KTH ROYAL INSTITUTE
OF TECHNOLOGY

Degree Project in Financial Mathematics

Second Cycle, 30 credits

Pricing and Modeling Heavy Tailed Reinsurance Treaties

A Pricing Application to Risk XL Contracts

ORMIA ABDULLAH MOHAMAD

ANNA WESTIN

Stockholm, Sweden 2023

Abstract

To estimate the risk of a loss occurring for insurance takers is a difficult task in the insurance industry. It is an even more difficult task to price the risk for reinsurance companies which insures the primary insurers. Insurance that is bought by an insurance company, the cedent, from another insurance company, the reinsurer, is called treaty reinsurance. This type of reinsurance is the main focus in this thesis. A very common risk to insure, is the risk of fire in municipal and commercial properties which is the risk that is priced in this thesis. This thesis evaluates Länsförsäkringar AB's current pricing model which calculates the risk premium for Risk XL contracts. The goal of this thesis is to find areas of improvement for tail risk pricing. The risk premium is commonly calculated by using one of three different types of pricing models, experience rating, exposure rating and frequency-severity rating. This thesis focuses on frequency-severity pricing, which is a model that assumes independence between the frequency and the severity of losses, and therefore splits the two into separate models. This is a very common model used when pricing Risk XL contracts. The risk premium is calculated with the help of loss data from two insurance companies, a Norwegian and a Finnish insurance company. The main focus of this thesis is to price the risk with the help of extreme value theory. The method of moments method is used to estimate the parameters in the distribution for the frequency of losses, and peaks over threshold model to model the severity of the losses. In order to model the frequency of losses, two distributions are compared, the Poisson and the negative binomial distribution. There are different distributions that can be used to model the severity of losses. In order to evaluate which distribution is optimal to use, two different goodness of fit tests are applied, the Kolmogorov-Smirnov and the Anderson-Darling test. The Peaks over threshold model is a model that can be used with the Pareto distribution. With the help of the Hill estimator we are able to calculate a threshold u, which regulates the tail of the Pareto curve. To estimate the rest of the in-going parameters in the generalized Pareto distribution, the maximum likelihood and the least squares method are used. Lastly, the bootstrap method is used to estimate the uncertainty in the price which was calculated with the help of the estimated parameters. From this, empirical percentiles are calculated and set as guidelines to where the risk premium should lie between, in order for both the data sets to be considered fairly priced.

Keywords: Reinsurance, Extreme Value Theory, POT-model, Hill estimator, Risk XL contracts, Generalized Pareto distribution, Method of Moments.

Sammanfattning

Svensk titel: Prisättning och modellering av långsvansade återförsäkringsavtal - En prissättningstillämpning på Risk XL kontrakt

Att uppskatta risken för att en skada ska inträffa för försäkringstagarna är svår uppgift i försäkringsbranschen. Det är en ännu svårare uppgift är prissätta risken för återförsäkringsbolag som försäkrar direktförsäkrarna. Den försäkring som köps av direkförsäkrarna, cedenten, från återförsäkrarna kallas avtalad återförsäkring. Denna typ av återförsäkring är den som behandlas i denna rapport. En vanlig risk att prisätta är brandrisken för kommunala och industriella byggnader, vilket är risken som prissätts i denna rapport. Denna rapport utvärderar Länsförsäkringar AB's nuvarande prissättning som beräknar riskpremien för Risk XL kontrakt. Målet med denna rapport är att hitta förbättringsområden för långsvansad affär. Riskpremien kan beräknas med hjälp av tre vanliga typer av prissättningsmodeller, experience rating, exposure rating och frequency-severity raring. Denna rapport fokuserar endast på frequency-severity rating, vilket är en modell som antar att frekevensen av skador och storleken av dem är oberoende, de delas därmed upp de i separata modeller. Detta är en väldigt vanlig modell som används vid prissättning av Risk XL kontrakt. Riskpremien beräknas med hjälp av skadedata från två försäkringsbolag, ett norskt och ett finskt försäkringsbolag. Det huvudsakliga fokuset i denna rapport är att prissätta risken med hjälp av extremevärdesteori, med hjälp av momentmetoden skattas parametrarna i fördelningen för frekvensen av skador och peaks over threshold metoden används för att estimera storleken av de skadorna. För att modellera den förväntade frekvensen av skador så jämförs två fördelingar, Poissonfördelingen och den negativa binomialfördelningen. Det finns ett antal fördelningar som kan användas för att modellera storleken av skadorna. För att kunna avgöra vilken fördeling som är bäst att använda så har två olika goodness of fit test applicerats, Kolmogorov-Smirnov och Anderson-Darling testet. Peaks over threhsold modellen är en modell som kan användas med Paretofördelningen. Med hjälp av Hillestimatorn så beräknas en tröskel u som regulerar paretokurvans uteseende. För att beräkna de resterande parametrarna i den generaliserade Paretofördelningen används maximum likliehood och minsta kvadratmetoden. Slutligen används bootstrap metoden för att skatta osäkerheten i riskpremien som satts med hjälp av de skattade parametrarna. Utifrån den metoden så skapas percentiler som blir en riktlinje för vart riskpremien bör ligga för de datamängder som undersöks skall kunna anses vara rättvist prissatt.

Nyckelord: Återförsäkring, Extremevärdesteori, POT-modellen, Hill estimatorn, Risk XL kontrakt, generella Paretofördelningen, Momentmetoden.

Acknowledgements

We would like to thank our supervisor Camilla Landén Johansson for her great feedback and inputs when writing this thesis. We would also like to send out a huge thanks to Maximilian Alamgir, analyst and actuary and our supervisor at LF, for all his help, inputs and support. A special thanks to Vidar Elsin, data scientist, for his help with understanding important concepts and support when writing this thesis. We would also like to thank Stefan Schneider, manager of the International Reinsurance team at LF, for giving us the opportunity to write this thesis.

Last but not least we would like to thank out friends and family for their ongoing support throughout this thesis.

Ormia Abdullah Mohamad and Anna Westin
Stockholm, May 2023

Authors

Ormia Abdullah Mohamad <ormia@kth.se>
Anna Westin <amgashi@kth.se>
SCI - Financial Mathematics
KTH Royal Institute of Technology

Place of Project

Länsförsäkringar AB
Stockholm, Sweden

Examiner and Supervisor at KTH

Camilla Johansson Landén <landen@kth.se>
Mathematical Statistics
KTH Royal Institute of Technology

External Supervisor

Maximilian Alamgir <maximilian.alamgir@lansforsakringar.se>
Länsförsäkringar AB

Contents

List of Figures

Nomenclature

$\hat{\alpha}$	Tail Index
$\hat{H}(k)$	Hill Estimator
μ	Mean
σ	Scale
θ	Location
ξ	Shape
A_n	Test Statistic, A-D
cv	Critical Value
D	Deductible
D_n	Test Statistic, K-S
E_N	Empirical Distribution Function
F	Distribution Function
f	Density Function
H_0	Null Hypothesis
H_a	Alternative Hypothesis
k	Tail Parameter
L	Likelihood Function
u	Threshold
$X_1, \ldots X_n$	Set of Independent and Identically Distributed Random Variables
X_o	Original Loss
X_{DI}	Loss Deductible

X_{RI} Amount Paid by Reinsurer

Y_i i Number of Ordered Data Points

Acronyms

A-D Anderson-Darling

EVT Extreme Value Theory

GEV Generalized Extreme Value Distribution

GPD Generalized Pareto Distribution

HE Hill Estimator

i.i.d. Independently and Identically Distributed

K-S Kolmogorov-Smirnov

LF Länsförsäkringar AB

LSQ Least Squares Estimation

MLE Maximum Likelihood Estimation

NB Negative Binomial

PD1 Type 1 Pareto Distribution

Po Poisson

POT Model Peaks Over Threshold Model

RP Risk Premium

XL Excess of Loss

xs Excess of

Glossary

Underwriting: An agreement to take on a financial risk or damage in exchange for payment.
Reinsurance: Insurance for insurance companies
Cedent/Cedant: Insurance taker/buyer. An insurance company gives or "cedes" a part of its to other reinsurance companies.
Layer: The risk is placed within sections which are called layers in the insurance industry. A reinsurer offers to cover losses up to a certain limit.
Excess point: The deductible, where the cover starts and the self retention ends.
Treaty: Treaty reinsurance is the insurance purchased by the cedent from the reinsurer.

Chapter 1

Introduction

This section will introduce the concept of reinsurance, why it exists and how it works. This part will also cover the goals and purpose of this thesis.

1.1 Background

When signing an insurance contract the insurer agrees to cover the costs for the insurance taker (policy holder) in case of a loss, or anything else that the contract concerns. If there is a large fire, storm, natural disaster or flood in an area, it could effect a large amount of insurance contracts that cover different types of losses. If something like this occurs, it could result in a total loss cost that is too large for the direct insurer to cover, which is why reinsurance companies exist. Reinsurance companies come in to protect the direct insurer in case of large loss events. Reinsurance is simply insurance for insurance companies.

The direct insurer, also called the cedent, spreads out their risk between a number of reinsurance companies so that in case of a large loss, they are able to pay the policy holders for their losses. The baseline for this system is that no company on the insurance market can survive alone in case of a large loss, especially when insuring a large number of policy holders. One type of reinsurance is retrocession insurance, which is where reinsurers model their whole portfolio and seek insurance for it. This thesis however will only focus on reinsurance contracts, which insure individual policies.

How paid premiums and covers offered is transferred between the actors in the insurance market is presented visually in Figure 1.1.

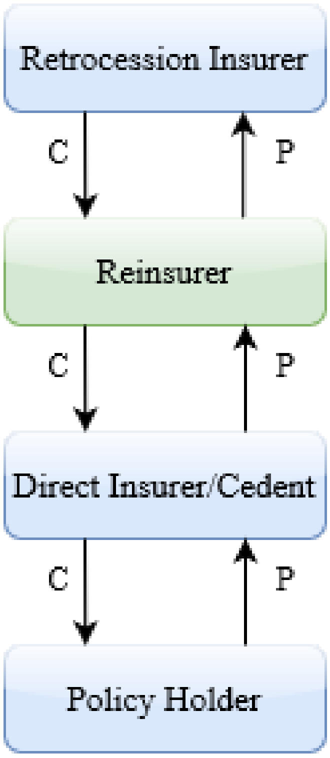


Figure 1.1: Cover, C, and Premium, P, transfer chart

In Figure 1.1, C, represents the cover and refers to bearing or sharing risks. The premium P refers to the premium being paid and the arrows show to whom.

1.2 Problems and Goals

The goal of this thesis is to bring forward a better way to price the risks that are being transferred among reinsurance companies by a ceding reinsurance company. Another goal is to show what methods are the most suitable when wanting to price contracts that cover extreme events and large losses. A way to accomplish this will be by looking at different distributions that are used for modeling in reinsurance, and analyze how they effect the price of the risk.

We will also look at the outcomes of estimating the parameters of the generalized Pareto distribution by means of the least squares method, and further how changing methods will ultimately effect the risk premium. Furthermore, we will look at how to improve and change the behavior of the tails of distribution functions, in for example the generalized Pareto distribution, defined in Section 2.5.1. This will be important when pricing the risk since it entails looking at extreme value data which is common in reinsurance. Lastly we will be looking at a way to re-sample the sparser data sets we will be working with. It is common in reinsurance to have smaller data sets since it mainly covers rare extreme events that affect multiple policy holders. These extreme events have a low probability of occurring, but when they do they can become very costly. We will therefore utilize the bootstrapping method, defined in Section 2.9.1, in order to get a pricing interval within which the risk premiums for each layer should fall within in order for the price to be considered fair. The fair price is defined in section 3.3.1.

We will evaluate and review if Länsförsäkringar AB's current pricing model is reasonable and investigate ways to improve it. We will use different distributions to estimate the price of the risk and the parameters used when pricing the risk. The end goal will be to create a percentile interval within which the price of the risk should fall within in order to be considered reasonable and fair. We will primarily use data from a Norwegian and Finnish Insurance company and the pricing will be done with a similar contractual structure as the cedent's proposed contracts of 2022. The data will consist of mainly historical occurrences of fires occurred in municipal and commercial properties.

1.3 Purpose

The purpose of this thesis is to better price independent risks in reinsurance and to produce a reasonable percentage interval for the risk premiums.

1.4 Delimitations

This thesis only handles data provided by Länsförsäkringar AB, from two Scandinavian companies and all results are based on that data. The data is limited to property fire claims. There are other risks that are priced by reinsurance companies that have not been considered in this thesis. The data provided are actual occurrences and their respective registered claims, which provides "real life" results. Because of this, for privacy reasons, we have pseudonymized the information involving the registered claims, as well as the structure of the contracts. The code used for the analysis in this thesis will therefore not be published either.

The methodology is based on an already implemented pricing model by LF, which follows a common method in the reinsurance industry. As mentioned, the methods will only be tested for two specific data sets. The methods implemented should however theoretically work well for other data sets as well.

1.5 Outline

This thesis includes five chapters excluding appendices which follows the structure:

The first chapter gives an introduction to the topic of the thesis, the second chapter includes all relevant theoretical background to understand the work that is done. The third chapter covers the methodology that was used to obtain the results. Thereafter, in Chapter 4 the results are presented. In the final chapter, Chapter 5, the results are discussed, and conclusions based on the results are drawn. Ideas for future work are also presented in Chapter 5.

Chapter 2

Theoretical Background

This section covers different parts of reinsurance and pricing models. First, the concept of reinsurance is introduced in more detail, and second the most common types of reinsurance are explained. Third, the most common pricing models in reinsurance are described. Thereafter, the data evaluation models as well as how to model different parts of reinsurance are presented. Finally, model evaluations and determinations are covered.

2.1 Definition of Reinsurance

According to [35] reinsurance refers to insurance for insurance companies. The two main reasons for purchasing reinsurance are; to minimize the amount of fluctuations in the losses that one insurance company must carry on their own, and to protect the insurance company in the case of a catastrophic event [35]. Reinsurance enables insurance companies to share the risk from claims that would have been too costly for a single insurance company to take up on their own. An insurance company will purchase insurance policies from multiple reinsurers in order to protect themselves from rare, but often costly events. The insurance company, also referred to as the cedent, will pay premiums to all the reinsurance companies involved in the policy. By spreading the risk between multiple insurance companies, the cedent company can still be able to offer insurance against more catastrophic events, such as natural disasters, or terrorism, to their direct customers. Reinsurance is an instrument that can help insurance companies decrease their risk of ruin in the case of unforeseen and costly events.

2.2 Types of Reinsurance

There are two main types of contracts that all reinsurance can be divided into, these are facultative reinsurance, and obligatory reinsurance. As stated by [35] facultative reinsurance uses a case-by-case approach, meaning that the cedent, the insurance taker, always has the option to accept or reject the individual policy offered to them. Obligatory reinsurance on the other hand, covers an entire insurance portfolio, and as the name suggests, the cedent would have to purchase a contract that covers the entirety of multiple insurance policies.

2.2.1 Proportional and Non-Proportional Reinsurance

Both facultative and obligatory reinsurance can be further divided into two subgroups, proportional and non-proportional treaties. Treaty reinsurance is the insurance an insurance company, the cedent, buys from another insurance companies, the reinsurer. In proportional reinsurance, the cedent and the reinsurer will share certain proportions, either in terms of percentages or set amounts of the premiums and losses. Non-proportional insurance on the other hand, does not have a pre-determined proportion that is set between the cedent and the reinsurer. Two of the most common forms of non-proportional reinsurance is Stop-Loss insurance and Risk Excess of Loss (XL) insurance. An example of where Risk XL contracts are used, are for accidents related to fires.

2.2.2 Risk Excess of Loss Insurance

The following theory can be found in [1]. The structure of an XL, works in such a way that the loss amount from the cedent must exceed a certain amount, known as the deductible, D, in order to be recovered from the reinsurer. There also exists a cover limit, C, which represents up to what amount the reinsurer will cover a claim. It holds that $D < C$. The losses that the cedent is required to pay their own direct customers can be defined as the the original loss X_0, and will constitute the two parts as: $X_0 = X_{DI} + X_{RI}$. Where X_{DI} is the loss deductible, which is the amount that is needed to be paid by the cedent, and X_{RI} is the amount paid by the reinsurer. The XL contract is then given by

$$[X_{RI} = max(0, X_0 - D) - max(0, X_0 - C - D), \quad X_{DI} = X_0 - X_{RI}]. \tag{2.1}$$

If $X_0 > D$ then the cedent will be compensated by the reinsurer. Often these contracts are divided into separate layers, each with different deductibles and cover limits. If divided into limits, the contract can then be denoted as: $C \quad xs \quad D$, which reads C excess D. To illustrate, let for example $C = \$850,000$, and $D = \$250,000$. This would mean that the cedent has a deductible of $\$250,000$, and anything above that amount up until $850,000 is paid by the reinsurer. Anything above the cover limit would go into a new layer, that has another deductible and cover limit. Therefore, the losses in each layer for the reinsurer can be expressed as a function of the original loss as:

$$L_{D,C}(X) = max(0, X_0 - D) - max(0, X_0 - C - D). \tag{2.2}$$

Thus, $L_{D,C}(X)$ denotes how much will be paid for by the reinsurer for that particular layer. The division of layers is usually set by the ceding company themselves or their reinsurnace brokers.

To illustrate the layer structure of reinsurance contracts, Figure 2.1 shows what a typical XL contract can look like. Retention refers to the deductible, the amount that the cedent themselves are responsible to pay for. Treaty refers to the total cost that the cedent has reported to be covered by the reinsurer. It is this amount that is then divided into the different layers as shown.

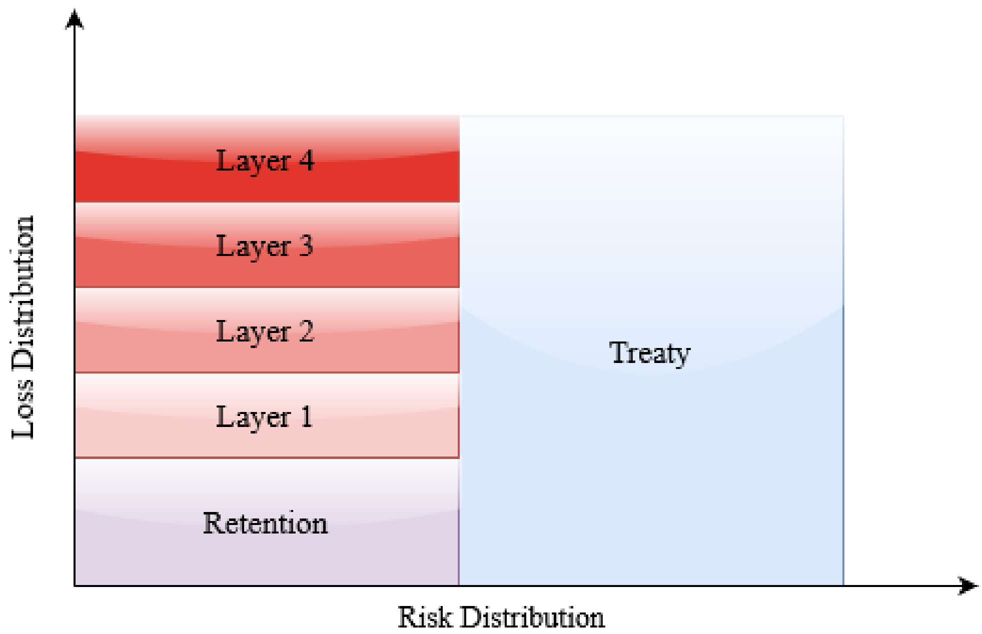


Figure 2.1: Loss and risk distribution for a treaty

2.3 Pricing Models

An essential part of reinsurance is to be able to correctly price risks. A reinsurance company needs to know what price they should charge for each of the different layers in an reinsurance contract. A risk premium is therefore determined for each layer. There exists several different methods for doing so, depending on what and how much previous data is available from the cedent. Pricing correctly is very important for reinsurers so that in the event of a loss they are able to regain the lost capital over a certain period of time. The factors that determine the price of a risk premium include, the degree of risk that is transferred, magnitude of capacity, financial aid offered, and the reinsurers expenses, such as broker fees and other internal fees. This however is not enough when deciding on a price of the risk premium. The reinsurance market is often influenced by competition and the current state of the world. This thesis will not focus on all these extra factors that influence the price. The "fair price" will be the main part being investigated. Pricing a risk premium in general starts off with gathering information about the ceding company, exposure, expenses and other information about the area being insured.

The different models can be divided into three different categories; experience ratings, exposure ratings, and the frequency-severity rating. Experience ratings, is defined as a methods strictly based on historical loss information. The basic idea of experience rating is that the historical experience, adjusted properly, is the best predictor of future expectations. The historical data utilized to price the risk premium, is adjusted for rate changes and inflation each year. Exposure ratings on the other hand, are often utilized when there does not exist enough or adequate historic data from the cedent. In such a case, an insurance portfolio with similar attributes will be utilized in order to determine the risk premium. The third method is called the frequency- severity rating, and it is considered a hybrid between the two previous models. The model relies on historic events for parameter determination, but once determined, it can be used to forecast losses in layers without prior data. The frequency-severity rating separates the frequency of claims into one distribution and the severity of claims into another.

Experience Rating

Experience rating, is a pricing method where the risk is calculated based on loss experiences for the insurance company. The reinsurer uses the losses that have occurred for the ceding company and use that data to approximate future losses for the same type of risk. This method is limiting, since it is common in reinsurance to say that the largest claim has not yet occurred [13].

Exposure Rating

Exposure rating is the second pricing method, that, contrary to the previously mentioned rating, models the risk profiles instead of using historical claims to price the risk. The method is based on using claims from a similar company with similar exposure, and it uses that data to create a model that estimates the expected losses for the ceding insurance company. The exposure rating method uses premium and limits suggested to estimate a loss cost and losses for each layer. The layer losses are specifically estimated by using a severity distribution.

Exposure rating does not require the reinsurer to have any historical experience with a specific risk for a specific cedent. Therefore, it is useful when collaborating with a new cedent [11][13].

Frequency-Severity Rating

The frequency-severity rating is a third pricing method. The model is based on developing a stochastic model of the specific risk being insured. The potential loss is simulated based on the reinsured losses and the given terms for the specific contract. This type of rating is very useful and versatile since the stochastic model can be modeled with general industry data; experience data for a specific contract or risk or with a combination of both [13]. This is the model that will be used in this thesis, due to it being versatile and for the fact that it can be applied to general industry data.

The frequency severity rating makes the assumption that the number of claims N are independent from the size of the claims X_i, for $i = 1, ..., N$. This assumption is done so that N and X can be modeled separately with different distributions, from which the expected values $E[N]$ and $E[X]$ can be determined. The expected values $E[N]$ and $E[X]$ will be used to compute the risk premium, which will be further explained in Section 3.3.
Therefore, if we let N be a random variable modeling the frequency of claims, that is independent of the X_i's, then the expected claim cost over the total amount of claims, is

$$E\left[\sum_{i=1}^{N} X_i\right] = E\left[E\left[\sum_{i=1}^{N} X_i \mid N\right]\right] = \sum_{i=1}^{\infty}\left(E\left[\sum_{i=1}^{N} X_i \mid N\right] P(N=n)\right), \tag{2.3}$$

but

$$E\left[\sum_{i=1}^{N} X_i \mid N=n\right] \quad = E\left[\sum_{i=1}^{n} X_i\right] = nE[X], \tag{2.4}$$

where the last step follows from the fact that the X_i's are i.i.d. Here $E[X]$ is the common expectation of the severity distribution of the claims. This implies that

$$\begin{aligned} \sum_{i=1}^{\infty}\left(E\left[\sum_{i=1}^{N} X_i \mid N\right] P(N=n)\right) &= \sum_{i=1}^{\infty} nE[X]P(N=n) = \\ &= E[X]\sum_{i=1}^{\infty} nP(N=n) = E[X]E[N]. \end{aligned} \tag{2.5}$$

For this thesis, both data sets involve fire claims. It is important to determine whether it is reasonable to assume that the number of claims N is independent from the size of the claims X_i for this type of data. It is a strong assumption to make, since it is easy to imagine factors that could influence both the frequency and the severity of fires, for example an extremely dry season. However, this is a standard assumption to make for reinsurance contracts for frequency-severity ratings, and this is also done for LF's pricing models. We will therefore treat N and X as if they were independent in this thesis as well [31].

2.4 Modeling the Frequency of Claims

For the frequency of claims, a common distribution used is the Poisson distribution. The negative binomial distribution is another distribution that can also be suitable to use. In this thesis, we will compare the effects on the risk premium from using both the Poisson distribution and also the negative binomial distribution to estimate the frequency of claims. Following is a brief description of both of the distributions.

2.4.1 Poisson Distribution

The Poisson distribution is commonly utilized in insurance settings to estimate the number of claims N. The Poisson distribution is a discrete probability distribution that describes the probability that a certain number of events will occur in a fixed time interval. According to [20], the Poisson distribution is considered beneficial to estimate the number of claims if the following assumptions hold:
(1) It is assumed that each event in the time interval is independent,
(2) Only one event at a time occurs,
(3) The probability that a specific event occurs at a specific time point is zero.
If a stochastic variable X is Poisson distributed, the probability mass function is given by

$$p_X(x) = P(X = x) = \frac{\mu^x e^{-\mu}}{x!}, \quad x = 0, 1, 2, ... \tag{2.6}$$

in which k is defined as the number of occurrences, and μ is the expected value of X, and also the variance, such that $\mu = E[X] = Var[X]$.

2.4.2 Negative Binomial Distribution

The negative binomial distribution (NB) is based on a sequence of independent Bernoulli trials. For a sequence of n Bernoulli trials, each trial will have two possible outcomes, with probability p for success and $1 - p$ for failure. The NB distribution is the distribution of the number n of trials it takes to get a fixed number of r successes.

Let X denote the total number of trials it takes to obtain r successes.

For the outcome x to occur, the first $x - 1$ trials will result in $r - 1$ successes, which yields,

$$\binom{x-1}{r-1} \cdot p^{(r-1)}(1-p)^{(x-1)-(r-1)}, \tag{2.7}$$

which means that the x^{th} trial results in a success with a probability p.

Thus the probability that the r^{th} success occurs on the x^{th} trial will be,

$$P(X=x) = \binom{x-1}{r-1} \cdot p^r(1-p)^{(x-r)}, \tag{2.8}$$

where $x = r, r+1, \ldots,$. Then we can say that the random variable X has a negative binomial (r, p)-distribution, i.e. $X \sim NB(r, p)$ distributed.

The mean μ, for the number of trials, can therefore be defined as:

$$E[X] = \mu = \frac{r}{p}, \tag{2.9}$$

and the variance σ^2 as

$$V[X] = \sigma^2 = \frac{r \cdot (1-p)}{p^2}. \tag{2.10}$$

If we instead define the random variable Y as the number of failures before the r^{th} success, where $Y = X - r$, then one can define an alternative form of the negative binomial distribution as [29]:

$$P(Y=y) = \binom{r+y-1}{y} \cdot p^r(1-p)^{(y)}. \tag{2.11}$$

From this alternative definition, in Equation (2.11), one can define the mean of the number of losses as:

$$E[Y] = \mu = r \cdot \frac{(1-p)}{p}, \tag{2.12}$$

and the variance as:

$$V[Y] = V[X] = \sigma^2 = r\frac{(1-p)}{p^2}. \tag{2.13}$$

If $r < \infty$ the variance exceeds the mean for the alternative definition of the negative binomial distribution [28][4].

The alternative model defined in Equation (2.11), together with the definitions of the mean and variance seen in Equations (2.12) and (2.13), will be the version that is implemented in the model used in this thesis.

2.4.3 Method of Moments

The Method of Moments (MoM) is a common method for parameter estimation for many distributions, for example the Poisson and the negative binomial distribution. It works by matching the sample moments to the theoretical (distribution) moments.

In order to find the MoM estimator for a given distribution we can start off by defining X as a random variable with $X_1, \ldots, X_n$ be n number of samples from a distribution, depending on some parameter θ.

We define the theoretical moment of a distribution about the origin as:

$$E(X^k; \theta) = \mu^k(\theta), \tag{2.14}$$

where μ^k is the definition of the k^{th} theoretical moment, with $k = 1, 2, \ldots$.

We can also define the k^{th} theoretical moment of the distribution about the mean as:

$$E[(X - \mu(\theta)^k; \theta]. \tag{2.15}$$

We now define the k^{th} sample moment about the origin, M_k as:

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k = \bar{X}^k, \tag{2.16}$$

which is just the sample mean of the k^{th} moment, $\bar{X}^k$.

We also define the k^{th} sample moment about the mean, M_k^* as:

$$M_k^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k, \tag{2.17}$$

which for $k = 2$ is just the sample variance.

To estimate parameters, we recall that if $k = 1$ we get that the first sample moment M_1 and set it equal to the mean $E[X]$ of the distribution such as

$$M_1 = \bar{X} = E[X^1] = \mu^1(\theta). \tag{2.18}$$

We recall that the variance $V[X;\theta] = \sigma^2(\theta)$ can be written as:

$$V[X;\theta] = E[X^2;\theta] - (E[X;\theta])^2, \tag{2.19}$$

which is expressed in terms of the first and second theoretical moment. From Equations (2.19) and (2.18) we can express the sample moment M_k^* about the mean, for $k = 2$ as:

$$M_2^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sigma^2(\theta). \tag{2.20}$$

From this definition of the MoM, we can use the sample moments to estimate the theoretical moments of a distribution and get the MoM estimators of θ. In this thesis, the MoM was used to estimate the parameters for the Poisson and negative binomial distribution.

For the Poisson distribution the first sample moment is equal to the $E[X]$, which is μ, since $X \sim Po(X)$. Whilst for the negative binomial distribution, the first and second moment is used to find the parameters p and r [21] [6].

2.5 Modeling the Severity of Claims

The severity of claims refers to the cost associated with a specific loss. There are often many losses of lower cost, but only a few of extremely large costs for reinsurance contracts. It is can therefore be difficult to find one distribution that can capture the entire range of losses accurately. Because of this, it is important to select the most suitable distribution based on how large the claim is, and how likely they are to occur. As mentioned earlier, reinsurance often involves rare but catastrophic events that no single insurance company is able to take on by themselves. These types of claims are of high cost, and occur with small probabilities, and will therefore as a result fall on the heavy-tails of a distribution. The main priority for a reinsurer is to find a distribution that captures these events efficiently. Possible distributions that work well for heavy-tailed distributions could be the lognormal, Weibull, and different types of Pareto distributions. We will give a brief description of the distributions just mentioned as well as the normal distribution, as it is used for comparison for goodness of fit tests in a following chapter.

2.5.1 The Pareto Distribution

A distribution that is commonly used to model the severity of claims in reinsurance is the Pareto distribution. One reason that the Pareto distributions are utilized is because it is able to capture the heavy tails that often exist in reinsurance data. The heavy tailed distribution is defined as the probability distributions where the tails of the distributions are not bounded exponentially. As mentioned by [14] *"if we restrict the fit on the very large losses the Pareto distribution or variants thereof frequently seem the best choice"*. The Pareto distribution is suitable for modeling the heavy tails because of its quality of parameter invariance. As [14] describes it, parameter invariance refers to that the distribution is so called *"memory-less"*, meaning that the distribution forgets the scale parameter in the tails of the distribution. This means that in order to work with the tail of a distribution, one does not need to know where the tail begins, which is a feature that will be useful for reinsurance contracts. It should be noted that usually these heavy-tailed distributions are not able to capture the events below the tail well. However, this might not be an issue if the focus is for reinsurance policies. There exist many different versions of the Pareto distribution with varying amounts of parameters used. Below is an introduction to the most common types used in reinsurance.

Generalized Pareto Distribution

The standard cumulative distribution function of the generalized Pareto distribution (GPD) for the stochastic variable X, is given by:

$$F_X(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\theta)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\theta}{\sigma}\right) & \text{for } \xi = 0 \end{cases}. \tag{2.21}$$

The corresponding probability density function is given by:

$$f_X(x) = \frac{1}{\sigma}\left[1 + \frac{\xi(x-\theta)}{\sigma}\right]^{-1/\xi-1}, \tag{2.22}$$

in which $x \geq 0$ for $\xi \geq 0$, and $0 \leq x \leq 1/\xi$ for $\xi < 0$. The parameter θ is the location parameter, ξ the shape parameter, and σ the scale parameter. The expected value of the GPD can be calculated using the equation

$$E[X] = \theta + \frac{\sigma}{1-\xi}. \tag{2.23}$$

In this thesis, $E[X]$ will be used to calculate the risk premium of each layer in the reinsurnce contracts.

The above information and further details about the generalized Pareto distribution can be found in [27].

Type I Pareto Distribution

A specific version of the generalized Pareto distribution is the type I Pareto distribution (PD1). The type I Pareto distribution differs from the GPD in that it only uses two parameters instead of three, as it does not include a location parameter. The distribution has a shape parameter α and a scale parameter x_m.The density function of the PD1 is

$$f_X(x) = \begin{cases} \frac{\alpha x_m^\alpha}{x^{\alpha+1}} & x \geq x_m \\ 0 & x < x_m. \end{cases} \tag{2.24}$$

It is assumed that both α and x_m are positive. The cumulative distribution function is given by:

$$F_X(x) = \begin{cases} 1 - (\frac{x_m}{x})^\alpha & x \geq x_m \\ 0 & x < x_m. \end{cases} \tag{2.25}$$

Since the PD1 is a special case of the GDP, there is a relationship between the two distributions which in terms of the parameters for GDP is such that; the location parameter is $\theta = x_m$, the shape parameter is $\xi = \frac{1}{\alpha}$, and scale parameter is $\sigma = \frac{x_m}{\alpha}$.

The above and more specific details can be found in [36].

2.5.2 Weibull Distribution

The Weibull distribution, is a distribution that is sometimes used to model the severity distribution in reinsurance contracts. It is suitable to model extreme data since it is a heavy-tailed distribution. The density function for the Weibull distribution (WD) is given by:

$$f_X(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \geq 0 \\ 0, & x < 0 \end{cases}, \tag{2.26}$$

in which $k > 0$ is a shape parameter, and $\lambda > 0$, is a scale parameter. The cumulative distribution function is given by $F_X(x) = 1 - e^{-(x/\lambda)^k}$.

2.5.3 Normal Distribution

The normal distribution will be used for comparison in the goodness of fit tests, and is therefore included here. This distribution is not considered a heavy tailed distribution. The normal distribution has the following density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.27}$$

and the cumulative distribution function is given by:

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt. \tag{2.28}$$

The parameter μ is the mean of the distribution, and σ is the standard deviation.

2.5.4 Lognormal Distribution

The lognormal is a suitable distribution to model the insurance claim costs, since the costs cannot fall below zero and the costs can become very large. The lognormal distribution has the following density function

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \tag{2.29}$$

and the cumulative distribution function is given by:

$$F_X(x) = \Phi\left(\frac{(\ln x) - \mu}{\sigma}\right). \tag{2.30}$$

Here Φ is the cumulative distribution of the standard normal distribution. The log-normal distribution is classified as a heavy tailed distribution.

The distributions can be found in [15].

2.6 Extreme Value Theory (EVT)

When modeling a sample of data, the statistical models, will often be able to capture the common events well, i.e. the events in the center of the distribution, while not being able to accurately capture its tails. As mentioned, since reinsurance often involves extreme values that occur with a low probability, methods that focus on these tails are important. Extreme Value Theory (EVT) does exactly this. The theory focus on the tails of the distributions, and states that regardless of what underlying the data comes from, the tail of the distribution will belong to one of three different distribution types. Therefore given a sample, the EVT is able to determine the probability of a random event happening that is more extreme than that has previously been observed. EVT is the study of extreme values that deviate significantly from the median of the probability distribution, and these events can involve minimums or maximums, which represent the tails. Reinsurance is mainly concerned with being able to model the upper tails.

In EVT, two different models, the block maxima models and threshold exceedance models are common and often utilized. Block maxima models involve the largest values from samples of identically distributed variables, whereas the threshold exceedances models, are often practical to use when the data on extreme outcomes is limited.

2.6.1 Generalized Extreme Value Distribution

To begin the discussion of EVT and the different types of models that are used within the theory, it is important to look at the general behavior of an extreme value. Suppose that $X_1, X_2, ..., X_n$ are independent random variables with the same probability distribution, ordered from smallest to largest. Denote their sum by $S_n = X_1 + X_2 + ... + X_n$. The central limit theorem states that appropriately normalized sums $(S_n - a_n)/b_n$ will converge to a normal distribution as n goes to infinity, as:

$$\lim_{n\to\infty} P\left(\frac{S_n - a_n}{b_n} \leq X\right) = \Phi(x), \tag{2.31}$$

in which $a_n > 0$, $b_n > 0$ represent normalizing constants, and are defined as

$$a_n = nE[X_1] \quad b_n = \sqrt{nVar(X_1)}. \tag{2.32}$$

Analogous to the central limit theorem, as $n \to \infty$ the largest observations from identically distributed samples will tend to the Generalized Extreme value distribution (GEV), which is covered next.

Block Maxima

The Block Maxima method follows directly from EVT. Similar to the central limit theorem, but for extreme values, we assume that we have some $M_n = (X_1, X_2, ..., X_n)$ of random i.i.d. variables, which represents normalized maximas from n observations. These are called block maxima, and represent the extremas from non overlapping time periods of equal size, such as: $P(M_n \leq x) = P(X_1 \leq x, ..., X_n \leq x) = F_X^n(x)$, where F_X^n is the cumulative distribution function of the block maxima.
Then as $n \to \infty$, similarly to the central limit theorem, M_n will converge in distribution to the Generalized Extreme value distribution (GEV), as shown below.

$$\mathrm{P}\left(\frac{M_n - d_n}{c_n} \leq x\right) = F_X^n\left(c_n x + d_n\right) \to G_X(x), \tag{2.33}$$

in which c_n and d_n are real constants, and $c_n > 0$. Here, $G_X(x)$ represents the GEV, and is given by:

$$G_X(x) = \exp\left\{-\left(1 + \xi\frac{x-\theta}{\sigma}\right)_+^{-1/\xi}\right\}, \tag{2.34}$$

in which σ is the scale parameter, θ is the location parameter, and ξ is the shape parameter. There are three different types of the GEV, and the shape parameter ξ determines which form the distribution will have. The parameter ξ is defined as $\xi = \frac{1}{\hat{\alpha}}$, in which $\hat{\alpha}$ is defined as the tail index. The lower the value of the tail index $\hat{\alpha}$ is , the heavier the tail will be. The three types of distributions are:

i.) Fréchet distributions, These involve any distribution with thick tails, such as the different Pareto distributions, here $\xi > 0$.

ii.) Gumbel distributions, characterized by distributions that decline exponentially, for example the lognormal distribution, here $\xi = 0$.

iii.) Weibull distributions, which are characterized by distributions with a thin tail and a finite endpoint, here $\xi < 0$.

All the theory and mathematical models from this section were found in [24], [8], and [3].

Peaks Over Threshold

A drawback of the block maxima method is that relatively large samples are needed for sufficient estimation. The method of Peaks Over Threshold (POT) is much more practical to use when working with smaller sample sizes, which is common in reinsurance. The POT method is a method that can be used to estimate tail probabilities and quantiles. It is especially useful when data is sparse, since it will only look at all values that exceed a certain threshold u.

The advantage of this method in reinsurance compared to other methods is that POT is very good for extreme values and to model the tail of a distribution, and in particular for the generalized Pareto distribution (GPD) [32].

We begin by defining observations of i.i.d. random variables $X_1, \dots X_n$ that have the same unknown distribution function F. Suppose that the distribution function F has a varying right tail, with tail index $\hat{\alpha}$, as in the GPD, defined as:

$$\bar{F}(x) = P(X_k > x). \tag{2.35}$$

Previously, we have defined the standard CDF $F_X(x)$ for the GPD for $\sigma > 0$ and $\xi > 0$ as:

$$F_X(x) = 1 - \left(1 + \frac{\xi(x-\theta)}{\sigma}\right)^{-1/\xi}, \tag{2.36}$$

where $x \geq 0$. To make this easier to follow we define the CDF for the GPD as

$$G_{\xi,\sigma} = F_X(x). \tag{2.37}$$

Now we define the excess distribution function of X over the threshold u as:

$$F_u(x) = P(X - u \leq x | X > u) \quad \text{for } x \geq 0. \tag{2.38}$$

We also have the general relation that

$$\bar{F}_u(x) = \frac{\bar{F}(u+x)}{\bar{F}(u)} = \frac{\bar{F}(u(1+\frac{x}{u}))}{\bar{F}(u)}, \tag{2.39}$$

in which $\bar{F}$ is regularly varying with the tail index $-\hat{\alpha} < 0$ as in the GPD.

From Equation (2.39) we get the following expression:

$$\lim_{u \to \infty} \sup_{x \geq 1} |\bar{F}_u(x) - \bar{G}_{\xi,\sigma(u)}(x)| = 0, \tag{2.40}$$

where $\xi = \frac{1}{\hat{\alpha}}$ and $\sigma(u) \sim \frac{u}{\hat{\alpha}}$ as $u \to \infty$, where σ is the scale parameter from the GPD, as defined in Section 2.5.1, that depends on the threshold u.

Now with the help of Equations (2.36), (2.37), (2.38), (2.39) and 2.40 we show how we can get the natural tail and quantile estimators. We start of with the i.i.d. observations $X_1 \dots, X_n$. By choosing a high threshold u, or mathematically chose one as explained in Section 2.6.2, we let the number of exceedences of the threshold be defined as

$$N_u = \#\{i \in \{1, \dots, n\} : X_i > u\}. \tag{2.41}$$

From previously defined in Equation (2.39) we know that

$$\bar{F}(x+u) = \bar{F}(u)\bar{F}_u(x). \tag{2.42}$$

It is important to mention that if the threshold is not far into the tail, then we know that the empirical approximation, defined as,

$$\bar{F}(u) \approx \bar{F}_n(u) = \frac{N_u}{n}, \tag{2.43}$$

is valid. This also gives us the approximation:

$$\bar{F}_u(x) \approx \bar{G}_{\xi,\sigma(u)}(x) \approx \bar{G}_{\hat{\xi},\hat{\sigma}(u)}(x) = \left(1 + \frac{\hat{\xi} \cdot x}{\hat{\sigma}}\right)^{\frac{-1}{\hat{\xi}}}, \tag{2.44}$$

where $\hat{\xi}$ and $\hat{\sigma}$ are estimated parameters. With the help of Equation (2.44) we get the following expression for the estimator of the quantile $F^{-1}(p)$:

$$\hat{F}^{-1}(p) = min\{x : \hat{\hat{F}}(x) \leq 1 - p\} = \cdots = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\xi}} - 1\right). \tag{2.45}$$

This concludes the derivation for the POT method for estimating tail probabilities and quantiles [19].

2.6.2 Hill Estimator

The choice of the location parameter θ is very important when applying the POT method and estimating the rest of the GPD parameters, i.e. the shape ξ and scale σ. In reinsurance the location parameter theta is sometimes chosen manually, based on the layer structure of a contract. Even though this method is common practice, it does not guarantee that the optimal value for the parameter is chosen. The Hill estimator (HE) is a theoretically supported method that is used to estimate θ. The location parameter θ is often denoted as the threshold u, where all random variables $X > u$, in the tail are Pareto distributed. The HE calculates after which data point the values are Pareto distributed. Since HE is a method that helps decide the optimal u, it can significantly change the shape of the Pareto curve.

If we assume we have a data set that only contains positive values, Then HE can be calculated as a function of the tail parameter k, which has nothing to do with the behavior of the tail. The tail parameter k is the order statistic and is used to find the right threshold. We assume that the data set is i.i.d., and are defined as $X_1 \dots X_n$, with corresponding distribution function $F_X(x)$. We also assume that the data is sorted in an decreasing order, such that $X^{(1)} \geq \cdots \geq X^{(n)}$.

The Hill estimator $\hat{H}(k)$ for a suitable $k = k(n)$ is then given by

$$\hat{H}(k) = \frac{1}{\hat{\alpha}} = \frac{1}{k}\sum_{i=1}^{k}(ln(X^{(i)} - ln(X^{(k)})), \tag{2.46}$$

where $\hat{\alpha}$ is the tail index, which decides how heavy the tail is. It is important to mention that the threshold can be a suitable input for the GPD even if it is not the optimal value depending on how the data set looks. If u, i.e. k are too small, we have much data however the approximation defined in Equation (2.44) will be uncertain. If u is too large we will miss out on important data points [19]. A rule of thumb is that a higher tail index gives a thinner tail, and vice versa [7][18]. Once the optimal threshold is found, the rest of the Pareto parameters can be estimated.

2.7 Model Determination and Evaluation

Goodness of fit tests are used in order to determine which distribution is the most suitable to use when pricing large loss contracts. Two of the most well-known ones are the Kolmogorov-Smirnov test and the Anderson-Darling test. Both tests are hypothesis tests with a null hypothesis H_0, stating that the data follows a specific distribution, and an alternative hypothesis, H_a, stating that the data does not follow the specific distribution. The resulting test statistics will be used to determine whether H_0 is rejected or not. If the probability of observing the test statistic given that H_0 is true, is below the predefined significance level, then the H_0 is rejected. The p-value is the probability of observing a test statistic as extreme or more extreme than the one observed from the data. If the p-value is below our significance level, we reject the null hypothesis. We will be using a significance level of 5% in this thesis [3].

2.7.1 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (K-S) test is based on the finding the maximum distance D_n, between the theoretical cumulative distribution function $F_T(x)$ and the sample empirical distribution function $F_S(x)$. An important criteria on $F_T(x)$ is that it must be a continuous distribution. The test statistic D_n for a sample size n, is given by

$$D_n = \sup_x |F_S(x) - F_T(x)|. \tag{2.47}$$

The empirical distribution , $F_S(x)$ is defined as

$$F_S(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{x_i \leq x}, \tag{2.48}$$

where the sample is $S = \{x_1, ..., x_n\}$ from F_T. Since the theoretical cumulative distribution function has to be evaluated, one needs to choose parameters for the distribution of interest. One way of choosing parameters is to estimate them from the sample data. This distribution could for example be the generalized Pareto distribution, the lognormal distribution, or any other distribution that is of interest in reinsurance, [12].

The test statistic D_n is independent of what distribution is being tested, and follows a K-S distribution. The K-S distribution, according to [26], can loosely be defined as $F_n(x) = \mathrm{P}\left[D_n \leqslant x \mid H_0\right]$, for $x \in [0, 1]$. $F_n(x)$ is the cdf of D_n, and H_0 again, is the null hypothesis.

There exists multiple methods to calculate the exact K-S distribution, they are however slow and suffer from computational cancellations between very large terms, which makes the methods computationally costly for large n [26]. Kolmogorov-Smirnoff proved in 1933, that the K-S distribution can be approximated by the limiting form of the distribution function. The limiting form, taken from [22], is

$$\lim_{n\to\infty} \mathrm{P}\left(\sqrt{n}D_n \leq x\right) = L(x) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{i=1}^{\infty} e^{-(2i-1)^2\pi^2/\left(8x^2\right)}. \tag{2.49}$$

.

2.7.2 Anderson-Darling Test

The Anderson Darling (A-D) test, in contrast to the K-S test, gives more weight to the tails of a distribution, making it more suitable for extreme value distributions. The A-D test belongs to a class of quadratic empirical distribution functions, called EDF statistics, which refers to tests based on the empirical distribution function. In the test, the squared area between the theoretical distribution function $F_T(x)$ and the sample empirical distribution function $F_S(x)$, times a weight function is $w(x)$, is computed to give the test statistic A_n. More precisely, A_n is defined as

$$A_n = n\int_{-\infty}^{\infty} \left(F_S(x) - F_T(x)\right)^2 w(x)dx, \tag{2.50}$$

with n being the number of data points in the sample. The weight function is given by $w(x) = [F_T(x)(1 - F_T(x)]^{-1}$. In order to make computations easier, the formula can be rewritten as $A_n \approx -n - S$, in which S is defined as

$$S = \sum_{i=1}^{n} \frac{(2i-1)}{n}[lnF_T(Y_i) + ln(1 - F_T(Y_{n+1-i}))]. \tag{2.51}$$

Y_i represents the ordered data, with $i = 1, 2, ..., n$, and $Y_1 < Y_2 < ... < Y_n$. In this case it is assumed that the parameters of $F_T(x)$ are known. When the parameters are not known, the parameters need to be estimated for $F_T(x)$ and the formula for A_n needs to be modified according to the tables given in [34]. The reference does not contain tables for the GPD and Weibull distribution.

Similarly to the K-S distribution, finding the distribution for the test statistic A_n is also complicated. The limiting form of the A-D distribution, according to [23], was given by Anderson and Darling as

$$\lim_{n\to\infty} \mathrm{P}\left(A_n < x\right) = \frac{\sqrt{2\pi}}{x}\sum_{j=0}^{\infty}\begin{pmatrix} -\frac{1}{2} \\ j \end{pmatrix}(4j{+}1)e^{-(4j+1)^2\pi^2/(8x)}\int_0^{\infty} e^{\frac{x}{8(1+w^2)}-w^2(4j+1)^2\pi^2/(8x)}dw. \tag{2.52}$$

This, and more details can be found in [23] and [2] and [34]

2.8 Parameter Determination

Maximum Likelihood (MLE) and the Least Squares Model (LSQ) are both common methods to determine the parameters in a probability distribution. In this thesis, both methods will be used and compared in order to estimate the parameters in the Pareto distribution.

2.8.1 Maximum Likelihood Method

In order to determine the MLE, we let X_i where $i = 1, \ldots, n$ are n random variables which are independent and have a joint probability likelihood function,

$$L(x, \theta) = \prod_{i=1}^{n} f_i(x_i; \theta) = f_i(x_1, \ldots, x_n | \theta), \tag{2.53}$$

where θ denotes some unknown parameter with the parameter space Ω_θ in the distribution function. Then the logarithm of the likelihood function is:

$$log(L(x, \theta)) = \sum_{i=1}^{n} log(f(x_i | \theta)). \tag{2.54}$$

The logarithmic likelihood function is a functions which is increasing because of the sum. Therefore, in order to calculate the desired parameters to be used in the Pareto distribution, we maximize the logarithmic likelihood function.

2.8.2 Method of Least Squares

The Least Squares Method (LSQ) is a method that minimizes the sum of squared residuals. This means that it squares the difference between an observed value and a predicted value which is given by a model of choice. It fits a probability distribution to a given data set.

We let $X_1, \ldots, X_n$ again denote n independent stochastic variables which have the outcomes $x_1, \ldots, x_n$. We assume that we know the means of X_i, where $i = 1, \ldots, n$, so we define $E(X_i) = \mu_i(\theta)$ and $V(X_i) = \sigma^2$ where θ is an unknown parameter in Ω_θ.

Then the least squares estimation for θ is given by the $\hat{\theta}$ that minimizes

$$Q(\theta) = \sum_{i=1}^{n} [x_i - \mu_i(\theta)]^2. \tag{2.55}$$

So if all the functions $\mu_i(\theta)$ are identical then we can take the derivative of $Q(\theta)$ with respect to θ and get:

$$\frac{dQ(\theta)}{d\theta} = -2\mu'(\theta) \sum_{1}^{n} [x_i - \mu(\theta)], \tag{2.56}$$

which shows that Q is minimized if $\mu(\theta) = \bar{x}$, which gives the estimate $\hat{\theta} = \mu^{-1}(\bar{x})$ [3].

LSQ, is a method that can be used to help decide some of the parameters in the Pareto distribution [3].

2.9 Method Evaluation

One goal of the thesis is to evaluate LF's current pricing model, to see if the risk premiums determined from it seem reasonable or not. In order to do so, we will create distributions of bootstrap estimates and compute the empirical percentiles from these bootstrapped distributions. This will allow us to estimate the degree of uncertainty in our risk premium estimates. Due to inaccurate results, the results of the methods introduced in Section 2.9.3 and in Section 2.9.4, have not been included in the results, and can be found in Appendix A.

2.9.1 Bootstrapping

Bootstrapping is a resampling method that can be used when the data is sparse. It also allows us to estimate the statistical uncertainty in our parameter estimates. The variant of bootstrap used in this thesis is non-parametric bootstrap. It makes no assumptions about the underlying distribution of the sample, and it creates bootstrap samples by drawing with replacement from the original sample. The procedure described in [17] is as follows.

We assume that we have an observed sample of data $\mathbf{x} = (x_1, ..., x_n)^T$ with $x_i \overset{\text{i.i.d.}}{\sim} F$, for some distribution function F. We also suppose that $\theta = t(F)$ is the parameter of interest, and that $\hat{\theta} = s(F)$ is the test statistic that will be used to estimate θ.

1.) To begin, sample x_i^* from $x_1, ..., x_n$ with replacement for $i = 1, ..., n$. It is assumed that each x_i^* is sampled with equal probability from $\mathbf{x}$, $P(x_i^* = x_j) = 1/n$ for all i, j.

2.) From this, $\hat{\theta}^* = s(\mathbf{x}^*)$ can be calculated in which $\mathbf{x}^* = (x_1^*, ..., x_n^*)^T$ is the resampled data.

3.) Repeat the steps a total of M times in order to form the bootstrap distribution of $\hat{\theta}$.

The resulting bootstrap distribution will be $(\hat{\theta}, \hat{\theta}_1^*, \ldots, \hat{\theta}_n^*)$, which consists of M estimates of θ^* and the original estimate $\hat{\theta}$. This distribution can be used to calculate properties of $\hat{\theta}$, such as estimating the mean and empirical percentiles for $\hat{\theta}$.

2.9.2 Confidence Intervals

A confidence region can be defined as a region of space such that the probability that the true value of a parameter lies in that region is equal to some predefined level. In this thesis, we are looking at confidence intervals for one-dimensional parameters, that take on values on the real number line. Confidence intervals are useful since they can give an estimate of the uncertainty in a parameter estimation. In the case of this thesis, the confidence intervals would give an uncertainty estimate of the risk premiums in each layer. The general definition of a confidence interval is stated by [10] and is as follows.

Suppose a data set $x_1, \ldots, x_n$ is given, modeled as realization of random variables $X_1, \ldots, X_n$. Let θ be the parameter of interest, and γ a number between 0 and 1 . If there exist sample statistics $L_n = g(X_1, \ldots, X_n)$ and $U_n = h(X_1, \ldots, X_n)$ such that

$$\mathrm{P}(L_n < \theta < U_n) = \gamma, \tag{2.57}$$

for every value of θ. Then (l_n, u_n), where $l_n = g(x_1, \ldots, x_n)$ and $u_n = h(x_1, \ldots, x_n)$, is called a $100\gamma\%$ confidence interval for θ. The number γ is called the confidence level.

2.9.3 Basic Bootstrap Confidence Intervals

One method of creating confidence intervals is via the basic bootstrap confidence intervals. As mentioned earlier, bootstrapping will allow the creation of a bootstrapped distribution $(\hat{\theta}, \hat{\theta}_1^*, \ldots, \hat{\theta}_n^*)$. One metric that provides an estimate of the uncertainty in the parameter estimate $\hat{\theta}$ is the empirical distribution of the bootstrap statistics themselves. For a large value of M, the 2.5 percentiles and 97.5 percentiles of the bootstrapped statistic will approximate the percentiles of the underlying distribution. This would result in an empirical 95 percent confidence interval. We call this method the empirical percentiles in this thesis. However, the more commonly used way of computing a 95 percent confidence interval using the percentiles of the bootstrapped statistics is a method called basic bootstrapped confidence intervals. According to [9], let $L_n = 2\hat{\theta} - \hat{\theta}^*_{0.975}$ and let $U_n = 2\hat{\theta} - \hat{\theta}^*_{0.025}$, where $\hat{\theta}^*_{0.975}$ is the 97.5^{th} bootstrapped percentile, and $\hat{\theta}^*_{0.025}$ is the 2.5^{th} bootstrapped percentile. Then the interval (L_n, U_n) is the basic bootstrap interval.

2.9.4 Pivot Confidence Intervals

One way method of constructing confidence interval is by using Pivots. A pivot, can according to [25] be defined as follows. Let a pivot $Z(\theta) = Z(\theta, X_1, ..., X_n)$ be a function of the sample X_1,...X_n and an unknown parameter θ and has a completely known probability distribution. Common pivots are the Z-test or the student's t-test. student's t-test is useful for smaller sample sizes, while the Z-test holds when $n \to\infty$. From the pivot $Z(\theta)$, one can obtain confidence interval for set fixed confidence level γ. From the pivot $Z(\theta)$, and its distribution, select two constants c_1 and c_2 such that

$$\mathrm{P}\left(c_1 < Z(\theta) < c_2\right) = \gamma. \tag{2.58}$$

From which solving for θ in the inequalities $Z(\theta) < c_2$ and $Z(\theta) > c_1$, we obtain equivalent inequalities in the form $\theta < U_n$ and $\theta > L_n$.

Chapter 3

Method

This section is split into three parts. First, the software used for this thesis is presented. Thereafter, the data handled in this thesis is introduced and explained. Lastly, LF's current pricing model is explained.

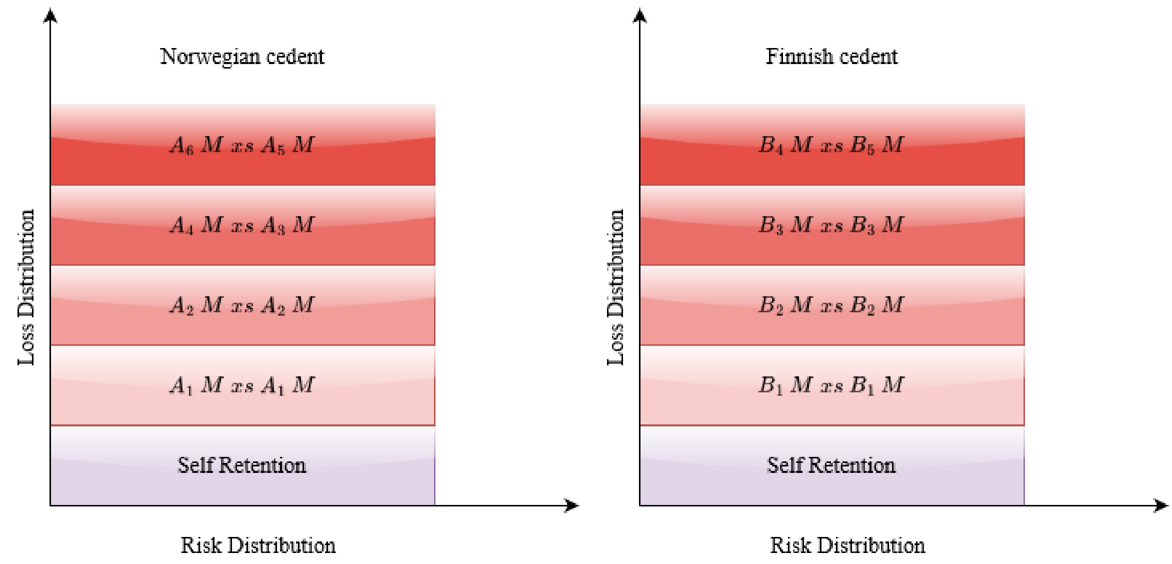
3.1 Software

The main softwares used in this thesis are MATLAB R2022b, Microsoft Excel, R, Lime Easy and QlikView. MATLAB has been used to simulate data. Excel has been used as a user friendly front for the MATLAB code. Lime Easy is a CRM (Customer Relationship Management) system and management application used at LF. All information that passes through the International Reinsurance department at LF is saved in Lime Easy. QlikView is a data visualization system where all active, canceled and expired affairs are visually presented. It is mainly in QlikView the required information for this thesis is found. If more elaborate information is required, Lime Easy is the application one can find it in.

R will be used for statistical programming. It is a useful program that has built in packages specifically that can be used in reinsurance. Data will be imported to R in excel format or as text files. R will also be used to visualize the data.

3.2 Data

We will be working with two different data sets, containing registered claims, provided by LF, one is from a Norwegian insurance company and one from a Finnish insurance company. Both data sets pertain to XL reinsurance contracts related to fire losses. Both companies insure commercial and private properties. For privacy reasons the exact layer structures have been pseudonymized with letters, as shown in Figure 3.1.



(a) Layer structure of the Norwegian cedent (b) Layer structure of the Finnish cedent

Figure 3.1: The figures show the layer structures for both cedents, A_i and B_j represent numerical values, where $i = 1 - 6$ and $j = 1 - 5$.

3.2.1 Experience Data

The data provided from the insurance companies contains information about the claims, such as what year the loss incurred, how much it cost, what area or property that was damaged, and what type of damage the loss involved, etc. Both the insurance companies also provide the proposed layers that they find suitable to divide their incurred losses into. The data from the Norwegian insurance company contains 162 losses incurred between the years 1999-2022, and the Finnish insurance company contains 278 losses from the years 1995-2022. In order to be able to relate and use the losses together, inflation need to be accounted for. This can either be done by indexing up the losses based on the prior risk premium development, or else by indexing up using the inflation rate today of the country that that cedent is operating in. For this thesis, indexing based on inflation was used.

To visually get a better idea of what the data looks like, both the data sets were graphed as histograms as seen in Figure 3.2, and 3.3. The claim costs are in millions of Euro. What can be seen from both the data sets, is that the majority of the data fall within the first few bars, and with a small amount of data on the bars to the right in the graphs, representing extremely high costs. Both the histograms, because of these extremely large claims, form a tail-like shape. Next, data sets were plotted as box-plots, in order to to better able to compare the spread of the data between the two data sets. This can be seen in Figure 3.4. We see that the Finnish data set has more extreme values than the Norwegian data set, but that the extreme values in the Norwegian data set are more spread apart from each other. The Finnish data set has a more defined tail than the Norwegian data set.

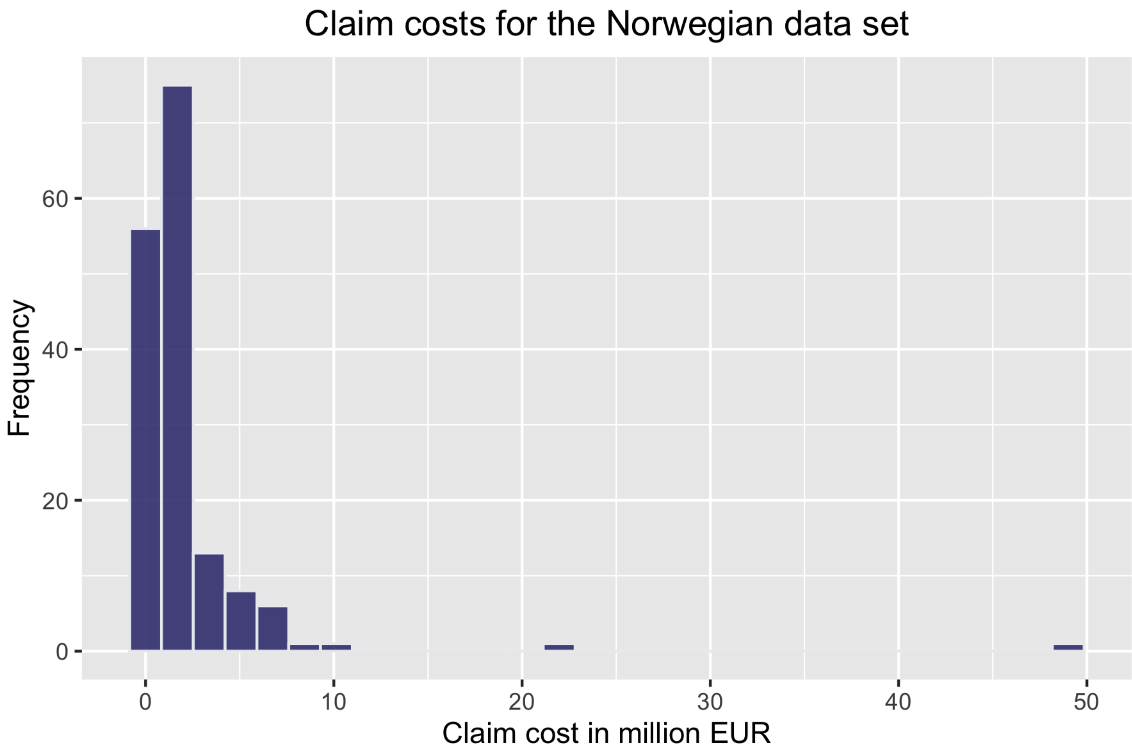


Figure 3.2: Histogram of the claim costs for the Norwegian data set, for the years 1999-2022.

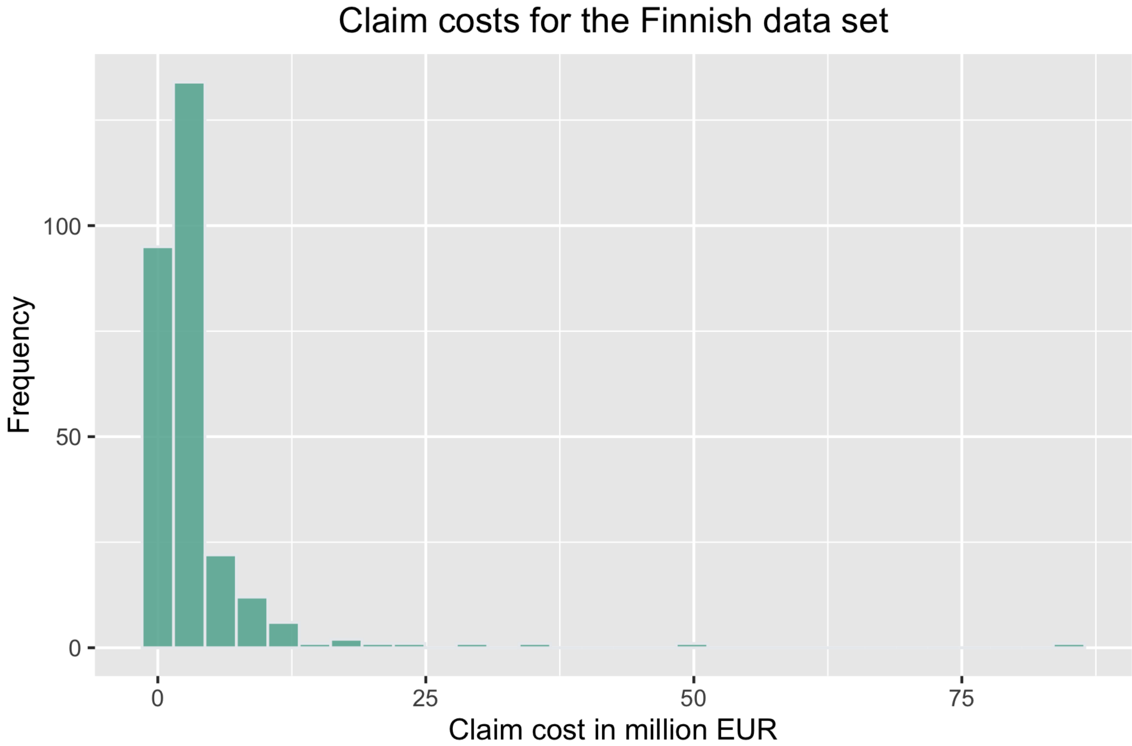


Figure 3.3: Histogram of the claim costs for the Finnish data set, for the years 1995-2022.

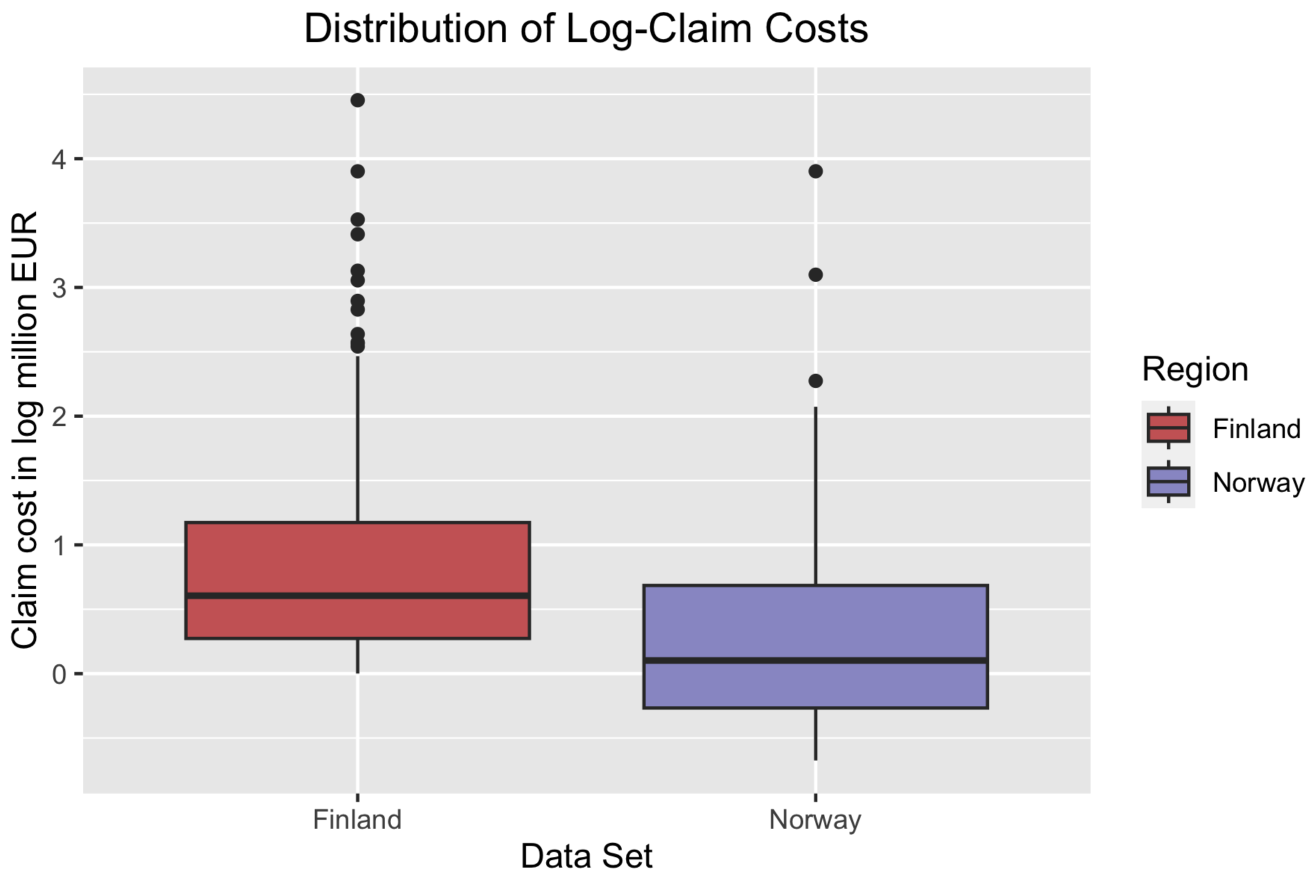


Figure 3.4: Boxplots of the claim costs for both the data sets in EUR.

Note that the data in the Norwegian data set was originally presented in NOK and was therefore converted into Euro in order for us to be able to compare the two data sets. The Finnish data set was already in Euro. The conversions were based on the exchange rate for NOK to EUR from December 31st for every years claim cost. The rates were found from [30].

3.3 LF's Current Pricing Model

To begin the analysis, both the data sets were analyzed using the conditions and procedures that LF already has in place to determine their risk premium. This in order to see how big of a difference parameter estimation methods, and changing distribution models will have on the data sets. We want to see whether these changes lead to a better fit for the data and how much the risk premium ultimately is effected by these changes. Analyzing the data sets with LF's current standards and practices will act as a baseline for comparison for the methods we will utilize in this analysis. The goal will be to improve their pricing model for their risk premiums.
Initially, from the data from the two insurance companies, the expected value $E[N]$ for

the frequency distribution, can be determined by dividing the number of losses over the time in years that they occurred over. The frequency distribution is modeled by the Poisson distribution. The severity of the claims X, is modeled using the generalized Pareto distribution, from which the expected value $E[X]$ is calculated. The parameters ξ and σ in the GPD are determined using MLE. LF also occasionally uses LSQ for parameter estimation. LF currently sets the location parameter θ without using a theoretically backed method. They use the first layer in the XL contract as a guideline for choosing the threshold. By multiplying $E[N]$ from the Poisson distribution, with $E[X]$ from the GPD, and dividing by the product by the duration of the contract, the risk premium can be calculated. To determine the risk premium for each layer, the risk premium is divided by the given limit of each layer in the contract.

3.3.1 Fair Price

The goal of the project is to find the the most accurate representation of the fair price. The meaning of fair price in this context, is the price of the risk that will cover the total cost of the expected losses. The risk premium or the price is calculated based on the common POT-model used to model extreme values.

A method to evaluate the risk premiums, is to use bootstrapping. By using bootstrapping, we will generate 1000 new samples for each of the data sets of the same size as the Norwegian and the Finnish data sets. We will create distributions of bootstrap estimates and then compute the empirical percentiles for these bootstrapped distributions. This will allow us to estimate the degree of uncertainty in our risk premium estimates, and by this be able to evaluate if LF's current pricing model yields a fair price or not.

Chapter 4

Results

In this section the results for all combination of the models are presented with and without modifications.

4.1 Determining Suitable Frequency Distributions

To begin, the variance and expected value for the Norwegian and Finnish data sets were calculated for the negative binomial distribution according to Equations 2.9, and 2.10. They were calculated in order to determine, in addition to the Poisson distribution, to see if also the NB distribution could be used to model the frequency of losses. The table below gives the calculated variance, and mean for each data set for the NB distribution, as well as the conditions on the variance and mean that are needed for the NB distribution to be suitable, according to the theory in Section 2.4.2. Table 4.1 and 4.3 have also been included below, displaying what the mean and variance are for each data set.

Norwegian Data Set

Parameters	**Norwegian data set**
$\mu = \sigma^2$	4.137

Table 4.1: The mean and variance of the Poisson distribution for the Norwegian data set.

Parameters	**Norwegian data set**
r	7.3289
p	0.63919
μ	4.1370
σ^2	6.4723
$\sigma^2 > \mu$	Yes
Suitable distribution	Negative Binomial

Table 4.2: Determination of what distribution is the most suitable for the Norwegian data set.

Here the mean has been calculated using Equation 2.12, which represents the mean defined for the alternative definition of the negative binomial model. In Table 4.2 one can see that $r < \infty$, which means that the mean does not exceed the variance. Therefore the negative binomial distribution is a suitable distribution to model the frequency of losses.

Since the the frequency part of the frequency-severity rating is estimated with the help of the Method of Moments, as defined in Section 2.4.3, the Poisson and the negative binomial distribution are not effected by the MLE and LSQ methods that are only used when modeling the severity of claims.

Finnish Data Set

Parameters	**Finnish data set**
$\mu = \sigma^2$	1.8082

Table 4.3: The mean and variance of the Poisson distribution for the Finnish data set.

Parameters	**Finnish data set**
r	3.0803
p	0.63012
μ	1.8081
σ^2	2.8695
$\sigma^2 > \mu$	Yes
Suitable distribution	Negative Binomial

Table 4.4: Determination of what distribution is the most suitable for the Finnish data set.

In Table 4.2 and 4.4 r and p have been calculated with the Method of Moments as defined in Section 2.4.3, and from that the variance and mean have been calculated analytically according to Equations (2.9) and (2.10).

Once again as for the Norwegian data set, one can see that the mean in Table 4.4 has been calculated using Equation (2.12).

In Table 4.4 one can once again see that $r < \infty$, which means that the mean does not exceed the variance. The same conclusion can be drawn for the Finnish data set as the Norwegian data set, which is that the negative binomial distribution is a suitable distribution to use for modeling the frequency.

It is seen from Tables 4.2 and 4.4, that the negative binomial distribution is suitable to model the frequency of losses for both the data sets since the variance does not exceed the mean. Both data sets will still however be analyzed using the Poisson distribution as well as the NB distribution for further comparison.

4.2 Goodness of Fit Tests for the Severity Distribution

LF currently uses the generalized Pareto distribution to model the severity of claims, which literature studies on modeling severity distributions in reinsurance agrees upon [20],[14], [5], and [1]. Despite this, it is still a good idea to analyze the data against other distributions to verify whether this is the most suitable distribution for both the data sets. Anderson-Darling test and the Kolmogorov-Smirnov test are therefore run in R to analyze what distributions are the most suitable for the data sets.

When running the K-S and A-D tests, we first estimate the parameters that will be used in the tests based on the data that we have. This is needed since we cannot assume that our data is centered around zero. Therefore, for all the K-S and the A-D tests, we first estimate the parameters of the distribution we are testing against, using the data sets. For example, we will use MLE to estimate the parameters for the cumulative distribution for the GPD from our data. The shape,scale and location parameter was approximated with MLE for these tests. Then the tests are run using our data with that cumulative distribution function.

4.2.1 Anderson-Darling Test

The Anderson-Darling test was run in R using *ad.test* to test if the two data sets fit the normal distribution, the lognormal distribution, the Weibull distribution, the generalized Pareto distribution and the Pareto type 1 distribution. The null hypothesis H_0 for each of the tests was: the data follows the specific distribution. The results showed that for the Norwegian data set, H_0 was rejected for all the distributions except for PD1 and GPD at a 5% significance level. As seen from Table 4.5, the p-value for the PD1 is larger than for the GPD. As mentioned, the PD1 is a more specific form of the GPD, and should therefore more likely have a lower p-value than the GPD. A potential reason for this result could be due to the Norweigan data set being fairly small, with only 162 data points, which could effect the results of the test. Also as seen from the boxplot in Figure 3.4, the Norwegian data set has a less defined tail than the Finnish data set, making it possibly more challenging to fit for a Pareto distribution in general.

For the Finnish data set, we see that both results for the PD1 and GPD are significant, and we do not reject H_0 for either distributions. The p-value for the GPD is significantly larger than for the PD1, which means that the data likely follows a GPD.

Norwegian Data Set

Anderson-Darling Test			
Distribution tested for	**Test statistic (A_n)**	**p-value**	**Conclusion**
Normal distribution	33.912	<2.2e-16	Reject H_0
Lognormal distribution	5.8249	2.129e-14	Reject H_0
Weibull distribution	13.185	3.704e-06	Reject H_0
PD1	0.36964	0.8781	Do not reject H_0
GPD	0.41532	0.8332	Do not reject H_0

Table 4.5: Anderson-Darling test tested for five different distributions.

Finnish Data Set

Anderson-Darling Test			
Distribution tested for	**Test statistic (A_n)**	**p-value**	**Conclusion**
Normal distribution	53.584	<2.2e-16	Reject H_0
Lognormal distribution	10.409	2.2e-16	Reject H_0
Weibull distribution	22.806	2.158e-06	Reject H_0
PD1	0.37248	0.8754	Do not reject H_0
GPD	0.22216	0.9831	Do not reject H_0

Table 4.6: Anderson-Darling test tested for five different distributions.

4.2.2 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test was run in R using *ks.test*. The results of the Kolmogorov-Smirnov test can be seen in Tables 4.7 and 4.8. The test statistic D_n is the distance between the empirical distribution function for our data set and the cumulative distribution function, of the reference distribution we are testing for. This means that the smaller the value is for D_n the better the empirical distribution matches the reference distribution. For the Kolmogorov-Smirnov test, the p-values for each distribution is compared to a significance level of 5%. The results for the K-S test shows that for the Norwegian data set, both for PD1 and GPD, we fail to reject H_0, meaning that the data set, according to the test, likely fits a GPD or a PD1. We can see from Table 4.7 that, similarly to the A-D test, the PD1 gives a larger p-value than the GPD, despite it being a more specific form of the GPD. Again, this could be due to the data set being small and not being as well defined in the tail as the Finnish data set.

For the Finnish data set, as seen in Table 4.8, we did not reject H_0 for PD1 and GPD. The GPD has a larger p-value than the PD1, indicating that the GPD is the most suitable distribution for the data set.

LF currently uses the GPD for the severity of claims, which both the A-D and K-S gave a large p-value for. Therefore, also since the GPD is a more general model than the PD1, we conclude that the GPD is in general the most suitable distribution to model the severity of claims for XL contracts in reinsurance. All future analysis will therefore include the GPD to model the severity of claims.

Norwegian Data Set

Kolmogorov-Smirnov Test			
Distribution tested for	**Test statistic (D_n)**	**p-value**	**Conclusion**
Normal distribution	0.36824	$<$ 2.2e-16	Reject H_0
Lognormal distribution	0.13781	0.004252	Reject H_0
Weibull distribution	0.252	2.318e-09	Reject H_0
PD1	0.039635	0.9610	Do not reject H_0
GPD	0.044061	0.9116	Do not reject H_0

Table 4.7: Kolmogorov-Smirnov test, tested for five different distributions.

Finnish Data Set

Kolmogorov-Smirnov Test			
Distribution tested for	**Test statistic (D_n)**	**p-value**	**Conclusion**
Normal distribution	0.35052	$<$2.2e-16	Reject H_0
Lognormal distribution	0.14003	3.684e-05	Reject H_0
Weibull distribution	0.24794	2.887e-15	Reject H_0
PD1	0.030404	0.9041	Do not reject H_0
GPD	0.029024	0.9733	Do not reject H_0

Table 4.8: Kolmogorov-Smirnov test, tested for five different distributions.

4.3 Analysis of the Severity Distribution

To begin, we looked at the effects of determining the parameters for the the GPD using MLE and LSQ. As mentioned earlier, LF mainly uses MLE for parameter estimation for the GPD, but occasionally also uses LSQ. We will therefore compare both methods. The parameters of interest is the shape parameter ξ and the scale parameter σ. The location parameter θ is set manually by LF, and is therefore not approximated by either of the methods. ξ and σ will be estimated with both MLE and LSQ. We will also look at what the effects of changing the frequency distribution has on the risk premium. In Section 4.1 it was determined that the negative binomial distribution is a suitable replacement for the Poisson distribution for modeling the loss frequency. Therefore, we will use both distributions to calculate the risk premiums.

4.3.1 Parameter Determination for the Generalized Pareto Distribution Using MLE and LSQ

The Tables 4.9, and 4.10 show the estimates for the parameters ξ, and σ for the GPD, when using MLE and LSQ for the Norwegian and the Finnish data set. As mentioned θ, is set manually by LF, and is not estimated by MLE and LSQ. the parameter θ is still included in the tables so that all the parameters of the GPD are included.
The estimates for the shape parameter ξ, and the scale parameter σ are significantly different depending on if MLE or LSQ is used. This is particularly true for the Norwegian data set. LSQ can be beneficial when the distribution is not fully known, but the method does however not perform as well as MLE for small data sets in general. This is potentially why we see larger differences in parameter estimation between MLE and LSQ for the Norwegian data set.

Norwegian Data Set

Generalized Pareto Distribution			
Parameter	**MLE**	**LSQ**	$\%\Delta$
θ **(Set manually)**	12.5 M	12.5 M	0.0
σ	12.45 M	9.18 M	-26.27
ξ	0.6043	0.7360	21.794

Table 4.9: Parameter comparison between using MLE and LSQ for the GPD for the Norwegian data set. The percentage change, $\%\Delta$, of changing the method of parameter estimation to LSQ from MLE was calculated.

Finnish Data Set

Generalized Pareto Distribution			
Parameter	**MLE**	**LSQ**	%Δ
θ **(Set manually)**	6.25 M	6.25 M	0.00
σ	7.21 M	7.01 M	-2.77
ξ	0.38407	0.43799	14.039

Table 4.10: Parameter comparison between using MLE and LSQ for the GPD for the Finnish data set. The percentage change, %Δ, of changing the method of parameter estimation to LSQ from MLE was calculated.

4.3.2 Comparison of Risk Premium

The next step in this thesis was to see what the effects would be of changing the parameter estimation method and the distribution to model the frequency of claims would have on the risk premiums. All calculations have therefore been compared to LF's current pricing model of using MLE for parameter estimation for the GPD, and using the Poisson distribution for the frequency of claims. The percentage change of changing methods from MLE to LSQ was calculated in Tables 4.11 and 4.13 but keeping the Poisson distribution to model the frequency. This was followed by changing the frequency distribution from the Poisson distribution to the negative binonmial distribution, using MLE and LSQ for parameter estimation for the GPD. Again the percentage change compared to LF's current pricing model was calculated for each change implemented. The results can be seen in Tables 4.12, and 4.12. For each of the tables **LF RP** represents LF's risk premiums determined from their current pricing model. The risk premiums are calculated according to the description in Section 3.3. In the tables, **L1**, **L2**, **L3** and **L4** represent the different layers of the contracts for the Norwegian and the Finnish data set. **L1** is the smallest layer with the lowest limit, followed by **L2** and so on. The Figures 4.1 and 4.2 show the GPD with MLE and LSQ for both the data sets.

The results show that for both the data sets, using LSQ results in significantly larger risk premiums in the highest layer, **L4** of the contracts. This can only be seen slightly from the graphs. We therefore conclude that changing parameter estimation method from MLE to LSQ does have significant effects on the risk premiums. MLE is more useful when the underlying distribution is known, and the method is more precise than LSQ. Also, LSQ is preferable when there is less information about the data sets, and for larger data sets in general. Therefore, since LSQ gave significantly different values for the risk premiums than MLE, we determine MLE to be more suitable for the data sets.

We compared the effects of changing the frequency distribution from the Poisson distribution to the negative binomial distribution. Initially while still using MLE for parameter estimation of the GPD, and then also with changing the parameter estimation method to LSQ as well. This was compared to LF's current risk premiums. When only changing frequency distribution, we only see small changes in the risk premiums for both the data sets. Overall, the negative binomial distribution gives a better fit to the frequency distribution than the Poisson distribution, since the constraint that the expectation is the same as the variance does not have to be met. Therefore we conclude that if the negative binomial distribution is suitable for a data set, according to the theory in Section 2.4.2, then it would be able to price the risk premiums more fairly than the Poisson distribution. When changing also the parameter estimation method to LSQ, we see significant changes in the risk premium compared to LF's current risk premium.

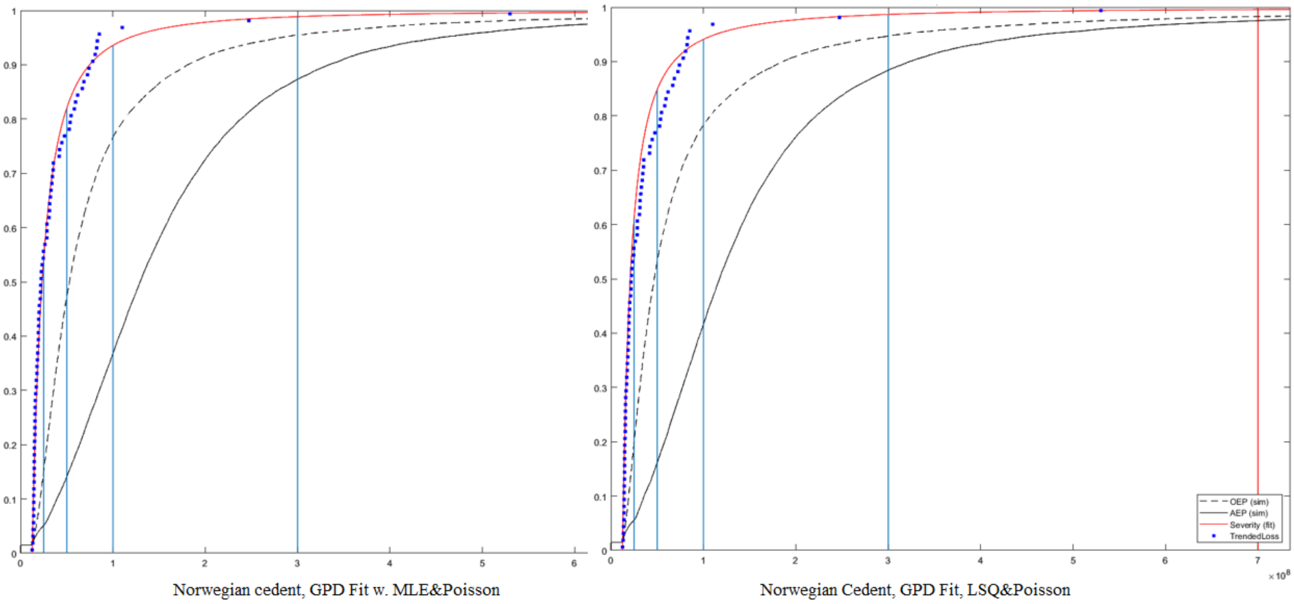
Norwegian Data Set

Risk Premiums with LSQ			
Layer	**LF RP**	**RP (LSQ)**	%Δ
L1	70.715	64.212	-9.196
L2	36.323	32.418	-10.751
L3	10.465	10.626	1.538
L4	2.319	2.970	28.072

Table 4.11: Risk premium calculated from the GPD with the Poisson distribution using MLE and LSQ for the Norwegian data set. The percentage change of using LSQ instead of MLE, as in LF's current pricing model, was calculated for comparison. RP is an abbreviation for the risk premium. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Risk Premiums with NB and MLE, LSQ					
Layer	**LF RP**	**RP(NB,MLE)**	%Δ	**RP(NB,LSQ)**	%Δ
L1	70.715	67.696	-4.269	61.621	-12.860
L2	36.323	35.335	-2.720	31.588	-13.036
L3	10.465	10.471	0.057	10.625	1.529
L4	2.319	2.322	0.129	3.000	29.366

Table 4.12: Risk premium calculated with the GPD and the negative binomial distribution using MLE and LSQ for the Norwegian data set. The percentage change of using LSQ instead of MLE, as in LF's current pricing model was calculated for comparison. The percentage change of using the NB to model the frequency as well as using LSQ for parameter estimation compared to LF's current pricing model, was also calculated. RP is an abbreviation for the risk premium. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.



(a) GPD Fit with MLE (b) GPD Fit with LSQ

Figure 4.1: GPD fit for the Norwegian data set. In the graphs, the blue dots represent the data points (claims), and the red line is the generalized Pareto curve.

In the Figure 4.1, one can see the data points and their fit to the Pareto curve, i.e. the empirical and theoretical CDFs. It is also possible to see how the data is spread out and any outliers there may be in the data set. In the same figure the x-axis represents the claims and the y-axis represents the probability $P(X \leq x)$, which refers to the probability of an event loss being less than or equal to x.

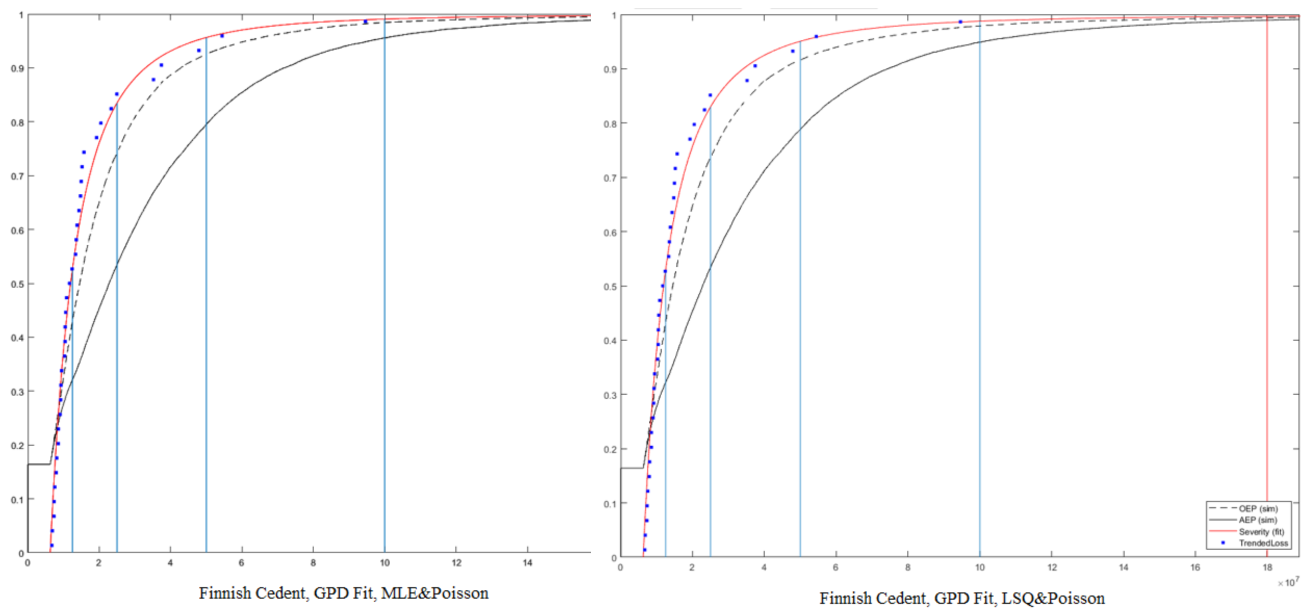
Finnish Data Set

Risk Premiums with LSQ			
Layer	**LF RP**	**RP(LSQ)**	%Δ
L1	40.250	40.464	0.532
L2	14.365	15.299	6.502
L3	3.647	4.375	19.962
L4	0.822	1.137	38.321

Table 4.13: Risk premium calculated from the GPD with the Poisson distribution using MLE and LSQ for the Finnish data set.The percentage change of using LSQ instead of MLE, as in LF's current pricing model, was calculated for comparison. RP is an abbreviation for the risk premium. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Risk Premiums with NB and MLE, LSQ					
Layer	**LF RP**	**RP(NB,MLE)**	%Δ	**RP(NB,LSQ)**	%Δ
L1	40.250	38.554	-4.214	38.746	-3.737
L2	14.365	14.466	0.703	15.375	7.031
L3	3.647	3.621	-0.713	4.354	19.386
L4	0.822	0.803	-2.311	1.132	37.713

Table 4.14: Risk premium calculated with the GPD and the negative binomial distribution using MLE and LSQ for the Finnish data set. The percentage change of using LSQ instead of MLE, as in LF's current pricing model was calculated for comparison. The percentage change of using the NB to model the frequency as well as using LSQ for parameter estimation compared to LF's current pricing model, was also calculated. RP is an abbreviation for the risk premium. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.



(a) GPD Fit with MLE (b) GPD Fit with LSQ

Figure 4.2: GPD fit for the Finnish data set. In the graphs, the blue dots represent the data points (claims), and the red line is the generalized Pareto curve.

As explained for Figure 4.1, one also in Figure 4.2 sees the fit of the Pareto curve, where the majority of data points lay and the outlook of the data set. The x-axis here, as previously mentioned for the Pareto plot, represents the claims and the y-axis represents the probability $P(X \leq x)$, which means the probability of an event loss being less than or equal to x.

4.4 Hill Estimator and its Effect on the Risk Premium

In this section the results for implementing the Hill estimator and deciding the optimal threshold u for the data sets are presented and elaborated on.

4.4.1 Hill Estimator Determination

In order to choose a tail optimized threshold for our data set we will use the Hill estimator. In LF's current pricing model, the threshold has been set by choosing a value for the threshold which is not too high, nor too low based on the excess point in the first layer, **L1**. A midpoint was therefore chosen, since choosing an excess point or anything larger than the excess point as the threshold, would result in loosing to much frequency in the tail of the distribution. Thus, choosing a threshold that is mathematically motivated will always be better than an arbitrarily chosen threshold. In order to optimally choose this threshold one can use a Hill estimator which is known to improve the tail behavior in distributions, such as the GPD. The Hill estimator $\hat{\alpha}$ is calculated as a function of the tail parameter k based on a given data set. It is important to choose a k where $\hat{\alpha}$ is not too high or too low. Then, $\hat{\alpha}$ will converge to a constant as $k \to \infty$, and will be too high if k is too close to origin. Using *hill.kopt* in *R* we can receive an optimal k as the output and find the respective claim in the data set which will be the new threshold input. We can also find the $\hat{\alpha}$ that corresponds to that k and from that calculate the shape parameter in GPD [18].

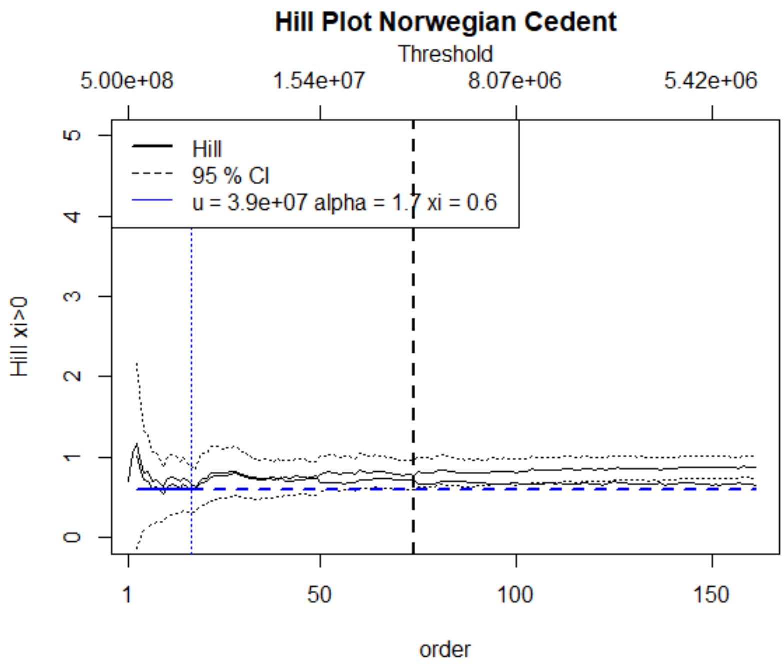


Figure 4.3: Hill Plot for finding the optimal threshold (Norwegian Cedent).

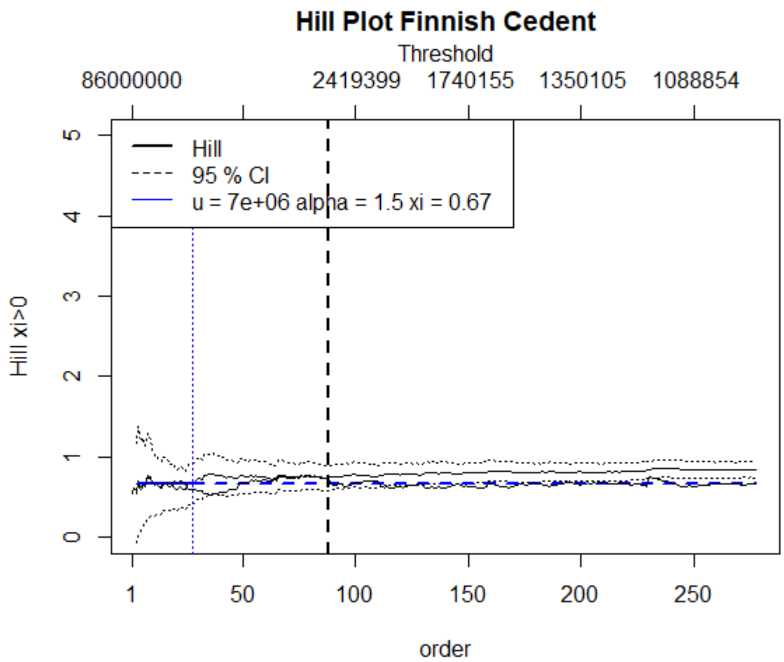


Figure 4.4: Hill Plot for finding the optimal threshold (Finnish Cedent).

The Hill plots determines the Hill estimator for $k \geq 1$, in which the blue dotted vertical line represents the 90th percentile of the claim costs. The small, black, dotted lines in Figures 4.3 and 4.4 represent the $100(1-\alpha)\%$ symmetric normal confidence interval, where $\alpha = 0.05$. The black line in between the CI is the HE and the optimal HE. The thicker dotted lines represent at what k the optimal threshold u is. The estimated threshold u at each tail parameter $\hat{\alpha}$ are plotted by the blue thicker horizontal dashed line for all higher thresholds. It is good to mention that since the HE is similar to the thicker blue horizontal, dashed line, it is possible to chose a lower threshold and still have reasonable values for the remaining GPD parameters [33].

From using *hill.kopt* in R, we get that the optimal threshold u, for the Norwegian data set, is obtained at $k = 74$ and $\hat{\alpha} = 0.7220632$. The corresponding data point is the minimum loss $u = 10065184$ which is a bit lower than LF's currently chosen threshold. It is as expected, still below where the cover starts/the first excess point is set, but should in theory give a better price than it. We can see in Figure 4.3 that the threshold u lies at a reasonable corresponding k.

For the Finnish data set, we determine the optimal threshold u to be at $k = 88$, where $\hat{\alpha} = 0.7337153$. This k corresponds to the data point $u = 2419399$, which is lower than the original threshold set by LF. We can see in Figure 4.4 that the threshold u lies at a reasonable corresponding k.

4.4.2 Comparison of Risk Premium with and without the Hill Estimator

The location parameter θ was determined with the Hill estimator, and used in calculating the risk premiums. These risk premiums have been summarized in the Tables 4.15, 4.16, 4.17, and 4.18. The results have been compared to LF's risk premiums determined by their current pricing model. The percentage change of using the Hill estimator with MLE, LSQ as well as for both frequency distributions have been calculated. The results show that the largest change in the risk premiums of using the Hill estimator occurs in the highest layer **L4** for both the data sets. This is regardless if MLE or LSQ is used, and regardless of choice of frequency distribution. Figures 4.5, and 4.6 show the GDP with using MLE and LSQ with the new location parameter. If comparing these figures to Figure 4.1 and 4.2, it is difficult to visually see the changes of using the Hill estimator, despite them being significant.

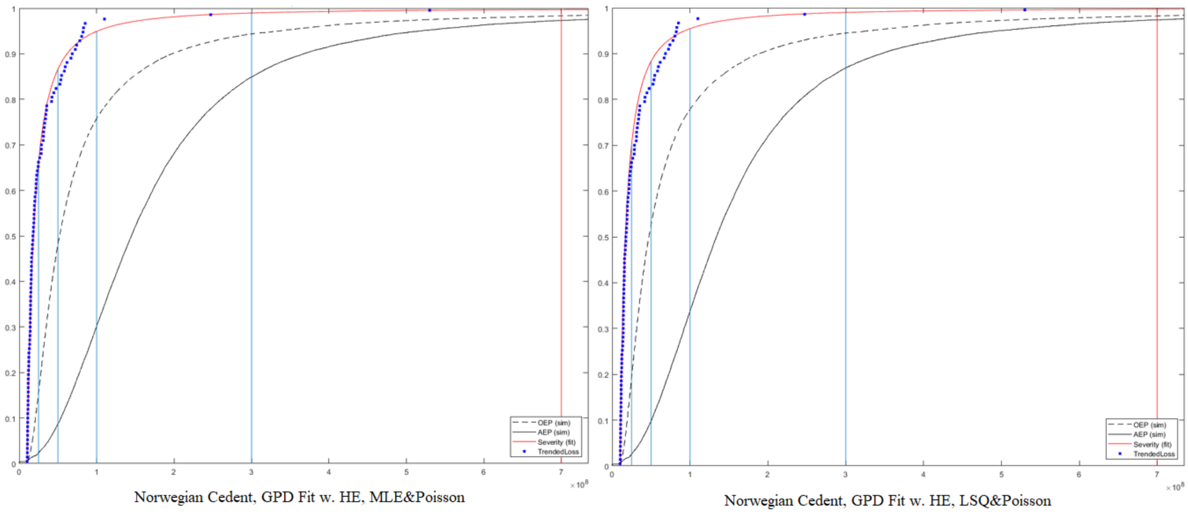
Norwegian Data Set

Risk Premiums with Hill estimator, MLE and LSQ					
Layer	**LF RP**	**RP(MLE,HE)**	%Δ	**RP (LSQ,HE)**	%Δ
L1	70.715	69.663	-1.488	64.687	-8.524
L2	36.323	36.317	-0.166	32.809	-9.674
L3	10.465	11.637	11.199	10.871	3.880
L4	2.319	3.044	31.263	3.078	32.730

Table 4.15: Comparison of the the risk premium for the different layers using MLE and LSQ, with the Hill estimator. The risk premiums have been calculated using the GPD and the Poisson distribution. RP is an abbreviation for the risk premium and HE is an abbreviation for the Hill estimator. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Risk Premiums with the Hill Estimator, NB, MLE and LSQ					
Layer	**LF RP**	**RP(NB,MLE,HE)**	%Δ	**RP(NB,LSQ,HE)**	%Δ
L1	70.715	68.725	-2.814	63.941	-9.579
L2	36.323	36.090	-0.641	32.661	-10.082
L3	10.465	11.733	12.117	10.965	4.778
L4	2.319	2.968	27.986	3.003	29.495

Table 4.16: Comparison of the the risk premium for the different layers using MLE and LSQ, with the Hill estimator. The risk premiums have been calculated using the GPD and the negative binomial distribution. RP is an abbreviation for the risk premium and HE is an abbreviation for the Hill estimator. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.



(a) GPD Fit with MLE

(b) GPD Fit with LSQ

Figure 4.5: GPD fit with modified threshold u (Norwegian data set). In the graphs, the blue dots represent the data points (claims), and the red line is the generalized Pareto curve.

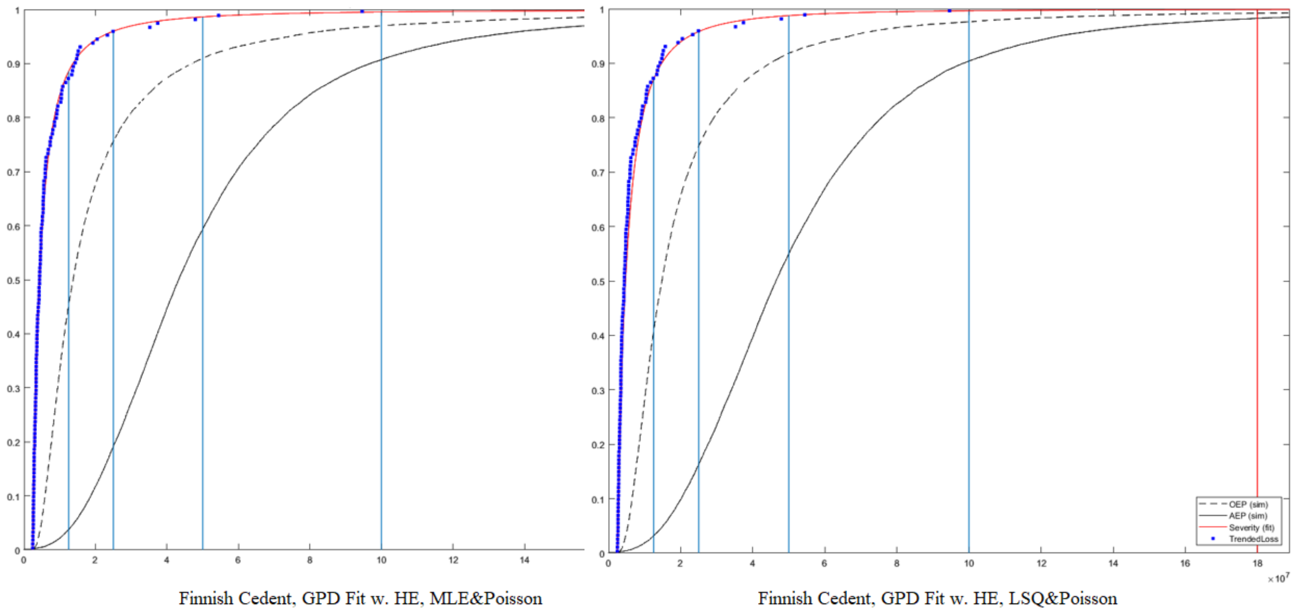
Finnish Data Set

Risk Premiums with Hill estimator, MLE and LSQ					
Layer	**LF RP**	**RP(MLE,HE)**	%Δ	**RP(LSQ,HE)**	%Δ
L1	40.250	37.696	-6.345	40.551	0.748
L2	14.365	15.039	4.692	14.763	2.771
L3	3.647	5.232	43.460	4.418	21.141
L4	0.822	1.879	128.59	1.335	62.409

Table 4.17: Comparison of the the risk premium for the different layers using MLE and LSQ, with the Hill estimator. The risk premiums have been calculated using the GPD and the Poisson distribution. RP is an abbreviation for the risk premium and HE is an abbreviation for the Hill estimator. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Risk Premiums with the Hill Estimator, NB, MLE and LSQ					
Layer	**LF RP**	**RP(MLE,HE)**	%Δ	**RP(LSQ,HE)**	%Δ
L1	40.250	35.752	-11.175	38.311	-4.817
L2	14.365	14.736	2.583	14.479	0.794
L3	3.647	5.201	42.610	4.383	20.181
L4	0.822	1.889	129.805	1.371	66.788

Table 4.18: Comparison of the the risk premium for the different layers using MLE and LSQ, with the Hill estimator. The risk premiums have been calculated using the GPD and the negative binomial distribution. RP is an abbreviation for the risk premium and HE is an abbreviation for the Hill estimator. The values are given in percentages.



(a) GPD Fit with MLE (b) GPD Fit with LSQ

Figure 4.6: GPD fit with modified threshold u (Finnish data set). In the graphs, the blue dots represent the data points (claims), and the red line is the generalized Pareto curve.

4.5 Comparison of Risk Premium with Bootstrapping

Initially, we implemented the basic and pivot studentized intervals to both the data sets as described in Sections 2.9.3, and 2.9.4. The results for both methods were highly inaccurate for both the data sets. The main reason for this is because both methods gave intervals that included negative values. The results, because of their inaccuracy, can be found in Appendix A. There could be multiple reasons for these results. One reason could be because the methods cannot guarantee that the test statistic, which in our case is the risk premiums, will fall within a certain range. The risk premiums are bounded, and can obviously not be zero or negative. Another issue is that when computing the intervals, we are assuming that the layers in the contract are independent, which they are not. Further, bootstrapping is usually done on the test statistic, however in this case, bootstrapping is done to the data samples that then being used to calculate the test statistics. Because of these results, we will continue with the most basic version of bootstrapping, and just report the empirical percentiles of the bootstrapped estimates of the risk premiums for each layer.

Based on the results in the previous section, we have seen that in general, MLE gives better estimates of the parameter for the GPD and the risk premiums. There are a large difference in parameter estimation of ξ and σ between LSQ and MLE. The resulting risk premiums also differed significantly. In general, MLE gives parameter estimates that are more precise with smaller estimated variance than LSQ. MLE is also favorable when the underlying distribution is known, which is the case here. Therefore, in order for LSQ to be deemed a suitable method for parameter estimation, we would expect it to give similar estimates as MLE. Since this was not the case, we therefore determine that MLE is more suitable to use for these data sets. The analysis will therefore continue with bootstrapping only for the data sets in which MLE is used for parameter estimation for the GPD.

Both the Norwegian and the Finnish data sets were bootstrapped using non-parametric bootstrapping with replacement. From bootstrapping, 1000 new samples for each data set were created. The risk premiums were then calculated by modeling the severity of claims with the GPD, and the claims frequency with the Poisson distribution or the negative binomial distribution. From bootstrapping the means of the risk premiums as well as empirical percentiles could be determined. The 2.5^{th} percentile as well as the 97.5^{th} percentile were calculated. This means out of 1000 simulations, that 95% of the risk premiums in each layer were between these percentiles. The goal of bootstrapping was to evaluate whether LF's current risk premiums would fall within the percentiles or not. This would be of value since it would be a way to determine whether LFs current pricing model seem reasonable and yields a fair price or not.

4.5.1 Evaluating LF's Current Risk Premiums

LF currently uses the GPD to model the severity of claims, and the Poisson distribution to evaluate the claims frequency. Bootstrapping was used for both the data sets to calculate the 2.5^{th} and 97.5^{th} empirical percentiles, which could be compared to LF's current risk premiums. The results can be seen in Tables 4.19, and 4.20.

For the Norwegian data set, All the risk premiums except the risk premium in **L1** fall between the 2.5^{th} and 97.5^{th} percentiles.

For the Finnish data set, all the risk premiums determined by LF fall within the 2.5^{th}

and 97.5^{th} percentile interval. Again a reason for the risk premiums of the Finnish data set comparing more favorably to the empirical intervals could be due to the Finnish data set being significantly larger than the Norwegian data set.

Norwegian Data Set

LF RP Empirical Percentiles			
Layer	LF RP	2.5^{th} %-ile	97.5^{th} %-ile
L1	70.715	92.440	158.601
L2	36.323	32.990	62.398
L3	10.465	8.076	26.041
L4	2.319	1.764	10.898

Table 4.19: LF's risk premium, and empirical percentiles. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Finnish Data Set

LF RP Empirical Percentiles			
Layer	LF RP	2.5^{th} %ile	97.5^{th} %ile
L1	40.250	20.631	83.896
L2	14.365	1.780	32.439
L3	3.647	0.033	7.131
L4	0.822	0.000	3.013

Table 4.20: LF's risk premium, and empirical percentiles. The values are given in percentages. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

4.5.2 Percentiles with the Hill Estimator

In this thesis, the Hill estimator was used in order to determine the optimal value for θ. We therefore bootstrapped 1000 new samples and calculated the 2.5^{th} and 97.5^{th} empirical percentiles of the risk premiums which included the new location parameter, in order to see the effects of using the Hill estimator. These calculations were made for both the negative binomial, and Poisson distribution with GPD using MLE. The results can be seen in Tables 4.21, 4.22,4.23, and 4.24. We can see that for both the data sets, all the risk premiums except for **L1** for both the data sets fall within the 2.5^{th} and 97.5^{th} percentiles.

Norwegian Data Set

LF RP with HE, Po and Empirical Percentiles				
Layer	LF RP	RP(Po,HE)	2.5^{th} %ile	97.5^{th} %ile
L1	70.717	69.663	84.011	151.357
L2	36.323	36.317	24.771	67.849
L3	10.465	11.637	3.428	23.379
L4	2.319	3.044	0.341	7.705

Table 4.21: The risk premiums determined using the Hill estimator, the Poisson distribution and MLE for parameter estimation of the GPD (except for the location parameter), and the corresponding bootstrapped empirical percentiles. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model. The values are given in percentages. RP is an abbreviation for risk premium, and HE is an abbreviation for the Hill estimator. The values are given in percentages.

Risk Premiums with HE, NB, and Empirical Percentiles				
Layer	LF RP	RP(NB,HE)	2.5^{th} %ile	97.5^{th} %ile
L1	70.715	68.725	77.397	150.354
L2	36.323	36.090	22.811	63.627
L3	10.465	11.733	2.824	21.225
L4	2.319	2.968	0.181	7.067

Table 4.22: The risk premiums determined using the Hill estimator, the negative binomial distribution and MLE for parameter estimation of the GPD (except for the location parameter), and the corresponding bootstrapped empirical percentiles. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model. The values are given in percentages. RP is an abbreviation for risk premium, and HE is an abbreviation for the Hill estimator. The values are given in percentages.

Finnish Data Set

LF RP with HE, Po and Empirical Percentiles				
Layer	LF RP	RP(Po,HE)	2.5^{th}%ile	97.5^{th}%ile
L1	40.250	37.696	15.859	70.958
L2	14.365	15.039	2.250	28.167
L3	3.647	5.232	0.287	12.022
L4	0.822	1.879	0.034	5.356

Table 4.23: The risk premiums determined using the Hill estimator, the Poisson distribution and MLE for parameter estimation of the GPD (except for the location parameter), and the corresponding bootstrapped empirical percentiles. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model. The values are given in percentages. RP is an abbreviation for risk premium, and HE is an abbreviation for the Hill estimator. The values are given in percentages.

Risk Premiums with HE, NB, and Empirical Percentiles				
Layer	LF RP	RP (NB,HE)	2.5^{th} %ile	97.5^{th} %ile
L1	40.250	35.752	22.682	67.231
L2	14.365	14.736	4.150	27.049
L3	3.647	5.201	0.764	11.799
L4	0.822	1.889	0.162	5.437

Table 4.24: The risk premiums determined using the Hill estimator, the negative binomial distribution and MLE for parameter estimation of the GPD (except for the location parameter), and the corresponding bootstrapped empirical percentiles. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model. The values are given in percentages. RP is an abbreviation for risk premium, and HE is an abbreviation for the Hill estimator. The values are given in percentages.

Chapter 5

Discussion

This final chapter includes a summary of the results as well as a discussion with conclusions. This chapter also introduces some ideas for future work.

5.1 Summary and Discussion of the Results

The goal of the thesis was to evaluate and validate LF's current pricing model for Risk XL reinsurance contracts, as well as make changes to the model in order to see if the risk premium could be more fairly priced. The negative binomial distribution was used instead of the Poisson distribution to model the frequency of claims, and the method of least squares instead of maximum likelihood was used to estimate the parameters for the Generalized Pareto distribution. We used the Hill estimator to determine the location parameter θ in the GPD, instead of using LF's previous practice of setting θ manually. Further, we performed two goodness of fit tests, the Kolmogorov-Smirnov test, and the Anderson-Darling test, in order to confirm that the generalized Pareto distribution was the most suitable distribution to model the severity of claims. Lastly we introduced the bootstrapping resampling method to determine the 2.5th and 97.5th empirical percentiles. Bootstrapping would be a way to evaluate whether LF's current risk premiums seem reasonable or not.

The goodness of fit tests, confirmed that for the Finnish data set, the most suitable distribution was the GPD. For the Norwegian data set, the most suitable distribution was the PD1, however with the GPD still having a large p-value. As mentioned in the results section, this could be due to the Norwegian data set is much smaller than the Finnish data set, which possibly makes fitting the data to a distribution more difficult. From Boxplot 3.4 we saw that the Norwegian data set had a less defined tail than the Finnish data set, which also possibly makes the fitting of any Pareto distribution less reliable.

Bootstrapping showed that all the risk premiums from LF's current pricing models fell between the 2.5th and the 97.5th percentiles for both data sets, except for the first layer **L1** for the Norwegian data set. It should however be noted that the percentiles determined by bootstrapping should in theory likely be lower than what the true values are. The reason for this is that when randomly selecting the samples in the bootstrapping method, there is a big chance that some of the extreme values in the data sets will not be included. For example for the Norwegian data set, the probability of a data point not being included in the bootstrap sample is around 36% each draw. In other words there is a large probability of leaving out the data points that are both rare and extreme, that would be found in the tails of the distribution. This could lead to bias. As we have seen from this thesis, the extreme values are of high importance in reinsurance contracts, and need to be included when modeling the severity of claims. For this reason, and based on the results of all the implemented bootstrapping techniques, we have determined that classical bootstrapping is not an appropriate method to estimate statistical uncertainty for our data sets, and data sets containing extreme values in general.

When looking at the results we can see that with the tail optimization with the Hill estimator the results do not seem to change significantly from LF's current model in the lower layers. However, when looking at the higher layers, e.g. **L4**, we see a significantly large change in the risk premiums. As mentioned previously, the HE infers the behavior of the tail of a distribution, which is of importance since this is where eventual extreme losses fall, which need to be included in the calculation of the risk premiums. We see that using a threshold u, which is theoretically motivated, shows that LF's current algorithm looses many potential losses in the top layers when pricing their risk premiums. The Hill estimator therefore yields a significant difference, especially when pricing reinsurance XL contracts. The computational cost for this is minimal compared to the effect it might have on the risk premiums.

5.2 Conclusion

In terms of improving LF's current pricing model, we conclude that, if the negative binomial distribution is applicable, then it beneficial to use it instead of the Poisson distribution for modeling the frequency of events. We did not see very large changes in the risk premiums from changing frequency distributions, despite this, the negative binomial distribution is favored, since is more flexible than the Poisson distribution. This is because it does not have the restriction that the expectation has to be the same as the variance, which could lead to an overall better fit for the data. The method of least squares gave different risk premiums especially in the highest layers of the risk premiums for both the data sets compared to using MLE. Parameter estimation with MLE is in general more precise than LSQ, and since LSQ deviated significantly from the values that MLE gave, we therefore do not deem LSQ to be suitable for these data sets. Also, LSQ is a useful method when the distribution function is not known, and would therefore be more useful for data sets where it is not as clear what the underlying distribution function is. In the case of the Norwegian and Finnish data sets, both goodness of fit tests indicated with high probability that both data sets most likely follows a PD1 or GPD distribution, making MLE a better choice for parameter estimation. For data sets where the goodness of fit tests does not give as clear results to which distribution the data follows, LSQ might be a better option than MLE to use for parameter estimation. Overall, using the goodness of fit tests were useful, since it can give a clearer idea of what distribution ones data comes from, and it help rule out unlikely distributions. The tests are also cheap in terms of the computational costs, and are relatively simple to implement.

Using the Hill estimator to estimate the location parameter did give significantly different results in terms of the risk premium in the highest layers for both data sets, and in particular for the Norwegian data set. We found clear patterns a more fairly determined risk premium when using a theoretically motivated threshold, and therefore conclude that the HE is worth using. In terms of bootstrapping, both the basic bootstrapping interval and the pivot studentized bootstrapped intervals gave highly inaccurate results. These can be found in Appendix A. We therefore used the empirical percentiles to construct intervals instead. We however know that this method is also biased, since it will likely generate lower values than what the true values are. Because of this, we conclude that we do not have enough evidence to draw any conclusions as to whether the risk premiums determined by LF's pricing model is fairly priced or not via classical bootstrapping methods. However since LF is using current methods that are common practice for XL contracts, we do believe that the risk premiums from their current pricing model are considered fair.

5.3 Future Work

Possible improvements of the pricing model for reinsurance contracts are always of interest. Other ways to improve LF's pricing model that was not explored in this thesis, could be to use a piece wise Pareto distribution, that models each layer of the reinsurance contract with its own distribution. This could be interesting to look at since it might help prevent overestimating or underestimating claims, especially when the data has a large spread. We saw issues with the first layer for the Norwegian data set, and having a separate distribution for this layer might help resolve this issue. One could also look further into choosing theoretically motivated thresholds and how to better model the data in the tail of a distribution. It could also be of value to look into exposure based pricing, especially when pricing higher layers which include extreme values. Since our data sets had very few losses in the top layers, using that type of pricing could help estimating large future losses.

As we have seen from this thesis, classical bootstrapping methods do not work well when the data sets have extreme values. As stated by [16], *"Broadly speaking, proving consistency of the bootstrap in extreme value theory is a difficult issue because even if observations are drawn from a distribution that satisfies the extreme value conditions, the corresponding empirical distribution function does not satisfy the extreme value conditions"*. It is therefore not obvious that bootstrapping can be used in order to calculate test statistics that can provide valuable inference about the true probability distribution of some data in extreme value theory. The article therefore proposes a bootstrap version of the fundamental expansions for the tail quantile process both for the POT-method and for block maxima methods to overcome this issue. This could therefore be worth dedicating more research to.

Bibliography

[1] Antal, P. “Quantitive methods in reinsurance”. In: (2003), pp. 1–68.

[2] Arshad, M., Rasool, M., and Ahmad, M. “Anderson Darling and Modified Anderson Darling Tests for Generalized Pareto Distribution”. In: (2003). DOI: 10.3923/jas.2003.85.88.

[3] Blom, G. et al. *Sannolikhetsteori och statistikteori med tillämpningar*. 5.5. Författarna och Studentlitteratur, 2008.

[4] Bolin, D. *Föreläsning 4: Kontinuerliga fördelningar*. 2018. URL: http://www.math.chalmers.se/Stat/Grundutb/CTH/tma074/1819/lectures/F4slides.pdf.

[5] Butun, B. “Excess of Loss Reinsurance Pricing in the Lloyd’s market”. In: *Institute des Actuaries* (), pp. 1–107.

[6] *Chapter 7. Statistical Estimation*. URL: https://web.stanford.edu/class/archive/cs/cs109/cs109.1218/files/student_drive/7.3.pdf (visited on 05/15/2023).

[7] Danielsson, J. et al. “Tail Index Estimation: Quantile-Driven Threshold Selection”. In: (2019).

[8] Daníelsson, J. *Financial Risk Forecasting- The Theory and Practice of Forecasting Market Risk, with Implementation in R and Matlab*. 1st ed. Wiley, 2011, pp. 167–181.

[9] Davison, A. C. and Hinkley, D. V. *Bootstrap Methods and their Application*. Cambridge University Press, Oct. 1997. DOI: 10.1017/cbo9780511802843.

[10] Dekking, F. M. et al. *A Modern Introduction to Probability and Statistics*. 1st ed. Springer London, 2005, pp. 341–343. DOI: https://doi.org/10.1007/1-84628-168-7.

[11] Desment, S. et al. “Experience and exposure rating for property per risk excess of loss reinsurance revisited”. In: (Sept. 2012), p. 239. DOI: 10.2143/AST.42.1.2160742.

[12] Dodge, Y. *The Concise Encyclopedia of Statistics*. 1st ed. Springer, 2008, pp. 283–287.

[13] Flowe, M. et al. “Reinsurance Pricing: Practical Isuues Considerations”. In: *2006 GIRO “Reinsurance Matters!” Working Party* (Sept. 2006), p. 8.

[14] Frackler, M. “Reinventing Pareto: Fits for both small and large losses”. In: (2022), pp. 1–27.

[15] Gut, A. *An Intermediate Course in Probability*. 2nd ed. Springer, 1995, Appendix 2.

[16] Haan, L. de and Zhou, C. “Bootstrapping Extreme Value Estimators”. In: (2022).

[17] Helwig, N. E. *Nonparametric Bootstrap in R*. 2021. URL: `http://users.stat.umn.edu/~helwig/notes/npboot-notes.html?fbclid=IwAR0INmj8GwyazQ5DZTrayrGt_Mb1nExrjKgtOW9AGdlScSHRgVd3Tuv9KNI`.

[18] Hill, B. M. “A Simple General Approach To Inference About The Tail Of A Distribution”. In: Ann. Stat. 3 (1975).

[19] Hult, H. et al. *Risk and Portfolio Analysis - Principles and Methods*. 1st ed. Springer, 2012.

[20] Johansson, B. “Matematiska modeller inom sakförsäkring”. In: Matematisk Statistik Stockholm Universitet, 2014, pp. 1–154.

[21] *Lesson 1: Point Estimation, 1.4 - Method of Moments*. URL: `https://online.stat.psu.edu/stat415/lesson/1/1.4?fbclid=IwAR3GOVcnuHbOOCLAt5l3IiX7MBtkM2P9fm0z` `mkdgP4vgpV2j88` (visited on 05/15/2023).

[22] Marsaglia, G., Tsang, W. W., and Wang, J. “Evaluating Kolmogorov’s Distribution”. In: (2003). DOI: `10.18637/jss.v008.i18`.

[23] Marsaglia, G. and Marsaglia, J. “Evaluating the Anderson-Darling Distribution”. In: (2004). DOI: `10.18637/jss.v009.i02`.

[24] McNeil, A. J., Frey, R., and Embrechts, P. *Quantitative Risk Management - Concepts, Techniques and Tools*. Revised. Princeton University Press, 2015, pp. 269–322.

[25] Molina Peralta, I. and García-Portugués, E. *A First Course on Statistical Inference*. Version 1.4.8. ISBN 978-84-09-29680-4. 2023.

[26] MSimard, R. and L’Ecuyer, P. “Computing the Two-Sided Kolmogorov-Smirnov Distribution”. In: (2011). DOI: `10.18637/jss.v039.i11`.

[27] Nadarajah, S., Zhang, Y., and Pogány, T. K. “On Sums of Independent Generalized Pareto Random Variables with Applications to Insurnace and Cat Bonds”. In: (2017), pp. 1–10.

[28] *nbinstat*. Introduced before R2006a. URL: `https://se.mathworks.com/help/stats/nbinstat.html` (visited on 05/01/2023).

[29] *Negative Binomial Distribution*. URL: `http://www.math.ntu.edu.tw/~hchen/teaching/StatInference/notes/lecture16.pdf` (visited on 05/01/2023).

[30] *OFX OzForex Limited*. 2022. URL: `https://www.ofx.com/en-au/` (visited on 04/01/2023).

[31] Ross, S. M. *A First Course In Probability*. 8th ed. Pearson, 2008, p. 335.

[32] Rosso, G. “Extreme Value Theory for Time Series using Peak-Over-Threshold method”. In: (2015).

[33] Scarrott, C., Hu, Y., and Akbar, A. "Extreme Value Mixture Modelling, Threshold Estimation and Boundary Corrected Kernel Density Estimation". In: (2019).

[34] Stephens, M. A. *The Anderson-Darling Statistic*. Tech. rep. Standford Univeristy, Oct. 1979.

[35] Training, T. and Office, C. U. "An introduction to Reinsurance". In: 8th Edition (2002), pp. 1–34.

[36] Zyl, J. M. van. *Estimation of the shape parameter of a generalized Pareto distribution based on a transformation to Pareto distributed variables*. 2012. URL: `https://arxiv.org/pdf/1210.7642.pdf`.

Appendix A

Confidence Interval Results

This section presents the results of using bootstrapping with 95% Pivot confidence intervals for both the data sets. Because of inaccurate results, the confidence intervals were only computed using LF's current pricing model. LF's current model is to use MLE to estimate the parameters for the GPD distribution which models the severity of claims, and the Poisson distribution to model the frequency of claims. The location parameter θ is set manually by LF. Here LF's risk premiums using their current pricing model are compared to the intervals from basic bootstrapping as well as the pivot studentized bootstrapped intervals.

Basic Bootstrap Confidence Intervals

Norwegian data set

LF RP	Lower Bound, 2.5^{th} Percentile	Upper Bound, 97.5^{th} Percentile
70.715	-17.171	48.990
36.323	10.248	39.656
10.465	-5.111	12.854
2.319	-6.260	2.874

Table A.1: Basic Bootstrap Confidence Intervals. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Finnish data set

LF RP	Lower Bound, 2.5^{th} Percentile	Upper Bound, 97.5^{th} Percentile
40.250	-3.396	59.869
14.365	-3.709	26.950
3.647	0.163	7.261
0.822	-1.369	1.644

Table A.2: Basic Bootstrap Confidence Intervals. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Pivot Studentized Bootstrap Intervals

Norwegian data set

LF RP	Lower Bound, 2.5^{th} Percentile	Upper Bound, 97.5^{th} Percentile
70.715	80.194	158.983
36.323	31.426	68.832
10.465	4.865	26.257
2.319	-1.509	11.085

Table A.3: Pivot Studentized Bootstrap Intervals. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

Finnish data set

LF RP	Lower Bound, 2.5^{th} Percentile	Upper Bound, 97.5^{th} Percentile
40.250	21.448	73.045
14.365	1.568	26.322
3.647	-0.753	7.326
0.822	-0.591	2.185

Table A.4: Pivot Studentized Bootstrap Intervals. LF RP is an abbreviation for the risk premiums determined from LF's current pricing model.

TRITA – TRITA-SCI-GRU 2023:305