NIG distribution in modelling stock returns with assumption about stochastic volatility
Estimation of parameters and application to VaR and ETL

Master's Thesis in Financial Mathematics

Magdalena Kucharska    Jolanta Pielaszkiewicz

School of Information Science, Computer and Electrical Engineering
Halmstad University
NIG distribution in modelling stock returns with assumption about stochastic volatility
Estimation of parameters and application to VaR and ETL

Magdalena Kucharska
Jolanta Pielaszkiewicz

Halmstad University
Project Report IDE0923

Master’s thesis in Financial Mathematics, 15 ECTS credits

Supervisor: Ph.D. Jan-Olof Johansson
Examiner: Prof. Ljudmila A. Bordag
External referee: Prof. Daniel Ševčovič

July 28, 2009

Department of Mathematics, Physics and Electrical Engineering
School of Information Science, Computer and Electrical Engineering
Halmstad University
Acknowledgments

The authors would like to thank their supervisor Jan-Olof Johansson for all advices, support and guidance during this work and course coordinator Ljudmila A. Bordag for all her advices. The authors are grateful to Carl Lindberg who was kind enough to answer some of our questions.

Magdalena Kucharska and Jolanta Pielaszkiewicz

I would like to thank to my family and to the people without whom I could not study. I express special thanks to my fiance Michal Meller who provided me with constant support through all my studies and helped me to make decision to study in Halmstad.

Magdalena Kucharska

I would like to thank my family, especially to my God Son Pawel who makes my life happy and makes that I particularly yearn for home. The person to whom I am the most grateful is my boyfriend Adam, who had supported me for last 1971 days (taking into account the bissextile years). I would also like to thank co-author Magda and my friends especially Natalia, Patrycja and Asia for time we spent together in Halmstad.

Jolanta Pielaszkiewicz
Abstract

We model Normal Inverse Gaussian distributed log-returns with the assumption of stochastic volatility. We consider different methods of parametrization of returns and following the paper of Lindberg, [21] we assume that the volatility is expressed by \( \sigma^2(\cdot) = \theta z(\cdot) \), where \( \theta \) is a constant and \( z(\cdot) \) is the number of trades. In addition to the Lindberg’s paper, we suggest daily stock volumes and amounts as alternative measures of the volatility.

As an application of the models, we perform Value-at-Risk and Expected Tail Loss predictions by the Lindberg’s volatility model and by our own suggested model. These applications are new and not described in the literature.

For better understanding of our calculations, programmes and simulations, basic informations and properties about the Normal Inverse Gaussian and Inverse Gaussian distributions are provided.

Practical applications of the models are implemented on the Nasdaq-OMX®, where we have calculated Value-at-Risk and Expected Tail Loss for the Ericsson B stock data during the period 1999 to 2004.
## Contents

1 Introduction  
1.1 The IG and NIG distributions  
1.1.1 The Inverse Gaussian Distribution  
1.1.2 The Normal Inverse Gaussian  
1.2 The EM algorithm for the maximum likelihood estimation  
1.2.1 Initial values for the EM algorithm  
1.2.2 The log-likelihood function for the NIG distribution  
1.2.3 The steps of the EM algorithm  
1.3 The Lindberg method  
1.3.1 The Barndorff-Nielsen and Shephard’s model  
1.3.2 The Lindberg method  
1.3.3 The steps of the method  
1.4 The measure of risk – Value-at-Risk  
1.4.1 Definition  
1.4.2 Methods of VaR evaluation  
1.4.2.1 The historical method  
1.4.2.2 Monte Carlo method  
1.4.2.3 The variance-covariance approach  
1.4.3 The advantages and the disadvantages of VaR method  
1.4.4 The Expected Tail Loss method (ETL)  
1.5 Consistency tests  
1.5.1 The Kolmogorov-Smirnov’s test (K-S test)  
1.5.2 The Anderson-Darling test (A-D test)  
1.5.3 The Kuiper’s test  
1.6 The Lindberg method  
1.6.1 The Barndorff-Nielsen and Shephard’s model  
1.6.2 The Lindberg method  
1.6.3 The steps of the method  
1.7 The measure of risk – Value-at-Risk  
1.7.1 Definition  
1.7.2 Methods of VaR evaluation  
1.7.2.1 The historical method  
1.7.2.2 Monte Carlo method  
1.7.2.3 The variance-covariance approach  
1.7.3 The advantages and the disadvantages of VaR method  
1.7.4 The Expected Tail Loss method (ETL)  
1.8 Consistency tests  
1.8.1 The Kolmogorov-Smirnov’s test (K-S test)  
1.8.2 The Anderson-Darling test (A-D test)  
1.8.3 The Kuiper’s test
7 The study of the statistical data for the Ericsson B stock

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>The features of the statistical data</td>
<td>41</td>
</tr>
<tr>
<td>7.2</td>
<td>The EM algorithm</td>
<td>44</td>
</tr>
<tr>
<td>7.2.1</td>
<td>The estimation of the NIG parameters</td>
<td>45</td>
</tr>
<tr>
<td>7.2.2</td>
<td>The IG parameters</td>
<td>46</td>
</tr>
<tr>
<td>7.3</td>
<td>The Lindberg method</td>
<td>46</td>
</tr>
<tr>
<td>7.3.1</td>
<td>The estimation of the parameters ( \mu, \beta ) and ( \theta )</td>
<td>47</td>
</tr>
<tr>
<td>7.3.2</td>
<td>The normalized returns ( \xi )</td>
<td>48</td>
</tr>
<tr>
<td>7.3.3</td>
<td>The IG parameters estimation</td>
<td>49</td>
</tr>
<tr>
<td>7.3.4</td>
<td>The NIG parameters</td>
<td>51</td>
</tr>
<tr>
<td>7.4</td>
<td>The EM algorithm versus Lindberg method</td>
<td>52</td>
</tr>
<tr>
<td>7.5</td>
<td>The VaR and ETL methods</td>
<td>55</td>
</tr>
<tr>
<td>7.5.1</td>
<td>The case of the number of trades</td>
<td>55</td>
</tr>
<tr>
<td>7.5.2</td>
<td>The case of the volume of the traded assets</td>
<td>56</td>
</tr>
<tr>
<td>7.6</td>
<td>The VaR and ETL predictions</td>
<td>58</td>
</tr>
<tr>
<td>7.6.1</td>
<td>The case of the number of trades</td>
<td>58</td>
</tr>
<tr>
<td>7.6.2</td>
<td>The case of the volume of traded assets</td>
<td>59</td>
</tr>
</tbody>
</table>

8 Conclusions

Bibliography

Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The EM algorithm implementation</td>
<td>67</td>
</tr>
<tr>
<td>1.2</td>
<td>The Lindberg method implementation</td>
<td>70</td>
</tr>
<tr>
<td>1.3</td>
<td>The VaR and ETL implementation</td>
<td>71</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Distributions of financial data were analyzed in many papers over the past several years [25], [1], [16]. It was observed that this data is fat-tailed and very often skewed. In early studies scientists took into account distributions such as stable Pareto or mixture distributions. In the 1999th, Barndorff-Nielsen introduced a new family of distributions called Normal Inverse Gaussian, which fulfills fat-tails condition, is analytically tractable and easily estimated by likelihood method, cf. [4] and [3].

For practical reasons, we need efficient and effective tools to estimate the NIG parameters. One way of NIG parameters estimation is the Expectation Maximization (EM) algorithm, see [14]. The NIG distribution can be presented as a Normal variance-mean mixture with IG as a mixing distribution, see [19]. This fact can be used to simplify the algorithm, improve its speed and reduce efforts needed for implementation. However, the parametrization obtained in this way does not reflect information about volatility.

A method, which uses assumption about a stochastic volatility to estimate returns, was proposed by Carl Lindberg in [21]. The Stochastic volatility was introduced before, e.g., in Barndorff-Nielsen’s and Shephard’s paper [6]. However, their approach demanded the quadratic variation method, which is difficult to implement. Lindberg proposed a new look at the volatility problem. He found a way for a stable estimation of parameters by fitting parameters to volatility first and then to returns. This approach is easier to implement. The Lindberg’s examination showed that for turbulent stocks, a good fitting for long period of time can be obtained. To understand the behaviour of volatility \( \sigma^2 \) better, Lindberg assumed that this process is given by

\[
\sigma^2(\cdot) = \theta z(\cdot),
\]

where \( \theta \) is a constant coefficient and \( z(\cdot) \) is a measure of a trading intensity, such as the number of trades, cf. [21].
Apart from verifying the Lindberg’s result, our goal in this thesis is to use this framework for the evaluation of risks. For this purpose, we consider two risk measures: Value at Risk and Expected Tail Loss. The concept of Value-at-Risk is recent. It was first used by financial firms in the late 1980s to measure the risks of their trading portfolios. Since that time period, the use of VaR has exploded. Currently VaR is used by most major derivatives dealers to measure and manage market risk. A 1995 Institutional Investor survey found that 32% of firms use VaR as a measure of market risk, and 60% of pension funds responding to a survey by the New York University Stern School of Business reported using VaR, cf. [22]. The other type of the measure of risk, based on VaR, is Expected Tail Loss (ETL), cf. [11].

In this thesis, we estimate the parameters of the NIG-distributed returns directly from real data using both algorithms described above. Additionally, for Lindberg’s method we investigate different measures of the trading intensity \( z(\cdot) \): number of trades, volume and amount. The use of the stock volumes and the stock amount is new and we have not found any implementation referenced in the literature. Our next goal is to compare the obtained results and make Value-at-Risk and the Expected Tail Loss prediction for the period of one year. Comparison of real losses and VaR and ETL level is also included. The implementation of the model, where the volatility is calculated by using a number of trades and stock volumes is a new and not described in the literature.

The text is organized as follows. In Chapter 2, we introduce the Normal Inverse Gaussian and the Inverse Gaussian distributions with some emphasis on properties that can be useful in our investigation. We discuss and derive formulas of the EM algorithm for a maximum likelihood estimation of NIG parameters in Chapter 3. The Lindberg’s method as an alternative way of estimating NIG parameters is presented in Chapter 4. Chapter 5 describes the main measures of risk: Value-at-Risk and Expected Tail Loss. The statistical tests used in our work are contained in Chapter 6. Practical applications and the interpretation of our results can be found in Chapter 7. The implementation of all methods in R language is provided in the Appendix.
Chapter 2

The IG and NIG distributions

The hyperbolic distribution family was discovered in the 1940s during study of the sand dune movements carried out by R.A. Bagnold. The size of the sand grains was observed and their distribution was studied. Scholars noticed that an obtained logarithmic function did not have a parabolic shape (typical to gaussian distribution), but a hyperbolic one. The name of the distribution arises from the latter fact.

Using this family, excellent fits were also obtained for the log-sizes of diamonds in the South Africa, as well as for daily stock returns from a number of leading German enterprises, see [9]. It also turned out that the Hyperbolic distribution is good for stock price modeling and for a market risk measurement, because both demand tasks heavy tails by a distribution, cf. [9].

In 1977, Ole Barndorff-Nielesen proposed the family of Generalized Hyperbolic distributions (GH) as a class of infinitely divisible distributions and a mean-variance mixture of the Generalized Inverse Gaussian distribution (GIG), see [4]. The GIG distribution has its roots in study of a monthly water flow in hydroelectric power plants, observed by Etienne Halphen in 1946 and also belongs to the infinitely divisible family distributions.

The goal of this chapter is to present the definitions and to discuss the basics properties of these families of cumulative distribution functions. Since in our master thesis we will use both IG and NIG, we would like to focus our attention especially on these two distributions. However, we think that it is useful to provide the reader with some basic information about GH and GIG distribution. The more detailed discussion can be found in [19],[12].

Definition 2.1. The probability density function of GIG is defined as

\[ f_{GIG}(x, \lambda, \delta, \gamma) = \frac{1}{2} \left( \frac{\gamma}{\delta} \right)^{\lambda/2} K_\lambda^{-1}(\sqrt{\delta \gamma}) x^{\lambda-1} \exp \left( -\frac{1}{2} (\delta^{-1} x^{-1} + \gamma^2 x) \right), \ x > 0 \]
with the following ranges of parameters

- if \( \lambda > 0 \) then \( \delta \geq 0, \gamma > 0 \),
- if \( \lambda = 0 \) then \( \delta > 0, \gamma > 0 \),
- if \( \lambda < 0 \) then \( \delta > 0, \gamma \geq 0 \).

The function \( K_\lambda(x) \) denotes the Bessel function of third kind with index \( \lambda \)

\[
K_\lambda(x) = \frac{1}{2} \int_0^\infty u^{\lambda-1} \exp \left( -\frac{1}{2} x(u^{-1} + u) \right) du, \quad x > 0.
\]

**Definition 2.2.** The random variable \( Z \) has the generalized hyperbolic distribution with parameters \( \lambda, \alpha, \beta, \delta, \mu \) if the conditional probability is equal to

\[
Z|Y = y \sim N(\mu + \beta y, y), \quad (2.1)
\]

where \( Y \sim GIG(\lambda, \delta, \sqrt{\alpha^2 - \beta^2}) \) and \( N(\mu + \beta y, y) \) is the normal distribution with the mean \( \mu + \beta y \) and the variance \( y \).

**Remark 2.1.** The formula (2.1) is useful in generating the GH random variables.

From (2.1) it is easy to show that the probability density function (pdf) of GH is given for \( x \in \mathbb{R} \) by

\[
f_{GH}(\lambda, \alpha, \beta, \delta, \mu, x) = \int_0^\infty f_{N(\mu + \beta y, y)}(x)f_{GIG(\lambda, \delta, \sqrt{\alpha^2 - \beta^2})}(y)dy
\]

\[
= a_\lambda(\alpha, \beta, \delta)(\sqrt{\delta^2 + (x - \mu)^2})^{\lambda-1/2}K_{\lambda-\frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2}) \exp(\beta(x - \mu)),
\]

where \( a_\lambda(\alpha, \beta, \delta) \) is normalizing constant

\[
a_\lambda(\alpha, \beta, \delta) = (\sqrt{\alpha^2 - \beta^2})^\lambda/((\sqrt{2\pi})^{\lambda-\frac{1}{2}} \delta^\lambda K_{\lambda-\frac{1}{2}}(\delta \sqrt{\alpha^2 - \beta^2})).
\]

The parameters of the distribution take the values from

- if \( \lambda > 0 \) then \( \delta \geq 0, |\beta| < \alpha \),
- if \( \lambda = 0 \) then \( \delta > 0, |\beta| < \alpha \),
- if \( \lambda < 0 \) then \( \delta > 0, |\beta| \leq \alpha \).

The parameter \( \lambda \), which is an index of the Bessel function, is responsible for the characterization of the distribution (scale, skewness, shift, flatness), i.e. its value determines a kind of the subdistribution. For example, the NIG is a subclass of GH obtained for \( \lambda \) equal to \(-\frac{1}{2}\).


### 2.1 The Inverse Gaussian Distribution

Barndorff-Nielsen defined the IG distribution as below [5].

**Definition 2.3.** We say that a random variable $X$ is IG distributed with parameters $\delta \in \mathbb{R}^+$ and $\gamma \in \mathbb{R}^+$ when the probability density function has the following form

$$f_{IG}(x, \delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} x^{-3/2} \exp\left(\delta \gamma - \frac{\delta^2 x^{-1} + \gamma^2 x}{2}\right), \quad x > 0.$$  

The IG distribution has just two parameters $\delta$ and $\gamma$ which describe scale and shape, respectively. To analyze the role played by $\delta$ and $\gamma$ we use a graphical approach. As shown on Figure 2.1b, increasing the value of $\delta$ causes the density function to be flatter. The smaller the value of $\delta$ the smaller the range of values reached by the density. On the other hand if we fix $\delta$, we can observe how changing $\gamma$ influences the shape of pdf. The greater values of $\gamma$ imply the greater maximal value of pdf, see Figure 2.1a.

![Figure 2.1: Plot of the IG probability density for $\delta = 1$ and varying parameters $\gamma$ values in the left hand side, for $\gamma = 1$ and varying parameters $\delta$ in the right hand side.](image)

**Proposition 2.1.** The moment generating function of the IG is equal to

$$M_X(t) = \exp\left(\delta \gamma - \delta \sqrt{\gamma^2 - 2t}\right).$$
Proof

\[ M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx}f_{IG}(x, \delta, \gamma)dx \]

\[ = \int_0^\infty \frac{\delta}{\sqrt{2\pi}} x^{-3/2} \exp\left(\delta \gamma - \frac{\delta^2 x - 1 + \gamma^2 x}{2}\right) \exp(tx)dx \]

\[ = \frac{\delta}{\sqrt{2\pi}} x^{-3/2} \exp\left(\delta \gamma - \frac{\delta^2 x - 2tx + \gamma^2 x}{2}\right) \int_0^\infty \frac{\delta}{\sqrt{2\pi}} x^{-3/2} e^\left(\delta \sqrt{\gamma^2-2t} - \frac{\delta^2 x - 1 + \gamma^2 x}{2}\right)dx \]

\[ = \exp(\delta \gamma - \delta \sqrt{\gamma^2 - 2t}) \cdot \int_0^\infty \frac{\delta}{\sqrt{2\pi}} x^{-3/2} e^\left(\delta \sqrt{\gamma^2-2t} - \frac{\delta^2 x - 1 + \gamma^2 x}{2}\right)dx \]

\[ = \exp(\delta \gamma - \delta \sqrt{\gamma^2 - 2t}). \]

Using this function \( M_X(t) \) and cumulants it is easy to derive the mean, the variance, the skewness and the kurtosis. First let us calculate the cumulant generating function of IG

\[ g(t) = \log M_X(t) = \delta \gamma - \delta \sqrt{\gamma^2 - 2t}. \]

The first four cumulants, which are defined in general by

\[ cum_n = \frac{\partial^n g(t)}{\partial t^n}|_{t=0}, \]

are equal

\[ cum_1 = g'(t)|_{t=0} = \delta (\gamma^2 - 2t)^{-1/2}|_{t=0} = \frac{\delta}{\gamma}, \]

\[ cum_2 = g''(t)|_{t=0} = \delta (\gamma^2 - 2t)^{-3/2}|_{t=0} = \frac{\delta}{\gamma^3}, \]

\[ cum_3 = g'''(t)|_{t=0} = 3\delta (\gamma^2 - 2t)^{-5/2}|_{t=0} = \frac{3\delta}{\gamma^5}, \]

\[ cum_4 = g''''(t)|_{t=0} = 15\delta (\gamma^2 - 2t)^{-7/2}|_{t=0} = \frac{15\delta}{\gamma^7}. \]

Hence, we obtain that

the mean of \( X \) is equal to

\[ E(X) = cum_1 = \frac{\delta}{\gamma}, \] (2.2)
the variance of $X$ is equal to

$$\text{Var}(X) = \text{cum}_2 = \frac{\delta}{\gamma^3}, \quad (2.3)$$

the skewness, which is defined as the fraction $= \text{cum}_3/\text{cum}_2^{3/2}$ is equal to

$$\text{Skewness} = \frac{3\delta}{\gamma^3} \cdot \left(\frac{\delta}{\gamma^3}\right)^{-3/2} = \frac{3}{\sqrt{\delta\gamma}}, \quad (2.4)$$

the kurtosis, which is defined as the fraction $= \text{cum}_4/\text{cum}_2^2$ is equal to

$$\text{Kurtosis} = \frac{15\delta}{\gamma^5} \cdot \left(\frac{\delta}{\gamma^3}\right)^{-2} = \frac{15}{\delta\gamma}. \quad (2.5)$$

Observe that (2.4) implies that pdf is always positive skewed, i.e. the right tail is longer than the left one. If $\delta$ is fixed, then for $\gamma \to \infty$ the mean, the variance, the skewness and the kurtosis decrease. On the other hand, when $\delta$ increasing, the expected value and the variance rise, while the kurtosis and the skewness decrease.

A simple, but important and very useful property is

**Proposition 2.2.** If $X \sim IG(\delta, \gamma)$ then a random variable $Y = aX$ is $IG(a^{1/2}\delta, a^{-1/2}\gamma)$ distributed, where $a > 0$. [21]

**Proof** Let $Y = aX$. Then the cumulative distribution function of $Y$ has the following form

$$F_Y(t) = P(Y \leq t) = P(aX \leq t) = P \left( X \leq \frac{t}{a} \right) = F_X \left( \frac{t}{a} \right).$$

Since $f_X(t) = \frac{dF_X(t)}{dt}$, we yield

$$f_Y(t) = \frac{dF_X\left(\frac{t}{a}\right)}{d\left(\frac{t}{a}\right)} = \frac{1}{a} f_X \left( \frac{t}{a} \right),$$

$$f_Y(t) = \frac{1}{a} \frac{\delta}{\sqrt{2\pi}} \left(\frac{t}{a}\right)^{-3/2} \exp \left( \delta\gamma - \frac{\delta^2 \left(\frac{t}{a}\right)^{-1} + \gamma^2 \frac{t}{a}}{2} \right)$$

$$= \frac{a^{1/2}\delta}{\sqrt{2\pi}} t^{-3/2} \exp \left( \delta\gamma - \frac{a \delta^2 t^{-1} + \gamma^2 t}{2} \right)$$

$$= \frac{a^{1/2}\delta}{\sqrt{2\pi}} t^{-3/2} \exp \left( \delta\gamma - \frac{(a^{1/2}\delta)^2 t^{-1} + (a^{-1/2}\gamma)^2 t}{2} \right)$$

$$= f_{IG}(x, a^{1/2}\delta, a^{-1/2}\gamma).$$

\[\square\]
2.2 The Normal Inverse Gaussian

The special case of the General Hyperbolic Distribution on which we focus our attention in our thesis, is a Normal Inverse Gaussian distribution with four parameters \( \alpha, \beta, \mu, \delta \).

**Definition 2.4.** The random variable \( X \) is Normal Inverse Gaussian distributed \( NIG(\alpha, \beta, \mu, \delta) \) if its probability density function is given by

\[
f_{NIG}(x, \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp \left( \delta \sqrt{\alpha^2 - \beta^2} - \beta \mu \right) \phi(x)^{-1/2} K_1(\delta \alpha \phi(x)^{1/2}) \exp(\beta x),
\]

where \( \phi(x) = 1 + [(x - \mu)/\delta]^2 \) and \( K_r(x) \) denotes the modified Bessel function of the third kind of order \( r \) evaluated at \( x \). The conditions for the parameters are \( 0 \leq |\beta| \leq \alpha, \mu \in \mathbb{R}, 0 < \delta \).

Depending on the role parameters play, they can be attributed to one of the two groups. To the first group belong \( \alpha \) and \( \beta \), which affect the shape. To the second group belong \( \mu \) and \( \delta \), which scale the distribution. The parameter \( \alpha \), which may take nonnegative values, refers to flatness of the density function. The greater the alpha, the greater the concentration of the probability mass around \( \mu \). Additionally, the density function will reach a higher maximum value, see Figure 2.2a. The parameter \( \beta \) determines a kind of skewness of the distribution. The value \( \beta = 0 \) implies the symmetric distribution around mean. A negative value means heavier left tail and a positive value means otherwise, cf. Figure 2.2b. The third parameter \( \delta \) correspond to the scale of the distribution. Small values narrow the distribution down and larger ones make it wider. The last parameter – \( \mu \) is responsible for the shift of the density function.

![Figure 2.2: Plot of the NIG probability density.](image-url)
The NIG($\alpha, \beta, \mu, \delta$) distribution is a normal variance-mean mixture, i.e. it can be presented as the marginal distribution of $X$ in the pair $(X, Z)$, where the conditional probability $X | Z$ is given by

\[ X | Z = z \sim N(\mu + \beta z, z) \]

and the variable $Z$ has the inverse gaussian distribution with parameters $\delta$ and $\sqrt{\alpha^2 - \beta^2}$ for $0 \leq |\beta| \leq \alpha$.

Let $\gamma = \sqrt{\alpha^2 - \beta^2}$. The mean, the variance, the skewness and the kurtosis of $X$ are given by expressions

\[
\begin{align*}
E(X) &= \mu + \delta \frac{\beta}{\gamma}, \\
\text{Var}(X) &= \delta \frac{\alpha^2}{\gamma^3}, \\
\text{Skewness} &= 3 \frac{\beta}{\alpha (\delta \gamma)^{1/2}}, \\
\text{Kurtosis} &= 3 \left(1 + 4 \left(\frac{\beta^2}{\alpha^2}\right)\right) \frac{1}{\delta \gamma}.
\end{align*}
\]

Although the probability density function is fairly complicated, its moment generating function is simple one

\[ M_X(t) = \exp[t\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2})]. \]

Using this function one can show easily the following properties of NIG.

**Proposition 2.3.** (*Lillestol, [20]*)

1. If $X \sim \text{NIG}(\alpha, \beta, \mu, \delta)$, then $Y = kX \sim \text{NIG}(\alpha/k, \beta/k, k\mu, k\delta)$.

2. If $X_1 \sim \text{NIG}(\alpha, \beta, \mu_1, \delta_1)$ and $X_2 \sim \text{NIG}(\alpha, \beta, \mu_2, \delta_2)$ are independent, then the sum $Y = X_1 + X_2 \sim \text{NIG}(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2)$.

3. If $X_i \sim \text{NIG}(\alpha, \beta, \mu, \delta), (i = 1, \cdots, n)$ are independent, then the sample mean $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim \text{NIG}(n\alpha, n\beta, \mu, \delta)$.

4. If $X \sim \text{NIG}(\alpha, \beta, \mu, \delta)$, then variable $Y = (X - \mu)/\delta$ has the **Standard Normal Inverse Gaussian Distribution** $\text{NIG}(\alpha\delta, \beta\delta, 0, 1)$.

**Proof**
1. Let $Y = kX$. Then

$$M_Y(t) = M_{kX}(t) \overset{def}{=} \mathbb{E}(e^{tkX}) = M_X(tk)$$

$$= \exp\left(tk\mu + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + tk)^2}\right)\right)$$

$$= \exp\left(t(k\mu) + k\delta\left(\sqrt{(\alpha/k)^2 - (\beta/k)^2} - \sqrt{(\alpha/k)^2 - ((\beta/k) + t)^2}\right)\right)$$

$$\sim \text{NIG}(\alpha/k, \beta/k, k\mu, k\delta).$$

2. Let $Y = X_1 + X_2$. For independent variables $X_1, X_2$ we have $M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t)$. Using this property we get

$$M_Y(t) = \exp\left(t\mu_1 + \delta_1\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2}\right) \cdot \exp\left(t\mu_2 + \delta_2\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2}\right)$$

$$= \exp\left(\mu_1 + \mu_2 + (\delta_1 + \delta_2)\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2}\right)\right)$$

$$\sim \text{NIG}(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2).$$

3. Let $Y = \sum_{i=1}^n X_i$. From the first property $M_Y(t) = M_{\sum_{i=1}^n X_i}(t/n)$. Since $X_i, i = 1, \cdots, n$ are i.i.d we obtain

$$M_{\sum_{i=1}^n X_i}(t/n) = \exp n \cdot \left((t/n)\mu + \delta\left(\sqrt{(n\alpha)^2 - (n\beta)^2} - \sqrt{(n\alpha)^2 - (n\beta + t)^2}\right)\right)$$

$$= \exp\left(t\mu + \delta\left(\sqrt{(n\alpha)^2 - (n\beta)^2} - \sqrt{(n\alpha)^2 - ((n\beta) + t)^2}\right)\right)$$

$$\sim \text{NIG}(n\alpha, n\beta, \mu, \delta).$$

4. Let $Y = (X - \mu)/\delta$. Using the first property we obtain

$$M_Y(t) \overset{def}{=} \mathbb{E}(\exp\left(\left(X - \mu\right)\frac{t}{\delta}\right)) = \exp\left(-\frac{t\mu}{\delta}\right) \cdot M_X\left(\frac{t}{\delta}\right)$$

$$= \exp\left(-\frac{t\mu}{\delta}\right) \cdot \exp\left(t(\mu/\delta) + \left(\sqrt{(\alpha^2/\delta)^2 - (\beta^2/\delta)^2} - \sqrt{(\alpha^2/\delta)^2 - ((\beta^2/\delta) + t)^2}\right)\right)$$

$$= \exp\left(\sqrt{(\alpha^2/\delta)^2 - (\beta^2/\delta)^2} - \sqrt{(\alpha^2/\delta)^2 - ((\beta^2/\delta) + t)^2}\right)$$

$$\sim \text{NIG}(\delta\alpha, \delta\beta, 0, 1).$$

\qed
Chapter 3

The EM algorithm for the maximum likelihood (ML) estimation

We will use the EM algorithm (introduced by Dempste, Laird, Rubin in 1977, see [24]) for the maximum likelihood (ML) estimation of the Normal Inverse Gaussian distribution parameters. The theory behind this algorithm, for problems where some values are unobservable, was introduced in 1977 (formalization and proof of convergence) but even before the algorithm was used in practical applications (1952, Hartley).

The EM algorithm estimates unknown parameters using iterations, which stops when the difference between two steps start to be small enough. It is very important to remember that the maximum which we obtain using this method can only be a local maximum. For this reason it is important to choose proper initial values or to eliminate this problem by multiple running of the program with different initial values.

3.1 Initial values for the EM algorithm

The most popular way to find initial values for the EM algorithm is the Moments estimation. This method have not always an unique solution but in case which interest to us, i.e. the NIG distribution, we can find an analytical one, cf. [14].

\[ \hat{\gamma} = \frac{3}{s \sqrt{3\hat{\gamma}_2 - 5\hat{\gamma}_1^2}} \]
Using the Moments Method we obtain the following estimators of NIG parameters. By $\hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\alpha}$ we note estimators of NIG parameters, respectively $\beta, \delta, \mu, \alpha$.

$$\hat{\beta} = \frac{\bar{\gamma}_1 s^2}{3},$$
$$\hat{\delta} = \frac{s^2 \hat{\gamma}^3}{\beta^2 + \hat{\gamma}^2},$$
$$\hat{\mu} = \bar{x} - \hat{\beta} \frac{\hat{\delta}}{\hat{\gamma}},$$
$$\hat{\alpha} = \sqrt{\hat{\gamma}^2 + \beta^2},$$

where the mean is note by $\bar{x}$, the variance is $s^2$, the skewness is equal to $\bar{\gamma}_1 = \frac{\mu_3}{\mu_2^2}$ and the kurtosis is equal to $\bar{\gamma}_2 = \frac{\mu_4}{\mu_2^2} - 3$, where $\mu_k = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^k$.

We should remember that $3\bar{\gamma}_2 - 5\bar{\gamma}_1^2 \geq 0$ is a condition for the existence of moment estimates.

### 3.2 The log-likelihood function for the NIG distribution

Let us remind that the density function for the $NIG(\alpha, \beta, \mu, \delta)$ sample is given by the formula

$$g(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2} - \beta \mu) \phi(x)^{-\frac{3}{2}} K_1(\delta \alpha \phi(x)^{\frac{1}{2}}) \exp(\beta x),$$

where $K_1$ is the modified Bessel function of the third kind of the order $r$ evaluated at $x$ and the function $\phi(x)$ is define by

$$\phi(x) = 1 + \left(\frac{x - \mu}{\delta}\right)^2.$$
We calculate log-likelihood function

\[ \ln L(\theta, x) = \ln \left( \prod_{i=1}^{n} g(x_i; \alpha, \beta, \mu, \delta) \right) \]

\[ = \ln \left( \prod_{i=1}^{n} \frac{\alpha}{\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2 - \beta \mu}) \phi(x_i)^{-\frac{1}{2}} K_1(\delta \alpha \phi(x_i)^{\frac{1}{2}}) \exp(\beta x_i) \right) \]

\[ = \ln \left( \left( \frac{\alpha}{\pi} \right)^n \prod_{i=1}^{n} \exp(\delta \sqrt{\alpha^2 - \beta^2 - \beta \mu}) \phi(x_i)^{-\frac{1}{2}} K_1(\delta \alpha \phi(x_i)^{\frac{1}{2}}) \exp(\beta x_i) \right) \]

\[ = n \ln \alpha - n \ln \pi + n(\delta \sqrt{\alpha^2 - \beta^2 - \beta \mu}) - \frac{1}{2} \sum_{i=1}^{n} \ln \phi(x_i) + \]

\[ + \beta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \ln K_1(\delta \alpha \phi(x_i)^{\frac{1}{2}}) \]

We are using the log-likelihood function instead of the likelihood function. Due to the properties of the logarithmic function which is convex, the procedures of maximization of \( \ln g(x; \alpha, \beta, \mu, \delta) \) and \( g(x; \alpha, \beta, \mu, \delta) \) are equivalent. Unfortunately in our log-likelihood function the Bessel function is included so it is not evident how to find its maximum. To maximize the log-likelihood function we can apply the EM algorithm which, in comparison to the other method, is quite easy to implement. By the other methods we mean for example, methods developed by Blaesild and Sorensen (1992), i.e. the HYP program, which is mainly prepared for the generalized hyperbolic distribution and after some modification can also calculate the ML estimates of parameters for the NIG distribution, cf. [14].

Let us say that our data \( Y_i \) contains observable variables \( X_i \) and unobservable variables \( Z_i \). The vector of parameters is called \( \theta = (\alpha, \beta, \mu, \delta) \). Assume that the density function of \( Y_i \) can be expressed by the formula

\[ f(Y_i, \theta) = f(X_i, Z_i, \theta) = f(X_i, \theta) f(Z_i | X_i, \theta). \]

**Definition 3.1 (Exponential family).** We say that random variable \( X \) belongs to an exponential family if its density can be written in the form

\[ f_X(x|\theta) = a(\theta)b(x) \exp \sum_{i=1}^{n} c_i(\theta)d_i(x), \]

where \( \theta = (\theta_1, \theta_2, \ldots, \theta_k) \) is a vector of parameters, \( n > 0, a, b > 0 \) and \( \{c_i\}_i \), \( \{d_i\}_i \) are some functions.
Chapter 3. The EM algorithm for the maximum likelihood estimation

We know that the NIG probability distribution is a variance-mean mixture of a Gaussian distribution with an IG distribution. Let \( X \sim \text{NIG}(\alpha, \beta, \mu, \delta) \) and the variable \( Z | X \) is \( \text{GIG}(-1, \delta \sqrt{\phi(x)}, \alpha) \) so the density is given by (density of \( \text{GIG}(\lambda, \delta, \gamma) \), cf. formula (2.1))

\[
f(z, -1, \delta \sqrt{\phi(x)}, \alpha) = \frac{1}{2} \delta \sqrt{\phi(x) \alpha} K_{-1}^{-1}(\delta \sqrt{\phi(x) \alpha}) z^{-2} \exp \left( -\frac{1}{2} \left( \frac{\delta^2 \phi(x)}{z} + \alpha^2 z \right) \right).
\]

The density of \( Y \) is given by

\[
f(Y_i, \theta) = f(X_i, \theta) f(Z_i | X_i, \theta) \\
= \frac{\alpha}{\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2 - \beta \mu} \phi(x)^{-\frac{1}{2}}) K_{1}(\delta \alpha \phi(x)^{\frac{1}{2}}) \exp(\beta x) \\
= \frac{1}{2} \delta \sqrt{\phi(x) \alpha} K_{-1}^{-1}(\delta \sqrt{\phi(x) \alpha}) z^{-2} \exp \left( -\frac{1}{2} \left( \frac{\delta^2 \phi(x)}{z} + \alpha^2 z \right) \right) \\
= \frac{\delta}{2\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2 - \beta \mu} z^{-2}) \exp \left( -\frac{1}{2} \left( \frac{\delta^2 \phi(x)}{z} + \alpha^2 z \right) + \beta x \right) \\
= \frac{1}{2\pi z^2} e^{-\frac{\beta x}{z}} \exp \left( \beta x + \mu \frac{x}{z} - \alpha^2 \frac{z^2}{2} - (\delta^2 + \mu^2) \frac{1}{2z} - (\mu \beta - \delta \sqrt{\alpha^2 - \beta^2 - \ln \delta}) \right),
\]

and it is the density, which belongs to the exponential family of distributions introduced in Definition 3.1 with following parameters:

\[
a(\alpha, \beta, \mu, \delta) = e^{-(\mu \beta - \delta \sqrt{\alpha^2 - \beta^2} - \ln \delta)}, \\
b(z, x) = \frac{1}{2\pi z^2} e^{-\frac{\beta x}{2z}}, \\
c_1(\alpha, \beta, \mu, \delta) = \beta, \\
d_1(z) = x, \\
c_2(\alpha, \beta, \mu, \delta) = \mu, \\
d_2(z) = \frac{x}{z}, \\
c_3(\alpha, \beta, \mu, \delta) = \frac{\alpha^2}{2}, \\
d_3(z) = z, \\
c_4(\alpha, \beta, \mu, \delta) = -\frac{\delta^2 + \mu^2}{2}, \\
d_4(z) = \frac{1}{z}.
\]
Then the log-likelihood function is a linear function of a sufficient statistics. Hence, E step is simplified to calculate the expectation of the sufficient statistics of IG distributed variable \( Z \mid X \), cf. [9], [10].

The family of IG distributions is the exponential family of distributions with vector of parameters \( \theta = (\alpha, \beta, \mu, \delta) \)

\[
f(z, \alpha, \beta, \mu, \delta) = \frac{\delta}{\sqrt{2\pi}} \exp \left( \delta \sqrt{\alpha^2 - \beta^2} z^{-\frac{3}{2}} \exp \left( - \frac{1}{2} \left( \frac{\delta^2}{z} + (\alpha^2 - \beta^2) z \right) \right) \right)
\]

\[
= a(\alpha, \beta, \mu, \delta) b(z) \exp \left( \sum_{i=1}^{2} c_i(\alpha, \beta, \mu, \delta) d_i(z) \right),
\]

where

\[
a(\alpha, \beta, \mu, \delta) = \frac{\delta}{\sqrt{2\pi}} \exp \left( \delta \sqrt{\alpha^2 - \beta^2} \right),
\]

\[
b(z) = z^{-\frac{3}{2}},
\]

\[
c_1(\alpha, \beta, \mu, \delta) = -\frac{\delta^2}{2},
\]

\[
d_1(z) = \frac{1}{z},
\]

\[
c_2(\alpha, \beta, \mu, \delta) = -\frac{\alpha^2 - \beta^2}{2},
\]

\[
d_2(z) = z.
\]

We obtain the log-likelihood function as a linear function of the sufficient statistics \( \sum Z_i^{-1}, \sum Z_i \) (see e.g. [9], [14], [10]).

### 3.3 The steps of the EM algorithm

An EM algorithm contains two main steps:

- **E(expectation)-step**
  The unobserved (hidden, missing) data are estimated using the observed data and the current parameters estimation. For the NIG which is an exponential family of distributions it can be achieved by calculation of the conditional expectation of the sufficient statistics for the Inverse Gaussian distribution (\( \sum Z_i^{-1} \) and \( \sum Z_i \)), it means that we calculate

\[
s_i = E(z_i \mid x_i, \theta^{(k)}) = \frac{\delta^{(k)} \sqrt{\beta^{(k)}(x_i)}}{\alpha^{(k)}} K_0 \left( \delta^{(k)} \alpha^{(k)} \sqrt{\beta^{(k)}(x_i)} \right) \frac{1}{K_1 \left( \delta^{(k)} \alpha^{(k)} \sqrt{\beta^{(k)}(x_i)} \right)}
\]
and

\[ w_i = E(z_i^{-1}|x_i, \theta^{(k)}) = \frac{\alpha^{(k)}}{\delta^{(k)} \sqrt{\phi^{(k)}(x_i)}} \frac{K_2}{K_1} \left( \frac{\delta^{(k)} \alpha^{(k)}}{\sqrt{\phi^{(k)}(x_i)}} \right), \]

for each step \( k \), where \( \theta^{(k)} \) is the current values of the parameter \( \theta \).

- **M(maximization)-step**
  In this step we assume that unobservable data are known (we are using the expectations from the previous step) and we maximize the likelihood function. Let us introduce new variables

\[ \hat{M} = \sum_{i=1}^{n} \frac{s_i}{n}, \]
\[ \hat{\Lambda} = n \left( \sum_{i=1}^{n} (w_i - \hat{M}^{-1}) \right)^{-1}, \]

and then the estimation of \( \delta^{(k+1)} \) and \( \gamma^{(k+1)} \) is given by the expressions

\[ \delta^{(k+1)} = \sqrt{\hat{\Lambda}}, \]
\[ \gamma^{(k+1)} = \frac{\delta^{(k+1)}}{\hat{M}}. \]

Using the maximum likelihood method for \( X_i|Z_i \) which is normal distributed with the mean \( \mu^{(k+1)} + \beta^{(k+1)}z_i \) and the variance \( \sqrt{z_i} \), we find the estimators of the parameters \( \mu^{(k+1)} \) and \( \beta^{(k+1)} \). The log-likelihood function for \( X_i|Z_i \) is given by

\[ L_{X_i|Z_i}(\theta, x) = \ln \left( \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi z_i}} e^{-\frac{(x_i - (\mu + \beta z_i))^2}{2z_i}} \right) \]
\[ = \ln \left( \left( \frac{1}{\sqrt{2\pi z_i}} \right)^n \prod_{i=1}^{n} e^{-\frac{(x_i - (\mu + \beta z_i))^2}{2z_i}} \right) \]
\[ = n \ln \frac{1}{\sqrt{2\pi z_i}} + \ln e^{-\sum_{i=1}^{n} \frac{(x_i - (\mu + \beta z_i))^2}{2z_i}} \]
\[ = n \ln \frac{1}{\sqrt{2\pi z_i}} - \sum_{i=1}^{n} \frac{(x_i - \mu - \beta z_i)^2}{2z_i}. \]
and the partial derivatives by

\[
\frac{\partial L_{X_i|Z_i}(\theta, x)}{\partial \mu} = -\frac{\partial}{\partial \mu} \sum_{i=1}^{n} \left( x_i - \mu - \beta z_i \right)^2 2z_i
\]

\[
= -\sum_{i=1}^{n} \frac{\partial}{\partial \mu} x_i^2 - 2x_i(\mu + \beta z_i) + (\mu + \beta z_i)^2 2z_i
\]

\[
= -\sum_{i=1}^{n} \frac{\partial}{\partial \mu} x_i^2 + \mu^2 + 2\mu(\beta z_i - x_i) - 2x_i\beta z_i + \beta^2 z_i^2 2z_i
\]

\[
= -\sum_{i=1}^{n} \frac{\mu}{z_i} - \sum_{i=1}^{n} \frac{\beta z_i - x_i}{z_i} = -\sum_{i=1}^{n} \frac{\mu}{z_i} - n\beta + \sum_{i=1}^{n} \frac{x_i}{z_i},
\]

\[
\frac{\partial L_{X_i|Z_i}(\theta, x)}{\partial \beta} = -\frac{\partial}{\partial \beta} \sum_{i=1}^{n} \left( x_i - \mu - \beta z_i \right)^2 2z_i
\]

\[
= -\sum_{i=1}^{n} \frac{\partial}{\partial \beta} x_i^2 + \mu^2 + 2\mu(\beta z_i - x_i) - 2x_i\beta z_i + \beta^2 z_i^2 2z_i
\]

\[
= -n\mu + \sum_{i=1}^{n} x_i - \beta \sum_{i=1}^{n} z_i.
\]

Now we need to solve following system of equations for the \(\mu\) and \(\beta\)

\[
\begin{cases}
-\sum_{i=1}^{n} \frac{\mu}{z_i} - n\beta + \sum_{i=1}^{n} \frac{x_i}{z_i} = 0, \\
-n\mu + \sum_{i=1}^{n} x_i - \beta \sum_{i=1}^{n} z_i = 0.
\end{cases}
\]

From the second equation we get the estimation of \(\mu\)

\[
\mu^{(k+1)} = \bar{x} - \beta^{(k+1)} \bar{s},
\]

and then we obtain that the parameter \(\beta^{(k+1)}\) is given by expression

\[
\beta^{(k+1)} = \frac{\sum_{i=1}^{n} x_i w_i - \bar{x} \sum_{i=1}^{n} w_i}{n - s \sum_{i=1}^{n} w_i}.
\]

Finally, we find the estimator of \(\alpha^{(k+1)}\) in form

\[
\alpha^{(k+1)} = \sqrt{(\gamma^{(k+1)})^2 + (\beta^{(k+1)})^2},
\]

where

\[
\bar{s} = \frac{\sum_{i=1}^{n} s_i}{n}.
\]
Iterations of our EM algorithm are stopped when the difference between the successive steps in the algorithm is small enough. It means that for some fixed \( a \) we find a \( j \) such that

\[
\max \left( \frac{\delta^{(j+1)} - \delta^{(j)}}{\delta^{(j+1)}}, \frac{\mu^{(j+1)} - \mu^{(j)}}{\mu^{(j+1)}}, \frac{\beta^{(j+1)} - \beta^{(j)}}{\beta^{(j+1)}}, \frac{\alpha^{(j+1)} - \alpha^{(j)}}{\alpha^{(j+1)}} \right) < a. \quad (3.1)
\]
Chapter 4

The Lindberg method

The Carl Lindberg paper’s main goal [21] was to fit the theoretical model of Barndorff-Nielsen and Shephard [6] to a daily data and to find a stable estimation of the model’s parameters. He used 5-years (1999-2004) data from the Ericsson stock from OMX Stockholmsbörsen.

The theoretical model of Barndorff-Nielsen and Shephard is based on the idea that the stochastic volatility of the asset price dynamics can be given by a weighted sum of the Non-Gaussian Ornstein-Uhlenbeck processes which can be written in the form

\[ dy = -\lambda y(t)dt + dz(t), \]

where \( \lambda > 0, z \) is subordinator.

The biggest difficulty of this model lie in the estimation of model’s parameters from data. From the theoretical point of view we get that we can find parameters of the volatility process from a real data, but when we try to apply this model using a quadratics variation of the stock price process we discover the following drawbacks: our data are not continuous and even downloading every second data only give us a large sample and a discontinuity of the price will exist because of fact that the stock market is not working at night. Other problem is that the intensity of trading is growing during certain hours.

The first and main step of the Lindberg method is the discretization made by the assumption that

\[ \int_{t-\Delta}^{t} \sigma(s)dB(s) \approx \sigma(t)\epsilon, \]

where \( \epsilon \sim N(0,1) \).

In the Barndorff-Nielsen and Shephard [6] work very important feature is
that the centered return divided by volatility is, i.i.d., and standard normal
distributed. Lindberg was taking as a volatility the daily number of trades.
He focused on finding the most reasonable volatility data and after that he
estimated the parameters of returns and the volatility distribution. This way
is easier to implement, use much less data and obtain more stable estimation
of the parameters. The quadratic variation method is harder to implement.

4.1 The Barndorff-Nielsen and Shephard’s model

Let us consider a probability space \((\Omega, \mathcal{F}, P)\) with a natural filtration \(\{\mathcal{F}_t\}_{0 < t < \infty}\),
where we take a cadlag version of the \(Z_i\) m-independent subordinators (Levy
process with non-negative values).
Following Barndorff-Nielsen and Shephard [7] we take \(B\) as a Brownian motion
and Ornstein-Uhlenbeck stochastic process \(Y_j, j = 1, \ldots, m\) with dy-
namics given by the formula

\[
dY_j(t) = -\lambda_j Y_j(t) dt + dZ_j(\lambda_j t),
\]

where \(\lambda_j > 0\) means the rate of decay and The process \(Y_i\) can be given by
the expression

\[
Y_i(t) = \int_{-\infty}^{0} \exp (s) dZ_j(\lambda_j t + s), \quad t \geq 0,
\]
or equivalently by

\[
Y_j(t) = Y_j(0) \exp (-\lambda_j t) + \int_{0}^{t} e^{-\lambda_j (t-s)} dZ_j(\lambda_j s), \quad t \geq 0,
\]

where \(Y_j(0)\) is independent from \(Z_j(t) - Z_j(0)\), where \(t \geq 0\) and has a sta-
tionary marginal distribution.

We assume that \(Z_j(0) = 0\) for all \(j = 1, 2, \ldots, m\) and we define the
volatility process as a sum with weights \(\omega_j\), which are summing up to 1. Let
us write formula for \(\sigma^2(t)\)

\[
\sigma^2(t) = \sum_{j=1}^{m} \omega_j Y_j(t), \quad t \geq 0, \omega_j \geq 0.
\]
Let us notate the mean rate of return as $\mu$ and the skewness parameter as $\beta$. Then the stock price dynamics and the stock price process has the form

$$
dS(t) = S(t) \left( (\mu + \beta \sigma^2(t))dt + \sigma(t)dB(t) \right),
$$

$$
S(t) = S(0)e^{\int_0^t (\mu + (\beta - \frac{1}{2})\sigma^2(s))ds + \int_0^t \sigma(s)dB(s)}.
$$

Proof.

We show that the really stock price dynamics has a solution (a stock price process) in the form showed above. We use Itô’s formula in the following form.

**Theorem 4.1.** If $F(X_t, t) \in C^2$ and $(X_t)_{t \geq 0}$ is an Itô process then

$$
dX_t = a(\omega, t)dt + b(\omega, t)dB_t,
$$

where

$$
P\left( \int_0^t |a(\omega, t)|ds < \infty \right) = 1,
$$

$$
P\left( \int_0^t b^2(\omega, t)ds < \infty \right) = 1
$$

and $a(\omega, t)$ and $b(\omega, t)$ are non-anticipating processes, then

$$
dF(x, s) = \left( \frac{\partial F}{\partial s}(x, s) + a(\omega, t)\frac{\partial F}{\partial x}(x, s) + \frac{1}{2}b(\omega, t)\frac{\partial^2 F}{\partial x^2}(x, s) \right)ds + b(\omega, t)\frac{\partial F}{\partial x}(x, s)dB(s).
$$

In our case we consider process

$$
S(t) = S(0)\exp\left( \int_0^t (\mu + (\beta - \frac{1}{2})\sigma^2(s))ds + \int_0^t \sigma(s)dB(s) \right)_{X_t},
$$

and we choose function $F(x, s) = S_0e^x$, then derivatives of function $F(x, s)$ are in following form

$$
\frac{\partial F}{\partial s}(x, s) = 0, \quad \frac{\partial F}{\partial x}(x, s) = S_0e^x, \quad \frac{\partial^2 F}{\partial x^2}(x, s) = S_0e^x,
$$

$$
S(t) = F(X_t, s).
$$

Using the Itô’s formula we obtain from (4.2)

$$
\frac{dF(X_t)}{dS(t)} = \left( (\mu + (\beta - \frac{1}{2})\sigma^2(t))S_0e^{X_t} + \frac{1}{2}\sigma^2(t)S_0e^{X_t} \right)dt + \sigma(t)S_0e^{X_t}dB(t)
$$

$$
= S(t)\left( (\mu + \beta \sigma^2(t))dt + \sigma(t)dB(t) \right).
$$

$\square$
Chapter 4. The Lindberg method

4.2 The Lindberg method

The increments of the returns \( R^c(t) = \ln \frac{S(t)}{S(0)} \) are stationary, cf. [21], where we have equality in law of \( R^c(s) - R^c(t) = R^c(s - t) \). From \( R^c(s) - R^c(t) = \ln \frac{S(s)}{S(t)} - \ln \frac{S(t)}{S(0)} = \ln \frac{S(s)}{S(0)} \), we get the following equality \( R^c(s - t) \overset{d}{=} \ln \frac{S(s)}{S(0)} \).

In the discrete case with \( d + 1 \) observations of returns taken every period \( \Delta \), we get the identical distributed sequence

\[
R^c(\Delta), R^c(2\Delta), \ldots, R^c(d\Delta) - R^c((d - 1)\Delta).
\]

To simplify this expression we take \( \Delta = 1 \). In model we use the approximation that

\[
\int_{t-\Delta}^{t} \sigma(s)dB(s) \approx \sigma(t)\epsilon, \quad \epsilon \sim N(0,1).
\]

We get that the log-returns for all \( t = 1, 2, \ldots \) can be written in form:

\[
R(t) = \ln \frac{S(t)}{S(t-\Delta)} = \ln \frac{S(0)e^{\int_{t-\Delta}^{t}(\mu+(\beta-\frac{1}{2})\sigma^2(s))ds+\int_{t-\Delta}^{t} \sigma(s)dB(s)}}{S(0)e^{\int_{0}^{t}(\mu+(\beta-\frac{1}{2})\sigma^2(s))ds+\int_{0}^{\Delta} \sigma(s)dB(s)}}
\]

\[
= \ln e^{\int_{0}^{t}(\mu+(\beta-\frac{1}{2})\sigma^2(s))ds+\int_{0}^{\Delta} \sigma(s)dB(s)} - \int_{t-\Delta}^{t} \sigma(s)dB(s)
\]

\[
= \int_{t-\Delta}^{t} (\mu + (\beta - \frac{1}{2})\sigma^2(s))ds + \sigma(t)\epsilon
\]

\[
\Delta \equiv 1 \mu + \beta \sigma^2(t) + \sigma(t)\epsilon,
\]

where \( \epsilon(\cdot) \) is i.i.d. and \( N(0,1) \).

From the assumption that the volatility \( \sigma^2 \) is Inverse Gaussian distributed, and using the formula \( R = \mu + \beta \sigma^2(t) + \sigma(t)\epsilon \) we obtain that the returns \( R \) are Normal Inverse Gaussian distributed (cf. [21]).

Lindberg, [21], stressed that when we try to estimate parameters of the NIG distribution we need to deal with the fact of ‘almost’ overparametrization of NIG. From this fact we obtain an unstability of parameter values. Following Lindberg we show example which illustrate that for a group of NIG parameterizations for the returns which fit to data we obtain significantly different IG distribution parameterizations.
Example 4.1. We have two sets of the NIG parameters, which give us almost undistinguish fit to the return and to the corresponding IG parametrization, which are visibly different.

The NIG parameter sets are

\[ \alpha_1 = 73.8, \beta_1 = 14.7, \mu_1 = -0.0092, \delta_1 = 0.0591 \]  \hspace{1cm} (4.3)

\[ \alpha_2 = 211, \beta_2 = 108, \mu_2 = -0.0659, \delta_2 = 0.114 \]  \hspace{1cm} (4.4)

Figure 4.1: In (a) and (b): solid line corresponds to the set (4.3), dotted line corresponds to the set (4.4).

Easy to see is that for a lot of sets of the IG parameters we obtain a good NIG parametrization for returns.

To eliminate this problem Lindberg propose to use equation for the log-returns \( R \) of type

\[ R = \mu + \beta \sigma^2 + \sigma \epsilon, \]  \hspace{1cm} (4.5)

by the estimation of the parameters of the NIG-distributed returns.

4.3 The steps of the method

The method is presented in the three steps.

**Step 1**
Firstly we need to find parameters \( \mu, \beta \) and the additional parameter \( \theta \) by
using the maximum log-likelihood function. We use the maximum likelihood function for the normalized returns
\[
\xi(\cdot) = \frac{R(t) - (\mu + \beta \sigma^2(t))}{\sigma(t)} \sim N(0, 1),
\]
and assumption that
\[
\sigma^2(t) = \theta z(t),
\]
where \(\theta\) is constant coefficient and \(z(t)\) is a number of trades/ volume/ amount for each day \(t\).
Then
\[
\xi(t)\sigma(t) = R(t) - (\mu + \beta \sigma^2(t)) \sim N(0, \sigma^2(t)). \tag{4.6}
\]
We calculate the log-likelihood function
\[
\ln L(\mu, \beta, \theta) = \ln \left( \prod_{t=1}^d \frac{1}{\sqrt{2\pi\sigma(t)}} e^{-\frac{(R(t) - (\mu + \beta \sigma^2(t)))^2}{2\sigma^2(t)}} \right)
\]
\[
= \ln \left( \frac{1}{\sqrt{2\pi}} \right)^d - \frac{1}{2} \sum_{t=1}^d \ln \sigma^2(t) + \ln \left( \prod_{t=1}^d e^{-\frac{(R(t) - (\mu + \beta \sigma^2(t)))^2}{2\sigma^2(t)}} \right)
\]
\[
= d \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \sum_{t=1}^d \ln \sigma^2(t) + \ln \left( e^{-\frac{1}{2} \sum_{t=1}^d \frac{(R(t) - (\mu + \beta \sigma^2(t)))^2}{\sigma^2(t)}} \right)
\]
\[
= d \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \sum_{t=1}^d \left( \frac{(R(t) - (\mu + \beta \sigma^2(t)))^2}{\sigma^2(t)} + \ln \sigma^2(t) \right)
\]
\[
= d \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \sum_{t=1}^d \left( \frac{(R(t) - (\mu + \beta \theta z(t)))^2}{\theta z(t)} + \ln \theta z(t) \right).
\]
To maximize the log-likelihood function we need to take into account the relation between parameters \(\mu, \beta\) and \(\theta\). We consider the system of equations with the constraint in the following form
\[
\mathcal{L} = \ln L + \lambda (\mu + \beta \theta E z - ER),
\]
where
\[
\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial \beta} = \frac{\theta}{\phi} \ln L + \lambda \frac{\partial \phi}{\partial \beta}, \\
0 &= \frac{\partial \mathcal{L}}{\partial \mu} = \frac{\phi}{\theta} \ln L + \lambda \frac{\partial \phi}{\partial \mu}, \\
0 &= \frac{\partial \mathcal{L}}{\partial \theta} = \frac{\phi}{\theta} \ln L + \lambda \beta z,
\end{aligned}
\]
We calculate derivatives of the log-likelihood function along parameters $\mu$, $\beta$, $\theta$. We obtain

$$\frac{\partial}{\partial \mu} \ln L(\mu, \beta, \theta) = -\frac{1}{2} \frac{\partial}{\partial \mu} \sum_{t=1}^{d} \frac{\left( R(t) - (\mu + \beta \theta z(t)) \right)^2}{\theta z(t)}$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \frac{\partial}{\partial \mu} R^2(t) - 2R(t)(\mu + \beta \theta z(t)) + \mu^2 + 2\mu \beta \theta z(t) + \beta^2(\theta z(t))^2}{\theta z(t)}$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \frac{-2R(t) + 2\mu + 2\beta \theta z(t)}{\theta z(t)} = \sum_{t=1}^{d} \frac{R(t)}{\theta z(t)} - \mu \sum_{t=1}^{d} \frac{1}{\theta z(t)} - d\beta,$$

$$\frac{\partial}{\partial \beta} \ln L(\mu, \beta, \theta) = -\frac{1}{2} \frac{\partial}{\partial \beta} \sum_{t=1}^{d} \frac{\left( R(t) - (\mu + \beta \theta z(t)) \right)^2}{\theta z(t)}$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \frac{\partial}{\partial \beta} R^2(t) - 2R(t)(\mu + \beta \theta z(t)) + \mu^2 + 2\mu \beta \theta z(t) + \beta^2(\theta z(t))^2}{\theta z(t)}$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \frac{-2R(t) \theta z(t) + 2\mu \theta z(t) + 2\beta (\theta z(t))^2}{\theta z(t)}$$

$$= \sum_{t=1}^{d} R(t) - d\mu - \beta \sum_{t=1}^{d} \theta z(t),$$

$$\frac{\partial}{\partial \theta} \ln L(\mu, \beta, \theta) = -\frac{1}{2} \frac{\partial}{\partial \theta} \sum_{t=1}^{d} \left( \frac{\left( R(t) - (\mu + \beta \theta z(t)) \right)^2}{\theta z(t)} + \ln \theta z(t) \right)$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \frac{\partial}{\partial \theta} \left( R^2(t) - 2R(t)(\mu + \beta \theta z(t)) + \mu^2 + 2\mu \beta \theta z(t) + \beta^2(\theta z(t))^2 \right) + \ln \theta z(t)$$

$$= -\frac{1}{2} \sum_{t=1}^{d} \left( \beta^2 z(t) - \frac{\mu^2 - 2\mu R(t) + R^2(t)}{\theta^2 z(t)} + \frac{1}{\theta} \right).$$

Hence, we obtain the system of equations

$$\begin{cases}
\sum_{t=1}^{d} R(t) / \theta z(t) - \mu \sum_{t=1}^{d} \frac{1}{\theta z(t)} - d\beta + \lambda = 0, \\
\sum_{t=1}^{d} R(t) - d\mu - \beta \sum_{t=1}^{d} \theta z(t) + \lambda \theta z(t) = 0, \\
-\frac{1}{2} \sum_{t=1}^{d} \left( \beta^2 z(t) - \frac{\mu^2 - 2\mu R(t) + R^2(t)}{\theta^2 z(t)} + \frac{1}{\theta} \right) + \lambda \beta z(t) = 0
\end{cases}
(4.7)$$
From the first equation (4.7) we obtain \( \lambda \) in the form

\[
\lambda = d\beta + \mu \sum_{t=1}^{d} \frac{1}{\theta z(t)} - \sum_{t=1}^{d} \frac{R(t)}{\theta z(t)},
\]

(4.8)

and from the second equation (4.7) after the substitution (4.8) we obtain \( \mu \) in the form

\[
\sum_{t=1}^{d} R(t) - d\mu - \beta \sum_{t=1}^{d} \frac{\mathbb{E}R(t)z(t) - \mu z(t)}{\beta \mathbb{E}z(t)} + d(\mathbb{E}R(t) - \mu) + \mu \mathbb{E}z(t) \sum_{t=1}^{d} \frac{1}{z(t)} - \mathbb{E}z(t) \sum_{t=1}^{d} \frac{R(t)}{z(t)} = 0,
\]

so

\[
\mu \sum_{t=1}^{d} \frac{z(t)}{\mathbb{E}z(t)} - 2d\mu + \mu \mathbb{E}z(t) \sum_{t=1}^{d} \frac{1}{z(t)} = \mathbb{E}z(t) \sum_{t=1}^{d} \frac{R(t)}{z(t)}
\]

\[
- d\mathbb{E}R(t) + \frac{\mathbb{E}R(t)}{\mathbb{E}z(t)} \sum_{t=1}^{d} z(t) - \sum_{t=1}^{d} R(t)
\]

\[
\mu = \frac{\mathbb{E}z(t) \sum_{t=1}^{d} \frac{R(t)}{z(t)} - d\mathbb{E}R(t) + \frac{\mathbb{E}R(t)}{\mathbb{E}z(t)} \sum_{t=1}^{d} z(t) - \sum_{t=1}^{d} R(t)}{-2d + \frac{1}{\mathbb{E}z(t)} \sum_{t=1}^{d} z(t) + \mathbb{E}z(t) \sum_{t=1}^{d} \frac{1}{z(t)}}.
\]

Hence for the parameter \( \mu \) we obtain representation

\[
\mu = \frac{1}{d} \sum_{t=1}^{d} R(t) - \frac{\beta}{d} \sum_{t=1}^{d} \theta z(t).
\]
In the last equation we substitute $\lambda$ using the formula (4.8)

$$
-\frac{1}{2} \beta^2 \sum_{t=1}^{d} z(t) + \frac{\mu^2}{2 \theta^2} \sum_{t=1}^{d} \frac{1}{z(t)} - \frac{\mu}{\theta} \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \frac{1}{2 \theta^2} \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} - \frac{d}{2 \theta} + d \beta^2 \mu E(z(t))
$$

$$
+ \mu \beta \theta E(z(t)) \sum_{t=1}^{d} \frac{1}{z(t)} - \frac{\beta \theta}{2} \mu E(z(t)) \sum_{t=1}^{d} \frac{R(t)}{z(t)} = 0
$$

$$
\beta^2 \theta^2 \sum_{t=1}^{d} z(t) + \mu^2 \sum_{t=1}^{d} \frac{1}{z(t)} - 2 \mu \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} - d \theta + 2 \mu \theta \mu E(z(t)) \sum_{t=1}^{d} \frac{1}{z(t)} = 0
$$

$$
-2 \beta \theta \mu E(z(t)) \sum_{t=1}^{d} \frac{1}{z(t)} = 0
$$

Using the relation $\beta \theta = \frac{E(R(t) - \mu)}{E(z(t))}$ we get for the parameter $\theta$

$$
\theta = \frac{1}{d} \left( - \left( \frac{E(R(t) - \mu)}{E(z(t))} \right)^2 \sum_{t=1}^{d} z(t) + \mu^2 \sum_{t=1}^{d} \frac{1}{z(t)} - 2 \mu \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} \right)
$$

$$
+ 2d \left( \frac{E(R(t) - \mu)}{E(z(t))} \right)^2 E(z(t)) + 2 \mu \left( \frac{E(R(t) - \mu)}{E(z(t))} \right) E(z(t)) \sum_{t=1}^{d} \frac{1}{z(t)} - 2 \left( \frac{E(R(t) - \mu)}{E(z(t))} \right) E(z(t)) \sum_{t=1}^{d} \frac{R(t)}{z(t)} \right) = 0
$$

Hence this parameter can be represented in the form

$$
\theta = \frac{1}{d} \left( - \left( \frac{E(R(t) - \mu)}{E(z(t))} \right)^2 \sum_{t=1}^{d} z(t) + \mu^2 \sum_{t=1}^{d} \frac{1}{z(t)} - 2 \mu \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} \right)
$$

The parameter $\beta$ is calculated from $\mu + \theta \beta \mu E(z(t)) - E(R(t)) = 0$

$$
\beta = \frac{1}{\theta} \left( \frac{\sum_{t=1}^{d} R(t)}{\sum_{t=1}^{d} z(t)} - \frac{d \sum_{t=1}^{d} R(t)}{d \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} + \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)} \right)
$$

$$
= \frac{1}{\theta} \left( \frac{d \sum_{t=1}^{d} R(t)}{d \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} - \frac{1}{d} \sum_{t=1}^{d} R(t) \sum_{t=1}^{d} \frac{1}{z(t)} - \frac{d \sum_{t=1}^{d} R(t)}{d \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} + \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)} \right)
$$

$$
= \frac{1}{\theta} \left( \frac{d \sum_{t=1}^{d} R(t)}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} \right).
$$

Hence as the result we get the estimators of $\mu$ in the form

$$
\mu = \frac{d \sum_{t=1}^{d} R(t)}{d \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} - \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}.
$$

(4.10)
Then we easily calculate $\theta$ from the formula (4.9)

$$\theta = \frac{1}{d} \left( \mu^2 \sum_{t=1}^{d} \frac{1}{z(t)} - \left( \frac{E R(t) - \mu}{E z(t)} \right)^2 \sum_{t=1}^{d} z(t) - 2\mu \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} \right),$$

(4.11)

Finally we obtain $\beta$ in form

$$\beta = \frac{1}{\theta} \left( \frac{d}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} \right),$$

(4.12)

where

$$\lambda = d\beta + \mu \sum_{t=1}^{d} \frac{1}{\theta z(t)} - \sum_{t=1}^{d} \frac{R(t)}{\theta z(t)}$$

$$= \frac{1}{\theta} \left( \frac{d}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} \right)$$

$$+ \frac{d}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} \sum_{t=1}^{d} \frac{1}{\theta z(t)} \sum_{t=1}^{d} z(t) - \sum_{t=1}^{d} \frac{R(t)}{\theta z(t)}$$

$$= \frac{1}{\theta} \left( \frac{d^2 \sum_{t=1}^{d} \frac{R(t)}{z(t)} - d \sum_{t=1}^{d} R(t) \sum_{t=1}^{d} \frac{1}{z(t)} + d \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} \right)$$

$$+ \frac{- \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)} - d^2 \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}}{d^2 - \sum_{t=1}^{d} z(t) \sum_{t=1}^{d} \frac{1}{z(t)}} = 0.$$  

**Step 2**

In the second step we need to calculate the parameters $\alpha$ and $\delta$. We estimate these parameters by the estimation of parameters of an empirical distribution of $\sigma^2(t) = \theta z(t)$, which is $IG(\delta, \sqrt{\alpha^2 - \beta^2})$ distributed.

It means that using, for example, the $R$ function we obtain $\delta$ and $\gamma$ and using this parameters we get

$$\alpha = \sqrt{\gamma^2 + \beta^2}.$$
Chapter 5

The measure of risk – Value-at-Risk

Usually three kinds of risk are associated with company activities:

- the business risk (The risk that the company will fail to make a profit.),
- the credit risk (It is a risk that a borrower fails to pay the interest in a timely manner.),
- the financial risk (The risk related to the structure of the company capital.).

The first type of risk, the business risk, is the essence of running any business. The other two risks, although undesirable, are always unavoidable. Hence, the companies should learn to treat these risks in a proper way.

The idea of Value-at-Risk appeared when some of the biggest financial institutions began investigating the risk measurement and aggregating the risk in these institutions as a whole. One of the most popular system is the RiskMetrics, originally developed at the American investment bank J.P.Morgan. The concept of this system appeared when the president of this bank demanded daily reports about the risks and potential losses arisen in the upcoming 24 hours as a result of the price changes on the stock market and the currency market, as well as changes of the interest rate. The RiskMetrics approach measures the all type of risk of all assets in the institution, and then aggregates all these risks to the one measure of risk called Value-at-Risk, VaR.

5.1 Definition

Suppose that we are interested in answering the following question: *What is the maximum value of the potential loss which the company can suffer*
during the known time with determined probability from the investment of the known value?
The tool which is useful to give the answer is Value at Risk.

**Definition 5.1.** [8] Let \( R_{t,D} \) be a random variable, whose outcome determines the value of the next D-days return. Then

\[
\text{Var}_{\alpha,D} = - \inf \{ x \in \mathbb{R} : P(R_{t,D} \leq x) \geq \alpha \}. \tag{5.1}
\]

The minus sign in the formula (5.1) assures positive value of VaR. Usually, in calculations, we use one daily VaR, denoted shortly by \( \text{VaR}_\alpha \). The value \( \alpha \) is the probability level, very often assumed to be 1%. Sometimes, when we have to calculate the D-daily VaR, we can use the following approximation, (cf. [8]),

\[
\text{VaR}_{\alpha,D} \approx \sqrt{D} \times \text{Var}_\alpha.
\]

The above formula will hold exactly for additive returns, i.i.d. for the Gaussian returns with the zero mean.

### 5.2 Methods of VaR evaluation

There are many methods to calculate the Value at Risk. In our work we will describe shortly the three most popular methods

- the historical method,
- the Monte Carlo method,
- the variance-covariance approach.

#### 5.2.1 The historical method

In the historical simulation approach, we use returns calculated from historical data, for example from the last 200 or 250 days. This method assumes that the price changes will be the same in the future, what can be perceived as its weakness. Then the hypothetical future price is defined as the movements of the historical price to the current price

\[
S_{i,t+s} = S_{i,t+s-1}^* + \Delta S_{i,t+s-k},
\]

where \( t \) is the current time, \( s = 1, 2, \ldots, k \) is a length of backward-time horizon, \( S_{i,t+s}^* \) is a hypothetical value of price of \( i \)-th asset at time \( t + s \) and \( \Delta S_{i,t+s-k} = S_{i,t+s-k}^* - S_{i,t+s-1-k} \) is the historical price of \( i \)-th asset at time \( t \).
If we consider the case $i = 1$ and the time horizon equals to 1, then the return at the time $t + s$ is given by the following formula

$$X^*_p,t+s = X^*_p,t+s-1 - X^*_p,t,$$

where $X^*_p,t$ is the current return.

The strong feature of this method is that it belongs to the type of non-parametric methods. It means that we do not assume any distribution of returns, hence we avoid procedure of the estimations of parameters based on the historical data.

### 5.2.2 Monte Carlo method

In contrast to the previous method, Monte Carlo method is based on the assumptions about the statistical model. First, we have to estimate the unknown parameters from the historical data and then using a stochastic simulation we generate the possible future prices and returns. Finally the VaR is simulated from the distribution of the portfolio value.

Now we present the algorithm describing this method, (cf. [15]).

1. Specify the stochastic processes and the process of the parameters for the financial variables and correlations.
2. Simulate the hypothetical price trajectories for all variables of interest. The hypothetical price changes are obtained by simulations, draws from the specified distribution.
3. Obtain the asset prices at time $T$, $S_i,T$, from the simulated price trajectories. Compute the portfolio value $P_{p,T} = \sum w_i,T S_i,T$.
4. Repeat steps 2 and 3 many times to form the distribution of the portfolio value $P_{p,T}$.
5. Measure $\text{VaR}_T$ equals to minus the $(1 - \alpha)$-quantile of the simulated distribution for $P_{p,T}$.

The positive feature of the Monte Carlo method is that we can use a very large set of observations. On the other hand, the main disadvantage may lie in the assumptions done on the returns’ distribution. In many cases, the calculations are simplified if the distribution used is the univariate or the multivariate Gaussian distribution.
5.2.3 The variance-covariance approach

The third method – variance-covariance approach – belongs to the set of the parametrical methods and is mostly based on the Gaussian assumption. It is the well-known fact, that the Gaussian distribution does not describe returns well. A distribution with heavy tails, for example, the General Hyperbolical distribution, suits the real data better. In our master thesis we use the special case of this family – the normal inverse Gaussian distribution. Suppose that we have a continuous distribution with the cumulative function $F_{R_t}$ and a density function $f_{R_t}$.

Then we define VaR as below

$$1 - \alpha = \int_{-\infty}^{-\text{VaR}_\alpha} f_{R_t}(s) ds = F_{R_t}(-\text{VaR}_\alpha),$$

which can also be written in the equivalent way as

$$\text{VaR}_\alpha = -F_{R_t}^{-1}(1 - \alpha),$$

where the distribution function $F_{R_t}$ is usually inverted numerically.

In our study we assume that the returns have the following form

$$R_t = \mu + \sigma_t X_t,$$

where the Lévy process is denoted by $X_t$. Since the returns are dependent, we consider the devolatized returns. Let the volatility $\sigma_t > 0$. Then

$$X_t = \frac{R_t - \mu}{\sigma_t}$$

form the i.i.d. sequence. If we take the same $\alpha$–level for $X_t$ and $R_t$, we calculate both $\text{VaR}_{\alpha}^{R_t}$, $\text{VaR}_{\alpha}^{X_t}$ respectively and obtain that

$$\text{VaR}_{\alpha}^{R_t} = \mu + \sigma_t \text{VaR}_{\alpha}^{X_t}.$$

5.3 The advantages and the disadvantages of VaR method

The conception of the risk measure, which is called Value at Risk method, is quite attractive for the institutions. Although it has many positive points, this method also has many drawbacks. Let us examine the advantages and the disadvantages of the VaR method, (see [18]).

Advantages of the VaR method are
- Universality – the same conception is used to measure different kinds of risk, which simplifies comparison and creates the aggregated risk measure.

- It determines the probability of the event that the risk factor changes its value by a given amount.

- This measure admits a simple interpretation.

- It can be used to determine the risk exposition of the institution capital.

- It takes into consideration the effects of portfolio diversification.

- It is popular. The concept of VaR has now been incorporated in the Basel II Capital Accord.

Disadvantages of the VaR method are

- It describes the loss during the ‘normal’ market behaviour and under specified assumptions such as time horizon or tolerance level. Hence, in the abnormal behaviour of market the VaR utility can be limited.

- It does not determine how large a loss will be, if the value of VaR will be crossed.

- It is not a coherent measure of risk in a general case, i.e. when the returns of portfolio have another distribution than the multidimensional normal or other multidimensional elliptic distributions.

- An accurate estimation may be a problem, especially for a complicated portfolio.

- Results of the estimation are sensitive to the method of estimation.

- It can be calculated only for traded assets or liabilities, but not for credit risks, deposits or loans.

- Recently, an Expected Tail Loss is considered which is a better risk measure than VaR
5.4 The Expected Tail Loss method (ETL)

We want to find a risk measure, which reflects diversification effects and satisfies the sub-additivity condition. The measure, which contains those information, belongs to the class of coherent measures. Suppose that $X$ and $Y$ are future values of two risky positions. Then we say that the risk measure $\rho(\cdot)$ is a coherent measure, \cite{9}, if it satisfies the conditions listed below

- **monotonicity:**
  \[
  \text{if } X \leq Y \Rightarrow \rho(X) \geq \rho(Y),
  \]

- **homogeneity:**
  \[
  \rho(tX) = t\rho(X),
  \]

- **sub-additivity:**
  \[
  \rho(X + Y) \geq \rho(X) + \rho(Y),
  \]

- **risk-free condition:**
  \[
  \rho(X + n) = \rho(X) - n,
  \]

for $n \in \mathbb{R}$ and $t \geq 0$. The first two conditions – the monotonicity and the homogeneity – do not seem to be unreasonable and additionally imply that the risk function $\rho$ is a convex function, i.e.

\[
\rho(tX + (1-t)Y) \leq t\rho(X) + (1-t)\rho(Y).
\]

The risk-free condition means that the addition of an amount $n$ of a riskless asset to our position will increase the value of our portfolio at the end of the period and it will decrease the risk. The last condition, the sub-additivity condition, matches decentralised decisions.

The most popular and perhaps the most attractive coherent measure is ETL (Expected Tail Loss), which is also called conditional VaR, \cite{11}.

**Definition 5.2.** \cite{9} Let $X_t$ describe a distribution of returns from an financial instrument in one unit of time. Then for the significance level $1 - \alpha$ the value of $\text{VaR}_\alpha$ is known and ETL has the following form

\[
\text{ETL} = -\mathbb{E}(X_t|(X_t < -\text{VaR}_\alpha))
\]

at the end of the time unit.

In contrast to VaR, ETL gives us information about the expected average loss, if an extremal event occurs. The influence of extremal events cause that the value of ETL is greater than the corresponding value of VaR. ETL is a better risk measure than VaR, because

- ETL, in contrast to VaR, as a coherent measure which always satisfies the sub-additivity conditions.

- ETL takes into account a risk diversification, while VaR does it occasionally.
• The sub-additivity of ETL measure implies that the risk function is always convex and it means we know that this function has always a unique well-behaved optimum.
Chapter 6

Consistency tests

Let \( x_1, x_2, \ldots, x_n \) be a taken sample. Suppose that the cumulative distribution function of a random variable \( X \) (an observed characteristic) is unknown and we put forward a hypothesis

\[ H_0 : F_0(x) \text{ is the cumulative distribution function of } X. \]

With respect to the kind of distribution, the above hypothesis can be presented in the following equivalent forms

- in the case of a continuous random variable

  \[ H_0 : \text{the density function of the observed characteristic is equal to } f_0(x), \]

- and in the discrete case, for \( i \in \mathbb{N} \)

  \[ H_0 : \text{the probability function of the observed characteristic is equal to } P(X = x_i) = p_i. \]

If the hypothesis functions \( F_0, f_0, P \) are completely described (all parameters of the distribution are known), then we have a nonparametric hypothesis. Otherwise, the hypothesis is parametric and in this situation parameters should be estimated, [17].

If the goal of our study is to estimate the distribution of a sample, we want to be sure that we obtain a good result. In this situation, a statistical tool called consistency test [17] is very useful.

**Definition 6.1.** The test to verify the hypothesis corresponding to consistency between the distribution of a sample range and a theoretical one is called the consistency test.
Chapter 6. Consistency tests

The most popular tests are the Pearson’s chi-square test, the Kolmogorov’s test and, for the normal distribution, the Shapiro-Wilk test or the Kolmogorov-Lilliefors test. In our study, we need a consistency test for the NIG distribution. Hence, we choose tests for a continuous case. We focus on the three following tests: the Kolmogorov-Smirnov test, the Anderson-Darling test and the Kuiper’s test.

6.1 The Kolmogorov-Smirnov’s test (K-S test)

The aim of this test is to verify that a characteristic $X$ is $F_0$ distributed. Assuming that a cumulative distribution function of $X$ is a continuous function and the null hypothesis is a simple hypothesis (a hypothesis which completely specifies the distribution of the observed random variables [17]), the test statistic is given by

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x; \mathcal{X}) - F_0(x)|,$$

where $F_n(x; \mathcal{X})$ is the empirical distribution function defined on the basis of an ordered sample

$$F_n(x; \mathcal{X}) = \begin{cases} 0, & \text{for } x < x(1), \\ \frac{k}{n}, & \text{for } x(k) \leq x < x(k+1), \quad 1 \leq k \leq n - 1, \\ 1, & \text{for } x \geq x(n). \end{cases}$$

**Theorem 6.1** (Cantelli-Glivenko [17]). If the characteristic $X$ has a cumulative distribution function $F$, then for all $x$

$$P(\lim_{n \to \infty} F_n(x; \mathcal{X}) = F_0(x)|H \text{ is true}) = 1.$$  

This theorem implies that

$$P(\lim_{n \to \infty} D_n = 0|H \text{ is true}) = 1.$$  

Let us calculate the probability $P(D_n \leq d)$

$$P(D_n \leq d) = P \left( \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)| \leq d \right) =$$

$$= P \left( \sup_{0 < t < 1} \left| \sum_{i=1}^{n} \chi_{(-\infty,F^{-1}(t])}(X_i) - F(F^{-1}(t)) \right| \leq d \right)$$

$$= P \left( \sup_{0 < t < 1} \left| \sum_{i=1}^{n} \chi_{(1,t]}(F(X_i)) - F(F^{-1}(t)) \right| \leq d \right)$$

$$= P \left( \sup_{0 < t < 1} \left| \sum_{i=1}^{n} \chi_{(1,t]}(F(X_i)) - t \right| \leq d \right).$$
In this way we showed that, under the assumption that the hypothesis $H_0$ is true, the statistic $D_n$ is independent on the distribution $F$.

In practical applications we do as follows [17]

1. Put measurements in the ascending order

$$x(1) \leq x(2) \leq \cdots \leq x(n);$$

2. Calculate all differences

$$\frac{i}{n} - F_0(x(i)), \text{ for } i \in 1, \cdots, n,$$

and find the maximum of the absolute values of them, denoted by $d_n^+$

$$d_n^+ = \max_{1 \leq i \leq n} \left| \frac{i}{n} - F_0(x(i)) \right|;$$

3. Calculate all differences

$$F_0(x(i)) - \frac{i - 1}{n}$$

and find the maximum of the absolute values of them, denoted by $d_n^-$

$$d_n^- = \max_{1 \leq i \leq n} \left| F_0(x(i)) - \frac{i - 1}{n} \right|;$$

4. Choose the greater value of $d_n^+, d_n^-$

$$d_n = \max(d_n^+, d_n^-);$$

5. For an assumed significance level $\alpha$ and $n$, read the critical value $d_{n,1-\alpha}$ of the Kolmogorov statistic from the special table ([17], p. 298);

6. Make a decision

- if $d_n < d_{n,1-\alpha}$, then we do not have grounds for a rejection of the verified hypothesis $H_0$ at the significance level $\alpha$,

- if $d_n \geq d_{n,1-\alpha}$, then we reject our hypothesis at the significance level $\alpha$.

The critical range at the significance level $\alpha$ is $[d_{n,1-\alpha}, 1]$. 
Chapter 6. Consistency tests

6.2 The Anderson-Darling test (A-D test)

In 1952 Anderson and Darling, [2], proposed a new distance as a measure of closeness of a theoretical distribution $F_0$ with an empirical one $F_n$. This metric is defined as

$$D_{AD} = \sup_{x \in \mathbb{R}} \left( \sqrt{n} \left| F_n(x) - F_0(x) \right| \sqrt{\psi(x)} \right),$$

(6.2)

where $\psi(x)$ denotes a weight function and $n$ is the size of the sample. Perhaps the most common weight function is represented by

$$\psi(x) = \frac{1}{\sqrt{F_0(x)(1 - F_0(x))}}.$$  

(6.3)

Since $F_0 \approx 0$ and $F_0 \approx 1$ in the tails of pdf, the weight function (6.3) is large for the tails, [13]. The disadvantage of this test, in contrast to K-S test, is that the test depends on the theoretical distribution of a sample. It means that critical values must be calculated for each distribution independently.

6.3 The Kuiper’s test

The Kuiper’s test is similar to the K-S test, but it distinguishes the direction of the deviation. The test statistic has the following form

$$D_K = \max_{x \in \mathbb{R}} (F_0(x) - F_n(x)) + \max_{x \in \mathbb{R}} (F_n(x) - F_0(x)).$$

(6.4)

While in the K-S test we calculate the maximum of the absolute values of distance 6.1 between a theoretical and an empirical cumulative functions, in the Kuiper’s case we take the sum of the maximum of a distance upwards and downwards difference between the theoretical and the empirical cumulative functions. This small modification causes that the Kuiper’s test is more sensitive in the tails and also makes it invariant under the cyclic transformations of the independent variable, [13]. The last property gives us the possibility to use this test for a variation of time.
Chapter 7

The study of the statistical data for the Ericsson B stock

7.1 The features of the statistical data

In our investigation we perform statistical analysis on the basis of the Ericsson B stock data for the period of five years – from January 1st, 1999 to December 30th, 2004. Using the software package SIX Edge\textsuperscript{TM}, we downloaded the following data:

- the stock price of the Ericsson B asset (in SEK) for every trading day (see a plot of data in Figure 7.1).

![Stock prices of Ericsson B](image)

Figure 7.1: The stock price of the Ericsson B asset in SEK for every trading day.
Chapter 7. The study of the statistical data for the Ericsson B stock

- the stock price of the OMX Stockholm 30 (in SEK) (Figure 7.2)

Figure 7.2: The daily stock price OMX Stockholm 30 in SEK.

- the 22-days volatility of the Ericsson B stock (Figure 7.3)

Figure 7.3: The volatility of the Ericsson B stock calculated for 22-days period of times.

- the number of trades for the Ericsson B stock presented in Figure 7.4
Figure 7.4: The number of trades for the Ericsson B stock for every trading day for non hidden orders.

- the trading volume for the Ericsson B stock given in Figure 7.5

Figure 7.5: The volume of the Ericsson B stock trading in SEK.

- the amount of the Ericsson B stock presented in Figure 7.6.
  As the amount we mean turnover for the stock, i.e. amount is a stock price multiply by volume for each day.
Chapter 7. The study of the statistical data for the Ericsson B stock

Figure 7.6: The amount of the Ericsson B for every trading day in SEK.

To compare our results with Lindberg, [21], we chose the data from the same stock and for the same period as in the Lindberg’s paper. For the calculation purpose and for the further study, we have to introduce returns which are defined by the formula

\[ \text{return}_n = \log(price_n) - \log(price_{n-1}), \quad n = 1, 2, \ldots, N, \]

where \( N \) is the sample size.

In our study we use a linear function of some measure of the trading intensity \( z(t) \) (e.g. the volume of traded asset, number of trades or the amount) as a volatility, following Lindberg method. We assume that the volatility should be IG distributed. To make our data more convenient to numeric calculations we normalize \( z(t) \) by division it by \( \max_t z(t) \). This transformation does not influence the volatility distribution, were described in section 2.1.

7.2 The EM algorithm

The EM algorithm, as described in the theoretical part section 3, can be used to estimate parameters of the distribution by the MLE method. Following by the Lindberg method [21], we expect that the log-returns are NIG distributed, so we are interested in an estimation of the NIG parameters. In the case of the NIG distribution, the EM algorithm for the maximum likelihood (ML) estimation is simplified (Chapter 3).

We implemented the algorithm in R. The code of the program ME.R can be found in the Appendix.
7.2.1 The estimation of the NIG parameters

We have two possibilities to obtain the values of the parameters of the NIG distribution. Running the program ME.R for different precisions (see formula (3.1) for more details) give us the results, which are collected in the Table 7.1.

<table>
<thead>
<tr>
<th>error</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \delta )</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^{-5} )</td>
<td>24.471765587</td>
<td>0.262086811</td>
<td>0.048374357</td>
<td>-0.001346234</td>
</tr>
<tr>
<td>( 10^{-6} )</td>
<td>24.473673942</td>
<td>0.262132845</td>
<td>0.048376990</td>
<td>-0.001346313</td>
</tr>
<tr>
<td>( 10^{-7} )</td>
<td>24.471766718</td>
<td>0.262137495</td>
<td>0.048377256</td>
<td>-0.001346320</td>
</tr>
</tbody>
</table>

Table 7.1: The results of the application of the EM algorithm for the NIG distribution.

The results fulfil statistical tests what is show in the further part of work and with decreasing error became more precise.

The second method, which we apply to calculate the parameters of the NIG distribution was function \textit{nigFit} from the R package \textit{fBasics}. Hence, we get

\[
\alpha = 25.17190832, \quad \beta = 0.3236879, \quad \mu = 0.0499951, \quad \delta = -0.0013618. \quad (7.1)
\]

Since the values represented in the Table 7.1 and in equation (7.1) are close to each other, we can conclude that the two R estimation results are confirmed.

In Figure 7.7 we compare the distribution obtained using the EM algorithm for \( a = 10^{-7} \) with the distribution obtained using \textit{nigFit}. The functions are almost indistinguishable, what is a result of ‘almost’ overparametrization of the NIG distribution.

![Figure 7.7: The NIG distribution with the parameters obtained using the EM method and function \textit{nigFit}. The NIG 1 curve is the NIG(24.471766718, 0.262137495, 0.048377256, -0.001346320) and the NIG 2 curve is the NIG(25.171908324, 0.323687998, 0.049995180, -0.001361818).](image)
Chapter 7. The study of the statistical data for the Ericsson B stock

To focus our attention in the following analysis we consider the NIG parameters computed by the EM algorithm with the error level \( a = 10^{-7} \)

\[ \alpha = 24.471766718, \beta = 0.262137495, \delta = 0.048377256, \gamma = -0.001346320. \]

The Goodness-of-fits

Now we would like to test if our estimates of the NIG parameters fit to the log-returns. For this aim we use the statistical Kolmogorov-Smirnov test, which is described in Chapter 6. For the significance level at \( 1 - \alpha_{level} = 0.95 \) the value of statistic is equal to \( D = 0.0357 \) and the p-value less then 0.08172. It means that we do not have any reasons to reject the hypothesis about the NIG distribution with the set of parameters \( \alpha = 24.471766718, \beta = 0.262137495, \delta = 0.048377256, \mu = -0.001346320) \).

7.2.2 The IG parameters

Now we want to find the IG parameters implied from the NIG distribution with the parameters \( \alpha = 24.471766718, \beta = 0.262137495, \delta = 0.048377256, \mu = -0.001346320 \). We compute the parameters IG(\( \delta, \gamma \)) from the NIG(\( \alpha, \beta, \delta, \mu \)) distribution in the following way

\[ \delta_{IG} = \delta_{NIG}, \]
\[ \gamma_{IG} = \sqrt{\alpha_{NIG}^2 - \beta_{NIG}^2}. \]

We obtained the values

\[ \delta_{IG} = 0.04837726, \gamma_{IG} = 24.47036. \]

We see that the values of \( \gamma \) and \( \alpha \) are very close to each other, because the value \( \beta^2 \) is very small.

7.3 The Lindberg method

Another method allowing to obtain the NIG(\( \alpha, \beta, \delta, \mu \)) distribution parameters for the returns \( R \) indirectly is the Lindberg method, which we described in Chapter 4. This method is based on the estimation of the IG parameters for \( \sigma^2 \) and the following formula

\[ R(t) = \mu + \beta \sigma^2(t) + \sigma(t) \xi(t), \quad (7.2) \]
where $\xi \sim N(0, 1)$.

In our study we assume that the volatility can be represented as $\sigma^2(t) = \theta z(t)$, where $\theta$ is a constant coefficient and $z(t)$ is the number of trades, the traded volume or the amount. Converting equation (7.2) to the equation for the function $\xi(t)$ we obtain normalized return

$$
\xi(t) = \frac{R(t) - \mu - \beta \sigma^2(t)}{\sigma(t)},
$$

It means that normalized return is a standard normal distributed and has i.i.d returns. Using this fact, we can calculate the maximum likelihood function to get the values of parameters $\mu$, $\beta$ and $\theta$, what was described in section 4.3. In the next step we calculate the IG distribution parameters $\delta$ and $\gamma$ for our volatility $\sigma^2$ and then use the transformation $\alpha = \sqrt{\gamma^2 + \beta^2}$. In this way we obtain all the NIG parameters.

### 7.3.1 The estimation of the parameters $\mu$, $\beta$ and $\theta$ for NIG distribution

Now we calculate the parameters $\mu$, $\beta$ and $\theta$ using formulas (4.10), (4.11) and (4.12), which are in form

$$
\mu = \frac{d \sum_{t=1}^{d} \frac{R(t)}{z(t)} - \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}}{\sum_{t=1}^{d} \frac{d^2}{z(t)} - \sum_{t=1}^{d} \frac{d}{z(t)}},
$$

$$
\theta = \frac{1}{d} \left( \mu^2 \sum_{t=1}^{d} \frac{1}{z(t)} \left( \frac{ER(t) - \mu}{EZ(t)} \right)^2 \sum_{t=1}^{d} z(t) - 2\mu \sum_{t=1}^{d} \frac{R(t)}{z(t)} + \sum_{t=1}^{d} \frac{R^2(t)}{z(t)} \right),
$$

$$
\beta = \frac{1}{\theta} \frac{d \sum_{t=1}^{d} \frac{R(t)}{z(t)} - \sum_{t=1}^{d} \frac{R(t)}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}}{d^2 - \sum_{t=1}^{d} \frac{d^2}{z(t)} \sum_{t=1}^{d} \frac{1}{z(t)}}.
$$

The values for $\mu$, $\beta$ and $\theta$ are obtained by the maximization of log-likelihood function (4.7). This method was described early in the paragraph 4.3. We get the following results for the parameter values.
Chapter 7. The study of the statistical data for the Ericsson B stock

For different measures of the trading intensity we received different values of the parameters. Both the number of trades and the volume give the value of parameter $\beta$ less than zero and $\mu$ slightly greater than zero. In the case of the amount of traded assets the signs of these two parameters are inverse. For all types of measures the value of the parameter $\theta$ is positive. Additionally, the amount and the volume values are close to each other.

### 7.3.2 The normalized returns $\xi$

Now we would like to check if the obtained parameters $\mu$, $\beta$ and $\theta$ listed in Table 7.2 give the $N(0,1)$ distributed normalized returns $\xi$, formula for $\xi$ was noted as (4.6). If such set of the parameters cannot be calculated, the assumption that $\sigma^2(t) = \theta z(t)$ is incorrect and must be rejected.

The null hypothesis means that $\xi$ has the standard normal distribution was tested using the Kolmogorov-Smirnov and Jaque-Bare tests. The value of statistics and p-value are depicted in the Table 7.3 below.

<table>
<thead>
<tr>
<th></th>
<th>Jarque-Bare</th>
<th>Kolmogorov-Smirnov</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JB</td>
<td>p-value</td>
</tr>
<tr>
<td>number of trades</td>
<td>3.66726</td>
<td>0.1598</td>
</tr>
<tr>
<td>volume</td>
<td>4.5932</td>
<td>0.1006</td>
</tr>
<tr>
<td>amount</td>
<td>$2.2837 \cdot 10^3$</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.3: The value of statistics and p-values.

From the Table 7.3 follows that we cannot reject the null hypothesis for the number of trades and for the volume of trades assets at the confidence level $\alpha_{level} = 0.05$, because all p-values for both tests are greater than $\alpha_{level}$. Hence, we can assume that the volatility is a linear function of each of them. However for the amount, both tests fail. This means that this value should be dropped from further analysis. Nevertheless, we still carry out next steps for all the data and show that amount does not give any good return parametrization.
### 7.3.3 The IG parameters estimation

In the next step we estimate the parameters of the IG distribution for $\sigma^2 = \theta z$. For this aim we use the R function `mleig`. Next, we can easily compute the parameter $\alpha = \sqrt{\gamma^2 + \beta^2}$. The Table 7.4 shows the values of $\delta$, $\gamma$ and $\alpha$ obtained in this manner.

<table>
<thead>
<tr>
<th></th>
<th>$\delta$</th>
<th>$\gamma$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trades</td>
<td>0.07016653</td>
<td>44.26799</td>
<td>44.33701</td>
</tr>
<tr>
<td>volume</td>
<td>0.05292669</td>
<td>27.13098</td>
<td>27.16631</td>
</tr>
<tr>
<td>amount</td>
<td>0.06642841</td>
<td>20.91521</td>
<td>20.948</td>
</tr>
</tbody>
</table>

Table 7.4: The estimated parameters $\delta$, $\gamma$, $\alpha$ after the `mleig` procedure.

One can immediately see that the values of the parameter $\alpha$ are almost equal to the values of the parameter $\gamma$. This is caused by the $\beta$-values are very small in all studied cases.

**The Goodness-of-fit**

Using the Kolmogorov-Smirnov statistical test, we check the goodness-of-fit of our ”indirect” volatility given by the function $\sigma^2(t) = \theta z(t)$ to the obtained IG distributions.

- The number of trades
  
  The Kolmogorov-Smirnov statistic: $D = 0.023$, p-value= 0.5201.

![Histogram of theta*number of trades](image_url)

Figure 7.8: The histogram of the $\theta z$, where $z$ is number of trades and the correspond density.
Chapter 7. The study of the statistical data for the Ericsson B stock

- The volume of trades assets
  The Kolmogorov-Smirnov statistic: $D = 0.0342$, p-value$= 0.1073$.

Figure 7.9: The histogram of the $\theta z$, where $z$ is volume and the correspond density.

- The amount for trades asset
  The Kolmogorov-Smirnov statistic: $D = 0.0383$, p-value$= 0.05107$.

Figure 7.10: The histogram of the $\theta z$, where $z$ is amount of the trades assets and the correspond density.
In all cases the p-value is greater than the confidence level $\alpha_{level} = 0.05$, which implies that all measures of the trading intensity (the number of trades, the volume and the amount) are IG distributed. Histograms of each trading intensity measure and the corresponding IG probability density functions are shown on Figures 7.8 – 7.10. These histograms confirm the correct fit to our data.

### 7.3.4 The NIG parameters

In the previous steps we obtained all parameters for the NIG distribution, which are collected and depicted in the Table 7.5 below.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\delta$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trades</td>
<td>44.33701</td>
<td>-2.472929</td>
<td>0.07016653</td>
<td>0.003126057</td>
</tr>
<tr>
<td>volume</td>
<td>27.16631</td>
<td>-1.385037</td>
<td>0.05292669</td>
<td>0.001908274</td>
</tr>
<tr>
<td>amount</td>
<td>20.948</td>
<td>1.171596</td>
<td>0.06642841</td>
<td>-0.004514719</td>
</tr>
</tbody>
</table>

Table 7.5: The estimated parameters of the NIG distribution obtained by the Lindberg method.

We see that the greatest value of the parameter $\alpha = 44.33701$ is achieved for the number of trades, i.e. this case has the greatest concentration around $\mu$. The parameter $\beta$, which reflects the skewness of the distribution, is less than zero for both the number of trades and the volume of the trades assets. It implies that the NIG distribution is skewed to the left. In the case of the chosen parameter equal to the amount of trades assets, the distribution has the fatter right tail because of $\beta > 0$. Finally, the parameter $\mu$ says us that for the two first measures the suitable density functions are shifted to the right (since $\mu$ is greater than 0) and for the amount to the left.

![Histogram of return](image.png)

Figure 7.11: The histogram of the returns and the densities obtained by Lindberg method.
Chapter 7. The study of the statistical data for the Ericsson B stock

The Goodness-of-fit

We carried out the goodness-of-fit analysis using the Kolmogorov-Smirnov test. The following Table 7.6 contains the values of the test statistics and the corresponding p-values.

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trades</td>
<td>0.0279</td>
<td>0.2852</td>
</tr>
<tr>
<td>volume</td>
<td>0.0265</td>
<td>0.3442</td>
</tr>
<tr>
<td>amount</td>
<td>0.0643</td>
<td>6.42 \cdot 10^{-5}</td>
</tr>
</tbody>
</table>

Table 7.6: The Kolmogorov-Smirnov statistics and the p-values.

From the Table 7.6 we conclude that for the number of trades and for the volume of trades assets we obtain a good estimation of the NIG parameters, because the p-values are considerably greater than the confidence level \( \alpha_{level} = 0.05 \). It means that both measures reflect the behaviour of the volatility well. However in the case of the amount parameter, we obtain p-value less than 0.05, i.e. the amount of traded assets cannot be interpreted as a function \( z(t) \). The reason is that this measure does not normalize the returns, see the paragraph 7.3.2.

7.4 The EM algorithm versus Lindberg method

Using the two different methods, the Lindberg method and the EM algorithm, we obtain the following estimates of the NIG parameters:

- The EM algorithm give us following values for \( \alpha, \beta, \delta, \mu \)
  \[
  \alpha = 24.471766718, \quad \beta = 0.262137495, \quad \delta = 0.048377256, \quad \mu = -0.001346320, 
  \]

- The Lindberg method lead to the two sets of parameters listed in the Table 7.7

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \delta )</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trades</td>
<td>44.33701</td>
<td>-2.472929</td>
<td>0.07016653</td>
<td>0.003126057</td>
</tr>
<tr>
<td>volume</td>
<td>27.16631</td>
<td>-1.385037</td>
<td>0.05292669</td>
<td>0.001908274</td>
</tr>
</tbody>
</table>

Table 7.7: The estimated parameters of the NIG distribution after the Lindberg method.
We see that all the parameters reach the greatest absolute values if we take
in $\sigma^2(t) = \theta z(t)$ as the parameter the number of trades. For the both the
volume and the number of trades the parameter $\beta$ is less than zero, while for
the EM algorithm it is greater than 0. In case of the parameter $\mu$, it is oppo-
site i.e. $\mu > 0$ for Lindberg method and $\mu < 0$ for the EM approach. Despite
these differences, the statistical tests confirmed that all the sets estimate re-
turns well. The plots of the density functions of these three distributions are
almost undistinguishable, see Figure 7.12. It implies that all methods give
us the same information about the behaviour of the returns.

![NIG densities](image)

**Figure 7.12:** The NIG densities obtained using the EM method and the
Lindberg method with the number of trades and the volume of traded assets
taken as trading intensity.
Now we compare the IG distribution of the volatility obtained from these three cases. Recall, that the values of the parameters are

- for the EM algorithm
  \[ \delta_{IG} = 0.04837726, \quad \gamma_{IG} = 24.47036, \]

- for the Lindberg method

<table>
<thead>
<tr>
<th></th>
<th>( \delta )</th>
<th>( \gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trades</td>
<td>0.07016653</td>
<td>44.26799</td>
</tr>
<tr>
<td>volume</td>
<td>0.05292669</td>
<td>27.13098</td>
</tr>
</tbody>
</table>

Table 7.8: The estimated IG parameters obtained by the Lindberg method.

To test if the distributions of the volatility obtained by the EM algorithm and by the Lindberg method are the same, we used K-S test. The test statistic equals to 0.1409 and the p-value less than \( 2.2 \cdot 10^{-16} \). Hence, we are forced to reject the null hypothesis. Figure 7.13 compares all three pdf’s. Indeed, the density functions differ significantly.

Figure 7.13: The IG distributions obtained after the EM method and the Lindberg method with the number of trades and the volume of traded assets taken as trading intensity.
7.5 The VaR and ETL methods

In this section our main goal is to present the VaR and the ETL methods applied to stock prices of the Ericsson B asset in the same time period. Now we carry out the analysis of the Value-at-Risk (VaR) method and the Expected Tail Loss (ETL) method, which is more sensitive for the influence of extremal events, like underlined in the Chapter 5.

First our analysis will be based on the period of the five years from January 1st, 1999 to December 30th, 2004. Our results are presented in the paragraphs below.

7.5.1 The case of the number of trades as the measure of trading intensity

We present the obtained realization of VaR$_\alpha$ and ETL$_\alpha$, for confidence level $\alpha$ equal to 1% and 5%.

The Value-at-Risk is obtained by The Variance-Covariance approach, which we described in the Chapter 5.2.3. In our model we assume $\mu = 0$. The definition of the ETL method is presented in Section 5.4.

![Graph](image)

Figure 7.14: The losses, VaR and ETL for $\sigma^2 = \theta z$, with $z$ taken as the number of trades. Confidence levels are 1% and 5%. The bottom plot the losses, the middle plot presents the values of VaR, the top plot gives the values of ETL.
The ETL value as a measure of the risk is more sensitive of the extreme events. Using this measure we expect higher losses than using the Value-at-Risk measure, what is visible on Figure 7.14. We checked also how often our real losses cross VaR and ETL curves and the table 7.9 depicted the results.

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Method</th>
<th>number of crosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>VaR</td>
<td>8</td>
</tr>
<tr>
<td>1%</td>
<td>ETL</td>
<td>2</td>
</tr>
<tr>
<td>5%</td>
<td>VaR</td>
<td>44</td>
</tr>
<tr>
<td>5%</td>
<td>ETL</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 7.9: The number of crosses between losses and the VaR/ETL values in the case of the numbers of trades as the measure of trading intensity.

The analysis of the crosses confirmed that the ETL method allows to predict the expected losses better than the VaR method. However, the percentage of losses at the confidence level 1%, which cross the value VaR$_{1\%}$ is smaller than 0.7% ($= 8/1252 = 0.006389776$).

### 7.5.2 The case of the volume of the traded assets as the measure of trading intensity.

When instead of the number of trades we use the model with the traded volume as the measure of trading intensity, our VaR and ETL relations are still the same. Differences can be observed between the risk measures and the obtained losses. In the case of the last period of the one and half year, so in time of the high volatility, we can easily noticed a quite large difference between the losses and the VaR values. It allows us to put forward the hypothesis that the volume as a measure of the volatility gives the less risky VaR, the ETL levels than the number of trades.
Figure 7.15 presents results discussed above.

![Figure 7.15](image)

Figure 7.15: The losses, the VaR values and the ETL values for $\sigma^2 = \theta z$ with $z$ taken as the volume of the traded assets. The confidence levels are equal to 1% and 5%. The bottom plot are losses, the middle the VaR values and the top the ETL values.

The number of crosses of the VaR values and the ETL values are presented in the Table 7.10.

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Method</th>
<th>number of crosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>VaR</td>
<td>14</td>
</tr>
<tr>
<td>1%</td>
<td>ETL</td>
<td>1</td>
</tr>
<tr>
<td>5%</td>
<td>VaR</td>
<td>62</td>
</tr>
<tr>
<td>5%</td>
<td>ETL</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 7.10: The number of the crosses between the losses and the VaR/ETL values for the case as the $z$ taken as the volume.
The comparison between the crosses obtained in the number of trades and the volume case shows that the VaR method better express the risk of the losses when the volatility is a linear function of the number of trades, although the Kolmogorov-Smirnov test takes the volume case as better fitted to the log-returns, see Table 7.6.

### 7.6 The VaR and ETL predictions

In the previous part of our work we introduced and presented the estimation of stock prices as well as the stochastic volatility in a way, which includes the better understanding of the structure of the volatility process. After analysing the Var an ETL methods for period of the five years for which we have market data, we proceed to one year forward prediction (to end of the year 2004).

The value of VaR is calculated on the basis of the obtained parametrizations of the returns and the volatility. For the number of trades and the volume, we obtained other parametrizations so we present results in two cases.

#### 7.6.1 The case of the number of trades as the measure of trading intensity

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Method</th>
<th>number of crosses</th>
<th>number of crosses in last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>VaR</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>1%</td>
<td>ETL</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5%</td>
<td>VaR</td>
<td>53</td>
<td>9</td>
</tr>
<tr>
<td>5%</td>
<td>ETL</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 7.11: The number of the crosses between the losses and the VaR/ETL values in case of the number of trades with the prediction for one year forward.

One can easily see that the estimation of the risk by the ETL method still gives the better results, i.e. the number of the crosses is smaller than for the VaR method. For one year ahead we obtained a good prediction of the losses. Real losses cross the ETL prediction level in the last year only 2 times at the 1% confidence level and 3 times at the 5% level, what give us respectively $2/252=0.8\%$ and $3/252=1.2\%$ crosses.
Figure 7.16: The losses, the VaR and the ETL values for $\sigma^2 = \theta z$, with $z$ taken as the number of trades with the prediction of one year. The confidence level is equal to 1% and 5%, respectively. The bottom plot presents the losses, the middle plot the VaR values and the top plot the ETL values.

### 7.6.2 The case of the volume of traded assets as the measure of trading intensity

The Table 7.12 presents the number of the crosses between the losses and the ETL values and the VaR values.

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Method</th>
<th>number of crosses</th>
<th>no of crosses in the last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>VaR</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>1%</td>
<td>ETL</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5%</td>
<td>VaR</td>
<td>67</td>
<td>5</td>
</tr>
<tr>
<td>5%</td>
<td>ETL</td>
<td>15</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 7.12: The number of the crosses between the losses and the VaR/ETL values in case of the volume equal to $z$ with the prediction for one year ahead.
As well in this case we obtained that the ETL method is better measure of risk. For the both confidence levels the number of crosses in the last year divided by the number of trading day (252) are smaller than 1% and 5% respectively, which implies good one year forward prediction. Figure 7.17 shows the above described situation.

Figure 7.17: The losses, the VaR and ETL values for $\sigma^2 = \theta z$, with $z$ taken as the volume of traded assets with the prediction for one year. The confidence levels are equal to 1% and 5%. The bottom plot presents the losses, the middle plot the VaR values, the top plot the ETL values.
Chapter 8

Conclusions

In this master thesis, an application of the NIG distribution to the estimation of the returns was considered. We took into account two methods of estimations. The first method is based on the Maximum Likelihood approach. A numerically efficient approach to solve the Maximum Likelihood problem is to use the Expectation Maximization algorithm.

Our study shows that using the EM algorithm usually leads to a good fit for the returns. However, it does not allow one to obtain any meaningful and accurate information about the volatility.

The Lindberg method uses daily data to fit the simple model of a stochastic volatility. It was assumed that a volatility is a linear function of some measure of the trading intensity. The number of trades, the volume of traded assets and the amount were used for this purpose. However, the results obtained using the amount of the traded assets failed some statistical tests, which means that only the number of trades and the volume may be successfully used to describe the volatility. In case of the number of trades we obtained the same values of all the NIG-distributed returns parameters as Lindberg in his paper, [21]. Further analysis revealed that adopting the volume leads to the best fit to market data.

The next part of our work presented in Sections 7.5, 7.6 was analysis of the Value-at-Risk and Expected Tail Loss methods. These were computed using the Lindberg method under the assumption that a volatility process is a linear function of the volume and the number of trades.

The analysis of real stock data for the period from 1999 to 2004 presented in our work that we obtained less crosses between VaR_{1\%}, VaR_{5\%} and the losses in the case of the volume of traded assets than for corresponding VaR level in the case of the number of trades. It means that using the volume to estimate the volatility makes our investment less risky. Our research also confirmed that the ETL method is a better risk measure than the VaR method and
it reached greater values than VaR method, [18] and [11]. Again, we found that using the volume makes investment less risky than in the case of the number of trades.

In the final part of our study, the Value-at-Risk and Expected Tail Loss methods were used to predict behaviour of risk exposure for one year forward. The real losses and predictions based on the parametrization obtained by Lindberg method were compared and again it turned out that the ETL limits losses better than the VaR method. It means that this measure provide us with more information about the expected losses and how high they are.

We want to stress out the need to carry out a new analysis in case of a different data set. It means that when we choose other stock we cannot be sure that the volatility process can be expressed as a linear function of the trading intensity.
Bibliography


   Normal Inverse Gaussian processes and modelling of stock returns. Research Report no. 300, Department of Theoretical Statistics, University of Aarhus, Denmark.


   Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. Z. Wahrscheinlichkeitstheorie verwandte Gebiete, 38, 309–312.


   Value at Risk using stochastic volatility models. Master thesis, Department of Mathematical Statistics, Chalmers University of Technology and Gothenburg University.


*Measuring Market Risk.* John Wiley & Sons, LTD

*Hyperbolic distributions in finance.* Bernoulli, 1, 281–299.


*Value at Risk: recent advances.* University of California at Santa Barbara, Economics Working Paper Series, Department of Economics, UC Santa Barbara.


*Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Cz II statystyka matematyczna.* Warsaw: PWN, Poland.

*Koncepcja wartości zagrożonej VaR (Value at Risk).* StatSoft Polska  


Inverse Gaussian Distribution. Department of Economics, The City University of New York   
Available at: http://www.maxmatsuda.com/Papers/Intro/IGDistribution.pdf


Appendix

1.1 The EM algorithm implementation

```r
#BEGINING OF ME.R

# authors: Magdalena Kucharska and Jolanta Pielaszkiewicz
# Implemented the EM algorithm for estimation of NIG parameters

# Download real data
x <- read.table("D:/MGR/Documents/MASTER THESIS/DANE_Z_GIELDY/data3.csv", sep=";")
# mean of log-returns
X <- mean(x[,7])
# standard deviation of log-returns
s <- sd(x[,7])
# vector of log-returns
y <- c(x[,7])

# mi function
mi <- function(k, h, G) {
    z <- 0;
    for (i in 1:length(h)) z <- z + (h[i] - G)^k;
    return(z/length(h))
}

gamma1 <- mi(3, y, X)/(mi(2, y, X)^(3/2))
gamma2 <- mi(4, y, X)/(mi(2, y, X)^2) - 3
```

67
if (3*gamma2-5*gamma1^2>0)
{
  g<-3/(s*sqrt(3*gamma2-5*gamma1^2))
  b<-(gamma1*s*g^2)/3
  d<-(s^2*g^3)/(b^2+g^2)
  m<-X-b*d/g
  a<-sqrt(b^2+g^2)
}

## ERROR
tot=10^(-7)

## E STEP

# fi function
fi<- (function(m,d,z) {
  1+((z-m)/d)^2
})

# M function
M<-(function(a,b,d,m,y) {
  si<-0;
  for(i in 1:length(y))
    si<-si+d*sqrt(fi(m,d,y[i]))*besselK(d*a*sqrt(fi(m,d,y[i])),0)/(a*besselK(d*a*sqrt(fi(m,d,y[i])),1));
  return(si/length(y))
})

# wi function
wi<- (function(i,a,b,d,m,y) {
  wi<-a*besselK(d*a*sqrt(fi(m,d,y[i])),2)/(d*sqrt(fi(m,d,y[i])))*besselK(d*a*sqrt(fi(m,d,y[i])),1));
  return(wi)
})

# wi_sum function
wi_sum <- (function(i,a,b,d,m,y) {
  wi_sum<-0;
  for(i in 1:length(y)) wi_sum<-wi_sum+wi(i,a,b,d,m,y);
  return(wi_sum)
})
# wixi_sum function
wixi_sum <- function(i, a, b, d, m, y) {
  wixi_sum <- 0;
  for(i in 1:length(y)) wixi_sum <- wixi_sum + wi(i, a, b, d, m, y) * y[i];
  return(wixi_sum)
}

# Lambda function
Lambda <- function(i, a, b, d, m, y) {
  L <- 0;
  for(i in 1:length(y)) L <- L + wi(i, a, b, d, m, y) - (1/M(a, b, d, m, y));
  return(length(y) * L^(-1))
}

## M STEP
error <- tot + 1
j = 1
while(error > tot & j < 100000) {
  j <- j + 1
  print(j)
  d_spr <- d
  b_spr <- b
  a_spr <- a
  m_spr <- m
  d <- sqrt(Lambda(i, a, b, d, m, y))
  g <- d/M(a, b, d, m, y)
  b <- (wixi_sum(i, a, b, d, m, y) - X * wi_sum(i, a, b, d, m, y)) / (length(y) - M(a, b, d, m, y) * wi_sum(i, a, b, d, m, y))
  m <- X - b * M(a, b, d, m, y)
  a <- sqrt(g^2 + b^2)
  error <- max((d - d_spr)/d, (m - m_spr)/m, (b - b_spr)/b, (a - a_spr)/a)

  # output of parameters alpha, beta, delta, mu
  print(c(a, b, d, m))
}

###########################################################################END
1.2 The Lindberg method implementation

BEGINNING OF Lindberg.R

# authors: Magdalena Kucharska and Jolanta Pielaszkiewicz
# Implemented Lindberg method for estimation of NIG parameters

# Loading of libraries
library(HyperbolicDist)
library(fBasics)
library(ig)

# Reading our data
x <- read.table("D:/MGR/magisterka/DANE_Z_GIELDY/data3.csv", sep=",", skip=1)

# vector of log-returns
return <- c(x[,6])

# volatility
volatility <- c(x[,4])

# number of trades
nr_trades <- c(x[,17]/max(x[,17]))

# volume
volume <- c(x[,14]/max(x[,14]))

# amount
amount <- c(x[,13]/max(x[,13]))

# length of our sample
n <- length(return)

# we choose z as number of trades, volume or amount
z <- nr_trades

### mu, beta and theta by MAXIMUM LOG-LIKEHOOD FUNCTION of returns
mu <- (sum(return/z)-n*(sum(return)/sum(z)))/(sum(1/z)-n^2/sum(z))
print(mu)
th <- sum(return^2/z)-2*mu*sum(return/z)+mu^2*sum(1/z)-((mean(return)-mu)/70)
\begin{verbatim}
mean(z))^2*sum(z)
th=th/n
print(th)
beta<-(n^2*(sum(return)/sum(z))-n*sum(return/z))/(sum(z)*
(sum(1/z)-n^2/sum(z))+(sum(return)))/(sum(z))
beta<-beta/th
print(beta)

# normalized returns
xi=(return-mu-beta*th*z)/sqrt(th*z)

# tests of normality
shapiro.test(xi)
ks.test(xi,"pnorm",0,1)

# delta and gamma by estimation of IG distributed volatility=theta*z
w<-as.numeric(mleig(th*z))
delta<-sqrt(w[2])
gamma<-sqrt(w[2])/w[1]
alpha<-sqrt(gamma^2+beta^2)

# output of obtained parameters
print("NIG parameters")
print(alpha)
print(beta)
print(delta)
print(mu)
print(gamma)

1.3 The VaR and ETL implementation

# authors: Magdalena Kucharska and Jolanta Pielaszkiewicz
# Implemented Value-at-Risk and Expected Tail Loss
\end{verbatim}
# Loading of libraries
library(ghyp)

#reading our data
x<-read.table("D:/MGR/magisterka/DANE_Z_GIELDY/data3.csv",sep=";",skip=1)

#vector of log-returns
return<-c(x[,6])

#vector of number of trades
nr_trades<-c(x[,17]/max(x[,17]))

# theta for number of trades
t1=0.009373112

#vector of volume
volume<-c(x[,14]/max(x[,14]))

# theta for volume
t2=0.02098137

# we choose case number of trades or volume
volatility<-t1*nr_trades

# # # EXCESSIVE LOSSES SEQUENCE
losses=c(length(return), NA)
for (i in 1:length(return))
{
  if (return[i]<0)
    losses[i]=-return[i]
  if (return[i]>=0)
    losses[i]=0
}

iter1=0
iter5=0
iter_etl_1=0
iter_etl_5=0
temp<-c(length(return),NA)*NA

NIGfit<-fit.NIGuv(return/volatility)
var1<-c(length(return),NA)
var5<-c(length(return),NA)
etl1<-c(length(return),NA)
etl5<-c(length(return),NA)

for(i in 1:length(return))
{
  fitgh<-transform(NIGfit,0,volatility[i])
  lambda1=qghyp(0.01,fitgh)
  var1[i]=-lambda1
  lambda5=qghyp(0.05,fitgh)
  var5[i]=-lambda5
  if(var1[i]<losses[i])iter1<-iter1+1
  if(var5[i]<losses[i])iter5<-iter5+1
}

var1m<-var1
for(i in 1:length(return))
{
  for(k in 1:length(return))
  {
    if(return[k]<var1m[i])temp[k]<-return[k]
  }
  etl1[i]<-(-mean(temp, na.rm=TRUE))
temp<-c(length(return),NA)*NA
}

var5m<-var5
temp<-c(length(return),NA)*NA
for(i in 1:length(return))
{
  for(k in 1:length(return))
  {
    if(return[k]<var5m[i])temp[k]<-return[k]
  }
  etl5[i]<-(-mean(temp, na.rm=TRUE))
temp<-c(length(return),NA)*NA
}

for(i in 1:length(return))
if(etl1[i]<losses[i]) iter_etl_1<-iter_etl_1+1
if(etl5[i]<losses[i]) iter_etl_5<-iter_etl_5+1
}

# Output of number of crosses
print("Number of crosses for VaR at 0.01 confidence level")
print(iter1)
print("Number of crosses for VaR at 0.05 confidence level")
print(iter5)
print("Number of crosses for ETL at 0.01 confidence level")
print(iter_etl_1)
print("Number of crosses for ETL at 0.05 confidence level")
print(iter_etl_5)

# Plot of results (VaR and ETL together)
par(mfrow=c(2,1))
plot(losses, type="l", main="Confidence level 0.01")
lines(var1,t="l",col="brown")
lines(etl1,t="l",col="orange")
plot(losses, type="l", main="Confidence level 0.05")
lines(var5,t="l",col="brown")
lines(etl5,t="l",col="orange")