

Application of Generalized Pareto in Non-Life Insurance

Mohamed Hanafy

Department Statistics, Zhejiang Gongshang University, Hangzhou, China

Email: mhanafy102@gmail.com

How to cite this paper: Hanafy, M. (2020). Application of Generalized Pareto in Non-Life Insurance. *Journal of Financial Risk Management*, 9, 334-353.
<https://doi.org/10.4236/jfrm.2020.93018>

Received: August 24, 2020

Accepted: September 26, 2020

Published: September 29, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper focuses on the modeling and estimation of tail loss distribution parameters from Egyptian's commercial fire loss severities. Using theoretical extreme value, we use the generalized distribution of Pareto (GPD) and compare it to standard parametric modeling based on exp, Weibull, gumbel, frechet, lognormal and gamma distributions. The goodness-of-fit tests included Kolmogorov-Smirnov, Anderson and Cramer-von Mises test is carried out, and the calculation of the value-at-risk and expected shortfall are performed. We use the bootstrap approach to create confidence intervals for the estimates.

Keywords

Genpareto, Claim Size, Goodness of Fit, Non-Life Insurance

1. Introduction

In the case of a non-life insurance company, only a few individual claims made on a portfolio often constitute the majority of the allowances paid by the company. Among the largest insurance claims, commercial fire insurance is of the highest value. Gaining an understanding of the tail distribution of fire loss severity is therefore useful for the pricing and risk management of non-life insurance companies. Extreme events are occurrences which are rare, high in magnitude and lead to huge losses. Extreme event risk affects all aspects of risk assessment, modeling and management especially in the context of credit market, insurance market, and operational market. These extreme events are either naturally occurring or man-made inflicted at the time they are least expected. Some of these extreme risks are insured under general insurance policies. In fact, 10% of extreme claims paid out represent the largest share of the paid funds. This is

equivalent to significant percentage of the performances of companies.

1.1. Problem Statement

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

1.2. Objectives of This Paper

1.2.1. Main Objective

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

1.2.2. Specific Objectives

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

To estimate the measures of risk.

Historical data on insurance loss severity is often modeled using lognormal. Distributions of exponential, weibull and gamma. However, these distributions appear to overestimate or underestimate the probability of tails. In terms of fitting the tail of the loss function, the pioneering and well-known work of [Hogg and Klugman \(1984\)](#) focused on fitting the size of the loss distributions to the data. They used the truncated distribution of Pareto to fit the loss function. [Boyd \(1988\)](#) argued, however, that the tail region of the fitted loss distribution was seriously underestimated. Hogg and Klugman compared two estimation methods, namely the maximum probability estimation (MLE) and the moment method. The issue of whether Extreme Value Theory (EVT), and so, Generalized Pareto Distribution (GPD) is better for measuring loss severity was also discussed extensively in the literature. Several early studies argued that EVT could provide a number of sensible approaches to this issue. [Bassi et al. \(1998\)](#), [McNeil \(1997\)](#), [McNeil and Saladin \(1997\)](#) and [Embrechts et al. \(1997, 1999\)](#) suggested that it was preferable to use the GPD to calculate the tail loss data. [Beirlant et al. \(2004\)](#) pointed out that the data on insurance losses usually demonstrate heavy tailedness, They tested the method for a variety of simulated heavy-tailed distributions to show what types of thresholds are required and what sample sizes are needed to provide accurate quantile estimates. As a result, it is key to many risk management issues related to insurance, reinsurance and finance, as demonstrated by [Embrechts et al. \(1999\)](#).

Furthermore, many early researchers experimented with operational loss data on insurance. [Beirlant and Teugels \(1992\)](#) modeled large claims in non-life insurance using an extreme value model. [Dahen et al. \(2010\)](#) used extreme values

in business interruption insurance. Rootzen and Tajvidi (2000) used extreme value statistics to fit wind-storm losses. Moscadelli (2004) showed that the tails of loss distribution functions are, in the first approximation, of heavy-tailed Pareto type. Patrick et al. (2004) examined the empirical regularities in operational loss data and found that loss data by event type is quite similar across institutions. Nešlehová et al. (2006) used EVT and the overall quantitative risk management consequences of extremely heavy-tailed data. Chava et al. (2008) focused on modeling and predicting the loss distribution for credit-risky assets such as bonds or loans. They also analyzed the dependence between the default probabilities and recovery rates and showed that they are negatively correlated. Dahlen et al. (2010) analyzed US bank data and showed that US banks could suffer, on average, more than four major losses a year. They also used the extreme distribution to fit the operational losses and estimated annual insurance premiums. Lee and Fang (2010) focused on modeling and estimating the tail parameters of Taiwan's commercial bank operation loss severity. They also measured the capital for operational risk. In an early work on fire loss, Mandelbrot (1964) used the random walks concept and some tail distributions to model and discuss fire damage and related phenomena. Furman, Kuznetsov, & Miles (2020). Risk aggregation: A general approach via the class of Generalized Gamma Convolutions. Variance, in press considers exactly these distributions in the context of the P & C insurance to measure the loss severity of commercial fire insurance loss, we attempt to answer the following questions. Which techniques fit the loss data statistically and also result in meaningful capital estimates? How well does the method accommodate a wide variety of empirical loss data?

For the purposes of our empirical study, we measure commercial fire insurance loss using a data-driven loss distribution approach (LDA). By estimating commercial fire loss insurance risk on business-line and event-type levels, we are able to present the estimates in a more balanced fashion. The LDA framework has three essential components: a distribution of the annual number of losses, a distribution of the Egyptian pound amount of loss severity and an aggregate loss distribution that combines the two. Strictly speaking, we utilize EVT to analyze the tail behavior of commercial fire insurance loss. The results may help non-life insurance companies to manage their risk. For the purposes of comparison, we consider the following one- and two-parameter distributions to model the loss severity: exp, Weibull, gumbel, frechet, lognormal and gamma distributions. These were chosen due to their simplicity and applicability to other areas of economics and finance. Distributions such as the exponential, Weibull and gamma are unlikely to fit heavy-tailed data, but provide a nice comparison to heavier-tailed distributions such as the GPD and generalized extreme value (GEV) distribution.

We show that the GPD can be fitted to commercial fire insurance loss severity. When the loss data, the GPD is a useful method for estimating the tails of loss severity distributions. This means that the GPD is a theoretically well-supported

technique for fitting a parametric distribution to the tail of an unknown underlying distribution.

The remainder of the paper is organized as follows. Section 2 introduces EVT and goodness of fit. Section 3 gives some empirical results and analysis. Section 4 gives a few concluding remarks and ideas for future work.

2. Extreme Value Theory

We now proceed to use EVT to estimate the tail of a loss severity distribution. Extreme event risk is present in all areas of risk management. Whether we are concerned with market, credit, operational or insurance risk, one of the greatest challenges for a risk manager is to implement risk management models that allow for rare but damaging events and permit the measurement of their consequences. The oldest group of extreme value models is block maxima models. These are models for the largest observations collected from large samples of identically distributed observations. The asymptotic distribution of a series of maxima is modeled, and under certain conditions, the distribution of the standardized maximum of the series is shown to converge to the Gumbel, Frechet or Weibull distribution. The GEV distribution is a standard form of these three distributions.

The GPD was developed as a distribution for modeling tails of a wide variety of distributions. Suppose that $F_X(x)$ is the cumulative distribution function for a random variable X and that threshold u is a value of X on the right tail of the distribution, $x \geq u$ (small x is not random variable).

The probability that X lies between u and $u + y$, $y > 0$, is $F(u + y) - F(u)$. the probability of X being greater than u is $1 - F(u)$. Define $F_u(y)$ as the probability that x is between u and $u + y$, conditional on $x > u$. We have:

$$F_u(y) = \text{pr}\{x - u \leq y \mid x > u\} = \frac{F(u + y) - F(u)}{1 - F(u)} \quad (1)$$

Once the threshold is estimated, the conditional distribution F_u converges to the GPD. We can find a (This is [Pickands \(1975\)](#))

$$\lim_{y \rightarrow 0} F_u(y) \approx G_{\xi, \sigma(u)}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\sigma}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0 \\ 1 - e^{-\frac{y}{\sigma}}, & \text{if } \xi = 0 \end{cases} \quad (2)$$

where ξ is the shape parameter and determines the heaviness of the tail of the distribution, and σ is a scale parameter. When $\xi = 0$, the random variable X has a standard exponential distribution. As the tails of the distribution become heavier (or longer tailed), the value of ξ increases. The parameters can be estimated using MLE (for a more detailed description of the model, see [Neftci \(2000\)](#)).

One of the most difficult problems in the practical application of EVT is choosing the appropriate threshold for where the tail begins. The most widely

used methods for exploring the data are graphical methods, i.e., quantile-quantile (Q-Q) plots, Hill plots and the distribution of mean excess. These methods involve creating several plots of the data and using heuristics to choose the appropriate threshold.

In EVT and its applications, the Q-Q plot is typically plotted against the exponential distribution to measure the fat-tailedness of a distribution (e.g., an exponential distribution with a medium-sized tail). If the data is taken from an exponential distribution, the points on the graph would lie along a straight line. If the graph is concave, this indicates a fat-tailed distribution, whereas a convex shape is an indication of a short-tailed distribution. In addition, if the Q-Q plot deviates significantly from a straight line, then either the estimate of the shape parameter is inaccurate or the model selection is untenable.

Selecting an appropriate threshold is a critical problem with the peaks-over threshold method. There are two graphical tools used to choose the threshold: the Hill plot and mean excess plot. The Hill plot displays an estimate of ξ for different exceedance levels and is the maximum likelihood estimator for a GPD. Hill (1975) proposed the following estimator for ξ . The Hill estimator is the maximum likelihood estimator for a GPD since the extreme distribution converges to a GPD over a high threshold u .

Let $x_1 > \dots > x_n$ be the ordered statistics of independent and identically distributed random variables. We set $k < n$ and define the Hill estimator of the tail index $1/\xi$ based on upper-order statistics as:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln \left(\frac{X_{i,n}}{X_{k+1,n}} \right) \quad (3)$$

$$\xi \cong H_{k,n}^{-1} \text{ when } n \rightarrow \infty, k/n \rightarrow 0$$

The number of upper-order statistics used in the estimation is $k+1$ and n is the sample size. A Hill plot is constructed such that the estimated ξ is plotted as a function either of k upper-order statistics or of the threshold. More precisely, the Hill graph is defined by the set of points, and hopefully the graph is stable so that a value of ξ can be chosen. The Hill plot also helps us to choose the data threshold and the parameter value. The parameter should be chosen where the plot looks stable:

$$\left\{ \left(k, H_{k,n}^{-1} \right), 1 \leq k \leq n \right\} \quad (4)$$

The mean excess plot introduced by Davidson and Smith (1990) graphs the conditional mean of the data above different thresholds. The sample mean excess function (MEF) is defined as

$$en_u(u) = \frac{\sum_{i=1}^{n_u} (x_i - u)}{\sum_{i=1}^{n_u} I_u(x_i > u)} \quad (5)$$

where $I = 1$ if $\xi > u$, and 0 otherwise, and where n_u denotes the number of data points that exceed the threshold u . The MEF is the sum of the excesses over

the threshold u divided by n_u . It is an estimate of the MEF that describes the expected overshoot of a threshold once an exceedance occurs. If the empirical MEF has a positive gradient above a certain threshold u , it is an indication that the data follows the GPD with a positive shape parameter. On the other hand, exponentially distributed data would show a horizontal MEF, while short-tailed data would have a negatively sloped line.

Following Equation (2.2), the probability pound amount that $X > u + y$ conditional on $X > u$ is $1 - G_{\xi, \sigma(u)}(y)$, while the probability that $x > u$ is $1 - F(u)$, and the unconditional probability that $x > u + y$ is therefore:

$$P(X > u + y) = [1 - F(u)] [1 - G_{\xi, \sigma(u)}(y)] \quad (6)$$

If n is the total number of observations, calculated from the empirical data is $n_u = n$. The unconditional probability that $x > u + y$ is therefore:

$$\frac{n_u}{n} [1 - G_{\xi, \sigma(u)}(y)] = \frac{n_u}{n} \left(1 + \hat{\xi} \frac{y}{\sigma} \right)^{-1/\hat{\xi}} \quad (7)$$

which means that our estimator of the tail for the cumulative probability distribution is:

$$F(x) = 1 - \frac{n_u}{n} \left(1 + \hat{\xi} \frac{x - u}{\sigma} \right)^{-1/\hat{\xi}} \quad (8)$$

RISK MEASURE

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

VALUE AT RISK VaR

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

To calculate value-at-risk (VaR) with a confidence level q it is necessary to solve the equation

$$F(\text{var}) = q \quad (9)$$

For risk management q is usually taken to be greater than 0.95 and quantile in this case is referred to us Value at risk.

From Equation (8), we have:

$$q = 1 - \frac{n_u}{n} \left(1 + \hat{\xi} \frac{\text{Var} - u}{\sigma} \right)^{-1/\hat{\xi}} \quad (10)$$

The VaR is therefore

$$\text{var} = u + \frac{\sigma}{\hat{\xi}} \left(\frac{n_u}{n} (1 - q)^{-\hat{\xi}} - 1 \right) \quad (11)$$

EXPECTED SHORTFALL

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

Artzner et al. (1999) argue that VaR is not a coherent risk measure, but proved that ES is a coherent measure. Once we know the values of the parameter of the generalized Pareto distribution. We can use them to calculate the value at risk and expected shortfall.

Expected shortfall (ES) is a concept used in finance and, more specifically, in the field of financial risk measurement to evaluate the market risk of a portfolio. It is an alternative to VaR. The expected shortfall at the $q\%$ level is the expected return on the portfolio in the worst $q\%$ of the cases. For example, ES (0.05) is the expectation of the worst 5 out of 100 events. Expected shortfall is also called conditional value-at-risk and expected tail loss.

In our case, we define the excess shortfall as the expected loss size, given that VaR is exceeded

$$ES_q = E(L | L > \text{var}_q) \quad (12)$$

where $q(=1-p)$ is the confidence level. Furthermore, we obtain the following ES estimator

$$ES_q = \frac{\text{var}_q}{1-\xi} + \frac{\sigma - \xi u}{1-\xi} \quad (13)$$

One can attempt to fit any particular parametric distribution to data; however, only certain distributions will have a good fit. There are two ways of assessing this goodness of fit: either by using graphical methods or by using formal statistical goodness-of-fit tests. The former method (a Q-Q plot or a normalized probability-probability (P-P) plot, for example) helps an individual to determine whether a fit is very poor, but may not reveal whether a fit is good in the formal sense of statistical fit. Examples of the latter method are the Kolmogorov-Smirnov (KS) test or the likelihood ratio (LR) test. The Q-Q plot depicts the match or mismatch between the observed values in the data and the estimated value given by the hypothesized fitted distribution. The KS test is a nonparametric supremum test based on the empirical cumulative distribution.

3. Application

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

There are 939 observations in the data set. All commercial fire insurance loss data sets used in this study were obtained from a non-life insurance company in Egypt.

3.1. Scope of the Data

Secondary data from E.G. insurance company regarding fire industrial claims for

the period 2000-2011 was used in this study.

3.2. Actuarial Modeling Process

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

- 1) Selecting the model family of distributions.
- 2) Exploratory data analysis.
- 3) Estimating the parameters.
- 4) Goodness of fit test.

3.3. Selecting the Model Family

Here considerations were made of a number of parametric probability distributions as potential candidates for the data generating mechanism for extreme claims. Most data in general insurance is skewed to the right and therefore most distributions that exhibit these characteristics can be used to model the extreme claims. However, the list of potential probability distributions is enormous and it is worth noting that the choice of distributions is to some extent subjective.

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

- Prior knowledge and experience in curve fitting.
- Time constraint.
- Availability of computer software to facilitate the study.
- The volume and quality of data.

Therefore seven distributions were used including: genpareto, exp, Weibull, gumbel, lognormal gamma and frechet.

3.4. Exploratory Data Analysis

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

3.5. Computation and Interpretations

3.5.1. Specific Objectives

Testing for the appropriate statistical distribution for the claim amount test the goodness of fit of the chosen distribution.

3.5.2. Variable

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

Descriptive statistics may help to choose candidates to describe a distribution among a set of parametric distributions.

3.5.3. Descriptive Data Analysis (e.g. Pound)

$$\text{Min} = 0.7951386 \times 10^4$$

$$\text{Min} = 0.7951386 \times 10^4$$

$$\text{Max} = 47.9656 \times 10^4$$

$$\text{Mean} = 5.627495 \times 10^4$$

$$\text{Median} = 4.202978 \times 10^4$$

$$\text{Skewness} = 3.055893$$

$$\text{Kurtosis} = 17.95521$$

$$\text{Number of observations} = 939$$

The data according to descriptive statistics shown above indicates that the data is skewed to the right (skewedness coefficient of 3.055893) Right-skewedness means that the right tail is long relative to the left tail.

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

3.6. Fit of Distributions by Maximum Likelihood Estimation.

Once selected, one or more parametric distributions $f(\cdot | \theta)$ (with parameter $\theta \in \mathbb{R}^d$) d is a natural number may be fitted to the data set, one at a time Under the i.i.d. sample assumption, distribution parameters θ are by default estimated by maximizing the likelihood function defined as:

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

with x_i the n observations of variable X and $f(\cdot | \theta)$ the density function of the parametric distribution.

Estimated parameters of distributions by maximum likelihood estimation (MLE)

Table 2 shows that the estimation parameters seven distribution by performing maximum likelihood estimate method.

Model checking

I will provide four classical goodness-of-fit plots for each distribution to check the model

- a density plot representing the density function of the fitted distribution along with the histogram of the empirical distribution;
- a CDF plot of both the empirical distribution and the fitted distribution;
- a Q-Q plot representing the empirical quantiles (y-axis) against the theoretical quantiles (x-axis);
- a P-P plot representing the empirical distribution function evaluated at each data point (y-axis) against the fitted distribution function (x-axis).

The CDF plot may be considered as the basic classical goodness-of-fit plots. The two other plots are complementary and can be very informative in some cases. The Q-Q plot emphasizes the lack-of-fit at the distribution tails while the P-P plot emphasizes the lack-of-fit at the distribution center.

Table 1. Show properties of distributions.

Distributions	parameters	PDF	CDF
genpareto	$\mu \in (-\infty, \infty)$ location $\sigma \in (0, \infty)$ scale $\xi \in (-\infty, \infty)$ shape	$\frac{1}{\sigma} (1 + \xi z)^{-(1/\xi+1)}$ Where $z = \frac{x-\mu}{\sigma}$	$1 - (1 + \xi z)^{-\frac{1}{\xi}}$
Exponential	$\lambda > 0$, rate, or inverse scale	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$
Weibull	$\lambda \in (0, +\infty)$ scale $k \in (0, +\infty)$ shape	$f(x) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\begin{cases} 1 - e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$
Gumbel	μ location $\beta > 0$ scale	$\frac{1}{\beta} e^{-\left(z + e^{-z}\right)}$ Where $z = \frac{x-\mu}{\beta}$	$e^{-e^{-(x-\mu)/\beta}}$
Log-normal	$\mu \in (-\infty, +\infty)$ $\sigma > 0$	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$	$\frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\ln x - \mu}{\sqrt{2}\sigma}\right)$
Gamma	$k > 0$ scale $\theta > 0$ shape	$\frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$	$\frac{1}{\Gamma(k)} \gamma\left(k, \frac{x}{\theta}\right)$
Fréchet	$s \in (0, \infty)$ scale $m \in (0, \infty)$ location	$\frac{\alpha}{s} \left(\frac{x-m}{s}\right)^{-1-\alpha} e^{-\left(\frac{x-m}{s}\right)^{-\alpha}}$	$e^{-\left(\frac{x-m}{s}\right)^{-\alpha}}$

Table 2. Result of estimated parameters of distributions by maximum likelihood estimation.

Distributions	the parameter estimates	
1-Gen pareto	shape 1	2.9753648
	shape 2	13.4914804
	scale	0.8390074
2-exponential	rate	0.177699
3-Weibull	shape	1.372363
	scale	6.225291
4-gumbel	location	3.593531
	scale	2.437994
5-lognormal	Location	1.4759611
	scale	0.6850889
6-gamma	Shape	2.1385816
	Rate	0.3800284
7-frechet	Shape	7.413944
	Scale	1.300028

Table 1: Show the properties of seven distributions used in this study including: genpareto, exp, Weibull, gumbel, log-normal gamma and frechet.

Figure 1: The gen Pareto plots displayed succeeds in measure the tails of the distribution. Based on the good linear relationship in the plots, the gen Pareto

distribution is preferred over the gamma distribution.

Figure 2: Plot shows that the exponential distribution is not as heavy in the right tail as the data. The plots display a poor fit and suggest that data is not exponential distributed.

Figure 3: The weibull plots indicate a slightly better fit than the gumbel distribution as the points are even closer to a 45-degree line.

Figure 5: The gumbel plots indicate that the gumbel distribution fails to measure the extreme right tail of the data. Comparing the gumbel to the gamma plots, the gamma distribution seems to be a better choice of claim size distribution. However, the gumbel plots display a better fit than the plots for the frechet and log-normal distributions as shown in **Figure 4** and **Figure 6**.

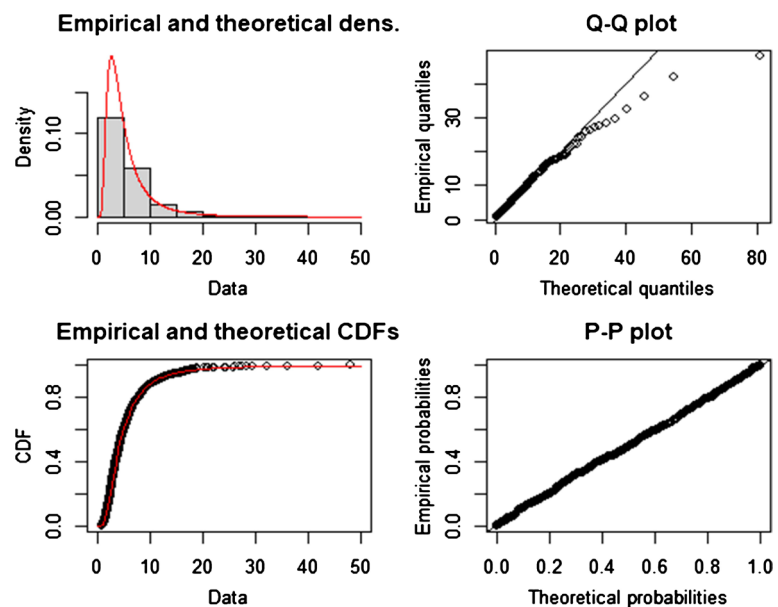


Figure 1. Four goodness of fit for gen pareto distribution.

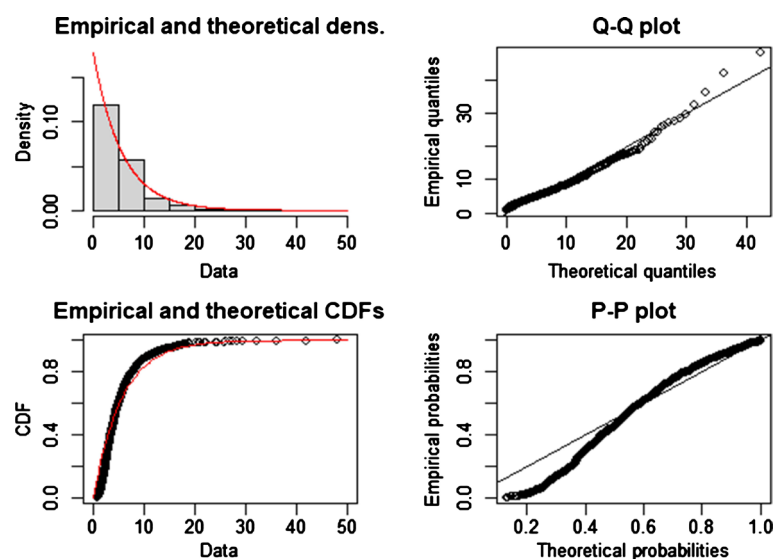


Figure 2. Four goodness of fit for exponential distribution.

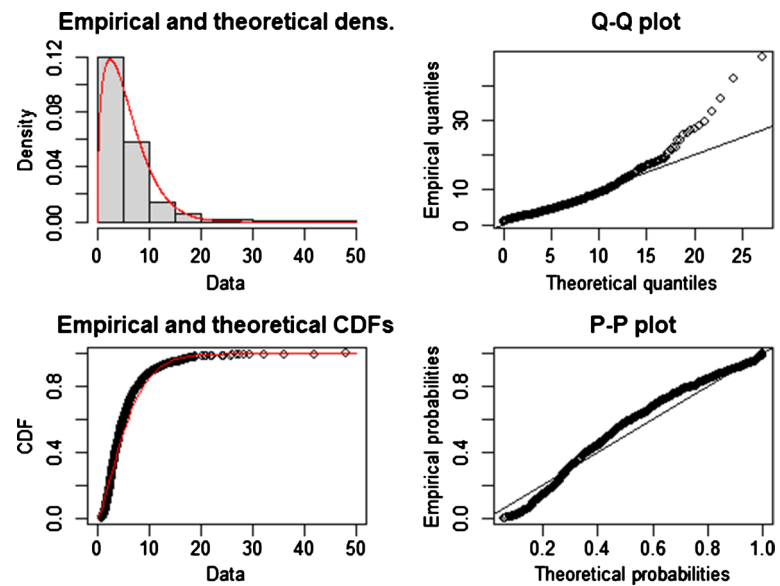


Figure 3. Four goodness of fit for Weibull distribution.

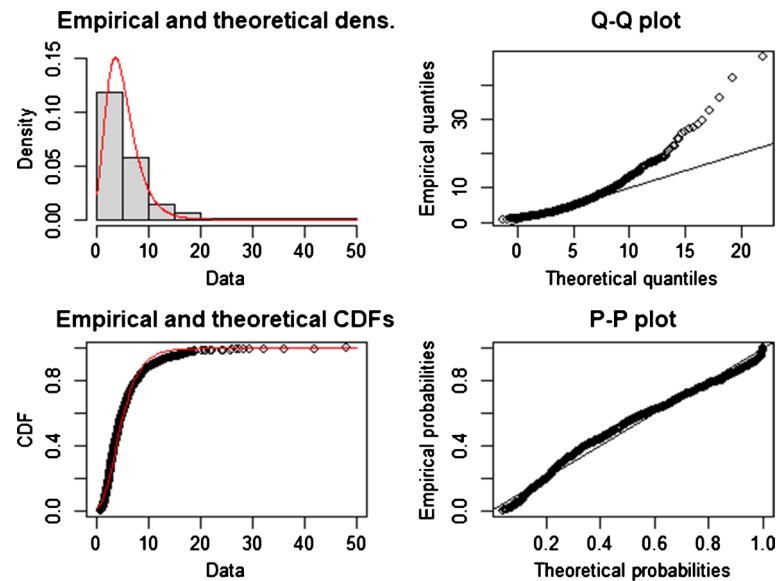


Figure 4. Four goodness of fit for gumbel distribution.

The relationship between the theoretical and the sample quantiles in the gamma plot implies a good fit and suggests that the gamma distribution is a good choice of claim size distribution as is shown in **Figure 7**.

I will report CDF values in a logscale so as to emphasize discrepancies on the tail as shown in **Figure 8**, the left-tail seems to be better described by the genpareto distribution.

The computation of different goodness-of-fit statistics in order to further compare fitted distributions. The purpose of goodness-of-fit statistics aims to measure the distance between the fitted parametric distribution and the empirical distribution: e.g., the distance between the fitted cumulative distribution function F and the empirical distribution function F_n . When fitting continuous

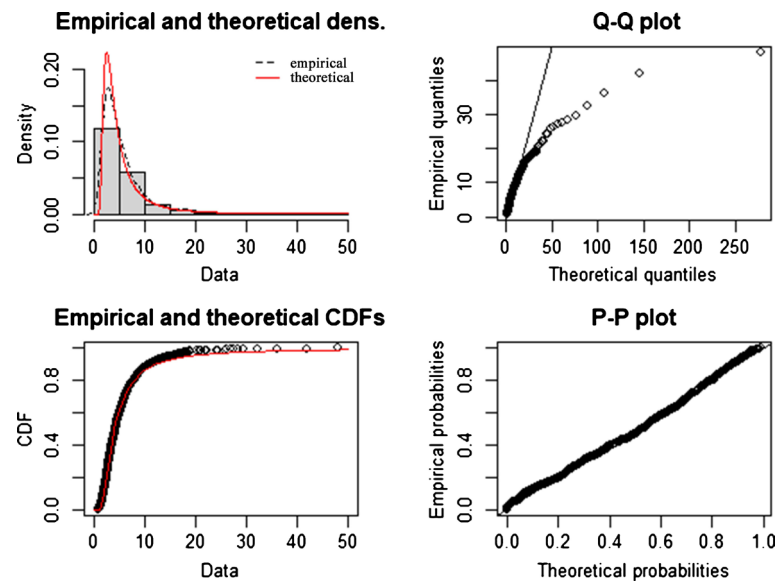


Figure 5. Four goodness of fit for frechet distribution.

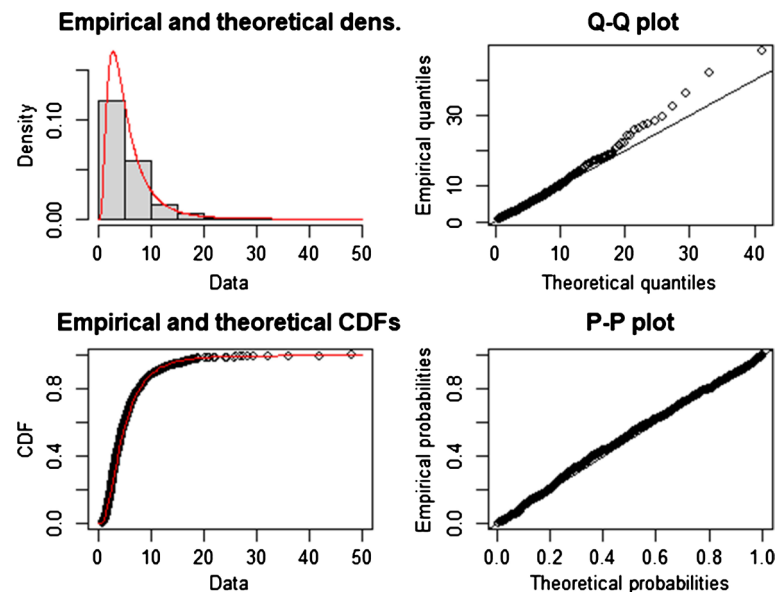


Figure 6. Four goodness of fit for lognormal distribution.

distributions, three goodness-of-fit statistics are classically considered: Cramer-von Mises, Kolmogorov-Smirnov and Anderson-Darling statistics.

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. And as shown in Table 3, genpareto has the lowest AIC and BIC, so genpareto is the best distribution for claim size.

As giving more weight to distribution tails, the Anderson-Darling statistic is of special interest when it matters to equally emphasize the tails as well as the main body of a distribution. This is often the case in risk assessment (Cullen and Frey, 1999; Vose, 2010). For this reason, these statistics are often used to select the best distribution among those fitted. Nevertheless, these statistics should be used

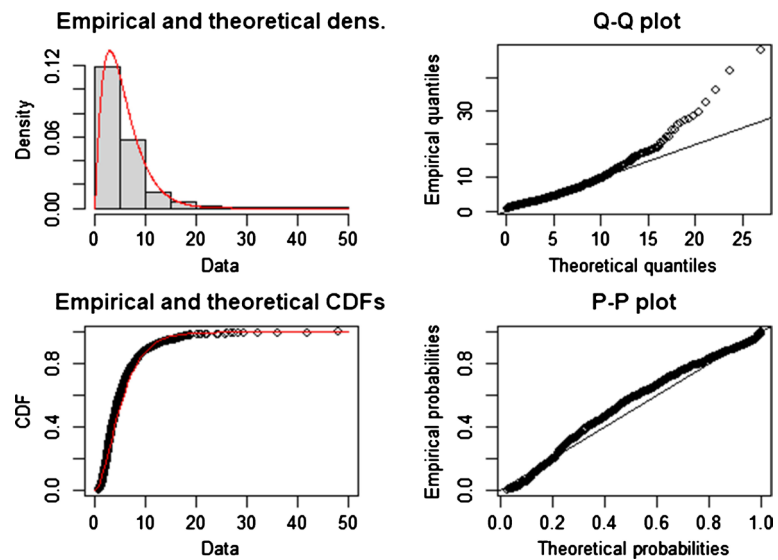


Figure 7. Four goodness of fit for gamma distribution.

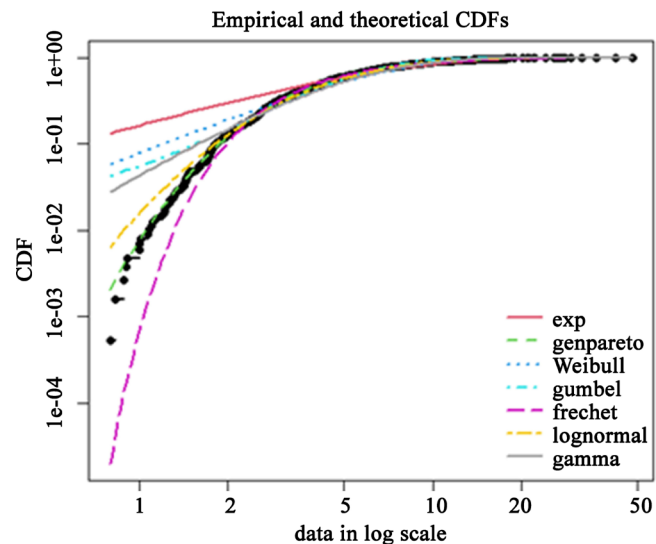


Figure 8. CDF plot to compare the fit of seven distributions data set, with CDF values in a logscale to emphasize discrepancies on the left tail.

Table 3. Goodness of fit for distributions.

Distributions	Akaike information criteria (AIC)	Bayesian information criteria (BIC)	Kolmogorov- Smirnov statistic	Cramer-von Mises statistic	Anderson-Darling statistic
1-Gen pareto	4714.542	4729.076	0.02229017	0.06304588	0.39154346
2-exponential	5124.554	5129.399	0.1966317	8.7821182	53.3074780
3-Weibull	4967.909	4977.599	0.09131398	2.90888129	19.45479034
4-gumbel	4982.206	4991.895	0.05571758	1.12757030	14.67465690
5-frechet	4730.349	4788.232	0.03063316	0.27309075	3.25132752
6-lognormal	4730.349	4740.039	0.03805363	0.27265821	1.76719310
7-gamma	4871.427	4881.117	0.08125536	2.07998426	12.32869779

cautiously when comparing fits of various distributions. Keeping in mind that the weighting of each CDF quadratic difference depends on the parametric distribution in its definition, Anderson-Darling statistics computed for several distributions fitted on a same data set are theoretically difficult to compare. Moreover, such a statistic, as Cramer-von Mises and Kolmogorov-Smirnov ones, does not take into account the complexity of the model (i.e., parameter number). It is not a problem when compared distributions are characterized by the same number of parameters, but it could systematically promote the selection of the more complex distributions in the other case. Looking at classical penalized criteria based on the loglikelihood (AIC, BIC) seems thus also interesting, especially to discourage overfitting. In the previous goodness of fit tests, all the goodness-of-fit statistics based on the CDF distance are in favor of the genpareto distribution, also AIC and BIC values respectively give the preference to the genpareto distribution.

Hypothesis Testing

The null and the alternative hypotheses are:

- H_0 : the data follow the genpareto distribution.
- H_A : the data do not follow the genpareto distribution.

Table 4 shows that p -value > 0.05 so we will accept the null hypothesis (the data follow the genpareto distribution).

3.7. Bootstrap Confidence Estimates

The paper goes ahead to find the confidence intervals of the genPareto distribution estimates. Bootstrapping is the practice of estimating properties of estimators (such as the variance) by measuring those properties when sampling from an approximating distribution. This can be implemented by constructing hypothesis test. The bootstrap method involves taking the original set of N heights and using a computer sampling from it to form a new sample called a resample or bootstrap sample that is also of size N . The bootstrap sample is taken from the original using sampling with replacement so it is not identical with original. This process is repeated a large number of times typically 1000 times. Then for each of the bootstrap sample we compute the mean and standard deviation. Using the estimated parameters we fit the parameters. We therefore obtain the histogram of genPareto distribution parameters which gives us the confidence intervals of those parameters. The uncertainty in the parameters of the fitted distribution can be estimated by parametric or nonparametric bootstraps. The bootstrapped

Table 4. Goodness of fit for the GPD model.

Test

distribution	Kolmogorov-Smirnov statistic		Cramer-von Mises statistic		Anderson-Darling statistic	
	statistic	p -value	statistic	p -value	statistic	p -value
Genpareto	0.018648	0.8997	0.038208	0.9422	0.29662	0.9407

values of parameters which can be plotted to visualize the bootstrap region. The medians and the 95 percent confidence intervals of parameters (2.5 and 97.5 percentiles). When inferior to the whole number of iterations (due to lack of convergence of the optimization algorithm for some bootstrapped data sets), the number of iterations for which the estimation converges. The plot of “bootdist” consists in a scatterplot or a matrix of scatterplots of the bootstrapped values of parameters providing a representation of the joint uncertainty distribution of the fitted parameters. Below is the bootdist with the previous fit of the genpareto distribution to data set (see **Figure 9**).

Parametric bootstrap medians and 95% percentile CI

	Median	2.5%	97.5%
shape 1	2.7924641	2.4095885	3.468161
shape 2	14.6077155	7.4036814	36.273665
scale	0.7200331	0.2486015	1.856251

The estimation method converged only for 734 among 1001 iterations.

Bootstrap samples of parameter estimates are useful especially to calculate confidence intervals on each parameter of the fitted distribution from the marginal distribution of the bootstrapped values. It is also interesting to look at the joint distribution of the bootstrapped values in a scatterplot (or a matrix of scatterplots if the number of parameters exceeds two) in order to understand the potential structural correlation between parameters as shown in (**Figure 9**). The

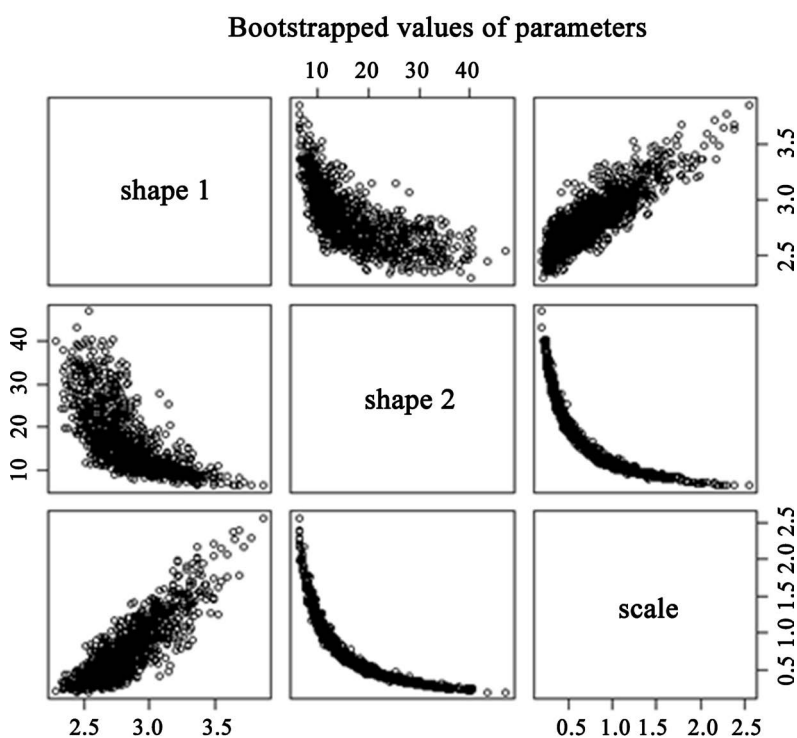


Figure 9. Bootstrapped values of parameters for a fit of the genpareto distribution characterized by three parameters.

bootstrap method can also be used to calculate confidence intervals on quantiles of the fitted distribution to the data set.

(original) estimated quantiles for each specified probability (non-censored data)

$$p = 0.05$$

estimate 1.526787

Median of bootstrap estimates

$$p = 0.05$$

estimate 1.526108

two-sided 95% CI of each quantile

$$p = 0.05$$

2.5% 1.422547

97.5% 1.630237

The estimation method converged only for 734 among 1001 bootstrap iterations.

3.8. Value at Risk and Expected Shortfall

We have calculated for the Value at risk and expected shortfall of 939 claims reported in an insurance company.

Table 5 shows that an increase in the quantile results to a decrease in the value at risk. Meaning that huge number of the claims if occurred will be borne by insurer himself. It also leads to a decrease in the expected shortfall. This is the probable loss that would result in case a risk happens at a certain quantile.

Figures in parentheses are standard deviation. VaR (90%), VaR (95%) and VaR (99%) denotes the value-at-risk at the 95%, 95% and 99% confidence levels, respectively. ES (95%) denotes the expected shortfall at the 95% level, and so on.

4. Summary and Conclusion

4.1. Summary

There are 939 observations in the data set. All commercial fire insurance loss data sets used in this study were obtained from a non-life insurance company in Egypt. The data is from 2000-2011 worth of fire losses. The major objective is to come up with one statistical distribution that fits the extreme claims data well and to test how well this statistical distribution fits those extreme claims so that this distribution can be used for modeling the extreme claims.

Table 5. Value at risk and expected shortfall of data.

Quantile	VaR	ES
0.90	0.06314551	2.122615
0.95	0.03126213	2.013327
0.99	0.008028979	1.933929

The data that was analyzed came from an Insurance company which were fire claims from 2000 to 2011. The data had 939 claims. I carried out the descriptive analysis of the data where the mean was found to be 5.627495, Skewness 3.055893 and kurtosis 17.95521. From the descriptive statistics it shows the data is heavy tailed hence extreme value theory was applicable.

The parameter estimates of the seven distributions were compared and genPareto distribution came up to be the best fitting distribution. The Q-Q plots indicate that most points of the genPareto distribution are lying along the reference line thus making it the best distribution family in preliminary stage. A histogram of claims and goodness of fit with loglikelihood value, AIC and BIC also pointed that genPareto distribution was the best fitting distribution among the three distributions, followed by frechet. The exponential function has the lowest value.

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

It is practically impossible to experiment with every possible parametric distribution that we know of. An alternative way of conducting such an exhaustive search could be to fit general class distributions to the loss data in the hope that the distributions will be flexible enough to conform to the underlying data in a reasonable way. For the purposes of comparison, we have used genpareto, exp, Weibull, gumbel, frechet, lognormal, and gamma distribution as a benchmark. preceding page shows the poor fit of the exponential, Weibull, and shows that other distributions fit the loss data much better, especially the genpareto distribution. The goodness-of-fit loglikelihood value, AIC and BIC show that the genpareto model is highest, followed by frechet. The exponential function has the lowest value.

4.2. Conclusion

In many applications of loss data distributions, a key concern is fitting the loss data in the tail. As mentioned above, good estimates of the tails of fire loss severity distributions are essential for pricing and risk management of commercial fire insurance loss. We first execute an exploratory loss data analysis using four classical goodness-of-fit plots densiplot, CDF, Q-Q plot and P-P plot of genpareto, exp, Weibull, gumbel, frechet, lognormal, and gamma distributions. The goodness of fit, loglikelihood function value, AIC and BIC revealed the genPareto distribution came up to be the best fitting distribution.

Last but not least, we showed that the genpareto can be fitted to commercial fire insurance loss severity. When the loss data exceeds high thresholds, the genpareto is a useful method for estimating the tails of loss severity distributions. It also means that the genpareto is a theoretically well-supported technique for fitting a parametric distribution to the tail of an unknown underlying distribu-

tion.

4.3. Recommendation

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

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9, 203-228. <https://doi.org/10.1111/1467-9965.00068>
- Bassi, F., Embrechts, P., & Kafetzaki, M. (1998). Risk Management and Quantile Estimation. In R. J. Adler, F. Feldman, & M. Taqqu (Eds.), *A Practical Guide to Heavy Tails* (pp. 111-130). Basel: Birkhäuser.
- Beirlant, J., & Teugels, J. L. (1992). Modeling Large Claims in Non-Life Insurance. *Insurance: Mathematics and Economics*, 11, 17-29. [https://doi.org/10.1016/0167-6687\(92\)90085-P](https://doi.org/10.1016/0167-6687(92)90085-P)
- Beirlant, J., Joossens, E., & Segers, J. (2004). Generalized Pareto Fit to the Society of Actuaries' Large Claims Database. *North American Actuarial Journal*, 8, 108-111. <https://doi.org/10.1080/10920277.2004.10596140>
- Boyd, V. (1988). Fitting the Truncated Pareto Distribution to Loss Distributions. *Journal of the Staple Inn Actuarial Society*, 31, 151-158. <https://doi.org/10.1017/S2049929900010291>
- Chava, S., Stefanescu, C., & Turnbull, S. (2008). *Modeling the Loss Distribution*. Working Paper.
- Cullen, A., & Frey, H. (1999). *Probabilistic Techniques in Exposure Assessment* (1st ed.). New York: Plenum Publishing Co.
- Dahen, H., Dionne, G., & Zajdenweber, D. (2010). A Practical Application of Extreme Value Theory to Operational Risk in Banks. *The Journal of Operational Risk*, 5, 1-16. <https://doi.org/10.21314/JOP.2010.074>
- Davidson, A. C., & Smith, R. L. (1990). Models for Exceedances over High Thresholds. *Journal of the Royal Statistical Society: Series B*, 52, 393-442. <https://doi.org/10.1111/j.2517-6161.1990.tb01796.x>
- Embrechts, P., Kluppelberg, C., & Mikosch, T. (1997). *Modeling Extreme Events for Insurance and Finance*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-33483-2>
- Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal*, 3, 30-41. <https://doi.org/10.1080/10920277.1999.10595797>
- Furman, E., Kuznetsov, A., & Miles, J. (2020). *Risk Aggregation: A General Approach via the Class of Generalized Gamma Convolutions*.

- Hill, B. M. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *Annals of Statistics*, 46, 1163-1173. <https://doi.org/10.1214/aos/1176343247>
- Hogg, R., & Klugman, S. (1984). *Loss Distributions*. Hoboken, NJ: John Wiley & Sons. <https://doi.org/10.1002/9780470316634>
- Lee, W. C., & Fang, C. J. (2010). The Measurement of Capital for Operational Risk of Taiwanese Commercial Banks. *The Journal of Operational Risk*, 5, 79-102. <https://doi.org/10.21314/JOP.2010.076>
- Mandelbrot, B. (1964). Random Walks, Fire Damage and Related Phenomena. *Operations Research*, 12, 582-585. <https://doi.org/10.1287/opre.12.4.582>
- McNeil, A. J. (1997). Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, 27, 117-137. <https://doi.org/10.2143/AST.27.1.563210>
- McNeil, A. J., & Saladin, T. (1997). *The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions*. Zurich: Department Mathematik, ETH Zentrum.
- Moscadelli, M. (2004). *The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee*. Working Paper No. 517, Bank of Italy. <https://doi.org/10.2139/ssrn.557214>
- Neftci, S. N. (2000). Value at Risk Calculations, Extreme Events and Tail Estimation. *Journal of Derivatives*, 7, 23-38. <https://doi.org/10.3905/jod.2000.319126>
- Nešlehová, J., Embrechts, P., & Chavez-Demoulin, V. (2006). Infinite-Mean Models and the LDA for Operational Risk. *The Journal of Operational Risk*, 1, 3-25. <https://doi.org/10.21314/JOP.2006.001>
- Patrick, D. F., Jordan, J. S., & Rosengren, E. S. (2004). *Implications of Alternative Operational Risk Modeling Techniques*. Working Paper W11103, National Bureau of Economic Research.
- Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics. *Annals of Statistics*, 3, 119-131. <https://doi.org/10.1214/aos/1176343003>
- Rootzen, H., & Tajvidi, N. (2000). Extreme Value Statistics and Wind Storm Losses: A Case Study. In P. Embrechts (Ed.), *Extremes and Integrated Risk Management* (70-94), London: Risk Books.
- Vose, D. (2010). *Quantitative Risk Analysis. A Guide to Monte Carlo Simulation Modeling*. Chichester: John Wiley & Sons.