Evaluation of the South African equity markets in a value-at-risk framework

by

Lesedi Mabitsela
(student no 20108542)

Dissertation submitted in partial fulfilment of the requirements for the degree of

Magister Scientiae

in **Financial Engineering**

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

May 2015

Declaration

I, the undersigned declare that the dissertation, which I hereby submit for the degree *Magister Scientiae in Financial Engineering* at the University of Pretoria, is my own independent work and has not previously been submitted by me or any other person for any degree at this or any other university.

....

Lesedi Mabitsela

May 2015

Abstract

The statistical distribution of financial returns plays a key role in evaluating Value-at-Risk using parametric methods. Traditionally, when evaluating parametric Value-at-Risk, the statistical distribution of the financial returns is assumed to be normally distributed. However, though simple to implement, the Normal distribution underestimates the kurtosis and skewness of the observed financial returns. This dissertation focuses on the evaluation of the South African equity markets in a Value-at-Risk framework. Value-at-Risk is estimated on five equity stocks listed on the Johannesburg Stock Exchange, including the FTSE/JSE TOP40 index and the S&P 500 index. The statistical distribution of the financial returns is modelled using the Normal Inverse Gaussian and is compared to the financial returns modelled using the Normal, Skew t-distribution and Student t-distribution. We then estimate Value-at-Risk under the assumption that financial returns follow the Normal Inverse Gaussian, Normal, Skew t-distribution, Student t-distribution and Extreme Value Theory and backtesting was performed under each distribution assumption. The results of these distributions are compared and discussed.

Acknowledgement

I would like to express my very great appreciation to Dr R Kufakunesu and Prof E Mare my research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work. I would also like to thank them for their valuable and constructive recommendations on this project.

Lesedi Mabitsela

May 2015

Contents

List of Figures

List of Tables

List of Symbols

α Determines the tail heaviness of the NIG distribution.

β Determines the asymmetry of the NIG distribution.

δ Scale parameter of the NIG.

η Scale parameter of the Generalised Pareto distribution.

Γ Gamma function.

$\mathbb{C}$ is the set of complex numbers.

$\mathbb{N}$ is the set of natural numbers.

$\mathbb{P}$ The real probability measure.

μ Location parameter of the peak of the NIG.

Ω Space of all the asset paths.

ζ Shape parameter of the Generalised Pareto distribution.

$arg(z)$ gives the argument of the complex number z.

k Degree of freedom for the t-distribution.

K_λ Modified Bessel function with index λ.

S_t Stock price at time t.

X_t Random variable at time t.

Glossary

Value-at-Risk (VaR) is defined as the maximum expected loss over a predefined horizon period and at a given confidence level.

Expected Shortfall (ES) is the expected value of losses greater than VaR.

Market risk is the risk of losses in investments positions arising from movements in market prices (market prices include stock prices, interest rates, commodities, bonds and currencies).

Market risk model is a technique using statistics to determine the potential losses in investment positions due to market risk. These techniques include for example volatility, Greeks, Value-at-Risk and Expected Shortfall.

Backtesting is a tool used to validate the VaR model, by periodically comparing the estimated VaR value to the observed profit and loss of an investment.

Logarithmic returns is the logarithmic ratio of the opening price to the closing price on an investment. In this dissertation the logarithmic returns are referred to as financial returns.

Statistical distribution is a function that assigns a probability to the random variables (financial return).

Normal distribution is a statistical distribution that depends on two parameters - the mean and the standard deviation. The mean determines the location of the center of the graph of the distribution and the standard deviation determines the height of the graph.

Normal Inverse Gaussian (NIG) distribution is a statistical distribution with four parameters - α determines the heaviness of the tails, β determines the skewness of the distribution and the mean and standard deviation are the same as in the Normal distribution.

Kurtosis measures the "peakedness", and as a result of heaviness of the tails of the probability distribution of random variables. The Normal distribution has kurtosis equal to three, therefore **excess kurtosis** equals to kurtosis minus three.

Skewness denotes to whether a probability distribution of a random variable is symmetrical. An asymmetrical probability distribution is referred to as skewed distribution and will either be positively (financial returns more likely to be negative) or negatively skewed (financial returns more likely to be positive).

Leptokurtic refers to a statistical distribution with heavier or fatter tails than the Normal distribution or distribution with positive excess kurtosis.

Chapter 1

Introduction

1.1 Overview

Value-at-Risk (VaR)[1] is defined as the worst expected loss over a given period at a specified confidence level [Ris96]. Jorion [Jor01] describe VaR as the quantile[2] of the projected distribution of losses and gains of an investment over a target horizon. VaR answers the question, "How much can I lose with $q\%$ probability over a certain holding period?" [Ris96]. The risk metric VaR, has become a widely used risk measure by financial institutions and regulatory authorities[3], as it attempts to provide a single number that summarizes the overall market risk in individual stocks and for portfolios [Hul10].

VaR is a tool used to measure market risk, where market risk is the potential for change in the value of an investment due to change in market risk factors [Ris96]. Market risk factors are interest rates, commodity prices, foreign exchange rates and stock and bond prices [Ris96]. Historically, market risk was measured by the standard deviation of unexpected outcomes or by simple indicators of the notional-amount of the individual stock [MEF05].

[1] A formal definition of Value-at-Risk is given in Chapter 4 of the dissertation.

[2] Also referred to as percentile.

[3] The Basel Committee imposing minimum capital requirements for market risk in the "1996 Amendment" [oBS96a].

McNeil *et al.*, in [MEF05] provides the pros and cons of each traditional measures of market risk.

When evaluating VaR for financial assets the distribution of the returns of the underlying asset play an important role. Methodology of estimating VaR can be classified into two groups, i.e. the *parametric* VaR and *non-parametric* VaR. The classification of VaR methodology is based on how the financial return distribution is modelled. Parametric VaR assumes that financial returns are modelled using a statistical distribution (e.g. Normal and Student t distribution). Whereas non-parametric VaR assumes that financial returns are modelled using the empirical distribution. The statistical distribution that is commonly assumed in parametric VaR is the Normal distribution, which is easy to implement as it depends on two parameters, i.e. the mean and standard deviation of historical returns.

However, a number of studies have shown that daily financial returns are non-normal, they display a leptokurtic and skewed distribution as noted by Mandelbrot [Man63] and Fama [Fam65]. A leptokurtic distribution has a higher peak and heavier tails than the Normal distribution [Ale08]. In other words, the frequency of financial returns near the mean will be higher and extreme movements are more likely than the Normal distribution would predict. For example, if we consider South African FTSE/JSE TOP40 Index[4] the largest decrease was roughly 14%, which occurred in 1997. The 14% decrease deviates by ten standard deviations from the mean and by modelling financial returns with the Normal distribution this decrease is practically impossible.

The quality of VaR is dependent on how well the statistical distribution

[4]FTSE/JSE TOP40 Index constitutes of the largest 40 companies (listed on the Johannesburg Stock Exchange (JSE)) ranked by full market value in the FTSE/JSE All-Share Index.

captures the leptokurtic behaviour of the financial returns [Ryd00]. Shortcoming of statistical distribution can result in incorrect estimation of risk and lead to serious mismanagement of risk, for example insufficient capital invested to limit the probability of extreme losses. Hence, finding a statistical distribution that represents the leptokurtic behaviour of financial returns in VaR estimation remains an important research topic.

The introduction of VaR as the market risk measure has seen a number of empirical studies being done to find alternative distributions to the Normal distribution. These studies include application of the Student t-distribution in VaR estimation for returns on US equities and bonds by Huisman, Koedijk, and Pownall [HKP98]. Application of the t-distribution in VaR estimation within the South African equity market was done by Milwidsky and Mare [MM10]. The t-distribution is also used by McNeil and Frey in [MF00] and Platen and Rendek [PR08]. Although the t-distribution addresses the issue of heavy tails, it fails to address the skewness present in financial returns because it is symmetrical about zero.

The lack of skewness in the t-distribution was first addressed by Hansen in [Han94], when he first proposed a skew extension to the t-distribution for modelling financial returns. Since then, several authors have studied the application of the Skew Student t-distribution (Skew t) to modelling financial returns, see for example, Azzalini and Capitanio [AC03], Aas and Haff [AH06], Jones and Faddy [JF03]. The other proposed distribution is the Extreme Value Theory, which only models the behaviour of losses and not the entire returns distribution. For application of the Extreme Value Theory refer to: Longin [Lon05], Danielsson and De Vries [DdV00], McNeil and Frey [MF00], Embrechts, Klüppelberg and Mikosch [EKM97], Gençay, Selçuk, and Ulugülyağci[GSU03] and Wentzel and Mare [WM07]. Other methods used include, modelling the returns and volatility process by the ARMA (1,1)-GARCH (1,1) time series and then fitting the residuals with the selected distribution, which this article will not be pursing, see for example Bhattachariya and Madhav [BM12], Schaumburg [Sch12] and Kuester et al [KMP06]. An outline of some of the statistical distributions fitted to financial returns in literature are also discussed in Chapter 5.

The focal point of this dissertation is to model financial returns of listed equity stocks using the Normal Inverse Gaussian (NIG) distribution and further estimate Value-at-Risk with the fitted NIG distribution. The NIG distribution is able to capture the skewness and kurtosis present in the financial returns. The tails of the NIG distribution are described as “semi-heavy”. Dependent on four parameters that affect the shape of the density function, with the NIG distribution one is able to create different shapes of the density function by adjusting the parameters, making the NIG distribution very flexible. Most of the authors have report an excellent fit to the financial returns. The application and reviews of the NIG distribution is given by Lillestøl [Lil00], Rydberg [Ryd00], Prause [Pra99], Barndorff-Nielsen [BN95], Venter, and de Jongh [VdJ01] and Bølviken and Benth [BB00]. The NIG distribution has an important property of being closed under *convolution* i.e. the sum of independent NIG random variables is also NIG distributed. This property is useful when working with a portfolio of shares and for time scaling of VaR.

1.2 Objectives

In this dissertation, we apply the Normal Inverse Gaussian distribution to the evaluation of the South African equity markets in a Value-at-Risk framework. This study is based on the work of Bølviken and Benth 2000 [BB00], who investigated the NIG distribution as a tool to evaluate the uncertainty in future prices of the shares listed on the Norwegian Stock Exchange in Oslo.

The initial objective of the dissertation is to demonstrate to the reader that the financial returns of selected shares are not normally distributed and present the concept of Value-at-Risk and NIG distribution. The final goal of the dissertation is to implement the Normal Inverse Gaussian distribution and compare the Value-at-Risk numbers to those under the Normal, t-distribution, Skew t and the Extreme Value Theory.

This study is restricted to market risk associated with price changes of individual equity stocks listed on the Johannesburg Stock Exchange in South Africa over 1-day time horizon. Therefore, VaR is estimated in linear positions in the underlying equity stocks.

The dissertation address the following:

(i) Analyse a defined list of shares and justify the selection criteria.

(ii) Demonstrate that the daily returns of the chosen shares are not normally distributed.

(iii) Discuss methods of estimating and validating Value-at-Risk.

(iv) Review the NIG distribution and alternative distributions.

(v) Examine how the distribution of financial returns fits the NIG distribution when compared to the Skew t, Normal and t-distribution.

(vi) Compare the daily VaR estimates across the daily returns for the Skew t, Normal, t-distribution, NIG distributed and EVT.

We expect the NIG and Skew t to adequately fit the financial returns both in the tails and in the center than the Normal and the Student's t distributions. The limitation of the Normal distribution when applied to financial data is the skewness and the kurtosis present in financial data, while the t-distribution captures the kurtosis of the financial returns it does not capture the skewness since it is symmetrical. The NIG captures both the skewness and the kurtosis of the financial returns. The tail of the NIG is described as "semi-heavy", therefore, we expect that the NIG to model skewness well in the case of not too heavy tail distribution. While the Skew t has heavy tails, we expect to fit data with heavy tail well and not handle extensive skewness [AH06].

1.3 Structure of the dissertation

In Chapter 2, we give a summary of the listed companies and indices that we will be using to evaluate Value-at-Risk and examine on the daily share data of these chosen companies. This will involve investigating the plots of the closing share prices and the log returns. We further demonstrate that the daily log returns are not normally distributed using the Q-Q plot, which compares the sample data to data that follows the Normal distribution. To validate the Q-Q plot findings we will use formal statistical tests such as the Shapiro-Wilk test, Jarque-Bera test and the Anderson-Darling test.

In Chapter 3, we define the concept of Value-at-Risk as a model to evaluate market risk. The advantages and disadvantages of Value-at-Risk are discussed and we further look at the three approaches for evaluating Value-at-Risk. Lastly, we give an overview of Expected Shortfall and the method of backtesting.

Chapter 4 considers the NIG distribution as the alternative distribution to modelling log returns as it possess features of both skewness and heavy tails. We start by defining the NIG process according to the theory of Lévy processes. We highlight how the different parameters of the NIG distribution contribute to the shape of the graph of the density function. This is done by plotting a graph of the density function for different values for each parameter of the distribution. We examine the properties of the NIG distribution, which can be useful for modelling of financial returns. If individual returns are NIG distributed then the sum of the returns will be NIG distributed. Finally we look at how to fit the NIG distribution by evaluating the maximum likelihood function.

Chapter 5 provides an overview of alternative distributions that have been studied in literature to model financial returns. These include the Stable Paretian, Skew t, t-distribution and Hyperbolic distribution. All these alternative distributions model the financial returns better than the Normal distribution. However, one of the shortfalls of the Stable Paretian distribution is that it does not have an explicit density function making it difficult to estimate parameters. The t-distribution does not address the skewness present in many financial returns and the Hyperbolic distribution does not possess the property of being closed under convolution. We also focus on the tails of the distribution, more importantly the tails of the losses through the study of Extreme Value Theory. This will be implemented using the Peak-over-Threshold method, which is discussed.

In Chapter 6, we fit the NIG, Skew t, t-distribution and Normal distribution to the chosen shares and indices and apply the Peak-over-Threshold method to the tails of the losses. We conclude the chapter by evaluating Value-at-Risk for the chosen shares and indices assuming that the underlying distribution is NIG, Skew t, Normal, t-distribution and Extreme value.

Finally, in Chapter 7, we discuss our findings and draw some conclusions.

Chapter 2

Empirical Study

In this chapter, we consider the listed equity data and investigate the normality assumption. This is done using the Q-Q plot and formal framework of hypothesis testing such as Shapiro-Wilk test, Jarque-Bera test and Anderson-Darling test. We first give a brief insight into the data using plots of the daily closing prices and daily log returns.

2.1 Description of shares

The empirical study is done using five South African stocks, FTSE/JSE TOP40 (J200) index and the S&P 500 index. The five shares (Standard Bank (SBK), African Bank (ABL), Merafe Resource (MRF), Grindrod (GND) and Anglo American (AGL)) are listed on Johannesburg Stock Exchange. Maximum available daily closing prices for each share were obtained resulting in varying periods ending 31 July 2014 [1]. The S&P 500 data is from 2 January 1991 to 31 July 2014 totalling 5831 daily returns.

These shares were randomly selected and they reflect different sub-sectors of the JSE main board.

[1] All the closing prices were obtained from I-Net Bridge.

(i) Merafe Resources is listed on the JSE under the General Mining sector. Merafe mines chrome, which they use to produce ferrochrome.

(ii) Standard Bank is listed under the Banking sector. The company provides services in personal, corporate, merchant and commercial banking, mutual fund and property fund management among other services.

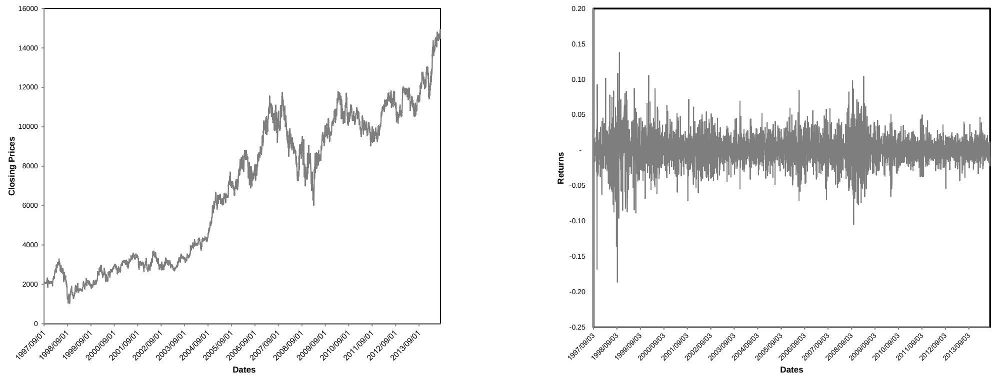
(iii) Grindrod is listed under Marine Transportation. Grindrod offers freighting, trading, shipping and financial services.

(iv) African Bank is listed under Consumer Finance. The bank provides unsecured credit, retail and financial services.

(v) Anglo American is listed under the General Mining sector and they mine platinum, diamonds, iron ore and thermal coal.

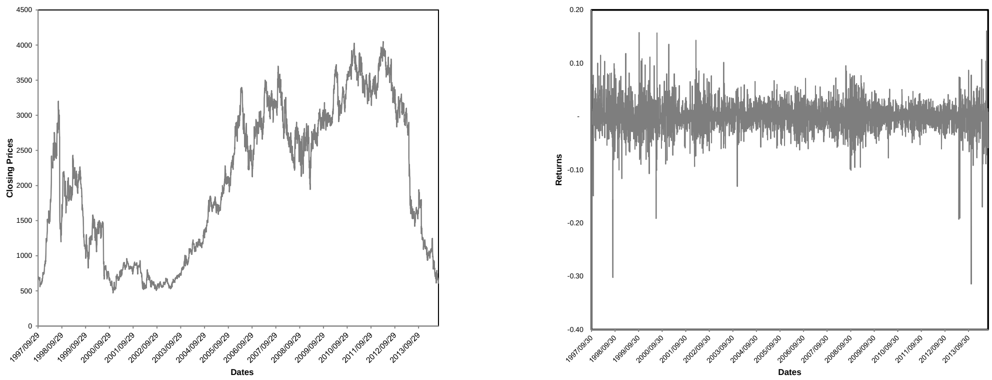
We included the FTSE/JSE TOP40 index, which consists of the 40 largest companies listed on the JSE in terms of market capitalisation [2]. The Index gives reasonable reflection of the entire South African stock market as these 40 top companies represent over 80% of the total market capitalisation of all the companies listed on the JSE [Cap14]. Standard Bank and Anglo American have large market capitalisation and we expect them to mimic the FTSE/JSE TOP40, while Merafe and African Bank are small and therefore would have an element of jump risk.

The daily log returns defined by $r_t = \ln(S_t/S_{t-1})$ were obtained totalling a series of $N - 1$ observations for South African listed shares, where N is the total number of closing prices observed for each share and $(S_t)_{t\geq 0}$ is the closing price/index level at day t. We assume the share prices and indices level to follow a random walk and exclude dividends. We start off with a statistical summary of the data for each share. Figure 2.1 to Figure 2.7 displays the daily closing price levels (a) and daily log returns (b) for each share and the indices over varying periods ending 31 July 2014. Table 2.1 shows the statistical summary of the data for each share, computed using



(a) (b)

Figure 2.1: Standard Bank closing prices (a) and log returns (b) for the period 1 September 1997 to 31 July 2014.



(a) (b)

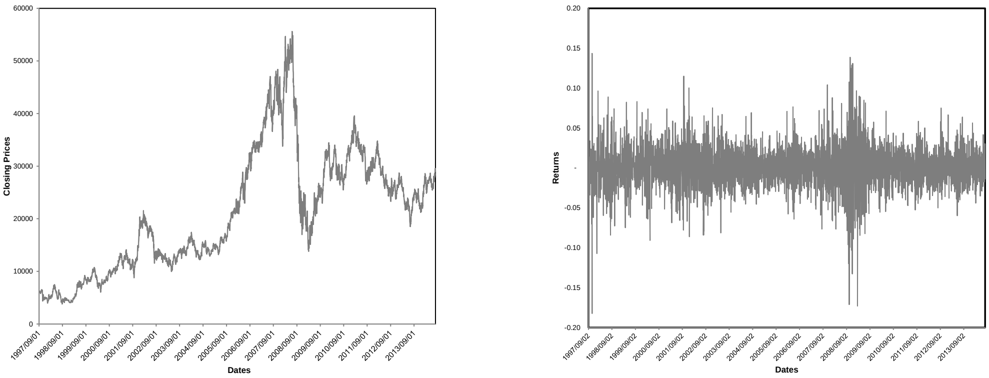
Figure 2.2: African Bank closing prices (a) and log returns (b) for the period 29 September 1997 to 31 July 2014.

the daily log returns. In Appendix A, the sample period is divided into three sub-samples and the analysis is performed using the entire sample period and the three sub-samples which are:

(i) Pre-crisis (from January 1991 - December 2007),

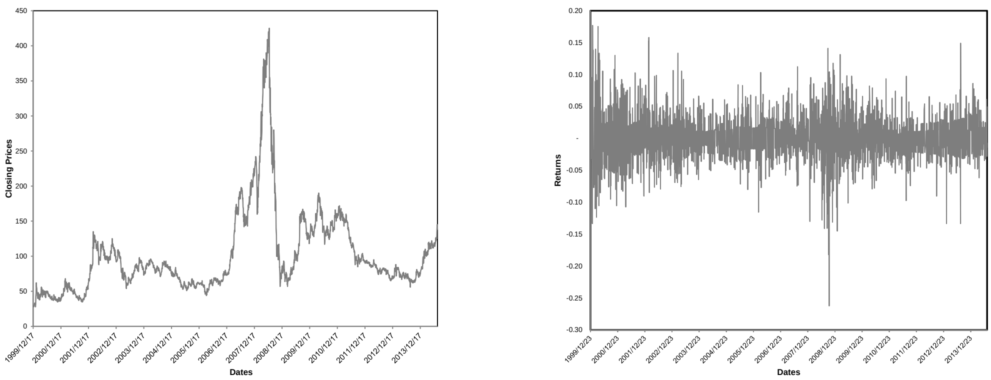
(ii) Crisis period (from January 2008 - December 2009),

(iii) Post-crisis (from January 2010 - July 2014).



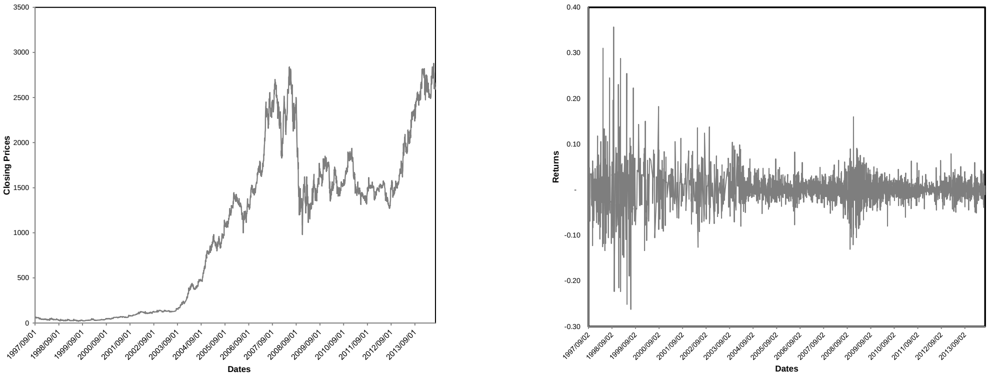
(a) (b)

Figure 2.3: Anglo American closing price (a) and log returns (b) for the period 1 September 1999 to 31 July 2014.



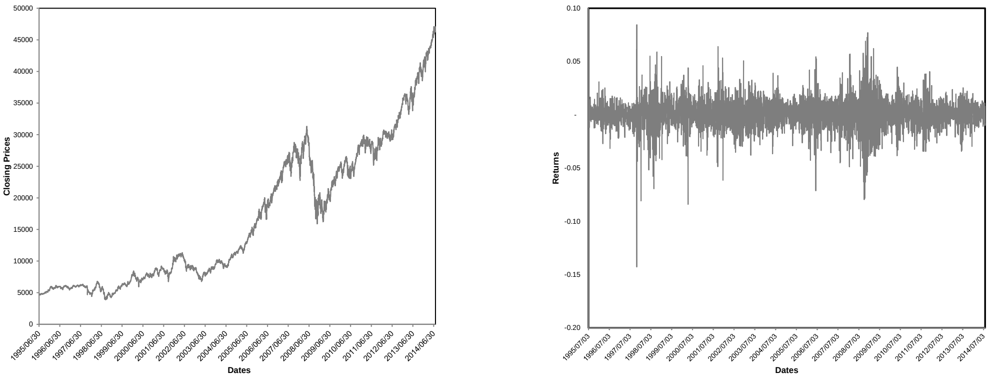
(a) (b)

Figure 2.4: Merafe Resources closing price (a) and log returns (b) for the period 17 December 1999 to 1 July 2013.



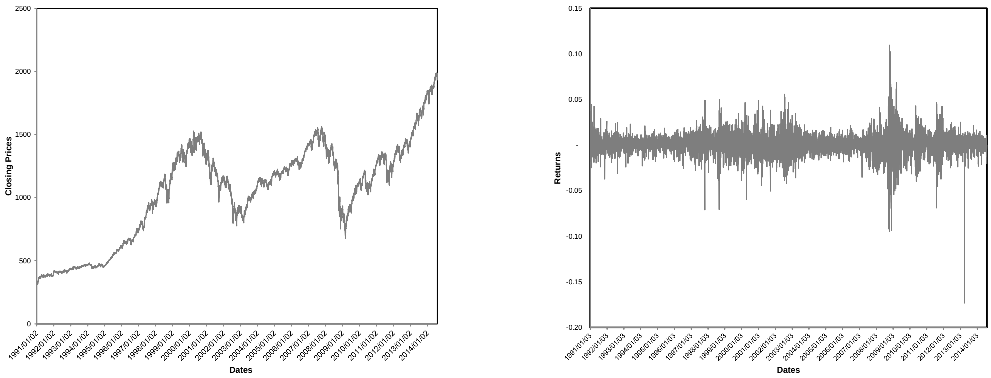
(a) (b)

Figure 2.5: Grindrod Limited closing prices (a) and log returns (b) for the period 1 September 1997 to 31 July 2014.



(a) (b)

Figure 2.6: FTSE/JSE Top40 Index closing levels (a) and log returns (b) for the period 30 June 1995 to 31 July 2014.



(a) (b)

Figure 2.7: S&P 500 closing prices (a) and log returns (b) for the period 2 January 1991 to 31 July 2014.

Statistical data of the empirical distribution					
	Mean	Variance	Skewness	Excess Kurtosis	No. OBS[3]
Standard Bank	0.00048	0.00047	−0.20369	4.86980	4090
African Bank	−0.00003	0.00082	−0.78087	10.8220	3944
Anglo American	0.00038	0.00064	−0.07927	3.93211	4162
FTSE/JSE TOP40	0.00048	0.00019	−0.40611	6.27868	4770
S&P 500	0.00029	0.00014	−0.73247	15.95379	6170
Merafe Resource	0.00054	0.00135	0.06429	2.52105	2806
Grindrod	0.00125	0.00118	0.69133	16.20972	3038

Table 2.1: Statistical data for each stock and the indices.

Based on the statistical results[4] of the empirical data in Table 2.1 the mean of African Bank, Anglo American and Merafe Resources is relatively small compared to the variance it is almost insignificant. The excess kurtosis for each stock and the indices is greater than zero. This indicates a higher peak and heavier tails meaning extreme loss and profit are more likely to occur than what the Normal distributed would predict.

The Indices, Standard Bank, African Bank and Anglo American have

	Percentile data	
	1st Percentile	5th Percentile
Standard Bank	−5.81%	−3.34%
African Bank	−7.35%	−4.32%
Merafe Resource	−9.02%	−5.41%
Grindrod	−9.53%	−4.65%
Anglo American	−6.60%	−3.87%
FTSE/JSE TOP40	−3.79%	−2.08%
S&P 500	−3.15%	−1.76%
Standard Normal	**-2.33%**	**-1.65%**

Table 2.2: Percentile data for each stock and the indices.

negative skewness that is the left tail is longer, indicating that losses occur more frequently than profits over the entire sample period. While Merafe has positive skewness implying more profits than losses were realised over the period. In general each stock and the indices display fatter tails and skewness in comparison to the Normal distribution as noted in literature [Fam65].

Further evidence from the Table 2.2 comes to support the claim of the heaviness of the tails of each stock and the indices, where the 1st and 5th percentile of each stock are compared to those of the Standard Normal distribution. These types of distributions with fatter tail behaviour are known as leptokurtic.

[2]Market capitalisation of a company is the number of outstanding shares multiplied by the current share price.

[4]The statistical data was computed using Microsoft Excel.

2.2 Normality test

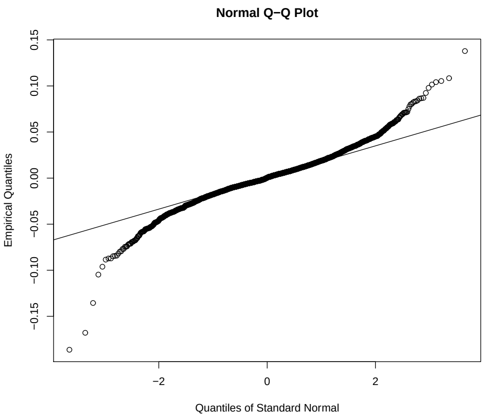
As already highlighted by the statistical data in Table 2.1, the process of the log returns deviates from the Normal distribution since the values of the skewness and kurtosis are not close to zero and three respectively. In this section, we further demonstrate how the process of the daily log returns of each stock and the indices deviate from the Normal distribution using the Q-Q plot and formal framework of hypothesis testing such as Shapiro-Wilk test, Jarque-Bera test and Anderson-Darling test.

Q-Q plot

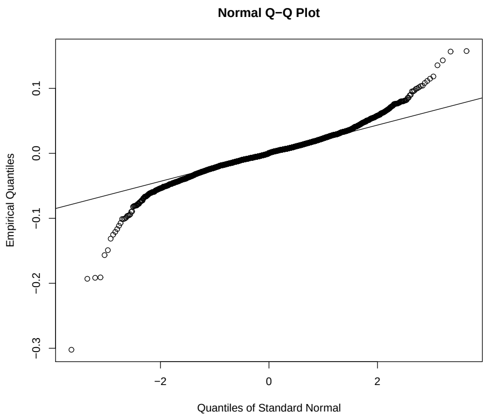
The Q-Q plot or quantile-quantile plot is a simple graphical method of testing the goodness of fit of observed returns to the Normal distribution. In Figure 2.8 the Q-Q plot demonstrates how the indices and each stock deviate from the straight line [5] at the tails of the distribution. Although Q-Q plots are easy to implement and can help reveal departure from normality, they are less formal and exclusive reliance on them can lead to erroneous conclusion [DS86]. Hence the use of formal numerical techniques is essential in order to avoid such errors. The formal numerical techniques quantify the information and evidence in the data or graphs and act as a verification of inferences suggested from these.

In the formal framework of hypothesis testing the null hypothesis H_0 is that the log returns follow the Normal distribution, while the alternative hypothesis H_1 is that the log returns are not normally distributed. There are a number of formal techniques applied to test the H_0 available in literature and statistical software, in this dissertation we will use the Shapiro-Wilk (SW) test, Jarque-Bera (JB) test and Anderson-Darling (A^2) test.

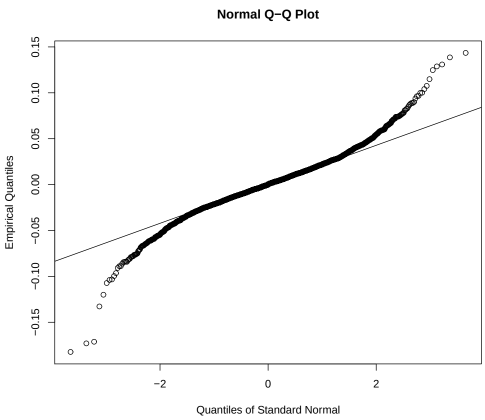
[5] The Q-Q plot we sketched using the R Statistical Program, which fits the theoretical line passing through the first and the third quartile of the empirical data.



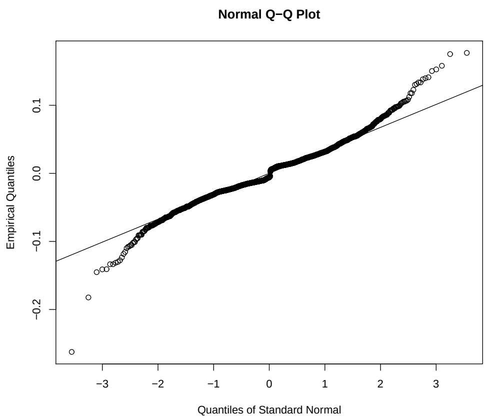
(a) Standard Bank



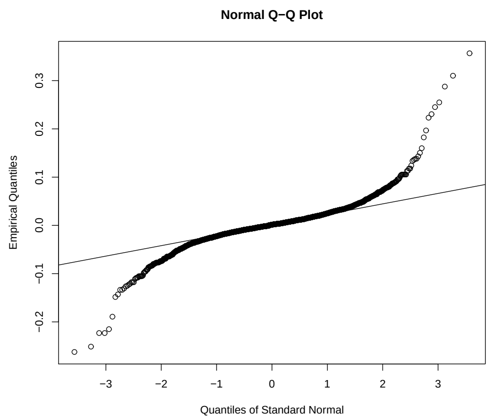
(b) African Bank



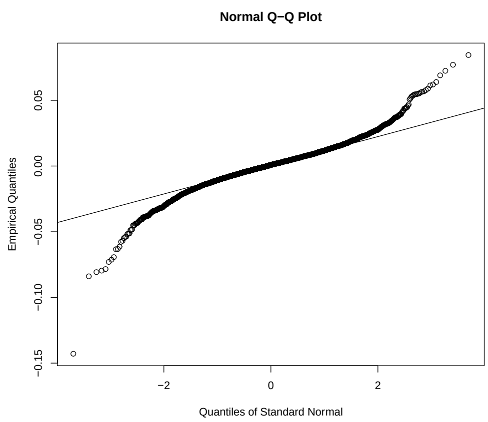
(c) Anglo American



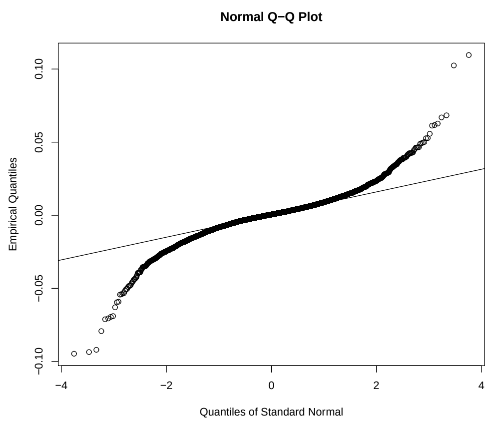
(d) Merafe Resources



(e) Grindrod Limited



(f) FTSE/JSE TOP40



(g) S&P 500

Figure 2.8: Q-Q plot of the daily log returns for each stock and the indices.

Shapiro-Wilk test

The Shapiro-Wilk test was introduced by Shapiro and Wilk in 1965 [SW65]. The Shapiro-Wilk test is considered a regression test. A test statistic is called a regression test if the ordered statistics $x_1 < x_2 < \cdots < x_n$, are plotted against the Standard Normal distribution and a straight line is then fitted to the points, the test is based on the statistics associated with the fitted line [DS86]. The null hypothesis is that the ordered sample, $x_1 < x_2 < \cdots < x_n$, is from a normally distributed population. The Shapiro-Wilk test statistic is defined as:

$$W = \frac{(\sum_{i=1}^{n} a_i x_i)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}, \tag{2.1}$$

where x_i is the i^{th} order statistic from the empirical sample, $\overline{x}$ is the sample mean,

$$\boldsymbol{a}_i = (a_1, \ldots, a_n) = \frac{\boldsymbol{m}^T \boldsymbol{V}^{-1}}{(\boldsymbol{m}^T \boldsymbol{V}^{-1} \boldsymbol{V}^{-1} \boldsymbol{m})^{1/2}}$$

and $\boldsymbol{m} = (m_1, \ldots, m_n)^T$ are the expected values of the standard normal ordered statistic and $\boldsymbol{V}$ is the covariance matrix. The values of a_i are appropriate constants for which Shapiro and Wilk [SW65] gave estimates for sample size less than equal to 50. In 1982 Royston [Roy82] presented an algorithm to broaden the sample size to 2000 and a method for obtaining the p-value of the test. However, Royston observed that Shapiro and Wilk approximation of the constant values of a_i were inadequate for sample size ≥ 50. Later, Royston [Roy95] gave an improvement to the estimates of a_i to extend to sample size of $3 \leq n \leq 5000$.

The Shapiro Wilk test statistic can be interpreted as an approximate measure of straightness of the Q-Q plot [Roy95]. That is, if a set of observations x_i come from the Normal distribution $N(\mu, \sigma^2)$, then under the null hypothesis the observations can be expressed as:

$$x_i = \mu + \sigma z_i. \qquad i = 1, \ldots, n$$

The slope of the regression line is an unbiased estimate to the standard deviation σ. Therefore, if the random observations x_i are not coming from a standard normally distributed sample, then the slope of the regression line will not be an estimate of σ. The numerator of W in Equation (2.1) is the estimate of σ which is the slope of the regression line and the denominator is the unbiased estimate of the standard deviation of the sample of random observations x_i given by $(n-1)s^2$. Hence if the x_i's are from the Normal distribution sample the value of W should be equal to one.

Jarque-Bera Test

The Jarque-Bera test is based on the third and fourth sample moments. The test matches the sample skewness and kurtosis to that of the Normal distribution. The test is defined as:

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right), \tag{2.2}$$

where S and K are sample skewness and kurtosis. If our data is from a normally distributed population then value of JB increases and the null hypothesis will be rejected. A sample from a Normal distribution has skewness of zero and kurtosis of three. Therefore by Equation (2.2) sample from a normally distributed data will have JB test close to zero. Furthermore, JB test asymptotically has chi-squared distribution with two degrees of freedom, therefore the null hypothesis is rejected if the calculated JB statistic is greater than the value of the critical value from the chi-square distribution with two degrees of freedom.

Anderson-Darling test

The Anderson-Darling test is regarded as the empirical distribution function (EDF) statistic. The EDF statistics are measures of the discrepancy between the EDF and a hypothesis cumulative distribution function. The test belongs to the *quadratic* class of EDF statistics [DS86], it is a refinement of the

Kolmogorov-Smirnov test and it gives more weight to the tails than the Kolmogorov-Smirnov test. The Anderson-Darling test is defined as:

$$A^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)\,(1 - F(x))} dF(x), \tag{2.3}$$

where the given a random sample of size n is $x_1, x_2, \ldots, x_n$ and $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$, are the ordered statistics and $F(x)$ is the hypothesized cumulative distribution of x. The empirical distribution function denoted by $F_n(x)$ is a step function, defined as:

$$F_n(x) = \frac{\text{number of observations} \leq x}{n}.$$

A suitable computing formula for the test is given as [DS86]:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \left[\ln F(x_i) + \ln(1 - F(x_{n+1-i})) \right]. \tag{2.4}$$

The test statistics A^2 is then compared to the critical value of the Normal distribution. If the value of A^2 is greater than the given critical value from the Normal distribution then the null hypothesis will be rejected, and this leads to the conclusion that log returns do not follow the Normal distribution.

The Kolmogorov-Smirnov test statistics is given by:

$$D_n = \sup_x |F_n(x) - F(x)| \tag{2.5}$$

where sup is the supremum [RW11]. The null hypothesis is rejected if D_n is greater than the critical value.

Results of the Normality test

For each test the null hypothesis is that the log returns are normally distributed. The test is considered to be statistically significant if the significant level of a given hypothesis test is less than the p-value. In general, a 0.05 or lower p-value is considered to be statistically significant, hence we choose a significant level to be 0.05 for our hypothesis test. The significant level value corresponds to the probability for which one chooses to reject or accept the null hypothesis. Therefore, if the p-value obtained is less than the significant level of 0.05, the null hypothesis will be rejected.

The results obtained for the three tests applied on the daily log returns are summarized in Table (2.3). These results were obtained using the statistical programme R package *'nortest'* and *'tseries'*. The p-value for the shares are less than 0.05 significant level in all the three tests, hence the H_0 is rejected and therefore we conclude that the log returns do not follow the Normal distribution.

The three test statistics uses different methods to test for normality, the Shapiro-Wilk tests the correlation in the Q-Q plot, Jarque-Bera test uses the third and fourth moments while the Anderson-Darling compares the empirical distribution to that of the Normal distribution. These three normality tests have indicated that the log returns of the selected equity data over varying periods as plotted in Figure 2.1 to 2.7 do not follow a Normal distribution.

	Shapiro Wilk test		Anderson-Darling test		Jarque-Bera test	
	W test	p-value	A^2 test	p-value	JB test	p-value [6]
Standard Bank	0.9567	$< 2.2e^{-16}$	29.0742	$< 2.2e^{-16}$	4057	$< 2.2e^{-16}$
African Bank	0.9252	$< 2.2e^{-16}$	37.7541	$< 2.2e^{-16}$	19750	$< 2.2e^{-16}$
Merafe Resource	0.9725	$< 2.2e^{-16}$	15.4049	$< 2.2e^{-16}$	844	$< 2.2e^{-16}$
Grindrod	0.8468	$< 2.2e^{-16}$	78.1954	$< 2.2e^{-16}$	33754	$< 2.2e^{-16}$
Anglo American	0.9638	$< 2.2e^{-16}$	22.2202	$< 2.2e^{-16}$	2677	$< 2.2e^{-16}$
FTSE/JSE Top 40	0.9454	$< 2.2e^{-16}$	41.0257	$< 2.2e^{-16}$	7946	$< 2.2e^{-16}$
S&P 500[7]	-	-	99.3793	$< 2.2e^{-16}$	65871	$< 2.2e^{-16}$

Table 2.3: Results of Normality tests from the Shapiro-Wilk, Anderson-Darling and the Jarque-Bera test statistics.

2.3 Chapter Summary

In this chapter, we introduced our empirical data and calculated the statistical data, which indicates that the empirical data exhibits skewness and heavier tails than that of the Normal distribution as commonly expected. We shall further examine the skewness and heavier tails of the empirical distribution by using the NIG distribution as it exhibits skewness and heavy tails.

We will also focus on modelling the tails of the empirical distribution. In this chapter we demonstrated using the Q-Q plot and formal numerical techniques that the empirical data is not normally distributed. The Q-Q plot shows that the data deviated at the tails, we will examine the tails in detail using the Extreme Value Theory.

Chapter 3

Value at Risk

In this chapter, the concepts of Value-at-Risk as a model to evaluate market risk is discussed. Value-at-Risk came about when several financial institutions started reporting internally their risk measurement and aggregate risk in the firm. It is generally reported that the J.P. Morgan CEO Dennis Weatherstone requested for a one-day summary report of the bank's overall market risk to be delivered ever afternoon at 4:15PM. Hence, VaR was introduced by J. P. Morgan in the first version of its RiskMetrics system and it became recognised by risk managers and regulators as an industry-wide standard [Ale08]. VaR has become the most used market risk measure.

Before defining Value-at-Risk we first start by defining market risk. The advantages and disadvantages of VaR as a market risk measure are also discussed in this chapter. We consider the different methods of evaluating VaR. Lastly, we make a comparison of VaR to Expected Shortfall.

3.1 Market risk

According to Jorion [Jor01], market risk is the risk due to movements in the level of market prices. Moreover, market risk can take two forms [Jor01]:

(i) **absolute risk** losses measured in rand terms or in the relevant currency.

(ii) **relative risk** losses measured relative to the benchmark index, which is also referred to as *tracking error*.

Jorion [Jor01] further classifies market risk into *directional risk* and *non-directional risk*.

(i) **Directional risk** involves exposure to the direction of movements in risk factors, such as equity and bond prices, interest rates, exchange rates and commodity prices.

(ii) **Non-directional risk** involves the remaining risk, consisting of exposures to hedged positions or to *volatility* and *basis risk*. *Basis risk* is created from unexpected movements in relative prices of assets in hedged position. For example, in a forward contract basis risk is the difference between the forward price and the spot price.

Crouhy *et al.* [CGM01] define market risk as the *risk that changes in financial market prices and rates will reduce the value of a security or a portfolio*. Furthermore, Crouhy, Galai and Mark in [CGM01] classify market risk into four major market risks:

(i) Interest rate risk,

(ii) Equity price risk,

(iii) Foreign exchange risk and

(iv) Commodity price risk.

The focus of this dissertation is on *equity price risk*, wherein Crouhy *et al.* [CGM01] divided it further into two components namely *General market risk* and *Specific risk*. *General market risk* refers to the sensitivity of an instrument or portfolio value to a change in the level of general stock market indices. *Specific* or *idiosyncratic* risk refers to firm specific risk, such as the quality of the firm's management, line of business or a breakdown in the firm's production process. Market risk is also referred to as **systematic risk** or **non-diversifiable risk**, it is the risk that is attributed to market wide risk sources [BKM09]. Although market risk can be hedged, it cannot be eliminated completely through diversification. Hence, market risk is therefore assumed by all investors.

Models for measuring market risk have evolved over the years, from the time where risk was measured using the face value or "notional" amount of the security, through to complex measures of price sensitivity for example duration and convexity for a bond and Greek measures for a derivative, to the latest model, Value-at-Risk [CGM01]. Many publications give a comprehensive presentation on traditional market risk measures e.g., Crouhy *et al.* [CGM01], Hull [Hul10], Dowd [Dow02] and Bessis [Bes10]. According to Crouhy, Galai and Mark [CGM01], VaR has proven to be a more powerful measure of the overall market risk of trading position over a short time horizon, such as 1-day or 10-day period and under "normal" market conditions. VaR strives to provide a single number that summarizes the overall market risk in individual stocks and for portfolios.

3.2 Definition of Value-at-Risk

The concept of VaR is defined here intuitively according to Crouhy, Galai and Mark [CGM01] as *the worst loss that might be expected from holding a security or portfolio over a given period of time (say a single day, or 10 days for the purpose of regulatory capital reporting), given a specified level of probability (known as the "confidence level").*

Crouhy *et al.* [CGM01] further explains that VaR does not provide the answer to the question: "How much can I lose on my portfolio over a given period of time?". The answer to this question is everything or almost the entire portfolio, should the entire market collapse all at once, in theory, the value of the portfolio may fall near to zero. Instead, what VaR offers is the answer to the question: "What is the maximum loss over a given time period such that there is a low probability, say 1% probability, that the actual loss over the given period will be larger?". Therefore, VaR offers a probability statement about the potential losses over a predetermined period of time.

McNeil *et al* in [MEF05] gives a mathematical definition of VaR as follows:

Definition 3.2.1. *Given some confidence level* $q \in (0, 1)$. *The VaR of the portfolio at the confidence level* q *is given by the smallest number* l *such that the probability that the loss* L *exceeds* l *is at most* $(1 - q)$. *This can be formulated as,*

$$\mathrm{VaR}_q = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - q\} = \inf\{l \in \mathbb{R} : F_L(l) \geq q\}, \tag{3.1}$$

where $F_L(l)$ *is the distribution function of losses and* L *is the loss of a portfolio.*

The main basic parameters used in evaluating VaR are:

(i) The confidence level (q);

(ii) The period of time over which VaR is measured also known as the risk or time horizon, denoted by h, and this is usually measured in trading days rather than calendar days.

The significance level is often set by the regulator body, for example in the banking industry they use the 1% significant level or 99% confidence level to assess their market capital requirement as encouraged by Basel II agreements [Ale08]. Basel II [oBS11] minimum requirement, is that VaR will be calculated at 99% percentile, one-tailed confidence level. Otherwise the significance or confidence level for VaR can be dependent on the risk appetite of an investor in the absence of a regulatory body. The higher the risk appetite of the investor the lower the confidence level and the lower the risk appetite the higher the confidence level [Ale08].

The risk horizon is the period over which the potential loss is measured. Under the Basel banking regulations [oBS11] VaR is assessed over a minimum of ten trading days. However, Banks may use VaR numbers calculated according to shorter risk horizon scaled up to ten days by the square root of time. Banks using this approach must periodically justify the reasonableness of its approach to the satisfaction of its supervisors [oBS11].

According to the Basel Committee [oBS11] each bank must meet on a daily bases a capital requirement expressed as the sum of:

- The maximum of the previous day's VaR estimate (VaR_{t-1}) and the average of daily VaR estimates of the previous sixty business days ($VaRavg$), multiplied by a scaling factor m_c.
- The maximum of the latest available stressed VaR estimate ($sVaR_{t-1}$) and an average of the stressed VaR estimates on each of the previous sixty business days ($sVaRavg$), multiplied by a scaling factor m_s.

Therefore, the capital requirement (c) is given by:

$$= \max\{VaR_{t-1} : m_c \times VaRavg\} + \max\{sVaR_{t-1} : m_s \times sVaRavg\}, \quad (3.2)$$

where the scaling factors m_c and m_s are based on the backtesting results and are subject to an absolute minimum of three.

Basically, VaR is a $(1-q)$ quantile of the distribution of the asset returns [MEF05]. McNeil [MEF05] further highlights that VaR is the maximum loss which is not exceeded with a given confidence level. In market risk management the values used for q are $q = 0.95$ or $q = 0.99$ with the time horizon usually being 1 or 10 days [MEF05]. VaR can be interpreted graphically as in Figure 6.8 under "normal" market conditions.

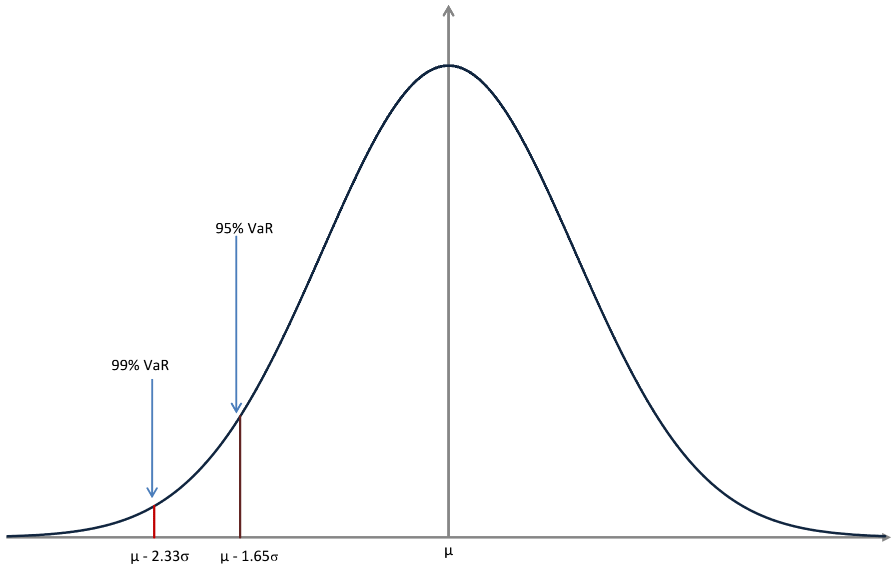


Figure 3.1: Value-at-Risk of a hypothetical return distribution (where μ is the mean of the returns and σ is the standard deviation of the returns when the returns are distributed according to a Normal distribution).

3.3 Standard Methods for evaluation of VaR

The first step of evaluating VaR is to determine the distribution of the portfolio future returns, at the chosen time horizon. The distribution of future returns is determined either by assuming the future returns have the same distribution as the historical returns or by assuming the future returns have some known distribution e.g. Normal, Student's t-distribution or NIG distribution (in the case of this dissertation). The second step is choosing the significant or confidence level. The three main methods to evaluating VaR are:

(i) Variance-Covariance Method.

(ii) Historical Simulation.

(iii) Monte Carlo Simulation.

The difference between the three methods is the way the distribution of future returns is constructed. However, the variance-covariance method is limited to linear securities in a portfolio, while the historical simulation method applies to portfolios that include options. The Monte Carlo method is the most flexible amongst the three. It may be used with a variety of return distributions and it is also applicable to portfolios with non-linear securities.

3.3.1 Variance-Covariance Method

The variance-covariance method has also been given many different names in publications for example *model-building approach, delta-normal, analytic VaR, and parametric VaR*. Under the variance-covariance method the returns of the assets are assumed to be normally distributed and their joint distribution is multivariate normal, and therefore the covariance matric of the asset returns is all that is required to capture the dependency between asset returns [Ale08]. Under these conditions, VaR is estimated by the following analytical formula:

$$\mathrm{VaR}_q = \mu + \sigma\,\Phi^{-1}(q),$$

where

(i) $\Phi^{-1}(q)$ is the inverse Normal distribution,

(ii) μ and σ are the estimated mean and standard deviation of the historical asset returns respectively.

The derivation of the analytical VaR formula can be found in Carol Alexander [Ale08] and also Simon Hubbert [Hub12]. The variance-covariance method is mostly suitable for portfolios with linear assets and not for portfolios that include non-linear assets such as derivatives. According to Hull

[Hul10], the variance-covariance method is mostly used for investment portfolios and not for trading portfolios of financial institutions because it does not work well when deltas are low. The other limitation of the variance-covariance method is that we assume the returns of the assets are normally distributed and therefore the method is not flexible on the distribution of the returns.

3.3.2 Historical Simulation

The historical simulation is the most popular method for estimating VaR amongst risk managers. The idea behind historical simulation is straightforward, the method assumes that historical returns will be observed again in the future. The first step is to identify the risk factors (in this case the equity prices) affecting the portfolio and collect data of the risk factors measured over a particular frequency (e.g. daily) over a given period of time. The portfolio under consideration is then re-evaluated under each historical simulated return. Secondly, evaluate the portfolio future returns for each simulated portfolio value. Lastly, given a confidence level $(1-q)$ VaR is estimated by sorting the financial returns and determining the $(1-q)$-quantile of the returns distribution.

The strength of the historical simulation is that it is easy to implement and it requires no parameter estimations because VaR is simply estimated by means of ordered observed profits and losses. The main disadvantage of the historical simulation method is it does not incorporate volatility-updating [HW98]. If the historical data is too large then irrelevant values from the distance past will have a misleading influence on the VaR estimate since all the data have equal weights. This issue is addressed in Hull and White's paper [HW98] where they incorporated volatility updates to the historical returns.

3.3.3 Monte Carlo Simulation

Monte Carlo simulation is based on simulating the future returns using an explicit parametric model. The basic algorithm of the Monte Carlo simulation is given according to Simon Hubbert in [Hub12] as follow:

(i) Choose an appropriate model for the daily returns.

(ii) Use the model to simulate potential future daily returns.

(iii) Repeat step 2 many times and use the simulated values to create an approximate distribution.

(iv) For a given confidence level evaluate VaR using the same approach as with the historical simulation method.

Monte Carlo accommodates other distributions as alternatives to the Normal distribution to allow for heavy tail distributions, where extreme events are expected to occur more commonly than the Normal distribution [CGM01]. The main disadvantage of the Monte Carlo method is that it is very slow because of the number of calculations required.

3.4 Advantages and Limitations of VaR

Although VaR is considered to be a simple concept to understand since it aggregates all the portfolio risk under one single number, it also has limitations. In this subsection, we discuss the positive points as well as some of VaR's limitations.

3.4.1 Advantages of VaR

Dowd [Dow02] describes the attraction of VaR as a risk measure under two characteristics:

(i) VaR provides a measure that is common and consistent for different portfolio positions and risk factors. For example, comparing the risk associated with fixed income position to risk associated with an equity position. VaR also provides institutions with an overall risk measure.

(ii) VaR takes into account the correlation of different risk factors in a portfolio. If the risk factors in a portfolio are negatively correlated the estimated VaR figure will be lower and if the risk factors were positively correlated the estimated VaR figure will be higher.

3.4.2 Limitations of VaR

The limitations of VaR according to Jorion [Jor01] are:

(i) The obvious limitation VaR has is it only provides us with the losses under "normal" market conditions with some confidence level i.e. VaR tells us the worst we can lose 95% of the time. VaR does not provide an estimate of the absolute possible losses under extreme conditions.

(ii) VaR is not a coherent risk measure since it is not sub-additive. A risk measure that is not sub-additive implies that portfolio diversification is a bad thing because the risk of the portfolio will be greater than the risk of the sum of individual assets.

(iii) VaR assumes the portfolio positions stays unchanged over the holding period. Hence, the adjustment of 1-day VaR to a multiple-day VaR using the square root of time does not accommodate for change in the portfolio positions. This ignores the possibility that trading position may change over time in response to change in market conditions.

(iv) VaR models rely on historical data and therefore assumes that the recent past is a good projection of the future randomness.

3.5 Properties of a Coherent Risk Measure

Suppose X and Y are random cash flows or future values such that $X < Y$, then a risk measure $\rho(.)$ is coherent if it satisfying the following condition in [Hul10] introduced by Artzner *et. al.* [ADEH99]. For $b > 0$ and $c > 0$ real numbers, the axioms are as follow:

Monotonicity: $\rho(Y) < \rho(X)$, $X < Y$ This condition states that if one portfolio produces better returns than the other, its risk must be less than the portfolio with worst returns.

Translation invariance: $\rho(X + b) = \rho(X) - b$ The condition says if a positive amount b is added into a portfolio then the portfolio risk is reduced by the added amount b.

Homogeneity: $\rho(cX) = c\rho(X)$ The third condition states that if we double our position then the risk will also double.

Subadditivity: $\rho(X+Y) \leq \rho(X)+\rho(Y)$ The last condition states that the portfolio risk should be less or equal to the sum of the individual risk factors.

VaR satisfy the first three conditions and the last conditions is not always satisfied. VaR of a portfolio is not always less than VaR of individual assets in the portfolio. This implies that having a diversified portfolio has no benefits since the portfolio risk can be greater than the individual risk factors. In contrast to VaR, the expected shortfall measure is coherent [Hul10].

3.6 Expected Shortfall

One of VaR's disadvantages is that despite the fact that it tells us of the size of the worst loss we can expect to suffer 100Xq% of the time, it does not tell us the size of the expected worst loss for the remaining $(100 - q)\%$. Expected shortfall (ES) is the risk measure that provides us with the magnitude of the expected loss $(100 - q)\%$ of the time. While VaR gives information on the magnitude of the worst loss under "normal" market condition, ES gives information on the size of worst expected loss under extreme market conditions. Like VaR, ES is a function of two parameters: over time horizon h and the confidence level q.

Dowd [Dow02] give the following definition of expected shortfall:

Definition 3.6.1. *The Expected Shortfall (ES) is the expected value of our losses, L, if we get a loss in excess of VaR:*

$$ES_q = \mathbb{E}[L|L > VaR_q]. \tag{3.3}$$

Expected shortfall seem to be a better risk measure than VaR, because it is a coherent measure and therefore always satisfy the subadditivity property and hence encourages diversification. Expected shortfall also goes by the name **Conditional VaR (CVaR)**, **Expected Tail Loss (ETL)**, and **Average Value-at-Risk (AVaR)**.

Dowd [Dow02] recommend the use of ES rather than VaR, unless it is more difficult to estimate ES. VaR is simpler to understand and VaR estimates can easily be back-tested unlike ES. Despite VaR limitations it is the most popular risk measure among both regulators and risk managers [Hul10].

3.7 Backtesting VaR

The aim of backtesting[1] is to validate the VaR model. It is a process of periodically comparing the daily projected losses (i.e. the VaR estimate) with the observed daily profits and losses [oBS96b]. This is to check how often the daily losses exceed the daily VaR estimate. The number of daily losses exceeding VaR estimate are referred to as *exceptions*. In this respect, the accurate VaR model should not have a number of exceptions that exceed a certain fraction of the daily profits and losses, where the fraction is determined by the confidence level of the VaR model [oBS96b]. If the number of exceptions exceeds a certain fraction of the returns, then the VaR model is considered not to measure risk accurately. Therefore, verifying the VaR model in this case is simply based on counting the number of exceptions over a given period and comparing it to the number of the given confidence level.

The verification of the model's accuracy is fundamental to the Basel Committee so as to prevent financial institutions understating their risk. The framework of backtesting is set out in [oBS96b]. Furthermore, the amount of capital required to be held by financial institutions depends on the outcome of the backtesting procedure[oBS96b]. The regulatory backtesting procedure is performed over the last 250 trading days with the 99% one-day VaR compared to the observed daily profits and losses over the period. For example, over 250 trading days, a 99% daily VaR model should have on average 2.5 exceptions out of 250.

The Basel Committee have classified the backtesting results into three zones which are green, yellow and red zones, these zones are linked to the capital requirement scaling factor. The zones are chosen in order to balance two types of statistical error, Type 1 and Type 2 errors. Type 1 error is the probability of rejecting an accurate risk model, while Type 2 error is the

[1] Backtesting in this dissertation is based on the Basel Committee framework [oBS96b].

probability of accepting an inaccurate risk model [oBS96b]. The description of the zones are [oBS96b]:

(i) The *green zone* consists of zero to four exceptions, indicating that the backtesting results do not suggest a problem with the quality or accuracy of the VaR model.

(ii) The *yellow zone* consists of exceptions from five to nine. The backtesting results in this zone could be generated by both accurate and inaccurate risk models. The number of exceptions in the yellow zone guide the size of the potential increase in the capital requirement scaling factor, as described in Table 3.1. However, the increase is not automatic since the results in the yellow zone do not always imply the risk model is inaccurate. If the financial institution is able to provide a reason for the number of exceptions, the supervisor will decide on applying an increase in the capital requirement factor. The Basel Committee [oBS96b] classifies the reason of the number of exceptions in to the following categories:

   (a) *Basic integrity of the model*: The risks of positions are not correctly captured or the model is not calculating volatility or correlation correctly.

   (b) *Model's accuracy could be improved*: The model does not measure the risk of some instruments with enough precision.

   (c) *Bad luck or markets moved in fashion not anticipated by the model*: Market volatility and correlation were significantly different than what the model predicted.

   (d) *Intra-day trading*: Large positions change during the day.

(iii) If the number of exception is ten or more, the backtesting results falls into the *red zone*. The red zone generally indicates a problem with the risk model and increase in the capital requirement scaling factor is automatic.

Zone	Number of Exceptions	Scaling Factor
Green	0 to 4	3
Yellow	5	3.4
	6	3.5
	7	3.7
	8	3.8
	9	3.9
Red	10 or more	4

Table 3.1: The Basel Committee [oBS96b] classification of backtesting outcomes with corresponding number of exceptions and their scaling factor.

3.8 Chapter Summary

Value-at-Risk is the core concept in this dissertation. The purpose of this chapter was to introduce the notation of Value-at-Risk as well as provide the reader with insight into the concepts. We discussed Value-at-Risk as a market risk measure and the three standard methods for calculating it, which are variance-covariance, historical simulation and Monte Carlo simulation.

We discussed the limitation and advantages of VaR and stated the properties of a coherent risk measure, where VaR satisfy only three conditions out of four. Hence we discuss the Expected Shortfall which is a coherent risk measure. Unlike Expected Shortfall, VaR is simpler to understand and easy to back-test.

Chapter 4

Normal Inverse Gaussian Distribution

In this chapter, we define the NIG distribution introduced by Barndorff-Nielsen [BN95] in 1995. Firstly, we give a brief overview of Lévy Processes. Secondly, we describe the decompositions of the Normal Inverse Gaussian process according to the theory of Lévy processes. Lastly, we discuss the parameters and properties of the NIG distribution.

4.1 Lévy Processes

In financial literature, uncertainty of the economy is represented by a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, where $\mathcal{F}_t$ is the information generated by history of assets up to time t, $\mathbb{P}$ is the real probability measure and Ω is a space of all asset paths.

Definition 4.1.1. [Kyp05] *A process $X = \{X_t : t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a Lévy process if it possesses the following properties:*

(i) The paths of X are $\mathbb{P}$-almost surely right continuous with left limits.

(ii) $\mathbb{P}(X_0 = 0) = 1$.

(iii) *For* $0 \leq s \leq t$, $X_t - X_s$ *is equal in distribution to* X_{t-s}.

(iv) *For* $0 \leq s \leq t$, $X_t - X_s$ *is independent of* $\{X_u : u \leq s\}$.

The first property implies that a Lévy process starts at position zero. The increments of the Lévy process are the differences $X_t - X_s$ between its values at different times $s < t$. The second property states that the increments $X_t - X_s$ of the process are stationary, that is increment with equal time length have identical distribution, hence increments are dependent only on the time length $t - s$. Independent increments implies no time intervals are overlapping each other, that is $X_t - X_s$ is independent of X_u for $u \leq s$. The continuous time viewpoint forces the distribution of the increments $X_t - X_s$ to belong to the infinitely divisible one [BB00].

Lévy process is a time-continuous stochastic process with independent and stationary increments. The most known example of Lévy process is the Brownian motion and Poisson process. A comprehensive summary of some examples of Lévy processes can be found in [Ben04], [Kyp05] and [Pap05]. The distribution of X_t is one of the things we will be modelling in this dissertation.

The other important results of a Lévy process is that for any $t > 0$, X_t is a random variable belonging to the class of *infinitely divisible* distributions [Kyp]. That is for any $n = 1, 2, \ldots,$ we have

$$X_t = X_{t/n} + (X_{2t/n} - X_{t/n}) + (X_{3t/n} - X_{2t/n}) + \cdots + (X_t - X_{(n-1)t/n}).$$

Therefore X_t can be written as the sum of n independent identically distributed random variables [Gem02].

The first statistical distribution that we model X_t with is the NIG. The NIG process is a Lévy process with stationary independent increments, where the increments follow a NIG distribution [Hen10]. The distribution of a Lévy process can be uniquely determined by a characteristic triplet (σ^2, ν, γ). This leads to the following Lévy-Khinchine theorem.

Definition 4.1.2. [Hen10] *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet $(\sigma^2\,, \nu\,, \gamma)$. Then*

$$\mathbb{E}[e^{izX_t}] = e^{t\psi(z)}, \quad z \in \mathbb{R} \tag{4.1}$$

with

$$\psi(z) = -\frac{1}{2} z\,\sigma^2\, z + i\,\gamma\, z + \int_{|x|<1} (e^{izx} - 1 - izx)\nu(dx) + \int_{|x|\geq 1} (e^{izx} - 1)\nu(dx), \tag{4.2}$$

where $\gamma \in \mathbb{R}$ is the drift constant, $\sigma \geq 0$ is the diffusion constant and ν is the Lévy measure of the process X_t on $\mathbb{R} \setminus \{0\}$ satisfying the conditions

$$\int_{|x|\leq 1} |x^2| \nu(dx) < \infty, \quad \int_{|x|>1} \nu(dx) < \infty.$$

By the theory of Lévy processes, a Lévy process X_t can be decompositioned into a drift term, Brownian motion term and jump terms [Hen10]. That is:

$$X_t = \gamma t + \sigma B_t + \int_{|x|<1} (e^{izx} - 1 - izx)\nu(dx) + \int_{|x|\geq 1} (e^{izx} - 1)\nu(dx), \tag{4.3}$$

where $\int_{|x|\geq 1}(e^{izx} - 1)\nu(dx)$ is the compensated Poisson random measure of X_t, $\nu(dx)$ is the rate of intensity of the process with jumps of size x and B_t is the Brownian motion [Gem02].

In the case of the NIG Lévy process, the Lévy- Khinchine formula for ψ is given by

$$\psi(z) = \gamma t + \int_{|x|<1} (e^{izx} - 1 - izx)\nu(dx) + \int_{|x|\geq 1} (e^{izx} - 1)\nu(dx),$$

for

$$\gamma = \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) dx, \quad \nu(dx) = \frac{\delta\alpha}{\pi|x|} \exp(\beta x) K_1(\alpha|x|) dx,$$

where $x \in \mathbb{R}$, $\delta > 0$, $\beta \leq |\alpha|$ and $K_1(x)$ is the modified Bessel function of the third kind with index 1 [BN95]. The derivation of this representation can be found in [BN95]. From this it is immediately clear that the NIG Lévy process is described by the characteristic triplet $(0, \nu, \gamma)$. The first term of the characteristic triplet is zero indicating that the NIG process has no diffusion component. This makes the NIG a pure jump process. From Equation (4.3) the NIG process X_t can be represented in terms of Poisson process in the form

$$X_t = \gamma t + \int_{|x|<1} (e^{izx} - 1 - izx)\nu(dx) + \int_{|x|\geq 1} (e^{izx} - 1)\nu(dx).$$

The Brownian motion is a popular model in finance, its Lévy process is given by

$$X_t = \gamma t + \sigma B_t, \tag{4.4}$$

and the corresponding characteristic function is given by

$$\mathbb{E}[e^{izX_t}] = e^{t(-\frac{1}{2}\sigma^2 z^2 + i\gamma z)}.$$

The Lévy process of the Brownian motion has no jump terms and therefore has stationary independent increments with continuous paths. Putting Equation (4.4) into the stock price model, we have $S_t = S_0 e^{(\gamma t + \sigma B_t)}$.

4.2 NIG Definition

The NIG distribution is a special case of the generalised hyperbolic distributions introduced by Barndorff-Nielsen in 1977 [BN95]. It is a continuous distribution defined on the entire real line.

Definition 4.2.1. [BN97] *The random variable $X : \Omega \to \mathbb{R}$ follows a NIG distribution with parameters α, β, μ and δ, if its probability density function, defined for all real $x \in \mathbb{R}$, is given by*

$$f_X(x;\alpha,\beta,\mu,\delta) = \frac{\delta\alpha}{\pi}\exp\left(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\right)\frac{K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\sqrt{\delta^2+(x-\mu)^2}} \tag{4.5}$$

where the function $K_1(x) = \frac{1}{2}\int_0^\infty \exp(-\frac{1}{2}x(\tau+\tau^{-1}))d\tau$ is a modified Bessel function of third order and index 1. In addition the parameters must satisfy $0 \leq |\beta| \leq \alpha$, $\mu \in \mathbb{R}$, $0 < \alpha$ and $0 < \delta$. If a random variable x follows a NIG distribution it can be denoted in short as $x \sim \mathrm{NIG}(\alpha,\beta,\mu,\delta)$ [BN97].

Barndorff-Nielsen [BN95] describe the NIG distribution and the Hyperbolic distributions as normal variance-mean mixtures. Here, the NIG distribution occurs as the marginal distribution of x for a pair of random variables (x, z), i.e. the conditional distribution of x given z is normally distributed with mean $\mu + \beta z$ and variance z. In symbols:

$$x|z \sim \mathrm{N}(\mu + \beta \mathrm{z}, \mathrm{z}). \tag{4.6}$$

Provided that the random variable z follows the Inverse Gaussian distribution with parameters δ and $\sqrt{\alpha^2-\beta^2}$ and its density distribution function is given by:

$$\mathrm{g}(\mathrm{z};\delta,\sqrt{\alpha^2-\beta^2}) = \frac{\delta}{\sqrt{2\pi}}\,\mathrm{z}^{-3/2}\exp\left(\delta\sqrt{\alpha^2-\beta^2} - \frac{\delta^2\,\mathrm{z}^{-1}+(\alpha^2-\beta^2)\,\mathrm{z}}{2}\right).$$

Proposition 4.2.1. [BN95] *The moment generating function of NIG distribution is given by:*

$$\mathrm{M}_x(t) = \exp\left[t\mu + \delta\big(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\right]. \tag{4.7}$$

Furthermore, the mean, variance, skewness and the kurtosis of random variable x can be obtained by successively differentiating the moment generating function.

Proposition 4.2.2. [Lil00] *The mean, variance and the kurtosis of the random variable x distributed according to* (4.5) *are given by the following expressions:*

$$\mathbb{E}[x] = \mu + \frac{\delta\,\beta}{\sqrt{\alpha^2-\beta^2}} \tag{4.8}$$

$$\mathrm{Var}[x] = \frac{\delta\,\alpha^2}{(\alpha^2-\beta^2)^{3/2}} \tag{4.9}$$

$$\mathrm{Skew}[x] = 3\alpha^{-1/4}\frac{\beta/\alpha}{(1-(\beta/\alpha)^2)^{1/2}} \tag{4.10}$$

$$\mathrm{Kurt}[x] = 3\,\frac{1+4(\beta^2/\alpha^2)}{\delta\sqrt{\alpha^2-\beta^2)}}. \tag{4.11}$$

Proof. The mean of x is obtained by finding the first derivative using the *Chain Rule* of the moment generating function as follow:

$$\mathrm{M}'_x(t) = \exp\left[t\mu + \delta\big(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\right]\left[\mu+\delta\big(\frac{-1}{2}(\alpha^2-(\beta+t)^2)^{-1/2}(-2(\beta+t))\big)\right]$$

simplifying and taking $t=0$ yields

$$\mathrm{M}'_x(0) = \mu + \frac{\delta\,\beta}{\sqrt{\alpha^2-\beta^2}}.$$

Thus the mean is given by

$$\mathbb{E}[x] = \mathrm{M}'_x(0) = \mu + \frac{\delta\,\beta}{\sqrt{\alpha^2-\beta^2}}.$$

Similarly, the variance is obtained by first finding the second derivative using both the *Chain Rule* and *Product Rule* of the moment generating function, which is a more complicated function:

$$\begin{aligned}
\mathrm{M}_x''(t) &= \exp\Big[t\mu + \delta\big(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big] \\
&\quad \Big[\mu + \delta\big(\frac{-1}{2}(\alpha^2-(\beta+t)^2)^{-1/2}(-2(\beta+t))\big)\Big]^2 \\
&\quad + \exp\Big[t\mu + \delta\big(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big] \\
&\quad \Big[\delta\big(\frac{-1}{2}(\alpha^2-(\beta+t)^2)^{-3/2}(-2(\beta+t)^2)\big) + \delta(\alpha^2-(\beta+t)^2)^{-1/2}\Big].
\end{aligned}$$

Simplifying and taking $t = 0$ yields

$$\mathrm{M}_x''(0) = \Big(\mu + \frac{\delta\,\beta}{\sqrt{\alpha^2-\beta^2}}\Big)^2 + \frac{\delta\,\beta^2}{(\alpha^2-\beta^2)^{3/2}} + \frac{\delta}{\sqrt{\alpha^2-\beta^2}}$$

and so

$$\mathbb{E}[x^2] = \mathrm{M}_x''(0) = \Big(\mu + \frac{\delta\,\beta}{\sqrt{\alpha^2-\beta^2}}\Big)^2 + \frac{\delta\,\beta^2}{(\alpha^2-\beta^2)^{3/2}} + \frac{\delta}{\sqrt{\alpha^2-\beta^2}}.$$

The variance of x is thus given by

$$\mathrm{Var}[x] = \mathbb{E}[x^2] - (\mathbb{E}[x])^2 = \frac{\delta\,\beta^2}{(\alpha^2-\beta^2)^{3/2}} + \frac{\delta}{\sqrt{\alpha^2-\beta^2}}.$$

Rationalizing the denominator of the second term as follow:

$$\begin{aligned}
&= \frac{\delta\,\beta^2}{(\alpha^2-\beta^2)^{3/2}} + \frac{\delta}{\sqrt{\alpha^2-\beta^2}}\frac{(\alpha^2-\beta^2)}{(\alpha^2-\beta^2)} \\
&= \frac{\delta\,\beta^2}{(\alpha^2-\beta^2)^{3/2}} + \frac{\delta(\alpha^2-\beta^2)}{(\alpha^2-\beta^2)^{2/3}},
\end{aligned}$$

implies

$$\mathrm{Var}[x] = \frac{\delta\,\alpha^2}{(\alpha^2-\beta^2)^{3/2}}.$$

Similar computation can be used to find the expressions of the kurtosis and the skewness for the random variable x.

□

4.3 NIG Parameters Description

The NIG distribution is characterized by four parameters α, β, μ and δ, each describing the overall shape of the density distribution [Lil98]. These parameters are usually categorized in one of the two groups. The first group of parameters affecting the scaling and location of the distribution are μ and δ. The second group of parameters affecting the shape of the distribution are α and β.

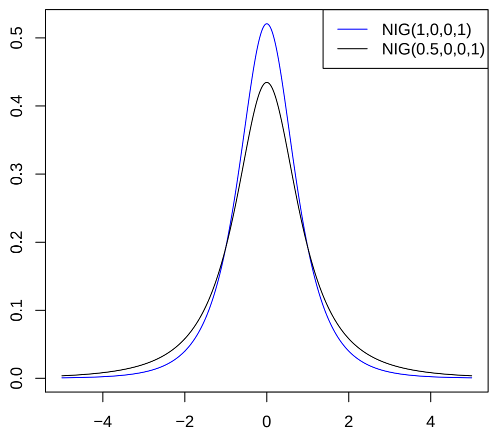


Figure 4.1: The NIG distribution with different α

.

The parameter α measures the tail heaviness of the distribution. The larger α the thinner the tails and the smaller α the fatter the tails, (see Figure 4.1). The skewness of NIG distribution is measured by β. When $\beta = 0$, the distribution is symmetric around μ. If $\beta > 0$, then the distribution is skewed to the right, whereas negative β gives skewness to the left [BB00] as demonstrated in Figure 4.2.

The parameter μ and δ have the same interpretation as the mean and

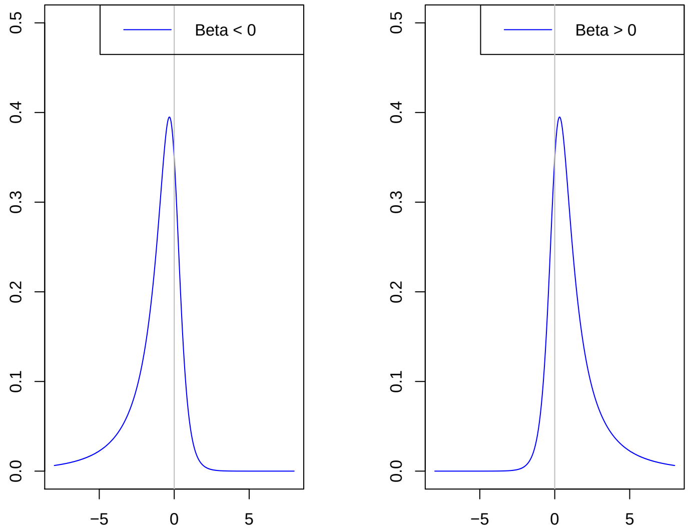


Figure 4.2: The NIG distribution with different β.

standard deviation on the Normal distribution. The parameter μ describes the location of the peak of the distribution or were the distribution is centered on the real number line [BB00]. The right-hand side of Figure 4.3 shows the different values of μ, the distribution of the blue graph is centered at $x = 2$ while the distribution of the orange graph is centered at $x = -2$. The scale parameter is δ, it describes the spread of the returns. The higher the value of δ the wider the distribution and the lower the value the narrower the distribution. The left hand side of Figure 4.3 demonstrates the different values of δ where the orange graph is when δ has a higher value and the blue graph is when δ is small.

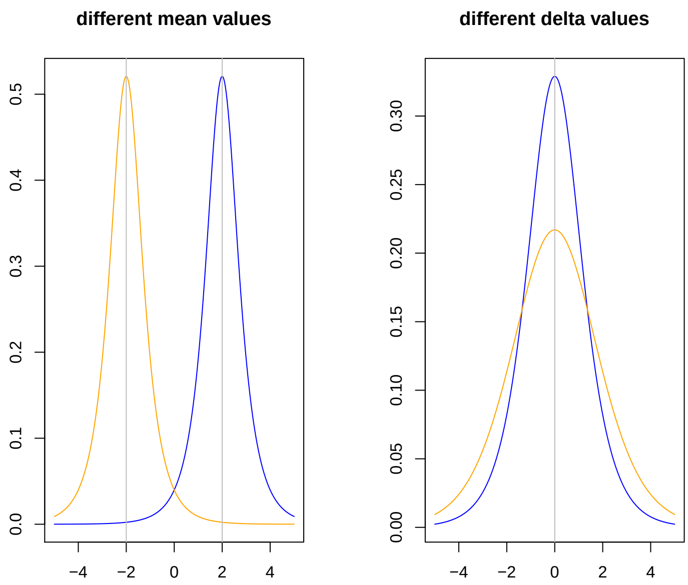


Figure 4.3: The NIG distribution with different δ and μ values.

4.4 NIG Distribution Properties

4.4.1 Convolution Property

One of the important property of the NIG distribution is that it is closed under convolution. This means, the sum of independent and identical random variables which are NIG distributed is also NIG distributed.

Proposition 4.4.1.1. [Ben04] *If X and Y are independent NIG random variables with common parameters α and β but having different location-scale parameters that is $X \sim \mathrm{NIG}(\alpha, \beta, \mu_{\mathrm{X}}, \delta_{\mathrm{X}})$ and $Y \sim \mathrm{NIG}(\alpha, \beta, \mu_{\mathrm{Y}}, \delta_{\mathrm{Y}})$, then $X + Y$ is NIG distributed with parameters $(\alpha, \beta, \mu_X + \mu_Y, \delta_X + \delta_Y)$.*

This property is proved using the moment generating function properties which are:[Ros03]

(i) *The moment generating function uniquely determines the distribution.*

(ii) *The moment generating function of the sum of independent random variables is just the product of the individual moment generating functions.*

Proof. The moment generating function of $X + Y$ is given by

$$\begin{aligned}
\mathrm{M}_{\mathrm{X+Y}}(t) &= \mathrm{M}_{\mathrm{X}}(t) \times \mathrm{M}_{\mathrm{Y}}(t) \\
&= \exp\Big(t\mu_{\mathrm{X}} + \delta_{\mathrm{X}}\big(\sqrt{\alpha^2+\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big) \\
&\quad \times \exp\Big(t\mu_{\mathrm{Y}} + \delta_{\mathrm{Y}}\big(\sqrt{\alpha^2+\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big) \\
&= \exp\Big(t\mu_{\mathrm{X}} + \delta_{\mathrm{X}}\big(\sqrt{\alpha^2+\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big) \\
&\quad + t\mu_{\mathrm{Y}} + \delta_{\mathrm{Y}}\big(\sqrt{\alpha^2+\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big) \\
&= \exp\Big(t(\mu_{\mathrm{X}}+\mu_{\mathrm{Y}}) + (\delta_{\mathrm{X}}+\delta_{\mathrm{Y}})\big(\sqrt{\alpha^2+\beta^2} - \sqrt{\alpha^2-(\beta+t)^2}\big)\Big),
\end{aligned}$$

which is the moment generating function of a NIG distributed with parameters $(\alpha, \beta, \mu_X+\mu_Y, \delta_X+\delta_Y)$. Hence $X+Y$ is NIG distributed with parameters $(\alpha, \beta, \mu_X + \mu_Y, \delta_X + \delta_Y)$, since the moment generating function uniquely determines the distribution.

□

4.4.2 Tail Behaviour

The NIG distribution has semi-heavy tail. In particular by using the following asymptotic formula of the Bessel function

$$K_1(s) \sim \sqrt{\pi/2}\, s^{-1/2}\, \mathrm{e}^{-s}, \quad \text{as} \quad s \to \infty$$

Barndorff-Nielsen [BN95] found that the tail of the NIG behaves as:

$$f(x;\alpha,\beta,\mu,\delta) \sim \mathrm{A}(\alpha,\beta,\mu,\delta)\mathrm{q}\left(\frac{x-\mu}{\delta}\right)^{-3/2} \exp\left[-\alpha\sqrt{\delta^2+(x-\mu)^2} + \beta(x-\mu)\right]$$

for $|x| \to \infty$, where $\mathrm{q}(x) = \sqrt{1+x^2}$ and the constant A is given by

$$\mathrm{A}(\alpha,\beta,\mu,\delta) = (2\pi)^{-1/2}(\alpha/\delta)^{1/2} \exp\left(\delta\sqrt{\alpha^2-\beta^2}\right).$$

This means the tails of the NIG distribution are always heavier than those of the Normal distribution. In addition, a proficient selection of parameters can create a wide range of density shapes, making the NIG distribution a very flexible tool to use for modelling financial returns.

4.4.3 Fitting the NIG model

Bølviken and Benth in [BB00] describe two methods for finding the parameter estimators. The first method they describe is the method of moments, since the expression of the mean, variance, skewness and kurtosis are described by the four parameters of the NIG distribution. To estimate the parameters the mean, variance, skewness and kurtosis are simply replaced by their sample version and the four equations are then solved for α, β, μ and σ.

The method of moments is straightforward since the Bessel function does not have to be evaluated [BB00]. However, Bølviken and Benth [BB00] argue that the method is based on higher moments and there is no guarantee that the sample moment satisfy the restrictions laid down by the NIG family, such that the moment equation has an actual solution. Bølviken and Benth [BB00] recommend the second method, which is the maximum likelihood estimation.

The maximum likelihood method requires the evaluation of the maximum likelihood function and it is obtained by maximizing the log-likelihood function and this will also require the evaluation of the Bessel function and its derivative as described in [BB00]. We prefer the maximum likelihood function since the density function of the NIG distribution is readily available to derive the maximum likelihood function. By maximizing the likelihood function we increase the chances of obtaining the best estimates. The statistical program R has a predefined function for the maximum likelihood estimation (MLE) of the NIG distribution. The log-likelihood function is obtained by taking the log of the density function of the NIG in Equation (4.5):

$$
\begin{aligned}
L(x_i,\theta) &= \ln\Big(\prod_{i=1}^{n} f(x_i;\alpha,\beta,\mu,\delta)\Big) \qquad (4.12)\\
&= \ln\Big(\prod_{i=1}^{n}\frac{\delta\alpha}{\pi}\exp\Big(\delta\sqrt{\alpha^2-\beta^2}+\beta(x_i-\mu)\Big)\frac{K_1\Big(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\Big)}{\sqrt{\delta^2+(x_i-\mu)^2}}\Big),
\end{aligned}
$$

simplifying the expression we get

$$
\begin{aligned}
&= \sum_{i=1}^{n}\ln\Big(\frac{\delta\alpha}{\pi}\Big)+\ln\Big(\prod_{i=1}^{n}\exp\Big(\delta\sqrt{\alpha^2-\beta^2}+\beta(x_i-\mu)\Big)\Big)\\
&\quad +\ln\Big(\prod_{i=1}^{n}K_1\Big(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\Big)\Big)-\frac{1}{2}\ln\Big(\prod_{i=1}^{n}(\delta^2+(x_i-\mu)^2)\Big)\\
&= n\ln(\delta\alpha)-n\ln(\pi)+n\big(\delta\sqrt{\alpha^2-\beta^2}-\beta\mu\big)+\beta\sum_{i=1}^{n}x_i\\
&\quad +\sum_{i=1}^{n}\ln\Big(K_1\big(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\big)\Big)-\frac{1}{2}\sum_{i=1}^{n}\ln\Big(\delta^2+(x_i-\mu)^2\Big).
\end{aligned}
\tag{4.13}
$$

To obtain the maximum likelihood parameter estimates we differentiate the log-likelihood function L in Equation (4.12) with respect to each parameter. For example the derivative of log-likelihood function with respect to β is:

$$
\begin{aligned}
\frac{\partial}{\partial\beta}L &= n\delta\left(\frac{1}{2\sqrt{\alpha^2-\beta^2}}\right)(-2\beta)-n\mu+\sum_{i=1}^{n}x_i\\
&= \frac{-n\delta\beta}{\sqrt{\alpha^2-\beta^2}}-n\mu+\sum_{i=1}^{n}x_i.
\end{aligned}
$$

Setting $\frac{\partial}{\partial\beta}L=0$ and solving for β we obtain the following likelihood estimate of β:

$$
\widehat{\beta}=\frac{\alpha(n\mu-\sum_{i=1}^{n}x_i)}{\sqrt{n^2\delta^2+(n\mu-\sum_{i=1}^{n}x_i)^2}}. \tag{4.14}
$$

The derivative of the likelihood function (4.12) with respect to β is easy to obtain as it is not dependent on the Bessel function. We apply the properties of derivatives for the modified Bessel function states in Appendix A to obtain the derivative of the log-likelihood function with respect to the remaining parameters for example μ is:

$$\begin{aligned}\frac{\partial}{\partial\mu}L &= -n\beta + \sum_{i=1}^{n}\frac{x_i-\mu}{\sqrt{\delta^2+(x_i-\mu)^2}} \text{ x } \Big[\frac{2}{\sqrt{\delta^2+(x_i-u)^2}} \\ &\quad + \frac{\alpha K_0(\alpha\sqrt{\delta^2+(x_i-u)^2})}{K_1(\alpha\sqrt{\delta^2+(x_i-u)^2})}\Big].\end{aligned}$$

Setting $\frac{\partial}{\partial\mu}L = 0$ and solving μ we obtain the likelihood estimates for μ to be [Pra99]:

$$\widehat{\mu} = -\frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}} + \frac{1}{n}\sum_{i=1}^{n}x_i. \tag{4.15}$$

The derivatives of the likelihood function (4.12) with respect to δ and α are given by [Pra99]:

$$\frac{\partial}{\partial\delta}L = \frac{n}{\delta} + n\sqrt{\alpha^2-\beta^2} - 2\sum_{i=1}^{n}\frac{\delta}{\delta^2+(x_i-\mu)^2} - \sum_{i=1}^{n}\frac{\alpha\delta}{\sqrt{\delta^2+(x_i-\mu)^2}} \tag{4.16}$$

and

$$\frac{\partial}{\partial\alpha}L = \frac{n\delta\alpha}{\sqrt{\alpha^2-\beta^2}} - \sum_{i=1}^{n}\sqrt{\delta^2+(x_i-\mu)^2}\frac{K_0(\alpha\sqrt{\delta^2+(x_i-u)^2})}{K_0(\alpha\sqrt{\delta^2+(x_i-u)^2})}. \tag{4.17}$$

According to Prause in [Pra99], it is preferable to maximize the $\ln(\alpha)$ and $\ln(\delta)$ to maintain the positivity condition of α and δ.

4.5 Chapter Summary

This chapter served as an introduction to the NIG distribution and its properties. It is intended to provide the background understanding required to implement the NIG distribution. The flexibility of the NIG distribution in terms of the skewness and kurtosis was explained by adjusting various parameters. The Lévy process was also presented where we discussed the NIG Lévy process and represented it in terms of the drift and the jump terms.

Chapter 5

Other distributions

In this chapter we provide an overview of alternative distributions that have been developed in literature to model financial returns. This is to familiarise the reader with some of the distributions that have been used in literature. We discuss the Stable Paretian distribution (Stable distribution), Skew t and t-distribution amongst the alternatives. The other goal of this dissertation is to fit some of these alternative distributions to our data and make a comparison to the NIG distribution.

5.1 The Stable Paretian distribution

Mandelbrot in [Man63] proposed that the financial returns can be characterized by the Stable distribution. The Stable distribution is described by four parameters: location parameter, scale parameter, skewness parameter and the characteristic exponent. The scale parameter is always greater than zero, while the skewness parameter is taken in the interval $[-1; 1]$. The location parameter is the mean or expected value of the distribution when the characteristic exponent is equal to one [Fam65]. The characteristic exponent measures the height of the extreme tails of the distribution and is taken in the interval $(0; 2]$ [Fam65]. The Stable distribution exhibit heavier tails than those of the Normal distribution when the characteristic exponent is taken in the interval $(0; 2]$, and the height of the extreme tails increases as the characteristic exponent [1] moves away from 2 and towards 0 [Fam63]. Mandelbrot's Stable distribution for financial returns is specified for the characteristic exponent in the interval $(1; 2]$, such that the distribution has a finite expected value but the variance is infinite [2][Fam63].

The infinite variance of the Stable Paretian distribution has "extreme consequence" from a statistical point of view, if the population variance of the distribution is infinite the sample variance will be a meaningless measure of dispersion Fama [Fam65]. Fama in [Fam63] argues that if the variance is infinite, other statistical tools for example, the least square regression, which is based on the assumption of finite variance or tools based on the sample second moments, will produce meaningless results. Consequently, by Fama in [Fam63] the Stable Paretian distribution seem to fit the data better than the Normal distribution. Mandelbrot's work may be "doubtful", since past research on "speculative prices" has been based on statistical tools which assume the existence of a finite variance.

The other shortcoming of the Stable Paretian distribution is the non-existence of an explicit density function, [3] making it difficult to analyse the sampling behaviour of the estimation of the Stable Paretian distribution parameters [Fam65]. The implementation problem of the Stable distribution has also been noted by Huisman in [HKP98].

[1] The Normal distribution is obtained when the characteristic exponent equals 2. Therefore, the Normal distribution is a special case of the Stable Paretian distribution [Fam63].

[2] If the characteristic exponent is in the interval $(0; 1)$ there is no obvious interpretation of the expected value [Man63]. The expected value of the Stable Paretian distribution is finite when the characteristic exponent is greater than 1 and the variance exists only when the characteristic exponent is equal to 2 [Fam63].

[3] The Stable Paretian distribution has an explicit density function in only two cases: the Cauchy (when the characteristic exponent equals 1) and the Normal distributions (when the characteristic exponent equals 1).

For further applications of the Stable Paretian distribution in financial returns see for example, Mandelbrot [Man63] and Blattberg et al [BG74]. Fama in [Fam63] goes into a detailed analysis of Mandelbrot's proposed Paretian Stable distribution as a model for financial returns.

5.2 The Student's t-distribution

The t-distribution suggested by Blattberg and Gonedes [BG74] as an alternative distribution to model financial returns, is characterised by the shape-defining parameter known as the degree of freedom $k > 0$. The density function is given by

$$t_k(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\,\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \tag{5.1}$$

where x is the random variable, k is the degree of freedom and Γ is the gamma function. The t-distribution is similar to Normal distribution, it is symmetrical about the mean and exhibits fatter tails. The mean of the distribution is zero when the $k > 1$, the mean is undefined and the variance is given by $k/(k-2)$ for $k > 2$, the variance is infinite for $0 < k \leq 2$. The degree of freedom k controls the fat tails of the distribution. The smaller the value of k the fatter the tails of the distribution. When k increases the variance approaches 1 and therefore the t-distribution converges the Normal distribution [4].

Unlike the Stable Paretian distribution, the t-distribution has a well-defined density function described in Equation 5.1, and therefore the maximum likelihood estimators of the distribution parameters can be easily obtained [BG74]. However, the disadvantage with the t-distribution is that it

[4]When the degree of freedom $k = 1$ we have the Cauchy distribution [Pra72]

does not address the skewness present in many financial returns, since, like the Normal distribution, it is symmetrically distributed about the mean.

Application of the Student's t-distribution in VaR estimation is given by Huisman et al [HKP98] and a most recent application within the South African equity market is by Milwidsky and Mare [MM10]. Blattberg et al in [BG74] make a comparison of the t-distribution to the Stable distribution as a statistical model for financial returns and they conclude that the prior is far easier to implement since it has a well-defined density function and a finite variance.

5.3 The Skew t-distribution

The Skew t is the skew extension of the t-distribution. It was first proposed by Hansen in [Han94] as an alternative distribution to model financial returns. There are several definitions of the density function of the Skew t given in literature. We apply the one proposed by Azzalini and Capitanio in [AC03] and it is given by:

$$st(x:k,\beta) = t_k(x) 2 t_{k+1}\left(\beta x \sqrt{\frac{k+1}{x^2+k}}\right), \tag{5.2}$$

where $t_k(.)$ is the density function of the t-distribution given in Equation (5.1) with degrees of freedom k and β is the skewness parameter. When $\beta = 0$, the Equation (5.2) reduces to the t-distribution. The expected value of the distribution is given by

$$\mathbb{E}[x] = \beta \sqrt{\frac{k}{2}} \frac{\Gamma((k-1)/2)}{\Gamma(k/2)}, \quad k > 1, \tag{5.3}$$

else the expected value does not exist when $k < 1$. The variance of the distribution is given by

$$\mathrm{Var}[x] = \frac{k(1+\beta^2)}{k-2} - \frac{\beta^2 k}{2}\left(\frac{\Gamma((k-1)/2)}{\Gamma(k/2)}\right)^2, \quad k > 2, \tag{5.4}$$

else the variance does not exist when $k \leq 2$. The third moment does not exist when $k < 3$, which makes the usual skewness to not be a good measure of asymmetry for the distribution. According to Aas in [AH06], Skew t has heavy tails, which means that it should model data with heavy tails well but may not handle extensive skewness.

5.4 Generalised Hyperbolic distributions

The most recent alternative to the Normal distribution is a class of Generalised Hyperbolic (GH) distributions and its sub-classes: Hyperbolic and the NIG are a particular cases of the GH. The GH distribution possesses fatter tails and can be both symmetric and skew [Pra99]. Generalised Hyperbolic distributions were first introduced by Barndorff-Nielsen [BN97] in 1977 to model grain size distribution of sand blown by wind. The GH distribution may be defined as a *Normal variance-mean mixture*[5] where the Generalised Inverse Gaussian maybe defined as a mixing distribution [Pra97]. The density function of the Generalised Hyperbolic distribution is given by [Pra97]:

$$\begin{aligned} gh(x;\lambda,\alpha,\beta,\delta,\mu) &= a(\lambda,\alpha,\beta,\delta)(\delta^2+(x-\mu)^2)^{\,(\lambda-1/2)/2} \\ &\quad \times\ \mathrm{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)\exp(\beta(x-\mu)), \\ a(\lambda,\alpha,\beta,\delta) &= \frac{(\alpha^2-\beta^2)^{\,\lambda/2}}{\sqrt{2\pi}\,\alpha^{\lambda-1/2}\,\delta^{\lambda}\,\mathrm{K}_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)};\quad x,\mu\in\mathbb{R}, \end{aligned} \tag{5.5}$$

where K_λ is a modified Bessel function of the third order with index λ

[5] A distribution of a random variable x is a *Normal variance-mean mixture* with a mixing distribution G, if x for a given $z \geq 0$ is normally distributed with mean $\mu + \beta z$ and variance z, given that z follows a probability distribution $G \in [0, \infty)$ [BNKS82].

and the integral representation is given by [Ebe01]:

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty \tau^{\lambda-1}\exp\left(-\frac{1}{2}x\left(\tau+\tau^{-1}\right)\right)d\tau.$$

The *gh* density function depends on five parameters, namely α, β, μ, δ and λ. The α determines the shape and is greater than 0, β in the interval $0 \leq |\beta| < \alpha$ determines the skewness and $\mu \in \mathbb{R}$ is the location parameter [Ebe01]. The scaling parameter is given by $\delta > 0$ and it is similar to the σ in the Normal distribution. Lastly, $\lambda \in \mathbb{R}$ determines the heaviness of the tails and it also describes certain sub-classes of the GH distribution. For $\lambda = -0.5$ we get the NIG distribution discussed in Chapter 4 and the Hyperbolic distribution is described for $\lambda = 1$ [Ebe01]. The Normal distribution is obtained as a limiting case of the GH distribution for $\delta \to \infty$ and $\delta/\alpha \to \sigma^2$. Unlike the Normal distribution defined by two parameters μ and σ, the class of Generalised Hyperbolic distributions is very flexible and therefore fits the financial returns in an optimal way [Ebe01].

Eberlein and Keller [EK95] proposed the Hyperbolic distribution in 1995 as a better fitting model for financial returns. Küchler et al fitted the Hyperbolic distribution to the daily returns of the German stocks in [KNSS99], in particular they demonstrated that skewness and kurtosis can be modelled much better than the Normal distribution. However, according to Barndorff-Nielsen [BN95] the Hyperbolic distribution lacks the property of being closed under *convolution*.

Later Barndorff-Nielsen extended on Eberlein and Keller's work by introducing the NIG distribution[6] as a realistic model for financial returns [BN95].

[6] A detailed discussion on the Normal Inverse Gaussian distribution is presented in Chapter 4 of this dissertation.

Barndorff-Nielsen found that the NIG distribution was a better fit to the daily empirical financial returns and is able to portray the heavy tails and skewness of financial returns [BN95]. Unlike the Hyperbolic distribution, the NIG distribution possesses an appealing property of being closed under convolution i.e. the sum of two independent random variables following an NIG distribution is again NIG distributed [BN95]. This property is important in forecasting application [HP07], for example time scaling of risk model such as VaR, when computing the 10-day VaR estimate from the daily VaR estimate.

The NIG distribution determines a process with stationary independent increments, which Barndorff-Nielsen in [BN95] refers to as the NIG processes. This type of process is known as the Lévy process, which is discussed in Chapter 4. Further reading on the implementation of the classes of generalised hyperbolic distributions in finance can be found in a number literatures. (See, for example Rydberg 1997 [Ryd97], Prause 1997 [Pra97], Bølviken and Benth 2000 [BB00], and Venter and Pieter [VdJ01]).

5.5 Extreme Value Theory

Extreme Value Theory (EVT) is the discipline of modelling the tails of a distribution. According to Hull [Hul10], EVT is a way of smoothing and extrapolating the tails of the probability distribution of daily returns calculated using historical simulation. Hull [Hul10] further explains that, EVT leads to VaR estimates that reflect the whole shape of the tail of the distribution, not just the positions of a few losses in the tails [Hul10]. Therefore, EVT can be used to assist in situations where risk managers want to estimate VaR with a very high confidence level [Hul10]. There are two types of approaches for Extreme Value Theory the *block maxima (minima)* model and *Peak Over Threshold* (POT) [MEF05].

The block maxima method involves modelling the largest observations collected from a large sample of identically distributed observations [MEF05]. The analysis involves partitioning identically independent random variables $(X_i)_{i\in\mathbb{N}}$ into intervals of equal length $j \in \mathbb{N}$. For each interval, the maximum observed value is obtained i.e. $BM_n = max\{X_1, X_2, \ldots, X_n\}$, BM_n is referred to as the block maxima. By the results of the *Fisher-Tippett, Gnedenko* [MEF05], the limiting distribution of the normalized maxima BM_n are in the Generalized Extreme Value (GEV) family, which includes the following distributions: Fréchet, Gumbel and Weibull [MEF05]. The density function of the GEV distribution is given by:

$$H_\zeta(x) = \begin{cases} \exp\left(-(1+\zeta\, x)^{1/\zeta}\right), & \zeta \neq 0, \\ \exp\left(-e^{-x}\right), & \zeta = 0, \end{cases}$$

where $1 + \zeta\, x > 0$ and ζ is the shape parameter [MEF05]. The type of distribution is defined by H_ζ depending on the values of ζ [MEF05]: when $\zeta > 0$, H_ζ is a Fréchet distribution, when $\zeta = 0$, H_ζ is a Gumbel distribution and when $\zeta < 0$, H_ζ is a Weibull distribution.

The POT model is more "modern" and "powerful", it models all large observations exceeding a certain threshold. McNeil *et al.*, in [MEF05] argue that the model is most efficient in using the data on extreme outcomes, which is often limited. Whereas the block maxima is very wasteful of data, since it models only the largest observation from a collection of large sample of identically distributed observations.

McNeil in [McN99], sub-classify the analysis within the POT model into two styles. The semi-parametric model based on the Hill estimator and the full-parametric model based on the Generalised Pareto distribution (GPD). The Generalised Pareto distribution is the probability distribution that models the exceedances over thresholds. Danielsson and Vries in [DdV00] work with the semi-parametric model to evaluate conditional VaR. In this section, we focus on the full-parametric model as described by McNeil in [McN99].

5.5.1 Generalised Pareto distribution

The distribution function of GPD is given by:

$$G_{\zeta,\eta}(x) = \begin{cases} 1 - (1 + \zeta x/\eta)^{-1/\zeta}, & \zeta \neq 0, \\ 1 - \exp(-x/\eta), & \zeta = 0, \end{cases} \tag{5.6}$$

where $\eta > 0$, and $x \geq 0$ when $\zeta \geq 0$ and $0 \leq x \leq -\eta/\zeta$ when $\zeta < 0$. The variable x represents daily negative log returns, which are converted to positive log returns. The GPD is depended on two parameters: ζ and η are referred to the shape and scale parameters respectively. The GPD, like the GEV includes a number of other distributions depending on the values of ζ: when $\zeta > 0$ the distribution function of GPD is that of an ordinary Pareto distribution, in this case the GPD is heavy-tailed. When $\zeta = 0$ the GPD function corresponds to the exponential distribution and when $\zeta < 0$ the distribution is short-tailed known as the Pareto type II distribution.

Definition 5.5.1. *[GR10] The distribution of the exceedances over a threshold u for a given random variable X with a distribution function F is defined by:*

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}. \tag{5.7}$$

for $0 \leq y < x_0 - u$, where $x_0 \leq \infty$ is the right endpoint of F [McN99].

Therefore, F_u describe the distribution of the losses above the threshold u, given that the threshold u has been exceeded or the probability that a loss exceeds the threshold by at most y, given that it exceeds the threshold [McN99]. In the POT model, the excess distribution $F_u(y)$ is best approximated by the Generalised Pareto distribution i.e.

$$F_u(y) = G_{\zeta,\eta}(y), \tag{5.8}$$

[GR10] see also [McN99].

5.5.2 The choice of threshold

The estimate of the parameters are dependent on the threshold u. The mean excess plot is used to determine a suitable threshold. In general if the data follows a Generalized Pareto distribution, then the plot will follow a reasonable straight line above a certain value of u [MEF05]. More specifically, a linear upward trend indicates a GPD model with a positive shape parameter ζ, a flat linear trend indicates a GPD model with a zero shape parameter or the exponential distribution and a linear downward trend indicates a GPD model with a negative shape parameter [MEF05].

When selecting the value of threshold it is important to note that it should be high enough for a better fit and not too high to result in an inadequate fit. On the other hand, the threshold value should not be too low, though this will result in a better fit but risks losing the extreme behaviour of the tail [Hub12] and [WM07]. Hull [Hul10] recommends choosing a threshold value close to the 95% percentile point of the empirical distribution.

Once the threshold u has been chosen, we now estimate the GPD parameters ζ and η using the maximum likelihood estimation method. In the maximum likelihood estimation method, the parameter values are chosen such that they maximize the joint probability density of the losses. Therefore, if we have that a total of N_u out of n losses have exceeded the threshold u, then the GPD is fitted to the N_u exceedances by estimating the parameters ζ and η.

5.5.3 Estimation of value-at-risk and expected shortfall

By setting $x = y + u$ in Equation (5.7), an approximation of $F(x)$ where $F_u(y) = G_{\zeta,\eta}(y)$ for $x > u$ is given by [McN99]

$$F(x) = (1 - F(u))G_{\zeta,\eta}(y) + F(u), \tag{5.9}$$

and $F(u)$ is estimated by using the empirical cumulative distribution function $(n - N_u)/n$.

By substituting Equation (5.6) and the estimate of $F(u)$ into Equation (5.9), we obtain the following estimate for $F(x)$

$$\widehat{F(x)} = 1 - \frac{N_u}{n}\left(1 + \widehat{\zeta}\,\frac{x-u}{\widehat{\eta}}\right)^{(-1/\widehat{\zeta})}, \tag{5.10}$$

where $x > u$ [McN99].

For a given probability $q > F(u)$ the Value-at-Risk estimate is obtained by solving for x in Equation (5.10) [McN99]

$$\mathrm{VaR}_q = u + \frac{\widehat{\eta}}{\widehat{\zeta}}\left(\left(\frac{n}{N_u}(1-q)\right)^{-\widehat{\zeta}} - 1\right). \tag{5.11}$$

Assuming that $\zeta < 1$ the corresponding expected shortfall is estimated as:

$$\mathrm{ES}_q = \frac{1}{1-q}\int_q^1 \mathrm{VaR}_x dx = \frac{\mathrm{VaR}_q}{1-\zeta} + \frac{\eta - \zeta u}{1-\zeta}. \tag{5.12}$$

5.6 Chapter Summary

In this chapter, we discussed some of the distributions that have been used to model financial returns in literature. We have noted how each of the discussed distributions fit financial returns better than the Normal distribution. We also noted how these distributions compare to the NIG distribution. For example, though the t-distribution has heavy tails it does not accommodate the skewness found in the financial returns, as it is centered about zero. The Hyperbolic distribution lacks the property of being closed under convolution, making it not useable when modelling a portfolio of financial assets.

When focusing on the tails of the probability distribution of financial returns, it is natural to consider the Extreme Value Theory. The inclusion of this Extreme Value Theory is so that we can compare Value-at-Risk values estimated by just focusing on modelling the tails of the distribution to modelling the entire returns distribution.

The objective was to provide the reader with the literature review of some of the distributions used to model financial returns and compare these distributions to the NIG so that to justify our choice of the NIG. We also provided basic background of concepts of the Extreme Value Theory, including how Value-at-Risk is estimated under the EVT assumptions.

Chapter 6

Distribution fitting and evaluating Value-at-Risk

In this chapter, we fit the NIG model to the empirical data described in Chapter 2 and compare the fit of the NIG to the fit of the Skew t, Normal and t-distribution. We then model the tails for the empirical data using the Generalised Pareto distribution and this is then followed by evaluate VaR under the assumption that the empirical data follows the NIG distribution. We compare the VaR estimates to those under the Normal, Skew t, t-distribution and the Generalised Pareto distribution in the final section.

6.1 Distribution fitting

Fitting the NIG distribution is straightforward using the statistical program R with the package fBasisc [De06] because the program has a predefined function for the maximum likelihood estimation (MLE) of the NIG distribution.

Table 6.1 shows the maximum likelihood parameter estimates results for the fitted NIG distribution. According to the NIG parameters in Table 6.1

FTSE/JSE Top40, S&P 500 and African Bank display negative skewness, where $\widehat{\beta}$ is -5.1090, -5.0934 and -0.6613 respectively. The skewness of Standard Bank, Merafe, Grindrod and Anglo American is positive, which indicates that log returns above zero occur more frequently than negative log returns. The location parameter $\widehat{\mu}$ for each stock and the indices is very small as noted from the empirical statistics in Chapter 2. The tail parameter $\widehat{\alpha}$ is in the range 30-69 for each stock and the indices displaying a higher peak and heavy tails.

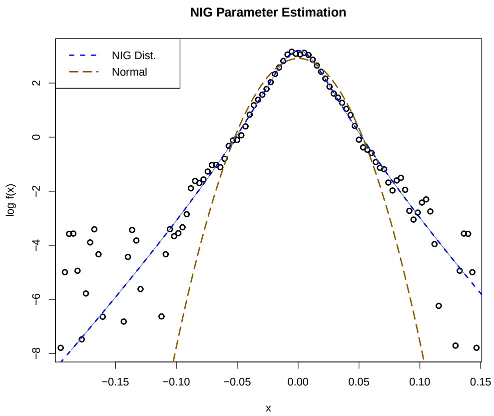
Estimated parameters of the NIG				
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	$\widehat{\mu}$
FTSE/JSE TOP 40	67.8875	-5.1090	0.0123	0.0014
S&P 500	58.3541	-5.0934	0.0075	0.0009
Standard Bank	44.7345	0.3524	0.0208	0.0003
African Bank	32.3720	-0.6613	0.0250	0.0005
Anglo American	43.1323	0.9557	0.0270	-0.0002
Merafe Resources	41.3981	4.5965	0.0527	-0.0054
Grindrod Limited	16.8123	0.4198	0.0185	0.0008

Table 6.1: Maximum likelihood parameter estimates resulting when fitting the NIG distribution.

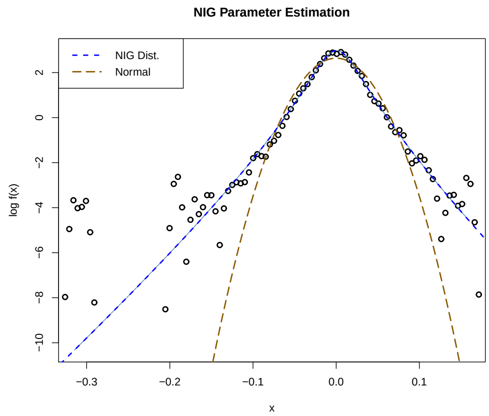
Statistical data of the fitted Normal Inverse Gaussian distribution				
	Mean	Variance	Skewness	Kurtosis
FTSE/JSE TOP 40	0.0005	0.0002	-0.2474	3.684585
S&P 500	0.0002	0.0001	-0.3966	7.090656
Standard Bank	0.0005	0.0005	0.0245	3.225051
African Bank	0.0000	0.0008	-0.0681	3.71387
Anglo American	0.0004	0.0006	0.0616	2.581746
Merafe Resources	0.0005	0.0013	0.2262	1.451873
Grindrod Limited	0.0013	0.0011	0.1343	9.672519

Table 6.2: Statistical data of the fitted NIG distribution.

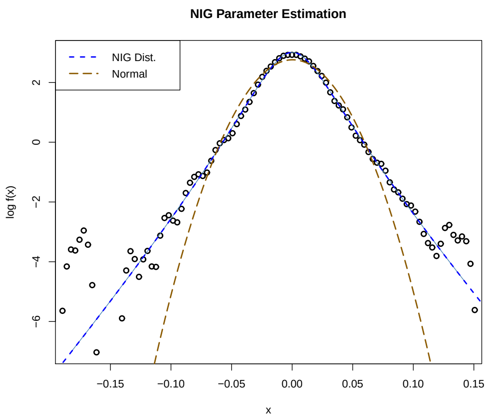
The statistical data estimated using the NIG distribution is displayed in Table 6.2. The plots of the fitted NIG log-density based on the estimated parameters from Table 6.1 together with the empirical log-density of each stock and the indices are compared to the Normal distribution and are displayed in Figure 6.1. Figure 6.1 shows how the Normal distribution has a shape of a parabola and how it deviates from the tails of the empirical log-density. Though the tails of the empirical log-density are heavier than those of the NIG log-density, the approximation in the center is very good.



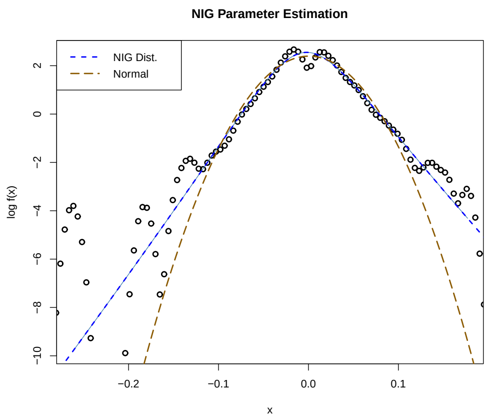
(a) Standard Bank



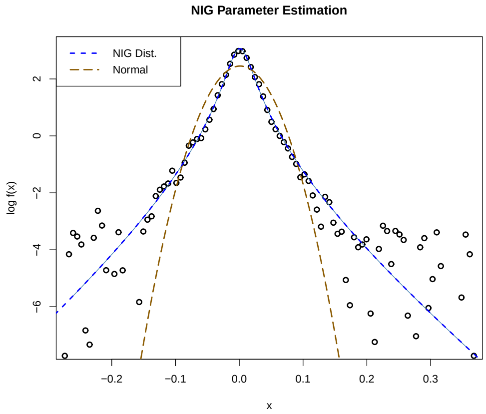
(b) African Bank



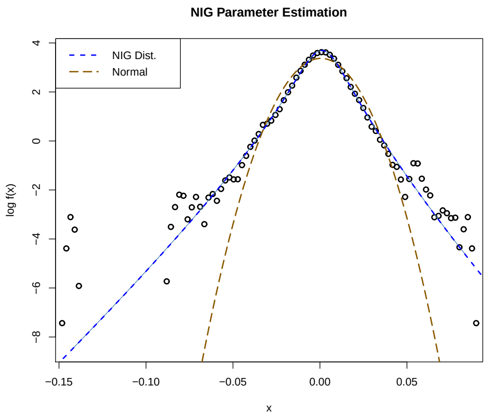
(c) Anglo American



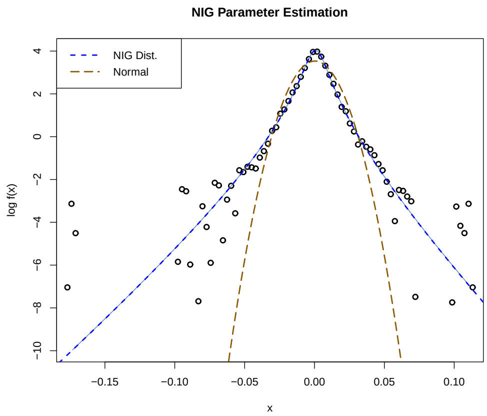
(d) Merafe Resources



(e) Grindrod Limited



(f) FTSE/JSE TOP40



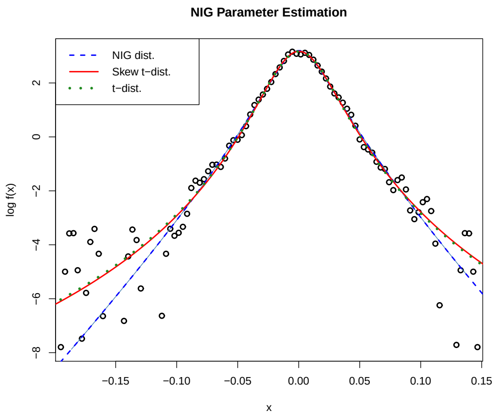
(g) S&P 500

Figure 6.1: The log-density of the empirical data with the fitted NIG and Normal distribution.

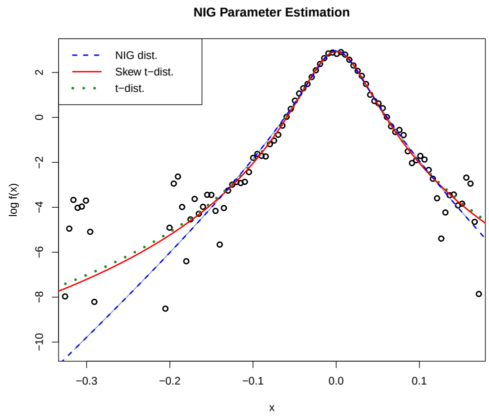
Table 6.3 shows the parameter estimates resulting from fitting the t-distribution and Skew t. The tail parameter or degree of freedom is in the range of 2.4-6.3 for each stock and indices. Figure 6.2 compares the log-density of the NIG, Skew t and t-distribution to the empirical data, the overall observation is that the Skew t and t-distribution have fatter tailed compared to the NIG distribution. In addition, the NIG, Skew t and t-distribution seem to adequately fit the center of the empirical distribution quite well compared to the Normal distribution.

Estimated parameters of the Skew t and t-distribution							
	t-distribution parameters			Skew t parameters			
	mean	std dev.	k	mean	std dev.	k	β
FTSE/JSE TOP 40	0.0007	0.0138	4.00	0.0005	0.0139	3.84	0.95
S&P 500	0.0006	0.0129	2.80	0.0003	0.0129	2.80	0.95
Standard Bank	0.0005	0.0222	4.00	0.0006	0.0222	4.00	1.01
African Bank	0.0001	0.0285	3.84	0.0001	0.0282	4.00	1.00
Anglo American	0.0003	0.0254	4.61	0.0004	0.0254	4.61	1.02
Merafe Resources	0.0000	0.0361	6.52	0.0008	0.0363	6.26	1.15
Grindrod Limited	0.0010	0.0430	2.41	0.0012	0.0430	2.41	1.01

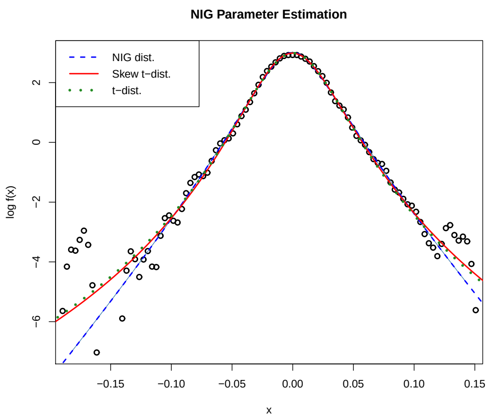
Table 6.3: Parameter estimations for the Skew t and t-distribution where k is the degrees of freedom and β is the skewness parameter.



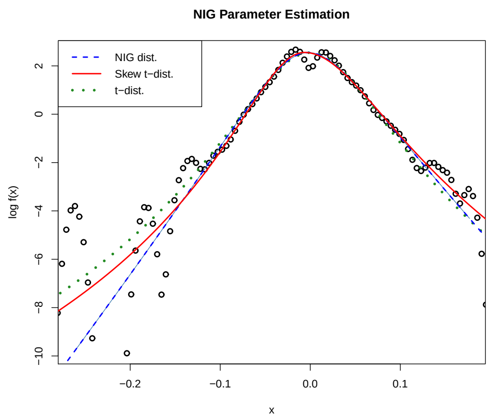
(a) Standard Bank



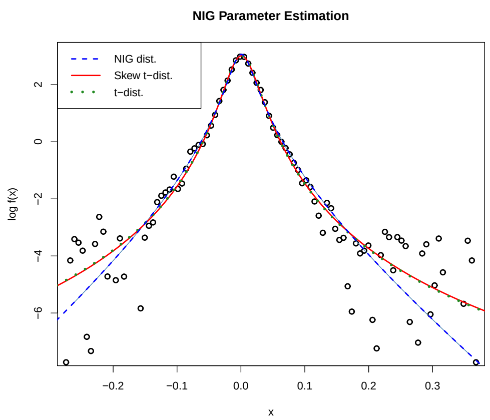
(b) African Bank



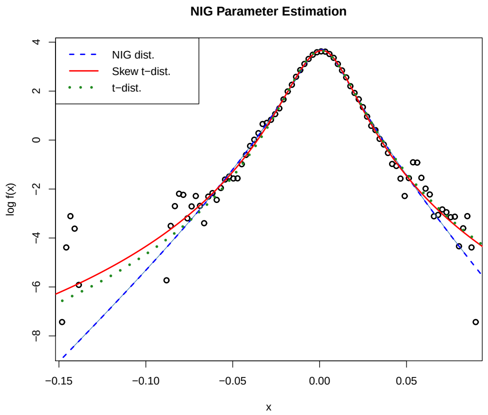
(c) Anglo American



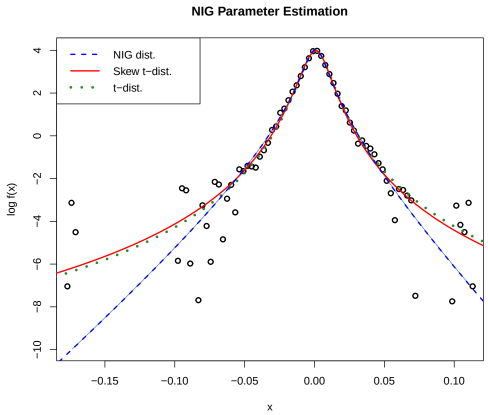
(d) Merafe Resources



(e) Grindrod Limited



(f) FTSE/JSE TOP40

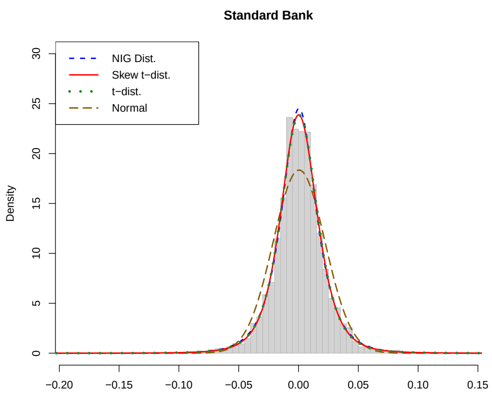


(g) S&P 500

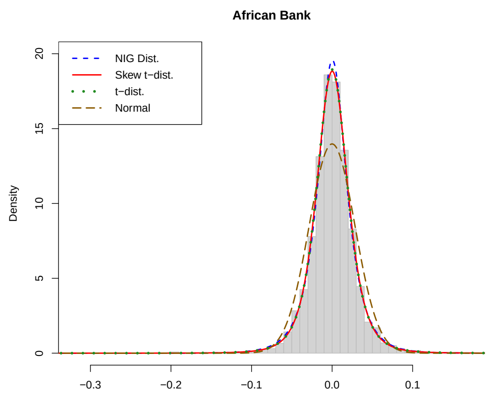
Figure 6.2: The log-density of the empirical data with the fitted NIG, Skew t and the t-distribution.

From Figure 6.3, it is observed that the NIG distribution fits the empirical histogram better than that of the Normal distribution and t-distribution. The NIG distribution also matches the empirical distribution better with respect to the skewness and the peak compared to the Normal distribution, while the t-distribution show a higher peak and fatter tails. The maximum likelihood estimated parameters are given in table 6.1.

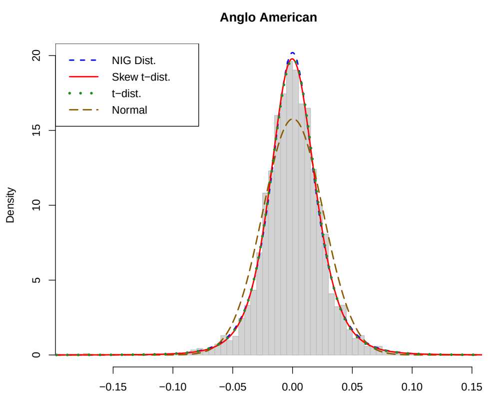
The fatter tails of the t-distribution is more prevalent in the zoomed left-tail of the distributions in Figure 6.4. Though the NIG distribution fits the data much better, financial institutions are mostly concerned with the extreme movements of the market as this could lead to huge losses. The extreme movements resulting in losses occur at the left-tails of the distribution. Hence, the need to study the tails distribution of the empirical data and this is done through the Extreme Value Theory (EVT).



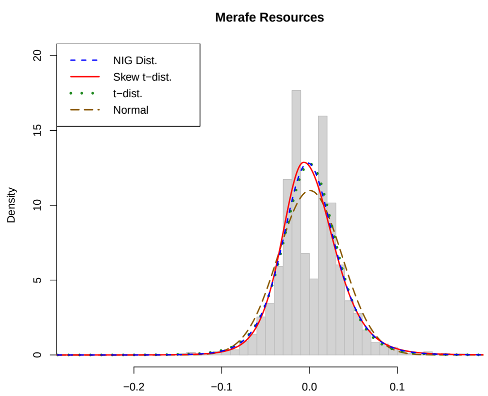
(a) Standard Bank



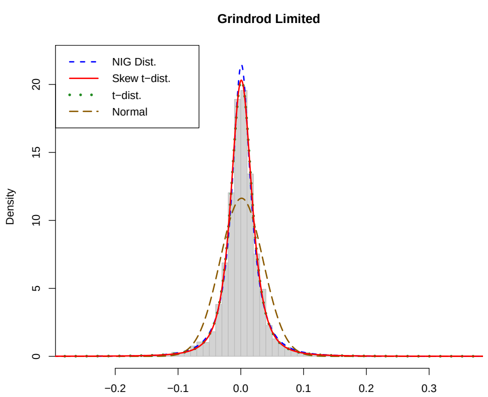
(b) African Bank



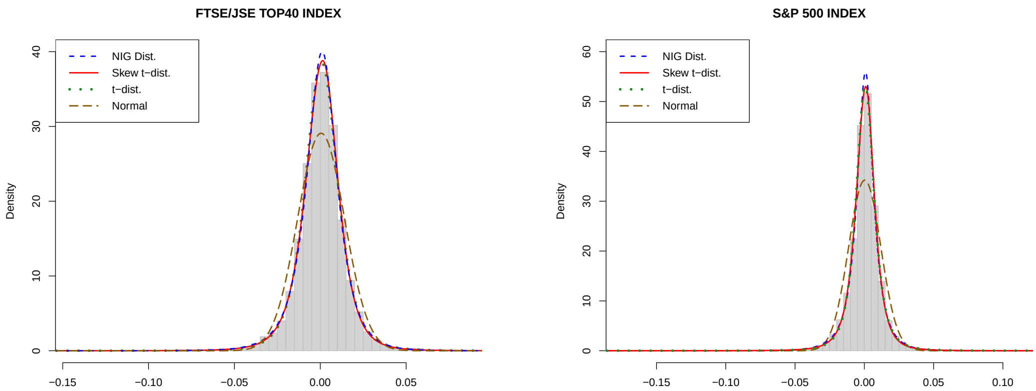
(c) Anglo American



(d) Merafe Resources

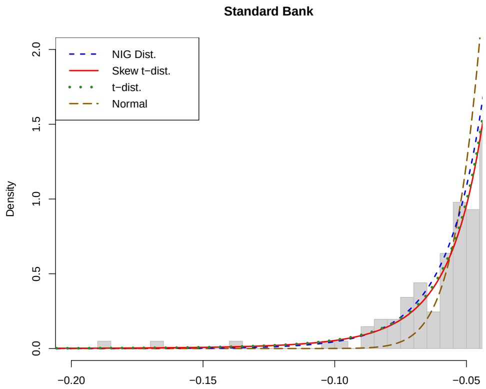


(e) Grindrod Limited

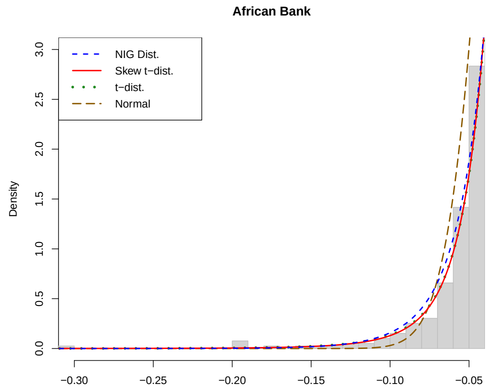


(f) FTSE/JSE TOP40 (g) S&P 500

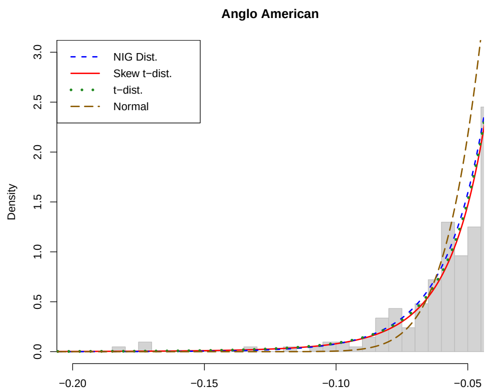
Figure 6.3: Comparison of each stock and the indices histograms with the fitted NIG, Skew t, t-distribution and Normal distribution.



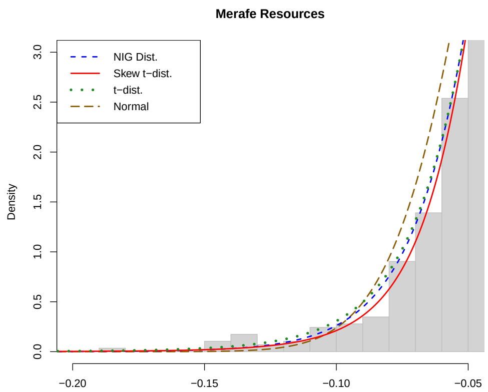
(a) Standard Bank



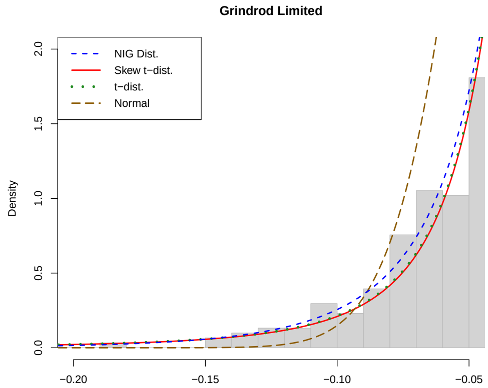
(b) African Bank



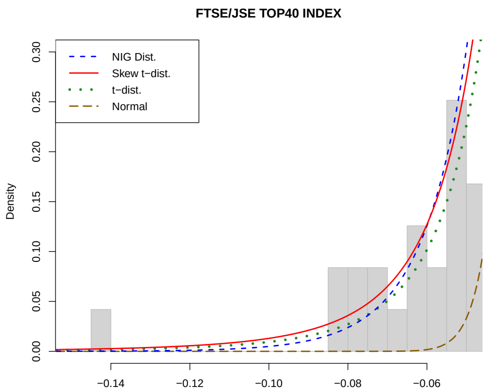
(c) Anglo American



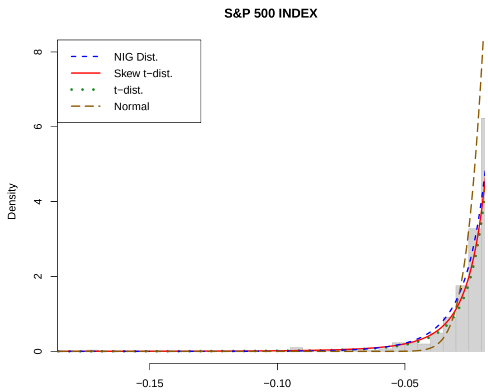
(d) Merafe Resources



(e) Grindrod Limited



(f) FTSE/JSE TOP40



(g) S&P 500

Figure 6.4: Comparison of each stock and the indices left-tail of the histograms with the NIG, Skew t, Normal and t-distribution.

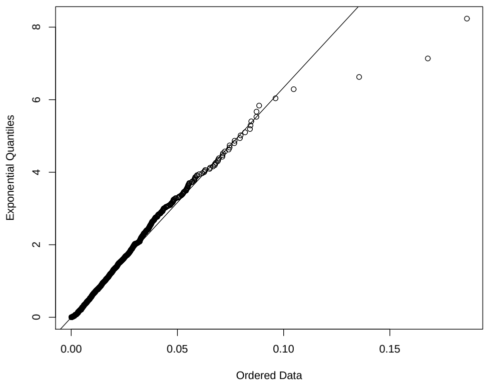
6.2 Extreme Value Theory

To explore the heavy left-tail of the empirical distribution, the statistical program R is used with the Extreme Value package in R (EVIR) [PMS12]. The aim is to model the behaviour of positive losses that exceed a pre-defined threshold with the Generalized Pareto distribution [McN99]. This method is called the Peaks over Threshold (POT), described in Chapter 5.

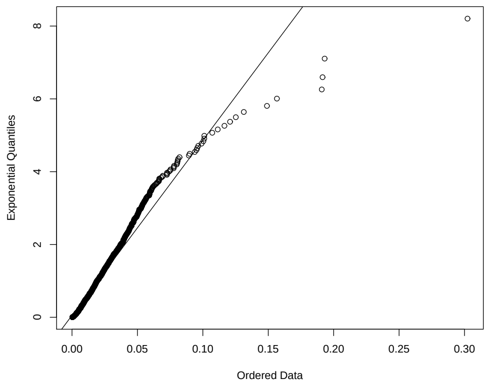
The first step of the analysis commences with the Q-Q plot of exponential distribution against the left tail of the empirical distributions. Some of the empirical quantiles deviate from the exponential quantile line, this suggests that the empirical data exhibits heavier tails than that of the exponential distribution see Figure 6.5.

The mean excess plot of the positive losses in Figure 6.6 is used to find the threshold, which we denote by u. The chosen thresholds are listed in Table 6.4, losses above these thresholds are modelled by the Generalized Pareto distribution and the estimated parameters are also listed in Table 6.4.

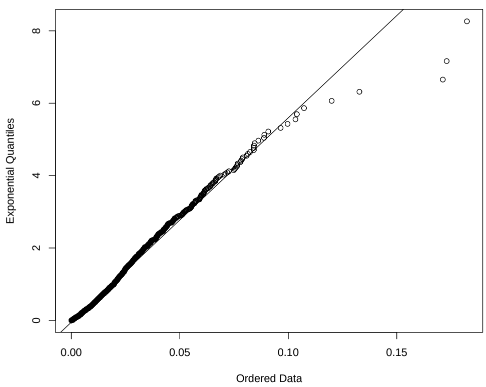
If the mean excess plot has a linear upward trend, it indicates a GPD model with a positive shape parameter ζ. A flat linear trend indicates a GPD model with a zero shape parameter or the exponential distribution. A linear downward trend indicates a GPD model with a negative shape parameter [MEF05]. Hull [Hul10] recommends choosing a threshold value close to the 95% percentile point of the empirical distribution. The horizontal line in the mean excess plot indicates the 95% percentile point of the empirical distribution.



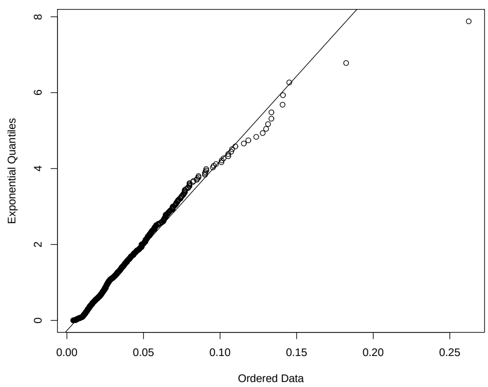
(a) Standard Bank



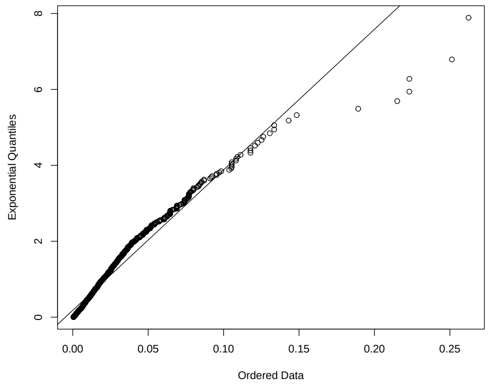
(b) African Bank



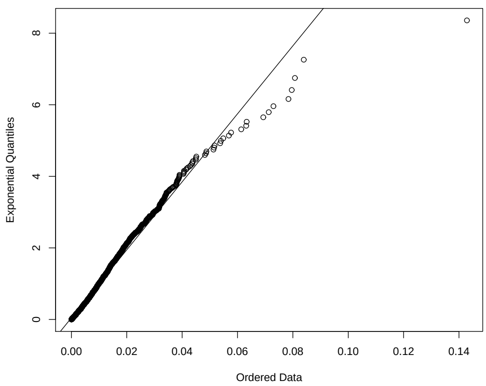
(c) Anglo American



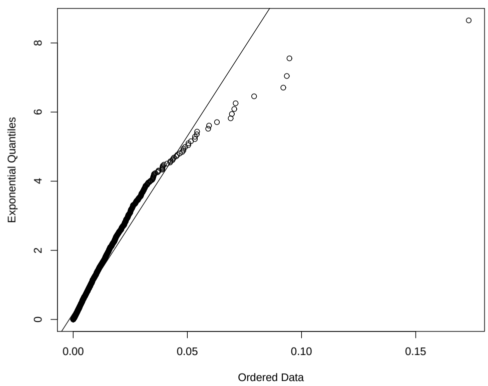
(d) Merafe Resources



(e) Grindrod Limited

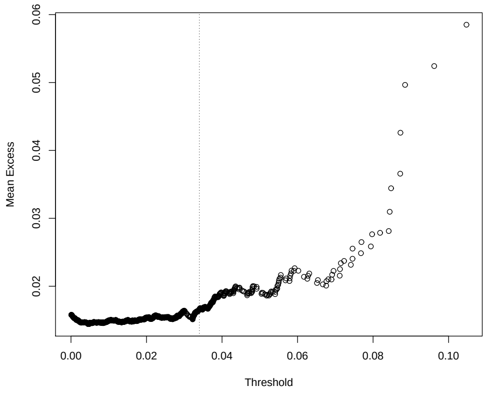


(f) FTSE/JSE TOP40

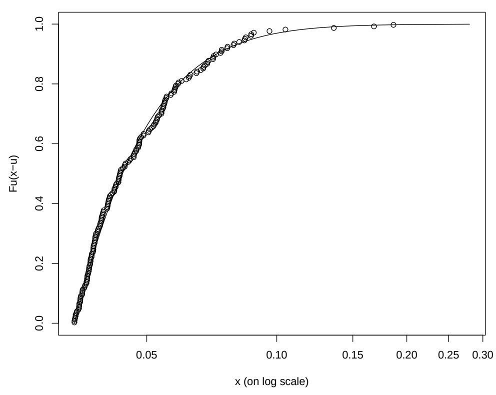


(g) S&P 500

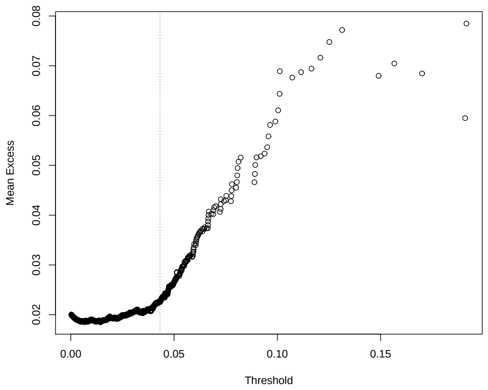
Figure 6.5: Q-Q plot against the exponential distribution.



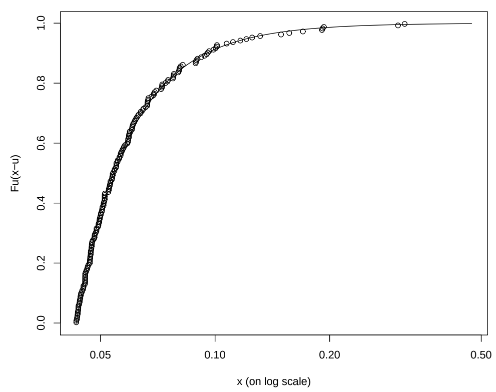
(a) Standard Bank



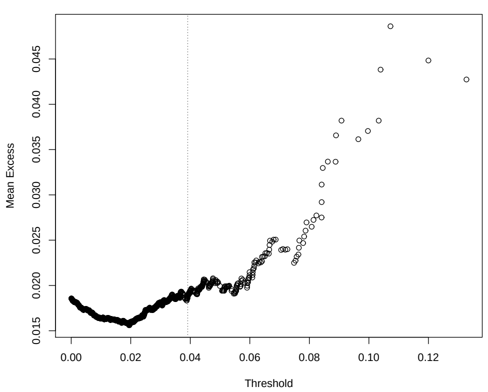
(b) Standard Bank



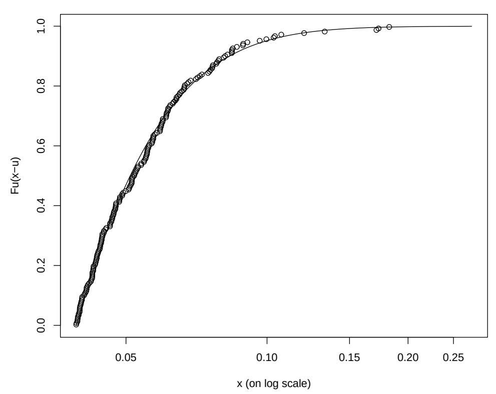
(a) African Bank



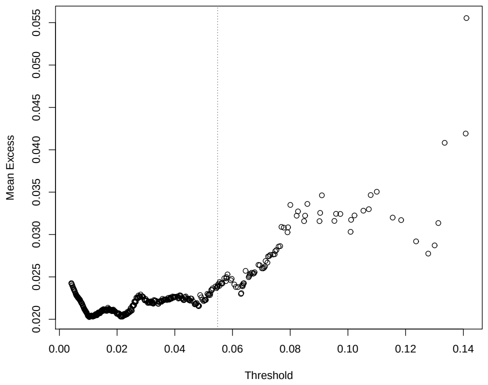
(b) African Bank



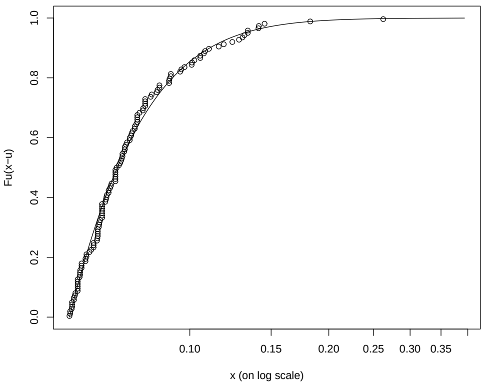
(a) Anglo American



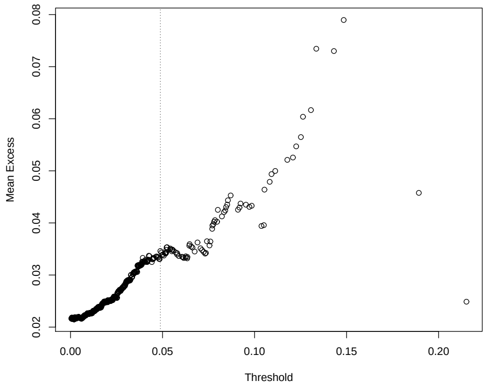
(b) Anglo American



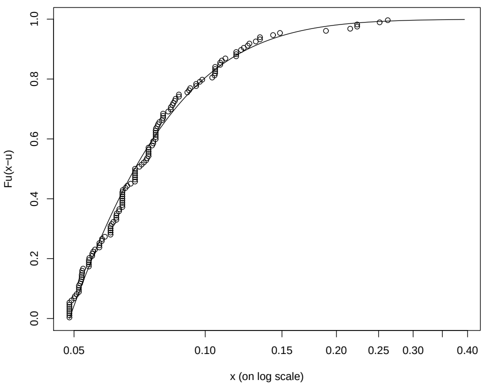
(a) Merafe Resources



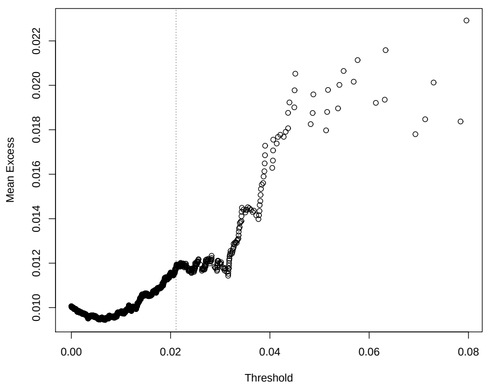
(b) Merafe Resources



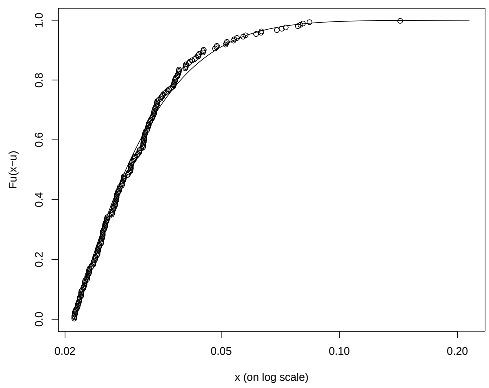
(a) Grindrod Limited



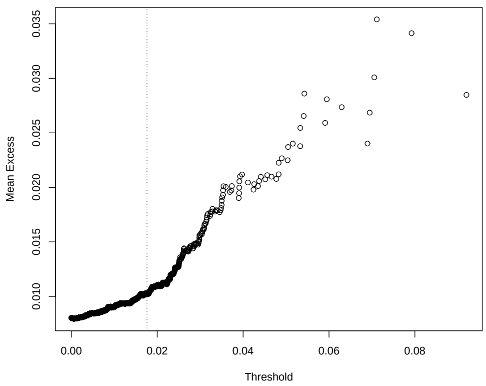
(b) Grindrod Limited



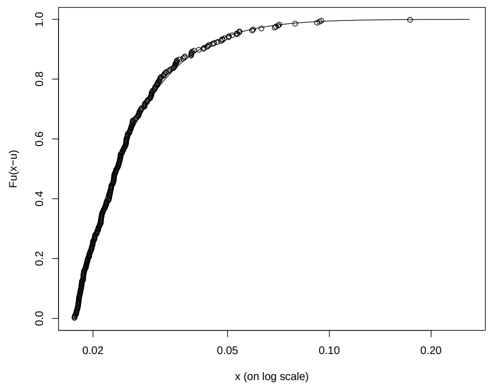
(a) FTSE/JSE TOP40



(b) FTSE/JSE TOP40



(a) S&P 500



(b) S&P 500

Figure 6.6: Figure (a) shows the mean excess plot and (b) shows the empirical distribution of excess and GPD fit.

Estimated parameters of the Generalized Pareto distribution					
	Threshold	$\widehat{\zeta}$	$\widehat{\eta}$	n	N
FTSE/JSE TOP 40	0.0211	0.1160	0.008	2252	229
S&P 500	0.0176	0.2650	0.004	2864	310
Standard Bank	0.0334	0.2060	0.008	2014	204
African Bank	0.0432	0.4060	0.005	1960	197
Anglo American	0.0387	0.1700	0.010	2070	209
Merafe Resources	0.0541	0.1420	0.015	1418	138
Grindrod Limited	0.0465	0.1760	0.018	1454	155

Table 6.4: Maximum likelihood parameter estimators of the Generalized Pareto distribution for each stock and the indices, where n is the total number of daily negative log returns and N is number of threshold exceedances.

6.3 Risk Measure

In this section, we present the comparison of Value-at-Risk estimates under the NIG, Skew t, Normal, t-distribution and Generalised Pareto distributions to the empirical distribution of the stocks and indices described in Chapter 2. We further verify the correctness of the VaR models using the backtesting technique discussed in Chapter 3.

6.3.1 Value-at-Risk estimates

The VaR and ES estimates obtained on 31 July 2014 under the NIG, Skew t, Normal, t-distribution and EVT assumption for a one-day holding period at 99% confidence level are shown in Table 6.5 and Table 6.6 respectively. The VaR and ES estimates under the NIG, Normal and t-distribution assumption were calculated using the Monte Carlo simulation as detailed in Chapter 3. Under the EVT assumption, the VaR estimates were calculated using Equation (5.11) and the ES estimates were obtained using Equation (5.12).

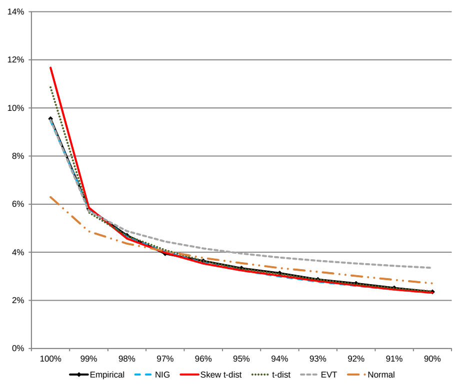
Figure 6.7 shows VaR estimates for a one-day holding period with respect to different confidence levels. The higher the confidence level, the higher the VaR and ES estimates are under the NIG, Skew t, t-distribution and EVT assumption compared to the Normal distribution.

Estimates of one-day VaR at a 99% confidence level on 31 July 2014.						
	Non-parametric	NIG	Skew t	t-dist.	EVT	Normal
FTSE/JSE TOP 40	3.79%	3.75%	3.84%	3.56%	4.21%	3.11%
S&P 500	3.15%	3.31%	3.45%	3.04%	3.14%	2.68%
Standard Bank	5.81%	5.81%	5.86%	5.65%	5.71%	4.87%
African Bank	7.35%	7.53%	7.54%	7.73%	6.82%	6.37%
Anglo American	6.60%	6.38%	6.46%	6.73%	6.76%	5.97%
Merafe Resources	9.02%	8.98%	8.13%	9.36%	9.39%	8.36%
Grindrod Limited	9.53%	9.82%	9.62%	10.14%	9.88%	7.73%

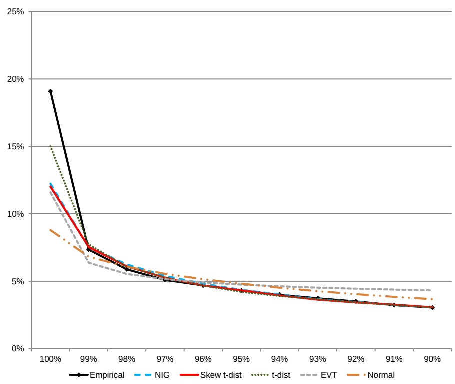
Table 6.5: Comparison of the Value-at-Risk estimates, the non-parametric estimates are calculated using the Historical Simulation approach.

Estimates of one-day Expected Shortfall at a 99% confidence level on 31 July 2014.						
	Non-parametric	NIG	Skew t	t-dist.	EVT	Normal
FTSE/JSE TOP 40	5.20%	4.96%	5.41%	4.89%	5.38%	3.50%
S&P 500	4.82%	4.69%	5.48%	4.88%	4.21%	3.01%
Standard Bank	7.96%	7.49%	8.03%	7.81%	7.33%	5.55%
African Bank	11.58%	9.62%	9.73%	11.06%	8.68%	7.76%
Anglo American	8.93%	8.33%	8.28%	8.71%	8.58%	6.76%
Merafe Resources	12.22%	10.96%	10.28%	11.70%	11.77%	9.55%
Grindrod Limited	13.61%	13.53%	16.99%	20.26%	13.15%	9.18%

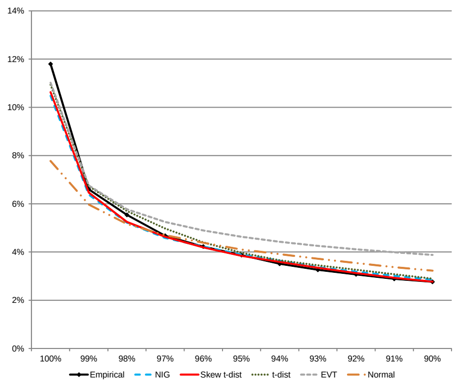
Table 6.6: Comparison of the Expected Shortfall estimates, the non-parametric estimates are calculated using the Historical Simulation approach.



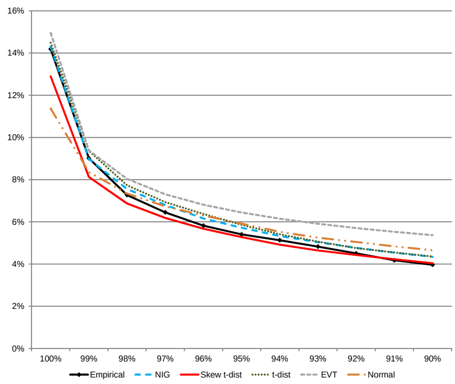
(a) Standard Bank



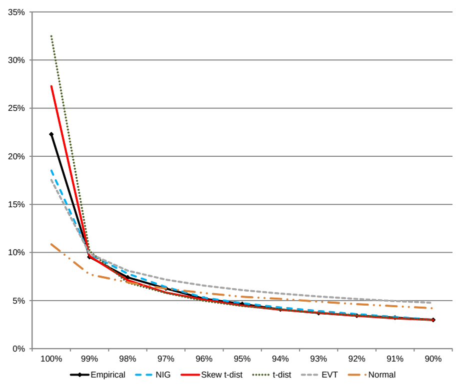
(b) African Bank



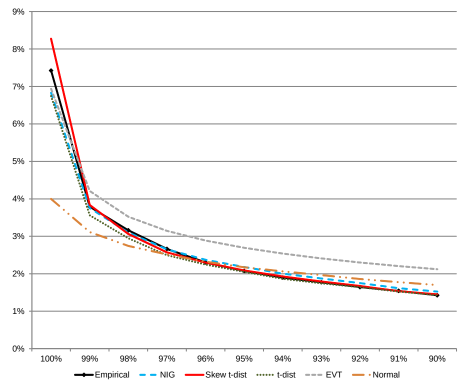
(c) Anglo American



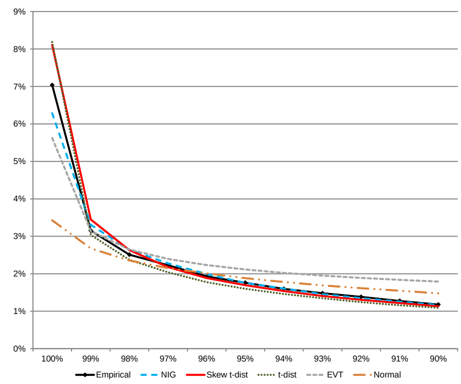
(d) Merafe Resources



(e) Grindrod Limited



(f) FTSE/JSE TOP40



(g) S&P 500

Figure 6.7: Value-at-Risk estimates at different confidence level.

Backtesting the model

To verify the correctness of the VaR models we perform backtesting method described in Section 3.7 of Chapter 3. We compare the VaR estimates obtained on the 31 July 2014 to the actual observed returns over the period 1 August 2013 to 31 July 2014, the results are presented in Table 6.7. We recorded the number of times the VaR estimates on 31 July 2014 exceeded the observed returns over the period 1 August 2013 to 31 July 2014 and classified the results into the three zones defined by the Basel Committee. For Example the results in Table 6.7 show that African Bank is the only stock that had a higher number of exceptions with the VaR estimates for NIG, Skew t and t-distribution recording 6 violations. Under the same example the VaR estimate for EVT recorded 9 violation and the Normal distribution recorded 11.

In Table 6.8 and Table 6.9 we show the likelihood of decreases on the given interval likely to be realised once every number of years, calculated under the NIG, Normal, Skew t, t-distribution and EVT assumptions of decrease in a given interval. The results in Table 6.8 were calculated by fitting the NIG, Normal, Skew t, t-distribution and EVT over the period 1 September 1997 to 31 July 2014 for Standard Bank, Grindrod and Anglo American. For African Bank the period is 29 September 1997 to 31 July 2014, Merafe the period is 17 December 1999 to 31 July 2014, for the FTSE/JSE Top 40 the period is 30 June 1995 to 31 July 2014 and S&P 500 is 2 January 1991 to 31 July 2014. These results under the NIG, Normal, Skew t, t-distribution and EVT assumptions are compared to the actual observed returns. For example in the case of Standard Bank a decrease in the interval 8.75% to 10% was observed once every 2.7 years over the period 1 September 1997 to 31 July 2014, the NIG estimates that such decrease would occur once every 4.7 years, the Skew t and t-distribution estimates 4.3 years and 4.1 years respectively. The EVT estimates the losses in the interval 8.75% to 10% to occur once every 3.3 years, while the Normal distribution estimates the losses to occur

Backtesting results for 99% daily-VaR over the most recent 250 days of our data.							
		Non parametric	NIG	Skew t	t-dist	EVT	Normal
FTSE/JSE Top40	No. of ex	0	0	0	0	0	0
	Zone	Green	Green	Green	Green	Green	Green
S&P 500	No. of ex	1	1	1	1	1	1
	Zone	Green	Green	Green	Green	Green	Green
Standard Bank	No. of ex	0	0	0	0	0	0
	Zone	Green	Green	Green	Green	Green	Green
African Bank	No. of ex	6	6	6	6	9	11
	Zone	Yellow	Yellow	Yellow	Yellow	Yellow	Red
Anglo American	No. of ex	0	0	0	0	0	0
	Zone	Green	Green	Green	Green	Green	Green
Merafe Resource	No. of ex	0	0	0	0	0	0
	Zone	Green	Green	Green	Green	Green	Green
Grindrod Limited	No. of ex	0	0	0	0	0	0
	Zone	Green	Green	Green	Green	Green	Green

Table 6.7: Backtesting results for one-day VaR at 99% a confidence level over the most recent 250 days of our data.

once every 136 years.

Table 6.9 shows the likelihood of decreases on the given interval likely to be realised once every number of years by fitting the NIG, Normal, Skew t, t-distribution and EVT using the empirical data over the first half of the original period for each stock and indices. For example, African Bank the distributions were fitted using data over the period 29 September 1997 to 29 June 2006. These results were compared to the actual observed returns over the second half of the data, i.e. 30 June 2006 to 31 July 2014 for African Bank. The results in Table 6.9 shows that the losses in the interval 15% to 20% for African Bank were realised once every 2.6 years over the period 30 June 2006 to 31 July 2014. The NIG predicts that losses in the interval 15% to 20% will occur once every 8.3 years, the Skew t and t-distribution predicts

Interval (decreases)	The likelihood of decreases on the given interval likely to be realised once every number of years for the period ending 31 July 2014					
FTSE/JSE Top40	Observed	NIG	Skew t	t-dist.	EVT	Normal
0 to 2.5%	0.0084	0.0091	0.0091	0.0090	0.0043	0.0084
2.5% to 5%	0.1176	0.1267	0.1411	0.1564	0.0676	0.1010
5% to 8.75%	0.9962	1.1863	1.1271	1.4119	0.5058	20.2065
8.75% to 10%	18.9286	38.2812	17.8197	24.2281	10.1798	1.68E+07
10% to 15%	18.9286	63.0238	14.3920	20.4024	10.7853	7.44E+09
15% to 20%	-	2.51E+03	76.3979	115.7666	109.5912	-
S&P 500	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0109	0.0089	0.0089	0.0088	0.0041	0.0081
2.5% to 5%	0.1521	0.1905	0.2246	0.2507	0.1729	0.2335
5% to 8.75%	1.2886	1.5236	1.3375	1.5406	1.0761	396.4572
8.75% to 10%	24.4841	37.6071	15.2851	17.7828	13.8078	1.08E+11
10% to 15%	24.4841	51.8338	10.0172	11.6906	10.0371	-
15% to 20%	-	1.30E+03	37.5097	4.39E+01	45.8336	-
Standard Bank	Observed	NIG	Skew t	t-dist.	EVT	Normal
0 to 2.5%	0.0081	0.0100	0.0099	0.0100	0.0044	0.0105
2.5% to 5%	0.0434	0.0524	0.0543	0.0536	0.0556	0.0335
5% to 8.75%	0.2319	0.2650	0.2988	0.2875	0.2699	0.3495
8.75% to 10%	2.7050	4.6775	4.3308	4.1112	3.3265	136.4962
10% to 15%	4.0575	5.3219	3.4828	3.2886	2.4932	1.68E+03
15% to 20%	8.1151	85.8005	18.8816	17.7106	12.6540	-
African Bank	Observed	NIG	Skew t	t-dist.	EVT	Normal
0 to 2.5%	0.0080	0.0111	0.0110	0.0110	0.0043	0.0125
2.5% to 5%	0.0293	0.0387	0.0378	0.0381	0.0763	0.0269
5% to 8.75%	0.1252	0.1285	0.1475	0.1465	0.2742	0.1140
8.75% to 10%	0.6260	1.4867	1.8286	1.7623	2.3536	6.1621
10% to 15%	1.0434	1.2498	1.3973	1.3133	1.3237	25.6393
15% to 20%	3.1302	10.2350	7.1777	6.4388	4.0124	1.26E+05

that losses for the same interval will occur once every 5.7 years and 4.7 years respectively.

Interval (decreases)	The likelihood of decreases on the given interval likely to be realised once every number of years for the period ending 31 July 2014					
Anglo American	Observed	NIG	Skew t	t-dist.	EVT	Normal
0 to 2.5%	0.0080	0.0107	0.0106	0.0107	0.0046	0.0116
2.5% to 5%	0.0341	0.0404	0.0405	0.0402	0.0394	0.0283
5% to 8.75%	0.1544	0.1731	0.1916	0.1836	0.1707	0.1614
8.75% to 10%	1.2705	2.7994	2.8792	2.6915	2.0719	16.2080
10% to 15%	2.0645	3.0635	2.4844	2.2990	1.5932	97.0204
15% to 20%	5.5053	45.9263	15.7555	14.3929	8.8006	2.41E+06
Merafe Resources	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0079	0.0137	0.0132	0.0140	0.0051	0.0157
2.5% to 5%	0.0184	0.0265	0.0255	0.0272	0.0253	0.0242
5% to 8.75%	0.0626	0.0679	0.0754	0.0667	0.0799	0.0494
8.75% to 10%	0.3840	0.8528	1.0562	0.7564	0.8083	0.7074
10% to 15%	0.5302	0.8699	1.0229	0.6556	0.5760	1.1668
15% to 20%	11.1349	12.3783	9.2258	5.0981	2.9498	164.3330
Grindrod	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0083	0.0115	0.0113	0.0114	0.0056	0.0147
2.5% to 5%	0.0303	0.0457	0.0451	0.0448	0.0217	0.0238
5% to 8.75%	0.0943	0.1218	0.1411	0.1389	0.0543	0.0553
8.75% to 10%	0.4024	0.9978	1.2265	1.2014	0.4485	0.9991
10% to 15%	0.5248	0.6084	0.7063	0.6904	0.2752	1.9909
15% to 20%	12.0714	2.4948	2.2211	2.1661	1.0823	545.4484

Table 6.8: Comparison between observed likelihood of decreases on the given interval likely to occur once every number of years to the likelihoods calculated under the EVT, Skew t, t-distribution, NIG and Normal distribution for the ending 31 July 2014.

Interval (decreases)	The likelihood of decreases on the given interval likely to be realised once every number of years for the period ending 31 July 2014, using half the original data.					
FTSE/JSE Top40	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0086	0.0088	0.0087	0.0087	0.0213	0.0083
2.5% to 5%	0.0977	0.1423	0.1769	0.1664	0.0263	0.1221
5% to 8.75%	1.0525	1.6164	1.6920	1.5448	0.0227	39.1579
8.75% to 10%	-	65.4568	29.6196	26.6067	0.0841	1.16E+08
10% to 15%	-	122.7782	25.0948	22.3691	0.0271	9.14E+10
15% to 20%	-	6.90E+03	143.3808	126.2540	0.0408	-
S&P 500	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0112	0.0086	0.0085	0.0085	0.0458	0.0079
2.5% to 5%	0.1262	0.2335	0.2824	0.2927	0.0502	0.3526
5% to 8.75%	1.3598	3.1520	3.0754	3.2051	0.0375	1 826.47
8.75% to 10%	-	148.6019	55.9240	58.3445	0.1235	-
10% to 15%	-	301.0705	47.8661	4.99E+01	0.0346	-
15% to 20%	-	2.07E+04	276.4179	2.88E+02	0.0415	-
Standard Bank	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0080	0.0106	0.0104	0.0105	0.0468	0.0111
2.5% to 5%	0.0511	0.0460	0.0476	0.0473	0.0513	0.0293
5% to 8.75%	0.3249	0.2001	0.2374	0.2225	0.0383	0.2003
8.75% to 10%	8.1230	3.1012	3.4007	2.9952	0.1261	30.1740
10% to 15%	8.1230	3.2491	2.7744	2.3480	0.0353	226.2969
15% to 20%	-	43.5798	15.6583	12.3680	0.0422	-
African Bank	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0080	0.0116	0.0114	0.0117	0.0594	0.0135
2.5% to 5%	0.0303	0.0362	0.0361	0.0355	0.0638	0.0247
5% to 8.75%	0.1565	0.1133	0.1324	0.1203	0.0466	0.0752
8.75% to 10%	0.7825	1.2700	1.5678	1.3555	0.1501	2.1725
10% to 15%	1.9563	1.0481	1.1645	0.9859	0.0410	6.0698
15% to 20%	2.6085	8.2802	5.6955	4.7072	0.0473	5.88E+03

Interval (decreases)	The likelihood of decreases on the given interval likely to be realised once every number of years for the period ending 31 July 2014, using half the original data.					
Anglo American	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0079	0.0108	0.0106	0.0108	0.0199	0.0113
2.5% to 5%	0.0336	0.0402	0.0414	0.0400	0.0250	0.0287
5% to 8.75%	0.1332	0.1923	0.2236	0.1998	0.0221	0.1825
8.75% to 10%	0.9175	3.7993	4.1050	3.4787	0.0835	23.3041
10% to 15%	1.6516	4.8608	4.0876	3.3942	0.0274	159.6729
15% to 20%	4.1290	107.7421	34.2507	27.6963	0.0426	7.42E+06
Merafe Resources	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0079	0.0146	0.0140	0.0152	0.0279	0.0161
2.5% to 5%	0.0199	0.0238	0.0228	0.0261	0.0328	0.0239
5% to 8.75%	0.0679	0.0595	0.0665	0.0559	0.0267	0.0450
8.75% to 10%	0.3095	0.9944	1.1995	0.6216	0.0939	0.5675
10% to 15%	0.3980	1.4715	1.5565	0.5916	0.0284	0.8433
15% to 20%	5.5714	65.5644	29.7320	6.5581	0.0387	84.4776
Grindrod	Observed	NIG	Skew t	t-dist	EVT	Normal
0 to 2.5%	0.0080	0.0135	0.0134	0.0133	0.0448	0.0178
2.5% to 5%	0.0400	0.0402	0.0384	0.0384	0.0495	0.0239
5% to 8.75%	0.1830	0.0888	0.0958	0.0972	0.0372	0.0361
8.75% to 10%	1.2079	0.6407	0.7550	0.7741	0.1231	0.3265
10% to 15%	1.5099	0.3540	0.4188	0.4316	0.0347	0.3576
15% to 20%	-	1.2103	1.2696	1.3171	0.0420	13.9601

Table 6.9: Comparison between observed likelihood of decreases on the given interval likely to occur once every number of years to the likelihood calculated under the EVT, Skew t, t-distribution, NIG and Normal distribution for the ending 31 July 2014.

6.4 Chapter Summary

In this chapter, we fitted NIG distribution to the empirical data and we compared the fit to that of the Normal, Skew t and t-distribution. The NIG, Skew t and t-distribution fitted the empirical data better and were able to capture certain empirical data features like the heavy tails and skewness. For heavier tailed data the Skew t and t-distribution provided better fit than the NIG distribution. We also focused on the tails of the empirical data by modelling the losses using the Extreme Value Theory.

We calculated VaR under the NIG, Normal, Skew t, t-distribution and Extreme Value Theory assumptions. The results obtained showed that the VaR calculated under the four distributions outperformed those under the Normal distribution. The backtesting results clearly showed that large negative returns were more likely to occur under the EVT, t-distribution, Skew t and NIG model as compared to the Normal distribution. Making the NIG, t-distribution, Skew t and EVT better distributions to use for estimating VaR.

The results in Appendix A show that we fail to reject the null hypothesis for the NIG, Skew t and t-distribution for the stocks and the indices with exception to Merafe Resources. We calculated VaR under the NIG, Normal, Skew t and t-distribution assumptions. The results obtained showed that the VaR calculated under the three distributions outperformed those under the Normal distribution for the three sample periods.

The Kupiec LR test further showed that the NIG provided better VaR estimates over different sample periods as the null hypothesis was not rejected at the 5% significant level over the different sample period. However, the t-distribution VaR estimates were rejected for the S&P 500 over the pre-crisis period and the Skew t VaR estimates were rejected for the S&P 500 and Standard Bank during the pre-crisis and post-crisis period respectively.

Chapter 7

Conclusion

In this dissertation, we studied how risk measures such as Value-at-Risk can be estimated using the NIG, Normal, Skew t, t-distribution and Extreme Value Theory. We have done our study in terms of distributional fit and estimating Value-at-Risk of listed equity stocks and indices to determine the distribution that best models the returns of the listed stocks and indices.

We demonstrated using five stocks listed on the Johannesburg Stock Exchange, the FTSE/JSE TOP 40 index and the S&P 500 that the NIG, Skew t and t-distribution approximates log-return data reasonably well in the center, out-performing the Normal. We have seen that the NIG distribution estimates the tail behaviour better than the Normal. The Skew t and t-distribution provided a better fit for data with heavy tails. The NIG distribution has four parameters that capture characteristics like kurtosis or semi heavy tails and skewness as observed in financial data. The Normal and t-distribution have similar characteristics such as symmetry about the mean, with the t-distribution providing the kurtosis displayed in financial data.

The next part of our work was to analyse Value-at-Risk model using the Monte Carlo method. We estimated one-day VaR and ES at a 99% confidence level for all the five stocks and indices using the Normal Inverse Gaussian, Normal, Skew t, t-distribution, EVT and the non-parametric model. The results obtained clearly show, that the NIG, Skew t and t-distribution gave values that are closer to the non-parametric VaR estimate. Whereas, the Normal distribution gave values that are constantly too optimistic compared to the other four distribution. This is expected as seen in the thin tails of the Normal distribution. On the other hand the Skew t, t-distribution and the Extreme Value Theory model competed very favourable with the Normal Inverse Gaussian.

We used the backtesting method to compare VaR estimates based on returns from the early half of the period against returns that actually occurred on the second half of the period. The likelihood of an average move within a given interval using the NIG, Normal, Skew t, t-distribution, EVT and the actual observed likelihood based on frequency analysis were used. The intervals were randomly selected to represent the negative returns of the stocks and the indices. From these results we observed that the EVT method is not the best method to predict estimates that do not occur in the extreme tails of the distribution. The results show that large negative returns are more likely to occur under the EVT, t-distribution, Skew t and NIG model as compared to the Normal distribution.

There are a number of well-founded statistical distributions that can model the distribution of financial returns for individual stocks. These statistical distributions can be used to model specific parts of the financial returns distribution. For example, when modelling the extreme events, the Pareto distribution from the Extreme Value Theory would be more appropriate as well as the t-distribution and Skew t. These give a good estimate of movements that can be seen when external events cause market turmoil. The NIG distribution can be used to best model the center and the tails of the financial distribution. However, this dissertation focused on linear positions of stocks, while on the non-linear level there is still a limited number of statistical distributions that can be used to model financial returns. There is an open field of research in the future for understanding the distribution of non-linear positions for example portfolio risk management and asset allocation decisions. The problem in non-linear positions is in understanding the dependency of individual stocks in a portfolio. The other challenge with the higher dimension NIG model is the parameter estimation. Further research could be done in incorporating the ARMA (1,1)-GARCH (1,1) time series to the returns and volatility over the different sample periods in Appendix A and estimating VaR.

Appendix A

VaR estimates over different sample periods

In this Appendix, we split the sample period into three sub-samples and fit the NIG, Skew t, t-distribution and Normal distribution to Standard Bank, African Bank, Merafe Resource, Anglo American, FTSE/JSE TOP40 (J200) index and the S&P 500 index. We then estimate VaR over the different sample periods for each share and indices.

A.1 Empirical Data

The empirical study is done using four South African equity stocks, FTSE/JSE TOP40 (J200) index and the S&P 500 index. The four shares (Standard Bank (SBK), African Bank (ABL), Merafe Resource (MRF) and Anglo American (AGL)) are listed on Johannesburg Stock Exchange. Maximum available daily closing prices for the equity stocks were obtained resulting in varying periods ending July 31, 2014. The entire sample period is split into three sub-samples, that is:

(i) Pre-crisis (from 1991 January - December 2007);

(ii) Crisis period (from January 2008 - December 2009);

(iii) Post-crisis (from January 2010 - July 2014).

The analysis is performed using the entire sample period and the three sub-samples. We start off with a statistical summary of the data for each share and the indices over varying periods ending July 31, 2014. Standard Bank and Anglo American have large market capitalisation and we expect them to mimic the FTSE/JSE TOP40, while Merafe and African Bank are small and therefore would have an element of jump risk. Table A.1 below, shows the statistical summary of the data for each share, computed using the daily log returns. Based on the statistical results of the empirical data in Table A.1 the mean of the each stock and the index is relatively small compared to the variance it is almost insignificant. The excess kurtosis for each stock and the index is greater than zero. This indicates a higher peak and heavier tails meaning extreme loss and profit are more likely to occur than what the Normal distributed would predict.

The Indices, Standard Bank, African Bank and Anglo American have negative skewness that is the left tail is longer, indicating that losses occur more frequently than profits over the entire sample period. While Merafe has positive skewness implying more profits than losses were realised over the period. In general each stock and the indices display fatter tails and skewness in comparison to the Normal distribution as noted in literature [Fam65].

A.2 Fitting the distribution

In this section, we fit the NIG, Skew t, and t-distribution and perform the goodness of fit using the Kolmogorov-Smirnov test over the different sample periods. Table A.2 shows the maximum likelihood parameter estimates results for the fitted distributions and the graphical representation of the parameters is presented in Figure A.1, where for example *alpha, beta, delta* and *mu* are the NIG parameter estimates of the entire sample period and

Statistical data of the empirical distribution over the period 1991 January - July 2014.					
	Mean	Variance	Skewness	Excess Kurtosis	No. OBS[1]
S&P 500	0.00029	0.00014	−0.73247	15.95379	6170
FTSE/JSE TOP40	0.00048	0.00019	−0.40611	6.27868	4770
Standard Bank	0.00048	0.00047	−0.20369	4.86980	4090
African Bank	−0.00003	0.00082	−0.78087	10.8220	3944
Anglo American	0.00038	0.00064	−0.07927	3.93211	4162
Merafe Resource	0.00054	0.00135	0.06429	2.52105	2806
Pre-crisis (from 1991 January - December 2007) statistical data.					
	Mean	Variance	Skewness	Kurtosis	No. obs
S & P 500	0.0004	0.0001	-0.0746	3.8997	4262
FTSE/JSE TOP40	0.0006	0.0002	-0.6571	8.1379	3124
Standard Bank	0.0006	0.0006	-0.3163	4.9000	2463
African Bank	0.0007	0.0009	-0.4599	7.9791	2338
Anglo American	0.0008	0.0006	-0.0394	3.0429	2520
Merafe Resources	0.0014	0.0014	0.4452	1.3903	1485
Crisis period (from January 2008 - December 2009) statistical data.					
	Mean	Variance	Skewness	Kurtosis	No. obs
S & P 500	-0.0005	0.0005	-0.0963	4.3544	505
FTSE/JSE TOP40	-0.0001	0.0005	0.0408	1.3969	501
Standard Bank	0.0000	0.0007	0.1997	1.3906	490
African Bank	-0.0002	0.0009	0.0418	0.4575	491
Anglo American	-0.0005	0.0016	-0.1295	1.8202	499
Merafe Resources	-0.0011	0.0022	-0.6017	2.7368	457
Post- crisis (from January 2010 - July 2014) statistical data.					
	Mean	Variance	Skewness	Kurtosis	No. obs
S & P 500	0.0004	0.0001	-3.2713	49.1255	1403
FTSE/JSE TOP40	0.0005	0.0001	-0.1593	1.3757	1145
Standard Bank	0.0003	0.0002	-0.1336	1.1070	1137
African Bank	-0.0014	0.0006	-2.5109	29.7154	1115
Anglo American	-0.0001	0.0003	0.1624	0.5933	1143
Merafe Resources	-0.0000	0.0008	0.1183	1.8172	864

Table A.1: Statistical data for each stock and indices.

alpha1, beta1, delta1 and *mu1* represent the NIG parameter estimates of the Pre-crisis period.

Table A.3 shows the Kolmogorov-Smirnov test statistic values and critical values for different confidence levels. The test statistic values are the distance between the empirical cumulative distribution and the fitted cumulative distribution. The results of the table shows that we do not reject the null hypothesis for the NIG, Skew t and t-distribution. However, the null hypothesis is rejected for the Normal distribution for all the shares with exception of Merafe Resources and in some cases the other three distributions are rejected for Merafe Resources as well.

A.3 Value-at-Risk

In this section, we present the comparison of Value-at-Risk estimates under the NIG, Skew t, Normal and t-distribution to the empirical distribution of the stocks and indices over the four sample periods. We further verify the correctness of the VaR models using the backtesting technique.

A.3.1 Value-at-Risk estimates

The VaR and ES estimates obtained under the NIG, Skew t, Normal and t-distribution assumption for a one-day holding period at 99% confidence level are shown in Table A.4 and Table A.5 respectively. The VaR and ES estimates under the NIG, Normal, Skew t and t-distribution assumption were calculated using the Monte Carlo simulation. The results for these two tables shows that the VaR estimates under the Normal distribution underestimates the VaR and ES values, as it is well known in literature. The VaR estimates under the NIG, Skew t and t-distribution better values when compared to the Historical VaR and ES values.

Parameter estimates over the period 1991 January - July 2014							
	NIG				t-dist.	Skew t	
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	$\widehat{\mu}$	$\widehat{k}$	$\widehat{k}$	$\widehat{\beta}$
S&P 500	58.3541	-5.0934	0.0075	0.0009	2.7962	2.8038	0.9519
FTSE/JSE TOP 40	67.8875	-5.109	0.0123	0.0014	4.0000	3.8419	0.9533
Standard Bank	44.7345	0.3524	0.0208	0.0003	4.0000	4.0006	1.0145
African Bank	32.372	-0.6613	0.025	0.0005	3.8414	4.0000	1.0000
Anglo American	43.1323	0.9557	0.027	-0.0002	4.6143	4.6127	1.0208
Merafe Resources	41.3981	4.5965	0.0527	-0.0054	6.5225	6.2649	1.1460
Pre-crisis (from 1991 January - December 2007) parameter estimates							
	NIG				t-dist.	Skew t	
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	$\widehat{\mu}$	$\widehat{k}$	$\widehat{k}$	$\widehat{\beta}$
S & P 500	82.3073	-2.7308	0.0084	0.0006	4.0000	3.5163	0.9765
FTSE/JSE TOP40	77.3477	-5.8151	0.0128	0.0015	4.0000	4.2290	0.9605
Standard Bank	44.1135	-0.1645	0.0238	0.0007	4.0000	4.3920	1.0121
African Bank	32.4499	0.9420	0.0274	-0.0001	4.0105	4.0001	1.0373
Anglo American	53.2541	0.9913	0.0306	0.0002	5.6917	5.6988	1.0134
Merafe Resources	58.7370	15.7660	0.0737	-0.0192	9.2786	9.2746	1.2459
Crisis period (from January 2008 - December 2009) parameter estimates							
	NIG				t-dist.	Skew t	
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	$\widehat{\mu}$	$\widehat{k}$	$\widehat{k}$	$\widehat{\beta}$
S & P 500	27.9336	-3.9019	0.0137	0.0014	2.5545	2.5254	0.9124
FTSE/JSE TOP40	59.4900	-0.3225	0.0270	0.0000	5.7367	5.7149	0.9874
Standard Bank	45.3549	4.4290	0.0323	-0.0031	5.5645	5.6319	1.0618
African Bank	78.1936	5.9075	0.0715	-0.0056	14.7400	14.4505	1.0546
Anglo American	29.5059	-0.0252	0.0474	-0.0005	5.2692	5.2457	1.0149
Merafe Resources	29.4521	-4.0030	0.0611	0.0073	5.6407	5.8777	0.9300
Post- crisis (from January 2010 - July 2014) parameter estimates							
	NIG				t-dist.	Skew t	
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	$\widehat{\mu}$	$\widehat{k}$	$\widehat{k}$	$\widehat{\beta}$
S & P 500	70.5049	-13.0538	0.0070	0.0017	2.9116	2.9563	0.9094
FTSE/JSE TOP40	122.4542	-13.7630	0.0132	0.0020	5.6707	5.7635	0.9264
Standard Bank	128.9099	-7.2170	0.0251	0.0017	8.8356	4.0000	1.0000
African Bank	34.4911	-7.7071	0.0175	0.0026	3.3022	3.3084	0.8747
Anglo American	123.3751	11.4136	0.0420	-0.0040	12.8391	13.2898	1.0454
Merafe Resources	72.0322	11.2690	0.0558	-0.0088	10.1389	8.8798	1.2196

Table A.2: Maximum likelihood parameter estimates.

A.3.2 Backtesting the model

Table A.7 shows the test statistic values according to Kupiec likelihood ratio (LR) test [Kup95]. The Kupiec LR test is given by:

$$-2\ln[(1-p)^{n-x}p^x] + 2\ln[(1-x/n)^{n-x}(x/n)^x], \tag{A.1}$$

where p is the probability under the VaR model, x is the number of violations and n the sample period. Under the null hypothesis the Kupiec LR test follows a chi-square distribution with one degree of freedom. The values of Kupiec LR test are high for either very low or very high numbers of violations [Hul10]. The null hypothesis is not reject when the Kupiec LR test is less than the critical values. At a 5% significance level the critical value is given by 3.8415, the null hypothesis is rejected for S&P 500 index under the Skew t and t-distribution VaR model during the pre-crisis period. Standard Bank Kupiec LR test is rejected during the post-crisis period, the number of violations is to low compared to the expected number of violations. The null hypothesis is rejected under the Normal VaR model for most of the different sample periods. The null hypothesis is not reject for all sample periods and for all shares and indices at the 5% significance level, this seems to be a better model for risk managers, given the results of the Kupiec LR test.

A.4 Conclusion

In this Appendix, we implement the NIG distribution and compare the fitting to that of the Skew t, t-distribution and Normal distribution over four sample periods defined in Appendix A.1. The NIG, Skew t and t-distribution fitted the financial returns better both in the center and tails as compared to the classic Normal distribution, with the Skew t and t-distribution showing heavier tails than the NIG semi-heavy tails. We failed to reject the null hypothesis for the NIG, Skew t and t-distribution for the stocks and the indices with exception to Merafe Resources. We calculated VaR under the

NIG, Normal, Skew t and t-distribution assumptions. The results obtained showed that the VaR calculated under the three distributions outperformed those under the Normal distribution.

The Kupiec LR test further showed that the NIG provided better VaR estimates over different sample periods as the null hypothesis was not rejected at the 5% significant level over the different sample period. However, the t-distribution VaR estimates were rejected for the S&P 500 over the pre-crisis period and the Skew t VaR estimates were rejected for the S&P 500 and Standard Bank during the pre-crisis and post-crisis period respectively. Further research could be done in incorporating the ARMA (1,1)-GARCH (1,1) time series to the returns and volatility over the different sample periods and estimating VaR.

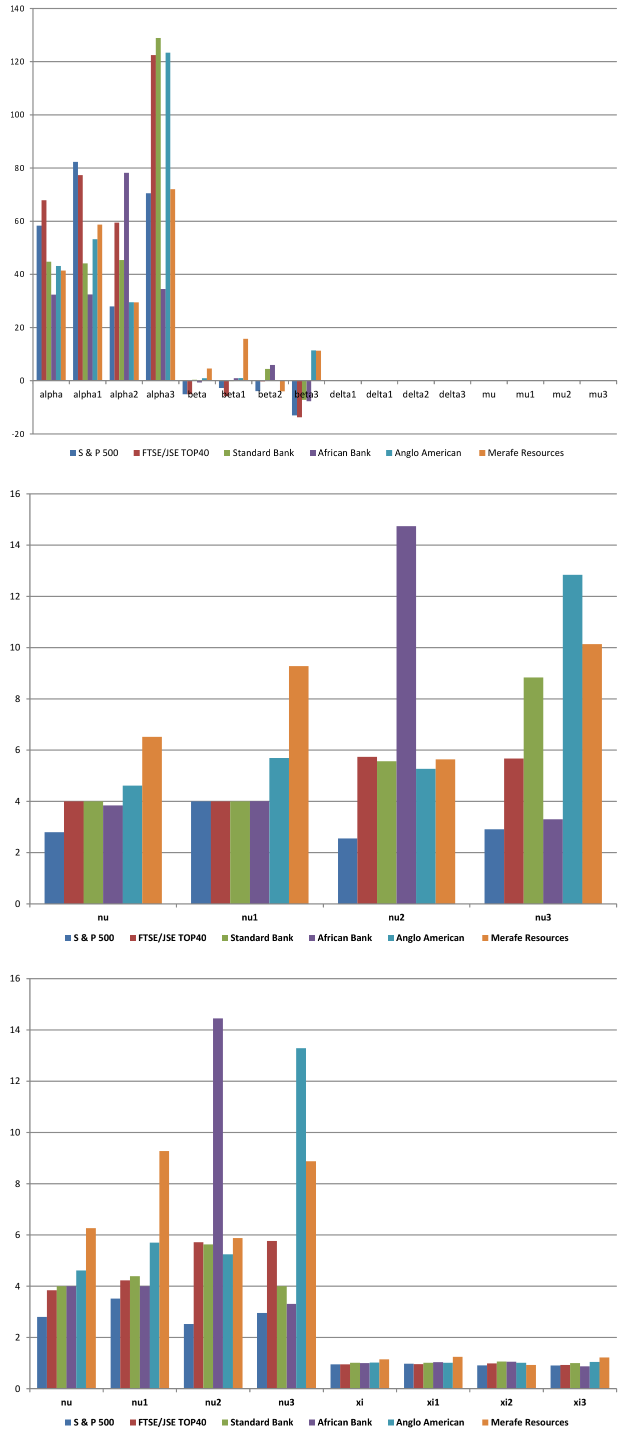


Figure A.1: Graphical representation of the NIG, t-distribution and Skew t-distribution maximum likelihood parameters estimates.

Parameter estimates for the period 1991 January - July 2014								
	Kolmogorov-Smirnov				Critical Value			
	NIG	t-dist.	Skew t	Normal	$C_{0.9}$	$C_{0.95}$	$C_{0.975}$	$C_{0.99}$
S & P 500	0.0058	0.0135	0.0113	0.0875	0.0156	0.0173	0.0188	0.0207
FTSE/JSE TOP40	0.0081	0.0082	0.0046	0.0580	0.0177	0.0197	0.0214	0.0236
Standard Bank	0.0201	0.0190	0.0184	0.0560	0.0191	0.0212	0.0231	0.0255
African Bank	0.0173	0.0150	0.0149	0.0657	0.0195	0.0216	0.0236	0.0259
Anglo American	0.0114	0.0101	0.0088	0.0443	0.0190	0.0211	0.0229	0.0252
Merafe Resources	0.0809	0.0862	0.0828	0.0765	0.0232	0.0257	0.0280	0.0308
Pre-crisis (from 1991 January - December 2007) parameter estimates								
	Kolmogorov-Smirnov				Critical Value			
	NIG	t-dist.	Skew t	Normal	$C_{0.9}$	$C_{0.95}$	$C_{0.975}$	$C_{0.99}$
S & P 500	0.0117	0.0187	0.0163	0.0626	0.0187	0.0208	0.0227	0.0249
FTSE/JSE TOP40	0.0116	0.0103	0.0092	0.0535	0.0219	0.0243	0.0265	0.0291
Standard Bank	0.1934	0.0288	0.0272	0.0546	0.0247	0.0274	0.0298	0.0328
African Bank	0.0245	0.0247	0.0224	0.0641	0.0253	0.0281	0.0306	0.0337
Anglo American	0.0112	0.0104	0.0105	0.0349	0.0244	0.0271	0.0295	0.0324
Merafe Resources	0.0885	0.0994	0.0968	0.0982	0.0318	0.0352	0.0384	0.0422
Crisis period (from January 2008 - December 2009) parameter estimates								
	Kolmogorov-Smirnov				Critical Value			
	NIG	t-dist.	Skew t	Normal	$C_{0.9}$	$C_{0.95}$	$C_{0.975}$	$C_{0.99}$
S & P 500	0.0278	0.0369	0.0288	0.0860	0.0545	0.0604	0.0659	0.0724
FTSE/JSE TOP40	0.0162	0.0177	0.0181	0.0410	0.0547	0.0607	0.0661	0.0727
Standard Bank	0.0235	0.0285	0.0287	0.0618	0.0553	0.0614	0.0669	0.0735
African Bank	0.0261	0.0297	0.0237	0.0279	0.0552	0.0613	0.0668	0.0735
Anglo American	0.0322	0.0342	0.0336	0.0676	0.0548	0.0608	0.0663	0.0729
Merafe Resources	0.0596	0.0557	0.0606	0.0594	0.0573	0.0635	0.0692	0.0761
Post- crisis (from January 2010 - July 2014) parameter estimates								
	Kolmogorov-Smirnov				Critical Value			
	NIG	t-dist.	Skew t	Normal	$C_{0.9}$	$C_{0.95}$	$C_{0.975}$	$C_{0.99}$
S & P 500	0.0160	0.0212	0.0165	0.1066	0.0327	0.0363	0.0395	0.0435
FTSE/JSE TOP40	0.0135	0.0166	0.0171	0.0504	0.0362	0.0401	0.0437	0.0481
Standard Bank	0.0175	0.0160	0.0254	0.0293	0.0363	0.0403	0.0439	0.0483
African Bank	0.0279	0.0236	0.0152	0.0942	0.0367	0.0407	0.0443	0.0487
Anglo American	0.0150	0.0138	0.0142	0.0266	0.0362	0.0402	0.0438	0.0481
Merafe Resources	0.0945	0.0932	0.1072	0.0899	0.0416	0.0462	0.0504	0.0554

Table A.3: Kolmogorov-Smirnov test statistics and estimated critical values for the fitted distributions.

VaR estimates over the period 1991 January - July 2014					
	Historical	NIG	t-dist	Skew t	Normal
S&P 500	3.15%	3.31%	3.04%	3.45%	2.68%
FTSE/JSE TOP 40	3.79%	3.75%	3.56%	3.84%	3.11%
Standard Bank	5.81%	5.81%	5.65%	5.86%	4.87%
African Bank	7.35%	7.53%	7.73%	7.54%	6.37%
Anglo American	6.60%	6.38%	6.73%	6.46%	5.97%
Merafe Resources	9.02%	8.98%	9.36%	8.13%	8.36%
Pre-crisis (from 1991 January - Dec 2007) VaR estimates					
	Historical	NIG	t-dist	Skew t	Normal
S & P 500	2.6213%	2.8307%	2.5053%	2.8592%	2.3426%
FTSE/JSE TOP40	3.7136%	3.5412%	3.5011%	3.6038%	2.9309%
Standard Bank	6.4264%	6.1871%	6.0822%	6.2585%	5.2607%
African Bank	7.6811%	8.2170%	7.9237%	7.7267%	7.0160%
Anglo American	6.1856%	6.2482%	6.0138%	5.9658%	5.5577%
Merafe Resources	8.0043%	8.3464%	9.2559%	7.9108%	8.3902%
Crisis period (from January 2008 - December 2009) VaR estimates					
	Historical	NIG	t-dist	Skew t	Normal
S & P 500	6.2799%	7.2567%	6.3789%	8.4283%	5.1633%
FTSE/JSE TOP40	5.3725%	5.6579%	5.4216%	5.5736%	4.8682%
Standard Bank	6.5636%	6.6214%	6.9057%	6.3490%	6.1787%
African Bank	6.8522%	7.2559%	7.6928%	7.2836%	6.9659%
Anglo American	9.9742%	10.5935%	10.9597%	10.1201%	9.4863%
Merafe Resources	13.5510%	12.4084%	11.6611%	12.8104%	10.9410%
Post- crisis (from January 2010 - July 2014) VaR estimates					
	Historical	NIG	t-dist	Skew t	Normal
S & P 500	2.8858%	2.8198%	2.9767%	3.2944%	2.5355%
FTSE/JSE TOP40	2.8897%	2.7751%	2.8675%	2.8675%	2.3807%
Standard Bank	3.6093%	3.6265%	3.3549%	4.1969%	3.1541%
African Bank	6.8760%	7.7128%	6.5781%	7.1226%	6.1529%
Anglo American	4.3435%	4.4079%	4.5477%	4.4322%	4.2903%
Merafe Resources	6.4198%	6.6560%	7.0918%	6.3123%	6.3751%

Table A.4: Comparison of the Value-at-Risk estimates, the non-parametric estimates are calculated using the Historical Simulation approach.

Estimates of one-day Expected Shortfall.					
	Historical	NIG	t-dist.	Skew t	Normal
FTSE/JSE TOP 40	5.20%	4.96%	4.89%	5.41%	3.50%
S&P 500	4.82%	4.69%	4.88%	5.48%	3.01%
Standard Bank	7.96%	7.49%	7.81%	8.03%	5.55%
African Bank	11.58%	9.62%	11.06%	9.73%	7.76%
Anglo American	8.93%	8.33%	8.71%	8.28%	6.76%
Merafe Resources	12.22%	10.96%	11.70%	10.28%	9.55%
Pre-crisis (from 1991 January - Dec 2007) Expected shortfall estimates.					
	Historical	NIG	t-dist.	Skew t	Normal
S & P 500	3.47%	3.64%	3.39%	4.19%	2.67%
FTSE/JSE TOP40	5.21%	4.57%	4.85%	4.94%	3.39%
Standard Bank	8.76%	7.99%	8.20%	8.25%	5.99%
African Bank	11.35%	10.28%	11.17%	10.22%	8.01%
Anglo American	7.96%	7.79%	8.04%	7.44%	6.35%
Merafe Resources	9.77%	10.01%	11.37%	9.51%	9.67%
Crisis period (from January 2008 - December 2009) Expected shortfall estimates.					
	Historical	NIG	t-dist.	Skew t	Normal
S & P 500	8.20%	9.49%	11.42%	21.09%	5.88%
FTSE/JSE TOP40	6.58%	6.80%	7.06%	7.26%	5.49%
Standard Bank	7.96%	8.20%	8.94%	8.20%	7.11%
African Bank	8.86%	8.47%	8.99%	8.51%	7.95%
Anglo American	14.01%	13.35%	14.02%	13.44%	10.78%
Merafe Resources	17.44%	15.43%	14.84%	16.01%	12.50%
Post- crisis (from January 2010 - July 2014) Expected shortfall estimates.					
	Historical	NIG	t-dist.	Skew t	Normal
S & P 500	4.74%	3.91%	4.72%	5.17%	2.90%
FTSE/JSE TOP40	3.29%	3.69%	3.62%	3.69%	2.71%
Standard Bank	4.37%	4.37%	4.06%	6.01%	3.64%
African Bank	12.55%	10.28%	9.82%	10.50%	7.06%
Anglo American	4.95%	5.08%	5.53%	5.13%	4.93%
Merafe Resources	9.10%	8.00%	8.61%	7.78%	7.43%

Table A.5: Comparison of the one-day Expected Shortfall estimates. The non-parametric estimates are calculated using the Historical Simulation approach.

Number of violations for 99% daily-VaR.						
	Historical	NIG	t-dist.	Skew t	Normal	Expected violations
S&P 500	63	54	72	51	102	62
FTSE/JSE TOP 40	49	51	56	43	98	48
Standard Bank	41	41	46	39	72	41
African Bank	40	37	37	37	60	39
Anglo American	42	48	38	45	63	42
Merafe Resources	28	30	25	36	34	28
Pre-crisis (from 1991 January - December 2007) Number of violations.						
	Historical	NIG	t-dist	Skew t	Normal	Expected violations
S & P 500	43	32	56	30	74	43
FTSE/JSE TOP40	32	35	36	34	63	31
Standard Bank	25	27	28	27	45	25
African Bank	24	17	22	24	29	23
Anglo American	26	24	28	28	37	25
Merafe Resources	17	13	8	18	13	15
Crisis period (from January 2008 - December 2009) Number of violations.						
	Historical	NIG	t-dist	Skew t	Normal	Expected violations
S & P 500	6	4	5	3	11	5
FTSE/JSE TOP40	6	5	5	5	9	5
Standard Bank	5	5	4	6	7	5
African Bank	5	5	4	4	5	5
Anglo American	5	4	4	5	7	5
Merafe Resources	5	8	9	7	9	5
Post - crisis (from January 2010 - July 2014) Number of violations.						
	Historical	NIG	t-dist	Skew t	Normal	Expected violations
S & P 500	15	17	13	8	20	14
FTSE/JSE TOP40	12	15	12	12	26	11
Standard Bank	12	12	17	5	23	11
African Bank	12	9	14	10	17	11
Anglo American	12	10	8	10	12	11
Merafe Resources	9	8	7	9	9	9

Table A.6: Number of violations for each VaR model and the expected violations at 99% confidence level.

Kupiec LR test statistic.					
	Historical	NIG	t-dist.	Skew t	Normal
S&P 500	0.0275	1.0133	1.6484	1.9920	22.2150
FTSE/JSE TOP 40	0.0355	0.2255	1.3817	0.4838	41.0648
Standard Bank	0.0002	0.0002	0.6175	0.0906	19.4767
African Bank	0.0080	0.1557	0.1557	0.1557	9.3361
Anglo American	0.0035	0.9414	0.3276	0.2701	9.5849
Merafe Resources	0.0001	0.1325	0.3499	2.0832	1.1898
Pre-crisis (from 1991 January - December 2007) Kupiec LR test statistic.					
	Historical	NIG	t-dist	skew t	Normal
S & P 500	0.0034	2.9251	3.8616	4.2101	19.1317
FTSE/JSE TOP40	0.0185	0.4400	0.6983	0.2394	25.1881
Standard Bank	0.0056	0.2234	0.4461	0.2234	13.6734
African Bank	0.0165	1.9429	0.0839	0.0165	1.2677
Anglo American	0.0254	0.0586	0.3033	0.3033	4.8774
Merafe Resources	0.3004	0.2430	3.8349	0.6321	0.2430
Crisis period (from January 2008 - December 2009) Kupiec LR test statistic.					
	Historical	NIG	t-dist	Skew t	Normal
S & P 500	0.1703	0.2375	0.0005	0.9837	5.2982
FTSE/JSE TOP40	0.1859	0.0000	0.0000	0.0000	2.5964
Standard Bank	0.0020	0.0020	0.1781	0.2328	0.8026
African Bank	0.0017	0.0017	0.1819	0.1819	0.0017
Anglo American	0.0000	0.2129	0.2129	0.0000	0.7268
Merafe Resources	0.0397	2.1249	3.3823	1.1226	3.3823
Post- crisis (from January 2010 - July 2014) Kupiec LR test statistic.					
	Historical	NIG	t-dist	Skew t	Normal
S & P 500	0.0662	0.5949	0.0783	3.0980	2.2671
FTSE/JSE TOP40	0.0263	1.0129	0.0263	0.0263	13.7331
Standard Bank	0.0346	0.0346	2.4442	4.5606	9.2683
African Bank	0.0639	0.4483	0.6807	0.1241	2.6714
Anglo American	0.0283	0.1887	1.1616	0.1887	0.0283
Merafe Resources	0.0149	0.0491	0.3362	0.0149	0.0149

Table A.7: Kupiec likelihood ration test statistic results.

Appendix B

Modified Bessel Function

The objective behind this chapter is to summaries some of the properties of the modified Bessel function found in the Normal Inverse Gaussian distribution.

Definition B.1. *The solutions to the Bessel Equation*

$$x^2\frac{d^2w}{dx^2} + x\frac{dw}{dx} - (x^2 + \lambda^2)w = 0 \tag{B.1}$$

are called the modified Bessel function and denoted by $K_\lambda(x)$ *and* $I_{\pm\lambda}(x)$*. These solutions are regular function of* $x \in \mathbb{C}$ *throughout the* $x -$ *plane cut along the negative real axis, and for fixed* $x \neq 0$ *each is an entire function of* λ*.* $K_\lambda(x)$ *tends to* 0 *as* $|x| \to \infty$ *in the sector* $|arg\ (x)| < \pi/2$ *and for all* λ*.* $K_\lambda(x)$ *and* $I_{\pm\lambda}(x)$ *are real and positive when* $\lambda > -1$ *and* $x > 0$*.*

[Pra99],[AS64]

Theorem B.2. (Basic Properties)

$$K_\lambda(x) = K_{-\lambda}(x) \tag{B.2}$$

$$K_{\lambda+1}(x) = \frac{2\lambda}{x}K_\lambda(x) + K_{\lambda-1}(x) \tag{B.3}$$

[Pra99]

Theorem B.3. (Integral Representation)

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty \tau^{\lambda-1}\exp\left(-\frac{1}{2}x\left(\tau+\tau^{-1}\right)\right)d\tau. \tag{B.4}$$

[Pra99]

Theorem B.4. (Asymptotic Formula)

$$K_1(x) \sim \sqrt{\pi/2}\,x^{-1/2}\,\mathrm{e}^{-x}, \quad \text{as} \quad x\to\infty. \tag{B.5}$$

[BN95]

Theorem B.5. (Derivatives)

$$K_0'(x) = -K_1(x) \tag{B.6}$$

$$K_\lambda'(x) = -\frac{1}{2}(K_{\lambda+1}(x)+K_{\lambda-1}(x)) \tag{B.7}$$

$$= -\frac{\lambda}{x}K_\lambda(x) - K_{\lambda-1}(x) \tag{B.8}$$

$$(\ln K_\lambda(x))' = \frac{\lambda}{x} - \frac{K_{\lambda+1}(x)}{K_\lambda(x)}, \quad x>0. \tag{B.9}$$

[Pra99],[AS64]

Appendix C

R source codes

Figure C.1: R source codes for the closing price and daily log returns graphs.

```
> J200file<- read.csv(file="J200.csv", header=T, dec=".", sep=",")
> J200_Price=J200file[,2]
> summary(J200_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3903    6801   10690   15330   24720   37600
> J200_returns=diff(log(J200file[,2]))
> summary(J200_returns)
     Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.1429000 -0.0067860  0.0008756  0.0004453  0.0079440  0.0844700
> par(mfrow=c(2,1))
> plot(J200_Price,type="l",xlab="Time",ylab="Daily Closing Level")
> plot(J200_returns,type="l",xlab="Time",ylab="log-returns")
> SBKfile<- read.csv(file="SBK.csv", header=T, dec=".", sep=",")
> SBK_Price=SBKfile[,2]
> summary(SBK_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1035    3000    6938    6487    9930   12030
> SBK_returns=diff(log(SBKfile[,2]))
> summary(SBK_returns)
Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.1863000 -0.0109100  0.0004463  0.0004385  0.0122500  0.1379000
> plot(SBK_Price,type="l",xlab="Time",ylab="Daily closing prices")
> plot(SBK_returns,type="l",xlab="Time",ylab="log-returns")
> ABLfile<- read.csv(file="ABL.csv", header=T, dec=".", sep=",")
> ABL_Price=ABLfile[,2]
```

```
> summary(ABL_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   471    1080    2300    2144    3050    4049
> ABL_returns=diff(log(ABLfile[,2]))
> summary(ABL_returns)
     Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.3023000 -0.0143900  0.0003448  0.0002339  0.0148200  0.1574000
> plot(ABL_Price,type="l",xlab="Time",ylab="Daily closing prices")
> plot(ABL_returns,type="l",xlab="Time",ylab="log-returns")
> Merafefile<- read.csv(file="Merafe.csv", header=T, dec=".", sep=",")
> Merafe_Price=Merafefile[,2]
> summary(Merafe_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  28.0    66.0    87.0   106.9   133.0   425.0
> Merafe_returns=diff(log(Merafefile[,2]))
> summary(Merafe_returns)
     Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.2624000 -0.0226600 -0.0059000  0.0003251  0.0228100  0.1769000
> plot(Merafe_Price,type="l",xlab="Time",ylab="Daily closing prices")
> plot(Merafe_returns,type="l",xlab="Time",ylab="log-returns")
> Anglofile<- read.csv(file="Anglo.csv", header=T, dec=".", sep=",")
> Anglo_Price=Anglofile[,2]
> summary(Anglo_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3800   11970   18500   20930   29420   55600
> Anglo_returns=diff(log(Anglofile[,2]))
> summary(Anglo_returns)
```

```
Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.1823000 -0.0140200  0.0002237  0.0002956  0.0146400  0.1435000
> plot(Anglo_Price,type="l",xlab="Time",ylab="Daily closing prices")
> plot(Anglo_returns,type="l",xlab="Time",ylab="log-returns")
 > Grindrodfile<- read.csv(file="Grindrod.csv", header=T, dec=".", sep=",")
> Grindrod_Price=Grindrodfile[,2]
> summary(Grindrod_Price)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    20     468    1400    1201    1645    2840
> Grindrod_returns=diff(log(Grindrodfile[,2]))
> summary(Grindrod_returns)
   Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
   -0.262400 -0.013330  0.001348  0.001291  0.015870  0.356700 >
plot(Grindrod_Price,type="l",xlab="Time",ylab="Daily closing prices")
> plot(Grindrod_returns,type="l",xlab="Time",ylab="log-returns")
> AngloUKfile<- read.csv(file="AngloUK.csv", header=T, dec=".", sep=",")
> AngloUK_Price=AngloUKfile[,2]
> summary(AngloUK_Price)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 674.9  1230.0  1933.0  1929.0  2582.0  3552.0
> AngloUK_returns=diff(log(AngloUKfile[,2]))
> summary(AngloUK_returns)
     Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.2246000 -0.0142400  0.0009344  0.0002219  0.0151700  0.2053000
> plot(AngloUK_Price,type="l",xlab="Time",ylab="Daily closing prices")
```

Figure C.2: R source codes for QQ-plots.

```
> qqnorm((J200_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(J200_returns)
> qqnorm((SBK_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(SBK_returns)
> qqnorm((ABL_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(ABL_returns)
> qqnorm((Anglo_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(Anglo_returns)
> qqnorm((AngloUK_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(AngloUK_returns)
> qqnorm((Merafe_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(Merafe_returns)
> qqnorm((Grindrod_returns),xlab="Quantiles of Standard Normal",ylab="Empirical Quantiles")
> qqline(Grindrod_returns)
```

Figure C.3: R source codes for fitting the NIG and t-distribution.

```
> library(fBasics)
> nigFit(J200_returns, alpha=1, beta=0, delta=1, mu=mean(J200_returns), doplot=TRUE)
> nigFit(SBK_returns, alpha=1, beta=0, delta=1, mu=mean(SBK_returns), doplot=TRUE)
> nigFit(ABL_returns, alpha=1, beta=0, delta=1, mu=mean(ABL_returns), doplot=TRUE)
> nigFit(Anglo_returns, alpha=1, beta=0, delta=1, mu=mean(Anglo_returns), doplot=TRUE)
> nigFit(AngloUK_returns, alpha=1, beta=0, delta=1, mu=mean(AngloUK_returns), doplot=TRUE)
>nigFit(Merafe_returns, alpha=1, beta=0, delta=1, mu=mean(Merafe_returns), doplot=TRUE)
> nigFit(Grindrod_returns, alpha=1, beta=0, delta=1, mu=mean(Grindrod_returns), doplot=TRUE)

> library(fGarch)
> tFit(J200_returns)
> tFit(SBK_returns)
> tFit(ABL_returns)
> tFit(Anglo_returns)
> tFit(AngloUK_returns)
> tFit(Merafe_returns)
> tFit(Grindrod_returns)

> x=seq(-0.2,0.2,length=4500)
> x1=x*100
> hist(J200_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="FTSE/JSE
TOP40 INDEX", ylim=c(0,40))
> lines(x,dnig(x,alpha=67.077482734, beta=-5.030372307, delta=0.012680344, mu=0.001398949),
lty=1, col="black")
> lines(x,dnorm(x, mean(J200_returns), sd(J200_returns)), lty=2, col="blue")
> lines(x,dt(x1, df=3.984366, sd(J200_returns))*100, lty=5, col="red")
> legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue",
"red"))
```

```
> hist(SBK_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="Standard Bank", ylim=c(0,40))
> lines(x,dt(x1, df= 1.768262, sd(SBK_returns))*100, lty=5, col="red")
> lines(x,dnig(x,alpha=43.837700000, beta=0.390100300, delta=0.021332960, mu=0.000248665), lty=1, col="black")
> lines(x,dnorm(x, mean(SBK_returns), sd(SBK_returns)), lty=2, col="blue")
> legend("topleft", c("NIG Dist.", "Normal Dist.", "t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))

> hist(ABL_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="African Bank", ylim=c(0,40))
> lines(x,dt(x1, df= 1.375823, sd(ABL_returns))*100, lty=5, col="red")
> lines(x,dnig(x,alpha=34.914860000, beta=0.207370400, delta=0.025836300, mu=0.000080426), lty=1, col="black")
> lines(x,dnorm(x, mean(ABL_returns), sd(ABL_returns)), lty=2, col="blue")
> legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))

> hist(Anglo_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="Anglo American", ylim=c(0,40))
> lines(x,dt(x1, df= 1.45802, sd(Anglo_returns))*100, lty=5, col="red")
> lines(x,dnig(x,alpha=42.265046320, beta=0.709176938, delta=0.027389859, mu=-0.000164029), lty=1, col="black")
> lines(x,dnorm(x, mean(Anglo_returns), sd(Anglo_returns)), lty=2, col="blue")
 > legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))

> hist(AngloUK_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="Anglo American UK", ylim=c(0,40))
> lines(x,dt(x1, df= 1.351801, sd(AngloUK_returns))*100, lty=5, col="red")
> lines(x,dnig(x,alpha=31.951865726, beta=-2.479612212, delta=0.024460343, mu=0.002125874), lty=1, col="black")
> lines(x,dnorm(x, mean(AngloUK_returns), sd(AngloUK_returns)), lty=2, col="blue")
> legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))

> hist(Merafe_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="Merafe Resources", ylim=c(0,40))
> lines(x,dt(x1, df= 0.9775525)*100, lty=5, col="red")
> lines(x,dnorm(x, mean(Merafe_returns), sd(Merafe_returns)), lty=2, col="blue")
> lines(x,dnig(x,alpha=42.329333877, beta=4.525437812 , delta=0.056368911, mu=-0.005736016), lty=1, col="black")
> legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))

> hist(Grindrod_returns, n=50, probability=TRUE, border="grey", col="lightgrey", xlab="",main="Grindrod Limited", ylim=c(0,40))
> lines(x,dt(x1, df= 1.195069, sd(Grindrod_returns))*100, lty=5, col="red")
> lines(x,dnorm(x, mean(Grindrod_returns), sd(Grindrod_returns)), lty=2, col="blue")
> lines(x,dnig(x,alpha=16.618440000, beta=0.376379500, delta=0.019603710, mu=-0.000846981), lty=1, col="black")
>  legend("topleft", c("NIG Dist.", "Normal Dist.", "Student t-Dist."),lty = c(1,2,4), col = c("black","blue", "red"))
```

Figure C.4: R source codes for generating EVT parameters.

```
> J200EVTfile<- read.csv(file="J200.csv", header=T, dec=".", sep=",")
> J200EVT_returns=J200EVTfile[,1]
> summary(J200EVT_returns)
    Min.   1st Qu.   Median     Mean   3rd Qu.      Max.
-1.429e-01 -1.336e-02 -7.215e-03 -1.006e-02 -3.370e-03 -3.660e-06
> SBKEVTfile<- read.csv(file="StandardBank.csv", header=T, dec=".", sep=",")
> SBKEVT_returns=SBKEVTfile[,1]
> summary(SBKEVT_returns)
    Min.   1st Qu.   Median     Mean   3rd Qu.      Max.
-1.863e-01 -2.119e-02 -1.115e-02 -1.585e-02 -5.552e-03 -9.148e-05
> ABLEVTfile<- read.csv(file="AfricanBank.csv", header=T, dec=".", sep=",")
> ABLEVT_returns=ABLEVTfile[,1]
> summary(ABLEVT_returns)
   Min.  1st Qu.  Median    Mean  3rd Qu.    Max.
-0.30230 -0.02600 -0.01450 -0.01969 -0.00690 -0.00030
> AngloEVTfile<- read.csv(file="Anglo.csv", header=T, dec=".", sep=",")
> AngloEVT_returns=AngloEVTfile[,1]
> summary(AngloEVT_returns)
    Min.   1st Qu.   Median     Mean   3rd Qu.      Max.
-0.1823000 -0.0245900 -0.0141000 -0.0186000 -0.0066950 -0.0000431
> AngloUKEVTfile<- read.csv(file="AngloUK.csv", header=T, dec=".", sep=",")
> AngloUKEVT_returns=AngloUKEVTfile[,1]
> summary(AngloUKEVT_returns)
   Min.  1st Qu.  Median    Mean  3rd Qu.    Max.
-0.22000 -0.03000 -0.01000 -0.02001 -0.01000  0.00000



> MerafeEVTfile<- read.csv(file="Merafe.csv", header=T, dec=".", sep=",")
> MerafeEVT_returns=MerafeEVTfile[,1]
> summary(MerafeEVT_returns)
   Min.  1st Qu.  Median    Mean  3rd Qu.    Max.
-0.26000 -0.04000 -0.02000 -0.02833 -0.01000  0.00000
> GrindrodEVTfile<- read.csv(file="Grindrod.csv", header=T, dec=".", sep=",")
> GrindrodEVT_returns=GrindrodEVTfile[,1]
> summary(GrindrodEVT_returns)
   Min.  1st Qu.  Median    Mean  3rd Qu.    Max.
-0.26000 -0.03000 -0.01000 -0.02198 -0.01000  0.00000

> qplot(J200EVT)
> qplot(SBKEVT)
> qplot(ABLEVT)
> qplot(AngloEVT)
> qplot(AngloUKEVT)
> qplot(MerafeEVT)
 > qplot(GrindrodEVT)

> meplot(J200EVT, omit=3)
> abline(v=0.0211, col="black",lty=3)
> meplot(SBKEVT, omit=3)
> abline(v=0.033997, col="black",lty=3)
> meplot(ABLEVT, omit=3)
> abline(v=0.041942, col="black",lty=3)
> Meplot(AngloEVT, omit=3)
> abline(v=0.03915, col="black",lty=3)
```

```
> meplot(AngloUKEVT, omit=3)
> abline(v=0.04340, col="black",lty=3)
> meplot(MerafeEVT, omit=3)
> abline(v=0.05485, col="black",lty=3)
> meplot(GrindrodEVT, omit=3)
> abline(v=0.04879, col="black",lty=3)

> outSBKEVT<-gpd(SBKEVT, 0.033997)
> outABLEVT<-gpd(ABLEVT, 0.041942)
> plot(outABLEVT)
> outAngloEVT<-gpd(AngloEVT, 0.03915)
> plot(outAngloEVT)
> outMerafeEVT<-gpd(MerafeEVT,0.05485)
> plot(outMerafeEVT)
```

Figure C.5: R source codes for computing cumulative probabilities of the fitted NIG and t-distribution.

```
> a<-c(0, -0.0125000, -0.0250000, -0.0375000, -0.0500000, -0.0625000, -0.0750000, -0.0875000, -
0.1000000, -0.1500000, -0.2000000, -0.2500000)
> a1=a*100
> pnig(a, alpha=67.077482734, beta=-5.030372307, delta=0.012680344, mu=0.001398949)
> pt(a1,df=3.984366)

> pnig(a,alpha=43.837700000, beta=0.390100300, delta=0.021332960, mu=0.000248665)
> pt(a1,df=1.768262)

> pnig(a,alpha=34.914860000, beta=0.207370400, delta=0.025836300, mu=0.000080426)
> pt(a1,df=1.375823)

> pnig(a,alpha=42.329333877, beta=4.525437812 , delta=0.056368911, mu=-0.005736016)
> pt(a1,df=0.9775525)

> pnig(a,alpha=16.618440000, beta=0.376379500, delta=0.019603710, mu=-0.000846981)
> pt(a1,df=1.195069)

> pnig(a,alpha=42.265046320, beta=0.709176938, delta=0.027389859, mu=-0.000164029)
> pt(a1,df=1.45802)

> pnig(a,alpha=31.951865726, beta=-2.479612212, delta=0.024460343, mu=0.002125874)
> pt(a1,df=1.351801)
```

Bibliography

[AC03] A. Azzalini and A. Capitanio. Distributions generated by pertubation of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society*, 65(2): 367–389, May 2003.

[ADEH99] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3): 203–228, July 1999.

[AE01] F. M. Aparicio and J. Estrada. Empirical distributions of stock returns: European securities markets, 1990-95. *The European Journal of Finance*, 7: 1–21, 2001.

[AH06] K. Aas and D. H. Haff. The generalised hyperbolic skew student's t-distribution. *Journal of Financial Econometrics*, 4(4): 275–309, 2006.

[Ale08] C. Alexander. *Market Risk Analysis Value-at-Risk Models*, volume IV. John Wiley & Sons Ltd, 2008.

[App04] D. Applebaum. Lévy processes - from probability to finance and quantum groups. *Notices of the American Mathematical Society*, 51(11): 1336–1347, 2004.

[AS64] M. Abramowitz and I. A Stegun. *Handbook of Mathematical Functions*. Dover Publ. Inc. New York, 1964.

[Bau00] C. Bauer. Value at risk using hyperbolic distributions. *Journal of Economics and Business*, 52: 455–467, 2000.

[BB00] E. Bølviken and F. E. Benth. Quatification of risk in norwegian stocks via the normal inverse gaussian distribution. *Proceedings of the 10th AFIR Colloquium, Tromso, Norway*, pages 87–98, 2000.

[BdH74] A. A. Balkema and L. de Haan. Residual life time at great age. *The Annals of Probability*, 2(5): 792–804, Oct 1974.

[Ben04] F. E. Benth. *Option Theory with Stochastic Analysis: An Introduction to Mathematical Finance*. Springer-Verlag Berlin Heidelberg, 2004.

[Ber09] R. Berry. Back testing value-at-risk. Technical report, J.P Morgan Investment Analytics and Consulting, September 2009.

[Bes10] J. Bessis. *Risk Management in Banking*. Wiley Finance, third edition, 2010.

[BG74] R. C. Blattberg and N. J. Gonedes. A comparison of the stable and student distribution as statistical models for stock prices. *Journal of Business*, 47(2): 244–280, April 1974.

[BKM09] Z. Bodie, A. Kane, and A. J Marcus. *Investments*. McGraw-Hill, eight edition, 2009.

[BM12] M. Bhattachariya and S. Madhav. A comparison of var estimation procedures for leptokurtic equity index returns. *Journal of Mathematical Finance*, 3:13–30, August 2012.

[BN95] O. E. Barndorff-Nielsen. Normal inverse gaussian processes and the modelling of stock returns. *Research Reports*, (300), March 1995.

[BN97] O. E. Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modelling. *Board of the Foundation of the Scandinavian Journal of Statistics*, 24: 1–13, 1997.

[BN98] O. E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Journal of Finance and Stochastics*, 2: 41–68, 1998.

[BNKS82] O. Barndorff-Neilsen, J. Kent, and M. Sørensen. Normal variance-mean mixtures and z distributions. *International Statistical Review*, 50(2): 145–159, August 1982.

[Cap14] Courtney Capital. Jse top 40 shares. www.courtneycapital.co.za/jse-top-40-shares/, March 2014.

[CGM01] M. Crouhy, D. Galai, and R Mark. *Risk Management*. McGraw-Hill, 2001.

[Coc05] S. D. Cochran. *The subfigure Package*. Digital Mapping Laboratory, School of Computer Science, Carnegie-Mello University, 5000 Forves Avenue Pittsburgh, PA 15213-3890, USA, March 2005.

[DdV00] J. Danielsson and C. G de Vries. Value-at-risk and extreme returns. *Annals of Economics and Statistics*, (60): 239–270, Oct-Dec 2000.

[De06] Wuertz D and *et al. THE fBasics Package: Rmetrics - Markets and Basic Statistics*, December 2006.

[DH08] J. A. Doornik and H. Hansen. An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, 70: 927–936, November 2008.

[Dow02] K. Dowd. *An Introduction to Market Risk Measurement*. Wiley Finance, 2002.

[DS86] R. B. D'Agostino and M. A. Stephens. *Goodness-of-fit techniques*, volume 68. Marcel Dekker, INC, 1986.

[Ebe01] E. Eberlein. Application of Generalized Hyperbolic Lévy Motions to Finance. In O. E. Barndorff-Nielsen, S. I. Resnick, and T. Mikosch, editors, *Lévy Processes: Theory and Applications*, pages 319–336. Birkhäuser Boston, 2001.

[EK95] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1(3): 281–299, 1995.

[EKM97] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, 1997.

[Fam63] F. E. Fama. Mandelbrot and the stable paretian hypothesis. *The Journal of Business*, 36(4): 420–429, October 1963.

[Fam65] F. E. Fama. The behavior of stock-market prices. *The Journal of Business*, 38(1): 34–105, January 1965.

[FF03] S. M. Focardi and F. J Fabozzi. Fat tails, scaling, and stable laws: A critical look at modeling extremal events in financial phenomena. *The Journal of Risk Finance*, 5(1): 5–26, 2003.

[Gem02] H. Geman. Pure jump lévy processes for asset modelling. *Journal of Banking & Finance*, 26: 1297–1316, 2002.

[GR10] S. Ghosh and S. Resnick. A discussion on mean excess plots. *Stochastic Processes and their Applications*, 120: 1492–1517, 2010.

[GSU03] R. Gençay, F. Selçuk, and A. Ulugülyağci. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathametics and Economics*, 33: 337–356, May-July 2003.

[Hai10] J. Hain. Comparison of common tests of normality. Master's thesis, Julius-Maximmilians-Universität Würzburg, August 2010.

[Han94] B. E. Hansen. Autoregressive conditional density estimation. international economic review. *International Economic Review*, 35(3): 705–730, August 1994.

[HCCSH14] C-K. Huang, K. Chinhamu, Huang C-S, and J. Hammujuddy. Generalized hyperbolic distributions and value-at-risk estimation for the south african mining index. *International Business & Economics Research Journal*, 13(2): 320–328, March/April 2014.

[Hen10] P. N. Henriksen. *Computational Methods with Appications in Finance and Insurance*. PhD thesis, Department of Mathematics, University of Oslo, August 2010.

[HKP98] R. Huisman, K. Koedijk, and R. Pownall. Var-x: Fat tails in financial risk management. *Journal of Risk*, 1(1): 47–60, 1998.

[HP07] M. Haas and C. Pigorsch. Financial economics, fat-tailed distributions. Technical report, German Research Foundation, 2007.

[HPW61] D. Hogben, R. S. Pinkham, and M. B. Wilk. The moments of the non-central t-distribution. *Biometrika Trust*, 48(3/4):465–468, December 1961.

[Hub12] S. Hubber. *Essential mathematics for market risk management*. Wiley Finance, 2012.

[Hul09] J. C. Hull. *Option Futures, and Other Derivatives*. Pearson, 2009.

[Hul10] J. Hull. *Risk Management and Financial Institutions*. Pearson, second edition, 2010.

[HW98] J. Hull and A. White. Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(1), 1998.

[JF03] C. M. Jones and M. J. Faddy. A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society*, 65(1): 159–174, February 2003.

[Jor01] P. Jorion. *Value at Risk: The New Benchmark for Measuring Financial Risk*. New York: McGraw-Hill, 2nd edition, 2001.

[KMP06] K. Kuester, S. Mittnik, and M. S. Paolella. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2006.

[KNSS99] U. Küchler, K. Neumann, M. Sørensen, and Streller. Stock returns and hyperbolic distributions. *Mathematical and Computer Modelling*, 29: 1–15, 1999.

[Kup95] P. H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2):73–84, 1995.

[Kyp] A. E. Kyprianou. Lévy processes and continuous-state branching processes: Part I. Technical report, Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY.

[Kyp05] A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, 2005.

[Lil98] J. Lillestøl. Fat and skew: Can nig cure? on the prospects of using the normal inverse gaussian distribution in finance. *The Norwegian School of Economics and Business Adminstration*, 1998.

[Lil00] J. Lillestøl. Risk analysis and the nig distribution. *Journal of Risk*, 2(4): 41–56, 2000.

[Lon05] F. Longin. The choice of the distribution of asset returns: How extreme value theory can help? *Journal of Banking & Finance*, 29: 1017–1035, 2005.

[Man63] B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36(4): 394–419, 1963.

[McN99] A. J. McNeil. Extreme value theory for risk managers. May 1999.

[MEF05] A. J. McNeil, P. Embrechts, and R. Frey. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2005.

[MF00] A. J. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic finance time series: an extreme value approach. *Journal of Empiriacl Finance*, 7: 271–300, November 2000.

[Mil11] C. Milwidsky. Credit valuation adjustments with application to credit default swaps. Master's thesis, Department of Mathematics, University of Pretoria, August 2011.

[MM10] C. Milwidsky and E. Mare. Value at risk in the south african equity market: A view from the tails. *SAJEMS NS*, 13(3): 345–361, 2010.

[MRT09] V. Marimoutou, B. Raggad, and A. Trabelsi. Extreme value theory and value at risk: Application to oil market. *Energy Economics*, (31): 519–530, 2009.

[oBS96a] Basel Committee on Bank Supervision. Amendment to the capital accord to incorporate market risk, 1996.

[oBS96b] Basel Committee on Bank Supervision. Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements, 1996.

[oBS11] Basel Committee on Bank Supervision. Revisions to the basel ii market risk framework, February 2011.

[Ona10] O. Onalan. Financial risk management with normal inverse gaussian distributions. *Journal of Finance and Economics*, (38): 104–115, 2010.

[Pap05] A. Papapantoleon. An introduction to lévy processes with applications in finance. Technical report, 2005.

[Pei99] A. Peiro. Skewness in financial returns. *Journal of Banking & Finance*, 23:847–862, 1999.

[Pic75] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1): 119–131, 1975.

[PMS12] B. Pfaff, A. McNeil, and A. Stephenson. *Extreme Values in R*, July 2012.

[PR08] E. Platen and R. Rendek. Empirical evidence on student-t log-returns of diversified world stock indices. *Journal of Statistical Theory and Practice*, 2(2):233–251, 2008.

[Pra72] P. D. Praetze. The distribution of share price changes. *Journal of Business*, 45(1): 49–55, January 1972.

[Pra97] K. Prause. Modelling financial data using generalised hyperbolic distributions. (48), September 1997.

[Pra99] K. Prause. The generalized hyperbolic model: Estimation, financial derivatives, and risk measures. Master's thesis, Albert-Ludwigs-Universitat Freiburg, October 1999.

[Ris96] RiskMetrics. Technical document, fourth edition. Technical report, Morgan Guaranty Trust Company of New York, December 1996.

[Ros03] S. M. Ross. *Introduction to Probability Models*. 8 edition, 2003.

[Roy82] J. P Royston. An extension of Shapiro and Wilks' W test for Normality to large samples. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 31(2): 115–124, 1982.

[Roy95] J. P Royston. Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44(4): 547–551, 1995.

[RW11] N. M. Razali and Y. B. Wah. Power comparisons of shapiro-wilk, kolmogorovo-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analystics*, 2(1): 21–33, 2011.

[Ryd97] T. H. Rydberg. The normal inverse gaussian levy process: Simulation and approximation. *Communications in Statistics. Stochastic Models*, 13(4): 887–910, 1997.

[Ryd00] T. H. Rydberg. Realistic statistical modelling of financial data. *International Statistical Review*, 68(3): 233–258, December 2000.

[Sch12] J. Schaumburg. Predicting extreme value at risk: Nonparametric quantile-regression with refinements from extreme value theory. *Computational Statistics and Data Analysis*, 56:4081–4096, 2012.

[SW65] S. S. Shapiro and M. B. Wilk. An anlysis of variance test for normality (complete samples). *Biometrika*, 52(3/4): 591–611, December 1965.

[Tho02] H. C Thode. *Testing for Normality*. Marcel Dekker, Inc, 2002.

[VdJ01] J. H. Venter and P. J. de Jongh. Risk estimation using the normal inverse gaussian distribution. *Journal of Risk*, 4(2), February 2001.

[Wik12] *Lévy Processes*. Wikipedia, 2012.

[WM07] C. Wentzel and E. Mare. Extreme value theory - an application to the south african equity market. *Investment Analysts Journal*, (66): 73–77, 2007.