# TUM

## Technische Universität München

## Department of Mathematics

Bachelor's Thesis

# Value-at-Risk Estimation
# in the Basel III Framework

Johanna Mager

Supervisor: Prof. Dr. Claudia Klüppelberg

Advisor: Prof. Dr. Claudia Klüppelberg

Submission Date: 2012-09-01

I hereby declare that I have written the Bachelor's Thesis on my own and have used no other than the stated sources and aids.

Garching,

# Zusammenfassung

In dieser Bachelorarbeit wird die Value-at-Risk (VaR) Schätzung im Basel III Market Risk Framework, basierend auf verschiedenen Extremwert Methoden analysiert. Neben dem VaR wird in Basel III nun zusätzlich der so genannte gestresste VaR (SVaR) eingeführt, damit die Kapitalrückstellungen auch hohe Verluste absorbieren können. Zuerst werden die mathematischen Grundlagen dargelegt. Dafür wird der VaR definiert und eine Einführung in die Extremwerttheorie gegeben. Darauf basierend werden dann zwei verschiedene Schätzmethoden für den VaR erklärt: Die erste Methode ist die so genannte Peaks over Threshold Methode, bei der man die Überschreitungen einer Folge von Zufallsvariablen $X_i$ über einen bestimmten Grenzwert $u$ betrachtet und einer Generalized Pareto Verteilung anpasst. Die zweite Methode basiert auf dem Hill-Schätzer.

Dann wird eine statistische Analyse der VaR- Schätzung der letzten 5 Jahre (27.04.2007 bis 23.05.2012) durchgeführt. Als Datensatz werden die täglichen Gewinne/Verluste der Goldman Sachs Group inc. benutzt. Zuerst wird die Verteilung der Daten untersucht. Das Ergebnis davon ist, dass unser gesamter Datensatz einer Exponentialverteilung zugrunde liegen. Schätzen wir aber den 99%-VaR für einen bestimmten Tag, nehmen wir immer nur die Verluste der 260 vorhergehenden Arbeitstage. So müssen wir also unsere Verteilung jeweils für diese 260 Tage schätzen. Diese unterliegen jeweils einer Extremwertverteilung mit verschiedenen Paramtern. Basierend auf den oben genannten Schätzmethoden (POT und Hill) sowie einer Kombination dieser beiden Methoden wird dann der 99%-VaR für diesen Zeitraum geschätzt. Das Ganze wird mit dem selben aber skalierten Datensatz wiederholt. Backtesting ergibt, dass sowohl die POT als auch die Hill-Methode basierend auf den nichtskalierten Daten die besten Ergebnisse gibt. Anschließend werden noch die einzelnen Terme der Kapitalanforderungen aus den Basel III Richtlinien behandelt.

# Contents

# 1 Introduction

The financial and insurance industry is naturally exposed to different risks, such as high losses or even bankruptcy, earthquakes, legal risks, etc. just to name a few examples. In the financial sector we can generally differ between three risk types: Market risk, "the risk of a change in the value of a financial position due to changes in the value of the underlying on which that position depends, such as stock and bond prices, exchange rates, commodity prices etc", (see [2]), credit risk, the risk of default of the obligor, and operational risk, the risk of losses due to a fail of people, processes or systems.

Among these three risk categories, the focus in this thesis will be on market risk or more specifically, the Value-at-risk (VaR) estimation in the Basel III framework, as described in the Revision to the Basel II market risk framework, published of the Basel Committee of Banking Supervision in 2010 [3]. The VaR is widely used as a measure of financial risk and is nothing else than a high quantile. In a non mathematical way, the $\text{VaR}_\alpha$ is the value, at which the probability of exceeding this value is equal to $(1 - \alpha)$, for a given time horizon and portfolio. This is quite intuitive, as we just find a threshold for exceeding a high loss which happens with a small probability $\alpha$. Another reason why this risk measure is so widespread, is the fact that it is so easy to calculate. But there are some criticisms of the VaR which should be discussed shortly.

First of all it does not fulfill the property of subadditivity, and thus is not a coherent risk measure (for a detailed discussion of coherent risk measures, see [2] p. 238-248). Subadditivity is, if we have two portfolios with loss random variables $L_1$ and $L_2$ and one merged portfolio with loss random variable $L = L_1 + L_2$, then it holds for the risk measure $q_\alpha : q_\alpha(L) \leq q_\alpha(L_1) + q_\alpha(L_2)$. The VaR violates this equation, even if we would intuitively say, that the combination of two portfolios should be less risky than the sum of the risk of the single portfolios, i.e. there should be a diversification benefit. This is exactly what subadditivity states. Another point is, that we do not know what happens above the probability $\alpha$. There could be significantly higher losses than the VaR or just losses which are a little higher than the VaR. Moreover, when modeling risk we always have to make some assumptions which are not met in practice. One of them is the fact that we assume a completely liquid market, which is not the case in reality, as for example sales are impossible if there is no counterparty who wants to buy. There is also always the risk of using a misspecified model for the loss distribution.

The Basel III market risk framework says, that every bank can decide which statistical model they use to calculate the Value-at-risk. It is very common to calculate the Value-at-risk based on the normal distribution, thus it is assumed that the movement of stocks is normally distributed. But investigations will show that the data fits better to a distribution which is more heavy tailed than the normal distribution. More heavy tailed means that the probability of large values is higher. In other words, if we fit our data to the normal distribution, the probability of high losses will be underestimated. This would be a big flaw, as our goal is to estimate a high quantile, which falls exactly in the part where the normal distribution is too low. This leads to the part of statistics which is dealing with the probability of rare events: Extreme Value statistics. Extreme value theory focusses on maxima of samples, and describes the behaviour of maxima in a probabilistic way as the generalized extreme value distribution.

The main emphasis of this work will be to calculate the Value-at-risk based on different

extreme value methods, the Peaks over Threshold (POT) method, the Hill-estimator and a combination of both. As many statistic models, these methods require data which is independent and identically distributed, and it is common to assume this for the data of stock returns. But one should be aware of the fact that this is not exactly the case for daily stock returns, as high losses occur in clusters frequently. This is one of the assumptions which we have to make, but is not met in reality. Thus our estimations can be defective at some points. Moreover, our whole further analysis is done with the stock price of one single stock. In reality the VaR is calculated based on the whole portfolio of a bank. This simplification can lead to some other results than in practice.

This work will begin with a short introduction to the quantitative standards in the Basel III market risk framework in section 2. The exact description and formula of the capital requirements, and the rules given for the calculation of the VaR are summarized there. The mathematical definition of the VaR will be treated in the following section, as well as the mathematical basics we need to apply the extreme value methods mentioned above. This will be followed by the analysis of the capital requirements in the Basel III market risk framework (section 4).

The data set used are the daily prices of the Goldman Sachs Group, Inc. (NYSE) obtained from `http://finance.yahoo.com`. Having this data there will be first a general analysis of the daily returns of the stock, using standard statistical tools, as the Quantile-Quantile plot. Then the VaR will be estimated for every day of the last five years (the period of 2007-04-27 until 2012-05-23) using the different methods in section 4.2. In section 4.3 the different terms of the formula of the capital requirements will be applied. All the statistical work is done with the `R` Project for statistical computing [6], and all the plots which follow are made with this software. The main R package used is **evir** [7], besides a few exceptions where **POT** [8] or **evd** [9] are used. They are all for doing extreme value statistics and can be obtained for free on `http://cran.r-project.org`. For all the statistical analysis in section 4, the functions used in R will be presented in a short way in the text and can be found in full length in the appendix.

## 2 Basel III market risk framework - Quantitative standards

Now, five years after the financial crisis in 2007-2008 financial markets have more or less recovered, but it is still of high demand to avoid crises of this magnitude. The low quality of capital and the resulting limited liquidity of many banks were among the reasons which made the crisis so severe. There is clearly a need for a new regulation defining bank's capital buffers to strengthen the whole financial sector against uncertain losses. On that account, the Basel Committee on Banking supervision introduces a reform of the regulatory framework for the bankings sector, named "Basel III".

It is not the first time that regulations are intensified. In 1988 Basel I, the first international guidelines for banks' capital buffers, was introduced and updated in 1996. Whereas in 1988 it only contained capital requirements for credit risk, in 1996 the risk category market risk was amended. The amendment was a reaction to the bankruptcy of Barings, the oldest merchant bank in UK, which was eventually caused by a bad trade of the Singapore-based trader, Nick Leeson. The next enhancement was the part operational risk, which was implemented in the Basel II framework in 2004. As response to the latest crises, the reform Basel III will help to raise the resilience of individual banking institutions in periods of stress. It will be implemented stepwise between 2013 and 2019. The basic goals are to strengthen the equity basis and tighten liquidity regulations of banks. This is achieved in a raise of the capital quality and also a raise in the level of the minimum capital requirements. More precisely, the common equity requirements increase from 2% to 4.5 % and a capital conservation buffer of 2.5 % is introduced. Furthermore, the risk coverage of the capital framework for trading activities, securitisation, etc. will be increased. Moreover, an internationally harmonised leverage ratio is implemented as well as the promotion of building up capital buffers in good times that can be drawn down in periods of stress. Also minimum global liquidity standards are introduced and the supervisory review process is strengthened.

Especially regarding the Basel II market risk framework a factor which caused high losses was the excessive build up leverage. A big problem was also that there were no capital charges which include default risk and migration risk for unsecuritised credit products, which are implemented now. Moreover, the minimum capital requirements were simply too low to capture the losses during the crises. The answer to this problem is the introduction of the so called stressed Value-at-risk (SVaR). Whilst before capital requirements were only based on the Value-at-risk (VaR), now the SVaR is added to increase the capital requirements and thus absorb higher losses.

According to the Basel III framework the banks still have flexibility in the calculation of their capital requirements, but there are some minimum standards which they have to adhere. The basis of the calculation is the VaR computed on a daily basis, using a 99th percentile, one tailed confidence interval. The 10-day-returns of the portfolio must be used, which can be approximated by using the 1-day-returns and scale them according to a normal distribution by $\sqrt{10}$ up to ten days. Furthermore, the length of the sample period underlying the calculation must be at least one year. Besides that the banks are free to choose between models based on variance-covariance matrices, historical simulations or Monte Carlo simulations. Moreover, the Basel III market risk framework defines a SVaR,

which is calculated on the same basis as the VaR, but in a period of significant stress for the bank's portfolio. The length of the period must be twelve months and the choice of the period has to be approved by the supervisor as well as reviewed regularly. As an example, a period relating to significant losses in 2007/2008 would adequately reflect a period of such stress, although other periods may be considered.

Applying these standards, the banks' capital requirements are the sum of

- The higher of (1) its previous day's Value-at-Risk estimate $\text{VaR}_{0.99,t-1}$ and (2) an average of the daily Value-at-Risk estimate on each of the preceding sixty business days $\text{VaR}_{0.99,avg}$, multiplied by a factor $m_c$.

- The higher of (1) its latest available Stressed-Value-at-Risk estimate $\text{SVaR}_{0.99,t-1}$ and (2) an average of the Stressed-Value-at-Risk estimates on each of the preceding sixty business days $\text{SVaR}_{0.99,avg}$, multiplied by a factor $m_s$.

The minimum for factors $m_c$ and $m_s$ is 3. To this minimum an amount between 0 and 1 has to be added, which is directly related to the ex-post performance of the bank's model. The backtesting results for calculating this "plus" are based on VaR only and not SVaR. This results in the following formula for the capital requirement at time $t$:

$$C_t = \max\{\text{VaR}_{0.99,t-1}; m_c\text{VaR}_{0.99,avg}\} + \max\{\text{SVaR}_{0.99,t-1}; m_s\text{SVaR}_{0.99,avg}\} \qquad (2.0.1)$$

The whole second term of the equation is new compared to the capital requirements in the Basel II market risk framework. That means that a large amount of money has to be added. As the SVaR as well as the average SVaR should be larger than the VaR by definition, we expect the capital requirements to be more as twice as high as before.

# 3  Mathematical Basics

Before starting a statistical analysis of the VaR, it is obligatory to know the basic mathematical results on which our chosen estimation methods are based on. Therefore we start with introducing the VaR as a risk measure of financial losses. Then we present the basic ideas of extreme value theory and end with two statistical extreme value methods, the POT method and the Hill-estimator. The contents and results of this section can all be found in P. Embrechts, C. Klüppelberg, and T. Mikosch [1] and A. J. McNeil, R. Frey, and P. Embrechts [2].

## 3.1  The Value-at-Risk

The Value-at-Risk (VaR) is probably the most popular risk measure in financial institutions. One of the purposes of risk measurement is to estimate possible future losses of a bank's portfolio and calculate minimal capital requirements to be able to absorb these losses. In the Basel III market risk framework the VaR is specified as risk measure for this calculation. To define the VaR we need some basic statistic definitions. For the following results we refer to [2], p. 37-43.

**Definition 3.1.1.** (Generalized inverse and Quantile function)

(i). *Let h: $\mathbb{R} \to \mathbb{R}$ be a non-decreasing function (i.e. $h(x) \leq h(y)$ for $x \leq y$). The generalized inverse is given by*

$$h^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : h(x) \geq t\}.$$

(ii). *Given some distribution function F, the generalized inverse $F^{\leftarrow}$ is called the quantile function of F. For $\alpha \in (0,1)$ the $\alpha$-quantile of F is given by*

$$q_{\alpha} := F^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

If $F$ is continuous and strictly non-decreasing, the quantile function is equal to the ordinary inverse of $F$: $q_{\alpha}(F) = F^{-1}(\alpha)$. We can also write $q_{\alpha}(F) = q_{\alpha}(X)$, where $X \sim F$. Now consider a random variable $L$, which is defined as the loss of a portfolio of risky assets, with distribution function $F_L(l) = P(L \leq l)$. Comparing two portfolios one could ask the question which portfolio is riskier than the other. We want to find a way to describe the different risks, i.e. we want to define a risk measure. A first idea could be the maximum loss possible, which is $\inf\{l \in \mathbb{R} : F_L(l) = 1\}$. This measure could be used in insurance cases, where the maximum loss is a finite number, but as we look at financial losses which are not bounded in general, this is not an appropriate risk measure. Therefore, we introduce the VaR as risk measure:

**Definition 3.1.2.** (Value-at-Risk (VaR))
*Given some confidence level $\alpha \in (0,1)$. The* VaR *of a portfolio at the confidence level $\alpha$ is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1 - \alpha)$. Formally:*

$$\mathrm{VaR}_{\alpha}(L) = \inf\left\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\right\} = \inf\left\{l \in \mathbb{R} : F_L(l) \geq \alpha\right\}.$$
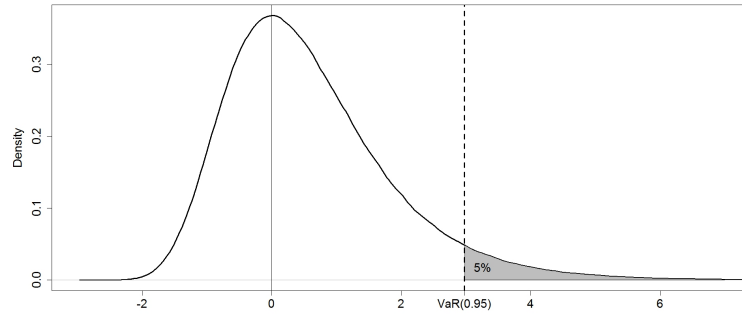
Figure 1: The density of the loss rv $L$ and its 95% value-at risk

In other words, the probability that the loss random variable $L$ is bigger than a certain threshold (i.e. the $(1 - \alpha)$ quantile, $\text{VaR}_\alpha$) is smaller than $\alpha$. It means that, if we choose $\alpha$ very small you get high values for $\text{VaR}_\alpha$. In the financial sector values of $\alpha = 0.95$ or $\alpha = 0.99$ are common.

## 3.2  Basics in Extreme Value Theory

Extreme Value Theory is the part of statistics which is dealing with the probability of rare events. It considers events that exceed all the events that have happened in the past. To study extreme value theory we assume that we have a sample $X_1, ..., X_n$ of independent identically distributed (iid) random variables (rvs) with distribution function (df) $F$. Now the main part of interest is to determine the df for the maximum of the sample $M_n = \max(X_1, ..., X_n)$.

In this section the basic results for finding this distribution will be presented. First, we will use the Poisson approximation to examine asymptotic behaviour of the maxima of our sample. This will lead us to the Fisher-Tippet Theorem, a precise representation of the so called extreme value distributions. To study the behaviour of the tail of our distributions we will come to the Pickands-Balkema de Haan theorem. What follows is a short summery of [1], pages 113-168. The proofs which are not presented can be found there.

### 3.2.1  Limit Probabilities for maxima

As said before we want to focus on the maximum $M_n = \max(X_1, ..., X_n)$ of a sample of iid rvs $X_1, ..., X_n$. First, we determine the distribution function of $M_n$ by using the independence of the $X_i$ and the fact that, if $M_n \leq x$, it follows $X_i \leq x, \forall i = 1, .., n$. Thus,

$$P(M_n \leq x) = P(X_1 \leq x, ..., X_n \leq x) = F^n(x).$$

As the right endpoint of $F$ we define: $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$ and we want to examine the behaviour of $M_n$ towards $x_F$. We achieve the probabilities as $n \to \infty$:

$$P(M_n \leq x) \to 0 \qquad x < x_F,$$
$$P(M_n \leq x) = 1 \qquad \forall x \geq x_F, \quad x_F < \infty,$$
$$P(M_n \geq x) \to 1 \qquad \forall x, \quad x_F = \infty.$$

This means that $M_n \xrightarrow{P} x_F$, and with the fact that $M_n$ is non-decreasing in $n$ we deduce that the maximum of a sample converges almost surely to the right endpoint of the df $F$, i.e. $M_n \xrightarrow{a.s.} x_F$.

This is a quite intuitive result which does not give us much information about the probabilistic behaviour of $M_n$. It is much more interesting to find a non-degenerate distribution to describe the behaviour of a samples maximum, i.e. we look for a df $H$ such that there exist constants $c_n > 0, d_n \in \mathbb{R}$ with

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim H.$$

Thus, we have to investigate the distribution

$$P(c_n^{-1}(M_n - d_n) \leq x)$$

or in other words

$$P(M_n \leq u_n) \quad \text{where} \quad u_n = u_n(x) = c_n x + d_n,$$

where $c_n$ is the normalising and $d_n$ is the centering constant. It follows an elementary result which brings us a step closer to the distribution of maxima of a sample.

**Proposition 3.2.1.** (Poisson Approximation)
*For given $\tau \in [0, \infty]$ and an increasing sequence $(u_n)$ of real numbers, for $n \to \infty$, it holds*

$$n\overline{F}(u_n) \to \tau \quad \Longleftrightarrow \quad P(M_n \leq u_n) \to e^{-\tau}. \tag{3.2.1}$$

It remains the question what the possible (non-degenerate) limit laws for the maxima $M_n$ are when properly normalised and centered. The answer to this question is given by the limit property of max-stable laws.

**Definition 3.2.2.** (Max-stable distribution)
*A df is max-stable if $\forall n \geq 2$ and iid $X_i \sim F$, there exist constants $c_n > 0$ and $d_n \in \mathbb{R}$, so that*

$$c_n^{-1}(M_n - d_n) \overset{d}{=} X_1 \quad \Longleftrightarrow \quad M_n \overset{d}{=} c_n x + d_n. \tag{3.2.2}$$

**Proposition 3.2.3.** (Limit property of max-stable laws)
*The max-stable distributions are the only non-degenerate limit distributions for normalised and centred maxima of iid rvs $X_1, ..., X_n$.*

However, we still do not know how these max-stable distributions look like and what properties they have. Thus, we want to present the limit distributions of normalised and centred maxima of iid rvs, the so called *extreme value distributions*:
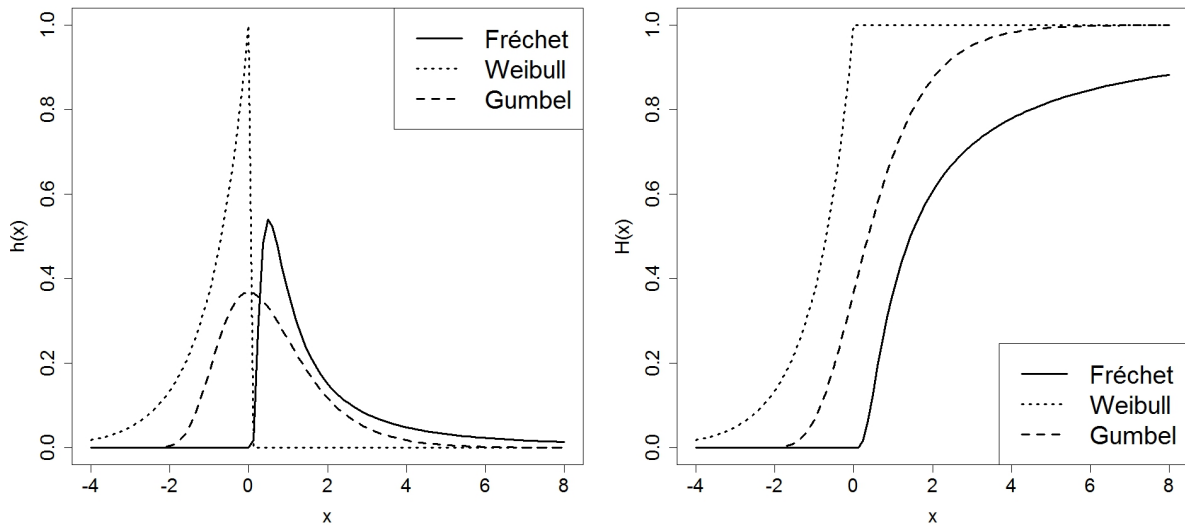
Figure 2: On the right hand side, the distribution functions of the three extreme value distributions. The solid line for Fréchet: $H(x) = \Phi_\alpha(x)$ where $\alpha = 1$, the dotted line for Weibull: $H(x) = \Psi_\alpha(x)$ where $\alpha = 1$ and the dashed line for Gumbel: $H(x) = \Lambda(x)$. On the left hand side are the corresponding densities.

**Fréchet distribution:** $\qquad \Phi_\alpha(x) \quad = \quad \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \qquad \alpha > 0$

**Weibull distribution:** $\qquad \Psi_\alpha(x) \quad = \quad \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \qquad \alpha > 0$

**Gumbel distribution:** $\qquad \Lambda_\alpha(x) \quad = \quad \exp\{-e^{-x}\}, \quad x \in \mathbb{R}.$

The corresponding rvs to these distributions are called extremal rvs. It is easy to show that the distributions are max-stable and (3.2.2) holds for all of them with:

$$\text{Fréchet} \qquad M_n \overset{d}{=} n^{\frac{1}{\alpha}} X_1$$

$$\text{Weibull} \qquad M_n \overset{d}{=} n^{-\frac{1}{\alpha}} X_1$$

$$\text{Gumbel} \qquad M_n \overset{d}{=} X_1 + \ln n$$

We will present an exemplary proof in the case of a Fréchet distribution, the other two cases can be proven analogously. Let $X_i \, iid \sim \Phi_\alpha$, then it holds:

$$P(M_n \leq x) = (\Phi_\alpha(x))^n = \exp\{-nx^{-\alpha}\} = \exp\{-(n^{-\frac{1}{\alpha}}x)^{-\alpha}\}$$

$$\implies \quad P(n^{-\frac{1}{\alpha}} M_n \leq x) = \exp\{-x^{-\alpha}\} = \Phi_\alpha(x)$$

$$\implies \quad n^{-\frac{1}{\alpha}} M_n \overset{d}{=} X_1$$

Actually, they are the only max-stable distributions one can find. In other words, the class of max-stable distributions is equivalent to the class of extreme value distributions. This is stated in the most important theorem in classical extreme value theory:

**Theorem 3.2.4.** (Fisher-Tippett theorem, limit laws for maxima)
*Let $(X_n)$ be a sequence of iid rvs. If there exist norming constants $c_n > 0, d_n \in \mathbb{R}$ and some non-degenerate df H such that*

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H,$$

*then H belongs to the type of one of the following three dfs:*

$$Frechet: \quad \Psi_\alpha, \qquad Weibull: \quad \Phi_\alpha, \qquad Gumbel: \quad \Lambda.$$

This theorem for maxima can be compared with the central limit theorem for sums. The Fisher-Tippett theorem describes the limit distribution of maxima $M_n$ of a sample of iid rvs $X_1, ..., X_n$ as one of the extreme value distributions. The central limit theorem does the same for the sum $S_n = X_1 + ... + X_n$ of a sample of iid rvs $X_1, ..., X_n$, where the $X_i$ have a finite variance. With norming constants $\mu = \mathbb{E}(X_1)$ and $\sigma = \sqrt{var(X_1)}$ you get the limit distribution for sums:

$$\lim_{n \to \infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x), \quad x \in \mathbb{R},$$

Where $\Phi(x)$ is the standard normal distribution $\mathbb{N}(0, 1)$. Here we only have one limit distribution, compared to the three limit distribution for maxima.

### 3.2.2   Maximum Domain of Attraction

If we only know that we have an extremal rv $X$ we still do not know to which extreme value distribution it belongs or, in other words, given an extreme value distribution $H$, what conditions on the distribution function $F$ imply that the normalised and centered maxima converge weakly to $H$.

**Definition 3.2.5.** (Maximum domain of attraction (MDA))
*Consider a sequence of iid rvs $X_i \sim F$ and an extreme value distribution H. If there exist constants $c_n > 0$, $d_n \in \mathbb{R}$ such that*

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim H,$$

*then F is in the maximum domain of attraction of H: $F \in \mathrm{MDA}(H)$.*

Now we will present how to decide which distribution to choose, i.e. which is the maximum domain of attraction of our sample.

**Proposition 3.2.6.** (Characterization of MDA(H))
*F belongs to the MDA(H) with norming constants $c_n > 0$, $d_n \in \mathbb{R}$ if and only if*

$$\lim_{n \to \infty} n\overline{F}(c_n x + d_n) = -\ln H(x), \quad x \in \mathbb{R},$$

*where $\overline{F}(x) = 1 - F(x)$. When $H(x) = 0$ the limit is interpreted as $\infty$.*

*Proof.* The proposition is easy to prove by using the continuity of $H$ on $\mathbb{R}$ and the Poisson approximation (3.2.1):

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim H$$

$$\Longleftrightarrow \quad \lim_{n \to \infty} P(M_n \leq c_n x + d_n) = \lim_{n \to \infty} F^n(c_n x + d_n) = H(x), \quad x \in \mathbb{R}$$

$$\Longleftrightarrow \quad \lim_{n \to \infty} n\overline{F}(c_n x + d_n) = -\ln H(x), \quad x \in \mathbb{R}$$

$\square$

This means that having the distribution of $X$ we can decide which of the extreme value distributions is the underlying distribution of our maxima. To characterize this exactly we need some technical help:

**Definition 3.2.7.** (Slowly varying and regularly varying functions)

*(i). A positive, Lebesgue measurable function $L$ on $(0, \infty)$ is slowly varying at $\infty$ if*

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1, \quad t > 0.$$

*(ii). A positive, Lebesgue measurable function $h$ on $(0, \infty)$ is regularly varying at $\infty$ with index $\rho \in \mathbb{R}$ if*

$$\lim_{x \to \infty} \frac{h(tx)}{h(x)} = t^{\rho}, \quad t > 0.$$

One can say that slowly varying functions change relatively slowly for $x \to \infty$ while regularly varying functions behave like power law functions. Regularly varying functions can be represented by $h(x) = x^{\rho} L(x)$, where $L(x)$ is a slowly varying function. This will be needed in the characterisation of the maximum domain of attraction of the Fréchet distribution. It should be mentioned that for our further purposes, the analysis of one-day losses of stocks(or indices), the Fréchet distribution is most appropriate and, thus, the results are presented in more detail for the Fréchet distribution.

**Theorem 3.2.8.** (Maximum domain of attraction of $\Phi_\alpha$)
$F \in \text{MDA}(\Phi_\alpha)$ *for some $\alpha > 0$ if and only if $\overline{F}(x) = x^{-\alpha} L(x)$ for some slowly varying function $L$. Then*

$$c_n^{-1} M_n \xrightarrow{d} Y \sim \Phi_\alpha,$$

*where $c_n = F^{\leftarrow}(1 - \frac{1}{n})$.*

*Proof.* We assume that we have a sequence of iid rvs $X_1, ..., X_n$ with distribution function $F$ and regularly varying tail distribution function $\overline{F}(x) = x^{-\alpha} L(x)$ for some slowly varying $L$ and $\alpha > 0$. With $c_n = F^{\leftarrow}(1 - \frac{1}{n})$ we get $\overline{F}(c_n) \sim \frac{1}{n}$ and then as $n \to \infty$:

$$\begin{aligned} \text{for} \quad x > 0: \quad n\overline{F}(c_n x) \quad &\sim \quad \frac{n\overline{F}(c_n x)}{\overline{F}(c_n)} \quad \to \quad x^{-\alpha} \\ \text{for} \quad x \leq 0: \quad n\overline{F}(c_n x) \quad &\geq \quad n\overline{F}(0) \quad \to \quad \infty \end{aligned}$$

Where we use that $\overline{F}(0) > 0$, since it is regularly varying. Summarizing the two cases above we have

$$n\overline{F}(c_n x) \to -\ln\Phi_\alpha(x) \quad \forall x$$

and with the Poisson approximation (3.2.1) we can conclude

$$c_n^{-1} M_n \xrightarrow{d} Y \sim \Phi_\alpha.$$

The other direction of the proof can be found in [1], p. 131f.                              $\square$

There we used that the tail of $\Phi_\alpha$ is regularly varying with index $-\alpha$. This can easily be shown by Taylor expansion.

$$1 - \Psi_\alpha(x) = 1 - \exp\{-x^{-\alpha}\} \sim x^{-\alpha}, \quad x \to \infty.$$

As you can see, $d_n = 0$ and we can choose $c_n = F^{\leftarrow}(1 - \frac{1}{n})$. For studying the extreme behaviour of a sample one can replace the distribution function $F$ by the empirical distribution function $F_n = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leq x\}}$, $x \in \mathbb{R}$ and gets $c_n = F_n^{\leftarrow}(1 - \frac{1}{n}) = F_n^{\leftarrow}(\frac{n-1}{n}) = X_{(2)}$, where $X_{(1)} \geq ... \geq X_{(n)}$ is the ordered sample of $X_1, ..., X_n$. Because the tail is regularly varying the distribution has an infinite right end point $x_F$. This is one of the properties which fits very good for the daily losses of a stock. Even if we can assume that the losses will not exactly be $\infty$, the loss of one day can be higher than any observed loss before. Another important property of $\Phi_\alpha$ which will not be shown here, is

$$\mathbb{E}(X_+^\delta) \begin{cases} = \infty & \text{for} \quad \delta > \alpha \\ < \infty & \text{for} \quad \delta < \alpha. \end{cases}$$

Which means that if you just have $\alpha < 2$, you get infinite second moments, i.e. an infinite variance. Consequently $F$ is very heavy tailed if $F \in \text{MDA}(\Phi_\alpha)$. In particular, it is more heavy tailed than the normal distribution.

**Theorem 3.2.9.** (Maximum domain of attraction of $\Psi_\alpha$)
$F \in \text{MDA}(\Psi_\alpha)$ *for some* $\alpha > 0$ *if and only if* $x_F < \infty$ *and* $\overline{F}\left(x_F - \frac{1}{x}\right) = x^{-\alpha}L(x)$ *for some slowly varying function L. Then*

$$c_n^{-1}(M_n - x_F) \xrightarrow{d} Y \sim \Psi_\alpha,$$

*where* $c_n = x_F - F^{\leftarrow}\left(1 - \frac{1}{n}\right)$.

This is quite similar to the characterization of $\text{MDA}(\Phi_\alpha)$, but it has a finite right endpoint. For example chose $F$ as uniform distribution $U(0, 1)$. The maximum of a sample $X_i \sim F$ is at most one a.s. and the right endpoint of the distribution is 1. It is easy to check that $n(M_n - 1) \xrightarrow{d} Y \sim \Psi_1$. Clearly this is not a considerable distribution for our daily stock returns, because they have a distribution with infinite right endpoint.

The $\text{MDA}(\Lambda)$ will not be discussed in detail here, as the distribution is lighter tailed than any regular varying tail. It consists of a wide range of distribution functions such as the lognormal distribution, the exponential distribution or the normal distribution. The characterization of $\text{MDA}(\Lambda)$ is more difficult than the regular variation of $\text{MDA}(\Phi_\alpha)$ or $\text{MDA}(\Psi_\alpha)$, but it can be characterized by the so called von Mises functions and their tail-equivalent dfs. For further discussion see [1], p.138-152.

To summarize the results above: As Figure 2 shows, the Weibull distribution has a finite right end point, while the Fréchet and Gumbel have an infinite one. If we compare the two latter distributions $\Phi_\alpha$ is more heavy tailed than $\Lambda$.

### 3.2.3   Generalized extreme value distribution

Now as the three different extreme value distributions are characterized, we still have to decide to which of the three distributions our sample belongs, before we can estimate parameters. For statistical methods it is mostly desirable that you find a parametric family of distributions that constitute the three extreme value distribution. Thus, we define:

**Definition 3.2.10.** (Jenkinson-von Mises representation of the extreme value distributions: Generalised extreme value distribution (GEV))
*The GEV is defined as*

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

*Hence, the support of $H_\xi$ corresponds to*

$$x \in \mathbb{R} \quad for \quad \xi = 0$$

$$x < -\frac{1}{\xi} = \alpha \quad for \quad \xi < 0$$

$$x > -\frac{1}{\xi} = \alpha \quad for \quad \xi > 0.$$

*By defining $H_{\xi,\mu,\sigma} := H_\xi(\frac{x-\mu}{\sigma})$ we obtain a three-parameter family with location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$ and shape parameter $\xi \in \mathbb{R}$.*

The shape parameter $\xi$ plays an important role as it characterizes the distribution of $H_\xi$ as one of the three extreme value distributions named above:

$$\begin{aligned} \text{Fréchet:} \quad & \xi = \frac{1}{\alpha} > 0 \\ \text{Weibull:} \quad & \xi = -\frac{1}{\alpha} < 0 \\ \text{Gumbel:} \quad & \xi = 0. \end{aligned}$$

$H_\xi$ is nothing else than a shifted and scaled version of these three distributions, which makes it easier to estimate parameters of a distribution of a sample of extremal rvs. Note that $H_\xi$ is a continuous function of $\xi$ if $x$ is fixed.

### 3.2.4   Generalized Pareto distribution

As we characterized the MDA(H) above, we can do the same for the generalized extreme value distribution $H_\xi$.

**Theorem 3.2.11.** (Pickands-Balkema-de Haan Theorem(1977))
*$F \in \mathrm{MDA}(H_\xi)$ for some $\xi \in \mathbb{R}$ if and only if $\exists a(x) > 0$, $a(\cdot)$ measurable, such that for $1 + \xi x > 0$ it holds:*

$$\lim_{x \uparrow x_F} \frac{\overline{F}(u + xa(u))}{\overline{F}(u)} \quad = \quad \begin{cases} (1+\xi x)^{-\frac{1}{\xi}} & for \quad \xi \neq 0 \\ e^{-x} & for \quad \xi = 0 \end{cases}$$

*Proof.* We only prove the result for Fréchet case.

First, we assume that $F \in \mathrm{MDA}(\Phi_\alpha)$, i.e. $\xi = \frac{1}{\alpha} > 0$ which means that we have $H_\xi(x) = \Phi_\alpha(\alpha^{-1}(x + \alpha))$. As in theorem 3.2.8 this is equivalent to $\overline{F}(x) = x^{-\alpha} L(x)$ where $L$ is slowly varying. We chose $a(u) = \frac{u}{\alpha}$ and get

$$\frac{\overline{F}(u + xa(u))}{\overline{F}(u)} = \frac{\overline{F}(u(1 + \frac{x}{\alpha}))}{\overline{F}(u)} = \frac{(u(1 + \frac{x}{\alpha}))^{-\alpha}}{u^{-\alpha}} \cdot \frac{L(u(1 + \frac{x}{\alpha}))}{L(u)}$$

$$= \left(1 + \frac{x}{\alpha}\right)^{-\alpha} (1 + o(1)) \to (1 + \xi x)^{-\frac{1}{\xi}}$$

Second, we assume that $\exists\, a(u) > 0$ such that

$$\frac{\overline{F}(u + xa(u))}{\overline{F}(u)} \xrightarrow{u \to x_F} (1 + \xi x)^{-\frac{1}{\xi}}.$$

We choose $d_n$ such that $\overline{F}(d_n) \sim \frac{1}{n}$ and get

$$n\overline{F}(d_n + xa(d_n) \to \left(1 + \frac{x}{\alpha}\right)^{-\alpha}$$

$$\stackrel{(3.2.1)}{\Longrightarrow} (a(d_n))^{-\alpha}(M_n - d_n) \xrightarrow{d} Y \sim H_\xi$$

And with $\xi = \frac{1}{\alpha}$ it holds that

$$H_\xi(x) = \exp\left\{-\left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right\} = \Phi_\alpha(\alpha^{-1}(x + \alpha))$$

Thus, we have $F \in \mathrm{MDA}(\Phi_\alpha)$.                                                          $\square$

To clarify what is meant by the left expression in the preceededing theorem we first define.

**Definition 3.2.12.** (Excess distribution function, mean excess function)
*For a rv X with df F and a right endpoint $x_F$ the excess distribution function for $u < x_F$ is given by:*
$$F_u(x) = P(X - u \le x | X > u), \quad x > 0.$$
*And the "corresponding expected value", i.e. the mean excess function is given by:*
$$e_F(u) = \mathbb{E}(X - u | X > u).$$

For the tail of $F_u$ we get

$$\overline{F}_u(x) = P(X - u > x | X > u) = \frac{P(X > u + x)}{P(X > u)} = \frac{\overline{F}(u + x)}{\overline{F}(u)}. \tag{3.2.3}$$

Therefore, $\overline{F}_u(xa(u))$ is equal to the left expression in theorem 3.2.11. Remark that we have:

$$e_F(u) = \int_0^{x_F} y \, dF_u(y) = \int_0^{x_F} \frac{1}{\overline{F}(u)} y \, dP(X \le u + x | X > u)$$

$$= \frac{1}{\overline{F}(u)} \int_u^{x_F} (Z - u) \, dF(Z) = \frac{1}{\overline{F}(u)} \int_u^{x_F} \overline{F}(x) \, dx. \tag{3.2.4}$$

These two expressions will be relevant for the Peaks over threshold method in section 3.3 and, thus, also for the further data analysis in section 4.

**Definition 3.2.13.** (The generalized Pareto distribution (GPD))
*The GPD is defined as*

$$G_\xi(x) \quad = \quad \begin{cases} 1 - (1 + \xi x)^{-\frac{1}{\xi}}, & if \quad \xi \neq 0 \\ 1 - e^{-x}, & if \quad \xi = 0 \end{cases}$$

*with*

$$x \geq 0 \qquad for \quad \xi \geq 0$$
$$0 < x \leq -\frac{1}{\xi} \quad for \quad \xi < 0.$$

*By defining $G_{\xi,\mu,\beta} := G_\xi(\frac{x-\mu}{\beta})$ we get a three-parameter family with location parameter $\mu \in \mathbb{R}$, scale parameter $\beta > 0$ and shape parameter $\xi \in \mathbb{R}$.*
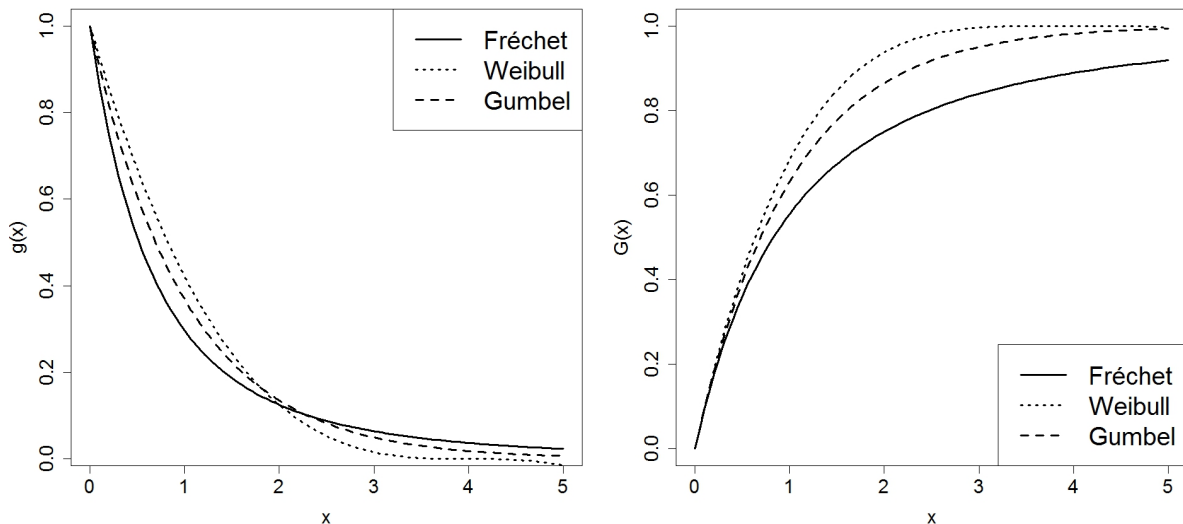


Figure 3: On the right hand side, the dfs of three generalized Pareto distributions, corresponding to the three different extreme value distributions. The solid line for Fréchet, with $\xi = 0.5$, the dotted line for Weibull, with $\xi = -0,25$ and the dashed line for Gumbel, with $\xi = 0$. On the left hand side are the corresponding densities.

In practice you mostly assume $\mu = 0$ and therefore you get

$$G_{\xi,\beta}(x) \quad = \quad 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}} \quad x \in D = \begin{cases} [0, \infty) & if \quad \xi \geq 0 \\ [0, -\frac{\beta}{\xi}] & if \quad \xi < 0. \end{cases}$$

Some examples of the GPD are the standard exponential distribution (for $\xi = 0$), the uniform distribution (for $\xi = -1$) and the pareto distribution (for $\xi > 0$). The relation to the general extreme value distribution follows from the Pickand-Balkema-de Haan theorem 3.2.11.

**Proposition 3.2.14.** (Relation of GPD and GEV)
$F \in \mathrm{MDA}(H_\xi)$ *for some* $\xi \in \mathbb{R}$ *if and only if* $\exists \, a(u) > u$ *so that*

$$\lim_{x \uparrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_\xi \left( \frac{x}{a(u)} \right)| = 0.$$

This means that the GPD is an appropriate approximation for the excess df $F_u$, if $F \in \mathrm{MDA}(H_\xi)$. It is a very important tool for tail and high quantile estimation. The POT method, which will be described below is based on the GPD approximation of excesses over a high threshold.

## 3.3  The POT method

Now that we have all the technical devices we can come to the main topic of our further analysis. A part of Extreme value statistics which refers to excesses of high thresholds: the Peaks over threshold (POT) method. The basic idea of the POT method is that you assume a sample $X_1, ..., X_n$ of iid rvs, which have a GPD as excess distribution function. For the statistical analysis we only look at excesses of the sample over a certain threshold $u$. Thereby, it is possible that you have any distribution in the "middle" of the sample and an extreme value distribution on the "right margin" of the sample, where "middle" denotes the values close to the mean of the sample and "right margin" denotes the values which lie above a certain threshold. This is a big advantage over using the whole sample.

The statistical basis of the POT method is the assumption, that the excesses of a high threshold $u$ are generalized pareto distributed. Our goal is to estimate the parameters of the GPD and certain quantiles or percentiles based on this distribution. The presented results can be found in [1], p. 352-358.
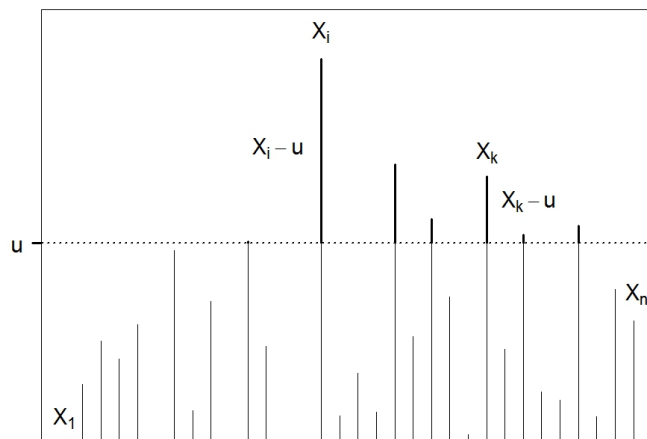


Figure 4: A sample $X_1, ... X_n$ of rvs and its excesses over the threshold.

### 3.3.1   Excess distribution function

First, we assume that we have chosen a threshold $u$. How to choose $u$ will be discussed in more detail in section 3.3.4. So we search for the tail of the distribution function of the excesses over $u$, i.e. $\overline{F}(u+x)$. From (3.2.3) it follows that

$$\overline{F}(x+u) = \overline{F}_u(x)\overline{F}(u).$$

Now we want to estimate the two expressions on the right hand side, based on the sample. Therefore, we look at the number of excesses over $u$ which we define by

$$N_u = \#\{i \leq n : X_i > u\}.$$

And we denote the excesses $X_k - u$ by $Y_1, ..., Y_{N_u}$. They are illustrated in Figure 4. They can be seen as a new sample of rvs and again we assume independence of the rvs. Having defined the number of excesses we can replace the tail of $F$ by its empirical distribution function:

$$\widehat{\overline{F}}(u) = \overline{F}_n(u) = \frac{1}{n}\#\{i \leq n : X_i > u\} = \frac{N_u}{n}$$

A direct result of (3.2.14) is that we can approximate

$$\overline{F}_u(x) \approx \overline{G}_{\xi,\beta}(x)$$

for high values of $u$. Note that we have $\beta = a(u)$ and $a(u)$ is a function of $u$. Thus, $\beta$ as well as $\xi$ depend on the choice of $u$. By estimating $\beta$ and $\xi$ we get:

$$\widehat{\overline{F}}_u(x) = \overline{G}_{\widehat{\xi},\widehat{\beta}}(x),$$

where again $\widehat{\xi}$ and $\widehat{\beta}$ depend on the number of exceedances and, thus, the choice of $u$. Combining the above yields

$$\widehat{\overline{F}}(u+x) = \widehat{\overline{F}}(u)\widehat{\overline{F}}_u(x) = \frac{N_u}{n}\overline{G}_{\widehat{\xi},\widehat{\beta}}(x).$$

### 3.3.2   Estimation of parameters

Estimating the parameters $\xi$ and $\beta$ can be done by maximum likelihood estimation (MLE). As MLE requires a sample of iid rvs, we have to make the assumption that $Y_1, ..., Y_{N_u}$ are iid. This assumption could be criticised in practice, as in reality large losses frequently occur in clusters which implies that they are dependent. But here we will assume independence of $Y_1, ..., Y_{N_u}$ with $Y_i \sim F_u$. For the density we get

$$g_{\xi,\beta}(x) = \frac{\xi}{\beta}\left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}-1}, x \in D = \begin{cases} [0, \infty) & \text{if } \xi \geq 0 \\ [0, -\frac{\beta}{\xi}] & \text{if } \xi < 0 \end{cases}.$$

Accordingly, the likelihood function is

$$L(\xi, \beta, Y_1, ...Y_{N_u}) = \prod_{i=1}^{N_u} g_{\xi,\beta}(Y_i) = \prod_{i=1}^{N_u} \frac{\xi}{\beta}\left(1 + \frac{\xi Y_i}{\beta}\right)^{-\frac{1}{\xi}-1}$$

and the log-likelihood function

$$l(\xi, \beta, Y_1, ... Y_{N_u}) = \ln L(\xi, \beta, Y_1, ... Y_{N_u}) \tag{3.3.1}$$

$$= N_u(\ln \xi - \ln \beta) - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{N_u} \ln\left(1 + \frac{\xi Y_i}{\beta}\right).$$

Differentiation with respect to $\xi$ and $\beta$ and setting to zero provides the results. The equations have to be solved numerically, and they yield the estimates $\widehat{\xi}$ and $\widehat{\beta}$.

### 3.3.3 Estimation of tail and quantile

Having the estimates, we obtain a representation of the tail $\overline{F}(u + x)$:

$$\widehat{\overline{F}}(u + x) = \frac{N_u}{n} \overline{G}_{\widehat{\xi}, \widehat{\beta}}(x) = \frac{N_u}{n}\left(1 + \frac{\widehat{\xi}x}{\widehat{\beta}}\right)^{-\frac{1}{\widehat{\xi}}}.$$

The estimated p-quantile is immediately obtained by setting $\widehat{\overline{F}}(\widehat{x}_p) = 1 - p$

$$\widehat{x}_p = u + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{n}{N_u}(1 - p)\right)^{-\widehat{\xi}} - 1\right). \tag{3.3.2}$$

This equation will be important when estimating the $\text{VaR}_\alpha$ with the POT method, as we have $\text{VaR}_\alpha = x_\alpha$. Moreover, if $p = 1$, we get the estimated right endpoint of the distribution.

$$\widehat{x}_F = u - \frac{\widehat{\beta}}{\widehat{\xi}}.$$

### 3.3.4 Choice of the threshold

Remember that we did the steps above under the assumption that we have already chosen a threshold $u$. A very important and also quite difficult task is the right choice of the threshold. We always have to take care that the threshold is neither "too high" nor "too low". If we pick a relatively low $u$, we will get a lot of exceedances of $u$ which seems to be good at first glance. But as mentioned above the sample could be not extreme value distributed in the "middle". This data will then lead to a wrong estimation of parameters. Opting for a high value of $u$, the problem occurs that it is just too less data to provide appropriate estimates. Especially the variance of the estimators becomes too large. So how do we chose an appropriate $u$? Therefore, we take a look at the mean excess function of the generalised pareto distribution, which we can calculate by using (3.2.4) where $\overline{F}(x) = \overline{G}_{\xi,\beta}(x)$ :

$$e(u) = \frac{1}{\overline{G}_{\xi,\beta}(u)} \int_u^{x_G} \overline{G}_{\xi,\beta}(x)dx$$

$$= \left(1 + \frac{\xi u}{\beta}\right)^{\frac{1}{\xi}} \left(-\frac{1}{\xi} + 1\right)^{-1} \frac{\beta}{\xi}\left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi} + 1}\Bigg|_u^{x_G}$$

$$= \frac{\beta + \xi u}{1 - \xi}. \tag{3.3.3}$$

As the empirical mean excess function we have

$$e_n(u) = \frac{1}{\overline{F}_n(u)} \int_u^\infty (x - u) \, dF_n(x)$$

$$= \frac{1}{N_u/n} \sum_{i=1}^n (X_i - u) \, \mathbf{1}_{\{X_i > u\}} \frac{1}{n}$$

$$= \frac{1}{N_u} \sum_{i:X_i > u} (X_k - u). \tag{3.3.4}$$

So for large $u$

$$\overline{F}_u(y) \approx \overline{G}_{\xi,\beta} \rightarrow e_n(u) \approx e(u). \tag{3.3.5}$$

As $e(u)$ is a linear function of u, and we know that $e_n(u) \approx e(u)$ for large $u$, we just have to look at the mean excess plot

$$\left\{ \left( X_{(k)}, e_n(X_{(k)}) \right) : k = 2, ...., n \right\} \tag{3.3.6}$$

and choose the smallest $u$ where it is linear.

## 3.4   The Hill estimator

Besides the POT method there are different ways of estimating the shape parameter of the GPD, for example the Hill-estimator ([1], p. 330-339). It estimates $\xi$ under the assumption that the underlying sample $X_1, ..., X_n$ is iid with df $F \in \text{MDA}(\Phi_\alpha)$ where $\xi = \frac{1}{\alpha}$, i.e. the sample is obeying a Fréchet distribution. It should be noticed that the estimator can be obtained by different approaches, e.g. the MLE approach, the regular variation approach or the mean excess function approach. This assures its functionality. The MLE approach will be shown here in a few steps. Again, we denote $X_{(1)} \geq ... \geq X_{(n)}$ as the ordered sample. From $F \in \text{MDA}(\Phi_\alpha)$ we know that the tail of $F$ is regularly varying and we assume that $c = 1$, then it holds for $X \sim F$:

$$P(X > x) = \overline{F}(x) = x^{-\alpha}, x \geq 1, \quad \Longrightarrow \text{for} \quad Y = \ln X \quad \text{we have} \quad P(Y > y) = e^{-\alpha y}.$$

This can be generalized by

$$P(X > x) = \overline{F}(x) = cx^{-\alpha} = \left(\frac{x}{u}\right)^{-\alpha} \quad \text{for} \quad x \geq u > 0, \, c = u^\alpha \, \text{and} \, Y = \ln\left(\frac{x}{u}\right).$$

Then Y is exponentially distributed and by MLE we get:

$$\widehat{\xi} = \frac{1}{\widehat{\alpha}} = \frac{1}{n} \sum_{j=1}^n \ln X_{(j)} - \ln u$$

But as $X \sim F$ often only holds for the upper extremes, we have to do a MLE on the $k$-th upper order statistics. If we choose either $u = X_{(k)}$ or $k = \#\{i \leq n : X_i \geq u\}$, we obtain the Hill-estimator for $\xi$:

$$\widehat{\xi}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_{(j)} - \ln X_{(k)} \tag{3.4.1}$$

Moreover, we get the estimation for $c$ by $\widehat{c} = \frac{k}{n} X_{(k)}^{1/\widehat{\xi}^{(H)}}$, so the tail is given by

$$\widehat{\overline{F}}(x) = \frac{k}{n} \left( \frac{x}{X_{(k)}} \right)^{-\frac{1}{\widehat{\xi}^{(H)}}}$$

and the corresponding p-quantile by

$$\widehat{x}_p = \left( \frac{n}{k}(1-p) \right)^{-\widehat{\xi}^{(H)}} X_{(k)}. \qquad (3.4.2)$$

As above, the last formula is used to estimate the $\mathrm{VaR}_{0.99}$ with $\mathrm{VaR}_\alpha = x_\alpha$. For analysis of the threshold $u$, the so called Hill-plot can be applied. For $\xi$ it is defined as

$$\left\{ \left( k, \xi^{(H)} \right) : k = 2, ..., n \right\}$$

The value of $k$ should be chosen out of a region where the plot is "almost constant", since the estimator is stable in this area.

# 4    Analysis of the market risk in the Basel III framework



Figure 5: The Goldman Sachs stock price from 2007-04-28 until 2012-05-23.
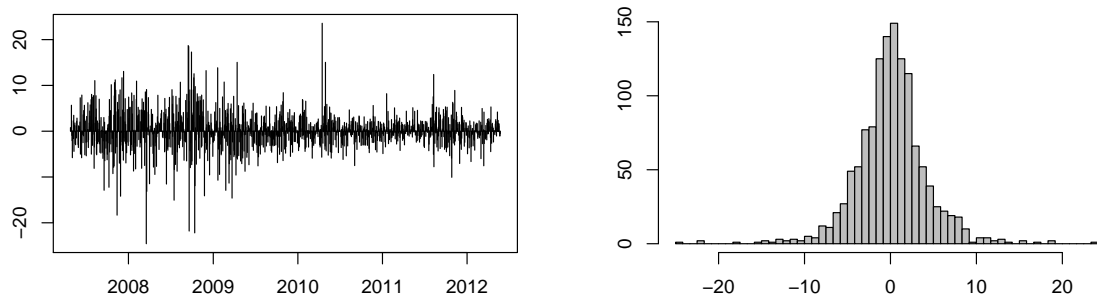


Figure 6: The Goldman Sachs stock negative logarithmic returns (left) in 2007-2012 and the corresponding histogram.

Having all the mathematical devices we can start with the statistical analysis. The quantitative standards of the Basel III market risk frameworks define some basic rules for the calculation of minimum capital requirements, but the banks still have flexibility in choosing the underlying statistical model. Here, the VaR estimation is done with extreme value statistics as explained in the previous section. To analyse the different ways of estimation as well as the historical development of the VaR, the daily returns of Goldman Sachs Group inc. are used as the underlying data. In Figure 5 you can see the chart of

the stock prices. In Figure 6 the negative returns and the corresponding histogram of the daily negative returns between 2007 and 2012 are depicted. The negative returns are given by $(S_{t-1} - S_t)$ where $S_t$ is the stock price at time $t$ and $S_{t-1}$ the stock price one day before. The negative returns are used instead of the positive because we want to assess the big losses and our statistical method used is based on the maxima of a sample, not the minima. As you can also see in Figure 5 the Goldman Sachs stock is varying a lot, which is related to the financial problems during the crisis, contrary to normal market behaviour. In this section the daily VaR of the last 5 years is estimated via the POT method, the Hill-estimator and a method combining both. The data we use can be downloaded from `http://finance.yahoo.com/q/hp?s=GS&a=00&b=3&c=2005&d=04&e=23&f=2012&g=d`. We begin the analysis in `R`:

```
gs<-read.csv("C:/.../gs.csv")
date.gs<-strptime(gs[,1],format="%Y-%m-%d")

head(gs)
        Date    Open   High    Low Close  Volume Adj.Close
1 2012-05-23  96.75  98.51  95.51 98.04 5469700     97.57
2 2012-05-22  96.88  99.95  96.59 97.53 5839700     97.07
3 2012-05-21  94.80  97.55  94.00 96.51 5314700     96.05
4 2012-05-18  97.53  97.90  95.00 95.49 7184700     95.04
5 2012-05-17  98.57  99.94  97.08 97.08 6375200     96.62
6 2012-05-16 100.55 102.14  98.20 98.20 5797200     97.73

ret.gs<-diff(gs[,5])
ret.gs0<-ret.gs[ret.gs>0]

head(ret.gs)
[1] -0.51 -1.02 -1.02  1.59  1.12  1.67
```

For the further analysis the difference of the closing prices of each day (`ret.gs`) are the negative returns of the stock. Thus, by `ret.gs0` the one day losses of the stock are given. Based on the data analysis we achieve a knowledge about the statistical behaviour of the sample. Subsequently the daily VaR for every day of the period between 2007-04-27 and 2012-05-23 will be estimated based on past values. Afterwards we will choose a period of significant stress to calculate the SVaR. Finally we will accomplish the different terms of the capital requirements of the Basel III framework calculated for the last 5 years, but by means of analysing, it is done with the one-day-VaRs instead of the 10-day-VaRs which is suggested in the Basel III framework.

## 4.1 General Data Analysis

We start the statistical analysis with an explorative analysis of the underlying data. As previously mentioned, we use extreme value methods, because our quantile estimation is based on the large observations of our sample. An empirical cumulative distribution function of the losses only provides a rough shape of the distribution. A more valuable tool for a detailed examination is the Quantile-Quantile plot. Thereby, we look at the

whole set of the data and not only on the data of one year, as we will do later to calculate the VaR. Having found an appropriate distribution of the whole sample we conclude that the data of one year has the same or similar distribution. We are aware that this is questionable, as we have to assume that our data is iid, which is actually not the case. Having a sample $X_1, ..., X_n$ and the corresponding ordered sample is $X_{(1)}, ...., X_{(n)}$ where $X_{(1)} \leq ... \leq X_{(n)}$ the QQ-plot is given by

$$\{(F^{\leftarrow}(\tfrac{k}{n+1}), X_{(k)}) : k = 1, ..., n\},$$

i.e. theoretical values of the chosen distribution $F$ are plotted against the values of the ordered sample. If the plot is roughly linear we can assume that $F$ is the distribution of our sample. This is stated in the Glivenko-Cantelli Theorem:

**Theorem 4.1.1.** (Glivenko-Cantelli theorem)
*Let $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq x\}}$, $x \in \mathbb{R}$ be the empirical distribution function of a sample $X_1, ..., X_n$, then it holds for $n \to \infty$*

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{a.s.} 0.$$

Note that, even if our sample is not appropriate scaled and centered, it does not have an impact on the linearity of the plot. That is because the location and scale parameter do not change the shape of our distribution. You can even estimate the intercept and slope watching the plot. This is, of course, not a very exact way of estimation.

In Figure 7 the quantiles of the whole sample are plotted versus different theoretical quantiles: the normal quantiles, exponential quantiles and Fréchet quantiles with $\xi = 0.1$ and $\xi = 0.5$. A QQ-plot of the threshold data (using 180 upper values) against the generalized Pareto distribution with different $\xi$'s is presented in Figure 8. The corresponding R code is given above each plot.

```
qqnorm(ret.gs)
qqplot(rexp(10000),ret.gs0)
qqplot(rfrechet(10000,shape=10),ret.gs0)
qqplot(rfrechet(10000,shape=2),ret.gs0)
```
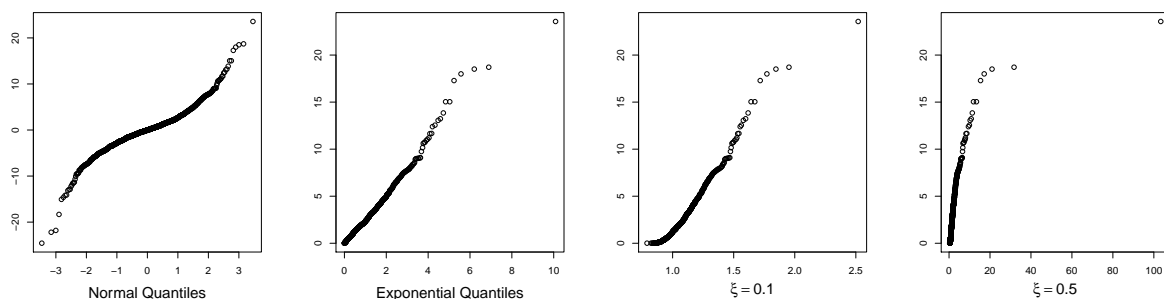


Figure 7: QQ-plots of the Goldman Sachs daily negative returns during 2007-2012, with quantiles of the normal, exponential and generalized extreme value distribution with $\xi = 0.1$ and $\xi = 0.5$ (from left to right).

```
threshold<-findthresh(ret.gs0,ne=180)
x<-c(0,0.1,0.2,0.5)
for (i in 1:4)
qplot(ret.gs0,xi=x[i],threshold=threshold)
```
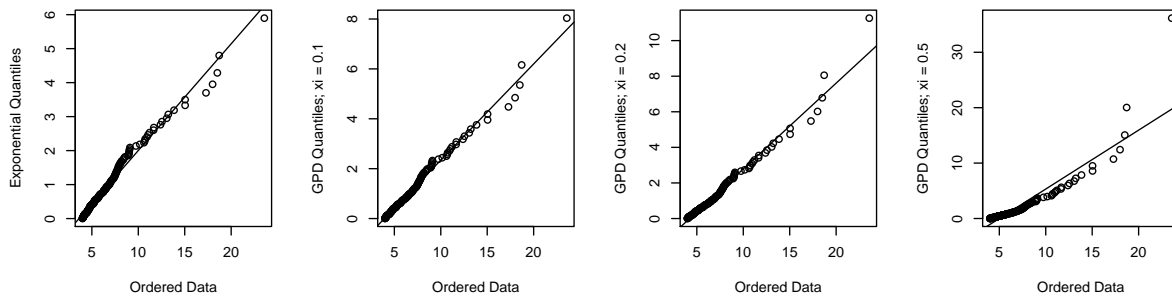


Figure 8: QQ-plots of the exceedances of u of Goldman Sachs daily negative returns during 2007-2012, with quantiles of the generalized pareto distribution with different choices of $\xi$: $\xi = 0$, $\xi = 0.1$, $\xi = 0.2$, $\xi = 0.5$ (from left to right).

The first thing we can observe from the left graph in Figure 7 is, whether our sample is heavy or light tailed compared to a normal distribution. As we can see, the right tail values are more increasing than the middle values, which means that the tail-distribution of the data is heavier tailed than a normal distribution. In the second graph from the left the data is plotted against the exponential distribution, where we can observe a more or less straight line, despite some points which are a little higher. Then the sample quantiles are plotted against Fréchet quantiles where $\xi = 0.1$, which is also a good fit. For the Fréchet quantiles of the right graph, we chose $\xi = 0.5$, which is too heavy tailed for our sample. But one should keep in mind that here we examine the behaviour of all positive values of our sample, or all values in the case of the Normal quantiles.

However, our goal is to determine a distribution especially of the large values of our sample. Thus, the logical step is to actually contemplate only these values. In other words, we look at the excess distribution function of our sample. Thus, we choose a threshold so that 180 of 1861 values lie above the threshold, i.e. approximately 9.7 % of our whole data set, or 180 out of 929 positive values, i.e. approximately 19.4 %. The threshold is then given by $u = 3.99$. Now this threshold data is plotted against the GPD with different $\xi$, where $\xi = 0.5$, $\xi = 0.2$, $\xi = 0.1$, and $\xi = 0$ (see Figure 8). Comparing the four different graphs with increasing $\xi$ from left to right, $\xi = 0$ is the best fit for our threshold data. That means that we can assume that the tail of our sample is obeying an exponential distribution. We should be aware of the fact that we used the whole sample and get a good fit for a certain $\xi$. Later on, when we look only at a part of our sample, the estimation of $\xi$ will vary for each period. Nevertheless, we can conclude that, when we examine the excesses of a certain high threshold, we have a sample which is generalized pareto distributed.

## 4.2   Estimation of the daily Value at Risk

In the previous section we made a general data analysis with the whole range of our data. Now we want to calculate the daily one-day VaR. For each particular date we will use the negative returns of the 260 preceding business days to estimate the VaR. The estimation will be based on different methods of extreme value statistics. All methods require an appropriate choice of numbers of extreme values which we include in our estimation. This is an important questions, and will be discussed more in detail in section 4.2.1. Having chosen a threshold, we will describe the different methods and compare their results over the last five years. This will be done by using the 1-day-log-returns of the Goldman Sachs stock and compared to the VaR estimation based on the normal distribution.

### 4.2.1   Choice of the threshold

First of all, we will discuss how to decide which threshold could be appropriate or, in other words, how many upper values do we assume to be extreme value distributed. To investigate the right choice of a threshold value, as described in section 3.3.4 we will first use the standard tool: The mean excess plot, which is given by $\left\{ \left( X_{(k)}, e_n(X_{(k)}) \right) : k = 2, ...., n \right\}$, cf. (3.3.6), i.e. we plot the sample against the empirical mean excess function (3.3.4). In Figure 9, we depict the mean excess plot of the positive values of our sample. We omitted the last 7 values, because they are affected by boundary effects which are not representative. This is because upper plotting points are the average of only a handful of extremes, and the last point is always on the x-axis. This downwards trend at the end should not be taken into account.
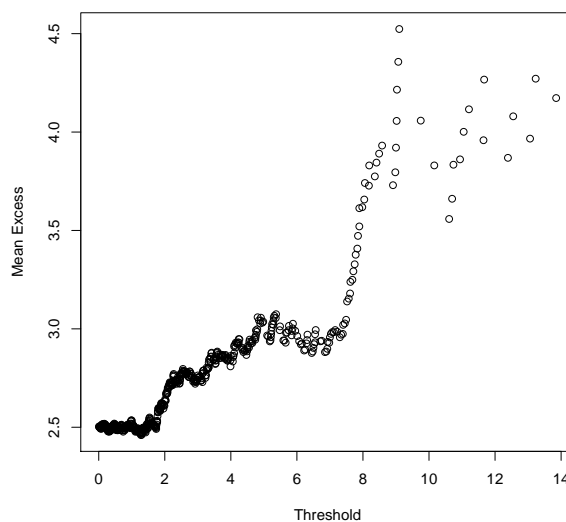
```
meplot(ret.gs0,omit=7)
```



Figure 9: Mean excess plot of the positive values of our sample, which consists of the negative returns of the Goldman Sachs stock during 2007-2012.

We observe that our plot fluctuates a lot, but is more or less increasing beginning from $u \approx 2.5$ and one could add a linear regression line which approximates the behaviour of the plot. This upwards trend shows heavy tailed behaviour of our sample. But still the points have large variation from the straight line for higher values of $u$. The next step of our threshold analysis is to observe the change of the shape parameter and the change of the quantile with a change of the threshold. In Table 1 we compare the outcomes of the GPD fit to our data when we choose different numbers of extremes. The comparison is performed for the VaR of 2010-07-01, which means that the preceding 260 business days are used as our sample to estimate the VaR (cf. 3.1), i.e. the negative returns from 2009-06-19 until 2010-06-30. We can see that the estimated shape parameters $\xi$ and scale parameters $\beta$ vary with different choices of the threshold, but the quantile which we estimate on basis of (3.3.2) does not change significantly. We show this exemplarily for the number of extremes $k = 26$.

```
date<-which(gs[,1]=="2010-07-01")
ret10.gs<-ret.gs[(date+1):(date+260)]

gpd.26<-gpd(ret10.gs,nextremes=26)
> gpd.26$p.less.thresh
[1] 0.9
> gpd.26$threshold
[1] 3.74
> gpd.26$par.ests
       xi      beta
0.4158144 1.5129573

> riskmeasures(gpd.26,0.99)
        p quantile    sfall
[1,] 0.99 9.580002 16.32669
```

|  | Number of extremes | | |
|---|---|---|---|
|  | 18 | 26 | 34 |
| Percentage above threshold | 6.9% | 10% | 13.1% |
| Threshold | 4.1 | 3.74 | 3.33 |
| $\xi$ | 0.2935115 | 0.4158144 | 0.414545 |
| $\beta$ | 2.2614619 | 1.5129573 | 1.367307 |
| $\text{VaR}_{0.99}$ | 9.990852 | 9.580002 | 9.60661 |

Table 1: Comparison of the estimates of a GPD fit to the negative returns from 2009-06-19 to 2010-06-30 choosing a different number of extremes.

In Figure 10, the shape parameter as well as the $\text{VaR}_{0.99}$ is plotted versus the number of extremes we use for fitting the GPD. The left column is based on the data for the VaR of 2008-07-01, a period of relatively normal market behaviour before the crisis. The middle two graphs are based on the data of 2009-07-01, which is right after a year of high volatility on the market (see Figure 6). The data for the VaR-estimation of 2010-07-01 is used for the two right graphs and this date coincides with the date of Tabular 1. This is

after the market has more or less calmed down. For all dates we can observe that based on the POT method the 99%-VaR does not really depend on the choice of the threshold $u$. However, what we want to emphasize is that the 95% confidence boundaries of our estimates for $\xi$ and the VaR are very large (see the red lines). In the upper graphs of Figure 10 we see that the estimated $\xi$ depends considerably on the date of the estimated VaR. But all three lines are more or less constant around a number of extremes $k = 26$.

```
date<-which(gs[,1]=="2009-07-01")
ret09.gs<-ret.gs[date:(date+260)]

shape(ret09.gs,start=17,end=45, reverse=FALSE)
quant(ret09.gs,p=0.99,start=17,end=45, reverse=FALSE)
```
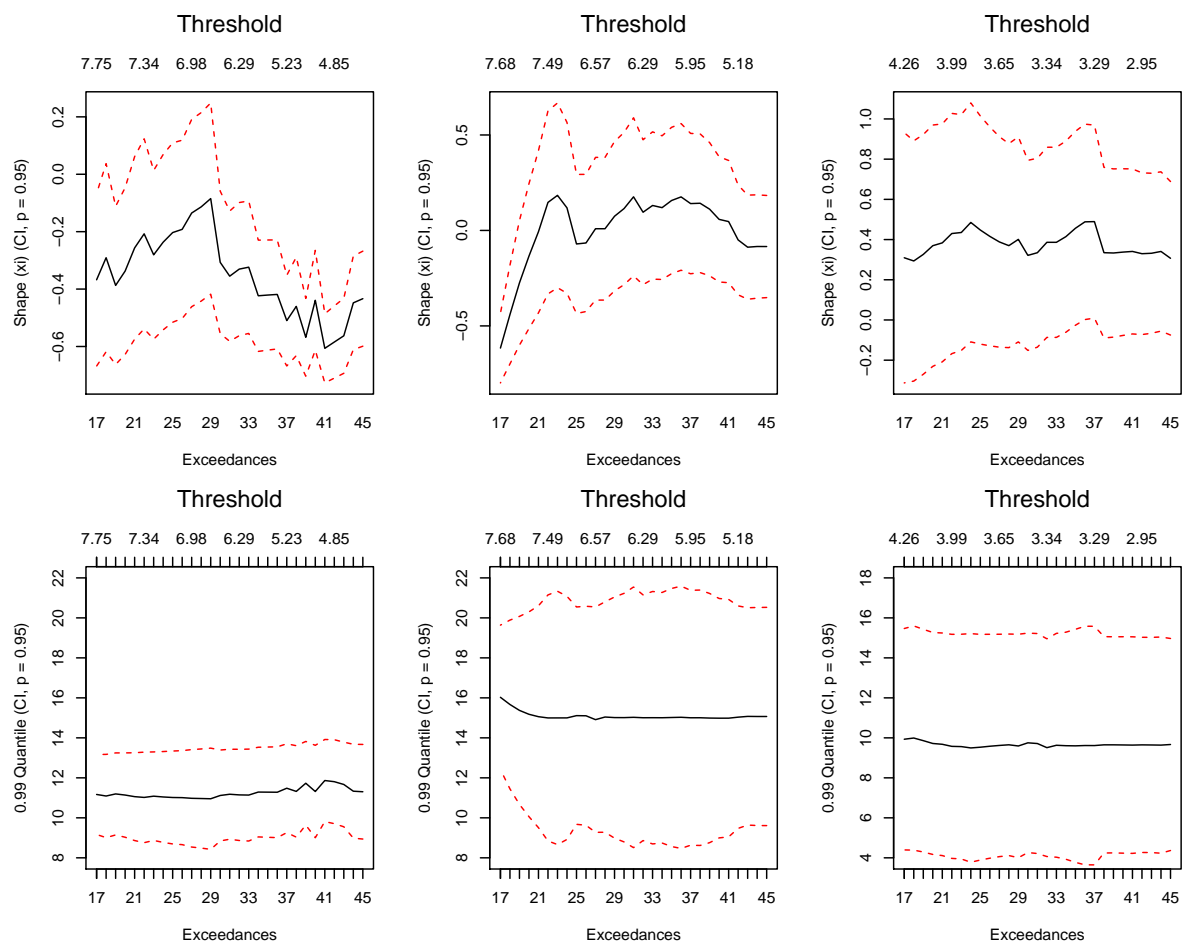


Figure 10: The shape parameter $\xi$ of the GEV and quantile when fitting the negative returns to a generalised pareto distribution dependent on the number of extremes used for estimation, based on the data for the VaR estimation from the following dates 2008-07-01 (left), 2009-07-01 (middle), 2010-07-01 (right).

Now we want to look at a different way to estimate the quantile: The Hill estimator. As described in 3.4 the shape parameter can be estimated by (3.4.1) and subsequently the quantile by (3.4.2). Using the function `hill` we can choose whether we want to plot $\xi$,

$\alpha$ or the quantile as a function of the threshold or the number of extremes. For a fixed date, this is done by

```
date<-which(gs[,1]=="2009-07-01")
ret09.gs<-ret.gs[date:(date+260)]

hill(ret09.gs,option="xi",start=15,end=45)
hill(ret09.gs,option="quantile",start=15,p=0.99,end=45)
```
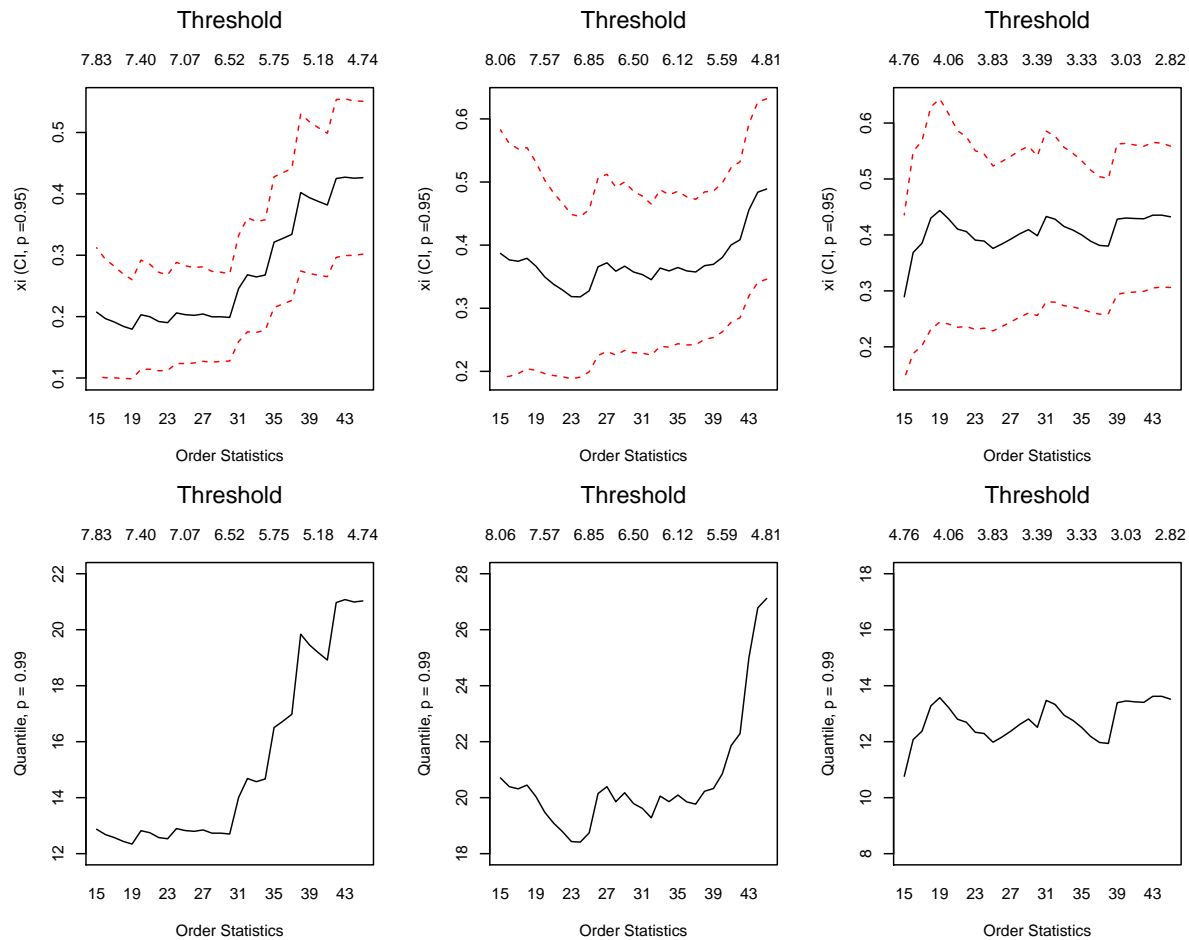


Figure 11: The shape parameter $\xi$ of the GEV and the quantiles estimated with the Hill-estimator dependent on the number of extremes used for estimation, based on the data for the VaR estimation from the following dates 2008-07-01(left), 2009-07-01(middle), 2010-07-01(right).

Note that the Hill estimator fits the data only to a Fréchet distribution not to the larger class of GEV distribution. In Figure 11 the shape parameter $\xi$, estimated by the Hill-estimator, as well as the estimated quantile as a function of the threshold are plotted for different dates. The same dates are chosen as in Figure 10. To make an appropriate choice of threshold, we should look at the upper three plots. In order to find an appropriate threshold with the help of the Hill plots, we search for a region where the graph is approximately constant. The middle and the right graph can be seen as more or less
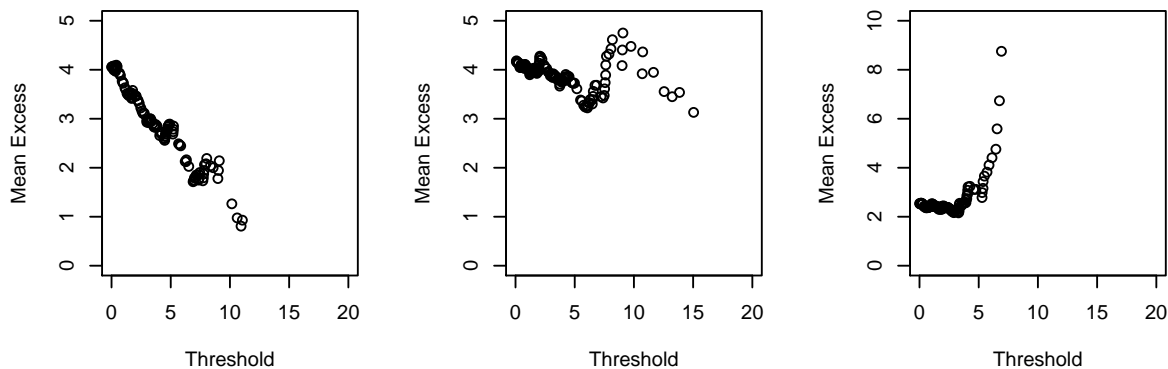
Figure 12: Three mean excess plots based on the data for the VaR estimation of the following dates: 2008-07-01 (left), 2009-07-01 (middle), 2010-07-01 (right).

constant in the whole range of our plot. The left graph, however, has a step around a number of extremes $k = 29$ and we might say that it is constant between 17 and 29. We could chose 26 exceedances as a value in the middle of the constant area. Contrary to the lower graphs in Figure 10, the estimated quantile depends on the number of extremes. Taking into account more excesses provides us a higher VaR estimate as we can see in the two left plots of Figure 11. But in the region around 26 extremes they are all consistent. It should be argued that we want to estimate the 99% VaR which means that only one percent of the data, here around "2.6" values, should lie above the VaR. Thus, we should also look at range of data around this 1% quantile. Consequently, the upper 10% (i.e. 26 of 260 values) of the ordered sample is better than using more values.

Another tool to examine the threshold or the behaviour of the sample is the mean excess plot which is depicted for the whole sample in Figure 9. Now it is plotted again for the three different dates in Figure 12. The plots coincide with the results of Figure 10: The left graph is decreasing, which means the underlying sample is light tailed, and a negative $\xi$ is estimated. The graph in the middle is more or less constant, which indicates a sample which is exponentially distributed. As the right graph is increasing, the underlying data is clearly heavy tailed. This is based on the data for the VaR of 2010-07-01, which means the data from 2009-06-19 until 2010-06-30. Based on the 26 largest values of this sample we get $\widehat{\xi} = 0.416$ and based on this estimate Figure 13 depicts the excess distribution function, the tail of the distribution function and a QQ-Plot of the data against general pareto quantiles.

We can obtain the plots in Figure 13 by

```
date<-which(gs[,1]=="2010-07-01")
ret10.gs<-diff(gs[date:(date+260),5])

k<-26
threshold<-findthresh(ret10.gs,ne=k)
gpd.10<-gpd(ret10.gs,nextremes=k)
plot.gpd(gpd.10)
```
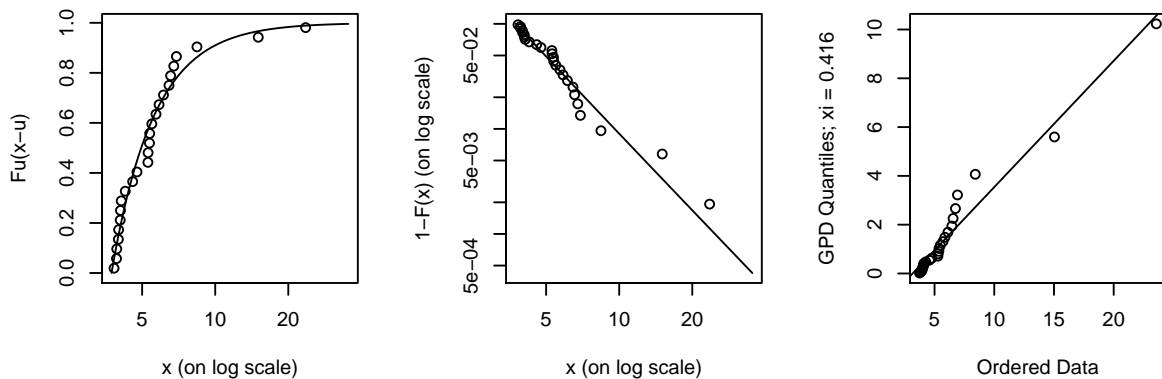
```
1
2
0
qplot(ret10.gs,xi=0.416,threshold=threshold)
```



Figure 13: The excess distribution function (left), the tail of the distribution function (middle) and the QQ-plot with Fréchet quantiles with $\widehat{\xi} = 0.416$ (right) based on the 26 largest values of the data from 2009-06-19 until 2010-06-30, the data for the VaR of 2010-07-01.

### 4.2.2   Parameter and VaR estimation with different methods

In the following section the shape parameter $\xi$ as well as the scale parameter $\beta$ will be estimated by different means of extreme value statistics, described in the mathematical basics, section 3. The estimation for the daily VaR is based on the largest 26 values out of a set of 260 values, the returns of the 260 preceding business days. As the analysis in 4.2.1 has shown, this is an appropriate choice of number of extremes. Based on these selected extremes, the parameter estimates can be determined. Subsequently, the $\text{VaR}_{0.99}$ will be estimated for each day of the last 5 years (the period of 2007-04-27 until 2012-05-23). The VaR of the first 260 days is estimated based on the data which is only partly depicted in Figures 14-16 since we always need 260 preceeding daily returns for the estimation. Therefore, we used returns beginning from 2006-01-18. The different VaRs are compared to the estimated $\text{VaR}_{0.99}$ based on the normal distribution.

For a fit of the normal distribution, the empirical mean and variance were estimated based on the previous 260 days. Since the mean is always close to zero, setting $\mu = 0$ would not change the results significantly, but we will nevertheless use the empirical mean. Consequently, we assume that the returns of each of the 260 days preceding day $i$ to be $\mathcal{N}(\mu_i, \sigma_i^2)$ distributed and we can estimate the 99% normal quantile as $\text{VaR}_{0.99}^N(i)$. The R code below shows how to estimate each VaR for one day $i$. Figure 14 - 16 have been obtained by using a loop of $i$ over the whole period (2007-04-27 until 2012-05-23). Therefore, we defined $i$ as the date of the VaR, `ret` as the negative returns of the

preceding 260 business days of $i$, `retsort` as the decreasing sorted `ret` and `k` as the number of extremes used for the estimation.

```
i<-which(gs[,1]=="2009-07-01")
ret<-ret.gs[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
k<-26
```

And the normal $\text{VaR}_{0.99}^N(i)$ can then be derived with:

```
VaRnorm<-qnorm(0.99,mean=mean(ret),sd=sd(ret))
```

### Method 1: Classical POT method

We begin with the first extreme value method, where $\xi$ and $\beta$ are estimated based on the POT method, which was presented in detail in section 3.3. Thereby $\widehat{\xi}$ and $\widehat{\beta}$ are determined by maximum likelihood estimation and put in formula (3.3.2) to calculate the $\text{VaR}_{0.99}$. In Figure 14 the POT-VaR (blue) is plotted as well as the Normal-VaR (black) and the daily negative returns. The outcomes of the VaR based on the POT method are higher than the outcomes of the normal distribution. This is due to the fact that the POT method uses the GEV distribution, which is more heavy tailed than the normal distribution and thus delivers higher quantiles. Moreover, it was observed that in some periods, the POT method provides negative $\widehat{\xi}$, which means that the estimated distribution has a finite right end point. But the quantiles are still appropriate. This will be discussed in more detail at the end of this chapter. The POT-VaR is estimated by:

```
gpd<-gpd(ret, nextremes=k)
q<-riskmeasures(gpd,0.99)[2]
```
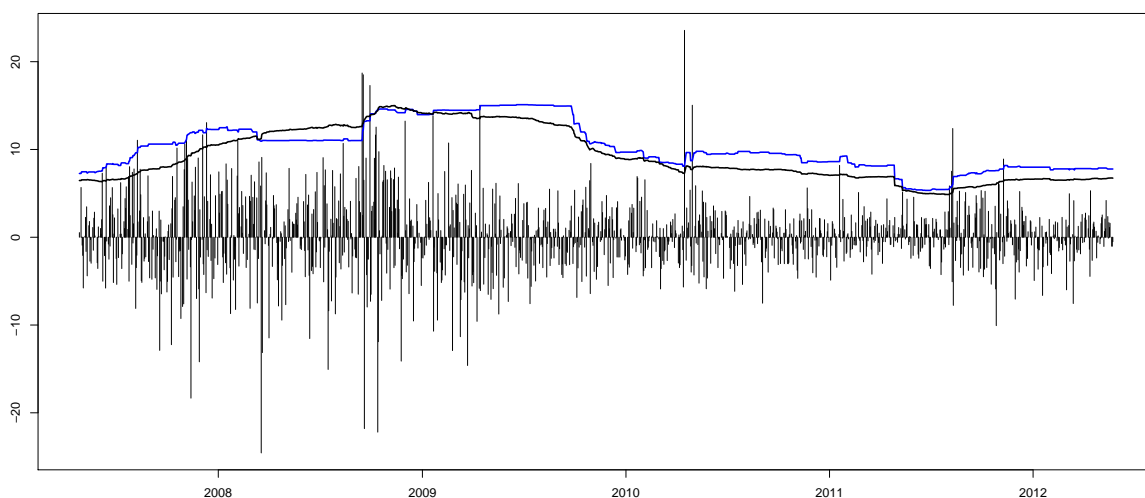


Figure 14: The dynamical development of the VaR estimated by the POT-method (blue) and the normal distribution (black).

**Method 2: Hill estimation**

Another classical extreme value method to estimate the VaR is the Hill-estimator, as explained in section 3.4. We will first estimate the shape parameter $\widehat{\xi}^{(H)}$ of the GEV distribution by formula (3.4.1). Then the VaR can be derived from formula (3.4.2). In R there is no function where the shape parameter is estimated by the Hill-estimator, so we use the formulas itself. It is important to note that no scale parameter is included in this estimation and the quantile is just based on the estimated $\widehat{\xi}^{(H)}$ and the threshold, which is the $k$-th largest value of the ordered sample $u = X_{(k)}$. Especially in the beginning of our observed period the outcomes of the Hill estimation are significantly better than the Normal-VaR. This can be observed in Figure 15, where the VaR by the Hill-estimator (red line) is higher than many losses, while the Normal-VaR is not.

```
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
q<-((260/k*(1-0.99))^(-shape))*retsort[k]
```
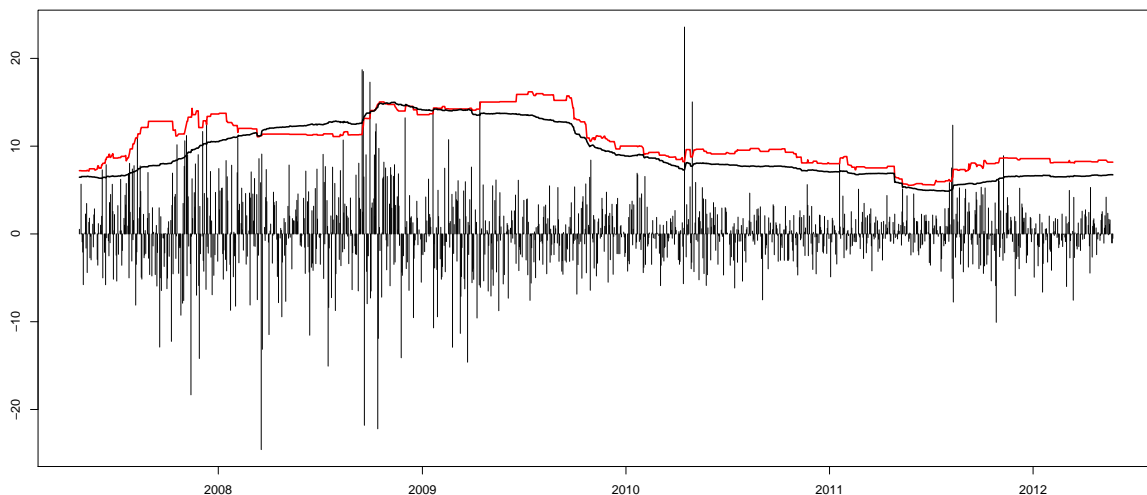


Figure 15: The dynamical development of the VaR, where $\xi$ is estimated by Hill-estimator (red) and by the normal distribution (black).

**Method 3: Combining Hill-estimation and POT method**

As in method 2, we will again use the Hill-estimator to determine the shape parameter $\xi$. Additionally, the shape parameter $\beta$ is estimated by the POT method. Therefore the estimate $\widehat{\xi}^{(H)}$ is put in the log-likelihood function of the GPD, $l(\widehat{\xi}^{(H)}, \beta, Y_1, ..., Y_{N_u})$ (3.3.1) and $\beta$ is then estimated by maximum likelihood estimation given $\xi = \widehat{\xi}^{(H)}$. This estimation is again compared to the normal distribution in Figure 16. The values of the VaR fluctuate more than the Normal-VaR. Especially at the beginning of 2007, the VaR is even higher than all occurred losses. When we determined the shape parameter by the Hill-estimator, we use the function `fitgpd` instead of the function `gpd`, because one parameter, in this case $\xi$, can be fixed.

```
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
gpd<-fitgpd(ret,threshold=retsort[k],est="mle",shape=shape)
scale<-gpd$scale
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
```
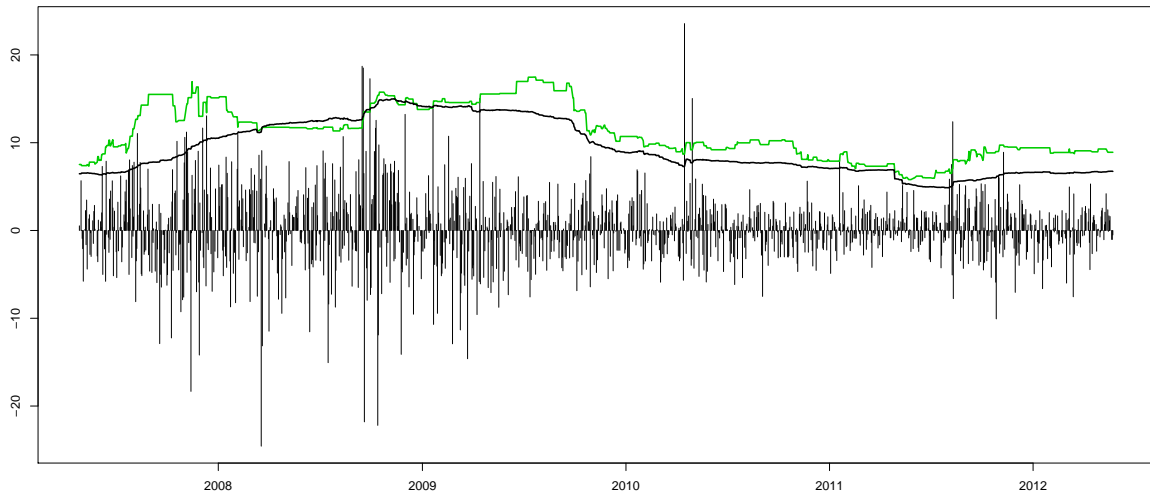


Figure 16: The dynamical development of the VaR, where $\xi$ is estimated by the Hill-estimator and $\beta$ by the POT-method, given $\hat{\xi}^{(H)}$ (green) and the normal VaR (black).

**Comparison of the different estimated shape parameters**

It is intressting to observe the different outcomes of the estimated shape parameter $\widehat{\xi}$ based on the different methods. In Figure 17 the different outcomes for the last five years are depicted. First, $\xi$ is estimated by the classical POT method (blue) and, secondly, $\xi$ is estimated by the Hill estimator (red). It is hard to explain why the POT-method delivers $\widehat{\xi}$'s with such a high variation. We would rather expect a more continuous line with a few steps. Moreover, it is negative in some areas, which indicates that we have a Weibull distribution. But we do not assume losses of a stock to have a certain finite threshold. This could be explained by the fact that the high losses, which are taken into the estimation, are all more or less the same in this areas and thus the MLE estimates a negative $\widehat{\xi}$ (see Figure 6). The shape parameter received from the Hill-estimator does not vary as much. But as we saw above both estimation methods provide relatively good estimates for the VaR. Which means that the VaR is estimated appropriate, as it would have absorbed most of the occurred losses. This will be discussed in more detail in section 4.2.4.
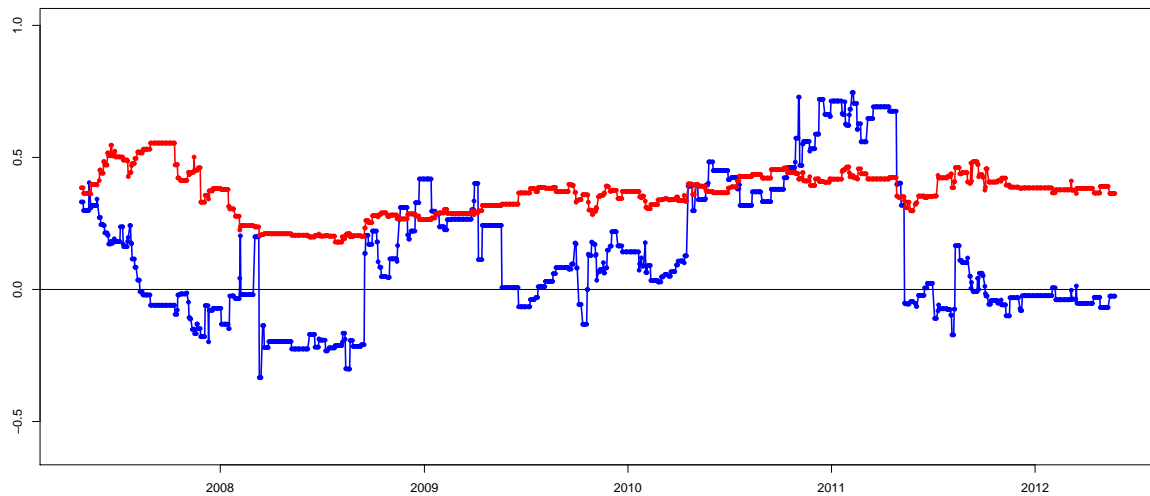
Figure 17: The dynamic development of the shape parameter $\xi$, where $\xi$ is estimated by the classical POT-method (blue dots) and the Hill-estimator (red dots).

### 4.2.3 Parameter and VaR estimation with scaled data

In the previous section we have assumed that 260 data points have the same distribution, and we have estimated $\xi_i$ and $\beta_i$ by methods 1 - 3 based on this assumption. But in reality every data point has a different distribution, so it could be problematic for statistical work to make this assumption. To take this into account, we scale the whole data set and derive:

$$\widetilde{X}_i = \frac{X_i - \widehat{\mu}_i}{\widehat{\beta}_i} \quad \text{where} \quad \widehat{\mu}_i := 0 \quad \text{and} \quad \widehat{\beta}_i := \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \widehat{\mu}_i)^2}.$$

Here $\widehat{\mu}_i$ is set to zero, because the empirical mean is so small that it is negligible and would not change the results significantly, and $\widehat{\beta}_i$ is determined by the empirical standard deviation. In R we get the scaled data set by

```
sigma1<-c()
for (i in 1: (length(ret.gs)-260)){
ret.help<-ret.gs[(i+1):(i+260)]
sigma<-sd(ret.help)
sigma1<-rbind(sigma1,sigma)
}
ret.neu<-ret.gs[1:(length(sigma1))]/sigma1
```

Now methods 1-3 are applied on $\widetilde{X}_i$ to estimate the VaR based on the scaled data $\widetilde{\text{VaR}}$. Therefore, we define ret as the scaled negative returns of the preceding 260 business days of $i$.

```
ret<-ret.neu[(i+1):(i+260)]
```

Now the different methods are applied as above. Only for method 1 we use the function `fitgpd` instead of the function `gpd`. The dynamic development of the scaled VaRs is depicted in Figure 18. Now we take this VaR and scale it backwards with $\mathrm{VaR}(i) = \widetilde{\mathrm{VaR}}(i)\widehat{\beta_i}$. This is done in Figure 19. We can observe that all three methods deliver more or less the same results.



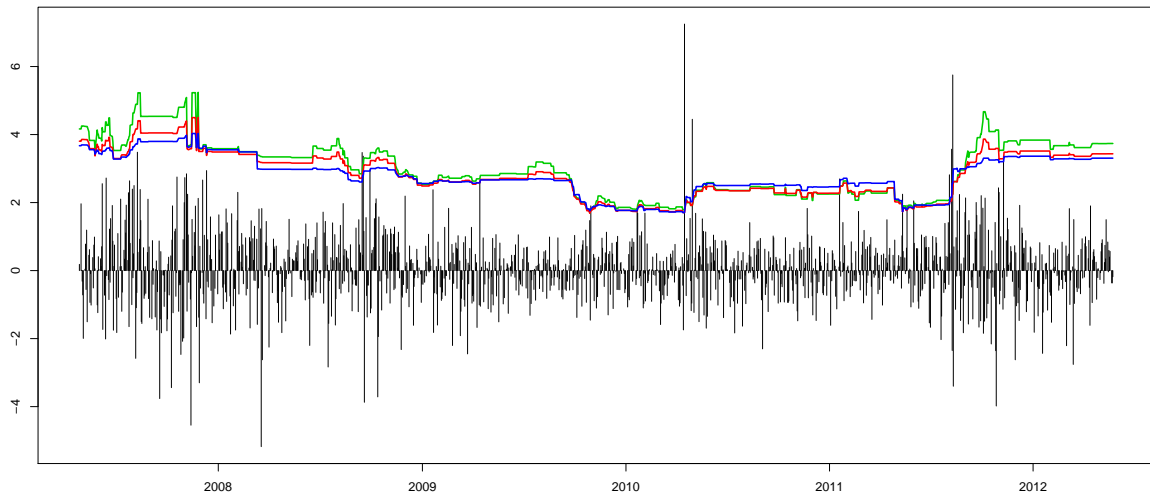Figure 18: The dynamical development of the VaR of the scaled data, based on the POT-method(blue), the Hill-estimator(red) and a combination of Hill-estimator and the POT-method(green).
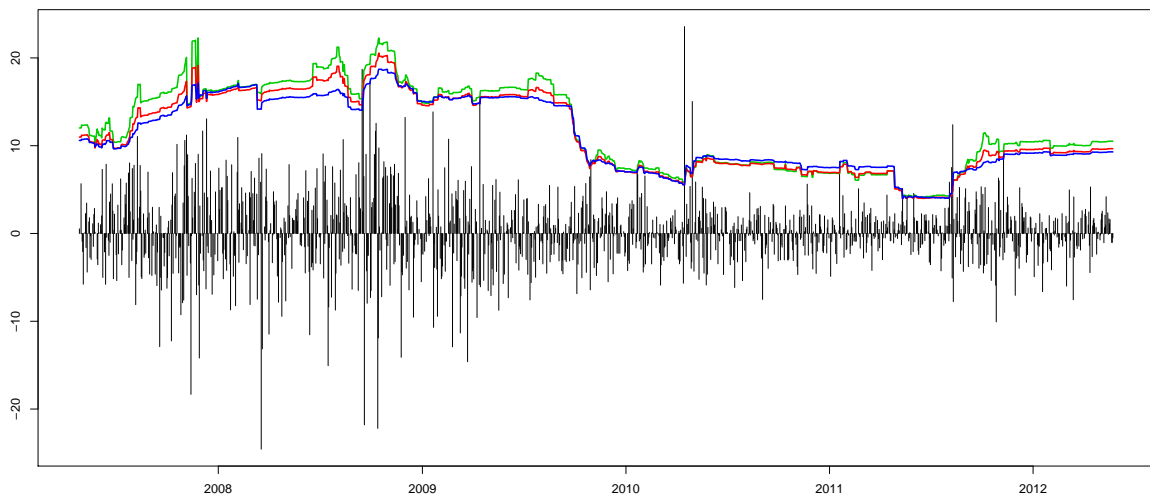


Figure 19: The dynamical development of the VaR, based on the POT-method (blue), the Hill-estimator (red) and a combination of Hill-estimator and the POT-method(green) of the scaled data, and scaled back after the estimation.

Again the different estimated shape parameters are plotted in Figure 20. Comparing the development of $\xi$ based on the scaled data with the development of $\xi$ in Figure 17, we observe that the trend is the same. But the $\xi$ estimates by the POT method fluctuates even more, when we use the scaled data for estimation.
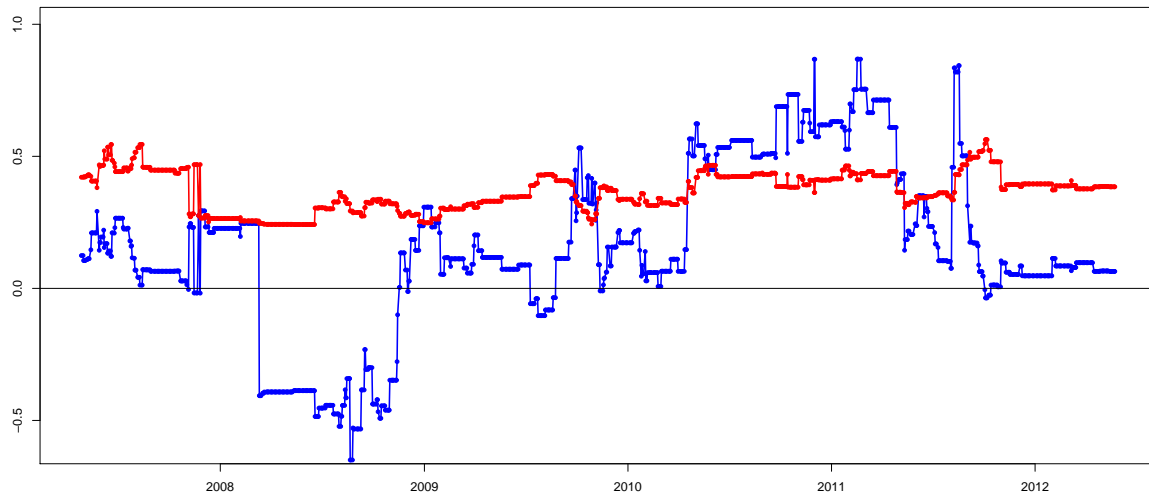


Figure 20: The dynamical development of the shape parameter $\xi$, where $\xi$ is estimated by the POT-method (blue dots) and the Hill-estimator (red dots) based on the scaled data set

### 4.2.4   Backtesting

Finally, we apply a simple backtesting procedure to our estimates. As we estimate the 99% VaR, the number of losses above the estimated VaR should be around 1% of all data, which cover 5 years, or exactly 1340 days. In Table 2 the different extreme value methods, as well as the method of using the normal distribution are compared. Using method 1 or 2 on the unscaled data, we have 13 excesses of the losses over the VaR, which is 1.02% of our data. This is the closest we get to 1%. This is very interesting, as we can see that the two classical methods in extreme value theory work very good. Even if we would expect that the scaling of the data would provide us better results, this is not the case. Nevertheless, all the methods deliver appropriate results. Moreover, it should be emphasised that regarding the percentage of losses exceeding the Normal-VaR, the estimated 99%-VaR is more a 98%-VaR. This means that the probability that we exceed the VaR is twice as high as it should be. This again confirms that the data is in fact more heavy tailed than the Normal distribution and, thus, it is not a good way of estimating the VaR.

|  | Method | # neg returns >VaR | in % |
|---|---|---|---|
| unscaled data | 1 POT | 13 | 1.015625 |
|  | 2 Hill | 13 | 1.015625 |
|  | 3 Hill & POT | 9 | 0.703125 |
| scaled data | 1 POT | 15 | 1.171875 |
|  | 2 Hill | 15 | 1.171875 |
|  | 3 Hill & POT | 12 | 0.9375 |
|  | Normal distr. | 25 | 1.953125 |

Table 2: Comparison of the backtesting results of the different methods for estimating the VaR during 2007-04-28 until 2012-05-23.

## 4.3   Additions to the daily VaR

The Basel III framework does not only require the calculation of the VaR. In addition to the VaR a so called stressed VaR (SVaR) is required, which is a new announcement compared to the Basel II framework. Besides this, a 60-days average VaR has to be calculated. This results in the formula for capital requirements in the Basel III framework:

$$C_t = \max\{\text{VaR}_{0.99,t-1}; m_c\text{VaR}_{0.99,avg}\} + \max\{\text{SVaR}_{0.99,t-1}; m_s\text{SVaR}_{0.99,avg}\} \quad \text{cf. (2.0.1)}$$

The different terms in the formula will now be described in more detail.

### 4.3.1   Estimation of the stressed Value at Risk

First of all, we consider the estimation of the stressed VaR. The Basel III framework requires that the SVaR is estimated based on a period of significant stress. It suggests a period around 2007/2008 if this adequately represents a stressed period. In the case of Goldman Sachs we observe in Figures 5 and 6 that the stock might have the highest volatility around 2008/2009. To find an adequate period, the maximal VaR of each method above is searched. A comparison of the outcomes and the corresponding dates of maximus are listed in Table 3. The 260-days data used for the estimation of the VaR of these exact day should be a period of significant stress. As we can see, the Normal-VaR reaches its maximum earlier than the VaRs of methods 1-3. The data used for the normal SVaR thus begins on 2007-10-31. In Figure 21 the different methods are plotted during the period of stress (2007-10-31 to 2009-07-23). We observe that all the extreme value methods deliver more or less the same results, but method 3 provides the highest.

| Method | max. VaR | dates of max. VaR |
|---|---|---|
| 1 POT | 15.10394 | 2009-06-19 to 2009-07-09 |
| 2 Hill | 16.19147 | 2009-07-10 to 2009-07-17 |
| 3 Hill & POT | 17.47609 | 2009-07-20 to 2009-07-23 |
| Normal. distr | 15.00873 | 2008-11-11 |

Table 3: Comparison of the maximal VaRs of the different methods and the corresponding dates where the maxima are reached.
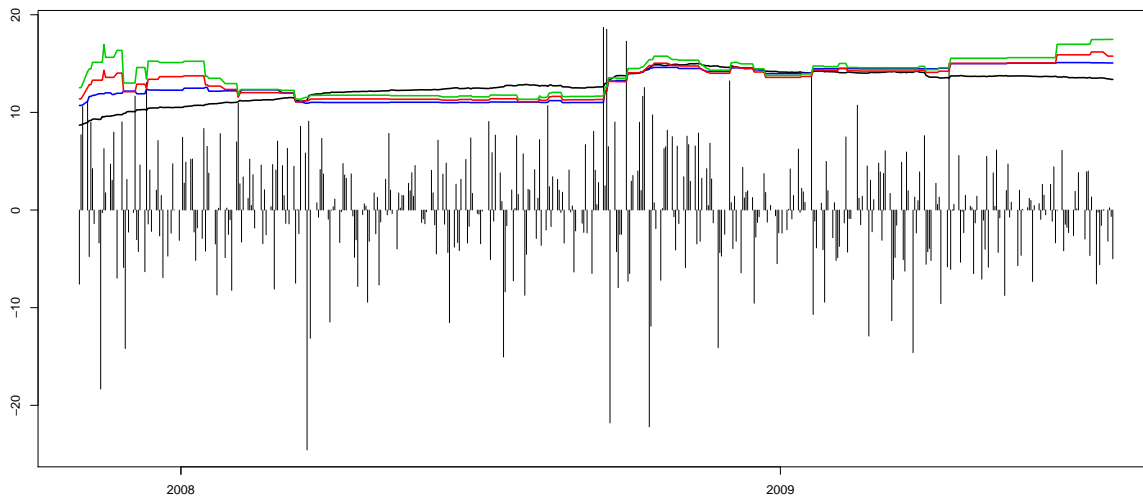
Figure 21:   The period of stress of Goldman Sachs and the VaRs by the POT-method(blue), the Hill-estimator(red) and the combination of Hill-estimator and POT-method(green) as well es the normal VaR(black).

### 4.3.2   Average of VaR and sVaR

In the following only the POT-method is used. Regarding formula (2.0.1), besides the VaR and the stressed VaR, an average VaR has to be calculated for both of them in order to determine the capital requirements in Basel III. As we did our whole analysis based on the 1-day and not 10-day log-returns, we can not determine the actual capital requirements, but we can analyse the formula based on our estimations of the 1-day-VaR. First, we calculate the average VaR, which is the average of the VaRs of the preceeding 60 business days, given by the blue line in Figure 22. Then we have to determine $\max\{\text{VaR}_{0.99,t-1}; m_c\text{VaR}_{0.99,avg}\}$, where $m_c = 3$. In fact $m_c$ can be chosen between 3 and 4, depending on the backtesting results. But here it is chosen as 3. The red line in this Figure represents $3\text{VaR}_{0.99,avg}$ and we can observe that at each point of time it holds that $\max\{\text{VaR}_{0.99,t-1}; m_c\text{VaR}_{0.99,avg}\} = m_c\text{VaR}_{0.99,avg}$. Thus, the same holds for the SVaR. This means that the capital requirements at time $t$ would be the sum of three times the $\text{VaR}_{0.99,avg}$ and three times the $\text{SVaR}_{0.99,avg}$. Regarding the capital requirements we would choose a certain SVaR , here for example from t=2009-07-09. The SVaR is given by $\text{SVaR}_{0.99,avg} = 15.02944$. This and $3\text{SVaR}_{0.99,avg} = 45.08832$ is marked as a point in Figure 22. Then $\max\{\text{SVaR}_{0.99,t-1}; m_c\text{SVaR}_{0.99,avg}\}$ is a constant which is added in the capital requirements of each day.
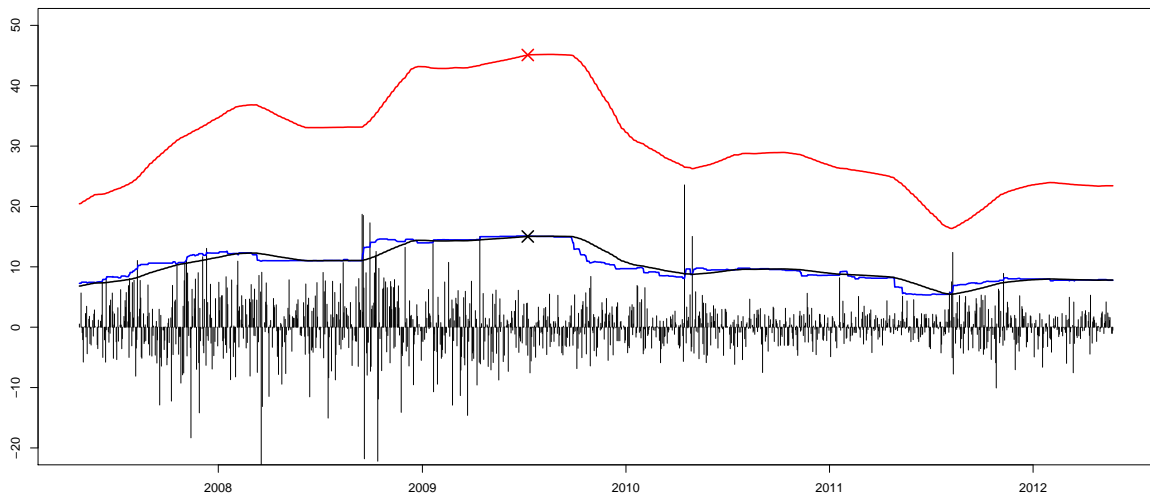
Figure 22: The dynamical development of the VaR estimated by the POT method(black), combined with the 60-days-average-VaR(blue) and 3 times the average VaR (red).

# 5   Conclusion

The goal of this thesis was to estimate the daily $VaR_{0.99}$ of the period 2008-04-27 until 2012-05-23 by different means of extreme value statistics. Moreover, the formula of the capital requirements in the Basel III framework was analysed based on this VaR estimation. Here we have to mention that the whole analysis was based on 1-day losses and, thus, we determined the 1-day VaR and not the 10-day VaR as it is required in the framework. Moreover, we did the estimation based on one single stock, and not a whole portfolio, how banks would do it. Thus, calculating the actual capital requirements would not make sense here. We would for example get higher capital requirements than the stock is worth, what is, of course, senseless. In reality banks have a very large portfolio and the variation is small in comparison with its value. But for our purpose to do a statistical analysis, it is satisfying to use one stock.

We first gave an overview of the mathematical topics the estimation is based on. The VaR and the basic steps in extreme value theory which lead to the estimation methods used were discussed.

Then the analysis was performed with the daily returns of the Goldman Sachs Group inc. stock which was chosen, because the stock has a high volatility which distinguishes the functionality of extreme value methods. We started with a general analysis of the sample, where we learned that the whole sample is exponential distributed. We calculated the VaR for each day in the sample period looking at the preceeding 260 business days. Based on these samples we used tools as the mean excess plot and the hill plot to examine over which threshold our data is extreme value distributed. We concluded that the 26 largest values can be fitted to a GEV. Then we firstly fitted the excesses over this threshold to a generalized pareto distribution by finding an appropriate shape parameter $\xi$ and scale parameter $\beta$ using the POT method. Secondly, we used the Hill estimator to estimate

a positive shape parameter $\xi$, which fits the data to a Fréchet distribution. The POT method delivers besides positive also negative $\xi$'s. This is actually not what we accepted, because a negative $\xi$ means that the underlying distribution is a Weibull distribution which has a finite right endpoint. Based on theses estimates we estimated the VaR with the POT method and the Hill estimator as well as with a combination of both. We saw that despite the completely different outcomes of $\xi$ the estimated VaR is similar for all methods used. In particular, we can conclude that the classical extreme value methods deliver very good results. In our observed period 1.02 % of losses lie above the VaR derived from the POT method and Hill estimator. This is very close to the 99% VaR, what we wanted to obtain.

Furthermore every data point was scaled by the empirical standard deviation of the preceeding 260 days, to derive a rather identically distributed dataset. The extreme value methods were again applied on the scaled sample. Actually, the work of our estimators was a little bit worse, but still delivers appropriate results. Lastly, we saw that fitting the data to the normal distribution is not appropriate and our sample is indeed more heavy tailed, which means that we were right when assuming a geneneral pareto distribution for the excesses of our sample values over a certain threshold.

# References

[1] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin, 1997.

[2] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press, Princeton and Oxford, 2005.

[3] Basel Committee on Banking Supervision. *Revisions to the Basel II market risk framework* (Updated as of 31 December 2010). 2011. http://www.bis.org/publ/bcbs193.pdf

[4] Basel Committee on Banking Supervision. *Basel III: A global reulatory framwork for more resilient banks and banking systems* (December 2010, revised June 2011). 2011. http://www.bis.org/publ/bcbs189.pdf

[5] Basel Committee on Banking Supervision. *The Basel Committee's response to the financial crisis: report to the G20.* October 2010. http://www.bis.org/publ/bcbs179.pdf

[6] R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

[7] S original by Alexander McNeil and R port by Alec Stephenson. (2011). evir: Extreme Values in R. R package version 1.7-2. http://CRAN.R-project.org/package=evir

[8] Mathieu Ribatet (2011). POT: Generalized Pareto Distribution and Peaks Over Threshold. R package version 1.1-1. http://CRAN.R-project.org/package=POT

[9] A. G. Stephenson. evd: Extreme Value Distributions. R News, 2(2):31-32, June 2002. URL: http://CRAN.R-project.org/doc/Rnews/

# A  R codes

For the whole section we have

```
gs<-read.csv("C:/.../gs.csv")
```

as the data downloaded from
http://finance.yahoo.com/q/hp?s=GS&a=00&b=3&c=2005&d=04&e=23&f=2012&g=d

## A.1  Figure 1

```
dens<-density(rgumbel(1000000))
plot(dens,xlim=c(-3,7),main="",sub="",xlab="",cex.axis=1.5,
lwd=3,cex.lab=1.5)
VaR95<-qgumbel(0.95)

abline(v=0)
abline(v=VaR95,lty=2,lwd=3)

x1 <- min(which(dens$x >= VaR95))
x2 <- max(which(dens$x <  7))

with(dens, polygon(x=c(x[c(x1,x1:x2,x2)]), y= c(0, y[x1:x2], 0),
col="gray"))
axis(side=1,at=VaR95,"VaR(0.95)",cex.axis=1.5)
text(locator(1),"5%",cex=1.5)
```

## A.2  Figure 2

```
library("evd")
par(mfrow=c(1,2))
x <- seq(-4, 8, length=100)
hx <- dweibull(-x,shape=1)
plot(x, hx, type="l",lty=3,lwd=3,ylab="h(x)",cex.axis=1.5,
cex.lab=1.5)
lines(x,dgumbel(x),type="l",lty=2,lwd=3)
lines(x,dfrechet(x,shape=1),type="l",lwd=3)
legend("topright",legend=c("Fr chet","Weibull","Gumbel"),
lwd=3,lty=c(1,3,2),cex=2)

x <- seq(-4, 8, length=100)
hx <- pweibull(-x,shape=1)
plot(x, (hx*(-1)+1), type="l",lty=3,lwd=3,ylim=c(0,1),
ylab="H(x)",cex.axis=1.5,cex.lab=1.5)
lines(x,pgumbel(x),type="l",lty=2,lwd=3)
lines(x,pfrechet(x,shape=1),type="l",lwd=3)
legend("bottomright",legend=c("Fr chet","Weibull","Gumbel"),
lwd=3,lty=c(1,3,2),cex=2)
```

## A.3   Figure 3

```
library("evd")
library("evir")
par(mfrow=c(1,2))
x <- seq(0, 5, length=100)
hx <- dgpd(x, xi=0.5)
plot(x, hx, type="l",lty=1,lwd=3,cex.axis=1.5, ylab="g(x)",
cex.lab=1.5)
lines(x,dgpd(x,xi=-0.25),type="l",lty=3,lwd=3)
lines(x,dgpd(x,xi=0.0000000001),type="l",lty=2, lwd=3)
legend("topright",legend=c("Fr chet","Weibull","Gumbel"),
lwd=3,lty=c(1,3,2),cex=2)

x <- seq(0, 5, length=100)
hx <- pgpd(x, xi=0.5)
plot(x, hx, type="l",lty=1,lwd=3,ylim=c(0,1),cex.axis=1.5,
ylab="G(x)",cex.lab=1.5)
lines(x,pgpd(x,xi=-0.25),type="l",lty=3,lwd=3)
lines(x,pgpd(x,xi=0.0000000001),type="l",lty=2, lwd=3)
legend("bottomright",legend=c("Fr chet","Weibull","Gumbel"),
lwd=3,lty=c(1,3,2),cex=2)
```

## A.4   Figure 4

```
library("evir")
set.seed(1)
gev<-rgev(50,xi=0.8)
gev0<-gev[6>gev&gev>0]
plot(gev0,ylim=c(0.2,5.2),type="h",xlab=" ", ylab=" ",xaxt="n",
yaxt="n",lwd=1)
abline(h=2.5,lwd=2,lty=3)
gev25<-gev0[gev0>2.5]
ind<-which(gev0>2.5)
for (i in 1:length(gev25)){
polygon(x=c(ind[i],ind[i]),y=c(2.5,gev25[i]),lwd=3)
}
axis(side=2,at=2.5,"u",cex.axis=1.5,las=1,lwd=3)
text(locator(1),expression(X[i]),cex=1.5)
text(locator(1),expression(X[i]-u),cex=1.5)
text(locator(1),expression(X[1]),cex=1.5)
text(locator(1),expression(X[n]),cex=1.5)
text(locator(1),expression(X[k]),cex=1.5)
text(locator(1),expression(X[k]-u),cex=1.5)
```

## A.5   Figure 5

```
date.gs<-strptime(gs[,1],format="%Y-%m-%d")
chart.gs<-gs[,5]

dateplot<-date.gs[1:(length(date.gs)-582)]
chartplot<-chart.gs[1:(length(chart.gs)-582)]

plot(dateplot,chartplot,type="l",ylab=" ", xlab=" ",
main="Goldman Sachs Group, Inc.(NYSE) Basic Chart")
```

## A.6   Figure 6

```
date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])

dateplot<-date.gs[1:(length(date.gs)-582)]
retplot<-ret.gs[1:(length(ret.gs)-581)]

par(mfrow=c(1,2))
plot(dateplot,retplot,ylab=" ", xlab=" ",main="", type="h")
hist(retplot,br=seq(-25,25,length=61),col="grey",labels=FALSE,
xlab="",ylab="",main="")
```

## A.7   Figure 7

```
library("evd")

ret.gs<-diff(gs[,5])
ret.gs0<-ret.gs[ret.gs>0]

par(mfrow=c(1,4))

#qqnorm
qqnorm(ret.gs,xlab="Normal Quantiles", ylab=" ",main=" ",
cex.lab=1.5)

##qqexp
qqplot(rexp(10000),ret.gs0,xlab="Exponential Quantiles",
ylab=" ",cex.lab=1.5)

#qqfrechet
qqplot(rfrechet(10000,shape=10),ret.gs0,
xlab=expression(xi==0.1),ylab=" ",cex.lab=1.5)
qqplot(rfrechet(10000,shape=2),ret.gs0,
xlab=expression(xi==0.5), ylab=" ",cex.lab=1.5)
```

## A.8   Figure 8

```
library("evir")

ret.gs<-diff(gs[,5])
ret.gs0<-ret.gs[ret.gs>0]
threshold<-findthresh(ret.gs0,ne=180)

par(mfrow=c(1,4))

x<-c(0,0.1,0.2,0.5)
for (i in 1:4)
qplot(ret.gs0,xi=x[i],threshold=threshold)
```

## A.9   Figure 9

```
library("evir")

ret.gs<-diff(gs[,5])
ret.gs0<-ret.gs[ret.gs>0]

meplot(ret.gs0,omit=7)
```

## A.10   Figure 10

```
library("evir")

date.gs<-gs[,1]
ret.gs<-diff(gs[,5])

date<-which(date.gs=="2008-07-01")
ret08.gs<-ret.gs[(date+1):(date+260)]
date<-which(date.gs=="2009-07-01")
ret09.gs<-ret.gs[(date+1):(date+260)]
date<-which(date.gs=="2010-07-01")
ret10.gs<-ret.gs[(date+1):(date+260)]

par(mfrow=c(2,3))
shape(ret08.gs,start=17,end=45, reverse=FALSE)
shape(ret09.gs,start=17,end=45, reverse=FALSE)
shape(ret10.gs,start=17,end=45, reverse=FALSE)
quant(ret08.gs,p=0.99,start=17,end=45, reverse=FALSE,
auto.scale=FALSE,ylim=c(8,22))
quant(ret09.gs,p=0.99,start=17,end=45, reverse=FALSE,
auto.scale=FALSE,ylim=c(8,22))
quant(ret10.gs,p=0.99,start=17,end=45, reverse=FALSE,
auto.scale=FALSE,ylim=c(4,18))
```

## A.11   Figure 11

```
library("evir")

date.gs<-gs[,1]
ret.gs<-diff(gs[,5])

date<-which(date.gs=="2008-07-01")
ret08.gs<-ret.gs[(date+1):(date+260)]
date<-which(date.gs=="2009-07-01")
ret09.gs<-ret.gs[(date+1):(date+260)]
date<-which(date.gs=="2010-07-01")
ret10.gs<-ret.gs[(date+1):(date+260)]

par(mfrow=c(2,3))
hill(ret08.gs,option="xi",start=15,end=45)
hill(ret09.gs,option="xi",start=15,end=45)
hill(ret10.gs,option="xi",start=15,end=45)
hill(ret08.gs,option="quantile",start=15,p=0.99,end=45,
auto.scale=FALSE, ylim=c(12,22))
hill(ret09.gs,option="quantile",start=15,p=0.99,end=45,
auto.scale=FALSE, ylim=c(18,28))
hill(ret10.gs,option="quantile",start=15,p=0.99,end=45,
auto.scale=FALSE, ylim=c(8,18))
```

## A.12   Figure 12

```
library("evir")

ret.gs<-diff(gs[,5])
date.gs<-gs[,1]

date<-which(date.gs=="2008-07-01")
ret08.gs<-diff(gs[(date+1):(date+260),5])
date<-which(date.gs=="2009-07-01")
ret09.gs<-diff(gs[(date+1):(date+260),5])
date<-which(date.gs=="2010-07-01")
ret10.gs<-diff(gs[(date+1):(date+260),5])

par(mfrow=c(1,3))
meplot(ret08.gs[ret08.gs>0],xlim=c(0,20),ylim=c(0,5))
meplot(ret09.gs[ret09.gs>0],xlim=c(0,20),ylim=c(0,5))
meplot(ret10.gs[ret10.gs>0],xlim=c(0,20),ylim=c(0,10))
```

## A.13   Figure 13

```
library("evir")

date<-which(gs[,1]=="2010-07-01")
ret10.gs<-diff(gs[date:(date+260),5])

par(mfrow=c(1,3))
k<-26
threshold<-findthresh(ret10.gs,ne=k)
gpd.10<-gpd(ret10.gs,nextremes=k)
plot.gpd(gpd.10)
1
2
0
qplot(ret10.gs,xi=0.416,threshold=threshold)
```

## A.14   Figure 14

```
library("evir")

date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])
k<-26


#######################
# VaR with POT method #
#######################

VaR.P<-c()

for (i in 1:(length(ret.gs)-521)) {

ret<-ret.gs[(i+1):(i+260)]
gpd<-gpd(ret, nextremes=k)
q<-riskmeasures(gpd,0.99)[2]
VaR.P<-rbind(VaR.P,q)
}

###############################
# VaR with Normal distribution #
###############################

VaRnorm<-c()

for (i in 1:(length(ret.gs)-521)) {
```

```
ret<-ret.gs[(i+1):(i+260)]
q<-qnorm(0.99,mean=mean(ret),sd=sd(ret))
VaRnorm<-rbind(VaRnorm,q)
}


dateplot<-date.gs[1:(length(date.gs)-582)]
retplot<-ret.gs[1:(length(ret.gs)-581)]
VaR.P.plot<-VaR.P[1:(length(VaR.P)-60)]
VaRnormplot<-VaRnorm[1:(length(VaRnorm)-60)]


plot(dateplot,retplot,type="h",xlab="",ylab="")
lines(dateplot,VaR.P.plot,lwd=2, col="blue")
lines(dateplot,VaRnormplot,col="black",lwd=2)

##############
# Backtesting #
##############

number.days<-length(retplot)
number.excesses.P<-length(which(retplot>VaR.P.plot))
number.excesses.P
number.excesses.P/number.days


###############
# Stressed VaR #
###############

date.gs[which(VaR.P.plot==max(VaR.P.plot))]
max(VaR.P.plot)
```

## A.15   Figure 15

```
date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])
k<-26


#########################
# VaR estimation by Hill #
#########################

VaR.H<-c()

for (i in 1:(length(ret.gs)-521)) {
```

```
ret<-ret.gs[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
q<-((260/k*(1-0.99))^(-shape))*retsort[k]
VaR.H<-rbind(VaR.H,q)
}


#############################
# VaR with normal distribution #
#############################

VaRnorm<-c()

for (i in 1:(length(ret.gs)-521)) {

ret<-ret.gs[(i+1):(i+260)]
q<-qnorm(0.99,mean=mean(ret),sd=sd(ret))
VaRnorm<-rbind(VaRnorm,q)
}

dateplot<-date.gs[1:(length(date.gs)-582)]
retplot<-ret.gs[1:(length(ret.gs)-581)]
VaR.H.plot<-VaR.H[1:(length(VaR.H)-60)]
VaRnormplot<-VaRnorm[1:(length(VaRnorm)-60)]

plot(dateplot,retplot,type="h",xlab="",ylab="")
lines(dateplot,VaR.H.plot,lty=1,lwd=2,col="red")
lines(dateplot,VaRnormplot,col="black",lwd=2)

##############
# Backtesting #
##############

number.days<-length(retplot)
number.excesses.H<-length(which(retplot>VaR.H.plot))
number.excesses.H
number.excesses.H/number.days

##############
# Stressed VaR #
##############

date.gs[which(VaR.H.plot==max(VaR.H.plot))]
max(VaR.H.plot)
```

## A.16   Figure 16

```
library("POT")

date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])
k<-26

###############################################
# VaR estimated xi by hill and beta by POT #
###############################################

VaR.HP<-c()

for (i in 1:(length(ret.gs)-521)) {

ret<-ret.gs[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
gpd<-fitgpd(ret,threshold=retsort[k],est="mle",shape=shape)
scale<-gpd$scale
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.HP<-rbind(VaR.HP,q)
}

##############################
# VaR with normal distribution #
##############################

VaRnorm<-c()

for (i in 1:(length(ret.gs)-521)) {

ret<-ret.gs[(i+1):(i+260)]
q<-qnorm(0.99,mean=mean(ret),sd=sd(ret))
VaRnorm<-rbind(VaRnorm,q)
}

dateplot<-date.gs[1:(length(date.gs)-582)]
retplot<-ret.gs[1:(length(ret.gs)-581)]
VaR.HP.plot<-VaR.HP[1:(length(VaR.HP)-60)]
VaRnormplot<-VaRnorm[1:(length(VaRnorm)-60)]

plot(dateplot,retplot,type="h",xlab="",ylab="")
lines(dateplot,VaR.HP.plot,lty=1,lwd=2,col="green3")
lines(dateplot,VaRnormplot,col="black",lwd=2)
```

```
##############
# Backtesting #
##############

number.days<-length(retplot)
number.excesses.HP<-length(which(retplot>VaR.HP.plot))
number.excesses.HP
number.excesses.HP/number.days

number.excesses.norm<-length(which(retplot>VaRnormplot))
number.excesses.norm
number.excesses.norm/number.days

###############
# Stressed VaR #
##############

date.gs[which(VaR.HP.plot==max(VaR.HP.plot))]
max(VaR.HP.plot)

date.gs[which(VaRnormplot==max(VaRnormplot))]
max(VaRnormplot)
```

## A.17   Figure 17

```
library("evir")

date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])
k<-26

shape.P<-c()

for (i in 1:(length(ret.gs)-581)) {

ret<-ret.gs[(i+1):(i+260)]
gpd<-gpd(ret, nextremes=k)
shape<-as.vector(gpd[["par.ests"]][1])
shape.P<-rbind(shape.P,shape)
}

shape.H<-c()

for (i in 1:(length(ret.gs)-581)) {
```

```
ret <-ret.gs [(i+1):(i+260)]
retsort <-sort(ret ,decreasing=TRUE)
threshold <-retsort [k]
retneu <-retsort [1:k]
shape <-(1/k)*sum(log(retneu))-log(retneu[k])
shape.H<-rbind(shape.H,shape)
}

dateplot <-date.gs [1:(length(date.gs)-582)]

plot(dateplot ,shape.P,type="p",col="blue",pch=20,xlab="",
ylab="",ylim=c(-0.6,1))
lines(dateplot ,shape.P, col="blue",lwd=2)
points(dateplot ,shape.H, col="red",pch=20)
lines(dateplot ,shape.H, col="red",lwd=2)
abline(h=0)
```

## A.18   Figure 18

```
library("POT")

date.gs <-strptime(gs[,1],format="%Y-%m-%d")
ret.gs <-diff(gs [,5])
k<-26

sigma1 <-c()

for (i in 1: (length(ret.gs)-260)){
ret <-ret.gs [(i+1):(i+260)]
sigma <-sd(ret)
sigma1 <-rbind(sigma1 ,sigma)
}

ret.neu <-ret.gs [1:(length(sigma1))]/sigma1

##########################
# VaR estimation by Hill #
##########################

VaR.H<-c()

for (i in 1:(length(ret.neu)-261)) {

ret <-ret.neu [(i+1):(i+260)]
retsort <-sort(ret ,decreasing=TRUE)
shape <-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
```

```
q<-((260/k*(1-0.99))^(-shape))*retsort[k]
VaR.H<-rbind(VaR.H,q)
}


###############################
# VaR estimation by POT method #
###############################

VaR.P<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
gpd<-fitgpd(ret,threshold=retsort[k],est="mle")
shape<-gpd[[1]][2]
scale<-gpd[[1]][1]
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.P<-rbind(VaR.P,q)
}


#############################################
# VaR estimated xi by hill and beta by POT #
#############################################

VaR.HP<-c()
shape.HP<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
gpd<-fitgpd(ret,threshold=retsort[k],est="mle",shape=shape)
scale<-gpd$scale
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.HP<-rbind(VaR.HP,q)
}

dateplot<-date.gs[1:(length(date.gs)-582)]
ret.neuplot<-ret.neu[1:(length(ret.neu)-321)]
VaR.H.plot<-VaR.H[1:(length(VaR1)-60)]
VaR.P.plot<-VaR.P[1:(length(VaR2)-60)]
VaR.HP.plot<-VaR3[1:(length(VaR3)-60)]

plot(dateplot, ret.neuplot, type="h",xlab="",ylab="")
```

```
lines(dateplot, VaR3plot, col="green3",lwd=2)
lines(dateplot, VaR1plot, col="red",lwd=2)
lines(dateplot, VaR2plot, col="blue",lwd=2)


###############
# Backtesting #
###############

number.days<-length(ret.neuplot)
number.excesses.H<-length(which(ret.neuplot>VaR.H.plot))
number.excesses.H
number.excesses.H/number.days

number.excesses.P<-length(which(ret.neuplot>VaR.P.plot))
number.excesses.P
number.excesses.P/number.days

number.excesses.HP<-length(which(ret.neuplot>VaR.HP.plot))
number.excesses.HP
number.excesses.HP/number.days
```

## A.19   Figure 19

```
library("POT")

date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])

sigma1<-c()

for (i in 1:(length(ret.gs)-260)){
ret<-ret.gs[(i+1):(i+260)]
sigma<-sd(ret)
sigma1<-rbind(sigma1,sigma)
}

ret.neu<-ret.gs[1:(length(sigma1))]/sigma1

##########################
# VaR estimation by Hill #
##########################

VaR.H<-c()

for (i in 1:(length(ret.neu)-261)) {
```

```
ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
q<-((260/k*(1-0.99))^(-shape))*retsort[k]
VaR.H<-rbind(VaR.H,q)
}

############################
# VaR estimation by POT method #
############################

VaR.P<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
gpd<-fitgpd(ret,threshold=retsort[k],est="mle")
shape<-gpd[[1]][2]
scale<-gpd[[1]][1]
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.P<-rbind(VaR.P,q)
}

##############################################
# VaR xi estimated by hill and beta by POT #
##############################################

VaR.HP<-c()
shape.HP<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
gpd<-fitgpd(ret,threshold=retsort[k],est="mle",shape=shape)
scale<-gpd$scale
q<-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.HP<-rbind(VaR.HP,q)
}

dateplot<-date.gs[1:(length(date.gs)-582)]
ret.neuplot<-ret.neu[1:(length(ret.neu)-321)]
VaR.H.plot<-VaR.H[1:(length(VaR1)-60)]
VaR.P.plot<-VaR.P[1:(length(VaR2)-60)]
```

```
VaR.HP.plot<-VaR.HP[1:(length(VaR3)-60)]

sigma2<-sigma1[1:(length(sigma1)-321)]

plot(dateplot, ret.neuplot*sigma2, type="h",xlab="",ylab="")
lines(dateplot, VaR.HP.plot*sigma2, col="green3",lwd=2)
lines(dateplot, VaR.H.plot*sigma2, col="red",lwd=2)
lines(dateplot, VaR.P.plot*sigma2, col="blue",lwd=2)
```

## A.20   Figure 20

```
library("POT")

date.gs<-strptime(gs[,1],format="%Y-%m-%d")
ret.gs<-diff(gs[,5])
k<-26

sigma1<-c()

for (i in 1: (length(ret.gs)-260)){
ret.help<-ret.gs[(i+1):(i+260)]
sigma<-sd(ret.help)
sigma1<-rbind(sigma1,sigma)
}

ret.neu<-ret.gs[1:(length(sigma1))]/sigma1

shape.H<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
shape<-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
shape.H<-rbind(shape.H,shape)
}

shape.P<-c()

for (i in 1:(length(ret.neu)-261)) {

ret<-ret.neu[(i+1):(i+260)]
retsort<-sort(ret,decreasing=TRUE)
gpd<-fitgpd(ret,threshold=retsort[k],est="mle")
shape<-gpd[[1]][2]
shape.P<-rbind(shape.P,shape)
```

```
}

dateplot <-date.gs[1:(length(date.gs)-582)]

plot(dateplot,shape.P[1:(length(shape.P)-60)],col="blue",
type="l",lwd=2,xlab="",ylab="",ylim=c(-0.6,1.0))
points(dateplot,shape.P[1:(length(shape.P)-60)],col="blue",
pch=20)
points(dateplot,shape.H[1:(length(shape.H)-60)],col="red",
pch=20)
lines(dateplot,shape.H[1:(length(shape.H)-60)],col="red",
lwd=2)
abline(h=0)
```

## A.21 Figure 21

```
library("evir")
library("POT")

date.gs <-strptime(gs[,1],format="%Y-%m-%d")
ret.gs <-diff(gs[,5])
k <-26

######################
# VaR with POT method #
######################

VaR.P <-c()

for (i in 1:(length(ret.gs)-521)) {

ret <-ret.gs[(i+1):(i+260)]
gpd <-gpd(ret, nextremes=k)
q <-riskmeasures(gpd,0.99)[2]
VaR.P <-rbind(VaR.P,q)
}

########################
# VaR estimation by Hill #
########################

VaR.H <-c()

for (i in 1:(length(ret.gs)-521)) {

ret <-ret.gs[(i+1):(i+260)]
```

```
retsort <-sort(ret,decreasing=TRUE)
shape <-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
q <-((260/k*(1-0.99))^(-shape))*retsort[k]
VaR.H <-rbind(VaR.H,q)
}


################################################
# VaR xi estimated by hill and beta by POT #
################################################

VaR.HP <-c()

for (i in 1:(length(ret.gs)-521)) {

ret <-ret.gs[(i+1):(i+260)]
retsort <-sort(ret,decreasing=TRUE)
shape <-(1/k)*sum(log(retsort[1:k]))-log(retsort[k])
gpd <-fitgpd(ret,threshold=retsort[k],est="mle",shape=shape)
scale <-gpd$scale
q <-((260/k*(1-0.99))^(-shape)-1)*(scale/shape)+retsort[k]
VaR.HP <-rbind(VaR.HP,q)
}

dateplot <-date.gs[1:(length(date.gs)-582)]
retplot <-ret.gs[1:(length(ret.gs)-581)]

plot(dateplot[716:1150],retplot[716:1150],type="h",xlab="",ylab="")
lines(dateplot[716:1150],VaRnorm[716:1150],col="black",lwd=2)
lines(dateplot[716:1150],VaR.P[716:1150],lwd=2, col="blue")
lines(dateplot[716:1150],VaR.H[716:1150],lty=1,lwd=2,col="red")
lines(dateplot[716:1150],VaR.HP[716:1150],lty=1,lwd=2,col="green3")
```

## A.22   Figure 22

```
library("evir")

date.gs <-strptime(gs[,1],format="%Y-%m-%d")
ret.gs <-diff(gs[,5])
k <-26

#######################
# VaR with POT method #
#######################

VaR.P <-c()
```

```
for (i in 1:(length(ret.gs)-521)) {

ret<-ret.gs[(i+1):(i+260)]
gpd<-gpd(ret, nextremes=k)
q<-riskmeasures(gpd,0.99)[2]
VaR.P<-rbind(VaR.P,q)
}

##############
# average VaR #
##############

a<-60
VaR.avg<-c()

for (i in 1:(length(VaR.P)-a)){

avg<-mean(VaR.P[i:(i+a)])
VaR.avg<-rbind(VaR.avg,avg)
}


dateplot<-date.gs[1:(length(date.gs)-582)]
retplot<-ret.gs[1:(length(ret.gs)-581)]
VaR.P.plot<-VaR.P[1:(length(VaR.P)-60)]

plot(dateplot,retplot,type="h",xlab="",ylab="",ylim=c(-20,50))
lines(dateplot,VaR.P.plot,lwd=2, col="blue")
lines(dateplot,VaR.avg,col="black",lwd=2)
lines(dateplot,3*VaR.avg,col="red",lwd=2)

max<-which(gs[,1]=="2009-07-09")
points(dateplot[max],VaR.avg[max],pch=4,cex=2,lwd=2)
points(dateplot[max],3*VaR.avg[max],pch=4,cex=2,lwd=2,col="red")
```