

How to deal with Extreme Observations in Empirical Finance: An  
Application to Capital Markets

Josué de Sousa e Silva

A Dissertation presented in partial fulfilment of the Requirements for the Degree of -Master  
in Finance

Supervisor:

Prof. Doutor José Dias Curto, Associate Professor, ISCTE Business School, Quantitative  
Methods Department

April 2011

## Abstract

In the last few years, Extreme Value Theory (EVT) has gained increased importance in modeling extreme observations in all social sciences. This is especially true in finance, since EVT is a tool used to consider probabilities associated with extreme and rare events with catastrophic consequences, as happened in the Sub-prime crisis in 2007.

To model extreme observations, we use two different statistical distribution families in this thesis: Generalized Extreme Value (GEV) and Generalized Pareto Distribution (GPD).

In this thesis, EVT methods were used to investigate and fit the empirical distribution of the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices to the theoretical GEV and GPD distributions. We have applied two approaches of extreme value theory, the *Block Maxima* and the *Peaks Over Threshold (POT)* approach, as well as the parametric approach of the Maximum Likelihood Estimate Method (MLE) for the distribution parameter estimation and the non-parametric approach of the Hill estimator.

As a result of the application, we have seen that in the GEV distribution application, our data was well represented by the Fréchet and Weibull distributions. On the other hand, in the GPD distribution, using the parametric approach MLE, our data was mostly well represented by the Exponential and Beta distributions. However, applying the GPD using the non-parametric approach of the Hill estimator for the tail index, we have seen that the monthly maximum returns of our indices are well represented by the Pareto distribution.

JEL Classification: C13, C14.

Key Words: Extreme Value Theory, Generalized Extreme Value (GEV), Generalized Pareto Distribution (GPD), Stock Market Returns.

## **Resumo**

Nos últimos anos, a Teoria de Valores Extremos (TVE) tem ganho uma importância crescente no estudo de observações extremas em todas as ciências. Isto é especialmente verdade em finanças, uma vez que a TVE é uma ferramenta utilizada para analisar as probabilidades associadas a eventos extremos e raros com consequências catastróficas, como a crise do Sub-Prime em 2007.

Para modelar observações extremas, usamos duas famílias de distribuição estatísticas: Distribuição Generalizada de Valores Extremos (GEV) e a Distribuição Generalizada de Pareto (GPD).

Nesta tese, a TVE foi utilizada para investigar e ajustar a distribuição empírica dos retornos máximos e mínimos mensais dos índices bolsistas FTSE 100, NIKKEI 225 e do S&P500 às distribuições teóricas da GEV e GPD.

Aplicamos duas abordagens na aplicação da TVE, o método do Block Maxima e o método dos excessos de nível (POT), onde para a estimação dos parâmetros da distribuição recorremos ao método paramétrico da Máxima Verosimilhança, bem como ao método não-paramétrico através do estimador Hill.

Como resultado do estudo empírico na aplicação da GEV, verificamos que as séries são bem representadas pela distribuição de Fréchet e Weibull. Por outro lado, na aplicação da GPD, utilizando a abordagem paramétrica para o cálculo dos parâmetros da distribuição, as séries são bem representadas pelas distribuições exponencial e Beta. No entanto, a aplicação do GPD utilizando a abordagem não-paramétrica, verificou-se que a série dos retornos máximos mensais dos índices são bem representados pela distribuição de Pareto.

Classificação JEL: C13, C14.

Palavras-chave: Teoria de Valores Extremos, Distribuição Generalizada de Valores Extremos, Distribuição Generalizada de Pareto, Retornos de Mercados de Capitais.

## Contents

1. Introduction.....	1
2. Literature Review.....	2
3. Theoretical Background.....	7
3.1 The Block Maxima Approach .....	8
3.2 Peaks-Over-Threshold Approach (POT) .....	12
3.3 Peaks-Over-Random Threshold Approach (PORT).....	15
3.4 Parameter Estimation Methods.....	15
3.4.1. The parametric approach .....	15
3.4.2. The non-parametric approach.....	17
4. Empirical Application.....	19
4.1 Methodology.....	19
4.2 Empirical Results.....	20
4.3 Fitting the GEV Distribution .....	27
4.4 Fitting the GPD.....	35
4.5 Non-parametric approach – Hill’s estimator .....	44
5. Conclusions.....	50
6. Final Comments .....	52
References .....	53
Appendix .....	59

**List of Tables**

**Table 1:** FTSE 100 Explanatory data analysis..... 20

**Table 2:** NIKKEI 225 Explanatory data analysis ..... 20

**Table 3:** S&P 500 Explanatory data analysis ..... 21

**Table 4:** LR Test and Log-Likelihood Results - GEV..... 27

**Table 5:** FTSE 100 Maximum/Minimum Parameters - GEV..... 28

**Table 6:** NIKKEI 225 Maximum/Minimum Parameters - GEV ..... 30

**Table 7:** S&P 500 Maximum/Minimum Parameters - GEV ..... 32

**Table 8:** FTSE 100 LR Test and Log-Likelihood Results - GPD ..... 35

**Table 9:** Maximum FTSE 100 Parameters - GPD..... 36

**Table 10:** Minimum FTSE 100 Parameters - GPD..... 37

**Table 11:** NIKKEI 225 LR Test and Log-Likelihood Results - GPD..... 38

**Table 12:** Maximum NIKKEI 225 Parameters - GPD ..... 39

**Table 13:** Minimum NIKKEI 225 Parameters - GPD..... 40

**Table 14:** S&P 500 LR Test and Log-Likelihood Results - GPD ..... 41

**Table 15:** Maximum S&P 500 Parameters - GPD..... 41

**Table 16:** Minimum S&P 500 Parameters - GPD ..... 43

**Table 17:** FTSE 100 LR Test and Log-Likelihood Results – GPD (Hill Estimator) ..... 46

**Table 18:** Maximum FTSE 100 Parameters - GPD (Hill Estimator) ..... 46

**Table 19:** NIKKEI 225 LR Test and Log-Likelihood Results – GPD (Hill Estimator)..... 47

**Table 20:** Maximum NIKKEI 225 Parameters - GPD (Hill Estimator)..... 47

**Table 21:** S&P 500 LR Test and Log-Likelihood Results – GPD (Hill Estimator) ..... 48

**Table 22:** Maximum S&P 500 Parameters - GPD (Hill Estimator)..... 48

## List of Figures

<b>Figure 1:</b> Asset returns distribution versus normal distribution	7
<b>Figure 2:</b> Block-maxima graph .....	9
<b>Figure 3:</b> Densities for the Fréchet, Weibull and Gumbel functions .....	11
<b>Figure 4:</b> Excesses over a threshold $u$ graph.....	12
<b>Figure 5:</b> Distribution Function and Distribution Over Threshold .....	12
<b>Figure 6:</b> Densities for the Beta, Pareto and Exponential functions .....	14
<b>Figure 7:</b> FTSE 100 Daily Returns Histogram.....	23
<b>Figure 8:</b> FTSE 100 Max Returns Histogram .....	23
<b>Figure 9:</b> FTSE 100 Min Returns Histogram .....	23
<b>Figure 10:</b> NIKKEI 225 Daily Returns Histogram .....	23
<b>Figure 11:</b> NIKKEI 225 Max Returns Histogram .....	23
<b>Figure 12:</b> NIKKEI 225 Min Returns Histogram.....	23
<b>Figure 13:</b> S&P 500 Daily Returns Histogram.....	24
<b>Figure 14:</b> S&P 500 Max Returns Histogram .....	24
<b>Figure 15:</b> S&P 500 Min Returns Histogram.....	24
<b>Figure 16:</b> FTSE 100 Adjust Close .....	24
<b>Figure 17:</b> FTSE 100 Daily Return .....	24
<b>Figure 18:</b> FTSE 100 Monthly Maximum Return.....	24
<b>Figure 19:</b> FTSE 100 Monthly Minimum Return .....	25
<b>Figure 20:</b> NIKKEI 225 Adjust Close.....	25
<b>Figure 21:</b> NIKKEI 225 Daily Return.....	25
<b>Figure 22:</b> NIKKEI 225 Monthly Maximum Return .....	25

<b>Figure 23:</b> NIKKEI 225 Monthly Minimum Return.....	25
<b>Figure 24:</b> S&P 500 Adjust Close .....	25
<b>Figure 25:</b> S&P 500 Daily Return .....	26
<b>Figure 26:</b> S&P 500 Monthly Maximum Return .....	26
<b>Figure 27:</b> S&P 500 Monthly Minimum Return .....	26
<b>Figure 28:</b> FTSE100 Max Return Fréchet Probability Plot.....	29
<b>Figure 29:</b> FTSE100 Max Return Fréchet Quantile Plot.....	29
<b>Figure 30:</b> FTSE100 Max Return Density Plot.....	29
<b>Figure 31:</b> FTSE100 Max Return Density Plot II .....	29
<b>Figure 32:</b> FTSE100 Min Return Weibull Probability Plot .....	29
<b>Figure 33:</b> FTSE100 Min Return Weibull Quantile Plot .....	29
<b>Figure 34:</b> FTSE100 Min Return Density Plot .....	29
<b>Figure 35:</b> FTSE100 Min Return Density Plot II.....	29
<b>Figure 36:</b> NIKKEI 225 Max Return Fréchet Probability Plot .....	31
<b>Figure 37:</b> NIKKEI 225 Max Return Fréchet Quantile Plot .....	31
<b>Figure 38:</b> NIKKEI 225 Max Return Density Plot .....	31
<b>Figure 39:</b> NIKKEI 225 Max Return Density Plot II.....	31
<b>Figure 40:</b> NIKKEI 225 Min Return Weibull Probability Plot.....	31
<b>Figure 41:</b> NIKKEI 225 Min Return Weibull Quantile Plot.....	31
<b>Figure 42:</b> NIKKEI 225 Min Return Density Plot .....	31
<b>Figure 43:</b> NIKKEI 225 Min Return Density Plot .....	31
<b>Figure 44:</b> S&P 500 Max Return Fréchet Probability Plot .....	34
<b>Figure 45:</b> S&P 500 Max Return Fréchet Quantile Plot .....	34

<b>Figure 46:</b> S&P 500 Max Return Density Plot.....	34
<b>Figure 47:</b> S&P 500 Max Return Density Plot.....	34
<b>Figure 48:</b> S&P 500 Min Return Probability Plot.....	34
<b>Figure 49:</b> S&P 500 Min Return Quantile Plot.....	34
<b>Figure 50:</b> S&P 500 Min Return Density Plot .....	34
<b>Figure 51:</b> S&P 500 Min Return Density Plot .....	34
<b>Figure 52:</b> FTSE 100 Max Return Density Plot for a threshold of 1 % - Exponential Distribution .....	37
<b>Figure 53:</b> FTSE 100 Max Return Density Plot for a threshold of 2.5% - Pareto Distribution .....	37
<b>Figure 54:</b> NIKKEI 25 Max Return Density Plot for a threshold of 1% - Exponential Distribution .....	39
<b>Figure 55:</b> NIKKEI 225 Max Return Density Plot for a threshold of 2.5% - Exponential Distribution .....	40
<b>Figure 56:</b> NIKKEI 225 Max Return Density Plot for a threshold of 5% - Exponential Distribution .....	40
<b>Figure 57:</b> S&P 500 Max Return Density Plot for a threshold of 1% - Exponential Distribution .....	42
<b>Figure 58:</b> S&P 500 Max Return Density Plot for a threshold of 2.5% - Exponential Distribution .....	42
<b>Figure 59:</b> S&P 500 Max Return Density Plot for a threshold of 5% - Exponential Distribution .....	42
<b>Figure 60:</b> FTSE 100 Max Return – Hill Estimator.....	45
<b>Figure 61:</b> NIKKEI 225 Max Return – Hill Estimator.....	45
<b>Figure 62:</b> S&P 500 Max Return – Hill Estimator.....	45



**Figure 63:**FTSE 100 Max Return Density: Hill Estimator - Pareto Distribution ..... 47

**Figure 64:**NIKKEI 225 Max Return Density: Hill Estimator - Exponential Distribution..... 48

**Figure 65:**S&P 500 Max Return Density: Hill Estimator - Pareto Distribution ..... 49

## **Abbreviations**

BM - Block Maxima

EDF – Empirical Distribution Function

EVT - Extreme Value Theory

FTSE 100 – Financial Times Stock Exchange Index

GEV - Generalized Extreme Value

GPD - Generalized Pareto Distribution

i.i.d. – independent and identically distributed

MLE - Maximum Likelihood Estimate Method

NIKKEI 225 - Tokyo Stock Exchange

POT - Peaks Over Threshold

PORT - Peaks Over Random Threshold

S&P500 – Standard and Poor's Index

## Commonly used notations

$\mu$  – Location parameter

$\sigma$  – Scale parameter

$\xi$  – Shape parameter

$\alpha$  - Tail index

$u$  - Threshold

$\Phi_\alpha(x)$  - Fréchet distribution

$\Psi_\alpha(x)$  - Weibull distribution

$\Lambda_\alpha(x)$  - Gumbel distribution

$H_\xi(x)$  - Generalized Extreme Value distribution

$G_\xi(x)$  - Generalized Pareto distribution

## Executive Summary

The last years have been characterized by significant instabilities in financial markets. This has led to numerous criticisms about the existing risk management systems and motivated the search for more appropriate methodologies able to deal with rare events with catastrophic consequences, as happened in 1929 with the Great Depression crisis and the Sub-prime crisis in 2007 which originated the biggest crisis since the Great Depression.

The typical question one would like to answer is: If things go wrong, how wrong can they go? Then the problem is how to model the rare phenomena. The answer can be found in the Extreme value theory (EVT), which provides a strong theoretical foundation on which we can build statistical models describing extreme events. One important example of an extreme event is the convulsion in financial markets that shows that asset prices can display extreme movements beyond those captured by the normal distribution. One of the solutions to deal with this problem is the Extreme Value Theory.

To model extreme observations using Extreme Value Theory (EVT), we can use two different statistical distribution families: Generalized Extreme Value (GEV) and Generalized Pareto Distribution (GPD).

In this thesis, EVT methods were used to investigate and fit the empirical distribution of the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices to the theoretical GEV and GPD distributions.

We apply two approaches of extreme value theory, the *Block Maxima* and the *Peaks Over Threshold (POT)* approach, as well as the parametric approach of the Maximum Likelihood Estimate Method (MLE) for the distribution parameter estimation and the non-parametric approach of the Hill estimator.

We can say that the goal is to try understanding which theoretical GEV and GPD distributions better fit our data, monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500.

## 1. Introduction

In the last few years, Extreme Value Theory (EVT) has gained increased importance in modeling extreme observations in all social sciences (e.g. hydrology). This is especially true in finance, since EVT is a tool used to consider probabilities associated with extreme and thus rare events. Assessing the probability of rare and extreme events is an important issue in the risk management of financial portfolios and EVT provides the solid fundamentals needed for the statistical modeling of such events and the computation of extreme risk measures.

EVT is useful in modeling the impact of crashes or situations of extreme stress on investor portfolios.

To model extreme observations, we use two different statistical distribution families in this thesis: Generalized Extreme Value (GEV) distribution, which has the *Gumbel*, *Fréchet* and *Weibull* distributions as particular cases; and Generalized Pareto Distribution (GPD) distributions, which has the *Exponential*, *Pareto* and *Beta* distributions as particular cases.

In this thesis, EVT methods are used to investigate and fit the empirical distribution of the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices to the theoretical GEV and GPD distributions. We apply two approaches of extreme value theory to our data, the *Block Maxima* and the *Peaks Over Threshold (POT)*, as well as the parametric approach of the Maximum Likelihood Estimate Method (MLE) for the distribution parameter estimation and the non-parametric approach of the Hill estimator.

For the application of the methodology, we use a diversity of tools such as the R Programming Language with an extRemes toolkit, as well as the Easy Fit 5.5 Professional Software and the EViews Software.

This thesis is organized as follows: Section 2 presents a literature review, section 3 provides EVT's theoretical background, section 4 provides an empirical study and results, section 5 provides conclusions and section 6 provides some final comments and future directions of our work and research.

## 2. Literature Review

The last years have been characterized by significant instabilities in financial markets. This has led to numerous criticisms about the existing risk management systems and motivated the search for more appropriate methodologies able to deal with rare events with catastrophic consequences, as happened in 1929 with the Great Depression crisis; Oil crises in 1973 and 1979; the dot-com crisis in 2000 and the Sub-prime crisis in 2007 which originated the biggest crisis since the Great Depression.

The typical question one would like to answer is: If things go wrong, how wrong can they go? Then the problem is how to model the rare phenomena that lie outside the range of available observations. In such a situation it seems essential to rely on a well founded methodology. Extreme value theory (EVT) provides a strong theoretical foundation on which we can build statistical models describing extreme events.

One of the first papers that dealt with extreme value problems dates back to 1709, when Nicholas Bernoulli discussed the mean of the largest distance among points lying at random on a line. The notion of the distribution of the largest value is more modern, and it was first introduced by von Bortkiewicz (1922).

In the next year Von Mises (1923) evaluated the expected value of this distribution, and Dodd (1923) calculated its median, also studying some non-normal related distributions.

The period of the 1920s and 1930s was an important period in which several authors wrote papers dealing with practical applications of extreme value statistics in distributions of human lifetimes, strength of materials, flood analysis and seismic analysis.

One of these works, proposed and developed by Fréchet (1927), was the analysis of asymptotic distribution of the largest values. Fréchet identified one possible limit distribution could only be one of three types, and the independent analysis of Fisher and Tippet (1928) for the same problem released the paper which is considered the foundation of EVT and showed that the distribution of normalized maxima can only be one of three types: type I or Fréchet, type II or Weibull and type III or Gumbel. Von Mises (1936) presented some simple and useful sufficient conditions for the weak convergence of the larger order statistic for each of the three types of limiting distributions given earlier by Fischer and Tippet (1928). Gnedenko (1943) presented the foundations for extreme value theory providing necessary and sufficient

conditions for Fischer and Tippett's "three types theorem", the sufficient conditions for the weak convergence of the extreme order statistics. Mejlzer (1949), Marcus and Pinsky (1969) and de Haan (1970, 1971) refined the work of Gnedenko.

One of the empirical applications in this period was the paper of Weibull (1939) on metallurgical failures. This paper led Gumbel (1941, 1958) to propose a statistical methodology for studying extremes based on fitting the extreme value distributions to data consisting of maxima or minima of some random process over a fixed block or interval of time. Contemporary methods derived from this early work involve fitting block maxima and minima with the generalized extreme value (GEV) distribution, which combines Fisher-Tippett and Gnedenko's three types of distributions into a single, three-parameter distribution.

In 1955 Jenkinson proposed the Generalized Extreme Value distribution (GEV), with the three asymptotic distributions mentioned before as particular cases. As important as GEV, we have the Generalized Pareto Distribution (GPD). While both distributions model extreme events, GEV fits maximum (minimum) from blocks of data through *Block Maxima Method* (Fisher and Tippett Theorem, 1928), while GPD fits exceedances over high threshold  $u$  through *Peaks Over Threshold (POT)* method (Pickands, Balkema-de-Haan Theorem, 1975). A pioneer in the applications of POT was Pickands (1975), who showed that excesses over a high threshold, follow asymptotically a generalized Pareto distribution. By taking into account all exceedances over (shortfalls below) an appropriately high (low) threshold, these methods make more efficient use of data by incorporating information on all extreme events in any given block rather than only the maximum among them.

From the three parameters of both GEV and GPD,  $\mu$ ,  $\sigma$  and  $\xi$  are the location, scale and shape parameters, respectively. The parameter  $\xi$  is of particular relevance because it is closely related to the tail heaviness of the distributions. Although with Block Maxima the definition of extreme observation is straightforward, that is not the case of POT method, on which a threshold  $u$  (or alternatively, the choice of the number of  $q$  extreme observations) has to be considered. Some authors have shown how the threshold selection influences the parameter estimation, among them Smith (1987), Coles and Tawn (1994, 1996), Embrechts et al. (1997) and Davidson and Smith (1990).

The most popular estimator for the tail index is the Hill (1975) estimator. However, due to the weaknesses of the Hill estimator, some alternative estimators have been proposed in the literature. Beirlant et al. (1996), for instance, suggest an optimal threshold choice by

minimizing bias-variance of the model, whereas DuMouchel (1983) and Chavez-Demoulin (1999) suggest the use of 10 percent of the sample to estimate the parameters of the distribution. For the latter threshold selection, Chavez-Demoulin and Embrechts (2004) provide a sensitivity analysis which shows that small changes in threshold  $u$  have nearly no impact on the estimation results. Mendes and Lopes (2004) suggest a procedure to fit by quasi-maximum likelihood a mixture model where the tails are GPD and the center of the distribution is normal. Bermudez and Turkman (2001) suggest an alternative method of threshold estimation by choosing the number of upper order statistics.

Other approaches are quantile-quantile plot and mean excess function (Embrechts *et al.*, 1997). The Pickands (1975) estimator has also been mentioned frequently and Dekkers and de Haan (1989) showed the consistency and asymptotic normality of this estimator. However, the Pickands estimator is highly sensitive to the number of  $q$  statistics used and its asymptotic variance is large. Refinements to this estimator have been proposed by Falk (1994), Alves (1995), Drees (1995), Yun (2000, 2002) and Castillo and Hadi (1997). Dekkers *et al.* (1989) presented the moment estimator.

Another advance was presented in Gomes and Martins (2002), where a tail estimator is developed through a maximum likelihood approach based on scaled log-spacing. There has been some promising work on using bootstrap methods to determine the optimal number of upper order statistics (Danielsson and de Vries, 1997), but further validation of such methods is still required.

In regards to estimation methods, the most widely used is the Maximum Likelihood Estimation Method (MLE). Smith (1985) described such method in detail, providing its numerical solvency and validity of properties (efficiency, invariance under changes of the data in location and scale and extendibility to various regression models) for  $\xi > 0.5$ . Hosking and Wallis (1987) derived a simple method of moments which works unless  $\xi < 0.5$ . They also apply a variant with probability weighted moments (PWM). Castillo and Hadi (1997) proposed an elemental percentile method (EPM) that does not impose any restrictions on the value of  $\xi$ , whereas Coles and Powell (1996) applied Bayesian methods. According to the extensive simulation studies by Matthys and Beirlant (2003), maximum likelihood provides the best estimator for  $\xi > 0$ , whereas EPM is preferred if  $\xi$  is estimated to be less than zero.

Within the GEV context, the testing problem of the so-called Gumbel hypothesis ( $H_0: \xi = 0$ ) has received much attention in the literature. The correct choice of the GEV is very important,



since the three types differ considerably in their right tails. Authors who studied this matter in detail include Van Montfort (1970), Bardsley (1977), Otten and Van Montfort (1978), Tiago de Oliveira (1981), Gomes (1982), Tiago de Oliveira (1984), Tiago de Oliveira and Gomes (1984), Hosking (1984), Marohn (1994), Wang *et al.* (1996) and Marohn (2000). Somehow related to this matter are goodness-of-fit tests for the Gumbel model, studied by Stephens (1976, 1977 and 1986).

The tests therein considered are based on Empirical Distribution Function (EDF) -statistics such as Kolmogorov, Cramér-Von-Mises and Anderson-Darling statistics. The hypothesis of nullity of  $\xi$  has also been tested in the POT setup. Considering the papers related with this testing problem, we mention Van Montfort and Witter (1985), Gomes and Von Montfort (1986) and more recently Brilhante (2004).

Further developments on the extreme value theory include serially dependent observations provided that the dependence is weak (Berman, 1964 and Leadbetter, 1939) and extension from univariate to multivariate analysis. In different fields of application like finance, it is important to be able to model joint extreme events such as large losses in several stock returns or large changes in several rates, simultaneously. The first works were presented by de Haan and Resnick (1977), followed by Tawn (1988) for bivariate extremes and later Coles and Tawn (1991). Recent work has been developed by de Haan and Ferreira (2006) and Beirlant *et al.* (2004).

Despite the extensive literature review, it is still possible to mention some additional references for general knowledge of the theory. Castillo (1988) has successfully updated Gumbel (1958) and presented many statistical applications of extreme value theory with emphasis on engineering. Galambos (1978, 1987), Tiago de Oliveira (1984), Resnick (1987) and Leadbetter *et al.* (1983) presented elaborated treatments of the asymptotic theory of extremes. Related approaches with application to insurance are to be found in Beirlant *et al.* (1996), Reiss and Thomas (1997) and the references therein. Interesting case studies using up to date EVT methodology are McNeil (1997) and Resnick (1997). The various steps needed to perform a quantile estimation within the above EVT context are reviewed in de Haan *et al.* (1994) as well as in McNeil and Saladin (1997), where a simulation study is also to be found. Embrechts *et al.* (1997) give a detailed overview of the EVT as a risk management tool. Muller *et al.* (1998) and Pictet *et al.* (1998) studied the probability of exceedances and compare them with GARCH models for the foreign exchange rates. Recently, McNeil and

Frey (2000) and Chavez-Demoulin et al. (2004) address the use of EVT for the estimation of the conditional distribution for non-stationary financial series.

To end this literature review, it is important to mention that nowadays extreme value theory is well established in many fields of modern science, engineering, insurance and finance. Recently, numerous research studies have analyzed the extreme variations that financial markets are subject to, mostly because of currency crises, stock market crashes and large credit defaults, where it is possible to include this thesis. The tail behavior of financial time series distribution has, among others, been discussed in Koedijk et al. (1990), Dacorogna et al. (1995), Loretan and Phillips (1994), Longin (1996, 2004), Danielsson and de Vries (2000), Kuan and Webber (1998), Straetmans (1998), McNeil (1999), Jondeau and Rockinger (1999), Kläupelberg (1999), Neftci (2000), McNeil and Frey (2000), Gençay et al. (2003b) and Tolikas and Brown (2006). In Longin (1996), where the author presents a study of extreme stock market price movements, using data from the New York Stock Exchange for the period 1885 – 1990, the author empirically shows that the extreme returns of the New York Stock Exchange follow a Fréchet distribution. Another example is the paper from Tolikas and Brown (2006), where the authors use the Extreme Value Theory (EVT) to investigate the asymptotic distribution of the lower tail of daily returns in the Athens Stock Exchange (ASE) over the period 1986 to 2001. In terms of empirical findings, they discovered that the Generalized Logistic (GL) distribution provides an adequate description of the ASE index daily returns minima. Its asymptotic convergence was found to be relatively stable, especially when large selection intervals were used. This is an important finding since current EVT applications in finance focus exclusively on either the GEV or GPD distributions. These implications for investors could be important since the GL is fatter tailed than its GEV and GPD counterparts, which implies higher probabilities of the extremes occurring.

The next section introduces a theoretical background of EVT - Extremes Values Theory, the different approaches and the different parameter estimation methods.

### 3. Theoretical Background

This section introduces the EVT – Extreme value theory and its reasoning. EVT can be understood as a theory that provides methods for modeling extremal events. Extremal events are the observations that take values from the tail of the probability distribution. EVT provides the tools to estimate the parameters of a distribution of the tails through statistical analysis of the empirical data.

One important example is the finance field, where EVT can be applied. The convulsion in financial markets has been evidence that asset prices can display extreme movements beyond those captured by the normal distribution. In the literature, one of the solutions to deal with this situation at that point is the Extreme Value Theory.

However, before starting EVT, it is important to mention some properties of the financial data. In particular, it is by now well-known that returns on financial assets typically exhibit higher than normal kurtosis as expressed by both higher peaks and fatter tails than can be found in normal distribution.

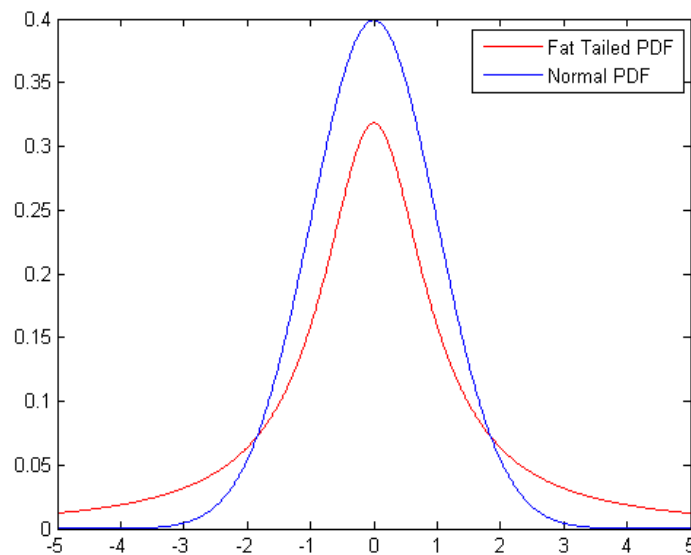


Figure 1: Asset returns distribution versus normal distribution

In figure 1, the red graph (Fat Tailed PDF) refers to the returns of a financial asset. The graphic tells us that the probabilities that normal distribution assigns to extreme events (tail events) are less than is required. In other words, the tails of the normal distribution are too thin to address the extreme losses or gains, assuming normality will lead to systematic

underestimation of the riskiness of an asset and increase the chance of having a hit. Thus, in order to get rid of such kind of problems, tails of the distribution must be modeled. This can be done by using EVT.

Within the EVT context, there are three approaches to modeling the extremal events: the first one for directly modeling the distribution of maximum/minimums returns named *Block Maxima*, the second one for modeling the exceedances of a particular estimated threshold named *Peak-Over-Threshold* (POT), and the third one denominated *Peaks-Over-Random-Threshold* (PORT).

The *Block Maxima*, which is used for the largest observations that are collected from large samples of identically and distributed observations (i.i.d), is a method that provides a model that may be appropriate for the monthly or annual maximum/minimums of such samples. The second method is the *Peaks-Over-Threshold*, which is used for all large observations that exceed a high threshold. They are considered to be the most useful methods for practical applications because of their more efficient use of the data on extreme values. The third method is the most recent of all and it is denominated *Peaks-Over-Random-Threshold*, which is a small variant of POT, where the threshold is a random variable.

When modeling the maxima/minima of a random variable, extreme value theory follows the same fundamental role as the central limit theorem follows when modeling sums of random independent variables. In both cases, the theory tells us what the limiting distributions are.

### 3.1 The Block Maxima Approach

Let's consider a random variable representing monthly maximum/minimum returns, which takes in successive periods. These selected observations constitute the extreme events, also called block (or per period) maxima/minima. In figure 2, the observations  $X_2$ ,  $X_5$ ,  $X_7$  and  $X_{11}$  represent the block maxima for four periods of three observations each<sup>1</sup>.

---

<sup>1</sup> Theoretical background formulas are presented for only the maximum returns, since we can obtain the formulas for the minimum returns with the opposite:  $H_n(r) = \text{Min}(r_1, \dots, r_n) = -\text{Max}(-r_1, \dots, -r_n)$

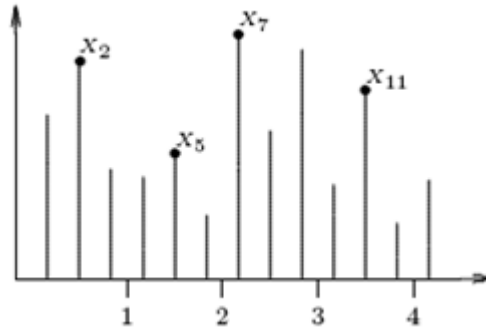


Figure 2: Block-maxima graph

The block maxima method is the traditional method used to analyze data with seasonality, as for instance, hydrological data. However, the threshold method uses data more efficiently, and for that reason, seems to have become the method of choice in recent applications.

According to the theorems of Fisher and Tippett (1928) and Gnedenko (1943), regardless of the specific distribution, the appropriately scaled maxima converge to one of three possible limit laws (parametric distributional forms). Under certain conditions, a standardized form of the three limit laws is called the *generalized extreme value (GEV)* distribution. Additionally, as is explained below by the theorems of Balkema and de Haan (1974) and Pickands (1975), the distribution function of the excesses above a high threshold converges to the generalized Pareto distribution (GPD).

Therefore, the block maxima model can be presented under a single family which is known as the *generalized extreme value (GEV)* distribution. The theory deals with the convergence of maxima, that is, the limit law for the maxima. To illustrate this, suppose that  $r_t, t=1, \dots, n$ , is a sequence of independent and identically distributed (i.i.d) observations with distribution function  $H(x) = \Pr\{r_t \leq x\}$ , not necessarily known, and let the sample maximum be denoted by  $M_n = \max\{r_1, \dots, r_n\}$  where  $n \geq 2$ , and  $\mathfrak{R}$  denote the real line. More generally, the generalized extreme value distribution (GEV) represented by  $H_\xi(x)$  describes the limiting distribution of suitably normalized maxima. The random variable  $X$  may be replaced by  $(X - \mu)/\sigma$  to obtain a standard GEV with a distribution function that is specified as shown below, where  $\mu$ ,  $\sigma$  and  $\xi$  are the location, scale and shape parameters, respectively.

The function  $H$  can belong to one of the three standard extreme value distributions:

$$\text{Fréchet: } \Phi\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad \alpha > 0 \quad (1)$$

$$\text{Weibull : } \Psi\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0 \quad (2)$$

$$\text{Gumbel : } \Lambda(x) = \exp(-\exp)^{-x}, x \in \mathfrak{R} \quad (3)$$

However, instead of having three functions, Jenkinson (1955) and Von Mises (1954) suggested the following distribution:

$$H_\xi(x) = \begin{cases} \exp^{-(1+\xi x)^{\frac{-1}{\xi}}} & \text{if } \xi \neq 0 \\ \exp^{-(\exp)^{-x}} & \text{if } \xi = 0 \end{cases} \quad (4)$$

This generalization, known as the *generalized extreme value (GEV)* distribution, is obtained by setting  $\xi = \alpha^{-1} > 0$  for the Fréchet distribution,  $\xi = -\alpha^{-1} < 0$  for the Weibull distribution and by interpreting the Gumbel distribution as the limit case for  $\xi = 0$ . (Note that  $\alpha$  refers to the tail index, which is the inverse of the shape parameter, which is defined as  $\alpha = \xi^{-1}$ ).

The shapes of the probability density functions for the standard Fréchet, Weibull and Gumbel distributions are given in figure 3.

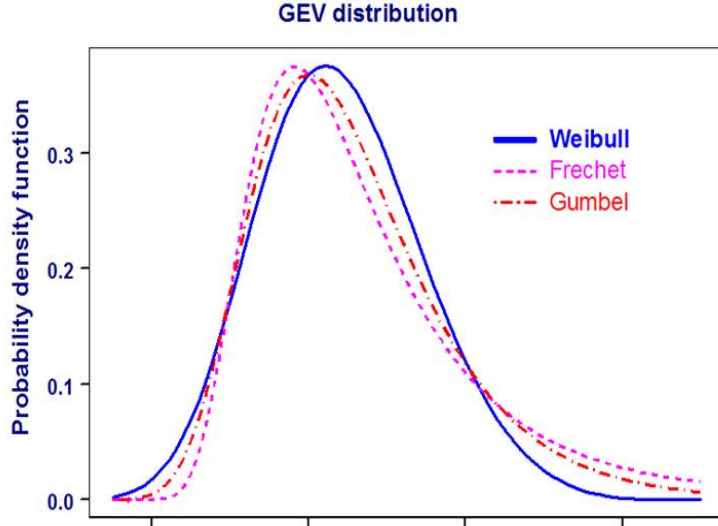


Figure 3: Densities for the Fréchet, Weibull and Gumbel functions

We observe that the Fréchet distribution has a polynomial decaying tail and therefore suits heavy-tailed distributions well. The exponentially decaying tails of the Gumbel distribution characterize thin-tailed distributions. Finally, the Weibull distribution is the asymptotic distribution of finite endpoint distributions.

As in general it is not possible to know in advance the type of limiting distribution of the sample maxima, the generalized representation is particularly useful when maximum likelihood estimates have to be computed. Moreover, the standard GEV defined in (4) is the limiting distribution of normalized extremes, given that in practice the true distribution of the returns is not known, and as a result, we used the three parameter specifications

$$H_{\xi, \sigma, \mu}(x) = H_{\xi}\left(\frac{x - \mu}{\sigma}\right) \quad x \in D, \quad D = \begin{cases} \left[ -\infty, \mu - \frac{\sigma}{\xi} \right] & \xi < 0 \\ \left[ -\infty, +\infty \right] & \xi = 0 \\ \left[ \mu - \frac{\sigma}{\xi}, +\infty \right] & \xi > 0 \end{cases} \quad (5)$$

of the GEV, which is the limiting distribution of the normalized maxima.

The class of distribution of  $H(x)$  where the Fisher-Tippett theorem holds is quite large. One of the conditions is that  $H(x)$  has to be in the domain of attraction of the Fréchet distribution ( $\xi > 0$ ), which in general holds for the financial time series. Gnedenko (1943) shows that if the tail of  $H(x)$  decays like a power function, then it is in the domain of attraction of the

Fréchet distribution. The class of distributions whose tails decay like a power function is large and includes Pareto, Cauchy, Student- $t$ , which are the well-known heavy-tailed distributions.

### 3.2 Peaks-Over-Threshold Approach (POT)

The second approach focuses on the realizations exceeding a given (high) threshold. The observations  $X_1, X_2, X_7, X_8, X_9$  and  $X_{11}$  in the figure 4, all exceed the threshold  $u$  and constitute extreme events.

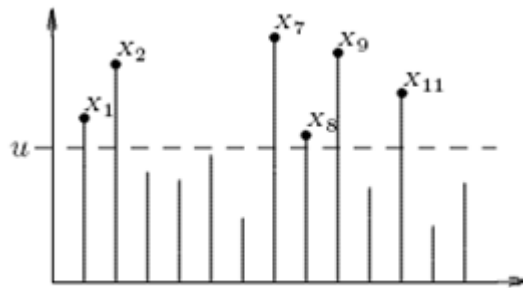


Figure 4: Excesses over a threshold  $\mu$  graph.

In general, we are not only interested in the maxima/minima of observations, but also in the behavior of large observations that exceed a high threshold. One method of extracting extremes from a sample of observations,  $r_t, t = 1, 2, \dots, n$  with a distribution function  $F(x) = \Pr\{r_t \leq x\}$  is to take exceedances over a deterministic predetermined high-threshold  $u$  (figures 4 and 5). Exceedances of a threshold  $u$  occur when  $r_t > u$  for any  $t$  in  $t = 1, \dots, n$ . An excess over  $u$  is defined by  $y = r_t - u$ .

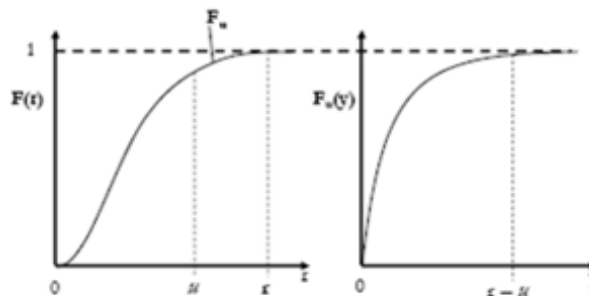


Figure 5: Distribution Function and Distribution Over Threshold



Given a high threshold  $u$ , the probability distribution of excess values of  $r_t$  over threshold  $u$  is defined by

$$F_u(y) = \Pr\{r - u \leq y \mid r > u\} \quad (6)$$

This represents the probability that the value of  $r$  exceeds the threshold  $u$  by at most an amount  $y$  given that  $r$  exceeds the threshold  $u$ . This conditional probability may be written as

$$F_u(y) = \frac{\Pr\{r - u \leq y, \quad r > u\}}{\Pr(r > u)} = \frac{F(y + u) - F(u)}{1 - F(u)} \quad (7)$$

Since  $x = y + u$  for  $r > u$ , we have the following representation:

$$F(x) = [1 - F(u)]F_u(y) + F(u) \quad (8)$$

Notice that this representation is valid only for  $r > u$ .

A theorem due to Balkema and de Haan (1974) and Pickands (1975) shows that for sufficiently high threshold  $u$ , the distribution function of the excesses may be approximated by GPD, because as the threshold gets large, the excess  $F_u(y)$  converges to the GPD.

The GPD in general is defined as

$$G_{\xi, \sigma, u}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x - u}{\sigma}\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp^{-(x - u)/\sigma} & \text{if } \xi = 0 \end{cases} \quad (9)$$

With

$$x \in \begin{cases} [u, \infty], & \text{if } \xi \geq 0 \\ \left[u, u - \frac{\sigma}{\xi}\right], & \text{if } \xi < 0 \end{cases} \quad (10)$$

Where  $\sigma$  and  $\xi$  are the scale and shape parameters, respectively. There is a simple relationship between the standard GPD  $G_\xi(x)$  and  $H_\xi(x)$  such that  $G_\xi(x) = 1 - \log H_\xi(x)$  if  $\log H_\xi(x) > -1$ .

The GPD embeds a number of other distributions; the  $\xi$  (shape parameter) determines the original distribution. When  $u = 0$  and  $\sigma = 1$ , and  $\xi > 0$ , the tail of the distribution function  $F$  of  $x$  decays like a power function  $x^{-1/\xi}$ . In this case,  $F$  belongs to a family of heavy-tailed distributions that includes, among others, the Pareto, log-gamma, Cauchy and t-distributions; it takes the form of the ordinary *Pareto distribution*. This particular case is the most relevant for financial time-series analysis, since it is a *heavy-tailed* distribution.

For  $\xi = 0$ , the tail of  $F$  decreases exponentially, and belongs to a class of medium-tailed distributions that includes the normal, exponential, gamma and log-normal distributions.

Finally, for  $\xi < 0$ , the underlying distribution  $F$  is characterized by a finite right endpoint, whose class of short-tailed distributions includes the uniform and beta distributions.

The shapes of the probability density functions for the standard Beta, Pareto and Exponential distributions are given in figure 6.

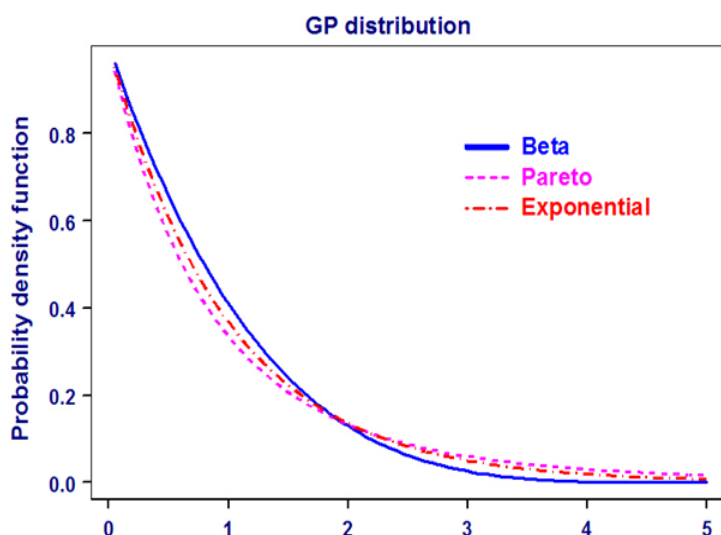


Figure 6: Densities for the Beta, Pareto and Exponential functions

The importance of the Balkema and de Haan (1974) and Pickands (1975) results is the fact that the distribution of excesses may be approximated by the GPD by choosing  $\xi$  and setting a high threshold  $u$ . The GPD model can be estimated by using the parametric methods, explained below.

### 3.3 Peaks-Over-Random Threshold Approach (PORT)

Another method, and the most recent one, is denominated *Peaks-Over-Random-Threshold (PORT)*. It is a semi-parametric method that is a small variant of POT, a method which is conditioned by a random threshold.

As we do not have an accurate choice of the threshold level, this new method introduces a new approach to dealing with this problem. Theory tells us that the threshold level should be high in order to satisfy the Pickands-Balkema-de Haan theorem, but the higher the threshold, the fewer observations are left for the estimation of the parameters of the tail distribution function.

In this method, any inference concerning the tail of the underlying distribution is based exclusively on the observations above a random threshold. This method compares with the alternative method of the number of observations that exceed a given high increasing deterministic level  $u$ , an approach named POT method.

Recently, Neves *et al.* (2006) and Neves and Fraga Alves (2007) have introduced two testing procedures that are based on the sample observations above a random threshold, procedures that consist in a reformulation of the asymptotic properties of the Hasofer and Wang (1992) test statistic.

### 3.4 Parameter Estimation Methods

In this section we explain the statistical parameter estimation of the distributions. Two approaches are considered. The first one, where the parameters of the distribution, including the tail index, are directly estimated by classical methods such as the maximum likelihood estimate method (MLE), is named the parametric approach. The second, where no parametric distribution is assumed for the extremes, is named the non-parametric approach.

#### 3.4.1. The parametric approach

The parametric approach assumes that maximum/minimum returns selected over a given period are exactly drawn from the extreme value distribution given by GEV (see formula 4), or alternatively, that those positive/negative return exceedances under or above a given threshold are exactly drawn from the distribution given by the GPD formula (see formula 9).

With either definition of extremes, the asymptotic distribution contains three parameters:  $\mu$ ,  $\sigma$  and  $\xi$  for extremes defined as maximum/minimum returns selected from a period containing  $n$  returns.

Under the assumption that the limit distribution holds, the maximum likelihood method gives unbiased and asymptotically normal estimators. The system of non-linear equations can be solved numerically using the Newton–Raphson iterative method. Both distributions are parameterized by the scale, location and shape parameters with the same interpretation in both cases. This method maximizes the likelihood function over the location, scale and shape parameters:  $\mu$ ,  $\sigma$  and  $\xi$ .

In practice, with MLE the parameters of the extreme value distributions can be estimated with different values of the number of returns contained in the selection period  $n$  (for maximum/minimum returns) and alternatively, with different values of the threshold  $\mu$  (for positive/negative return exceedances).

Mathematically, this solves the general formula presented below:

$$L(\mu, \sigma, \xi) = \frac{1}{\sigma^n} \prod_{i=1}^n \left( 1 + \xi \frac{x_i - \mu}{\sigma} \right)^{-(1/\xi + 1)} \times \exp \left\{ - \sum_{i=1}^n \left( 1 + \xi \frac{x_i - \mu}{\sigma} \right)^{-1/\xi} \right\} \quad (11)$$

The estimates of  $\mu$ ,  $\sigma$  and  $\xi$  – say  $\hat{\mu}$ ,  $\hat{\sigma}$  and  $\hat{\xi}$  – are taken to be those values which maximize the likelihood  $L$ .

However, this methodology presents some limitations. Maximum likelihood methods perform better when tails are thicker, providing greater observations exceeding the threshold. Any weaknesses in these assumptions directly affect the significance of the results. Jansen and de Vries (1991) show that in the Fréchet domain of attraction this includes most distributions of financial returns. Maximum likelihood methods are consistent, but not the most efficient. Therefore, in some situations we should use parametric methods that are alternatives to the Maximum Likelihood Estimate Method (MLE), such as the method of Probability Weight Moments (PWM). However, in this thesis we focus only on the first one.

### 3.4.2. The non-parametric approach

The parametric approach method assumes that the parameters are drawn exactly from the extreme value distribution. However, the non-parametric approach assumes that estimator for the shape parameter  $\xi$ , which do not assume that the observations of extremes follow exactly the extreme value distribution, have been developed by Pickands (1975) and Hill (1975).

Pickands's estimator is given by

$$\xi^{Pickands} = -\frac{1}{\ln 2} \ln \left( \frac{R'_{n-q+1} - R'_{n-2q+1}}{R'_{n-2q+1} - R'_{n-4q+1}} \right) \quad (12)$$

where  $(R'_t)_{t=1, N}$  is the series of returns ranked in an increasing order and  $q$  is an integer depending on the total number of returns contained in the database  $n$ . Pickands's estimator is consistent if  $q$  increases at a suitably rapid pace with  $n$ . Pickands's statistic is asymptotically normally distributed with mean  $\xi$  and variance  $\xi^2$ .

Hill's estimator is given by

$$\xi_{Hill} = \left( \frac{1}{q-1} \right) \sum_{i=1}^{q-1} (\ln R'_{n-i} - \ln R'_{n-q}) \quad (13)$$

Where  $R'_{N-1}$  denotes decreasing order statistics of the series and  $q$  denotes the  $q$ th smallest order statistic of  $n$  observations included in the calculation of the shape parameter. Intuitively the Hill index measures the distance with which the average extreme observation exceeds a specific threshold.

However, the Hill estimator has some weaknesses. Dekkers and De Haan (1990), who take into account the sensitivity of the Hill index to the initial tail size, propose an extension to the Hill estimator incorporating its second moment.

$$\xi_{Dekker} = \xi_{Hill1} + 1 - \frac{1}{2} \left( 1 - \left\{ \frac{(\xi_{Hill1})^2}{\xi_{Hill2}} \right\} \right)^{-1} \quad (14)$$

Where  $\xi_{Hill1}$  is the standard Hill estimate defined in equation (12) and

$$\xi_{Hill2} = \left( \frac{1}{q-1} \right) \sum_{i=1}^{q-1} [\ln R'_{n-i} - \ln R'_{n-q}]^2 \quad (15)$$

The authors prove the consistency and normality of  $\xi_{\text{Dekkers}}$ .

In summary, the problem consists of choosing the correct method to obtain the correct value for the tail index parameter. As largely discussed in the extreme value theory literature, the choice of its value is a critical issue. The origin of this problem comes from the fact that the database contains a finite number of maximum/minimum return observations; the number of extreme returns,  $q$ , used for the estimation of the model is finite.

On one hand, choosing a high value for  $q$  leads to few observations of extreme returns and implies inefficient parameter estimates with large standard errors. On the other hand, choosing a low value for  $q$  leads to many observations of extreme returns but induces biased parameter estimates as observations not belonging to the tails are included in the estimation process.

This sensitive tradeoff can be dealt with in a variety of ways, as was described above. However, in recent years new approaches have been developed. One of them is when the levels at which the tail index begins to plateau is in some sense the ‘optimal’ choice of  $n$  where it is no longer as sensitive to individual observations. Other approaches to making this choice include the quantile–quantile plot, regression methods (Huisman et al., 2000) and the subsample bootstrap (Danielsson and de Vries 2001).

## 4. Empirical Application

In this section we proceed to the application of the GEV and GPD distributions to our database, which is composed of the monthly maximum and minimum returns of the FTSE100, NIKKEI225 and S&P500 indices. The goal is to fit our maximum and minimum return series to the GEV and GPD distributions.

There will be a description of the methodology that we adopted, an empirical analysis of our database, and finally, the application of the GEV and GPD distributions.

### 4.1 Methodology

In this thesis, we applied some of the theory presented above, such as the *Block Maxima* and *POT* approaches. We also applied parameter estimation methods such as the MLE method, which uses a parametric approach, and the Hill estimator, which uses a non-parametric approach. The goal is to fit our maximum and minimum return series to the GEV and GPD distributions.

For the application of the methodology, we used a diversity of tools such as the language R with an extRemes toolkit, whose outputs and code can be seen in the appendix. Another tool that we used was the Easy Fit 5.5 Professional Software, which is a powerful tool in terms of graphs. Finally, we used the EViews Software.

The data consist of the monthly maximum and minimum log returns from the FTSE100, NIKKEI225 and S&P500 indices. The monthly maximum and minimum refers to the maximum and minimum return, extracted from the daily log return, in each month for each index.

In terms of log return series, assuming that  $P_t$  is the closing value on day  $t$ , continuous daily logarithmic returns are given by:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (16)$$

The data that we considered were downloaded from the Yahoo Finance webpage. The period under analysis for the S&P 500 index is from January 3, 1950 to March 11, 2011. The FTSE

100 index corresponds to a period between February 4, 1984 and March 11, 2011, and the NIKKEI 225 index corresponds to the period between January 4, 1984 and March 11, 2011.

## 4.2 Empirical Results

In this section, we proceed to an exploratory data analysis, followed by the application of the *Block Maxima* and *POT* approaches.

To start, tables 1, 2 and 3 show the summary statistics of the daily log returns and monthly maximum (Max) and minimum (Min) log return series of the FTSE100, NIKKEI 225 and S&P500 indices.

Data	FTSE 100 - Daily	FTSE 100 - Min	FTSE 100 - Max
Observations	6807	324	324
Mean	0.0002	-0.0191	0.0189
Median	0.0006	-0.0161	0.0150
Maximum	0.0938	-0.0045	0.0938
Minimum	-0.1303	-0.1303	0.0057
Std. Dev.	0.0112	0.0126	0.0117
Skewness	-0.3850	-3.4470	2.8954
Kurtosis	11.7930	24.6915	14.8249
Jarque-Bera	22097.09	6993.67	2340.40
Probability	0.0000	0.0000	0.0000

Table 1: FTSE 100 Explanatory data analysis

Data	NIKKEI225 - Daily	NIKKEI225 - Min	NIKKEI225 - Max
Observations	6684	327	327
Mean	0.0000	-0.0254	0.0256
Median	0.0004	-0.0231	0.0218
Maximum	0.1323	-0.0036	0.1323
Minimum	-0.1614	-0.1614	0.0062
Std. Dev.	0.0146	0.0164	0.0158
Skewness	-0.2214	-2.8612	2.6781
Kurtosis	11.2980	20.0629	15.1732
Jarque-Bera	19231.05	4412.97	2409.92
Probability	0.0000	0.0000	0.0000

Table 2: NIKKEI 225 Explanatory data analysis



Data	S&P500 - Daily	S&P500 – Min	S&P500 – Max
<b>Observations</b>	15396	735	735
<b>Mean</b>	0.0003	-0.0165	0.0167
<b>Median</b>	0.0005	-0.0137	0.0144
<b>Maximum</b>	0.1096	-0.0018	0.1096
<b>Minimum</b>	-0.2290	-0.2290	0.0029
<b>Std. Dev.</b>	0.0097	0.0136	0.0100
<b>Skewness</b>	-1.0577	-6.7697	2.8213
<b>Kurtosis</b>	32.1044	88.7136	18.4995
<b>Jarque-Bera</b>	546263.20	230610.30	8332.26
<b>Probability</b>	0.0000	0.0000	0.0000

**Table 3:** S&P 500 Explanatory data analysis

Looking at table 1, it is possible to see that in the daily return the stock index FTSE100 has a mean of 0.02% and a standard deviation of 1.12%, with a maximum of 9.38% and a minimum of -13.03%. In terms of the maximum returns, the series has a mean of 1.89% and a standard deviation of 1.17%, with a maximum return of 9.38% and a minimum of 0.57%. In the minimum return FTSE 100, the series has a mean of -1.91% with a standard deviation of 1.26%, and a maximum of -0.45% and a minimum of -13.03%.

In the other stock index, the NIKKEI 225, we can see that in the daily return the series has a mean of 0.00% and a standard deviation of 1.46%, with a maximum of 13.23% and a minimum of -16.14%. In terms of the maximum returns, we can see that the series has a mean of 2.56% and a standard deviation of 1.58%, with a maximum return of 13.23% and a minimum of 0.62%. In the minimum return of the NIKKEI 225, the series has a mean of -2.54% with a standard deviation of 1.64%, a maximum of -0.36% and a minimum of -16.14%.

In the last stock index, the S&P 500, in terms of daily return the index has a mean of 0.03% and a standard deviation of 0.97%, with a maximum of 10.96% and a minimum of -22.90%. In terms of the maximum returns, it has a mean of 1.67% and a standard deviation of 1.00%, with a maximum return of 10.96% and a minimum of 0.29%. Finally, the minimum S&P 500 return has a mean of -1.65% with a standard deviation of 1.36%, a maximum of -0.18% and a minimum of -22.90%.

On the other hand, these tables show that the kurtosis and skewness values are very different from 3 and 0, respectively, for the distribution to be considered a normal distribution. However, in addition to this method, another way to prove the normality of a distribution is the Jarque-Bera (JB) test, which depends on the skewness and kurtosis estimates. Based on the JB test, it is possible to conclude if the normality assumption is rejected or not, depending on the probability associated with the test. If it is lower than 0.05 (the default significance level assumed for all tests in this thesis), we reject the null and the normality assumption; otherwise, we can admit the normality based on the considered sample.

$$JB \text{ Test} = \begin{cases} H_0 = \text{Normal Distribution} \\ H_1 \neq \text{Normal Distribution} \end{cases} \quad (17)$$

For the FTSE 100, NIKKEI 225 and S&P 500 series we can see that daily log returns and the monthly maximum and minimum log return series distributions do not follow a normal distribution, because in all cases we reject the null hypothesis in the JB test, as a consequence of the associated probability of the JB test being lower than 0.05.

In the figures below, it is possible to see that none of the histograms of the daily log returns and the monthly maximum and minimum log returns of the FTSE 100 (figures 7 – 9), NIKKEI 225 (figures 10-12) and S&P 500 (figures 13 – 15) seem to follow the normal distribution, which is confirmed by the rejection of the null hypothesis in the JB test.

In the figures below it is also possible to see and analyze the graph of the adjust close of the FTSE 100 (figure 16), NIKKEI 225 (figure 20) and S&P 500 (figure 24), as well as the daily log return and the monthly maximum and minimum returns graphs.

# How to deal with Extreme Observations in Empirical Finance

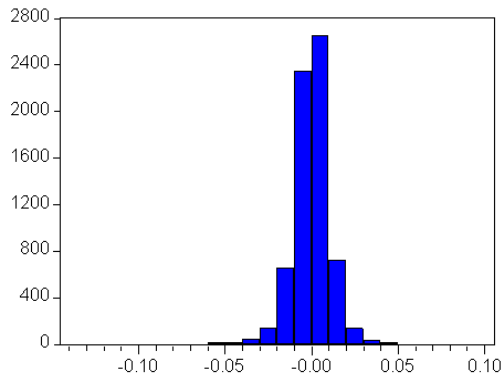


Figure 7: FTSE 100 Daily Returns Histogram

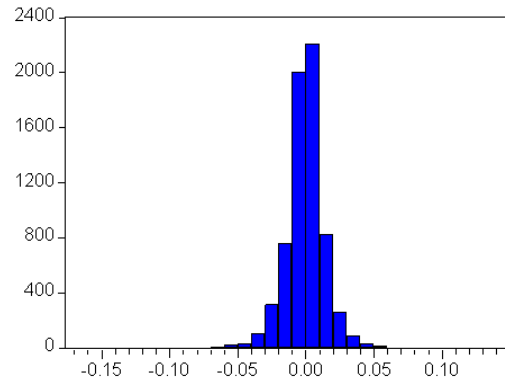


Figure 10: NIKKEI 225 Daily Returns Histogram

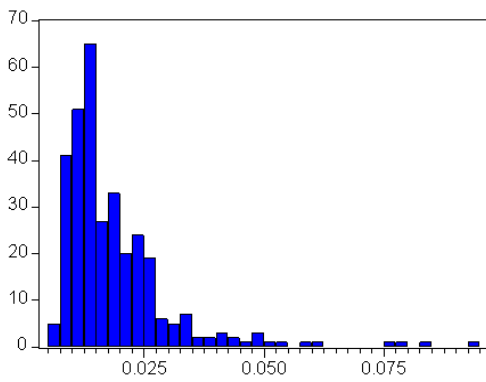


Figure 8: FTSE 100 Max Returns Histogram

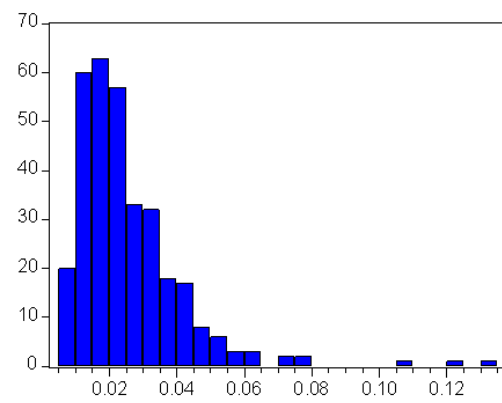


Figure 11: NIKKEI 225 Max Returns Histogram

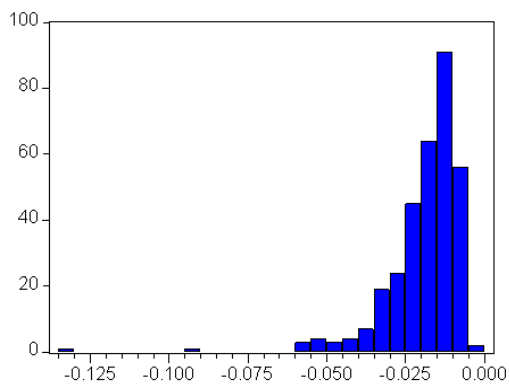


Figure 9: FTSE 100 Min Returns Histogram

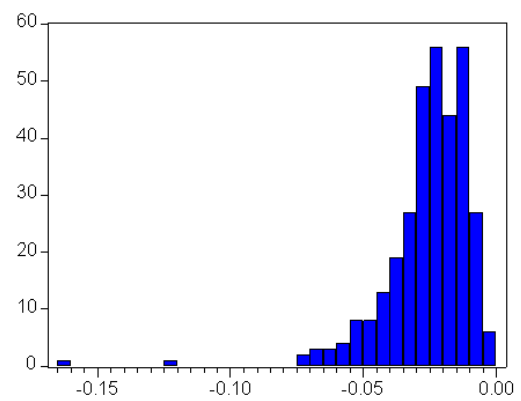


Figure 12: NIKKEI 225 Min Returns Histogram

# How to deal with Extreme Observations in Empirical Finance

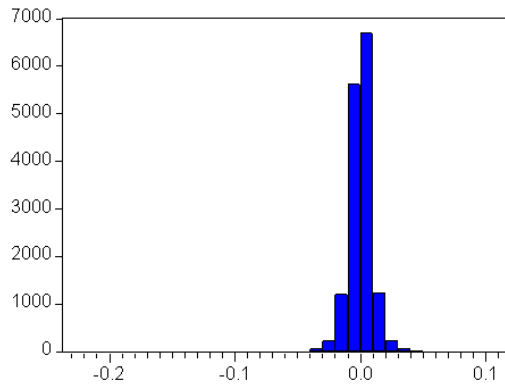


Figure 13: S&P 500 Daily Returns Histogram

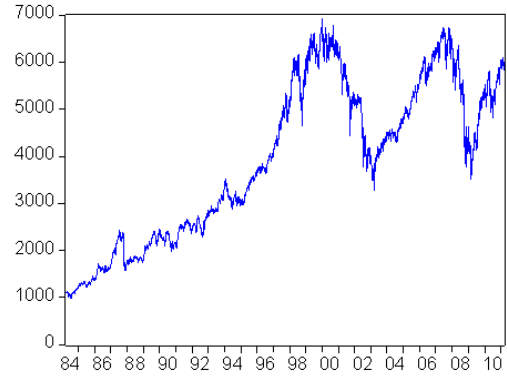


Figure 16: FTSE 100 Adjust Close

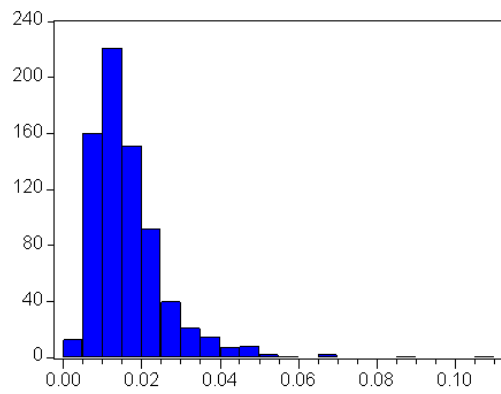


Figure 14: S&P 500 Max Returns Histogram

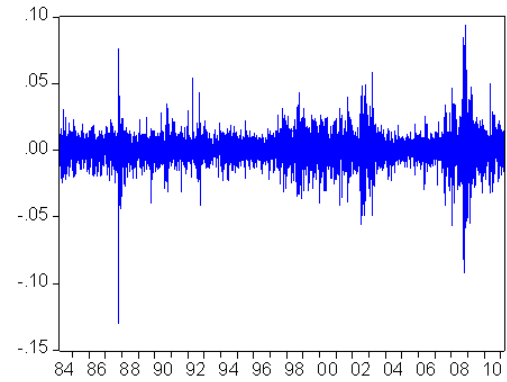


Figure 17: FTSE 100 Daily Return

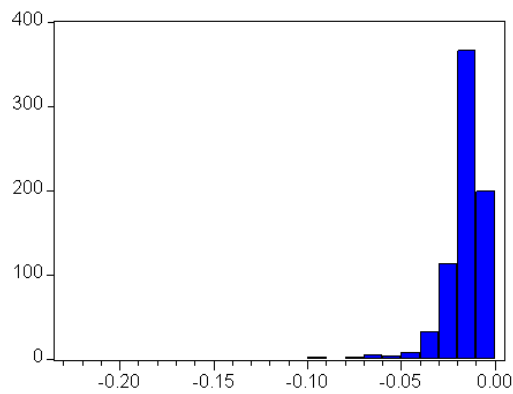


Figure 15: S&P 500 Min Returns Histogram

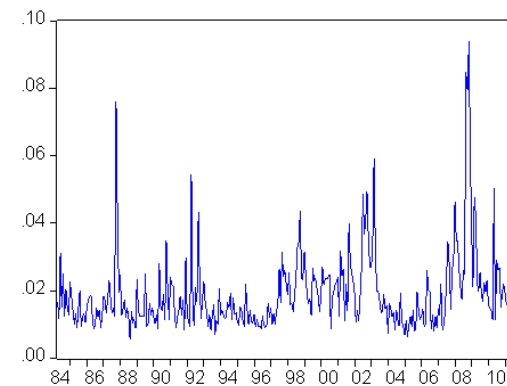


Figure 18: FTSE 100 Monthly Maximum Return

## How to deal with Extreme Observations in Empirical Finance

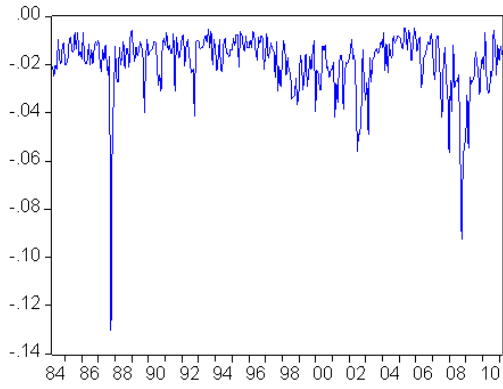


Figure 19: FTSE 100 Monthly Minimum Return

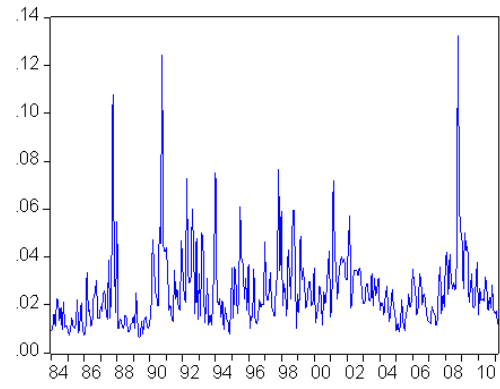


Figure 22: NIKKEI 225 Monthly Maximum Return

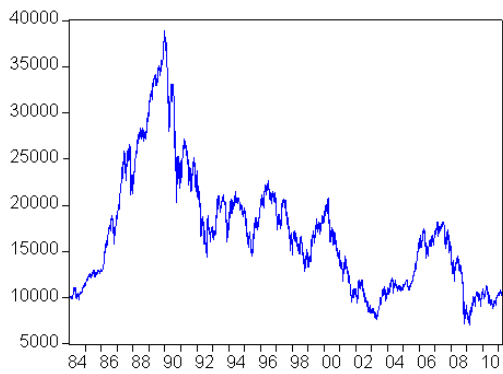


Figure 20: NIKKEI 225 Adjust Close

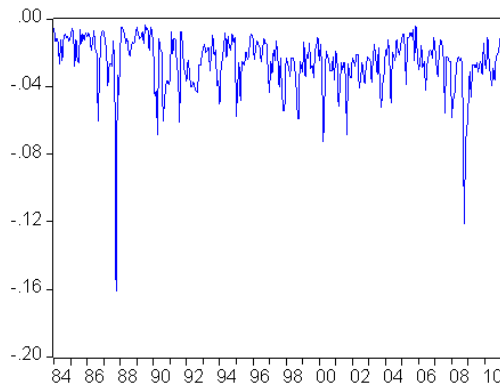


Figure 23: NIKKEI 225 Monthly Minimum Return

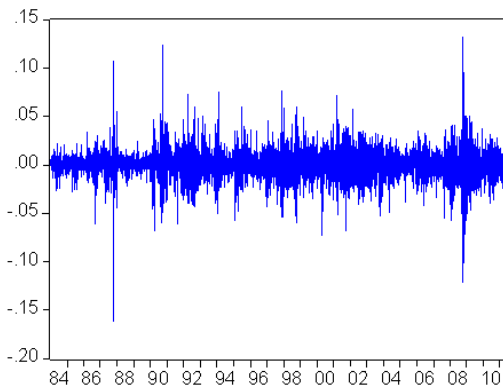


Figure 21: NIKKEI 225 Daily Return

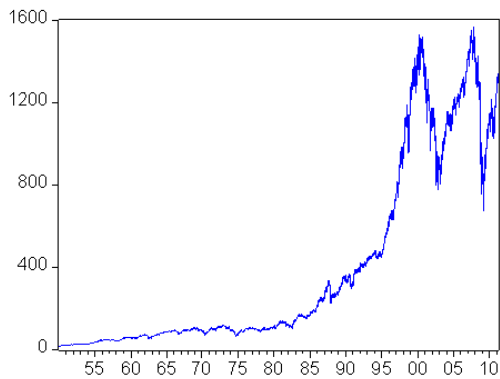


Figure 24: S&P 500 Adjust Close

## How to deal with Extreme Observations in Empirical Finance

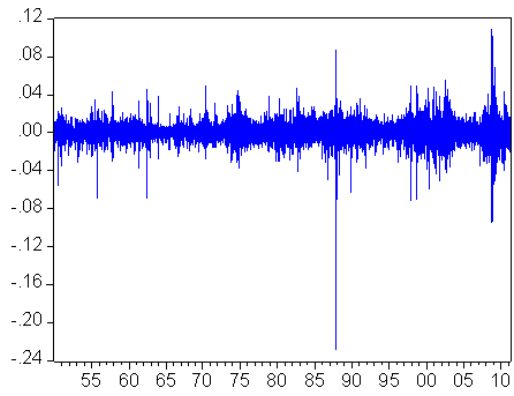


Figure 25: S&P 500 Daily Return

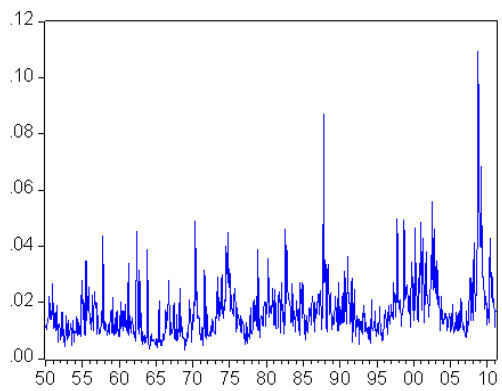


Figure 26: S&P 500 Monthly Maximum  
Return

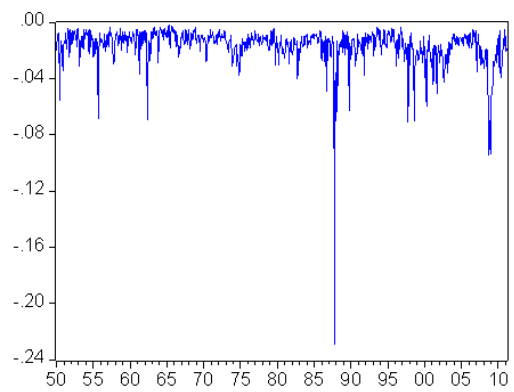


Figure 27: S&P 500 Monthly Minimum  
Return

### 4.3 Fitting the GEV Distribution

In this section we will determine which one of the extreme value distributions is more appropriate for modeling the monthly maximum/minimum returns of the FTSE 100, NIKKEI 225 and S&P500 indices under analysis, applying the GEV distribution using the MLE parametric approach.

We start with the LR test (Likelihood Ratio test) and the Log-Likelihood value for each of the series.

Based on the LR test, it is possible to conclude if the Gumbel distribution assumption is rejected or not, depending on the probability associated with the test. If it is lower than 0.05, we reject the null and the Gumbel distribution assumption; otherwise, we can admit the Gumbel distribution based on the considered sample.

$$LR \text{ Test} = \begin{cases} H_0 = \text{Gumbel Distribution} \\ H_1 \neq \text{Gumbel Distribution} \end{cases} \quad (18)$$

	Likelihood Ratio Test	p-value	Negative Log-likelihood
<b>FTSE100 -Max Return</b>	65.42	6.051E-16	-1125.03
<b>FTSE100 -Min Return</b>	603.54	2.850E-133	-1064.76
<b>NIKKEI 225 -Max Return</b>	30.51	3.318E-08	-993.17
<b>NIKKEI 225 -Min Return</b>	512.95	1.443E-113	-956.08
<b>S&amp;P 500 -Max Return</b>	56.84	4.734E-14	-2564.37
<b>S&amp;P 500 -Min Return</b>	1995.90	0.000E+00	-2426.38

**Table 4:** LR Test and Log-Likelihood Results - GEV

For a LR test, we reject the null hypothesis. This means that we do not assume that the series follows a Gumbel process. Therefore, we can conclude that none of the maximum/minimum return series have been generated by a Gumbel distribution. This means that for the considered series, only the Fréchet or Weibull particular cases of the more general GEV distribution remain.

Studying each series individually now, and starting with the monthly maximum and minimum series of the FTSE 100 index, using the MLE, we found the optimal values for the parameter estimates, which we present next:

	Obs.	MLE	Stand. Error
<b>FTSE100 -Max Return</b>	<b>324</b>		
Location - $\mu$		0.0136	0.0003
Scale - $\sigma$		0.0054	0.0002
Shape - $\xi$		0.3146	0.0511
<b>FTSE100 -Min Return</b>	<b>324</b>		
Location - $\mu$		-0.0201	0.0000
Scale - $\sigma$		0.0125	0.0000
Shape - $\xi$		-0.8015	0.0000

**Table 5:** FTSE 100 Maximum/Minimum Parameters - GEV

For the FTSE 100 maximum returns, we conclude that the Fréchet distribution (a common result in empirical finance) seems to be the most appropriate to describe that kind of data, as the shape ( $\xi$ ) estimate is higher than zero. Consequently, we can conclude that FTSE 100 maximum return empirical distribution has fat tails. The larger the estimate values for the shape parameter, the more fat-tailed the distribution. We can see the representation of the distribution in figure 30.

For the maximum return series, the Probability Plot (PP) (figure 28) and the Quantile Plot (QQ) (figure 29) seem to confirm that the data has been generated by the Fréchet distribution, as the points are near the straight line. In both graphs we are comparing the returns series (data) against the Fréchet distribution.

For the FTSE 100 minimum returns, it is possible to see that we have a Weibull distribution, because the shape estimate ( $\xi$ ) is lower than zero. We can see the histogram with the Weibull density function in figure 34.

For the minimum return series, we can see in the Probability Plot (PP) (figure 32) that the series could be generated by a Weibull process, since the graph is approximately linear. On the other hand, the Quantile - Quantile Plot (QQ) (figure 33) confirms the PP plot, since the observed data (series) tend to concentrate around the straight line, which confirm the Weibull goodness-of-fit. In both graphs we are comparing the series (data) empirical distribution against the theoretical Weibull distribution.

Therefore, in terms of monthly maximum and minimum FTSE 100 series, we can admit the series have been generated by Fréchet and Weibull process distributions, respectively.



# How to deal with Extreme Observations in Empirical Finance

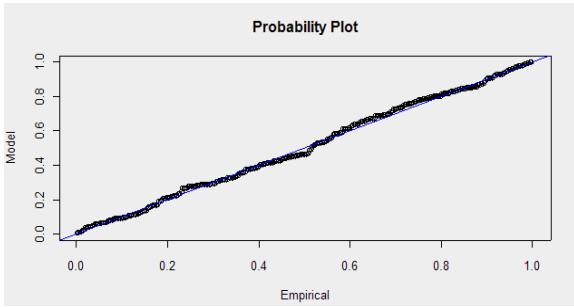


Figure 28: FTSE100 Max Return Fréchet Probability Plot

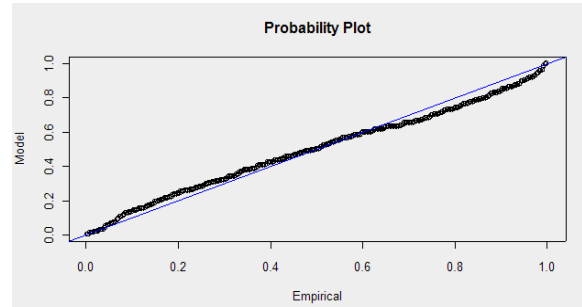


Figure 32: FTSE100 Min Return Weibull Probability Plot

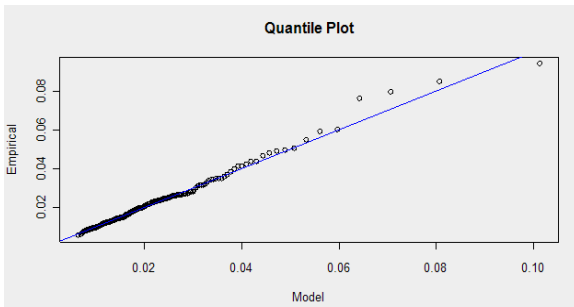


Figure 29: FTSE100 Max Return Fréchet Quantile Plot

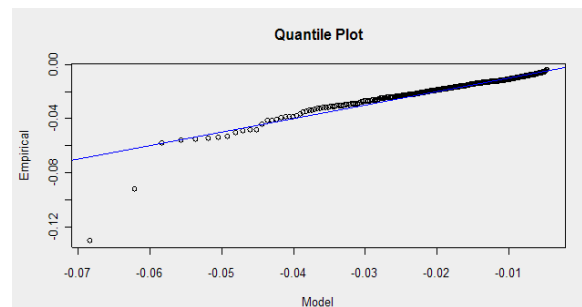


Figure 33: FTSE100 Min Return Weibull Quantile Plot

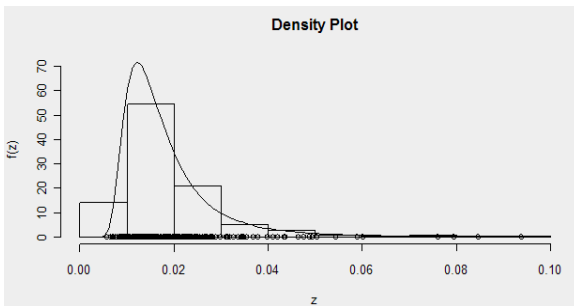


Figure 30: FTSE100 Max Return Density Plot

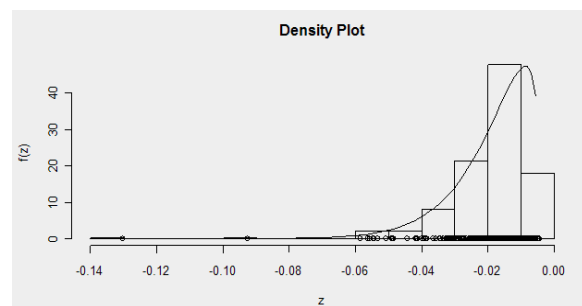


Figure 34: FTSE100 Min Return Density Plot

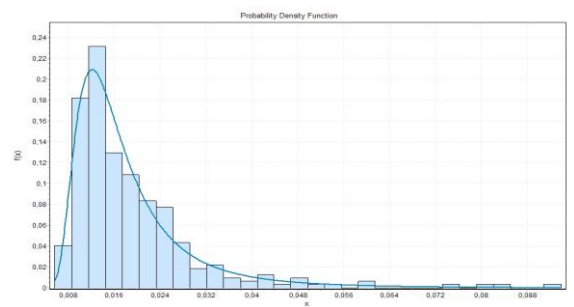


Figure 31: FTSE100 Max Return Density Plot II

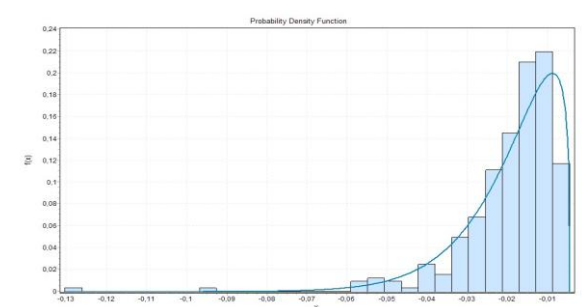


Figure 35: FTSE100 Min Return Density Plot II

Now analyzing the monthly maximum and minimum return series of the NIKKEI 225 index, we found the optimal values for the parameter estimates, which we present next:

	Obs.	MLE	Stand. Error
<b>NIKKEI 225 -Max Return</b>	<b>327</b>		
Location - $\mu$		0.0183	0.0006
Scale - $\sigma$		0.0087	0.0004
Shape - $\xi$		0.2251	0.0508
<b>NIKKEI 225 -Min Return</b>	<b>327</b>		
Location - $\mu$		-0.0278	0.0000
Scale - $\sigma$		0.0166	0.0000
Shape - $\xi$		-0.6799	0.0051

**Table 6:** NIKKEI 225 Maximum/Minimum Parameters - GEV

For the NIKKEI 225 maximum returns, we conclude that the Fréchet distribution seems to be the most appropriate to describe that kind of data, as the shape ( $\xi$ ) estimate is higher than zero. Therefore, we can conclude that NIKKEI 225 maximum return empirical distribution has fat tails. We can see the representation of the distribution in figure 38.

For the maximum return series, the Probability Plot (PP) (figure 36) and the Quantile Plot (QQ) (figure 37) seem to confirm that the data has been generated by the Fréchet distribution, as the points are near the straight line. In both graphs we are comparing the returns series (data) against the Fréchet distribution.

For the NIKKEI 225 minimum returns, it is possible to see that we have a Weibull distribution, because the shape estimate ( $\xi$ ) is lower than zero ( $\hat{\xi} = -0.67366$ ). We can see the histogram with the Weibull density function in figure 42.

For the minimum return series, we can see in the Probability Plot (PP) (figure 40) that the series could be generated by a Weibull process, since the graph is approximately linear. On the other hand, the Quantile - Quantile Plot (QQ) (figure 41) confirms the PP plot, since the observed data (series) tend to concentrate around the straight line, which confirm the Weibull goodness-of-fit. In both graphs we are comparing the series (data) empirical distribution against the theoretical Weibull distribution.

Consequently, in terms of monthly maximum and minimum NIKKEI 225 series, we can admit the series have been generated by Fréchet and Weibull process distributions, respectively, as happened with the monthly maximum and minimum FTSE 100.

# How to deal with Extreme Observations in Empirical Finance

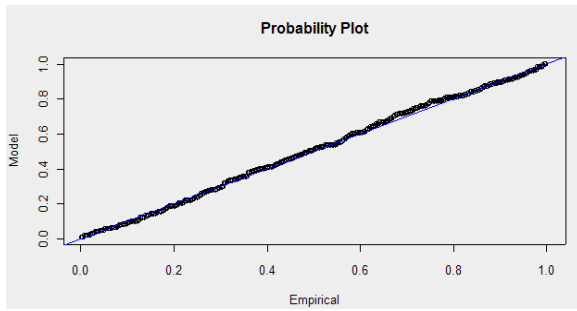


Figure 36: NIKKEI 225 Max Return Fréchet Probability Plot

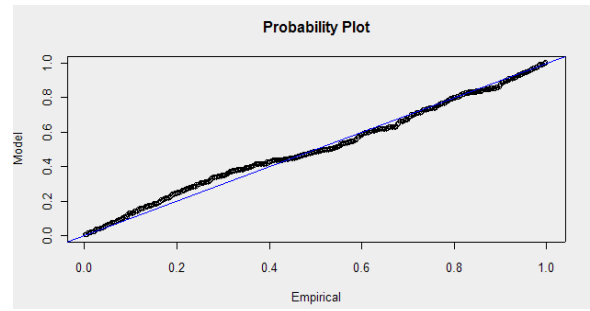


Figure 40: NIKKEI 225 Min Return Weibull Probability Plot

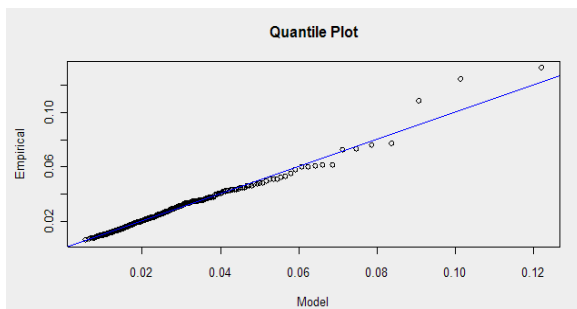


Figure 37: NIKKEI 225 Max Return Fréchet Quantile Plot

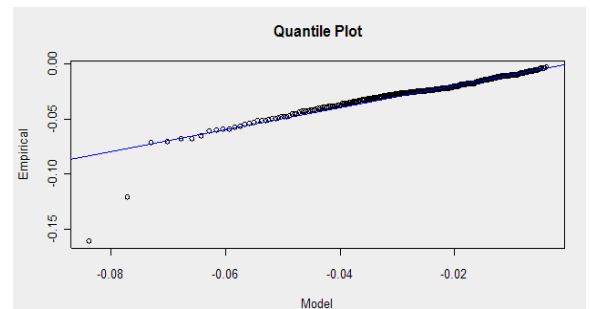


Figure 41: NIKKEI 225 Min Return Weibull Quantile Plot

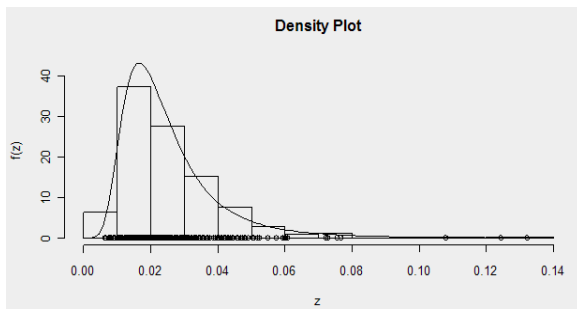


Figure 38: NIKKEI 225 Max Return Density Plot

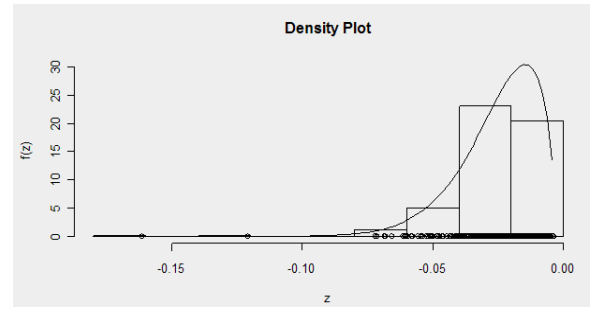


Figure 42: NIKKEI 225 Min Return Density Plot

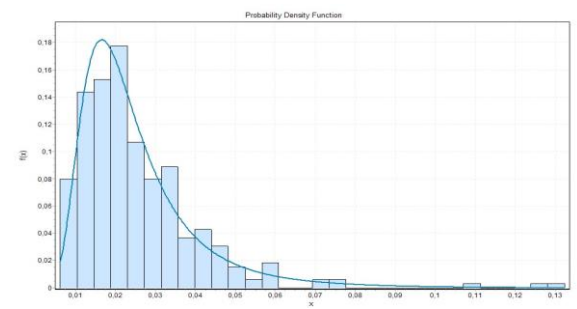


Figure 39: NIKKEI 225 Max Return Density Plot II

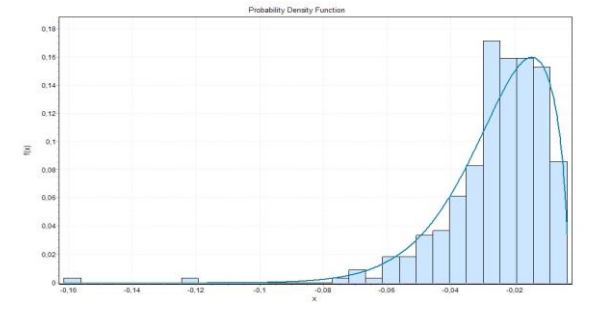


Figure 43: NIKKEI 225 Min Return Density Plot

To finish the GEV distribution application, we now proceed to the study of the maximum and minimum series of the S&P 500 index. Using the MLE, we found the optimal values for the parameter estimates, which we present next:

	Obs.	MLE	Stand. Error
<b>S&amp;P 500 -Max Return</b>	<b>735</b>		
Location - $\mu$		0.0122	0.0002
Scale - $\sigma$		0.0057	0.0002
Shape - $\xi$		0.1864	0.0301
<b>S&amp;P 500 -Min Return</b>	<b>735</b>		
Location - $\mu$		-0.0182	0.0000
Scale - $\sigma$		0.0119	0.0000
Shape - $\xi$		-0.7223	0.0000

**Table 7:** S&P 500 Maximum/Minimum Parameters - GEV

For the S&P 500 maximum returns, we conclude that the Fréchet distribution (a common result in empirical finance) seems to be the most appropriate to describe that kind of data, as the shape ( $\xi$ ) estimate is higher than zero. Therefore, we can conclude that S&P 500 maximum return empirical distribution has fat tails. The larger the estimate values for the shape parameter, the more fat-tailed the distribution. We can see the representation of the distribution in figure 46.

For the maximum return series, the Probability Plot (PP) (figure 44) and the Quantile Plot (QQ) (figure 45) seem to confirm that the data has been generated by the Fréchet distribution, as the points are near the straight line. In both graphs we are comparing the returns series (data) against the Fréchet distribution.

To finish the application of GEV distribution, for the S&P 500 minimum returns, it is possible to see that we have a Weibull distribution, because the shape estimate ( $\xi$ ) is lower than zero ( $\hat{\xi} = -0.69958$ ). We can see the histogram with the Weibull density function in figure 50.

For the minimum return series, we can see in the Probability Plot (PP) (figure 48) that the series could be generated by a Weibull process, since the graph is approximately linear. On the other hand, the Quantile - Quantile Plot (QQ) (figure 49) confirms the PP plot, since the observed data (series) tend to concentrate around the straight line, which confirm the Weibull goodness-of-fit. In both graphs we are comparing the series (data) empirical distribution against the theoretical Weibull distribution.

## How to deal with Extreme Observations in Empirical Finance

In conclusion, for monthly maximum and minimum S&P 500 series, as happens with the FTSE 100 and NIKKEI 225 series, we can admit the series have been generated by Fréchet and Weibull process distributions, respectively.

# How to deal with Extreme Observations in Empirical Finance

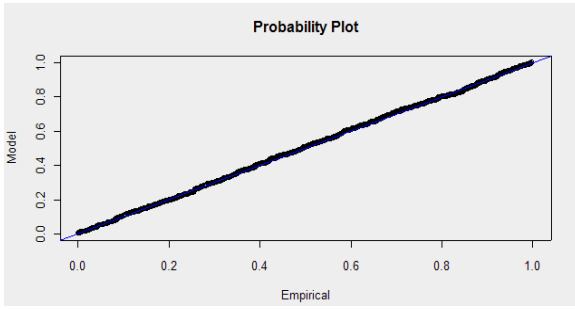


Figure 44: S&P 500 Max Return Fréchet Probability Plot

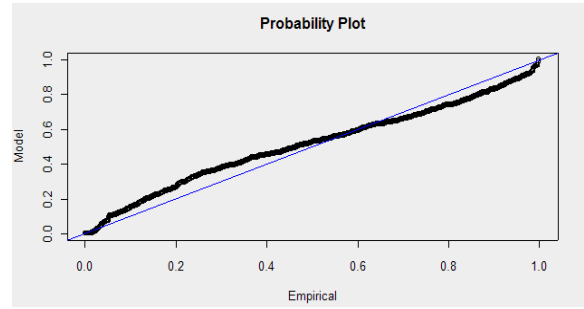


Figure 48: S&P 500 Min Return Probability Plot

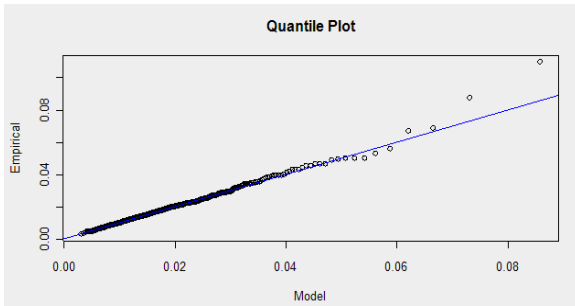


Figure 45: S&P 500 Max Return Fréchet Quantile Plot

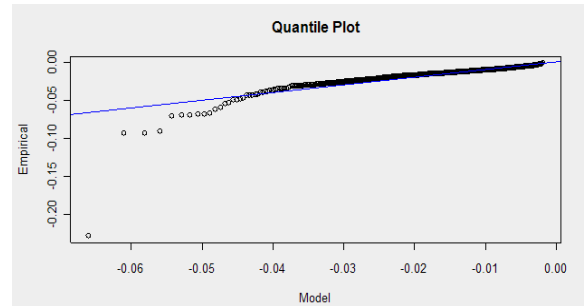


Figure 49: S&P 500 Min Return Quantile Plot

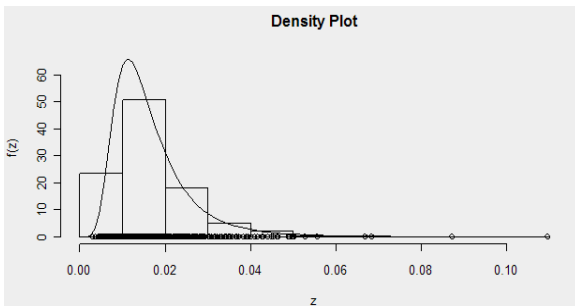


Figure 46: S&P 500 Max Return Density Plot

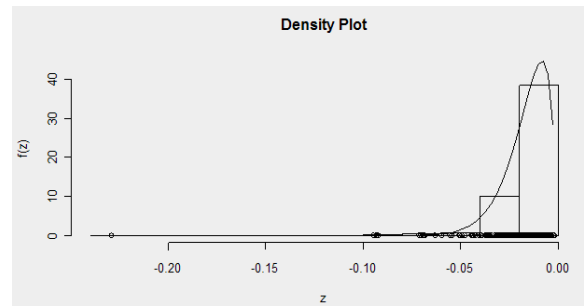


Figure 50: S&P 500 Min Return Density Plot

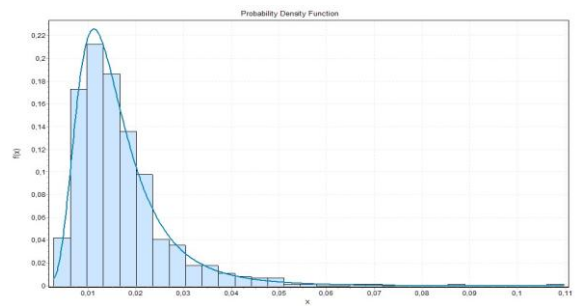


Figure 47: S&P 500 Max Return Density Plot

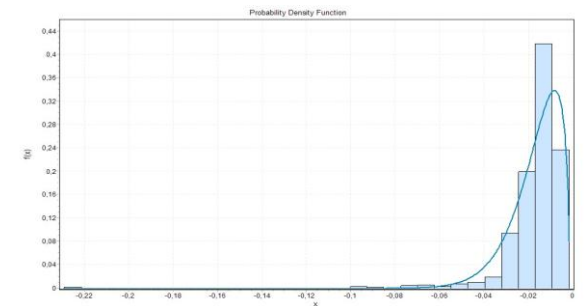


Figure 51: S&P 500 Min Return Density Plot

After fitting the GEV distribution to all our maximum and minimum series, and in accordance to the shape estimates, we can conclude that the maximum return series are well represented by the Fréchet distribution. This was already expected due to the fact that the same conclusion was reached by other studies in financial research related to maximum and minimum returns, as the financial series normally follows this distribution. However, the surprise happens in the minimum return series which follows the Weibull distribution instead of the Fréchet distribution.

As we already applied the GEV distribution, in the next section, we will apply the GPD distribution to the same database.

#### 4.4 Fitting the GPD

In this section we will consider the GPD distribution to model the monthly maximum and minimum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, applying the POT approach, considering a random threshold ( $u$ ) of 1%, 2.5% and 5%, and using the MLE parametric method to estimate the parameters.

To start we compute the LR test (Likelihood Ratio test) and the Log-Likelihood value for each of the series, where in the null we assume the monthly maximum/minimum returns have been generated by an exponential distribution:

$$LR \text{ Test} = \begin{cases} H_0 = \text{Exponential Distribution} \\ H_1 \neq \text{Exponential Distribution} \end{cases} \quad (19)$$

		Likelihood Ratio Test	p-value	Negative Log-likelihood
FTSE100 - Max Return	Threshold = 0.01	2.3773	0.1231	-988.1627
	Threshold = 0.025	4.5286	0.0333	-198.4939
	Threshold = 0.05	12.0860	0.0005	-29.4550
FTSE100 - Min Return	Threshold = 0.01	13.5067	0.0002	-306.8009
	Threshold = 0.025	174.2542	0.0000	-1003.6751
	Threshold = 0.05	290.0608	0.0000	-911.2157

Table 8: FTSE 100 LR Test and Log-Likelihood Results - GPD

For the LR test (significance level of 5%), we cannot assume the series have been generated by an exponential distribution when we reject the null hypothesis. In the case of the monthly maximum and minimum FTSE 100 series, we do not reject the null hypothesis for the FTSE

## How to deal with Extreme Observations in Empirical Finance

100 maximum return with a threshold ( $u$ ) of 1%, so we can consider the exponential distribution to represent the returns.

For the FTSE 100 monthly maximum return with a threshold of 2.5% and 5%, as well as the minimum return FTSE 100 (all thresholds), we reject the null hypothesis, since the probability associate is lower than 0.05. Consequently, the series do not follow an exponential distribution.

Analyzing the monthly maximum return series of the FTSE 100 index and using the MLE, the estimates for the parameters are presented next:

FTSE100 - Max Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.01</b>	278		
Scale - $\sigma$		0.0097	0.0008
Shape - $\xi$		0.0792	0.0568
<b>Threshold = 0.025</b>	58		
Scale - $\sigma$		0.0083	0.0020
Shape - $\xi$		0.3751	0.2198
<b>Threshold = 0.05</b>	8		
Scale - $\sigma$		0.0508	0.0000
Shape - $\xi$		-1.1588	0.0010

(a) Nr. Exceedances of threshold

**Table 9:** Maximum FTSE 100 Parameters - GPD

For the FTSE 100 maximum return with a threshold of 1%, we have already seen that we can assume the series follows an exponential distribution, with a scale estimate ( $\sigma$ ) equal to 0.0097 and a shape estimate ( $\xi$ ) equal to 0.0792. If we now consider the same series with a threshold of 2.5%, as the estimate for the shape parameter ( $\xi$ ) is higher than zero ( $\hat{\xi} = 0.3751$ ), we can assume the returns have been generated by a Pareto distribution. Finally, in considering a 5% threshold, we have seen that the series does not follow an exponential distribution, and as the shape parameter ( $\xi$ ) is lower than zero, we can assume the returns have been generated by a beta distribution.

For the FTSE 100 monthly maximum return, threshold of 1% and 2.5%, we can see in figures 52 and 53 the exponential and Pareto distribution, respectively, for each series.



## How to deal with Extreme Observations in Empirical Finance

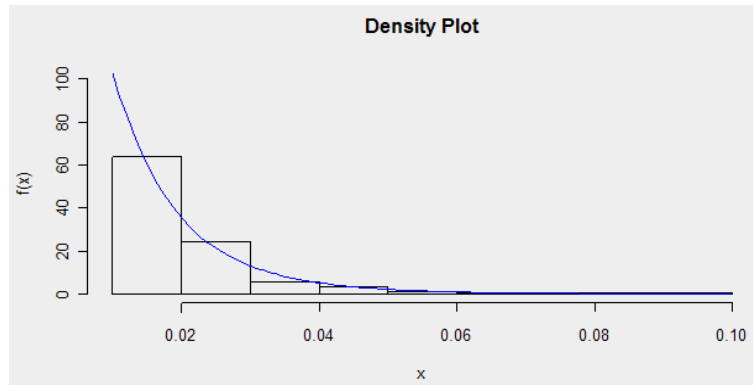


Figure 52: FTSE 100 Max Return Density Plot for a threshold of 1 % - Exponential Distribution

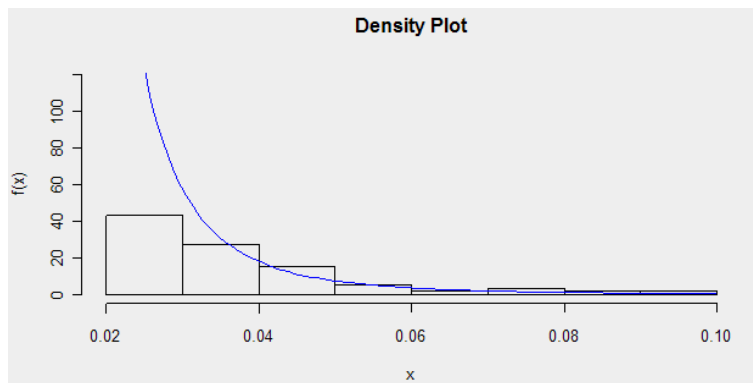


Figure 53: FTSE 100 Max Return Density Plot for a threshold of 2.5% - Pareto Distribution

Analyzing the monthly minimum return series of the FTSE 100 index and using the MLE, the estimates for the parameters are presented next:

FTSE100 - Min Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.01</b>	58		
Scale - $\sigma$		0.0033	0.0000
Shape - $\xi$		-0.5657	0.0317
<b>Threshold = 0.025</b>	258		
Scale - $\sigma$		0.0190	0.0000
Shape - $\xi$		-0.9241	0.0000
<b>Threshold = 0.05</b>	315		
Scale - $\sigma$		0.1827	0.0000
Shape - $\xi$		-4.0156	0.0000

(a) Nr. Exceedances of threshold

**Table 10:** Minimum FTSE 100 Parameters - GPD

## How to deal with Extreme Observations in Empirical Finance

As we presented above for the FTSE 100 monthly minimum return (all thresholds), we reject the null hypothesis in the LR test, so the series do not follow an exponential distribution.

Looking more specifically at the series, and considering all thresholds, as the shape estimate ( $\xi$ ) is lower than zero for all, we can assume that the returns have been generated by a beta distribution, for a threshold of 1%, 2.5% and 5%.

Therefore, in terms of monthly maximum and minimum FTSE 100 series, we can admit the maximum series with a threshold of 1% have been generated by an exponential distribution and with a threshold of 2.5% generated by a Pareto distribution. Finally, we can admit that the maximum series with a threshold of 5% and the minimum series (all thresholds) have been generated by a Beta distribution.

In now analyzing the LR (Likelihood Ratio test) results for the NIKKEI 225 monthly maximum and minimum return (table 11), we do not reject the null hypothesis for the NIKKEI 225 maximum series (all thresholds). Consequently, we can consider the exponential distribution to represent the maximum returns.

For the NIKKEI 225 monthly minimum return, for all thresholds (1%, 2.5% and 5%), we reject the null hypothesis, as the probability associate is lower than 0.05. Therefore, the series do not follow an exponential distribution.

		Likelihood Ratio Test	p-value	Negative Log-likelihood
NIKKEI 225 -Max Return	Threshold = 0.01	0.7986	0.3715	-944.9793
	Threshold = 0.025	1.4025	0.2363	-409.0120
	Threshold = 0.05	1.3433	0.2465	-55.9751
NIKKEI 225 -Min Return	Threshold = 0.01	15.4013	0.0001	-166.7374
	Threshold= 0.025	72.6604	0.0000	-732.6747
	Threshold= 0.05	207.2326	0.0000	-893.4935

**Table 11:** NIKKEI 225 LR Test and Log-Likelihood Results - GPD

Analyzing the monthly maximum return series of the NIKKEI 225 index, using the MLE, the estimates for the parameters are presented next:

NIKKEI 225 - Max Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.01</b>	306		
Scale - $\sigma$		0.0174	0.0012
Shape - $\xi$		-0.0374	0.0403
<b>Threshold = 0.025</b>	127		
Scale - $\sigma$		0.0135	0.0016
Shape - $\xi$		0.0862	0.0816
<b>Threshold = 0.05</b>	19		
Scale - $\sigma$		0.0134	0.0057
Shape - $\xi$		0.3701	0.3801

(a) Nr. Exceedances of threshold

**Table 12:** Maximum NIKKEI 225 Parameters - GPD

As was already mentioned, for the NIKKEI 225 monthly maximum return, for all thresholds (1%, 2.5% and 5%), we do not reject the null hypothesis. Therefore, we can consider the exponential distribution to represent the maximum returns, with a negative estimate parameter shape ( $\xi$ ) for the 1% threshold (figure 54) and positive estimate parameter ( $\xi$ ) for the other two (figures 55 and 56).

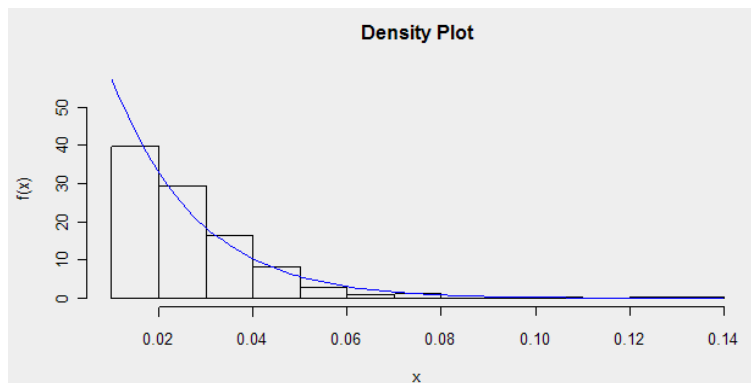


Figure 54: NIKKEI 25 Max Return Density Plot for a threshold of 1% - Exponential Distribution

## How to deal with Extreme Observations in Empirical Finance

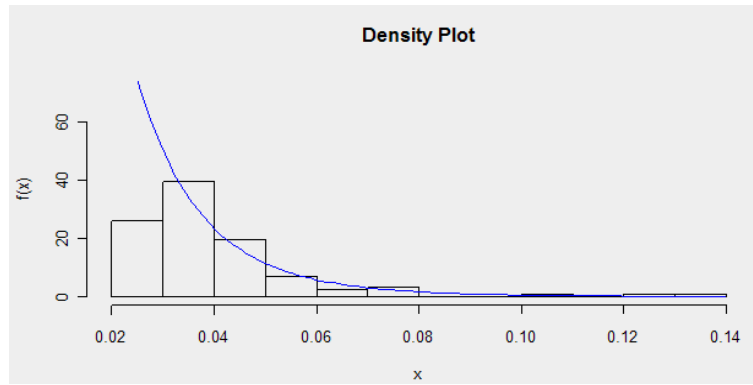


Figure 55: NIKKEI 225 Max Return Density Plot for a threshold of 2.5% - Exponential Distribution

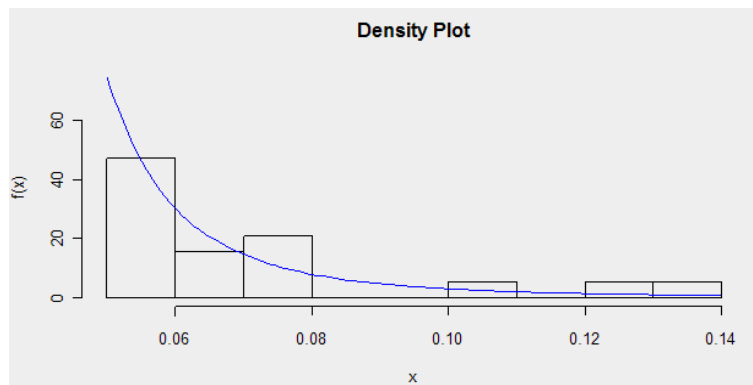


Figure 56: NIKKEI 225 Max Return Density Plot for a threshold of 5% - Exponential Distribution

In now analyzing the NIKKEI 225 monthly minimum return, considering all thresholds, we can assume that the minimum returns have been generated by a Beta distribution, as the shape estimate ( $\xi$ ) is lower than zero.

NIKKEI 225 - Min Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.01</b>	33		
Scale - $\sigma$		0.0060	0.0000
Shape - $\xi$		-0.9359	0.0197
<b>Threshold= 0.025</b>	189		
Scale - $\sigma$		0.0162	0.0000
Shape - $\xi$		-0.7527	0.0055
<b>Threshold= 0.05</b>	305		
Scale - $\sigma$		0.1077	0.0000
Shape - $\xi$		-2.3213	0.0000

(a) Nr. Exceedances of threshold

**Table 13:** Minimum NIKKEI 225 Parameters - GPD

Consequently, in terms of monthly maximum and minimum NIKKEI 225 series, we can admit the maximum series (all thresholds) have been generated by an exponential distribution. On the other hand, we can assume that the returns in the minimum series (all thresholds) have been generated by a beta distribution, as the shape estimate ( $\xi$ ) is lower than zero for all thresholds.

To finish the GPD distribution application, we now analyze the LR (Likelihood Ratio test) results for the S&P 500 monthly maximum and minimum return (Table 14). Since we do not reject the null hypothesis for the S&P500 maximum series (all thresholds), we can consider the exponential distribution to represent the maximum returns.

For the S&P500 monthly minimum return, for all thresholds (1%, 2.5% and 5%), we reject the null hypothesis, as the probability associate is lower than 0.05. Therefore, the series do not follow an Exponential distribution.

		Likelihood Ratio Test	p-value	Negative Log-likelihood
S&P 500 -Max Return	Threshold = 0.01	0.7079	0.4001	-2058.1463
	Threshold = 0.025	1.9681	0.1606	-346.3881
	Threshold = 0.05	0.0023	0.9615	-20.3554
S&P 500 -Min Return	Threshold = 0.01	41.5227	0.0000	-979.8349
	Threshold = 0.025	442.8829	0.0000	-2395.6007
	Threshold = 0.05	1112.9300	0.0000	-2251.1191

Table 14: S&P 500 LR Test and Log-Likelihood Results - GPD

Analyzing the monthly maximum return series of the S&P 500 index, using the MLE, the estimates for the parameters are presented next:

S&P 500 - Max Return	Nr. Exce. (a)	MLE	Std. Err.
Threshold = 0.01	562		
Scale - $\sigma$		0.0092	0.0005
Shape - $\xi$		0.0281	0.0352
Threshold = 0.025	98		
Scale - $\sigma$		0.0094	0.0013
Shape - $\xi$		0.1285	0.1078
Threshold = 0.05	7		
Scale - $\sigma$		0.0208	0.0163
Shape - $\xi$		-0.0339	0.6941

(a) Nr. Exceedances of threshold

Table 15: Maximum S&P 500 Parameters - GPD

## How to deal with Extreme Observations in Empirical Finance

As was already mentioned, for the S&P 500 monthly maximum return, for all thresholds (1%, 2.5% and 5%), we do not reject the null hypothesis. As a result, we can consider the exponential distribution to represent the maximum returns, with a positive estimate parameter shape ( $\xi$ ) for the 1% and 2.5% thresholds (figures 57 and 58) and a negative estimate parameter ( $\xi$ ) for the 5% threshold (figure 59).

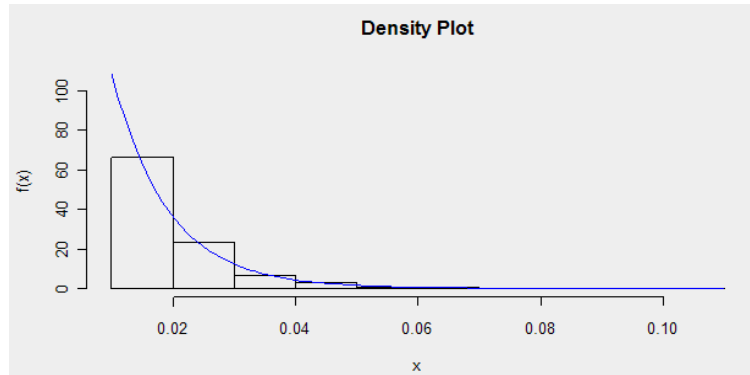


Figure 57: S&P 500 Max Return Density Plot for a threshold of 1% - Exponential Distribution

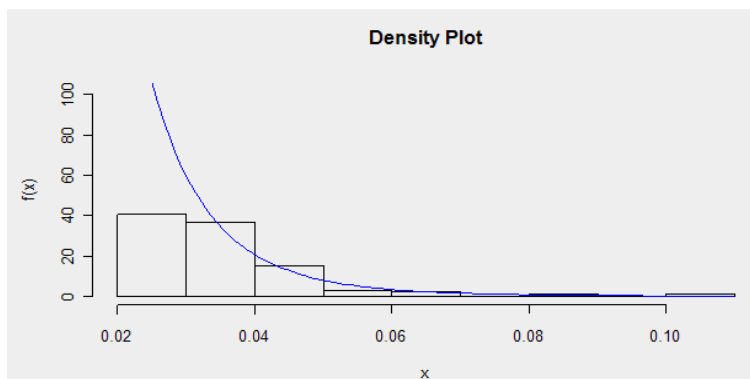


Figure 58: S&P 500 Max Return Density Plot for a threshold of 2.5% - Exponential Distribution

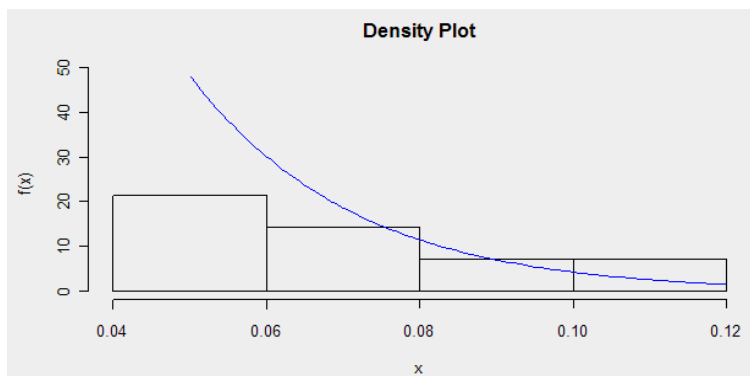


Figure 59: S&P 500 Max Return Density Plot for a threshold of 5% - Exponential Distribution

As we presented above for the S&P 500 monthly minimum return (all thresholds), we reject the null hypothesis in the LR test. Therefore, the series do not follow an exponential distribution.

In looking more specifically at the series, considering all thresholds, we can assume that the returns have been generated by a Beta distribution, for a threshold of 1%, 2.5% and 5%, as the shape estimate ( $\xi$ ) is lower than zero for all (Table 16).

S&P 500 - Min Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.01</b>	197		
Scale - $\sigma$		0.0047	0.0000
Shape - $\xi$		-0.5694	0.0061
<b>Threshold = 0.025</b>	635		
Scale - $\sigma$		0.0201	0.0000
Shape - $\xi$		-0.8668	0.0000
<b>Threshold = 0.05</b>	719		
Scale - $\sigma$		0.0721	0.0000
Shape - $\xi$		-1.4958	0.0006

(a) Nr. Exceedances of threshold

**Table 16:** Minimum S&P 500 Parameters - GPD

Consequently, in terms of monthly maximum and minimum S&P 500 series, we can admit the maximum series (all thresholds) have been generated by an exponential distribution. On the other hand, we can assume that the returns in the minimum series (all thresholds) have been generated by a Beta distribution, as the shape estimate ( $\xi$ ) is lower than zero.

After applying the GPD distribution to monthly maximum and minimum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, using the POT approach with a random threshold level of 1%, 2.5% and 5%, for the positive (maximum returns) and negative (minimum returns) tail, we can assume, in accordance to the shape estimates, that the maximum and minimum returns series are well represented mostly by the exponential distribution or by the Beta distribution. The only exception was in the FTSE 100 monthly maximum return with a threshold of 2.5%, which is well represented by the Pareto distribution.

This is a result that we were not expecting, since the Pareto distribution is the most common distribution for financial series, as we can see in other studies in financial research related to maximum and minimum returns. Because of this, it is a result that could be studied in more detail in future research.

However, one of the possible reasons for this result was the random threshold that we chose. As largely discussed in the extreme value theory literature, this value assumes a huge importance for two main reasons. The first one is the fact that choosing a high value for the threshold leads to few extreme observations and implies inefficient parameter estimates. The second reason is the fact that choosing a low value for the threshold level leads to many observations of extreme observations, but induces biased parameter estimates as observations not belonging to the tails are included in the estimation process.

For this reason, in this thesis we proceed to a calculation of the Hill estimator to optimize the tradeoff between bias and inefficiency. This process is presented in the next section.

### **4.5 Non-parametric approach – Hill’s estimator**

After we applied the POT approach using the parametric approach, this is the MLE method for the calculation of the parameters and choosing a random threshold. Now we proceed to the calculation of the Hill estimator, which is a non-parametric approach. This means that although we have tried several thresholds to apply the GPD distribution, we will now proceed to the calculation of the Hill estimator to obtain a threshold.

In this section we will consider the GPD distribution to model the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, applying the POT approach. For the threshold parameter we consider the values obtained from Hill’s estimator.

The Hill estimator is defined in formula (13). The number of  $q$  upper statistics to be considered should be chosen from the region where the line is roughly horizontal, which means stability around a constant for  $\xi$ . Despite the simplicity of the Hill’s estimator calculation, it has some weaknesses. This means that it presents some sensitivity to the initial tail size, or initial threshold, and therefore to the number of extreme observations.

Based on the results that we get from the Hill Plots, (figures 60 - 62) for the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, we should choose the values for  $q$  around 84; 85 and 175, respectively. The corresponding value of shape parameter would be approximately 0.35; 0.25 and 0.2, as it is possible to see in the figures below.



# How to deal with Extreme Observations in Empirical Finance

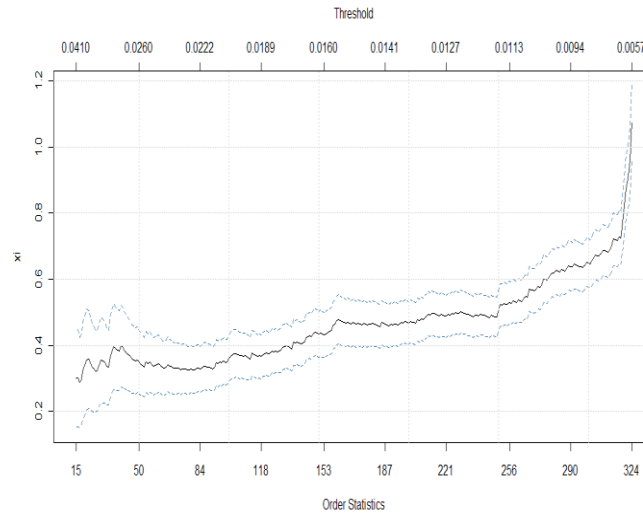


Figure 60: FTSE 100 Max Return – Hill Estimator

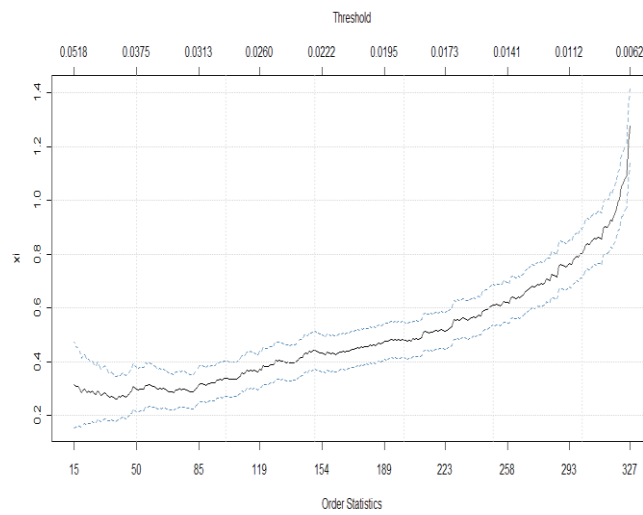


Figure 61: NIKKEI 225 Max Return – Hill Estimator

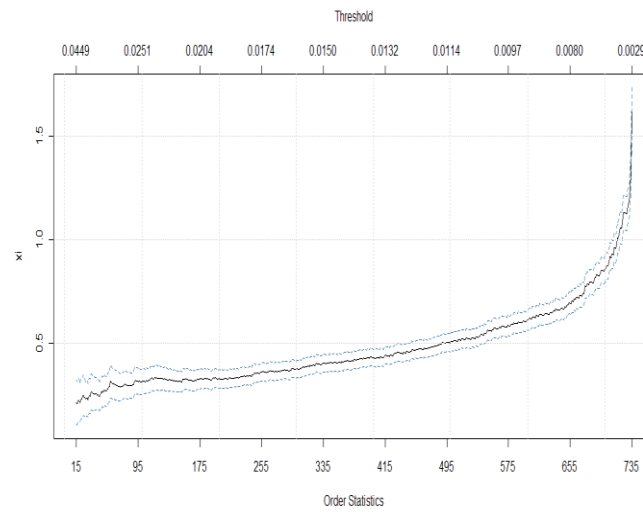


Figure 62: S&P 500 Max Return – Hill Estimator

To start, and as we did in the POT approach above, we compute the LR test (Likelihood Ratio test) and the Log-Likelihood value for each of the series, where in the null we assume the monthly maximum returns have been generated by an exponential distribution:

$$LR \text{ Test} = \begin{cases} H_0 = \text{Exponential Distribution} \\ H_1 \neq \text{Exponential Distribution} \end{cases} \quad (20)$$

FTSE100 -Max Return	Likelihood Ratio Test	p-value	Negative Log-likelihood
Threshold = 0.0222	8.4317	0.0037	-299.2261

**Table 17:** FTSE 100 LR Test and Log-Likelihood Results – GPD (Hill Estimator)

For the LR test (significance level of 5%), we cannot assume the series have been generated by an exponential distribution when we reject the null hypothesis. For the FTSE 100 monthly maximum returns with a threshold of 2.22%, we reject the null hypothesis, as the probability associate is lower than 0.05. Consequently, the series do not follow an exponential distribution.

Analyzing the monthly maximum return series of the FTSE 100 index and using the MLE, the estimates for the parameters are presented next:

FTSE100 -Max Return	Nr. Exce. (a)	MLE	Std. Err.
Threshold = 0.0222	84		
Scale - $\sigma$		0.0072	0.0013
Shape - $\xi$		0.3679	0.1605

(a) Nr. Exceedances of threshold

**Table 18:** Maximum FTSE 100 Parameters - GPD (Hill Estimator)

As we have already seen, the FTSE 100 maximum return series do not follow an exponential distribution. Considering the estimated parameters, as the estimate for the shape parameter ( $\xi$ ) is higher than zero ( $\hat{\xi} = 0.3679$ ), we can assume the returns have been generated by a Pareto distribution.

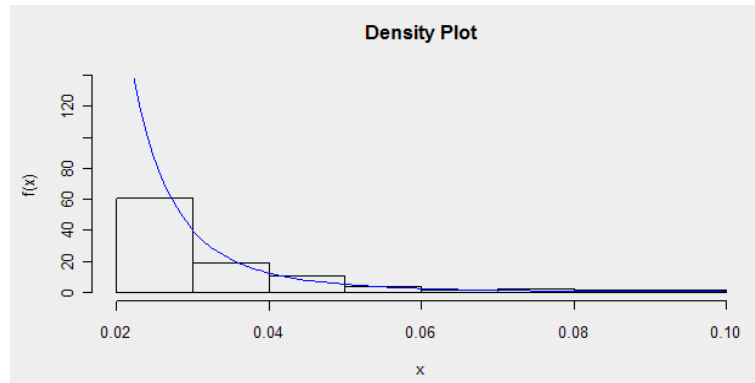


Figure 63: FTSE 100 Max Return Density: Hill Estimator - Pareto Distribution

In now analyzing the LR (Likelihood Ratio test) results for the NIKKEI 225 monthly maximum returns (Table 19), we do not reject the null hypothesis. Therefore, we can assume that the series follow an exponential distribution.

NIKKEI 225 -Max Return	Likelihood Ratio Test	p-value	Negative Log-likelihood
<b>Threshold = 0.0375</b>	3.1614	0.0754	-159.4130

Table 19: NIKKEI 225 LR Test and Log-Likelihood Results – GPD (Hill Estimator)

Analyzing the monthly maximum return series of the NIKKEI 225 index, using the MLE, the estimates for the parameters are presented next:

NIKKEI 225 -Max Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.0375</b>	50		
Scale - $\sigma$		0.0119	0.0025
Shape - $\xi$		0.2466	0.1705

(a) Nr. Exceedances of threshold

Table 20: Maximum NIKKEI 225 Parameters - GPD (Hill Estimator)

As we presented above for the NIKKEI 225 monthly maximum return, we do not reject the null hypothesis in the LR test. Consequently, the series do follow an exponential distribution, with a positive estimate parameter shape.

## How to deal with Extreme Observations in Empirical Finance

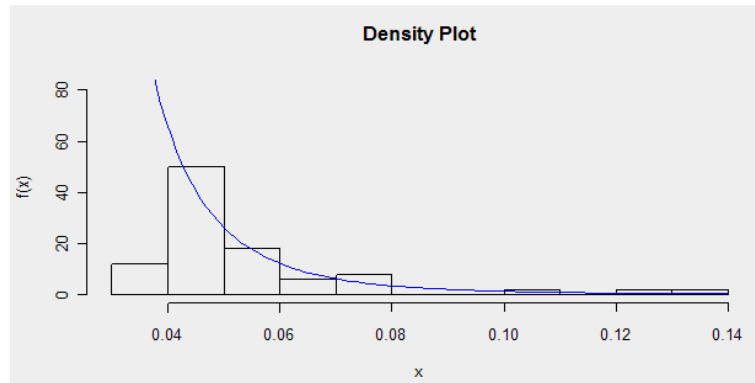


Figure 64: NIKKEI 225 Max Return Density: Hill Estimator - Exponential Distribution

To finish the GPD distribution application, we now analyze the LR (Likelihood Ratio test) results for the S&P 500 monthly maximum returns using the non-parametric approach and calculating the Hill estimator. Since we reject the null hypothesis for the S&P500 maximum series, the series do not follow an exponential distribution.

S&P 500 -Max Return	Likelihood Ratio Test	p-value	Negative Log-likelihood
<b>Threshold = 0.0204</b>	6.1894	0.0129	-646.8429

**Table 21:** S&P 500 LR Test and Log-Likelihood Results – GPD (Hill Estimator)

Analyzing the monthly maximum return series of the S&P 500 index, using the MLE, the estimates for the parameters are presented next:

S&P 500 -Max Return	Nr. Exce. (a)	MLE	Std. Err.
<b>Threshold = 0.0204</b>	176		
Scale - $\sigma$		0.0078	0.0009
Shape - $\xi$		0.1830	0.0009

(a) Nr. Exceedances of threshold

**Table 22:** Maximum S&P 500 Parameters - GPD (Hill Estimator)

In looking more specifically at estimated parameters, we can assume that the returns have been generated by a Pareto distribution, as the shape estimate ( $\xi$ ) is higher than zero, (figure 65).

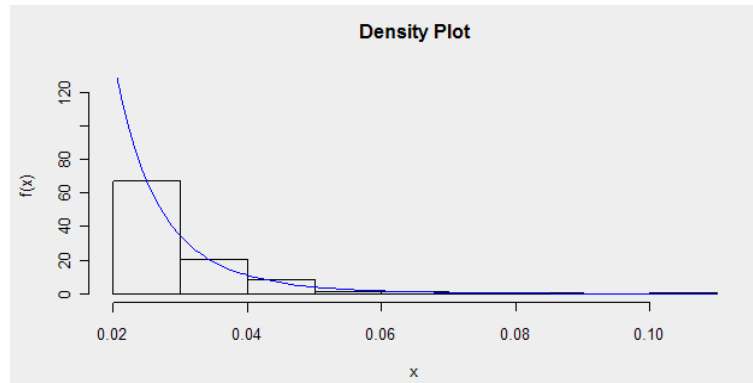


Figure 65: S&P 500 Max Return Density: Hill Estimator - Pareto Distribution

After applying the GPD distribution to the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, using the Hill estimator, for the positive tail, we can assume, and in accordance to the shape estimates, that the maximum return series are well represented mostly by the Pareto distribution. The only exception was in the NIKKEI 225 monthly maximum returns which are well represented by the exponential distribution.

This is a result that we were expecting, since the Pareto distribution is the most common distribution for the financial series, as we can see in other studies in financial research related to maximum and minimum returns.

This is a difference to the first approach where we use random thresholds of 1%, 2% and 5%, and where the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices are well represented mostly by the exponential distribution.

## 5. Conclusions

In this thesis, EVT methods were used to investigate and fit the empirical distribution of the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices to the theoretical GEV and GPD distributions. We have applied two approaches of extreme value theory to our data, the *Block Maxima* and the *POT* approach, as well as the MLE method's parametric approach for the distribution parameter estimation, and the Hill estimator's non-parametric approach.

The goal of this thesis was to fit the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices to the GEV and GPD distributions.

For the application of the methodology we used a diversity of tools such as the R Programming Language with an extRemes toolkit (the outputs can be found in the appendix), as well as the Easy Fit 5.5 Professional Software and the EViews Software.

After fitting the GEV distribution to all our maximum and minimum series, we conclude that the maximum return series are well represented by the Fréchet distribution. This was already expected due to the same conclusion reached by other studies in financial research related to maximum and minimum returns, as financial series normally follow this distribution. However, for the minimum returns we conclude that the series was well represented by the Weibull distribution instead of the Fréchet distribution.

After applying the GPD distribution to our data, using the POT approach with a random threshold level of 1%, 2.5% and 5%, for the positive (maximum returns) and negative (minimum returns) tail, another conclusion was the fact that we can assume that the maximum and minimum return series are well represented mostly by the exponential distribution or by the Beta distribution. The only exception was in the FTSE 100 monthly maximum return with a threshold of 2.5%, which is well represented by the Pareto distribution.

This was a result that we were not expecting, since other studies in financial research related to maximum and minimum returns conclude that the considered series was well represented by the Pareto distribution. Because of this difference, our result is one that could be studied in more detail in future research. However, one of the possible reasons for this result was the random threshold that we chose. The value of the threshold assumes a huge importance, since choosing a high value for the threshold leads to few extreme observations and implies

inefficient parameter estimates. On the other hand, choosing a low value for the threshold level leads to many observations of extreme observations, but induces biased parameter estimates as observations not belonging to the tails are included in the estimation process.

Due to this reason, we applied the GPD distribution to the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices, using the Hill estimator, for the positive tail, and in accordance to the shape estimates. The maximum returns series are well represented mostly by the Pareto distribution. The only exception was in the NIKKEI 225 monthly maximum return which is well represented by the exponential distribution. This was slightly different from the first approach, where the maximum return series are well represented mostly by the exponential distribution.

In conclusion, we have seen that in the GEV distribution application, our data was well represented by the Fréchet and Weibull distributions. On the other hand, in the GPD distribution, using the parametric MLE approach, we have seen that our data was mostly well represented by the exponential and Beta distributions. However, applying the GPD using the non-parametric Hill estimator approach for the tail index, we have seen that the monthly maximum returns of the FTSE 100, NIKKEI 225 and S&P500 indices are well represented by the Pareto distribution.

## 6. Final Comments

In terms of future directions for our work and the EVT theory in general, it is possible to highlight some points. The continuation of this thesis includes the estimation of the extreme value index using different non-parametric approaches such as the Pickands estimator, the Dekkers and de Haan estimator, as well as the bootstrap method developed by Danileson & De Vries (1997), in which the asymptotic mean square error is minimized.

Beyond the calculation of the new non-parametric approaches, an interesting research project would be the study of the daily or monthly maximum and minimum returns of some indices using the Generalized Logistic (GL) distribution instead of the GEV and GPD distributions. Tolikas, K. and Brown, R.A, (2006), show that in the investigation of the asymptotic distribution of the lower tail for daily returns in the Athens Stock Exchange (ASE) over the period 1986 to 2001, the Generalized Logistic (GL) distribution provides an adequate description of the stochastic behavior of the ASE index extreme minima over the period studied. This research could be important, because the GL distribution is fatter tailed than the GEV and GPD distributions.

In terms of the future of the EVT, one method that can have importance and developments is the *Peaks-Over-Random-Threshold (PORT)* approach. As we do not have an accurate choice of the threshold level, this new method introduces a new approach to dealing with this problem. However, there is no unanimity robust algorithm or test that can define the optimal threshold for this approach.

In conclusion, it is possible that the future of the EVT could be involved in the development of the PORT approach, and by consequence, the discovery of the optimal threshold. On the other hand, the EVT could try finding other distributions beyond the GEV and GPD distributions, which could allow us to manage safety risks and better deal with extreme events such as financial crises.



## References

- Alvarado, E., Sandberg, D. V. and Pickford, S.G. (1998). Modeling Large Forest Fires as Extreme Events, *Journal of Northwest Science* 72;
- Alves, I.F. (1995). Estimation of the Tail Parameter in the Domain of Attraction of an Extremal Distribution, Extreme Value Theory and Application, *Journal of Statistical Planning and Inference* 45(1-2),143-173;
- Alves, I.F.(2001). A location invariant Hill-type estimator, *Extremes* 4:3,199-217;
- Alves, I.F. (2007). Acerca de teste estatísticos para valores extremos, *Boletim da Sociedade Portuguesa de Estatística, Primavera de 2007*, 20-26;
- Alves, I.F. and Neves, C. (2008). Testing Extreme Value Conditions – An overview and recent approaches, *REVSTAT - Statistical Journal* 6(1), 83-100;
- Baesley, W. (1977). A test for distinguishing between extreme value distributions, *Journal of Hydrology* 34(3-4), 377-381;
- Bautista, E.A. (2002). *A distribuição Generalizada de Valores Extremos no Estudo da Velocidade Máxima do Vento em Piracicaba, SP*. Tese de Mestrado em Agronomia, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo;
- Beirlant, J., Teugels, J. and Vynckier, P. (1997). Practical Analysis of Extreme Value, *Journal of Insurance: Mathematics and Economics* 21(1), 91-92;
- Beirlant J. and Matthys G. (2000). Adaptive Threshold Selection in Tail Index Estimation, *Working Paper*, Catholic University Leuven;
- Beirlant J. and Matthys G. (2000). Adaptive Threshold Selection in Tail Index Estimation, *Working Paper*, University Center for Statistics - Catholic University Leuven;
- Beirlant J. and Matthys G. (2003). Estimating the Extreme value index and High Quantiles with Exponential regression Models, *Working Paper*, Katholieke Universiteit Leuven;
- Beirlant J. Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. New York: John Wiley & Sons;
- Beirlant.J and et alli. (2006). A goodness-of-fit statistic for Pareto-type behaviour, *Journal of Computational and Applied Mathematics* 186, 99-116;
- Berman, S. (1964). Limit theorems for the maximum term in stationary sequences, *Annals of Mathematical Statistics* 35(2), 502-516;
- Bermudez, P. Z., Turkman, M.A. and Turkman, K.F. (2001). A predictive approach to tail probability estimation, *Extremes* 4, 295-314;
- Bermudez, P. Z., Turkman, M.A., Turkman, K.F. (2001). A predictive approach to tail probability estimation, *Extremes* 4(4), 295-314;

Brilhante, M.F. (2004). Exponentiality versus generalized Pareto – A Resistant and Robust Test, *REVSTAT - Statistical Journal* 2 (1), 1-14;

Byström, H.N.E (2005). Extreme value theory and extremely large electricity price changes, *International Review of Economics and Finance* 14, 41-55;

Castillo, E. (1988). *Extreme Value Theory in Engineering*. Academic Press, Boston;

Castillo, E. and Hadi, A. (1997). Fitting the Generalized Pareto distribution to Data, *Journal of the American Statistical Association*, 92 (440), 1609-1620;

Chavez – Demoulin, V. (1999). *Two Problems in Environmental Statistics: Capture-Recapture Analysis of smooth Extremal Models*. Ph.D. Thesis, Department of Mathematics, Swiss Federal Institute of Technology, Lausanne;

Chavez – Demoulin, V. and Embrechts, P. (2004). Smooth Extremal Models in Finance, *The Journal of Risk and Insurance* 71 (2), 183-199;

Chavez – Demoulin V., Davison, A.C. and McNeil, A.J. (2004). *Estimating value-at-risk for financial time series: an approach combining self-exciting processes and extreme value theory*, Working Paper, Swiss Federal Institute of Technology, Lausanne;

Coles, S.G. and Tawn, J.A. (1991). Modelling Extreme Multivariate Events, *Journal of the Royal Statistical Society* 53(2), 377-392;

Coles, S.G. and Tawn, J.A. (1994). Statistical Methods for Multivariate Extremes: An application to Structural Design, *Journal of the Royal Statistical Society* 43, 1-48;

Coles, S.G. and Tawn, J.A. (1996). Modelling Extremes of the Areal rainfall Process, *Journal of the Royal Statistical Society* 58(2), 329-347;

Coles, S.G. and Powell, E.A. (1996). Bayesian methods in extreme value modelling: review and new developments. *International Statistical Review* 64, 119-136;

Cotter, J. (2000). Margin Exceedences for European Stock Index Futures using Extreme Value, *Working Paper*, University College Dublin;

Danielson, J. and de Vries, C. (1997). Beyond the sample: Extreme Quantile and Probability Estimation, *Working Paper*, London School of Economics and Institute of Economics;

Danielson, J. and Morimoto, Y. (2000). Forecasting Extreme Financial Risk: A critical Analysis of Practical Methods for the Japanese Market, *Monetary and Economic Studies*;

Davidson, A. and Smith, R. (1990). Models for Exceedances over High Thresholds, *Journal of the Royal Statistical Society* 52 (3), 393-442;

de Haan, L., Jansen, D.W., Koedijk, K. and de Vries, C.G. (1994). *Safety first portfolio selection, extreme value theory and long run asset risks*. Extreme value theory and applications. (J. Galambos et al., Eds.), Dordrecht;

- de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An introduction*. Springer;
- Dekkers, A. and de Haan, L. (1989). On the Estimation of the Extreme-Value index and Large Quantile Estimation, *Annals of Statistics* 17(4), 1795-1832;
- Dekkers, A., Einmahl, J. and de Haan, L. (1989). A Moment Estimator for the Index of an Extreme-Value Distribution, *Annals of Statistics* 17(4), 1833-1855;
- Diebold, F., Schuermann, T. and Stroughair, J. (1988). Pitfalls and Opportunities in the Use of Extreme Value theory in Risk Management, *Warton Financial Institutions Center – The Warton School Working Paper*, University of Pennsylvania;
- Dodd, E.L. (1923). The Greatest and Least Variation under General Laws of Error, *Transactions of the American Mathematical Society* 25(4), 525-539;
- Drees, H. (1995). Refined Pickands estimators for the extreme value index, *Annals of Statistics* 23 (6), 2059-2080;
- Drees, H., Ferreira, A. and de Haan, L. (2000). On maximum likelihood estimation of the extreme value index, *Working Paper*, Saarland University;
- Drees, H., de Haan, L. and Li, D. (2006). Approximation to the tail empirical distribution function with application to testing extreme value conditions, *Journal of Statistical Planning and Inference* 136, 3498-3538;
- DuMouchel, W.H (1983). Estimating the stable index  $\alpha$  in order to measure tail thickness: A critique, *Annals of Statistics* 11, 1019-1031;
- Embrechts, P. And Schmidli, H. (1994). Modelling of Extremal Events in Insurance and Finance, *ZOR - Mathematical Methods of Operational Research* 39, 1-34;
- Embrechts, P., Klüppelberg, C. and Mikosh, T. (1997). *Modelling Extremal events for Insurance and Finance*. Springer;
- Falk, M., (1994). Efficiency of convex combinations of Pickands estimator of the extreme value index, *Journal of Nonparametric Statistics* 4(2), 133-147;
- Falk, M., Husler, J. and Reiss, R. (2004). *Laws of Small numbers: extremes and Rare Events*. Birkhäuser;
- Gençay, R and Selçuk, F (2005). Overnight borrowing, interest rates and extreme value theory, *European Economic Review* 50, 547-563;
- Gilli, M. and Küllezi, E. (2003). An Application of Extreme Value Theory for Measuring Risk, *Computational Economics* 27(1), 1-23;
- Gomes, M.I. and Martins, M.J. (2002). “Asymptotically unbiased” estimators for the tail index based on extremal estimation of the second order parameter, *Extremes* 5, 5-31;
- Gomes, M.I.(2007), Memorial da Escola, *Boletim da Sociedade Portuguesa* 37-51;

- Gumbel, E.J. (1958). *Statistics of Extremes*. Columbia University Press, New York;
- Hall, A. (2007). Não há caminhos, há que caminhar, *Boletim da Sociedade Portuguesa de Estatística*, Primavera de 2007, 15-19;
- Hosking, J. (1984). Modelling persistence in hydrological time series using fractional differencing, *Water Resources* 20, 1898 – 1908;
- Hosking, J. (1984). Testing whether the shape parameter is zero in the generalized extreme – distribution, *Biometrika* 71(2), 367-374;
- Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985). Estimation of the Generalized Extreme-Value Distribution by the Method of Probability Weight Moments, *Technometrics* 27(3), 251-261;
- Hosking, J. and Wallis, J. (1987). Parameter and Quantile estimation for the Generalized Pareto Distribution, *Technometrics* 29(3), 339-349;
- Jenkinson, A..F. (1995). The frequency distribution of the annual maximum (or minimum) of meteorological elements, *Quarterly Journal of the Royal Meteorological Society* 81(348), 158-171;
- Klüppelberg, C. (2002). Risk Management with Extreme Value theory, *Working Paper*, Universität München;
- Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences, *Probability Theory and Related Fields* 65(2), 291-306;
- LeBaron, Blake and Samanta, Ritirupa (2005). Extreme value Theory and Fat Tails in Equity Markets, *Society for Computational Economics in its series computing in Economics and Finance* 2005,140;
- Lindskog, F. The Mathematics and Fundamental Ideas of Extreme Value Theory, *Working Paper*, Royal Institute of Technology, Stockholm.
- Longin.F (2004). The choice of the distribution of asset returns: How extreme value theory can help?, *Journal of Banking & Finance* 29, 1017-1035;
- Marohn, F. (1994). On testing the exponential and Gumbel distribution, *Working Paper*, Katholische Universität Eichstätt;
- Marohn, F. (2000). Testing Extreme Value Models. *Extremes* 3(4), 363-384;
- McNeil, A. J. and Saladin, T. (1997). The Peaks Over Threshold Method for estimating High Quantiles of Loss distributions, *Working Paper*, Department Mathematik, ETH Zentrum;
- McNeil, A. J. (1997). Estimating the Tails of Loss severity Distributions using Extreme Value theory. *ASTIN Bulletin* 27, 117-137;

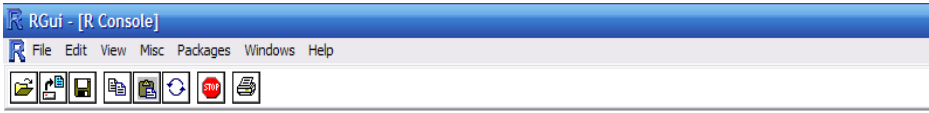
- McNeil, A. J. (1998). Calculating Quantile Risk Measures for Financial return series Using Extreme value Theory, *Working Paper*, Department Mathematik, ETH Zentrum;
- McNeil, A. J and Frey, R. (2000). Estimation of tail-related risk measures for heteroskedastic financial time series: An Extreme Value Approach, *Journal of Empirical Finance* 7(3-4), 271-300;
- Muller, U.A., Dacorogna, M. M. and Pictet, O.V. (1998). *Heavy Tails in high-frequency financial data*. In R.J. Adler, R.E. Feldman and M.S., Taqqu (Eds.), *A practical guide to heavy tails: Statistical techniques and applications*. Birkhäuser;
- Oliveira.J.T (1990). Perspectivas sobre a Estatística de Extremos. Resultados básicos e probelamas em aberto, *Working Paper*, Conferência expuesta en la Academia;
- Otten, A. and Montfort, M. (1978). The power of two tests on the type of distribution of extremes. *Journal of Hydrology* 37(1-2), 195-199;
- Pickands III, J. (1975).Statistical Inference Using Extreme Order Statistics. *Annals of Statistics* 3(1), 119-131.
- Pictet, O. V.,Dacorogna, M. M. and Muller, U.A. (1998). Hill, bootstrap and jackknife estimators for heavy tails, *A practical guide to heavy tails: Statistical Techniques and applications*, 283-310.
- Rachev, S. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. Wiley;
- Rachev, S.T., Mittnik, S., Fabozzi, F.J., Focardi, S.M. and Jasic,T. (2007). *Financial Econometrics: From Basics to Advanced Modeling Techniques*. Wiley;
- Reiss, R.D., Thomas, M. and Reiss,R.D. (2001). *Statistical Analysis of extreme Values:From Insurance, Hydrology and Other Fields*. Birkhäuser Verlag;
- Reiss, R.D.and Cormann, U. (2008). An example of Real-Life Data where the Hill estimator is correct. *Statistics for Industry and Technology* 5, 209-216;
- Resnick, S.I. (1997). Discussion of the Danish Data on Large Fire Insurance. *ASTIN Bulletin* 27, 139-151;
- Resnick, S.I. (1997). Heavy Tail Modelling and Teletraffic Data. *Annals of Statistics* 25 (5), 1805-1869;
- Rios, A.S. (2008). *Extreme Value Theory: An Application to Capital Markets*. Master Thesis in Finance, ISCTE Business School;
- Rootzén, H. and Tajvidi, N. (1997). Extreme Value Statistics and Wind Storm Losses: A Case Study, *Journal of Insurance: Mathematics and Economics* 20(3), 257;
- Segers, J.(2002), Generalized Pickands estimators for the Extreme Value Index, *Working Paper*, Chalmers University of Technology, Gothenburg;

- Smith, R.L. (1985). Maximum likelihood estimation in a class of non-regular cases, *Biometrika* 72, 67-90;
- Smith, R.L. (1987). Estimating Tails of Probability Distributions, *Annals of Statistics* 15 (3), 1174-1207;
- Smith, R. and Weissman, I. (1994). Estimating the Extremal Index, *Journal of the Royal Statistical Society* 56(3), 515-528;
- Stephens, M.A. (1976). Asymptotic results for goodness-of-fit Statistics with unknown parameters, *Annals of Statistics* 4 (2), 357-369;
- Stephens, M.A. (1977). Goodness-of-fit for the extreme value distribution, *Biometrika* 64(3), 583-588;
- Stephens, M.A. (1986). *Tests for the exponential distribution*. Goodness-of-fit Techniques, Marcel Dekker - New York;
- Tiago de Oliveira, J. (1984). *Bivariate extremes: Statistical choice*. In Tiago de Oliveira, J., ed. *Statistical Extremes and Applications* 91-107, Dordrecht;
- Tolikas, K. and Brown, R.A. (2006). The Distribution of the Extreme daily Share Returns in the Athens Stock Exchange, *The European Journal of Finance* 12(1), 1-22;
- Van Montfort, M.A. (1970). On testing that the distribution of extremes is of type I when type II is the alternative, *Journal of Hydrology* 11(4), 421-427;
- Van Montfort, M.A. and Witter, J. V. (1985). Testing exponential against Generalized Pareto distribution, *Journal of Hydrology* 78, 305-315;
- Yun, S. (2000). A class of Pickands-type estimators for the extreme value index, *Journal of Statistical Planning and Inference* 83, 113-124;
- Yun, S. (2002). On a Generalized Pickands estimator of the extreme value index, *Journal of Statistical Planning and Inference* 102(2), 389-409.

## Appendix

In the figures below it is possible to find the outputs from the R Programming Language, which we obtained from our calculations that support all the conclusions.

The next figure shows the R Programming Language screen and the extremes toolkit search in the program's library.



```
R version 2.13.0 (2011-04-13)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

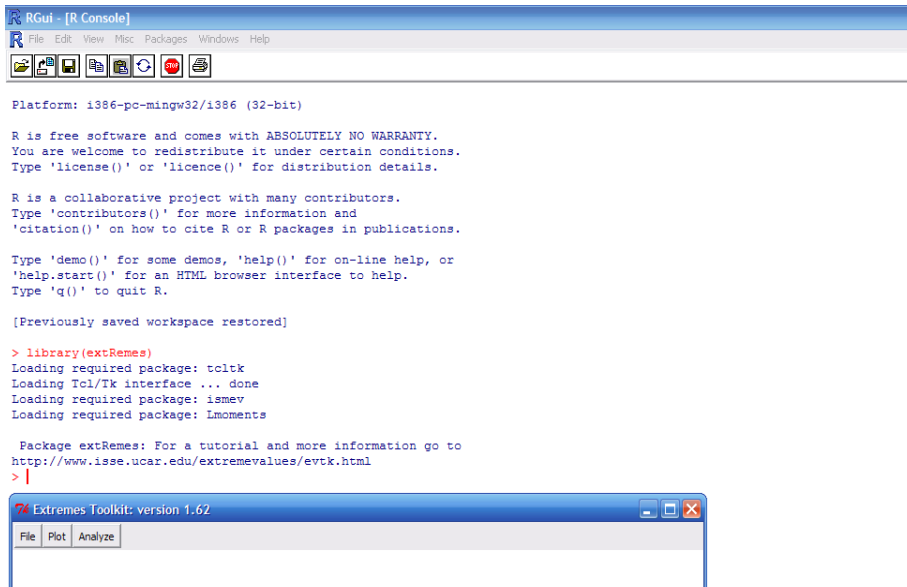
[Previously saved workspace restored]

> library(extRemes)
Loading required package: tcltk
Loading Tcl/Tk interface ... done
Loading required package: ismev
Loading required package: Lmoments

Package extRemes: For a tutorial and more information go to
http://www.isse.ucar.edu/extremevalues/evtk.html
> |
```

## How to deal with Extreme Observations in Empirical Finance

The next figure shows the R Programming Language screen and the extremes toolkit opened:



The next figures show the data analysis outputs from our data, the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices.

```
[1] "Successfully opened file: Daily FTSE.txt"
      Daily
N      6.807000e+03
mean   2.436756e-04
Std.Dev. 1.122324e-02
min    -1.303000e-01
Q1     -5.400000e-03
median  6.000000e-04
Q3     6.400000e-03
max    9.380000e-02
missing values 0.000000e+00

Saving workspace (may take a few moments for large workspaces) ...
```



## How to deal with Extreme Observations in Empirical Finance

```
[1] "Successfully opened file: Max and Min FTSE.txt"
```

	Max	Min
N	324.00000000	324.00000000
mean	0.01890278	-0.01907191
Std.Dev.	0.01166791	0.01260198
min	0.00570000	-0.13030000
Q1	0.01220000	-0.02340000
median	0.01500000	-0.01615000
Q3	0.02255000	-0.01140000
max	0.09380000	-0.00450000
missing values	0.00000000	0.00000000

```
Saving workspace (may take a few moments for large workspaces) ...
```

```
Workspace saved.
```

```
[1] "Successfully opened file: daily return.txt"
```

	DailyReturnNIKKEI225
N	6.684000e+03
mean	4.937163e-06
Std.Dev.	1.461738e-02
min	-1.614000e-01
Q1	-6.900000e-03
median	4.000000e-04
Q3	7.300000e-03
max	1.323000e-01
missing values	0.000000e+00

```
Saving workspace (may take a few moments for large workspaces) ...
```

```
Workspace saved.
```

```
[1] "Successfully opened file: Max and MIN.txt"
```

	Max	Min
N	327.00000000	327.00000000
mean	0.02563211	-0.02534832
Std.Dev.	0.01584578	0.01636964
min	0.00620000	-0.16140000
Q1	0.01525000	-0.03120000
median	0.02180000	-0.02310000
Q3	0.03245000	-0.01440000
max	0.13230000	-0.00360000
missing values	0.00000000	0.00000000

```
Saving workspace (may take a few moments for large workspaces) ...
```

```
Workspace saved.
```

## How to deal with Extreme Observations in Empirical Finance

```
[1] "Successfully opened file: daily.txt"
      Daily
N          1.539600e+04
mean       2.827358e-04
Std.Dev.   9.716366e-03
min        -2.290000e-01
Q1         -4.100000e-03
median     5.000000e-04
Q3         4.925000e-03
max        1.096000e-01
missing values 0.000000e+00

Saving workspace (may take a few moments for large workspaces) ...

Workspace saved.
[1] "Successfully opened file: max and min.txt"
      Max      Min
N          7.350000e+02 735.00000000
mean       1.672476e-02 -0.01654558
Std.Dev.   9.984292e-03  0.01358781
min        2.900000e-03 -0.22900000
Q1         1.030000e-02 -0.01925000
median     1.440000e-02 -0.01370000
Q3         2.015000e-02 -0.00960000
max        1.096000e-01 -0.00180000
missing values 0.000000e+00 0.00000000

Saving workspace (may take a few moments for large workspaces) ...

Workspace saved.
```

The next figures show the GEV and GPD distribution outputs from our data, the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices.

## How to deal with Extreme Observations in Empirical Finance

```
GEV fit
-----
Response variable: Max

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): 0.01360707
Scale (sigma): 0.005448654
Shape (xi): 0.288852

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 65.42049 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 6.050651e-16

Convergence successful![1] "Convergence successful!"
[1] "Maximum Likelihood Estimates:"
              MLE Stand. Err.
MU: (identity) 0.01356      0.00031
SIGMA: (identity) 0.00535    0.00018
Xi: (identity) 0.31464      0.05105

[1] "Negative log-likelihood: -1125.0290232532"

Parameter covariance:
              [,1]      [,2]      [,3]
[1,] 9.747275e-08 3.140148e-08 -4.857693e-06
[2,] 3.140148e-08 3.168810e-08 1.083268e-06
[3,] -4.857693e-06 1.083268e-06 2.606549e-03
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit1

```

.....

```
GEV fit
-----
Response variable: Min

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): -0.01937543
Scale (sigma): 0.01174781
Shape (xi): -0.9390378

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 603.535 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 2.850483e-133

Convergence successful![1] "Convergence successful!"
[1] "Maximum Likelihood Estimates:"
              MLE Stand. Err.
MU: (identity) -0.02013      0
SIGMA: (identity) 0.01254     0
Xi: (identity) -0.80147      0

[1] "Negative log-likelihood: -1064.76115905691"

Parameter covariance:
              [,1]      [,2]      [,3]
[1,] 3.995622e-12 -1.123522e-16 -1.549115e-18
[2,] -1.123522e-16 3.995607e-12 -1.584788e-18
[3,] -1.549115e-18 -1.584788e-18 3.995745e-12
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit2
```

## How to deal with Extreme Observations in Empirical Finance

```
*****

GEV fit
-----
Response variable: Max

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): 0.01843403
Scale (sigma): 0.009063659
Shape (xi): 0.1817563

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 30.51234 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 3.317504e-08

Convergence successfull![1] "Convergence successfull!"
[1] "Maximum Likelihood Estimates:"
      MLE Stand. Err.
MU: (identity) 0.01826 0.00056
SIGMA: (identity) 0.00871 0.00041
Xi: (identity) 0.22506 0.05083

[1] "Negative log-likelihood: -993.174705332302"

Parameter covariance:
      [,1] [,2] [,3]
[1,] 3.085523e-07 1.375824e-07 -9.736467e-06
[2,] 1.375824e-07 1.695890e-07 -2.162541e-06
[3,] -9.736467e-06 -2.162541e-06 2.584132e-03
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit1
. |

*****

GEV fit
-----
Response variable: Min

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): -0.02704945
Scale (sigma): 0.01612951
Shape (xi): -0.7554838

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 512.9549 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 1.443135e-113

Convergence successfull![1] "Convergence successfull!"
[1] "Maximum Likelihood Estimates:"
      MLE Stand. Err.
MU: (identity) -0.02781 0.00000
SIGMA: (identity) 0.01655 0.00000
Xi: (identity) -0.67988 0.00513

[1] "Negative log-likelihood: -956.077570067717"

Parameter covariance:
      [,1] [,2] [,3]
[1,] 3.996134e-12 -4.377946e-17 -9.089685e-12
[2,] -4.377946e-17 3.996115e-12 -9.442212e-12
[3,] -9.089685e-12 -9.442212e-12 2.632304e-05
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit2
> |
```

## How to deal with Extreme Observations in Empirical Finance

```
*****

GEV fit
-----
Response variable: Max

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): 0.01219956
Scale (sigma): 0.005685543
Shape (xi): 0.1829611

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 56.83743 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 4.73371e-14

Convergence successful![1] "Convergence successful!"
[1] "Maximum Likelihood Estimates:"
      MLE Stand. Err.
MU: (identity) 0.01219 0.00023
SIGMA: (identity) 0.00566 0.00015
Xi: (identity) 0.18639 0.03008

[1] "Negative log-likelihood: -2564.36688672795"

Parameter covariance:
      [,1] [,2] [,3]
[1,] 5.446451e-08 1.848580e-08 -2.122074e-06
[2,] 1.848580e-08 2.135868e-08 -6.485434e-08
[3,] -2.122074e-06 -6.485434e-08 9.050537e-04
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit1

*****

GEV fit
-----
Response variable: Min

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments (stationary case) estimates (used to initialize MLE optimization routine):
Location (mu): -0.01650406
Scale (sigma): 0.01088485
Shape (xi): -1.009023

Likelihood ratio test (5% level) for xi=0 does not accept Gumbel hypothesis.
likelihood ratio statistic is 1995.898 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0

Convergence successful![1] "Convergence successful!"
[1] "Maximum Likelihood Estimates:"
      MLE Stand. Err.
MU: (identity) -0.01821 0
SIGMA: (identity) 0.01187 0
Xi: (identity) -0.72229 0

[1] "Negative log-likelihood: -2426.37652145714"

Parameter covariance:
      [,1] [,2] [,3]
[1,] 3.990077e-12 -2.129956e-16 -4.455367e-18
[2,] -2.129956e-16 3.990021e-12 -4.373545e-18
[3,] -4.455367e-18 -4.373545e-18 3.990315e-12
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gev.fit2
~|
```

## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.01041136

shape: 0.01448497

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 2.377294 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.1231107

Convergence successfull!

[1] "Threshold = 0.01"

[1] "Number of exceedances of threshold = 278"

[1] "Exceedance rate (per year)= 313.393518518519"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
--	-----	-----------

Scale (sigma):	0.009719459	0.0007778749
----------------	-------------	--------------

Shape (xi):	0.079239531	0.0568129494
-------------	-------------	--------------

[1] "Negative log-likelihood: -988.162736748929"

Parameter covariance:

	[,1]	[,2]
--	------	------

[1,]	6.050893e-07	-2.746508e-05
------	--------------	---------------

[2,]	-2.746508e-05	3.227711e-03
------	---------------	--------------

[1] "Convergence code (see help file for optim): 0"

NULL

Model name: gpd.fit1

Model name: gpd.fit1No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.00816764

shape: 0.3457766

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 4.528587 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.03333314

Convergence successfull!

[1] "Threshold = 0.025"

[1] "Number of exceedances of threshold = 58"

[1] "Exceedance rate (per year)= 65.3842592592593"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
--	-----	-----------

Scale (sigma):	0.008250072	0.002003806
----------------	-------------	-------------

Shape (xi):	0.375148817	0.219756858
-------------	-------------	-------------

[1] "Negative log-likelihood: -198.493882847962"

Parameter covariance:

	[,1]	[,2]
--	------	------

[1,]	4.015239e-06	-0.0003132063
------	--------------	---------------

[2,]	-3.132063e-04	0.0482930768
------	---------------	--------------

[1] "Convergence code (see help file for optim): 0"

NULL

Model name: gpd.fit2

## How to deal with Extreme Observations in Empirical Finance

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.02110654
```

```
shape: -0.07071844
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 12.08604 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 0.0005080073
```

```
Convergence successful!
```

```
[1] "Threshold = 0.05"
```

```
[1] "Number of exceedances of threshold = 8"
```

```
[1] "Exceedance rate (per year)= 9.01851851851852"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
MLE Std. Err.  
Scale (sigma): 0.0507568 1.999966e-06  
Shape (xi): -1.1588310 9.781363e-04
```

```
[1] "Negative log-likelihood: -29.4550391635451"
```

```
Parameter covariance:
```

```
 [,1] [,2]  
[1,] 3.999862e-12 -2.456208e-14  
[2,] -2.456208e-14 9.567506e-07
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

```
No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.00332619
```

```
shape: -0.595691
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 13.50665 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 0.0002377194
```

```
Convergence successful!
```

```
[1] "Threshold = -0.01"
```

```
[1] "Number of exceedances of threshold = 58"
```

```
[1] "Exceedance rate (per year)= 65.3842592592593"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
MLE Std. Err.  
Scale (sigma): 0.003266919 1.999679e-06  
Shape (xi): -0.565696117 3.165639e-02
```

```
[1] "Negative log-likelihood: -306.800928401117"
```

```
Parameter covariance:
```

```
 [,1] [,2]  
[1,] 3.998717e-12 -5.124668e-11  
[2,] -5.124668e-11 1.002127e-03
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.0270818  
shape: -1.56926  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 174.2542 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 8.710997e-40

Convergence successfull!

```
[1] "Threshold = -0.025"
[1] "Number of exceedances of threshold = 258"
[1] "Exceedance rate (per year)= 290.847222222222"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma):  0.01895177 1.998975e-06
Shape (xi):     -0.92413744 1.998997e-06
```

```
[1] "Negative log-likelihood: -1003.67508843623"
```

```
Parameter covariance:
      [,1]      [,2]
[1,] 3.995902e-12 -1.628196e-18
[2,] -1.628196e-18 3.995988e-12
[1] "Convergence code (see help file for optim): 0"
NULL
```

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.02077799  
shape: -0.237387  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 0.7986209 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.371506

Convergence successfull!

```
[1] "Threshold = 0.01"
[1] "Number of exceedances of threshold = 306"
[1] "Exceedance rate (per year)= 341.793577981651"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma):  0.01738846 0.001201076
Shape (xi):     -0.03742295 0.040330461
```

```
[1] "Negative log-likelihood: -944.979332285516"
```

```
Parameter covariance:
      [,1]      [,2]
[1,] 1.442584e-06 -2.994523e-05
[2,] -2.994523e-05 1.626546e-03
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gpd.fit1
```



## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.01464452  
shape: 0.008606658  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 1.402498 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.2363058

Convergence successful!

```
[1] "Threshold = 0.025"
[1] "Number of exceedances of threshold = 127"
[1] "Exceedance rate (per year)= 141.855504587156"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma): 0.01347421 0.001596997
Shape (xi):    0.08621541 0.081643905
```

```
[1] "Negative log-likelihood: -409.0119938077"
```

```
Parameter covariance:
      [,1] [,2]
[1,] 2.550401e-06 -7.860838e-05
[2,] -7.860838e-05 6.665727e-03
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gpd.fit2
```

> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.01242843  
shape: 0.3793952  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 1.343264 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.2464593

Convergence successful!

```
[1] "Threshold = 0.05"
[1] "Number of exceedances of threshold = 19"
[1] "Exceedance rate (per year)= 21.2224770642202"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma): 0.01335175 0.005745437
Shape (xi):    0.37010576 0.380106798
```

```
[1] "Negative log-likelihood: -55.9750852868811"
```

```
Parameter covariance:
      [,1] [,2]
[1,] 3.301005e-05 -0.001551109
[2,] -1.551109e-03 0.144481178
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gpd.fit3
```

## How to deal with Extreme Observations in Empirical Finance

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.006057998
```

```
shape: -1.039938
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 15.40126 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 8.693051e-05
```

```
Convergence successful!
```

```
[1] "Threshold = -0.01"
```

```
[1] "Number of exceedances of threshold = 33"
```

```
[1] "Exceedance rate (per year)= 36.8600917431193"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
                MLE      Std. Err.  
Scale (sigma):  0.006021363 1.999825e-06  
Shape (xi):     -0.935921166 1.973908e-02
```

```
[1] "Negative log-likelihood: -166.737374806797"
```

```
Parameter covariance:
```

```
          [,1]      [,2]  
[1,] 3.999299e-12 -1.544095e-11  
[2,] -1.544095e-11 3.896314e-04
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

```
No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.01567376
```

```
shape: -0.6965466
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 72.66044 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 1.539899e-17
```

```
Convergence successful!
```

```
[1] "Threshold = -0.025"
```

```
[1] "Number of exceedances of threshold = 189"
```

```
[1] "Exceedance rate (per year)= 211.107798165138"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
                MLE      Std. Err.  
Scale (sigma):  0.01618281 1.999257e-06  
Shape (xi):     -0.75273357 5.486682e-03
```

```
[1] "Negative log-likelihood: -732.674693605637"
```

```
Parameter covariance:
```

```
          [,1]      [,2]  
[1,] 3.997027e-12 -6.396114e-12  
[2,] -6.396114e-12 3.010367e-05
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.09983507

shape: -2.616568

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 207.2326 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 5.516484e-47

Convergence successfull!

[1] "Threshold = -0.05"

[1] "Number of exceedances of threshold = 305"

[1] "Exceedance rate (per year)= 340.676605504587"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
Scale (sigma):	0.1077104	1.999073e-06
Shape (xi):	-2.3213459	1.999090e-06

[1] "Negative log-likelihood: -893.493484242828"

Parameter covariance:

	[,1]	[,2]
[1,]	3.996292e-12	-2.565814e-18
[2,]	-2.565814e-18	3.996359e-12

[1] "Convergence code (see help file for optim): 0"

NULL

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.01005926

shape: -0.06427147

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 0.7078934 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.400144

Convergence successfull!

[1] "Threshold = 0.01"

[1] "Number of exceedances of threshold = 562"

[1] "Exceedance rate (per year)= 279.279591836735"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
Scale (sigma):	0.009183139	0.000487185
Shape (xi):	0.028061343	0.035243100

[1] "Negative log-likelihood: -2058.14631860591"

Parameter covariance:

	[,1]	[,2]
[1,]	2.373493e-07	-1.054351e-05
[2,]	-1.054351e-05	1.242076e-03

[1] "Convergence code (see help file for optim): 0"

NULL

Model name: gpd.fit1

## How to deal with Extreme Observations in Empirical Finance

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:
scale: 0.009436275
shape: 0.1294785
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for xi=0 does not reject Exponential hypothesis.
likelihood ratio statistic is 1.968136 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.1606458

Convergence successfull!

[1] "Threshold = 0.025"

[1] "Number of exceedances of threshold = 98"

[1] "Exceedance rate (per year)= 48.7"

[1] "Maximum Likelihood Estimates:"
      MLE Std. Err.
Scale (sigma): 0.009437015 0.001347397
Shape (xi):    0.128494054 0.107773083

[1] "Negative log-likelihood: -346.388116887375"

Parameter covariance:
      [,1] [,2]
[1,] 1.815479e-06 -9.258213e-05
[2,] -9.258213e-05 1.161504e-02
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gpd.fit2

> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:
scale: 0.01252458
shape: 0.3764434
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for xi=0 does not reject Exponential hypothesis.
likelihood ratio statistic is 0.002331815 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.961486

Convergence successfull!

[1] "Threshold = 0.05"

[1] "Number of exceedances of threshold = 7"

[1] "Exceedance rate (per year)= 3.47857142857143"

[1] "Maximum Likelihood Estimates:"
      MLE Std. Err.
Scale (sigma): 0.02076846 0.01629644
Shape (xi):    -0.03387567 0.69410398

[1] "Negative log-likelihood: -20.3553910480236"

Parameter covariance:
      [,1] [,2]
[1,] 0.0002655738 -0.01003382
[2,] -0.0100338168 0.48178033
[1] "Convergence code (see help file for optim): 0"
NULL
Model name: gpd.fit3
```

## How to deal with Extreme Observations in Empirical Finance

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:
scale: 0.004738576
shape: -0.6759417
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.
likelihood ratio statistic is 41.5227 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 1.165102e-10

Convergence successfull!

[1] "Threshold = -0.01"

[1] "Number of exceedances of threshold = 197"

[1] "Exceedance rate (per year)= 97.8969387755102"

[1] "Maximum Likelihood Estimates:"
                MLE      Std. Err.
Scale (sigma): 0.004713081 1.998983e-06
Shape (xi):    -0.569394062 6.054497e-03

[1] "Negative log-likelihood: -979.834922305416"

Parameter covariance:
                [,1]      [,2]
[1,] 3.995932e-12 -6.451906e-12
[2,] -6.451906e-12 3.665693e-05
[1] "Convergence code (see help file for optim): 0"
NULL

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:
scale: 0.03615583
shape: -2.016073
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.
likelihood ratio statistic is 442.8829 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 2.552788e-98

Convergence successfull!

[1] "Threshold = -0.025"

[1] "Number of exceedances of threshold = 635"

[1] "Exceedance rate (per year)= 315.55612244898"

[1] "Maximum Likelihood Estimates:"
                MLE      Std. Err.
Scale (sigma): 0.02011531 1.997563e-06
Shape (xi):    -0.86682241 1.997608e-06

[1] "Negative log-likelihood: -2395.60067595695"

Parameter covariance:
                [,1]      [,2]
[1,] 3.990260e-12 -3.914302e-18
[2,] -3.914302e-18 3.990438e-12
[1] "Convergence code (see help file for optim): 0"
NULL
```

## How to deal with Extreme Observations in Empirical Finance

```
No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:
scale: 0.249699
shape: -6.166808
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.
likelihood ratio statistic is 1112.93 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 5.111826e-244

Convergence successful!

[1] "Threshold = -0.05"

[1] "Number of exceedances of threshold = 719"

[1] "Exceedance rate (per year)= 357.298979591837"

[1] "Maximum Likelihood Estimates:"
              MLE      Std. Err.
Scale (sigma): 0.07209925 1.997697e-06
Shape (xi):     -1.49583509 5.694931e-04

[1] "Negative log-likelihood: -2251.11908502505"

Parameter covariance:
              [,1]      [,2]
[1,] 3.990793e-12 -7.057877e-13
[2,] -7.057877e-13 3.243224e-07
[1] "Convergence code (see help file for optim): 0"
NULL
```

The next figure shows the code to obtain the Hill estimator:

```
> z<-scan("MaxReturnNIKKEI.txt")
Read 327 items
> z
 [1] 0.0091 0.0108 0.0163 0.0110 0.0225 0.0219 0.0134 0.0191 0.0095 0.0213 0.0110 0.0116 0.0099 0.0075 0.0099 0.0147 0.0111 0.0103 0.0117 0.0222
[22] 0.0129 0.0210 0.0094 0.0076 0.0083 0.0338 0.0174 0.0179 0.0112 0.0125 0.0243 0.0247 0.0304 0.0146 0.0144 0.0193 0.0173 0.0248 0.0260 0.0175
[43] 0.0386 0.0147 0.0291 0.1078 0.0410 0.0171 0.0548 0.0100 0.0140 0.0139 0.0106 0.0104 0.0158 0.0147 0.0091 0.0087 0.0115 0.0128 0.0154 0.0103
[64] 0.0133 0.0062 0.0066 0.0130 0.0078 0.0137 0.0152 0.0108 0.0100 0.0125 0.0203 0.0472 0.0375 0.0256 0.0232 0.0193 0.0525 0.0456 0.1243 0.0444
[85] 0.0438 0.0344 0.0180 0.0199 0.0144 0.0132 0.0345 0.0225 0.0265 0.0191 0.0178 0.0469 0.0349 0.0247 0.0211 0.0728 0.0327 0.0291 0.0315 0.0603
[106] 0.0162 0.0480 0.0138 0.0329 0.0157 0.0503 0.0422 0.0127 0.0184 0.0311 0.0116 0.0157 0.0114 0.0202 0.0428 0.0755 0.0226 0.0202 0.0228 0.0148
[127] 0.0099 0.0190 0.0144 0.0135 0.0079 0.0161 0.0352 0.0144 0.0358 0.0278 0.0164 0.0225 0.0608 0.0397 0.0367 0.0225 0.0313 0.0179 0.0371 0.0102
[148] 0.0148 0.0348 0.0131 0.0122 0.0140 0.0222 0.0189 0.0205 0.0200 0.0460 0.0241 0.0217 0.0254 0.0336 0.0189 0.0177 0.0152 0.0272 0.0328 0.0766
[169] 0.0593 0.0198 0.0288 0.0246 0.0153 0.0430 0.0331 0.0225 0.0518 0.0599 0.0404 0.0101 0.0247 0.0170 0.0489 0.0305 0.0352 0.0285 0.0187 0.0222
[190] 0.0299 0.0196 0.0209 0.0355 0.0125 0.0162 0.0190 0.0278 0.0236 0.0116 0.0255 0.0205 0.0258 0.0353 0.0423 0.0150 0.0214 0.0722 0.0430 0.0346
[211] 0.0330 0.0377 0.0403 0.0376 0.0392 0.0308 0.0307 0.0458 0.0574 0.0186 0.0249 0.0345 0.0343 0.0345 0.0322 0.0354 0.0333 0.0231 0.0207 0.0209
[232] 0.0290 0.0221 0.0216 0.0333 0.0195 0.0311 0.0239 0.0269 0.0311 0.0178 0.0207 0.0206 0.0159 0.0237 0.0276 0.0186 0.0131 0.0199 0.0265 0.0187
[253] 0.0093 0.0122 0.0089 0.0116 0.0221 0.0128 0.0094 0.0165 0.0176 0.0246 0.0190 0.0215 0.0352 0.0292 0.0259 0.0159 0.0196 0.0330 0.0303 0.0203
[274] 0.0225 0.0165 0.0139 0.0129 0.0122 0.0192 0.0172 0.0162 0.0114 0.0141 0.0296 0.0360 0.0162 0.0244 0.0192 0.0402 0.0418 0.0245 0.0412 0.0298
[295] 0.0293 0.0260 0.0369 0.1323 0.0608 0.0508 0.0481 0.0269 0.0503 0.0431 0.0445 0.0212 0.0270 0.0329 0.0194 0.0185 0.0287 0.0377 0.0160 0.0270
[316] 0.0229 0.0216 0.0319 0.0272 0.0174 0.0231 0.0190 0.0282 0.0179 0.0164 0.0176 0.0122

> par(mfrow = c(1, 1))
> hillPlot(z, plottype = "xi")
> grid()
> |
```

The next figures show the GEV and GPD distribution outputs from our data, the monthly maximum and minimum return series of the FTSE 100, NIKKEI 225 and S&P500 indices, using the threshold obtained from the Hill estimator:

## How to deal with Extreme Observations in Empirical Finance

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.007195408
```

```
shape: 0.3444531
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 8.431742 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 0.003687269
```

```
Convergence successfull!
```

```
[1] "Threshold = 0.0222"
```

```
[1] "Number of exceedances of threshold = 84"
```

```
[1] "Exceedance rate (per year)= 94.69444444444444"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
MLE Std. Err.
```

```
Scale (sigma): 0.007225312 0.00131235
```

```
Shape (xi): 0.367857609 0.16053551
```

```
[1] "Negative log-likelihood: -299.226126921109"
```

```
Parameter covariance:
```

```
[,1] [,2]
```

```
[1,] 1.722264e-06 -0.0001359432
```

```
[2,] -1.359432e-04 0.0257716503
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

```
Model name: gpd.fit25
```

```
> |
```

```
> No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)
```

```
L-moments estimates for (stationary) GPD are:
```

```
scale: 0.02285673
```

```
shape: -1.50283
```

```
These L-moments estimators were used as initial parameter estimates.
```

```
Likelihood ratio test (5% level) for xi=0 does not accept Exponential hypothesis.  
likelihood ratio statistic is 148.6803 > 3.841459 1 df chi-square critical value.
```

```
p-value for likelihood-ratio test is 3.368351e-34
```

```
Convergence successfull!
```

```
[1] "Threshold = -0.0228"
```

```
[1] "Number of exceedances of threshold = 238"
```

```
[1] "Exceedance rate (per year)= 268.300925925926"
```

```
[1] "Maximum Likelihood Estimates:"
```

```
MLE Std. Err.
```

```
Scale (sigma): 0.0159878 1.999026e-06
```

```
Shape (xi): -0.8726928 1.999047e-06
```

```
[1] "Negative log-likelihood: -953.971955724593"
```

```
Parameter covariance:
```

```
[,1] [,2]
```

```
[1,] 3.996103e-12 -1.398494e-18
```

```
[2,] -1.398494e-18 3.996187e-12
```

```
[1] "Convergence code (see help file for optim): 0"
```

```
NULL
```

## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.01166101

shape: 0.2553635

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not reject Exponential hypothesis.  
likelihood ratio statistic is 3.161399 < 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.07539856

Convergence successful!

[1] "Threshold = 0.0375"

[1] "Number of exceedances of threshold = 50"

[1] "Exceedance rate (per year)= 55.848623853211"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
Scale (sigma):	0.01185576	0.002548933
Shape (xi):	0.24657163	0.170525589

[1] "Negative log-likelihood: -159.412978790928"

Parameter covariance:

	[,1]	[,2]
[1,]	6.497058e-06	-0.0002666448
[2,]	-2.666448e-04	0.0290789766

[1] "Convergence code (see help file for optim): 0"

NULL

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:

scale: 0.01649261

shape: -0.8315653

These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 73.07882 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 1.245741e-17

Convergence successful!

[1] "Threshold = -0.0239"

[1] "Number of exceedances of threshold = 172"

[1] "Exceedance rate (per year)= 192.119266055046"

[1] "Maximum Likelihood Estimates:"

	MLE	Std. Err.
Scale (sigma):	0.01527037	1.999315e-06
Shape (xi):	-0.74759489	7.540978e-03

[1] "Negative log-likelihood: -674.661825328779"

Parameter covariance:

	[,1]	[,2]
[1,]	3.997262e-12	-1.127541e-11
[2,]	-1.127541e-11	5.686636e-05

[1] "Convergence code (see help file for optim): 0"

NULL



## How to deal with Extreme Observations in Empirical Finance

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.007632976  
shape: 0.1955666  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 6.189374 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 0.01285197

Convergence successful!

```
[1] "Threshold = 0.0204"
[1] "Number of exceedances of threshold = 176"
[1] "Exceedance rate (per year)= 87.4612244897959"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma): 0.007762864 0.0008712303
Shape (xi):    0.182989037 0.0907805931
```

```
[1] "Negative log-likelihood: -646.842911626347"
```

```
Parameter covariance:
      [,1]      [,2]
[1,] 7.590422e-07 -5.190823e-05
[2,] -5.190823e-05 8.241116e-03
[1] "Convergence code (see help file for optim): 0"
NULL
```

No optimization method selected. Using "Nelder-Mead" (use 'help( optim)' for more details)

L-moments estimates for (stationary) GPD are:  
scale: 0.01291066  
shape: -1.079004  
These L-moments estimators were used as initial parameter estimates.

Likelihood ratio test (5% level) for  $\xi=0$  does not accept Exponential hypothesis.  
likelihood ratio statistic is 187.9828 > 3.841459 1 df chi-square critical value.

p-value for likelihood-ratio test is 8.763125e-43

Convergence successful!

```
[1] "Threshold = -0.0169"
[1] "Number of exceedances of threshold = 479"
[1] "Exceedance rate (per year)= 238.033673469388"
[1] "Maximum Likelihood Estimates:"
      MLE      Std. Err.
Scale (sigma): 0.009669169 1.997911e-06
Shape (xi):    -0.638196944 1.873320e-03
```

```
[1] "Negative log-likelihood: -2049.07348841385"
```

```
Parameter covariance:
      [,1]      [,2]
[1,] 3.99165e-12 -1.81709e-12
[2,] -1.81709e-12 3.50933e-06
[1] "Convergence code (see help file for optim): 0"
NULL
```