UNIVERSITY OF GHANA

MODELING LARGE INSURANCE CLAIMS USING EXTREME VALUE THEORY: A CASE STUDY OF THE 37 MILITARY HOSPITAL.

By

ABEKA COLLINS

(10701682)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF MASTER OF PHILOSOPHY IN STATISTICS DEGREE.

NOVEMBER, 2020

DECLARATION

This is the result of research work undertaken by Abeka Collins under the supervision of Dr. Richard Minkah and Dr. Ezekiel Nii Noye Nortey. I declare that the entity of the work contained therein is my own work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and I have not previously in its entirety or in part submitted it for obtaining any qualification.

Abeka Collins (10701682) Signed: Date: 07/06/2022

(Student)

Dr. Richard Minkah Signed: Date: 07-06-2022

(Principal Supervisor)

Dr. Ezekiel N. N. Nortey. Signed: Date: 07-06-2022

(Co Supervisor)

ABSTRACT

The private health insurance industry is one of the vital components in nation-building. It complements government's efforts in reducing "out-of-pocket" payment for healthcare services in the country. However, some private health insurance companies face severe insolvency issues due to accumulation of unanticipated huge claim amounts. The Extreme Value Theory (EVT) is a statistical tool proven to help solve or mitigate some of these challenges since it focuses mainly on the behaviour of severe but rare occurrence. In this study, we employ the EVT approaches to model large insurance claims from the 37 Military hospital; and to estimate financial risk indicators such as Value-at-Risk (VaR) and Expected Shortfall (ES) among other extreme quantiles. Conclusions drawn from analysis established that the Weibull class of distributions is more appropriate for the data at hand and for this reason, it is not likely for the 37 Military hospital to submit claim amount exceeding 24,618 cedis for any given day. In addition, private health insurance firms can be assured at a confidence level of 99%, 99.5% and 99.9% that within a day, the hospital is not likely to submit a claim amount exceeding 2,910 cedis, 3,938 cedis and 7,946 cedis respectively. Finally, it was recommended that the NHIA could replicate this study using the claims received by the public health insurance scheme (i.e. NHIS) since it can go a long way to strengthen the financial sustainability of the scheme.

DEDICATION

This thesis is dedicated to my lovely wife, Mrs. Valentina Abeka-Arthur, to an angel of a daughter, Annette Aseda Abeka-Arthur, and to my caring parents, Mr. and Mrs. Abeka-Arthur.

ACKNOWLEDGEMENT

My foremost appreciation goes to Jehovah God Almighty for His grace and mercies in my life and for the strength He gave me to accomplish this thesis. My deepest gratitude goes to my supervisors, Dr. Richard Minkah and Dr. Ezekiel N.N. Nortey, for their immense time, support and guidance towards the completion of this thesis. I also recognise Dr. Samuel Iddi, Mr. Stephen Nkrumah, Mr. Eric Ocran and Emmanuel Kojo Aidoo of the Statistics and Actuarial Science Department of the University of Ghana for their various ways of assisting me to complete this thesis. I acknowledge Major Richard Osei-Boateng, Officer-In-Charge of the Health Information System Department and Mr. Isaac Addo, my immediate superior, at the 37 Military Hospital for their support and motivation in combing office duties and the writing of this thesis. Next in line are Lt. Emmanuel Dakubu, Mrs. Mary Adu-Brempong, Mr. Ebenezer Nortey and the entire Claims Unit of the 37 Military Hospital. I thank them for their diverse support in completion of this thesis. Many more thanks to my parents Mr. and Mrs. Abeka-Arthur, to my siblings, to my friend and confidant, Pastor Bismark Okai, for their extreme support in my life and also in completion of this thesis. Last and very important is my 'God-sent' wife, Mrs. Valentina Abeka-Arthur. I appreciate and thank her for being the "wind in my sail".

May God richly bless you all.

Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

BMM	Block Maxima Method
ES	Expected Shortfall
EVD	Extreme Value Distribution
EVI	Extreme Value Index
EVT	Extreme Value Theory
GEV	Generalised Extreme Value
GEVD	Generalised Extreme Value Distribution
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimation (or Estimates)
PL-CI	Profile Likelihood-based Confidence Interval
POTM	Peak Over Threshold Method
PSP	Parameter Stability Plot
PWM	Probability Weighted Moments
PWME	Probability Weighted Moments Estimation (or Estimates)
VaR	Value-at-Risk

Chapter 1

INTRODUCTION

Classical data analysis is mostly concerned with the characteristics of the centre of a distribution for a data set. It usually deals slightly with outliers (extrema) and assumes the normal distribution as a suitable asymptotic fit for the data, given a sufficiently large sample size even when the individual observations may not be normally distributed. Thus, little or sometimes no attention is given to the tail behaviour of the distribution and as a result, the model formulated to fit the data mathematically underestimates statistical inferences for rare but severe observations. Consequently, room is given to inadequate preparedness in preventing or mitigating damages that may arise from the occurrence of such extreme events.

Taking the health insurance industry in Ghana for instance, the presence of high and increasing claim cost have placed the potency of the country's public (national) health insurance scheme under severe financial pressure; and this has contributed to the scheme's inability to pay claims on time to healthcare providers (NHIS Review, 2020). As a result of the scheme's indebtedness, some of the hospitals end up demanding "out-of-pocket" payment from insured patients for medical services provided (Donkor, 2014; Ghanaweb Network, 2015). Consequently, an increase in the financial burden on healthcare can be expected from the residents of the country. Another example of the catastrophic effect of severe

but rare observations can be seen in St. Helens Insurance company limited, in the United Kingdom (Massey, 2003). Massey narrated that the St. Helens insurance company ceased underwriting after incurring large losses from Hurrican Bestsy in 1965, and eventually folded up in 1967 mainly as a result of huge unanticipated claims and insufficient reinsurance.

Other tragic effects of extreme happenings include: increased rate in Cerebrospinal Meningitis (CSM) infection due to occurrence of high temperature levels in dry seasons (Codjoe and Nabie, 2014); destruction of infrastructure and farms by floods due to presence of high rainfall levels (Nkrumah, 2017); and the Global Financial Crisis caused by the ripple effect of extremely low prices in the US housing market that led to high mortgage loan defaulters, which also resulted into insolvency by banks and investors who gave out the loans (Reserve Bank of Australia, 2018).

Considering the damages left behind by the occurrence of extreme realisations, it becomes more appropriate to separately model the tails of a distribution when accurate extrapolations of maxima or minima observations is of concern; and this is where Extreme Value Theory (EVT) comes to 'play'.

EVT is a statistical framework that explicitly deals with methods for modelling and inference of severe but infrequent events (Nortey et al., 2017). It focuses its attention on the behaviour of a distribution's tails by examining very high and very low observations in a data set (Minkah, 2016). The applicability of EVT spans from hydrology (Minkah, 2016; Nortey et al., 2017) , climatology (Nkrumah, 2017), finance (Danielsson and De Vries, 2000, Allen et al., 2011; Uppal, 2013; Nortey et al., 2015), insurance (Embrechts et al., 2013, Pérez-Fructuoso and García Pérez, 2010; Adesina et al., 2016; Weru and Waititu, 2019) and sports (Sérgio, 2012).

The concern of this study is to employ the EVT methods to assess and infer the behaviour of large insurance claims in the private health insurance industry using the 37 Military Hospital as a case study.

1.1 Background of Study

Health insurance is the protection a person secures against financial detriment as a result of an untimely illness that is capable of causing monetary loss to the person. Boadu et al. (2014) records that the insurance system in Ghana started in 1924 with Enterprise Insurance Company Limited (formerly, Royal Guardian Enterprise). He continued that the insurance system is divided into life insurance, non-life insurance, and composite insurance (a combination of life and non-life insurance). In Ghana, the National Insurance Commission (NIC) is the sole institution that has the authority to regulate and supervise insurance activities in the country under the Insurance Act, 2006 (Act 724) of the 1992 constitution of Ghana (National Insurance Commission, 2019). Currently, there are 142 regulated insurance entities made up of 24 life insurance companies, 29 non-life insurance companies, 3 reinsurance companies and 85 insurance brokers and loss adjusters (National Health Insurance Authority, 2020).

Over the years, one component of the non-life insurance that has seen a facelift by successive governments in the country is the health insurance. Though with different approaches, the implementation of the Health Insurance Scheme is one which all the major political parties have openly embraced in line with the country's pursuance of reducing financial hardship in accessing healthcare delivery. Prior to the first republic, healthcare financing was predominantly by "out-of-pocket" payments at the point of service (Arhinful, 2003). It then became almost free as with other social services during the first republic and later saw a "turn-around" with the "cash and carry" system following the overthrow of the

then government (Wagstaff, 2010). The "cash-and-carry" system of healthcare financing requires patients to pay part or full amount of the cost of health services in order to reduce government's finances in the health sector (Owusu-Sekyere and Chiaraah, 2014). However, the "cash-and-carry" system became highly ostracized among the citizens and was eventually replaced by the National Health Insurance Scheme (NHIS) - a social intervention by government to reduce the cost of health for her residents (National Health Insurance Scheme, 2020).

National Health Insurance Authority (2020) states that the NHIS was established under the National Health Insurance Act, Act 650 in 2003, and later amended in 2012 by Act 852. It continued that the NHIS was instituted by the government of Ghana to finance basic healthcare services to residents in the country through public and private health insurance schemes. The National (public) Health Insurance Scheme (NHIS) is managed and largely funded by the government through tax revenues, statutory deductions (such as the National Health Insurance Levy (NHIL) and SSNIT contributions), and other government subventions. It is also funded by premiums - an agreed amount borne by an insured for a contract of insurance - from informal sector subscribers (National Health Insurance Authority, 2020). The NHIS benefit package covers about 95 percent of diseases in the country, and is accepted in all government health facilities and some accredited private healthcare centres (National Health Insurance Scheme, 2020).

The Private Health Insurance Scheme (PHIS), on the other hand, offers financial protection on a wider range of health issues and allows coverage of certain healthcare services that are excluded under the public health insurance scheme. Examples include VIP ward accommodation, photography, elective heart and brain surgeries, dialysis treatment, and treatments overseas.

National Health Insurance Authority (2018) asserts that there are two types of private health

insurance schemes in Ghana - Private Mutual Health Insurance Scheme (PMHIS) and Private Commercial Health Insurance Scheme (PCHIS). The former operates exclusively for the benefit of its members and can be formed by any identifiable group of persons in the country; while the latter is a limited liability insurance company that is operated on a profit motive and is purely funded by premiums paid by its subscribers. Currently, there are about 14 companies licensed by the NHIA to operate as Private Commercial Health Insurance Scheme (PCHIS) (National Health Insurance Scheme, 2020). Notably among them are Nationwide Medical Insurance Company Limited, Premier Health Insurance Company Limited, Acacia Health Insurance Company and Glico Healthcare Company Limited.

The health insurance system comprises three key stakeholders - the insurer (insurance company), the insured (subscriber), and the service provider (healthcare facility) whose role is crucial in the overall quest for sustainable delivery of quality and accessible healthcare service. In Ghana, the Health Facility Regulatory Agency (HEFRA) in collaboration with the National Health Insurance Authority (NHIA) are responsible for respectively supervising and licensing healthcare facilities that wish to avail their services to the country's health insurance industry. Aside pharmaceutical and government facilities, there are about 1,370 healthcare facilities licensed to provide services in the country as at March 2020 (Health Facilities Regulatory Agency, 2020).

According to Ministry of Health (2016), categories of healthcare facilities that can be credentialed by the NHIA to provide services to residents include licensed chemical shops, diagnostic centres, Community-based Health Planning and Services (CHPS), maternity homes, clinics, and hospitals. The hospital is one of the health facilities that provides a broad range of healthcare services to a wider populace. Ghana's health system classifies various hospitals on the basis of management and administration to be a private hospital, a government hospital, or a quasi-government hospital. In addition, the hospitals can also be grouped into

primary, secondary, or tertiary based on the level of healthcare delivery. One such hospital that has proven to be of significant value in healthcare delivery to the country is the 37 Military Hospital.

An interview at the 37 Military Hospital with Major Richard Osei-Boateng (Officer-In-Charge of the Health Information System Department) and Mr. Isaac Addo (Officer-In-Charge of the Health Insurance Unit) on same accounts attest that the 37 military hospital was initially instituted in July 1941 to provide services for troops injured in the second world war. In addition, it is classified as a tertiary quasi-government hospital established by the Ghana Armed Forces to primarily provide healthcare to its serving and ex-service personnel and their families, to civilian employees of the Ministry of Defence and their families, and also to the general public as a referral healthcare facility.

The 37 Military hospital is made up of several departments - surgical, medical, obstetrics, gynaecology, paediatrics, dental, physiotherapy, radiology, pathology, and a chemist department. In addition to being a teaching hospital, the 37 Military Hospital serves as a National Disaster and Emergency Response health facility and also comes to the rescue in times of industrial strikes by government healthcare professionals (Tv3 Network, 2015). The administration of the hospital is headed by the Commander of the unit and assisted by the Commanding Officer and other high-ranking military officers. Healthcare financing at the 37 Military Hospital by the general public is either through "out-of-pocket payment" (cash or cheque), through insurance (NHIS or accredited PHIS), or by the social welfare organisation.

1.2 Problem Statement

One of the targets of universal health coverage is to ensure that people obtain the healthcare services without suffering financial hardship when paying for them (De Wolf and Toebes, 2016). The private health insurance industry plays an important role in achieving this goal since they complement government's efforts in financing healthcare services and thus reduce "out-of-pocket" payment for healthcare needs. However, the sustainability of healthcare financing through insurance is sometimes very challenging for health insurance companies especially when vetted claim amounts are very large and are not easily anticipated due to their rare occurrence. As a result, health insurance companies are likely to find themselves in insolvency issues when these claim amounts accumulate and remains unpaid. The additional effect is for healthcare providers to suspend their services on insurance basis and revert to the "cash and carry" system of healthcare delivery.

Based on the records from the private health insurance unit of the 37 Military Hospital, the following were brought to light:

- In the last quarter of 2016, the facility suspended services to four partnered health insurance companies - Med-X Health Systems Limited, First Fidelity Health Insurance, Oval health insurance and Empire Health Insurance Limited; mainly as a consequence of high indebtedness to the hospital.
- In 2018, services to companies like GN Medicals Limited and Beige Healthcare Limited were suspended as a result of insolvency issues.
- There were records of long-overdue unpaid claims amounts which in the long run can bring a possible reversion to the "cash-and-carry" system for the general public.

In light of this, the researcher employs the extreme value theory framework to model and predict the behaviour of very large claim sizes that are likely to occur when dealing with the

37 Military Hospital. This we believe could help the health insurance companies to position themselves better on financing medical care for their clients at the hospital. as the above mentioned problems are curtailed.

1.3 Research Objectives

- To identify and fit an appropriate limiting distribution to describe the tail behaviour of the distribution of large health insurance claims.
- To approximate the value of the maximum possible claim amount that can be generated by the hospital in a day.
- To forecast with a given probability the claims amount that could be exceeded once every 5, 10, 20, 50 and 100 months respectively.
- To project with some level of confidence some extreme claims that could not be exceeded in the next 24 hours.

1.4 Significance of Study

Universal health coverage (UHC) ensures that people have access to healthcare services whilst being protected from more than 30% health expenditure on their household income (Akinola and Dessislava, 2019). The World Health Organization (WHO) reports that though governments are spending more on health, residents are still paying much out of their own pocket (World Health Organization, 2019). This brings to light the need and sustainability of health insurance schemes that is capable of providing substantive financial protection against health expenditure. As part of the prudent measures in ensuring financial soundness in the sustainability of these schemes, it is important for health insurance companies to be able to anticipate the likelihood of receiving very large claims that could lead to large losses

and eventually cause financial instability or even disrupt their operations with healthcare facilities. In this regard, the study is significant in the following ways:

- The study will help private health insurance companies to better predict the likelihood of large insurance claims and the frequency with which they occur when it comes to dealing with the 37 Military Hospital.
- The study will highly aid insurance companies in the calculation of premium volumes needed to cover future huge claims for their clients who visit the 37 Military hospital.
- The goodwill between the hospital and the insurance companies will be boosted since companies will be able to pay claims on time due to adequate liquidity reserves and reinsurance measures.
- The study will also benefit the hospital by serving as a working document for future partnership agreement with any health insurance firm.
- The study will in the long run help to prevent the need for "out-of-pocket" payment of healthcare delivery due to the readiness of health insurance companies to reimbursed the hospital.
- The replication of the study in other healthcare facilities or insurance firms across the country can help ensure that the health insurance sector is stronger and well positioned to serve the growing demands of the public.

1.5 Contributions of the Study

The main contributions of this thesis are as follows:

- Academically, the study will contribute to present literature as a ghanaian perspective on the practice of EVT in the insurance industry, particularly, the private health insurance industry.

- Financially, the 37 Military hospital stands to gain immensely from this work as it is able to anticipate the maximum possible revenue it can generate from its partnered private health insurance companies at a particular point in time.

- Findings of this paper will greatly contribute to the dealings of the Private Health Insurance Association of Ghana (PHIAG) with healthcare facilities across the country as huge claims amount are easily anticipated and later indemnified in the periods they occur.

- In terms of integrity, both private health insurance companies and healthcare facilities can use this work to investigate certain large claim amounts that occurs in unexpected time periods.

1.6 Outline of the Study

Chapter 1 of the study captures the motive for the research and introduces the statistical tool of interest for the study. It further provides a background investigation of the study and identifies the problems the study is meant to resolve. It then captures the specific goals of the study, its significance, and contribution to both literature and stakeholders involved. Chapter 2 unfolds the origin and advancement of the Extreme Value Theory (EVT) and also peruses some of the scholarly contributions and diverse applications of EVT in real life. Chapter 3 is centred on the EVT methodology and discusses the parametric EVT-approaches that can be employed to describe the characteristics of extreme occurrences. Chapter 4 captures the data analysis and gives the description of the data at hand. This chapter also covers the GEV model and the GP model that were used to describe the large claims through the estimation of parameters and extreme quantiles under both MLE and PWME. Chapter 5 draws conclusions from the findings obtained in Chapter 4, and then makes some recommendations for further research.

Chapter 2

LITERATURE REVIEW

2.1 The Origin and Advancement of EVT

Extreme Value Theory (EVT) is that field of science that deals with the extraordinary deviations from the mean of a probability distribution (Allen et al., 2011). In particular, extreme value analysis focuses on the estimation and probabilities of events that are greatly different than any ever observed in samples of some specified size. (Coles, 2001).

The origin of the study of extreme value statistics, as stated by Kinnison (1985), begun with the early astronomers who were confronted with either to use or exclude observations that seemed to be very distant from the rest in a data set. He added that though these astronomers were able to specify the problem of extreme deviations in statistics, mathematical tools available were too crude to solve it. Kinnison further stated that, in 1922, L. von Bortkiewicz founded good numerical approximations which indicated that the sample maxima from a Gaussian population becomes new variables with separate distributions. He acknowledged that Bortkiewicz was the first to clearly state the extreme value problem in statistical terms. Kinnison continued that, in the subsequent year, von Mises introduced the mathematical concept of the expected largest value in a Gaussian sample; and this became the history on

the study of asymptotic distribution of extreme values in samples from Gaussian distributions.

The advancement of EVT, as narrated by Sérgio (2012), started in 1925 when Leonard Tippett studied material resistance by examining the strength of wool for the cotton industry in Britain. Tippet discovered that the strength of wool largely relied on the strength of the weakest fibres and not the average fibre. Sergio added that with the help of Sir Ronald Fisher, Tippet created a probabilistic theory by means of tables of large values and founded corresponding probabilities to serve as a framework to be applied for sample maxima from a normal distribution. Kinnison (1985), again narrated that Fisher and Tippet published the paper which is now known as the foundation of the asymptotic theory of extreme value distributions and consequently suggested the three domains of attraction - Gumbel, Fréchet, and Weibull domains of attraction - to characterise the distribution of sample extremes. Fisher and Tippet also indicated that there is an extremely slow convergence of the distribution of a gaussian sample maxima towards its asymptotic distribution.

Kinnison (1985) recorded that von Mises in 1936, made an improvement by establishing some simple and sufficient conditions for the weak convergence of the largest order statistic to each of the three family of distributions. However, Boris Vladimirovich Gnedenko in 1943, formalized all the previous knowledge about extremes and provided a much more rigorous foundation for the necessary and sufficient conditions for the weak convergence of the extreme order statistics; and this is known as the first theorem (i.e. the Fisher-Tippett-Gnedenko theorem) of Extreme Value Theory (Sérgio, 2012). Later on, an extensive bibliography of the literature and early development of statistical analysis of extremes from a theoretical and practical point of view was presented by Emil Julius Gumbel in 1958, in his book "Statistics of Extreme" (Kinnison, 1985). Since then, EVT has found its way in explaining the underlying mechanisms in many areas were the behaviour of extreme happenings is of utmost concern.

2.2 Applications of EVT and Scholarly Contributions

In Ghana, reference can be made to the application of EVT in modelling climatological and hydrological data. Nkrumah (2017) applied the EVT in estimating the occurrence of extreme temperature and extreme rainfall in Ghana for the period January 1960 to December 2012 using data from some selected regions of the country. Employing both the Generalized Extreme Value (GEV) model and the Generalized Pareto (GP) model, he found that the extreme occurrences of temperature and rainfall can be modelled using Fréchet and Weibull classes of distributions respectively. This allowed him to predict temperature maxima of 34.7, 34.6, and 39.6 degree Celsius, to be anticipated in Accra, Ashanti and Northern regions respectively, once every five years. He also expected an amount of 88.38mm, 85.61mm, and 81.98mm rainfall levels in Accra, Ashanti, and Northern regions respectively, once every 5 years.

Also, Minkah (2016) and Nortey et al. (2017), both studying the water levels of the Akosombo dam in Ghana, made use of EVT methods in their analyses of the dam's critical water levels. By means of the Generalized Pareto Distribution (GPD), Minkah (2016) estimated the probability and return period of very low water-levels that can lead to a complete shutdown of the dam's operations and very high water-levels that leads to spillage of excess water and cause flooding. While he found a negligible probability of a complete shutdown of the dam - since the left endpoint estimate is greater than the dam's critical level of 226ft - he also established the need for extension of the dam in order to reduce the probability of a flood occurring, or to even prolong the return period of a flood - since the right endpoint estimate was greater than the dam's maximum operating level of 278ft. Employing a different approach of the EVT, Nortey et al. (2017) arrived at the same results by using the

Generalized Extreme Value Distribution (GEVD) to study the tail behaviour of the dam's water level. However, they discovered evidence of the water level falling below the dam's minimum operation level in the future.

In the aspect of risk-management strategies in the financial industry, Nortey et al. (2015) modelled the outlier values of the Ghana Stock Exchange all-shares index by applying the conditional approach of the EVT to model the tails of the daily stock returns data. Applying the Peaks-Over-Threshold method, they fitted an ARMA-GARCH model to the data to correct for the effects of autocorrelation and conditional heteroscedastic terms present in the data. They also employed the Maximum Likelihood (ML) method to estimate the model's parameters. They established that the EVT model provided a better fit to the tail distribution of the returns because the data set contained more severe observations. Moreover, the observed volatility in the daily returns data justified the conditional EVT approach as the preferred statistical technique for the study. In addition, the Maximum Likelihood (ML) estimates for the parameters of the GPD were also found to be associated with lower standard errors as compared to the Probability Weighted Moment (PWM) estimates when applied to the large data set. In the end, their findings established the GPD as a best fit for dealing with behaviour of extreme share index in the stock market.

Elsewhere around the world, we see the employment of EVT in decisions of selection and pricing of portfolio. Danielsson and De Vries (2000) proposed a semi-parametric EVT-based method for Value-at-Risk (VaR) estimation for a portfolio of stocks. They stated that while there exist several methods for VaR estimation - some based on the use of conditional volatilities like the RiskMetrics method and others relying on unconditional historical simulation method; the EVT-based VaR for estimation of tail probabilities performs better. They illustrated that the GARCH-based RiskMetrics method underpredicts VaR and the Historical simulation also suffers from a high variance and discrete sampling in the tails of the distribution.

In addition, the historical simulation is not able to capture losses outside the sample it analyses. They concluded that the reason for the improved performance of the extreme value method is in its combination of some of the advantages of the other competing methods. Gencay and Selcuk (2004) also arrived at the same conclusion that, the dynamic-EVT based approach for calculating VaR gave more accurate forecast than other competing approach and also has the advantage of reacting to changing market conditions.

Again, by using the EVT framework, Allen et al. (2011) computed Value-at-Risk (VaR) measure in order to accurately assess the risk of severe losses on investment, especially when the distribution of negative returns is characterized by heavy realisations that sometimes fall outside the bounds of a normal distribution. In a similar vein, Uppal (2013) examined the performance of market risk measures based on Dynamic EVT to estimate the probabilities of large deviations of returns in financial markets for five developed and five emerging markets before and during the Global Financial Crisis (GFC). In his work, he established that, the Value at Risk (VaR) measure based on the Dynamic EVT procedure performed best among the other VaR estimation methods. He further indicated that the GPD model fits the observed distribution of extreme values well, in both pre-crisis and the crisis periods, with the exception of the US markets where the crisis originated from.

Another field where EVT can "play" is sports science. Sérgio (2012) applied the EVT to estimate the maximal amount of oxygen needed to break records set by top athletes. Using both the parametric and semi-parametric EVT approaches, he examined the probability of exceeding Usain Bolt's minimum 9.58 seconds speed in the 100-meters race and the (96 ml-kg-min) maximum oxygen uptake for speed by Bjorn Daehlie and ESpen Harald Bjerke - world's fastest snow-skiers. In the maximum oxygen ($V0^2max$) snow-skiing case analysis, he indicated that the semi-parametric approach rendered a more concordant findings with those obtained under the Peaks-Over-Threshold (POT) method when the Probability Weighted

Moment (PWM) method is used for parameter estimation. He also demonstrated that the estimates for the right endpoint of the distribution functions are asymptotic to the sample maximum. In other words, the record of Bjorn Daehie and Espen Harald Bjerke can hardly be broken. On the other hand, in the analysis involving the 100-meters race, the semi-parametric approach yielded similar results as that of the POT method when the Maximum Likelihood (ML) method is used for parameter estimation. Furthermore, the analysis revealed that, in the present circumstances, a 100-meters-race athlete can improve his speed, with a likelihood of breaking the current record of Usain Bolt under 9 seconds.

EVT procedures have also be utilised in the insurance industry to provide appropriate tools to mitigate losses (Embrechts et al., 2013). Pérez-Fructuoso and García Pérez (2010) demonstrated one of the performances of EVT in the Spanish motor liability insurance industry. They employed the EVT framework to perform a parametric estimation by fitting loss severities to the Generalized Pareto Distribution. They illustrated how EVT improves classicalist's adjustments by going beyond the average to consider outlier behaviour. Furthermore, they demonstrated that adjustments with EVT-based distributions in modelling the tails of loss severities significantly improve tail-distribution inference. They concluded that the EVT deals well with behaviour of extreme events and allows both the insurer and reinsurer to evaluate the expectation of losses over a certain level, and the premiums sufficient to cover such losses.

In Africa, particularly Kenya and Nigeria, mention can be made of Wainaina and Waititu (2014), Adesina et al. (2016), and Weru and Waititu (2019) who applied the EVT to the insurance sector of their economy. Wainaina and Waititu (2014) employed the extreme value theory approach to model huge losses arising from Kenya's fire insurance industry. Empirical analysis carried out by them revealed that the GPD model is the best fit to the returns since the distribution of returns was heavy-tailed. In addition, they performed an EVT-based

quantile estimation of the distribution of excesses over a seemly high threshold at different confidence intervals. Their results demonstrated how well the EVT applies to the insurance industry in terms of dealing with large but infrequent claims.

In a related work, Weru and Waititu (2019) also centred his study on the tail behaviour of claims in the motor insurance sector using Kenindia Assurance Company limited. He employed the EVT framework as a study practise to help in determining company's future maximum loss and the frequency with which they occur. He fitted a GPD model to large claims above a certain threshold, and estimated risk measures such as Value-At-Risk and Expected Shortfall based on extreme value analysis. He then compared the results to other competing methods of VaR estimation, and concluded that though the other methods of estimation perform well, the EVT approach to modelling seems to perform better with fat-tailed data. A replicated study of Weru and Waititu (2019) was carried out by Adesina et al. (2016) in Nigeria. Adesina et al. (2016) estimated the Value-at-Risk quantile by using the EVT methods to model large losses of Nigeria's motor insurance business. They chose the GPD approach over the GEV approach because of their interest in the overall behaviour of the tail area of the distribution of returns, rather than just the maxima loss in each year. They also compared EVT-based VaR estimates to VaR based on Historical and Gaussian methods at 5% confidence interval. and concluded that the EVT-based VaR is most suitable to calculate VaR as long as outlier behaviour of a dataset is of concern.

Coles (2001) listed a few other areas where extreme value analysis have been applied: alloy strength prediction, ocean wave modelling, memory cell failure, wind engineering, biomedical data processing, thermodynamics of earthquakes, non-linear beam vibrations, and food science. In this study, the focus is on applying the EVT methods to the health insurance sector of the Ghanaian economy in which we find and fit an appropriate limiting distribution to large claims amount submitted by the 37 Military Hospital. We also examine the likelihood

of very large claims occurring above a fixed high amount, and the time interval at which these extreme claims occur.

Chapter 3

METHODOLOGY

One of the main purposes of statistics is to take and analyse a portion of a population in order to make conclusive statements about the characteristics of the population as a whole. Usually, if focus is on the behaviour of the central part of a distribution, then Central Limit Theory (CLT) helps in addressing related questions. However, if the subject of interest is explicit about the behaviour of the largest (or lowest) realisations, it becomes statistically prudent to employ Extreme Value Theory (EVT) - a set of statistical techniques for modeling and estimation of rare events (Minkah, 2016; Nortey et al., 2017) - to address related questions. We begin by first considering the convergence of maxima.

3.1 Convergence of Maxima

One of the key concerns in the financial sustainability of any private health insurance scheme lies in making accurate projections maximum claims and when they occur.

Consider $X_1, \ldots, X_n$ as a sample of independent and identically distributed claims having a common continuous but unknown distribution F, then we can define the maximum value of

the sequence as:

$$M_n = max\{X_1, \ldots, X_n\}. \tag{3.1}$$

but as the sample size, $n \to \infty$, $M_n \to x^F$, where x^F is the upper endpoint of F and is defined as $x^F = sup\{x \in \Re : F(x) < 1\}$. Thus, when we increase the sample size and we consider the largest observation, it will keep moving to the right and finally converge to the upper endpoint.

Also, the probability distribution of M_n is defined as:

$$\begin{aligned} P(M_n \leq x) &= P(X_1 \leq x, \ldots, X_n \leq x) = P(X_1 \leq x) \times \cdots \times P(X_n \leq x) \\ &= F(x)^n = F^n(x). \end{aligned} \tag{3.2}$$

But as $n \to \infty, F^n(x) \to 1$ for $x \geq x^F$ and $F^n(x) \to 0$ for $x < x_F$ (where x^F and x_F represent the upper and lower endpoints respectively) so that M_n has a degenerate limiting distribution. But for mathematical and statistical implementation, we need a non-degenerate function. Therefore, in order for the function F^n not to degenerate on the upper endpoint, as n increases, we rescale M_n by applying normalising constants for the location and scale. This leads us to:

$$M_n^* = \frac{M_n - b_n}{a_n}, \text{ where } a_n \text{ and } b_n \text{ are normalising constants..}$$

We remark that even though the focus of this paper is on modelling the behaviour of sample maxima, results for sample minima can be obtained from:

$$min(X_1, \ldots, X_n) = -max(-X_1, \ldots, -Xn).$$

3.2 Maximum Domain of Attraction

Usually, it is difficult and sometimes practically impossible to know the true probability distribution of a set of data in real life. We therefore use limiting (asymptotic) distribution to approximate the true distribution of the data. Let us consider the limiting distribution for a large number of M_n, where M_n is as defined in (1).

Given $a_n > 0$ and b_n as any normalising constants such that $\frac{M_n - b_n}{a_n}$ converges in distribution, then from (2),

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= P(M_n \leq a_n x + b_n) \\ &= F^n(a_n x + b_n) \approx G_\xi(x), \text{ as } n \to \infty. \end{aligned} \tag{3.3}$$

where G_ξ is a non-degenerate distribution function. Thus, the unknown distribution F for the maxima, is said to be in the Maximum Domain of Attraction (MDA) of some distribution G_ξ; symbolically stated as, $F \in MDA(G_\xi)$. In other words, for a sufficiently large sample, the rescaled sample maxima of a data with unknown distribution F is attracted to the limiting distribution G_ξ which also belongs to a unique family of distributions - Gumbel family, Fréchet family, and Weibull family (Fisher and Tippett, 1928).

3.2.1 The Gumbel Family of Distributions

This class of distributions is usually characterised by density functions whose tails decay exponentially (or light-tailed) as we move towards the upper endpoint of the distribution. They have distribution functions characterised by a shape parameter, $\xi = 0$, and are of the form:

$$G_\xi(x) = e^{-e^{-\left(\frac{x-b}{a}\right)}, \ -\infty<x<\infty}, \tag{3.4}$$

where a and b are normalising constants (Coles, 2001). Though both $x^F < \infty$ and $x^F = \infty$ are possible for the upper endpoint, all moments exist for distributions in the Gumbel family. The normal, log-normal,exponential, χ^2, gamma, and gumbel are examples of some of the distributions attracted to the MDA of the Gumbel family.

3.2.2 The Fréchet Family of Distributions

The Fréchet family is made up of heavy-tailed distributions with shape parameter $\xi > 0$ and an infinite upper endpoint. A distribution is said to be heavy-tailed if we observe a polynomial decay in the tails of its density function as we move towards the upper endpoint of the distribution. This class of distributions is very often encountered in finance and insurance data. The Fréchet class of distributions take the form:

$$G_\xi(x) = \begin{cases} e^{-\left(\frac{x-b}{a}\right)^{-\frac{1}{\xi}}}, & x > b, \text{ for } \xi > 0 \\ 0, & x \leq b \end{cases}, \tag{3.5}$$

where a and b are normalising constants and ξ is the shape parameter (Coles, 2001). Also, $E(X^k) < \infty$ if $k < -\frac{1}{\xi}$ and $E(X^k) = \infty$ if $k \geq -\frac{1}{\xi}$. Examples of distributions that are attracted to the maximum domain of the Fréchet family are Cauchy, Pareto, Fréchet, Student t, and Log-gamma.

3.2.3 The Weibull Family of Distributions

This class of distribution always has a finite upper endpoint and a short decay in the tails of its density function (i.e. light-tailed distribution). It has a negative shape parameter and an upper bounded tail at: $b - \frac{a}{\xi}$. The distribution function of the Weibull class is of the form:

$$G_\xi(x) = \begin{cases} e^{-\left(-\frac{x-b}{a}\right)^{-\frac{1}{\xi}}}, & x < b, \text{ for } \xi < 0 \\ 1, & x \geq b \end{cases}, \tag{3.6}$$

where a and b are normalising constants and ξ is the shape parameter (Coles, 2001). Examples of distributions in the maximum domain of attraction of the Weibull family are Beta distribution, Weibull distribution, and Uniform distribution.

The above families of distribution mentioned are the only possible extreme value limiting distributions for the normalised sample maxima, regardless of the underlying distribution function from which the samples are drawn from (Fisher and Tippett, 1928; Gnedenko, 1943; Coles, 2001). As noticed, each of the three families of extreme value distribution give different representation of extreme value behaviour. These representations are primarily determined by the tail characteristic of the distribution function (i.e. the shape parameter ξ); and this can be inferred from the behaviour of the upper endpoint of the underlying distributions.

Whiles the upper endpoint for the Weibull class of distributions is finite and that of Fréchet class is infinite, the Gumbel class of distributions can assume both finite and infinite upper endpoint. However, the density function of the Fréchet class of distributions decays faster than that of Gumbel class of distributions. These three standard extreme value distribution classes can be put into practice by EVT through two main approaches - Block Maxima Method (BMM) and Peaks-Over-Threshold Method (POTM). Let us consider the former in the subsequent section.

3.2.4 The Block Maxima Method (BMM)

The BMM works by splitting the data into groups (blocks) of equal length, say yearly, monthly, weekly, or even daily. Depending on the available data, the groups can also be formed according to strength, weakness, or even speed (Sérgio, 2012). The BMM then considers the maximum of each block, sets them as a new data (maxima) and then runs analysis by fitting the GEV distribution to the normalised maxima. This approach operates

very well for the environmental sciences to the natural blocking (seasons) in nature. However, it can also be used for financial and insurance data even though not highly recommended (Embrechts et al., 2013).

One of the main problems associated with the BMM is the choice of block length - a compromise on "bias - variance" relationship. Thus, small block lengths generate more sample maxima for model estimation and consequently reduces the variance and improves the accuracy of the model. However, choosing small block lengths may include intermediate observations which could increase the biasness of parameter estimation in the model.

In contrast, choosing large block lengths leads to small number of maxima which satisfy the asymptotic assumptions underlying the model. This eventually results in small bias but at the expense of large variance. For this reason, the use of block length of one year has become more pragmatic when deciding on block length for EVT analysis in climatological and hydrological data, and hence the name Annual Maxima Method (AMM).

3.2.5 The Generalised Extreme Value Distribution (GEVD)

In practice, it follows then to select one of the three limiting distribution families and to estimate the parameters of that distribution when applying the EVT framework. However, there is the challenge of choosing a technique for the right selection of the distribution that is best for the data at hand. In order to simplify statistical implementation, Mises (1936) and Jenkinson (1955) unified these three families of distribution into a single "three-parameter" family of distributions called the Generalised Extreme Value Distribution (GEVD). The advantage of this new model is its ability to determine the domain of attraction by using only the shape parameter. This means that, through inference on the shape parameter, there is no need for subjective prior selection of the best limiting distribution to adopt for the data.

Definition: Suppose that $M_n = max\{X_1, \ldots, X_n\}$ for a set of i.i.d. random variable X, and given any normalising constants $a > 0$ and b such that $P\left(\frac{M_n - b}{a} \leq x\right) \approx G_\xi(x)$ as $n \to \infty$, where $G_\xi(x)$ is a non-degenerate distribution function, then according to Jenkinson (1955), $G_\xi(x)$ belongs to the family of distributions which take the form:

$$G(x; \mu, \sigma, \xi) = \begin{cases} e^{-[1+\xi\left(\frac{x-\mu}{\sigma}\right)]^{-\frac{1}{\xi}}}, & \text{for, } \xi \neq 0, \ 1 + \frac{\xi(x-\mu)}{\sigma} > 0 \\ e^{-e^{-\left(\frac{x-\mu}{\sigma}\right)}}, & \text{for, } \xi = 0, \ -\infty < x < \infty \end{cases}, \tag{3.7}$$

where $-\infty < \mu < \infty$, $\sigma > 0$ and ξ represent location, scale and shape parameters respectively. Furthermore, the distribution function G is in the Gumbel domain if $\xi = 0$; Fréchet domain if $\xi > 0$ and Weibull domain if $\xi < 0$. In literature, ξ is usually referred to as the Extreme Value Index (EVI) or tail index as it is a key feature in determining the heaviness of the tail of an underlying unknown distribution function, and consequently its asymptotic distribution.

Sometimes, we can have models from different classes of distributions as a plausible fit to a dataset. It then becomes necessary to run more objective tests in order to select the model that best fits the data. For instance, given that the shape parameter of a proposed model belonging to the Weibull class of distributions has a confidence interval that includes zero, or has an upper confidence interval which is very close to zero; it means that the data at hand can also be fitted with another model from the gumbel class of distributions. Hence, we test for the feasibility of the "gumbel-class" model in such circumstances.

The first point of call will be the diagnostic plots for the "Gumbel-fit". Here, an "eye-ball" inspection of the linearity in both the probability and quantile plot can be a good indicator of which model is superior given the available data. Thus, we choose the model in which such linearity is better achieved.

We can also perform the following hypothesis test to determine if the "Gumbel-class" model is appropriate for the data or not by setting H_0: The "Gumbel-class" model is fit for the data *vs.* H_1: The "Gumbel-class" model is not fit for the data. We reject H_0 if the test statistic is greater than the critical value (or if the p-value is less than the significance level of the test). Here, the test statistic used is called the Gumbel statistic (Sérgio, 2012) and is given by:

$$GS_m = \frac{Y_{m:m} - Y_{\frac{m}{2}+1:m}}{Y_{\frac{m}{2}+1:m} - Y_{1:m}}.$$

Under the test, the following holds for the validity of H_0:

$$GS_m^* = \frac{GS_m - b_m}{a_m} \rightarrow_{m\rightarrow\infty} \Lambda$$

where b_m and a_m are coefficients of attraction to the Gumbel class of distributions (Sérgio, 2012), and therefore we reject H_0 if $GS_m^* \leq G_\alpha$ at a significance level α, where G_α represents the standard Gumbel quantile.

Last but not least, we can also compare the negative loglikelihood, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) of the two models and choose the model with the smallest score (Gilleland and Katz, 2016).

When comparing models fitted by maximum likelihood to the same data, the negative log-likelihood can be used as a measure of model fit. Assuming $F_\theta(x)$ is the distribution function of the random variable X, then the likelihood function is given as $L(\theta|x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\theta)$ where x is the observed value and θ is the parameter of interest and the log-likelihood as $\ln L(\theta|x_1, \ldots, x_n) = \sum_{i=1}^{n} lnf(x_i|\theta)$. The MLE then tends to maximise the likelihood $L(\theta)$ over all possible values of x; and the higher the value, the "better". Consequently, a lower negative log-likelihood value (i.e. close to 0) is an indication of a better model fit.

As explained by Pinheiro and Grotjahn (2015) and also by Gilleland and Katz (2016), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are statistical measures that estimates the relative quality of a model by assessing the amount of information lost by the model taking into account the model's structure and its goodness-of-fit. Thus, these two statistical measures are aimed at achieving parsimony in the model by penalising models with more parameters. The smaller the AIC / BIC (i.e. the less information lost), the more preferable the model is. The AIC is given as $2k - 2ln(L)$, whilst the BIC is given as $[ln(n) - ln(2\pi)]k - 2ln(L)$ where k is the number of parameters, L is the maximised value of the likelihood function and n is the number of data points.

3.2.6 Estimation of Parameters

Usually, it is difficult to know the exact numerical value that summarises the properties of an underlying unknown distribution; we can only approximate it with a "near-value" by estimating the parameters of the corresponding limiting distribution. In the case of fitting the Generalised Extreme Value Distribution (GEVD) to the sample maxima, there is the need to estimate the value of the location (μ), scale (σ), and shape (ξ)) parameters if we desire to make any meaningful inference about the behaviour of extreme events.

In literature, there exist a number of estimation methods for the parameters of the GEVD. These include Maximum Likelihood Estimation, L-Moments, Probability Weighted Moments, and the Bayesian estimation. In this section, we shall consider the two most popular ones used in several EVT applications - Maximum likelihood Estimation (MLE) and Probability Weighted Moments Estimation (PWME).

The method of maximum likelihood is a routine procedure for obtaining the estimators for

unknown parameters of a distribution of a data set. It assumes that the data are realisations of independent and identically distributed random variables and then defines a likelihood function in terms of the observed data and the parameters. The ML estimate is then obtained by maximising the likelihood function or the corresponding log-likelihood function - provides a monotone transformation by changing the products into sums - with respect to the unknown parameter of interest. Though for smaller sample size (say, $n < 50$), the MLE is unstable and can give unrealistic estimates for the shape parameter (Coles and Dixon, 1999), nevertheless, the adaptability of the Maximum Likelihood to changes in model structure such as missing values and non-stationarity, as compared to other parameter estimation methods, makes it more preferable (Coles, 2001; Sérgio, 2012).

Now, suppose that $x_1, \ldots, x_m$ are independent realisations of sample maxima with a common GEVD, the log-likelihood function in terms of the GEVD parameters when $\xi \neq 0$ is of the form: $\ell(\mu, \sigma, \xi | x_1, \ldots, x_m) =$

$$-m \ln \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{m} \ln \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^{m} \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \tag{3.8}$$

provided

$$1 + \xi \left(\frac{x_i - \mu}{\sigma}\right) > 0, \text{ for } i = 1, \ldots, m \tag{3.9}$$

and when $\xi = 0$, the log-likelihood is given as:

$$\ell(\mu, \sigma, \xi | x_1, \ldots, x_m) = -m \ln \sigma - \sum_{i=1}^{m} \left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{m} e^{-\left(\frac{x_i - \mu}{\sigma}\right)} \tag{3.10}$$

The maximum likelihood estimators $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ for the unknown parameters (μ, σ, ξ) is found by maximising (8) or equivalently (10) with respect to the parameter of interest.

As stated by Coles (2001), the standard MLE regularity conditions for asymptotic efficiency and consistency are not automatically applicable in EVT. This is because the endpoints of the GEVD are functions of the parameter values, where $\mu - \frac{\sigma}{\xi}$ is the upper and lower endpoint for GEVD if $\xi < 0$ and if $\xi > 0$ respectively. This limitation makes the common regularity conditions that is associated with the MLE to be invalid. Smith (1985) studied this problem and stated the following restrictions on the basis of the shape parameter ξ:

- The MLE is regular, and have standard asymptotic properties when $\xi > -\frac{1}{2}$
- The MLE is unlikely attainable when $\xi < -1$
- The MLE is attainable, but do not have standard asymptotic properties when $-1 < \xi < -\frac{1}{2}$

The second most popular estimation method employed in EVT is the Probability Weighted Moments Estimator (PWME). Introduced by Landwehr et al. (1979) and Greenwood et al. (1979), the PWME is a generalised version of the Method of Moments Estimator (MME) technique for parameter estimation. However, the PWME assigns more weight to tail observations of the distribution. According to Diebolt et al. (2008), the PWME is commonly employed in hydrological and climatological data due to its easy implementation and performance for distributions encountered in the geosciences. Similar to the classical Method of Moments Estimation (MME), the underlying notion of the PWME is to match the moments with their empirical counterparts. Thus,

$$M_{p,r,s} = E[Y^p (F(Y))^r (1 - F(Y))^s] \tag{3.11}$$

where $p, r,$ and s are real numbers and Y is a random variable with distribution function F. In the context of a GEVD for a sample maxima $(y_1, \ldots, y_p)$, parameter estimation by use of PWME is given by

$$\beta = M_{1,r,0} = E[Y(F(Y))^r], \; r = 0, 1, \ldots \tag{3.12}$$

As demonstrated by Landwehr et al. (1979), suppose we have a non-decreasing order statistics $Y_{1:p} \leq \cdots \leq Y_{p:p}$ associated with the random sample maxima, $(Y_1, Y_2, \ldots, Y_p)$, which has been divided into p blocks of size q, and assuming the sample follows a GEVD, then the empirical estimator $\hat{\beta}_r$ of the PWME is of the form:

$$\begin{aligned} \hat{\beta}_r = \hat{M}_{1,r,0} &= \frac{1}{p}\sum_{i=1}^{p}\left(\frac{(i-1)(i-2)\cdots(i-r)}{(p-1)(p-2)\cdots(p-r)}\right)Y_{i:p} \\ &= \frac{1}{p}\sum_{i=1}^{p}\left(\prod_{k=1}^{r}\frac{i-k}{m-k}\right)Y_{i:p} \end{aligned} \tag{3.13}$$

Hosking et al. (1985) studied and estimated the parameters of the GEVD for the sample described using the PWME and proved that when $\xi \neq 0$, the PWME is of the form:

$$\begin{aligned} \beta &= M_{1,r,0} = E[Y(F(Y))^r] \\ &= \frac{1}{r+1}\left[\mu - \frac{\sigma}{\xi}(1-(r+1)^{\xi}\Gamma(1-\xi)\right] \text{ for } \xi < 1 \text{ and } \xi \neq 0 \end{aligned}$$

where $r = 0, 1, \ldots$ (3.14)

Hence for $r = 0, 1, 2$, the PWM estimators $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ for the unknown parameters (μ, σ, ξ) is obtained by solving the following system of equations

$$\left.\begin{array}{rll} \beta_0 &= M_{1,0,0} & = \mu - \frac{\sigma}{\xi}(1-\Gamma(1-\xi)) \\ 2\beta_1 - \beta_0 &= 2M_{1,1,0} - M_{1,0,0} & = \frac{\sigma}{\xi}\Gamma(1-\xi)(2^{\xi}-1) \\ \frac{3\beta_2-\beta_0}{2\beta_1-\beta_0} &= \frac{3M_{1,2,0}-M_{1,0,0}}{2M_{1,1,0}-M_{1,0,0}} & = \frac{3^{\xi}-1}{2^{\xi}-1} \end{array}\right\} \tag{3.15}$$

so that,

$$\left.\begin{array}{rl} \hat{\mu} &= \hat{M}_{1,0,0} + \frac{\hat{\sigma}}{\hat{\xi}}(1-\Gamma(1-\xi)) \\ \hat{\sigma} &= \frac{\hat{\xi}(2\hat{M}_{1,1,0}-\hat{M}_{1,0,0})}{\Gamma(1-\hat{\xi}(2^{\xi}-1))} \\ \frac{3\hat{M}_{1,2,0}-\hat{M}_{1,0,0}}{2\hat{M}_{1,1,0}-\hat{M}_{1,0,0}} &= \frac{3^{\xi}-1}{2^{\xi}-1} \end{array}\right\} \tag{3.16}$$

where the moments β_r is replaced by their empirical counterparts in (13), and $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ corresponds to $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$. It worth noting that ξ is obtained by solving the last equation of (16) numerically.

For $\xi = 0$, the PWME $(\hat{\mu}, \hat{\sigma})$ is given by the solution to the equation systems:

$$\left.\begin{aligned} \hat{M}_{1,0,0} &= \mu + \sigma\Gamma(1) \\ 2\hat{M}_{1,1,0} - \hat{M}_{1,0,0} &= \ln 2\sigma \end{aligned}\right\} \tag{3.17}$$

where $\hat{M}_{1,r,0} = \frac{1}{p}\sum_{i=1}^{p}\left(\prod_{k=1}^{r}\frac{i-k}{m-k}\right)Y_{i:p}$ so that

$$\left.\begin{aligned} \hat{\mu} &= \hat{M}_{1,0,0} - \sigma\Gamma(1) \\ \hat{\sigma} &= \frac{2\hat{M}_{1,1,0} - \hat{M}_{1,0,0}}{\ln 2} \end{aligned}\right\} \tag{3.18}$$

Hosking et al. (1985) indicated that for small sample sizes, the PWME outperforms the MLE in terms of minimum variance. He continued that this advantage has made the PWME more preferable in recent analysis of extreme values because of sparse data (few observations) which is usually available for extrapolation of rare events. However, he mentioned that unlike the MLE, the PWME has difficulty in estimations with complex model structures.

3.2.7 Estimation of Extreme Quantiles

Aside the location, scale, and shape parameters that are usually estimated when fitting the GEVD for inferences on the behaviour of extreme occurrences; there is the need to also estimate other equally important parameters such as return levels and their corresponding return periods and exceedance probabilities. These are extreme quantiles that form part of the characterisation of the behaviour of uncommon events. Stated differently, the expected time interval (i.e. return period) within which extreme events are likely to occur or reoccur,

and the magnitude (i.e. return level) with which they are expected to occur, at a given probability (i.e. exceedance probability) are of key importance for the behavioural study of extreme events.

Assuming an i.i.d. random variable Y with a distribution function $F(Y) = P(Y \leq y)$, then for the probability $0 < \alpha < 1$, the α^{th} quantile of Y is given as the inverse cumulative distribution function:

$$Q(\alpha) = F^{-1}(\alpha) = inf(y : F(y) \geq \alpha) \tag{3.19}$$

Now, let $y_{1-\alpha}$ be the extreme quantile of order $(1-\alpha)$ of the GEVD for a random maxima variable Y. With a sufficiently small probability α, we can obtain estimates of extreme quantiles by inverting the GEVD function in (7). Thus,

$$y_{1-\alpha} = G_{\xi}^{-1}(1-\alpha) = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - (-\ln(1-\alpha))^{-\xi}], & \text{if } \xi \neq 0 \\ \mu - \sigma \ln(-\ln(1-\alpha)), & \text{if } \xi = 0 \end{cases} \tag{3.20}$$

where the parameters (μ, σ, ξ) is replaced by their estimators $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$.

Now, setting $y_{\alpha} = -\ln(1-\alpha)$, we can rewrite $eqn(20)$ in terms of the GEV parameters as:

$$y_{1-\alpha} = G_{\xi}^{-1}(1-\alpha) = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - y_{\alpha}^{-\xi}], & \text{if } \xi \neq 0 \\ \mu - \sigma \ln(y_{\alpha}), & \text{if } \xi = 0 \end{cases} \tag{3.21}$$

Recall that for $\xi < 0$, the limiting GEVD has a finite upper endpoint, x^F which can be estimated as:

$$x^F = \mu - \frac{\sigma}{\xi}, \tag{3.22}$$

where μ, σ, and ξ are replaced by their respective ML or PWM estimator.

Following the BMM, suppose $M_p = max(Y_1, \ldots, Y_p)$ denotes the maximum of a block for i.i.d. random variable Y with unknown distribution F; and assuming the obtained p sample maxima $(M_{p1}, \ldots, M_{pp})$ follows a GEVD, $G_{\mu,\sigma,\xi}$, then from $eqn(3)$, we can write

$$F_Y \equiv F_{M_p} = F^p \approx G_{\mu,\sigma,\xi} \tag{3.23}$$

so that the $(1-\alpha)$ quantile of the original data Y, denoted $y_{1-\alpha}$, can be written using the relations in (20) as:

$$F^{-1}(y_{1-\alpha})^p = (1-\alpha)^p \approx G^{-1}_{\mu,\sigma,\xi}(y_{1-\alpha})$$

so that

$$y_{1-\alpha} = G^{-1}_{\mu,\sigma,\xi}(1-\alpha)^p = \begin{cases} \mu - \frac{\sigma}{\xi}[1-(-\ln(1-\alpha)^p)^{-\xi}], & \text{if } \xi \neq 0 \\ \mu - \sigma \ln(-\ln(1-\alpha)^p), & \text{if } \xi = 0 \end{cases} \tag{3.24}$$

Again, the parameters (μ, σ, ξ) is replaced by their ML or PWM estimators $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$.

Now, given that the distribution of sample maxima observed over a non-overlapping and equal time periods follows a GEVD G_ξ, the return level can be defined as:

$$R^k_n = \left(1 - \frac{1}{k}\right),$$

where R^k_n is the return level expected to be exceeded with the probability $\frac{1}{k}$, once every k intervals (or return periods) of length n. Hence from (20) we can write the return level estimate as:

$$\hat{R}^k = G^{-1}\left(1 - \frac{1}{k}\right) = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}[1-(-\ln(1-\frac{1}{k}))^{-\hat{\xi}}], & \text{if } \xi \neq 0 \\ \hat{\mu} - \hat{\sigma} \ln(-\ln(1-\frac{1}{k})), & \text{if } \xi = 0 \end{cases} \tag{3.25}$$

On a logarithmic scale, a return level plot (i.e. a plot of $(1-\alpha)$ against $\ln(y_\alpha)$) should be linear when $\xi = 0$, or should be convex with a finite endpoint when $\xi < 0$, and concave with an infinite endpoint when $\xi > 0$. Coles (2001) holds that care should be taken when interpreting inferences on the return levels corresponding to long return periods because precision measures of an estimator are based on an assumption that the model is correct.

3.3 Peaks-Over-Threshold Method (POTM)

The Peaks-Over-Threshold Method (POTM) is the second main approach used for analysing behaviour of rare events in EVT. It considers all available data and focuses on the observations that exceed a given fixed high level (threshold). It then fits an appropriate parametric model to the excesses (exceedances) over that fixed high threshold with an asymptotic distribution (i.e. the Generalised Pareto Distribution (GPD)) given that the requisite assumptions are satisfied.

Suppose $X_1, X_2, \ldots, X_n$ are a sequence of independent and identically distributed random variables with an unknown distribution function $F(x) = P(X \leq x)$. Given a designated high threshold u, we consider those X beyond u as being exceedances of u. That is to say, an exceedance of a threshold u occurs whenever $X > u$. In this sense, we can define the excess value, y, over a given threshold u as $y = x - u$ where $0 < y < x^F - u$ with x^F being the upper endpoint of the distribution function F. We then define the probability distribution (also called the conditional excess distribution function) of Y as:

$$\begin{aligned} F_u(y) &= P(X - u \leq y | X > u) \\ &= \frac{P(X \leq y+u, X > u)}{P(X > u)} \\ &= \frac{F(y+u) - F(u)}{1 - F(u)} \\ F_u(y) &= \frac{F(x) - F(u)}{1 - F(u)} \text{ since } x = y + u \end{aligned} \tag{3.26}$$

so that

$$F(x) = F_u(y)[1 - F(u)] + F(u) \tag{3.27}$$

In practise, it is not possible to know the exact distribution of F in order to know the corresponding distribution of the excesses, hence we use asymptotic distributions. Cited in Coles (2001), Balkema and de Hann (1974) and Pickands (1975) provided that, for a sufficiently large threshold u, the distribution of exceedances can be approximated by the Generalised Pareto Distribution (GPD). Thus, Balkema and de Haan (1974) and Pickands (1975) showed that if F is a distribution function of a random sample $(X_1, \ldots, X_n)$ with an upper endpoint $(x^F = sup[x : F(x) < 1])$; and F_u is the distribution for the excesses $(y : y = x - u)$ over a threshold u, where $y \in [0, x^F - u)]$ then

$$\lim_{u \to x^F} |F_u(y) - H_{\xi,\sigma_u}(y)| = 0 \tag{3.28}$$

where $\sigma_u > 0$ and ξ are the scale and shape parameter of the GPD function H. In other words, the GPD family is the limiting distribution for normalised excesses over an increasingly large threshold.

3.3.1 Selection of Threshold

The POTM works by analysing observations that exceed a predetermined high threshold level. Similar to the issue on the choice of block length for estimation purposes in the BMM, there is also the issue of choice of threshold level for estimation purposes in the POTM and that decision still remains a subjective matter - implying a compromise between bias and variance.

Stated clearly, fixing a low threshold level increases the number of observations (exceedances)

for model estimation and leads to more precise model (i.e. a decrease in model's variability). Nonetheless, it introduces some less extreme observations from the centre of the distribution and this leads to an increase in estimation bias. On the other hand, selecting a high threshold level only gives off few exceedances and renders the estimator to be highly sensitive (i.e. a high variance estimator) to sample size . However, it ensures that the estimated value is close to the true parameter value.

As a result of this "trade-offs", a careful combination of some techniques are considered in determining a suitable threshold that provides a reasonable approximation for the limiting model. Two graphical methods were proposed by Coles (2001) in that regard: the first method is based on fitting the model at a range of different thresholds and then selecting the threshold level where there seems to be stability in the parameter estimates; and the second method is based on the mean excess function (or mean excess plot) - a tool generally used to help in the selection of threshold level and also to determine the suitability of the GPD model. This tool is quiet common in the field of insurance.

Considering the latter tool, suppose we assume a random variable X with a distribution F and given a fixed high threshold u, then the distribution of excesses of X over u, i.e. $Y = X - u$, is given as:

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}$$

for $0 \leq y \leq x^F - u$. Also, if $E[X]$ is finite, then the Mean Excess Function (MEF) - the mean w.r.t. F_u - is defined as:

$$e(u) = E(X - u | X > u) = \int_0^{x^F - u} \bar{F}_u(y)dy = \frac{\int_u^{x^F} \bar{F}(x)dx}{\bar{F}(u)} \tag{3.29}$$

Additionally, suppose the mean of a random variable X having a GPD with parameter σ

and ξ is given as $E(X) = \frac{\sigma}{1-\xi}$, then the distribution of excesses of X over a fixed threshold u also follows a GPD with parameter $\sigma + \xi u$ and ξ, with a mean excess function defined as:

$$e(u) = \frac{\sigma_u + \xi u}{1 - \xi}, \text{ for all } u : \sigma_u + \xi u > 0 \tag{3.30}$$

Now, assuming the distribution of the random variable Y (excesses) over a threshold level u is $F_u(y) \approx H_{\sigma_u,\xi}$, where H is a GPD with parameters ξ and $\sigma_u > 0$, then for any higher threshold level $v \geq u$, the distribution of excesses over the higher threshold v remains a GPD with same shape parameter, but a varying scale parameter with a linear growth in v. Thus,

$$F_v(y) \approx H_{\sigma_u + \xi(v-u),\xi}$$

Therefore, for $\xi < 1$, the MEF for the higher threshold $v \geq u$ is given as:

$$e(v) = E[H_{\sigma_u+\xi(v-u),\xi}] = \frac{\sigma_u + \xi(v-u)}{1-\xi} = \frac{\xi v}{1-\xi} + \frac{\sigma_u - \xi u}{1-\xi}, \; v \in [u, x^F] \tag{3.31}$$

where $x^F = u - \frac{\sigma_u}{\xi}$ if $\xi < 0$ and $x^F = \infty$ if $0 \leq \xi < 1$.

Both (30) and (31) implies the fact that the MEF (also known as Mean Residual Life Function) is a linear function of the threshold. The diagnostic tool for examining this linearity is the sample mean excess plot.

Given the data $X_1, \ldots, X_n$, the MEF can empirically be estimated by its sample mean excess function as the sum of excesses over the threshold u divided by the exceedances (i.e. number these excesses). Thus,

$$e(u) = \frac{\sum_{i=1}^{n}(X-u)I_{X>u}}{\sum_{i=1}^{n} I_{X>u}}$$

where I is an indicator function and equal to 1 for $x > u$, or otherwise equals to 0. The

Sample Mean Excess Plot (SMEP) is therefore a plot of

$$\{(X_{(i,n)}, e_n[X_{(i,n)}]) : 1 \leq i \leq n-1\},$$

where $X_{(i,n)}$ denotes the i^{th} order statistic of the sample size n.

Coles (2001) stated that the estimates for the mean excess sample should change linearly with the threshold for which the GPD model is deemed appropriate. Concordantly, Davison and Smith (1990) used the property that the linearity of the mean excess function for increasing thresholds is an indication that the GPD model is valid for the underlying data. Gilleland and Katz (2016) stated that, in using the MEP to aid in threshold selection, the focal point is to locate the smallest possible threshold whereby a straight line can be deduced from that point to higher thresholds for a given uncertainty bounds.

Now, when the graph of the SMEP is a fairly positive straight line, then a Pareto Type I distribution (i.e. a GPD model with a positive shape parameter) is assumed to be a suitable fit for the underlying data. On the other hand, a fairly negative straight line is an indication that a Pareto Type II distribution (i.e. a GPD model with a shape parameter less than 0) will best fit the underlying data, and a fairly horizontal line would mean that data is exponentially distributed and hence the appropriate GPD model should have a shape parameter equal to 0. Also, if a GPD model for a threshold level u fits a data, then for higher threshold levels $v \geq u$, we apply the same GPD model but with a linear increase in the mean excess function $e_n(v)$.

Coles (2001) and Sérgio (2012) independently indicated that the interpretation of the mean excess plot (or mean residual plot) is not always straightforward especially when there are too few exceedances to make meaningful inferences. As a result, fixing of the threshold level

should be made at the point to the right where there is a roughly linearly pattern in the plot.

The next graphical technique aside the MEP that aids in the selection of threshold is the Parameter Stability Plot (PSP). This technique involves estimating the GPD parameters at a range of thresholds (or order statistics) and choosing that threshold where the estimated parameters becomes fairly stable (Coles, 2001; Beirlant et al., 2006). Coles (2001) indicated that above a threshold level u for which the GPD is a valid model, estimates of the shape parameter ξ should be approximately constant while that of the scale parameter σ_u should be linear in u. In sections that follows, we look at the model fitting and the estimation of the key parameters of interest.

3.3.2 The Generalised Pareto Distribution (GPD)

Suppose the probability, $P\left(\frac{M_n - b_n}{a_n} \leq x\right)$, of a normalised maxima, M_n, converges in distribution to a non-degenerate GEVD G_ξ as $n \to \infty$, where $a_n > 0$ and b_n are normalising constants; then for sufficiently high threshold u, the distribution function of the excesses, $(X - u)$, above u can be approximated by the Generalised Pareto distribution class (Coles, 2001; Sérgio, 2012 given as:

$$H(x; u, \sigma_u, \xi) = \begin{cases} 1 - \left(1 + \frac{\xi(x-u)}{\sigma_u}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0, \text{ for } 1 + \frac{\xi(x-u)}{\sigma_u} \geq 0 \\ 1 - e^{-\left(\frac{x-u}{\sigma_u}\right)}, & \text{if } \xi = 0, \text{ for } x \geq 0 \end{cases} \tag{3.32}$$

where ξ and σ_u are the shape and scale parameters respectively. Furthermore, for $\xi < 0$, $\xi > 0$, and $\xi = 0$, the GPD turns to the Weibull class of distributions (with an upper endpoint of $u - \frac{\sigma_u}{\xi}$), the Pareto (or Fréchet) class of distributions (with infinite upper endpoint), and the Gumbel class of distributions respectively. Likewise the GEVD, the shape parameter ξ is dominant in determining the characteristics of the underlying GPD.

3.3.3 Estimation of Parameters

In literature, there are several methods of estimating the GPD parameters. However, we will consider just the Maximum Likelihood Estimation (MLE) and the Probability Weighted Moment Estimation (PWME) due to their popularity in usage under EVT.

Likewise under the GEVD, using the MLE to approximate parameter values for the GPD requires the formation of the likelihood function, or equivalently, the log-likelihood function and then maximising that function with respect to the parameter of interest in order to get their estimates.

Suppose the i.i.d. random variables $X_1, \ldots, X_n \sim F$ where $F \in MDA(G_\xi)$, for some $\xi \in \Re$ and $X_1, \ldots, X_p$ are p exceedances of a threshold u, such that $Y_i = X_i - u,\ i \in (1, \ldots, p)$ are the corresponding excesses. Then with the excesses $y_1, \ldots, y_p$ independent and roughly distributed as a GPD function, $H(x; u, \sigma_u, \xi)$, the log-likelihood function is given as:

$$\ell(\sigma_u, \xi | y_1, \ldots, y_p) = \begin{cases} -p \ln \sigma_u - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{p} \ln\left(1 + \frac{\xi y_i}{\sigma_u}\right), & \text{if } \xi \neq 0 \\ -p \ln \sigma_u - \frac{1}{\sigma_u} \sum_{i=1}^{p} y_i, & \text{if } \xi = 0 \end{cases} \tag{3.33}$$

To simplify arithmetic, Coles (2001) recorded that Davison introduced a reparameterisation of the log-likelihood function to get $\hat{\xi}$ explicitly as a function of $\hat{\zeta}$ by defining $\zeta = \frac{\xi}{\sigma_u}$ so that the log-likelihood function can be rewritten as:

$$\ell(\sigma_u, \xi | y_1, \ldots, y_p) = -p \ln \xi + p \ln \zeta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{p} \ln(1 + \zeta y_i) \text{ if } \xi \neq 0 \tag{3.34}$$

The maximum likelihood estimators $(\hat{\sigma_u}, \hat{\xi})$ or $(\hat{\zeta}, \hat{\xi})$is then obtained by maximising (33) or respectively (34) with respect to the parameter of interest. For $\xi = 0$, we have the case of a classical exponential distribution.

We now look at the Probability Weighted Moment Estimation (PWME). We recall that for a random variable X, the PWM is defined as

$$M_{p,r,s} = E[(X^p(F(X)^r(1-F(X)))^s]$$

where $p, r, s \in \Re$. Hosking et al. (1985) proposed that to estimate the GPD parameters using the PWME, we take $M_{p,r,s}$ with $p = 1,\ r = 0,$ and $s = 0, 1, \ldots$ so that the moment is defined as:

$$M_{1,0,s} = \frac{\sigma_u}{(s+1)(s+1-\xi)}, \text{ for } \xi < 1 \tag{3.35}$$

Also, the empirical estimate of the moment $M_{1,0,s}$ is given by:

$$\hat{M}_{1,0,s} = \frac{1}{p}\sum_{i=1}^{p}\left(\prod_{j=1}^{s}\frac{p-i-j+1}{p-j}\right)X_{i:p} \tag{3.36}$$

where $X_{i:p}$ is the i^{th} order statistic of the p number of exceedances over a fixed threshold.

Therefore replacing $M_{1,0,s}$ with its estimator $\hat{M}_{1,0,s}$ and solving for $s = 0$ and $s = 1$ with respect to the parameters ξ and σ_u respectively, yields the PWM estimator for the shape and scale parameter as:

$$\left.\begin{aligned}\hat{\xi} &= 2 - \frac{\hat{M}_{1,0,0}}{\hat{M}_{1,0,0}-2\hat{M}_{1,0,1}} \\ \hat{\sigma_u} &= \frac{2\hat{M}_{1,0,0}\hat{M}_{1,0,1}}{\hat{M}_{1,0,0}-2\hat{M}_{1,0,1}}\end{aligned}\right\} \tag{3.37}$$

Hosking et al. (1985) established that for small samples sizes, the PWME is found to have smaller variance and hence performs better than the MLE. However, Sérgio (2012) indicated that for $\xi \geq 1$, the PWME do not exist, and in some cases, the observations may fall beyond the estimated right endpoint x^F.

3.3.4 Estimation of Extreme Quantiles

Similar to the BBM, useful indicators such as quantile estimates for extreme value analysis can also be obtained under the POTM by inversion of the GPD function. In other words, by inverting the GPD function in (32) and substituting for ξ and σ_u with their respective ML of PWM estimates, the $(1-\alpha)$ quantile of exceedances over a fixed threshold can be estimated as:

$$y_{1-\alpha} = H^{-1}_{\sigma_u,\xi}(1-\alpha) = \begin{cases} \frac{\sigma_u}{\xi}(\alpha^{-\xi}-1), & \text{if } \xi \neq 0 \\ -\sigma_u \ln \alpha, & \text{if } \xi = 0. \end{cases} \tag{3.38}$$

where $0 \leq \alpha \leq 1$. Also, if $\xi < 0$, the associated GPD model has a finite upper endpoint whose estimate is given as:

$$\hat{x}^F = u - \frac{\hat{\sigma}_u}{\hat{\xi}}. \tag{3.39}$$

Also, the quantile (or tail) estimators can be expressed in terms of their original random variable X with the unknown distribution F using relations in (26) and (27). Thus, given that $F(x) = F_u(y)[1-F(u)] + F(u)$ and $F_u(y)$ is asymptotic to the GPD, $H_{\sigma_u,\xi}(y)$, for sufficiently large u, where $y = x - u$ for $x > u$. Then

$$F(x) = H_{\sigma_u,\xi}(x-u)[1-F(u)] + F(u).$$

Replacing $F(u)$ with its empirical estimator, $\frac{n-e_u}{n}$ Smith (1985), where n is the sample size and e_u is the number of exceedances over the threshold u, we can write the following tail

estimator:

$$\begin{aligned}
\hat{F}(x) &= H_{\sigma_u,\xi}(x-u)[1-\tfrac{n-e_u}{n}]+\tfrac{n-e_u}{n} \\
&= \left\{1-\left(1+\tfrac{\xi(x-u)}{\sigma_u}\right)^{-\frac{1}{\xi}}\right\}\left[\tfrac{e_u}{n}\right]+\tfrac{n}{n}-\tfrac{e_u}{n} \\
&= \tfrac{e_u}{n}\left[1-\left(1+\tfrac{\xi(x-u)}{\sigma_u}\right)^{-\frac{1}{\xi}}-1\right]+1 \\
\hat{F}(x) &= 1-\tfrac{u}{n}\left(1+\tfrac{\xi(x-u)}{\sigma_u}\right)^{-\frac{1}{\xi}}
\end{aligned} \tag{3.40}$$

where ξ and σ_u is substituted with their respective ML or PWM estimates. Inverting the last line of (40) and solving for a given probability α yields the $(1-\alpha)$ extreme quantile estimator as:

$$F^{-1}(\alpha) = \begin{cases} u+\frac{\sigma_u}{\xi}[(\frac{n(1-\alpha)}{e_u})^{-\xi}-1], & \text{if } \xi \neq 0 \\ u+\sigma_u \ln(\frac{n(1-\alpha)}{e_u}), & \text{if } \xi = 0 \end{cases} \tag{3.41}$$

where ξ and σ_u is substituted with their respective MLE or PWME. Thereafter, measures such as Value-at-Risk and Expected Shortfall can be computed.

3.4 Value-at-Risk (VaR) and Expected Shortfall (ES)

One of the most frequent concerns of risk assessment in finance and insurance is the indicative values of extreme quantiles such as Value-at-Risk (VaR) and Expected Shortfall (ES). Most regulators in today's financial and insurance systems have come to accept these measures as the benchmark tool for market risk measurement and as a basis of capital requirements for risk exposure. Basically, VaR is a high quantile (usually a 99th or 99.9th percentile) of the distribution of returns that offers an estimated boundary which we cannot exceed within a specified period, for a given level of confidence (Uppal, 2013; Allen et al., 2011). Though the BMM can be employed in its estimation, the POTM, through the GPD, provides a more suitable means for estimating these high quantiles .

In terms of the GPD parameters, we thus define VaR_α as the α^{th} extreme quantile of the distribution which is estimated as:

$$\hat{VaR}_\alpha = u + \frac{\hat{\sigma}_u}{\hat{\xi}}\left[\left(\frac{n(1-\alpha)}{e_u}\right)^{-\hat{\xi}} - 1\right], \tag{3.42}$$

where $0 < \alpha < 1,\ x \geq u,\ \sigma_u > 0,\ \xi \in \Re$ and $\hat{F}(u) = \frac{n-e_u}{n}$.

Though the VaR has the advantage of easy interpretation and allows direct comparison of risk in a diverse portfolio, it ignores the potential size of the return in case the calculated VaR is exceeded at given probability (Szubzda and Chlebus, 2020). This brings to light the Expected Shortfall.

The Expected Shortfall (also called the Conditional Value-at-Risk (CVaR)) is a measure that takes into consideration the potential sizes of the returns expected to exceed a given VaR. Stated differently, it is the average value of the returns above the VaR. Thus,

$$ES_\alpha = \frac{1}{\alpha}\int_\alpha^1 VaR_\alpha. \tag{3.43}$$

According to Nortey et al. (2015) the Expected Shortfall can also be defined as:

$$\hat{ES}_\alpha = \hat{VaR}_\alpha + E[X - \hat{VaR}_\alpha | X > \hat{VaR}_\alpha]. \tag{3.44}$$

But the expression $E[X - \hat{VaR}_\alpha | X > \hat{VaR}_\alpha]$ is simply the expected value of excesses over the threshold level of VaR_α, which is similar to the conditional mean excess function. Then by definition of (30), we can write such expression as:

$$E[X - \hat{VaR}_\alpha | X > \hat{VaR}_\alpha] = e(VaR_\alpha) = \frac{\hat{\sigma_u} + \hat{\xi}(\hat{VaR}_\alpha - u)}{1 - \hat{\xi}} \tag{3.45}$$

Substituting (45) into (44), we have:

$$\hat{ES}_\alpha = \hat{VaR}_\alpha + \frac{\hat{\sigma_u} + \hat{\xi}(\hat{VaR}_\alpha - u)}{1 - \hat{\xi}} = \frac{\hat{VaR}_\alpha}{1 - \hat{\xi}} + \frac{\hat{\sigma_u} - \hat{\xi}u}{1 - \hat{\xi}} \tag{3.46}$$

3.5 Block Maxima Method (BMM) Vs. Peaks-Over-Threshold Method (POTM)

Modelling normalised sample maxima with only BMM is a wasteful approach for analysing extreme value behaviour if all observations are available in the data. Thus, it is statistically more efficient to disregard the BMM and instead, use the Peaks-Over-Threshold Method (POTM) if an entire time series of observations is accessible to the practitioner. Coles (2001) explained the wastefulness as this - given a model that describes variations in maxima from one period to another, realisations that are lower in one block but more extreme than the maxima of other blocks are excluded from the analysis. He added that rather than artificially blocking the data, say annually, and extracting the annual maximum from each block, it is more efficient to define a realisation as being extreme if it falls above some predetermined high level. Consequently, efficiency is improved since all information about the observed values, in terms of exceeding a high fixed threshold, can be used in model fitting.

Nonetheless, Diebolt et al. (2008) stated that the BMM may still be preferred for the following reasons. Foremost, it is computationally easier to only focus on block maximums especially when the practitioner is often faced with very high number of time series data to analyse, as with an hourly climatological data. Secondly, working with blocks of a given size, say one year, makes interpretability of the estimated parameters to the environmentalist more appreciable due to natural (or seasonal) blocking of environmental data. Lastly, the block maxima may be the only measurements available to the practitioner at the point in time when faced with incomplete or missing data. Stated differently, because the approach

considers only the maximum of each block, there is little or no need for the rest of the observation below the maxima. For these reasons among others, modelling block maxima with a GEVD remains a very frequent procedure in hydrology and climatology.

Chapter 4

DATA ANALYSIS AND FINDINGS

The data was obtained from the records of the private health insurance unit of the 37 Military Hospital and comprises daily claims that are submitted monthly to its partnered private health insurance companies - currently 13 in number. The data is from a secondary source and is made up of 32,964 claims spanning a period of 8 years (2012 - 2019). The source of the data also recorded "no-submission" in some months, and "multiple-submissions" - claims from preceding and current months - in other months. As a result, there are 90 (instead of 96) monthly recorded submissions for the 8-year period.

The variables contained in the data include "Diagnosis", "Amount of Medical Services", "Amount of Drugs" and "Total Amount". However, the prime variable of interest is the "Amount of Medical Services" - all healthcare services other than drugs, provided to a private health insurance patient. The choice of variable of interest was influenced by the substantive missing data on the "Amount of Drugs" for the period 2018 - 2019. Consequently, the "Total Amount" for the entire period will be understated and hence becomes unattractive for use in analysis.

Furthermore, records from the hospital reveal that most new companies cover for only

medical services in the early stage of their partnership and later renew their contracts to cover drugs. This is in line with the overall aim of the study - to provide insights on the likely occurrence of large claim amounts particularly for prospective and new health insurance firms interested in partnering the 37 Military hospital. Therefore, for the purpose of this study, the word "claims" specifically refers to "Amount of Medical Services".

In this chapter, we run preliminary statistical analysis to describe the data, and then employ the parametric EVT approach to describe the behaviour of large claim amounts. Based on the proposed model and its parameter estimates, we forecast the likelihood of surpassing the maximum claim amount in the data and also project the probability of exceeding of certain large claim amounts at a particular period of time. The **R**-software was used for all statistical computations, with emphasis on "packages" such as *evd, extRemes, extremefit, fExtremes, ismev,* and *POT* for the core analysis.

4.1 Preliminary Analysis

Table 4.1.1: Summary Statistics

Claims	Min. Amt	Max. Amt	Std. Er	Mean	Skewness	Kurtosis
32963	5	15,740	3.06	215.5	7.79	100.7

Table 4.1.1 reveals that on a particular day, the hospital can claim a minimum amount of 5 cedis and a maximum amount of 15,740 cedis for services rendered to patients on the private health insurance basis. These amounts correspond to a minor wound dressing and a tumour surgery respectively. Furthermore, an insurance company can expect an average claim amount of 216 cedis from the hospital at any particular day. Also, the two "so-called" shape statistics - skewness and kurtosis - give signals of more extreme deviations from the centre of the distribution - an indication that the data has a long right-tailed distribution. However, these two measures are not sufficient enough to determine if the underlying distribution for

data is heavily tailed or light tailed. Below is a pictorial description of the data at hand. The results were obtained using the **R**-Software (see Appendix B2)

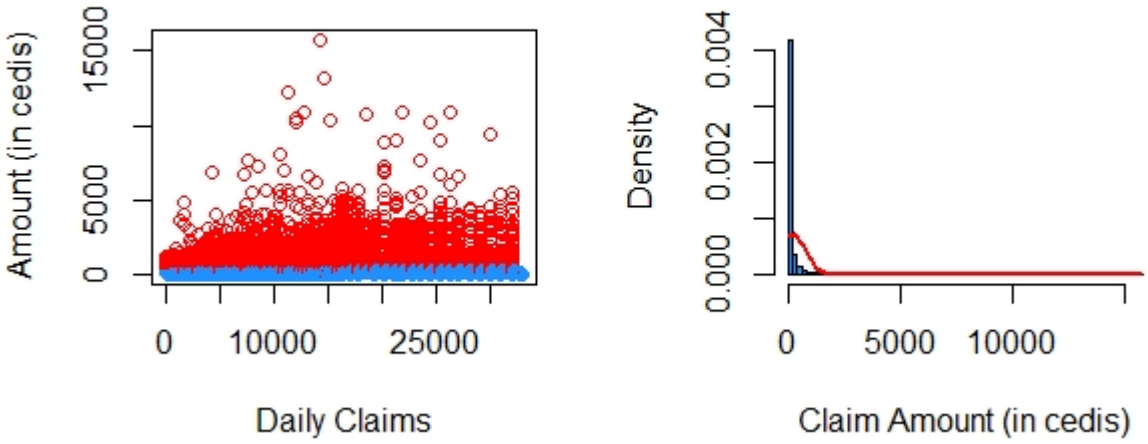

Figure 4.1.1: Scatter Plot and Histogram of Claims Submitted for the Period 2012 - 2019.

3

The scatterplot from Figure 4.1.1 exhibits no clear pattern or trend and hence gives an indication of non-associativity between the variables. Also, the plot shows that more of the claims amount are larger than the average claim size of 216 cedis. This is in concord with the long right-tailed distribution depicted by the histogram.

In order to make fair projections from any time series data, we require that the data be stationary. Here, we employ the "Augmented Dickey-Fuller (ADF)" test by setting H_0: the data is non-stationary against H_1: the data is stationary, at a significance level of 0.05 (Nkrumah, 2017). The results were obtained using the **R**-Software (see Appendix B3).

Table 4.1.2: ADF Stationarity Test for Claim Amounts

Variable	Test Statistic	P-value
Claims	-22	0.01

From the test, the p-value (0.01) is less than the significance level (0.05). Therefore, we reject the null hypothesis and conclude that the data is stationary.

Another preliminary analysis needed to admit the use of Extreme Value Distribution (EVD) in this work is the Exponential QQ-Plot. Using the **R**-Software (see Appendix B4) we obtain the Exponential QQ-plot for the sample maxima.

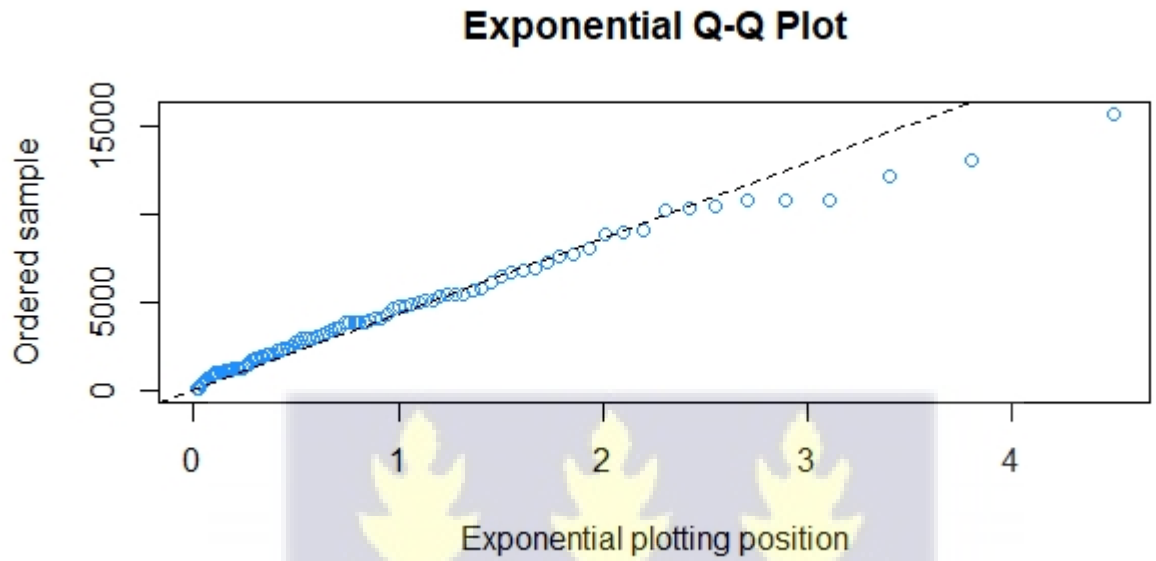

Figure 4.1.2: Exponential QQ-Plot of Claims

The Exponential QQ-plot in Figure 4.1.2 depicts a roughly linear pattern for lower and middle values of the sample but a convex curvature as we approach the upper values. This is an indication that the underlying limiting distribution for the data is light-tailed than expected from an Exponential distribution (Sérgio, 2012). Hence, we consider the plausibility of the extreme value distributions as a limiting distribution the data. With the help of the EVT framework, In this regard, we employ the "GEVD" and the "GPD" fits respectively for the Block Maxima Method and Peaks-Over-Threshold Method of EVT.

4.2 The Block Maxima Method (BMM)

In employing the Block Maxima approach, we split the data into monthly blocks and obtain the maximum from each block. The choice of block length is based on the fact that the hospital submits its claims monthly; and more importantly, we obtain a larger sample size (89 observations) for analysis than when the data is blocked yearly (i.e. 8 observations). Using the MLE and the PWME we approximate the parameters for the appropriate limiting distribution and then use it to estimate important extreme quantiles such as return levels, Value-at-Risk (VaR) and Expected Shortfall (ES) with their corresponding exceedance probabilities.

In pursuance of the parametric approach of the EVT under the BMM, we fit the GEVD as a limiting distribution to the monthly claims maxima under the BMM. We also obtain both the MLE and the PWME for the GEVD parameters - location (μ), scale (σ) and shape (ξ). The **R**-Software (see Appendix A1 and B5) was used to obtain the results .

Captured in Appendix A1, the standard errors for the estimated GEVD parameters are unreasonably large - an indication of a less precise model. Taking a "square-root" transformation of the underlying variable results in a more precise model. It is worth noting that though the "log-transformation" of the underlying variable gives much smaller standard errors for the parameter estimates (see Appendix A2), its curvilinear nature inhibits accurate results for the data at hand. This can be seen in the diagnostic plots captured in Appendix A3. Thus, both the probability plot and quantile plot under the "log-transformation" is less linear as compared to that of the "square-root transformation".

The figure below gives the diagnostic plots for fitting the GEV distribution to the sample maxima after the "square-root transformation".

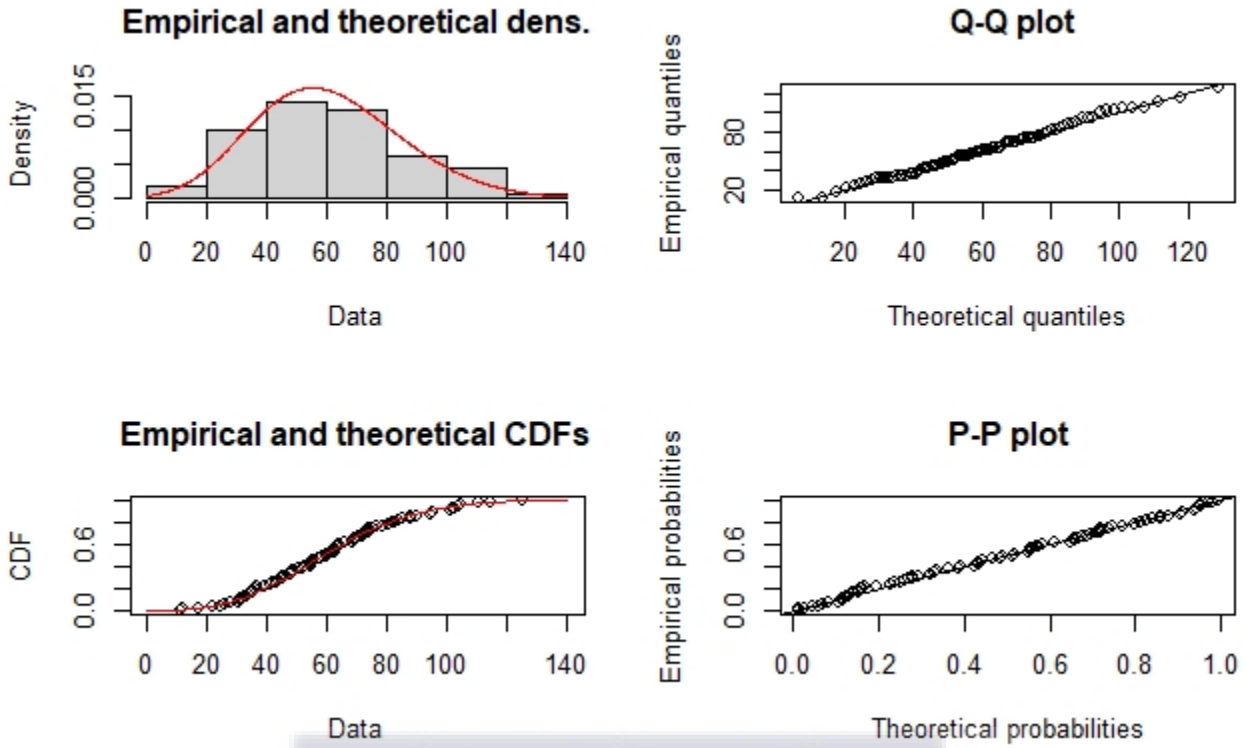

Figure 4.2.1: Diagnostic Plots for the Proposed GEV Model

From Figure 4.2.1, both the probability plot and the quantile plot exhibits a very strong linear pattern between the quantiles of the sample and that of the theoretical GEVD - a confirmation that the proposed GEV model is a better fit to the data as compared to the gumbel model. Furthermore, both the theoretical density function and the theoretical CDF of the "GEV-fit" superimposed on the observed sample maxima affirms that the model is appropriate for the data at hand. These results indicate that the limiting distribution for the sample maxima belongs to the GEV families of distributions.

4.2.1 Estimation of Parameters

Below is the table for the ML and PWM estimates, and the 95% Profile Likelihood-based Confidence Interval (PL-CI) - a more accurate measure than the normal approximation

confidence interval in terms of allowing for skewed intervals for the estimates (Coles, 2001). The **R**-Software (see Appendix B8 and B9) was used to obtain the results below.

Table 4.2.1: Parameter Estimates For The GEV Model

Parameter	MLE	PWME	95% PL-CI
Location (μ)	50.69	50.28	[45.29, 56.23]
Scale (σ)	23.17	23.56	[19.74, 27.75]
Shape (ξ)	-0.18	-0.16	[-0.32, -0.0004]

Table 4.2.1 shows a very slight difference between the parameter estimates under the MLE and the PWME. Furthermore, the table suggests that the distribution of the sample maxima is attracted to the Weibull class of distributions because the shape parameter in both estimations is less than zero.

However, a careful examination of Table 3 indicates a "near-zero" value for the upper bound of the PL-CI for the shape parameter. This can imply a plausibility of the data being attracted to the Gumbel class of distributions. Using the same data at hand, we can perform various statistical test to choose the appropriate GEV model.

Having two families of distributions as a plausible fit to the data requires more objective tests in order to select the model that best fits the data. We therefore test for the validity of the gumbel model, just as we did the GEV model using the same data at hand.

We first run the diagnostic plots for the "Gumbel-fit". As shown in Appendix A4, an "eye-ball" inspection of the diagnostic plots gives an indication that the proposed GEV model (belonging to the Weibull class of distributions) is superior to the "Gumble-class" model for the data at hand. This is because linearity in both the probability and quantile plots are more achieved in the "Weibull-class" model than the Gumbel model.

Employing the gumbel statistic, we also perform the following hypothesis test by setting H_0: The "Gumbel-class" model is fit for the data *vs.* H_1: The "Gumbel-class" model is not fit for the data. We obtain the following results with the **R**-software (see Appendix B10)

Table 4.2.2: Test for the Validity of the Gumbel Model

Test Statistic (GS_m)	Critical Val.(GS_m^*)	P-value
0.6116	-2.395	0.00001724

From the above results, we reject the null hypothesis at 0.05 significance level and conclude that the gumbel model is not fit for the data.

Last but not least, we can also compare the negative log-likelihood, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) of the two models and choose the model with the smallest score (Pinheiro and Grotjahn, 2015; Gilleland and Katz, 2016. The results were obtained using the **R**-Software (see Appendix B11 and B11).

Table 4.2.3: Model Comparison of the "Gumbel-class" vs. The "Weibull-class"

	Gumbel-Class	Weibull-Class
Negative log-likelihood	414	412
AIC score	831	829
BIC score	836	836

Regardless of the minute differences in the scores of the two model, the above table establishes that the GEV model (i.e. Weibull class of distributions) is more suitable for the data at hand because it has the lower score in general. It is worth noting that we cannot completely discard the plausibility of the "gumbel-class" model to the data since we only used a few of the many parametric test available for a situation like this. This is an entirely different area that can be researched into at a later time.

4.2.2 Estimation of Extreme Quantiles

Having decided on the proposed model (i.e. the GEV model) and its parameter estimates, we turn to approximate for some selected return levels with their corresponding return periods and exceedance probabilities. We also determine the likelihood of surpassing the highest daily claim amount and also estimate the upper endpoint (since the shape parameter is less than zero) for the given data. The **R**-Software (see Appendix B12 and B13) was used to obtained the results . For the PWME, see Appendix A5 for the results. We remark that for reporting purposes, the resulting values from the table must be squared due to the initial square-root transformation of the underlying variable.

Table 4.2.4: Return Period, Return Level and Exceedance Probability Estimates

Return Period	Return Level	Exceedance Probability
5 Months	81.16	0.1999
10 Months	93.60	0.09977
20 Months	104.05	0.04981
50 Months	115.73	0.01985
100 Months	154.76	0.0001029
	Value	Exceedance Probability
Maximum Claim	125.46	0.007948
Upper Endpoint (X^F)	179.6	0

In general, both the MLE and PWME yielded similar results for the data at hand. Furthermore, given the respective probabilities, the hospital on average is likely to submit a claim amount exceeding 6,587 cedis, 8,761 cedis, 10,826 cedis, 13,393 cedis, and 23,951 cedis once every 5, 10, 20, 50, and 100 months respectively. Also, Table 6 establishes that for any given day, there is a very low chance for the hospital to submit a claim that exceeds the maximum claim (i.e. 15,740 cedis) contained in the data. However, it is not likely for the hospital to submit a claim exceeding 32,256 cedis for any given day.

We now consider other equally important tools of interest for this study - Value-at-Risk

(VaR) and Expected Shortfall (ES). These are popular measures trusted by most insurance companies in devising policies to mitigate losses resulting from the occurrence of very large claims. Given that these tools are basically the extreme quantiles of a distribution, the EVT provides a suitable technique for their estimation.

Here, we assess the 99%, 99.5% and 99.9% 1-day-VaR and its corresponding ES since that the underlying data is made up of daily claims. The results were obtained using the **R**-Software (see Appendix B14). We take note that the reporting values from the table must be squared due to the initial square-root transformation of the underlying variable.

Table 4.2.5: Loss Mitigating Tools: 1-day-VaR Estimates

Probability	Value-at-Risk (VaR)
0.990	123.172
0.995	129.791
0.999	142.285

Observing from table 4.2.5, we are assured at a confidence level of 99%, 99.5% and 99.9% that in the next 24 hours, the hospital is not likely to submit a claim amount exceeding 15,178 cedis, 16,848 cedis and 20,249 cedis respectively. However, given the chances that these thresholds are exceeded in a day, the expected claim amounts that might occur beyond these breakpoints are 4,247 cedis, 5,345 cedis and 8, 012 cedis respectively. It is also imperative to know that the class of disease associated with these extreme quantiles include fibrosarcoma, open fractures, severe head injuries, pelvic and liver abscess, and deep lacerations.

4.3 Peaks-Over-Threshold Method (POTM)

In employing the Peaks-Over-Threshold approach, we consider all available observations in the data and assume an unknown distribution F. We then determine a sufficiently

high threshold and concentrate on those realisations that exceeds this designated threshold. Fitting a limiting parametric model (in this case the Generalized Pareto Distribution) to these exceedances, we estimate the model's parameters and approximate for some return levels and other extreme quantiles of interest - the Value-at-Risk and Expected Shortfall. Likewise under the BMM, we also undertake a square-root transformation of the underlying variables, primordially to reduce the error margin involved when estimating the model's parameters.

4.3.1 Threshold Selection

We recall under chapter 3 that the choice of threshold, u, is not always straight forward - a 'trade-off' between bias and variance of the model parameters. Thus, the higher the threshold, the fewer observations left for analysis, the better the bias, but the higher the variability in the parameter estimates. The converse is also true. We also recall some graphical tools - the Mean Excess Plot (MEP), also known as the Mean Residual Life (MRL) Plot, and the Parameter Stability Plot (PSP) - that can help us in deciding an optimum threshold.

We established that a good "compromise" of u is usually the point to the right where an approximately linear pattern is observed in the MEP; or the point at which the asymptotic properties for the GP model holds and the parameter estimates are approximately stable as we move towards higher thresholds (Coles, 2001). We also recall that an upward pattern of the MEP indicates a heavy-tailed distribution function (in this case "Pareto Type I"); a downward trend indicates a light-tailed distribution function (in this case "Pareto Type II"); whereas a fairly horizontal pattern indicates an exponential tail distribution function (Gilleland and Katz, 2016; Coles, 2001). The MEP below was obtained using the **R**-Software (see Appendix B15).

Figure 4.3.1: Mean Excess Plot of Claims Submitted, with Gray Lines Indicating 95% Confidence Intervals for the Mean Excess Values

Pictured above, the mean excess value increases gradually for the lower order statistics, followed by a decrease as we approach the middle-order statistics. It then increases again, and eventually falls as we move towards the top-order statistics. This is an indication that a light-tailed distribution with a Weibull domain of attraction (i.e. "Pareto Type II") is suitable for the data at hand (Coles, 2001; Beirlant et al., 2006).

The main goal in Figure 4.3 is to help select an appropriate threshold such that the plot becomes approximately linear as we approach the top-order statistics (Coles, 2001; Beirlant et al., 2006). Observing from the plot, we can trace a reasonably linear pattern from points 20 to 30. Outside these ranges, the plot keeps fluctuating and hence becomes difficult to choose any point as a suitable threshold.

In working with the preceding argument, we explore the Parameter Stability Plot respectively for the scale and shape parameters over the whole sample values. However, we 'cut off' a margin towards both the largest and the smallest order statistics. This is because within those regions, any threshold selected will not result in a good "compromise" - either a very high bias or a very high variance in the parameter estimates. We then choose the threshold value where these plots exhibit reasonable stability. Below is stability plot for the shape and scale parameters respectively. The results were obtained using the **R**-Software (see Appendix B16 and B17).

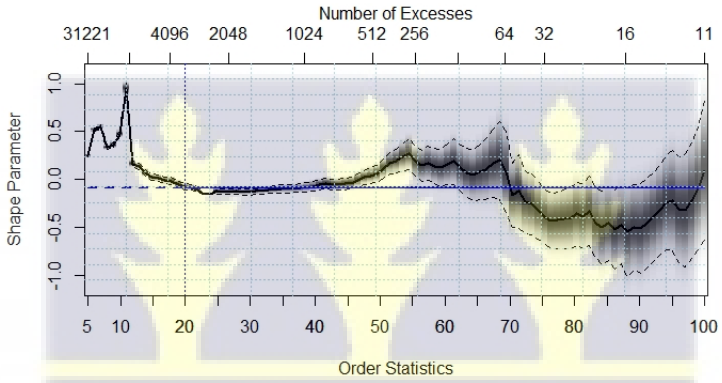

Figure 4.3.2: Shape Parameter Stability Plot

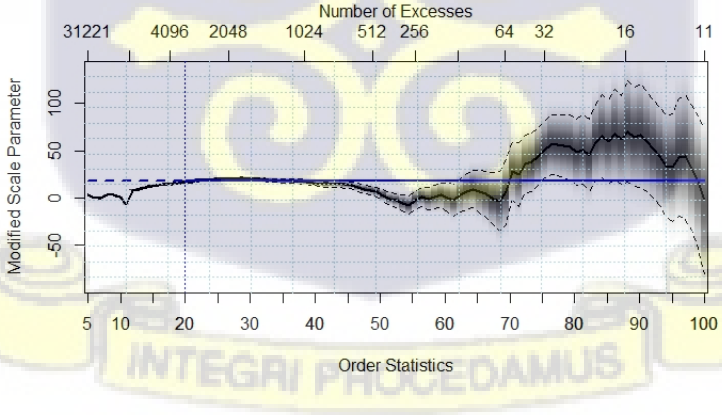

Figure 4.3.3: Scale Parameter Stability Plot

Both Figure 4.3.2 and 4.3.3 exhibit reasonably stable parameters at a threshold range of 20 to 35. Though possible, it will be very difficult to select a suitable threshold outside this range since the plots will be fluctuating. We remark from chapter 3 that since the scale parameter for the GPD is a function of the threshold, a modified scale parameter is rather plotted so that it becomes independent of the threshold. Thus, for the scale parameter, we use $\sigma^* = \sigma(u) - \xi u$ where σ^* is the modified scale parameter, ξ is the shape parameter and u is the threshold.

We also take note that the 10^{th} largest claim (i.e. 101.4) and beyond are dropped since any threshold within that region will result in a very large variance in the model's parameter estimates due to very few exceedances left for analysis. Similarly, the 32953^{th} largest claim (i.e. 4.0) and below are also dropped since any threshold within that region will result in a very high bias in the model's parameter estimates due to the inclusion of less extreme exceedances in the analysis. Thus, we make sure that a balance is kept such as to limit both bias and variance from getting too large.

With the suitability of the threshold range narrowed down, we can move further to specify the most appropriate threshold by zooming into the range and exploring the parameter stability plot one more time.

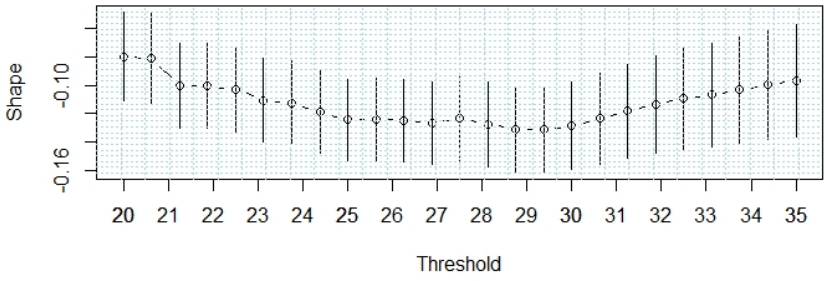

Figure 4.3.4: Shape Parameter Stability Plot (zoomed)

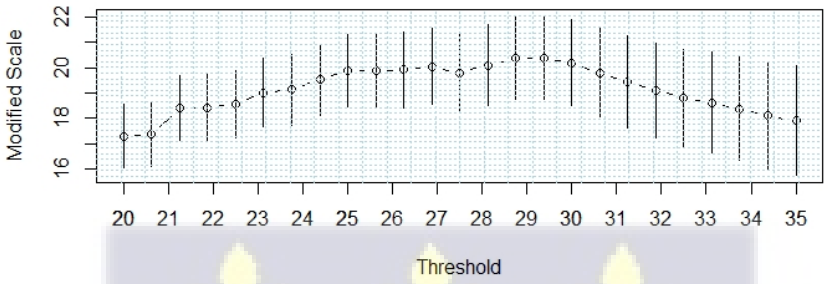

Figure 4.3.5: Scale Parameter Stability Plot (zoomed)

A careful look at Figure 4.3.4 and 4.3.5, shows a more stable parameter at the threshold range of 25 to 30. We subsequently fit the GPD at these potential thresholds and estimate their respective parameters.

4.3.2 Estimation of Parameters

Employing the MLE and the PWME, we approximate the model's parameters using the exceedances of the prospective thresholds. The results are tabulated below in 4.3.1 to control sequence. The **R**-Software (see Appendix B18) was used to obtain the results. We remark that, due to baseline approximations between packages, the PWM estimates have slightly different results. Moreover, of the different packages available, we place more emphasis on

the package "fExtremes" - a group of written functions for analysing and modelling extreme events especially in Finance and Insurance (Wuertz et al., 2009).

Table 4.3.1: Estimates of Model Parameters and Upper Endpoints for the Prospective Thresholds

Threshold	Exceedances	Shape (Std. Er.)	Scale (Std. Er.)	X^F
25	2272	-0.124 (0.014)	16.778 (0.425)	160.307
26	2130	-0.175 (0.015)	19.353 (0.435)	136.589
27	2012	-0.174 (0.015)	19.145 (0.443)	137.029
28	1893	-0.177 (0.015)	19.080 (0.452)	135.797
29	1754	-0.134 (0.015)	16.726 (0.467)	153.821
30	1667	-0.129 (0.016)	16.371 (0.472)	156.907

A careful look at Table 4.3.1 shows that the standard error increases as the threshold increases. This is due to the decrease in the number of exceedances as the threshold increases and hence the so-called 'trade-off' associated with threshold selection. Furthermore, the table confirms the proposition that the distribution of the data is asymptotic to the Weibull class of distributions (i.e. "Pareto Type II"), as previously expressed by the MEP. Hence, we can estimate the upper endpoint (X^F) of the data using the respective parameter estimates as depicted in the table.

The choice of GPD model is based on the choice of threshold. Having two or more models as a plausible fit to the data requires further analysis in order to select the model that 'best' fits the data. Moreover, the choice of model can be entirely subjective to the interest of the researcher. Thus, while a low threshold might produce a theoretically more accurate model, we might still not choose it if the interest lies in the behaviour of claims above some higher amount.

Likewise under the BMM, we run the diagnostic plots for the 'GPD-fit' at the various threshold and also compute the negative log-likelihood, the AIC and the BIC scores for

the corresponding models in order to select the most appropriate model for the data at hand. The diagnostic plots for the GPD-fit at the various thresholds are shown in Appendix A7 - A12, and were obtained using the **R**-Software (see Appendix B20).

Comparing the diagnostic plots for the various GPD models, the GPD-fit at a threshold of 30 is adjudged the the most suitable. This is because the plot reveals a stronger linear pattern between the quantiles of the exceedances and that of the theoretical GPD in both the probability plot and the quantile plot and hence making it a better fit for the data. Below is the diagnostic plot for fitting the GPD at threhold of 30.

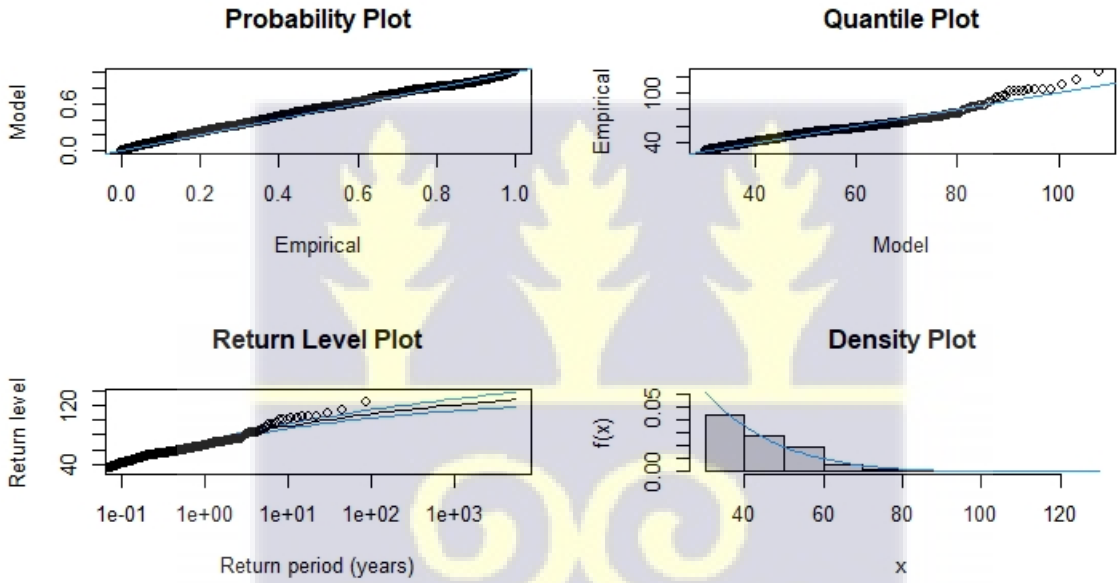

Figure 4.3.6: Diagnostic Plots for the GPD-fit at Threshold of 30

Also, using the **R**-Software (see Appendix B18), we obtain the negative log-likelihood, AIC and BIC scores for the potential GPD models as shown in the table below.

Table 4.3.2: Negative Log-likelihood, AIC and BIC scores for the Potential GPD Models

Threshold	Negative Log-likelihood	AIC	BIC
25	8397.4	16798.9	16810.3
26	7865.7	15735.4	15746.7
27	7409.3	14822.7	14833.9
28	6957.0	13918.1	13929.2
29	6456.7	12917.4	12928.3
30	6108.9	12221.9	12232.8

From Table 4.3.2, we realise that given the available data, the negative log-likelihood, the AIC and the BIC scores decreases for the models with higher thresholds. Also, the GPD model at the threshold of 30 produces the lowest scores and hence retains more information when used for the available data. Therefore, the proposed GPD model for this study has a threshold of 30 and has shape and scale parameter estimates of -0.129 and 16.371 respectively, and an upper-bound estimate of 156.907.

4.3.3 Estimation of Extreme Quantiles

With the parameter estimates at hand, we approximate for the return levels of some specified periods with their corresponding exceedance probabilities for the available data. We also determine the likelihood of exceeding the maximum claim amount of the entire data. The **R**-software (see Appendix B21, and B22) was used to obtain the results. See Appendix A8 for the results when PWME is used. We remark that for the purpose of reporting, the resulting values must be squared due to the initial square-root transformation of the underlying variable.

Table 4.3.3: Return Period, Return Level and Exceedance Probability Estimates

Return Period	Return Level	Exceedance Probability
5 Months	53.746	0.037
10 Months	62.553	0.022
20 Months	70.610	0.013
50 Months	80.217	0.007
100 Months	86.768	0.005
	Value	Exceedance Probability
Maximum Claim	125.46	0.0005
Upper Endpoint (X^F)	156.9	0

From Table 4.3.3, given the respective probabilities, the hospital on average is likely to submit a claim amount exceeding 2,889 cedis, 3,913 cedis, 4,986 cedis, 6,435 cedis, and 7,529 cedis once every 5, 10, 20, 50, and 100 months respectively, for medical services provided to clients of its partnered private health insurance companies. Also, Table 10 establishes that there is a very low chance for the hospital to submit a claim that exceeds the maximum amount (i.e. 15,740 cedis) contained in the data, for any given day. However, it is not likely for the hospital to submit a claim exceeding 32,256 cedis for any given day.

The next tools of interest in our estimation is the Value-at-Risk (VaR) and Expected Shortfall (ES). Compared to the BMM, the POTM offers a more suitable means for estimating the VaR and the ES (Allen et al. 2011). Using the GPD model, we assess the 99%, 99.5% and 99.9% 1-day-VaR and its corresponding ES using the available data. The results were obtained using the **R**-software. Again, we remark that for the purpose of reporting, the obtained values must be squared due to the initial square-root transformation of the underlying variable.

Table 4.3.4: Loss Mitigating Tools: 1-day-VaR and 1-day-ES Estimates

Probability	Value-at-Risk (VaR)	Expected Shortfall (ES)
0.990	53.944	65.707
0.995	62.750	73.508
0.999	80.401	89.141

Observing from Table 4.3.4, we are assured at a confidence level of 99%, 99.5% and 99.9% that in the next 24 hours, the hospital is not likely to submit a claim amount exceeding 2,910 cedis, 3,938 cedis and 7,946 cedis respectively. However, given the chances that these thresholds are exceeded in a day, the expected claim amounts that might occur beyond these breakpoints are 4,317 cedis, 5,403 cedis and 7,946 cedis respectively. It is also imperative to know that the class of disease associated with these extreme quantiles include fibrosarcoma, open fractures, severe head injuries, pelvic and liver abscess, and deep lacerations.

Chapter 5

CONCLUSIONS AND REMARKS

We have shown by use of Extreme Value Theory that the monthly submitted claims by the 37 Military hospital can be modelled by the Weibull class of distributions. This implies that partnered health insurance firms can be assured of claim amounts not exceeding a certain amount. In particular, it is not likely for the 37 Military hospital to submit claim amount exceeding 24,618 cedis for any given day. In addition, private health insurance firms can be assured at a confidence level of 99%, 99.5% and 99.9% that within a day, the hospital is not likely to submit a claim amount exceeding 2,910 cedis, 3,938 cedis and 7,946 cedis respectively.

Also, given the data at hand, the ML estimator and the PWM estimator yielded almost similar results. Moreover, since the MLE performs better than the PWME for large sample sizes (Coles and Dixon, 1999; Hosking et al., 1985; Sérgio, 2012), the Generalised Probability Weighted Moments (GPWM) can be employed in place of the PWME to yield more accurate results.

5.1 RECOMMENDATIONS

The EVT approaches used in the study are purely parametric and hence becomes insufficient when some of the requisite assumptions are not met. In this regard, we recommend a look into the semi-parametric approach of EVT given the available data.

Furthermore, VaR and ES are two simple but yet powerful tools in risk management. Their estimates should not be taken upfront but should be viewed as an idea of how much risk is involved in a financial instrument. Furthermore, considering their estimation under the paramteric approach of EVT, the BMM gives a more conservative estimates than POTM.

In the spirit of improving the operations between the 37 Military Hospital and its partnered health insurance companies, a further look into this study is recommended for the Statistical Units of these institutions in order for management to make more meaningful and appropriate decisions on the treatment tariffs of certain classes of diseases that are likely to have rarely huge cost of services.

Finally, it will be very beneficial to the National Health Insurance Authority, if this study could be replicated for for claims received by the public health insurance scheme (i.e. NHIS), since it can go a long way to strengthen the financial sustainability of the scheme.

Bibliography

Adesina, O. S., Adeleke, I., and Oladeji, T. F. (2016). Using extreme value theory to model insurance risk of nigeria's motor industrial class of business. *The Journal of Risk Management and Insurance*, 20(1):40–51.

Akinola, S. and Dessislava, D. (2019). 7 ways the private sector can contribute to universal health coverage — world economic forum. `https://www.weforum.org/agenda/2019/09/7-ways-the-private-sector-can-contribute-to-universal-health-coverage/`.

Allen, D. E., Singh, A. K., and Powell, R. J. (2011). *Extreme market risk-an extreme value theory approach*. The Econometric Society.

Arhinful, D. K. (2003). *The solidarity of self-interest: Social and cultural feasibility of rural health insurance in Ghana*, volume 71. African Studies Centre, Leiden.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L. (2006). *Statistics of extremes: theory and applications*. John Wiley & Sons.

Boadu, F., Dwomo-Fokuo, E., Boakye, J. K., and Frimpong, A. O. (2014). Assessing the life insurance industry in ghana. *European Journal of Business and Management*, 6(21):14–23.

Codjoe, S. N. A. and Nabie, V. A. (2014). Climate change and cerebrospinal meningitis in the ghanaian meningitis belt. *International journal of environmental research and public health*, 11(7):6923–6939.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer-Verlag.

Coles, S. G. and Dixon, M. J. (1999). Likelihood-based inference for extreme value models. *Extremes*, 2(1):5–23.

Danielsson, J. and De Vries, C. G. (2000). Value-at-risk and extreme returns. *Annales d'Economie et de Statistique*.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3):393–425.

De Wolf, A. H. and Toebes, B. (2016). Assessing private sector involvement in health care and universal health coverage in light of the right to health. *Health and Human Rights*, 18(2):79.

Diebolt, J., Guillou, A., Naveau, P., and Ribereau, P. (2008). Improving probability-weighted moment methods for the generalized extreme value distribution. *REVSTAT-Statistical Journal*, 6(1):33–50.

Donkor, B. K. (2014). Knust hospital suspends services to nhis subscribers. `https://www.graphic.com.gh/news/general-news/knust-hospital-suspends-services-to-nhis-subscribers.html`.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.

Gencay, R. and Selcuk, F. (2004). Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of forecasting*, 20(2):287–303.

Ghanaweb Network (2015). Private hospitals reject nhis cards. `https://www.ghanaweb.com/GhanaHomePage/NewsArchive/Private-hospitals-reject-NHIS-cards-373150`.

Gilleland, E. and Katz, R. W. (2016). extremes 2.0: An extreme value analysis package in r. *Journal of Statistical Software*, 72(8):1–39.

Gnedenko, B. (1943). On the limited distribution of the term of a random series. annals of mathematics 44, 423–53.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water resources research*, 15(5):1049–1054.

Health Facilities Regulatory Agency (2020). Licensed facilities. `http://hefra.gov.gh/index.php/licensed-facilities/`.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

Kinnison, R. R. (1985). *Applied extreme value statistics*. Battelle Press Columbus, OH.

Landwehr, J. M., Matalas, N., and Wallis, J. (1979). Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water resources research*, 15(5):1055–1064.

Massey, R. (2003). Why do insurance companies fail. `https://www.globalreinsurance.com/why-do-insurance-companies-fail/1312806.article`.

Ministry of Health (2016). Ghs standard hospitals. `http://www.moh.gov.gh/wp-content/uploads/2016/02/Standard-Hospital.pdf`.

Minkah, R. (2016). An application of extreme value theory to the management of a hydroelectric dam. *SpringerPlus*, 5(1):96.

Mises, R. v. (1936). The distribution of the largest of n values. *Rev. Math. Interbalcanic Union*, 1:141–160.

National Health Insurance Authority (2018). Guidelines for private health insurance schemes in ghana. `http://www.nhis.gov.gh/files/PHISREVISEDGUIDELINES.pdf`.

National Health Insurance Authority (2020). National health insurance act, 2012 (act 852). `http://nhis.gov.gh/files/ACT852.pdf`.

National Health Insurance Scheme (2020). Benefits package. `http://www.nhis.gov.gh/benefits.aspx`.

National Insurance Commission (2019). New minimum capital requirements (mcr) for insurance entities. `http://nicgh.org/news/press-release-new-minimum-capital-requirements-mcr-for-insurance-entities/`.

NHIS Review (2020). Nhis: Terms of reference for defining options for national health insurance scheme reforms. `http://www.nhis.gov.gh/nhisreview.aspx`.

Nkrumah, S. (2017). *Extreme Value Analysis of Temperature and Rainfall: Case Study of Some Selected Regions in Ghana.* Department of Statistics and Actuarial Science, University of Ghana.

Nortey, E. N., Asare, K., and Mettle, F. O. (2015). Extreme value modelling of ghana stock exchange index. *SpringerPlus*, 4(1):696.

Nortey, E. N. N., Doku-Amponsah, K., and Ocran, E. (2017). *Estimating Exceedance Probability of Extreme Water Levels of the Akosombo dam.* Department of Statistics and Actuarial Science. University of Ghana.

Owusu-Sekyere, E. and Chiaraah, A. (2014). Demand for health insurance in ghana: what factors influence enrollment?

Pérez-Fructuoso, M. J. and García Pérez, A. (2010). Analyzing solvency with extreme value theory: an application to the spanish motor liability insurance market. *Innovar*, 20(36):35–48.

Pinheiro, M. and Grotjahn, R. (2015). An introduction to extreme value statistics.

Reserve Bank of Australia (2018). The global financial crisis. `https://www.rba.gov.au/education/resources/explainers/the-global-financial-crisis.html`.

Sérgio, L. G. V. (2012). *Extreme Value Theory: An Application to Sports*. Department of Statistics and Operational Research, University of Lisbon.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.

Szubzda, F. and Chlebus, M. (2020). Comparison of block maxima and peaks over threshold value-at-risk models for market risk in various economic conditions. *Central European Economic Journal*, 6(53):70–85.

Tv3 Network (2015). Doctors' strike: No pressure on us 37 military hospital. `http://tv3network.com/health/doctors-strike-no-pressure-on-us-37-military-hospital`.

Uppal, J. Y. (2013). Measures of extreme loss risk–an assessment of performance during the global financial crisis. *Journal of Accounting and Finance*, 13(3):105–117.

Wagstaff, A. (2010). Social health insurance reexamined. *Health economics*, 19(5):503–517.

Wainaina, H. and Waititu, A. (2014). Modeling insurance returns with extreme value theory (a case study for kenya's fire industrial insurance class of business). *Mathematical theory and modeling*, 4.

Weru, S. k. and Waititu, A. (2019). Modelling of large insurance claims using extreme value theory: A case study of kenindia assurance company limited motor business. *International Journal of Mathematics and Physical Sciences Research*, 6(2):10–20.

World Health Organization (2019). Countries are spending more on health, but people are still paying too much out of their own pockets. `https://www.who.int/news-room/detail/20-02-2019-countries-are-spending-more-on-health-but-people-are-still-paying-too-much-out-of-their-own-pockets.`

Wuertz, D., Setz, T., and Chalabi, Y. (2009). Rmetrics - modelling extreme events in finance.

APPENDIX A

A1 Estimation of GEVD Parameters (without variable transformation)

Table 5.1.1: GEV Parameter Estimates For The Claim Maxima

Parameter	Estimate	Std. Error
Location (μ)	2680.45	280.47
Scale (σ)	2216.10	243.78
Shape (ξ)	0.22	0.13

A2 Comparison of the "Log" and "Square-root" Variable Transformations

Table 5.1.2: Parameter Estimates After Variable Transformation

Parameter	Square Root Transformation Estimate (Std. Error)	Log Transformation Estimate (Std. Error)
Location (μ)	50.69 (2.779)	7.79 (0.117)
Scale (σ)	23.17 (1.998)	1.04 (0.089)
Shape (ξ)	-0.18 (0.083)	-0.53 (0.057)

A3 Diagnostic Plots for the GEVD-fit After Log-Transformation

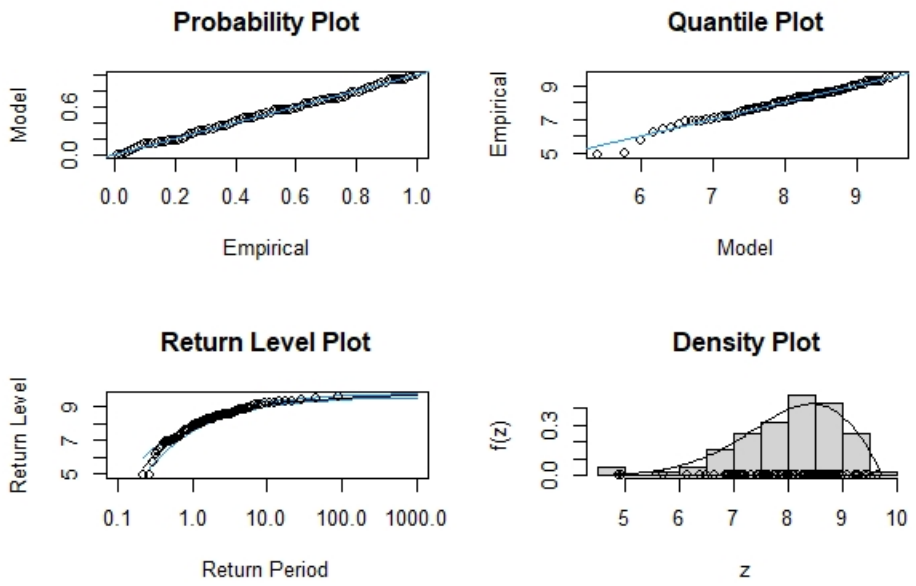

Figure 5.1.1: Diagnostic Plots for the GEVD-fit After Log-Transformation

A4 Diagnostic Plots for the Prospective Gumbel-fit

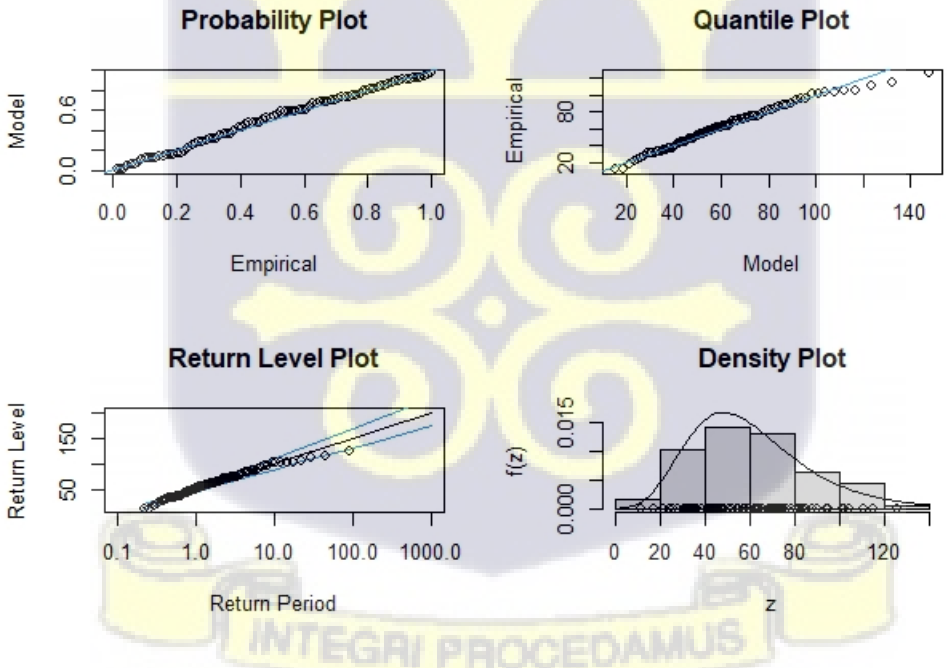

Figure 5.1.2: Diagnostic Plots for the Prospective Gumbel-fit

A5 **PWM Estimates of the GEVD Quantiles**

Table 5.1.3: Return Period, Return Level and Exceedance Probability Estimates

Return Period	Return Level	Exceedance Probability
5 Months	81.78	0.1992
10 Months	94.96	0.0991
20 Months	106.23	0.04917
50 Months	119.04	0.01941
100 Months	127.48	0.00958
	Value	Exceedance Probability
Maximum Claim	125.46	0.01143
Upper Endpoint (X^F)	179.6	0

A6 **PWM Estimates of VaR an ES**

Table 5.1.4: Loss Mitigating Tools: 1-day-VaR and 1-day-ES Estimates

Probability	Value-at-Risk (VaR)
0.990	126.995
0.995	134.425
0.999	148.767

A7 PWM Estimates of the GPD Paramters

Table 5.1.5: Potential Thresholds And The PWM Parameter Estimates

Threshold	Exceedances	Shape	Scale
25	3276	-0.186	17.746
26	2821	-0.208	18.032
27	2130	-0.211	17.889
28	1754	-0.223	17.936
29	1754	-0.285	18.998
30	1754	-0.277	18.556

A8 Diagnostic Plots for the GPD-fit at Threshold of 25

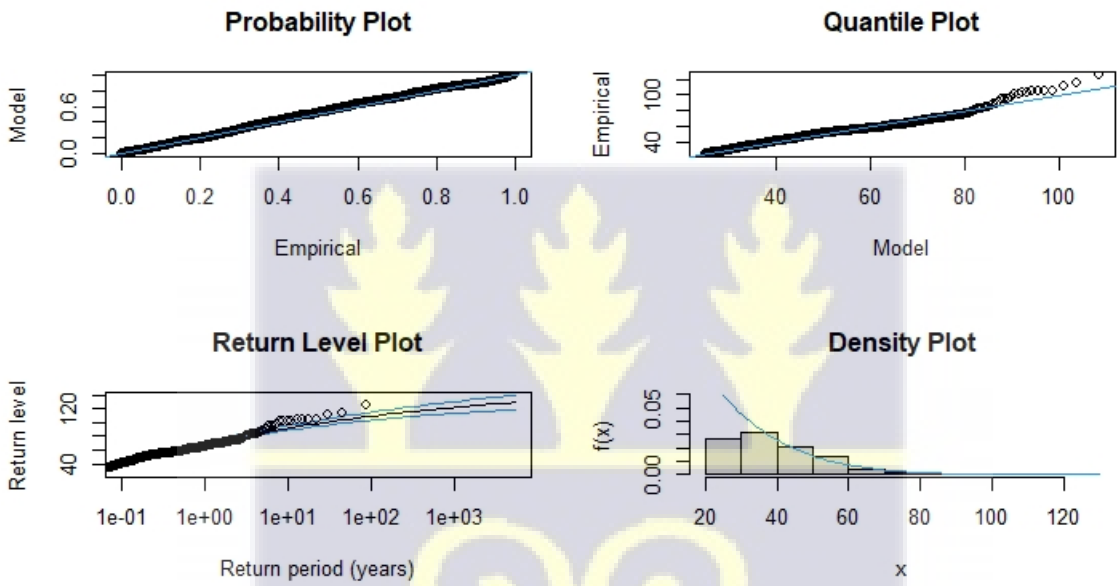

Figure 5.1.3: Diagnostic Plots for the GPD-fit at Threshold of 25

A9 Diagnostic Plots for the GPD-fit at Threshold of 26

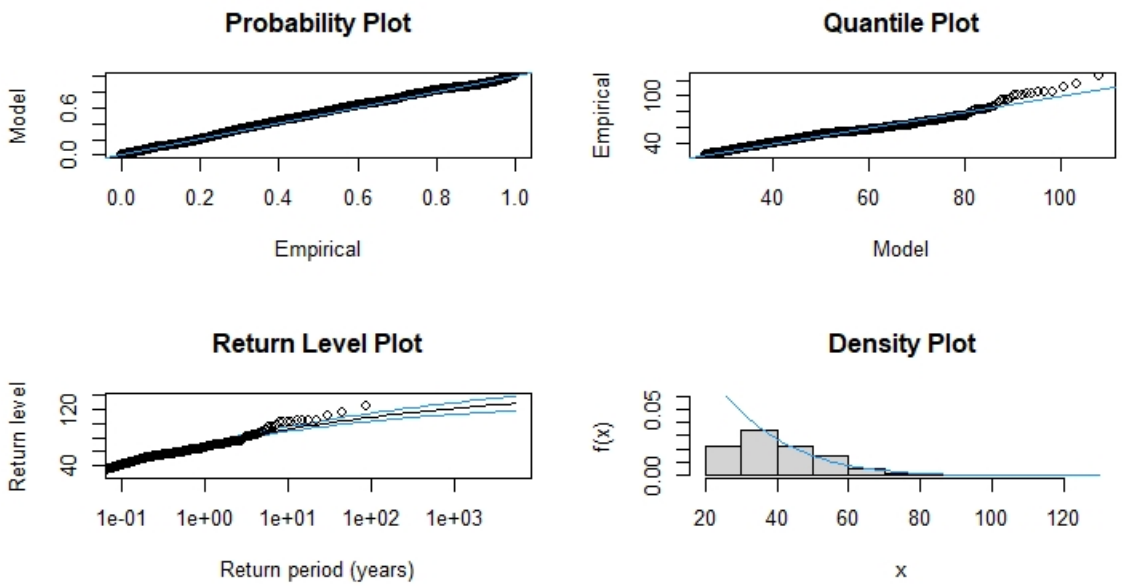

Figure 5.1.4: Diagnostic Plots for the GPD-fit at Threshold of 26

A10 Diagnostic Plots for the GPD-fit at Threshold of 27

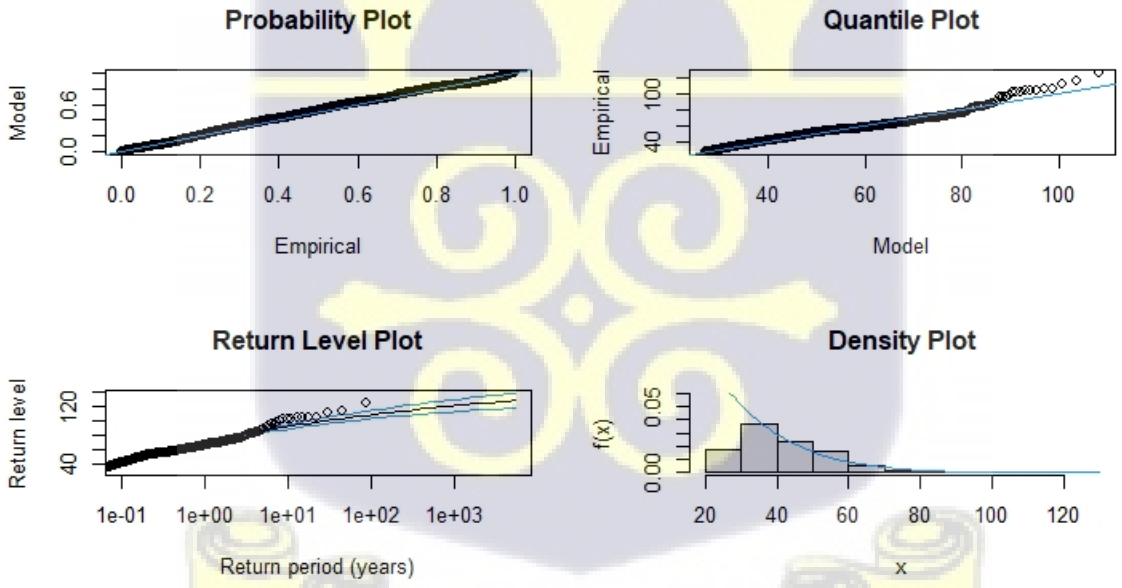

Figure 5.1.5: Diagnostic Plots for the GPD-fit at Threshold of 27

A11 Diagnostic Plots for the GPD-fit at Threshold of 28

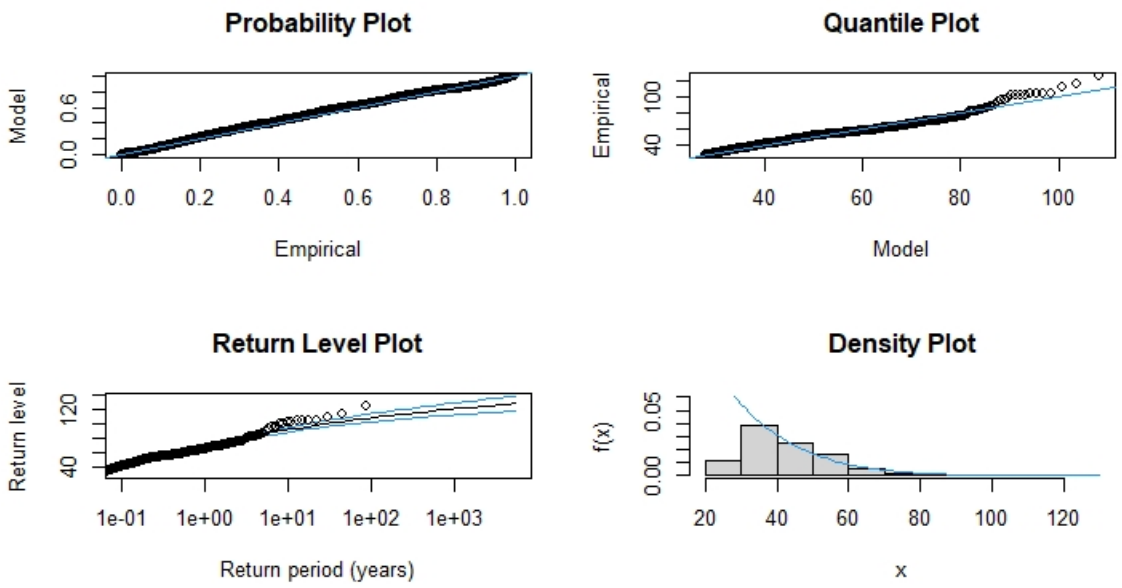

Figure 5.1.6: Diagnostic Plots for the GPD-fit at Threshold of 28

A12 Diagnostic Plots for the GPD-fit at Threshold of 29

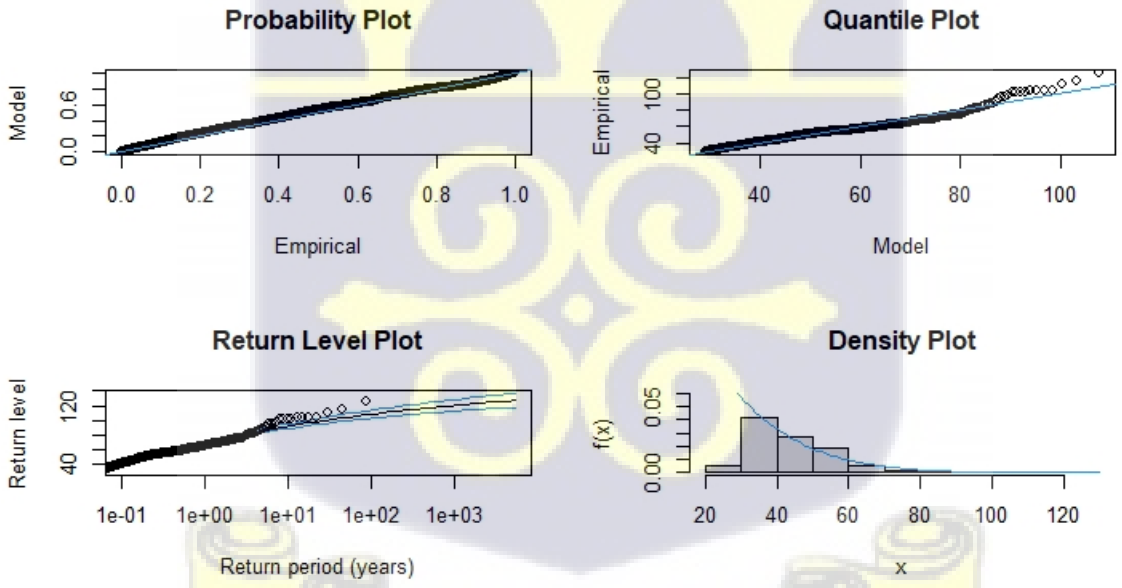

Figure 5.1.7: Diagnostic Plots for the GPD-fit at Threshold of 29

A13 Diagnostic Plots for the GPD-fit at Threshold of 30

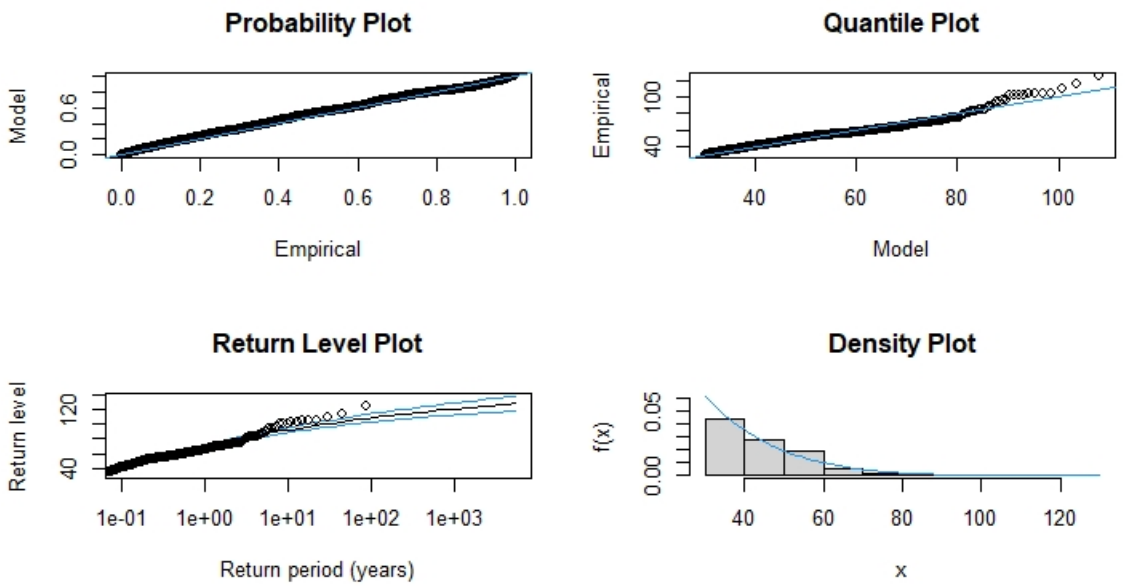

Figure 5.1.8: Diagnostic Plots for the GPD-fit at Threshold of 30

A14 PWM Estimates of the GPD Quantiles

Table 5.1.6: Return Period, Return Level and Exceedance Probability Estimates

Return Period	Return Level	Exceedance Probability
5 Months	54.097	0.053
10 Months	61.595	0.036
20 Months	67.786	0.026
50 Months	74.344	0.018
100 Months	78.311	0.014
	Value	Exceedance Probability
Maximum Claim	125.46	0.001
Upper Endpoint (X^F)	156.9	0

A15 **PWM Estimates of VaR and ES Under GPD**

Table 5.1.7: Loss Mitigating Tools: 1-day-VaR and 1-day-ES Estimates

Probability	Value-at-Risk (VaR)	Expected Shortfall (ES)
0.990	54.232	63.509
0.995	61.703	69.360
0.999	74.40	79.303

APPENDIX B

B1 Descriptive Statistics:

```
summary(d2$service)
describe(d2$service)
```

B2 Scatter plot and Histogram

```
plot(d2$service,xlab = "Daily Claims", ylab="Amount",

col=ifelse(d2$service>216,"red","dodgerblue"))

x<-d2$service
hist(x,breaks=90,main="", xlab="Claim Amount (in cedis)",
col="dodgerblue",freq=FALSE)
xfit<-seq(min(x),max(x),length=50)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
lines(xfit, yfit, col="red", lwd=2)
```

B3 Augmented Dickey-Fuller (ADF) Test for Stationarity

```
library(tseries)
adf.test(d2$service)
```

B4 Exponential QQ-Plot:

```
library(SMPracticals)
qqexp(d5$service,line=TRUE,col="dodgerblue",main="Exponential Q-Q Plot")
```

B5 Estimation of GEVD Parameters After Log Transformation

```
library(ismev)
l<-log(d5$service)
fevd(l, method = "MLE", type = "GEV")
```

B6 Diagnosis Plot for the GEVD-fit After Log Transformation

```
library(ismev)
l<-log(d5$service)
gev_mle<-gev.fit(l)
gev.diag(gev_mle)
```

B7 Diagnosis Plot for the GEVD-fit After Square-root Transformation

```
library(ismev)
s<-sqrt(d5$service)
gev_mle<-gev.fit(s)
gev.diag(gev_mle)
```

B8 Estimation of GEVD Parameters After Square-root Transformation

```
library(fExtremes)
s<-sqrt(d5$service)
gev_mle<-gevFit(s,type="mle")
gev_mle
gev_pwm<-gevFit(s,type="pwm")
gev_pwm
```

B9 Profile Likelihood-based Confidence Interval Estimation

```
s<-sqrt(d5$service)
gev_mle<-fgev(s)
confint(profile(gev_mle),level=0.95)
```

B10 Gumbel Test Statistic:

```
s<-sqrt(d5$service)
m<-length(s)
i<-c(1:m)
GS<-(max(s)-s[floor(m/2)+1])/(s[floor(m/2)+1]-min(s))
bn0<-(log(m)+log(log(2)))/(log(log(m))-log(log(2)))
an0<-1/(log(log(m)))
library(evd)
pvaluewn<-pgumbel((GS-bn0)/an0)
cat("[1] gs_m=",GS," gs*_m=",(GS-bn0)/an0," p-value=",pvaluewn,"\n")
```

B11 Negative Log-likelihood, AIC and BIC for the GEV and Gumbel models

```
library(extRemes)
#gev(weibull)
s<-sqrt(d5$service)
fevd(s, type = "GEV", method = "MLE", time.units = "months" )
#gumbel
fevd(s, type = "Gumbel", method = "MLE", time.units = "months" )
```

B12 Estimation of Return Levels Under GEVD

```
library(extRemes)
#mle
s<-sqrt(d5$service)
gev_mle<-fevd(s, type = "GEV", method = "MLE", time.units = "months" )
ci(gev_mle, return.period = c(5,10,20,50,100), alpha = 0.05)
#pwm
```

```
gev_pwm<-fevd(s, type = "GEV", method = "PWM", time.units = "months" )
ci(gev_pwm, return.period = c(5,10,20,50,100),alpha = 0.05)
```

B13 Estimation of Exceedance Probabilities for the GEVD

```
library(fExtremes)
#mle
pgev(81.16,50.69,23.17,-0.18, lower.tail = FALSE)
pgev(93.60,50.69,23.17,-0.18, lower.tail = FALSE)
pgev(104.05,50.69,23.17,-0.18, lower.tail = FALSE)
pgev(115.73,50.69,23.17,-0.18, lower.tail = FALSE)
pgev(154.76,50.69,23.17,-0.18, lower.tail = FALSE)
#max. claim
pgev(125.46,50.69,23.17,-0.18, lower.tail = FALSE)
#upper endpoint
pgev(179.6,50.69,23.17,-0.18, lower.tail = FALSE)

#pwm
pgev(81.78,50.28,23.56,-0.16, lower.tail = FALSE)
pgev(94.96,50.28,23.56,-0.16, lower.tail = FALSE)
pgev(106.23,50.28,23.56,-0.16, lower.tail = FALSE)
pgev(119.04,50.28,23.56,-0.16, lower.tail = FALSE)
pgev(127.48,50.28,23.56,-0.16, lower.tail = FALSE)
#max. claim
pgev(125.46,50.28,23.56,-0.16, lower.tail = FALSE)
#upper endpoint
pgev(179.6,50.28,23.56,-0.16, lower.tail = FALSE)
```

B14 Estimation of VaR and ES Under GEVD

```
library(fExtremes)
#mle
qgev(0.010,50.69,23.17,-0.18, lower.tail = FALSE)
qgev(0.005,50.69,23.17,-0.18, lower.tail = FALSE)
qgev(0.001,50.69,23.17,-0.18, lower.tail = FALSE)

#pwm
qgev(0.010,50.28,23.56,-0.16, lower.tail = FALSE)
qgev(0.005,50.28,23.56,-0.16, lower.tail = FALSE)
qgev(0.001,50.28,23.56,-0.16, lower.tail = FALSE)
```

B15 Mean Excess Plot

```
library(extRemes)
claims<-sqrt(d2$service)
claims<-sort(claims)
mrlplot(claims,main = "Mean Excess Plot",xlab = "Order Statistics (u)",
ylab = " Mean Excess Value");axis(1, at=seq(0,125,by=10));
grid(25,25,lwd=0.0005,col = "lightblue")
```

B16 Stability Plot for the Shape Parameter

```
library(evd)
tshapeplot(claims,tlim=c(5,100),nt=97,alpha=0.05,
legend.loc = NULL,xlab = "Order Statistics",main = "");
axis(1, at=seq(5,100, by=5));grid(15,15,lwd=0.0005,col = "lightblue")
```

B17 Stability Plot for the Scale Parameter

```
library(evd)
tscaleplot(claims,tlim=c(5,100),nt=97,alpha=0.05,
legend.loc = NULL,xlab = "Order Statistics",main = "");
axis(1, at=seq(5,100, by=5));grid(15,15,lwd=0.0005,col = "lightblue")
```

B18 **Parameters Estimation, Negative Log-likelihood, AIC, and BIC Scores for the Potential GPD Models**

```
library(fExtremes)
claims<-sqrt(d2$service)
#mle
gpdFit(claims,u=25,type="mle",information = "observed")
gpdFit(claims,u=26,type="mle",information = "observed")
gpdFit(claims,u=27,type="mle",information = "observed")
gpdFit(claims,u=28,type="mle",information = "observed")
gpdFit(claims,u=29,type="mle",information = "observed")
gpdFit(claims,u=30,type="mle",information = "observed")

#pwm
gpdFit(claims,u=25,type="pwm",information = "observed")
gpdFit(claims,u=26,type="pwm",information = "observed")
gpdFit(claims,u=27,type="pwm",information = "observed")
gpdFit(claims,u=28,type="pwm",information = "observed")
gpdFit(claims,u=29,type="pwm",information = "observed")
gpdFit(claims,u=30,type="pwm",information = "observed")
```

B19 **Upper endpoint Estimation for the Potential GPD Models**

```
upperend<-function(u,scale,shape){
output<-u-scale/shape
return(output)
}
upperend(25,16.778,-0.124)
upperend(26,19.353,-0.175)
upperend(27,19.145,-0.174)
upperend(28,19.080,-0.177)
upperend(29,16.726,-0.134)
upperend(30,16.371,-0.129)
```

B20 Diagnosis Plot for the GPD-fit at Various Thresholds

```
library(ismev)
claims<-sqrt(d2$service)
a<-gpd.fit(claims,threshold=25,show=T)
gpd.diag(a)
b<-gpd.fit(claims,threshold=26,show=T)
gpd.diag(b)
c<-gpd.fit(claims,threshold=27,show=T)
gpd.diag(c)
d<-gpd.fit(claims,threshold=28,show=T)
gpd.diag(d)
e<-gpd.fit(claims,threshold=29,show=T)
gpd.diag(e)
f<-gpd.fit(claims,threshold=30,show=T)
gpd.diag(f)
```

B21 Estimation of Return Levels Under GPD

```
library(POT)
claims<-sqrt(d2$service)
#mle
mle<-fitgpd(claims,30,"mle")
mle.rl<-retlev(mle)
mle.rl(5)
mle.rl(10)
mle.rl(20)
mle.rl(50)
mle.rl(100)

#pwm
pwm<-fitgpd(claims,30,"pwmu")
pwm.rl<-retlev(pwm)
pwm.rl(5)
pwm.rl(10)
pwm.rl(20)
pwm.rl(50)
pwm.rl(100)
```

B22 Estimation of Exceedance Probabilities Under GPD

```
library(fExtremes)
#mle
pgpd(53.746,-0.129,16.371,lower.tail = FALSE)
pgpd(62.553,-0.129,16.371,lower.tail = FALSE)
pgpd(70.610,-0.129,16.371,lower.tail = FALSE)
pgpd(80.217,-0.129,16.371,lower.tail = FALSE)
```

```
pgpd(86.768,-0.129,16.371,lower.tail = FALSE)
#max. claim
pgpd(125.46,-0.129,16.371,lower.tail = FALSE)
#upper endpoint
pgpd(179.6,-0.129,16.371,lower.tail = FALSE)

#pwm
pgpd(54.097,-0.277,18.556,lower.tail = FALSE)
pgpd(61.595,-0.277,18.556,lower.tail = FALSE)
pgpd(67.786,-0.277,18.556,lower.tail = FALSE)
pgpd(74.344,-0.277,18.556,lower.tail = FALSE)
pgpd(78.311,-0.277,18.556,lower.tail = FALSE)
#max. claim
pgpd(125.46,-0.277,18.556,lower.tail = FALSE)
#upper endpoint
pgpd(200.9,-0.277,18.556,lower.tail = FALSE)
```

B23 Estimation of VaR and ES Under GPD

```
library(fExtremes)
claims<-sqrt(d2$service)
#mle
model<-gpdFit(claims,u=30,type="mle",information = "observed")
gpdRiskMeasures(model,prob=c(0.99,0.995,0.999))

#pwm
model<-gpdFit(claims,u=30,type="pwm",information = "observed")
gpdRiskMeasures(model,prob=c(0.99,0.995,0.999))
```