
$$z_{i:n} = (x_{i:n} - \bar{x})/s$$

where  $\bar{x}$  and  $s$  are the sample mean and standard deviation, respectively. Standard M-M plot is focusing on the shape of the data after excluding the location and scale parameters.

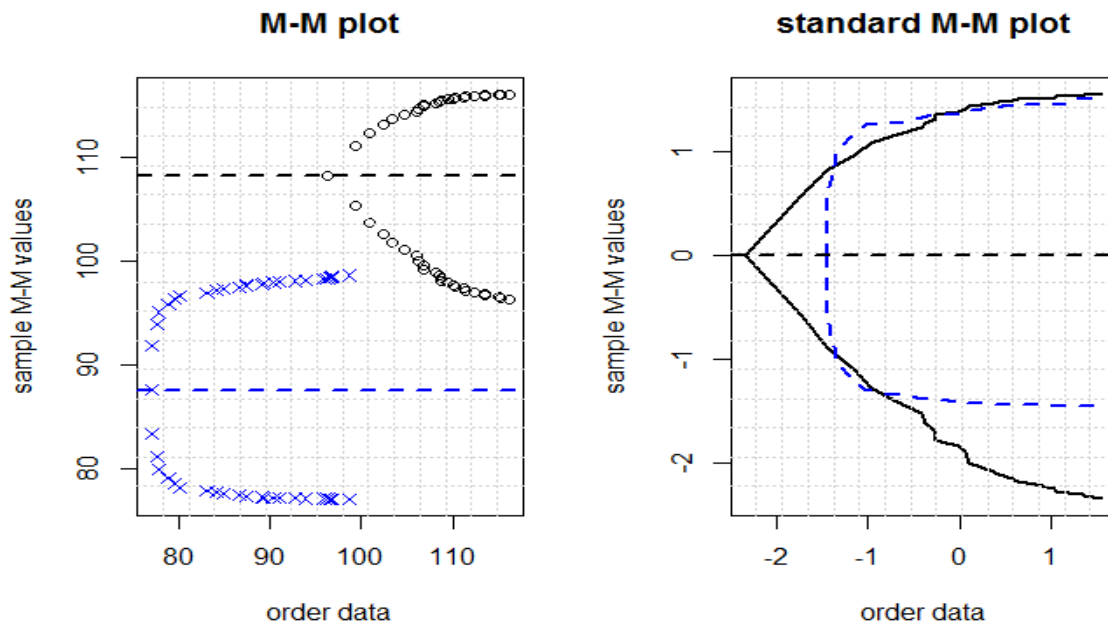
Figure 8 shows M-M and standard M-M plot theoretical data from uniform and normal distributions versus simulated data from same distributions.



**Figure 6.** M-M and Q-Q plots for simulated normal (100,10) versus simulated Beta (0.5,0.5) in (a) and (b), and simulated normal (100,10) versus simulated Pareto (10,2.5) in (c) and (d) and  $n = 50$



**Figure 7.** M-M and Q-Q plot for simulated data from lognormal distribution (5,0.5) and  $n = 500$



**Figure 8.** M-M and standard M-M plots for simulated data from uniform (75,100), normal (110,5) distributions and  $n = 25$

## 4. Application

M-M, standard M-M and MM-line plots can be used to compare among several variables. For example, in the analysis of variance the data are assumed to be normally distributed and have equal variances. An article in *Nature* describes an experiment to investigate the effect of consuming chocolate on cardiovascular health (“Plasma Antioxidants from Chocolate,” *Nature*, Vol. 424, 2003). The experiment consisted of using three different types of

chocolates: 100g of dark chocolate, 100g of dark chocolate with 200mL of full-fat milk, and 200g of milk chocolate. Twelve subjects were used, 7 women and 5 men, with an average age range of  $32.2 \pm 1$  years, an average weight of  $65.8 \pm 3.1$  kg, and body-mass index of  $21.9 \pm 0.4$  kg m<sup>-2</sup>. On different days a subject consumed one of the chocolate-factor levels and one hour later the total antioxidant capacity of their blood plasma was measured in an assay; see, Montgomery (2013). Data is summarized in Table 1.



Figure 9 presents M-M plots for the data from this experiment. The result is an indication that the blood antioxidant capacity one hour after eating the dark chocolate is higher than for the other two treatments. The variability in the sample data from all three treatments seems very similar. Also the shape for three treatments looks similar to normal distribution. Moreover, Table 2 gives some measures can be estimated from the graph.

Figure 10 shows M-M line plot using theoretical standard normal distribution versus the three treatments for blood plasma levels one hour following chocolate consumption data with linear regression line added to each line based on

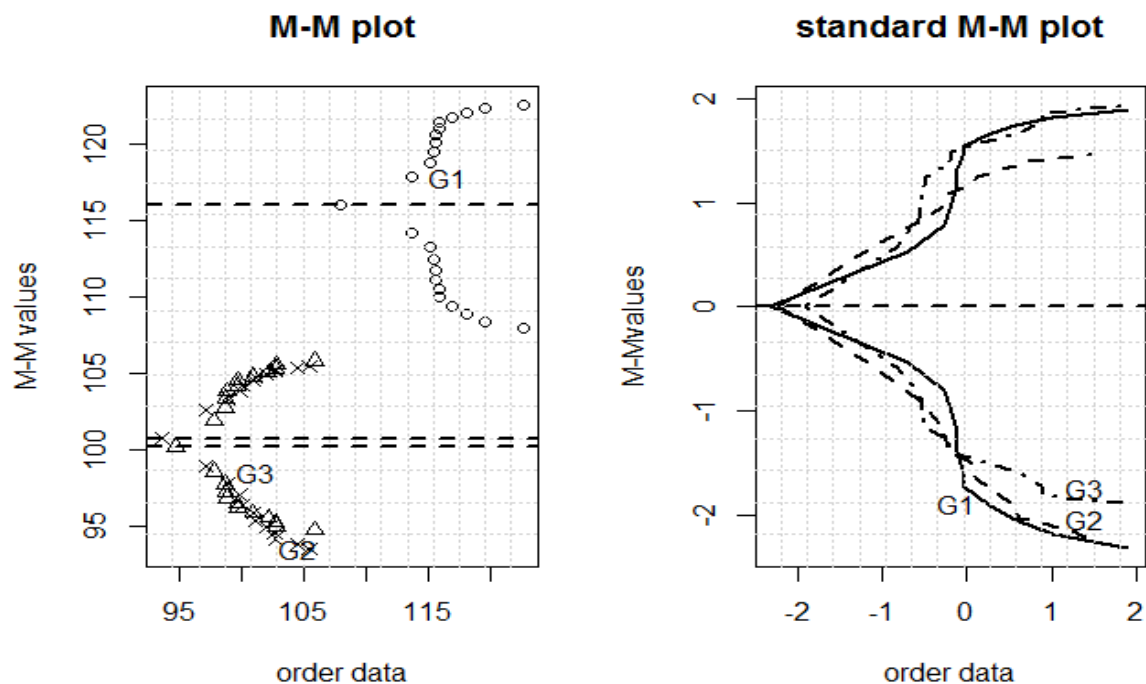
Sample MM values =  $a + b$  Theoretical MM values

The same conclusion can be made from this plot as M-M plot.

**Table 1.** Blood plasma levels one hour following chocolate consumption\*

Factor	observations											
DC (G1)	118.8	122.6	115	113.6	119.5	115.9	115.8	115.1	116.9	115.4	115.6	107.9
DC+MK (G2)	105.4	101.1	102.7	97.1	101.9	98.9	100.0	99.8	102.6	100.9	104.5	93.5
MC (G3)	102.1	105.8	99.6	102.7	98.8	100.9	102.8	98.7	94.7	97.8	99.7	98.6

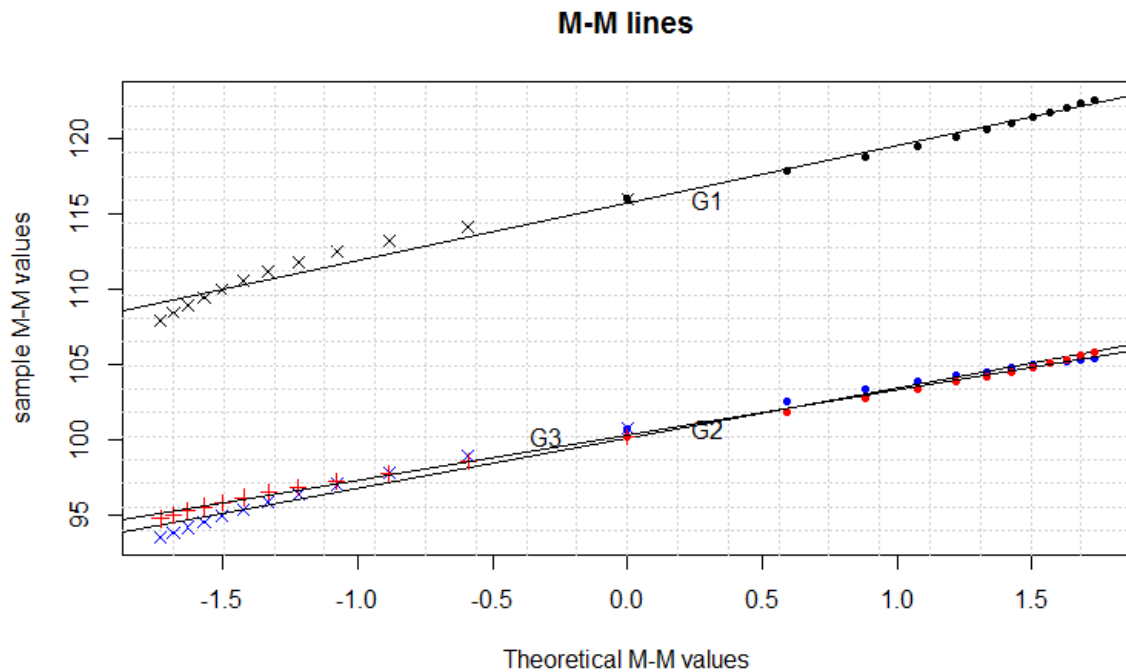
\* "Plasma Antioxidants from Chocolate" *Nature*, Vol. 424, 2003, pp. 1013



**Figure 9.** M-M and standard M-M plots for blood plasma levels one hour following chocolate consumption data

**Table 2.** Estimated measures from figure 9 for blood plasma levels data

	G1	G2	G3
	117.8	102.5	101.8
	116	100.7	100.2
	114	98.8	98.5
Gini= (e22 – e12)	3.8	3.7	3.3



**Figure 10.** M-M line plot based on theoretical standard normal distribution versus the three treatments for blood plasma levels one hour following chocolate consumption data

## 5. Conclusions

Minima-maxima plot depends on extreme order statistics and it is introduced to capture all information not only about the tails of the distribution but also about the whole distribution of the data. The pattern of points in the M-M plot is used to compare the shapes of distributions and providing a graphical view of how properties such as location, scale and skewness. Also, it used to plot the sample data against the theoretical extreme order statistics or sample extreme order statistics and then a visual check is made to see whether or not the points are close to a straight line. The main advantage of M-M plot over Q-Q plot is its stability at the extreme ends of the distribution. Actually, the main idea behind M-M line plot is that the average is a good representative for data at the middle but the extreme tails must be given most attention or weights.

One limitation of M-M plot is when the extreme order statistics are not defined such as the Cauchy distribution where few extreme order statistics are not defined. But the M-M plot may still be plotted by using the available information and ignoring undefined values. Of course, in this case some information will be lost.

## Appendix A: R program for M-M plots and Q-Q plot

**## function to estimate lower extreme order statistics**

```
LGd=function(x,t){
  n=length(x); i=1:n; x=sort(x)
```

```
  c1=1/choose(n,t)
```

```
  t1=choose(n-i,t-1)*x
```

```
  c1*sum(t1)}
```

**## function to estimate upper extreme order statistics**

```
LGo=function(x,t){
  n=length(x); i=1:n; x=sort(x)
```

```
  c1=1/choose(n,t)
```

```
  t1=choose(i-1,t-1)*x
```

```
  c1*sum(t1) }
```

```
wd1=0;wo1=0;wdz=0;woz=0;wdz1=0;woz1=0;wdy1=0;
woy1=0
```

**## graph function**

```
GPH.MM0=function(y,y1){## y first variable ## y1
second variable
```

```
  z=(y-mean(y))/sd(y)
```

```
  z1=(y1-mean(y1))/sd(y1)
```

```
  for (i in 1:(n)) {
```

```
    wd1[i]=LGd(y,i); wo1[i]=LGo(y,i); wdz[i]=LGd(z,i);
```

```
    woz[i]=LGo(z,i)
```

```
    wdy1[i]=LGd(y1,i); woy1[i]=LGo(y1,i);
```

```
    wdz1[i]=LGd(z1,i); woz1[i]=LGo(z1,i)
```

```
  }
```

```
  ## graphs
```

```
  par(mfrow=c(2,2))
```

**##MM plot**

```
  plot(sort(y),wd1,ylim=c(min(y,y1),max(y,y1)),
```

```
  xlim=c(min(y,y1),max(y,y1)),typ="p",pch=1,lwd=1,
```

```
  main="M-M plot",xlab="order data",ylab="M-M
values")
```

```
  points(sort(y),wo1,pch=1,lwd=1)
```

```

abline(h=mean(y),lwd=2,lty=2)
points(sort(y1),wdy1,col="blue",pch=20)
points(sort(y1),woy1,col="blue",pch=20)
abline(h=mean(y1),col="blue",lwd=2,lty=2)

plot(sort(z),wdz,ylim=c(min(wdz,wdz1),max(woz,woz1)),
),

xlim=c(min(z,z1),max(z,z1)),type="l",pch=20,lwd=2,
main="standard M-M plot",xlab="order
data",ylab="M-M values")
lines(sort(z),woz,pch=18,lwd=2)
abline(h=mean(z),lwd=2,lty=2)
lines(sort(z1),wdz1,col="red",pch=20,lty=2,lwd=2)
lines(sort(z1),woz1,col="red",pch=20,lty=2,lwd=2)
abline(h=mean(z1),lwd=2,lty=2)
##M-M line
wdo=c(wd1,wo1)
wdoy1=c(wdy1,woy1)
plot(wd1,wdy1,main="M-M line
plot",xlab="Theoretical M-M values",
ylab="M-M
values",xlim=c(min(wd1,wo1),max(wd1,wo1)),
ylim=c(min(wdy1,woy1),max(wdy1,woy1)))
points(wo1,woy1, col="blue",pch=3)
## qq plot
qqplot(y,y1,main="Q-Q plot",xlab="theoretical
quantile",ylab="sample quantile")
grid(5,5)
}

```

## REFERENCES

- [1] Alpay, O. (2016) On the local dependence function of extreme order statistics. *Ege University of Faculty of Sciences*, 40, 10-16.
- [2] Arnold, B.C., Balakrishnan, N. and Nagataja, H.N. (2008) *A first course in order statistics*. 2<sup>nd</sup> Ed., Society for Industrial and Applied Mathematics, SIAM.
- [3] Arnold, B.C. (1983) Pareto Distributions. International Co-operative Publishing House.
- [4] Chambers, J., Cleveland, W., Kleiner, B. and Tukey, P. (1983) *Graphical methods for data Analysis*. 1<sup>st</sup> Ed., Wadsworth.
- [5] Chan, L.K. (1967) On a characterization of distributions by expected values of extreme order statistics. *Am. Math. Monthly*, 74, 950-951.
- [6] Cleveland, W.S. (1994) *The Elements of Graphing Data*. 1<sup>st</sup> Ed., Hobart Press.
- [7] Cleveland, W. and McGill, M. (1988) *Dynamic Graphics for Statistics*. Wadsworth.
- [8] Cowell, F. (2011) *Measuring inequality*. 1<sup>st</sup> Ed., OUP UK.
- [9] David, H. A. (1981) *Order Statistics*. 2nd ed., Wiley, New York.
- [10] Downton, F. (1966) Linear estimates with polynomial coefficients. *Biometrika*, 53, 129-141.
- [11] Elamir, E.A.H. and Seheult, A. (2003) Trimmed L-moments. *Computational Statistics and Data Analysis*. 43, 299-314.
- [12] Elamir, E.A.H. and Seheult, A. (2004) Exact variances of sample L-moments. *Journal of Statistical Planning and Inference*, 124, 337-359.
- [13] Elamir, E. A. H. (2013) On estimation of some abbreviated social welfare measure. *Quality & Quantity: International Journal of Methodology*, 63, 245-268.
- [14] Elamir, E.A.H. (2016) Extreme order statistics plot versus quantile quantile plot: Nonparametric visualization for a data. *International Journal of Applied Mathematics & Statistical Sciences*, 5, 47-62.
- [15] Haung, J.S. (1989) Moment problem of order statistics. A review. *International Statistical Review*, 57, 59-66 (1989)
- [16] Makkonen, L. (2008) Bringing closure to the plotting position controversy. *Communications in Statistics - Theory and Methods*, 37, 460-467.
- [17] Malik H. J. (1966) Exact Moments of Order Statistics from Pareto Distribution. *Skand. Aktuarie Tidskr*, 49, 144-157.
- [18] Michael, H. (2010) Pareto's Law. *Mathematical Intelligencer*. 32, 38-43.
- [19] Montgomery, D.C. (2013) *Design and Analysis of Experiments*. 8<sup>th</sup> Ed., John Wiley & Sons.
- [20] Scott, D. (1992) *Multivariate Density Estimation: Theory, Practice and Visualization*. John Wiley and Sons.
- [21] Pearson, K. (1895) Contributions to the mathematical theory of evolution. II. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 186, 343-414.
- [22] Wicklin, R. (2011) Modeling the distribution of data? Create a Q-Q plot. *The Do Loop*. OCT. 28.
- [23] Wilk, M.B.; Gnanadesikan, R. (1968) Probability plotting methods for the analysis of data. *Biometrika*, 55, 1-17.