# MODELLING EXTREME CLAIMSUSING GENERALISED PARETO DISTRIBUTIONS FAMILY IN AN INSURANCE COMPANY

**REG NO I56/80543/2012**

**A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT FOR THE REQUIREMENTS OF THE MASTERS OF SCIENCE IN ACTUARIAL SCIENCE IN THE SCHOOL OF MATHEMATICS, IN UNIVERSITY**

**JUNE 2015**

# DECLARATION

I, the undersigned,declare that this is my original work and has not been submitted to any other college, institution or University other than the University of Nairobi for academic credit.

Signature:……………………………………. Date:…………………………

ROBERT JOACKIM KAROBIA

I56/80543/2012

This project has been submitted for review with my approval as the university supervisor.

Signature:…………………………………… Date………………………..

Prof.J.A.M OTIENO

# ACKNOWLEDGEMENT

I would like to thank God for seeing me through my Masters programme.

The success of this work was as a result of the valuable help of various persons.First I would like to extend my sincere gratitude to my able supervisor Prof J.A.M Ottieno assisted by RachelSarguta.Their supervision and through analysis of my work widened the scope of understanding.I also want to extend my appreciation to the school of mathematics, University): to all the lectures who taught me throughout my course work.It is because of the knowledge they impacted in me that I was able to complete this work.

Special thanks goes to my dear mother Mary Wambui Anthony for her continued support and immeasurable dedication to this work. Her words of encouragement, moral, financial and physical support ensured completion of this work.

I am also indebted to my sibling Peter Maina, Dominic Wahome, Teresa Mumbi, Sarah Muthui,

and Adamson Mbugua for their contribution, love and support throughout the study.

I would like to acknowledge Samwel Wambui who tirelessly assisted me with the secondary data I needed for this project.

Lastly I would like to thank all my friends, classmates, collegues and relatives who have made this work success.

# ABSTRACT

The project focuses on modeling and estimating loss claims from  Insurance company.

Generalized Pareto distributions family was used and compared to come with a best fitting distribution. These distributions include Exponential, Pareto and Uniform distribution.

In the methodology the project shows how to develop the distributions from one distribution. Three methods for estimating parameters of the distributions were used i.e .the maximum likelihood method, the method of moments and L moment method. Properties of each distribution were shown. Then measures of risk were derived.

In application, Using the three methods of estimations to come up with the best fitting distribution. I'll also compare the three methods of estimations and come up with the best method. In addition .Using the best two distributions I'll plot histogram and QQ plot to come up with the best distribution. Thereafter I will estimate the confidence intervals of the chosen distribution parameter estimate using the bootstrap method.

# Contents

# CHAPTER ONE
## GENERAL INTRODUCTION

### 1.1 BACKGROUND INFORMATION

Extreme events are occurrences which are rare, high in magnitude and lead to huge losses. Extreme event risk affects all aspect of risk assessment, modeling and management especially in the context of credit market, insurance market, and operational market. These extreme events are either naturally occurring or man-made inflicted at the time they are least expected. If expected to happen there is very little or nothing can be done on their prevention. Some of these extreme risks are insured under general insurance policies. In non-life insurance a few large claims hitting a portfolio usually represent the most part of the indemnities paid out by the company. In fact 10% of extreme claims paid out represent the largest share of the paid funds. This is equivalent to significant percentage on the performances of companies.

Examples of such events include; The terrorist activities in Kenya like 1998 American embassy bomb blast; the Westgate terrorist attack (2014)wheremany lives and millions of money were lost;2007 post-election violence where a lot of property was destroyed; Kenya airways plane crash in Douala Cameroon; J.K.I.A fire 2014. Others include; The terrorist attack on the world trade center on 11 September 2001; The Hurricane Andrew earthquake in 1992 which cost USD 16billion on insurance payout; Ebola outbreak in West Africa(2014) and Earthquakes in Japan(2014); Gikomba market fire breakout(2005)

### 1.2 EXTREME VALUE THEORY

Extreme value theory (EVT) is a statistical methodology that is well suited for assessing, modeling and managing such catastrophic events. EVT offers a framework for assessing the uncertainties' inherent to describing the loss or claims distributions associated with these rare or extreme events.

Extreme value theory provides a good tool for handling these extremes. Broadly speaking there are two kinds of models for extreme values. The oldest group of models is the block maxima models. These are the models for the largest observations collected from large samples of identical distributed observation, for example. If we record daily or hourly losses and profits from trading a particular instrument or group of instrument. The block maxima method provides a model which may be appropriate for the quarterly or annual maximum of such values.

A more modern group of models are the peaks over threshold (POT). 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 their more efficient use of data in extreme values. Within the POT class of models one may further distinguish two styles of analysis. There are the semi parametric models built around the Hill estimator and the fully parametric models based on the generalized Pareto distribution (GPD). All these approaches are theoretically justified and empirically useful when used correctly. However all these models that have been used before are not working hence the need to look for a better alternative which is easy to use and error free .We therefore favor the latter style of analysis for reasons of simplicity both of exposition and implementation. GPD obtains simple parametric formulae for measures of extreme risk for which is relatively easy to give estimate of statistical error using the technique of maximum likelihood.

Traditional statistical methods focus on the probability laws governing averages of sums, whereas EVT concerns the properties of the largest observation in a sample and the probability laws governing these extremal values.

## 1.3 PROBLEM STATEMENT

Over the last 25 years there have been an increasing large number of extreme events in the financial and insurance market in Kenya leading to huge losses and claims. These extreme events affect the day to day operation of the individuals or company, and hence the economy of the country at large making it unable to achieve its core business function. If such events are insured one indemnification can lead to an insurance company going under-receivership if not properly reinsured. It can also lead to winding up a company if no business insurance was done. Therefore there is need to study those extreme risk and advice the insurance companies on how to cushion them in case of a risk covered happening.

## 1.4 OBJECTIVES OF STUDY

### 1.4.1 MAIN OBJECTIVE

The main objective is to model extreme claims in Insurance Company.

### 1.4.2 SPECIFIC OBJECTIVES

To fit fire claims data using the family of generalized Pareto distribution.

To estimate the measures of risk.

## 1.5JUSTIFICATION OF THE STUDY

This information enables the company make important decisions regarding premium rates and premium loading which it is supposed to charge policyholders to ensure high profitability.

The study is of interest to the actuaries since they help in the pricing of reinsurance agreement such as excess of loss reinsurance treaties under which the insurer has to identify to contact re insurer for all his expenditure associated with given claim as soon as that expenditure exceeds a fixed limit

The study is also geared to enlighten the insurance companies on the maximum amount of reserve that should be held to cover the cost of claims.

# CHAPTER TWO
# LITERATURE REVIEW

## 2.1 INTRODUCTION

This chapter reviews thematically the relevant literature regarding the different ways of modelling extreme valuerisks .Extreme value theory has become a key point in the performances of most financial institutions. Many researchers have developed different models to fit different kinds of claims .The most studied and applied being the generalized Pareto distribution.

### 2.2 MODELS PREVOUSLY USED

Hogg and Klugman(1984) focused on fitting the size of loss distribution to the data. They used a truncated Pareto distribution to fit the loss function. However Boyd (1988) argued that they seriously underestimated the tail region of the fitted loss distribution. Hogg and klugman compared two methods of estimation namely maximum likelihood estimation and methods of moments. They proved that the generalized Pareto distribution is better for measuring loss severity.

Rootzen and Tajvidi(1999)used extreme value statistics to estimate loss due to windstorm. They used Swedish insurance group data Lansforsakring during the period 1928-1993.They described uncertainty by presenting several different quintiles of distribution of maximum loss over several different periods .They computed the distribution of excess loss, the conditional distribution of loss given it exceeds the upper reinsurance limit. They concluded that statistical extreme value theory provides a flexible and theoretically well motivated approach to study large losses.

WO-Chiang Lee (2012) focused on modeling and estimating tail parameters of loss distribution from Taiwanese commercial fire loss severity .Using extreme value theory he employed the Generalized Pareto distribution and compared it with standard parametric modeling based on Lognormal, Exponential, Gamma and Weibull distributions. In an empirical study he determined the threshold of the GPD using mean excess plots and hill plots. Kolmogorov –Smirnov and likelihood ratio goodness of fit are conducted, value at risk and expected shortfall are calculated. He also constructed confidence intervals for the estimates using bootstrap method .This study set the reference point in my research.

Li-Hua Lai and Pei-Hsuan Wu(2005 estimating the threshold value and loss distribution on rice damaged by typhoons in Taiwan. They applied extreme value theory to determine the threshold peaks of the data and then used the Kolmogorv-Smirnov and Anderson-Darling goodness of fit tests to show that the Generalized Pareto distribution fits the heavy-tailed distribution better than the Lognormal, Gamma, Weibull and Normal distributions in rice damaged by typhoons. The appropriate of the threshold value and probable maximum loss was calculated as one of reference indexes on risk retention and crop insurance associated with the natural systematic risk of major agricultural disasters. The properties were found to be useful in crop loss assessment and in the decision making of government's risk financing for major agricultural disasters. They concluded that the method can be applied to other disasters in other countries.

Rodney Coleman (2002) looked at probability models appropriate for modeling extreme losses. The methodology was applied to extreme banking losses. He used the hill plot to estimate the mean excess estimate. He used data for 12 monthly maxima of the losses from fraud during the 1995 fraud in UK retail bank.

McNeil (1996) studied the tails of loss severity distribution using extreme value theory. He used the Danish data of 2157 claims. He described parametric curve fitting methods for modeling extreme historic losses. This method revolves around the generalized Pareto distribution and is supported by extreme value theory. He concluded the Picklands –Balkema –de Haantheorem (Balkema& de Hann 1974 Picklands 1975).which essentially says that for a wide class of distributions losses which exceed high enough threshold follow the generalized Pareto distribution.

Beatriz Vaz de Medo (2005)modeled extreme losses from an excess of loss reinsurance contract assuming there exist catastrophic process generating sequences of large claims. He arranged the claims in clusters. The number of clusters was modeled using the usual probability discrete probability and the severity of the sum of excess within clusters is modeled using a flexible extension of a generalized Paretodistribution. He studied this using the Danish fire insurance claims data set. Maximum likelihood estimates and bootstrap confidence interval were obtained for the parameter and statistical premium. He concluded that this cluster approach may provide a better fit for the extreme tail of the annual excess loss amount when compared to classical models of the risk theory.

McNeil and Saladin (1997) looked at peaks over threshold method to estimating high quartiles of loss distribution. They reviewed the peaks over threshold method of modeling tails of loss severity distributions and discussed the use of this technique for estimating high quantiles and the possible relevance of this to excess of loss insurance in high layers. They tested the method on a variety of simulated heavy-tailed distributions to show what kind of thresholds is required and what sample sizes are necessary to give accurate estimates of quantiles. They used Pareto Log-gamma, Studen- t and Log-normal distributions.

Cebrian, Denuitand  Lambert (2003)looked at Generalized Pareto fit  to the society of actuaries large claims database. They discussed a statistical modeling strategy based on extreme value theory to describe the behavior of an insurance portfolio, with particular emphasis on large claims. The strategy was illustrated using the 1991-92 group medical claims database maintained by the Society of Actuaries. Using extreme value theory, the modeling strategy focused on the “excesses over threshold" approach to fit Generalized Pareto distributions. The proposed strategy was compared to standard parametric modeling based on Gamma, Log-normal and Log-gamma distributions. They concluded that Extreme value theory outperforms classical parametric fits and allows the actuary to easily estimate high quantiles and the probable maximum loss from the data.

Leonardo (2002) looked at Bayesian analysis of extreme values by use of mixture modeling. He modeled extreme values in the presence of heterogeneity. He considered losses for several related categories for each category he viewed exceedances over a given threshold as generated by Poisson process whose intensity is regulated by a specific location, shape and scale parameter using Bayesian approach. He developed a hierarchical mixture priori with unknown number of components for each of the above parameters. He performed the computation using the reversible jump MCMC.This model accounted for possible clustering effects and take advantage of the similarities across categories' both for estimation and prediction. The method illustrated throughout using a data set on large claims against a well-known insurance company for 15 yrs.

Hosking and Wallis (1987)looked at parameter and quantile estimation for the generalized Pareto distribution. They used a two parameter Pareto distribution which contains Exponential, Uniform and Pareto distribution. They estimated the parameter using three methods of estimation i.e,Maximum likelihood method, Method of moment and Probability weighted moment.They showed that unless the sample size is 500 or moreestimators derived by the method of moment or the method of probability weighted moments are more reliable.They used computer simulation to assess the accuracy of confidence intervalfor the parameter.Their main conclusion was that maximum likelihood estimation although asymptotically the most efficient method does not clearly display efficiency even in samples as large as 500and that the method of moment is generally reliable except when $k<-2$, and that PWM may be recommended if it seems likely that $k<0$.Where k is the scale parameter.They finally applied the results to estimation of extreme floodsusing an example of a series of flood peaks for river Nidd at Hunsingore England..

Eric Brodin (2005) studied the univariate and bivariate GPD method of predicting windstorm losses. He aimed to select, tailor and develop extreme value methods for use in windstorm insurance. Themethod was applied to the 1982-2005 losses for the largest Swedish insurance company the Lansforsakrigargroup. Both a univariate and a bivariate generalizedPareto distribution gave models which fitted the data well. TheBivariate model led to lower estimates of risk except for extreme cases but taking statistical uncertainty into account the two models lead to qualitative similar results, He concluded that the bivariate model provided the most realistic picture in real uncertainities. It additionally made it possible to explore the effects of changes in the insurance portfolio

MeelisKaarik (2012) studied the estimation of loss distribution and risk measures.The research used extreme value theory and Generalized Pareto distribution in relation to heavy tailed data. He studied third party liability claims from EstonTraffic Insurance Fund.The fitting consisted of two parts fitting a Lognormal distribution and a Generalized Pareto distributionwhich was used for the tail resulting in a certain composite Lognormal/Generalized Pareto model.He emphasized on the proper threshold selection.He sought for stability of parameter estimates and study the behavior of risk measures at wide range of thresholds. He finally provided an alternative threshold selection method which is based on the risk measures (quantiles)and recommended it to be used in insurance data analysis.

Ramchadran(2010) presented the generalized extreme value technique for making use of all large losses. He investigated the problem of assessing the relative contributions of various factors to fire losses. He did multiple regressions with extreme observations of given ranks. He took into consideration the biases due to use of extremes and the difference between catergory of risk in regard to the frequency of fire claims. He illustrated this using the largest and second largest losses in textile industries in the United Kingdom during the six year period of 1965-1970.The technique enables different estimates to be obtained for each regression parameter for different ranks. He developed a second model for performing a regression analysis combining observations pertaining to a number of ranks covariance of the residual were also taken into account in this model.

Patricia de Zea Bermudez(2011)looked at importance of Generalized Pareto distribution in the extreme value theory. Patricia used EVT on Forest Fires that occurred in Portugal between 1984-2004.Patricia reviewed the methods to estimate the parameter of Generalized Pareto distribution. She used Bayesian hierarchical models and Non-linear time series models.She went ahead and incorporated Uniform, Beta, Exponential and Pareto in analysis.To get the best fit she plotted the QQ plots.

Matthew J Pocernich(2002) looked at the application of extreme value theory and threshold models to Flood events. The aim of the paper was to quantify the probability of very high stream flows. The paper examined the use of extreme value statistics in predicting the probability of rare flood events. He used maximum likelihood estimate to fit parameter in a point process model .He also used L-moment method with annual maximum values. He therefore considered time as a covariate in order to determine if the magnitude and frequency of high flows events change with time. He applied this method in a data from Cherry Creek at Franktown,Colorado. It was found that the parameter estimate did not vary significantly with time.

Gonzalo and Olmo(2004)looked at establishing which extreme values are really extreme.They defined the extreme values of any random sample of size n from a distribution function F as the observations exceeding a threshold and following a type of Generalized Pareto distribution involving the tail index F.The threshold was defined as the order statistic that minimized distribution of the corresponding largest observation and the corresponding GPD. To formalize the definition they used semi parametric bootstrap to test the corresponding GPD approximation. Finally in the methodology they estimated the tail index and value at risk of some financial indexes of major markets. They compared data from Dax, Ftse, Ibex, Nikkie and Dow Jones.

Carl (2012) reviewed of the extreme value threshold estimation and uncertainty quantification. He argued that from a statistical perspective the threshold is loosely defined such that the population tail can be well approximated by an extreme value model (generalized Pareto distribution)obtaining a balance between the bias due to the asymptotic tail approximation and parameter approximation due to inherent sparcity of threshold

## 2.3 A SUMMARY OF LITERATURE REVIEW

Generalized Pareto distribution has been extensively been used with a combination of different estimation method depending on the event being analysed.This is because it easy to use .The Generalized Pareto distribution has also been compared with other parametric distributions like Lognormal, Gamma, Exponential, Student-t, Uniform and has resulted to be the best distribution in fitting extreme values. Previous studies have modeled data from fire insurance, floods, windstorms and medical claims.

Generalized Pareto distribution parameter has been estimated by a number of methods including Maximum likelihood method, Method of moment,L moments, Percentile weighted method, Empirical Bayesian method and Probability Weighted Method.

In this study we are going to apply Generalized Pareto distribution whose parameters has been estimated bythree methods namely maximum likelihood estimation method, Method of moment and Linear combination of moments (L-Moment method). It will also compare three distributions include, Pareto, Exponential and Uniform.

# CHAPTER THREE
# CONSTRUCTION AND PROPERTIES OF GPD

## 3.1 INTRODUCTION

In this chapter,We have constructed generalized Pareto distribution and its members using transformation technique, mixtures and mixed distribution technique.

## 3.2 DEFINITION AND MEAN OF GPD

The probability density function of a Generalized Pareto distribution with parameters (μ,δ,ξ) is defined as where ξ>0,x>0,

$$f(x) = \frac{1}{\delta}^{(1+\frac{\xi(x-\mu)}{\delta})^{\frac{-1-\xi}{\xi}}} \qquad \xi \neq 0, \xi = 0$$

And its cumulative distribution function is therefore

$$F(x) = \{1-\left(1+\frac{\xi}{\delta}(x-\mu)\right)^{\frac{-1-\xi}{\xi}} \qquad for\ \xi \neq 0\}$$

when μ = 0,we have $f(x) = \frac{(1+\frac{\xi x}{\delta})^{\frac{-1-\xi}{\xi}}}{\delta}$ $\xi \neq 0$

Mean of GPD $E(X) = \int_{\mu}^{\infty} xf(x)dx$

$$= \int_{\mu}^{\infty} x\frac{1}{\delta}[1+\frac{\xi(x-\mu)}{\delta}]^{-1-\frac{1}{\xi}}dx$$

$$\text{Let } z/\mu = 1+\frac{\xi(x-\mu)}{\delta}$$

$$x= \frac{\delta}{\xi}(\frac{z}{\mu} - 1) + \mu$$

$$= \int_u^\infty [\frac{\delta}{\xi}(\frac{z}{u} - 1) + \mu] \frac{1}{\delta}(z/\mu)^{\frac{-1-\xi}{\xi}} \frac{\delta}{\xi\mu} dz$$

$$= \int_u^\infty [\frac{\delta}{\xi}(\frac{z}{u} - 1) + \mu] z^{(-1-\xi)/\xi} \mu^{(1+\xi)/\xi} dz$$

$$= \int_u^\infty \frac{1}{\xi} \mu^{1/\xi} [\frac{\delta}{\xi u} z^{1/\xi} - \frac{\delta}{\xi} z^{\frac{-1-\xi}{\xi}} + \mu z^{(-1-\xi)/\xi}] dz$$

Opening the brackets and simplifying you get

$$= [\frac{\delta}{1-\xi} \mu^{(-1+1/\xi)} z^{(1-1/\xi)} + \frac{\delta}{\xi} \mu^{1/\xi} z^{-1/\xi} - z^{-1/\xi} \mu^{(1+1/\xi)}]$$

Therefore you substitute for z with u and ∞

$$= -\delta\mu^{(-1+1/\xi)} \mu^{(1-1/\xi)} - \frac{\delta}{\xi} \mu^{\frac{1}{\xi}} \mu^{-\frac{1}{\xi}} + \mu^{-1/\xi} \mu^{(1+1/\xi)}$$

$$= \frac{-\delta}{\xi(1-\xi)} - \frac{\delta}{\xi} + \mu$$

$$= \mu + \frac{\delta - \delta(1-\xi)}{\xi(1-\xi)}$$

$$= \mu + \frac{\delta}{1-\xi}$$

## 3.3 CONSTRUCTING GENERALIZED PARETO DISTRIBUTION

### 3.3.1 USINGTRANSFORMATION METHOD

Uniform distribution can be transformed to a generalized Pareto distribution by the change of variable technique.

If U is uniformly distributed on [0, 1] then

$$X = \mu + \frac{\delta}{\xi}(u^{-\xi} - 1)$$

Solving for the distribution using the change of variable technique we have

$$X - \mu = \frac{\delta}{\xi}(u^{-\xi} - 1)$$

$$\xi( x - \mu) = \delta\ (u^{-\xi} - 1)$$

$$\frac{\xi(x-\mu)}{\delta} = (u^{-\xi} - 1)$$

$$\frac{\xi(x-\mu)}{\delta} + 1 = u^{-\xi}$$

$$U = \frac{(\xi(x-\mu)+1)^{\frac{-1}{\xi}}}{\delta}$$

$$\frac{du}{dx} = \frac{1}{\delta}\left(\frac{\xi(x-\mu)}{\delta} + 1\right)^{\frac{-1-\xi}{\xi}}$$

$$f(x) = 1|J|$$

$$f(x) = \frac{1}{\delta}\left(\frac{\xi(x-\mu)}{\delta} + 1\right)^{\frac{-1-\xi}{\xi}}$$

This is the Generalized Pareto Distribution function.

### 3.3.2 CONSTRUCTION OF GPD BASED ON INVERSE METHOD

F(x) = Prob[ X≤ x]

$= \text{Prob}\{\mu + \frac{\delta(u^{-\xi}-1)}{\xi} \leq x\}$

$= \text{Prob}\{\frac{\delta(u^{-\xi}-1)}{\xi} \leq x-\mu\}$

$= \text{Prob}\{u^{-\xi} - 1 \leq \frac{\xi(x-\mu)}{\delta}\}$

$= \text{Prob}\{u^{-\xi} \leq 1+\frac{\xi(x-\mu)}{\delta}\}$

$= \text{Prob}\{\frac{1}{u^{\xi}} \leq [1 + \frac{\xi(x-\mu)}{\delta}]\}$

$= \text{Prob}\{\frac{1}{1+\frac{\xi(x-\mu)}{\delta}} \leq u^{\xi}\}$

$=\text{Prob}\{u^{\xi} \geq (1 + \frac{\xi(x-\mu)}{\delta})^{-1}\}$

$=\text{Prob}\ \{u \geq [1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}}\}$

$F(x) = 1- \text{Prob}\{u \leq [1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}}\}$

$f(x) = 1- G[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}}$

Therefore $f(x) = \frac{1}{\xi}\{G'[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}}[[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1-\xi}{\xi}}]$

$f(x) = \frac{1}{\delta}\{G'[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}}\}[[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1-\xi}{\xi}}]$

But $G'[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1}{\xi}} =1$ Uniform (0,1)

Therefore $f(x) = \frac{1}{\delta}[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1-\xi}{\xi}}$

From $x = \mu + \frac{\delta(u^{-\xi}-1)}{\xi}$

u= 0   x = ∞       u= 1 , x = μ

$$f(x) = \frac{1}{\delta}[1 + \frac{\xi(x-\mu)}{\delta}]^{\frac{-1-\xi}{\xi}} \qquad \text{for } \mu \le x < \infty \quad \xi \neq 0$$

### 3.3.3 CONSTRUCTING GPD BASED ON MIXTURES

We shall consider a gamma –gamma mixtures where the first gamma distribution is given by

$$f(x/ß) = \frac{ß^{\alpha}}{\Gamma\alpha} e^{-ßx} x^{\alpha-1} \qquad x>0,\ \alpha>0,\ ß>0$$

which is the conditional distribution.

The second gamma which is the mixing distribution is given by

$$g(ß) = \frac{\lambda^{v}}{\Gamma v} e^{-\lambda ß} ß^{v-1} \qquad ß>0,\ \lambda>0, v>0$$

Thus the mixed distribution is

$$F(x) = \int_0^{\infty} f(x/ß) g(ß) dß$$

$$= \int_0^{\infty} \frac{ß^{\alpha}}{\Gamma\alpha} e^{-ßx} x^{\alpha-1} \frac{\lambda^{v}}{\Gamma v} e^{-\lambda ß} ß^{v-1} dß$$

$$= \frac{x^{\alpha-1}\lambda^{v}}{\Gamma\alpha\,\Gamma v} \int_0^{\infty} ß^{\alpha+v-1} e^{-(x+\lambda)ß} dß$$

$$= \frac{x^{\alpha-1}\lambda^{v}}{\Gamma\alpha\,\Gamma v(x+\lambda)^{\alpha+v}} \Gamma(\alpha+v)$$

$$= \frac{\Gamma(\alpha+v)}{\Gamma\alpha\,\Gamma v} \frac{\lambda^{v} x^{\alpha-1}}{(x+\lambda)^{\alpha+v}} \qquad x>0,\ \lambda>0,\ \alpha>0$$

When α = 1 we get

$$F(x) = \frac{v\lambda^{v}}{(x+\lambda)^{v+1}} \qquad x>0,>0,v>0$$

Which is Pareto ii( Lomax) distribution

Further put $v = \frac{1}{\xi}$ and $\lambda = \frac{\delta}{\xi}$

Then $f(x) = \frac{\frac{1}{\xi}(\frac{\delta}{\xi})^{\frac{1}{\xi}}}{(x+\frac{\delta}{\xi})^{(\frac{1+\xi}{\xi})}}$

$$= \frac{\frac{1}{\xi}\frac{(\delta)}{\xi}^{\frac{1+\xi}{\xi}}}{\frac{\delta}{\xi}(x+\frac{\delta}{\xi})^{\frac{1+\xi}{\xi}}}$$

$$= \frac{\frac{1}{\xi}}{\frac{\delta}{\xi}(1+\frac{\xi x}{\delta})^{1+\frac{1}{\xi}}}$$

$$= \frac{1}{\delta(1+\frac{\xi x}{\delta})^{1+\frac{1}{\xi}}}$$

Replace x by x-μ we have

$$f(x) = \frac{1}{\delta(1+\frac{\xi x}{\delta})^{1+\frac{1}{\xi}}} \qquad \text{for } x>\mu, \delta>0, \xi>0$$

Which is GPD (μ,δ,ξ)

## 3.3.4 OTHER APROACH USING MIXTURES

We consider the case when k<0, δ>0 and x>0

Proposition

Castillo (1997) let $x_\Theta \Longmapsto$ G(x, Θ) = 1- $e^{\frac{\Theta k x}{\delta}}$ with x>0, δ>0, k<0 and let Θ>0.a sample of the random variable Θ ⟾ Gamma (0,1,-$\frac{1}{k}$)

Then the random variable x obtained from mixing the two random variable $x_\Theta$ and Θ is GPD (K, δ) random variable.

PROOF

$x_\Theta \Longmapsto 1\text{-}e^{\frac{\Theta k x}{\delta}}$ $\qquad \Theta \rightarrow \frac{1}{\Gamma(-\frac{1}{K})} e^{-\Theta}\Theta^{\frac{-1-k}{K}}$

We calculate the cumulative distribution function of the random variable x

$$F(x) = P(X<x) = \int_0^{\infty}(1 - e^{\frac{\Theta kx}{\delta}}) \frac{1}{\Gamma(-\frac{1}{K})} e^{-\Theta}\Theta^{\frac{-1-k}{K}} \, d\Theta$$

$$= \frac{1}{\Gamma(\frac{-1}{k})}\int_0^{\infty} e^{-\Theta}\, \Theta^{-1/k-1} \, d\Theta - \frac{1}{\Gamma(-1/k)}\int_0^{\infty} e^{-\Theta}\left(1 - \frac{kx}{\delta}\right)\left[\Theta\left(1 - \frac{kx}{\delta}\right)\right]^{-1/k-1} d\left[\Theta(1 - \frac{kx}{\delta})\right]$$

$$= \frac{1}{\Gamma(-1/k)}\int_0^{\infty} e^{-\Theta}\, \Theta^{-1/k-1} \, d\Theta - \frac{1}{\Gamma(-1/k)}\int_0^{\infty} e^{-\Theta}\left(1 - \frac{kx}{\delta}\right)\left[\Theta\left(1 - \frac{kx}{\delta}\right)\right]^{-1/k-1} d\left[\Theta(1 - \frac{kx}{\delta})\right]$$

$$\frac{1}{\Gamma(-1/k)}\int_0^{\infty} e^{-\Theta}\left(1 - \frac{kx}{\delta}\right)\left[\Theta\left(1 - \frac{kx}{\delta}\right)\right]^{-1/k-1} d\left[\Theta(1 - \frac{kx}{\delta})\right] = (1 - \frac{kx}{\delta})^{\frac{-1}{k}}$$

$$= (1 - \frac{kx}{\delta})^{\frac{-1}{k}} \cdot \frac{1}{\Gamma(-1/k)}\int_0^{\infty} e^{-\Theta}\left(1 - \frac{kx}{\delta}\right)\left[\Theta\left(1 - \frac{kx}{\delta}\right)\right]^{-1/k-1} d\left[\Theta(1 - \frac{kx}{\delta})\right]$$

$$=(1 - \frac{kx}{\delta})^{\frac{-1}{k}}$$

## 3.4SPECIAL CASES OF GPD

### 3.4.1 TRANFORMED PARETO DISTRIBUTION

A two parameter Pareto distribution with the shape parameter ξ and the scale parameter δ, denoted by GPD (ξ , δ) is the distribution of the random variable where

$$X = \frac{\delta}{\xi}(1- e^{-\xi y})$$

Where y is a random variable with the standard exponential distribution.

The GPD (ξ, δ) has the following distribution.

$$F_{\xi,\delta}(X) = \begin{cases} 1 - (1 - \frac{\xi}{\delta}x)^{\frac{1}{\xi}} & \delta > 0 \\ 1 - exp^{\left(\frac{-x}{\delta}\right)} & \end{cases}$$

1 When ξ = 0 the GPD (ξ, δ) reduces to an exponential distribution with mean Exp (δ)

2 When ξ=1 the GPD (ξ, δ) becomes uniform U (0, δ)

3 When ξ<0 the GPD (ξ, δ) reduces to a Pareto (ξ, a, c) distribution of the second kind.

Which imply that a family of generalized Pareto distribution comprises of Uniform, Exponential and Pareto distribution which needs to be studied and compared into detail in this project

PROOF

$X = \frac{\delta}{\xi}(1 - e^{\xi y})$ which is a GPD random variable.

$F(x) = p\,(X<x)$

$=P\,(\frac{\delta}{\xi}(1 - e^{-\xi y}) < x)$

$=p\,(1 – e^{-\xi y}) < \frac{\xi x}{\delta})$

$=P\,(1 - e^{-\xi y} < \xi\frac{x}{\delta})$

$=P\,(1- \frac{\xi x}{\delta} < e^{-\xi y})$

$=P\,(\ln(1- \frac{\xi x}{\delta}) < -\xi y)$

$=P\,(y < \frac{-\ln(1-\frac{\xi x}{\delta})}{\xi})$

$=P\,(y < -\ln(1-\frac{\xi x}{\delta})^{\frac{1}{\xi}}$

When $y \rightarrow$ Exp (1) it's $f(x) = 1 - e^{-x}$

Replacing for x with $-\ln(1-\frac{\xi x}{\delta})^{\frac{1}{\xi}}$ in f(x) above we get $=1- e^{\ln(1-\frac{\xi x}{\delta})^{\frac{1}{\xi}}}$

$= 1 - (1-\frac{\xi x}{\delta})^{\frac{1}{k}}$ For $x \in (0, \frac{\delta}{k})$

i. When ξ = 0 in the CDF of GPD $1-(1-\frac{\xi x}{\delta})^{\frac{1}{\xi}}$

The F(x) becomes $\lim_{x\to\infty} 1- \ (1+\frac{0}{\delta})^{\frac{1}{0}}$

$$= 1 - e^{\delta x}$$

Which is the CDF of an exponential distribution with mean of δ and a pdf is $\delta e^{-\delta x}$

ii. When ξ = 1 $\lim_{x\to\infty} 1-(1-\frac{x}{\delta})^{1}$

$=1-(1-\frac{x}{\delta})$

$=\frac{x}{\delta}$

Which is a uniform distribution with mean of $\frac{x}{\delta}$ and the limits are u [0, δ]

iii. When ξ< 0 it reduces to a Pareto distribution (k,a,c)distribution of the second kind.

$$1-(1--\frac{x}{\delta})^{-1} = \lim_{x\to\infty} 1-(1+\frac{x}{\delta})^{-1}$$

$$= 1-(1+\frac{x}{\delta})^{-1}$$

$$= 1-\frac{k}{(x+c)^{a}}$$

### 3.4.2 RELATION OF PARETO TO EXPONTIAL DISTRIBUTION.

If x is a Pareto distributed with parameters u and a then y =log($\frac{x}{u}$) is Exponential distributed with intensity a

Proof,

P(y<x)

P(y <log $\frac{x}{u}$)

P ($e^{y} < \frac{x}{u}$)

P ($\frac{x}{u} < u\, e^{y}$)

The F(x) of a Pareto Distribution with parameters u and a has CDF of 1 -$(\frac{u}{x})^{a}$

The substituting for x in the CDF of a Pareto we get

$= 1\text{-}(\frac{u}{uey})^{a}$

$=1 -(\frac{1}{ey})^{a}$

$=1 - e^{-ya}$

= Exp (a)

Which is an Exponential distribution withparameters a.

### 3.4.3 DERIVATION OF EXPONENTIAL TO PARETO DISTRIBUTION

If Y is an Exponential distribution with intensity a then $X=ue^{y}$ is Pareto distribution.

Using the change of variable technique.

$$=P(X<x)$$

$$=P(Y<y)$$

$$=P(X< ue^{y})$$

$$=P(X/U<e^{y})$$

$$=P(e^{y} < x/u)$$

$$=P(\ln e^{y} < \ln(x/u))$$

$$=P(y<\ln (x/u))$$

The CDF of exponential distribution is $1\text{-}e^{-ya}$ substituting it in our equation we get

$$=1\text{-}e^{-a(lnx/u)}$$

$$=1\text{-}e^{-}(\ln x/u)^{a}$$

$= 1\text{-}(x/u)^{a}$ which is a Pareto distribution with parameter (x, a)

## 3.5 PROPERTIES OF DISTRIBUTIONS

### 3.5.1 PARETO DISTRIBUTION

$$f(x) = \frac{\alpha\,\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}\boldsymbol{\lambda}>0,\alpha>0,x>0$$

Mean

$$E(X) = x\int_0^{\infty} f(x)dx$$

$$=\int_0^{\infty} \frac{x\alpha^{\alpha}}{(\alpha+x)^{\alpha+1}}\,dx$$

$= \int_0^{\infty} \alpha\lambda^{\alpha}\, x\, (\lambda + x)^{-(\alpha+1)}$ dx

=let **λ**+x= t , x= t –**λ**, dx = dt

$=\int_{\lambda}^{\infty} \alpha\lambda^{\alpha}$ (t-**λ**) $t^{-(\alpha=1)}$ dt

$=\int_{\lambda}^{\infty} \alpha\lambda^{\alpha}$ t, $-t^{-(\alpha+1)}$dt $-\int_{\lambda}^{\infty} \alpha\lambda + \lambda\, t^{-(\alpha+1)}$ dt

$=\int_{\lambda}^{\infty} \alpha\lambda^{\alpha}\, t^{-\alpha}$dt $-\int_{\lambda}^{\infty} \alpha\lambda^{\alpha+1}\ t^{-\alpha-1}$ dt

$=-\alpha\lambda^{\alpha}[\frac{t^{-\alpha}}{\alpha-1}] + \alpha\lambda^{\alpha+1}\, [\frac{t^{-\alpha}}{\alpha}]$

$=-\ \alpha\lambda^{\alpha}\, \frac{\infty^{-\alpha}}{\alpha-1} + \alpha\lambda^{\alpha+1}\, \frac{\infty^{-\alpha}}{\alpha-1} - [-\alpha\frac{\lambda^{\alpha}}{\alpha-1}\, \lambda^{-\alpha} + \alpha\frac{\lambda^{(\alpha+1-\alpha)}}{\alpha}\ ]$

$= \frac{\alpha\lambda}{\alpha-1} - \frac{\alpha\lambda}{1}$

$= \frac{\alpha\lambda - \lambda(\alpha-1)}{\alpha-1}$

$E(X) = \frac{\lambda}{\alpha-1}$

E ($X^2$) of Paretodistribution.

$E(X^2) = \int_0^{\infty} x^2\, \frac{\alpha\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}$ dx

Let t=**λ**+x dt=dx

X=t-h

$x^2 = (t - h)^2$

$= t^2 - 2t\lambda + \boldsymbol{\lambda^2}$

Replacing of $x^2$ in our intergrationwe have

$$=\int_{\lambda}^{\infty}(t^2 - 2t\lambda + \lambda^2)\alpha\lambda^{\alpha}[t]^{-(\alpha+1)}dt$$

$$=\int_{\lambda}^{\infty} \alpha\lambda^{\alpha}\, t^{-\alpha+1} - 2\alpha\lambda^{\alpha+1}\, t^{-\alpha} + \alpha\lambda^{\alpha+2} t^{-\alpha-1}dt$$

$$=-\alpha\lambda^{\alpha}[\frac{t^{-\alpha+2}}{\alpha-2}] + 2\alpha\lambda^{\alpha+1}\,[\frac{t^{-\alpha+1}}{\alpha-1}] + \lambda^{\alpha+2}\,[t^{-\alpha}]$$

$$=\frac{\alpha\lambda^2}{\alpha-2} -- \frac{2\alpha\lambda^2}{\alpha-2} + \frac{\lambda^2}{1} = \frac{\alpha\lambda^2(\alpha-1)-2\alpha\lambda^{2(\alpha-2)+\lambda^2(\alpha-2)(\alpha-1)}}{(\alpha-1)(\alpha-1)}$$

$$=\alpha^2\lambda^2 - \alpha\lambda^2 - 2\alpha\alpha^2\lambda^2 + \mathbf{4\alpha\lambda^2} + \boldsymbol{\lambda^2\alpha^2} - \mathbf{3\alpha\lambda^2} + \mathbf{2\lambda^2}$$

$$E\,(X^2) = \frac{2\lambda^2}{(\alpha-2)(\alpha-1)}$$

Variance (X) =E $(x^2)$ –$[E(x)]^2$

$$=\frac{\alpha\lambda^2}{(\alpha-2)(\alpha-1)} - \frac{\lambda^2}{(\alpha-1)^2}$$

$$=\frac{2\lambda^2(\alpha-1) - \lambda^2(\alpha-1)}{(\alpha-1)^2\,(\alpha-2)}$$

$$=\frac{2\lambda^2(\alpha-1) - \lambda^2(\alpha-2)}{(\alpha-1)^2\,(\alpha-1)}$$

Variance (X) = $\frac{\lambda^2\alpha}{(\alpha-1)^2\,(\alpha-1)}$

The CDF of Pareto distribution

$$F(X) = \int_0^m fx dx$$

$$\int_0^m \frac{\alpha\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}} = \int_0^m \alpha\lambda^{\alpha}\,(\lambda+1)^{-(\alpha+1)}\, dx$$

Let $\boldsymbol{\lambda}$+x = t $\int_0^m \lambda^{\alpha} t^{-\alpha-1}$ dx = $\alpha\lambda^{\alpha} \frac{(\lambda+x)^{-x}}{\alpha}$

$=\alpha(\frac{\lambda}{\lambda+x})\alpha$

$=1-(\frac{\lambda}{\lambda+1})$

**3.5.2 EXPONENTIAL DISTRIBUTION**

f(x)= $\lambda e^{-\lambda x}$ $x>0$

F(x) =$\int_0^x f(t)$dt

$=\int_0^x \lambda e^{-\lambda t}$ dt

$=\lambda \int_0^x e^{-\lambda t}$dt

$=-\boldsymbol{\lambda}\frac{\lambda^t}{\lambda}$dt

= $[e^{-\lambda x}$ -1]

$=1-e^{-\lambda x}$ The cdf of Exponential.

Mean E(X) = $\int_0^\infty xfx\, dx$

$=\int_0^\infty \lambda e^{-\lambda x}\ dx$

$=\boldsymbol{\lambda}\int_0^\infty xe^{-\lambda x}$ dx using intergration by partswe have$\int udvdx$ =u v –$\int v$ du/dx

Let u =x dv = $e^{-\lambda x}$

$\frac{dy}{dx}$ =1     v = $\frac{-e^{-\lambda x}}{\lambda}$

= [-x$\frac{e^{-\lambda x}}{\lambda}$ +$\int_0^\infty \frac{e^{-\lambda x}}{\lambda}$dx ]

= [-x$\frac{e^{-\lambda x}}{\lambda}$ +$\frac{-1e^{-\lambda x}}{\lambda}$ ]

= [-x$e^{-\lambda x - \frac{1e^{-\lambda x}}{\lambda}}$]

=-∞$e^{-\infty}$ -$\frac{-1e^{-\infty}}{\lambda}$

=0-[-$\frac{-1e^{0}}{\lambda}$]

=$\frac{1}{\lambda}$

E ($X^2$) =$\int_0^\infty x^2\, fx\; dx$

=$\int_0^\infty x^2 \boldsymbol{\lambda} e^{-\lambda x}$ dx     let u= $x^2$   dv = $e^{-\lambda x}$

du = 2x   v = $\frac{-e^{-\lambda x}}{\lambda}$

$\frac{udv}{dx}$ = u. v $\frac{vdu}{d\pi}$

=$\boldsymbol{\lambda}$[-$x^2 \frac{e^{-\lambda x}}{\lambda} - \int_0^\infty \frac{e^{-\lambda x}}{\lambda}\; 2xdx$]

=$\boldsymbol{\lambda}$[ -$x^2 \frac{e^{-\lambda x}}{\lambda} + \frac{2}{\lambda}\; x\; e^{-\lambda x}$]

$$=\boldsymbol{\lambda}[x^2\text{-}\frac{e^{-\lambda x}}{\lambda}+\frac{2}{\lambda}\ [-\frac{2e^{-\lambda x}}{\lambda}-\ \frac{1^{-\lambda x}}{\lambda}]]$$

$$=\boldsymbol{\lambda}\{-x^2\ \frac{e^{-\lambda x}}{\lambda}-\ \frac{2xe^{-\lambda x}}{\lambda^2}-\ \ \frac{2e^{-\lambda x}}{\lambda^2}\}$$

$$=[-x^2e^{-\lambda x-\frac{2xe^{-\lambda x}}{\lambda}}-\ \frac{2xe^{-hx}}{\lambda}\ -\frac{2^{e-\lambda x}}{\lambda}]$$

$$=\infty\text{-}2\infty\ \text{-}\frac{2}{\infty}=0$$

$$=[-x^2e^{-\lambda}\ \text{-}\frac{2xe^{-\lambda x}}{\lambda}\text{ - }2\frac{xe^{-\lambda 2}}{\lambda}]$$

$$=\frac{2}{\lambda}$$

Variance (X) = E $(x^2)$ - $(Ex)^2$

$$=\frac{2}{\lambda}\text{-}\frac{1}{\lambda 2}=\frac{2\lambda-1}{\lambda^2}$$

$$=\frac{2\lambda-1}{\lambda^2}$$

**3.5.3 UNIFORM DISTRIBUTION**

$$f(x)=\frac{1}{b-a} \qquad a<x<b$$

MEAN

$$E(X)=\int_a^b xfxdx$$

$$=\int_a^b \frac{x}{b-a}dx$$

$$=\frac{1}{2(b-a)}[x^2]$$

$$=\frac{1}{2(b-a)}(b^2 - a^2)$$

$$=\frac{(b+a)(b-a)}{2(b-a)}$$

$$=\frac{a+b}{2}$$

E$(X^2)$ = $\int x^2 f(x)$dx

$$=\int_a^b \frac{x^2}{b-a}\text{dx}$$

$$= \frac{1}{b-a}\int_a^b x^2\text{dx}$$

$$= \frac{1}{3(b-a)}\ [x^3]$$

$$=\frac{b^3-a^3}{3(b-a)}$$

Variance(X) = E $(x^2)$-$(Ex)^2$

$$= \frac{b^3-a^3}{3(b-a)}-\frac{(a+b)^2}{4} = \frac{4\left(b^3-a^3\right)-(b-a)\left(a+b^2\right)}{12(b-a)}$$

$$=\frac{4b^3-4a^3-3[(b-a)(a+b)^2]}{12(b-a)}$$

$$=\frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12}$$

$$=\frac{b^2+a^2+2ab}{12}$$

$$=\frac{(b-a)^2}{12}$$

## 3.6 THE CDF OF GENERALIZED PARETO DISTRIBUTION USING THE TRANSFORMATION METHOD

$$f(x) = \frac{\{1+\frac{\xi(x-u)}{\delta}\}^{\frac{-1-\xi}{\xi}}}{\delta}$$

Let $z/u = 1+\frac{\xi(x-u)}{\delta}$

$$=\int_u^x \frac{z/u^{\frac{-1-\xi}{\xi}}}{\delta}\frac{\delta}{\xi u}\,dz$$

$$=\int_u^x z^{\frac{-1-\xi}{\xi}} u^{\frac{1+\xi}{\xi}} \frac{1}{\xi u} dz$$

$$= \int_u^x [u/z^{1/\xi}]$$

Replacing for 1/ξ with α we get

$$= 1- (u/z(xp))^{\alpha}$$

# CHAPTER 4

# ESTIMATION OF GENERALIZED PARETO DISTRIBUTION

## 4.1 INTRODUCTION

There are three methods of estimating parameters of GPD , namely method of moment, L-moment method and maximum likelihood method.

In this chapter we shall discuss the three methods and also derive the parameters of the three distributions .I shall also discuss order statistics.We shall also describe the quantiles of the generalized Pareto distribution .

## 4.2 MAXIMUM LIKELIHOOD ESTIMATION.

It involves estimation of parameters of each of the sampled probability distributions .Once the parameters of a given distribution are estimated then a fitted distribution is available for further analysis.

The maximum likelihood estimates are used because they have several desirable properties which include consistency,efficiency,asymptotic normality and invariance.

The advantage of using the maximum likelihood estimation is that it fully uses all the available information about the parameters contained in the data and that if highly flexible (Denuit 2007)

The steps involved in finding the maximum likelihood estimators are as follows.

(1) Write down the likelihood function for the available data. If the likelihood is based on a set of known values $xn$1, 2 ,…, , then the likelihood function will take the form $f(x_1)\, f(x_2)$,,,$f(x_n)$ , where $f(x)$ is the PDF (or PF if it's discrete) of the distribution that is to be fitted.
(2) Take logs. This will usually simplify the algebra.
(3) Maximize the log-likelihood function. This usually involves differentiating the Log-likelihood function with respect to each of the unknown parameters, and setting the resulting expression(s) equal to zero.
(4) Solve the resulting equation(s) to find the MLEs of the parameters.

### 4.2.1 MLE OF PARETO DISTRIBUTION

Let $x_{1,}x_2,,,x_n$ be a random variable from a Pareto distribution.Then its MLE is given by

$$f(x) = \alpha\frac{\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}$$

$$\prod_{k=1}^{n} fx = \alpha^{n}\lambda^{\alpha n}\sum_{i=1}^{n}(\lambda + x)^{-(\alpha+1)}$$

$$\ln\prod_{k=1}^{n} fx = n\ln\alpha + \alpha n\ln\boldsymbol{\lambda} - (\alpha+1)\sum_{i=1}^{n}(\lambda + xi)$$

$$\frac{d}{d\alpha}\ln\prod_{k=1}^{n} fx = \frac{n}{\alpha} + n\ln\boldsymbol{\lambda} - \sum_{i=1}^{n}(\lambda + xi)$$

$$\frac{n}{\alpha} + n\ln\boldsymbol{\lambda} - \sum_{i=1}^{n}(\lambda + xi) = 0$$

$$\frac{n}{\alpha} = -n\ln\boldsymbol{\lambda} + \sum_{i=1}^{n}(\lambda + xi)$$

$$\alpha = \frac{n}{\sum_{i-1}^{n}(\lambda+xi) - nln\lambda}$$

Then differentiating with respect to $\boldsymbol{\lambda}$

$$\frac{dln(x,\alpha\lambda)}{d\lambda} = \frac{\alpha n}{\lambda} - \alpha\sum_{i=1}^{n}(\lambda + xi) - \sum_{i=1}^{n}(\lambda + xi)$$

$$= \frac{\alpha n}{\lambda} - \frac{d}{d\lambda}\{\alpha n\boldsymbol{\lambda} + \alpha nxi - n\boldsymbol{\lambda} + nx\}$$

$$= \frac{\alpha n}{\lambda} - \alpha n - n$$

Equating to zero we have

$$0 = \frac{\alpha n}{\lambda} - \alpha n - n$$

$$\frac{\alpha n}{\lambda} = n + \alpha n$$

$$\frac{\lambda}{\alpha n} = \frac{1}{n + \alpha n}$$

$$\boldsymbol{\lambda} = \frac{\alpha}{1+\alpha}$$

**4.2.2 MLE USING THE METHOD OF TRANSFORMATION**

$$f(x) = \frac{1}{\delta}\left(\frac{\xi(x-u)}{\delta} + 1\right)^{\frac{-1-\xi}{\xi}}$$

Using the transformation to a Pareto distribution we let $\alpha = 1/\xi$

And using the transformation $z/u = \frac{\xi(x-u)}{\delta} + 1$

Therefore $z = u\left\{\frac{\xi(x-u)}{\delta} + 1\right\}$

And $dz = \frac{\xi u dx}{\delta}$

$\Longrightarrow dx = \frac{\delta dz}{\xi u}$

Therefore the Jacobian of the transformation is f (z) = g(z)|J|

$$= \frac{1}{\delta}(Z/U)^{(-1-\xi)/\xi}\,\frac{\delta}{\xi U}$$

$$= \frac{1}{\xi U}(Z/U)^{(-1-\xi)/\xi}$$

$$= \frac{Z^{(-1-1/\xi)}}{\xi}\,U^{1/\xi}$$

Replacing $1/\xi$ with $\alpha$ we get $= \alpha z^{-(\alpha+1)} u^{\alpha}$

$= \frac{\alpha u^{\alpha}}{z^{\alpha+1}}$ Which is a Pareto distribution

Finding its MLE

$$L(x, u, \alpha) = \prod_{k=1}^{n} fx$$

$$= \prod_{k=1}^{n} \frac{\alpha u^{\alpha}}{z^{\alpha+1}}$$

$$= \alpha^{n} u^{\alpha n} \sum_{i=1}^{n} z^{-(\alpha+1)}$$

$$\text{Ln} \prod_{k=1}^{n} fx = \text{n ln}\alpha + \alpha\text{n ln u} - (\alpha+1)\text{ln}z_i$$

$$\frac{d\text{Ln} \prod_{k=1}^{n} fx}{d\alpha} = \frac{n}{\alpha} + \text{n ln u} - \sum_{i=1}^{n} z_i$$

$$\frac{n}{\alpha} + \text{nln u} - \sum_{i=1}^{n} z_i = 0$$

$$\frac{n}{\alpha} = \sum_{i=1}^{n} z_i - \text{nln u}$$

$$\frac{n}{\alpha} = n \sum_{i=1}^{n} \ln(\frac{z_i}{u})$$

$$\frac{1}{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \ln(\frac{z_i}{u})$$

Substituting back 1/ξ=α.We get

$$\xi = \frac{1}{n} \sum_{i=1}^{n} \ln(\frac{z_i}{u})$$

**4.2.3 MLE OF UNIFORM DISTRIBUTION**

Suppose $x_1 x_2 x_3 ,,, x_n$ form a random sample from a uniform distribution on the interval (0,Θ) where Θ> 0 and is unknown.

$$\text{f(x)} = \frac{1}{b-a}$$

$$= \frac{1}{\Theta}$$

$$\text{L(fx,}\Theta) = \frac{1}{\Theta^n}$$

Taking natural logs

$$\text{LnL(fx,}\Theta) = \ln\Theta^{-1}$$

$$\text{LnL(fx,}\Theta) = -\ln\Theta$$

Differentiating with respect to Θ

$$\frac{dlnlf(x.\Theta)}{d\Theta} = -\frac{1}{\Theta}$$

Equating to zero we have

Θ=min of $x_1 x_2 x_{3,,,} x_n$

### 4.2.4 MLE OF EXPONENTIAL DISTRIBUTION

Let $x_1 x_2 x_3, , , x_n$be a random variable from an exponential distribution.

Then the likelihood function is given by,

$$\text{f(x)} = \boldsymbol{\lambda} e^{-\lambda x}$$

$$\text{L(x,fx)} = \lambda^n e^{-\lambda \sum_{i=0}^{n} x_i}$$

$$\text{Ln (L(x,fx))} = \text{nln}\boldsymbol{\lambda} - \boldsymbol{\lambda} \sum_{i=0}^{n} x_i$$

$$\frac{d}{d\lambda} \text{Ln (L(x,fx))} = \frac{n}{\lambda} - \sum_{i=0}^{n} x_i$$

$$\frac{d}{d\lambda} \text{Ln (L(x,fx))} = 0$$

$$\frac{n}{\lambda} - \sum_{i=0}^{n} x_i = 0$$

$$\frac{n}{\lambda} = \sum_{i=0}^{n} x_i$$

$$\boldsymbol{\lambda} = \frac{1}{x}$$

## 4.3METHODS OF MOMENT

The basic idea behind this form of the method is to equate the first sample moment about the origin $\mu_1 = \frac{1}{n}\sum_{i=1}^{n} xi$ to the theoretical moment E(x)

Equate the second moment about the origin $\mu_2 = \frac{1}{n}\sum_{i=1}^{n} xi^2$ to the second theoretical moment E($X^2$)

Continue equating the sample moments about the origin $u_k$ with the corresponding theoretical moments E($X^K$)

Solve for the parameters

### 4.3.1 MOM OF PARETO DISTRIBUTION

Let $x_1, x_2, x_3, , , x_n$be  a random variable from a Pareto distribution.Then its moments are

$$f(x)=\frac{\propto\lambda^{\alpha}}{x^{\alpha+1}} \qquad x>0, \lambda>0, \alpha>0$$

$$E(X)=x\int_0^{\infty} f(x)dx$$

$$=x\int_{\lambda}^{\infty} \alpha\lambda^{\alpha}x^{-(\alpha+1)}dx$$

$$=\alpha x\lambda^{\alpha}x^{-(\alpha+1)}dx$$

$$=-\alpha\lambda^{\alpha}\frac{x^{-\alpha+1}}{\alpha-1}$$

$$=\alpha\lambda^{\alpha}\left\{\frac{x^{-\alpha+1}}{\alpha+1}\right\}$$

$$=\alpha\lambda^{\alpha}\frac{\lambda^{-\alpha}}{\alpha-1}$$

$$\mu_1=\frac{\alpha\lambda}{\alpha-1}$$

$$=\sum_{i=1}^{n} xi\frac{1}{n}$$

$$E(X^2)= x^2 f(x)dx$$

$$=x^2\int_{\lambda}^{\infty}\frac{\alpha\lambda^{\alpha}}{x^{\alpha+1}}dx$$

$$=\alpha\lambda^{\alpha}\int_{\lambda}^{\infty} x^{2-\alpha+1}dx$$

$$=-\alpha\lambda^{\alpha}\left\{\frac{x^{-\alpha+2}}{\alpha-2}\right\}$$

$=\alpha\lambda^{\alpha}\frac{\lambda^{-\alpha+2}}{\alpha-2}$

$\mu_2=\frac{\alpha\lambda^2}{\alpha-2}$

Using the methods of moment we can solve the parameters **λ** and α by Equating the first and second moments $\mu_1$ and $\mu_2$

$\mu_1=\frac{\sum_{i=1}^{n} xi}{n}=\frac{\alpha\lambda}{\alpha-1}$

$\mu_1$(α-1)=α**λ**

$\mu_1$α-u1=α**λ**

$\mu_1$α-α**λ**=$\mu_1$

Factoring α and solving for it we have

$\alpha=\frac{\mu_1}{\mu_1-\lambda}$ *1

Solving for **λ** in the same equation we have $\mu_1=\frac{\alpha\lambda}{\alpha-1}$

$\mu_1$(α-1)=α**λ**

Solving for **λ** we have **λ**$=\frac{\mu_1(\alpha-1)}{\alpha}$ *2

Then with the use of *1 and *2 we can get the individual estimate by replacing either of the two in our second moment equation.

$\mu_2=\frac{\alpha\lambda^2}{\alpha-2}$

Using *1 replace α and solve

$(\frac{\mu_1\lambda^2}{\mu_1-\lambda})/(\frac{\mu_1}{\mu_1-\lambda}-2)=\mu_2$

$(\frac{\mu_1\lambda^2}{\mu_1-\lambda})/(\frac{\mu_1-2(\mu_1-\lambda)}{\mu_1-\lambda})=\mu_2$

$(\frac{\mu_1\lambda^2}{\mu_1-\lambda})/(\frac{\mu_1-2\mu_1+2\lambda}{\mu_1-\lambda})=\mu_2$

$(\frac{\mu_1\lambda^2}{\mu_1-\lambda})/(\frac{2\lambda-\mu_1}{\mu_1-\lambda})=\mu_2$

Opening brackets and simplifying we have$\mu_1$

$$\frac{\mu_1\lambda^2}{2\lambda-\mu_1}=\mu_2$$

Cross multiplying we have $\mu_1\lambda^2$-2$\mu_2\boldsymbol{\lambda}$+$\mu_2$=0

This result is a quadratic equation where we can solve for the values of **λ** by completing squares.

$$\lambda^2\text{-}\frac{2\mu_2\lambda}{\mu_1} =\text{-}\ \mu_2$$

$$\lambda^2\text{-}\ \frac{2\mu_2\lambda}{\mu_1}+(\frac{\mu_2}{\mu_1})^2 = (\frac{\mu_2}{\mu_1})^2 -\mu_2$$

$$(\lambda-\frac{\mu_2}{\mu_1})^2=\frac{{\mu_2}^2}{{\mu_1}^2}-\mu_2$$

$$\boldsymbol{\lambda}\text{-}\ \frac{\mu_2}{\mu_1} = \sqrt{(\frac{{\mu_2}^2}{{\mu_1}^2}-\mu_2)}$$

$$\boldsymbol{\lambda} = \frac{\mu_2}{\mu_1}+\ \sqrt{(\frac{{\mu_2}^2}{{\mu_1}^2}-\mu_2)}$$

We can now get the α estimate by replacing the **λ** in our *1

$$\alpha = \frac{\mu_1}{\mu_1-\frac{\mu_2}{\mu_1}+\sqrt{(\frac{{\mu_2}^2}{{\mu_1}^2}-\mu_2)}}$$

### 4.3.2 MOM OF EXPONENTIAL DISTRIBUTION

let$x_1, x_2, x_3, , , x_n$be a random variable from an exponential distribution assumed to be independent and identically distribution with exponential (**λ**)

$$\text{f(x)} = he^{-hx} \qquad \text{x>0}$$

$$\text{E(x)} = xhe^{-hx^-}\ \text{dx} \quad = \frac{T2}{\lambda} = \frac{1}{\lambda}$$

$$\text{E}(x^2) = \int_0^x x^2\ he^{-hx}\ \text{dx} = \frac{T3}{\lambda^2} = \frac{2}{\lambda^2}$$

The method of moment of **λ** based on the second moment

$$\lambda^- = \frac{1}{x} = \frac{n}{\xi x1}$$

$$\lambda^- = \sqrt{\frac{2n}{\xi x^2}}$$

### 4.3.3 MOM OF UNIFORM DISTRIBUTION

Parameter estimation of uniform distribution using the method of moment.

$$\mu_1 = \frac{a+b}{2} \qquad *1$$

$$\mu 2 = \frac{a^2 + b + b^2}{3} \qquad *2$$

Solving for a in *1 we get a=2μ1 –b, replacing μ in *2 we get

$$\mu 2 = (2\mu_1 - b)^2 + \frac{b + (2\mu_1 - b)}{3} + b^2 \boldsymbol{\lambda}$$

$$\mu 2 = 4\ \mu^2\ \frac{2\mu_1\ b + b^2 + 2\mu_1 b - b^2 + b^2}{3}$$

$$\mu = 4{\mu_1}^2 - \frac{2\mu 1 b}{3} + b^2$$

$$3\mu 2 = 4{\mu_1}^2 - 2\mu_1 b + b^2$$

Solving the equation by completing the squares we have

$$b^2 - 2\mu_1 b = 3\mu 2 - 4{\mu_1}^2$$

$$b^2 = 2\mu_1 b + {\mu_1}^2 = 3\mu 2 - 4{\mu_1}^2 + {\mu_1}^2$$

$$(b - \mu_1)^2 = 3\mu_2 - 3{\mu_1}^2$$

$$= 3(\mu 2 - \mu_1 \mu_2 - \mu_1^2)$$

Therefore $b - \mu_1 = \sqrt{3(\mu_2 - \mu_1^2)}$

Therefore $b^\lambda = \mu_1 + \sqrt{3(\mu_2 - \mu_1^2)}$

Replacing for $b^\lambda$ in $a^\lambda = 2\mu_1 - b$ we get

$$a^\lambda = 2\mu_1 - \mu_1 - \sqrt{3(\mu_2 - \mu_1^2)}$$

$$a^\lambda = \mu_1 - \sqrt{3(\mu_2 - \mu_1^2)}$$

## 4.4 ORDER STATISTICS

The order statistics of a random sample $x_1 x_2$ ………$x_n$ are the sample values placed in ascending order. They are denoted by -$x_1 x_2$ … … . . $x_n$.

Let $x_1 x_{2,} x_n$ be a random sample of size n from a distribution with pdf f(x) and cdf .

Then the cdf of the j th order statistic is given by,

$$Fj(x) = \sum_{k=j}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-j}$$

Let $x_1 x_2$……….$x_n$ be a random sample size and form a distribution of continuous population with pdff(x)andcdf f(x).

Then the pdf of the jth order statistic is given by,

$$Fj(x) = j\binom{n}{j} f(x)\left[F(X)^{j-1}\right][1 - F(x)]^{n-j}$$

## 4.5 L MOMENTS

Definition,

Let x be a real valued random variable with cumulative distribution function F(x) and quartile x(f) and let $x_1$ . n ≤ $x_2$ n ≤ …$x_n$ .n be the order statistic of a random sample size n drawn from the distribution of x.

Define the L moments of x to be the quantities

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (\_ - 1)^k \binom{r-1}{k} \text{ Exr-k;r} \qquad r = 1,2,\ldots\ldots$$

The L in the moments emphasizes that $\lambda_r$ is a given function of the expected order statistics furthermore .

The natural estimation of $\lambda_r$ based on an observed sample of data is linear combination of the ordered data values.

From the previous page the pdf of jth order statistic is given by,

$$Fj(x) = j\binom{r}{j} f(x)\ [F(x)]^{j-1} [1 - F(x)]^{r-j}$$

$$= \frac{r!}{(r-1)!(r-j)!} [F(x)]^{j-1} [1 - F(x)]^{r-j} \text{fx}$$

The expectation of an order statistic may be written as

$$\text{Exj.r} = \int_{-\infty}^{\infty} xfi(x)dx$$

$$= \int_{-\infty}^{\infty} x \frac{r\,!}{(j-1)!(r-j)!} [F(x)]^{j-1}[1-F(x)]^{r-j} \text{ f(x) dx}$$

Hence,

$$\text{E}x_j\text{.r} = \frac{r!}{(j-1)(r-j)!} \int_0^1 x\,[F(x)]^{j-1}\,[1-F(x)]^{r-j} \text{ dfx}$$

Lemma 2

A finite mean implies finite expectation of all order statistics

Proof

Assume that the mean u=$\int_0^1 xudu$ is finite so x(u) is intergrable in the interval (0,1)

$$\int_0^1 u^{j-1}\,[1-u]^{r-j}\text{du} = \text{B (j, r-j+1)} = \frac{(j-1)!(r-1)!}{r\,!}$$

Is finite then $u^{j-1}[1-u]^{r-j}$ is intergrable in the internal (0,1)

Hence x(u) $u^{j-1}(1-u)^{r-j}$ is intergrable in the interval (0,1) because the product of the intergrable function on any interval is an intergrable function function on this interval and so

$$\int_0^1 x(u)\,u^{r-j}\text{du is finite}$$

$$\text{E}x_j\text{.r} = \frac{r!}{(j-1)!(r-j)!} \int_0^1 x(u)u^{j-1}\,(1-u)^{r-j} \text{ du is finite}$$

Therefore a finite mean implies a finite expectation all order statistics.

i.e. $\lambda_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k}\text{E}x_{r-k.r}$ r=1,2,…………..2.1.2

to a simple form that is easy to use

Change variables u=F(x) let Q be the inverse of function F i.e. Q(fx) = x or F(Quj) =u

$$\text{E}x_{r-k:r} = \frac{r!}{(j-1)!(r-j)!} \int_0^1 Q(u)\,u^{r-k-1}(1-u)^k \text{ du} \qquad 2.1.3$$

Substitute from equation 2.1.3 into eqn, 2.1.1

$$\lambda_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \frac{r!}{(r-1-k)!k!} \int_0^1 Quu^{r-k-1}\,(1-u)^k \text{ du}$$

For convenience consider $\lambda_r$+1 ………………..of $\lambda_r$

$$\lambda_r{+}1 = (r+1)^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r}{k}\frac{(r+1)!}{(r-k)!k!}\int_0^1 Qu\, u^{r-k}(1-u)^x \text{ du}$$

Make that $\lambda_{r+1}$ =$(r+1)^{-1}$ (r+1)! =r! and rearrange terms

$$\lambda_{r+1} = \int_0^1 \sum_{k=0}^{r}(-1)^k \binom{r}{k}^2 u^{r-k}(1-u)^k \text{Q(u)du} \qquad 2.1.4.$$

Expand $(1-u)^k$ in powers of u!

$$\lambda_{r+1} = \int_0^1 \sum_{k=0}^{r}(-1)^k \binom{r}{k}^2 u^{r-k} \sum_{k=0}^{r}(u)^{r-k}\binom{k}{j} u^{k-j}\text{Q(u) du}$$

$$= \int_0^1 \sum_{k=0}^{r}\sum_{j=0}^{r}(-1)^j \binom{r}{k}^2 \binom{k}{j} u^{r-j}\text{Q(u)du.}$$

Interchange order of summation over j and k.

$$\lambda_{r+1} = \int_0^1 \sum_{k=0}^{r}\sum_{j=0}^{r}(-1)^j \binom{r}{k}^2 \binom{k}{j} u^{r-1}\text{Q(u)du}$$

Reverse order of summation over j and k

$$\lambda_{r+1} = \int_0^1\int_0^1 \sum_{m=0}^{r}\sum_{n=0}^{r}(-1)^{r-m} \binom{r}{r-n}^2 \binom{r-n}{r-m} u^m\text{Q(u)du.}$$

$$\lambda_{r+1} = \int_0^1\int_0^1 \sum_{n=0}^{r}(-1)^{r-m} \left\{\sum_{n=0}^{m}\binom{r}{r-n}^2 \frac{r-n}{r-m}\right\} u^m\text{Qudu} \qquad 2.1.5.$$

Note that,

$$\binom{r}{r-n}^2\binom{r-n}{r-m}\binom{r}{n} = \binom{r}{m}\binom{m}{n}$$

Expand the binomial in terms of factorial and the

$$\sum_{n=0}^{m}\binom{r}{n}\binom{m}{n} = \sum_{n=0}^{m}\binom{r}{r-n}\binom{m}{n} = \binom{r+m}{r} = \binom{r+m}{m} \quad 2.1.7$$

Second equality follow because to choose r items form r+m we can choose from the first m items and r-n from the remaining r items for any n in 0.1,…….m from 2.2.5 and 2.2.6 we have

$$\sum_{n=0}^{m}\left(\binom{r}{r-n}^2\right)\binom{r-n}{r-m} = \binom{r}{m}\binom{r+m}{m} \quad 2.1.8$$

And substituting into 2.1.5 gives,

$$\lambda_{r+1} = \int_0^1 \underset{m=0}{\overset{m}{\xi}} (-1)^{r-m}\binom{r}{m}\binom{r+m}{m}\text{x(f) Fm df} \qquad 2.1.9$$

Let $p^*_{r,m} = (-1)^{r-m}\binom{r}{m}\binom{r+m}{m}$ 2.1.10

And $p_r^*$(f) $=\sum_{m=0}^{r} p^*{}_{r,m} f^m$ 2.1.11

Substituting (2.1.11) into 2.1.9 we have

$\lambda_r = \int_0^1 x(f)\, p_{r-1}^*$ (f) df r=1,2,…… 2.1.12

To find $\lambda_r = \frac{1}{2}\sum_{k=0}^{1}(-1)^k \binom{1}{k}$ E $x_{2-k:2}$

$$=\frac{1}{2}\left[(-1)^0\binom{1}{0}\ EX_{2:2}\ +\ (-1)^1\binom{1}{1}\ Ex_{1;2}\right]$$

$$=\frac{1}{2}\left[Ex_{2,2}\ - EX_{1;2}\right] = \frac{1}{2}\,\mathrm{E}\left(X_{2;2}\ -\ X_{1;2}\right)$$

And we can substitute r=2 in equation 2.1.12

$$\lambda_2 = \int_0^1 2(F)\, p *_1(\mathrm{F})\ \mathrm{df}$$

$$=\int_0^1 x)(f)\left[\sum_{m=0}^{1} p_{1,m}{}^*\ \ Fm\right]\mathrm{df}$$ from equation 2.1.12.

$$=\int_0^1 x(F)\ [p *_{1,0}\, F^0 + p *_{1.1}\, F^1]$$

$$=\int_0^1 x\ [(F)\ (-1)^1\binom{1}{2}\binom{1}{0} + (-1)^0\binom{1}{1}\binom{2}{1}\mathrm{F}]\ \mathrm{dy}$$ from 2.1.10

$$=\int_0^1 xF(2F-1)\mathrm{df}$$

$$\lambda_2 = \frac{1}{2}\mathrm{E}(X_{2.2} - X_{1.2}) = \int_0^1 x\ (2\mathrm{F} - 1)\ \mathrm{df}$$

The first four **λ** moments are $\lambda_1$= E(X)=$\int_0^1 xdf$

$$\lambda_2 = \mathrm{E}(X_{2.2} -\ X_{1.2}) = \int_0^1 x\ (2\mathrm{f} - 1)\ \mathrm{df}$$

$$\lambda_3 = {}^1\!/_3\mathrm{E}(X_{3.3} - X_{2.3} + X_{1.3}) = \int_0^1 x\ (6F^2 - 6\mathrm{F}+1)\ \mathrm{df}$$

$$\lambda_4 = \frac{1}{4}\mathrm{E}(X_{4.4} - 3X_{3.4} + 3X_{2.4} - X_{1.4}) = \int_0^1 X\ (20F^3 - 30F^2 - 12\mathrm{F})\mathrm{df}$$

The use of **λ** moments to describe probability distribution is justified by the next theorem.

$\lambda_2$is a measure of the scale or disperse on of the random variable x. It is often convenient to standardize the higher moments$\lambda_r$ r≥ 3 so that they are independent of the units of measurement of x.

### 4.5.1 L MOMENTS OF UNIFORM DISTRIBUTION.

In this section we find the L moment for the uniform distribution. The uniform distribution has the probability density.

$$F(x) = \frac{1}{b-x}Fx = \int_a^x \frac{1}{b-a}\,dx = \frac{x-a}{b-a} \qquad F= \frac{x-a}{b-a}$$

$$XF = F(b – a )+a$$

And its quartile function x(f) = α + (ß - α)F

We are about to find the first four L moments of the uniform distribution. Beforedoing so we have to determine the first PWM of the uniform distribution.

$$ß_r = \int_0^1 xf\ f^r df \qquad r = 0, 1, 2…$$

$$=\int_0^1 [\alpha + (b-\alpha)F]\,F^r dF \ = \int_0^1 \alpha F^r dF + \int_0^1 (ß - \alpha)F^{r+1}\,dF$$

$$=\left[\alpha\frac{F^{r+1}}{r+1}\right] + \left[\frac{(ß-\alpha)F^{r+2}}{r+2}\right]$$

$$=\frac{\alpha}{r+1} + \frac{ß+\alpha}{r+2}$$

$$ß_r = \frac{\alpha}{r+1} + \frac{ß-\alpha}{r+2}$$

$$\lambda_1 = ß_0 = \alpha + \frac{ß-\alpha}{2} = \frac{1}{2}(ß + \alpha)$$

$$\lambda_2 = 2ß_1 - ß_0 = 2\left[\frac{\alpha}{2} + \frac{ß-\alpha}{3}\right] - \frac{\alpha-ß}{2} = \frac{1}{6}(ß - \alpha)$$

$$\lambda_3 = 6ß_2 – 6\ ß_1 + ß_0 = 6\left[\frac{\alpha}{3} + \frac{ß-\alpha}{4}\right] – 6\left[\frac{\alpha}{2} + \frac{ß-\alpha}{3}\right] + \frac{\alpha+ß}{2} = 0$$

$$=20\left[\frac{\alpha}{4} + \frac{ß+\alpha}{5}\right] – 30\left[\frac{\alpha}{3} + \frac{ß-\alpha}{4}\right] + 12\left[\frac{\alpha}{2} + \frac{ß-\alpha}{3}\right] - \frac{\alpha-ß}{2} = 0$$

Hence $T_3 = \lambda_3/\lambda_2 = 0$

$T_4 = \lambda_4/\lambda_2 = 0$

### 4.5.2 L MOMENTS FOR EXPONENTIAL DISTRIBUTION.

In this section we find the four L moments for the exponentials distribution.

$$F(x) = 1 - \exp(- \frac{x-\xi}{\alpha}) \quad \text{where} \quad \xi \leq x < \infty$$

Firstly we want to find the quartile function of the exponential distribution so we replace x(f) with x and f with F(x) we have

$$F = 1 - \exp\{-x(f) - \xi/\alpha\} \quad \text{then } 1 - F \exp\{-(F) - \xi)/\alpha\}$$

Hence in $(1 - F) = - (x(F) - \xi) /\alpha$

Therefore $x (F) = \xi - \alpha\ln (1 - F)$

Secondly we want to find the rth PWM for the exponential distribution.

$$ß_r = \int_0^1 x(f)F^{\alpha}\, df \qquad \alpha = 0,1,2,\ldots\ldots\ldots..$$

$$= \int_0^1 [\xi - \alpha ln(1-F)]\, F^{\alpha}\, df = \frac{\xi}{\alpha+1}\alpha\int_0^1 F^{\alpha}\ln (1 - F)df$$

Now we find $\int_0^1 F^{\alpha} \ln (1 - F)df$

Integration by part we get,

$$\int_0^1 F^{\alpha}\ln (1 - f)\, df = \left[\frac{F^{\alpha+1}}{\alpha+1}\ ln(1-F)\right] + \frac{1}{\alpha+1}\int_0^1 \frac{F^{\alpha+!}}{1-F}\, df$$

$$= \frac{1}{\alpha+1}\int_0^1 \frac{F^{\alpha+1}}{1-F}df$$

Let $z = 1 - F$ Then $dz = -df \quad F = 1 - 2$

So $F^{\alpha+1} = (1-Z)^{\alpha+1} = \sum_{k=0}^{r+1}(-1)^k \binom{r+1}{k} z^k$

Therefore

$$\frac{F^{\alpha+1}}{1-F} = \sum_{k=0}^{r+1}(-1)^k \binom{r+1}{k} z^{k-1} = \frac{1}{z} + \sum_{k=1}^{r+1} z^{k+1} \binom{r+1}{k}$$

$$\int_0^1 F^r \text{in } (1 - F)df = \frac{-1}{r+1}\int_0^1 [\frac{1}{z} + \sum_{k=1}^{r+1}(-1)^k \left(\frac{r+1}{k}\right) Z^{k+1}]df$$

$$= \frac{-1}{r+1}\left[lnz + \sum_{k=1}^{r+1}(-1)^k \frac{1}{z}\binom{r+1}{k} z^k\right]$$

$$= \frac{1}{r+1}\left[\sum_{k=1}^{r+1}(-1)^k \frac{1}{z}\binom{r+1}{k}\right]$$

$=\frac{1}{r+1}\,[\sum_{k=1}^{r+1}(-1)^k \frac{1}{z}\binom{r+1}{k}$

$ß_r = \frac{\xi}{r+1} - \frac{\alpha}{r+1}\sum_{k=1}^{r+1}(-1)^k \frac{1}{z}\binom{r+1}{k}$

$ß_0 = \xi - \alpha\,(-1) = \xi + \alpha$

$\lambda_1 = ß_0 = \xi + \alpha$

$ß_1 = \frac{\xi}{z} - \frac{\alpha}{z}\sum_{k=1}^{2}(-1)^k \frac{1}{k}\binom{2}{k} = \frac{\xi}{2} - \frac{\alpha}{2}\left[-\binom{2}{1} + \frac{1}{2}\binom{2}{2}\right]$

$=\frac{\xi}{z} + \frac{3\alpha}{4}$

Therefore,

$\lambda_2 = 2ß_1 - ß_0 = 2\left(\frac{\xi}{2} + \frac{3\alpha}{4}\right) - (\xi + \alpha) = \frac{\alpha}{2}$

$ß_2 = \frac{\xi}{3} - \frac{\alpha}{3}\sum_{k=1}^{2}(-1)^k \frac{1}{k}\binom{3}{k} = \frac{\xi}{3} - \frac{\alpha}{3}\left[-\binom{3}{1} + {}^1\!/_2\binom{3}{2} - {}^1\!/_3\binom{3}{3}\right]$

$=\frac{\xi}{3} + \frac{11\alpha}{18}$

Thus $\lambda_3 = 6ß_2 - 6ß_1 + ß_0$

$=6\left(\frac{\xi}{3} + \frac{11\alpha}{18}\right) - 6\left(\frac{\xi}{2} + \frac{3\alpha}{4}\right) + \alpha + \xi = \frac{\alpha}{6}$

Then,

$T_3 = \lambda_3/\lambda_2 = {}^1\!/_3$

$ß_3 = \frac{\xi}{4} - \frac{\alpha}{4}\sum_{k=1}^{2}(-1)^k \frac{1}{k}\binom{4}{k} = \frac{\xi}{4} - \frac{\alpha}{4}\left[-\binom{4}{1} + \frac{1}{3}\binom{4}{2} - \frac{1}{3}\binom{4}{3} + \frac{1}{4}\binom{4}{4}\right]$

$=\frac{\xi}{4} + \frac{25\alpha}{48}$

$\lambda_4 = 20ß_3 - 30ß_2 + 12ß_1 - ß_0$

$=20\left(\frac{\xi}{4} + \frac{25\alpha}{48}\right) - 30\left(\frac{\xi}{3} + \frac{11\alpha}{18}\right) + \left(\frac{\xi}{2} + \frac{3\alpha}{4}\right)(\xi + \alpha)$

$=5\xi + \frac{125\alpha}{12} - 10\xi - \frac{55\alpha}{3} + 6\xi + 9\alpha - \xi - \alpha = \frac{\alpha}{12}$

$T_4 = \frac{\lambda_4}{\lambda_2} = {}^1\!/_6$

### 4.5.3. L MOMENTS FORGENERALIZED PARETO DISTRIBUTION

The generalized Pareto distribution has the probability density function.

$$F(x) = \alpha^{-1} e^{-(1-k)[-k^{-1} \log 1 - k(x-\xi)\alpha]}$$

And has the quartile,

$$X(f) = \xi + \alpha\{1-(1-f)^k\} \text{ (k=}\xi + \frac{\alpha}{k} - \frac{\alpha}{k}(1-f)^k$$

Now we will find $ß_\alpha$ for the generalized Pareto df

$$ß_r = \int_0^1 xfF^r = \int_0^1 (\xi + \frac{\alpha}{k})F^r df - \frac{\alpha}{k}\int_0^1 (1-F)^k F^r df$$

$$= \frac{1}{r+1}(\xi + \frac{\alpha}{k}) - \frac{\alpha}{k}\int_0^1 (1-F)^k F^r df$$

Let u=1-f⟹du –dfF = 1-u

Then,

$$\int_0^1 (1-F)^k f^r df = -\int_1^0 u^k (1-u)^r du = -\int_1^0 u^k \sum_{j=0}^{j} (-1)^j \binom{r}{j} u^j du$$

$$= -\sum_{j=0}^{r} (-1)^j \binom{r}{j} \int_0^1 u^{k+j} du = \sum_{j=0}^{r} (-1)^j \frac{1}{k+j+1}\binom{r}{j}$$

$$ß_r = \frac{1}{r+1}(\xi + \frac{\alpha}{k}) - \frac{\alpha}{k}\sum_{j=0}^{r} (-1)^j - \frac{1}{k+J+1}\binom{r}{j}$$

Now,

$$ß_0 = \xi + \frac{\alpha}{k} - \frac{\alpha}{k}\left(\frac{1}{k+1}\right) = \alpha + \frac{\alpha}{k} - \frac{\alpha}{k(k+1)}$$

$$= \xi + \frac{\alpha}{k+1}$$

Thus, $\lambda_1 = ß_0 = \xi + \frac{\alpha}{k+1}$

Furthermore,

$$ß_1 = \frac{1}{2}(\xi + \frac{\alpha}{k}) - \frac{\alpha}{k}\left[\frac{1}{k+1}\binom{1}{0} - \frac{1}{k+2} \ (\binom{1}{1})\right]$$

$$= \frac{1}{2}\xi + \frac{\alpha}{2k} - \frac{\alpha}{k}\left[\frac{1}{k+1-} - \frac{1}{k+21}\right] = \frac{1}{2}\xi + \frac{\alpha(k+3)}{2(k+1)(k+2)}$$

$$\lambda_2 = 2ß_1 - ß_0 = 2\frac{1}{2}\xi + \frac{\alpha(k+3)}{2(k+1)(k+2)}$$

$$= \frac{\alpha}{(k+2)(k+2)}$$

Moreover,

$$ß_2 = \frac{1}{3}(\xi + \frac{\alpha}{k}) - \frac{\alpha}{k}\,\xi\,(-1)^j \frac{1}{k+j+1}\binom{2}{1}$$

$$= \frac{1}{3}\xi + \frac{1}{3}\frac{\alpha}{k} - \frac{\alpha}{k}\left[\frac{1}{k+1}\binom{2}{0} - \frac{1}{k+1}\binom{2}{1} + \frac{1}{k+3}\binom{2}{2}\right]$$

$$= \frac{1}{3} + \frac{1}{3}\frac{\alpha}{k} - \frac{\alpha}{k}\left[\frac{1}{k+1} - \frac{2}{k+2} + \frac{1}{k+3}\right]$$

$$= \frac{1}{3}\xi + \frac{1}{3}\frac{\alpha}{k} - \frac{2\alpha}{k(k+1)(k+2)(k+3)}$$

$$= \frac{1}{3}\xi + \frac{k^3+6k\ 11}{3(k+1)(k+2)(k+3)}$$

Then,

$$\lambda_3 = 6ß_2 - 6ß_1 + ß_0$$

$$= 6\frac{1}{3}\xi + \frac{k^2+6k+11}{3(k+1)(k+2)(k+3)}$$

Then

$$\lambda_3 = 6ß_2 - 6ß_1 + ß_0$$

$$= 6\left[\frac{1}{3}\xi + \frac{k^2+6k+11}{3(k+1)(k+2)(k+3)}\right] - 6\left[\frac{1}{2}\xi + \frac{\alpha(k+3)}{2(k+1)(k+2)}\right] + \xi + \frac{\alpha}{k+1}$$

$$= 2`\xi + \frac{2k^2+12k+22}{(k+1)(k+2)(k+3)} - 3\xi - \frac{3\alpha(k+1)}{(k+1)(k+2)} + \xi + \frac{\alpha}{k+1}$$

$$= 2k^2 + \frac{12k+\alpha 2-3(k+3)^2+(k+2)(k+3)}{(k+1)(k+2)(k+3)}\alpha$$

$$= \frac{1-k}{(k+1)(k+2)(k+3)}\alpha$$

$$T_3 = \lambda_{3/}\lambda_2 = \frac{1-k}{(k+1)(k+2)(k+3)}\alpha / \frac{\alpha}{(k+1)(k+2)} \quad = \frac{1-k}{k+3}$$

Finally

$$ß_3 = \frac{1}{4}\left(\xi + \frac{\alpha}{k}\right) - \frac{\alpha}{k}\,\xi(-1)^j \frac{1}{k+1+j}\binom{3}{j}$$

$$= \frac{1}{4}\xi + \frac{\alpha}{4k} - \frac{\alpha}{k}\left[\frac{1}{k+1}\binom{3}{0} = \frac{1}{k+2}\binom{3}{1} + \frac{1}{k+3}\binom{3}{2} - \frac{1}{k+4}\binom{3}{3}\right]$$

$$= \frac{1}{4}\xi + \frac{\alpha}{4k} - \frac{\alpha}{k}\left[\frac{1}{k+1} - \frac{3}{k+2} + \frac{3}{k+3} - \frac{1}{k+4}\right]$$

$=\frac{1}{4}\xi + \frac{1}{4}\alpha/k - \frac{\alpha}{k}\left[\frac{6}{(k+1)(k+2)(k+3)(k+4)}\right]$

$=\frac{1}{4}\xi + \frac{k^3+10k^2+35k+5}{4(k+1)(k+2)(k+3)(k+4)}\alpha$

$\lambda_4$ = 20ß$_3$ - 30ß$_2$ + 12ß$_1$ -ß$_0$

$=20\left[\frac{1}{4}\ \xi + \frac{k^3\ +10k^2+35k+50}{4(k+1)(k+2)(k+3)(k+4)}\ \alpha\right] - 30\frac{1}{3}\xi + \frac{k^2+6k+11}{3(k+1)(k+2)(K+3)} + 12\left[\frac{1}{2}\ \xi + \frac{\alpha(k+3)}{2(k+1)(k+2)}\right] - \left[\xi + \frac{\alpha}{k+1}\right]$

$=\frac{k^2-3k+2}{(k+1)(k+2)(1+3)}\alpha = \frac{(1-k)(2\_k)}{(k+1)(k+2)(k+3)(k+4)}\alpha$

Therefore,

$$T_4 = \lambda_4/\lambda_2 = \frac{(1-k)(2-k)}{(k+1)(k+2)(k+3)(k+4)}\alpha \;/\; \frac{\alpha}{(k+1)(k+2)} = \frac{(1-k)(2-k)}{(k+3)(k+4)}$$

## PARAMETERS OF DISTRIBUTION USING THE L MOMENTS

### 4.5.1.1 UNIFORM DISTRIBUTION

$\frac{(\alpha+\text{ß})}{2} = \mu_1$ *1

$\frac{(-\alpha+\text{ß})}{6} = \mu_2$ *2

Solving for α and ß we have

$2\mu_1$= α + ß

α = $2\mu_1$-ß *3

ß = $2\mu_1$ - α *4

Substituting *3 in *2 we have

$\mu_2 = \frac{-(2\mu_1 - \text{ß}) + \text{ß}}{6}$

Solving we have

$6\mu_2$ = -$2\mu_1$ + 2ß2ß = $6\mu_2$ + $2\mu_1$

ß= $3\mu_2$ + $\mu_1$

We can also solve for α by replacing ß in our *3

α = $2\mu_1$- ($3\mu_2$ +$\mu_1$)= $2\mu_1$- $3\mu_2$ –$\mu_1$

= $\mu_1$ -$3\mu_2$

### 4.5.2.1 EXPONENTIAL DISTRIBUTION

$\mu_1 = \xi + \alpha$

If $\xi = 0$ then $\mu_1 = \alpha$

Hence $\alpha = \mu_1 = \frac{1}{x}$

### 4.5.2.1GENERALIZED PARETO DISTRIBUTION

The first and second L moments for the generalized Pareto distribution are

$\mu_1 = \xi + \frac{\alpha}{1+k}$

$\mu_2 = \frac{\alpha}{(1+k)(2+k)}$

If $\xi = 0$ then we need to solve for the parameters

$\mu_1 = \frac{\alpha}{1+k}$

$\alpha = \mu_1(1+k)$

Replacing for α in our other equation we have

$\mu_2 = \frac{\alpha}{(1+k)(2+k)}$

$\mu_2 = \frac{u1(1+k)}{(1+k)(2+k)}$

Cross multiplying we have

$\mu_2(1+k)(2+k) = \mu_1(1+k)$

Simplifying by 1+ k on both sides we have

$\mu_2(2+K) = \mu_1$

$2+K = \frac{\mu_1}{\mu_2}$

$K = \frac{\mu_1}{\mu_2} - 2$

We can now get the parameter estimate α in $\alpha = \mu_1(1+k)$

$\alpha = \mu_1(1+\frac{\mu_1}{\mu_2} - 2)$

$$\alpha = \mu_1(\frac{\mu_2+\mu_1-2\mu_2}{\mu_2})$$

$$\alpha = \frac{\mu_1\mu_2-2\mu_1+ {\mu_1}^2}{\mu_2}$$

$$\alpha = \frac{\mu_1 u2+ \mu_1}{\mu_2}$$

## 4.6 RISK MEASURE

Some of the most frequently used measure of risk in extreme quantile estimation includes value at risk (VaR)and Expected shortfall (ES)and return level. This corresponds to the determination of the value at a given variable exceed with a given probability .This risk measure will be discussed into detail.

### 4.6.1 VALUE AT RISK

VaR is generally defined as the risk capital sufficient, in most instances to cover losses from portfolio over a holding period of a fixed number of days. Suppose a random variable X with a distribution function F describe negative returns on a certain financial instrument over a certain time horizon. Then VaR can be defined as the qth quantile of the distribution F.

$$VaR_q = F^{-1}(q)$$

Where F-1 is the inverse of the distribution(q).The inverse of the distribution at a particular probability level is the called quantile. For risk management q is usually taken to be greater than 0.95and quantile in this case is referred to us Value at risk.

### 4.6.2 EXPECTED SHORTFALL

Another informative measure of risk is the expected shortfall (ES)or the tail conditional expectation which estimates the potential size of loss that exceed VaRq.

$$ES_q = E\lfloor X / X > VaR_q \rfloor$$

Artzner et al (1999)argue that VaR is not a coherent risk measure,but proved that ES is a coherent measure.

Once we know the values of the parameter of the generalized Pareto distribution. We can use them to calculate the value at risk and expected shortfall.

PROOF

VaRq

We start from the fact that the GPD is a good approximation of the excess distribution function.

$$G_{\xi,\delta U}(\text{x}) = F_U(X) = \frac{F(x+\mu)-F(\mu)}{1-f(\mu)}$$

In addition F(u) can be numerically be approximated by

$$\text{F(U)} = \frac{N-N_u}{N}$$

Where N denotes the total number of data and Nu denotes the number of excedances over the threshold u

$$G_{\xi,\delta}(\text{x}) = \frac{F(X+\mu)-F(\mu)}{1-F(u)} \quad = \frac{F(x)-F(u)}{1-F(u)}$$

$$G_{\xi,\delta}(\text{x})[1\text{-F}(\mu)] = \text{F}(\text{x}+\mu) - \text{F}(\mu)$$

$$\text{F}(\text{x}+\mu) = G_{\xi,\delta}(\text{x})\,[1\text{-F}(\mu)] + \text{F}(\mu)$$

And replacing Fu(X) by GPD which is $\left\{1+\frac{\xi(x-\mu)}{\delta}\right\}^{-\frac{1}{\xi}}$ *and*

$$F(u) = \frac{N-N_u}{N} \quad we\ therefore$$

$$= \frac{N_u}{N} G_{\xi,\delta}(x) + \{1\text{-}\frac{N_u}{N}\}$$

$$= \frac{N_u}{N}\{1\text{-}(1+\frac{\xi(X-\mu)}{\delta})^{-1/\xi}\} + \{1\text{-}\frac{N_u}{N}\}$$

$$= \frac{N_u}{N} \text{-} \frac{N_u}{N}(1+\frac{\xi(X-\mu)}{\delta})^{-1/\xi} + 1\text{-}\frac{N_u}{N}$$

$$\text{F(x+u)} = 1\text{-}\frac{N_u}{N}(1+\frac{\xi(x-\mu)}{\delta})^{-1/\xi}$$

Let us denote the probability level of 1 day VaR by α .We want to calculate VaR(1-α) given that the random variable x denote 1 day losses taken with the positive sign we have.

$$P\ (x \leq VaR\ (1-\alpha)) = 1-\alpha$$

$$\Longrightarrow F[VaR\ (1-\alpha)] = 1-\alpha\ = F(x)$$

In addition let xp be such that

$$X\ p + u = VaR(1-\alpha)$$

Hence we have $1-\alpha = F(\ xp + u) \Longrightarrow F(u + x) = q$

Therefore $1-\frac{N_u}{N}(1+\frac{\xi(x-u)}{\delta})^{-1/\xi} = q$

Then replacing for $VaR_q$=x we have

$$Q=1-\frac{N_u}{N}(1+\frac{\xi(VaR\ q-\mu)}{\delta})^{-1/\xi}$$

then solving for VaR q We have

$$(q-1)\frac{N}{N_u} = (1+\frac{\xi(VaR\ q-\mu)}{\delta})^{-1/\xi}$$

$$(1-q)\frac{N}{N_u} = (1+\frac{\xi(VaR\ q-\mu)}{\delta})^{-1/\xi}$$

$$(\frac{qN}{N_u})^{-\xi} = 1+\frac{\xi(VaR\ q-\mu)}{\delta}$$

$$(\frac{qN}{N_u})^{-\xi} - 1 == \frac{\xi(VaR\ q-\mu)}{\delta}$$

$$VaR_q - \mu = \frac{\delta}{\xi}((\frac{qN}{N_u})^{-\xi} - 1)$$

$$VaR_q = \mu + \frac{\delta}{\xi}((\frac{qN}{N_u})^{-\xi} - 1)$$

**EXPECTED SHORTFALL**

$$ES_q = VaR_q + E(X - VaR_q / x > VaR_q)$$

The second term on the right is the expected exceedances over threshold VaRq.This is also known as the mean excess function for GPD with parameter ξ<1 which corresponds to the following expression.

$$E(u) = E(x-u/x>u) = \frac{\delta + \xi\mu}{1-\xi} \qquad \text{for } \delta + \xi u > 0$$

Therefore we get

$$ESq = VaR_q + \frac{\delta + \xi(VaRp - \mu)}{1-\xi}$$

$$= \frac{VaRP(1-\xi) + \delta + \xi VaRp + \xi\mu}{1-\xi}$$

$$= (VaRp - \xi\ VaRp + \delta + \xi\ VaRp + \xi\mu)/(1-\xi)$$

$$ESq = \frac{VaRp + \delta - \xi\mu}{1-\xi}$$

The function gives the average of the excesses of X over varying values of the threshold x.

# CHAPTER FIVE

## APPLICATION

This section presents the procedure which was used in the study .It explains in detail the steps that were encountered in the modeling process which includes the data processing and analysis.

### 5.1 SCOPE OF THE DATA

Secondary data from U.A.P insurance company regarding fire industrial claims for the period 2002-2013 was used in this study. Three assumptions were made on the data before use.

1 All claims came from the same distribution.ie they were independent and identically distributed.

2 There were no zero claims for any fire policy sold.

3 Future claims were to be generated from the same distribution.

### 5.2 ACTUARIAL MODELING PROCESS

This section will describe the steps that were followed in fitting a statistical distribution to the extreme claim severity .These steps include

1) Selecting the model family of distributions.

2) Exploratory data analysis.

3) Choosing the threshold.

4) Estimating the parameters.

5) Goodness of fit test.

### 5.3 SELECTING THE MODEL FAMILY

Here considerations were made of a number of parametric probability distributions as potential candidates for the data generating mechanism for extreme claims. Most data in general insurance is skewed to the right and therefore most distributions that exhibit these characteristics can be used to model the extreme claims.

However the list of potential probability distributions is enormous and it is worth noting that the choice of distributions is to some extent subjective.

For this study the choice of the sample distributions was with regard to

-Prior knowledge and experience in curve fitting.

-Time constraint

-Availability of computer soft-ware to facilitate the study.

-The volume and quality of data.

Therefore three distributions were used including:Pareto,Exponential and Uniform

**5.4 EXPLORATORY DATA ANALYSIS**

It was necessary to do some descriptive analysis of the data to obtain the salient features .This involves the Mean,Median,Mode,Standard Deviation,Skewness and Kurtosis. This was done using MATLAB programming language and also manual calculation..

**5.5 CHOOSING A THRESHOLD**

A good choice of threshold is important. Balancing act was applied where if a lower threshold is chosen it results in a large model error with more data and smaller parameter error. High threshold results in smaller model error and a less data with larger parameter error.

**5.6 QUANTILE-QUANTILE (Q-Q) PLOTS**

The quantile-quantile plots are graphical techniques used to check whether or not a sampled data set could have come from some specific target distribution .i.e. to determine how well a theoretical distribution models the set of sampled data provided. This study used the Q-Q plot to check for the distribution that fit the sample.

The first q stands for the quantile of the sampled data set and the second q stands for the quantile of the distribution being checked whether it fits the data. Q-Q plot is a plot of the target population quantile (y)against the respective sample quantile (x).

If the sample data follow the distribution suspected ,then the quantile from the sample data would lie close to where they might be expected and the points on the plot would straggle about the line y=x.

In addition. If the Q-Q plot deviates significantly from a straight line,then either the shape parameter is inaccurate or the model selection is untenable. If the graph is concave this indicates a fat tailed distribution, whereas a convex shape is an indication of a short –tailed distribution.

Theoretically,In order to calculate the quantile of the distribution,this target distribution must first be specified .i.e.Its population mean and standard deviation but in practice,the sample estimates are used, thereforesample mean and standard deviation of the distribution were estimated to be same as ones of the sampled data set.

### *5.6.1ADVANTAGES OF Q-Q PLOT*

1The sample sizes do not need to be equal.

2Many distributional aspects can be simultaneously tested for example shifts in locations, shifts from scale, changes in symmetry and the presence of outliers. This is important because if the claims amount and the claim count of the data set comes from population whose distributions differ only in location, the points should lie along a straight line that is placed either up or down from the 45-degree reference line..

### 5.7RESOURCES

The study required a computer preferably 250GB hard disk,1 GB Ram,1-73GHZ  dual processor any model.It also required a Microsoft office package,Matlab Program,Microsoft office especially Excel and R program.

## 5.8COMPUTATION AND INTERPRETATIONS

### 5.8.1SPECIFIC OBJECTIVES

Testing for the appropriate statistical distribution for the claim amount

Test the goodness of fit of the chosen distribution

### 5.8.2VARIABLE

The random variables used in the study were the fire claim amount reported and claimed at UAP Insurance.

### 5.8.3DESCRIPTIVE DATA ANALYSIS

Mean = 6.2966e+005

Variance =1.0395e+013

Median =5.4276e+004

Mode = 27840

Skewness= 17.0740

Kurtosis =381.5717

Number of observations =902

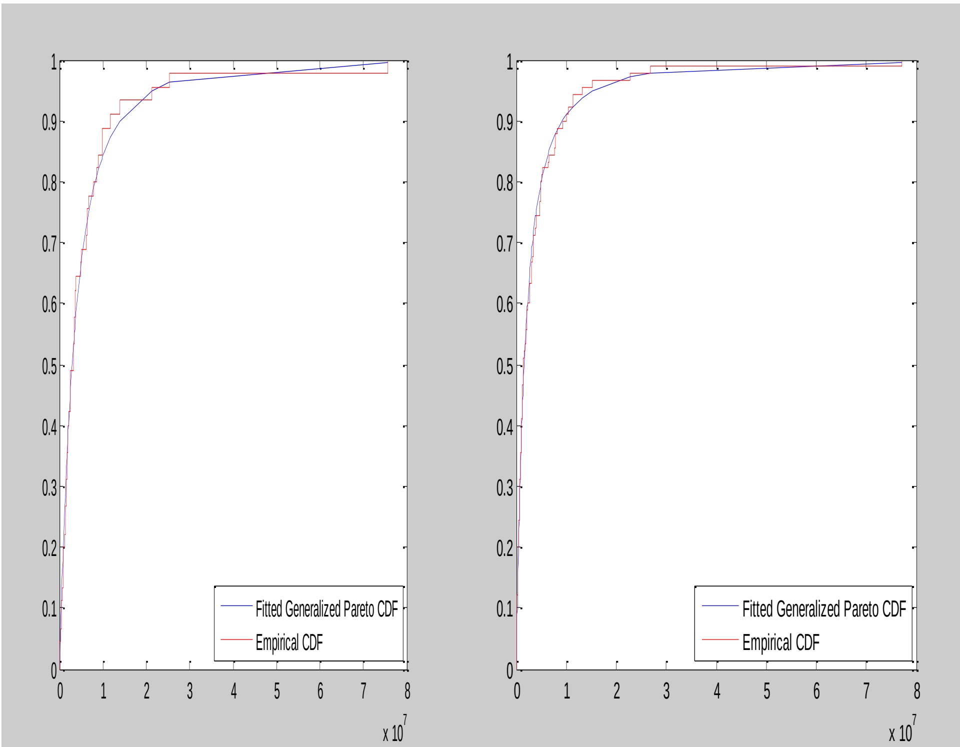
The data according to descriptive statistics shown above indicates that the data is skewed to the right(skewedness coefficient of 17.07)Right –skewedness means that the right tail is long relative to the left tail.

Kurtosis is a measure of whether the data is peaked or flat relative to a normal distribution. The loss data set with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails.

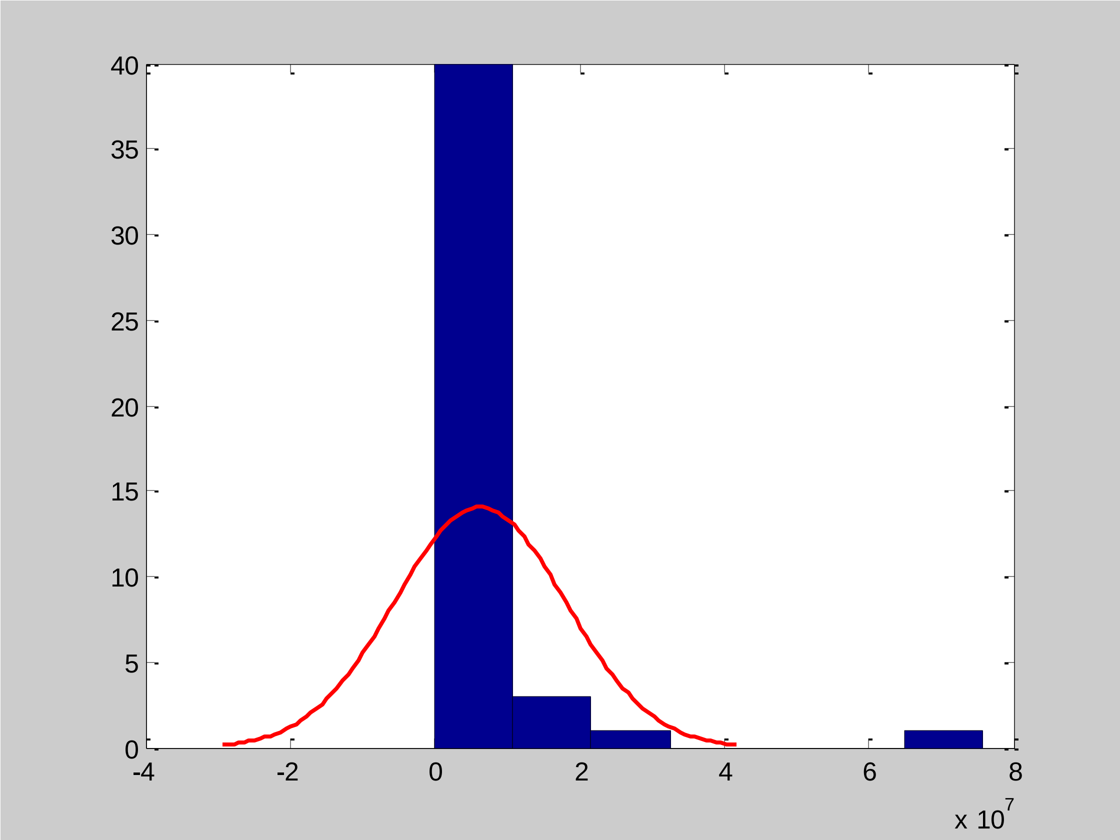
### 5.8.4CHOICE OF THE THRESHOLDS

The study sampled data using two thresholds the 90% and 95%quantile.A comparison of the two samples was done by plotting a generalized Pareto distribution.A threshold of 95% quantile was chosen which has 45 claims.

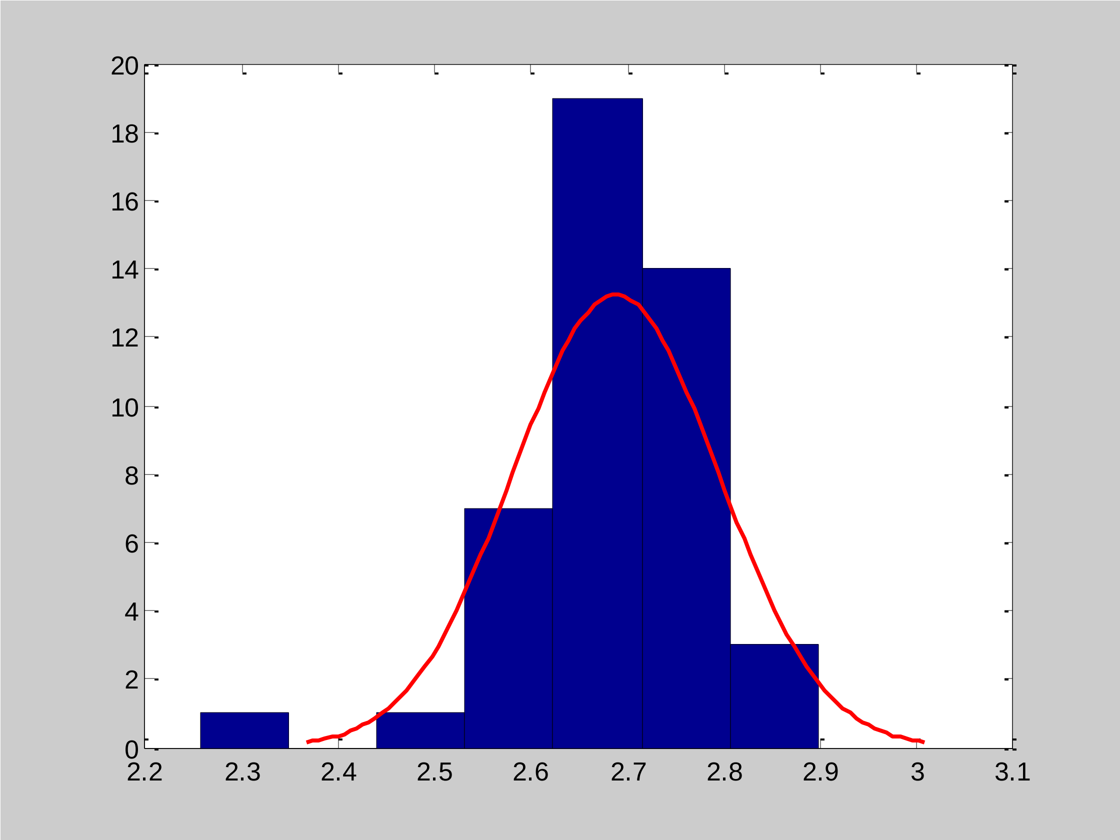
(a) ( b)




Graph b from a threshold of 95%quantile is more fitting as compared to (a)which is from a threshold of 90%




The histogram plot of claims with a normal distribution shows that the data highly skewed and hence for any analysis we need to take the first ,second and third logarithm to reduce skewness and normalize the data.




Histogram of claims with a normal distribution superimposed on it shows that the data is normalized. This is after taking the third natural logarithms.

**5.8.5 THE PARAMETER ESTIMATION**

The parameters of the three distributions were estimated and compared for the three methods of estimation.Its clear that Pareto distribution had the least estimate of the three distribution followed by exponential and finally uniform distribution meaning that Pareto distribution was the best fitting distribution.

| DISTRIBUTION | MLE | MOM | L-MOMENT |
|---|---|---|---|
| Exponential | Mu=0.9878 | Mu=0.9878 | Mu = 0.987803 |
| Uniform | a = 0.8136<br>b = 1.0641 | a=0.9174<br>b=1.058 | a=3.92<br>b= -1.944 |
| Pareto | k = -1.94<br>Sigma = 2.065 | K= - 1.9309<br>Sigma= 2.08 | K= -2.1893<br>Sigma= 1.9893 |
| Generalized Pareto | K= -1.94<br>Sigma = 2.065 | K= 1.0309<br>Sigma= 2.08 | k = -.9893<br>Sigma = 1.9893<br>Theta = 0 |

The maximum likelihood estimate of Exponetial distribution was equal for the three methods of estimation.For uniform distribution the parameter estimate was in the range of 0.8136 and 1.06 for the method of moment and the maximum likelihood method. The L moment method gives results that are in the range of -1.944 to 3.92.meaning it does not give the best estimate with the least error.

For Pareto distribution the three methods of estimation gives result in the range of -2.1893 to 2.08. The three methods are almost giving results that are almost equal. Method of maximum likelihood gives the least estimate hence the best estimation method. Therefore with the two distributions that is exponential and Pareto distribution we shall carry out Q-Q test to come up with the best fit.

Therfore we conclude that Pareto distribution is the best fitting distribution of the three distribution but further analysis needs to be done to check confirm that Pareto is the best distribution.

**5.8.6Q-Q PLOTS**

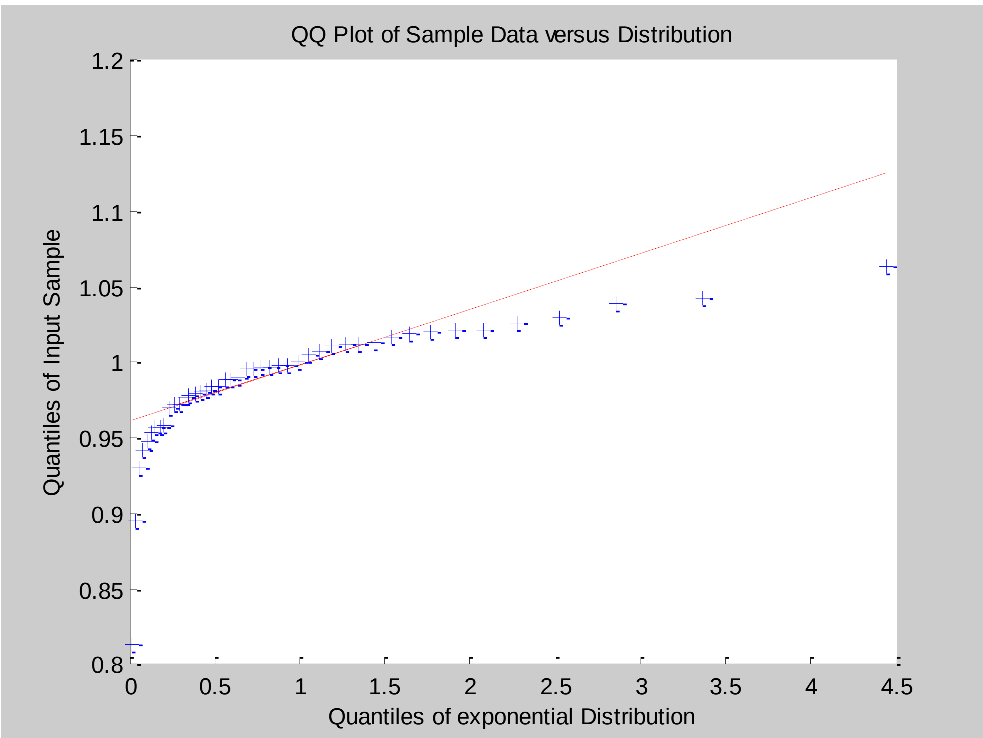
HO:The statistical distribution provides the best fit for extreme claims data.

H1:The statistical distribution does not provide the best fit for the extreme claims data.

***5.8.6.1EXPONENTIAL***

H0:The exponential distribution provides the correct statistical model for the extreme claims data.

H1:The exponential does not provide the correct statistical model for the extreme claims.




This figure shows the 70% of the points are lying close to the abline.The plot depicts that the exponential distribution has light tails on both ends as most points are not falling on the reference line.

Conclusion

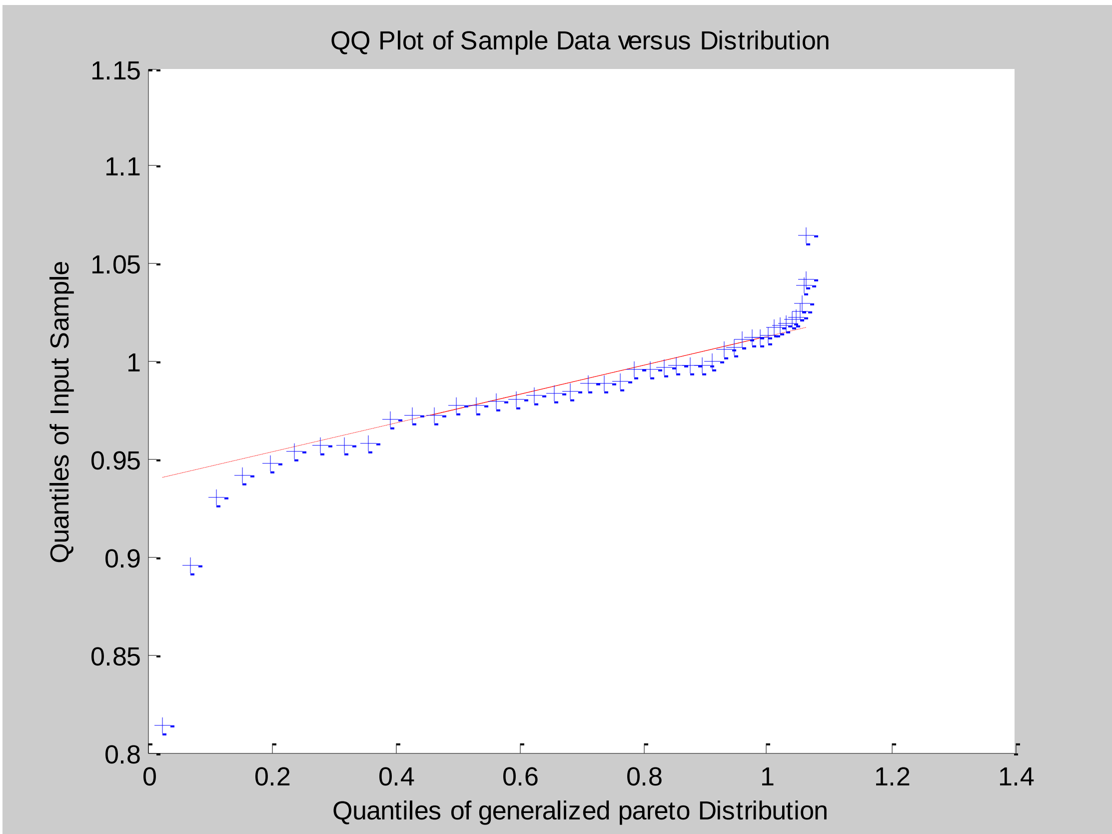
The null hypothesis was therefore rejected and a conclusion was made that the exponential distribution does not provide the correct statistical model at 99% confidence level.

**5.8.6.2GENERALIZED PARETO DISTRIBUTION**

H0: Generalized Pareto distribution provides the correct distribution for extreme claim data

H1: Generalized Pareto distribution does not provide the correct statistical distribution for extreme claims data

.




The Q-Q plot shows the best distribution where almost all the points are lying on the line,except few that are outliers but are very close to the line if you keenly compare the points they lie closest to the line.

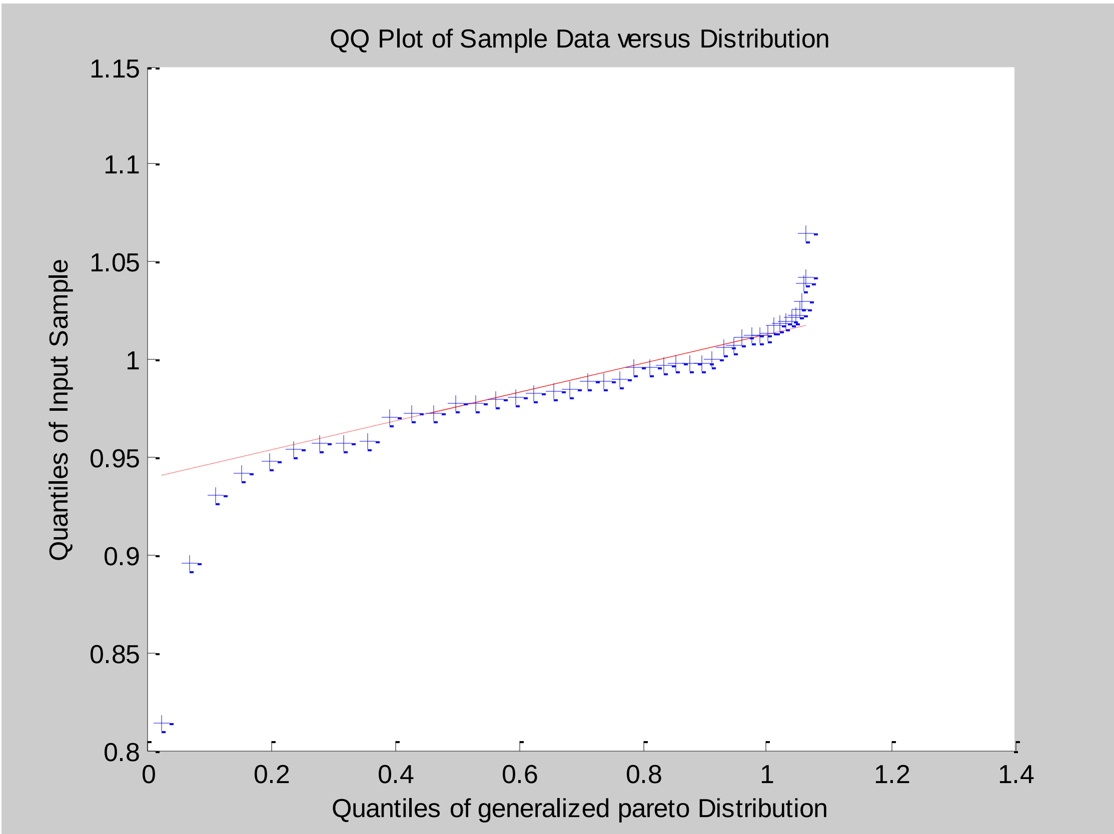
CONCLUSION

We do not reject the hypothesis and hence we conclude that generalized Pareto distribution provides the correct distribution for extreme claims data at 99% confidence level.

**5.8.6.3 Q-Q PARETO DISTRIBUTION**

HO: Pareto distribution provides the correct claims data set

H1: Pareto distribution does not provide the correct claims data set.

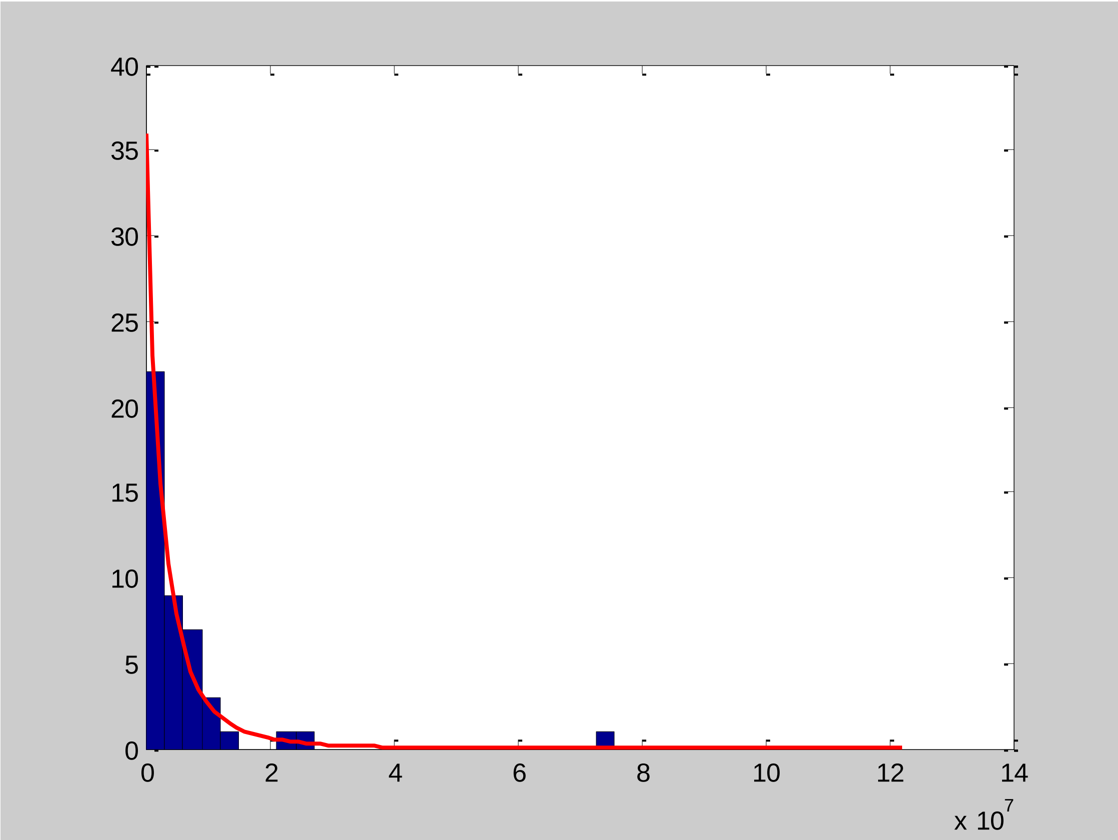


### 5.8.7Histogram Plot

To come up with the best distribution of the two selected distributions histogram plot of data with a distribution were plotted.

HO: Claims data comes from Pareto distribution

H1:Claims data does not come from Pareto distribution

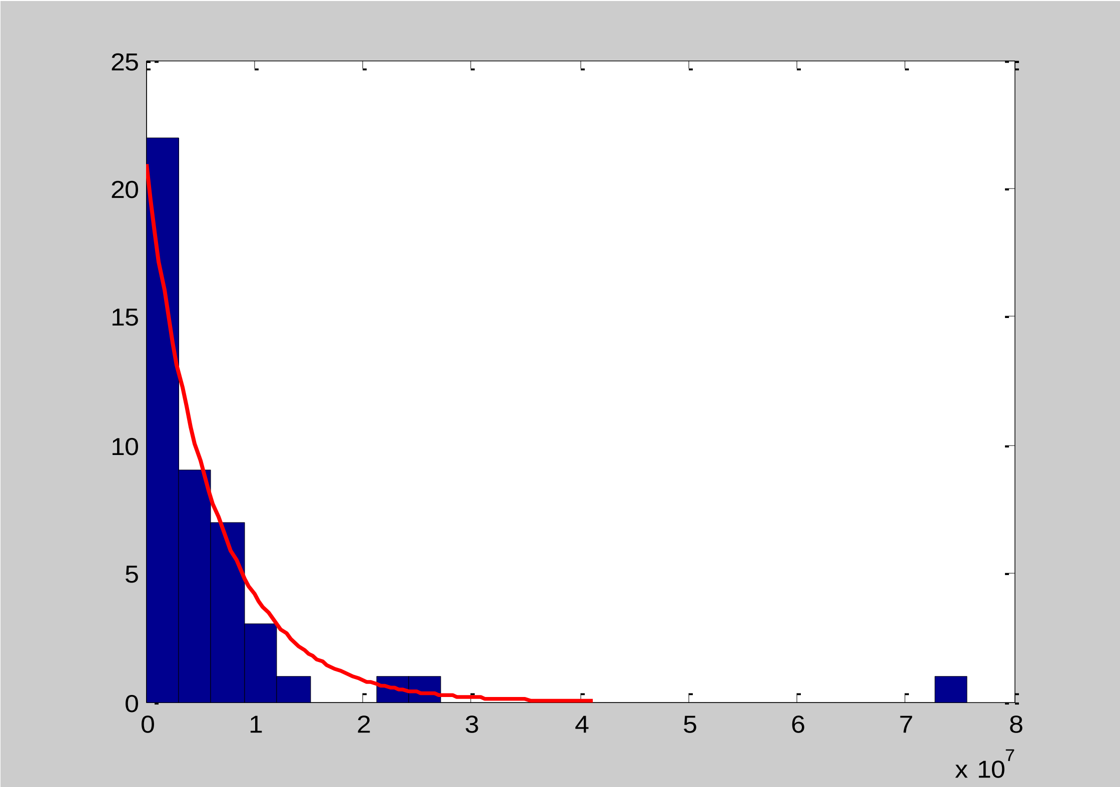



This histogram shows that about 90% of the points are lieing on the distribution line.Meaning that,about 10% of the points are lieing outside the distribution line. These points lieing outside the distribution curve are the ouliers, in this case they are the extremely high claims. Thus making Pareto to be the best fitting distribution

### 5.8.8Exponential Histogram Plot

H0:Claims data comes from exponential distribution

H1:Claims data does not come from exponential distribution




This histogram shows that about 80% of the points are lieing on the distribution curve.Meaning about 20% are lieing outside the distribution curve .Thus making it to be the least fitting distribution

CONCLUSION

The two histogram shows clearly that Pareto distribution is the best fitting distribution of the claims data.

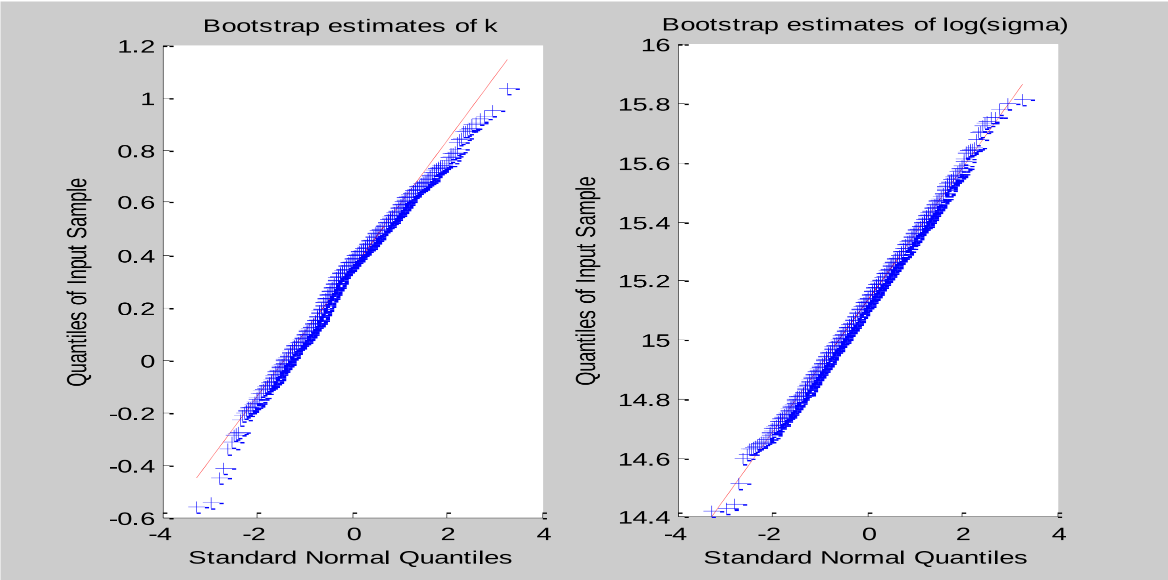
## 5.8.9BOOTSTRAP CONFIDENCE ESTIMATES

The study went ahead to find the confidence intervals of the generalized Pareto distribution estimates.

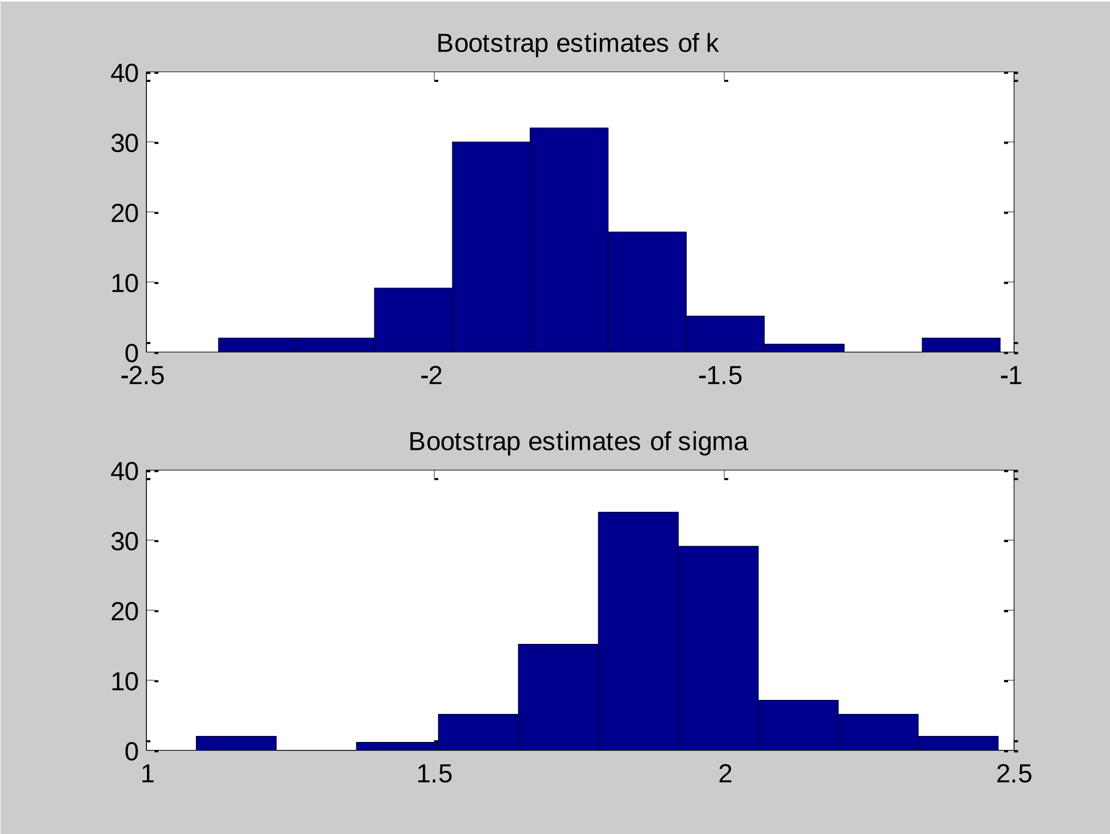
Bootstrapping is the practice of estimating properties of estimators (such as the variance)by measuring those properties when sampling from an approximating distribution. This can be implemented by constructing hypothesis test.

The bootstrap method involves taking the original set of N heights and using a computer samplingfrom it to form a new sample called a resample or bootstrap sample that is also of size N.The bootstrap sample is taken from the original using sampling with replacement so it is not identical with original.This process is repeated a large number of times typically 1000 times. Then for each of the bootstrap sample we compute the mean and standard deviation. Using the estimated parameters we fit a Q-Q plot of the parameters. We therefore obtain the histogram of Generalized Pareto distribution parameters which gives us the confidence intervals of those parameters.

### *5.8.9.1QQ PLOTS OF BOOTSTRAP ESTIMATES*



*5.8.9.2 HISTOGRAMS*




## 5.8.10 VALUE AT RISK AND EXPECTED SHORTFALL

We have calculated for the Value at risk and expected shortfall using a threshold of Ksh 2450300 giving me 45 samples exceeding threshold from a sample of 902 claims reported in an insurance company.

| Quntile | VaR | ES |
|---|---|---|
| .90 | 24500O9 | 2450201 |
| .95 | 2449977 | 2450188 |
| .99 | 2449950 | 2450181 |

The table shows that an increase in the quantile results to a decrease in the value at risk.Meaning that, Huge number of the claims if occurred will be borne by insurer himself.It also leads to a decrease in the expected shortfall. This is the probable loss that would result incase a risk happens at a certain quantile.

# CHAPTER SIX

## SUMMARY AND CONCLUSION

### 6.1 SUMMARY

The major objective is to come up with one statistical distribution that fits the extreme claims data well and to test how well this statistical distribution fits those extreme claims so that this distribution can be used for modeling the extreme claims.

The data that wasanalyzed came from an Insurance company which were fire claims from 2002 to 2013.The data had 902 claims. I carried out the descriptive analysis of the data where the mean was found to be 6.2966e+5 , The variance is 1.0395e+13, Skewnessis 17.0740 and kurtosis 381.5717.From the descriptive statistics it shows the data is heavy tailed hence extreme value theory was applicable. Two different thresholds were selected to sample the extreme claims and the normal claims. The thresholds chosen were beyond 95% percentile and beyond 90% percentile. The peaks of threshold graph shows that the 95% data is the best to use hence I went ahead to use it discarding data below the 95% percentile. A threshold of 2,450,300Ksh was used.

The parameter estimates of the three distributions were compared and Pareto distribution came up to be the best fitting distribution.The Q-Q plots indicate that most points of the Pareto distribution are lying along the reference line thus making it the best distribution family in preliminary stage. A histogram of claims with two selected distributions also pointed that Pareto distribution was the best fitting distribution among the three distributions, followed by Exponential distribution, while the Uniform distribution was the worst distribution with the claim points lying very far away from the abline of the Uniform distribution..

Bootstrap method was carried out where I got a another sample with replacement then estimated the parameters of the generalized Pareto distribution, plotted the QQ plots of the sample then estimated the confidence intervals of the parameters to be for scale parameter and shape parameter

### 6.2 CONCLUSIONS

We have shown that GPD can be fitted to fire insurance loss severity. When the data exceeds a high threshold, the GPD is a useful method for estimating the tails of loss severity distributions. It also means that the GPD is theoretically well supported technique for fitting a parametric distribution to the tails of unknown distribution.

### 6.2.1 RECOMMENDATION

I would like to recommend future researchers to model the tail forms of other forms of insurance. Secondly, from a risk management viewpoint,constructing a useful management technique for avoiding large claims would be an interesting line of research. In addition to, I would like to recommend a similar research using the extreme value distributions family still employing  the three methods of estimation or come up with other methods of estimation.

### 6.2.2CHALLENGES ENCOUNTERED

The scarcity of observations from the tail region of the distribution is small compaired to the sample size required to estimate the form of the tail region with significant power.There is a bias –variance tradeoff issues when deciding the number of upper order statistics to use in the analysis.

# REFERENCES

.

CT 6, society of actuaries

Castillo, E. and A.S.Hadi (1997). Fitting the Generalized Pareto distribution to data

Castillo J.Daoudib J. (2008).Estimation of the generalized Pareto distribution, statistics & probability letters.

Cebrian, Denuitand Lambert (2003) .Generalized Pareto fit to the society of actuaries large claims database.

Dress, H., L. de Haan, and S. Resnick. 1998. *How to Make a Hill Plot*.

Embrechts, P. 2000. Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool.

Embrechts, P., L. de Haan, and X. Huang. 1999. Modelling Multivariate Extremes.

Embrechts, P. S. Resnick, and G. Samorodnitsky. 1998. Extreme Value Theory as a Risk Management Tool.

Embrechts, P. C. Kluppelberg, and T. Mikosch. 1997. *ModellingExtremal Events for Insurance And Finance.*

Embrechtsand Kafetzaki. 1997. A Survival Kit on Quantile Estimation.

Greenwood, J.A.Landwehr.j.mMatalas, N.C and Wallis J.R (1979) probability weighted moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse form.

Hosking, J.R.M (1985). Maximum likelihood estimation of parameters of the generalized extreme value distributions.

Huisman, R., G.K. Koedijk, and A.J.R. Pownall. 1999. *Asset Allocation in a Value at Risk Framework.*Erasmus University, Rotterdam, Faculty of Business Administration.

Jorion, P. 1996. *Value at Risk—The New Benchmark for Controlling Derivatives Risk*.

Jesus Gonzalo and Jose Olmo (2004). Which extreme values are really extremes.

JelenaJocovic (2012). Quantile estimation for the Generalized Pareto distribution

Hosking &Wallis (1987).Parameter and quantile estimation for the generalized Pareto distribution.

JMartin Van Zyl(2010). Estimation of the shape parameter of a generalized Pareto distribution based on transformation to Pareto distributed variables.

Kluppelberg, C. and A. May. 1998. “The Dependence Function for Bivariate Extreme Value Distributions.
Longuin, M.F. 1999. “From Value at Risk to Stress Testing: The Extreme Value Approach.”

Li-Hua Lai and Pei-HsuanWu(2005)” Estimating the threshold value and loss distribution:Rice damaged by typhoons in Taiwan.
McNeil, A.J. 1999. “Extreme Value Theory for Risk Managers.”*Risk* special volume, to be published.
McNeil, A.J. and T. Saladin. 1997. The Peak over Thresholds Method for Estimating High extreme values.
McNeil (1999).Extreme Value theory for risk managers

Mohammed Soliman (2009). The properties of L-moments compared to conventional moments.

Matthew J.Pocernich (2002). Application of extreme value theory and threshold model to hydrological events.

MeelisKariik (2002).Estimation of loss distribution and risk measures.
Theory

McNeil & Saladin (1997). The peaks over threshold method for estimating high quantile of loss distribution

McNeil 1999).Quantiles of Loss Distributions.
Mierlus-Mazilu (2010). On Generalized Pareto distribution
Mendes and Lima (2005). Modeling clusters of Extreme losses

Patricia de ZeaBermudes (2011) .The importance of the Generalized Pareto distribution in the extreme value

Resnick, S. and C. Starica. 1996. Tail Index for Dependent Data.
Smith, L.R. 1987. Estimating Tails of Probability Distributions.

# APPENDIX

## MATLAB PROGRAM CODE

```
helpxlsread
g=xlsread('C:\Users\ROBERTS\Desktop\roba\uaplingm.xls',1,'C1:C903');
g
% decriptive data statistics
A=mean(g)
B=var(g)
C=median(g)
D=mode(g)
E=skewness(g);
F=kurtosis(g)
%Selecting the threshold at 95%extreme
q=quantile(g,0.95);
q
%Mean Excess Function
mydata=g(g>q)-q;
mydata
klop=mean(mydata)
c=length(mydata)
J1=mydata
J1
y=mydata
%Transforming the data to reduce it skewness by taking los
lx=log(mydata);
llx=log(lx);
lllx=log(llx);
figure(9)
hist(mydata)
myalpha=0.01;
%evaluate the parameters
[expparms,expci]=expfit(lllx,myalpha);
[gamparms,gamci]=gamfit(lllx,myalpha);
[wblparms,wblci]=wblfit(lllx,myalpha);
%log likehood function is calculated
o=expfit(lllx,myalpha)
o
k=wblfit(lllx,myalpha)
k
t=gpfit(lllx,myalpha)
r=gpfit(lllx,0.95)
r
p=gpfit(lllx,0.995)
p;
m=histfit(mydata)
m
figure (12)
histfit(llx)
figure (11)
```

```
histfit(lllx)
figure (10)
llllx=log(lllx)
histfit(llllx)
paramEsts = gpfit(y)
kHat      = paramEsts(1)  % Tail index parameter
a=kHat
sigmaHat  =paramEsts(2);   % Scale parameter
b=sigmaHat
bins = 0:10:25
h = bar(bins,histc(y,bins)/(length(y)*.25),'histc')
set(h,'FaceColor',[.1 .1 .1])
ygrid = linspace(0,1.1*max(y),1)
figure (40)
line(ygrid,gppdf(ygrid,kHat,sigmaHat))
xlim([0,10]); xlabel('Exceedance'); ylabel('Probability Density')
[F,yi] = ecdf(y)
plot(yi,gpcdf(yi,kHat,sigmaHat),'-')
hold on; stairs(yi,F,'r'); hold off;
legend('Fitted Generalized Pareto CDF','EmpiricalCDF','location','southeast')
replEsts = bootstrp(100,@gpfit,lllx)
figure (6)
subplot(2,1,1), hist(replEsts(:,1)); title('Bootstrap estimates of k');
subplot(2,1,2), hist(replEsts(:,2)); title('Bootstrap estimates of sigma');
figure(5)
subplot(1,2,1), qqplot(replEsts(:,1)); title('Bootstrap estimates of k');
subplot(1,2,2), qqplot(log(replEsts(:,2))); title('Bootstrap estimates of log(sigma)');
[paramEsts,paramCI] = gpfit(lllx);
h = lillietest(lllx,alpha)
alpha1=0.001
h = lillietest(lllx,alpha1)
alpha2=0.001
h = lillietest(lllx,alpha2,'ev')
alpha3=0.05
s=gpcdf(lllx)
P = gpcdf(lllx,a,b,1)
figure (45)
nbins=25
histfit(mydata,nbins,'exponential')
figure (46)
histfit (mydata,nbins,'gp')
figure (47)
PD3 = fitdist(lllx,'exponential')
PD3
PD4 = fitdist(lllx,'gp')
PD4
[m,v]=unifstat(A,B)
helpqqplot
figure (16)
qqplot(lllx,PD4)
figure (18)
qqplot(lllx,PD3)
figure (19)
v=gpfit(lllx,0.01)
z=unifit(lllx,0.01)
U=expfit(lllx,0.001)
```

```
ALPHA5=0.05
H = KSTEST(mydata,PD4,ALPHA5,'unequal')
K1 = mle(lllx,'distribution','exponential')
K2 = mle(lllx,'distribution','gp')
K5 = mle(lllx,'distribution','uniform')
nlogL=-explike(expfit(mydata),mydata)
nlogL=-gplike(gpfit(mydata),mydata)
[p,ci] = gpfit(lllx)
figure (22)
qqplot(lllx,K5)
paramEsts=K2
[nll,acov] = gplike(paramEsts,mydata);
StdErr = sqrt(diag(acov))
[ahat,bhat] = unifit(mydata)
figure (27)
plot(lllx)
figure (28)
plot(llx)
figure(29)
plot(lllx,'o')
PD4
figure (30)
hist (lllx)
bo=moment(mydata,1)
bo
b1=moment(mydata,2)
b1
b2=moment(mydata,3)
b2
lo=mean(lllx)
lo
lp=var(lllx)
lp
```