FINDING APPROPRIATE LOSS DISTRIBUTIONS TO INSURANCE DATA

Case Study: Kenya (2010-2014)

NDUWAYEZU FLORENT – 077862

Submitted in partial fulfillment of the requirements for the degree of Bachelors of Business Science - Actuarial Science at Strathmore University

Strathmore Institute of Mathematical Sciences

Strathmore University

Nairobi, Kenya

JULY 2016

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Proposal contains no material previously published or written by another person except where due reference is made in the Research Proposal itself.

Nduwayezu Florent

…………….. [*Signature*]

…………….. [*Date*]

This Research Proposal has been submitted for examination with my approval as the Supervisor.

John Ocheche

…………….. [*Signature*]

…………….. [*Date*]

School of Finance and Applied Economics

Strathmore University

ACKNOWLEDGEMENT

I would like to thank God for the favour, strength and ability to do this research study. I extend my sincere gratitude to the family, especially my parents, and friends who have encouraged me to do this study.

To my supervisor, John Ocheche, I highly appreciate your patience, guidance, time and effort you have put in to help me through this research project from the start. Thank you.

Finally, to the School of Finance and Applied Economics, gratitude for having me as your student.

CONTENTS

CHAPTER 1: INTRODUCTION

1.1. Background

In the insurance industry, specific probability distributions are used to model the individual claims arising from the policies that the insurance company has issued: These are referred to as loss distributions (Achieng, 2010).

Obtaining the total amount of claims for a specific period is a vital part of the daily work of insurance companies. This will help in various ways the management in running the company (Jouravlev, 2009). For instance, the insurance company will be able to calculate the premium for a type of policy by the use of the claim experience. Moreover, it will be able to reserve a certain amount of money to cover the cost of future claims. Premium computation and Reserving are not the only reasons for which loss distributions are needed. Loss distributions are also utilised in reviewing reinsurance arrangements and also in testing for solvency. This explicitly highlights the importance of loss distribution in the insurance industry.

This paper therefore aims to determine the most suitable loss distributions for various sort of insurance contracts being general or life insurance in the Kenyan market industry. The following distributions will be compared: the exponential distribution, the Pareto distribution, the Generalised Pareto distribution, the lognormal distribution, the Weibull distribution & the Burr distribution. We will see how these distributions can be tailored in order to suit the observed data. Afterwards, a test of goodness-of-fit will be used to determine the level of robustness of the distribution in fitting the given data. The loss distributions will also be used in order the probabilities of future events happening.

1.2. Problem Statement

For different types of insurance policies, specific distributions are commonly used to model data corresponding to the given data available. For instance, for fire insurance, discrete distributions such as the Poisson and the Negative Binomial distributions are preferred due to their believed suitability to that type of data. (Boadi et al., 2015). For other types of insurance policies, such as motor insurance, considered as a highly risky type of insurance, requires thoughtful measures in order to handle them (IRA, 2013). According to Achieng (2010), claims arising from this type of insurance can be modelled using a lognormal distribution given that it suits better the claims than any other distribution. Additionally, there are other applications of loss distributions in other insurance sectors such as health insurance. In fact, Paranaiba et al.

(2011) showed that an improved Burr distribution can have an application in the health business.

However, in reality, it is not that easy to find the appropriate distribution to use for a given data. This is due to the fact that the loss distributions, commonly used for certain types of insurance, have limitations. For example, for the lognormal distribution, it is unclear which base of logarithm should be chosen. Moreover, the nature of the distribution makes it difficult to interpret its skewness and kurtosis (Limpert et al., 2001).

In a fast growing industry such as the Kenyan insurance market (IRA, 2015), it is hence substantial to determine the adequate loss distributions that can be used in the insurance industry.

In his assessment of the challenges that the Kenyan insurance companies are facing, Kiragu (2014) concluded that the development of suitable products constitutes a considerable obstacle to the Kenyan market industry. The use of better loss distributions can help in reducing the impact of that barrier.

There is therefore need to determine the different types of loss distributions that would be suitable for the various types of insurance policies that are offered and also determine to which such as a model should be trusted in modelling the levels of incoming claims in a given period of time.

1.3. Research Objectives

The objective of this research is to determine for the adequate loss distribution for the claim amounts by applying a goodness-of-fit test on the selected loss distribution

1.4. Research Questions

Throughout the research, the paper will be seeking answers to the following questions:

1. What should be the loss distribution for the given claim amounts?
2. How well does this distribution fit the data?

1.5. Justification

The purpose of the study is to determine an appropriate loss distribution for a given set of data. This is of relevance in the current insurance world because it provides the insurance company with better models that are used in their day-to-day activities. The advantage of this study is that it will help the insurance companies in obtaining more accurate results when dealing with various forms of claims and hence they will manage to minimise risk in a better way. Moreover, the research will benefit policyholders: Given that risk associated with claims has been minimised, the premiums that the policyholder pays will consequently be reduced.

CHAPTER 2: LITERATURE REVIEW

2.1. The Burr XII (BXII) distribution

This section will try to highlight how various researchers have tried to approach the concept of loss distributions. We will start with Thomas Wright. In 2005, Wright, in his studies, tried to find the most approximate loss distribution for data consisting of 490 claims. These claims had been collected over a period of seven years. The approach used was the Maximum likelihood approach for each single year. The statistical distributions used were the inverse Pareto, Pareto, burr, Pearson VI, inverse burr log-normal, Burr XII (BXII) and the restricted benktander families. The benktander, the Pareto and the exponential families have a special property: if their truncation is taken separately, it will give another distribution in the same family. The benktander family mainly emphasises on excess claim amounts given a certain number of claims. However, for his study, he was dealing with whole claim data set and hence the benktander family could not be used for the given study.

One of the realisations was that the cumulative distribution function (cdf) and the reliability of the BXII distribution can be written in a closed form and this hence greatly shortens the computation of the percentiles and the likelihood function for a given set of censored data. Some of the main characteristics of this distribution is that tails that are most the time algebraic. Zimmer et al (1998) had already established that the BXII distribution can be significantly relevant for modelling failure time data. This is due to its characteristic of an algebraic tail. Shao (2004a) confirmed the findings Zimmer et al had come up with earlier. Shao et al (2004b) continued the study of this distribution by analysing the distribution in relation with the good frequency analysis. Other researchers focused on the same distribution: Soliman (2005) stated that the distribution can be used as a provisional distribution for a given set data whose original distribution is not identified. Another group of researchers who also spent time on this distribution are Wu et al. (2007) who tried to find a solution the estimation problems arising from the use of this distribution. To achieve this, they used the progressive type II censoring where removals at each time were random and followed a discrete uniform distribution.

Mathematically, the followings are respectively the cdf and the pdf of the BXII distribution:

$$F(y; v, l, s) = \left(1 - (1 + \left(\frac{y}{s}\right)^{v}\right)^{-l}$$

$$\text{And } f(y; v, l, s) = vls^{-v}y^{v-1}\left[1+\left(\frac{y}{s}\right)^{v}\right]^{-l-1}$$

Where l > 0 and v > 0 are shape parameters and s > 0 is a scale parameter. The n^{th} moment about zero of the BXII distribution (for n < lv) is given by:

$$\partial'_n = s^n * l * B\left(l - \frac{n}{c}, \frac{n}{c} + 1\right)$$

2.2. Exponentiated Inverted Weibull distribution

In the past years, the Exponential Inverted Weibull Distribution ((EIW) has been used by various researchers. The reason for that is the ability of this distribution to approach different distributions when its shape parameter changes (KAN, PASHA, and H. PASHA, 2008).

Flaih et al. (2012) discussed various aspects of the standard Exponential Inverted Weibull, the ancestor of the EIW distribution, such as its moments, its median, MLE… In their study, they tried to determine of the distributions, between EIW and the Inverted Weibull would fit better a certain set of data. As results, they arrived to the conclusion that EIW distribution would be fit better the given data.

Other researchers such as Aljouharah Aljuaid (2013) also dealt with this distribution. He used the Bayes and classical estimators in order to estimate the parameters of the EIW distribution. Under type I censoring, Hassan (2013) used the EIW distribution in order to model the optimal designing of failure step-stress partially accelerated life tests with two stress levels. Another Notable use of the EIW distribution can be found in the studies of Hassan et al. (2014). In this study, an estimation of the population parameters for the EIW distribution was carried out based on grouped data with equi and unequi-spaced grouping.

Mathematically, the followings are respectively the cdf and the pdf of the EIW distribution:

$$F(x) = \left(e^{-x^{-\mu}}\right)^{\theta}$$

Where x, θ, μ > 0

$$f(x) = \mu * \theta * \left(e^{-x^{-\mu}}\right)^{\theta} * x^{-\mu-1}$$

Where x, θ, μ > 0

2.3. Data

2.3.1. Data Collection

A major issue while determining the appropriate loss distribution is the data collection. If the data used (i.e. the sample) is not a good representative of the whole set that is needed to be analysed, the researcher might have difficulties in choosing the most suitable loss distribution or might even a model that is not consistent with the data that needs to be analysed. Various methods can be used to connect the sample data to the population: These are the conditional probabilities and also the maximum likelihood distribution. Various researchers (Baud, Frachot and Roncalli (2002) and Fontnouvelle et al. (2003)) have acknowledged the need to make this link between the sample and the population data. Additionally, Baud, Frachot and Roncalli (2002) went a step further and managed to demonstrate that neglecting reporting bias can cause considerable low or inadequate estimates of the severity distributions. Failing to ascertain an unbiased sample will result in strongly biased severity distribution; hence increasing the level of risk associated with the data. In their study, Baud, Frachot and Roncalli (2002) opted to treat the entire data as stochastic. However, this constitutes a serious burden in the sense that the maximum likelihood resulting from such a model would be extremely difficult to generate and also greatly impractical to use (Frachot et al., 2003).

Mathematically, the maximum likelihood, aforementioned, would be of the form:

$$\max_{(\mu,\sigma)} l_n(\mu,\sigma) = \langle \textstyle\sum_{i=1}^{n} l(\tau_i,\mu,\sigma)\,|H_i\rangle$$

Where:

- n is the number of losses;
- $\langle \sum_{i=1}^{n} l(\tau_i,\mu,\sigma)\,|H_i\rangle$: This is the summation of the log likelihoods of the n losses (reported subject to the threshold H_i).

2.3.2. *Aggregate Losses*

In certain situations, losses are reported in a group that will then constitute the aggregate losses. This also constitute a problem in that the losses in a given group might have different loss distributions. Extracting information about the individual loss distributions is tiresome tedious task. Frachot, Georges and Roncalli (2001) suggested the Generalized Method of Moment (GMM) as a way to handle this issue. This method proved to be more efficient in dealing with both single-losses and aggregate losses simultaneously as Maximum Likelihood might be limited in some circumstances.

One of the assumptions (and also the main assumption) of the GMM is that the data can be generated by a weakly stationary **ergodic** stochastic process. A stochastic process is assumed ergodic if its statistical properties can be deduced from a single, sufficiently long, random sample of the process.

Mathematically, let us consider that the available data consists of T observations $\{Y_t\}_{t=1 \ldots T}$, where Y_t our random variable representing each observation. The aim of the estimation problem is to find at least a reasonably close estimate of θ. The reasoning behind the GMM is substitute the theoretical expected value with its empirical analog—sample average:

$$\widehat{m}(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(Y_t, \theta)$$

The norm of this expression will then be minimized in light of θ. The minimized value of θ will then be used as the estimate of θ.

CHAPTER 3: METHODOLOGY

0. Introduction

This chapter details the actual process that is going to be used in finding plausible solutions to the research questions that have been formulated earlier on. The chapter is divided in various sections which are: Section 3.1 will give a picture of how the population and the sample design will be carried out; Section 3.2 explains the context in which the data collection is going to be carried out; The last part, the section process, will then conclude by showing how an adequate will be used to determine the appropriate distribution loss for a given type form of data.

3.1. Population and Sample design

As the goal of this study is to find the appropriate loss distribution for a certain type of data, the population will hence been all the insurance companies in Kenya, for both life insurance and nonlife insurance (or general insurance) companies. For life noninsurance companies, we will be dealing with the various policies that are offered in the Kenyan market such as: Auto insurance, Health insurance, Income protection insurance, Casualty insurance and so on.

Given that there are forty seven (47) insurance companies offering both general insurance and life insurance products, there is thus a need to diminish that number to one that we can work with. For the purpose of this study, five (5) insurance companies will be considered as a sample for our research.

3.2. Data Collection

Required data will be obtained from secondary sources. The main source of data will be the Kenya insurance companies as we would want to identify the adequate loss distributions for various types of insurance products. The insurance company has to be registered by the Insurance Regulatory Authority, which is the Kenya body in charge of such activities.

The data used will be annually (annual claims for that product), ranging from the year 2010 to 2014 due to the availability of the Kenyan data and also the limited amount of time required to proceed all the claims associated with the various insurance products offered by the diverse companies. Both General insurance and life insurance companies will be part of the study that is being carried out. Reinsurance companies will not be part of the study due to the difficulty involved with collecting the amount of claims that have actually been settled.

3.3. Model Used

Step 1

The first step in developing the model will consist of calculating the descriptive parameters using the data that we will have obtained. For instance, the mean will be calculated using the following formula:

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Where:

- x_i is the claim amount for the i^{th} year ;
- n is the total number of years.

Other parameters such as the variance, the median will also be calculated. All the descriptive parameters will then be compared later with the parameters found using the likelihood function of various loss distributions.

Step 2

The second step will consist of calculating the parameters using estimates obtained from the maximum likelihood estimation method. This method has several advantages such as its desirable properties: consistency, efficiency, asymptotic normality and invariance. Another advantage of this method is the fact it incorporates all information provided about the parameters contained in the data.

Consider Xi be the ith claim amount, where $1 \leq i \leq n$. n being claims the total number of years, L is the likelihood function, θ is the parameter and f(x) is the probability distribution function of a specific loss distribution.

The likelihood function of the claims data is given by:

$$L = \prod_{i=1}^{n} f(X_i)$$

Where L is the likelihood function.

To get maximum likelihood, differentiation of the previous equation is done:

$$MLE = \frac{dl}{d\theta}$$

To solve the value of the parameter, we will have to equate MLE to zero and hence get the value of the estimate.

Step 3

The last step will consist of comparing the descriptive parameters obtained and the estimate obtained from the various loss distributions. We will then choose the loss distribution that has an estimate that describes better the data. To achieve the above, we will be using a goodness-of-fit test. The various steps will be:

State the hypothesis

The null hypothesis and the alternative hypothesis will be as follows:

H_0: The data are consistent with a specified distribution.

H_a: The data are not consistent with a specified distribution.

Sample Data Analysis

The degrees of freedom, the expected frequency counts, test statistic, and the P-value are going to be calculated:

- Degrees of freedom (DF): This will be equal to the number of years (n) of the categorical variable minus 1: DF = k - 1.
- Test statistic: The test statistic is a chi-square random variable (X^2) will be calculated using the following equation

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Where:

- O_i is the observed value in the data for the i^{th} year;
- E_i is the value obtained from the loss distribution.

- The P-value. This is the probability of observing a sample statistic as extreme as the test statistic. Since the test statistic is a chi-square, use the Chi-Square Distribution Calculator to assess the probability associated with the test statistic. The degrees of freedom computed above will be used.

Interpretation

If the P value is less than the significance level, the null hypothesis will be rejected. We will then compare the significance levels of the various loss distributions whose null hypothesizes were not rejected and the best of them will be chosen. The corresponding distribution will hence be the appropriate distribution for that data.

CHAPTER 4: DATA ANALYSIS

The hypotheses to be tested in this study are stated below in their null form:

H_0: The data are consistent with a specified distribution;

H_a: The data are not consistent with a specified distribution.

We will first start by calculating the descriptive parameters using the data that we have obtained.

The following table represents the required descriptive parameters for our data:

Parameter	Value
1. Mean	3.385
2. Variance	72.343
3. Standard Deviation	8.505
4. Median	1.778
5. Skewness	18.74983
6. Kurtosis	-483.764
7. Number	2167

All these values will be used when comparing the descriptive parameters obtained and the estimate obtained from the various loss distributions. We will now proceed to the second step of our methodology. This will consist of calculating the parameters using estimates of the maximum likelihood estimation method. This hence requires us to derive all these estimates from the distributions that were stated in the beginning of this study. These are the following:

- Exponential distribution;
- Pareto distribution;
- Lognormal distribution;
- Gamma distribution;

Derivation of the Estimates

Exponential distribution

The following is the probability distribution of the exponential distribution:

$$f(x) = \frac{1}{\lambda} e^{-\lambda x}$$

Where x > 0 and λ a parameter.

The mean and variance of the exponential distribution are the following:

$E(X) = \frac{1}{\mu}$ and $Var(X) = \frac{1}{\mu^2}$

Let us now obtain the estimate of this function using the Maximum Likelihood Estimation (MLE) Method.

The likelihood of obtaining the estimate from an exponential distribution with parameter λ is:

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \Sigma x_i} = \lambda^n e^{-\lambda n \bar{x}}$$

To find the MLE, we need to find the value of λ that maximises the likelihood, or, alternatively, the value that maximises the log-likelihood:

$$\log L = n \log \lambda - \lambda n \bar{x}$$

Differentiating to look for stationary points:

$$\frac{\partial}{\partial \lambda} \log L = \frac{n}{\lambda} - n \bar{x}$$

Setting this to zero gives:

$$\hat{\lambda} = 1 / \bar{x}$$

The second derivative is

$$\frac{\partial^2}{\partial \lambda^2} \log L = -\frac{n}{\lambda^2} < 0$$

,

This shows that this is a maximum.

Pareto Distribution

The following is the probability distribution of the Pareto distribution:

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

Where:

Parameters: $\alpha, \lambda \ (\alpha > 0, \lambda > 0)$

The mean and variance of the Pareto distribution are the following:

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

To find the MLE, we need to find the value of λ that maximises the likelihood, or, alternatively, the value that maximises the log-likelihood:

$$\log L = log \prod_{i=1}^{n} f(X_i|\theta) = \sum_{i=1}^{n} logf(X_i|\theta)$$

Where θ is the unknown parameter. It can either be one of the two parameters: α or λ.

After carrying out the appropriate computations, the estimate of the parameter α will be:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(\frac{X_i}{\lambda})}$$

$\ln\left(\frac{X_i}{\lambda}\right)$ is exponentially distributed with mean value $\frac{1}{\alpha}$

Lognormal distribution

The following is the probability distribution of the Pareto distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{x}\exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\},\ x > 0$$

Where:

Parameters: $\mu,\ \sigma^2\ (\sigma > 0)$

The mean and variance of the lognormal distribution are the following:

$$E(X) = e^{\mu + \frac{1}{2}\sigma^2},\quad \mathrm{var}(X) = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

To find the MLE, we need to find the values of μ and σ that maximise the likelihood, or, alternatively, the value that maximises the log-likelihood:

$$\log L = log \prod_{i=1}^{n} f(X_i|\theta) = \sum_{i=1}^{n} log f(X_i|\theta)$$

Where θ is the unknown parameter. It can either be one of the two parameters: μ or σ.

After carrying out the appropriate computations, the estimate of the parameters μ and σ will be:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \ln(X_i)}{n}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(\ln(X_i) - \frac{\sum_{i=1}^{n}\ln(X_i)}{n}\right)^2}{n}$$

Discussion

After obtaining the different estimates formula of the loss distributions that we are supposed to use, we will then proceed to compute the expected values from each of the distributions. The first step consists of calculating the estimates using the formula. After that, we will then generate the various expected values of each distribution. These are the values that are going to be compared to the observed from our data. Afterwards, a hypothesis test is going to be carried to determine whether we have enough evidence to reject the null hypothesis. Our null hypothesis is that the data, which we are using in this study, are consistent with a specified distribution. This is going to be achieved by comparing a t statistic value and a t critical value. These values were obtained in Microsoft Excel using t-Test: Two-Sample Assuming Unequal Variances. This study was carried out using a confidence interval of 95 %. This is the most used level of confidence interval.

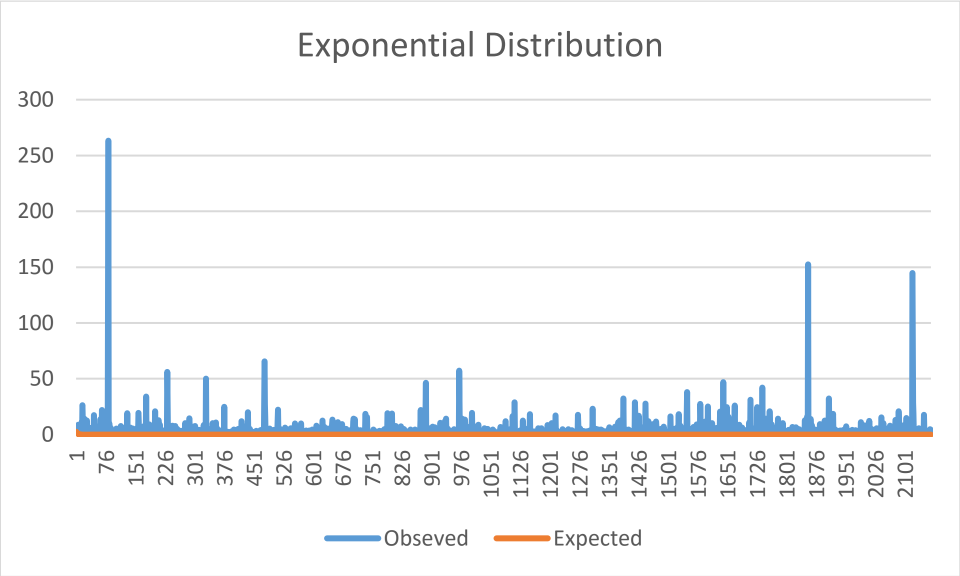
We will start by the exponential distribution.

Exponential Distribution

t-Test: Two-Sample Assuming Unequal Variances		
	Observed	*Expected*
Mean	3.385	0.0045
Variance	72.377	0.0065
Observations	2167	2167
Hypothesized Mean Difference		
df	2166	
t Stat	18.497	
P(T<=t) one-tail	0	
t Critical one-tail	1.646	
P(T<=t) two-tail	0	

t Critical two-tail	1.961	

The results for this distribution show that on average the mean of the observed is way greater than the mean of the expected. It is also apparent that there is a significant difference between the variances of the observed and the expected values. The results of the t statistics are hence not surprising. The t-stat is higher than the critical value and the p-value is not significant at 5%. We hence have enough evidence to reject the null hypothesis.

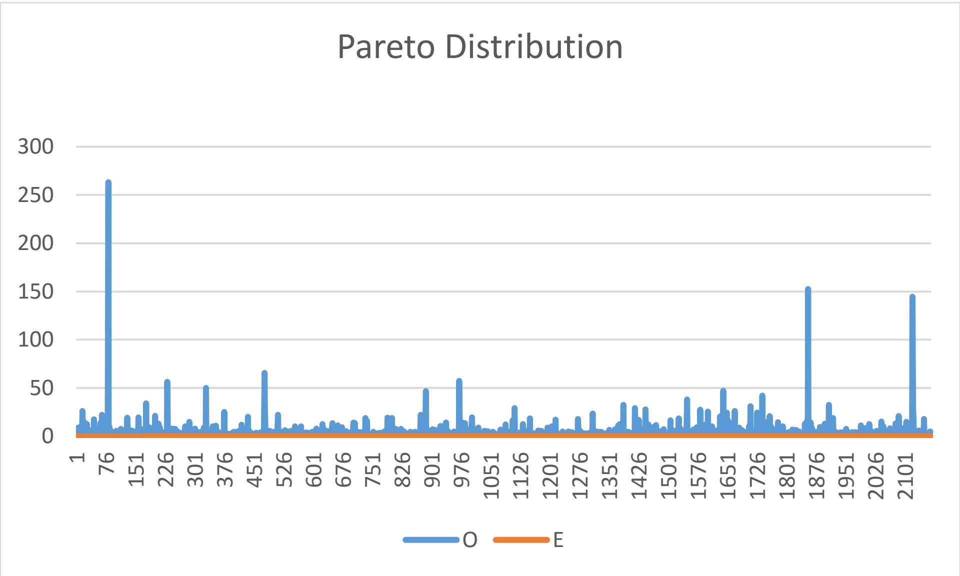


The graph shows that the movement of the observed and the expected values have significant differences. The values of the observed values exhibit much larger variations than the values of the expected values. Once again, this confirms the results from the t statistic that the data, which we are using in this study, are consistent with the exponential distribution.

Pareto Distribution

t-Test: Two-Sample Assuming Unequal Variances		
	Observed	*Expected*
Mean	3.385088304	0.000269545
Variance	72.37674016	2.23985E-05
Observations	2167	2167
Hypothesized Mean Difference	0	
Df (Degrees of freedom)	2166	
t Stat	18,52103696	
P(T<=t) one-tail	1.64949E-71	
t Critical one-tail	1.645557424	
P(T<=t) two-tail	3.29898E-71	
t Critical two-tail	1.961059818	

Once again, the results for this distribution show that on average the mean of the observed is way greater than the mean of the observed. We can also see that there is a significant difference between the variances of the observed and the expected values; this way justifying the results of the t statistics. The t-stat is higher than the critical value and the p-value is not significant at 5%. We hence have enough evidence to reject the null hypothesis.

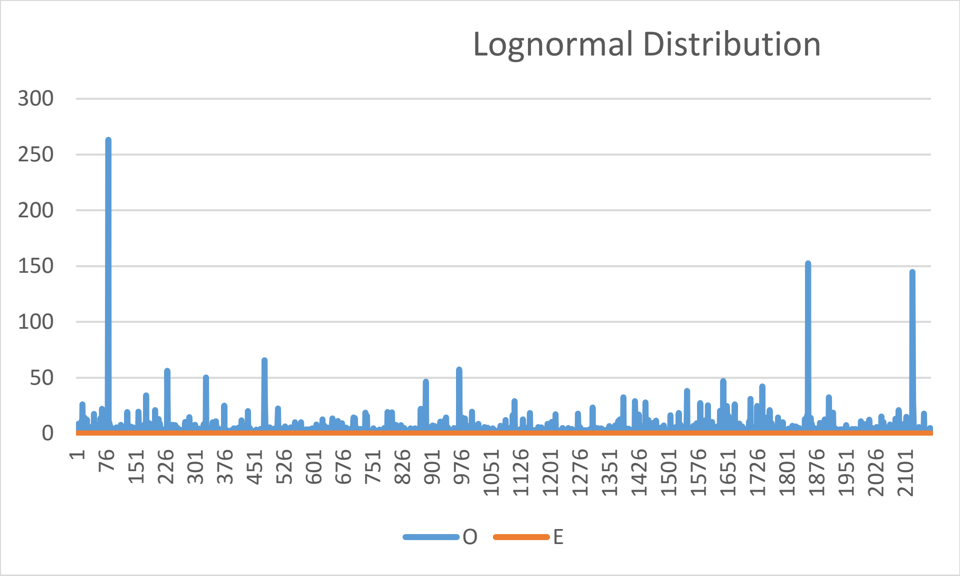


For this distribution, just like the previous one, the graph shows that the movement of the observed and the expected values have significant differences. The values of the observed values exhibit much larger variations than the values of the expected values. Once again, this confirms the results from the t statistic that the data of this study are not consistent with the Pareto distribution.

Lognormal Distribution

t-Test: Two-Sample Assuming Unequal Variances		
	Observed	*Expected*
Mean	3.385088	3.91E-07
Variance	72.37674	2.94E-14
Observations	2167	2167
Hypothesized Mean Difference	0	
df	2166	
t Stat	18.52251	
P(T<=t) one-tail	1.61E-71	
t Critical one-tail	1.645557	
P(T<=t) two-tail	3.22E-71	
t Critical two-tail	1.96106	

The results obtained for this distribution have similarities with results from previous distributions. They show that on average the mean of the observed is way greater than the mean of the expected. It is also apparent that there is a significant difference between the variances of the observed and the expected values. The results of the t statistics are hence not surprising. The t-stat is higher than the critical value and the p-value is not significant at 5%. We hence have enough evidence to reject the null hypothesis.



Once again, the graph shows that the movement of the observed and the expected values have significant differences. The values of the observed values exhibit much larger variations than the values of the expected values. We can then confirm that the results from the t statistic are not consistent with the lognormal distribution.

CHAPTER 5: CONCLUSION

In this paper, the hypothesis was supposed to be tested using Kenya Insurance data. However, due to lack of it, we used other insurance data. This does not constitute a major issue as the goal of the study was to show clearly the various steps that are required in order to obtain the most suitable distribution for a set of data.

In the study, estimates from various distributions were generated and these were compared to the main data to determine which distribution suits most the data that was being used. The comparability was achieved by use carrying a hypothesis test. For all the distributions that were used in the study, the t-stat value was greater than the critical value. We will then choose the distribution with the lowest t-stat value; this is the *Exponential Distribution.* Henceforth, this is the distribution that is most suitable for the data that we had. Further studies could be carried out to determine the distributions that are most suitable for each class of insurance.

REFERENCES

1. Wright T. (2005). *Challenge Round 3*. FIA, Deloitte, UK. COTOR, available at www.casact.org/cotor/

1. Achieng, O.M. (2010). *Actuarial modelling for insurance claim severity in motor comprehensive policy using industrial statistical distributions*. International Congress of Actuaries, Cape Town, 7–12 March 2010

2. Aljuaid, A. (2013). *Estimating the Parameters of an Exponentiated Inverted Weibull Distribution under Type-II Censoring*. Applied Mathematical Sciences, **7**, 35, 1721 – 1736

3. Baud, N., A. Frachot, & T. Roncalli (2002). *Internal data, external data and consortium data for operational risk measurement: How to pool data properly?*. Working Paper, Crédit Lyonnais, Groupe de Recherche Opérationnelle.

4. Baud, N., A. Frachot, & T. Roncalli (2003). *How to avoid over-estimating capital charge for operational risk?*. Operational Risk – Risk's Newsletter.

5. Flaih A., H. Elsalloukh, E. Mendi & M. Milanova (2012). *The Exponentiated Inverted Weibull Distribution*. Applied Mathematics & Information Sciences, **6**, 2, 167 - 171.

6. Frachot, A., P. Georges & T. Roncalli (2001). *Loss Distribution Approach for operational risk*. Working Paper. Crédit Lyonnais, Groupe de Recherche Opérationnelle.

7. Frachot, A., Moudoulaud, O., & Roncalli, T. (2003). *Loss distribution approach in practice*. The Basel handbook: A guide for financial practitioners, 369-396.

8. Frees, E.W. (2013). *Regression modelling with actuarial and financial implications*. Cambridge University Press.

9. De Fontnouvelle, P., V. DeJesus-Rueff, J. Jordan, & E. Rosengren (2003). *Using loss data to quantify operational risk*. Working Paper, Federal Reserve Bank of Boston, Department of Supervision and Regulation.

10. Harris, J. W. & Stocker, H. (1998). *Maximum Likelihood Method*. §21.10.4 in Handbook of Mathematics and Computational Science. New York: Springer-Verlag, p. 824.

11. Hassan, A.S. (2013). *On the Optimal Design of Failure Step-Stress Partially Accelerated Life Tests for Exponentiated Inverted Weibull with Censoring*. Australian Journal of Basic and Applied Sciences, **7**, 1, 97 - 104.

12. Hassan, A.; Marwa, A.; Zaher, H. & Elsherpiny, E. (2010). *Comparison of Estimators for Exponentiated Inverted Weibull Distribution Based on Grouped Data*. Int. Journal of Engineering Research and Applications, 4, 4, 77 - 90.

13. Hoel, P. G. (1962). *Introduction to Mathematical Statistics, 3rd ed.* New York: Wiley, p. 57.

14. Jouravlev, I. (2009). *Loss Distribution Generation in Credit Portfolio Modelling*. MMF. Walden University, USA

15. Kiragu, S. M. (2014). Assessment of challenges facing insurance companies in building competitive advantage in Kenya: A survey of insurance firms. International Journal of Social Sciences and Entrepreneurship, 1 (11), 467-490.

16. Limpert, E., Stahel, W. A., & Abbt, M. (2001). *Log-normal distributions across the sciences: Keys and clues on the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question*. BioScience, 51(5), 341-352.

17. Miller, I.; Miller, M (2013). *John E. Freund's Mathematical statistics with applications*..; [Freund, J. E.] 8th ed. Prentice Hall International.

18. Moses, V. & Kuloba, R. (2013). *2013 Kenya Insurance Industry Outlook*. Insurance Regulatory Authority. Retrieved May 30th, 2016 from https://www.google.com/search?q=INSURANCE+REGULATORY+AUTHORITY%2C+2013+KENYA+INSURANCE+INDUSTRY+OUTLOOK%2C+VICTOR+MOSE+AND+ROBERT+KULOBA+Policy+Research+and+Development+Division&ie=utf-8&oe=utf-8&client=firefox-b-ab

19. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; &Vetterling, W. T. (1992). *Least Squares as a Maximum Likelihood Estimator*. §15.1 in Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, pp. 651-655.

20. Shao, Q. (2004a*). Notes on maximum likelihood estimation for the three-parameter Burr XII distribution*. Computational Statistics and Data Analysis, 45, 675–687. http://dx.doi.org/10.1016/S0167-9473(02)00367-5

21. Shao, Q., Wong, H., & Xia, J. (2004b). *Models for extremes using the extended three parameter Burr XII system with application to flood frequency analysis*. Hydrological Sciences (Journal des Sciences Hydrologiques), 49, 685–702. http://dx.doi.org/10.1623/hysj.49.4.685.54425

22. Soliman, A. A. (2005). *Estimation of Parameters of Life from Progressively Censored Data using Burr-XII Model*. IEEE Transactions on Reliability, 54, 34–42. http://dx.doi.org/10.1109/TR.2004.842528
23. Wright T. (2005). *Challenge Round 3*. FIA, Deloitte, UK. COTOR, available at www.casact.org/cotor/

24. Wu, S. J., Chen, Y. J., & Chang, C. T. (2007). *Statistical inference based on progressively censored samples with random removals from the Burr type XII distribution*. Journal of Statistical Computation and Simulation, 77, 19–27. http://dx.doi.org/10.1080/10629360600569204

2. Flaih A., H. Elsalloukh, E. Mendi & M. Milanova (2012). *The Exponentiated Inverted Weibull Distribution*. Applied Mathematics & Information Sciences, **6**, 2, 167 - 171.
3. Aljuaid, A. (2013). *Estimating the Parameters of an Exponentiated Inverted Weibull Distribution under Type-II Censoring*. Applied Mathematical Sciences, **7**, 35, 1721 – 1736.
4. Hassan, A.; Marwa, A.; Zaher, H. & Elsherpiny, E. (2010). *Comparison of Estimators for Exponentiated Inverted Weibull Distribution Based on Grouped Data*. Int. Journal of Engineering Research and Applications, 4, 4, 77 - 90.
5. Hassan, A.S. (2013). *On the Optimal Design of Failure Step-Stress Partially Accelerated Life Tests for Exponentiated Inverted Weibull with Censoring*. Australian Journal of Basic and Applied Sciences, **7**, 1, 97 - 104.

6. Baud, N., A. Frachot, & T. Roncalli (2002). *Internal data, external data and consortium data for operational risk measurement: How to pool data properly?*. Working Paper, Crédit Lyonnais, Groupe de Recherche Opérationnelle.
7. De Fontnouvelle, P., V. DeJesus-Rueff, J. Jordan, & E. Rosengren (2003). *Using loss data to quantify operational risk*. Working Paper, Federal Reserve Bank of Boston, Department of Supervision and Regulation.
8. Baud, N., A. Frachot, & T. Roncalli (2002). *Internal data, external data and consortium data for operational risk measurement: How to pool data properly?*, Working Paper, Crédit Lyonnais, Groupe de Recherche Opérationnelle.

9. Frachot, A., P. Georges & T. Roncalli (2001). *Loss Distribution Approach for operational risk*. Working Paper. Crédit Lyonnais, Groupe de Recherche Opérationnelle.
10. Baud, N., A. Frachot, & T. Roncalli (2003). *How to avoid over-estimating capital charge for operational risk?*. Operational Risk – Risk's Newsletter.
11. I. Jouravlev (2009). *Loss Distribution Generation in Credit Portfolio Modelling*. MMF. Walden University, USA
12. V. Moses & R. Kuloba (2013). *2013 Kenya Insurance Industry Outlook*. Insurance Regulatory Authority. Retrieved May 30th, 2016 from https://www.google.com/search?q=INSURANCE+REGULATORY+AUTHORITY%2C+2013+KENYA+INSURANCE+INDUSTRY+OUTLOOK%2C+VICTOR+MOSE+AND+ROBERT+KULOBA+Policy+Research+and+Development+Division&ie=utf-8&oe=utf-8&client=firefox-b-ab
13. Limpert, E., Stahel, W. A., & Abbt, M. (2001). *Log-normal distributions across the sciences: Keys and clues on the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question*. BioScience, 51(5), 341-352.
14. Kiragu, S. M. (2014). Assessment of challenges facing insurance companies in building competitive advantage in Kenya: A survey of insurance firms. International Journal of Social Sciences and Entrepreneurship, 1 (11), 467-490.
15. Frachot, A., Moudoulaud, O., & Roncalli, T. (2003). *Loss distribution approach in practice*. The Basel handbook: A guide for financial practitioners, 369-396.
16. Harris, J. W. & Stocker, H. (1998). *Maximum Likelihood Method*. §21.10.4 in Handbook of Mathematics and Computational Science. New York: Springer-Verlag, p. 824.

17. Hoel, P. G. (1962). *Introduction to Mathematical Statistics, 3rd ed.* New York: Wiley, p. 57.
18. Press, W. H. Flannery, B. P.; Teukolsky, S. A.; &Vetterling, W. T. (1992). *Least Squares as a Maximum Likelihood Estimator*. §15.1 in Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, pp. 651-655.

25. Zimmer, W. J., Keats, J. B., & Wang, F. K. (1998). *The Burr XII distribution in reliability analysis*. Journal of Quality Technology, 30, 386–94