THRESHOLD DETERMINATION FOR MAXIMUM PRODUCT OF SPACING METHODOLOGY WITH TIES FOR EXTREME EVENTS

Peter G. Murage[1], Joseph K Mung'atu[2] and Evelyne Odero[3]

Abstract

Extreme events are defined as values of the event below or above a certain value called threshold. A well chosen threshold helps to identify the extreme levels. Several methods have been used to determine threshold so as to analyze and model extreme events. Maximum Product of spacing is one of these methods. However, there is a problem encountered while modeling data through this method in that the method breaks down when there is a tie in the exceedances. This study improved MPS method in order to determine an optimal threshold for extreme values in a data set containing ties, estimated the GPD parameters with the optimal threshold derived and then applied the method to determine the GPD parameters for a real market data that could be containing ties . The study applied a method to determine optimal threshold based on improved maximum product of spacing method and used Generalized Pareto Distribution (GPD) and Peak over threshold (POT) methods as the basis of identifying extreme. The peaks-over-threshold (POT) models are models for all large observations which exceed a high threshold. The POT models are generally considered to be the most useful for practical applications. The study used the method developed to deal with the ties to model the market volume data. This study will help the Statisticians in different sectors of our economy to model extreme events involving ties. To Statisticians, the structure of the extreme levels which exist in the tails of the ordinary distributions is very important in analyzing, predicting and forecasting the likelihood of an occurrence of extreme event.

KEY WORDS: Generalized Pareto distribution, Peak Over Threshold, Improved MPS.

[1] PhD student; Email: gitmurage@gmail.com; Mobile: 0710507266

[2] Department of Statistics and Actuarial Science JKUAT; Email: j.mungatu@fsc.jkuat.ac.ke, Mobile: 0720824708

[3] Department of Mathematics-MMUST; Email: oedero@mmust.ac.ke, Mobile: 0713145508

Introduction

Certain values in the tails of any distribution represent extreme events and they are pointers to eventuality. The values in the tails are rare, few, but can have great impact on the conclusion arrived at by the analysts. Different sectors of our life experience Extreme events and here we mention just but a few. According to (Butterfield, 2009) Extreme low production in agriculture results to famine if the agriculture depends on rainfall. This means that the amount of rain experienced in that region was too low that crops dried up or very high rainfall that it destroyed all crops that had been planted. (Prudhome, 1999) studying extreme rainfall in mountainous region and (Geremy, 2012) studying extreme rainfall in west Africa did observe that, how low or high the amount of rainfall depends on the threshold attached to the rainfall in that region. In insurance industries Box [10] while discussing tools in finance and insurance, noted that extreme high claims by the customers that can be very dangerous for the company while extreme low claims by the customers can be very beneficial for the company's profit. This means that there is a critical level that the insurance company would wish it is not surpassed and if it is, according to (Embretchet, 1997) it must be prepared for this eventuality. Very high emissions of the waste products from the manufacturing industries are detrimental to the environment and ozone layer. However, countries must continue to industrialize or expand their industries for economic prosperity. Certain level of emissions must not be exceeded otherwise the environment and ozone layer would be destroyed. The critical value for which if exceeded an eventuality occurs is called threshold. The events beyond this threshold are called extreme events and they happen to be at the tails of the distribution. Extreme value theory (EVT) proposed by (Fisher, 1928) is a tool which attempts to provide us with the best possible estimate of the tail area of the distribution. In work on the importance of tail dependence in Bivariate frequency analysis, there are two principal kinds of model for extreme values. The oldest group of models is the block maxima models; these are models for the largest observations collected from large samples of identically distributed observations. For example, if we record daily or hourly losses and profits from trading a particular instrument or group of instruments, the block maxima or minima method provides a model which may be appropriate for the quarterly or annual maximum of such values. According to (Balkema, 1974) and (Pickands, 1975) the block maxima/minima methods are fitted with the generalized extreme value (GEV) distribution. A more modern group of models is the peaks-over-threshold (POT) models; these are models for all large observations which exceed a high threshold. The POT models are generally considered to be the most useful for practical applications, due to a number of reasons. First, by taking all exceedances over a suitably high threshold into account, they use the data more efficiently. Second, they are easily extended to situations where one wants to study how the extreme levels of a variable Y depend on some other variable X for instance, Y may be the level of tropospheric ozone on a particular day and X a vector of meteorological variables for that day. This kind of problem is almost impossible to handle through the annual maximum method. In order to identify the extreme values, one must understand how the sample data collected is analyzed to determine the extreme cases. This involves determination of the threshold and the exceedances. These exceedances are then modeled to understand the behavior of the data in the tails. Many methods of determining an optimal threshold have been developed. The most common one is the graphical method proposed by (Hill, 1975) This method is however subjective and requires experts to determine the threshold. The most successful method is the Maximum Product of Spacing (MPS) proposed by (Cheng, 1983). This method however encounters a problem whenever the exceedances have a tie. To study this problem, the study used simulated data containing ties and real data from Nairobi Securities exchange (NSE). Events in the area outside three standard deviation in a normal distribution are termed extreme events. Our study was based on these extreme events. Extreme events can be either beneficial or destructive. One of the greatest challenges to a risk manager according to (Yuejian, 2002) is to implement risk management tools which allow for modeling rare but damaging events, and permit the measurement of their consequences. Extreme value theory (EVT) plays a vital role in these activities. (Coles, 2001) in his book on extreme modeling of extreme values emphasized that, Extreme value theory relates to the asymptotic behavior of extreme observations of a random variable. It provides the fundamentals for the statistical modeling of rare events, and is used to compute tail risk measures.

Methodology

Improved MPS Methodology

The MPS allows efficient estimators in non-regular cases where MLE may not exist. This is especially relevant to the GEV distribution in which the MLE does not exist when $\varepsilon < -1$, Smith [38]. Let $x_1, x_2, ..., x_n$ be a random sample of independent observations from a continuous distribution F_{θ_0} belonging to $F_{\theta,}\ \theta \in \Theta$ Applying the probability transform $F_0\left(.\right)$ to the order $x_{1,n} \le x_{2,n} \le \ldots \le x_{n,n}$ yields $0 \equiv F_\theta\left(x_{0,n}\right) \le F_\theta\left(x_{1,n}\right) \le \ldots \le FF_{\theta_0}\left(x_{n+1,n}\right) \equiv 1$. We define the spacing's as the gaps between the values of the distribution function at adjacent ordered points $D_i\left(\theta\right) = F_\theta\left(x_i\right) - F_\theta\left(x_{i-1}\right)$, where $i = 1,2,\ldots,n+1$ The maximum spacing estimator of θ_0 is defined as value $\hat{\theta} = \arg\max_{\theta\in\Theta} S_n\left(\theta\right)$ that maximizes the logarithm of the geometric mean of sample spacing's.

$$S_n\left(\theta\right) = \ln \sqrt[n+1]{\left(D_1\left(\theta\right).D_2\left(\theta\right)\ldots D_{n+1}\left(\theta\right)\right)} \tag{1}$$

$$= \frac{1}{n+1}\sum_{i=1}^{n+1} \ln D_i\left(\theta\right)$$

The maximum spacing estimator as defined is sensitive to closely spaced observations and especially the ties. That is, for any $x_{i+m} = x_{i+m} = \ldots = x_i$

Then

$$D_{i+m}\left(\theta\right) = D_{i+m=1}\left(\theta\right) = \ldots = D_i\left(\theta\right)$$

This therefore collapses the method. The modified MPS method proposed here is to use grouped data frequency table. Let $x_1, x_2, \ldots, x_n$ occur $f_1, f_2, \ldots, f_n$ times respectively. The geometric mean is given by

$$G = \left(x_1^{f_1}.x_2^{f_2}\ldots.x_n^{f_n}\right)^{\frac{1}{N}}$$

$$= \left[\prod_{i=1}^{n} x_i^{f_i}\right]^{\frac{1}{N}}$$

$$\ln G = \frac{1}{N}\sum_{i=1}^{n} f_i \ln x_i \tag{2}$$

This leads to the modified MPS method as

$$S_n\left(\theta\right) = \ln \sqrt[n+1]{\left(D_1^{f_1}\left(\theta\right)D_2^{f_2}\left(\theta\right)\ldots D_{n+1}^{f_{n+1}}\left(\theta\right)\right)}$$

$$= \frac{1}{n+1}\sum_{i=1}^{n+1} f_i \ln D_i\left(\theta\right) \tag{3}$$

If $f_1 = f_2 = ... = f_{n+1} = 1$ then we go back to the standard MPS. The Spacing's are such that $\sum_{i-1}^{n} D_i(\theta) = 1$: Under MPS, the $D_i(\theta)$ are defined as:

$$\begin{aligned} D_1(\theta) &= F(x_{1;n}, \theta) \\ D_i(\theta) &= F(x_{i;n}, \theta) - F(x_{i-1;n}, \theta) \\ D_{n+1}(\theta) &= 1 - F(x_{n;n}, \theta) \end{aligned} \tag{4}$$

Therefore, Equation 3 can be partitioned as:

$$S_n(x_i;\theta,\varepsilon,\sigma) = \frac{1}{n+1}\left\{ f_1 \ln D_1(\theta) + \sum_{i=2}^{n} f_i \ln D_i(\theta) + f_{n+1} \ln D_{n+1}(\theta) \right\} \tag{5}$$

To estimate the parameters of a Generalized Pareto distribution, we use the equation

$$G(x;\theta,\varepsilon,\sigma) = \begin{cases} 1 - \left[1 + \varepsilon\left(\dfrac{x-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}, & \varepsilon \neq 0 \\ 1 - \exp\left[-\dfrac{x-\theta}{\sigma}\right], & \varepsilon = 0 \end{cases} \tag{6}$$

Equation 6 was substituted in equation 5

Case 1: $\varepsilon \neq 0$

We define,

$$\begin{aligned} D_1(\theta) &= 1 - \left[1 + \varepsilon\left(\frac{x_1-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}} \\ D_i(\theta) &= \left(1 - \left[\left[1 + \varepsilon\left(\frac{x_i-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right]\right) - \left(1 - \left[1 + \varepsilon\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right) \\ D_{n+1}(\theta) &= 1 - \left(1 - \left[1 + \varepsilon\left(\frac{x_n-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right) \end{aligned} \tag{7}$$

Which leads to:

$$S_n\left(x_i;\theta,\varepsilon,\sigma\right)=\frac{1}{n+1}\left\{f_1\ln\left(1-\left[1+\varepsilon\left(\frac{x_1-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right)+\sum_{i=2}^{n}f_i\ln\left(\left[1+\varepsilon\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}-\left[1+\varepsilon\left(\frac{x_i-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right)+f_{n+1}\ln\left[1+\varepsilon\left(\frac{x_n-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right\} \tag{8}$$

If we let ,

$$\begin{aligned}
K_1 &= \ln\left(1-\left[1+\varepsilon\left(\frac{x_1-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right)\\
K_2 &= \ln\left\{\left[1+\varepsilon\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}-\left[1+\varepsilon\left(\frac{x_i-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}\right\}\\
K_3 &= \ln\left[1+\varepsilon\left(\frac{x_n-\theta}{\sigma}\right)\right]^{-\frac{1}{\varepsilon}}
\end{aligned} \tag{9}$$

Through partial differentiation of equation 9 and substituting in equation 8, we get

$$\begin{aligned}
S_\theta^{'} &= \frac{1}{n+1}\left\{f_1\frac{\partial K_1}{\partial\theta}+\sum_{i=2}^{n}f_i\frac{\partial K_2}{\partial\theta}+f_{n+1}\frac{\partial K_3}{\partial\theta}\right\}\\
S_\sigma^{'} &= \frac{1}{n+1}\left\{f_1\frac{\partial K_1}{\partial\sigma}+\sum_{i=2}^{n}f_i\frac{\partial K_2}{\partial\sigma}+f_{n+1}\frac{\partial K_3}{\partial\sigma}\right\}\\
S_\varepsilon^{'} &= \frac{1}{n+1}\left\{f_1\frac{\partial K_1}{\partial\varepsilon}+\sum_{i=2}^{n}f_i\frac{\partial K_2}{\partial\varepsilon}+f_{n+1}\frac{\partial K_3}{\partial\varepsilon}\right\}
\end{aligned} \tag{10}$$

Which we optimize by setting them to zero to get the GPD parameter estimates $\hat{\theta}, \hat{\sigma}\ \ and\ \ \hat{\varepsilon}$

Case2: $\varepsilon = 0$

We define

$$D_1(\theta) = 1 - \exp\left[-\left(\frac{x_1-\theta}{\sigma}\right)\right]$$

$$D_i(\theta) = \exp\left[-\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right] - \exp\left[-\left(\frac{x_i-\theta}{\sigma}\right)\right] \quad (11)$$

$$D_{n+1}(\theta) = \exp\left[-\left(\frac{x_n-\theta}{\sigma}\right)\right]$$

$$S_n(x_i;\theta,\varepsilon,\sigma) = \frac{1}{n+1}\left\{f_1 \ln\left(1-\exp\left[-\left(\frac{x_1-\theta}{\sigma}\right)\right]\right) + \sum_{i=2}^{n} f_i \ln\left(\exp\left[-\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right] - \exp\left[-\left(\frac{x_i-\theta}{\sigma}\right)\right]\right) - f_{n+1}\left(\frac{x_n-\theta}{\sigma}\right)\right\}$$

(12)

We define

$$K_1 = \ln\left(1-\exp\left[-\left(\frac{x_1-\theta}{\sigma}\right)\right]\right)$$

$$K_2 = \ln\left\{\exp\left[-\left(\frac{x_{i-1}-\theta}{\sigma}\right)\right] - \exp\left[-\left(\frac{x_i-\theta}{\sigma}\right)\right]\right\} \quad (13)$$

$$K_3 = \left(\frac{x_n-\theta}{\sigma}\right)$$

Through partially differentiating equation 10 and substituting in equation 12, we obtained

$$S'_{\theta^*} = \frac{1}{n+1}\left\{f_1\frac{\partial K_1^*}{\partial\theta} + \sum_{i=2}^{n} f_i\frac{\partial K_2^*}{\partial\theta} + f_{n+1}\frac{\partial K_3^*}{\partial\theta}\right\}$$

$$S'_{\sigma^*} = \frac{1}{n+1}\left\{f_1\frac{\partial K_1^*}{\partial\sigma} + \sum_{i=2}^{n} f_i\frac{\partial K_2^*}{\partial\sigma} + f_{n+1}\frac{\partial K_3^*}{\partial\sigma}\right\} \quad (14)$$

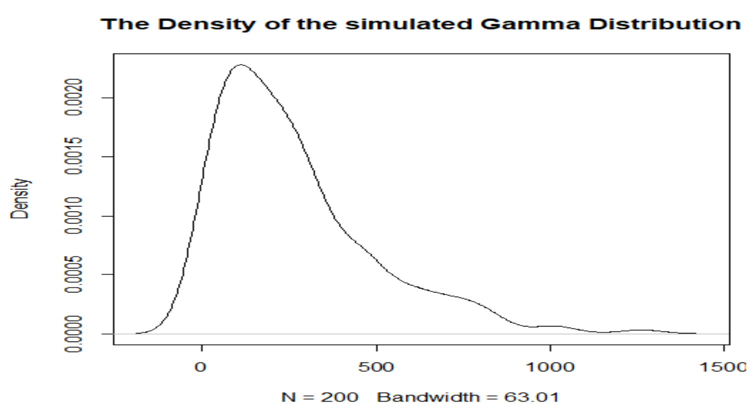
Equation 14 was then set to zero in order to optimize and obtain gpd parameters

Results and discussion

We developed an R-code for the standard and improved MPS model where the method of optimization was SANN.

We simulated data from a gamma distribution with the parameters shape=2.6, scale=1:1000. Repetitions were later introduced in the order of 0, 20, 40 and 60. The repeated values gave rise to situations of ties. Gamma distribution is known to have fairly heavy tails. To determine our threshold, we simulated a set of data constituting of 300 values. 100 values did not have a repetition while 100 values had each a repetition making them to have a frequency of 2 each. This set of data was used in the improved MPS model . After the simulation, this set of data was reorganized in such a way that the 300 values had a frequency of 1 each regardless of whether it was repeated or not. This set of data was used in the standard MPS model. The normal equations derived above were used as the model for the improved MPS methodology. For our improved three parameter MPS method, each tie formed a frequency f_i . When the values have not tied, the frequency f_i is 1. The frequency of the first value is f_1 while that of the last value is f_{n+1}. The simulated data was used to optimize the model equation 10 for the three parameter and equation 14 for the two parameter. The threshold ,scale and shape of the GPD parameters were therefore determined through the optimization of model 10 and 14 using the simulated data.

The simulated values had a distribution with the density shown in the figure 1



.

Figure 1:Gamma density

The density is skewed towards right figure 1. Meaning that the distribution had some extreme values. The performance of the standard and improved MPS model was then compared using the obtained values from the optimized results.

The table gives the results of the performance of the two parameter model for the standard and improved MPS model.

Table 1 .Two Parameter model

	Location	**scale**
improved	736.476	13.72969
standard	725.5767	16.31062

From Table 1, the threshold of the improved model was higher than the threshold of obtained through the standard model. The scale of the improved model was lower than the scale of the standard model.

The same data was used to compare the performance of the three parameter improved and standard models and the results are as in Table 2

Table 2:Three parameter

	location	**scale**	**shape**
improved	738.1303	9.483573	-0.84884
standard	726.3707	13.33941	-5.49648

In Table 2, the threshold obtained from the improved MPS model was 738.130 as compared to that of the standard MPS model which was 726.370. The improved MPS model performed better than the standard model. The scale parameter of the improved model was 9.48373 while that of the standard MPS model was 13.33941. The scale parameter of the standard MPS model was higher than that of the improved model.. The shape parameter of the improved model was 0.8488 and that of the standard model was 5.49648. The standard MPS model had a higher shape parameter compared to that of improved model.

To investigate how the parameters were behaving, we created repetitions within our simulated data of 300 values. The repetitions were in the order of 0 repetitions, 20 repetitions, 40 repetitions and 60 repetitions. The results are shown in Table 3 for two parameter model and Table 4 for the three parameter model.

Table 3	location			
Repetitions	0	20	40	60
Improved	1111.954	1129.368	1133.003	1139.156
Standard	1111.473	1118.298	1120.145	1121.647

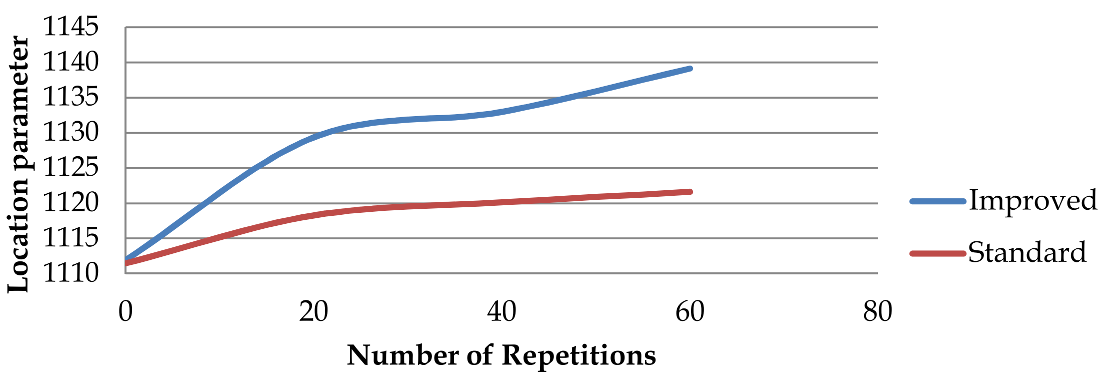


Figure.1:Locatition parameter and Repetitions

In Table 3 and figure 1,both locations were improving as the number of repetitions increased but the improved model indicated a higher increase in the location as the number of repetitions increases. As compared to standard model.

Table4:scale

Repetitions	0	20	40	60
Improved Standard	4.097801 9.993496	7.830554 9.99465	5.119941 2.056949	4.395648 3.373637

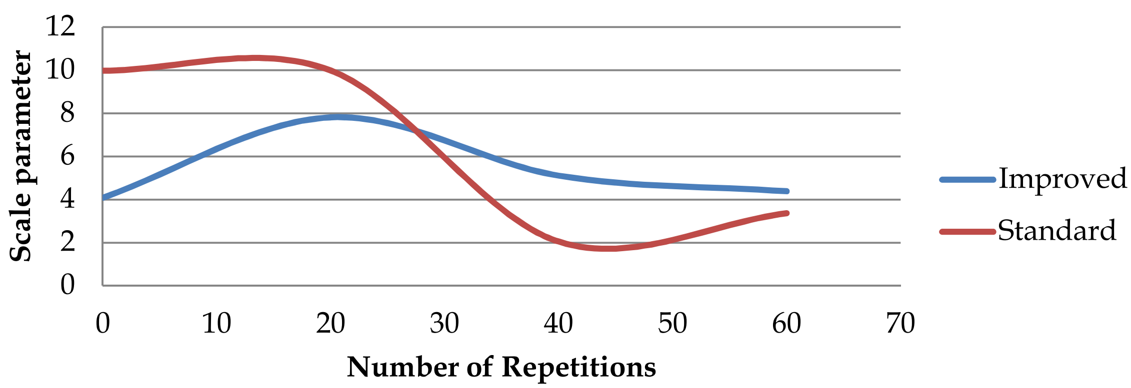


Figure. 2;Scale parameter and Repetitions

In Table 4 and figure 2 , the scale of the standard model indicates a decrease up to the 40th repetition and then shows some increase.

The three parameter MPS model exhibited the behavior shown in figure 3, figure 4 and figure 5

Table5	location			
	0	20	40	60

Improved	1111.954	1129.3679	1133.003	1139.156
Standard	1111.473	1118.29819	1120.145	1121.647

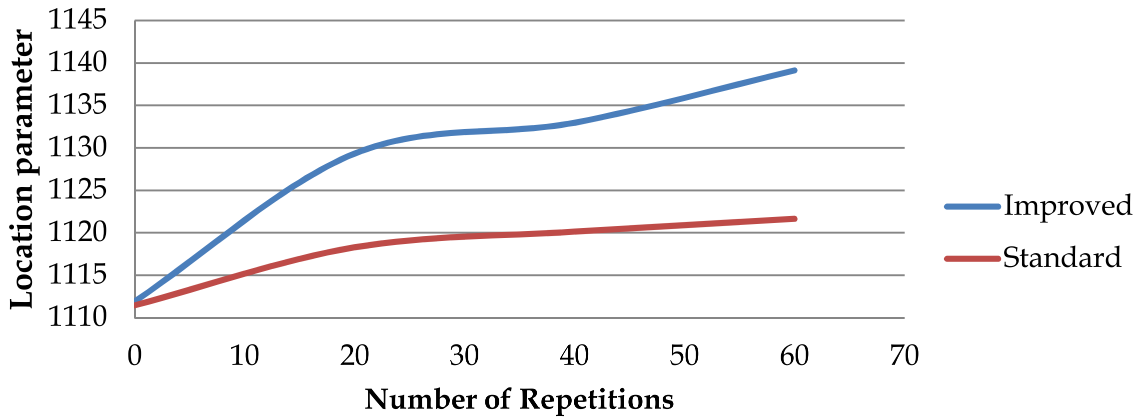


Figure.3 :Location parameter and Repetitions

Both location parameters showed an improvement as repetitions increases but the improved MPS model in Table 5 and figure 3 indicated a higher increase. Meaning that it had a more optimal threshold than the standard model.

Table 6	scale			
	0	20	40	60
Improved	15.4904	2.0094715	8.335507	21.42994
Standard	17.44149	12.0415209	4.716338	3.950339

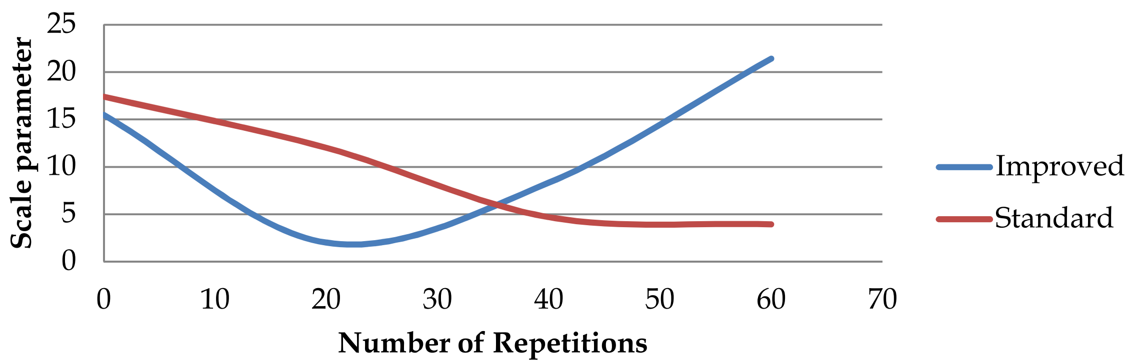


Figure.4:Scale parameter and Repetitions

The scale parameter of the standard model in Table 6 and figure 4 indicate a decrease all through until the 40th repetition when it indicate some stability. The improved model initially indicated a decrease up to the 20th repetition where it started increasing as the repetitions increased. The two models had the same scale at the 35th repetition.

Table7	shape			
	0	20	40	60
Improved	0.05365	1.586433	5.54892	-4.74385
Standard	-0.05918	-0.0409754	-1.90142	-6.59819

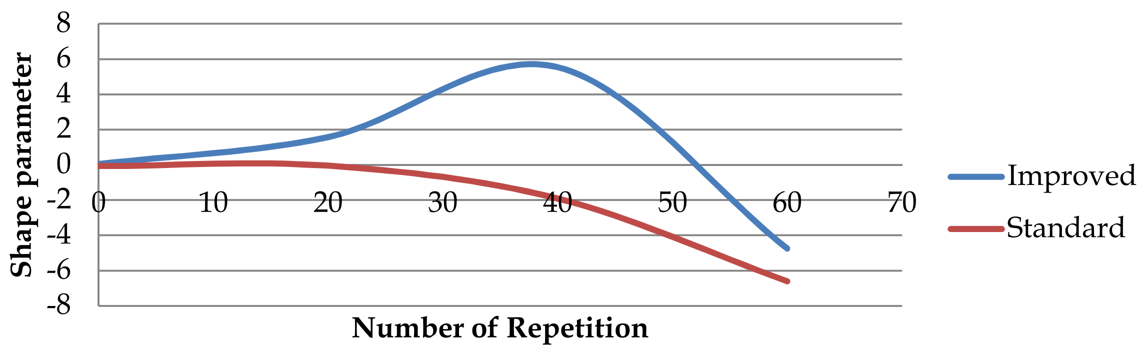


Figure.5:Shape parameter and Repetitions

In Table 7 and figure 5, at 0 repetition, both MPS model had the same shape . The standard model then showed a decrease of the size of the shape as repetitions increased. The improved model indicated a slow upward trend up to the 40th repetition after which it indicated a downward trend.

A real market data was used to investigate whether the data we simulated has some similar behavior with market data. A data was sought from NSE in the sector of Investment service called Nairobi securities exchange limited. The density of this data was plotted and compared to that of Gamma distribution. The data was then analyzed through both the standard MPS and the improved MPS model. Both two parameter and three parameter models were used.

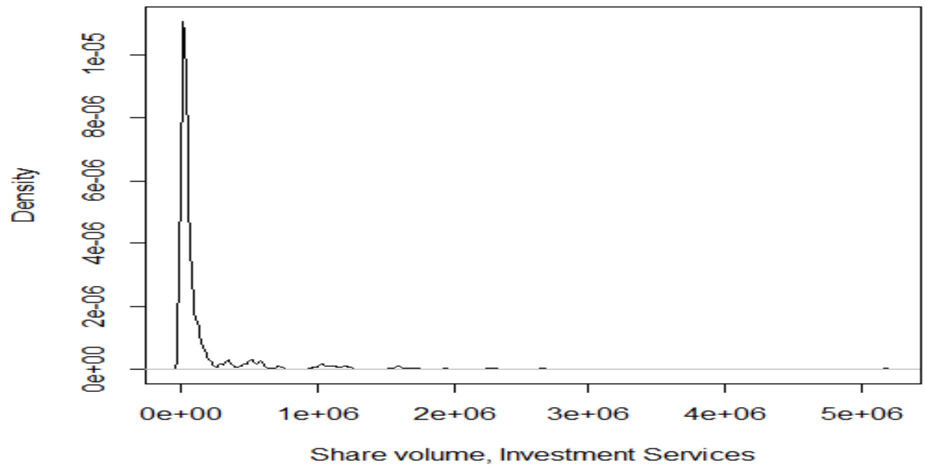


Figure. 6: Density of investment service

Figure 6 shows that there were many trading volumes concentrated between 0 and about 400000 However there were several other values that were extreme. This figure is very much similar to figure 1. Meaning that real data are skewed and sometimes contains heavy tails.

Table 8 :MLE and MPLE Estimates-NSE

	Two parameter		Three parameter	
Estimate	Standard	Improved	Standard	Improved
Threshold	573942.1463	1025881.322	573933.6022	1025908.743
Number above threshold	30	21	30	21
Proportion above	0.0508	0.0515	0.0508	0.0515
Scale Estimate	8.14E+05	6.06E+05	8.14E+05	6.06E+05
Scale standard error	5933.39	2.07E+05	2.12E+05	2.07E+05
Shape Estimate	7.76E-03	2.24E-01	2.09E-02	2.24E-01
Shape standard error	0.108	2.67E-01	1.86E-01	2.67E-01
Asymptotic var cov for scale	3.52E+07	4.28E+10	4.51E+10	4.28E+10
Asymptotic var cov for shape	1.17E-02	7.14E-02	3.47E-02	7.14E-02
Deviance	876.5707	599.1953	876.5713	5.99E+02
Penalized Deviance	8.77E+02	599.668	876.5999	599.6657

AIC	880.5707	603.1953	8.81E+02	603.1929
Penalized AIC	603.1953	603.668	880.5999	603.6657

From Table 8, the threshold for the two parameter standard model was 573942.146 while that of the improved model was 1025881.322. The improved model was better. The three parameter standard model had a threshold of 573933.602 while the improved model had a threshold of 1025908.743. The threshold in the improved MPS model is better than that of the standard MPS model. The Proportions above these threshold indicates that the cases of standard models had more exceedances than the improved model. The AIC criterion used to select the competing models indicates that the improved model was better than the standard model in both cases of two and three parameter model.

Conclusion

The results of this study indicate that The improved MPS model performed well in determining a more optimal threshold than the standard MPS model. When the threshold improves, it is expected that the values above this threshold will reduce. This will in turn impact on the scale parameter and shape parameter. Depending on how many the exceedances are, the scale may decrease or increase. The case of standard model is to drop the values that exhibit ties and only one value is left among the ones that had tied. The improved model does not drop of the values even if they have a tie. It takes care of them through the frequencies. In both models, the sample size will reduce but in case of improved model, all the values will still be taken care of by frequencies. The case of standard model drops the values out of the sample otherwise, it would fail. Many practical situations particularly in trading sector, exhibits the situations of ties. Therefore, in order to model such a data, the improved model is the best tool to help in preparing for full impact of any extreme event. This study therefore attains its objective by improving MPS methodology so that it takes care of all situations whether they contain ties or not. Further study need to be done to establish the actual at about 35th repetition, there was a drop in the threshold and that at this point both scale and the shape had a changing behavior.

References

Balkema, A. (1974). *Residual Life Time at a great age.* London: JMASM.

Beirlant, J. T. (1996). Excess Function and Estimation of the extreme values index. *Bernoulli*, 293-318.

Box, G. a. (1951). Experimental Attainment of optimum conditions. *J.R Statistic Society*, 1-45.

Butterfield, R. (2009). Dfid Economic Impacts of climate change in Kenya,Rwanda and Burundi. *ICPAC*, 1-45.

Cheng, R. a. (1983). Estimating parameters in a continuous univariate distribution with a shifted origin. *Royal Statistical Society*, 394-403.

Coles, S. (2001). *Introduction to Statistical modeling of extreme values.* London: Springer-Verlag.

Embretchet, P. M. (1997). *Modelling extremal events for insurance and Finance.* New York: Springer.

Fisher, R. a. (1928). *Limiting forms of the frequency distributions of the largest and smallest member.* Cambridge: Cambridge Philosophical Society.

Geremy, P. V. (2012). Extreme rainfall in west Africa region modeling. *American Geographical Union*, 1029.

Hill, B. (1975). Asimple general approach to inference about the tail of a distribution. *Annals of Statistics*, 341.

Pickands, J. (1975). Statistical Inference Using Order Statistics. *Annals of Statistics*, 3:119.

Prudhome, C. D. (1999). Mapping extreme rainafall inmountanious region using geostatistcal techniques. *Jpornal of Climatology*.

Yuejian, Z. a. (2002). *Extreme weather events and their probabilisticprediction.* Washngton DC: