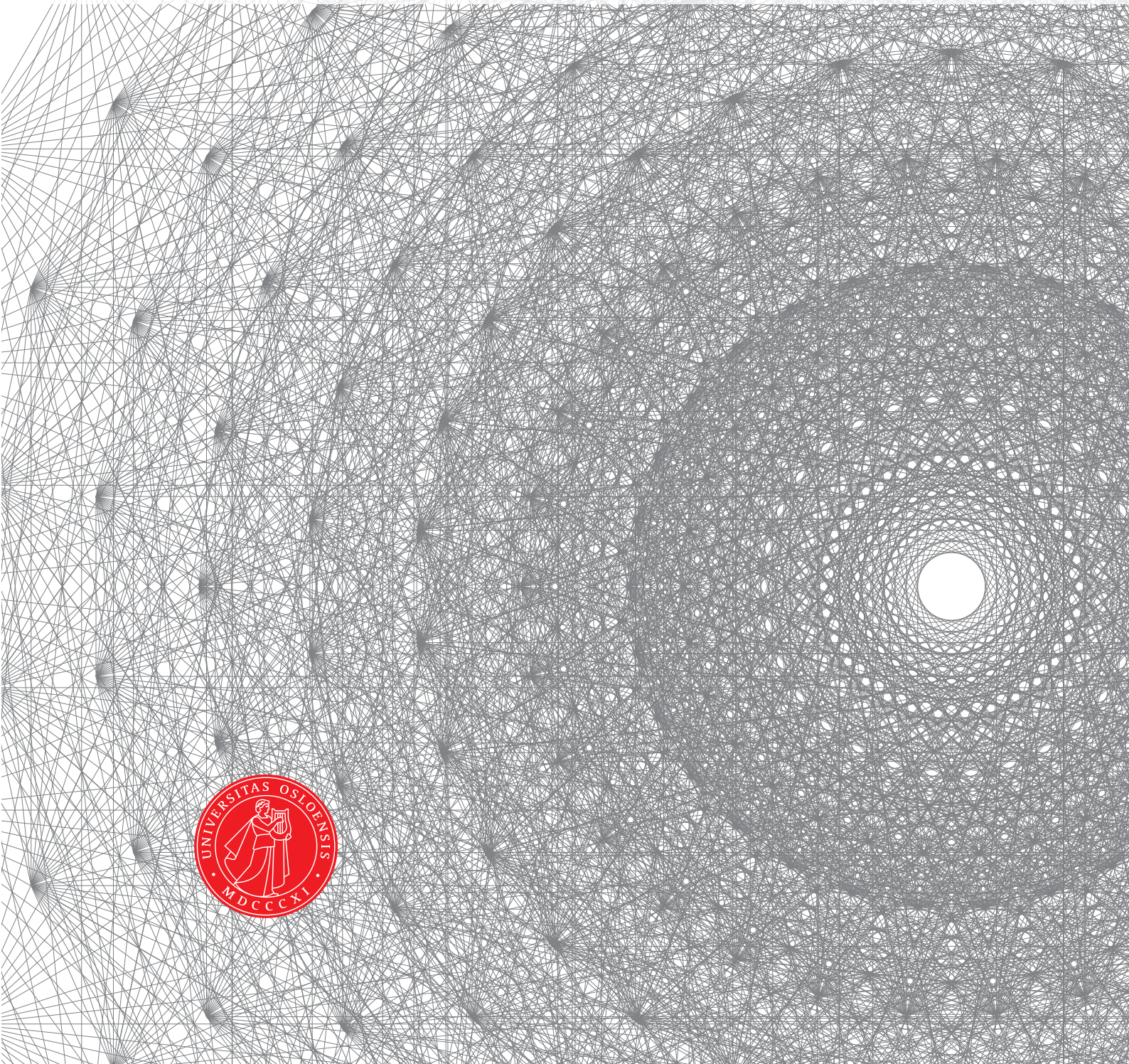
UiO : Department of Mathematics
University of Oslo

Selection of claim size distribution in property insurance

Kristina Hauger

Master's Thesis, Spring 2017



Cover design by Martin Helsø

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

This thesis has assessed whether a six-parameter extension of the Pareto distribution may be used in an automatic procedure for selecting the claim size distribution in property insurance. This fitted distribution was compared to eight of its special case distributions with regards to the estimation of the 99% and 99.5% reserve for data consisting of property insurance claims on office and industrial buildings from Gjensidige forsikring as well as a subset of the data. As the data contained extreme claims, the claims above a certain threshold were modeled with a two-parameter Pareto distribution according to Pickand's theorem. The threshold was set at the 96.7% and 92.0% quantile for the total and reduced data, respectively. As a result, the two-parameter Pareto over-threshold distribution dominated the modeling of the reserve leading to virtually no significant differences in the estimated reserves for the different under-threshold models combinations, especially for the smaller data set, for which the threshold was lower.

The Extended Pareto and the four-parameter Pareto distributions were selected as preferred under-threshold distributions for the total and reduced data, respectively. The two-parameter Pareto distribution was chosen as the preferred under-threshold distribution to estimate the 99% and 99.5% reserves for both the total and the reduced data. The six-parameter extension of the Pareto distribution failed to optimize the parameter γ, which lead to the four-parameter Pareto distribution. However, the four-parameter Pareto distribution is more flexible than the two-parameter Pareto and the Extended Pareto distribution, and since there are little difference in the estimation of the reserve for the three models, the four-parameter Pareto distribution may automatically selected to estimate the 99% and 99.5% reserves.

However, if the quantity of interest was further left in the distribution, such as the the premium or the 95% quantile, the under-threshold distribution would have more influence in the modeling and perhaps the Extended Pareto distribution or an other distribution would have been chosen.

KEY WORDS: Property insurance, Claim size distribution, Threshold, Pareto distributions, Heavy-tailed data, Monte Carlo methods, Bootstrap, Maximum likelihood estimation, Model selection, Reserve estimation.

Acknowledgements

I would first like to express my very great appreciation to my supervisor Ingrid Hobæk Haff. You have been a great support and motivated me during this period. Thank you for all your help and engagement.

I would like to express my appreciation to my fellow students for your support and good times in "Parameter rommet". I would also like to thank my friends outside the university for interesting discussions, and my boyfriend for your support and encouragement.

I would like to offer my special thanks to "Statisikk og Analyse" in Finans Norge for giving me work experience, providing me with information about the actuarial industry and a positive work environment.

Finally I would like to thank my family, especially my father for your motivation and interest in my thesis.

Oslo, February 2017
Kristina Hauger

Contents

1 Introduction **1**

2 Methods **3**
2.1 Maximum likelihood estimation 3
2.2 Evaluating the model 5
2.3 Bootstrap and Monte Carlo methods 7
2.4 Reserve 10
2.5 Estimating and bootstrapping the reserve 11
2.6 Claim frequency 13
2.7 Claim size distributions 15
2.7.1 The Exponential distribution 15
2.7.2 The Weibull distribution 16
2.7.3 The Gamma distribution 18
2.7.4 The Log-gamma distribution 19
2.7.5 The Log-normal distribution 21
2.7.6 The two-parameter Pareto distribution 24
2.7.7 The Extended Pareto distribution 26
2.7.8 The four-parameter Pareto distribution 27
2.7.9 The six-parameter Pareto distribution 28
2.7.10 Modeling extreme claim sizes 31
2.7.11 Claims below the threshold 32

3 Results **34**
3.1 The claim size distributions 35
3.1.1 The under-threshold claim size data 36
3.1.2 The over-threshold claim size distribution 44
3.1.3 The reserve 46
3.2 Reducing the sample size 54
3.3 The reduced claim size data 54
3.3.1 The reduced under-threshold claim size data 55
3.3.2 The reduced over-threshold claim size distribution . . . 64
3.3.3 The reduced data reserve 65

4 Conclusion **74**

A Appendix **80**
A.1 The Extended Pareto distribution 80
A.2 From Extended Pareto to Gamma 82

A.3 From two-parameter Pareto to Exponential 83
A.4 The Log-gamma distribution 84
A.5 The four-parameter Pareto moments 84

B R-script **86**
B.0.1 Reading the data 86
B.0.2 Descriptive statistics for the total data 86
B.0.3 Selecting the threshold 86
B.0.4 Descriptive statistics for the under-threshold data . . . 87
B.0.5 Codes used in optimization 87
B.0.6 Fitting the over-threshold distribution given the total data 87
B.0.7 Selecting the reduced data 89
B.0.8 Fitting the over-threshold distribution given the reduced data 90
B.0.9 The Exponential under-threshold distribution 92
B.0.10 The Weibull under-threshold distribution 97
B.0.11 The Gamma under-threshold distribution 102
B.0.12 The Log-gamma under-threshold distribution 108
B.0.13 The Log-normal under-threshold distribution 113
B.0.14 The two-parameter Pareto under-threshold distribution 119
B.0.15 The Extended Pareto under-threshold distribution . . . 124
B.0.16 The four-parameter Pareto under-threshold distribution 132
B.0.17 The six-parameter Pareto under-threshold distribution 139

List of Figures

1 Mean excess plot. 36
2 Q-Q plots for all fitted under-threshold distributions. 40
3 Two-parameter Pareto over-threshold Q-Q plot. 45
4 Probability density functions of the 99% Monte Carlo simulated reserves with 95% confidence intervals for the nine models 49
5 Probability density functions of the 99.5% Monte Carlo simulated reserves with 95% confidence intervals for the nine models 50
6 Mean excess plot for the reduced data. 55
7 Q-Q plots for all fitted under-threshold distributions given the reduced data. 59
8 Two-parameter Pareto over-threshold Q-Q plot for the reduced data. 65
9 Probability density functions of the 99% Monte Carlo simulated reserves with 95% confidence intervals for the nine models, given the reduced data 69
10 Probability density functions of the 99.5% Monte Carlo simulated reserves with 95% confidence intervals for the nine models, given the reduced data 70

List of Tables

1 Five-number summary of the data without threshold. 35
2 Five-number summary of the under-threshold data. 36
3 Maximum likelihood estimated parameters with 95% confidence intervals for the under-threshold distributions. 38
4 Statistics and model evaluation for the under-threshold data. . 42
5 Maximum likelihood estimated parameters with 95% confidence intervals for the over-threshold data. 45
6 Kolmogorov-Smirnov goodness-of-fit test for the over-threshold data. 45
7 Estimated 99% and 99.5% reserves with 95% confidence intervals. 47
8 Five-number summary of the reduced data without threshold. 54
9 Five-number summary of the reduced under-threshold data. . 55
10 Maximum likelihood estimated parameters with 95% confidence intervals for the reduced data under-threshold distributions. 58
11 Statistics and model evaluation for the reduced under-threshold data. 62
12 Maximum likelihood estimated parameters with 95% confidence intervals for the over-threshold reduced data. 64
13 Kolmogorov-Smirnov goodness-of-fit test for the reduced over-threshold data. 64
14 Estimated 99% and 99.5% reserves with 95% confidence intervals for the reduced data. 67

1 Introduction

When dealing with heavy-tailed data, the claim severity modeling may be divided into two parts; an under-threshold distribution and an over-threshold distribution, modeling claims below and above the selected threshold b, respectively [Bølviken, 2014, p. 328]. Generally the data is fitted to several claim size distributions, the results are compared and the best fitted model is selected and its parameters estimated from historical data [Bølviken, 2014, p. 314]. However, selecting an appropriate model using this method may be difficult and time-consuming. A different way to model the claim size is to adapt more complex families of distributions, with more frequently used distributions as special cases. The six-parameter Pareto distribution, which will be introduced in the next chapter, is constructed based on a parameterized power transformation inspired by Box and Cox (1964), now with six parameters instead of two. This thesis intends to assess whether the six-parameter Pareto distribution may be used in an automatic procedure for selecting the claim size distribution compared to its one, two, three and four parameter special case distributions in order to make the process less time-consuming and efficient.

EU insurance legislation requires insurance companies to reserve a sufficient amount of capital to cover their liabilities over a one-year time perspective with a 99.5% probability [European Comission, 2009]. This thesis aims to study how the uncertainty in the fitted under-threshold distributions increase with the number of parameters and the impact of uncertainty on the estimated 99% and 99.5% reserves.

This study will be performed on data consisting of property insurance claims on office and industrial buildings from January 2001 to October 2015 from Gjensidige forsikring. The aim is to repeat the procedure on a subset of the data to study how the sample size influences the uncertainty in the reserve.

In chapter 2, the methods used in this thesis are presented. These include model evaluation methods, selecting a threshold for the data, and the estimation and bootstrapping of the reserve.

Chapter 3 will aim to give the results for the different under-threshold distributions, the over-threshold distribution and the reserves estimated with the

nine different models. The results will then be presented for a subset of the data.

Chapter four will aim to give a conclusion.

2 Methods

2.1 Maximum likelihood estimation

The most common method for parameter estimation is the Maximum likelihood estimation [Gray and Pitts, 2012, p. 59]. The idea is to select estimated parameters maximizing the probability of observing the sample given the distribution.

Given k parameters θ_1,θ_2,...,θ_k, assume the sample Z_1,...,Z_n, consists of iid observations with density function f(z;θ_1,θ_2,...,θ_k) [Jong and Heller, 2008, p. 40].

Their joint density function is:

$$f(z_1, ..., z_n; \theta_1, \theta_2, ..., \theta_k) = \prod_{i=1}^{n} f(z_i, \theta_1, \theta_2, ..., \theta_k). \tag{1}$$

Taking the logarithm of equation (1) leads to the log-likelihood function:

$$\mathcal{L}(\theta_1, \theta_2, ..., \theta_k) = \sum_{i=1}^{n} \log\left(f(z_i, \theta_1, \theta_2, ..., \theta_k)\right). \tag{2}$$

The Maximum likelihood estimator of the parameter θ_j will be denoted $\hat{\theta}_j$.

In a few straightforward situations, the parameters can be estimated by differentiating the log-likelihood function with respect to the parameter of interest and setting the equation equal to zero:

$$\frac{\partial \sum_{i=1}^{n} \log(f(z_i, \theta_1, \theta_2, ..., \theta_k))}{\partial \theta_j} = 0, \qquad j = 1, ..., k. \tag{3}$$

In general, optimization requires numerical maximization of the log-likelihood, or equivalently minimization of the negative log-likelihood function [Millar, 2011, p. 101].

Important properties of Maximum likelihood estimation

Invariance:
Let g be a one-to-one function and $\hat{\theta}_j$ the Maximum likelihood estimate (MLE) of θ_j . Then g($\hat{\theta}_j$) equals the MLE of g(θ_j).

Consistency:
As the number of observations increases, the estimated parameter gets closer to the true value: $\hat{\theta}_j \xrightarrow{p} \theta_j$, as n $\to \infty$.

Asymptotically unbiased and normally distributed:
The following limits are only valid under certain conditions and assumptions and the interested reader is referred to [Lehmann, 1999, p. 469].

The level of bias goes to zero as the sample size increases:
E($\hat{\theta}_j$) $\to \theta_j$ as n $\to \infty$.

As the sample sizes becomes large, $\sqrt{n}\left(\hat{\theta}_j - \theta_j\right)$ becomes normally distributed with mean zero, and variance $I^{-1}(\theta)$, where I(θ) is the Fisher information matrix given by:

$$I(\theta) = \left(E\left[-\frac{\partial^2 \mathcal{L}(\theta_j)}{\partial^2 \theta_j}\right]\right).$$

For the multi-parameter case, the vector of Maximum likelihood estimates becomes asymptotically multivariate normally distributed, cf. [Klugman et al., 2008, p. 395] and [Lehmann, 1999, p. 498].

The following equation holds for all θ_1,θ_2,...,θ_k, as n $\to \infty$:

$$\sqrt{n}\left((\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k)^T - (\theta_1, \theta_2, ..., \theta_k)^T\right) \xrightarrow{d} N(\mathbf{0}, I^{-1}(\theta)), \tag{4}$$

where $\mathbf{0}$ = (0,...,0).

The Fisher information matrix for the multi-parameter case is given:

$$I(\theta) = E\left[-\begin{pmatrix} \frac{\partial^2}{\partial\theta_1^2} & \frac{\partial^2}{\partial\theta_1\partial\theta_2} & \cdots & \frac{\partial^2}{\partial\theta_1\partial\theta_k} \\ \frac{\partial^2}{\partial\theta_2\partial\theta_1} & \frac{\partial^2}{\partial\theta_2^2} & \cdots & \frac{\partial^2}{\partial\theta_2\partial\theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial\theta_k\partial\theta_1} & \frac{\partial^2}{\partial\theta_k\partial\theta_2} & \cdots & \frac{\partial^2}{\partial\theta_k^2} \end{pmatrix} \mathcal{L}(\theta_1, \theta_2, ..., \theta_k)\right]. \quad (5)$$

2.2 Evaluating the model

The goodness-of-fit of the different claim size distributions will be assessed by the Kolmogorov-Smirnov hypothesis test, Q-Q plots and the Akaike (AIC) and the Bayesian information criteria (BIC).

AIC is one of the most favored methods for model selection [Claeskens et al., 2008, p.22]. Given a statistical model of a set of data,

$$\text{AIC} = -2\mathcal{L} + 2p, \quad (6)$$

where $\mathcal{L}$ is the log-likelihood in equation (2) and p is the number of estimated parameters.

The goodness-of-fit, measured by $\mathcal{L}$, is awarded with the negative sign, and the penalty is an increasing function of p.

The Bayesian Information Criteria (BIC) was originally developed by Gideon E. Schwarz to assess models with regard to a given posterior probability. However, the criteria is applicable as a general model selection criteria under certain assumptions. The interested reader is referred to [Sadanori Konishi, 2007, p. 211-212]. BIC exercise a heavier penalty for for p, and favors models with fewer estimated parameters than AIC.

$$\text{BIC} = -2\mathcal{L} + p\log(n), \quad (7)$$

where $\mathcal{L}$ is the log-likelihood in equation (2), n the sample size of the data and p is the number of estimated parameters.

When comparing statistical models, the model with the lowest AIC or BIC

is selected because both methods are based on a trade-off between goodness of fit and the complexity of the model. The model with the less parameters along with the best fit is chosen [Jong and Heller, 2008, p. 62-63].

The Q-Q plot is a graphical technique that plots the empirical quantiles against the corresponding quantiles of the distribution of interest. An approximately straight line indicates that the two data sets belongs to the same distribution [Walpole et al., 2007, p. 241].

The Kolmogorov-Smirnov test is not a tool to compare the different distributions against each other. The assessment is a hypothesis test and will be used independently for a distribution to determine whether the data set has a specific theoretical distribution. The test examines the fit between the observed frequency and the expected (fitted) frequency.

The hypothesis tested is:

H_o: The data follows the specified distribution
H_a: The data does not follow the specified distribution

The Kolmogorov-Smirnov test statistic is defined as:

$$D_n = \sup_{0 \leq Z \leq b} \mid F_n(z) - F(z) \mid, \tag{8}$$

where F(z) is the given distribution function of the observed observations, $F_n(z)$ the empirical cumulative distribution function of the data and b the threshold level. The model distribution function F(z) is assumed to be continuous for all values of z [Klugman et al., 2008, p. 448]. H_0 is rejected at an α level of significance when $D_n > K_\alpha$, where K_α is the critical value.

The Kolmogorov-Smirnov test is distribution free, which means that the distribution of the test statistic under H_0 is equal to the distribution of the sample, regardless of the distribution tested [Gray and Pitts, 2012, p. 67].

The cumulative distribution function of K_α is given by:

$$P\left(K_\alpha \leq \sqrt{n}D_n\right) = 1 - \frac{\sqrt{2\pi}}{\sqrt{n}D_n} \sum_{k=1}^{\infty} \left(e^{-(2k-1)^2 \frac{\pi^2}{8nD_n{}^2}} \right),$$

for every fixed $\sqrt{n}D_n \geq 0$ as n $\to \infty$ [Feller, 1948, p. 178].

The p-values in section 3.1.1 and section 3.3.1 will be performed on the under-threshold data. For this reason the p-values have been obtained using Monte Carlo to simulate D_n - the observed value 100,000 times from the distribution under H_o, where F is the distribution to be tested. The number of simulated D_ns $\geq$ the observed D_n is calculated (see Appendix (B)).

2.3 Bootstrap and Monte Carlo methods

Monte Carlo methods are frequently applied in the insurance industry. These methods use statistical sampling on a computer to find solutions to problems [Korn, 2010, p. 55].

Let $\psi = \psi(\boldsymbol{\theta})$ be the quantity of interest to estimate, for example the reserve, which may be estimated by $\hat{\psi} = \psi(\hat{\boldsymbol{\theta}})$. When no expression for $\hat{\psi}$ is available, ψ may be estimated by Monte Carlo. The Monte Carlo estimate of $\psi(\boldsymbol{\theta})$ is denoted $\psi_m^*(\boldsymbol{\theta})$. When $\psi(\boldsymbol{\theta})$ is estimated by Monte Carlo, error due to simulation occurs in addition to estimation error. Let $\hat{\psi}_m^* = \psi_m^*(\hat{\boldsymbol{\theta}})$ be the Monte Carlo approximation of $\psi(\hat{\boldsymbol{\theta}})$. Then:

$$\underbrace{\hat{\psi}_m^* - \psi}_{\text{total error}} = \underbrace{\hat{\psi}_m^* - \hat{\psi}}_{\text{Monte Carlo error}} + \underbrace{\hat{\psi} - \psi}_{\text{estimation error}}$$

The Monte Carlo error is the error generated due to the fact that ψ is approximated by ψ_m^* and the deviation between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$ gives rise to the estimation error [Bølviken, 2014, p.229-230].

A common result is $\mid \hat{\psi}_m^\star - \psi \mid \to 0$ as m $\to \infty$, and $\mid \hat{\psi} - \psi \mid \to 0$ as n $\to \infty$. In other words, the Monte Carlo error will decrease with the number of replications, and the estimation error decreases with the number of data observations. As the number of data observations usually is fixed, estimation error is normally more serious than the Monte Carlo error [Bølviken, 2014, p. 230-231]. The numbers of replications in this thesis are set large enough for the Monte Carlo error to be negligible compared to the estimation error.

Assume that $\psi(\boldsymbol{\theta})$ can be written in the form:

$$\psi(\boldsymbol{\theta}) = E\left(h(z)\right) = \int h(z) f(z;\boldsymbol{\theta}) dz. \tag{9}$$

The quantity $\psi(\boldsymbol{\theta})$ in equation (9) may be approximated by the Monte Carlo estimate:

$$\hat{\mathcal{X}}_m^* = \hat{\psi}_m^*(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m} h\left(z_i^*\right), \tag{10}$$

where equation (10) is obtained by sampling m iid samples $z_i^* \sim f(z;\hat{\boldsymbol{\theta}})$. Then the function h is calculated for all samples $z_1^*, ..., z_m^*$ and the average of the m samples is computed.

As an example, simply for illustration, consider the special case of equation (9) and (10) where h(z) = z:

$$\psi(\boldsymbol{\theta}) = E(Z) = \int z f(z;\boldsymbol{\theta}) dz,$$

The Monte Carlo estimate becomes:

$$\bar{\hat{\psi}}_m^*(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m} \left(z_i^*\right).$$

Based on the Strong law of large numbers, $\hat{\psi}_m^*(\boldsymbol{\theta}) \to \psi(\boldsymbol{\theta})$, almost surely when m is sufficiently large [Robert and Casella, 2004, p. 83].

If f has finite expectation and variance, the standard deviation can be estimated by

$$\hat{s}^* = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(z_i^* - \bar{\hat{\psi}}_m^*(\boldsymbol{\theta})\right)^2}.$$

Algorithm 1 illustrates the procedure of computing the estimated mean $\bar{\hat{\psi}}^*$ and standard deviation $\hat{s}^*$ when the true values of $\bar{\psi}$ and s are unknown. First input the distribution of Z, and the vector of estimated parameters $\hat{\boldsymbol{\theta}}$. Then m claim sizes are generated from the given distribution with estimated parameters $\hat{\boldsymbol{\theta}}$. Finally the estimated mean and standard deviation are computed based on the Monte Carlo sample.

The bootstrap method is a computer-based method for estimating the uncertainty in the estimated vector of parameters $\hat{\boldsymbol{\theta}}$.

Algorithm 1 Algorithm for estimating the mean and standard deviation.

1: Input f(z), $\hat{\boldsymbol{\theta}}$
2: For i = 1 to m draw $z_i^* \sim f(z, \hat{\boldsymbol{\theta}})$
3: Compute
$\bar{\hat{\psi}}_m^* = \sum_{i=1}^m \frac{z_i^*}{m}$,
$\hat{s}^* = \sqrt{\frac{1}{m-1}\sum_{i=1}^m \left(z_i^* - \bar{\hat{\psi}}_m^*\right)^2}$
4: Return $\bar{\hat{\psi}}^*$ and $\hat{s}^*$.

The non-parametric bootstrap applies resampling with replacement. Consider a data set with n observations $\boldsymbol{x} = (x_1, ..., x_n)$. Then m_b new data sets of size n, $\hat{\boldsymbol{x}}_{\boldsymbol{i}}^{\star} = (\hat{x}_{1i}^{\star}, ...\hat{x}_{ni}^{\star})$, are drawn from $\boldsymbol{x}$ with replacement. For each sample drawn, the vector of parameters $\boldsymbol{\theta_i}$ is estimated by $\hat{\boldsymbol{\theta}}_i^*$ with the same method of estimation used to estimate $\hat{\boldsymbol{\theta}}$ [Gray and Pitts, 2012, p. 322-323].

The parametric bootstrap method generates m_b new samples $\hat{\boldsymbol{x}}_{\boldsymbol{i}}^{\star} = (\hat{x}_{1i}^{\star}, ..., \hat{x}_{ni}^{\star})$ $\sim f(z; \hat{\boldsymbol{\theta}})$ of the same size n as the original sample. For each sample i, $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}$ is estimated by $\hat{\boldsymbol{\theta}}_i^*$, which in our case is obtained by Maximum likelihood estimation [Miranda et al., 2011, p. 14-17].

The nonparametric bootstrap is not as precise as the parametric bootstrap, given that the parametric model is correct. For this reason the parametric bootstrap method will be applied in this thesis [Shao and Tu, 1995, p. 16].

For the parametric bootstrap, the Maximum likelihood estimates $\hat{\boldsymbol{\theta}}_{\mathbf{1}}^{\star}$, $\hat{\boldsymbol{\theta}}_{\mathbf{2}}^{\star}$,...,$\hat{\boldsymbol{\theta}}_{\boldsymbol{m_b}}^{\star}$ are returned from the generated samples $\hat{\boldsymbol{x}}_{\mathbf{1}}^{\star}$, $\hat{\boldsymbol{x}}_{\mathbf{2}}^{\star}$,...,$\hat{\boldsymbol{x}}_{\boldsymbol{m_b}}^{\star}$, respectively. A confidence interval (CI) and standard deviation can be estimated to assess the uncertainty. Let the sample mean of the bootstrap estimated vector be:

$$\bar{\hat{\boldsymbol{\theta}}}^{\star} = \sum_{i=1}^{m_b} \frac{\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\star}}{m_b}.$$

The bootstrap estimate of the standard error of $\hat{\boldsymbol{\theta}}$ is:

$$\hat{s}_{\hat{\theta}}^{\star} = \sqrt{\frac{1}{m_b - 1}\sum_{i=1}^{m_b} \left(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\star} - \bar{\hat{\boldsymbol{\theta}}}^{\star}\right)^2},$$

which is the sample standard deviation of $\left(\hat{\boldsymbol{\theta}}_{\mathbf{1}}^{\star}, \hat{\boldsymbol{\theta}}_{\mathbf{2}}^{\star}, ..., \hat{\boldsymbol{\theta}}_{\boldsymbol{m_b}}^{\star}\right)$ [Devore and Berk,

2007, p. 339].

When $\boldsymbol{\theta}$ is being underestimated, $\bar{\hat{\boldsymbol{\theta}}}^{\star} < \hat{\boldsymbol{\theta}}$, and $\bar{\hat{\boldsymbol{\theta}}}^{\star} > \hat{\boldsymbol{\theta}}$ indicates overestimation of $\boldsymbol{\theta}$. An adjustment may be made for correction:

$$\hat{\boldsymbol{\theta}}^{\star bc} = \underbrace{\hat{\boldsymbol{\theta}}}_{estimate} - \underbrace{\left(\bar{\hat{\boldsymbol{\theta}}}^{\star} - \hat{\boldsymbol{\theta}}\right)}_{adjustment} = 2\hat{\boldsymbol{\theta}} - \bar{\hat{\boldsymbol{\theta}}}^{\star}, \tag{11}$$

where $\hat{\boldsymbol{\theta}}^{\star bc}$ is the bias corrected bootstrap estimate of $\boldsymbol{\theta}$ [Bølviken, 2014, p. 241-242]. The adjustment term in equation (11) will increase $\hat{\boldsymbol{\theta}}^{\star bc}$ when $\boldsymbol{\theta}$ is underestimated, and decrease $\hat{\boldsymbol{\theta}}^{\star bc}$ when $\boldsymbol{\theta}$ is overestimated.

An approximate 100(1 - α)% confidence interval for $\hat{\boldsymbol{\theta}}^{\star}$ is given by sorting the estimates in increasing order:

$$\hat{\boldsymbol{\theta}}^{\star}_{(1)} \leq \hat{\boldsymbol{\theta}}^{\star}_{(2)} \leq, ..., \leq \hat{\boldsymbol{\theta}}^{\star}_{(m_b)}. \tag{12}$$

The $\frac{\alpha}{2}$ and 1 - $\frac{\alpha}{2}$ empirical quantiles of $\hat{\boldsymbol{\theta}}^{\star}$ are given by:

$$\left(\hat{\boldsymbol{\theta}}^{\star}_{\left(\frac{\alpha}{2}m_b\right)}, \hat{\boldsymbol{\theta}}^{\star}_{\left(1-\frac{\alpha}{2}m_b\right)}\right),$$

which gives rise to the possibility of asymmetric confidence intervals [Rubinstein and Kroese, 2008, p. 115].

2.4 Reserve

Let the total loss on policy and portfolio level be X and $\mathcal{X}$, respectively, where

$$X = Z_1 + ... + Z_N,$$
$$\mathcal{X} = Z_1 + ... + Z_{\mathcal{N}},$$

Z_i being the size of each claim i, N the number of claims on policy level, and $\mathcal{N}$ the number of claims on portfolio level.

Given a portfolio with J policies, each with claim X_i, for i = 1,...,J. The overall claim of the portfolio is given by:

$$\mathcal{X} = X_1 + ... + X_J. \tag{13}$$

Equation (13) may also be written as:

$$\mathcal{X} = Z_1 + ... + Z_{\mathcal{N}},$$

which is the representation used in the calculation of the reserve in this thesis.

The probability that the actual loss $\mathcal{X}$ will be greater than a given percentile q_ϵ is formulated as:

$$P(\mathcal{X} > q_\epsilon) = \epsilon. \tag{14}$$

The solvency capital, or the reserve, is denoted by q_ϵ [Bølviken, 2014, p. 6]. EU insurance legislation requires insurance companies to reserve enough capital to cover their liabilities over the following 12 months with a 99,5 % probability [European Comission, 2009]. This is equal to saying that the insurance company must have sufficient capital to cover a potential loss up to the 99.5 percentile. In this thesis, the 99% and the 99.5% reserves will be used to select the claim size distribution and study how the uncertainty in the under-threshold distributions increase with the number of parameters fitted to the distributions. The impact of the uncertainty on the reserve will also be studied.

2.5 Estimating and bootstrapping the reserve

Estimating an upper quantile q_ϵ by Monte Carlo is done by simulating z_1,...,z_n quantities of interest, sorting them in decreasing order $z_{(1)} \geq ... \geq z_{(m)}$ and letting the estimated quantile $\hat{q}^*_{\epsilon,m} = z_{(\epsilon,m)}$.

Algorithm 2 describes the process of estimating the claim reserve by Monte Carlo simulation. First input the Maximum likelihood estimated parameters for the under-threshold distribution $\hat{\boldsymbol{\theta}}$, the Maximum likelihood estimated parameters for the over-threshold distribution $\hat{\alpha}^b$, $\hat{\beta}^b$, the estimated claim intensity $\hat{\mu}$, the distribution f(z), J, T and $p = n_{>b}/n$, where $n_{>b}$ is the

number of claims above the threshold b, and n is the total number of claims. Draw the total number of claims $\hat{\mathcal{N}}^*$ and the number of claims above the threshold $\hat{\mathcal{N}}^*_{>b}$ from the Binomial distribution. Simulate and sum the claims below and above the threshold. Claims are then sorted in decreasing order and the reserve is returned.

Algorithm 2 Algorithm for estimating the reserve.

1: Input $\hat{\boldsymbol{\theta}}, \hat{\alpha}^b, \hat{\beta}^b, \hat{\mu}, J, T, f(z), p = n_{>b}/n$
2: For j = 1,...,m
Draw $\hat{\mathcal{N}}^* \sim$ Poisson$(J\hat{\mu}T)$
Draw $\hat{\mathcal{N}}^*_{>b} \sim$ Binomial$(\hat{\mathcal{N}}^*, p)$
$\hat{\mathcal{N}}^*_{\leq b} \leftarrow \hat{\mathcal{N}}^* - \hat{\mathcal{N}}^*_{>b}$
Draw $\hat{Z}^*_1, ..., \hat{Z}^*_{\hat{\mathcal{N}}^*_{\leq b}} \sim f(z; \hat{\boldsymbol{\theta}})$,
$\hat{Z}^{*b}_1, ..., \hat{Z}^{*b}_{\hat{\mathcal{N}}^*_{>b}} \sim$ Pareto$(\hat{\alpha}^b, \hat{\beta}^b)$ + b
3: $\mathcal{X}^*_j \leftarrow \sum_{k=1}^{\hat{\mathcal{N}}^*_{\leq b}} \hat{Z}^*_k + \sum_{k=1}^{\hat{\mathcal{N}}^*_{>b}} \hat{Z}^{*b}_k$
4: Sort $\hat{\mathcal{X}}^*_1, ..., \hat{\mathcal{X}}^*_m$:
$\hat{\mathcal{X}}^*_{(1)} \geq ... \geq \hat{\mathcal{X}}^*_{(m)}$
5: Return $\hat{q}_\epsilon = \hat{\mathcal{X}}^*_{(\epsilon m)}$.

Algorithm 3 bootstraps the reserve. Step 2 - 4 involves bootstrap estimations of the model parameters $(\mu, \boldsymbol{\theta}, \alpha, \beta)$. For each set of bootstrap estimates $\left(\hat{\mu}^*_i, \hat{\boldsymbol{\theta}}^*_{\boldsymbol{i}}, \hat{\alpha}^*_i, \hat{\beta}^*_i\right)$, j number of Monte Carlo simulations are drawn to estimate the reserve, and based on the bootstrap estimates $\hat{q}^*_{\epsilon,1}, ..., \hat{q}^*_{\epsilon,m_b}$ of the reserve, the influence of the selected distribution on the uncertainty in the reserve can me examined [Bølviken, 2014, p. 248]. Step 6 draws the number of claims from a Poisson distribution with parameter J$\hat{\mu}^\star_i$T. Step 7 draws the number of claims over the threshold from the Binomial distribution with parameters $\hat{\mathcal{N}}^{\star\star}$ from step 6 and p = $n_{>b}/n$ being the number of claims above the threshold and n being the total number of claims. The number of claims below the threshold is denoted $\hat{\mathcal{N}}_{\leq} b^\star$. $\hat{\mathcal{N}}_{\leq} b^\star$ claim sizes are drawn from the claim size distribution with the parameters estimated with Maximum likelihood estimation in step 4. Then $\hat{\mathcal{N}} > b^\star$ claim sizes are drawn from the over-threshold distribution and the threshold b is added. Step 8 sums the number of claims below the threshold and the number of claims above the threshold and adds them together.

Algorithm 3 Algorithm for bootstrapping the reserve.

Input m, m_b, $n_{\leq b}$, $n_{>b}$, $\hat{\mu}, \hat{\boldsymbol{\theta}}$, f(z), $\hat{\alpha}^b$, $\hat{\beta}^b$ A, J, T, p = $n_{>b}/n$
For i = 1,...,m_b
  Draw $\hat{\mathcal{N}}^\star \sim \text{Poisson}(A\hat{\mu})$,
  $\hat{Z}_1^\star, ..., \hat{Z}_{n\leq b}^\star \sim f(z; \hat{\boldsymbol{\theta}})$,
  $\hat{Z}_1^{\star b}, ..., \hat{Z}_{n>b}^{\star b} \sim \text{Pareto}(\hat{\alpha}^b, \hat{\beta}^b)$
  $\hat{\mu}_i^\star \leftarrow \hat{\mathcal{N}}^\star / A$,
  $\hat{\boldsymbol{\theta}}_i^\star \xleftarrow{MLE} \hat{Z}_1^\star, ..., \hat{Z}_{n\leq b}^\star$,
  $\hat{\alpha}_i^{\star b}, \hat{\beta}_i^{\star b} \xleftarrow{MLE} \hat{Z}_1^{\star b}, ..., \hat{Z}_{n>b}^{\star b}$
  For j = 1,...,m
    Draw $\hat{\mathcal{N}}^{\star\star} \sim \text{Poisson}(\text{J}\hat{\mu}_i^\star\text{T})$
    Draw $\hat{\mathcal{N}}_{>b}^\star \sim \text{Binomial}(\hat{\mathcal{N}}^{\star\star}, p)$
    $\hat{\mathcal{N}}_{\leq} b^\star \leftarrow \hat{\mathcal{N}}^{\star\star} - \hat{\mathcal{N}}_{>b}^\star$
    Draw $\hat{Z}_1^{\star\star}, ..., \hat{Z}_{\hat{\mathcal{N}}\leq b^\star}^{\star\star} \sim f(z; \hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^\star)$
    $\hat{Z}_1^{\star\star b}, ..., \hat{Z}_{\hat{\mathcal{N}}>b^\star}^{\star\star b} \sim \text{Pareto}(\hat{\alpha}_i^{\star b}, \hat{\beta}_i^{\star b}) + \text{b}$
    $\mathcal{X}_j^{\star\star} \leftarrow \sum_{k=1}^{\hat{\mathcal{N}}\leq b^\star} \hat{Z}_k^{\star\star} + \sum_{k=1}^{\hat{\mathcal{N}}>b^\star} \hat{Z}_k^{\star\star b}$
    Sort $\hat{\mathcal{X}}_1^{\star\star}, ..., \hat{\mathcal{X}}_m^{\star\star}$:
    $\hat{\mathcal{X}}_{(1)}^{\star\star} \geq ... \geq \hat{X}_{(m)}^{\star\star}$
    Return $\hat{q}_{\epsilon_i}^\star = \hat{\mathcal{X}}_{(\epsilon m)}^{\star\star}$.

2.6 Claim frequency

The claim frequency equals the number of claims arriving within the time interval [0,T] divided by T. These events are treated as random variables, usually modeled by the Poisson or Negative binomial distribution [Gray and Pitts, 2012, p. 11].

The Poisson process is the most common model for claim frequency in general insurance. Claims are assumed to occur at random, one following the other, with a constant intensity. Let N_i denote the number of claims for policy i, and $\mathcal{N} = N_1 + N_2 + ... + N_J$ the number of claims on portfolio level. For $N_1, ..., N_n$ independent Poisson variables with parameters $\lambda_1, ..., \lambda_n$, $\mathcal{N} = N_1 + ... + N_n$ has a Poisson distribution with parameter $\lambda_1 + ... + \lambda_n$, for

proof see [Klugman et al., 2008, p. 103]. The claim numbers at policy and portfolio level follow the Poisson distribution with parameters $\lambda = \mu$T, and $\lambda = $ JμT, respectively. The claim intensity μ is the policy average within the time interval of exposure to risk T, and J is the number of policies [Gray and Pitts, 2012, p. 12]. In this thesis, the claim intensity μ will be set identical for all claims.

The Poisson point mass function is given by:

$$P\left(N = n\right) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2...$$

The mean, standard deviation and skewness are given as:

$$E(N) = \lambda,\ sd(N) = \sqrt{N},\ \text{and}\ skew(N) = \frac{1}{\sqrt{\lambda}}.$$

Consider n claims. The Maximum likelihood estimate of the claim intensity μ is given by:

$$\hat{\mu} = \frac{n}{A}, \tag{15}$$

where n is the total number of claims and A $= T_1 + ... + T_n$ is the total exposure to risk [Bølviken, 2014, p. 284].

The negative binomial distribution may be a better fit than the Poisson distribution for the claim frequency. Consider the mean and variance for the negative binomial distribution:

$$E[N] = \frac{kq}{p}, \qquad Var[N] = \frac{kq}{p^2}, \qquad p < 1.$$

Since p < 1, Var[N] $>$ E[N], which is not possible for the Poisson distribution. As a result, given the same mean as for the Poisson, the negative binomial distribution permits larger uncertainty and a heavier tail than the Poisson distribution. Therefore, the Poisson distribution may not be able to provide a good model for the number of claims. Also the negative binomial distribution may be a better choice for a heavy-tailed claim frequency distribution [Gray and Pitts, 2012, p. 16-17].

As the main focus in this thesis is the claim size distribution, the Poisson distribution will be the only choice of claim frequency distribution.

2.7 Claim size distributions

2.7.1 The Exponential distribution

The Exponential distribution is a special case of the Gamma distribution, see equation (24), when the shape parameter $\alpha = 1$ [Dickson, 2005, p. 6]. The Exponential density is defined as:

$$f(z) = \beta e^{-z\beta}, \qquad z > 0, \tag{16}$$

where $\beta > 0$. The expectation and standard deviation of the Exponential distribution are given by:

$$E[Z] = sd(Z) = \frac{1}{\beta}.$$

As the Exponential distribution only depends on one parameter, the rate, or inverse scale parameter β, it is less flexible than the Gamma and other distributions with more than one parameter. The Exponential distribution may be regarded as a heavy-tailed Gamma and a light-tailed two-parameter Pareto distribution [Bølviken, 2014, p. 322]. In fact, the Exponential distribution is the special case of many distributions (see Appendix $(A.3)$).

The log-likelihood function for the Exponential distribution is specified as:

$$\mathcal{L}(\beta) = n \log(\beta) - \sum_{i=1}^{n} z_i \beta, \qquad \beta > 0. \tag{17}$$

Taking the derivative of equation (17) w.r.t. β, setting the equation equal to zero and solving for $\hat{\beta}$ gives the estimate:

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=1}^{n} z_i \right)^{-1}.$$

2.7.2 The Weibull distribution

The Weibull distribution is a popular distribution in statistical literary work [Kleiber and Kotz, 2003, p. 174]. Its probability density function is given by:

$$f(z) = \frac{\alpha}{\beta}\left(\frac{z}{\beta}\right)^{\alpha-1} \exp(-z/\beta)^{\alpha}, \qquad z > 0, \tag{18}$$

where the parameters $\alpha > 0$ and $\beta > 0$.

The Weibull distribution is connected to the Exponential distribution in the following way:

$$Z = \beta Y^{1/\alpha}, \tag{19}$$

where Y is Exponentially distributed parameter 1.

The expectation and standard deviation are defined as:

$$E(Z) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right), \text{ and}$$
$$sd(Z) = \beta\sqrt{\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma\left(1 + \frac{1}{\alpha}\right)^2}.$$

The log-likelihood function for the Weibull distribution is defined as:

$$\mathcal{L}(\alpha, \beta) = n\Big(\log(\alpha) - \alpha\log(\beta)\Big) + (\alpha - 1)\sum_{i=1}^{n}\log(z_i)$$
$$-\sum_{i=1}^{n}\left(\frac{z_i}{\beta}\right)^{\alpha}, \qquad \alpha, \beta > 0. \tag{20}$$

Taking the derivative of equation (20) with respect to β yields the following equation:

$$\frac{\partial\mathcal{L}}{\partial\beta} = -n\frac{\alpha}{\beta} + \alpha\sum_{i=1}^{n} z_i^{\alpha}\frac{1}{\beta^{\alpha+1}}, \qquad \beta > 0. \tag{21}$$

Setting equation (21) equal to zero and solving for $\hat{\beta}$ gives the MLE estimate for β:

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=1}^{n} z_i^{\hat{\alpha}} \right)^{\frac{1}{\hat{\alpha}}}. \tag{22}$$

Inserting equation (22) in equation (20), gives the log-likelihood function of α:

$$\begin{aligned}
\mathcal{L}(\alpha, \hat{\beta}) = \mathcal{L}(\alpha) &= n \left(\log(\alpha) - \alpha \log \left(\left(\frac{1}{n} \sum_{i=1}^{n} z_i^{\alpha} \right)^{\frac{1}{\alpha}} \right) \right) \\
&\quad + (\alpha - 1) \sum_{i=1}^{n} \log(z_i) - \underbrace{\sum_{i=1}^{n} \left(\frac{z_i}{\left(\frac{1}{n} \sum_{i=1}^{n} z_i^{\alpha} \right)^{\frac{1}{\alpha}}} \right)^{\alpha}}_{=n} \\
&= n \left(\log(\alpha) - \alpha \log \left(\left(\frac{1}{n} \sum_{i=1}^{n} z_i^{\alpha} \right)^{\frac{1}{\alpha}} \right) \right) \\
&\quad + (\alpha - 1) \sum_{i=1}^{n} \log(z_i) - n.
\end{aligned} \tag{23}$$

The MLE estimate $\hat{\alpha}$ may be optimized numerically by maximizing the likelihood function given in equation (23) with respect to α.

The Weibull distribution as a special case

Letting $\tau = \gamma = \theta = 1$ and $\alpha \to \infty$ in equation (58), gives the following equation:

$$Z = \lim_{\theta \to 1} \lim_{\alpha \to \infty} \beta \left(\frac{G_\theta}{G_\alpha} \right)^{\eta} = \beta \Bigg(\underbrace{\lim_{\theta \to 1} G_\theta}_{Y \sim \exp(1)} \times \underbrace{\lim_{\alpha \to \infty} \frac{1}{G_\alpha}}_{=1} \Bigg)^{\eta} = \beta Y^{\eta},$$

where Y $\sim$ exp(1). Therefore Z $\sim$ Weibull(β,η).

2.7.3 The Gamma distribution

The probability distribution function of the Gamma distribution is given by:

$$f(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}, \qquad z > 0, \tag{24}$$

where the Gamma function is defined as:

$$\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz. \tag{25}$$

The parameters $\alpha > 0$ and $\beta > 0$, are the shape and rate parameters respectively [Kleiber and Kotz, 2003, p. 284].

The expectation and standard deviation are specified as:

$$E[Z] = \frac{\alpha}{\beta}, \text{ and } sd(Z) = \frac{\sqrt{\alpha}}{\beta}.$$

On the occasion of $\beta = \alpha$, equation (24) becomes a standard Gamma distribution:

$$f(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-z\alpha}, \qquad z > 0. \tag{26}$$

The log-likelihood function for the Gamma distribution may be written as:

$$\begin{aligned} \mathcal{L}(\alpha, \beta) = n\alpha \log(\beta) + (\alpha - 1) \sum_{i=1}^{n} \log(z_i) - \beta \sum_{i=1}^{n} z_i \\ - n \log\big(\Gamma(\alpha)\big), \end{aligned} \tag{27}$$

where α, $\beta > 0$. The partial derivative with respect to β is given as:

$$\frac{\partial \mathcal{L}}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^{n} z_i, \qquad \beta > 0. \tag{28}$$

Setting equation (28) equal to zero and solving for $\hat{\beta}$ gives the MLE estimate for β:

$$\hat{\beta} = \frac{n\alpha}{\sum_{i=1}^{n} z_i} = \frac{\alpha}{\bar{z}}, \qquad \bar{z} > 0. \tag{29}$$

Substituting equation (29) into equation (27), leads to the log likelihood function, depending only on α:

$$\mathcal{L}(\alpha, \hat{\beta}) = \mathcal{L}(\alpha) = n\alpha \log\left(\frac{\alpha}{\bar{z}}\right) + (\alpha - 1)\sum_{i=1}^{n} \log(z_i) - \frac{\alpha}{\bar{z}}\sum_{i=1}^{n} z_i - n\log\big(\Gamma(\alpha)\big), \tag{30}$$

where $\alpha > 0$ and for i = 1 to n, $z_i > 0$. The MLE $\hat{\alpha}$ may be obtained by maximizing equation (30) with respect to α, using numerical methods.

The Gamma distribution as a special case

Letting $\tau = \gamma = \eta = 1$, and $\alpha \to \infty$ in equation (58), yields:

$$Z = \lim_{\alpha \to \infty} \frac{\beta G_\theta}{G_\alpha} \to \beta G_\theta, \tag{31}$$

which is a Gamma distributed random variable with parameters, θ and β (see Appendix (A.1)).

2.7.4 The Log-gamma distribution

The Log-gamma distribution is a transformation of the Gamma distribution. When $\log(Z)$ is Gamma distributed, Z is Log-gamma distributed. The Log-gamma probability density function is given as (see Appendix $(A.4)$):

$$f(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} \log(z)^{\alpha-1} z^{-\beta-1}, \qquad z > 0, \tag{32}$$

where $\alpha > 0$ and $\beta > 0$ are both shape parameters [Kleiber and Kotz, 2003, p. 169].

The mean and standard deviation are given as:

$$E(Z) = \left(1 - \frac{1}{\beta}\right)^{-\alpha}, \qquad \beta > 1,$$

$$sd(Z) = \sqrt{\left(1 - \frac{2}{\beta}\right)^{-\alpha} - \left[\left(1 - \frac{1}{\beta}\right)^{-\alpha}\right]^2}, \qquad \beta > 2.$$

Substituting z_i by $\log(z_i)$ in equation (29) and (30) in section (2.7.3) gives the log-likelihood estimate of β and log-likelihood function of α:

$$\hat{\beta} = \frac{\alpha}{\frac{\sum_{i=1}^{n} \log(z_i)}{n}} = \frac{\alpha}{\bar{z}_{\log}}, \qquad \text{where } z_i > 0, \text{ for each i = 1 to n, and} \tag{33}$$

$$\mathcal{L}(\alpha, \hat{\beta}) = \mathcal{L}(\alpha) = n\alpha \log\left(\frac{\alpha}{\bar{z}_{\log}}\right) + (\alpha - 1)\sum_{i=1}^{n} \log(\log(z_i)) - \frac{\alpha}{\log(\bar{z})}\sum_{i=1}^{n} \log(z_i) - n \log\big(\Gamma(\alpha)\big), \tag{34}$$

where $\alpha > 0$ and for i = 1 to n, $z_i > 0$.

Numerical methods may be applied to maximize equation (34) with respect to α.

The Log-gamma distribution as a special case

Letting $\beta = \eta$ =1, and $\tau = \frac{\gamma}{\xi}$ and letting α and γ go to infinity in equation (58), gives the following equation:

$$Z = \lim_{\gamma\to\infty}\lim_{\alpha\to\infty}\left(1 + \frac{\xi G_\theta}{\gamma G_\alpha}\right)^{\gamma} - 1 \quad \to \lim_{\gamma\to\infty}\left(1 + \frac{\xi G_\theta}{\gamma}\right)^{\gamma} - 1 \to \exp(G_\theta \xi) - 1,$$

where ξG_θ is a Gamma distributed random variable. Therefore Z + 1 is a Log-gamma distributed random variable with parameters θ and ξ (see Appendix (A.4)).

2.7.5 The Log-normal distribution

The Log-normal distribution is a transformation of the Normal distribution. When $\log(Z)$ is normally distributed, Z is Log-normal. The Log-normal probability density function is given by:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{z}\exp\left(-\frac{1}{2}\left(\frac{\log z - \mu}{\sigma}\right)^2\right), \qquad z > 0. \tag{35}$$

On a logarithmic scale, the parameters μ and σ are the location and scale parameters, respectively [Gray and Pitts, 2012, p. 35]. The mean and standard deviation are given by:

$$E(Z) = \exp\left(\mu + \frac{1}{2}\sigma^2\right), \text{ and}$$
$$sd(Z) = E[Z]\sqrt{\exp\left(\sigma^2\right) - 1}.$$

The Log-normal log-likelihood function may be expressed in the following way:

$$\begin{aligned}\mathcal{L}(\mu,\sigma) = &-n\log(\sigma) - \frac{n}{2}\log(2\pi) - \sum_{i=1}^{n}\log(z_i) \\ &- \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\log(z_i) - \mu}{\sigma}\right)^2, \qquad \sigma, z_i > 0.\end{aligned} \tag{36}$$

By setting

$$\frac{\partial\mathcal{L}}{\partial\mu} = \sum_{i=1}^{n}\log(z_i) - n\mu = 0, \qquad z_i > 0, \text{ and}$$
$$\frac{\partial\mathcal{L}}{\partial\sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{1}{\sigma^3}\big(\log(z_i) - \mu\big)^2 = 0, \qquad \sigma, z_i > 0,$$

and solving for μ and σ, the following estimates are obtained:

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\log(z_i)}{n}, \text{ and } \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}\big(\log(z_i) - \hat{\mu}\big)^2}{n}}. \tag{37}$$

The Log-normal distribution as a special case

From equation (58) we have

$$\log(z+\beta) = \log(\beta) + \gamma \log\left(1 + \frac{X^\eta}{\tau}\right), \tag{38}$$

where X=$\dfrac{G_\theta}{G_\alpha}$ is the ratio of two independent standard Gamma variables. Now, let $\tau = \sqrt{\gamma}$, $\eta = 1$, and $\beta = \exp(-\sqrt{\gamma} + \xi + \frac{1}{2})$ in equation (38):

$$\log(z+\beta) = -\sqrt{\gamma} + \xi + \frac{1}{2} + \underbrace{\gamma \log\left(1 + \frac{X}{\sqrt{\gamma}}\right)}_{(*)}$$

Using a Taylor expansion on $(*)$ gives:

$$\gamma \log\left(1 + \frac{X}{\sqrt{\gamma}}\right) = X\sqrt{\gamma} - \frac{X^2}{2} + O\left(\frac{X^3}{\sqrt{\gamma}}\right).$$

Let $\gamma = \sigma^2\theta$:

$$z = \exp\left(\xi + \sigma\sqrt{\theta}(X-1) - \frac{1}{2}(X^2-1) + O\left(\frac{X^3}{\sigma\sqrt{\theta}}\right)\right) - \beta. \tag{39}$$

$$\sqrt{\theta}(X-1) = \sqrt{\theta}(G_\theta - 1) + \epsilon, \text{ where } \epsilon = \sqrt{\theta}G_\theta\left(\frac{1}{G_\alpha} - 1\right).$$

Let

$$U = \sqrt{\theta}(G_\theta - 1) = Y - \sqrt{\theta},$$

where Y is Gamma distributed with parameters θ and $\sqrt{\theta}$. Moreover

$$\begin{aligned} M_U(t) &= E[e^{tU}] = E[e^{t(Y-\sqrt{\theta})}] \\ &= e^{-t\sqrt{\theta}} E[e^{tY}] \\ &= e^{-t\sqrt{\theta}} \left(\frac{1}{1 - \frac{t}{\sqrt{\theta}}}\right)^\theta. \end{aligned} \tag{40}$$

Taking the logarithm of equation (40):

$$\begin{aligned}\log\left(M_U(t)\right) &= -t\sqrt{\theta} - \theta\log\left(1 - \frac{t}{\sqrt{\theta}}\right) \\ &= -t\sqrt{\theta} - \theta\left[-\frac{t}{\sqrt{\theta}} - \frac{1}{2}\frac{t^2}{\theta} - O\left(\theta^{-\frac{3}{2}}\right)\right] \\ &= \frac{1}{2}t^2 + O\left(\frac{1}{\sqrt{\theta}}\right).\end{aligned}$$

Thus $M_U(t) \to e^{\frac{t^2}{2}}$ which is the standard Normal moment generating function.

From equation (40), we have

$$\xi^2 = \theta G_\theta\left(\frac{1}{G_{\alpha^2}} - \frac{2}{G_\alpha} + 1\right).$$

Since G_θ and H_α are independent, it follows that:

$$E[\xi^2] = \theta E[G_\theta^2]\left(E[H_\alpha^2] - 2E[H_\alpha] + 1\right),$$

where H_α is an inverse-Gamma variable with parameters α and α. Hence,

$$\begin{aligned}E\left(\xi^2\right) &= \frac{(\theta+1)}{(\alpha-1)(\alpha-2)})\left(\alpha^2 - 2\alpha(\alpha-2) + (\alpha-1)(\alpha-2)\right) \\ &= \frac{(\theta+1)(\alpha+2)}{(\alpha-1)(\alpha-2)}. \qquad (41)\end{aligned}$$

Letting $\theta \to \infty$, and $\dfrac{\alpha}{\theta} \to \infty$, yield $\mathrm{E}[\xi^2] \to 0$.

As $\mathrm{E}[\xi^2] \to 0$, $(E[\xi])^2$ must also converge to zero as the variance cannot be negative. Hence ξ itself must converge to zero. This means that $\sqrt{\theta}\,(X-1) = U + \xi \to \mathrm{N}$, where N is standard normal.

Consider $\dfrac{1}{2}(X^2-1)$ in equation (39). As $\theta \to \infty$, $\alpha \to \infty$, $G_\theta \to 1$ and $G_\alpha \to 1$, we have

$$\frac{1}{2}(X^2 - 1) \to 0.$$

Finally, $\beta \to 0$ as $\gamma \to \infty$ and $O\left(\dfrac{X^3}{\sigma\sqrt{\theta}}\right)$ vanishes as $\theta \to \infty$.

What remains of equation (39) is

$$z = e^{\xi + \sigma N},$$

which is a Log-normal random variable.

2.7.6 The two-parameter Pareto distribution

The two-parameter Pareto distribution is commonly used in claim size modeling [Daykin et al., 1994, p. 89]. Due to its heavy tail, measured by α, the two-parameter Pareto distribution is a favored distribution in property insurance [Kleiber and Kotz, 2003, p. 59]. The probability density function of the two-parameter Pareto distribution is given by:

$$f(z) = \frac{\alpha\beta^{\alpha}}{(\beta + z)^{\alpha+1}}, \qquad z > 0, \tag{42}$$

where the parameters α and β are strictly positive.

The cumulative distribution function is defined as:

$$F(z) = 1 - \frac{\beta^{\alpha}}{(\beta + z)^{\alpha}}, \qquad z > 0. \tag{43}$$

For large values of α, the two-parameter Pareto distribution tends to the thin tailed Exponential distribution (see Appendix $(A.3)$):

$$\lim_{\alpha \to \infty} \frac{\alpha\beta^{\alpha}}{(\beta + z)^{\alpha+1}} = \beta \exp(-z\beta). \tag{44}$$

On the other hand, the smaller α, the heavier is the tail of the distribution. The skewness is also regulated by α [Kleiber and Kotz, 2003, p. 59], while β is the scale parameter [Gray and Pitts, 2012, p. 41].

The mean and standard deviation are given by:

$$E[Z] = \frac{\beta}{\alpha - 1}, \qquad \alpha > 1, \text{ and}$$
$$sd(Z) = E[Z]\sqrt{\frac{\alpha}{\alpha - 2}}, \qquad \alpha > 2.$$

The log-likelihood function for the two-parameter Pareto distribution is specified as:

$$\mathcal{L}(\alpha, \beta) = n\log(\alpha) + n\alpha\log(\beta) - (\alpha + 1)\sum_{i=1}^{n}\log(\beta + z_i),$$
$$\alpha, \beta > 0 \tag{45}$$

Setting

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left(\frac{\beta}{\beta + z_i}\right) = 0, \qquad \alpha > 0, \tag{46}$$

and solving for α yields the estimate:

$$\hat{\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n}\log\left(1 + \frac{z_i}{\beta}\right)\right)^{-1}. \tag{47}$$

Inserting equation (47) into equation (45) returns the log-likelihood function of β:

$$L(\hat{\alpha}, \beta) = \mathcal{L}(\beta) = n\log\left(\frac{n}{\sum_{i=1}^{n}\log\left(\frac{\beta + z_i}{\beta}\right)}\right)$$
$$+ (n-1)\sum_{i=1}^{n}\log(z_i + \beta) + n^2\log(\beta). \tag{48}$$

Equation (48) is a function of β only, and may be optimized with respect to β by the use of numerical methods.

The two-parameter Pareto distribution as a special case

Letting $\tau = \gamma = \theta = \eta = 1$ in equation (59), gives the following equation for the random variable Z:

$$Z = \alpha X,$$

where X is Extended Pareto distributed.

Using the following formula to calculate $f_z(z)$:

$$f_z(z) = \frac{1}{\alpha} f_x\left(\frac{z}{\alpha}\right), \tag{49}$$

where X is a continuous variable with probability density function $f_x(x)$ and cumulative distribution function $F_x(x)$ [Klugman et al., 2008, p. 62].

The following equation is obtained:

$$f_z(z) = \frac{1}{\beta} \underbrace{\frac{\Gamma(\alpha+1)}{\alpha\Gamma(\alpha)}}_{=1} \frac{1}{\left(1+\frac{z}{\alpha\beta}\right)^{\alpha+1}}. \tag{50}$$

Multiplying the numerator and denominator in equation (50) by $(\alpha\beta)^{\alpha+1}$ gives:

$$f_z(z) = \frac{\alpha(\alpha\beta)^{\alpha}}{(\alpha\beta+z)^{\alpha+1}} \qquad \text{where } \alpha, \beta, z > 0. \tag{51}$$

This is the two-parameter Pareto distribution with parameters $(\alpha, \alpha\beta)$.

2.7.7 The Extended Pareto distribution

The Extended Pareto is a generalized version of the two-parameter Pareto distribution with three parameters. The probability density function is given by:

$$f(z) = \frac{1}{\beta} \frac{\Gamma(\alpha+\theta)}{\Gamma(\theta)\Gamma(\alpha)} \frac{(z/\beta)^{\theta-1}}{(1+z/\beta)^{\alpha+\theta}}, \qquad x > 0, \tag{52}$$

where the shape parameter α, and the scale parameters β and θ are all strictly positive [Bølviken, 2014, p. 335].

When $\theta = 1$, the distribution function becomes the two-parameter Pareto distribution:

$$f(z) = \frac{\alpha/\beta}{(1 + z/\beta)^{1+\alpha}} = \frac{\alpha\beta^{\alpha}}{(\beta + z)^{1+\alpha}}, \qquad x > 0.$$

In Appendix $(A.2)$ it is shown that the Gamma distribution can be derived as a limit of the Extended Pareto distribution.

The mean and standard deviation of the Extended Pareto distribution are defined as (see Appendix $(A.1)$):

$$E(Z) = \frac{\theta\beta}{\alpha - 1}, \qquad \alpha > 1,$$
$$sd(Z) = E(Z)\left(\frac{\alpha + \theta - 1}{\theta(\alpha - 2)}\right)^{1/2}, \qquad \alpha > 2.$$

The estimated parameters $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$ may be derived by numerically optimizing the log-likelihood function given by:

$$\begin{aligned}&\mathcal{L}(\alpha, \beta, \theta) = \\ &n\left[\log\big(\Gamma(\alpha + \theta)\big) - \log\big(\Gamma(\alpha)\big) - \log\big(\Gamma(\theta)\big) - \theta\log\big(\beta\big)\right] \\ &+ (\theta - 1)\sum_{i=1}^{n}\log(z_i) - (\alpha + \theta)\sum_{i=1}^{n}\log\left(1 + \frac{z_i}{\beta}\right),\end{aligned} \tag{53}$$

where α, β, $\theta > 0$.

2.7.8 The four-parameter Pareto distribution

One obtains what is denoted as the four-parameter Pareto distribution in this thesis by letting $\gamma = \tau = 1$ in equation (64). The probability density function is given by:

$$f_z(z) = \left(\frac{\theta}{\alpha}\right)^{\theta} \frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)} \frac{1}{\beta\eta} \frac{\left(\frac{z}{\beta}\right)^{\frac{\theta}{\eta}-1}}{\left(1+\left(\frac{z}{\beta}\right)^{\frac{1}{\eta}} \frac{\theta}{\alpha}\right)^{\alpha+\theta}}. \tag{54}$$

The mean and standard deviation are given by (see Appendix A.5):

$$E[Z] = \beta \left(\frac{\alpha}{\theta}\right)^{\eta} \frac{\Gamma(\alpha-\eta)\Gamma(\theta+\eta)}{\Gamma(\alpha)\Gamma(\theta)}, \qquad \alpha-\eta>0$$

$$sd[Z] = \frac{\beta \left(\frac{\alpha}{\theta}\right)^{\eta}}{\Gamma(\alpha)\Gamma(\theta)} \times \sqrt{\Gamma(\alpha)\Gamma(\theta)\Gamma(\alpha-2\eta)\Gamma(\theta+2\eta) - (\Gamma(\alpha-\eta)\Gamma(\theta+\eta)^2}.$$

The log-likelihood function for the four-parameter Pareto distribution is defined as:

$$\mathcal{L}(\alpha,\theta,\beta,\eta) =$$

$$n\theta\left(log(\theta) - log(\alpha)\right) + n\log\left(\Gamma(\alpha+\theta)\right) - n\log(\beta) - n\log(\eta) \tag{55}$$

$$- n\log\left(\Gamma(\alpha)\right) - n\log\left(\Gamma(\theta)\right) + \left(\frac{\theta}{\eta}-1\right)\sum_{i=1}^{n}\log\left(\frac{z_i}{\beta}\right) \tag{56}$$

$$- (\alpha+\theta)\sum_{i=1}^{n}\log\left(1+\left(\frac{z_i}{\beta}\right)^{\frac{1}{\eta}}\frac{\theta}{\alpha}\right), \tag{57}$$

providing α, β, η, $\theta > 0$.

2.7.9 The six-parameter Pareto distribution

The following distribution, that is denoted the six-parameter Pareto distribution in this thesis, is based on a parameterized power transformation [Box and Cox, 1964], with six parameters.

$$Z = \beta\left(\left(1+\frac{X^{\eta}}{\tau}\right)^{\gamma}-1\right), \tag{58}$$

where

$$X = \frac{G_{\theta}}{G_{\alpha}}.$$

G_{θ} and G_{α} are independent variables belonging to the standard Gamma distribution with parameters θ and α, respectively. The six-parameter Pareto distribution has many special cases, among those are the following distributions: the two-parameter Pareto, Extended Pareto, four-parameter Pareto, Gamma, Weibull, Log-Gamma, Log-normal and the Exponential distribution.

Let $X^{*} = \frac{\theta\beta G_{\theta}}{\alpha G_{\alpha}}$. This variable is Extended Pareto distributed [Bølviken, 2014, p. 324]. The random variable Z in (58) becomes:

$$Z = \beta\left\{\left(1+\frac{1}{\tau}\left(\frac{\alpha X^{*}}{\theta\beta}\right)^{\eta}\right)^{\gamma}-1\right\}. \tag{59}$$

The following formula will be used to find the distribution function of Z:

$$f_z(z) = f_x\Big(g^{-1}(z)\Big)\left|\frac{\partial g^{-1}(z)}{\partial z}\right|, \tag{60}$$

where Z = g(x) is a monotonic function (strictly increasing or decreasing) [Devore and Berk, 2012, p. 221].

The inverse function of Z is given by:

$$X = g^{-1}(z) = \frac{\theta\beta}{\alpha}\left\{\tau\left[\left(\frac{z}{\beta}+1\right)^{\frac{1}{\gamma}}-1\right]\right\}^{\frac{1}{\eta}}, \tag{61}$$

with derivative

$$\frac{\partial g^{-1}(z)}{\partial z} = \frac{\theta\tau}{\alpha\eta\gamma}\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}-1}\Big\{\tau\Big(\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}}-1\Big)\Big\}^{\frac{1}{\eta}-1}. \tag{62}$$

It is known that the random variable X^* is Extended Pareto distributed with distribution function given by equation (52) in section (2.7.7). Substituting equation (61) into equation (52) gives the distribution function of X^* $= g^{-1}(z)$:

$$f_x(g^{-1}(z)) = \frac{1}{\beta}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\frac{\Big\{\Big(\Big(\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}}-1\Big)\tau\Big)^{\frac{1}{\eta}}\frac{\theta}{\alpha}\Big\}^{\theta-1}}{\Big\{1+\Big[\Big(\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}}-1\Big)\tau\Big]^{\frac{1}{\eta}}\frac{\theta}{\alpha}\Big\}^{\alpha+\theta}}, \tag{63}$$

where $\beta, \alpha, \theta > 0$ and $z > 0$. For justification see Appendix $(A.1)$.

Inserting equation (62) and (63) into equation (60) gives the probability density function of Z:

$$f_z(z) = \Big(\frac{\theta}{\alpha}\Big)^{\theta}\frac{\tau^{\frac{\theta}{\eta}}}{\beta\eta\gamma}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\frac{\Big(\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}}-1\Big)^{\frac{\theta}{\eta}-1}}{\Big\{1+\Big[\Big(\Big(\frac{z}{\beta}+1\Big)^{\frac{1}{\gamma}}-1\Big)\tau\Big]^{\frac{1}{\eta}}\frac{\theta}{\alpha}\Big\}^{\alpha+\theta}\Big(\frac{z}{\beta}+1\Big)^{1-\frac{1}{\gamma}}}. \tag{64}$$

The log-likelihood function for the six-parameter Pareto distribution is given by:

$$
\begin{aligned}
&\mathcal{L}(\gamma, \eta, \tau, \alpha, \theta, \beta) = \\
&\left(\frac{\theta}{\eta} - 1\right) \sum_{i=1}^{n} \log\left(\left(\frac{z_i}{\beta} + 1\right)^{\frac{1}{\gamma}} - 1\right) + \left(\frac{1}{\gamma} - 1\right) \sum_{i=1}^{n} \log\left(\frac{z_i}{\beta} + 1\right) \\
&- n\Big(\log(\beta) + \log(\eta) + \log(\gamma) - \theta \log(\theta) + \theta \log(\alpha) - \frac{\theta}{\eta} \log(\tau) \\
&- \log\left(\Gamma\left(\alpha + \theta\right)\right) + \log\left(\Gamma\left(\alpha\right)\right) + \log\left(\Gamma\left(\theta\right)\right) \Big) \\
&- \left(\alpha + \theta\right) \sum_{i=1}^{n} \log\left(1 + \left(\left(\frac{z_i}{\beta} + 1\right)^{\frac{1}{\gamma}} - 1\right)^{\frac{1}{\eta}} \frac{\theta}{\alpha} \tau^{\frac{1}{\eta}}\right), \qquad (65)
\end{aligned}
$$

providing α, θ, β, η, $\tau > 0$, and $\gamma \geq 1$ for the distribution to be unimodal.

The parameters (α, θ, β, η, τ, γ) can be estimated by optimizing the log-likelihood function.

2.7.10 Modeling extreme claim sizes

The Pareto family has a unique performance in its extreme right tail. This is a very useful characteristic when dealing with excess claim sizes. Extreme claims exceeding a threshold b, are commonly modeled by the (Generalized) Pareto distribution [Beirlant et al., 2009]. The (Generalized) Pareto distribution has two special case distributions, the two-parameter Pareto distribution and the Exponential distribution. The interested reader if referred to [Hosking and Wallis, 1987]. To obtain a better estimate of the claim size distribution, the claims may be divided in two, where claims up to and including a threshold b, will be fitted to the six-parameter Pareto distribution, or one of its special cases. The remaining claims will be modeled as a (Generalized) Pareto distribution.

Let Z be a sequence of mutually independent random variables, representing insurance claims, with continuous distribution function F(z)<1, $\forall$ z. Setting some threshold b, gives the over-threshold distribution:

$$Z_b = Z - b \mid Z \geq b.$$

According to Pickand's theorem Z_b becomes (Generalized) Pareto distributed as b $\to \infty$ [Pickands III, 1975] and [Bølviken, 2014, p. 325-328]. Hence Z_b

= Z - b | $Z > b$, will be modeled by the (Generalized) Pareto distribution [Bølviken, 2014, p. 326]. The remaining question is where to set the threshold b. To be able to apply Pickand's theorem, b must be sufficiently large. However, as b $\to \infty$, the number of claims above b decreases and the parameters may become hard to estimate. Given an iid sample $Z_1, ..., Z_n$, a frequently used technique to determine the threshold in risk analysis is the mean excess plot. The mean excess function is given by:

$$\hat{M}(b) = \frac{\sum_{i=1}^{n} (Z_i - b)\, I_{[Z_i > b]}}{\sum_{i=1}^{n} I_{[Z_i > b]}},$$

for $0 \leq b < \max(Z_1, ..., Z_n)$. The points $\left\{ \left(Z_{(k)}, \hat{M}(Z_{(k)}) \right) : 1 < k \leq n \right\}$ are plotted, where $Z_{(1)} \geq ... \geq Z_{(n)}$ are the claim sizes sorted in decreasing order. The underlying data in the over-threshold distribution should show a pattern of linearity for large values of the threshold b [Ghosh and Resnick, 2011].

2.7.11 Claims below the threshold

For all the claim sizes below the threshold b, the following formulas will be used. Consider the data Z | Z $\leq$ b, where b is the threshold. The conditional probability density function is given by:

$$f(z \mid z \leq b) = \frac{f(z)}{P(z \leq b)} = \frac{f(z)}{F(b)}, \qquad z \leq b. \tag{66}$$

The conditional cumulative distribution function is given by:

$$F(z \mid z \leq b) = \frac{F(z)}{F(b)}, \qquad z \leq b, \tag{67}$$

where equation (66) and (67) are the truncated probability density and cumulative distribution functions, respectively.

The Maximum likelihood function for the truncated data is given by:

$$\begin{aligned} \mathcal{L}_b(\boldsymbol{\theta}; \boldsymbol{z} \leq b) &= \sum_{i=1}^{n} \log\left(f(z_i \mid z_i \leq b; \boldsymbol{\theta})\right) \\ &= \sum_{i=1}^{n} \log\left(f(z_i; \boldsymbol{\theta})\right) - n \log\left(F(b; \boldsymbol{\theta})\right) \\ &= \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{z}) - n \log\left(F(b; \boldsymbol{\theta})\right). \end{aligned} \tag{68}$$

The quantile function for the truncated distribution becomes:

$$F(z \mid z \leq b) = \frac{F(z)}{F(b)} = u,$$
$$F(z) = uF(b). \tag{69}$$

The inverse of equation (69):

$$z = F^{-1}\left(uF(b)\right),$$

is the quantile function used to compute the qq-plot.

Sections 2.7.1 - 2.7.9 have not taken truncation into account. The truncated log-likelihood function may be obtained by inserting the log-likelihood function of interest in equation (68). The non-truncated estimators may be used as starting values when optimizing equation (68).

3 Results

Estimation in R

The parameters for the under-threshold distributions are first fitted by Maximum likelihood estimation assuming the non-truncated distribution, where the start values are determined by the method of moments. The optimization is executed using the 'optim' function in R. For the Exponential distribution, the mean is used to find a start value for β. The two-parameter distributions use the mean and the variance to set the start values for the two parameters of interest. The Extended Pareto distribution uses the mean, variance and skewness. The four-parameter Pareto distribution takes the estimated parameters $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\beta}$ from the Extended Pareto distribution and sets $\eta = 1$ as start values. The six-parameter Pareto distribution uses $\hat{\alpha}$, $\hat{\theta}$, $\hat{\beta}$ and $\hat{\eta}$ estimated by the four-parameter Pareto distribution, with $\tau = \gamma = 1$.

The parameter estimates obtained for the non-truncated distributions are then used as start values for optimization of the truncated distributions. The optimization is performed using the 'optim' function in R. When using 'optim', the logarithm of the parameters is estimated. The reason for this is to enforce positive values. The parameters are then transformed back by taking the exponential of the estimates. The exception is the parameter γ in the six-parameter Pareto distribution that needs to be ≥ 1 for the six-parameter distribution to be unimodal. The function 'optim' estimates the transformed gamma parameter represented by $\hat{\phi} = -\log(\hat{\gamma}) - \log(1 - 1/\hat{\gamma})$. The parameter is then transformed back by the transformation $\hat{\gamma} = \dfrac{1}{\exp\left(\hat{\phi}\right) + 1}$.

The over-threshold distribution applies Maximum likelihood estimation using the mean and variance to find start values for α and β for the Pareto distribution, for $Z_b = (Z - b | Z > b)$. The function 'optim' is used for optimization.

Data and claim frequency

The data set consists of property insurance claims on office and industrial buildings from Januar 2001 to October 2015 from Gjensidige forsikring. The total exposure to risk A = 213,700 policy year. The resulting claim intensity estimate is $\hat{\mu} = \frac{n}{A} = 0.03$. After removing claims < 0 from the data set, 6,411 claim sizes remain.

Table 1 presents a summary of the total data. The data ranges from 1 to 1.055×10^8 NOK, which indicates a great spread in the total claim size data. The middle value is at 3.765×10^4 NOK, which reveals a heavy right tail in the data. The standard deviation $\sigma = 3.012\times10^6$ indicates a great variation in the data set. The skewness = 1.632×10, and the kurtosis = 3.995×10^2, which again reveal a heavy tail in the claim size distribution. The data appears to have extreme values.

Summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	1.095×10^4	3.765×10^4	5.605×10^5	1.249×10^5	1.055×10^8

Table 1: Five-number summary of the data without threshold.

3.1 The claim size distributions

The threshold

The estimated mean excess function of the sorted claims is plotted against the sorted claims in figure 1 below, as described in section 2.7.10. The threshold level b is determined from this plot. Figure 1 shows a roughly linear trend from around 4×10^6 NOK. Setting the threshold at 4×10^6, gives the

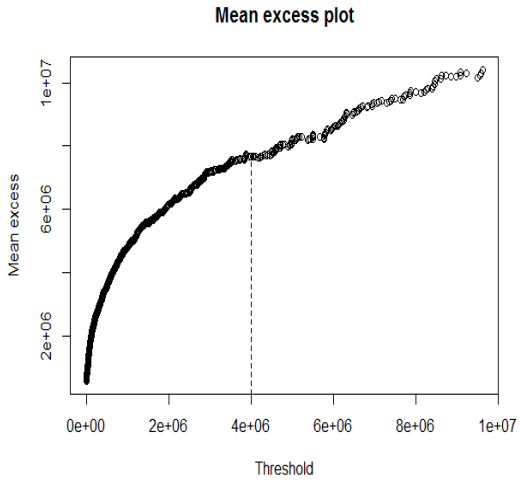


Figure 1: Mean excess plot.

over-threshold distribution 207 data points to be fitted to the (Generalized) Pareto distribution. Removing the 207 claims exceeding the threshold level, leaves 6,204 claims to be fitted to the under-threshold distribution.

3.1.1 The under-threshold claim size data

Table 2 shows a summary of the under-threshold claim size data.

Summary

Min.	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max.**
1.000	1.043×10^4	3.527×10^4	1.910×10^5	1.039×10^5	3.969×10^6

Table 2: Five-number summary of the under-threshold data.

Table 2 shows that the under-threshold empirical model has mean = 1.910×10^5. The standard deviation = 4.873×10^5 which reveals a great amount of uncertainty. The data does not appear to be Exponentially distributed as the standard deviation is greater than the mean. The skewness = 4.417, which

implies that the under-threshold model is more symmetric about its mean than the total empirical model. This is expected as the extreme claims have been removed. The kurtosis = 2.186×10 has been reduced compared to that of the total empirical model. The under-threshold model has a much lighter tail than the total empirical model, which is reasonable as the extreme right tail has been removed.

Comparing the spread between the minimum and maximum value, and the 1st quantile and the 3rd quantile in table 1, with table 2, shows that the spread has been notably reduced in the under-threshold model. The median is also more centered in table 2. However, this is expected as the extreme claims have been removed.

Table 3 displays the Maximum likelihood estimates for claims below the threshold b and a 95% confidence interval for the parameters based on 10,000 bootstrap samples to estimate the error in the fitted parameters. There is no need for adjustments for correction in the bootstrap estimates as $\bar{\hat{\boldsymbol{\theta}}}$ is very close to the vector of Maximum likelihood estimates $\hat{\boldsymbol{\theta}}$ for all distributions.

The lower and upper 95% confidence intervals lie quite close for the Exponential distribution which means that the Maximum likelihood estimate has little uncertainty. This is expected as the Exponential distribution has only one parameter to be estimated from 6,204 observations.

The confidence intervals for both parameters in all the two-parameter distributions are narrow, which indicates little estimation uncertainty. Again this is no surprise as there are only two parameters to be estimated from a large amount of data.

The Extended Pareto distribution displays narrow confidence intervals for all three bootstrap simulated parameters, which implies little uncertainty in the Maximum likelihood estimations. When $\theta = 1$, the Extended Pareto distribution equals the two-parameter Pareto distribution. Table 3 shows that the confidence interval for $\hat{\theta}$, $[1.462, 1.702]$, does not include 1 and $\hat{\theta}$ is therefore significantly different from 1. In other words, the fitted Extended Pareto distribution is significantly different from the two-parameter Pareto distribution.

Maximum likelihood estimation

Distribution	Parameters with confidence intervals			
Exponential	$\hat{\beta} = 5.236\times10^{-6}$	$[5.127 \times 10^{-6}, 5.347 \times 10^{-6}]$		
Weibull	$\hat{\alpha} = 5.429\times10^{-1}$	$[5.339 \times 10^{-1}, 5.522 \times 10^{-1}]$	$\hat{\beta} = 9.352\times10^{4}$	$[8.983 \times 10^{4}, 9.738 \times 10^{4}]$
Gamma	$\hat{\alpha} = 4.030\times10^{-1}$	$[3.936 \times 10^{-1}, 4.128 \times 10^{-1}]$	$\hat{\beta} = 2.109\times10^{-6}$	$[2.026 \times 10^{-6}, 2.196 \times 10^{-6}]$
Log-gamma	$\hat{\alpha} = 3.170\times10$	$[3.081 \times 10, 3.267 \times 10]$	$\hat{\beta} = 3.010\times10$	$[2.924 \times 10, 3.103 \times 10]$
Log-normal	$\hat{\mu} = 1.056\times10$	$[1.052 \times 10, 1.060 \times 10]$	$\hat{\sigma} = 1.832$	$[1.802, 1.862]$
2-par. Pareto	$\hat{\alpha} = 7.108\times10^{-1}$	$[6.725 \times 10^{-1}, 7.516 \times 10^{-1}]$	$\hat{\beta} = 2.242\times10^{4}$	$[2.053 \times 10^{4}, 2.450 \times 10^{4}]$
Ext. Pareto	$\hat{\alpha} = 6.051\times10^{-1}$	$[5.731 \times 10^{-1}, 6.395 \times 10^{-1}]$	$\hat{\theta} = 1.571$	$[1.462, 1.702]$
	$\hat{\beta} = 2.381\times10^{4}$	$[2.244 \times 10^{4}, 2.524 \times 10^{4}]$		
4-par. Pareto	$\hat{\alpha} = 4.793\times10^{-1}$	$[3.713 \times 10^{-1}, 6.318 \times 10^{-1}]$	$\hat{\theta} = 1.234$	$[9.481 \times 10^{-1}, 1.688]$
	$\hat{\beta} = 2.280\times10^{4}$	$[2.124 \times 10^{4}, 2.448 \times 10^{4}]$	$\hat{\eta} = 8.510\times10^{-1}$	$[7.070 \times 10^{-1}, 1.031]$
6-par. Pareto	$\hat{\alpha} = 4.795\times10^{-1}$	$[3.774 \times 10^{-1}, 7.024 \times 10^{-1}]$	$\hat{\theta} = 1.234$	$[0.950, 1.824]$
	$\hat{\beta} = 2.566\times10^{4}$		$\hat{\eta} = 8.512\times10^{-1}$	$[0.715, 1.085]$
	$\hat{\tau} = 1.125$		$\hat{\gamma} = 1.000$	$[1.000, 1.116]$
	$\frac{\hat{\beta}}{\hat{\tau}} = 2.281\times10^{4}$	$[2.028 \times 10^{4}, 2.462 \times 10^{4}]$		

Table 3: Maximum likelihood estimated parameters with 95% confidence intervals for the under-threshold distributions.

The confidence intervals for the four-parameter Pareto distribution show that the uncertainty in all the parameters has increased from the confidence intervals for the Extended Pareto distribution, which is normal as the uncertainty increases with the number of parameters. The Extended Pareto distribution is a special case of the four-parameter Pareto distribution. The Extended Pareto distribution occurs when $\eta = 1$. The 95% CI for $\hat{\eta}$, $[0.710, 1.026]$, indicates that $\hat{\eta}$ is not significantly different from 1. Including the additional parameter η is of no benefit. The other estimated parameters do not differ much from those estimated with the Extended Pareto distribution.

The confidence intervals for the six-parameter Pareto distribution indicate that the spread has increased in the confidence intervals for $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\eta}$ from the confidence intervals for the four-parameter Pareto distribution, which is expected as two more parameters have been fitted. The estimated parameter $\hat{\gamma}$ is not significantly different from 1, and once again $\hat{\eta}$ is not significantly different from 1. When $\gamma = 1$, the parameters β and τ are impossible to identify and equation (58) becomes:

$$Z = \frac{\beta}{\tau} X^{\eta}, \tag{70}$$

where the parameters β and τ are only represented as the ratio β/τ. For this reason the Maximum likelihood estimate $\hat{\beta}/\hat{\tau}$ and a 95% CI for the bootstrap simulated ratio $\hat{\beta}^*/\hat{\tau}^*$ is computed.

Setting $\hat{\gamma}$ and $\hat{\eta}$ equal to one, leads to the Extended Pareto distribution. Hence, the Extended Pareto distribution is again preferred.

Figure 2 presents Q-Q plots for all the nine under-threshold distributions to illustrate how well each distribution fits the data. The Exponential Q-Q plot overestimates small claims, which is no surprise as the Exponential distribution has a very light left tail. The Exponential distribution also fails to capture the extreme right tail of the sample distribution as the plot is located above the 45 degree line to the right. This indicates that the Exponential distribution is not heavy enough in the tails to model the data.

Figure 2 shows that the Weibull distribution underestimates claim sizes approximately above $4{\times}10^5$ NOK and does not have a sufficiently heavy right tail for the data. However, the Weibull Q-Q plot shows a slight improvement in fit from the Exponential Q-Q plot.

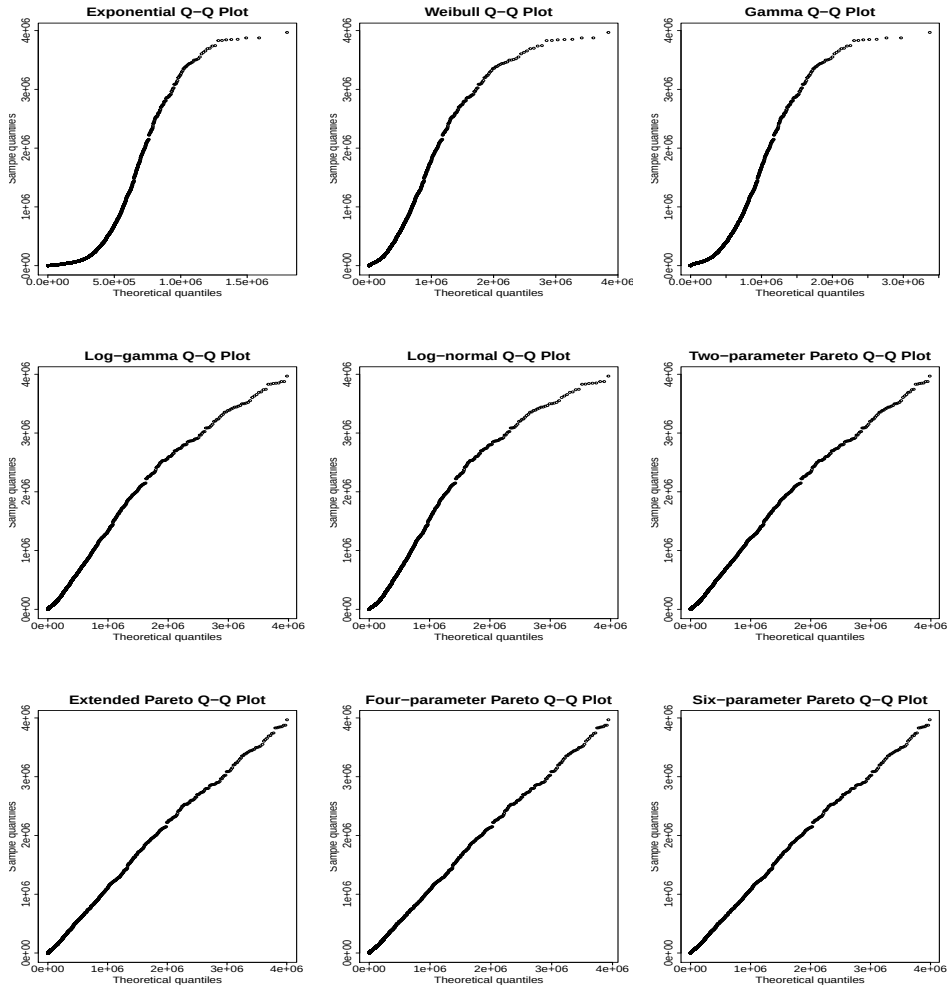


Figure 2: Q-Q plots for all fitted under-threshold distributions.

The right of the Gamma Q-Q plot shows that the Gamma distribution underestimates claims above approximately 7×10^5 NOK and is not as heavy in the right tail as the data. The plot displays a poor fit and suggests that the under-threshold data is not Gamma distributed. The Weibull Q-Q plot indicates the best fit so far.

The Log-gamma Q-Q plot shows a fairly linear relationship between the sample, and the theoretical quantiles. However, except from the tails, the Log-gamma distribution tends to underestimate the claims. The Log-gamma distribution is more heavy-tailed than the previous presented distributions and seems to give the best fit until now.

The Log-normal Q-Q plot indicates that the Log-normal distribution undervalues the claim sizes and fails to measure the extreme right tail of the data. Comparing the Log-gamma to the Log-normal Q-Q plot, the Log-gamma distribution seems to be a better choice of under-threshold claim size distribution. However, the Log-normal Q-Q plot displays a better fit than the Q-Q plots for the Exponential, Gamma and Weibull distributions.

The two-parameter Pareto Q-Q plot displayed in figure 2 succeeds in measure the tails of the distribution. This is expected as the Pareto is an extremely heavy-tailed distribution. Based on the good linear relationship in the Q-Q plot, the two-parameter Pareto distribution is preferred over the Log-gamma-distribution.

The Extended Pareto Q-Q plot indicates a slightly better fit than the two-parameter Pareto distribution as the points are even closer to a 45 degree line. This supports the claim that the Extended Pareto is significantly different from the two-parameter Pareto distribution.

The relationship between the theoretical and the sample quantiles in the four-parameter Pareto Q-Q plot implies a good fit and suggests that the four-parameter Pareto distribution is a good choice of under-threshold distribution. However the plot shows no improvement from the Extended Pareto Q-Q plot.

The six-parameter Pareto Q-Q plot illustrates a fit similar to that of the four-parameter Pareto and Extended Pareto distribution. There is very little difference between the Extended Pareto, the four- and the six-parameter Pareto Q-Q plots.

Statistics

Distribution	$\hat{\mu}$	$\hat{\sigma}_z$	AIC	BIC	P-value
Exponential	1.912×10^5	1.912×10^5	1.63×10^5	1.63×10^5	$< 10^{-4}$
Weibull	1.601×10^5	3.061×10^5	1.57×10^5	1.57×10^5	$< 10^{-4}$
Gamma	1.907×10^5	3.000×10^5	1.58×10^5	1.58×10^5	$< 10^{-4}$
Log-gamma	1.614×10^5	3.995×10^5	1.56×10^5	1.56×10^5	1.80×10^{-4}
Log-normal	1.571×10^5	3.643×10^5	1.55×10^5	1.56×10^5	$< 10^{-4}$
Two-parameter Pareto	1.749×10^5	4.366×10^5	1.55×10^5	1.55×10^5	$< 10^{-4}$
Extended Pareto	1.817×10^5	4.583×10^5	1.55×10^5	1.55×10^5	2.23×10^{-3}
Four-parameter Pareto	1.817×10^5	4.583×10^5	1.55×10^5	1.55×10^5	4.30×10^{-4}
Six-parameter Pareto	1.817×10^5	4.583×10^5	1.55×10^5	1.55×10^5	4.70×10^{-4}

Table 4: Statistics and model evaluation for the under-threshold data.

Table 4 gives AIC and BIC values and compares the model selection criteria between the different distributions. Also the Kolmogorov-Smirnov goodness-of-fit test is applied to each individual under-threshold distribution. The estimated mean and standard deviation for each distribution in table 4 and table 11 have been computed with Monte Carlo based on 1,000,000 samples.

The Exponential and the Gamma under-threshold distributions have estimated means equal to 1.912×10^5 and 1.907×10^5, respectively, which both lie the closest of all the means in table 4 to the under-threshold distributions empirical mean of 1.910×10^5. The Extended, four-parameter and six-parameter Pareto distributions underestimate the average claim size slightly. All the remaining distributions underestimate the mean even more than the Pareto distributions.

The Exponential, Gamma and Weibull under-threshold distribution have the smallest standard deviation. This is expected as they have the lightest tails of the distributions. The four different Pareto under-threshold distributions have the largest standard deviation, as expected since the Pareto distribution has the heaviest tails of the distributions. The Extended Pareto, four-parameter Pareto and six-parameter Pareto under-threshold distributions all have standard deviations of 4.583×10^5 which are closest to the empirical standard deviation. However all the under-threshold distributions underestimate the standard deviation.

It should be noticed that the Extended, four-parameter and six-parameter Pareto under-threshold distributions have the same estimated average claim size and estimated amount of uncertainty in the data. This supports the claim that there is no significant difference between the three models.

The best goodness-of-fit based on AIC are all the Pareto distributions and the Log-normal distribution, with AIC = 1.55×10^5 followed by the Log-gamma distribution with AIC = 1.56×10^5. Based on BIC, that exercises a heavier penalty for number of estimated parameters, all the Pareto distributions have the best goodness-of-fit with BIC = $1.55\ \times10^5$ followed by the Log-gamma and Log-normal distribution, both with BIC = 1.56×10^5. The AIC and BIC are not able to distinguish between the distributions with the lowest information criteria values.

The p-value from the Kolmogorov–Smirnov test is < 0.05 for all the distri-

butions. The null hypothesis is rejected and it can be concluded that the data does not follow any of the specific distributions. The reason for this may be the result of a large data set with little uncertainty in the parameter estimates, leading to the rejection of the null hypothesis, even when the Q-Q plot indicates the opposite. This hypothesis will be tested when performing the model selection methods on a subset of the data.

The Q-Q plots indicate that all the versions of the Pareto distributions fit the data well. Because the under-threshold claim size data have large uncertainty and a heavy right tail, one of the Pareto distributions were expected to be the best choice of model as it has a very heavy right tail. The fitted four-parameter Pareto and six-parameter Pareto distribution are not significantly different from the Extended Pareto distribution. However, the Extended Pareto is significantly different to the two-parameter Pareto as $\hat{\theta}$ differs significantly from 1. This indicates that the Extended Pareto distribution is the best choice of under-threshold distribution for the data. The question is whether the model with the Extended Pareto under-threshold distribution will be selected as the best fitted model for a reduced sample of the data.

3.1.2 The over-threshold claim size distribution

Because the over-threshold data, consisting of only 207 claims, has an extreme right tail, it is reasonable to model the data with the two-parameter Pareto distribution.

Table 5 shows the Maximum likelihood estimates for claims above the threshold b and a 95% confidence interval for the estimates based on 10,000 bootstrap simulations. The Maximum likelihood estimates are very close to the means of the bootstrap estimates and therefore no adjustments are needed. The confidence interval for $\hat{\beta}^{\star}$ in table 5 shows a great amount of uncertainty in $\hat{\beta}$, which is normal in the Pareto distribution and expected when modeling extreme claims.

Table 6 shows that there is insufficient evidence to reject H_0 and conclude that the over-threshold claim size data do not follow the two-parameter Pareto distribution. This is expected due to Pickand's theorem, when the threshold is set appropriate, and indicates that the chosen threshold is reasonable.

Maximum likelihood estimate

$\hat{\alpha}$	$\hat{\beta}$	95 % CI for $\hat{\alpha}^{\star}$	95 % CI for $\hat{\beta}^{\star}$
2.778	1.362×10^{7}	$[1.973, 5.185 \times 10]$	$[8.463 \times 10^{6}, 2.938 \times 10^{7}]$

Table 5: Maximum likelihood estimated parameters with 95% confidence intervals for the over-threshold data.

Goodness-of-fit test

D_n	P-value
3.769×10^{-2}	9.303×10^{-1}

Table 6: Kolmogorov-Smirnov goodness-of-fit test for the over-threshold data.

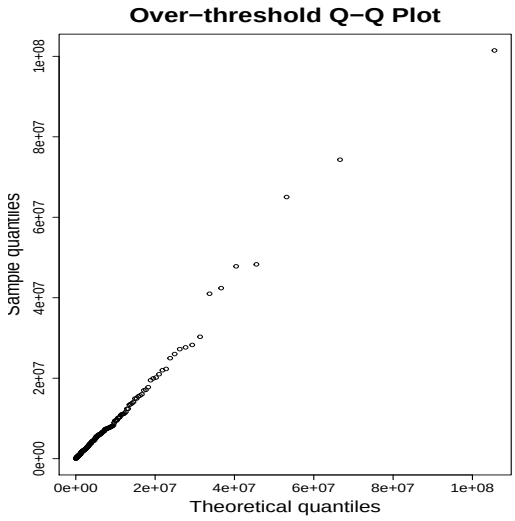


Figure 3: Two-parameter Pareto over-threshold Q-Q plot.

Figure 3 shows an approximate linear relationship for the 207 observations in the over-threshold model. The plot shows a good liner relationship, which indicates that the threshold level b = 4×10^{6} is sufficiently large for Pickand's theorem to be valid. It seems to be a fine balance between skewness and variance as the plot is approximately 45 degrees.

3.1.3 The reserve

The aim now is to estimate the 99% and the 99.5% reserve for next year (i.e. T = 1) for a portfolio of 30,000 policies. Table 7 shows the resulting estimates along with 95% confidence intervals. Density functions for the estimated 99% and the 99.5% reserves are plotted in figure 4 and 5, respectively, based on 1,000 bootstrap simulations and 100,000 Monte Carlo simulations.

The 99% estimated reserve with 95% confidence intervals in table 7 and figure 4 indicate that the width of the confidence intervals does not increase according to the number of parameters fitted to the under-threshold distributions. In other words, the uncertainty in the estimated reserve does not increase with the number of parameters fitted to the model. The estimated reserve with the Gamma under-threshold distribution has the most narrow confidence interval of all the estimates, and therefore the smallest amount of uncertainty in the 99% simulated reserve. The estimated reserve with the two-parameter Pareto under-threshold distribution has the second smallest confidence interval. The estimated reserves with the following under-threshold distributions follows with 95% confidence intervals in increasing order: the six-parameter Pareto, Extended Pareto, Weibull, Log-gamma, Log-normal, four-parameter Pareto and finally the Exponential distribution, with the widest interval, hence the greatest amount of uncertainty in the 99% estimated reserve of all the nine estimates. Having the smallest amount of uncertainty does not necessarily mean having the correct amount of uncertainty in the estimated reserve. As the Extended Pareto distribution was the favored under-threshold distribution, the lower and upper 95% confidence bounds for the 99% reserve estimated with the Extended Pareto under-threshold distribution will be compared with the respective values estimated with the other distributions.

Estimated Reserve

Distribution	99% reserve	99.5% reserve
Exponential	7.920×10^8 $[6.632 \times 10^8, 1.121 \times 10^9]$	8.584×10^8 $[6.913 \times 10^8, 1.348 \times 10^9]$
Weibull	7.712×10^8 $[6.373 \times 10^8, 1.047 \times 10^9]$	8.334×10^8 $[6.644 \times 10^8, 1.254 \times 10^9]$
Gamma	7.957×10^8 $[6.650 \times 10^8, 1.057 \times 10^9]$	8.615×10^8 $[6.922 \times 10^8, 1.249 \times 10^9]$
Log-gamma	6.291×10^8 $[4.898 \times 10^8, 9.047 \times 10^8]$	6.947×10^8 $[5.168 \times 10^8, 1.093 \times 10^9]$
Log-normal	7.676×10^8 $[6.352 \times 10^8, 1.079 \times 10^9]$	8.343×10^8 $[6.660 \times 10^8, 1.286 \times 10^9]$
Two-parameter Pareto	7.737×10^8 $[6.526 \times 10^8, 1.056 \times 10^9]$	8.164×10^8 $[6.819 \times 10^8, 1.264 \times 10^9]$
Extended Pareto	7.847×10^8 $[6.643 \times 10^8, 1.071 \times 10^9]$	8.533×10^8 $[6.957 \times 10^8, 1.275 \times 10^9]$
Four-parameter Pareto	7.924×10^8 $[6.609 \times 10^8, 1.105 \times 10^9]$	8.522×10^8 $[6.887 \times 10^8, 1.331 \times 10^9]$
Six-parameter Pareto	7.918×10^8 $[6.573 \times 10^8, 1.061 \times 10^9]$	8.519×10^8 $[6.833 \times 10^8, 1.261 \times 10^9]$

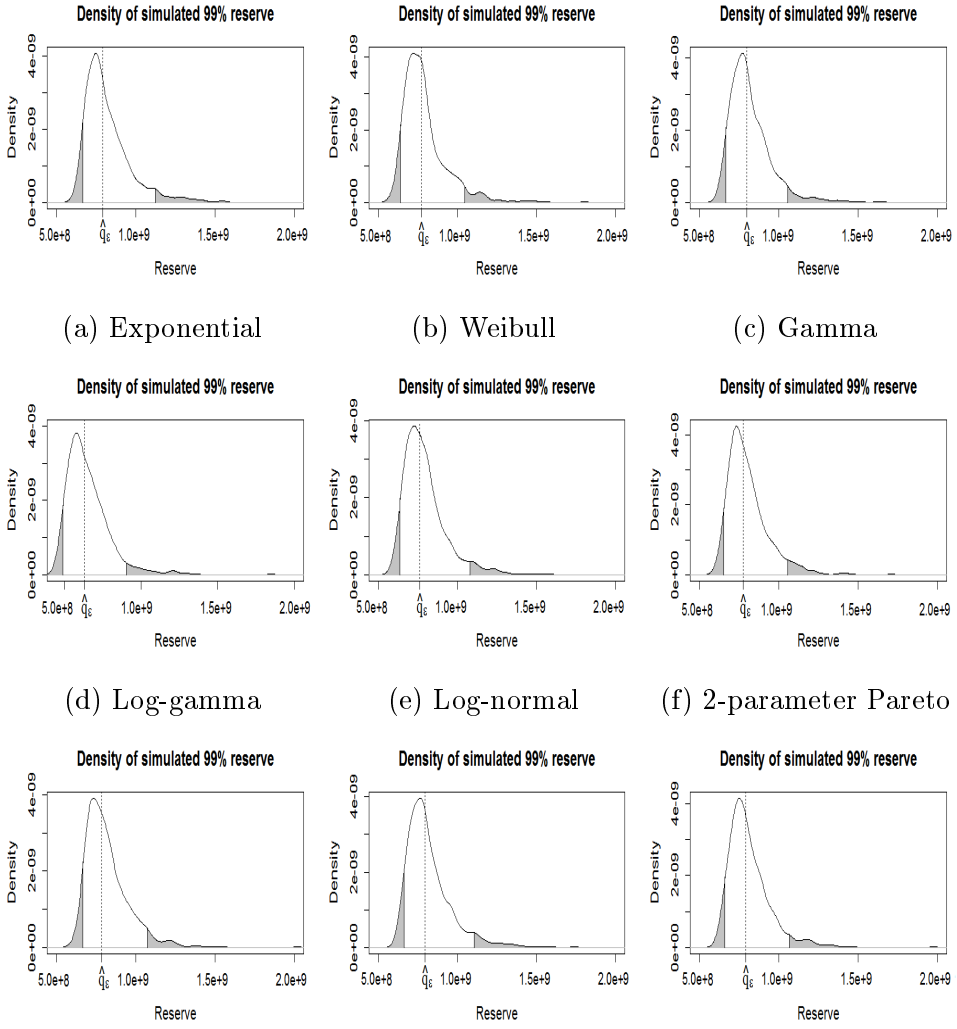
Table 7: Estimated 99% and 99.5% reserves with 95% confidence intervals.

Table 7 and figure 4 display that the estimated reserve with the Exponential under-threshold distribution has a wider confidence interval than the reserve estimated with the Extended Pareto under-threshold distribution. The estimated 99% reserve with the Exponential under-threshold distribution is higher than the corresponding value for the Extended Pareto under-threshold distribution which suggests that the reserve estimated with the Exponential under-threshold distribution is slightly overestimated. However, the 99% reserve estimated with the Exponential under-threshold distribution lies within the confidence interval for the reserve estimated by the Extended Pareto under-threshold distribution which indicates that the two estimated reserves are not significantly different.

Table 7 and figure 4 show that the reserve estimated with the Weibull under-threshold distribution has a smaller lower and upper confidence bound than the reserve estimated with the Extended Pareto under-threshold distribution and the 99% reserve which suggests that the 99% reserve estimated with the Weibull under-threshold distribution is somewhat underestimated. Yet the 99% reserve estimated with the Weibull under-threshold distribution is located within the confidence intervals estimated with the Extended Pareto under-threshold distribution. Hence the two estimates do not differ significantly.

The 99% reserve estimated with the Gamma under-threshold distribution has a lower confidence bound very close to that estimated with the Extended Pareto under-threshold distribution. The upper confidence bound is slightly lower than that estimated with the Extended Pareto under-threshold distribution. The 99% reserve estimated by the Gamma under-threshold distribution is larger than that estimated by the Extended Pareto under-threshold distribution, which indicates that the reserve estimated with the Gamma under-threshold distribution is slightly overestimated. However the 99% reserve estimated with the Gamma under-threshold distribution lies within the confidence interval estimated with the Extended Pareto under-threshold distribution. Hence the reserve estimated with the Gamma under-threshold distribution is not significantly different to the reserve estimated with the favored Extended Pareto under-threshold distribution.

Table 7 shows that the 99% reserve estimated with the Log-gamma under-threshold distribution has the smallest lower and upper confidence bound of all the models. Figure 4 illustrates that the density of the reserve estimated with the Log-gamma under-threshold distribution is shifted to the

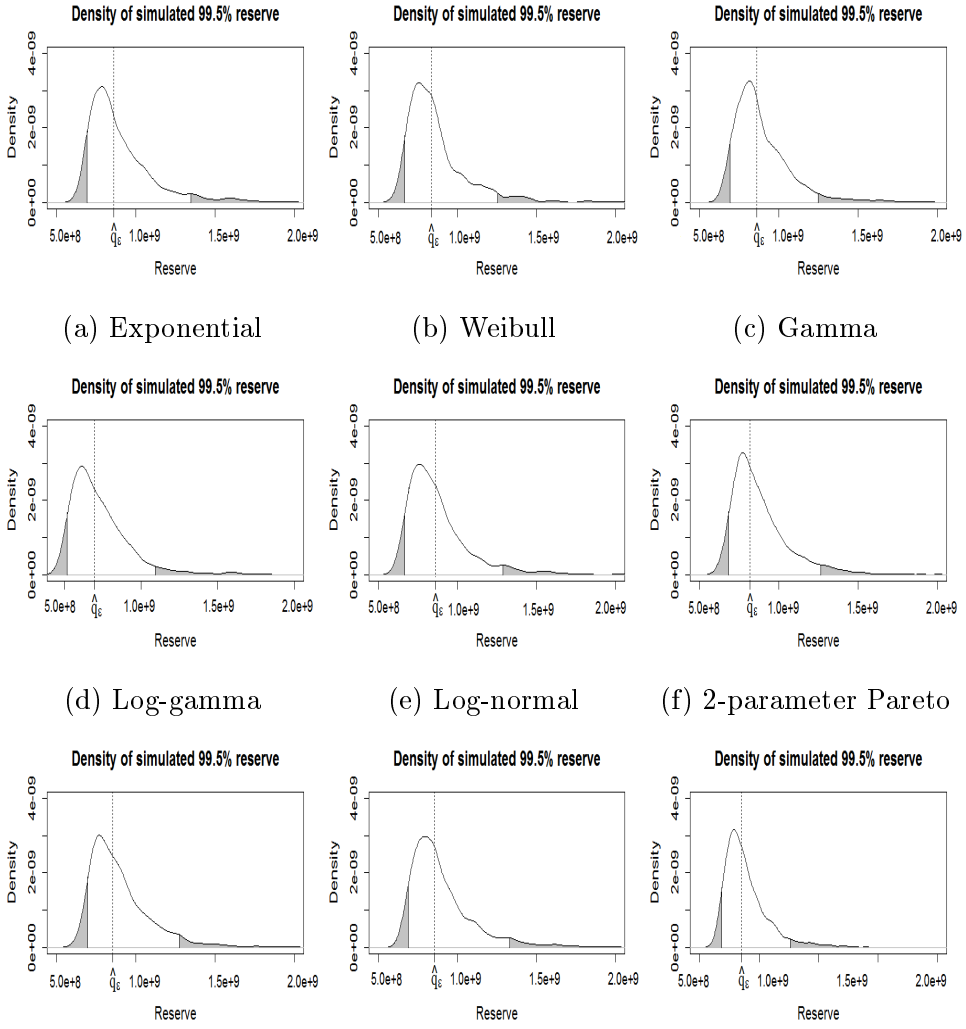


(a) Exponential (b) Weibull (c) Gamma

(d) Log-gamma (e) Log-normal (f) 2-parameter Pareto

(g) Extended Pareto (h) 4-parameter Pareto (i) 6-parameter Pareto

Figure 4: Probability density functions of the 99% Monte Carlo simulated reserves with 95% confidence intervals for the nine models. The 99% Monte Carlo simulated reserves are presented by the dashed line. The 95% confidence intervals are illustrated by the white area.



(a) Exponential (b) Weibull (c) Gamma

(d) Log-gamma (e) Log-normal (f) 2-parameter Pareto

(g) Extended Pareto (h) 4-parameter Pareto (i) 6-parameter Pareto

Figure 5: Probability density functions of the 99.5% Monte Carlo simulated reserves with 95% confidence intervals for the nine models. The 99.5% Monte Carlo simulated reserves are presented by the dashed line. The 95% confidence intervals are illustrated by the white area.

left of the density of the reserve estimated with the Extended Pareto under-threshold distribution. The 99% reserve estimated with the Log-gamma under-threshold distribution is considerably smaller than that of the Extended Pareto distribution, which is illustrated well in figure 4. It seems like the 99% reserve estimated with the Log-gamma under-threshold distribution is underestimated. The 99% reserve estimated with the Log-gamma under-threshold distribution is smaller than the lower 95% confidence bound estimated with the Extended Pareto under-threshold distribution. Hence at a 5% level of significance, the 99% reserve estimated with the Log-gamma under-threshold distribution is significantly different from the reserve estimated with the Extended Pareto under-threshold distribution.

The lower 95% confidence bound for the 99% reserve estimated with the Log-normal under-threshold distribution is lower than that estimated with the Extended Pareto under-threshold distribution. The upper confidence bound however, is almost equal to that estimated with the Extended Pareto under-threshold distribution. The 99% reserve estimated with the Log-normal under-threshold distribution is not significantly different from that estimated with the favored under-threshold distribution.

Table 7 and figure 4 display that both the lower and upper confidence bounds for the 99% reserve estimated with the two-parameter Pareto under-threshold distribution are smaller than those estimated with the Extended Pareto under-threshold distribution, and the 99% reserve is also smaller than that estimated with the preferred Extended Pareto under-threshold distribution. However, there is no significant difference between the two estimated reserves as the 99% reserve estimated with the two-parameter Pareto under-threshold distribution lies within the confidence interval estimated with the Extended Pareto under-threshold distribution.

The 95% lower confidence bound for the reserve estimated with the four-parameter Pareto under-threshold distribution is slightly lower than that estimated with the Extended Pareto under-threshold distribution. The upper bound however, is slightly larger compared to that estimated with the Extended Pareto under-threshold distribution. The 99% reserve estimated with the four-parameter Pareto under-threshold distribution is higher than that estimated with the preferred under-threshold distribution, but lies within the confidence interval estimated with the Extended Pareto under-threshold distribution, and it can be concluded that the two estimated 99% reserves do not differ significantly.

Table 7 and figure 4 show that the lower and upper 95% confidence bounds estimated with the six-parameter Pareto under-threshold distribution are somewhat smaller than those estimated with the Extended Pareto under-threshold distribution, and the 99% reserve is marginally larger. However, the 99% reserve estimated with the six-parameter Pareto under-threshold distribution lies within the confidence interval estimated with the Extended Pareto under-threshold distribution. Hence the two estimated reserves do not differ significantly.

Considering the estimated 99.5% reserves, table 7 and figure 5 again show that the uncertainty in the estimated reserves does not increase with the number of parameters fitted to the under-threshold distribution. Comparing the size of the 95% confidence intervals for the 99.5% reserves estimated with the different under-threshold distributions, the results are similar to those of the 99% estimated reserves. The exceptions are the confidence interval estimated with the Weibull under-threshold distribution now having the fourth largest interval, the confidence interval estimated with the Log-gamma under-threshold distribution now being the second smallest and the confidence interval estimated by the two-parameter Pareto under-threshold distribution now having the fifth smallest confidence interval. The differences are small and seem unimportant as there is still no apparent correlation between the width of the confidence intervals and the number of parameters fitted to the under-threshold distributions.

Comparing the lower and upper confidence bounds for the 99.5% reserve estimated with the favored Extended Pareto under-threshold distributions with the remaining estimates gives results consistent with those for the 99% estimated reserves. The exceptions being the 99.5% reserve estimated with the Gamma under-threshold distribution now having a slightly smaller lower confidence bound than that estimated with the Extended Pareto under-threshold distribution, and the 99.5% reserve estimated with the four-parameter and six-parameter Pareto under-threshold distributions now being slightly smaller than that estimated with the Extended Pareto under-threshold distribution.

The 99.5% reserve estimated with the Log-gamma under-threshold distribution is once again significantly different from the reserve estimated with the Extended Pareto under-threshold distribution. All the remaining estimates are once more not significantly different from that estimated with the Extended Pareto under-threshold distribution at a 5% level of significance.

Figure 4 and figure 5 illustrate a heavy right tail in both the 99% and 99.5% reserves estimated by all models. As a result of the heavy right tail all the estimated reserves represented by $\hat{q}_\epsilon$ are located towards the lower confidence bounds for all models.

Most of the 99% and 99.5% estimated reserves, and their respective 95% confidence bounds lie very close for all the different under-threshold distributions and only the reserves estimated with the Log-gamma under-threshold distribution are significantly different from those of the favored Extended Pareto under-threshold model. This model combination also seems to underestimate the reserve. Choosing this model may lead to Gjensidige forsikring not being able to cover their liabilities with the required 99% or 99.5% probability, and this model should not be selected. However, choosing one of the remaining model combinations to estimate the reserve shows no significant difference from that estimated with the preferred Extended Pareto under-threshold distribution if they are only to be used for estimation of the reserve.

The 99% and 99.5% reserves estimated with the two-parameter Pareto under-threshold distribution is not significantly different from those estimated with the Extended Pareto under-threshold distribution and the estimates are quite similar. Perhaps the two-parameter Pareto distribution is a better choice of under-threshold distribution to model the reserve than the Extended Pareto distribution as it has one less parameter to be fitted to the model and the difference in the reserve estimated with the two under-threshold distributions in combination with the claim frequency and over-threshold distribution is marginal.

From figure 1, the threshold was placed at 4×10^6 NOK. Table 7 shows that the 99% and 99.5% estimated reserves, and their respective 95% confidence intervals all have values above the threshold. This indicates that the over-threshold distribution largely controls the estimated reserves. This may explain why the estimated reserves and their respective 95% confidence intervals are so similar for all the estimates and why the uncertainty in the estimated reserve does not increase with the number of parameters fitted to the under-threshold distribution. Considering figure 2, the difference between the distributions would likely be larger if the focus was the quantiles further left in the distribution such as the 95% or the 90% quantile.

3.2 Reducing the sample size

A smaller sample of 500 claims is selected at random without replacement from the total data.

Claim frequency calculations for the reduced data

The claim frequency for the reduced sample size is calculated with exposure to risk A = 16,667 policy years leading to the same claim intensity $\hat{\mu} = 0.03$ as for the original data set.

3.3 The reduced claim size data

Table 8 shows that the spread has decreased with the reduced data, which is expected as a subset of the total claims have been selected. The median and the mean is larger than that for the total data without threshold. The standard deviation $\sigma = 3.712 \times 10^6$ indicates a large amount of uncertainty in the reduced sample. The skewness = 1.431×10, and the kurtosis = 2.417×10^2 are smaller than those of the total data, but still indicate a heavy tail and extreme values in the distribution.

Summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.980×10^2	1.104×10^4	4.350×10^4	6.700×10^5	1.717×10^5	6.903×10^7

Table 8: Five-number summary of the reduced data without threshold.

The threshold for the reduced data

The mean excess plot for the reduced data is illustrated below. According to the plot in figure 6, a reasonable threshold is $1{,}5 \times 10^6$ NOK as the plot

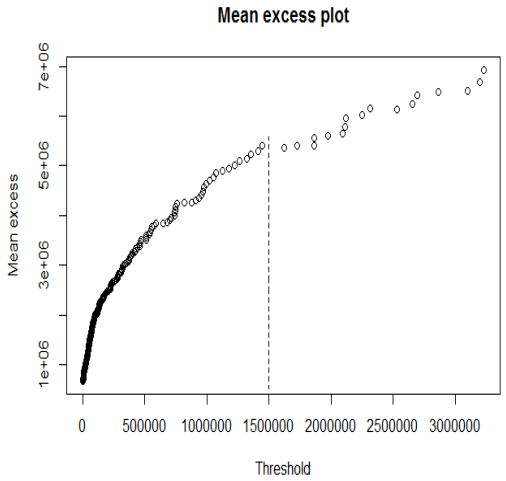


Figure 6: Mean excess plot for the reduced data.

displays an approximately linear relationship from this point. Also the threshold selected for the total data corresponds to the 96.7% quantile. Setting the threshold = $1{,}5\times10^6$ corresponds to the 92.0% quantile which is close to the quantile for the total data. There are now 460 claims under the threshold and 40 claims to be fitted to the over-threshold model.

3.3.1 The reduced under-threshold claim size data

Table 9 gives a summary of the under-threshold data. It shows that the spread has decreased for the reduced under-threshold data, which is expected as the extreme claims have been removed. However, table 9 still displays a great spread in the reduced under-threshold data.

Summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.980×10^2	9.321×10^3	3.768×10^4	1.325×10^5	9.936×10^4	1.448×10^6

Table 9: Five-number summary of the reduced under-threshold data.

The standard deviation $\sigma = 2.506 \times 10^5$ still exposes a large amount of uncertainty. The standard deviation is larger than the mean. This suggests that the data is far from Exponentially distributed. The skewness = 2.994 is smaller than the skewness for the total under-threshold model. The decreasing asymmetry is not extraordinary as the extreme claims have been removed. The kurtosis has also been reduced to 1.412×10. Even though the tail is less heavy, the kurtosis for the reduced sample still signals a heavy tail in the under-threshold claim size distribution. Hence, the reduced data also appears to have extreme values.

Table 10 gives the Maximum likelihood estimates and a 95% confidence interval for the estimated parameters based on 10,000 bootstrap simulations, given the reduced sample. Again the average of the bootstrap estimates lie close to the Maximum likelihood estimates and there is no need for adjustment in the bootstrap estimates.

Table 10 shows that the confidence intervals for all the parameters estimated from the reduced sample, are much wider than those for the total data in table 3. Hence the uncertainty in the Maximum likelihood estimates has increased with the reduced data. This is expected as the sample size now has been reduced to less than 1/10 of the original data, and the uncertainty in the estimates increases as the sample size decreases.

The confidence interval for the Exponential distribution is very narrow and indicates little uncertainty in the Maximum likelihood estimate, just like with the total data. The two-parameter distributions also have narrow confidence intervals, revealing little uncertainty, like with the total data. Again, this is expected as only two parameters are fitted to the data.

The confidence intervals for the Extended Pareto distribution in table 10 shows a great amount of uncertainty in the estimated parameter $\hat{\theta}$. The uncertainty in the respective confidence intervals for the total data in table 3 is significantly smaller, meaning that introducing a third parameter for the reduced sample creates significantly greater uncertainty in the Maximum likelihood estimate of θ compared to the respective uncertainty, given the total data. However, the other two parameters show little uncertainty in the Maximum likelihood estimates, as with the total data. As with the total data, the 95% confidence interval for $\hat{\theta}$ under the Extended Pareto distribution in table 10, $[1.719, 5.480]$, does not include 1, and is therefore significantly different from 1. Hence the Extended Pareto distribution is significantly dif-

ferent from the two-parameter Pareto distribution.

Table 10 shows that the confidence intervals for the four-parameter Pareto distribution reveals greater uncertainty in the estimated parameters than the Extended Pareto distribution, as with the total data, which again is expected as an additional parameter is fitted. Table 3 indicated that $\hat{\eta}$ was not significantly different from 1, and that introducing a fourth variable for the total data is of no use. Table 10 however, illustrates that the 95% confidence intervals for $\hat{\eta}$ under the four-parameter Pareto distribution, $[1.326, 4.125]$, does not include 1, and is consequently significantly different from 1. This means that the four-parameter Pareto distribution is significantly different from the Extended Pareto distribution for the reduced data.

Table 10 shows that the confidence intervals for $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\eta}$ are wider for the six-parameter Pareto distribution compared to the four-parameter Pareto distribution. This was also true for the total data, and is normal as two additional parameters have been fitted to the data. The six-parameter Pareto distribution in table 10 shows that $\hat{\gamma} = 1$ and the parameters τ and β can not be determined individually and will again be presented as the ratio β/τ. As with the four-parameter distribution, $\hat{\eta}$ is significantly different from 1, and as with the total data, the 95% confidence intervals for $\hat{\gamma}$, $[1.000, 2.504]$, includes one and is therefore not significantly different from 1, leading back to the four-parameter distribution. The four-parameter Pareto distribution is now favored.

Figure 7 contains Q-Q plots for all the under-threshold distributions fitted to the reduced sample. The Exponential Q-Q plot in figure 7 indicates that the Exponential distribution exaggerates small claims and underestimates large claims. The same result was present for the total data in figure 2. Again, the Exponential distribution is overly light-tailed and not able to model the data.

Figure 7 shows that the Weibull distribution fits the data better than the Exponential distribution, although it underestimates claims above approximately $2{\times}10^5$ NOK. The Q-Q plot for the Weibull distribution given the reduced data shows a better fit than that of the total data in figure 2. As with the total data, the Weibull distribution is not heavy enough in the right tail to capture the extreme claims but is favored over the Exponential distribution.

Maximum likelihood estimation

Distribution	Parameters with confidence intervals			
Exponential	$\hat{\beta} = 7.544\times10^{-6}$	$[6.998 \times 10^{-6}, 8.165 \times 10^{-6}]$		
Weibull	$\hat{\alpha} = 5.992\times10^{-1}$	$[5.621 \times 10^{-1}, 6.404 \times 10^{-1}]$	$\hat{\beta} = 8.551\times10^{4}$	$[7.439 \times 10^{4}, 9.824 \times 10^{4}]$
Gamma	$\hat{\alpha} = 4.828\times10^{-1}$	$[4.427 \times 10^{-1}, 5.311 \times 10^{-1}]$	$\hat{\beta} = 3.600\times10^{-6}$	$[3.117 \times 10^{-6}, 4.200 \times 10^{-6}]$
Log-gamma	$\hat{\alpha} = 3.788\times10$	$[3.422 \times 10, 4.237 \times 10]$	$\hat{\beta} = 3.615$	$[3.263 \times 10, 4.047 \times 10]$
Log-normal	$\hat{\mu} = 1.057\times10$	$[1.043 \times 10, 1.072 \times 10]$	$\hat{\sigma} = 1.785$	$[1.669, 1.906]$
2-par. Pareto	$\hat{\alpha} = 6.238\times10^{-1}$	$[8.599 \times 10^{-1}, 1.506]$	$\hat{\beta} = 2.016\times10^{4}$	$[2.542 \times 10^{4}, 5.465 \times 10^{4}]$
Ext. Pareto	$\hat{\alpha} = 4.163\times10^{-1}$	$[3.163 \times 10^{-1}, 5.361 \times 10^{-1}]$	$\hat{\theta} = 2.436$	$[1.719, 5.480]$
	$\hat{\beta} = 1.950\times10^{4}$	$[1.540 \times 10^{4}, 2.432 \times 10^{4}]$		
4-par. Pareto	$\hat{\alpha} = 1.721$	$[0.539, 5.279 \times 10^{-1}]$	$\hat{\theta} = 2.536\times10^{3}$	$[1.492 \times 10, 1.687 \times 10^{6},]$
	$\hat{\beta} = 2.394\times10^{4}$	$[1.742 \times 10^{4}, 3.110 \times 10^{4}]$	$\hat{\eta} = 2.440$	$[1.326, 4.125]$
6-par. Pareto	$\hat{\alpha} = 1.726$	$[5.598 \times 10^{-1}, 3.754 \times 10]$	$\hat{\theta} = 2.438\times10^{5}$	$[7.181, 2.415 \times 10^{13}]$
	$\hat{\beta} = 2.777\times10^{4}$		$\hat{\eta} = 2.444$	$[1.335, 9.883]$
	$\hat{\tau} = 1.160$		$\hat{\gamma} = 1.000$	$[1.000, 2.504]$
	$\frac{\hat{\beta}}{\hat{\tau}} = 1.136\times10^{4}$	$[8.462 \times 10^{3}, 3.485 \times 10^{4}]$		

Table 10: Maximum likelihood estimated parameters with 95% confidence intervals for the reduced data under-threshold distributions.

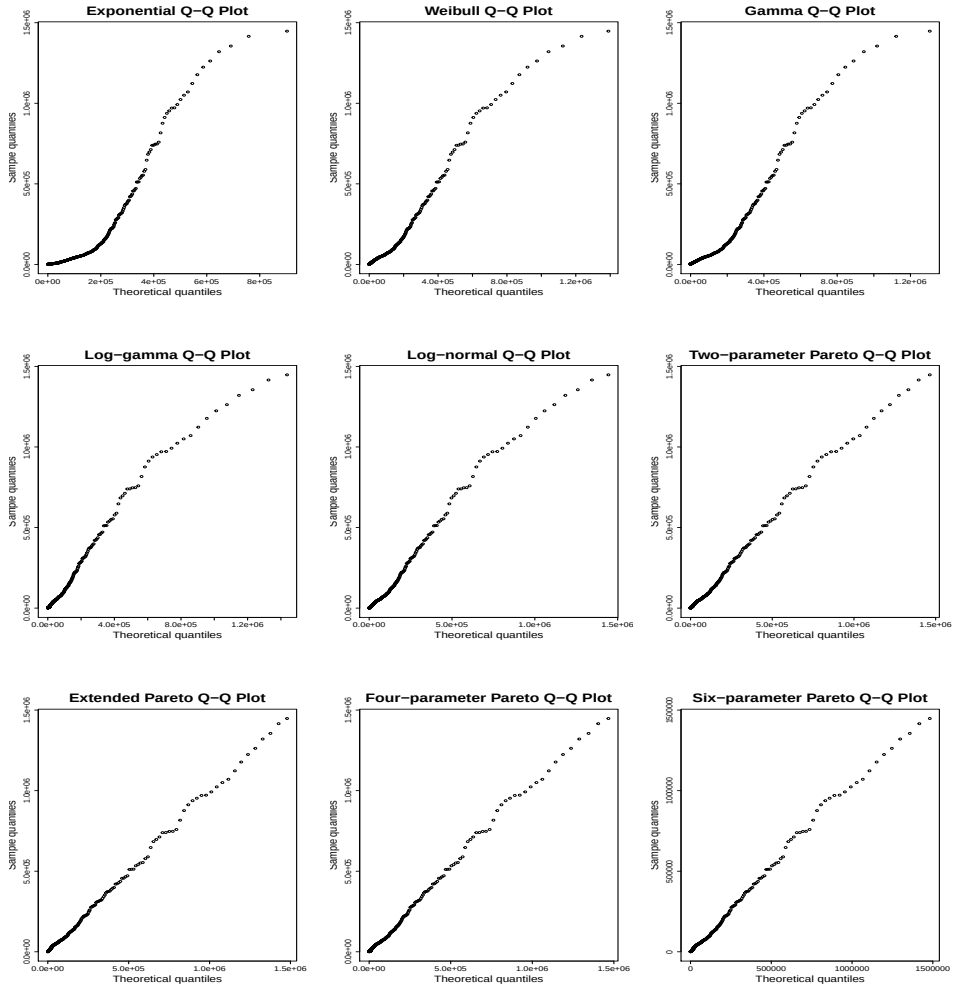


Figure 7: Q-Q plots for all fitted under-threshold distributions given the reduced data.

The Gamma Q-Q plot in figure 7 indicates an improvement in fit compared to the Gamma Q-Q plot in figure 2. However, figure 7 shows that the Gamma distribution underestimates claims even more than the Weibull distribution. Both figure 2 and figure 7 indicate that the Gamma distribution is not sufficiently heavy in the right tail to model the data. As with the total data, based on the Q-Q plot, the Weibull distribution is preferred over the Gamma distribution.

The Log-gamma Q-Q plot for the total data showed a slightly better fit and the plot was smoother than that of the reduced sample in figure 7. The reason for this may be that the total data consists of a significantly greater amount of data and may lead to a smoother curve and perhaps a better fit. The Log-gamma Q-Q plot for the total data displayed a better fit than the Exponential, Weibull and Gamma distributions in figure 2. This is also the case for the reduced data as figure 7 shows that the Log-gamma Q-Q plot has points lying closer to a 45 degree line than the three previously presented distributions. Hence, the Log-gamma distribution is now preferred.

The Log-normal Q-Q plot tends to underestimate claims from approximately 2×10^5 NOK but presents a better fit than the previous plots in figure 7. It is also a slight improvement in fit from that in figure 2. Based on the Q-Q plots the Log-normal distribution is now favored which was not the case for the total data, where the Log-gamma distribution was favored over the Log-normal distribution.

The two-parameter Pareto Q-Q plot in figure 7 shows an even better fit than that of the Log-normal as it captures the tails well. Compared to the two-parameter Pareto Q-Q plot in figure 2, it indicates a somewhat poorer fit. However, this may be due to the reduction in the sample size, as there are fewer data points to fit than before. Looking at figure 7, the two-parameter Pareto distribution is clearly preferred over the previous distributions, just as with the total data.

Figure 2 indicated a better fit for all the Pareto distributions compared to figure 7, which is expected as the reduced sample has only 500 observations to be fitted. All the Pareto Q-Q plots in figure 7 look very similar and do not reveal a significant difference between them. This was also the case with the Pareto Q-Q plots for the total data in figure 2, except from the Extended Pareto distribution that indicated a slightly better fit than the two-parameter Pareto distribution.

As with the total data, according to the Q-Q plots in figure 7, the two-parameter, Extended, four-parameter and six-parameter Pareto distributions all seem to be the best choice of under-threshold distribution to model the reduced data as it is not possible to distinguish between them.

Table 11 presents the simulated mean and standard deviation, the AIC and BIC values and a goodness-of-fit test for all the under-threshold distributions. The Exponential mean in table 11 is slightly larger than the Exponential standard deviation. This may be due to a tiny error in the Monte Carlo simulation despite the fact that 1,000,000 simulations have been computed. The Gamma distribution has like with the total data the closest simulated mean to that of the empirical under-threshold distribution, closely followed by the Extended Pareto, four-parameter and six-parameter Pareto distributions, who now unlike with the total claim size, slightly overestimate the average. In contrast to the total claim size comparison in table 4, the Exponential mean now deviates the most from the under-threshold empirical mean of all distributions in table 11. As in table 4, all the remaining distributions underestimate the mean. This was not the case for the total data where the Exponential distribution was the only distribution not to underestimate the mean.

As in table 4, the Exponential, Weibull and Gamma distribution have the smallest standard deviations, being the lightest tailed distributions. Equally to table 4, all the Pareto distributions in table 11 have the greatest standard deviations, being the most heavy-tailed distributions.

As in table 4, the Extended, four-parameter and six-parameter Pareto distributions all have standard deviations closest to the empirical standard deviation. Their standard deviations are now almost identical to the empirical standard deviation. The two-parameter Pareto distribution follows with a tiny underestimation in the standard deviation. The remaining distributions underestimate the standard deviation.

Equivalent to the statistics listed in table 4, the Extended, four-parameter and six-parameter Pareto distributions in table 11 have the same mean and standard deviation. Again, this suggests that there is not much of a difference between the three models.

Statistics

Distribution	$\hat{\mu}$	$\hat{\sigma}_z$	AIC	BIC	P-value
Exponential	1.325×10^4	1.323×10^4	1.18×10^4	1.18×10^4	$< 10^{-4}$
Weibull	1.234×10^5	1.930×10^5	1.16×10^4	1.16×10^4	3.700×10^{-4}
Gamma	1.328×10^5	1.868×10^5	1.14×10^4	1.14×10^4	$< 10^{-4}$
Log-gamma	1.035×10^5	1.958×10^5	1.14×10^4	1.14×10^4	1.765×10^{-1}
Log-normal	1.179×10^5	2.094×10^5	1.14×10^4	1.14×10^4	2.801×10^{-1}
Two-parameter Pareto	1.266×10^5	2.331×10^5	1.14×10^5	1.14×10^5	4.957×10^{-1}
Extended Pareto	1.344×10^5	2.508×10^5	1.14×10^4	1.14×10^4	2.584×10^{-1}
Four-parameter Pareto	1.344×10^5	2.508×10^5	1.14×10^4	1.14×10^4	5.209×10^{-1}
Six-parameter Pareto	1.344×10^5	2.508×10^5	1.14×10^4	1.14×10^4	5.203×10^{-1}

Table 11: Statistics and model evaluation for the reduced under-threshold data.

Based on both the AIC and BIC values in table 11, the Exponential distribution has the highest AIC and BIC values both equal to 1.18×10^4 and therefore the worst fit of all the distributions, which was also the case for the total data in table 4. The Weibull distribution has the second worst fit now with AIC and BIC both equal to 1.16×10^4. Table 4 also ranked the Weibull distribution as the second worst fit based on the AIC and BIC values. The rest of the distributions have equal AIC and BIC values and it is impossible to distinguish between them. However, this was not the case with the total data, where all the Pareto distributions stood out with lowest AIC and BIC values (except from the AIC value for the Log-normal distribution that was equal to the AIC values for the Pareto distributions).

The hypothesis presented for the total data set claiming that the large amount of data resulted in rejection of the null hypothesis is supported in table 11. The hypothesis for the Exponential, Weibull and Gamma distributions are rejected and it may be concluded that the data do not follow any of the specific distributions. However, the null hypothesis is not rejected for the other distributions and it cannot be concluded that the data are not significantly different from the distributions. This is consistent with the Q-Q plots for the reduced sample in figure 7 which display a poor fit for the Exponential, Weibull and Gamma distributions.

The Q-Q plots however indicate that one of the Pareto distributions should be selected. The 95% confidence intervals in table 10 shows that the four-parameter Pareto distribution is significantly different to the Extended Pareto distribution, and that the six-parameter Pareto distribution is not significantly different to the four-parameter distribution. Taking all the model selection tools into account, the result is not the same as with the total data. The four-parameter Pareto is the preferred under-threshold distribution for the reduced data. However, it is important to notice that the differences between the fitted models have been reduced with the sample size, and the choice is not straightforward. The information criteria values are more similar and the Q-Q plots for all the distributions do not differ as much between the distributions as they did with the total data, perhaps because the data size has been reduced. The next question will be how the number of parameters in the claim size distribution influence the uncertainty in the reserve for a smaller data set.

3.3.2 The reduced over-threshold claim size distribution

As with the total over-threshold data, the reduced over-threshold data has an extreme right tail, now with 460 claims to be fitted to the over-threshold distribution. For this reason the two-parameter Pareto distribution will be used to model the over-threshold claim size data.

Table 12 presents the Maximum likelihood estimates for claims above the threshold b and a 95% confidence interval for the estimates based on 10,000 bootstrap simulations, given the reduced sample. As with the total data, the mean of the bootstrap estimates is very close to the Maximum likelihood estimate for both estimated parameters, and there is no need for adjustment in the bootstrap simulations.

Maximum likelihood estimate

$\hat{\alpha}$	$\hat{\beta}$	95% CI for $\hat{\alpha}^{\star}$	95% CI for $\hat{\beta}^{\star}$
2.017	5.229×10^6	$[1.175, 3.868 \times 10]$	$[2.187 \times 10^6, 1.138 \times 10^8]$

Table 12: Maximum likelihood estimated parameters with 95% confidence intervals for the over-threshold reduced data.

As in table 5, the confidence interval for $\hat{\beta}^{\star}$ in table 12 displays large uncertainty. Both $\hat{\alpha}$ and $\hat{\beta}$ are located to the left in their respective confidence intervals in table 12, while they are both more centered in table 5. This may be a result of the reduced sample size, as only 40 observations are used to estimate the parameters, leading to large estimated values influencing the upper confidence bounds, as the parameters are strictly positive.

Goodness-of-fit test

D_n	P-value
7.596×10^{-2}	9.751×10^{-1}

Table 13: Kolmogorov-Smirnov goodness-of-fit test for the reduced over-threshold data.

Table 13 fails to reject the null hypothesis, using the Kolmogorov-Smirnov

goodness-of-fit test, that the reduced sample is Pareto distributed, as with the total data. This means that a reasonable threshold b has been selected. Figure 8 shows three extremely large values to the right having a great in-

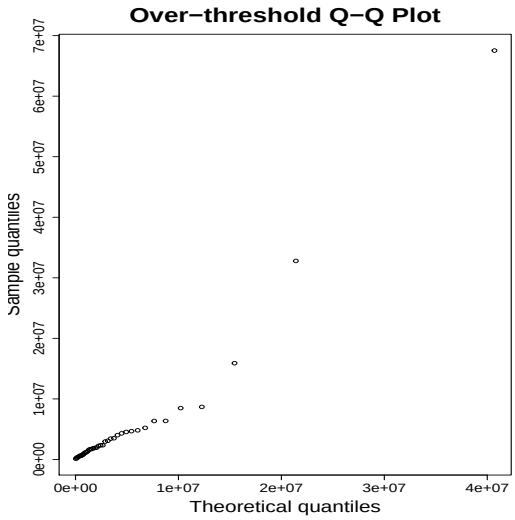


Figure 8: Two-parameter Pareto over-threshold Q-Q plot for the reduced data.

fluence on the Q-Q plot. This may be the result of the reduced sample, where only 40 observations have been fitted and a few extreme observations are expected. Disregarding the three extreme values, the Q-Q plot shows an approximately linear relationship between the theoretical quantiles and the sample quantiles. Figure 8 indicates that the over-threshold model is Pareto distributed, which indicates that the threshold b = $1{,}5\times10^6$ is determined correctly.

3.3.3 The reduced data reserve

The objective is to estimate the 99% and the 99.5% reserve for next year (i.e. T = 1) for a portfolio of 2,340 policies, which is comparable to the portfolio of 30,000 for the full data set. Table 14 shows the resulting estimates along with a respective 95% confidence interval. The density functions for the 99% and 99.5% estimated reserves are plotted in figure 9 and 10, respectively. The simulations are based on 1,000 bootstrap simulations and 100,000

Monte Carlo simulations. The right tails of the density functions have been cut off to obtain the same scale for all the plots in figure 9 and 10. The 99% estimated reserves with confidence intervals in table 14 and figure 9 show that the width of the confidence intervals does not increase with the number of parameters fitted to the under-threshold distributions for the reduced sample. This was also the case for the total data.

The amount of uncertainty for the different estimates shows no pattern consistent with the amount of uncertainty from the total data. The reserve estimated with the Extended Pareto under-threshold distribution has the smallest confidence interval of all the estimates. The 99% reserve estimated with the following under-threshold distributions follows with 95% confidence intervals in increasing order: the Gamma, two-parameter Pareto, Exponential, four-parameter Pareto, six-parameter Pareto, Weibull, Log-normal and at last the Log-gamma distribution.

Now, the reserve estimates and confidence bounds resulting from different under-threshold distributions will be compared to the ones for the four-parameter Pareto, which is the favored distribution for the reduced sample. Table 14 and figure 9 show that the reserve estimated with the Exponential under-threshold distribution has a lower confidence bound almost identical to that estimated with the four-parameter Pareto under-threshold distribution, but the upper confidence bound is smaller. The 99% reserve estimated with the Exponential under-threshold distribution is slightly larger than that estimated with the favored under-threshold distribution, but lies within its 95% confidence interval, and is therefore significant at a 5% level of significance.

The lower confidence bound for the 99% reserve estimated with the Weibull under-threshold distribution is smaller than that estimated with the preferred under-threshold distribution, and the upper confidence bound is larger. The 99% reserve estimated is marginally smaller, and lies within the confidence interval of the reserve estimated with the favored under-threshold distribution. Hence, there is no significant difference between the two 99% estimated reserves.

The 99% reserve estimated with the Gamma under-threshold distribution has a slightly larger lower confidence bound and a smaller upper confidence bound compared to that estimated with the four-parameter Pareto under-threshold distribution, and the 99% reserve is somewhat larger. Again, there is no significant difference between the two 99% estimated reserves.

Estimated Reserve

Distribution	99% reserve	99.5 % reserve
Exponential	1.697×10^8 $[7.492\times10^7, 5.809\times10^8]$	2.212×10^8 $[8.128\times10^7, 9.890\times10^8]$
Weibull	1.641×10^8 $[7.342\times10^7, 6.298\times10^8]$	2.092×10^8 $[7.941\times10^7, 1.116\times10^9]$
Gamma	1.690×10^8 $[7.512\times10^7, 5.526\times10^8]$	2.189×10^8 $[8.139\times10^7, 9.368\times10^8]$
Log-gamma	1.571×10^8 $[6.681\times10^7, 6.505\times10^8]$	2.036×10^8 $[7.291\times10^7, 1.129\times10^9]$
Log-normal	1.573×10^8 $[7.375\times10^7, 6.316\times10^8]$	2.051×10^8 $[8.107\times10^7, 1.082\times10^9]$
Two-parameter Pareto	1.649×10^8 $[7.564\times10^7, 5.663\times10^8]$	2.128×10^8 $[8.192\times10^7, 9.688\times10^8]$
Extended Pareto	1.671×10^8 $[7.755\times10^7, 5.168\times10^8]$	2.179×10^8 $[8.435\times10^7, 8.951\times10^8]$
Four-parameter Pareto	1.652×10^8 $[7.493\times10^7, 5.954\times10^8]$	2.160×10^8 $[8.167\times10^7, 1.031\times10^9]$
Six-parameter Pareto	1.653×10^8 $[7.717\times10^7, 6.287\times10^8]$	2.146×10^8 $[8.365\times10^7, 1.092\times10^9]$

Table 14: Estimated 99% and 99.5% reserves with 95% confidence intervals for the reduced data.

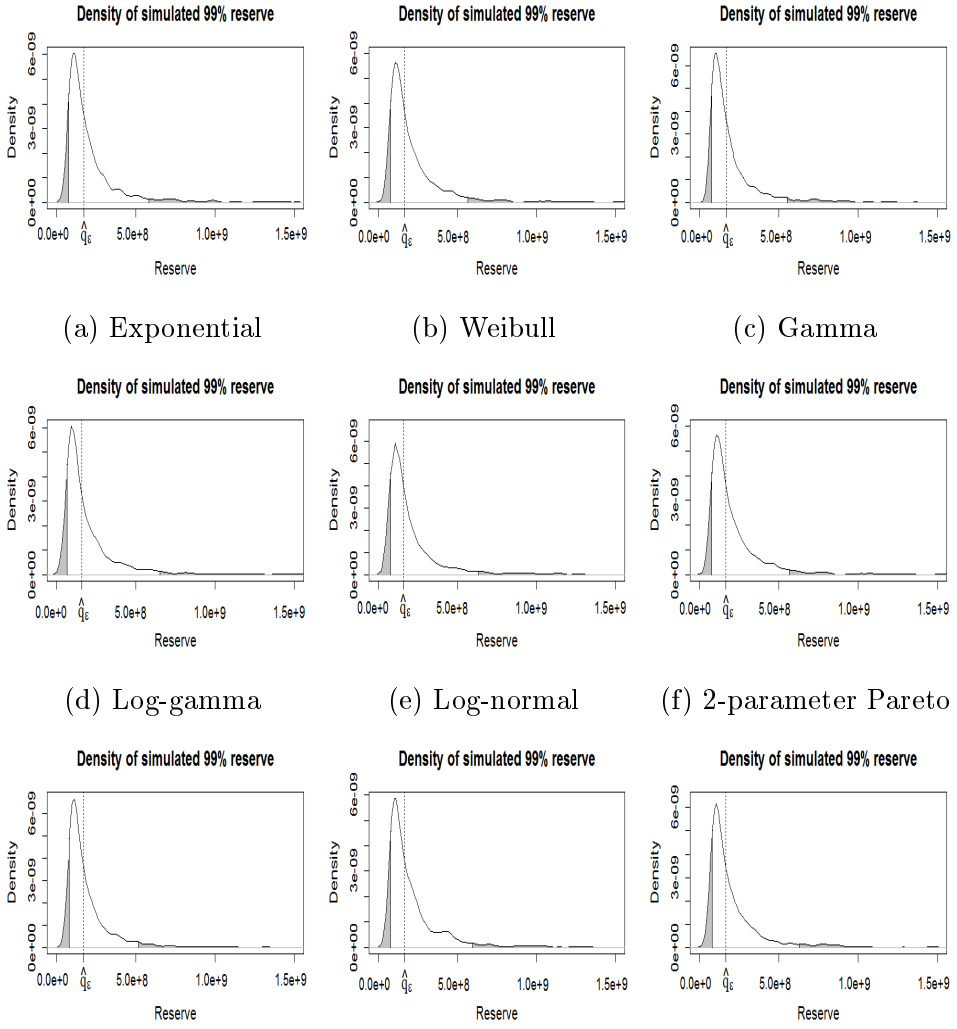
Table 14 and figure 9 display that the 99% reserve estimated with the Log-gamma and the Log-normal under-threshold distributions have a wider confidence interval than for the four-parameter Pareto under-threshold distribution and the 99% reserves are smaller. However the 99% reserves estimated with the Log-gamma and the Log-normal under-threshold distributions for the reduced sample are not significantly different from the 99% reserve estimated with the four-parameter Pareto under-threshold distribution, as both estimates lie within the 95% confidence interval estimated by the favored under-threshold distribution.

Table 14 and figure 9 illustrate that the lower confidence bound estimated with the two-parameter Pareto under-threshold distribution is larger, and the upper confidence bound smaller than that estimated with the favored under-threshold distribution, whereas the 99% reserves are almost identical. Once again the 99% reserve estimated by the two-parameter Pareto under-threshold distribution is not significantly different from the reserve estimated with the four-parameter under-threshold distribution.

The confidence interval estimated with the Extended Pareto under-threshold distribution is narrower than that of the favored under-threshold distribution, and the 99% reserve is marginally larger. Again it can be concluded that the 99% reserve estimated with the Extended Pareto under-threshold distribution does not differ significantly from that estimated with the four-parameter Pareto under-threshold distribution.

The upper and lower confidence bounds estimated with the six-parameter Pareto under-threshold distribution are both larger compared to those estimated with the four-parameter Pareto under-threshold distribution, whereas the 99% reserve is almost identical. Once more the 99% reserve estimated by the six-parameter Pareto under-threshold distribution is not significantly different from the reserve estimated by the four-parameter Pareto under-threshold distribution.

Table 14 and figure 10 reveal that the width of the 95% confidence intervals for the estimated 99.5% reserves does not increase with the number of parameters fitted to the under-threshold distributions. The ranking of the size of the 95% confidence intervals for the 99.5% estimated reserves are similar to the ranking of the 99% estimated reserves. The exceptions are the Weibull under-threshold now giving the second largest amount of uncertainty, the Log-normal giving the fourth largest amount of uncertainty and the six-

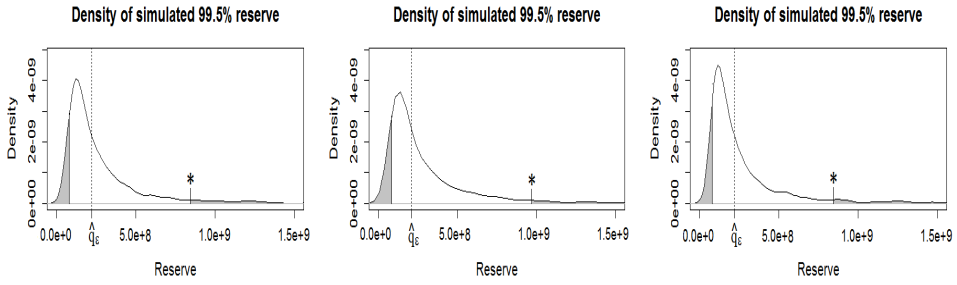


(a) Exponential (b) Weibull (c) Gamma

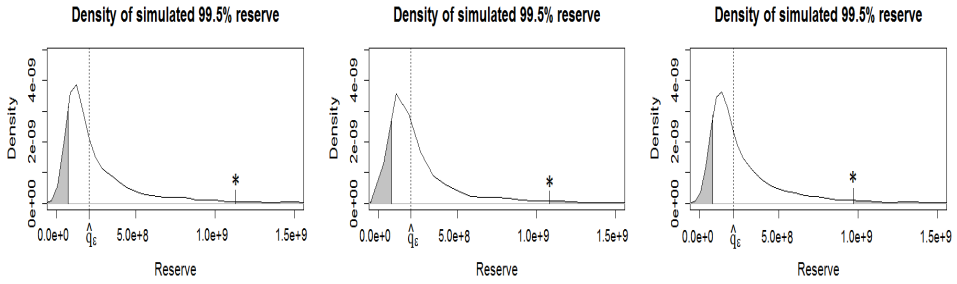
(d) Log-gamma (e) Log-normal (f) 2-parameter Pareto

(g) Extended Pareto (h) 4-parameter Pareto (i) 6-parameter Pareto

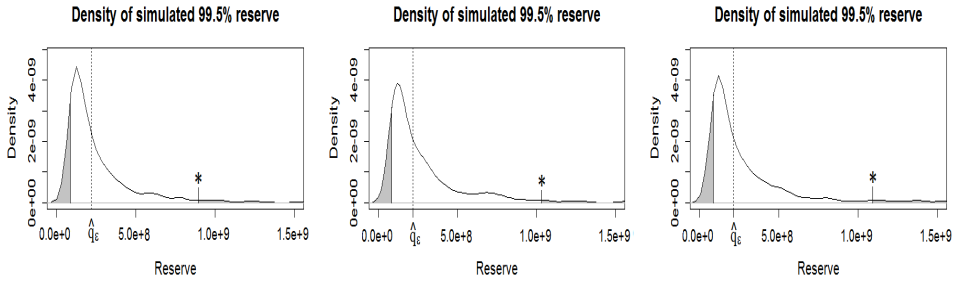
Figure 9: Probability density functions of the 99% Monte Carlo simulated reserves with 95% confidence intervals for the nine models, given the reduced data. The 99% Monte Carlo simulated reserves are presented by the dashed line. The 95% confidence intervals are illustrated by the white area.



(a) Exponential (b) Weibull (c) Gamma



(d) Log-gamma (e) Log-normal (f) 2-parameter Pareto



(g) Extended Pareto (h) 4-parameter Pareto (i) 6-parameter Pareto

Figure 10: Probability density functions of the 99.5% Monte Carlo simulated reserves with 95% confidence intervals for the nine models, given the reduced data. The 99.5% Monte Carlo simulated reserves are presented by the dashed line. The 95% confidence intervals are illustrated by the white area. The upper confidence intervals are marked by the symbol $*$.

parameter Pareto giving the third largest amount of uncertainty. As with the total data, the differences are marginal. For this reason it is reasonable to believe these differences to be irrelevant.

Comparing the lower and upper 95% confidence bounds for the 99.5% reserve estimated with the four-parameter Pareto under-threshold distribution to the confidence bounds for the reserves estimated with the remaining under-threshold distributions gives similar results to those for the 99% estimated reserves for the reduced sample. The exceptions are the lower confidence bounds for the reserve estimated with the Exponential and the Gamma under-threshold distributions now being slightly smaller than that of the favored under-threshold distribution. Comparing the 99.5% reserve estimated with the four-parameter Pareto under-threshold distribution with the remaining estimates gives similar results to those with the 99% estimated reserves. The only exceptions being the reserves estimated with the two-parameter and six-parameter Pareto under-threshold distributions now being slightly smaller. Once again all the 99.5% reserves estimated with the different under-threshold distributions are not significant different from the reserve estimated with the favored under-threshold distribution, and it can not be concluded that any of them differ significantly from the 99.5% reserve estimated with the four-parameter Pareto under-threshold distribution.

As with the total data, figure 9 and figure 10 display a heavy right tail for all the density plots given the reduced sample. Once again all the estimated reserves represented by $\hat{q}_\epsilon$ are located towards the lower confidence bounds for all the estimates. Once more all the 99% and 99.5% estimated reserves, and their respective 95% confidence bounds, are very similar for all the different under-threshold distributions and none of the estimates are significantly different from the reserve estimated with the four-parameter Pareto under-threshold distribution at a 5% level of significance.

As the 99% and 99.5% reserves estimated with the two-parameter Pareto and the Extended Pareto under-threshold distributions are not significantly different from those estimated with the favored model and the confidence intervals are very similar, it may be preferable to chose a model with an under-threshold distribution containing fewer parameters to model the reserve. Selecting the model with the two-parameter Pareto under-threshold distribution gives two less parameters to be estimated to the model, and the difference in the estimated reserve is not significant compared to the reserve estimated with the four-parameter Pareto under-threshold distribution. For this reason selecting the two-parameter Pareto under-threshold distribution

seems to be a better choice than the four-parameter or the Extended Pareto under-threshold distribution if the models are only to be used for the estimation of the reserve.

The uncertainty in the estimated reserve does not increase with the number of parameters fitted to the under-threshold distribution for the total and the reduced data. However, note that the width of the confidence intervals have increased significantly from the total data to the reduced sample, which is very clear in figure 9 and figure 10. The increased uncertainty in the reduced sample is likely a result from estimation error, which naturally increases as the sample size decreases. Also, table 3, 5, 10 and 12 shows that the uncertainty increases in the Maximum likelihood estimate with the number of parameters fitted to the distributions.

The threshold was placed at 1.5×10^6 NOK from figure 6 for the reduced data. As with the estimated reserves for total data in table 7, table 14 shows that the 99% and 99.5% estimated reserves, and their respective 95% confidence intervals all have values above the threshold, which means that the over-threshold distribution (largely) models the reserve for all the models. This may again explain why the results for the estimated reserves are very similar and why the uncertainty does not increase according to the number of parameters fitted to the under-threshold distribution.

The threshold was set at the 96.7% quantile for the total data, giving the under-threshold distribution a slightly greater influence in estimating the reserve compared to the models for the reduced data, and may explain why the 99% and 99.5% reserves estimated with the Log-gamma under-threshold distribution were significantly different from the model with the Extended Pareto under-threshold distribution.

The threshold was set at the 92.0% quantile for the reduced sample giving the over-threshold distribution more influence over the reserve and may explain why there was no significant difference between the reserve estimated with the four-parameter Pareto under-threshold distribution and the eight models for the reduced sample. The differences would perhaps have been greater if the focus had been on quantiles further left in the distribution, such as the 95% or the 90% quantile.

Given the total data, all the 99% and 99.5% reserves and their respective 95% confidence bounds were quite similar for all the estimates. However the 99% and 99.5% reserve estimated with the Log-gamma under-threshold

distribution were significantly different from those estimated with the Extended Pareto under-threshold distribution. The model combination with this under-threshold distribution seemed to underestimate the threshold and Gjensidige forsikring will not be able to cover their liabilities with a 99% and 99.5% probability, and takes on a greater risk if choosing this model. The other reserves estimated with the remaining under-threshold distributions were not significantly different to the 99% and 99.5% reserve estimated with the Extended Pareto under-threshold distribution.

4 Conclusion

This thesis has assessed whether a six-parameter extension of the Pareto distribution may be used in an automatic procedure for selecting the claim size distribution, compared to its one, two, three and four-parameter special case distributions on data consisting of property insurance claims on office and industrial buildings. The procedure was repeated on a subset of the data to study how the sample size influence the uncertainty in the estimated reserves.

The thresholds were set at 4×10^6 NOK and 1.5×10^6 NOK for the total and reduced data, respectively. The best fitting models were selected based on Q-Q plots, AIC and BIC values and the Kolmogorov-Smirnov hypothesis tests. A 95% confidence interval for the estimated parameters were simulated to test the significance of the parameters. The Extended Pareto and the four-parameter Pareto distribution were selected as the best fitting models for the under-threshold data, given the total and reduced data, respectively.

The 99% and 99.5% reserves with 95% confidence intervals were simulated by nested bootstrap simulations for the different combinations of the Poisson distributed claim frequency, over-threshold distribution and the nine different under-threshold distributions. The results were compared with that simulated with the favored under-threshold distribution. The 99% and 99.5% reserves were estimated by Monte Carlo simulation. The estimated reserves were compared with that estimated with the preferred under-threshold distribution. The estimates were also compared with the 95% confidence interval simulated with the favored under-threshold distribution.

Insurance companies do not want to bind an unnecessary high amount of money and miss the opportunity of investment return. However, not meeting the capital requirements to cover the claims may lead to economical difficulties and perhaps even bankruptcy. The 99% and 99.5% estimated reserves for the total data based on the model combination with the Log-gamma under-threshold distribution were significantly underestimated compared to the ones estimated with the Extended Pareto under-threshold distribution. There were however no significant differences between the remaining models in estimating the 99% and 99.5% reserve for the total data. For the reduced data set there was no significant difference between any of the reserve estimates. All the 99% and 99.5% reserves and their confidence intervals were

of much higher value than the thresholds established for both the total and reduced data. This means that the over-threshold distribution largely modeled the estimated reserves for both data sets. This can likely explain why the reserves and their confidence intervals were so similar for all the models. However, the threshold was set at the 96.7% quantile for the total data, as opposed to the 92% quantile for the reduced data set. This might have given the under-threshold distribution a slightly greater influence in estimating the reserve, for the full data set, which may explain why the Log-gamma under-threshold distribution was significantly different for this data set. It may also explain why the uncertainty in the reserve did not increase with the number of parameters even though the uncertainty in the Maximum likelihood estimations did.

The total data with 6,411 claims > 0 for property insurance claims on office and industrial buildings from Gjensidige forsikring is a great amount of data. Being the largest insurance company, Gjensidige forsikring had 25,3 % of the market share for non-life premiums in Norway in 2015 [Finans Norge, 2016]. Estimating the reserve with heavy-tailed data from a different Norwegian insurance company is likely to give a considerably smaller set of claims, which, based on the result from the reduced data, perhaps may lead to all the reserves estimated with different under-threshold distributions not being significantly different.

Note that the data given by Gjensidige forsikring has a very heavy right tail and the results may differ for a set of data with a lighter right tail, where the threshold is set to the right of the estimated reserve, or there are no extreme claims, making an over-threshold distribution unnecessary. Perhaps this would lead to greater differences between the estimated reserves and it may be that one of the Pareto distributions combined with the claim frequency (and the over-threshold distribution) no longer would be the best choice of model.

This thesis has focused on the estimation of the reserve, which was shown to be modeled (largely) by the over-threshold distribution as described above. However, if the focus was the center of the distribution, for instance the pure premium, the mean, or quantiles further left in the distribution, the results would likely differ. The means computed with the different under-threshold distributions had substantial variations for both the total and the reduced data. This indicated that focusing below the threshold, would lead to greater differences in results between the different combinations of claim frequency and under-threshold distributions. Also the standard deviation computed

in table 4 and 11 increased with number of parameters fitted to the model. Hence, the uncertainty in the reserve would likely increase with the number of parameters fitted to the model, if the focus had been further left in the distribution.

The Extended Pareto and the four-parameter Pareto were selected as under-threshold distributions for the total and reduced data, respectively. When estimating the reserves, the two-parameter Pareto was chosen as the favored under-threshold distribution to estimate the 99% and 99.5% reserves for both the total and reduced data. The six-parameter Pareto distribution failed to optimize the parameter γ and seemed to be too complex to model the data. However, the four-parameter Pareto distribution is a more flexible distribution than the Extended and two-parameter Pareto distributions and seemed to be a better option than the six-parameter Pareto distribution in estimating the 99% and 99.5% reserves. Hence the four-parameter Pareto distribution may be automatically selected to estimate the 99% and 99.5% reserves.

Introducing reinsurance as a threshold to cover the event of extreme claims is a common practice in the insurance industry. It could also allow for the data to be modeled with only one claim severity distribution. This could possibly lead to greater differences between the fitted models, and perhaps larger variation between the estimated reserves.

It could be interesting to estimate the pure premium, or other quantiles further left in the distribution than the reserve to see how it would influence the choice of under-threshold distribution

A final possible extension would be to perform the same methods on a different data set, perhaps one with a lighter right tail to assess how the uncertainty in the estimated reserves would differ between the nine model combinations.

References

[Abramowitz and Stegun, 1965] Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Dover Publications, New York.

[Beirlant et al., 2009] Beirlant, J., Joossens, E., and Segers, J. (2009). Second-order refined peaks-over-threshold modelling for heavy-tailed distributions. *Journal of Statistical Planning and Inference*, 139(8):2800–2815.

[Bølviken, 2014] Bølviken, E. (2014). *Computation and modelling in insurance and finance*. Cambridge University Press.

[Box and Cox, 1964] Box, G. and Cox, D. (1964). An analysis of transformations.

[Claeskens et al., 2008] Claeskens, G., Hjort, N. L., et al. (2008). *Model selection and model averaging*, volume 330. Cambridge University Press Cambridge.

[Daykin et al., 1994] Daykin, C. D., Pentikäinen, T., and Pesonen, M. (1994). *Practical risk theory for actuaries*, volume 53. Chapman and Hall, London, 1st edition.

[Devore and Berk, 2007] Devore, J. L. and Berk, K. N. (2007). *Modern Mathematical Statistics with Applications*.

[Devore and Berk, 2012] Devore, J. L. and Berk, K. N. (2012). *Modern mathematical statistics with applications*. Springer texts in statistics. Springer, New York, 2nd edition.

[Dickson, 2005] Dickson, D. C. M. (2005). *Insurance risk and ruin*. Cambridge University Press, Cambridge, UK.

[European Comission, 2009] European Comission (2009). Solvency ii directive (2009/138/ec).

[Feller, 1948] Feller, W. (1948). On the kolmogorov-smirnov limit theorems for empirical distributions.

[Finans Norge, 2016] Finans Norge (2016). Statistikk og nøkkeltall for skadeforsikring 2016.

[Ghosh and Resnick, 2011] Ghosh, S. and Resnick, S. I. (2011). When does the mean excess plot look linear? *Stochastic Models*, 27(4):705–722.

[Gray and Pitts, 2012] Gray, R. J. and Pitts, S. M. (2012). *Risk modelling in general insurance: from principles to practice*. International series on actuarial science. Cambridge University Press, Cambridge.

[Hosking and Wallis, 1987] Hosking, J. R. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349.

[Jong and Heller, 2008] Jong, P. d. and Heller, G. Z. (2008). *Generalized linear models for insurance data*. Cambridge University Press, Cambridge.

[Kleiber and Kotz, 2003] Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. Wiley, Hoboken, NJ.

[Klugman et al., 2008] Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2008). *Loss models: from data to decisions*. John Wiley and Sons, Hoboken, N.J., 3rd edition.

[Korn, 2010] Korn, R. (2010). *Monte Carlo Methods and Models in Finance and Insurance*.

[Lehmann, 1999] Lehmann, E. L. (1999). *Elements of large-sample theory*. Springer, New York.

[Millar, 2011] Millar, R. B. (2011). *Maximum likelihood estimation and inference: with examples in R, SAS, and ADMB*. Statistics in practice. Wiley, Chichester, West Sussex.

[Miranda et al., 2011] Miranda, M. D. M., Nielsen, B., Nielsen, J. P., and Verrall, R. (2011). Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers. *Astin Bulletin*, 41(01):107–129.

[Pickands III, 1975] Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

[Rice, 2007] Rice, J. A. (2007). *Mathematical statistics and data analysis*. Duxbury advanced series. Thomson/Brooks/Cole, Belmont, CA, 3rd edition.

[Robert and Casella, 2004] Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*.

[Rubinstein and Kroese, 2008] Rubinstein, R. Y. and Kroese, D. P. (2008). *Simulation and the Monte Carlo Method*. 2nd edition.

[Sadanori Konishi, 2007] Sadanori Konishi, G. K. (2007). *Information Criteria and Statistical Modeling*.

[Shao and Tu, 1995] Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*.

[Walpole et al., 2007] Walpole, R. E., Myers, R. H., Myers, S. L., and Ye, K. (2007). *Probability and statistics for engineers and scientists*. Pearson, 9th edition.

A Appendix

A.1 The Extended Pareto distribution

The Extended Pareto random variable is given as:

$$Z = \frac{\theta}{\alpha}\frac{G_\theta}{G_\alpha} = \frac{U}{V},$$

where $G_\theta > 0$ and $G_\alpha > 0$ are two independent standard Gamma distributed variables with shape parameters $\theta > 0$ and $\alpha > 0$, respectively [Bølviken, 2014, p. 335]. If G $\sim$ Gamma(λ, 1), then for a constant > 0, λG $\sim$ Gamma(λ, λ) [Gray and Pitts, 2012, p. 29]. Therefore U $\sim$ Gamma(θ,θ) and V $\sim$ Gamma(α,α) are two independent random variables with density functions $f_u(u)$ and $f_v(v)$, respectively [Bølviken, 2014, p. 324].

$$F(Z) = P(Z \leq z) = P(\theta G_\theta \leq z\alpha G_\alpha) = P(U \leq V)$$

The following formula will be used to calculate the Extended Pareto distribution function:

$$f(z) = \int_0^\infty v f_v(v) f_u(vz) du, \tag{71}$$

given the independence of U and V [Rice, 2007, p. 98]. Since U and V are independent variables, it follows that:

$$f_u(vz) = \frac{\theta^\theta}{\Gamma(\theta)} (vz)^{\theta-1} \exp(-vz\theta), \text{ and} \tag{72}$$

$$f_v(v) = \frac{\alpha^\alpha}{\Gamma(\alpha)} v^{\alpha-1} \exp(-v\alpha). \tag{73}$$

Inserting equation (72) and (73) into equation (71) provides the following equation:

$$f(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)}\frac{\theta^\theta}{\Gamma(\theta)} z^{\theta-1} \int_0^\infty v^{\alpha+\theta-1} \exp\left(-v(\alpha+\theta z)\right) dv. \tag{74}$$

Letting v(α+θz) = u and multiplying and dividing equation (74) by $\Gamma(\alpha+\theta)$, leads to the equation:

$$f(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)}\frac{\theta^\theta}{\Gamma(\theta)}\frac{\Gamma(\alpha+\theta)}{z^{\theta-1}}(\alpha+\theta z)^{\alpha+\theta}\underbrace{\int_0^\infty \frac{u^{\alpha+\theta-1}\exp(-u)}{\Gamma(\alpha+\theta)}du}_{=1}. \tag{75}$$

Dividing the numerator and denominator by $\alpha^{\alpha+\theta}$ and letting $\beta = \frac{\alpha}{\theta}$ one obtains:

$$\begin{aligned} f(z) &= \frac{1}{\beta^\theta}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\frac{z^{\theta-1}}{(1+z/\beta)^{\alpha+\theta}} \\ &= \frac{1}{\beta}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\frac{(z/\beta)^{\theta-1}}{(1+z\beta)^{\alpha+\theta}}, \qquad z>0. \end{aligned} \tag{76}$$

The mean and standard deviation of the Extended Pareto distribution can be calculated using the formula:

$$E(Z^i) = \int_0^\infty z^i f(z)dz = \frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\beta^i\int_0^\infty \frac{1}{\beta}\frac{(z/\beta)^{\theta-1+i}}{(1+z/\beta)^{\alpha+\theta}}dz. \tag{77}$$

Recognizing (77) as an Extended Pareto distribution with parameters α - i, and θ + i, the equations will be as below:

$$= \frac{\Gamma(\alpha-i)\Gamma(\theta+i)\beta^i}{\Gamma(\alpha)\Gamma(\theta)}\underbrace{\int_0^\infty \frac{1}{\beta}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha-i)\Gamma(\theta+i)}\frac{(z/\beta)^{\theta-1+i}}{(1+z/\beta)^{\alpha+\theta}}}_{=1}$$

$$= \frac{\Gamma(\alpha-i)\Gamma(\theta+i)\beta^i}{\Gamma(\alpha)\Gamma(\theta)}. \tag{78}$$

Using the recursion formula $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, for $\alpha > 1$ [Walpole et al., 2007, p. 195], the first and second moments yield:

$$\begin{aligned} E(Z) &= \frac{\Gamma(\alpha-1)}{\Gamma(\alpha)}\frac{\Gamma(\theta+1)}{\Gamma(\theta)}\beta = \frac{\Gamma(\alpha-1)}{(\alpha-1)\Gamma(\alpha-1)}\frac{\theta\Gamma(\theta)}{\Gamma(\theta)}\beta \\ &= \frac{\theta\beta}{\alpha-1}, \qquad \alpha>1. \end{aligned}$$

$$E(Z^2) = \frac{\Gamma(\alpha-2)\Gamma(\theta+2)\beta^2}{\Gamma(\alpha)\Gamma(\theta)} = \frac{\Gamma(\alpha-2)}{(\alpha-1)(\alpha-2)\Gamma(\alpha-2)}\frac{\theta(\theta+1)\Gamma(\theta)}{\Gamma(\theta)}\beta^2$$
$$= \frac{\theta(\theta+1)\beta^2}{(\alpha-1)(\alpha-2)}, \qquad \alpha > 2.$$

The standard deviation is then

$$sd(Z) = E(Z)\left(\frac{\alpha+\theta-1}{\theta(\alpha-2)}\right)^{1/2}, \qquad \alpha > 2.$$

A.2 From Extended Pareto to Gamma

The Gamma distribution can be derived as a limit of the Extended Pareto distribution. Letting

$$\beta = \frac{\alpha}{\theta\beta} \quad \text{and} \quad \alpha \to \infty \text{ in equation (52)},$$

gives [Bølviken, 2014, p. 335]:

$$\lim_{\alpha\to\infty} \frac{1}{\beta}\frac{\Gamma(\alpha+\theta)}{\Gamma(\theta)\Gamma(\alpha)}\frac{(z/\beta)^{\theta-1}}{(1+z/\beta)^{\alpha+\theta}} = \lim_{\alpha\to\infty} \underbrace{\frac{\theta\beta}{\alpha}\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}}_{\phi_1}\underbrace{\frac{(z\beta\theta/\alpha)^{\theta-1}}{(1+z\beta\theta/\alpha)^{\alpha+\theta}}}_{\phi_2}. \quad (79)$$

Using the fact that

$$\lim_{\alpha\to\infty} \frac{\Gamma(\alpha+\theta)}{\alpha^\theta\Gamma(\alpha)} = 1 \text{ [Abramowitz and Stegun, 1965, p. 257]},$$

we have

$$\lim_{\alpha\to\infty} \frac{\alpha^{\theta-1}\beta\theta}{\Gamma(\theta)}\underbrace{\frac{\Gamma(\alpha+\theta)}{\alpha^\theta\Gamma(\alpha)}}_{=1} = \lim_{\alpha\to\infty}\frac{\alpha^{\theta-1}\beta\theta}{\Gamma(\theta)}.$$

Setting ϕ_1 and ϕ_2 together yields:

$$\lim_{\alpha\to\infty} \frac{(\beta\theta)^\theta}{\Gamma(\theta)}\frac{z^{\theta-1}}{(1+z\beta\theta/\alpha)^{\alpha+\theta}} = \frac{z^{\theta-1}(\beta\theta)^\theta\exp(-z\beta\theta)}{\Gamma(\theta)}, \qquad z > 0, \quad (80)$$

where equation (80) is a Gamma distribution with shape parameter $\theta > 0$ and scale parameter $\beta\theta > 0$.

A.3 From two-parameter Pareto to Exponential

Dividing the numerator and denominator in (43) by $\beta^{\alpha+1}$ and inserting $\beta = \frac{(\alpha-1)}{\xi}$, the two-parameter Pareto density function leads to the following equation:

$$\underbrace{\frac{\alpha\xi}{\left(\alpha-1\right)}}_{\omega_1}\underbrace{\frac{1}{\left(1+\frac{x\xi}{(\alpha-1)}\right)^{1+\alpha}}}_{\omega_2}. \tag{81}$$

Dividing the numerator and denominator of each term in ω_1 by α and taking the limit as α goes to infinity gives:

$$\lim_{\alpha\to\infty}\omega_1 = \lim_{\alpha\to\infty}\frac{\xi}{(1-\frac{1}{\alpha})} = \xi.$$

Taking the limit of ω_2 as α goes to infinity gives us the following equation:

$$\begin{aligned}
&\lim_{\alpha\to\infty}\omega_2 = \lim_{\alpha\to\infty}\left(1+\frac{x\xi}{(\alpha-1)}\right)^{-\alpha-1}\\
&= \lim_{\alpha\to\infty}\exp\left\{\log\left(\left(1+\frac{x\xi}{(\alpha-1)}\right)^{-\alpha-1}\right)\right\}\\
&= \lim_{\alpha\to\infty}\exp\left\{(-\alpha-1)\log\left(1+\frac{x\xi}{(\alpha-1)}\right)\right\}\\
&= \exp\left\{\lim_{\alpha\to\infty}(-\alpha-1)\log\left(1+\frac{x\xi}{(\alpha-1)}\right)\right\}\\
&= \exp\left\{\lim_{\alpha\to\infty}\alpha\log\left(1+\frac{x\xi}{(\alpha-1)}\right)\left(\frac{-\alpha-1}{\alpha}\right)\right\}\\
&= \exp\left\{-\lim_{\alpha\to\infty}\alpha\log\left(1+\frac{x\xi}{(\alpha-1)}\right)\lim_{\alpha\to\infty}(1/\alpha)\right\}\\
&= \exp\left\{-\lim_{\alpha\to\infty}\frac{\log\left(1+\frac{x\xi}{(\alpha-1)}\right)}{1/\alpha}\right\}.
\end{aligned}$$

The following equation is obtained using L'Hôpital's rule:

$$\exp\left\{-\lim_{\alpha\to\infty}\frac{\frac{\partial\log\left(1+\frac{x\xi}{(\alpha-1)}\right)}{\partial\alpha}}{\frac{\partial\alpha^{-1}}{\partial\alpha}}\right\}=\exp\left\{-\lim_{\alpha\to\infty}\frac{\frac{x\xi}{(\alpha-1)^2(1+\frac{x\xi}{(\alpha-1)})}}{\frac{1}{\alpha^2}}\right\}$$
$$=\exp\left\{-\lim_{\alpha\to\infty}\frac{\alpha^2x\xi}{(\alpha^2-2\alpha+x\xi(\alpha-1)+1}\right\}$$
$$=\exp\left\{-\lim_{\alpha\to\infty}\frac{x\xi}{\left(1-\frac{2}{\alpha}+x\xi(\frac{1}{\alpha}-\frac{1}{\alpha^2})+\frac{1}{\alpha^2}\right)}\right\}=\exp(-x\xi).$$

The following formula results from setting ω_1 and ω_2 together:

$$\xi\exp(-x\xi).$$

A.4 The Log-gamma distribution

Let X be Gamma distributed with probability density function $f_x(x)$ as in equation (25). Using formula (60) with Z = g(X) = exp(X) following the Log-gamma distribution. The inverse function is g^{-1}(X) = g(Z) = log(Z).

$$f_z(z)=f_x\big(g(z)\big)\left|\partial\frac{g(z)}{\partial z}\right|=\frac{\beta^\alpha}{\Gamma(\alpha)}\log(z)^{\alpha-1}e^{-\beta\log(z)}\left|\frac{1}{z}\right|$$
$$=\frac{\beta^\alpha}{\Gamma(\alpha)}\log(z)^{\alpha-1}z^{-\beta-1},\qquad z>0, \tag{82}$$

where $\alpha > 0$ and $\beta > 0$ [Kleiber and Kotz, 2003, p. 169].

A.5 The four-parameter Pareto moments

Letting $\gamma = \tau = 1$, in equation (58) gives:

$$Z=\beta X^\eta,$$

where X = $\frac{G_\theta}{G_\alpha}$ is Extended Pareto distributed with parameters α, θ and $\frac{\alpha}{\theta}$.

The following is obtained from equation (78):

$$E[X^{\eta}] = \left(\frac{\alpha}{\theta}\right)^{\eta} \frac{\Gamma(\alpha-\eta)\Gamma(\theta+\eta)}{\Gamma(\alpha)\Gamma(\theta)}.$$

$$E[Z] = \beta E[X^{\eta}] = \beta \left(\frac{\alpha}{\theta}\right)^{\eta} \frac{\Gamma(\alpha-\eta)\Gamma(\theta+\eta)}{\Gamma(\alpha)\Gamma(\theta)}.$$

Hence,

$$E[X^{2\eta}] = \left(\frac{\alpha}{\theta}\right)^{2\eta} \frac{\Gamma(\alpha-2\eta)\Gamma(\theta+2\eta)}{\Gamma(\alpha)\Gamma(\theta)}.$$

$$E[Z^2] = \beta^2 E[X^{2\eta}] = \beta^2 \left(\frac{\alpha}{\theta}\right)^{2\eta} \frac{\Gamma(\alpha-2\eta)\Gamma(\theta+2\eta)}{\Gamma(\alpha)\Gamma(\theta)}$$

$$(E[Z])^2 = \beta^2 \left(\frac{\alpha}{\theta}\right)^{2\eta} \left(\frac{\Gamma(\alpha-\eta)\Gamma(\theta+\eta)}{\Gamma(\alpha)\Gamma(\theta)}\right)^2$$

$$sd[Z] = \frac{\beta\left(\frac{\alpha}{\theta}\right)^{\eta}}{\Gamma(\alpha)\Gamma(\theta)} \times \sqrt{\Gamma(\alpha)\Gamma(\theta)\Gamma(\alpha-2\eta)\Gamma(\theta+2\eta) - (\Gamma(\alpha-\eta)\Gamma(\theta+\eta)^2}.$$

B R-script

This section contains codes for calculations in section 3.

B.0.1 Reading the data

Reading the total data and removing claims < 0:

```
mydata <- read.csv("C:/Users/stina/Desktop/mydata.csv")
claims1 <- as.numeric(gsub('\\.','',mydata[,8]))
claims <- subset(claims1, claims1 > 0)
```

B.0.2 Descriptive statistics for the total data

Computing descriptive statistics in table 1:

```
summary(claims)
sd(claims)
library(moments)
library(fBasics)
skewness(claims)
kurtosis(claims)
```

B.0.3 Selecting the threshold

Computing the mean excess plot and selecting the data below the threshold b in subsection 3.1:

```
#mean excess plot
library(laeken)
meanExcessPlot(claims)
abline(v=4e+6, \dots)
#setting the threshold
b <- 4e6
z <- subset(claims, claims<=b)
n <- length(z)
#quantile for threshold
ecdf(claims)(b)
```

B.0.4 Descriptive statistics for the under-threshold data

Codes for table 2:

```
summary(z)
sd(z)
skewness(z)
kurtosis(z)
```

B.0.5 Codes used in optimization

```
zBar <- mean(z)
s2 <- var(z)
skewZ <- mean((z-zBar)^3)/s2^1.5
```

B.0.6 Fitting the over-threshold distribution given the total data

```
#setting the overtreshold
b <- 4e6
zc <- subset(claims, claims>b)
zc <- zc-b
n <- length(zc)

#Descriptive statistics used in optimization
zcBar <- mean(zc)
s2c <- var(zc)
skewZc <- mean((zc-zcBar)^3)/s2^1.5

#Estimation, non-truncated
mlpareto <- function(par,zc)
{beta <- exp(par)
alpha <- 1/mean(log(1+zc/beta))
-log(alpha/beta)+1/alpha+1}

alpha.start <- max(1.0001,2*s2c/(s2c-zcBar^2))
beta.start <- zcBar*(alpha.start-1)
start.par <- beta.start
log.start.par <- log(start.par)
```

```
est.pa <- optim(log.start.par,mlpareto,method="BFGS",z=
    zc)
est.beta <- exp(est.pa$par)
est.alpha <- 1/mean(log(1+zc/est.beta))
est.par.pa <- c(est.alpha,est.beta)
alpha.b <- est.par.pa[1]
beta.b <- est.par.pa[2]

ppareto <- function(zc,alpha,beta)
{1-(1+zc/beta)^(-alpha)}

qpareto <- function(u,alpha,beta)
{beta*((1-u)^(-1/alpha)-1)}

#QQ-plot
u <- ppoints(n)
qpar <- qpareto(u,est.par.pa[1],est.par.pa[2])
plot(qpar, sort(zc), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="2-parameter Pareto
    overtreshold Q-Q Plot")

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
F <- ppareto(sort(zc),est.par.pa[1],est.par.pa[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

#Bootstrap
avec=c()
bvec=c()
alpha.b=est.par.pa[1]
beta.b=est.par.pa[2]
m_b=10000
for (i in 1:m_b)

#simulate Pareto
{z_par=beta.b*(runif(n)**(-1/alpha.b)-1) ;
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z_
```

```
    par)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z_par/bvec[i]))
print(c(avec[i],bvec[i]))}

#Confidence intervals
qalp=sort(avec)
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet=sort(bvec)
c(qbet[0.05*m_b],qbet[0.95*m_b])
```

B.0.7 Selecting the reduced data

```
#non parametric bootstrap to select data with
    replacement
library(moments)
library(fBasics)
mb=1

for (i in 1:mb)
{x.star.t=sample(claims,500,replace=FALSE)
print(x.star.t)}

#write.csv(x.star.t, file ="C:/Users/stina/Desktop/Latex
    /smallsample.csv")

mydata.small <- read.csv("C:/Users/stina/Desktop/Latex/
    smallsample.csv")
claims.small <- as.numeric(gsub('\\.','',mydata.small
    [,2]))
sort(claims.small)
max(claims.small)
summary(claims.small)
sd(claims.small)
skewness(claims.small)
kurtosis(claims.small)

#mean exces plot
library(laeken)
```

```
meanExcessPlot(claims.small)
abline(v=2e+6, \dots)

#under-threshold distribution
b <- 1500000
z <- subset(claims.small,  claims.small<=b)
length(z)

#over-threshold distribution
zc <- subset(claims.small, claims.small>b)
length(zc)

#quantile for threshold
ecdf(claims.small)(b)

#descriptive statistic for the under-treshold data
summary(z)
sd(z)
skewness(z)
kurtosis(z)
#Codes used in optimization
zBar <- mean(z)
s2 <- var(z)
skewZ <- mean((z-zBar)^3)/s2^1.5
```

B.0.8 Fitting the over-threshold distribution given the reduced data

```
#setting the overtreshold
zt.h <- subset(claims.small, claims.small>b)
zt.h <- zt.h-b
n <- length(zt.h)

#Fitting the distributions
zcBar <- mean(zt.h)
s2c <- var(zt.h)
skewZc <- mean((zt.h-zcBar)^3)/s2c^1.5
b <- 1.5e6
```

```
# Estimation, non-truncated
mlpareto <- function(par, zt.h)
{beta <- exp(par)
alpha <- 1/mean(log(1+zt.h/beta))
-log(alpha/beta)+1/alpha+1}

alpha.start <- max(1.0001,2*s2c/(s2c-zcBar^2))
beta.start <- zcBar*(alpha.start-1)
start.par <- beta.start
log.start.par <- log(start.par)
est.pa <- optim(log.start.par, mlpareto, method="BFGS", z=
    zt.h)
est.beta <- exp(est.pa$par)
est.alpha <- 1/mean(log(1+zt.h/est.beta))
est.par.pa <- c(est.alpha, est.beta)
alpha.b <- est.par.pa[1]
beta.b <- est.par.pa[2]

ppareto <- function(zt.h, alpha, beta)
{1-(1+zt.h/beta)^(-alpha)}

qpareto <- function(u, alpha, beta)
{beta*((1-u)^(-1/alpha)-1)}

# QQ-plot
u <- ppoints(n)
qpar <- qpareto(u, est.par.pa[1], est.par.pa[2])
plot(qpar, sort(zt.h), xlab="Theoretical quantiles",
    ylab="Sample quantiles", main="Pareto overtreshold Q
    -Q Plot", cex.axis=1.5, cex.main=2.5, cex.lab=2)

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
F <- ppareto(sort(zt.h), est.par.pa[1], est.par.pa[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn, p.value)

#Bootstrap
```

```
avec=c()
bvec=c()
alpha.b=est.par.pa[1]
beta.b=est.par.pa[2]
m_b=100000
for (i in 1:m_b)

#simulate Pareto
{z_par=beta.b*(runif(n)**(-1/alpha.b)-1) ;
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z_
    par)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z_par/bvec[i]))}

#Confidence intervals
qalp=sort(avec)
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet =sort(bvec)
c(qbet[0.05*m_b],qbet[0.95*m_b])
```

The following codes have been used to fit the under-threshold distributions, bootstrap simulation of parameter, and estimate the reserves with 95% confidence intervals for the total data. For the reduced data let b = 1.5×10^6, J = 2340, change the estimated 99% and 99.5% with those estimated for the reduced sample and use the parameters fitted with the over-threshold distribution for the reduced sample.

B.0.9 The Exponential under-threshold distribution

```
b <- 4e6
n <- length(z)

# Estimation, non-truncated
est.par.exp <- 1/zBar

# Estimation truncated
mlexp.c <- function(log.par,z,b)
{beta <- exp(log.par)
b <- 4e6
```

```
#minus loglikelihood for truncated distribution divided
    by n
-log(beta)+beta*mean(z)+log(pexp(b,beta))}

#optimizing beta
est.exp.c <- optim(log(est.par.exp),mlexp.c,method="
   BFGS",z=z,b=b)

#beta
est.par.exp.c <- exp(est.exp.c$par)
betaexp <- est.par.exp.c

#Simulated mean and sd for the truncated distribution
options(max.print=999999)
m = 1000000
z.star=rexp(round(m*1.1),betaexp)
z.star <- (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

#QQ-plot
u <- ppoints(n)
#qexp produces a vector of quantiles for the
    exponential dist.
#pexp produces a vector of probabilites for the
    exponential dist. F(b)
#u*pexp = F(z) = u*F(b)
qpar <- qexp(u*pexp(b,est.par.exp.c),est.par.exp.c)
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Exponential Q-Q Plot"
    ,cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(dexp(z,est.par.exp.c,log=TRUE))-n*log(pexp(b,
    est.par.exp.c)))+2

#BIC
-2*(sum(dexp(z,est.par.exp.c,log=TRUE))-n*log(pexp(b,
    est.par.exp.c)))+log(n)

#Goodness-of-fit: Kolmogorov Smirnov
```

```
# GoF: KS
Fn <- (0:(n-1))/n
#Given distribution function F(z|z \geq b) = F(z)\F(b)
F <- pexp(sort(z),est.par.exp.c)/pexp(b,est.par.exp.c)
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
#Cdf of K_alpha produces the p value
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pexp(b,est.par.exp.c)
for(i in 1:m)
{
z.star <- rexp(2*n,est.par.exp.c)
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pexp(sort(z.star),est.par.exp.c)/normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)


#Bootstrap of uncertainty in beta
m_b=10000
beta.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{x.star=rexp(round(n*1.1),betaexp)
x.star <- (x.star[x.star <= b])[1:n]
optstarexp <- optim(log(1/mean(x.star)),mlexp.c,method=
    "BFGS",z=x.star)
beta.ml.star[i] <- exp(optstarexp$par)
print(beta.ml.star[i])}

#Confidence intervals
q=sort(beta.ml.star)
c(q[0.05*m_b],q[0.95*m_b])

#Reserve
options(max.print=999999)
```

```
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{N.star=rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star=rexp(round(n*2),betaexp)
z.star <- (z.star[z.star<=b])[1:n]
optstarexp <- optim(log(1/mean(z.star)),mlexp.c,method=
    "BFGS",z=z.star)
beta.ml.star[i] <- exp(optstarexp$par)
z.star.b beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar=rpois(1,J*mu.star[i]*T)
N.star.gtb=rbinom(1, N.star.mcstar, p)
N.star.leqb=N.star.mcstar - N.star.gtb
z.star.mcstar <- rexp(round(N.star.leqb*2),beta.ml.star
```

```
    [i])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar<=b])[1:N.
    star.leqb]
z.star.mcstar.b=bvec[i]*(runif(N.star.gtb)**(-1/avec[i
    ])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
print(sort(X, decreasing=TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(0e+8,2e+9),cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at=7.920e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.920e+08,lty=3)
abline(v=q[0.01*mb])
abline(v=q[0.99*mb])
plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",xlim=c(0e+8,2e+9),cex
    .axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at=8.584+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.584e+08,lty=3)
abline(v=q[0.005*mb])
abline(v=q[0.995*mb])

#confidence intervals
q=sort(qeps.star)
c(q[0.01*mb],q[0.99*mb])
c(q[0.005*mb],q[0.995*mb])

#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps=0.01
eps2=0.005
```

```
beta.hat=est.par.exp.c
X=rep(NA,m)
for (j in 1:m)
{N.star.mcstar=rpois(1,J*mu.hat*T)
N.star.gtb=rbinom(1, N.star.mcstar, p)
N.star.leqb=N.star.mcstar - N.star.gtb
z.star.mcstar <- rexp(round(N.star.leqb*1.1),beta.hat)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar<=b])[1:N.
    star.leqb]
z.star.mcstar.b=beta.b*(runif(N.star.gtb)**(-1/alpha.b)
    -1) + b
X[j]=sum(z.star.mcstar) + sum(z.star.mcstar.b)}
print(sort(X, decreasing=TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.10 The Weibull under-threshold distribution

```
b <- 4e6
n <- length(z)

#Estimation, non-truncated
mlweibull <- function(par,z)
{alpha <- exp(par)
-log(alpha)+log(sum(z^alpha))-(alpha-1)*mean(log(z))+1}

findAlphaWeibull <- function(alpha,zBar,s2)
{zBar^2*(gamma(1+2/alpha)/(gamma(1+1/alpha))^2-1)-s2}

alpha.start <- uniroot(findAlphaWeibull,interval=c
    (0.9,1e6),extendInt="yes",zBar=zBar,s2=s2)$root
beta.start <- zBar/gamma(1+1/alpha.start)
start.par <- alpha.start
log.start.par <- log(start.par)
est.we <- optim(log.start.par,mlweibull,method="BFGS",z
    =z)
```

```
est.alpha <- exp(est.we$par)
est.beta <- (mean(z^alpha.start))^(1/alpha.start)
est.par.we <- c(est.alpha,est.beta)

#Estimation truncated
mlweibull.c <- function(log.par,z,b)
{alpha <- exp(log.par[1])
beta <- exp(log.par[2])
b <- 4e6
-sum(dweibull(z,alpha,beta,log=TRUE))+n*log(pweibull(b,
    alpha,beta))}

est.we.c <- optim(log(est.par.we),mlweibull.c,method="
    BFGS",z=z,b=b)
est.par.we.c <- exp(est.we.c$par)
est.alphaw <- exp(est.we.c$par)[1]
est.betaw <- exp(est.we.c$par)[2]

#QQ-plot
u <- ppoints(n)
qpar <- qweibull(u*pweibull(b,est.par.we.c[1],est.par.
    we.c[2]),est.par.we.c[1],est.par.we.c[2])
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Weibull Q-Q Plot",cex
    .axis=2,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(dweibull(z,est.par.we.c[1],est.par.we.c[2],log=
    TRUE))-n*log(weibull(b,est.par.we.c[1],est.par.we.c
    [2])))+2*2

#BIC
-2*(sum(dweibull(z,est.par.we.c[1],est.par.we.c[2],log=
    TRUE))-n*log(pweibull(b,est.par.we.c[1],est.par.we.c
    [2])))+log(n)*2

#Goodness-of-fit: Kolmogorov Smirnov
# GoF: KS
Fn <- (0:(n-1))/n
F <- pweibull(sort(z),est.par.we.c[1],est.par.we.c[2])/
    pweibull(b,est.par.we.c[1],est.par.we.c[2])
```

```
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pweibull(b,est.par.we.c[1],est.par.we.c
    [2])
for(i in 1:m)
{
z.star <- rweibull(2*n,est.par.we.c[1],est.par.we.c[2])
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pweibull(sort(z.star),est.par.we.c[1],est.par
    .we.c[2])/normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)


#simulating mean and sd
m=1000000
z.star <- rweibull(round(m*1.1),shape=est.alphaw,scale=
    est.betaw)
z.star <- (z.star[z.star<=b])[1:m]
mean(z.star)
sd(z.star)

#Bootstrap of uncertainty in parameters
m_b=10000
beta.ml.star=rep(NA,m_b)
alpha.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{x.star <- rweibull(round(n*1.1),shape=est.alphaw,scale
    =est.betaw)
x.star <- (x.star[x.star<=b])[1:n]
par=optim(log(c(est.alphaw,est.betaw)),mlweibull.c,z=x.
    star)$par
alpha.ml.star[i]=par[1]
beta.ml.star[i]=par[2]
```

```
print(exp(c(alpha.ml.star[i],beta.ml.star[i])))}
mean(exp(alpha.ml.star))
mean(exp(beta.ml.star))

#Confidence intervals
qalp =sort(exp(alpha.ml.star))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet =sort(exp(beta.ml.star))
c(qbet[0.05*m_b],qbet[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{N.star=rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star <- rweibull(round(n*2),shape=est.alphaw,scale=
    est.betaw)
z.star <- (z.star[z.star<=b])[1:n]
par=optim(log(c(est.alphaw,est.betaw)),mlweibull.c,z=z.
    star)$par
alpha.ml.star[i]=exp(par[1])
```

```
beta.ml.star[i]=exp(par[2])
z.star.b=beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar=rpois(1,J*mu.star[i]*T)
N.star.gtb=rbinom(1, N.star.mcstar, p)
N.star.leqb=N.star.mcstar - N.star.gtb
z.star.mcstar=rweibull(round(N.star.leqb*2),alpha.ml.
    star[i],beta.ml.star[i])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar<=b])[1:N.
    star.leqb]
z.star.mcstar.b=bvec[i]*(runif(N.star.gtb)**(-1/avec[i
    ])-1) + b
X[j]=sum(z.star.mcstar) + sum(z.star.mcstar.b)}

#print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]

print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(5e+8,2e+9),cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.712e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.712e+08,lty=3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])
plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",ylim=c(0,4e-09),xlim=
    c(5e+8,2e+9),cex.axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at=8.334e+08,expression(hat(epsilon)),cex.axis
```

```
    =0.2)
abline(v=8.334e+08,lty=3)
abline(v=q1[0.05*mb])
abline(v=q1[0.95*mb])
q=sort(qeps.star)
c(q[0.05*mb],q[0.95*mb])
q1=sort(qeps.star2)
c(q1[0.05*mb],q1[0.95*mb])

#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps=0.01
eps2=0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar=rpois(1,J*mu.hat*T)
N.star.gtb=rbinom(1, N.star.mcstar, p)
N.star.leqb=N.star.mcstar - N.star.gtb
z.star.mcstar=rweibull(round(N.star.leqb*1.1),est.
    alphaw,est.betaw)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar<=b])[1:N.
    star.leqb]
z.star.mcstar.b=beta.b*(runif(N.star.gtb)**(-1/alpha.b)
    -1) + b
X[j]=sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.11 The Gamma under-threshold distribution

```
b <- 4e6
```

```
n <- length(z)
#Estimation, non-truncated
mlgamma <- function(par,z)
{alpha <- exp(par)
zBar <- mean(z)
alpha*(1-log(alpha/zBar))+lgamma(alpha)-(alpha-1)*mean(
    log(z))}

alpha.start <- zBar^2/s2
start.par <- alpha.start
log.start.par <- log(start.par)
est.ga <- optim(log.start.par,mlgamma,method="BFGS",z=z
    )
est.beta <- exp(est.ga$par)/zBar
est.par.ga <- c(exp(est.ga$par),est.beta)

#Estimation truncated
mlgamma.c <- function(log.par,z,b)
{b <- 4e6
alpha <- exp(log.par[1])
beta <- exp(log.par[2])
#dgamma is the density function of gamma
# negative LL divided by n
-sum(dgamma(z,alpha,beta,log=TRUE))+n*log(pgamma(b,
    alpha,beta))}

#optimizing the parameters with log(est.par.ga) = the
    parameters for the non-truncated data as starting
    values.
est.ga.c <- optim(log(est.par.ga),mlgamma.c,method="
    BFGS",z=z,b=b)
est.par.ga.c <- exp(est.ga.c$par)
alphag <- exp(est.ga.c$par)[1]
betag <- exp(est.ga.c$par)[2]

#QQ-plot
#generates set of probabilities at which to evaluate
    the inverse distribution
u <- ppoints(n)
#qgamma produces a vector of quantiles for the gamma
    function
```

```
#pgamma produces a vector of probabilites for the gamma
    function F(b)
#u*pgamma=F(z)=u*F(b)
qpar <- qgamma(u*pgamma(b,est.par.ga.c[1],est.par.ga.c
    [2]),est.par.ga.c[1],est.par.ga.c[2])
plot(qpar, sort(z),xlab="Theoretical quantiles",ylab="
    Sample quantiles",main="Gamma Q-Q Plot",cex.axis
    =1.5,cex.main=2.7,cex.lab=2)

#AIC
-2*(sum(dgamma(z,est.par.ga.c[1],est.par.ga.c[2],log=
    TRUE))-n*log(pgamma(b,est.par.ga.c[1],est.par.ga.c
    [2])))+2*2

#BIC
-2*(sum(dgamma(z,est.par.ga.c[1],est.par.ga.c[2],log=
    TRUE))-n*log(pgamma(b,est.par.ga.c[1],est.par.ga.c
    [2])))+log(n)*2

# GoF: KS
Fn <- (0:(n-1))/n
F <- pgamma(sort(z),est.par.ga.c[1],est.par.ga.c[2])/
    pgamma(b,est.par.ga.c[1],est.par.ga.c[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pgamma(b,est.par.ga.c[1],est.par.ga.c[2])
for(i in 1:m)
{
z.star <- rgamma(2*n,est.par.ga.c[1],est.par.ga.c[2])
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pgamma(sort(z.star),est.par.ga.c[1],est.par.
    ga.c[2])/normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)
```

```
#simulating mean and sd
m=1000000
z.star=rgamma(round(m*1.1),shape=alphag,rate=betag)
z.star <- (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

#Bootstrap of uncertainty in parameters
m_b=1000
beta.ml.star=rep(NA,m_b)
alpha.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{x.star=rgamma(round(n*1.1),shape=alphag,rate=betag)
x.star <- (x.star[x.star<=b])[1:n]
par=optim(log(c(alphag,betag)),mlgamma.c,z=x.star)$par
alpha.ml.star[i]=par[1]
beta.ml.star[i]=par[2]
print(exp(c(alpha.ml.star[i],beta.ml.star[i])))}
mean(exp(alpha.ml.star))
mean(exp(beta.ml.star))

#Confidence intervals
qalp=sort(exp(alpha.ml.star))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet=sort(exp(beta.ml.star))
c(qbet[0.05*m_b],qbet[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
```

```
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star=rgamma(round(n*1.1),shape=alphag,rate=betag)
z.star <- (z.star[z.star <= b])[1:n]
par=optim(log(c(alphag,betag)),mlgamma.c,z=z.star)$par
alpha.ml.star[i]=exp(par[1])
beta.ml.star[i]=exp(par[2])
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = rgamma(round(N.star.leqb*1.1),alpha.ml.
    star[i],beta.ml.star[i])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
```

```
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlim=c(0,5e+09),xlab="Reserve",cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at=7.957e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v= 7.957e+08,lty = 3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])
plot(density(qeps.star2),main="Density of simulated
    99.5% reserve",xlim=c(0,1e+10), xlab="Reserve",cex.
    axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 8.615e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.615e+08,lty = 3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])
q =sort(qeps.star)
c(q[0.05*mb],q[0.95*mb])

#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = rgamma(round(N.star.leqb*2),alphag,
    betag)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
```

```
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.12 The Log-gamma under-threshold distribution

```
b <- 4e6
n <- length(z)

#Estimation, non-truncated
mlgamma <- function(par,z)
{alpha <- exp(par)
zBar <- mean(z)
alpha*(1-log(alpha/zBar))+lgamma(alpha)-(alpha-1)*mean(
    log(z))}
zBarLog <- mean(log(z+1))
s2Log <- var(log(z+1))
alpha.start <- zBarLog^2/s2Log
start.par <- alpha.start
log.start.par <- log(start.par)
est.lga <- optim(log.start.par,mlgamma,method="BFGS",z=
    log(z+1))
est.beta <- exp(est.lga$par)/zBarLog
est.par.lga <- c(exp(est.lga$par),est.beta)

#Estimation truncated
mlgamma.c <- function(log.par,z,b)
{alpha <- exp(log.par[1])
beta <- exp(log.par[2])
b <- 4e6
-sum(dgamma(z,alpha,beta,log=TRUE))+n*log(pgamma(b,
    alpha,beta))}
est.lga.c <- optim(log(est.par.lga),mlgamma.c,method="
    BFGS",z=log(z+1),b=log(b+1))
est.par.lga.c <- exp(est.lga.c$par)
```

```
alphalg <- exp(est.lga.c$par)[1]
betalg <- exp(est.lga.c$par)[2]

#QQ-plot
#generates set of probabilities at which to evaluate
    the inverse distribution
u <- ppoints(n)
#qgamma produces a vector of quantiles for the gamma
    distribution
#pgamma produces a vector of probabilites for the gamma
     distribution F(b)
#Z = exp(W)-1, W is gamma(alpha, beta)
#W = log(Z+1)
#f_z(z) = (1\z+1)f_w(log(z+1)), where f_w(log(z+1)) is
    dgamma ( gamma distribution)
#F_z(z) = P(Z leq z) = P(e^w -1 leq z) = P(W leq log(z
    +1)) = F_w(log(z+1)), which is pgamma.
#u*pexp = F(z) = u*F(b)
#u = F_z(z) = F_w(log(z+1))
# log(z+1) = F_w^(-1)(u)
#z = e^(F_w^(-1)(u)) -1, where F_W^(-1)(u) is qgamma

qpar <- exp(qgamma(u*pgamma(log(b+1),est.par.lga.c[1],
    est.par.lga.c[2]),est.par.lga.c[1],est.par.lga.c[2])
    )-1
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Log-gamma Q-Q Plot",
    cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(dgamma(log(z+1),est.par.lga.c[1],est.par.lga.c
    [2],log=TRUE)-log(z+1))-n*log(pgamma(log(b+1),est.
    par.lga.c[1],est.par.lga.c[2])))+2*2

#BIC
-2*(sum(dgamma(log(z+1),est.par.lga.c[1],est.par.lga.c
    [2],log=TRUE)-log(z+1))-n*log(pgamma(log(b+1),est.
    par.lga.c[1],est.par.lga.c[2])))+log(n)*2

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
```

```
F <- pgamma(log(sort(z)+1),est.par.lga.c[1],est.par.lga
    .c[2])/pgamma(log(b+1),est.par.lga.c[1],est.par.lga.
    c[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pgamma(log(b+1),est.par.lga.c[1],est.par.
    lga.c[2])
for(i in 1:m)
{
z.star <- exp(rgamma(2*n,est.par.lga.c[1],est.par.lga.c
    [2]))-1
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pgamma(log(sort(z.star)+1),est.par.lga.c[1],
    est.par.lga.c[2])/normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)

#simulating mean and sd
m=1000000
z.star=exp(rgamma(round(m*1.1),shape=alphalg,rate=
    betalg))
z.star <- (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

#Bootstrap of uncertainty in parameters
m_b=10000
beta.ml.star=rep(NA,m_b)
alpha.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{x.star=rgamma(round(n*1.1),shape=alphalg,rate=betalg)
x.star <- (x.star[x.star <= b])[1:n]
par=optim(log(c(alphalg,betalg)),mlgamma.c,z=x.star)$
    par
```

```
alpha.ml.star[i]=par[1]
beta.ml.star[i]=par[2]
print(exp(c(alpha.ml.star[i],beta.ml.star[i])))}
mean(exp(alpha.ml.star))
mean(exp(beta.ml.star))

#Confidence intervals
qalp =sort(exp(alpha.ml.star))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet =sort(exp(beta.ml.star))
c(qbet[0.05*m_b],qbet[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star=rgamma(round(n*2),shape=alphalg,rate=betalg)
z.star <- (z.star[z.star <= b])[1:n]
par=optim(log(c(alphalg,betalg)),mlgamma.c,z=z.star)$
    par
```

```
alpha.ml.star[i]=exp(par[1])
beta.ml.star[i]=exp(par[2])
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = rgamma(round(N.star.leqb*2),alpha.ml.
    star[i],beta.ml.star[i])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(4.5e+8,2e+9),cex.
    axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 6.291e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=6.291e+08,lty = 3)
abline(v=q[0.01*mb])
abline(v=q[0.99*mb])

plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",xlim=c(4.5e+8,2.5e+9)
    ,cex.axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 6.947+08,expression(hat(epsilon)),cex.axis
```

```
    =0.2)
abline(v=6.947e+08,lty = 3)
abline(v=q[0.005*mb])
abline(v=q[0.995*mb])
q =sort(qeps.star)
c(q[0.01*mb],q[0.99*mb])
c(q[0.005*mb],q[0.995*mb])

#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = rgamma(round(N.star.leqb*1.1),alphalg,
    betalg)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.13 The Log-normal under-threshold distribution

```
b <- 4e6
n <- length(z)
```

```
# Estimation, non-truncated
est.par.ln <- c(mean(log(z)),sd(log(z)))

# Estimation truncated
mllognormal.c <- function(log.par,z,b)
{
mu <- log.par[1]
sigma <- exp(log.par[2])
b <- 4e6
-sum(dlnorm(z,mu,sigma,log=TRUE))+n*log(plnorm(b,mu,
    sigma))
}

est.ln.c <- optim(c(est.par.ln[1],log(est.par.ln[2])),
    mllognormal.c,method="BFGS",z=z,b=b)
est.par.ln.c <- c(est.ln.c$par[1],exp(est.ln.c$par[2]))

muln <- est.ln.c$par[1]
sigmaln <- exp(est.ln.c$par[2])

#simulating mean and sd
m=10000000
z.star=rlnorm(round(m*1.1),muln,sigmaln)
z.star <- (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

# QQ-plot
u <- ppoints(n)
qpar <- qlnorm(u*plnorm(b,est.par.ln.c[1],est.par.ln.c
    [2]),est.par.ln.c[1],est.par.ln.c[2])
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Log-normal Q-Q Plot",
    cex.axis=1.5,cex.main=2.7, cex.lab=2)

# AIC
-2*(sum(dlnorm(z,est.par.ln.c[1],est.par.ln.c[2],log=
    TRUE))-n*log(plnorm(b,est.par.ln.c[1],est.par.ln.c
    [2])))+2*2
```

```
# BIC
-2*(sum(dlnorm(z,est.par.ln.c[1],est.par.ln.c[2],log=
    TRUE))-n*log(plnorm(b,est.par.ln.c[1],est.par.ln.c
    [2])))+log(n)*2

# GoF: KS
Fn <- (0:(n-1))/n
F <- plnorm(sort(z),est.par.ln.c[1],est.par.ln.c[2])/
    plnorm(b,est.par.ln.c[1],est.par.ln.c[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

#Bootstrap of uncertainty in mu and sigma
m_b=10000
mu.ml.star=rep(NA,m_b)
sigma.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{
x.star=rlnorm(round(n*1.1),muln,sigmaln)
x.star <- (x.star[x.star <= b])[1:n]
par=optim(c(muln,log(sigmaln)),mllognormal.c,z=x.star)$
    par
mu.ml.star[i]=par[1]
sigma.ml.star[i]=par[2]
print(c(mu.ml.star[i],exp(sigma.ml.star[i])))
}

mean(mu.ml.star)
mean(exp(sigma.ml.star))
#CI
qmu =sort(mu.ml.star)
c(qmu[0.05*m_b],qmu[0.95*m_b])

qsig =sort(exp(sigma.ml.star))
c(qsig[0.05*m_b],qsig[0.95*m_b])

#RESERVE
options(max.print=9999999)
```

```
m= 100000
mb=1000
n= length(z)
nb=length(zc)
N= nb+ n
mu.hat= 0.03
eps = 0.01
eps2= 0.005
A=N/mu.hat
J= 30000
T = 1
b = 4e6
p = nb/N
mu.ml.star=rep(NA,mb)
sigma.ml.star=rep(NA,mb)
avec = c()
bvec = c()
mu.star=rep(NA,mb)
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
for (i in 1:mb)
{
N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A

z.star=rlnorm(round(n*1.1),muln,sigmaln)
z.star <- (z.star[z.star <= b])[1:n]
par=optim(c(muln,log(sigmaln)),mllognormal.c,z=z.star)$
    par
mu.ml.star[i]=par[1]
sigma.ml.star[i]=exp(par[2])

z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))

X=rep(NA,m)
for (j in 1:m)
{
```

```
N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb

z.star.mcstar = rlnorm(round(N.star.leqb*2),mu.ml.star[
    i],sigma.ml.star[i])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]

z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b

X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]

print(qeps.star[i])
print(qeps.star2[i])}
mean(X)

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(5e+8,2e+9),cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.676e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.676e+08,lty = 3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])

plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",xlim=c(5e+8,2e+9),
    ylim=c(0,4e-09),cex.axis=1.2,cex.main=1.5, cex.lab
    =1.3)
axis(1,at= 8.343e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.343e+08,lty = 3)
abline(v=q1[0.05*mb])
```

```
abline(v=q1[0.95*mb])


q =sort(qeps.star)
c(q[0.05*mb],q[0.95*mb])

q1 =sort(qeps.star2)
c(q1[0.05*mb],q1[0.95*mb])


#MC reserve
options(max.print=9999999)
m=1000000
n= length(z)
nb=length(zc)
N= nb+ n
mu.hat= 0.03
eps = 0.01
eps2= 0.005
A=N/mu.hat
J= 30000
T = 1
b = 4e6
p = nb/N
X=rep(NA,m)
for (j in 1:m)
{
N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb

z.star.mcstar = rlnorm(round(N.star.leqb*1.1),muln,
    sigmaln)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]

z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b

X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
```

```
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]

print(qeps.star)
print(qeps.star2)
```

B.0.14 The two-parameter Pareto under-threshold distribution

```
b <- 4e6
n <- length(z)

# Estimation, non-truncated
mlpareto <- function(par,z)
{beta <- exp(par)
alpha <- 1/mean(log(1+z/beta))
-log(alpha/beta)+1/alpha+1}

alpha.start <- max(1.0001,2*s2/(s2-zBar^2))
beta.start <- zBar*(alpha.start-1)
start.par <- beta.start
log.start.par <- log(start.par)
est.pa <- optim(log.start.par,mlpareto,method="BFGS",z=
    z)
est.beta <- exp(est.pa$par)
est.alpha <- 1/mean(log(1+z/est.beta))
est.par.pa <- c(est.alpha,est.beta)

#Estimation truncated
#pdf
dpareto <- function(z,alpha,beta)
{alpha*(1+z/beta)^(-(alpha+1))/beta}

#cdf
ppareto <- function(z,alpha,beta)
{1-(1+z/beta)^(-alpha)}
```

```
qpareto <- function(u,alpha,beta)
{beta*((1-u)^(-1/alpha)-1)}

mlpareto.c <- function(log.par,z,b)
{alpha <- exp(log.par[1])
beta <- exp(log.par[2])
b <- 4e6
-sum(log(dpareto(z,alpha,beta)))+n*log(ppareto(b,alpha,
    beta))}

est.pa.c <- optim(log(est.par.pa),mlpareto.c,method="
    BFGS",z=z,b=b)
est.par.pa.c <- exp(est.pa.c$par)

#simulating mean and sd
m=1000000
z.star = est.par.pa.c[2]*(runif(round(m*1.1))**(-1/est.
    par.pa.c[1])-1);
z.star <- (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

#QQ-plot
u <- ppoints(n)
qpar <- qpareto(u*ppareto(b,est.par.pa.c[1],est.par.pa.
    c[2]),est.par.pa.c[1],est.par.pa.c[2])
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Two-parameter Pareto
    Q-Q Plot",cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(log(dpareto(z,est.par.pa.c[1],est.par.pa.c[2]))
    )-n*log(ppareto(b,est.par.pa.c[1],est.par.pa.c[2])))
    +2*2

#BIC
-2*(sum(log(dpareto(z,est.par.pa.c[1],est.par.pa.c[2]))
    )-n*log(ppareto(b,est.par.pa.c[1],est.par.pa.c[2])))
    +log(n)*2

#Goodness-of-fit: Kolmogorov Smirnov
```

```
Fn <- (0:(n-1))/n
F <- ppareto(sort(z),est.par.pa.c[1],est.par.pa.c[2])/
    ppareto(b,est.par.pa.c[1],est.par.pa.c[2])
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- ppareto(b,est.par.pa.c[1],est.par.pa.c[2])
for(i in 1:m)
{
z.star <- est.par.pa.c[2]*(runif(round(2*n))^(-1/est.
    par.pa.c[1])-1)
z.star <- (z.star[z.star <= b])[1:n]
F.star <- ppareto(sort(z.star),est.par.pa.c[1],est.par.
    pa.c[2])/normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)

#Bootstrap
alpha.hat = est.par.pa.c[1]
beta.hat = est.par.pa.c[2]
alpha.boot=rep(NA,m_b)
beta.boot=rep(NA,m_b)
m_b = 10000
for (i in 1:m_b)
{z_par = beta.hat*(runif(round(n*1.1))**(-1/alpha.hat)
    -1);
z_par <- (z_par[z_par <= b])[1:n]
par <- optim(log(est.par.pa.c),mlpareto.c,method="BFGS"
    ,z=z_par)$par
alpha.boot[i] <- par[1]
beta.boot[i] <- par[2]
print(exp(c(alpha.boot[i],beta.boot[i])))}

mean(exp(alpha.boot))
mean(exp(beta.boot))
```

```
#Confidence intervals
qalp =sort(exp(alpha.boot))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qbet =sort(exp(beta.boot))
c(qbet[0.05*m_b],qbet[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star = beta.hat*(runif(round(n*2))**(-1/alpha.hat)-1)
    ;
z.star <- (z.star[z.star <= b])[1:n]
par <- optim(log(est.par.pa.c),mlpareto.c,method="BFGS"
    ,z=z.star)$par
alpha.ml.star[i] <- exp(par[1])
beta.ml.star[i] <- exp(par[2])
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
```

```
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = beta.ml.star[i]*(runif(round(N.star.
    leqb*2))**(-1/alpha.ml.star[i])-1)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(5e+8,2e+9),ylim=c
    (0,4e-09),cex.axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.737e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.737e+08,lty = 3)
abline(v=q[0.01*mb])
abline(v=q[0.99*mb])

plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",xlim=c(5e+8,3.1e+9),
    ylim=c(0,3.1e-09),cex.axis=1.2,cex.main=1.5, cex.lab
    =1.3)
axis(1,at= 8.164+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.164e+08,lty = 3)
abline(v=q[0.005*mb])
```

```
abline(v=q[0.995*mb])

q =sort(qeps.star)
c(q[0.01*mb],q[0.99*mb])
c(q[0.005*mb],q[0.995*mb])
#c(q[0.05*mb],q[0.95*mb])

#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar = beta.hat*(runif(round(N.star.leqb*2))**
    (-1/alpha.hat)-1);
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.15 The Extended Pareto under-threshold distribution

```
b <- 4e6
n <- length(z)
```

```
#Estimation, non-truncated
mlExtPareto <- function(log.par,z)
{par <- exp(log.par)
alpha <- par[1]
theta <- par[2]
beta <- par[3]
b <- 4e6
mloglik <- -sum(lgamma(alpha+theta)-lgamma(alpha)-
    lgamma(theta)+theta*(log(theta)-log(alpha)-log(beta)
    )+(theta-1)*log(z)-(alpha+theta)*log(1+z*theta/(
    alpha*beta)))
if(is.na(mloglik))
mloglik <- 1e30
else if(mloglik > 1e30)
mloglik <- 1e30
mloglik}

findAlpha <- function(alpha,zBar,s2,skewZ)
{theta <- zBar^2*(alpha-1)/(s2*(alpha-2)-zBar^2)
2*sqrt((alpha-2)/(theta*(alpha+theta-1)))*(alpha+2*
    theta-1)/(alpha-3)-skewZ}

alpha.start <- uniroot(findAlpha,interval=c(3.0000001,1
    e6),extendInt="yes",zBar=zBar,s2=s2,skewZ=skewZ)$
    root
if(alpha.start <= 3)
alpha.start <- 1000
theta.start <- max(zBar^2*(alpha.start-1)/(s2*(alpha.
    start-2)-zBar^2),1e-6)
beta.start <- zBar*(alpha.start-1)/alpha.start
start.par <- c(alpha.start,theta.start,beta.start)
log.start.par <- log(start.par)
est.ep <- optim(log.start.par,mlExtPareto,z=z)
est.par.ep <- exp(est.ep$par)

#Estimation truncated
rExtPareto <- function(m, par)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
```

```
Gtheta <- rgamma(m, theta , theta )
Galpha <- rgamma(m, alpha , alpha )
beta*Gtheta/Galpha}

dExtPareto <- function(z,par)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
exp(lgamma(alpha+theta)-lgamma(alpha)-lgamma(theta)+
    theta*(log(theta)-log(beta)-log(alpha))+(theta-1)*
    log(z)-(alpha+theta)*log(1+z*theta/(beta*alpha)))}

pExtPareto <- function(z,par)
{tmp <- try(integrate(dExtPareto,lower=0,upper=z,par=
    par,stop.on.error=FALSE)$value,silent=TRUE)
if(is.numeric(tmp))
tmp
else
1}

qExtPareto <- function(p,par,q.start,eps=0.1,tol=1e-9)
{q.c <- max(q.start,0)
q.l <- max(q.start-eps,0)
q.u <- max(q.start+eps,0)
p.c <- pExtPareto(q.c,par)
p.l <- pExtPareto(q.l,par)
p.u <- pExtPareto(q.u,par)

if(abs(p.c-p) > tol)
{while(abs(p.c-p) > tol)
{if(p.c > p)
{if(p.l > p)
{  while(p.l > p)
{q.l <- max(q.l-eps,0)
p.l <- pExtPareto(q.l,par)}}
q.u <- q.c
q.c <- (q.l+q.u)/2
p.c <- pExtPareto(q.c,par)
p.u <- pExtPareto(q.u,par)}
else
{if(p.u < p)
```

```
{while(p.u < p)
{q.u <- max(q.u+eps,0)
p.u <- pExtPareto(q.u,par)}}
q.l <- q.c
q.c <- (q.l+q.u)/2
p.c <- pExtPareto(q.c,par)
p.l <- pExtPareto(q.l,par)}}}
q.c}

mlextpareto.c <- function(log.par,z,b)
{par <- exp(log.par)
-sum(log(dExtPareto(z,par)))+n*log(pExtPareto(b,par))}

est.ep.c <- optim(log(est.par.ep),mlextpareto.c,method=
    "BFGS",z=z,b=b)
est.par.ep.c <- exp(est.ep.c$par)

#simulating mean and sd
m=1000000
u1=rgamma(round(m*2),est.par.ep.c[2],est.par.ep.c[2])
u2=rgamma(round(m*2),est.par.ep.c[1],est.par.ep.c[1])
z.star=(est.par.ep.c[3])*(u1/u2)
z.star = (z.star[z.star <= b])[1:m]
mean(z.star)
sd(z.star)

#QQ-plot
u <- ppoints(n)
z.sim <- rExtPareto(1e6,est.par.ep.c)
q.start <- quantile(z.sim,u*pExtPareto(b,est.par.ep.c))
qpar <- rep(0,n)
Fb <- pExtPareto(b,est.par.ep.c)
for(i in 1:n)
{qpar[i] <- qExtPareto(u[i]*Fb,est.par.ep.c,q.start[i],
    eps=0.1*q.start[i],tol=1e-3)}
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Extended Pareto Q-Q
    Plot",cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(log(dExtPareto(z,est.par.ep.c)))-n*log(
```

```
    pExtPareto(b,est.par.ep.c)))+2*3

#BIC
-2*(sum(log(dExtPareto(z,est.par.ep.c)))-n*log(
    pExtPareto(b,est.par.ep.c)))+log(n)*3

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
z.sort <- sort(z)
F <- pExtParetoVec(z.sort,est.par.ep.c)/pExtPareto(b,
    est.par.ep.c)
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pExtPareto(b,est.par.ep.c)
for(i in 1:m)
{
z.star <- rExtPareto(2*n,est.par.ep.c)
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pExtParetoVec(sort(z.star),est.par.ep.c)/
    normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)

#Bootstrap of uncertainty in parameters
m_b=10000
alpha.ml.star=rep(NA,m_b)
theta.ml.star=rep(NA,m_b)
beta.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{u1=rgamma(round(n*2),est.par.ep.c[2],est.par.ep.c[2])
u2=rgamma(round(n*2),est.par.ep.c[1],est.par.ep.c[1])
x.star=(est.par.ep.c[3])*(u1/u2)
x.star = (x.star[x.star <= b])[1:n]
par=optim(log(c(est.par.ep.c[1],est.par.ep.c[2],est.par
```

```
    .ep.c[3])),mlextpareto.c  ,z=x.star,  b=b)$par
alpha.ml.star[i]=par[1]
theta.ml.star[i]=par[2]
beta.ml.star[i]=par[3]
print(exp(c(alpha.ml.star[i],theta.ml.star[i], beta.ml.
    star[i])))}
mean(exp(alpha.ml.star))
mean(exp(theta.ml.star))
mean(exp(beta.ml.star))

#Confidence intervals
qalp =sort(exp(alpha.ml.star))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qthet = sort(exp(theta.ml.star))
c(qthet[0.05*m_b],qthet[0.95*m_b])
qbet =sort(exp(beta.ml.star))
c(qbet[0.05*m_b],qbet[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
```

```
{N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
u1=rgamma(round(n*2),est.par.ep.c[2],est.par.ep.c[2])
u2=rgamma(round(n*2),est.par.ep.c[1],est.par.ep.c[1])
z.star=(est.par.ep.c[3])*(u1/u2)
z.star = (z.star[z.star <= b])[1:n]
par=optim(log(c(est.par.ep.c[1],est.par.ep.c[2],est.par
    .ep.c[3])),mlextpareto.c  ,z=z.star,b=b)$par
alpha.ml.star[i]=exp(par[1])
theta.ml.star[i]=exp(par[2])
beta.ml.star[i]=exp(par[3])
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
u1=rgamma(round(N.star.leqb*2),theta.ml.star[i],theta.
    ml.star[i])
u2=rgamma(round(N.star.leqb*2),alpha.ml.star[i],alpha.
    ml.star[i])
z.star.mcstar=(beta.ml.star[i])*(u1/u2)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
```

```
plot(density(qeps.star),main="Density of simulated 99% reserve", xlab="Reserve",xlim=c(0,2e+9),cex.axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.847e+08,expression(hat(epsilon)),cex.axis=0.2)
abline(v=7.847e+08,lty = 3)
abline(v=q[0.01*mb])
abline(v=q[0.99*mb])
plot(density(qeps.star2),main="Density of simulated 99.5% reserve", xlab="Reserve",xlim=c(0,2.5e+9),cex.axis=1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 8.533+08,expression(hat(epsilon)),cex.axis=0.2)
abline(v=8.533+08,lty = 3)
abline(v=q[0.005*mb])
abline(v=q[0.995*mb])
q =sort(qeps.star)
c(q[0.01*mb],q[0.99*mb])
c(q[0.005*mb],q[0.995*mb])


#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
u1=rgamma(round(N.star.leqb*2),theta.hat,theta.hat)
u2=rgamma(round(N.star.leqb*2),alpha.hat,alpha.hat)
z.star.mcstar=(beta.hat)*(u1/u2)
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
```

```
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.16 The four-parameter Pareto under-threshold distribution

```
b <- 4e6
n <- length(z)
#Estimation, non-truncated
mlFourPar <- function(log.par,z)
{par <- exp(log.par)
alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
v <- z/beta
mloglik <- -sum(lgamma(alpha+theta)-lgamma(alpha)-
   lgamma(theta)+theta*(log(theta)-log(alpha))+(theta/
   eta-1)*log(v)-(alpha+theta)*log(1+theta*v^(1/eta)/
   alpha)-log(beta)-log(eta))
if(is.na(mloglik))
mloglik <- 1e30
else if(mloglik > 1e30)
mloglik <- 1e30
mloglik}

start.par <- c(est.par.ep,1)
log.start.par <- log(start.par)
est.fp <- optim(log.start.par,mlFourPar,z=z)
est.par.fp <- exp(est.fp$par)

#Estimation truncated
rFourPar <- function(m,par)
{alpha <- par[1]
theta <- par[2]
```

```
beta <- par[3]
eta <- par[4]
Gtheta <- rgamma(m, theta , theta )
Galpha <- rgamma(m, alpha , alpha )
x <- Gtheta/Galpha
fourParTransform(par , Gtheta/Galpha)}

fourParTransform <- function(par ,x)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
beta*x^eta}

fourParTransformInv <- function(par ,z)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
(z/beta)^(1/eta)}

dFourPar <- function(z ,par)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
v <- z/beta
exp(lgamma(alpha+theta)-lgamma(alpha)-lgamma(theta)+
    theta*(log(theta)-log(alpha))+(theta/eta-1)*log(v)-(
    alpha+theta)*log(1+theta*v^(1/eta)/alpha)-log(beta)-
    log(eta))}

pFourPar <- function(z ,par)
{tmp <- try(integrate(dFourPar ,lower=0,upper=z ,par=par ,
    stop.on.error=FALSE)$value ,silent=TRUE)
if(is.numeric(tmp))
tmp
else
1}

qFourPar <- function(p ,par ,q.start ,eps=0.1,tol=1e-9)
```

```
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
q.start.ep <- fourParTransformInv(par,q.start)
eps <- fourParTransformInv(par,eps)
par.ep <- c(par[1:2],1)
q.ep <- qExtPareto(p,par.ep,q.start.ep,eps,tol)
fourParTransform(par,q.ep)}

mlfourpar.c <- function(log.par,z,b)
{par <- exp(log.par)
-sum(log(dFourPar(z,par)))+n*log(pFourPar(b,par))}
est.fp.c <- optim(log(est.par.fp),mlfourpar.c,method="
    BFGS",z=z,b=b)
est.par.fp.c <- exp(est.fp.c$par)

#simulating mean and sd
m=1000000
u1=rgamma(round(m*1.1),est.par.fp.c[2],est.par.fp.c[2])
u2=rgamma(round(m*1.1),est.par.fp.c[1],est.par.fp.c[1])
x.star=(est.par.fp.c[3])*(u1/u2)**(est.par.fp.c[4])
x.star = (x.star[x.star <= b])[1:m]
mean(z.star)
sd(z.star)

#QQ-plot
u <- ppoints(n)
z.sim <- rFourPar(1e6,est.par.fp.c)
q.start <- quantile(z.sim,u*pFourPar(b,est.par.fp.c))
qpar <- rep(0,n)
Fb <- pFourPar(b,est.par.fp.c)
for(i in 1:n)
{qpar[i] <- qFourPar(u[i]*Fb,est.par.fp.c,q.start[i],
    eps=0.1*q.start[i],tol=1e-3)
}
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Four-parameter Pareto
    Q-Q Plot",cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(log(dFourPar(z,est.par.fp.c)))-n*log(pFourPar(b
```

```
    ,est.par.fp.c)))+2*4

#BIC
-2*(sum(log(dFourPar(z,est.par.fp.c)))-n*log(pFourPar(b
    ,est.par.fp.c)))+log(n)*4

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
z.sort <- sort(z)
F <- pFourParVec(z.sort,est.par.fp.c)/pFourPar(b,est.
    par.fp.c)
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pFourPar(b,est.par.fp.c)
for(i in 1:m)
{
z.star <- rFourPar(2*n,est.par.fp.c)
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pFourParVec(sort(z.star),est.par.fp.c)/
    normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)

#Bootstrap of uncertainty in parameters
m_b=10000
alpha.ml.star=rep(NA,m_b)
theta.ml.star=rep(NA,m_b)
beta.ml.star=rep(NA,m_b)
eta.ml.star=rep(NA,m_b)

for (i in 1:m_b)
{u1=rgamma(round(n*2),est.par.fp.c[2],est.par.fp.c[2])
u2=rgamma(round(n*2),est.par.fp.c[1],est.par.fp.c[1])
x.star=(est.par.fp.c[3])*(u1/u2)**(est.par.fp.c[4])
```

```
x.star = (x.star[x.star <= b])[1:n]
par=optim(log(c(est.par.fp.c[1],est.par.fp.c[2],est.par
    .fp.c[3], est.par.fp.c[4])),mlfourpar.c ,z=x.star, b
    =b)$par
alpha.ml.star[i]=par[1]
theta.ml.star[i]=par[2]
beta.ml.star[i]=par[3]
eta.ml.star[i]=par[4]
print(exp(c(alpha.ml.star[i],theta.ml.star[i], beta.ml.
    star[i], eta.ml.star[i])))}

mean(exp(alpha.ml.star))
mean(exp(theta.ml.star))
mean(exp(beta.ml.star))
mean(exp(eta.ml.star))

#Confidence intervals
qalp =sort(exp(alpha.ml.star))
c(qalp[0.05*m_b],qalp[0.95*m_b])
qthet = sort(exp(theta.ml.star))
c(qthet[0.05*m_b],qthet[0.95*m_b])
qbet =sort(exp(beta.ml.star))
c(qbet[0.05*m_b],qbet[0.95*m_b])
qeta =sort(exp(eta.ml.star))
c(qeta[0.05*m_b],qeta[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
```

```
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{ N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
u1=rgamma(round(n*2),est.par.fp.c[2],est.par.fp.c[2])
u2=rgamma(round(n*2),est.par.fp.c[1],est.par.fp.c[1])
z.star=(est.par.fp.c[3])*(u1/u2)**(est.par.fp.c[4])
z.star = (z.star[z.star <= b])[1:n]
par=optim(log(c(est.par.fp.c[1],est.par.fp.c[2],est.par
    .fp.c[3], est.par.fp.c[4])),mlfourpar.c ,z=z.star, b
    =b)$par
alpha.ml.star[i]=exp(par[1])
theta.ml.star[i]=exp(par[2])
beta.ml.star[i]=exp(par[3])
eta.ml.star[i]=exp(par[4])
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
u1=rgamma(round(N.star.leqb*2),theta.ml.star[i],theta.
    ml.star[i])
u2=rgamma(round(N.star.leqb*2),alpha.ml.star[i],alpha.
    ml.star[i])
z.star.mcstar=(beta.ml.star[i])*(u1/u2)**(eta.ml.star[i
    ])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
```

```
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density of simulated 99%
    reserve", xlab="Reserve",xlim=c(0,1.5e+9),cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.924e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.924e+08,lty = 3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])
plot(density(qeps.star2),main="Density of simulated
    99.5% reserve", xlab="Reserve",xlim=c(0,1.5e+9),ylim
    =c(0,4.85e-09),cex.axis=1.2,cex.main=1.5, cex.lab
    =1.3)
axis(1,at= 8.522e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.522e+08,lty = 3)
abline(v=q1[0.05*mb])
abline(v=q1[0.95*mb])

q =sort(qeps.star)
c(q[0.05*mb],q[0.95*mb])
q1 =sort(qeps.star2)
c(q1[0.05*mb],q1[0.95*mb])


#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
```

```
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
u1=rgamma(round(N.star.leqb*2),est.par.fp.c[2],est.par.
    fp.c[2])
u2=rgamma(round(N.star.leqb*2),est.par.fp.c[1],est.par.
    fp.c[1])
z.star.mcstar=(est.par.fp.c[3])*(u1/u2)**(est.par.fp.c
    [4])
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}
print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star)
print(qeps.star2)
```

B.0.17 The six-parameter Pareto under-threshold distribution

```
b <- 4e6
n <- length(z)

#Estimation, non-truncated
mlSixPar <- function(log.par,z)
{par <- exp(log.par)
alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
invgamma <- exp(log.par[6])/(1+exp(log.par[6]))
```

```
v <- z/beta+1
w <- v^invgamma
mloglik <- -sum(lgamma(alpha+theta)-lgamma(alpha)-
    lgamma(theta)+theta*(log(theta)+log(tau)/eta-log(
    alpha))-log(eta)+log(invgamma)-log(beta)+(invgamma
    -1)*log(v)+(theta/eta-1)*log(w-1)-(alpha+theta)*log(
    theta*tau^(1/eta)/alpha*(w-1)^(1/eta)+1))
if(is.na(mloglik))
mloglik <- 1e30
else if(mloglik > 1e30)
mloglik <- 1e30
mloglik}
start.par <- c(est.par.fp,rep(1,2))
start.par[6] <- start.par[6]+1e-5
log.start.par <- log(start.par)
log.start.par[6] <- -log(start.par[6])-log(1-1/start.
    par[6])
est.sp <- optim(log.start.par,mlSixPar,z=z)
est.par.sp <- exp(est.sp$par)
est.par.sp[6] <- 1/est.par.sp[6]+1.00001
gamma.sp <- 1/est.par.sp[6]+1

# Estimation truncated
rSixPar <- function(m,par)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- par[6]
Gtheta <- rgamma(m,theta,theta)
Galpha <- rgamma(m,alpha,alpha)
sixParTransform(par,Gtheta/Galpha)}

sixParTransform <- function(par,x)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- par[6]
```

```
beta*((1+x^eta/tau)^gamma-1)}

sixParTransform2 <- function(par,y)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- par[6]
x <- alpha*y/theta
beta*((1+x^eta/tau)^gamma-1)}

sixParTransformInv <- function(par,z)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- par[6]
v <- z/beta+1
w <- v^(1/gamma)
tau^(1/eta)*(w-1)^(1/eta)}

dSixPar <- function(z,par)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- par[6]
v <- z/beta+1
w <- v^(1/gamma)
exp(lgamma(alpha+theta)-lgamma(alpha)-lgamma(theta)-log
    (eta)-log(gamma)-log(beta)+theta*(log(theta)+log(tau
    )/eta-log(alpha))+(1/gamma-1)*log(v)+(theta/eta-1)*
    log(w-1)-(alpha+theta)*log(theta*tau^(1/eta)/alpha*(
    w-1)^(1/eta)+1))}

pSixPar <- function(z,par)
{tmp <- try(integrate(dSixPar,lower=0,upper=z,par=par,
    stop.on.error=FALSE)$value,silent=TRUE)
```

```
if(is.numeric(tmp))
tmp
else
1}

qSixPar <- function(p,par,q.start,eps=0.1,tol=1e-9)
{alpha <- par[1]
theta <- par[2]
beta <- par[3]
q.start.ep <- sixParTransformInv(par,q.start)
eps <- fourParTransformInv(par,eps)
par.ep <- c(par[1:2],1)
q.ep <- qExtPareto(p,par.ep,q.start.ep,eps,tol)
sixParTransform(par,q.ep)}

mlsixpar.c <- function(log.par,z,b)
{par <- exp(log.par)
alpha <- par[1]
theta <- par[2]
beta <- par[3]
eta <- par[4]
tau <- par[5]
gamma <- (1+exp(log.par[6]))/exp(log.par[6])
b <- 4e6
par <- c(alpha,theta,beta,eta,tau,gamma)
-sum(log(dSixPar(z,par)))+n*log(pSixPar(b,par))}

est.sp.c <- optim(c(log(est.par.sp[1:5]),-log(est.par.
    sp[6])-log(1-1/est.par.sp[6])),mlsixpar.c,method="
    BFGS",z=z,b=b)
est.par.sp.c <- exp(est.sp.c$par)
est.par.sp.c[6] <- 1/est.par.sp.c[6]+1
gamma.sp <- 1/est.par.sp.c[6]+1

#beta/theta
est.par.sp.c[3]/est.par.sp.c[4]

#simulating mean and sd
options(max.print=9999999)
m=1000000
x.star=rSixPar(round(1.1*m),est.par.sp.c)
```

```
x.star <- (x.star[x.star <= b])[1:m]
mean(z.star)
sd(z.star)

#QQ-plot
u <- ppoints(n)
z.sim <- rSixPar(1e6,est.par.sp.c)
q.start <- quantile(z.sim,u*pSixPar(b,est.par.sp.c))
qpar <- rep(0,n)
Fb <- pSixPar(b,est.par.sp.c)
for(i in 1:n)
{qpar[i] <- qSixPar(u[i]*Fb,est.par.sp.c,q.start[i],eps
    =0.1*q.start[i],tol=1e-3)}
plot(qpar, sort(z), xlab = "Theoretical quantiles",
    ylab="Sample quantiles", main="Six-parameter Pareto
    Q-Q Plot",cex.axis=1.5,cex.main=2.7, cex.lab=2)

#AIC
-2*(sum(log(dSixPar(z,est.par.sp.c)))-n*log(pSixPar(b,
    est.par.sp.c)))+2*6

#BIC
-2*(sum(log(dSixPar(z,est.par.sp.c)))-n*log(pSixPar(b,
    est.par.sp.c)))+log(n)*6

#Goodness-of-fit: Kolmogorov Smirnov
Fn <- (0:(n-1))/n
z.sort <- sort(z)
F <- pSixParVec(z.sort,est.par.sp.c)/pSixPar(b,est.par.
    sp.c)
Dn <- max(abs(F-Fn))
v <- sqrt(n)*Dn
p.value <- 1-sqrt(2*pi)/v*sum(exp(-0.125*seq(1,10000,2)
    ^2*pi^2/v^2))
c(Dn,p.value)

m <- 1e5
Dn.sim <- rep(0,m)
normconst <- pSixPar(b,est.par.sp.c)
for(i in 1:m)
{
```

```
z.star <- rSixPar(2*n,est.par.sp.c)
z.star <- (z.star[z.star <= b])[1:n]
F.star <- pSixParVec(sort(z.star),est.par.sp.c)/
    normconst
Dn.sim[i] <- max(abs(F.star-Fn))
}
p.value.sim <- mean(Dn.sim >= Dn)

#Bootstrap uncertainty
options(max.print=9999999)
m_b=10000
alpha.ml.star=rep(NA,m_b)
theta.ml.star=rep(NA,m_b)
beta.ml.star=rep(NA,m_b)
eta.ml.star=rep(NA,m_b)
tau.ml.star=rep(NA,m_b)
gamma.ml.star=rep(NA,m_b)
for (i in 1:m_b)
{x.star=rSixPar(round(2*n),est.par.sp.c)
x.star <- (x.star[x.star <= b])[1:n]
b <- 4e6
par=optim(c(log(est.par.sp[1:5]),-log(est.par.sp[6])-
    log(1-1/est.par.sp[6])),mlsixpar.c,z=x.star,b=b)$par
alpha.ml.star[i]=exp(par[1])
theta.ml.star[i]=exp(par[2])
beta.ml.star[i]=exp(par[3])
eta.ml.star[i]=exp(par[4])
tau.ml.star[i]=exp(par[5])
gamma.ml.star[i]=1/exp(par[6])+1
print(c(alpha.ml.star[i],theta.ml.star[i],beta.ml.star[
    i],eta.ml.star[i],tau.ml.star[i], gamma.ml.star[i]))
    }

#Confidence intervals
qalp =sort(alpha.ml.star)
c(qalp[0.05*m_b],qalp[0.95*m_b])
qthet =sort(theta.ml.star)
c(qthet[0.05*m_b],qthet[0.95*m_b])
qbet =sort(beta.ml.star)
c(qbet[0.05*m_b],qbet[0.95*m_b])
qeta =sort(eta.ml.star)
```

```
c(qeta[0.05*m_b],qeta[0.95*m_b])
qtau =sort(tau.ml.star)
c(qtau[0.05*m_b],qtau[0.95*m_b])
qgam =sort(gamma.ml.star)
c(qgam[0.05*m_b],qgam[0.95*m_b])

#Reserve
m=100000
mb=1000
n=length(z)
nb=length(zc)
N=nb+n
mu.hat=0.03
beta.hat=betaexp
eps=0.01
eps2=0.005
A=N/mu.hat
J=30000
T=1
b=4e6
p=nb/N
qeps.star=rep(NA,mb)
qeps.star2=rep(NA,mb)
beta.ml.star=rep(NA,mb)
avec=c()
bvec=c()
mu.star=rep(NA,mb)

for (i in 1:mb)
{ N.star= rpois(1,A*mu.hat)
mu.star[i] <- N.star/A
z.star=rSixPar(round(2*n),est.par.sp.c)
z.star <- (z.star[z.star <= b])[1:n]
b <- 4e6
par=optim(c(log(est.par.sp[1:5]),-log(est.par.sp[6])-
    log(1-1/est.par.sp[6])),mlsixpar.c,z=z.star,b=b)$par
alpha.ml.star[i]=exp(par[1])
theta.ml.star[i]=exp(par[2])
beta.ml.star[i]=exp(par[3])
eta.ml.star[i]=exp(par[4])
tau.ml.star[i]=exp(par[5])
```

```
gamma.ml.star[i]=1/exp(par[6])+1
z.star.b = beta.b*(runif(nb)**(-1/alpha.b)-1)
est.pa <- optim(log(beta.b),mlpareto,method="BFGS",z=z.
    star.b)
bvec[i] <- exp(est.pa$par)
avec[i] <- 1/mean(log(1+z.star.b/bvec[i]))
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.star[i]*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar=rSixPar(round(N.star.leqb*2),c(alpha.ml.
    star[i],theta.ml.star[i],beta.ml.star[i],eta.ml.star
    [i],tau.ml.star[i],gamma.ml.star[i]))
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = bvec[i]*(runif(N.star.gtb)**(-1/avec[
    i])-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star[i]=apply(X,2,sort)[m*(1-eps)]
qeps.star2[i]=apply(X,2,sort)[m*(1-eps2)]
print(qeps.star[i])
print(qeps.star2[i])}

#density
plot(density(qeps.star),main="Density␣of␣simulated␣99%␣
    reserve", xlab="Reserve",xlim=c(0e+8,2e+9),cex.axis
    =1.2,cex.main=1.5, cex.lab=1.3)
axis(1,at= 7.918e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=7.918e+08,lty = 3)
abline(v=q[0.05*mb])
abline(v=q[0.95*mb])
plot(density(qeps.star2),main="Density␣of␣simulated␣
    99.5%␣reserve", xlab="Reserve",xlim=c(0e+8,2.5e+9),
    ylim=c(0,4e-09),cex.axis=1.2,cex.main=1.5, cex.lab
    =1.3)
```

```
axis(1,at= 8.519e+08,expression(hat(epsilon)),cex.axis
    =0.2)
abline(v=8.519e+08,lty = 3)
abline(v=q1[0.05*mb])
abline(v=q1[0.95*mb])
q =sort(qeps.star)
c(q[0.05*mb],q[0.95*mb])
q1 =sort(qeps.star2)
c(q1[0.05*mb],q1[0.95*mb])



#Monte Carlo simulated reserve
options(max.print=999999)
m=100000
eps = 0.01
eps2 = 0.005
beta.hat=est.par.exp.c
X=rep(NA,m)

for (j in 1:m)
{N.star.mcstar = rpois(1,J*mu.hat*T)
N.star.gtb = rbinom(1, N.star.mcstar, p)
N.star.leqb = N.star.mcstar - N.star.gtb
z.star.mcstar=rSixPar(round(N.star.leqb*2),est.par.sp.c
    )
z.star.mcstar <- (z.star.mcstar[z.star.mcstar <= b])[1:
    N.star.leqb]
z.star.mcstar.b = beta.b*(runif(N.star.gtb)**(-1/alpha.
    b)-1) + b
X[j] = sum(z.star.mcstar) + sum(z.star.mcstar.b)}

print(sort(X, decreasing = TRUE))
X=matrix(X,m)
qeps.star=apply(X,2,sort)[m*(1-eps)]
qeps.star2=apply(X,2,sort)[m*(1-eps2)]

print(qeps.star)
print(qeps.star2)
```