OPTION PRICING : A GARCH MODEL WITH LEVY PROCESS INNOVATIONS

by

GRANT PETER MOOLMAN

A DISSERTATION

submitted in fulfilment of the
requirements for the degree

MAGISTER SCIENTIAE

in

MATHEMATICAL STATISTICS

in the

FACULTY OF SCIENCE

at the

UNIVERSITY OF JOHANNESBURG

SUPERVISOR: DR. H.F. VAN ROOY
CO-SUPERVISOR: PROF. F. LOMBARD
April 2008

Acknowledgements

To my supervisors, Professor Freek Lombard and Doctor Herrie F. van Rooy, I owe much gratitude for your unrelenting guidance, time and sheer persistence. I am indebted to you for this remarkable opportunity. It has been an absolute privilege working with you. This journey that we have enjoyed this past year and a half was not without obstacles, but the insight gained and the lessons learnt, in both life and academically, are invaluable. You have taught me that the best way to overcome the problems we face is to be patient and persevere. It was John Quincy Adams who said it best:

> *"Patience and perseverance have a magical effect before which difficulties disappear and obstacles vanish."*

Finally, thanks must go to my family and friends. Your love and support has been unwavering. Although the contents of this dissertation are foreign to you, your constant curiosity and encouragement has kept me going. Lastly, thanks to my mother and father for affording me this opportunity to further my studies.

Table of Contents

List of Figures

List of Tables

Notation

$\mathbb{P}, \mathbb{Q}$	Probability measures.
$\mathcal{F}_t$	The filtration generated up until time t.
$\mathbb{E}^{\mathbb{Q}}[X]$	The expected value of the random variable X w.r.t. measure $\mathbb{Q}$.
$\mathbb{VAR}^{\mathbb{Q}}[X]$	The variance of the random variable X w.r.t. measure $\mathbb{Q}$.
$\mathbb{SKEW}^{\mathbb{Q}}[X]$	The skewness of the random variable X w.r.t. measure $\mathbb{Q}$.
$\mathbb{KURT}^{\mathbb{Q}}[X]$	The kurtosis of the random variable X w.r.t. measure $\mathbb{Q}$.
$\mathbb{COV}^{\mathbb{Q}}[X, Y]$	The covariance of random variables X and Y w.r.t. measure $\mathbb{Q}$.
$RNVR$	Risk Neutral Valuation Relationship.
$LRNVR$	Locally Risk Neutral Valuation Relationship.
$f_X(x; \theta)$	The density function of the random variable X with parameter(s) θ.
$F_X(x; \theta)$	The distribution function of the random variable X with parameter/s θ.
$N(0, 1)$	Standard normal distribution.
$\Phi(\cdot)$	Normal cumulative distribution function.
$\Phi'(\cdot)$	Normal density function.
$IG(a, b)$	Inverse Gaussian distribution with parameters (a, b).
$NIG(\alpha, \beta, \delta)$	Normal inverse Gaussian distribution with parameters (α, β, δ).
$stdNIG(\alpha, \beta, \delta)$	Normal inverse Gaussian distribution with zero mean and unit variance.

$Meixner(\alpha, \beta, \delta)$	Meixner distribution with parameters (α, β, δ).
$stdMeixner(\alpha, \beta, \delta)$	Meixner distribution with zero mean and unit variance.
χ^2_k	Chi-squared distribution with k degrees of freedom.
$\chi^2_{k;\alpha}$	The $100\alpha^{th}$ percentile of a chi-squared distribution with k degrees of freedom.
$ARCH$	Auto Regressive Conditional Heteroskedasticity.
$GARCH$	Generalized Auto Regressive Conditional Heteroskedasticity.
$\Gamma(\cdot)$	The gamma function.
$\mathbb{R}$	The real numbers.
$\mathbb{I}(A)$	Indicator function. If A is true, then $\mathbb{I}(A) = 1$, otherwise $\mathbb{I}(A) = 0$.
$K_\lambda(\cdot)$	Modified Bessel function of the third kind with index λ.
$\phi_X(u)$	Characteristic function of the random variable X.
$\kappa_X(u)$	Cumulant generating function of the random variable X.
$\sim$	Asymptotically equivalent to. ($a_n \sim b_n$ if and only if $\frac{a_n}{b_n} \to 1$ as $n \to \infty$)
$\overset{d}{\sim}$	Distributed as.
$\hat{se}_{boot}(\theta)$	Bootstrapped standard error of the parameter θ.
$O(\cdot)$	asymptotically bounded. ($a_n = O(b_n)$ if and only if $\frac{a_n}{b_n}$ is a bounded sequence)
$a \wedge b$	The minimum of a and b.
càdlàg	Continuous from the right and has limits from the left.

Chapter 1

Introduction

The seminal papers of Black and Scholes [8] and Merton [39] provided the first analytical formula for pricing European options, termed the Black-Scholes model. Since then a vast amount of research has been dedicated to improving the imperfections of the Black-Scholes model. Two assumptions postulated by the Black-Scholes model have come under much scrutiny, namely the assumption of normality for the log-returns and the homoskedastic volatility. Empirical studies have shown that the Black-Scholes model cannot deal with the volatility clustering and leptokurtosis observed in asset prices. It is widely accepted that the distribution of log-returns is skewed, peaked around the mean and heavy tailed (see Anderson, et al. [3], Bollen and Inder [9], Carr, Geman, Madan and Yor [13] and Cont [16]). Another critical point in the Black-Scholes framework is the requirement that continuous trading be possible.

A Quantile-Quantile plot (Q-Q plot) of residuals, which are assumed to be standard normal random variates from our fitted model are plotted in Figure 1.1. The deviation from normality is easily seen in the left tail in this Q-Q plot.

In the late eighties and early nineties the use of Lévy processes was proposed to relax the assumption of lognormal asset returns. Among these proposed Lévy processes are the variance gamma process of Madan and Seneta [35], the normal inverse Gaussian process of Barndorff-Nielson [4], the Meixner process of Schoutens [48] and the CGMY process

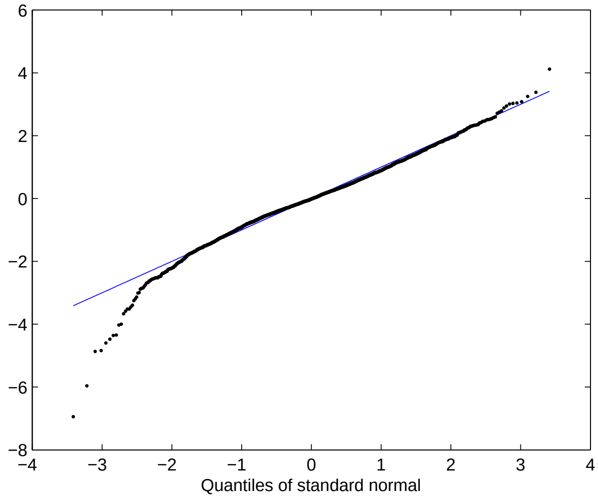

Figure 1.1: Q-Q plot of S&P 500 residuals.

of Carr, et al. [13]. Although these models provide adequate fits, one still had the assumption of homoskedastic volatility.

Numerous authors have dealt with heteroskedastic volatility models and discrete time processes in order to achieve a more realistic model.

The framework for the discrete time approach was provided by Rubinstein [44] and Brennan [12]. Rubinstein and Brennan introduced us to the Risk Neutral Valuation Relationship (RNVR). Rubinstein and Brennan assumed that all investors have the same characteristics as a representative investor and it is assumed that constant proportional risk aversion is exhibited by the representative investor.

Two important classes of volatility models are the continuous time stochastic volatility models and the generalised autoregressive conditional heteroskedastic (GARCH) models. A wide range of continuous time stochastic volatility models have been proposed. These include the jump diffusion model of Merton [40] and the bivariate diffusion models of Hull and White [32] and Heston [29]. Barndorff-Nielsen and Shephard [6] proposed the use of Ornstein-Uhlenbeck (OU) processes, driven by Lévy processes, to model volatility. These models are generally referred to as BNS models. A different procedure for incorporating stochastic volatility through the randomisation of time was proposed by Clark [15]. These models are referred to as time change models. Clark considered geometric Brownian motion subordinated by an independent Lévy subordinator (nonnegative nondecreasing Lévy process) for the stock price process. In Geman, Madan and Yor [25] the subordination or time change of Lévy processes was considered.

Engle [23] introduced the Auto-regressive Conditional Heteroskedastic (ARCH) process. The ARCH process was generalised by Bollerslev [10] and aptly named the Generalised Auto-regressive Conditional Heteroskedastic (GARCH) process. Since its introduction, the GARCH process has gained prominence for modeling financial time series. Many variants of the GARCH process have since been proposed, most notably the nonlinear asymmetric GARCH (NGARCH) process of Engle and Ng [24] which incorporates a leverage effect. The leverage effect refers to the negative correlation that exists between the asset return innovations and volatility innovations. The ability of the GARCH(1,1) process in modeling volatility was documented by Hansen and Lunde [28]. They compared over three hundred time series models and were unable to find conclusive evidence that the GARCH(1,1) model is outperformed by any of them.

Duan [21] provided the first rigorous theoretical foundation for option pricing using this powerful econometric model. Duan's model provided a connection between the heteroskedastic volatility approach and the discrete time approach. Duan proposed a GARCH process with Gaussian innovations for the volatility process. However, a more general form of the RNVR was required. Duan introduced us to the Locally Risk Neutral Valuation Relationship (LRNVR). The generalisation of the RNVR incorporated the condition that the conditional variances of the log-returns remain unchanged under a change from the real world measure to the risk neutral measure. Heston and Nandi [30] proposed a GARCH option pricing model (Hereafter, HN-GARCH) with a closed form solution for European options. Their model incorporated a very specific GARCH like process for the stochastic volatility. The HN-GARCH model contains a diffusion approximation equivalent to the diffusion model introduced by Heston [29]. Hence, Heston's stochastic volatility model is a continuous time limit of the HN-GARCH model. Although the HN-GARCH model obtains a closed form solution, an empirical comparison was performed by Hsieh and Ritchken [31] showing the HN-GARCH model is outperformed by a variant of Duan's GARCH model . The variant of Duan's GARCH model incorporated an NGARCH process for the stochastic volatility, thereby incorporating a leverage effect. However both models were capable of explaining the maturity and strike bias in the Black-Scholes model.

The main drawback to the GARCH process and Duan's GARCH model in volatility estimation and option pricing is the assumption of normality. Numerous papers in the volatility estimation literature deal with this assumption. These include Bollerslev [11], Barndorff-Nielsen [5] and Griebenow [26]. Recently, more general distributions have been proposed in an attempt to relax the assumption of normality in the GARCH option pricing literature. Duan, Ritchken and Sun [22] included jumps in the Duan model through a compound Poisson process (Poisson random sum of Gaussian random variables). Menn and Rachev [37], [38] proposed α-stable and smoothly truncated stable distributions. Other models proposed include modified tempered stable distributions (Kim, Rachev and Chung [33]) and Student-t and Paretian distributions (Curto, Pinto and Tavares [19]). Kim, Rachev and Chung [33] and Menn and Rachev [37] perform out-of-sample predictions using the maximum likelihood estimates from asset prices. Menn and Rachev [38] also perform in-sample calibrations in addition to the same out-of-sample predictions performed in the other two papers ([33], [37]). The empirical results in the three papers ([33], [37] and [38]) are very encouraging. Duan's GARCH model is regularly outperformed in the modeling of asset returns and the pricing of European options.

In this dissertation we attempt to relax the assumption of normality by incorporating infinitely divisible distributions, specifically the normal inverse Gaussian (NIG) and Meixner distributions, for the random innovations. These distributions have semi-heavy tails and can be skewed. Semi-heavy tails refers to the instance when the tails of a distribution are heavier than those of a Gaussian distribution but lighter than those of the non-Gaussian stable laws. The NIG and Meixner distributions provide much more flexibility through their three characterizing parameters. However, the incorporation of these distributions provides an additional restriction on the volatility process. This restriction is introduced in Chapter 6 and we will discuss the impact of this restriction in the empirical analysis in Chapter 7. We will term these models Lévy GARCH models.

The remainder of the dissertation is set out as follows: In Chapter 2 we introduce Duan's Gaussian GARCH option pricing model, the LRNVR and the risk-neutral GARCH model. In Chapter 3 we define the concept of Lévy processes and discusses their main properties. We define the NIG and Meixner distributions in Chapter 4 and Chapter 5 respectively. We discuss the properties of these distributions and describe methods of generating random numbers from these distributions. In Chapter 5 we formulate a new algorithm for generating Meixner random numbers using the rejection method (see Ross [43], p. 66). In Chapter 6 we formulate the Lévy GARCH model and more specifically the NIG-GARCH and Meixner-GARCH models. Their respective risk neutral versions are also introduced. In Chapter 7 we discuss matters regarding parameter estimation, goodness of fit and the calibration of option prices. This chapter is concluded with a presentation of results based on the data obtained from the S&P 500 and S&P 100.

Chapter 2

The GARCH Model

2.1 Introduction

Since the GARCH process of Bollerslev [10] was introduced, it has gained prominence for modeling financial time series. Duan [21] provided the first rigorous theoretical foundation for option pricing using this powerful econometric model. Due to the complex nature of the GARCH process, Duan [21] developed his GARCH option pricing model by extending the conventional risk neutralization in Rubinstein [44] and Brennan [12]. He called it the locally risk-neutral valuation relationship (LRNVR) (see Definition 2.3.1).

2.2 The Model

Duan [21] proposed the following model for the stock price process,

$$S_t = S_{t-\Delta t} \exp\left(r\Delta t - \frac{1}{2}h_t + \lambda\sqrt{h_t} + \sqrt{h_t}\epsilon_t\right) \tag{2.1}$$

where S_t denotes the stock price at time t and S_0 is known. Δt is the time unit (i.e. one minute, one day, etc.). The sequence $(\epsilon_t)_{t\in\mathbb{N}}$ consist of independent and identically distributed standard normal random variables, i.e. $\epsilon_t \overset{d}{\sim} N(0,1)\ \forall\, t \in \mathbb{N}$. λ is a positive real constant and denotes the market price of risk. r is the constant continuously compounded risk free interest rate. h_t is the conditional variance (squared stochastic volatility) process which follows a GARCH(1,1) process (see Bollerslev [10]),

$$h_{t+\Delta t} = \alpha_0 + \alpha_1 h_t \epsilon_t^2 + \beta_1 h_t \tag{2.2}$$

where α_0, α_1 and β_1 are non-negative and $\alpha_1 + \beta_1$ is assumed to be less than one to ensure covariance stationarity of the sequence.

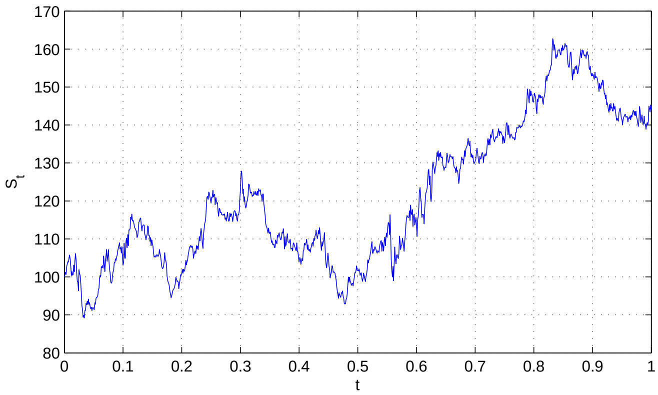

Figure 2.1: Simulated stock price path.

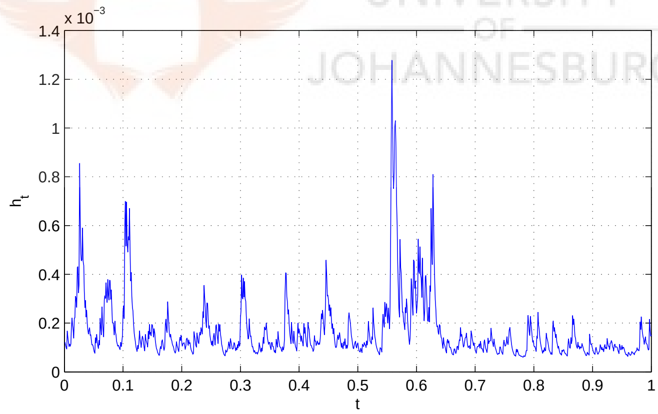

Figure 2.2: The variance path
which drives the stock price path in Figure 2.1.

Figures 2.1 and 2.2 plot a single realization of the GARCH model (equations 2.1 and 2.2). The parameter values are given by $(\alpha_0, \alpha_1, \beta_1, \lambda) = (1.524\text{x}10^{-5}, 0.188, 0.716, 0.007)$, $S_0 = 100$ and $r = 0$.

This GARCH model has two distinctive features. Firstly, in contrast to the general diffusion type models, which are Markovian, the GARCH model is non-Markovian with regards to the filtration generated by ϵ_t. Secondly, the GARCH model contrasts with standard preference-free option pricing, since the GARCH option price is a function of the risk premium embedded into the underlying asset.

Remark. If the coefficients α_1 and β_1 are zero. Then the GARCH model (equations (2.1) and (2.2)) reduces to the Black-Scholes discrete time model (see Black and Scholes [8] and Merton [39]). Hence, this includes the homoskedastic lognormal Black-Scholes model as a special case of the GARCH model.

2.3 The Risk Neutral Model

The model proposed in the previous section can not be applied to option pricing as it is not risk-neutral. We must therefore find the risk-neutral model. Before this is done, we first introduce Duan's [21] locally risk-neutral valuation relationship.

Definition 2.3.1 *A pricing measure $\mathbb{Q}$ is said to satisfy the locally risk-neutral valuation relationship ($LRNVR$) if the measures $\mathbb{Q}$ and $\mathbb{P}$ are mutually absolutely continuous and measure $\mathbb{Q}$ must also satisfy the following requirements:*

(i) *The following equation must hold for all $0 \leq t \leq T$*

$$\mathbb{E}^{\mathbb{Q}}[e^{-rt}S_t \mid \mathcal{F}_k] = e^{-rk}S_k$$

i.e. the discounted stock price process must be a $\mathbb{Q}$-martingale.

(ii) *The conditional variances of the logarithmic returns are unaffected by the change of measure.*

$$\mathbb{VAR}^{\mathbb{Q}}\left(\log\frac{S_{t+\Delta t}}{S_t}\middle|\mathcal{F}_t\right) \overset{\text{a.s.}}{=} \mathbb{VAR}^{\mathbb{P}}\left(\log\frac{S_{t+\Delta t}}{S_t}\middle|\mathcal{F}_t\right)$$

Duan extended the conventional risk-neutral valuation relationship in the aspect of variances, condition (ii). Under the LRNVR , the one-period ahead conditional variance, is invariant with respect to a change to the risk-neutral measure.

Lemma 2.3.2 *Define the stock price dynamics under measure $\mathbb{Q}$ as*

$$S_t \quad = \quad S_{t-\Delta t} \exp\left(r\Delta t - \frac{1}{2}h_t + \sqrt{h_t}\xi_t\right), \tag{2.3}$$

where $\xi_t = \epsilon_t + \lambda$ is a sequence of independent and identically distributed random variables whose distribution under $\mathbb{Q}$ equals that of ϵ_t under measure $\mathbb{P}$, namely $N(0,1)$. Note that $\xi_t \overset{d}{\sim} N(\lambda, 1)$ under measure $\mathbb{P}$. Then the conditional variance process, h_t, under measure $\mathbb{Q}$ has the following form,

$$h_{t+\Delta t} \quad = \quad \alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t \tag{2.4}$$

and then this stock price process satisfies the locally risk-neutral valuation relationship.

See Appendix 2.A.1 for the proof of this result.

In general, the conditional variance process is altered under an equivalent change of measure. Under the LRNVR the admissible equivalent measures are restricted to those under which the conditional variance of the log-returns remains unchanged. The squared volatility process, under the LRNVR, is not being driven by a chi-squared random variable ξ_t^2 but by a noncentral chi-squared random variable $(\xi_t - \lambda)^2$.

The GARCH process, after local risk-neutralization, is characterized by the following properties:

Theorem 2.3.3 *Under the risk-neutral measure $\mathbb{Q}$, if $|\lambda| < \sqrt{(1-\alpha_1-\beta_1)/\alpha_1}$, then*

(i) *The stationary variance of $\sqrt{h_t}\xi_t$ equals $\frac{\alpha_0}{1-(1+\lambda^2)\alpha_1-\beta_1}$.*

(ii) $\sqrt{h_t}\xi_t$ *is leptokurtic.*

(iii) $\mathbb{COV}^{\mathbb{Q}}(h_{t+\Delta t}, \xi_t) = \frac{-2\lambda\alpha_0\alpha_1}{1-(1+\lambda^2)\alpha_1-\beta_1}$.

Proof: See Appendix 2.A.2.

Under the original measure $\mathbb{P}$, the stationary variance of the GARCH return process is $\frac{\alpha_0}{1-\alpha_1-\beta_1}$ and the conditional variance, under measure $\mathbb{P}$, is uncorrelated with the lagged asset return (see Bollerslev [10]). Theorem 2.3.3(i), shows that the stationary variance is increased under local risk neutralization, since $\lambda > 0$. We also see that, under local risk neutralization, the conditional variance is negatively correlated with the lagged asset return.

Note that the unconditional variance or any conditional variance beyond one period is not invariant under a change in equivalent pricing measures. Thus, although the risk premium has been locally risk-neutralized under measure $\mathbb{Q}$, the latter measure still influences the conditional variance globally. In other words, local risk neutralization and global risk neutralization are not equivalent.

Due to the nature of European options (see Section 7.2), their pricing requires aggregating single period asset returns to obtain a random terminal asset price at some future date. We see that the terminal asset price of the GARCH model, S_T, can be expressed in terms of the initial asset price, S_0, by

$$S_T = S_0 \exp\left[rT - \frac{1}{2}\sum_i h_{i\Delta t} + \sum_i \sqrt{h_{i\Delta t}}\xi_{i\Delta t}\right], \tag{2.5}$$

where $i = 1, 2, ..., n$ and $n\Delta t = T$. This follows directly from equation (2.3).

2.A APPENDIX

2.A.1 PROOF OF LEMMA 2.3.2

To show this result holds, we must prove that the price dynamics defined by equations (2.3) and (2.4), satisfy the LRNVR conditions.

(i) Using the measurability of h_t with respect to $\mathcal{F}_{t-\Delta t}$, we get

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[e^{-rt}S_t|\mathcal{F}_{t-\Delta t}] &= \mathbb{E}^{\mathbb{Q}}[e^{-rt}S_{t-\Delta t}\exp(r\Delta t - \tfrac{1}{2}h_t + \sqrt{h_t}\xi_t)|\mathcal{F}_{t-\Delta t}] \\ &= e^{-r(t-\Delta t)-\frac{1}{2}h_t}S_{t-\Delta t}\mathbb{E}^{\mathbb{Q}}[\exp\left(\sqrt{h_t}\xi_t\right)|\mathcal{F}_{t-\Delta t}] \\ &= e^{-r(t-\Delta t)-\frac{1}{2}h_t}S_{t-\Delta t}e^{\frac{1}{2}h_t} \\ &= e^{-r(t-\Delta t)}S_{t-\Delta t}.\end{aligned}$$

Hence, S_t is a $\mathbb{Q}$-martingale.

(ii) To avoid any ambiguity, we will write h_t in equations (2.2) and (2.4) as h_t and h_t^* respectively.

$$\begin{aligned}\mathbb{VAR}^{\mathbb{Q}}\left(\log\frac{S_t}{S_{t-\Delta t}}\bigg|\mathcal{F}_{t-\Delta t}\right) &= \mathbb{VAR}^{\mathbb{Q}}\left(r\Delta t - \tfrac{1}{2}h_t^* + \sqrt{h_t^*}\xi_t|\mathcal{F}_{t-\Delta t}\right) \\ &= \mathbb{VAR}^{\mathbb{Q}}\left(\sqrt{h_t^*}\xi_t|\mathcal{F}_{t-\Delta t}\right) \\ &= \mathbb{VAR}^{\mathbb{Q}}\left(\sqrt{h_t}(\epsilon_t+\lambda)|\mathcal{F}_{t-\Delta t}\right) \\ &= h_t \end{aligned} \qquad (2.6)$$

where the third equality follows from setting $\xi_t = \epsilon_t + \lambda$. Similarly,

$$\begin{aligned}\mathbb{VAR}^{\mathbb{P}}\left(\log\frac{S_t}{S_{t-\Delta t}}\bigg|\mathcal{F}_{t-\Delta t}\right) &= \mathbb{VAR}^{\mathbb{P}}\left(r\Delta t + \lambda\sqrt{h_t} - \tfrac{1}{2}h_t + \sqrt{h_t}\epsilon_t|\mathcal{F}_{t-\Delta t}\right) \\ &= h_t\mathbb{VAR}^{\mathbb{P}}(\epsilon_t|\mathcal{F}_{t-\Delta t}) \\ &= h_t. \end{aligned} \qquad (2.7)$$

■

2.A.2 Proof Theorem 2.3.3

Before we prove the result, the following Lemma will play an important role.

Lemma 2.A.1 *Let h_t be defined as in equation (2.4). Then*

$$\begin{aligned} \mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t] &= \mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2] \\ &= \mathbb{E}^{\mathbb{Q}}[h_t]. \end{aligned} \tag{2.8}$$

Proof:

$$\begin{aligned} \mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t] &= \mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2] - \{\mathbb{E}^{\mathbb{Q}}[\sqrt{h_t}\xi_t]\}^2 \\ &= \mathbb{E}^{\mathbb{Q}}[h_t \underbrace{\mathbb{E}^{\mathbb{Q}}[\xi_t^2|\mathcal{F}_{t-\Delta t}]}_{=1}] - \{\mathbb{E}^{\mathbb{Q}}[\sqrt{h_t}\underbrace{\mathbb{E}^{\mathbb{Q}}[\xi_t|\mathcal{F}_{t-\Delta t}]}_{=0}]\}^2 \\ &= \mathbb{E}^{\mathbb{Q}}[h_t]. \end{aligned}$$

■

(i) For the variance of a process $\{X_t,\, t \geq 0\}$ to be stationary we require that

$$\mathbb{VAR}^{\mathbb{Q}}[X_t] = \mathbb{VAR}^{\mathbb{Q}}[X_s] \qquad \forall\, t \neq s. \tag{2.9}$$

Now, using Lemma 2.A.1

$$\begin{aligned} \mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t] &= \mathbb{E}^{\mathbb{Q}}[h_t] \\ &= \mathbb{E}^{\mathbb{Q}}[\alpha_0 + \alpha_1 h_{t-\Delta t}(\xi_{t-\Delta_t} - \lambda)^2 + \beta_1 h_{t-\Delta t}] \\ &= \alpha_0 + \alpha_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}(\xi_{t-\Delta t} - \lambda)^2] + \beta_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}] \\ &= \alpha_0 + \alpha_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}(\xi_{t-\Delta t}^2 - 2\xi_{t-\Delta t}\lambda + \lambda^2)] + \beta_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}] \\ &= \alpha_0 + \alpha_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}\mathbb{E}[\xi_{t-\Delta t}^2 - 2\lambda\xi_{t-\Delta t} + \lambda^2|\mathcal{F}_{t-\Delta t}]] + \beta_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}] \\ &= \alpha_0 + \alpha_1(1+\lambda^2)\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta_t}] + \beta_1\mathbb{E}^{\mathbb{Q}}[h_{t-\Delta t}] \\ &= \alpha_0 + \alpha_1(1+\lambda^2)\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_{t-\Delta t}}\xi_{t-\Delta t}] + \beta_1\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_{t-\Delta t}}\xi_{t-\Delta t}]. \end{aligned}$$

Now, setting

$$\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_{t-\Delta t}}\xi_{t-\Delta t}] = \mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t]$$

and solving for $\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t]$ yields

$$\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t] = \frac{\alpha_0}{1-(1+\lambda^2)\alpha_1-\beta_1}. \tag{2.10}$$

Since $\mathbb{VAR}^{\mathbb{Q}}[\sqrt{h_t}\xi_t] \geq 0$, it follows that

$$1-(1+\lambda^2)\alpha_1-\beta_1 > 0, \tag{2.11}$$

which implies

$$\begin{aligned}\lambda^2 &< \frac{1-\beta_1}{\alpha_1} - 1 \\ &= \frac{1-\beta_1-\alpha_1}{\alpha_1}.\end{aligned}$$

(ii) For $\sqrt{h_t}\xi_t$ to be leptokurtic we require that

$$\mathbb{E}^{\mathbb{Q}}[h_t^2\xi_t^4] \geq 3\{\mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2]\}^2. \tag{2.12}$$

Now,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[h_t^2\xi_t^4] &= \mathbb{E}^{\mathbb{Q}}[\mathbb{E}[h_t^2\xi_t^4|\mathcal{F}_{t-\Delta t}]] \\ &= \mathbb{E}^{\mathbb{Q}}[h_t^2\mathbb{E}[\xi_t^4|\mathcal{F}_{t-\Delta t}]] \\ &= 3\mathbb{E}^{\mathbb{Q}}[h_t^2] \end{aligned} \tag{2.13}$$

because ξ_t, conditional upon $\mathcal{F}_{t-\Delta t}$, is $N(0,1)$ under $\mathbb{Q}$. Furthermore,

$$\begin{aligned}3\{\mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2]\}^2 &= 3\{\mathbb{E}^{\mathbb{Q}}[\mathbb{E}[h_t\xi_t^2|\mathcal{F}_{t-\Delta t}]]\}^2 \\ &= 3\{\mathbb{E}^{\mathbb{Q}}[h_t\mathbb{E}[\xi_t^2|\mathcal{F}_{t-\Delta t}]]\}^2 \\ &= 3\{\mathbb{E}^{\mathbb{Q}}[h_t]\}^2. \end{aligned} \tag{2.14}$$

Therefore

$$
\begin{aligned}
& \mathbb{E}^{\mathbb{Q}}[h_t^2\xi_t^4] - 3\{\mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2]\}^2 \\
= \; & 3[\mathbb{E}^{\mathbb{Q}}[h_t^2] - \{\mathbb{E}^{\mathbb{Q}}[h_t]\}^2] \\
= \; & 3\mathbb{VAR}^{\mathbb{Q}}[h_t] \geq 0. \qquad (2.15)
\end{aligned}
$$

(iii) Since $\mathbb{E}^{\mathbb{Q}}[\xi_t] = \mathbb{E}^{\mathbb{Q}}[\xi_t^3] = 0$, we have

$$
\begin{aligned}
\mathbb{COV}^{\mathbb{Q}}(h_{t+\Delta t}, \xi_t) &= \mathbb{E}^{\mathbb{Q}}[h_{t+\Delta t}\xi_t] \\
&= \mathbb{E}^{\mathbb{Q}}[(\alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t)\xi_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\alpha_0\xi_t + \alpha_1 h_t(\xi_t^3 - 2\lambda\xi_t^2 + \lambda^2\xi_t) + \beta_1 h_t\xi_t] \\
&= \alpha_1\mathbb{E}^{\mathbb{Q}}[h_t\mathbb{E}[\xi_t^3 - 2\lambda\xi_t^2 + \lambda^2\xi_t|\mathcal{F}_{t-\Delta t}]] \\
&= -2\alpha_1\lambda\mathbb{E}^{\mathbb{Q}}[h_t].
\end{aligned}
$$

■

Chapter 3

Levy Processes

3.1 Introduction

Lévy processes are aptly named after French mathematician Paul Lévy (1886-1971), one of the founding fathers of probability theory and the modern theory of stochastic processes. Lévy made substantial contributions to the study of infinitely divisible laws and pioneered the study of processes with independent and stationary increments, now known as Lévy processes (see Loève [34]).

The Wiener process and the Poisson process are fundamental examples of Lévy processes. All Lévy processes are superpositions of a Wiener process and a number of independent Poisson processes.

More recently, the use of infinitely divisible distributions in modeling financial time series has been proposed. Madan and Senata [35] introduced the variance gamma distribution as a model for stock returns. Other distributions introduced to model stock returns include the normal inverse Gaussian distribution by Barndorff-Nielson [4] and the Meixner distribution by Schoutens and Teugels [51].

3.2 Definition and Properties

Definition 3.2.1 (Lévy Process) *A càdlàg stochastic process* $(X_t)_{t\geq 0}$ *on* $(\Omega,\mathcal{F},\mathbb{P})$ *with values in* $\mathbb{R}$ *such that* $X_0 = 0$*, is called a Lévy processes if it possesses the following properties*:

(i) $(X_t)_{t\geq 0}$ *has independent increments:* $X_t - X_s$ *is independent of* $\mathcal{F}_s$ *for all* $0 \leq s < t$, *where* $\mathcal{F}_s$ *is the history of the process up to time* s.

(ii) $(X_t)_{t\geq 0}$ *has stationary increments:* $X_t - X_s$ *has the same distribution as* X_{t-s} *for all* $0 \leq s < t$.

(iii) $(X_t)_{t\geq 0}$ *is stochastically continuous* (*continuous in probability*)*:* $\lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0$, *for every* $\epsilon > 0$ *and* $t > 0$.

Condition (iii) does not imply that the paths of the process $(X_t)_{t\geq 0}$ are continuous. In fact, it serves only to exclude processes that have jumps at nonrandom (fixed) times. Thus, given a time point t, the probability of a Lévy process jumping at t is zero.

The jump of the process $(X_t)_{t\geq 0}$ at time t, is defined as,

$$\Delta X_t = X_t - X_{t-}, \tag{3.1}$$

where

$$X_{t-} := \lim_{s\uparrow t} X_s. \tag{3.2}$$

Definition 3.2.2 (Infinite Divisibility) *A distribution* F *on* $\mathbb{R}$ *is said to be infinitely divisible if, for every* $n \geq 2$, *there exists* n *independent and identically distributed random variables* $X_1^{(n)}, ..., X_n^{(n)}$ *such that* $\sum_{i=1}^{n} X_i^{(n)}$ *has distribution* F.

Let X be a random variable with distribution function F. The characteristic function of X (equivalently, of F), $\phi(u)$, is defined as

$$\phi(u) = \mathbf{E}[e^{iuX}], \qquad u \in \mathbb{R}. \tag{3.3}$$

Infinite divisibility may then be reformulated in terms of characteristic functions.

Lemma 3.2.3 *F (or X) is infinitely divisible if and only if for every $n \geq 2$ there exists a characteristic function $\phi_n(u)$ such that*

$$\phi(u) = [\phi_n(u)]^n. \tag{3.4}$$

Common examples of distributions that are infinitely divisible are the Poisson distribution, the gamma distribution and the normal distribution. A random variable, Y, with any of these distributions can be written as $Y_1^{(n)}+...+Y_n^{(n)}$, where $Y_i^{(n)}$ $i \in \{1,..,n\}$ has the same distribution as Y but with modified parameters. For example if $Y \overset{d}{\sim} gamma(\alpha, \beta)$, then $Y = Y_1^{(n)} + ... + Y_n^{(n)}$ where $Y_i^{(n)} \overset{d}{\sim} gamma(\alpha/n, \beta)$.

Infinite divisibility has a strong relation to Lévy processes, namely

Proposition 3.2.4 *Let $(X_t)_{t\geq 0}$ be a Lévy process. Then, for every t, X_t has an infinitely divisible distribution. Conversely, if F is an infinitely divisible distribution, then there exists a Lévy process (X_t) such that the distribution of X_1 is given by F.*

For a proof of this result see Cont and Tankov [17], p. 69.

Define

$$N_t(A) = \sum_{s\leq t} \mathbb{I}[\Delta X_s \in A], \tag{3.5}$$

where A is a set bounded away from 0 (does not contain 0 as a limit point). The jump behavior of a Lévy process is dictated by its Lévy measure, which is defined next.

Definition 3.2.5 (Lévy measure) *Let $(X_t)_{t\geq 0}$ be a Lévy process. The Lévy measure ν on $\mathbb{R}$ is defined by*

$$\nu(A) = \mathbb{E}[N_1(A)] = \mathbb{E}[N_t(A)]/t. \tag{3.6}$$

The Lévy measure, ν, is a measure on $\mathbb{R}$, that satisfies the following conditions:

$$\int_{\mathbb{R}} (|x|^2 \wedge 1)\nu(dx) < \infty \quad \text{and} \quad \nu(\{0\}) = 0. \tag{3.7}$$

Thus we see that the Lévy measure of a set A is just the expected number of jumps per unit time whose size belongs to A . If the Lévy measure is of the form $\nu(dx) = u(x)dx$, then $u(x)$ is called the Lévy density.

For the proof of the integrability condition (3.7), see Appendix 3.A.1. But first an important theorem is required, namely the Lévy-Itô decomposition theorem.

Theorem 3.2.6 (Lévy-Itô decomposition theorem) *Let $(X_t)_{t\geq 0}$ be a Lévy process and let N_t be given by (3.5). Then N_t is a Poisson process and*

$$\begin{aligned} X_t &= \gamma t + \sigma W_t + \int_{0<|x|\leq 1} x[N_t(dx) - t\nu(dx)] \\ &\quad + \int_{|x|>1} x N_t(dx) \end{aligned} \tag{3.8}$$

where $\gamma \in \mathbb{R}$, $\sigma > 0$ and W_t is a standard Brownian motion, statistically independent of N_t.

See Sato [46], p. 125 for a proof of this result.

The Lévy-Itô decomposition theorem shows that the form of a general Lévy process consists of three parts (see (3.8)): a deterministic part (γt), a Brownian part (σW_t) and a pure jump part. The Lévy measure $\nu(dx)$ dictates the frequency and sizes of the jumps in the process.

Every infinitely divisible distribution has a triplet of Lévy characteristics or Lévy triplet.

Definition 3.2.7 (Lévy Triplet) *Let X_t be a Lévy process with decomposition (3.8). Then*

$$(\gamma, \sigma^2, \nu) \tag{3.9}$$

is called the Lévy triplet of X_t.

Lemma 3.2.8 *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy measure $\nu(dx)$ and $\sigma^2 = 0$. Then the paths of the process have finite variation if and only if*

$$\int_{-1}^{1} |x|\nu(dx) < \infty \tag{3.10}$$

and a finite number of jumps occur in every finite time interval if and only if

$$\int_{-1}^{1} \nu(dx) < \infty. \tag{3.11}$$

3.3 Lévy-Khintchine Representation

Theorem 3.3.1 (Lévy-Khintchine representation) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with Lévy triplet (γ, σ^2, ν). Then*

$$\phi_X(u) \;=\; e^{t\psi_X(u)}, \qquad u \in \mathbb{R} \tag{3.12}$$

where

$$\begin{aligned}\psi_X(u) \;&=\; iu\gamma - \frac{u^2\sigma^2}{2} + \int_{0<|x|\leq 1} (e^{iux} - 1 - iux)\nu(dx) \\ &\quad + \int_{|x|>1} (e^{iux} - 1)\nu(dx)\end{aligned} \tag{3.13}$$

Proof: See Appendix 3.A.2.

3.A Appendix

3.A.1 Proof of Condition (3.7)

It is sufficient to show that, for some $\epsilon > 0$, $\int_{|x|\leq\epsilon} |x|^2 \nu(dx) < \infty$, since the Lévy measure of any closed set not containing zero is finite.

Let $(X_t)_{t\geq 0}$ be a Lévy process and define

$$X_t^* = \int_{\delta\leq|x|\leq\epsilon} x[N_t(dx) - t\nu(dx)] \quad \text{and} \quad Y_t = X_t - X_t^*.$$

Then (X_t^*) and (Y_t) are Lévy processes (from the Lévy-Itô decomposition theorem), also (Y_t) and (X_t^*) are independent. We have,

$$\begin{aligned} |\mathbb{E}[e^{iuX_t}]| &= |\mathbb{E}[e^{iuY_t}]\mathbb{E}[e^{iuX_t^*}]| \\ &= |\mathbb{E}[e^{iuY_t}]| \cdot |\mathbb{E}[e^{iuX_t^*}]| \\ &\leq |\mathbb{E}[e^{iuX_t^*}]| \end{aligned}$$

because

$$\begin{aligned} |\mathbb{E}[e^{iuY_t}]| &\leq \mathbb{E}[|e^{iuY_t}|] \\ &= \mathbb{E}[1] \\ &= 1. \end{aligned} \tag{3.14}$$

For some u and t, $|\mathbb{E}[e^{iuX_t}]| > 0$ because a characteristic function cannot equal zero for all u and t. Thus $|\mathbb{E}[e^{iuX_t^*}]|$ is bounded below by some positive number $C < 1$ independent of δ.

The following identity plays an important role in the proof:

$$\frac{2(1-\cos at)}{a^2t^2} = \left(\frac{\sin\frac{at}{2}}{\frac{at}{2}}\right)^2; \tag{3.15}$$

see Chung [14], p. 138.

Now

$$\begin{aligned}
|\mathbb{E}[e^{iuX_t^*}]| &= |\mathbb{E}[\exp(iu\int_{\delta\le|x|\le\epsilon} x[N_t(dx)-t\nu(dx)])]| \\
&= |\exp(t\int_{\delta\le|x|\le\epsilon}[\exp(iux)-iux-1]\nu(dx))| \\
&= |\exp(t\int_{\delta\le|x|\le\epsilon}[\cos(ux)+i\sin(ux)-iux-1]\nu(dx))| \\
&= |\exp(t\int_{\delta\le|x|\le\epsilon}[\cos(ux)-1]\nu(dx))\exp(it\int_{\delta\le|x|\le\epsilon}[\sin(ux)-ux]\nu(dx))| \\
&= \exp(t\int_{\delta\le|x|\le\epsilon}[\cos(ux)-1]\nu(dx)),
\end{aligned}$$

where the second equality follows from the fact that

$$\mathbb{E}[\exp(iu\int_{\delta\le|x|\le\epsilon} x[N_t(dx)-t\nu(dx)])] = \exp(t\int_{\delta\le|x|\le\epsilon}[\exp(iux)-iux-1]\nu(dx)),$$

see Appendix 3.A.2 for proof of this result (equations (3.19), (3.21) and (3.23)), and the fifth equality follows from the identity

$$|e^{i\theta}| = 1 \text{ for all } \theta\in\mathbb{R}. \tag{3.16}$$

Therefore,

$$0 < C \le \exp(t\int_{\delta\le|x|\le\epsilon}[\cos(ux)-1]\nu(dx)) \le 1$$

because $\cos(ux)\le 1$ for all u and x. Taking the logarithm and multiplying throughout by $-\frac{1}{t}$ yields

$$0 \le \int_{\delta\le|x|\le\epsilon}[1-\cos(ux)]\nu(dx) \le \tilde{C},$$

where $\tilde{C} = -\frac{\log C}{t} > 0$ because $0 < C < 1$ which implies that $-\infty < \log C < 0$.

Since $\tilde{C}$ is independent of δ, the preceding inequalities imply that

$$0 \le \lim_{\delta\downarrow 0}\int_{\delta\le|x|\le\epsilon}[1-\cos(ux)]\nu(dx) \le \tilde{C},$$

is equivalent to

$$0 \leq \int_{0\leq|x|\leq\epsilon} [1-\cos(ux)]\nu(dx) \leq \tilde{C},$$

i.e. to

$$0 \leq \int_{0\leq|x|\leq\epsilon} \frac{u^2x^2}{2}\left(\frac{\sin\frac{ux}{2}}{\frac{ux}{2}}\right)^2\nu(dx) \leq \tilde{C},$$

from (3.15). Now, since

$$\left(\frac{\sin\frac{ux}{2}}{\frac{ux}{2}}\right)^2 \to 1 \quad \text{as} \quad \frac{ux}{2} \to 0, \tag{3.17}$$

there exists $a_0 > 0$ such that

$$\left(\frac{\sin\frac{ux}{2}}{\frac{ux}{2}}\right)^2 > \frac{1}{2} \quad \text{for all} \quad \frac{|ux|}{2} < a_0. \tag{3.18}$$

Choose u such that $\frac{2a_0}{x} > \epsilon$. Then

$$0 \leq \frac{u^2}{4}\int_{0\leq|x|\leq\epsilon} x^2\nu(dx) \leq \int_{0\leq|x|\leq\epsilon} \frac{u^2x^2}{2}\left(\frac{\sin\frac{ux}{2}}{\frac{ux}{2}}\right)^2\nu(dx) \leq \tilde{C}.$$

Therefore,

$$\int_{0\leq|x|\leq\epsilon} x^2\nu(dx) \leq \frac{4\tilde{C}}{u^2} < \infty.$$

■

Outlines of this proof are given in Cont and Tankov [17], p. 82.

3.A.2 PROOF OF THEOREM 3.3.1

We can write X_{t+s} as

$$X_{t+s} = (X_{t+s} - X_s) + X_s$$

Therefore,

$$\begin{aligned} \phi_{X_{t+s}}(u) &= \phi_{(X_{t+s}-X_s)+X_s}(u) \\ &= \phi_{X_{t+s}-X_s}\phi_{X_s}(u) \\ &= \phi_{X_t}\phi_{X_s}(u), \end{aligned}$$

where the second and third equalities follow from the properties of a Lévy process. It follows that

$$\log \phi_{X_{t+s}} = \log \phi_{X_t}(u) + \log \phi_{X_s}(u).$$

We have shown that $\phi_{X_t}(u)$ is linear in t, hence there exists a number $\psi(u)$ such that

$$\log \phi_{X_t}(u) = t\psi(u) \qquad \forall\, t \geq 0.$$

Therefore

$$\phi_{X_t}(u) = e^{t\psi(u)}.$$

From the Lévy-Itô decomposition theorem, we can represent $(X_t)_{t\geq 0}$ as

$$X_t = \gamma t + \sigma W_t + \int_{0<|x|\leq 1} x\{N_t(dx) - t\nu(dx)\} + \int_{|x|>1} xN_t(dx).$$

It then follows that

$$\begin{aligned}
\phi_{X_t}(u) &= \mathbb{E}[\exp(iuX_t)] \\
&= \mathbb{E}[\exp(iu\{\gamma t + \sigma W_t + \int_{0<|x|\leq 1} x\{N_t(dx) - t\nu(dx)\} + \int_{|x|>1} xN_t(dx)\})] \\
&= \mathbb{E}[\exp(iu\gamma t + iu\sigma W_t)]\mathbb{E}[\exp(iu\int_{0<|x|\leq 1} x\{N_t(dx) - t\nu(dx)\})] \\
&\quad \cdot\mathbb{E}[\exp(iu\int_{|x|>1} xN_t(dx))].
\end{aligned}$$

Now

$$\begin{aligned}
\mathbb{E}[\exp(iu\gamma t + iu\sigma W_t)] &= \exp(iu\gamma t)\mathbb{E}[\exp(iu\sigma W_t)] \\
&= \exp(iu\gamma t - u^2\sigma^2 t/2),
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[\exp(iu\int_{0<|x|\leq 1} x\{N_t(dx) - t\nu(dx)\})] \\
&= \lim_{n\to\infty} \mathbb{E}[\exp(iu\sum_{|k|=1}^{n-1} \frac{k}{n}\{N_t(\frac{k}{n}, \frac{k+1}{n}) - t\nu(\frac{k}{n}, \frac{k+1}{n})\})]
\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}
&\mathbb{E}[\exp(iu\int_{|x|>1} xN_t(dx))] \\
= &\lim_{n\to\infty} \mathbb{E}[\exp(iu\sum_{|k|=n}^{\infty} \frac{k}{n} N_t(\frac{k}{n}, \frac{k+1}{n}))],
\end{aligned} \tag{3.20}$$

where the expectations on the right hand side of equations (3.19) and (3.20) are given by

$$\begin{aligned}
&\mathbb{E}[\exp(iu\sum_{|k|=1}^{n-1} \frac{k}{n}\{N_t(\frac{k}{n}, \frac{k+1}{n}) - t\nu(\frac{k}{n}, \frac{k+1}{n})\})] \\
= &\prod_{|k|=1}^{n-1} \mathbb{E}[\exp(iu\frac{k}{n}\{N_t(\frac{k}{n}, \frac{k+1}{n}) - t\nu(\frac{k}{n}, \frac{k+1}{n})\})] \\
= &\prod_{|k|=1}^{n-1} \sum_{y\geq 0} \exp(iu\frac{k}{n}\{y - t\nu(\frac{k}{n}, \frac{k+1}{n})\}) \exp(-t\nu(\frac{k}{n}, \frac{k+1}{n}))[t\nu(\frac{k}{n}, \frac{k+1}{n})]^y/y! \\
= &\prod_{|k|=1}^{n-1} \exp(-t[iu\frac{k}{n} + 1]\nu(\frac{k}{n}, \frac{k+1}{n}))\sum_{y\geq 0}[\exp(iu\frac{k}{n})t\nu(\frac{k}{n}, \frac{k+1}{n})]^y/y! \\
= &\prod_{|k|=1}^{n-1} \exp(-t[iu\frac{k}{n} + 1]\nu(\frac{k}{n}, \frac{k+1}{n})) \exp(\exp(iu\frac{k}{n})t\nu(\frac{k}{n}, \frac{k+1}{n})) \\
= &\prod_{|k|=1}^{n-1} \exp(t[\exp(iu\frac{k}{n}) - iu\frac{k}{n} - 1]\nu(\frac{k}{n}, \frac{k+1}{n})) \\
= &\exp(t\sum_{|k|=1}^{n-1}[\exp(iu\frac{k}{n}) - iu\frac{k}{n} - 1]\nu(\frac{k}{n}, \frac{k+1}{n}))
\end{aligned} \tag{3.21}$$

and

$$\begin{aligned}
&\mathbb{E}[\exp(iu\sum_{|k|=n}^{\infty} \frac{k}{n} N_t(\frac{k}{n}, \frac{k+1}{n}))] \\
= &\prod_{|k|=n}^{\infty} \mathbb{E}[\exp(iu\frac{k}{n} N_t(\frac{k}{n}, \frac{k+1}{n}))] \\
= &\prod_{|k|=n}^{\infty} \sum_{y\geq 0} \exp(iu\frac{k}{n} y) \exp(-t\nu(\frac{k}{n}, \frac{k+1}{n}))[t\nu(\frac{k}{n}, \frac{k+1}{n})]^y/y!
\end{aligned}$$

$$
\begin{aligned}
&= \prod_{|k|=n}^{\infty} \exp(-t\nu(\frac{k}{n}, \frac{k+1}{n})) \sum_{y\geq 0}[\exp(iu\frac{k}{n})t\nu(\frac{k}{n}, \frac{k+1}{n})]^y/y! \\
&= \prod_{|k|=n}^{\infty} \exp(-t\nu(\frac{k}{n}, \frac{k+1}{n})) \exp(\exp(iu\frac{k}{n})t\nu(\frac{k}{n}, \frac{k+1}{n})) \\
&= \prod_{|k|=n}^{\infty} \exp(t[\exp(iu\frac{k}{n}) - 1]\nu(\frac{k}{n}, \frac{k+1}{n})) \\
&= \exp(t\sum_{|k|=n}^{\infty}[\exp(iu\frac{k}{n}) - 1]\nu(\frac{k}{n}, \frac{k+1}{n})) \qquad (3.22)
\end{aligned}
$$

respectively. Substituting (3.21) and (3.22) into equations (3.19) and (3.20) respectively and taking the limit yields the following:

$$
\begin{aligned}
&\mathbb{E}[\exp(iu\int_{0<|x|\leq 1} x\{N_t(dx) - t\nu(dx)\})] \\
&= \exp(t\int_{0<|x|\leq 1}[\exp(iux) - iux - 1]\nu(dx)) \qquad (3.23)
\end{aligned}
$$

and

$$
\mathbb{E}[\exp(iu\int_{|x|>1} xN_t(dx))] = \exp(t\int_{|x|>1}[\exp(iux) - 1]\nu(dx)) \qquad (3.24)
$$

by the definition of a Riemann-Stieljes integral.

Hence,

$$
\begin{aligned}
\phi_{X_t}(u) &= \exp(iu\gamma t - u^2\sigma^2 t/2 + t\int_{0<|x|\leq 1}[\exp(iux) - iux - 1]\nu(dx) \\
&\quad + t\int_{|x|>1}[\exp(iux) - 1]\nu(dx)) \\
&= \exp(t\{iu\gamma - u^2\sigma^2/2 + \int_{\mathbb{R}}[\exp(iux) - 1 - iux\mathbb{I}_{(0<|x|\leq 1)}]\nu(dx)\}).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\psi_X(u) &= iu\gamma - \frac{u^2\sigma^2}{2} + \int_{0<|x|\leq 1}(e^{iux} - 1 - iux)\nu(dx) \\
&\quad + \int_{|x|>1}(e^{iux} - 1)\nu(dx).
\end{aligned}
$$

■

Chapter 4

The Normal Inverse Gaussian Distribution

4.1 Introduction

In this chapter we present the normal inverse Gaussian distribution and some of the properties this distribution contains. The normal inverse Gaussian (NIG) distribution is a three parameter distribution, introduced by Barndorff-Nielsen [4]. This distribution often fits asset returns quite well (see Barndorff-Nielsen [4],[5] and Rydberg [45]). This chapter is concluded with a description on the method used to generate NIG random numbers.

4.2 Definition and properties

Before introducing the normal inverse Gaussian distribution, we need the following definition:

Definition 4.2.1 (IG density function, Schoutens [50], p. 53) *A random variable is inverse Gaussian distributed with parameters $\boldsymbol{\theta} = (a, b)$ if it has the probability density function*

$$f(x; \boldsymbol{\theta}) = \frac{a}{\sqrt{2\pi}} \exp(ab) x^{-3/2} \exp(-\frac{1}{2}(a^2 x^{-1} + b^2 x)), \tag{4.1}$$

where

$$x \in (0, \infty), \quad a \in (0, \infty), \quad b \in (0, \infty).$$

This distribution is denoted $X \overset{d}{\sim} IG(a, b)$.

The normal inverse Gaussian (NIG) distribution can be obtained as a mixture of independently distributed random variables from the normal and inverse Gaussian distributions.

Definition 4.2.2 *Let* $Z \overset{d}{\sim} N(0,1)$ *and* $Y \overset{d}{\sim} IG(\delta, \sqrt{\alpha^2 - \beta^2})$. *Then*

$$X = \beta Y + \sqrt{Y} Z \tag{4.2}$$

has a $NIG(\alpha, \beta, \delta)$ *distribution.This distribution is denoted as* $X \overset{d}{\sim} NIG(\alpha, \beta, \delta)$.

This is more commonly referred to as the normal-variance mixture representation of the normal inverse Gaussian distribution.

Proposition 4.2.3 (NIG density function, Schoutens [50], p. 60) *A random variable* X *which is normal inverse Gaussian distributed with parameters* $\boldsymbol{\theta} = (\alpha, \beta, \delta)$ *has the probability density function*

$$f(x; \boldsymbol{\theta}) = \frac{\alpha\delta}{\pi} \exp(\delta\sqrt{\alpha^2 - \beta^2} + \beta x) \frac{K_1(\alpha\sqrt{\delta^2 + x^2})}{\sqrt{\delta^2 + x^2}} \tag{4.3}$$

where

$$x \in \mathbb{R}, \quad \alpha \in (0, \infty), \quad \beta \in (-\alpha, \alpha), \quad \delta \in (0, \infty)$$

and $K_1(\cdot)$ *is the modified Bessel function of the third kind with index 1.*

Proof: See Appendix 4.A.1 for the proof of this result.

It follows from equation (4.2) that the conditional distribution of X given $Y = y$ is $N(\beta y, y)$. This result will play an important role in generating normal inverse Gaussian random numbers (see Section 4.3).

Lemma 4.2.4 *Let* $X \overset{d}{\sim} NIG(\alpha, \beta, \delta)$. *Then the expected value, variance, skewness and kurtosis of* X *are respectively given by*

$$\begin{aligned} \mathbb{E}[X] &= \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} && (4.4) \\ \mathbb{VAR}[X] &= \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{3/2}} && (4.5) \\ \mathbb{SKEW}[X] &= \frac{3\beta}{\alpha\sqrt{\delta}(\alpha^2 - \beta^2)^{1/4}} && (4.6) \\ \mathbb{KURT}[X] &= 3\left(1 + \frac{4\beta^2 + \alpha^2}{\delta\alpha^2\sqrt{(\alpha^2 - \beta^2)}}\right). && (4.7) \end{aligned}$$

See Schoutens [50], p. 60 for the above results.

Moments of all orders exist for the NIG distribution. Looking at the kurtosis of the NIG distribution, it can clearly be seen that it is always greater than the kurtosis of the Normal distribution, which equals 3.

Figures 4.1, 4.2 and 4.3 illustrate the effect of the parameters on the normal inverse Gaussian distribution. The kurtosis of the density function is influenced by all the parameters of the distribution. However, the skewness (symmetry) of the density function is described by the parameter β (see Figure 4.2). The distribution is skewed to the left for $\beta < 0$, skewed to the right for $\beta > 0$ and symmetric for $\beta = 0$.

Definition 4.2.5 *Let* X *be a random variable with characteristic function* $\phi(u)$. *Then the cumulant generating function,* $\kappa(u)$, *is defined as*

$$\kappa(u) = \log \mathbb{E}[e^{uX}] = \log \phi(-iu). \tag{4.8}$$

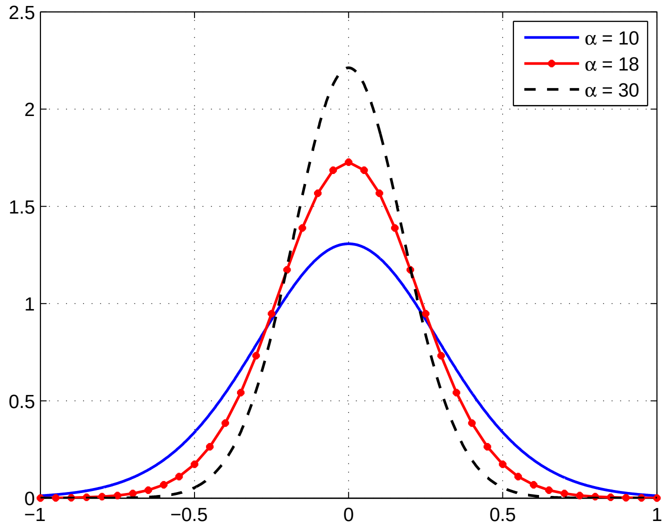

Figure 4.1: The effect of α in the NIG density.
$\alpha \in \{10, 18, 30\}, \beta = 0, \delta = 1.$

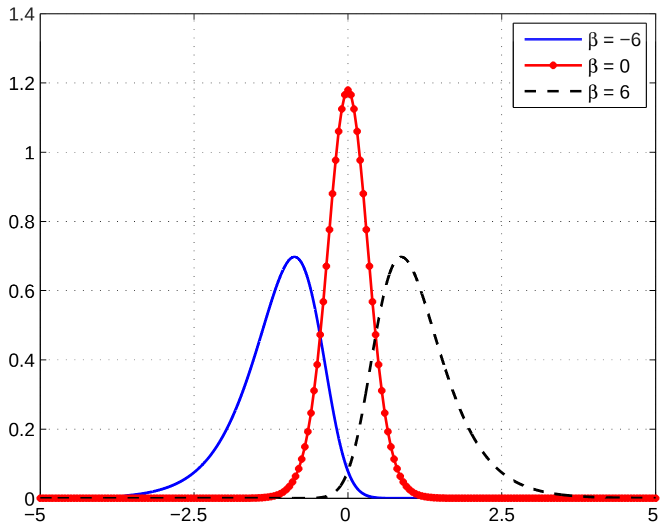

Figure 4.2: The effect of β in the NIG density.
$\alpha = 8, \beta \in \{-6, 0, 6\}, \delta = 1.$

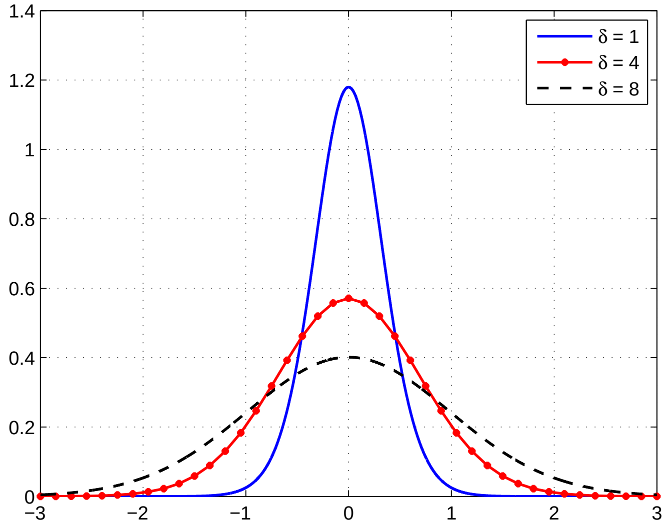

Figure 4.3: The effect of δ in the NIG density.
$\alpha = 8, \beta = 0, \delta \in \{1, 4, 8\}$.

Lemma 4.2.6 *Let* $X \overset{d}{\sim} NIG(\alpha, \beta, \delta)$. *Then the characteristic function,* $\phi(u)$, *and the cumulant generating function,* $\kappa(u)$, *are respectively given by*

$$\phi(u; \boldsymbol{\theta}) \quad = \quad \exp\left(\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2}\right)\right) \tag{4.9}$$

and

$$\kappa(u; \boldsymbol{\theta}) \quad = \quad \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}\right) \tag{4.10}$$

where $i = \sqrt{-1}$ *and* $u \in \mathbb{R}$.

Proof: See Appendix 4.A.2 and 4.A.3.

Considering the form of the characteristic function of a NIG random variable (4.9), we see that the $NIG(\alpha, \beta, \delta)$ distribution is infinitely divisible, since

$$\phi(u; \alpha, \beta, \delta) = [\phi(u; \alpha, \beta, \delta/n)]^n. \tag{4.11}$$

The tail behavior of the $NIG(\alpha, \beta, \delta)$ distribution is often referred to as 'semi-heavy'. Semi-heavy tails refers to the instance when the tails are heavier than those of the Gaussian distribution but lighter than those of the non-Gaussian stable laws. The $NIG(\alpha, \beta, \delta)$ distribution's semi-heavy tails are characterised by the following asymptotic relation:

$$f(x; \alpha, \beta, \delta) \sim C|x|^{-\frac{3}{2}} \exp(-\alpha|x| + \beta x) \quad \text{as } x \to \pm\infty, \tag{4.12}$$

for some $C \geq 0$. See Appendix 4.A.5 for proof of this result.

4.3 SIMULATING A NIG RANDOM VARIABLE

The Monte Carlo method discussed in Benth, et al. [7] was implemented in generating $NIG(\alpha, \beta, \delta)$ random variables.

Previously we saw that if $Y \overset{d}{\sim} IG(\delta, \sqrt{\alpha^2 - \beta^2})$ and $Z \overset{d}{\sim} N(0, 1)$, then

$$X = \beta Y + \sqrt{Y} Z \overset{d}{\sim} NIG(\alpha, \beta, \delta).$$

This relationship will be used in the simulation of NIG random variables.

The following result plays an important role:

Theorem 4.3.1 (see Schuster [52]) *Let $Z \overset{d}{\sim} IG(a, b)$. Then*

$$V = \frac{a^2(Z - a/b)^2}{(a/b)^2 Z} \overset{d}{\sim} \chi_1^2 \tag{4.13}$$

where χ_1^2 denotes a chi-squared random variable with 1 degree of freedom.

We follow the algorithm set out by Michael, Schucany and Haas [41] to generate random variables through transformations with multiple roots. χ_1^2 random variables are easily generated as squares of a standard normal random variable. Now, given a χ_1^2 observation v, we solve for z in (4.13) to obtain a inverse Gaussian observation. There are two roots associated with this quadratic equation. These roots can be expressed as

$$z_1 = \frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{\sqrt{4ab\upsilon + \upsilon^2}}{2b^2} \tag{4.14}$$

and

$$z_2 = \frac{a}{b} + \frac{\upsilon}{2b^2} + \frac{\sqrt{4ab\upsilon + \upsilon^2}}{2b^2}, \tag{4.15}$$

where $z_1 \geq 0$ and $z_2 \geq 0$. See Appendix 4.A.4 for a proof of this.

Define

$$z = z_1 \mathbb{I}(u \leq \frac{a}{a + z_1 b}) + z_2 \mathbb{I}(u > \frac{a}{a + z_1 b}) \tag{4.16}$$

where $u \overset{d}{\sim} uniform(0,1)$. Then z is an observation from an $IG(a,b)$ distribution (see Michael, et al. [41] for justification of the choice of z). Hence, we can generate a $NIG(\alpha, \beta, \delta)$ random variable with the following algorithm:

Generating a NIG random variable X

(i) Generate $\upsilon \overset{d}{\sim} \chi_1^2$.

(ii) Generate $u \overset{d}{\sim} uniform(0,1)$.

(iii) Calculate z in (4.16), with $a = \delta$ and $b = \sqrt{\alpha^2 - \beta^2}$.

(iv) Generate $y \overset{d}{\sim} N(0,1)$.

(v) Return $X = \beta z + \sqrt{z} y$.

Histograms of $NIG(\alpha, \beta, \delta)$ data sets of size 10^5, generated by the proceeding method, are plotted in Figures 4.4, 4.5 and 4.6. The parameter values are $(\alpha, \beta, \delta) = (8, 0, 1), (8, 6, 1)$ and $(8, -6, 1)$ respectively.

These histograms are directly comparable to the probability density functions plotted in Figure 4.2, where the effect of β is displayed. Figures 4.4, 4.5 and 4.6 have the NIG(8,0,1), NIG(8,6,1) and NIG(8,-6,1) density functions superimposed on the histograms respectively.

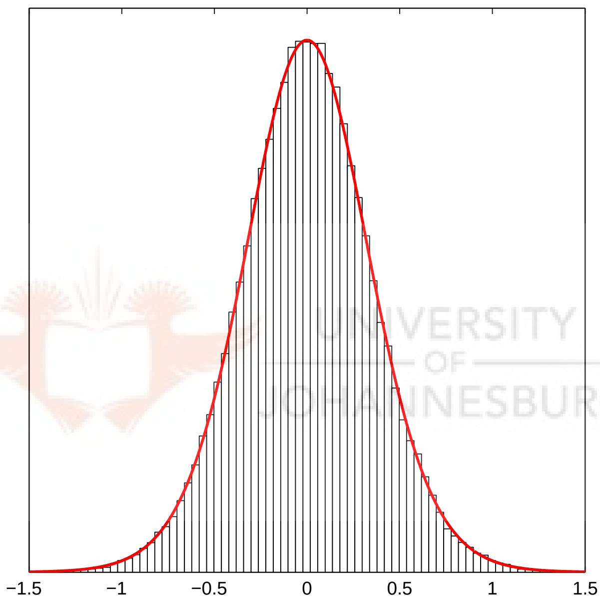

Figure 4.4: Histogram of a $NIG(8, 0, 1)$ generated data set with the theoretical density function superimposed.

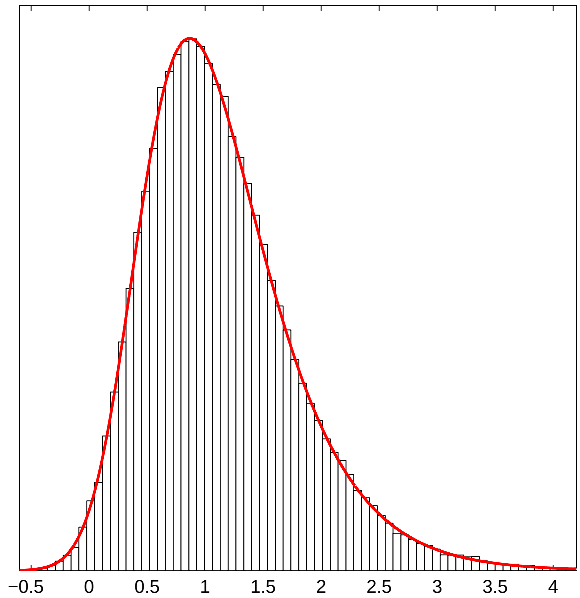

Figure 4.5: Histogram of a $NIG(8, 6, 1)$ generated data set with the theoretical density function superimposed.

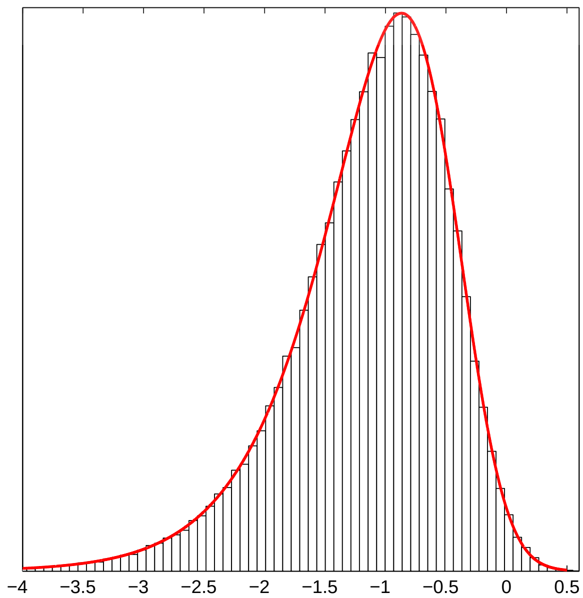

Figure 4.6: Histogram of a $NIG(8, -6, 1)$ generated data set with the theoretical density function superimposed.

4.A Appendix

4.A.1 Proof of Proposition 4.2.3

The following integral form of the modified Bessel function of the third kind (see Schoutens [50], p. 148) will play an important part:

$$K_{-1}(z) = K_1(z) = \frac{1}{2}\int_0^\infty u^{-2} \exp\left[-\frac{1}{2}z(u+\frac{1}{u})\right]du. \tag{4.17}$$

We will prove Proposition 4.2.3 using relationship (4.2). We have

$$\begin{aligned} P[X \leq x] &= P[\beta Y + \sqrt{Y}Z \leq 0] \\ &= \int_0^\infty \Phi(\frac{x-\beta y}{\sqrt{y}}) f_Y(y)dy \end{aligned}$$

and differentiating this with respect to x yields

$$f_X(x) = \int_0^\infty y^{-\frac{1}{2}}\Phi'(\frac{x-\beta y}{\sqrt{y}})f_Y(y)dy,$$

where $\Phi'(\cdot)$ denotes the normal density function.

Recalling Definition 4.2.1, we see that

$$f_Y(y) = \frac{\delta}{\sqrt{2\pi}}y^{-\frac{3}{2}}\exp\left[\delta\sqrt{\alpha^2-\beta^2} - \frac{1}{2}(\frac{\delta^2}{y} + (\alpha^2-\beta^2)y)\right]. \tag{4.18}$$

Hence,

$$\begin{aligned} & f_X(x) \\ = & \frac{\delta}{2\pi}\int_0^\infty y^{-2}\exp\left[-\frac{1}{2}(\frac{x-\beta y}{\sqrt{y}})\right]\exp\left[\delta\sqrt{\alpha^2-\beta^2} - \frac{1}{2}\left(\frac{\delta^2}{y} + (\alpha^2-\beta^2)y\right)\right]dy \\ = & \frac{\delta}{2\pi}\exp(\alpha\sqrt{\alpha^2-\beta^2})\int_0^\infty y^{-2}\exp\left[-\frac{1}{2}\left(\frac{x^2}{y} - \frac{2\beta xy}{y} + \frac{\beta^2y^2}{y} + \frac{\delta^2}{y} + \alpha^2 y - \beta^2 y\right)\right]dy \\ = & \frac{\delta}{2\pi}\exp(\alpha\sqrt{\alpha^2-\beta^2}+\beta x)\int_0^\infty y^{-2}\exp\left[-\frac{1}{2}((x^2+\delta^2)\frac{1}{y} + \alpha^2 y)\right]dy. \end{aligned} \tag{4.19}$$

Now, writing

$$(x^2+\delta^2)\frac{1}{y} + \alpha^2 y = \alpha\sqrt{x^2+\delta^2}\left(\frac{\sqrt{x^2+\delta^2}}{\alpha y}\right) + \alpha\sqrt{x^2+\delta^2}\left(\frac{\alpha y}{\sqrt{x^2+\delta^2}}\right) \tag{4.20}$$

and making the following transformation

$$u = \frac{\alpha y}{\sqrt{x^2+\delta^2}} \qquad du = \frac{\alpha}{\sqrt{x^2+\delta^2}}dy, \tag{4.21}$$

we have

$$f_X(x) = \frac{\delta}{2\pi}\exp(\alpha\sqrt{\alpha^2-\beta^2}+\beta x)\frac{\alpha}{\sqrt{x^2+\delta^2}}\int_0^\infty u^{-2}\exp\left[-\frac{1}{2}(\alpha\sqrt{x^2+\delta^2}(u+\frac{1}{u})\right]dy.$$

The result is achieved by using relationship (4.17) with $z = \alpha\sqrt{x^2+\delta^2}$.

■

4.A.2 *NIG* Cumulant Generating function

From Proposition 4.2.2,

$$\begin{aligned}
\mathbb{E}[e^{uX}] &= \int_{\mathbb{R}} e^{ux}\frac{\alpha\delta}{\pi}\exp(\delta\sqrt{\alpha^2-\beta^2}+\beta x)\frac{K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{\delta^2+x^2}}dx\\
&= \int_{\mathbb{R}} \frac{\alpha\delta}{\pi}\exp(\delta\sqrt{\alpha^2-\beta^2}+(\beta+u)x)\frac{K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{\delta^2+x^2}}dx\\
&= \exp(\delta\sqrt{\alpha^2-\beta^2})\int_{\mathbb{R}} \frac{\alpha\delta}{\pi}\exp((\beta+u)x)\frac{K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{\delta^2+x^2}}dx\\
&= \exp(\delta\sqrt{\alpha^2-\beta^2}-\delta\sqrt{\alpha^2-(\beta+u)^2})\cdot\\
&\quad \int_{\mathbb{R}} \frac{\alpha\delta}{\pi}\exp(\delta\sqrt{\alpha^2-(\beta+u)^2}+(\beta+u)x)\frac{K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{\delta^2+x^2}}dx\\
&= \exp(\delta\sqrt{\alpha^2-\beta^2}-\delta\sqrt{\alpha^2-(\beta+u)^2})\int_{\mathbb{R}} f(x;\alpha,\beta+u,\delta)dx\\
&= \exp\left(\delta(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})\right).
\end{aligned}$$

Hence,

$$\kappa(u) = \log\mathbb{E}[e^{uX}] = \delta(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2}).$$

■

4.A.3 *NIG* CHARACTERISTIC FUNCTION

Using the relationship, $\phi(u) = \exp\{\kappa(iu)\} = \mathbb{E}[e^{iuX}]$, we have

$$\phi(u) = \mathbb{E}[e^{iuX}] = \exp\left(\delta(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+iu)^2})\right).$$

■

4.A.4 ROOTS z_1 AND z_2

We are only required to show that $z_1 \geq 0$, since $z_2 \geq 0$ follows from the restrictions on a and b and the fact that a χ_1^2 random variable is always positive.

$$\begin{aligned}
\frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{\sqrt{4ab\upsilon+\upsilon^2}}{2b^2} &= \frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{\sqrt{4a^2b^2+4ab\upsilon+\upsilon^2-4a^2b^2}}{2b^2} \\
&= \frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{\sqrt{(2ab+\upsilon)^2-4a^2b^2}}{2b^2} \\
&\geq \frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{\sqrt{(2ab+\upsilon)^2}}{2b^2} \\
&= \frac{a}{b} + \frac{\upsilon}{2b^2} - \frac{(2ab+\upsilon)}{2b^2} \\
&= 0.
\end{aligned}$$

■

4.A.5 PROOF OF RELATION (4.12)

The following result is important in deriving the asymptotic relation (4.12) (see Abramowitz and Stegun [1], p. 378).

Let $\lambda = 1$. Then

$$K_1(x) \sim \sqrt{\frac{\pi}{2|x|}}\exp(-|x|) \tag{4.22}$$

as $|x| \to \infty$.

Using Proposition 4.2.3 and equation (4.22), it follows that

$$\begin{aligned} f(x) &= \frac{\alpha\delta}{\pi}\exp(\delta\sqrt{\alpha^2-\beta^2}+\beta x)\frac{K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{\delta^2+x^2}} \\ &\sim \frac{\alpha\delta}{\pi}\frac{\exp(\delta\sqrt{\alpha^2-\beta^2}+\beta x)}{\sqrt{\delta^2+x^2}}\sqrt{\frac{\pi}{2\alpha\sqrt{\delta^2+x^2}}}\exp(-\alpha\sqrt{\delta^2+x^2}). \qquad (4.23) \end{aligned}$$

Now

$$\begin{aligned} \exp(-\alpha\sqrt{\delta^2+x^2}) &= \exp\left(-\alpha|x|\sqrt{\frac{\delta^2}{x^2}+1}\right) \\ &= \exp\left(-\alpha|x|\left[1+\frac{\delta^2}{2x^2}+O(\frac{1}{x^2})\right]\right) \\ &= \exp\left(-\alpha|x|-\frac{\alpha\delta^2}{2|x|}+O(\frac{1}{|x|})\right]\right) \\ &\sim \exp(-\alpha|x|) \qquad \text{as } |x|\to\infty, \qquad (4.24) \end{aligned}$$

$$\begin{aligned} \exp(\delta\sqrt{\alpha^2-\beta^2}+\beta x) &= \exp\left(x\left[\frac{\delta\sqrt{\alpha^2-\beta^2}}{x}+\beta\right]\right) \\ &\sim \exp(\beta x) \qquad \text{as } |x|\to\infty, \qquad (4.25) \end{aligned}$$

and

$$\begin{aligned} \sqrt{\delta^2+x^2} &= |x|\sqrt{\frac{\delta^2}{x^2}+1} \\ &\sim |x| \qquad \text{as } |x|\to\infty. \qquad (4.26) \end{aligned}$$

Substituting (4.24), (4.25) and (4.26) into (4.23) yields

$$\begin{aligned} f(x) &\sim \frac{\alpha\delta}{\pi}\frac{\exp(\beta x)}{|x|}\sqrt{\frac{\pi}{2\alpha|x|}}\exp(-\alpha|x|) \\ &= C\exp(-\alpha|x|+\beta x)|x|^{-\frac{3}{2}} \end{aligned}$$

where $C=\sqrt{\frac{\alpha}{2\pi}}\delta$.

■

Chapter 5

The Meixner Distribution

5.1 Introduction

In this chapter we present the Meixner distribution and some of the properties this distribution contains. The Meixner distribution and the Meixner process were introduced by Schoutens and Teugels [51] (see also Schoutens [47]). The Meixner distribution is a three parameter distribution. The fitting of stock returns using this distribution was considered by Grigelionis [27] and Schoutens [48],[49]. We conclude this chapter with the proposal of a new method for generating Meixner random numbers.

5.2 Definition and properties

Definition 5.2.1 (Meixner density function, Schoutens [50], pp. 62) *A random variable X is Meixner distributed with parameters $\boldsymbol{\theta} = (\alpha, \beta, \delta)$, if it has the probability density function*

$$f(x;\boldsymbol{\theta}) = \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{\beta x}{\alpha}\right) \left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2 \tag{5.1}$$

where

$$\begin{aligned} \left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right| &= \left(\int_0^\infty \cos\left(\frac{x}{\alpha}\log y\right) y^{\delta-1} e^{-y}\right)^2 \\ &\quad + \left(\int_0^\infty \sin\left(\frac{x}{\alpha}\log y\right) y^{\delta-1} e^{-y}\right)^2 \end{aligned} \tag{5.2}$$

and

$$x \in \mathbb{R}, \quad \alpha \in (0,\infty), \quad \beta \in (-\pi,\pi), \quad \delta \in (0,\infty).$$

This distribution is denoted as $X \overset{d}{\sim} Meixner(\alpha,\beta,\delta)$.

Lemma 5.2.2 *Let* $X \overset{d}{\sim} Meixner(\alpha,\beta,\delta)$. *Then the expected value, variance, skewness and kurtosis of* X *are respectively given by*

$$\mathbb{E}[X] = \alpha\delta\tan(\beta/2) \tag{5.3}$$

$$\mathbb{VAR}[X] = \frac{\alpha^2\delta}{2}\sec^2(\beta/2) \tag{5.4}$$

$$\mathbb{SKEW}[X] = \sqrt{\frac{2}{\delta}}\sin(\beta/2) \tag{5.5}$$

$$\mathbb{KURT}[X] = 3 + \frac{2-\cos(\beta/2)}{\delta}. \tag{5.6}$$

See Schoutens [50], p. 63 for the above results.

Moments of all orders exist for the Meixner distribution. Looking at the kurtosis of the Meixner distribution, it can clearly be seen that it is always greater than the kurtosis of the Normal distribution, which always equals 3.

Figures 5.1, 5.2 and 5.3 show the effect of the parameters on the Meixner distribution. Like the normal inverse Gaussian distribution, the kurtosis of the density function is described by all the parameters of the distribution. However, the skewness (symmetry) of the density function is described by the parameter β (see Figure 5.2). The distribution is skewed to the left for $\beta < 0$, skewed to the right for $\beta > 0$ and symmetric for $\beta = 0$.

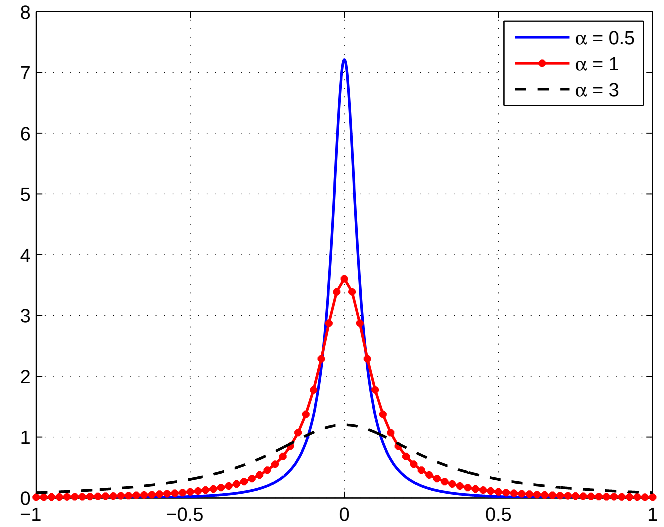

Figure 5.1: The effect of α in the Meixner density.
$\alpha \in \{0.5, 1, 3\}, \beta = 0, \delta = 0.1.$

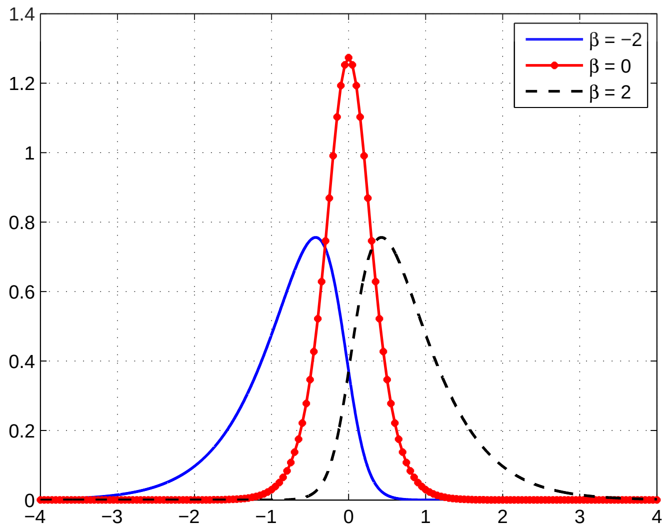

Figure 5.2: The effect of β in the Meixner density.
$\alpha = 0.5, \beta \in \{-2, 0, 2\}, \delta = 1.$

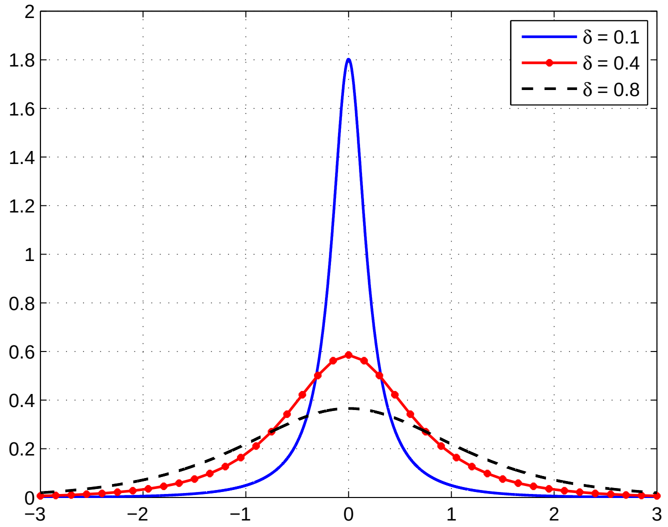

Figure 5.3: The effect of δ in the Meixner density.
$\alpha = 2, \beta = 0, \delta \in \{0.1, 0.4, 0.8\}$.

Lemma 5.2.3 *Let* $X \overset{d}{\sim} Meixner(\alpha, \beta, \delta)$. *Then the characteristic function,* $\phi(u)$, *and the cumulant generating function,* $\kappa(u) = \log \mathbb{E}[e^{uX}] = \log \phi(-iu)$, *are respectively given by*

$$\phi(u; \boldsymbol{\theta}) = \left(\frac{\cos(\beta/2)}{\cosh((\alpha u - i\beta)/2)} \right)^{2\delta} \tag{5.7}$$

and

$$\kappa(u; \boldsymbol{\theta}) = 2\delta \Big[\log(\cos(\beta/2)) - \log(\cos((\alpha u + \beta)/2)) \Big] \tag{5.8}$$

where $i = \sqrt{-1}$ *and* $u \in \mathbb{R}$.

Proof: See Appendix 5.A.1 and 5.A.2.

Looking at the form of the characteristic function of a Meixner random variable (5.7), we see that the $Meixner(\alpha, \beta, \delta)$ distribution is infinitely divisible, since

$$\phi(u; \alpha, \beta, \delta) = [\phi(u; \alpha, \beta, \delta/n)]^n. \tag{5.9}$$

The $Meixner(\alpha, \beta, \delta)$ distribution, like the NIG distribution, also has semi-heavy tails. The tail behavior is characterised by the following asymptotic relation:

$$f(x; \alpha, \beta, \delta) \sim \begin{cases} C|x|^{\rho} \exp(-\eta_{-}|x|) & \text{as } x \to \infty \\ C|x|^{\rho} \exp(-\eta_{+}|x|) & \text{as } x \to -\infty, \end{cases} \tag{5.10}$$

where

$$\rho = 2\delta - 1, \quad \eta_{-} = \frac{\pi - \beta}{\alpha}, \quad \eta_{+} = \frac{\pi + \beta}{\alpha}$$

and for some $C \geq 0$. See Appendix 5.A.3 for the proof of this asymptotic relation.

5.3 Simulating a Meixner random variable

There seems to be no published algorithm to generate random numbers from a Meixner distribution. One can attempt to generate Meixner random numbers using the inverse transform of the distribution function. However because there is no closed-form expression for the latter, the distribution function must be approximated numerically with a Riemman sum (see Robbertse [42], p. 47). We propose a different method, one using the rejection method (also referred to as the acceptance-rejection method) described by Ross [43], p. 66.

Suppose a method exists for generating a random variable from a continuous distribution with density function $g(x)$. We can then use this distribution as a basis for generating a random value from a continuous distribution with density function $f(x)$ by generating Y from g and accepting this value y with a probability proportional to $\frac{f(y)}{g(y)}$. Let c be a constant such that

$$\frac{f(x)}{g(x)} \leq c \qquad \forall\, x. \tag{5.11}$$

Typically, we choose

$$c = \max_{x} \frac{f(x)}{g(x)} \tag{5.12}$$

and accept y with probability $\frac{f(y)}{cg(y)}$. Note that this is accomplished by generating a $uniform(0,1)$ random number u and then accepting y if $u \leq \frac{f(y)}{cg(y)}$.

For the Meixner distribution with parameters $\boldsymbol{\theta} = (\alpha, \beta, \delta)$, we will use for $g(x)$ the NIG density, with parameters $\boldsymbol{\theta^*} = (\alpha^*, \beta^*, \delta^*)$. The parameters $(\alpha^*, \beta^*, \delta^*)$ are chosen such that the first three moments of the NIG distribution are equal to those of the Meixner distribution. This entails the solution of three non-linear equations. In some instances $\boldsymbol{\theta^*}$ contains complex values. After some simulation we noticed that this occurred mostly when $\alpha < |\beta| < \pi$. Looking at Figures 4.1, 4.2 and 4.3 we notice that the NIG distribution is peaked for large values of α and small values of δ. The Meixner distribution is peaked for small values of α and small values δ (Figures 5.1, 5.2 and 5.3). For the instances when the moments can not be matched, we set $\boldsymbol{\theta}^*$ equal to:

$$\alpha^* = \frac{1}{\alpha} \tag{5.13}$$

$$\beta^* = \alpha^* \frac{\beta}{\pi} \tag{5.14}$$

$$\delta^* = \delta. \tag{5.15}$$

Hence we can generate a $Meixner(\alpha, \beta, \delta)$ random variable with the following algorithm :

Generating a Meixner random variable X

(i) Given α, β, δ, calculate $\alpha^*, \beta^*, \delta^*$ in the manner indicated above.

(ii) Calculate c using (5.12) with f a $Meixner(\alpha, \beta, \delta)$ density and g a $NIG(\alpha, \beta, \delta)$ density.

(iii) Generate $u \stackrel{d}{\sim} uniform(0,1)$.

(iv) Generate $y \stackrel{d}{\sim} NIG(\alpha^*\beta^*, \delta^*)$.

(v) If $u \leq \frac{f(y)}{cg(y)}$, set $X = y$. Otherwise return to step (iii).

Histograms of $Meixner(\alpha, \beta, \delta)$ generated data sets, of size 10^5, are plotted in Figures 5.4, 5.5 and 5.6. The parameter values are given by $(0.5, 0, 1), (0.5, 2, 1)$ and $(0.5, -2, 1)$ respectively. These histograms are directly comparable to the probability density functions plotted in Figure 5.2, where the effect of β is displayed. Figures 5.4, 5.5 and 5.6 have the $Meixner(0.5, 0, 1)$, $Meixner(0.5, 2, 1)$ and $Meixner(0.5, -2, 1)$ density functions superimposed on the histograms respectively.

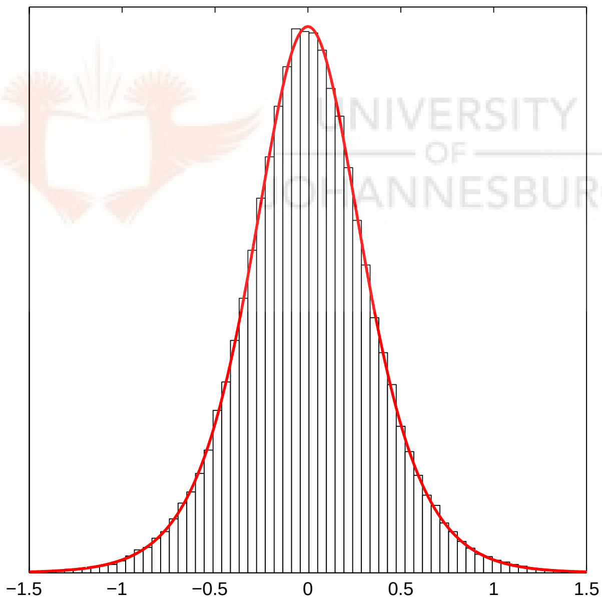

Figure 5.4: Histogram of a $Meixner(0.5, 0, 1)$ generated data set with the theoretical density function superimposed.

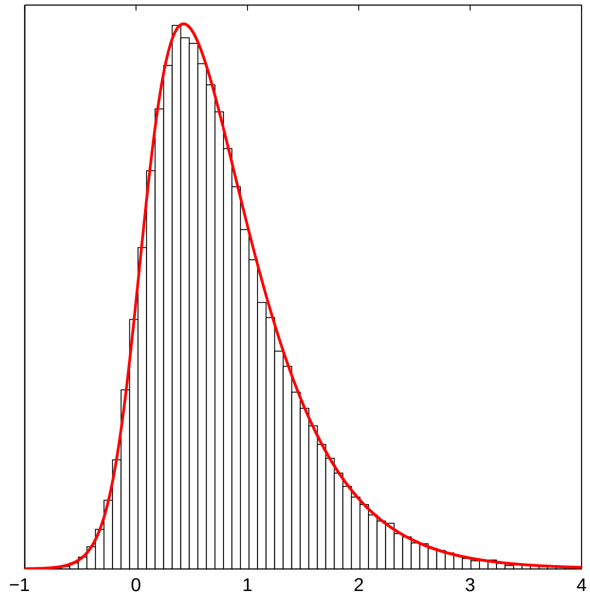

Figure 5.5: Histogram of a $Meixner(0.5, 2, 1)$ generated data set with the theoretical density function superimposed.

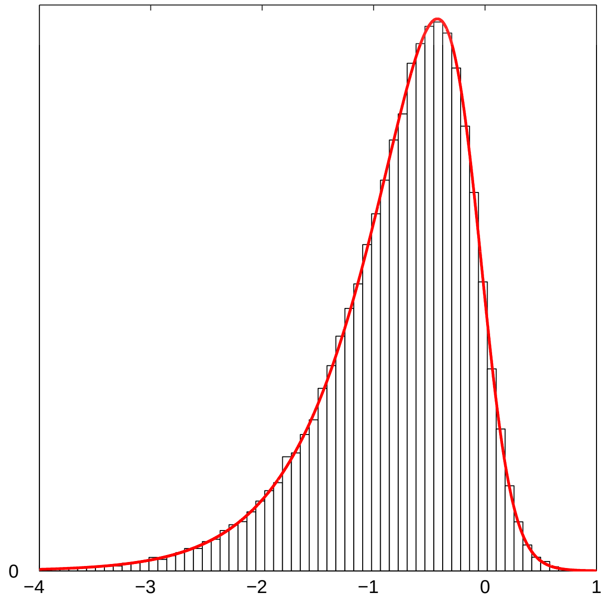

Figure 5.6: Histogram of a $Meixner(0.5, -2, 1)$ generated data set with the theoretical density function superimposed.

5.3.1 Simulation Study

We compared simulating Meixner random numbers using the rejection method (see Section 5.3) and the method which incorporates the inverse transform of the distribution function. The $Meixner(0.5, 0, 1)$ distribution was used in the study. We simulated[1] a 1000 $N \times M$ matrices of Meixner random numbers and calculated the average time(seconds) taken for the simulation. The standard deviations are included in brackets. The results are provided in the Table 5.1 below.

(N,M)	Method			
	Inverse Transform		Rejection	
(1,10)	0.640	(0.0452)	0.026	(0.0023)
(1,1000)	0.755	(0.0689)	0.056	(0.0038)
(10,10)	2.149	(0.0248)	0.033	(0.0042)
(100,100)	20.134	(1.1364)	0.230	(0.0071)
(1000,1000)	181.273	(2.2526)	8.722	(0.1898)

Table 5.1: Simulation times of a $N \times M$ matrix of Meixner random variates when moment matching is possible.

For the $Meixner(1, 2, 1)$ distribution one is unable to match the first three moments to that of the NIG distribution. Hence, the parameter set $\boldsymbol{\theta}^*$ contains complex values. One thousand $N \times M$ matrices of $Meixner(1, 2, 1)$ random numbers were simulated using the rejection method. Table 5.2 gives the average time(seconds) taken to generate a (N, M) matrix of Meixner random numbers (standard deviations included in brackets).

Looking at Tables 5.1 and 5.2 we see that the rejection method for generating Meixner random numbers is significantly quicker than the inverse transform method. We see that, at its worst, the rejection method is more than two times slower when the moments of the NIG and Meixner distributions cannot be matched compared to the case when they can be matched. However, when the moments cannot be matched, the rejection method

[1] Simulations run in Matlab 2007b on a AMD Turion(tm) 64x2 Mobile Technology TL-56 1.79 Ghz processor with 2 Gb of RAM and Windows XP operating system.

is still significantly quicker than the inverse transform method.

(N,M)	Time	
(1,10)	0.045	(0.0035)
(1,1000)	0.077	(0.0054)
(10,10)	0.046	(0.0043)
(100,100)	0.231	(0.0134)
(1000,1000)	19.826	(0.3693)

Table 5.2: Simulation times of a $N \times M$ matrix of Meixner random variates when moment matching is not possible.

Figure 5.7 below plots a histogram of a generated data set, size 10^5, of $Meixner(1, 2, 1)$ random numbers. The theoretical $Meixner(1, 2, 1)$ density function is superimposed on the histogram.

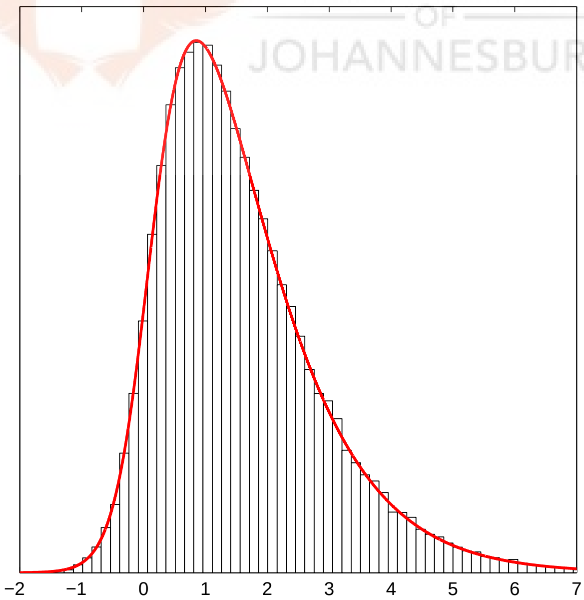

Figure 5.7: Histogram of a $Meixner(1, 2, 1)$ generated data set with the theoretical density function superimposed.

5.A Appendix

5.A.1 Meixner Cumulant Generating function

From equation (5.1),

$$\begin{aligned}
\mathbb{E}[e^{uX}] &= \int_{\mathbb{R}} e^{ux} \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{\beta x}{\alpha}\right)\left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= \int_{\mathbb{R}} \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{\beta x + \alpha u x}{\alpha}\right)\left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= (2\cos(\beta/2))^{2\delta} \int_{\mathbb{R}} \frac{1}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{(\beta + \alpha u)x}{\alpha}\right)\left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= \frac{(2\cos(\beta/2))^{2\delta}}{(2\cos((\beta+\alpha u)/2))^{2\delta}} \int_{\mathbb{R}} \frac{(2\cos((\beta+\alpha u)/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{(\beta + \alpha u)x}{\alpha}\right)\left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= \frac{(2\cos(\beta/2))^{2\delta}}{(2\cos((\beta+\alpha u)/2))^{2\delta}} \int_{\mathbb{R}} f(x;\alpha,\beta+\alpha u,\delta) dx \\
&= \left(\frac{\cos(\beta/2)}{\cos((\beta+\alpha u)/2)}\right)^{2\delta}
\end{aligned}$$

Hence,

$$\kappa(u) = \log \mathbb{E}[e^{uX}] = 2\delta\Big[\log(\cos(\beta/2)) - \log(\cos((\beta+\alpha u)/2))\Big].$$

■

5.A.2 Meixner Characteristic function

Using the relationship, $\phi(u) = \exp\{\kappa(iu)\} = \mathbb{E}[e^{iuX}]$ and the following relationship between complex trigonometric and hyperbolic functions

$$\cos(iz) = \cosh(z), \tag{5.16}$$

we have

$$\begin{aligned}
\phi(u) = \mathbb{E}[e^{iuX}] &= \left(\frac{\cos(\beta/2)}{\cos((\beta+i\alpha u)/2)}\right)^{2\delta} \\
&= \left(\frac{\cos(\beta/2)}{\cos(i(\alpha u - i\beta)/2)}\right)^{2\delta} \\
&= \left(\frac{\cos(\beta/2)}{\cosh((\alpha u - i\beta)/2)}\right)^{2\delta}.
\end{aligned}$$

■

5.A.3 Proof of Relation (5.10)

The following result is important in deriving relation (5.10) (see Copson [18], p. 224). Let x be finite. Then

$$|\Gamma(x+iy)|^2 \sim 2\pi|y|^{2x-1}\exp(-\pi|y|). \tag{5.17}$$

as $|y| \to \infty$.

Using Proposition 5.2.1 and equation (5.17), it follows that

$$\begin{aligned}
f(x) &= \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\exp\left(\frac{\beta}{\alpha}x\right)\left|\Gamma\left(\delta+\frac{ix}{\alpha}\right)\right|^2 \\
&\sim \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\exp\left(\frac{\beta}{\alpha}x\right)2\pi\left|\frac{x}{\alpha}\right|^{2\delta-1}\exp\left(-\pi\left|\frac{x}{\alpha}\right|\right) \\
&= C\exp\left(\frac{-\pi}{\alpha}|x|+\frac{\beta}{\alpha}x\right)|x|^{2\delta-1} \\
&= \begin{cases} C\exp\left(\frac{-\pi+\beta}{\alpha}|x|\right)|x|^{2\delta-1} & \text{as } x\to+\infty \\ C\exp\left(\frac{-\pi-\beta}{\alpha}|x|\right)|x|^{2\delta-1} & \text{as } x\to-\infty \end{cases} \\
&= \begin{cases} C\exp(-\eta_-|x|)|x|^{\rho} & \text{as } x\to+\infty \\ C\exp(-\eta_+|x|)|x|^{\rho} & \text{as } x\to-\infty \end{cases}
\end{aligned}$$

where

$$\rho = 2\delta - 1, \qquad \eta_- = \frac{\pi - \beta}{\alpha}, \qquad \eta_+ = \frac{\pi + \beta}{\alpha}$$

and $C \geq 0$.

■

Chapter 6

The Lévy-GARCH Model

6.1 Introduction

Duan [21] attempted to relax one of the two main assumptions in the Black-Scholes model, namely the assumption of constant volatility. As we have seen (Chapter 2), Duan proposed a GARCH(1,1) process for the variance (squared volatility) process. However, Duan still assumed the log returns of the asset to be normally distributed. Empirical studies show that the log returns are skewed and heavy-tailed. (see Anderson, et al. [3], Bollen and Inder [9], Carr, et al. [13] and Cont [16])

Thus, in this chapter we attempt to relax the assumption of normality and replace it with a more flexible Lévy process distribution, namely the NIG distribution or the Meixner distribution. We show that the properties of Duan's GARCH model still hold when more general distributions are utilized for the random innovations.

Duan's GARCH model will be referred to as the Gaussian-GARCH model in this chapter and the succeeding chapters.

6.2 Formulation of the model

For risk-neutrality we require a measure $\mathbb{Q}$ which is equivalent to the real world measure $\mathbb{P}$, such that the discounted stock price process $e^{-rt}S_t$ is a martingale, i.e.

$$\mathbb{E}^{\mathbb{Q}}[e^{-rt}S_t|\mathcal{F}_k] = e^{-rk}S_k. \tag{6.1}$$

The log returns model will be written in the following form

$$\log\frac{S_t}{S_{t-\Delta t}} = \mu_t\Delta t - \kappa(\sqrt{h_t};\boldsymbol{\theta}) + \sqrt{h_t}\epsilon_t, \tag{6.2}$$

$$h_{t+\Delta t} = \alpha_0 + \alpha_1 h_t\epsilon_t^2 + \beta_1 h_t \tag{6.3}$$

where $\boldsymbol{\theta}$ is the parameter set of random variable ϵ_t and $\kappa(\sqrt{h_t};\boldsymbol{\theta})$ is the cumulant generating function of ϵ_t. $\epsilon_t = \frac{Z_t-\mu_Z}{\sigma_Z}$ has a standardized distribution, i.e. $\epsilon \overset{d}{\sim} (0,1)$, and Z_t is an infinitely divisible random variable with mean μ_Z and variance σ_Z^2.

Note that the drift μ_t can be rewritten as

$$\begin{aligned}\mu_t &= r + \mu_t - r \\ &= r + \frac{\mu_t - r}{\sqrt{h_t}}\sqrt{h_t}\frac{\Delta t}{\Delta t} \\ &= r + \frac{\lambda}{\Delta t}\sqrt{h_t}\end{aligned} \tag{6.4}$$

where $\lambda = \frac{\mu_t - r}{\sqrt{h_t}}\Delta t$ is assumed to be constant. λ is defined as the *risk premium* per unit time. The risk premium is the expected return above the risk free rate per unit of volatility.

Hence, our model can be specified as

$$\log\frac{S_t}{S_{t-\Delta t}} = r\Delta t + \lambda\sqrt{h_t} - \kappa(\sqrt{h_t};\boldsymbol{\theta}) + \sqrt{h_t}\epsilon_t. \tag{6.5}$$

Taking the expectation of S_t conditional on the history up until time $t-\Delta t$, yields

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}[e^{-rt}S_t|\mathcal{F}_{t-\Delta t}] &= \mathbb{E}^{\mathbb{P}}[e^{-rt}S_{t-\Delta t}e^{r\Delta t+\lambda\sqrt{h_t}-\kappa(\sqrt{h_t};\boldsymbol{\theta})+\sqrt{h_t}\epsilon_t}|\mathcal{F}_{t-\Delta t}] \\ &= S_{t-\Delta t}e^{-r(t-\Delta t)+\lambda\sqrt{h_t}-\kappa(\sqrt{h_t};\boldsymbol{\theta})}\mathbb{E}^{\mathbb{P}}[e^{\sqrt{h_t}\epsilon_t}|\mathcal{F}_{t-\Delta t}] \\ &= S_{t-\Delta t}e^{-r(t-\Delta t)+\lambda\sqrt{h_t}-\kappa(\sqrt{h_t};\boldsymbol{\theta})}e^{\kappa(\sqrt{h_t};\boldsymbol{\theta})} \\ &= S_{t-\Delta t}e^{-r(t-\Delta t)+\lambda\sqrt{h_t}}.\end{aligned} \tag{6.6}$$

We see that the discounted stock price process $e^{-rt}S_t$ is not a martingale.

We now apply the following transformation in equation (6.5),

$$\xi_t = \epsilon_t + \lambda \tag{6.7}$$

and let $\mathbb{Q}$ be a measure such that $\xi_t = \epsilon_t + \lambda \overset{d}{\sim} (0,1)$ and such that measure $\mathbb{Q}$ is equivalent to measure $\mathbb{P}$. In the case where measure $\mathbb{P}$ generates a NIG or Meixner distribution, we will show that such a measure $\mathbb{Q}$ exists and further more generates respectively a NIG or Meixner distribution (see Appendix 6.A.1). Note that $\xi_t \overset{d}{\sim} (\lambda, 1)$ under measure $\mathbb{P}$.

Now substituting (6.7) into (6.2) and (6.3) yields,

$$S_t = S_{t-\Delta t} \exp\left(r\Delta t - \kappa(\sqrt{h_t};\boldsymbol{\theta}) + \sqrt{h_t}\xi_t \right) \tag{6.8}$$

$$h_{t+\Delta t} = \alpha_0 + \alpha_1 h_t (\xi_t - \lambda)^2 + \beta_1 h_t \tag{6.9}$$

Then under $\mathbb{Q}$, $\lambda = 0$ in equation (6.5), and it follows from (6.6) after replacing $\mathbb{P}$ by $\mathbb{Q}$ and λ by 0 that $e^{-rt}S_t$ is a martingale as required.

This model looks similar to that of the Gaussian-GARCH model. However, there is one more important distinction to point out that can not be seen explicitly above. The inclusion of the term $\kappa(\sqrt{h_t};\boldsymbol{\theta})$ creates a restriction on the parameters $\boldsymbol{\theta}$ and/or h_t, since $\kappa(\sqrt{h_t};\boldsymbol{\theta})$ must be real. $\kappa(\sqrt{h_t};\boldsymbol{\theta})$ contains a function $u(\sqrt{h_t};\boldsymbol{\theta})$ which is complex for certain combinations of values of $\boldsymbol{\theta}$ and/or h_t. Hence, we incorporate additional restrictions on $\boldsymbol{\theta}$ and/or h_t. These restrictions are dependent on the particular distribution chosen for the innovation ϵ_t.

For instance, let $Z_t \overset{d}{\sim} NIG(\alpha, \beta, \delta)$. Then using equation (4.10) we have

$$\begin{aligned}\kappa(\sqrt{h_t};\boldsymbol{\theta}) &= -\frac{\mu_Z}{\sigma_Z}\sqrt{h_t} + \kappa_{Z_t}(\frac{\sqrt{h_t}}{\sigma_Z};\boldsymbol{\theta}) \\ &= -\frac{\mu_Z}{\sigma_Z}\sqrt{h_t} + \delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2 - (\beta + \sqrt{h_t}/\sigma_Z)^2}\right) \end{aligned} \tag{6.10}$$

where $\boldsymbol{\theta} = (\alpha, \beta, \delta)$ and μ_Z and σ_Z are given by (4.4) and (4.5) respectively. It follows that $\epsilon_t = \frac{Z_t - \mu_Z}{\sigma_Z}$ follows a standardized NIG distribution, i.e. $\epsilon_t \overset{d}{\sim} std\ NIG(\alpha, \beta, \delta)$. For (6.10) to be real we require that,

$$
\begin{array}{lccc}
 & |\beta| \leq \alpha & \text{and} & |\beta + \sqrt{h_t}/\sigma_Z| \leq \alpha \\
\Leftrightarrow & -\alpha \leq \beta \leq \alpha & & -\alpha - \beta \leq \sqrt{h_t}/\sigma_Z \leq \alpha - \beta \\
\Leftrightarrow & -(\alpha + \beta) \leq 0 \leq \alpha - \beta & & -\sigma_Z(\alpha + \beta) \leq \sqrt{h_t} \leq \sigma_Z(\alpha - \beta)
\end{array}
$$

The first restriction and the fact that $\sigma_Z > 0$ implies that, the lower bound and the upper bound of the second restriction are always negative and positive respectively. Now since $h_t \geq 0$, we get the following restriction for h_t

$$0 \leq h_t \leq \sigma_Z^2(\alpha - \beta)^2.$$

We see that for the NIG distribution we are required to cap the value of h_t. We will denote this cap value by $g(\boldsymbol{\theta})$. Hence, for the NIG distribution $g(\boldsymbol{\theta})$ is given by

$$g(\boldsymbol{\theta}) = \sigma_Z^2(\alpha - \beta)^2. \tag{6.11}$$

For the Meixner distribution, the incorporation of the additional restrictions on $\boldsymbol{\theta}$ and/or h_t are of the same form as the NIG distribution, i.e. we are also required to cap h_t when $Z_t \overset{d}{\sim} Meixner(\alpha, \beta, \delta)$ (see Section 6.2.2). Our two Lévy-GARCH models, where Z_t is a NIG or Meixner distribution, can be generally specified as:

$$
\begin{aligned}
S_t &= S_{t-\Delta t} \exp\left(r\Delta t + \lambda\sqrt{h_t} - \kappa(\sqrt{h_t}; \boldsymbol{\theta}) + \sqrt{h_t}\epsilon_t\right) && (6.12) \\
h_{t+\Delta t} &= (\alpha_0 + \alpha_1 h_t \epsilon_t^2 + \beta_1 h_t) \wedge g(\boldsymbol{\theta}). && (6.13)
\end{aligned}
$$

The risk-neutral Lévy-GARCH models are given by

$$
\begin{aligned}
S_t &= S_{t-\Delta t} \exp\left(r\Delta t - \kappa(\sqrt{h_t}; \boldsymbol{\theta}) + \sqrt{h_t}\xi_t\right) && (6.14) \\
h_{t+\Delta t} &= (\alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t) \wedge g(\boldsymbol{\theta}). && (6.15)
\end{aligned}
$$

The effect of the restriction $g(\boldsymbol{\theta})$ will be discussed in the next chapter with the pricing results.

Remark. This model has the following properties:

- The process $h_{t+\Delta t}$ is $\mathcal{F}_t$ measurable and thus predictable.
- The expected price increment over one period conditional on the history, $\mathbb{E}[\frac{S_{t+\Delta t}}{S_t}|\mathcal{F}_t]$, equals $\exp(r\Delta t+\lambda\sqrt{h_t})$ and thus agrees with the interpretation that λ is the market price of risk.
- When ϵ_t follows a standard normal distribution for all $t \in \mathbb{N}$, the model is that introduced by Duan [21] (see Chapter 2).
- If, in addition, the coefficients α_1 and β_1 are zero then the model reduces to the Black-Scholes discrete time model (see Black and Scholes [8] and Merton [39]). Hence, this ensures the homoskedastic lognormal Black-Scholes model as a special case of the Lévy GARCH models.

Proposition 6.2.1 *Under the risk-neutral measure $\mathbb{Q}$, if $|\lambda| < \sqrt{(1-\alpha_1-\beta_1)/\alpha_1}$, then*

(i) *The stationary variance of $\sqrt{h_t}\xi_t$ equals $\frac{\alpha_0}{1-(1+\lambda^2)\alpha_1-\beta_1}$.*

(ii) *$\sqrt{h_t}\xi_t$ is leptokurtic.*

(iii) $\mathbb{COV}^{\mathbb{Q}}(h_{t+\Delta t}, \xi_t) = \frac{(s-2\lambda)\alpha_0\alpha_1}{1-(1+\lambda^2)\alpha_1-\beta_1}$

where s denotes the skewness coefficient of the distribution ξ_t.

Proof: See Appendix 6.A.2.

We see that the stationary variance of the Lévy-GARCH return process is equal to that of the normal GARCH return process. Also, the conditional variance, under risk-neutralization, is correlated with the lagged asset return for $s \neq 2\lambda$. However, unlike the normal GARCH return process, the correlation is positive for $s > 2\lambda$ and negative for $s < 2\lambda$.

As in the Gaussian-GARCH model, the pricing of European options requires aggregating single period asset returns to obtain a random terminal asset price at some future date. From (6.8) we see that the terminal asset price, S_T, can be expressed in terms of the initial asset price, S_0, by

$$S_T = S_0 \exp\left[rT - \sum_i \kappa(\sqrt{h_{i\Delta t}}; \boldsymbol{\theta}) + \sum_i \sqrt{h_{i\Delta t}}\xi_{i\Delta t}\right] \tag{6.16}$$

where $i = 1, 2, ..., n$ and $n\Delta t = T$.

6.2.1 *NIG*-GARCH Model

Let $Z_t \overset{d}{\sim} NIG(\alpha, \beta, \delta)$. Then, using equations (6.10) and (6.11), the stock price process is given by

$$\begin{aligned} S_t &= S_{t-\Delta t} \exp\left(r\Delta t + \frac{\mu_Z}{\sigma_Z}\sqrt{h_t} - \delta\left[\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \sqrt{h_t}/\sigma_Z)^2}\,\right]\right) \\ &\quad \cdot \exp\left(\sqrt{h_t}(\epsilon_t + \lambda)\right) \end{aligned}$$

and the variance process follows a restricted-GARCH(1,1) process,

$$h_{t+\Delta t} = (\alpha_0 + \alpha_1 h_t \epsilon_t^2 + \beta_1 h_t) \wedge \sigma_Z^2(\alpha - \beta)^2.$$

6.2.1.1 The *NIG*-GARCH Risk-Neutral Model

The risk neutral stock price process and volatility process are given by

$$\begin{aligned} S_t &= S_{t-\Delta t} \exp\left(r\Delta t + \frac{\mu_Z}{\sigma_Z}\sqrt{h_t} - \delta\left[\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \sqrt{h_t}/\sigma_Z)^2}\,\right]\right) \\ &\quad \cdot \exp\left(\sqrt{h_t}\xi_t\right), \qquad (6.17) \\ h_{t+\Delta t} &= (\alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t) \wedge \sigma_Z^2(\alpha - \beta)^2, \qquad (6.18) \end{aligned}$$

where $\xi_t \stackrel{d}{\sim} std\ NIG(\alpha, \beta, \delta)$. We will call this the NIG-GARCH (risk neutral) model.

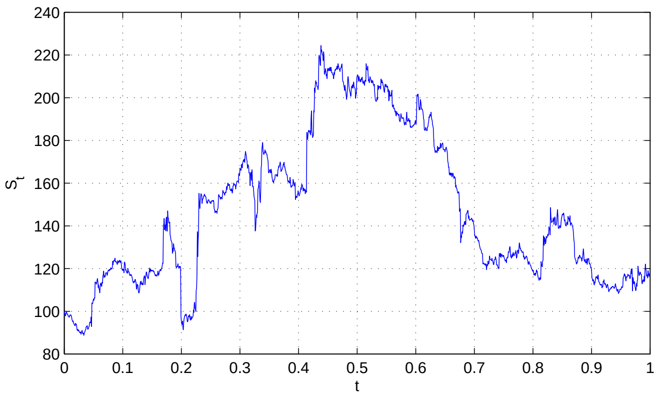

Figure 6.1: Simulated NIG-GARCH stock path.

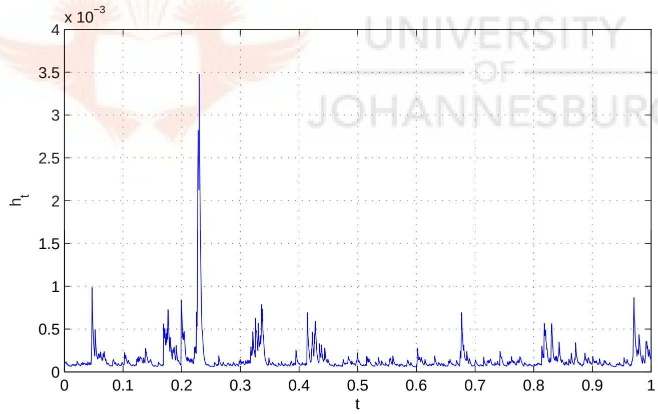

Figure 6.2: Simulated NIG-GARCH variance path.

Figures 6.1 and 6.2 plot a single realization of the NIG-GARCH model. The parameter values are given by $(\alpha_0, \alpha_1, \beta_1, \lambda) = (1.524\text{x}10^{-5}, 0.188, 0.716, 0.007)$, $(\alpha, \beta, \delta) = (1.8, 0.189, 1.62)$, $S_0 = 100$ and $r = 0$. For this parameter set $g(\boldsymbol{\theta}) = 2.3769$. Hence, looking at Figure 6.2, we see that the restriction $g(\boldsymbol{\theta})$ plays no role in this single realization. The impact of the restriction $g(\boldsymbol{\theta})$ in the empirical analysis will be looked at in Chapter 7.

6.2.2 Meixner-GARCH Model

Let $Z_t \overset{d}{\sim} Meixner(\alpha, \beta, \delta)$. Then, using equation (5.8) we have

$$\kappa(\sqrt{h_t};\boldsymbol{\theta}) = -\frac{\mu_Z}{\sigma_Z}\sqrt{h_t} + 2\delta\Big[\log(\cos(\beta/2)) - \log(\cos((\alpha\sqrt{h_t}/\sigma_Z + \beta)/2))\Big] \quad (6.19)$$

where $\boldsymbol{\theta} = (\alpha, \beta, \delta)$, μ_Z and σ_Z are given by (5.3) and (5.4) respectively. It follows that $\epsilon_t = \frac{Z_t - \mu_Z}{\sigma_Z}$ follows a standardized Meixner distribution, i.e. $\epsilon_t \overset{d}{\sim} stdMeixner(\alpha, \beta, \delta)$. For (6.19) to be real we require that

$$\begin{array}{cccc} & \left|\frac{\beta}{2}\right| \leq \frac{\pi}{2} & \text{and} & \left|\frac{\alpha\sqrt{h_t}/\sigma_Z+\beta}{2}\right| \leq \frac{\pi}{2} \\ \Leftrightarrow & -\pi \leq \beta \leq \pi & & -\pi - \beta \leq \alpha\sqrt{h_t}/\sigma_Z \leq \pi - \beta \\ \Leftrightarrow & -(\pi+\beta) \leq 0 \leq \pi - \beta & & -\sigma_Z(\pi+\beta)/\alpha \leq \sqrt{h_t} \leq \sigma_Z(\pi-\beta)/\alpha \end{array}$$

The first restriction and the fact that $\sigma_Z > 0$ and $\alpha > 0$ implies that, the lower bound and the upper bound of the second restriction are always negative and positive respectively. Now since $h_t \geq 0$, we get the following restriction for h_t

$$0 \leq h_t \leq \frac{\sigma_Z^2(\pi-\beta)^2}{\alpha^2}.$$

Hence, the restriction $g(\boldsymbol{\theta})$ is given by

$$g(\boldsymbol{\theta}) = \frac{(\pi-\beta)^2\sigma_Z^2}{\alpha^2}. \quad (6.20)$$

Then the stock price process is defined by

$$\begin{aligned} S_t &= S_{t-\Delta t}\exp\left(r\Delta t - 2\delta\Big[\log(\cos(\beta/2)) - \log(\cos((\alpha\sqrt{h_t}/\sigma_Z + \beta)/2))\Big]\right) \\ &\quad \cdot \exp\left(\frac{\mu_Z}{\sigma_Z}\sqrt{h_t} + \sqrt{h_t}(\epsilon_t + \lambda)\right) \end{aligned}$$

and the variance process follows a restricted-GARCH(1,1) process,

$$h_{t+\Delta t} = (\alpha_0 + \alpha_1 h_t \epsilon_t^2 + \beta_1 h_t) \wedge \frac{(\pi-\beta)^2\sigma_Z^2}{\alpha^2}.$$

6.2.2.1 The Meixner-GARCH Risk-Neutral Model

The risk neutral stock price process and volatility process are given by

$$S_t = S_{t-\Delta t} \exp\left(r\Delta t - 2\delta \Big[\log(\cos(\beta/2)) - \log(\cos((\alpha\sqrt{h_t}/\sigma_Z + \beta)/2)) \Big] \right) \cdot \exp\left(\frac{\mu_Z}{\sigma_Z}\sqrt{h_t} + \sqrt{h_t}\xi_t \right), \quad (6.21)$$

$$h_{t+\Delta t} = (\alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t) \wedge \frac{(\pi - \beta)^2 \sigma_Z^2}{\alpha^2}. \quad (6.22)$$

where $\xi_t \overset{d}{\sim} stdMeixner(\alpha, \beta, \delta)$. We will call this the Meixner-GARCH (risk neutral) model.

Figures 6.3 and 6.4 plot a single realization of the Meixner-GARCH model. The parameter values are given by $(\alpha_0, \alpha_1, \beta_1, \lambda) = (1.524\text{x}10^{-5}, 0.188, 0.716, 0.007)$, $(\alpha, \beta, \delta) = (1, 0.18, 1)$, $S_0 = 100$ and $r = 0$. For this parameter set $g(\boldsymbol{\theta}) = 0.71$. Hence, looking at Figure 6.4, we see that the restriction $g(\boldsymbol{\theta})$ plays no role in this single realization.

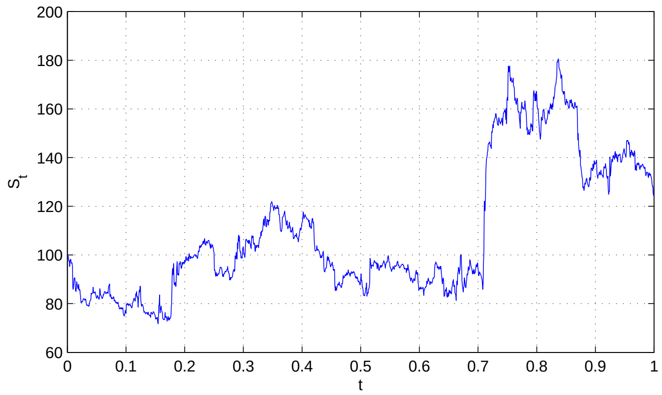

Figure 6.3: Simulated Meixner-GARCH stock path.

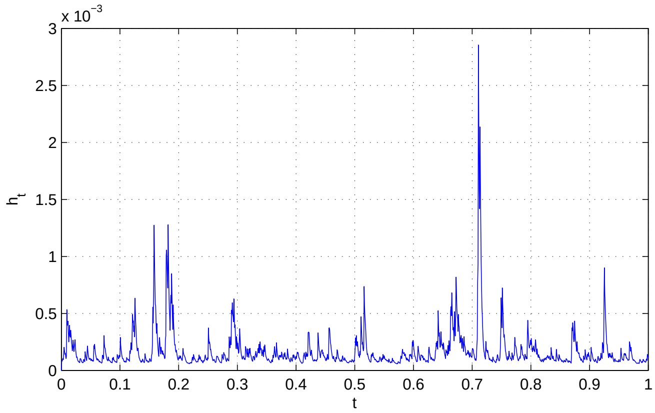

Figure 6.4: Simulated Meixner-GARCH variance path.

6.A Appendix

6.A.1 Construction of the Risk Neutral Measure $\mathbb{Q}$

We base our construction of the risk neutral measure $\mathbb{Q}$ on the following result, which gives necessary and sufficient conditions for the equivalence of the distributions for two infinitely divisible distributions.

Theorem 6.A.1 (Sato [46], Theorem 33.1) *Let $(X,\mathbb{P})$ and $(X,\mathbb{Q})$ be two infinitely divisible random variables on $\mathbb{R}$ with Lévy triplet (γ,σ^2,ν) and $(\tilde{\gamma},\tilde{\sigma}^2,\tilde{\nu})$ respectively. Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if the following conditions are satisfied:*

(i) $\sigma^2 = \tilde{\sigma}^2$.

(ii) *The Lévy measures are equivalent with*

$$\int_{-\infty}^{\infty} (e^{\psi(x)/2} - 1)^2 \nu(dx) < \infty, \tag{6.23}$$

where $\psi(x) = \log(\frac{\tilde{\nu}(dx)}{\nu(dx)})$.

(iii) *If $\sigma = 0$ then we must in addition have*

$$\tilde{\gamma} - \gamma = \int_{-1}^{1} x(\tilde{\nu} - \nu)dx. \tag{6.24}$$

Proof: See Sato [46], p. 218 for a proof of this result.

6.A.1.1 *NIG* Distribution

For the *NIG* distribution the Lévy triplet under $\mathbb{P}$ is given by (see Schoutens [50], p. 59)

$$\begin{aligned} \sigma &= 0 && (6.25)\\ \gamma(\beta) &= \frac{2\delta\alpha}{\pi}\int_0^1 \sinh(\beta x)K_1(\alpha x)dx && (6.26)\\ \nu(\beta) &= \frac{\delta\alpha}{\pi}\frac{\exp(\beta x)K_1(\alpha|x|)}{|x|}. && (6.27)\end{aligned}$$

Proposition 6.A.2 *Let X be distributed $NIG(\alpha,\beta,\delta)$ under measure $\mathbb{P}$ and $NIG(\alpha,\tilde{\beta},\delta)$ under measure $\mathbb{Q}$ with Lévy triplets $(\gamma(\beta),0,\nu(\beta))$ and $(\gamma(\tilde{\beta}),0,\nu(\tilde{\beta}))$ respectively. Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures.*

Proof: For convenience we write $\tilde{\gamma}$ and $\tilde{\nu}$ for $\gamma(\tilde{\beta})$ and $\nu(\tilde{\beta})$ respectively. We must show that conditions (ii) and (iii) from Theorem 6.A.1 hold, since condition (i) follows easy from equation (6.25).

(ii)

$$\begin{aligned}\psi(x) &= \log(\frac{\tilde{\nu}(dx)}{\nu(dx)}) \\ &= \log\left(\frac{\frac{\delta\alpha}{\pi}\frac{\exp(\tilde{\beta}x)K_1(\alpha|x|)}{|x|}}{\frac{\delta\alpha}{\pi}\frac{\exp(\beta x)K_1(\alpha|x|)}{|x|}}\right) \\ &= (\tilde{\beta}-\beta)x \end{aligned} \tag{6.28}$$

The integral in condition (ii) can be rewritten as

$$\begin{aligned}\int_{-\infty}^{\infty}(e^{\psi(x)/2}-1)^2\nu(dx) &= \int_{|x|<1}(e^{\psi(x)/2}-1)^2\nu(dx) \\ &+ \int_{|x|>1}(e^{\psi(x)/2}-1)^2\nu(dx). \end{aligned} \tag{6.29}$$

The second integral in (6.29) is finite since

$$\begin{aligned}\int_{|x|>1}(e^{\psi(x)/2}-1)^2\nu(dx) &= \int_{|x|>1}(\sqrt{\frac{\tilde{\nu}(dx)}{\nu(dx)}}-1)^2\nu(dx) \\ &= \int_{|x|>1}(\sqrt{\tilde{\nu}(dx)}-\sqrt{\nu(dx)})^2 \\ &\leq \int_{|x|>1}\tilde{\nu}(dx)+\int_{|x|>1}\nu(dx) \\ &< \infty. \end{aligned} \tag{6.30}$$

Now

$$\int_{|x|<1}(e^{\psi(x)/2}-1)^2\nu(dx) = \int_{|x|<1}(e^{(\tilde{\beta}-\beta)x/2}-1)^2\nu(dx).$$

For $\tilde{\beta} \neq \beta$ and $|x| < 1$

$$\begin{aligned}
\left(e^{(\tilde{\beta}-\beta)x/2} - 1\right)^2 &= \left(\frac{\tilde{\beta}-\beta}{2}x + O(x^2)\right)^2 \\
&= \left(\frac{\tilde{\beta}-\beta}{2}\right)^2 x^2 + O(x^3) \\
&\leq Cx^2.
\end{aligned}$$

for some positive constant C. Hence,

$$\begin{aligned}
\int_{|x|<1} (e^{(\tilde{\beta}-\beta)x/2} - 1)^2 \nu(dx) &\leq C \int_{|x|<1} x^2 \nu(dx) \\
&< \infty
\end{aligned}$$

for all Lévy measures ν.

(iii)

$$\begin{aligned}
\tilde{\gamma} - \gamma &= \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\tilde{\beta}x) K_1(\alpha x) dx - \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) dx \\
&= \frac{2\delta\alpha}{\pi} \int_0^1 [\sinh(\tilde{\beta}x) - \sinh(\beta x)] K_1(\alpha x) dx \\
&= \frac{\delta\alpha}{\pi} \int_0^1 (e^{\tilde{\beta}x} - e^{-\tilde{\beta}x} - e^{\beta x} + e^{-\beta x}) K_1(\alpha x) dx \\
&= \frac{\delta\alpha}{\pi} \int_0^1 (e^{\tilde{\beta}x} - e^{\beta x}) K_1(\alpha x) dx - \frac{\delta\alpha}{\pi} \int_0^1 (e^{-\tilde{\beta}x} - e^{-\beta x}) K_1(\alpha x) dx \\
&= \frac{\delta\alpha}{\pi} \int_0^1 (e^{\tilde{\beta}x} - e^{\beta x}) K_1(\alpha |x|) dx - \frac{\delta\alpha}{\pi} \int_{-1}^0 (e^{\tilde{\beta}x} - e^{\beta x}) K_1(\alpha |x|) dx \\
&= \frac{\delta\alpha}{\pi} \int_0^1 (e^{\tilde{\beta}x} - e^{\beta x}) \frac{x}{|x|} K_1(\alpha |x|) dx + \frac{\delta\alpha}{\pi} \int_{-1}^0 (e^{\tilde{\beta}x} - e^{\beta x}) \frac{x}{|x|} K_1(\alpha |x|) dx \\
&= \frac{\delta\alpha}{\pi} \int_{-1}^1 (e^{\tilde{\beta}x} - e^{\beta x}) \frac{x}{|x|} K_1(\alpha |x|) dx \\
&= \int_{-1}^1 x(\tilde{\nu} - \nu) dx. \qquad (6.31)
\end{aligned}$$

■

6.A.1.2 MEIXNER DISTRIBUTION

We use a different form of γ than stated in Schoutens [50]. γ can be calculated as follows:

$$\gamma = \lim_{n\to\infty} n \int_{-1}^{1} x dF_n(x), \tag{6.32}$$

where $dF_n(x)$ is a $Meixner(\alpha, \beta, \delta/n)$ distribution (see Marshall [36]). It follows from (6.32) that

$$\begin{aligned}
\gamma(\beta) &= \lim_{n\to\infty} n \int_{-1}^{1} x \frac{(2\cos(\beta/2))^{2\delta/n}}{2\alpha\pi\Gamma(2\delta/n)} \exp(\beta x/\alpha) \left| \Gamma\left(\frac{\delta}{n} + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= \lim_{n\to\infty} n \int_{-1}^{1} x \frac{(2\cos(\beta/2))^{2\delta/n}}{2\alpha\pi\Gamma(2\delta/n+1)n/2\delta} \exp(\beta x/\alpha) \left| \Gamma\left(\frac{\delta}{n} + \frac{ix}{\alpha}\right)\right|^2 dx \\
&= \int_{-1}^{1} x \frac{\delta}{\alpha\pi} \exp(\beta x/\alpha) \left| \Gamma\left(\frac{ix}{\alpha}\right)\right|^2 dx \\
&= \int_{-1}^{1} x \frac{\delta}{\alpha\pi} \exp(\beta x/\alpha) \frac{\pi\alpha}{x\sinh(\pi x/\alpha)} dx \\
&= \delta \int_{-1}^{1} \frac{\exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)} dx. \qquad (6.33)
\end{aligned}$$

For the Meixner distribution the Lévy triplet under $\mathbb{P}$ is given by (see Schoutens [50], p. 63)

$$\sigma = 0 \tag{6.34}$$

$$\gamma(\beta) = \delta \int_{-1}^{1} \frac{\exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)} dx \tag{6.35}$$

$$\nu(\beta) = \delta \frac{\exp(\beta x/\alpha)}{x\sinh(\pi x/\alpha)}. \tag{6.36}$$

Proposition 6.A.3 *Let X be distributed $Meixner(\alpha, \beta, \delta)$ under measure $\mathbb{P}$ and $Meixner(\alpha, \tilde{\beta}, \delta)$ under measure $\mathbb{Q}$ with Lévy triplets $(\gamma(\beta), 0, \nu(\beta))$ and $(\gamma(\tilde{\beta}), 0, \nu(\tilde{\beta}))$ respectively. Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures.*

Proof: For convenience we write $\tilde{\gamma}$ and $\tilde{\nu}$ for $\gamma(\tilde{\beta})$ and $\nu(\tilde{\beta})$ respectively. We must show that conditions (ii) and (iii) from Theorem 6.A.1 hold, since condition (i) follows easy from equation (6.34).

(ii)

$$\begin{aligned}\psi(x) &= \log(\frac{\tilde{\nu}(dx)}{\nu(dx)}) \\ &= \log\left(\frac{\delta\frac{\exp(\tilde{\beta}x/\alpha)}{x\sinh(\pi x/\alpha)}}{\delta\frac{\exp(\beta x/\alpha)}{x\sinh(\pi x/\alpha)}}\right) \\ &= \frac{\tilde{\beta}-\beta}{\alpha}x. \end{aligned} \tag{6.37}$$

Looking at equations (6.29) and (6.30), we are required to show

$$\int_{-1}^{1}(e^{(\tilde{\beta}-\beta)x/2\alpha}-1)^2\nu(dx) < \infty. \tag{6.38}$$

For $\tilde{\beta} \neq \beta$ and $|x| < 1$

$$\begin{aligned}\left(e^{(\tilde{\beta}-\beta)x/2\alpha}-1\right)^2 &= \left(\frac{\tilde{\beta}-\beta}{2\alpha}x + O(x^2)\right)^2 \\ &= \left(\frac{\tilde{\beta}-\beta}{2\alpha}\right)^2 x^2 + O(x^3) \\ &\leq Cx^2. \end{aligned}$$

for some positive constant C. Hence,

$$\begin{aligned}\int_{|x|<1}(e^{(\tilde{\beta}-\beta)x/2\alpha}-1)^2\nu(dx) &\leq C\int_{|x|<1}x^2\nu(dx) \\ &< \infty \end{aligned}$$

for all Lévy measures ν.

(iii)

$$\begin{aligned}\tilde{\gamma}-\gamma &= \delta\int_{-1}^{1}\frac{\exp(\tilde{\beta}x/\alpha)}{\sinh(\pi x/\alpha)}dx - \delta\int_{-1}^{1}\frac{\exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)}dx \\ &= \delta\int_{-1}^{1}\frac{\exp(\tilde{\beta}x/\alpha)-\exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)}dx \\ &= \delta\int_{-1}^{1}\frac{x}{x}\frac{\exp(\tilde{\beta}x/\alpha)-\exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)}dx \\ &= \int_{-1}^{1}x(\tilde{\nu}-\nu)dx. \end{aligned} \tag{6.39}$$

■

6.A.2 PROOF OF PROPOSITION 6.2.1

(i) This proof follows that in Appendix 2.A.2 (i).

(ii) Let $\mathbb{KURT}[\xi_t] = k$. Now,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[h_t^2\xi_t^4] &= \mathbb{E}^{\mathbb{Q}}[\mathbb{E}[h_t^2\xi_t^4|\mathcal{F}_{t-\Delta t}]] \\ &= \mathbb{E}^{\mathbb{Q}}[h_t^2\mathbb{E}[\xi_t^4|\mathcal{F}_{t-\Delta t}]] \\ &= \mathbb{E}^{\mathbb{Q}}[h_t^2\mathbb{KURT}[\xi_t]] \\ &= k\mathbb{E}^{\mathbb{Q}}[h_t^2] \qquad (6.40)\end{aligned}$$

and

$$\begin{aligned}3\{\mathbb{E}^{\mathbb{Q}}[h_t\xi_t^2]\}^2 &= 3\{\mathbb{E}^{\mathbb{Q}}[\mathbb{E}[h_t\xi_t^2|\mathcal{F}_{t-\Delta t}]]\}^2 \\ &= 3\{\mathbb{E}^{\mathbb{Q}}[h_t\mathbb{E}[\xi_t^2|\mathcal{F}_{t-\Delta t}]]\}^2 \\ &= 3\{\mathbb{E}^{\mathbb{Q}}[h_t]\}^2. \qquad (6.41)\end{aligned}$$

We know that

$$\mathbb{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 \geq 0 \qquad (6.42)$$

therefore

$$\mathbb{E}[X^2] \geq \{\mathbb{E}[X]\}^2. \qquad (6.43)$$

It then follows that

$$\mathbb{E}^{\mathbb{Q}}[h_t^2] \geq \{\mathbb{E}^{\mathbb{Q}}[h_t]\}^2. \qquad (6.44)$$

Now, since k is always greater than 3, we see that the condition for a leptokurtic random variable (2.12) is satisfied.

(iii) Let $\mathbb{SKEW}[\xi_t] = s$.

$$\begin{aligned}
\mathbb{COV}^{\mathbb{Q}}(h_{t+\Delta t}, \xi_t) &= \mathbb{E}^{\mathbb{Q}}[h_{t+\Delta t}\xi_t] - \mathbb{E}^{\mathbb{Q}}[h_{t+\Delta t}]\mathbb{E}^{\mathbb{Q}}[\xi_t] \\
&= \mathbb{E}^{\mathbb{Q}}[(\alpha_0 + \alpha_1 h_t(\xi_t - \lambda)^2 + \beta_1 h_t)\xi_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\alpha_0\xi_t + \alpha_1 h_t(\xi_t^3 - 2\lambda\xi_t^2 + \lambda^2\xi_t) + \beta_1 h_t\xi_t] \\
&= \alpha_1\mathbb{E}^{\mathbb{Q}}[h_t\mathbb{E}[\xi_t^3 - 2\lambda\xi_t^2 + \lambda^2\xi_t|\mathcal{F}_{t-\Delta t}]] \\
&= \alpha_1\lambda\mathbb{E}^{\mathbb{Q}}[h_t(\mathbb{SKEW}[\xi_t] - 2\lambda)] \\
&= \alpha_1(s - 2\lambda)\mathbb{E}^{\mathbb{Q}}[h_t]. \qquad (6.45)
\end{aligned}$$

■

Chapter 7
Option Pricing

7.1 Introduction

This chapter is devoted to the fitting of times series and the calibration of option prices, using the GARCH models discussed in previous chapters. We discuss aspects such as parameter estimation, calibration and goodness of fit. We give a brief description of the data. This chapter is then concluded with a presentation of the results based on the data analysis.

7.2 European Options

The importance of derivatives in the world of finance is forever increasing. Many different types of derivatives exist, these include forwards, options and swaps. Options are traded both on exchanges and in the over-the-counter market. Two basic types of option exist, namely a call option and a put option. A call (put) option gives the holder the right to buy (sell) the underlying. Once this has been distinguished, options are further categorized by other aspects, including the strike price, time to maturity, exercise times and payoff.

European call (put) options give the holder the right to buy (sell) the underlying for a given price (the strike price) when the option matures on a specified date, known as the expiry date. European options are the simplest of options. The payoff of a European call option, with strike price K, is given by:

$$payoff = \begin{cases} S_T - K & \text{if } S_T > K \\ 0 & \text{otherwise,} \end{cases} \tag{7.1}$$

this can be written more concisely as $(S_T - K)^+$.

7.2.1 Pricing Formula for European Call Options

Let $C_t(K,T)$ denote the value (price) of a European call option, with strike K and maturity T, at time t. For a European call option the value, $C_t(K,T)$, is given by the discounted expectation of the payoff under the risk-neutral measure $\mathbb{Q}$ (see Delbaen and Schachermayer [20]).

$$C_t(K,T) = \exp\{-r(T-t)\}\mathbb{E}^{\mathbb{Q}}\Big[\max\{S_T - K, 0\}\Big], \tag{7.2}$$

where r is the risk free interest rate.

Since no analytical formulas exist for the Gaussian-GARCH, NIG-GARCH and Meixner-GARCH option pricing models, the expectation in equation (7.2) is calculated using Monte Carlo simulation.

7.3 Parameter Estimation

In select cases it is an easy task to decide how to estimate a parameter and often intuition can lead us to good parameter estimates. However in complicated models, such as the option pricing models we have discussed, we need a more theoretical approach in estimating parameters. Methods of estimating parameters include the method of moments and maximum likelihood. Estimating parameters using maximum likelihood is by far the most popular technique and we also use it in our estimation procedure. We employ a two part parameter estimation procedure. Firstly we calculate the maximum likelihood estimates on the stock price series and then use the maximum likelihood estimates as initial estimates in our calibration procedure on option data. It is relevant to use the real world model to get initial parameter estimates because, when ϵ_t has a NIG or Meixner distribution, the distribution of ξ_t falls into the same class of distributions as ϵ_t (see Chapter 6). Calibration entails matching an option pricing model to observed market prices by minimizing the root mean square error (see Section 7.3.3.1) between the market and the model prices.

7.3.1 Maximum Likelihood Estimators

Before defining the concept of maximum likelihood estimators, we need the following definition:

Definition 7.3.1 (Likelihood function) *Let $x_1, x_2, ..., x_n$ be an i.i.d. sample from a population with pdf $f(x;\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)$. Then the likelihood function is defined by*

$$L(\boldsymbol{\theta};\mathbf{x}) = L(\theta_1, ..., \theta_k; x_1, ..., x_n) = \prod_{i=1}^{n} f(x_i;\theta_1, ..., \theta_k) \tag{7.3}$$

where $\mathbf{x} = \{x_1, x_2, ..., x_n\}$.

We now define the concept of maximum likelihood estimators.

Definition 7.3.2 (Maximum likelihood estimators) *For each sample point x, let $\hat{\boldsymbol{\theta}}(x)$ be a parameter value at which $L(\boldsymbol{\theta};x)$ attains its maximum as a function of $\boldsymbol{\theta}$, with x held fixed. A maximum likelihood estimator of the parameter $\boldsymbol{\theta}$ based on a sample $\mathbf{x} = (x_1, x_2, ..., x_n)$ is $\hat{\boldsymbol{\theta}}(\mathbf{x})$, where*

$$\hat{\boldsymbol{\theta}}(\mathbf{x}) = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}|\mathbf{x}) \tag{7.4}$$

However, in many cases it is easier to use the natural log of the likelihood function, defined as

$$l(\boldsymbol{\theta};\mathbf{x}) = \log L(\boldsymbol{\theta};\mathbf{x}) = \sum_{i=1}^{n} \log f(x_i;\theta_1, ..., \theta_k). \tag{7.5}$$

Since the log function is monotone increasing, maximizing the likelihood function is equivalent to maximizing the log-likelihood function. Hence we have the following relationship

$$\hat{\boldsymbol{\theta}}(\mathbf{x}) = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; \mathbf{x}) = \arg\max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}; \mathbf{x}). \tag{7.6}$$

The abbreviation MLE (maximum likelihood estimate) is used when referring to the realized value of the estimator. Intuitively, the MLE can be defined as the parameter value for which the observed sample is most likely. Also note, by the construction of the maximum likelihood estimator, the range of the MLE coincides with that of the parameter. Solving for the MLE analytically is impossible except in some special cases, for example the normal distribution. Often, in solving for the maximum likelihood estimates, it is best to rely on a highly efficient optimization package. In fact, this is one of the most important features of maximum likelihood estimation. If the likelihood function can be expressed explicitly, then there is hope of maximizing the likelihood function numerically. However a drawback to using optimization packages is that one generally requires an initial estimate for the MLEs and if the likelihood function has multiple local maxima, the MLEs are often dependent on the initial starting values.

Let $y_{i\Delta t} = \log \frac{S_{i\Delta t}}{S_{(i-1)\Delta t}}$ for $i = 1, ..., n.$, denote the log returns. Then our 'real world' model is given by

$$y_{i\Delta t} = r\Delta t + \lambda\sqrt{h_t} - \kappa(\sqrt{h_t}; \boldsymbol{\theta}) + \sqrt{h_{i\Delta t}}\epsilon_{i\Delta t}, \tag{7.7}$$

where $\epsilon_{i\Delta t} \overset{d}{\sim} (0, 1)$. Then, using equation (7.5), the log-likelihood functions are given by:

For $\epsilon_{i\Delta t} \overset{d}{\sim} N(0, 1)$

$$l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{n}\left[\left(\frac{y_{i\Delta t} - r\Delta t - \lambda\sqrt{h_{i\Delta t}} + \frac{1}{2}h_{i\Delta t}}{\sqrt{h_{i\Delta t}}}\right)^2 + \log h_{i\Delta t}\right]. \tag{7.8}$$

For $\epsilon_{i\Delta t} = \frac{Z_{i\Delta t} - \mu_Z}{\sigma_Z}$, where $Z_{i\Delta t} \overset{d}{\sim} NIG(\alpha, \beta, \delta)$ and μ_Z and σ_Z are given by (4.4) and (4.5) respectively,

$$l(\boldsymbol{\theta};\mathbf{y}) = \sum_{i=1}^{n} \log f_{NIG}\left(\frac{y_{i\Delta t} - r\Delta t - \lambda\sqrt{h_{i\Delta t}} + \kappa(\sqrt{h_{i\Delta t}}/\sigma_Z;\boldsymbol{\theta})}{\sqrt{h_{i\Delta t}}/\sigma_Z};\boldsymbol{\theta}\right) - \frac{1}{2}\sum_{i=1}^{n}\log\frac{h_{i\Delta t}}{\sigma_Z^2}. \tag{7.9}$$

For $\epsilon_{i\Delta t} = \frac{Z_{i\Delta t}-\mu_Z}{\sigma_Z}$, where $Z_{i\Delta t} \overset{d}{\sim} Meixner(\alpha,\beta,\delta)$ and μ_Z and σ_Z are given by (5.3) and (5.4) respectively,

$$l(\boldsymbol{\theta};\mathbf{y}) = \sum_{i=1}^{n} \log f_{Meixner}\left(\frac{y_{i\Delta t} - r\Delta t - \lambda\sqrt{h_{i\Delta t}} + \kappa(\sqrt{h_{i\Delta t}}/\sigma_Z;\boldsymbol{\theta})}{\sqrt{h_{i\Delta t}}/\sigma_Z};\boldsymbol{\theta}\right) - \frac{1}{2}\sum_{i=1}^{n}\log\frac{h_{i\Delta t}}{\sigma_Z^2}. \tag{7.10}$$

The likelihood function is dependant on the choice of the starting values ϵ_0 and h_0. However, for large samples the impact of the starting values on the estimation results is negligible. Therefore, we set $\epsilon_0 = 0$ and h_0 equal to the stationary variance of the return process $\sqrt{h_{i\Delta t}}\epsilon_{i\Delta t}$, i.e. $h_0 = \frac{\alpha_0}{1-(1+\lambda^2)\alpha_1-\beta_1}$.

Maximizing the likelihood functions (7.8), (7.9) and (7.10) leads to estimates $\hat{\boldsymbol{\theta}}(\mathbf{y})$ for the unknown model parameters $\boldsymbol{\theta}$. From the estimates $\hat{\boldsymbol{\theta}}(\mathbf{y})$ we can obtain the time series of empirical residuals $\hat{\epsilon}_{i\Delta t}$ for $i = 1, ..., n$. The empirical residuals are calculated using the following equation

$$\hat{\epsilon}_{i\Delta t} = \frac{y_{i\Delta t} - r\Delta t - \lambda\sqrt{\hat{h}_{i\Delta t}} + \kappa(\sqrt{\hat{h}_{i\Delta t}};\hat{\boldsymbol{\theta}})}{\sqrt{\hat{h}_{i\Delta t}}} \tag{7.11}$$

where $\hat{h}_{i\Delta t}$ is obtained from

$$\hat{h}_{i\Delta t}\hat{\alpha}_0 + \hat{\alpha}_1 h_{i\Delta t}\hat{\epsilon}^2_{i\Delta t} + \hat{\beta}_1\hat{h}_{i\Delta t} \tag{7.12}$$

and $\hat{\epsilon}_0 = \epsilon_0$ and $\hat{h}_0 = h_0$.

7.3.2 Goodness of Fit

To assess the goodness of fit of the Gaussian-GARCH and Lévy-GARCH models to a series of asset closing prices, we use the chi-squared test (χ^2 test).

Define the null hypotheses as follows:

H_0^{normal} : $\epsilon_{i\Delta t}$ follows the standard normal distribution.

H_0^{NIG} : $\epsilon_{i\Delta t}$ follows the standard normal inverse Gaussian distribution.

$H_0^{Meixner}$: $\epsilon_{i\Delta t}$ follows the standard Meixner distribution.

7.3.2.1 THE χ^2 TEST

A general method for comparing fact with theory, namely the chi-squared test or χ^2 test, was devised by Karl Pearson (1857 - 1936). The χ^2 test is the most well known test for the goodness of fit problem.

Let $A_1, A_2, ..., A_m$ denote the division of the sample space into m cells of equal width, for a random sample $\mathbf{x}$. Let $\hat{\boldsymbol{\theta}}_0$ denote the MLE of the parameter $\boldsymbol{\theta}$ of a distribution F under the null hypothesis. Then the expected number of observations in the i^{th} cell, $\hat{e}_i$, is given by

$$\hat{e}_i = n\hat{p}_i, \qquad i = 1, 2, ..., m, \tag{7.13}$$

where n denotes the sample size and $\hat{p}_i$, for $i = 1, 2, ..., m$, is given by

$$\hat{p}_i = P[X \in A_i] = F(A_i; \hat{\boldsymbol{\theta}}_0) - F(A_{i-1}; \hat{\boldsymbol{\theta}}_0), \tag{7.14}$$

Let o_i, for $i = 1, 2, ...m$, denote the number of observations from a random sample $\mathbf{x}$ falling into the i^{th} cell, A_i. Then the χ^2 statistic is given by

$$\hat{\chi}^2 = \sum_{i=1}^{m} \frac{(o_i - \hat{e}_i)^2}{\hat{e}_i}, \tag{7.15}$$

and if n is large, $\hat{\chi}^2$ has approximately a chi-squared distribution with $(m-1-k)$ degrees of freedom where k is the number of parameters in the null distribution that we must estimate. We therefore reject H_0 at significance level α if

$$\hat{\chi}^2 > \chi^2_{m-1-k;1-\alpha} \tag{7.16}$$

where $\chi^2_{m-1-k;1-\alpha}$ denotes the $100(1-\alpha)^{th}$ percentile of the chi-squared distribution with $(m-1-k)$ degrees of freedom.

7.3.3 Calibration

We calibrate our different models to a given set of option data. In the figures we denote the market prices by a circle ($\bigcirc$), the in-sample calibrated prices by plus sign (+) and the out-of-sample predictions by a star ($\star$). Our goal is to get the plus signs and stars as close to the circles as possible.

Except in the calculation of initial parameter estimates under the real world model, we do not explicitly use any historical data in the calibration process. All necessary information is contained in the option prices, which are observed in the market.

Option prices, from the GARCH models, are calculated using Monte Carlo simulation and equation (7.2). We simulate n stock paths, using the risk neutral stock path processes, and estimate the expectation by

$$\frac{1}{n}\sum_{i=1}^{n}\max\{S_T^i-K,0\} \tag{7.17}$$

where $i, \forall i=1,2,...,n.$, denotes the i^{th} stock path .

7.3.3.1 RMSE

For an estimate of the goodness of calibration, we calculate the root-mean-square error (RMSE):

$$\text{RMSE}=\sqrt{\sum_{\text{options}}\frac{(\text{market price - model price})^2}{\text{number of options}}} \tag{7.18}$$

In the calibration procedure we estimate the model parameters by minimizing the root-mean-square error (RMSE) between the model and market prices.

In-sample calibrations and out-of-sample predictions are calculated on the option price data. In-sample calibrations refers to the procedure discussed above whereby the parameters are estimated through the minimization of some criteria, in our case the RMSE. Out-of-sample refers to the predicting of the option prices using the parameters estimated by the in-sample calibration or maximum likelihood procedure. We then calculate the RMSE for these predicted prices.

7.4 DATA

7.4.1 S&P 500

The S&P 500 data set consists of both the Index series and a set of option prices.

The Index series consists of the closing prices from the 2^{nd} of January 1990 till the 18^{th} of April 2002 (see Figure 7.1). The option set consists of 75 mid-prices of a set of European call options on the S&P 500 Index (see Figure 7.2) at the close of the market on 18 April 2002. On this date the S&P 500 Index closed at 1124.47. The risk-free interest rate is given as 0.7%. For the exact option prices and their maturities see Table 7.8 in Appendix 7.A.1.

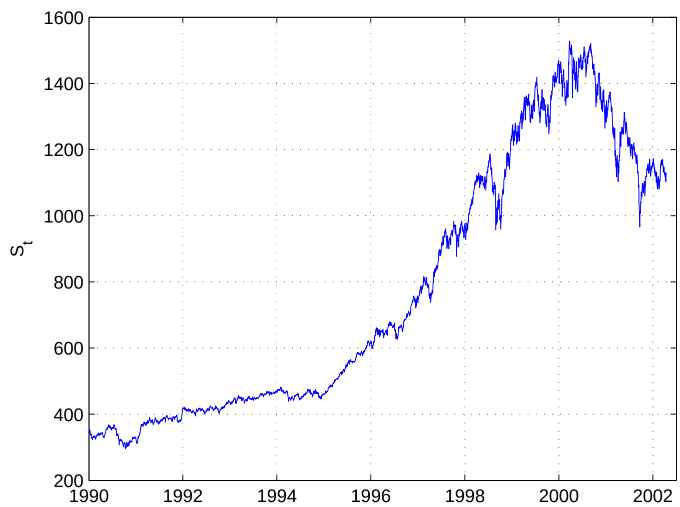

Figure 7.1: S&P 500 Index Series, 2 January 1990 - 18 April 2002.

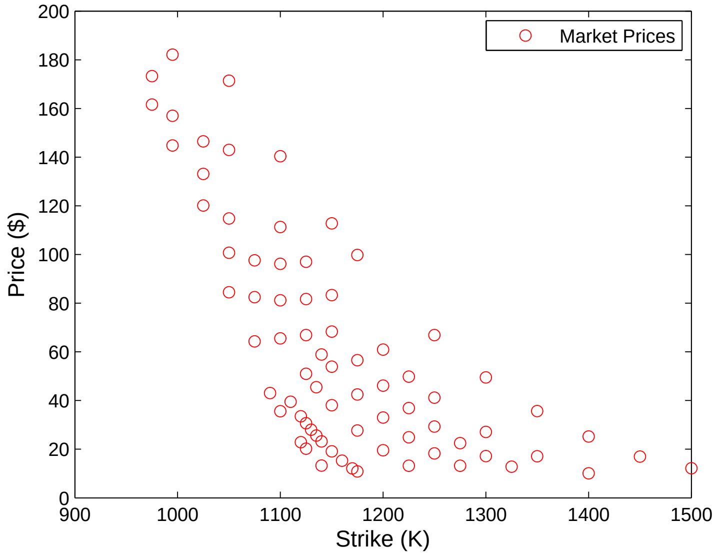

Figure 7.2: S&P 500 Option prices.

7.4.2 S&P 100

The S&P 100 data set also consists of both the Index series and a set of option prices.

The Index series consists of the closing prices from the 4^{th} of March 1998 till the 4^{th} of March 2008 (see Figure 7.3). The option set consists of mid-prices of a set of European call options on the S&P 100 Index (see Figure 7.4) at the close of the market on 4 March 2008. On this date the S&P 100 Index closed at 611.15. For the risk-free interest rate we used a zero coupon swap yield curve. For the exact option prices and their maturities see Tables 7.9 and 7.10 in Appendix 7.A.2.

Unlike the S&P 500 data, the risk-free interest rate we use in the S&P 100 data analysis is nonconstant. We thus replace r in our GARCH models by r_t. It then follows that the stock price processes under measure $\mathbb{P}$ and $\mathbb{Q}$ are given by

$$S_t = S_{t-\Delta t} \exp\left(r_t \Delta t + \lambda\sqrt{h_t} - \kappa(\sqrt{h_t};\boldsymbol{\theta}) + \sqrt{h_t}\epsilon_t\right) \tag{7.19}$$

and

$$S_t \;=\; S_{t-\Delta t}\exp\left(r_t\Delta t - \kappa(\sqrt{h_t};\boldsymbol{\theta}) + \sqrt{h_t}\xi_t\right) \tag{7.20}$$

respectively. The option price at time t with strike K and maturity T (see Equation (7.2)) then becomes

$$C_t(K,T) = \exp\{-\sum_{t}^{T} r_t\Delta t\}\mathbb{E}^{\mathbb{Q}}\Big[\max\{S_T - K, 0\}\Big]. \tag{7.21}$$

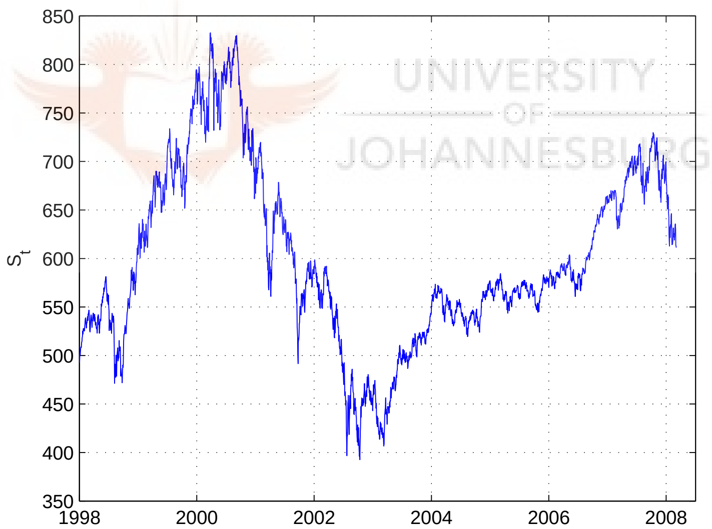

Figure 7.3: S&P 100 Index Series, 4 March 1998 - 4 March 2008.

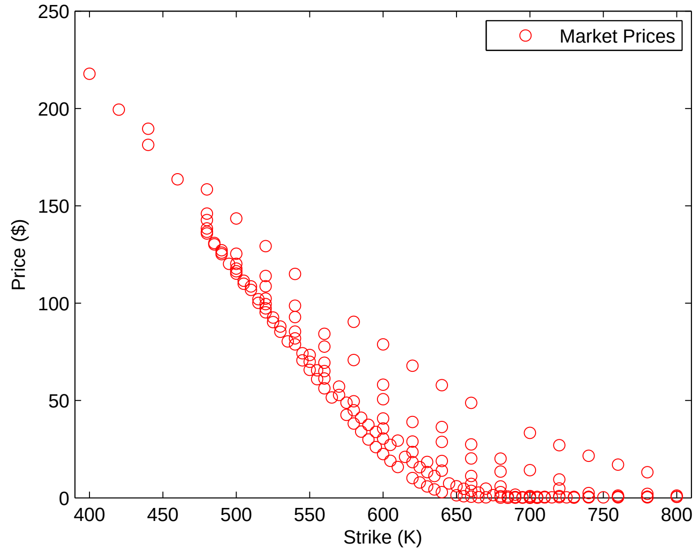

Figure 7.4: S&P 100 Option prices.

7.5 Results

The maximum likelihood estimates for the S&P 500 and S&P100 Index series are given in Tables 7.1 and 7.2 respectively. The χ^2 statistics and the respective p-values for the two data sets are given in Table 7.3. The chi-squared statistic was calculated on the range of five standard deviations either side of the mean or [-5,5] since the innovations are distributed with zero mean and standard deviation one. This range was then partitioned into 80 equal intervals.

Looking at the results in Table 7.3, we see that the hypothesis that $\epsilon_{i\Delta t}$ is distributed standard normal is rejected for the S&P 500 and S&P 100 index series. At a 5% significance level, we were unable to reject the hypotheses that $\epsilon_{i\Delta t}$ has a NIG or Meixner distribution for both data sets. With p-values of 0.75 and 0.69 respectively, we see that the NIG and Meixner distributions fit the innovations extremely well in the S&P 100 index. This is evident in the Q-Q plots (see Figures 7.5, 7.6 and 7.7).

	Model		
Parameters	Gaussian-GARCH	NIG-GARCH	Meixner-GARCH
α_0	5.8135x10^{-7}	1.0225x10^{-6}	7.6620x10^{-7}
α_1	0.0578	0.0367	0.0640
β_1	0.9376	0.9063	0.9233
α	-	1.4970	1.3376
β	-	-0.3590	-0.5524
δ	-	2.6038	1.0159
λ	0.0700	0.0578	0.0066
Log-Likelihood	13118.2315	10363.2894	10361.8063

Table 7.1: S&P 500 Maximum Likelihood Estimates.

	Model		
Parameters	Gaussian-GARCH	NIG-GARCH	Meixner-GARCH
α_0	9.9533x10^{-7}	9.6983x10^{-7}	6.9593x10^{-7}
α_1	0.0669	0.1048	0.1100
β_1	0.9254	0.8813	0.8868
α	-	4.3854	0.5105
β	-	-1.3997	-0.7752
δ	-	2.1590	3.4219
λ	0.0440	0.0025	0.0001
Log-Likelihood	10240.1408	8011.3395	8009.6526

Table 7.2: S&P 100 Maximum Likelihood Estimates.

Data	Model	χ^2	p-value
S&P 500	Gaussian-GARCH	7290.19	0
	NIG-GARCH	94.80	0.094
	Meixner-GARCH	93.66	0.095
S&P 100	Gaussian-GARCH	3950.93	0
	NIG-GARCH	68.26	0.751
	Meixner-GARCH	71.19	0.694

Table 7.3: Goodness of Fit Statistics.

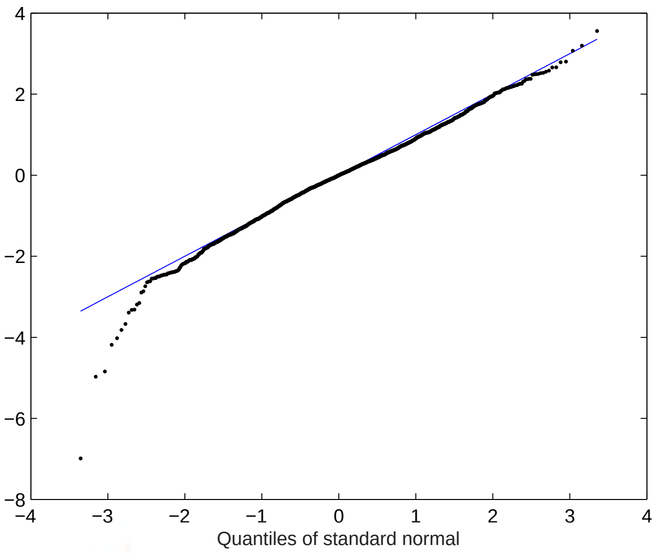

Figure 7.5: Gaussian-GARCH model: Q-Q plot of S&P 100 residuals.

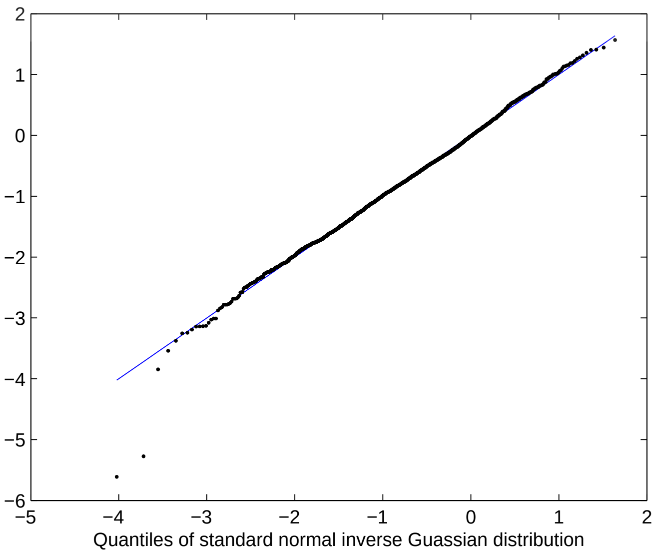

Figure 7.6: *NIG*-GARCH model: Q-Q plot of S&P 100 residuals.

Figure 7.7: Meixner-GARCH model: Q-Q plot of S&P 100 residuals.

From Figure 7.5 we see that the residuals from the Gaussian-GARCH model do not fit the normal distribution very well, especially in the left tail. In Figures 7.6 and 7.7 we see that the residuals for the NIG-GARCH and Meixner-GARCH models, except for a few points in the left tail, are close to their model-implied distributions.

The calibration results include the results for both the in-sample and out-of-sample tests. The S&P 500 in-sample test was calibrated on the first four maturities for all strikes (43 options), while the S&P 100 in-sample test was calibrated on the first three maturities for all strikes (117 options). The out-of-sample predictions where calculated on the remaining options. Tables 7.4 and 7.5 give the RMSE for the S&P 500 and S&P 100 options respectively. The S&P 500 and S&P 100 calibrated parameters are given in Tables 7.6 and 7.7 respectively.

The first couple of maturities were chosen for the in-sample calibrations. This was done so that the out-of-sample predictions would be a forecast of long dated options and future options that would become available in the market.

Models Calibrated	RMSE	
	In-sample	Out-of-sample
Black-Scholes	3.60	7.14
Gaussian-GARCH	2.83	6.35
NIG-GARCH	1.07	4.05
Meixner-GARCH	1.07	3.21

Table 7.4: S&P 500 Calibration Results: Measures of fit.

Models Calibrated	RMSE	
	In-sample	Out-of-sample
Black-Scholes	2.93	8.98
Gaussian-GARCH	2.81	7.39
NIG-GARCH	1.67	6.22
Meixner-GARCH	1.84	6.73

Table 7.5: S&P 100 Calibration Results: Measures of fit.

We see, from Tables 7.4 and 7.5, that the NIG-GARCH and Meixner-GARCH models outperform the Gaussian-GARCH model in both the in-sample calibrations as well as out-of-sample predictions. The in-sample RMSE for the Lévy Models is significantly lower than the Gaussian-GARCH models.

	Model		
Parameters	Gaussian-GARCH	NIG-GARCH	Meixner-GARCH
α_0	5.4390x10^{-7}	1.06x10^{-7}	3.5931x10^{-7}
α_1	0.0679	0.0063	0.0014
β_1	0.9271	0.9907	0.9951
α	-	0.2251	7.9649
β	-	-0.0301	-0.1800
δ	-	6.4955	0.8111
λ	5.2981x10^{-4}	3.5428x10^{-8}	7.3532x10^{-8}

Table 7.6: S&P 500 Calibration Results: Parameters.

	Model		
Parameters	Gaussian-GARCH	NIG-GARCH	Meixner-GARCH
α_0	$1.4977 \text{x} 10^{-6}$	$1.7236 \text{x} 10^{-6}$	$9.9509 \text{x} 10^{-7}$
α_1	0.0430	0.0162	0.0273
β_1	0.9515	0.9807	0.9706
α	-	1.7439	0.5609
β	-	-0.2066	-0.6237
δ	-	3.6055	4.0280
λ	$2.5659 \text{x} 10^{-4}$	$3.6055 \text{x} 10^{-6}$	$3.0736 \text{x} 10^{-6}$

Table 7.7: S&P 100 Calibration Results: Parameters.

Figures 7.8, 7.9 and 7.10 show the market prices of the S&P 500 options with the calibrated option prices superimposed for the Gaussian-GARCH, NIG-GARCH and Meixner-GARCH models respectively.

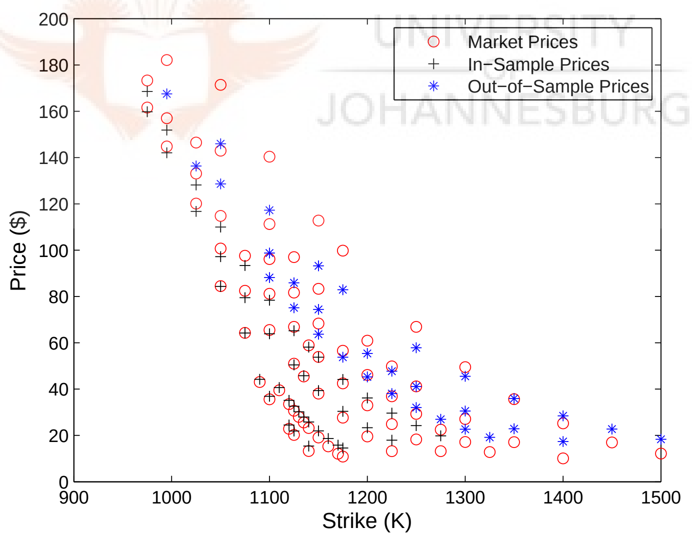

Figure 7.8: S&P 500 Gaussian-GARCH Calibration.

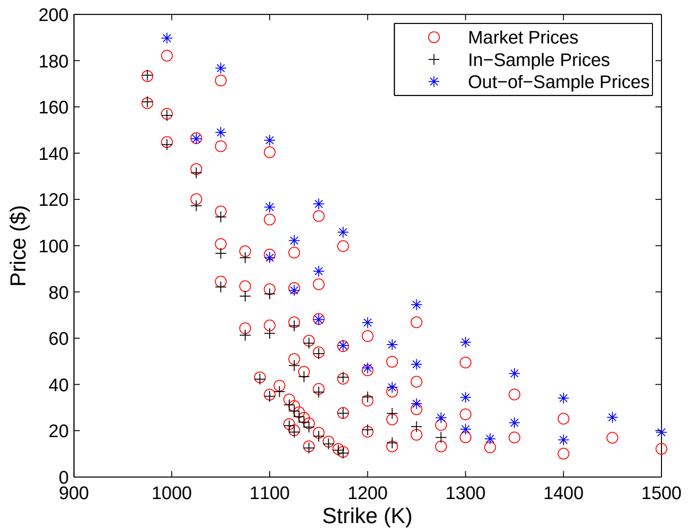

Figure 7.9: S&P 500 NIG-GARCH Calibration.

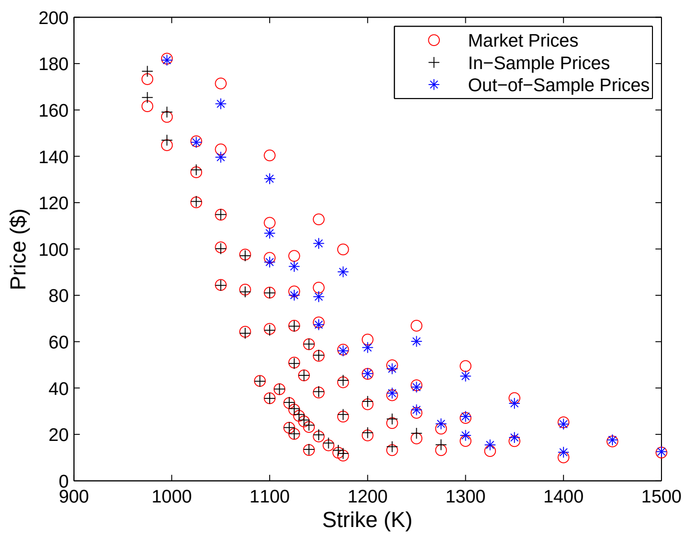

Figure 7.10: S&P 500 Meixner-GARCH Calibration.

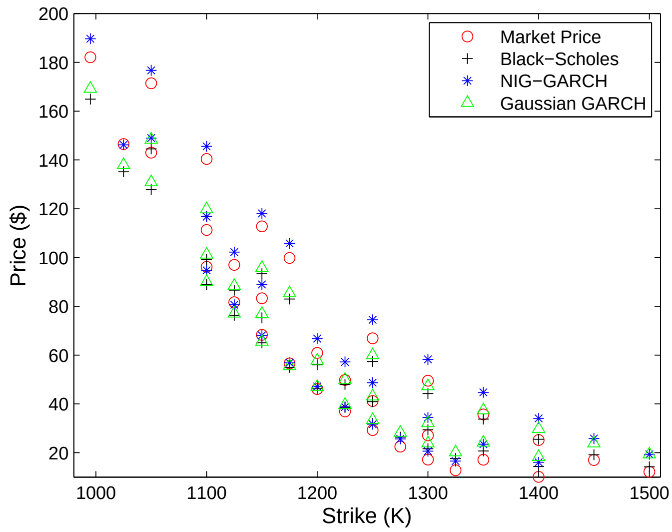

Figure 7.11: S&P 500 Comparison of out-of-sample prices.

Figure 7.11 plots the out-of-sample prices for the Black-Scholes, Gaussian-GARCH and NIG-GARCH models against the market prices. We left the Meixner-GARCH model prices out so that the figure is not cluttered. We can clearly see that the NIG-GARCH outperforms the Gaussian-GARCH and Black-Scholes models in predicting the market prices. This is most evident in the out of the money options, were the Black-Scholes and Gaussian-GARCH models greatly undervalue the option prices.

Figures 7.12, 7.13 and 7.14 plot the market prices of the S&P 100 options with the calibrated option prices superimposed for the Gaussian-GARCH, NIG-GARCH and Meixner-GARCH models respectively. Since the S&P 100 option prices are congested (see Figure 7.4) only the options with strikes between \$550 and \$750 are plotted. We see that the Lévy GARCH models provide better fits than the Gaussian-GARCH model.

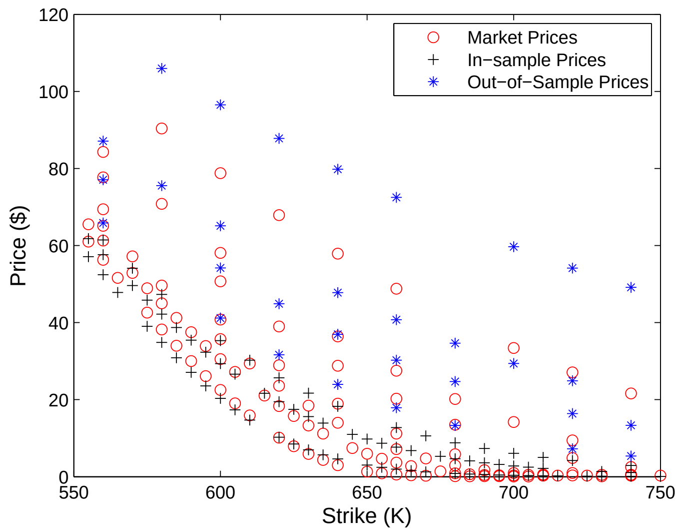

Figure 7.12: S&P 100 Gaussian-GARCH Calibration.

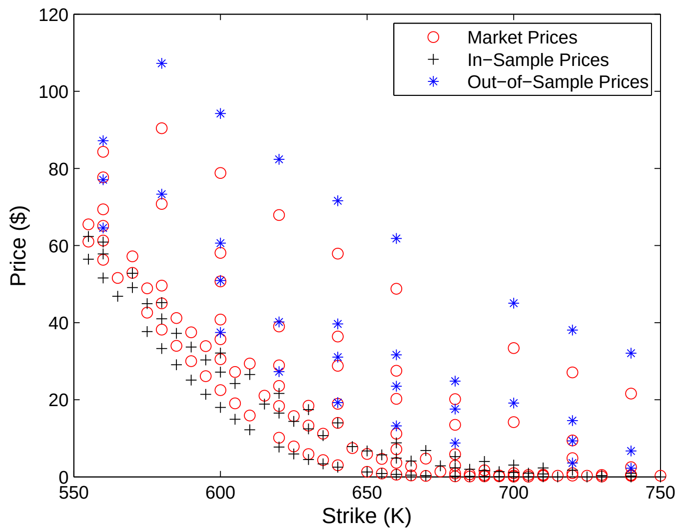

Figure 7.13: S&P 100 NIG-GARCH Calibration.

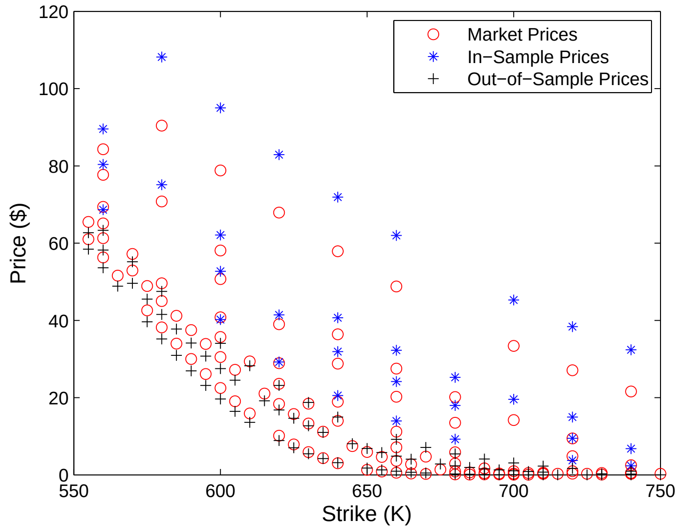

Figure 7.14: S&P 100 Meixner-GARCH Calibration.

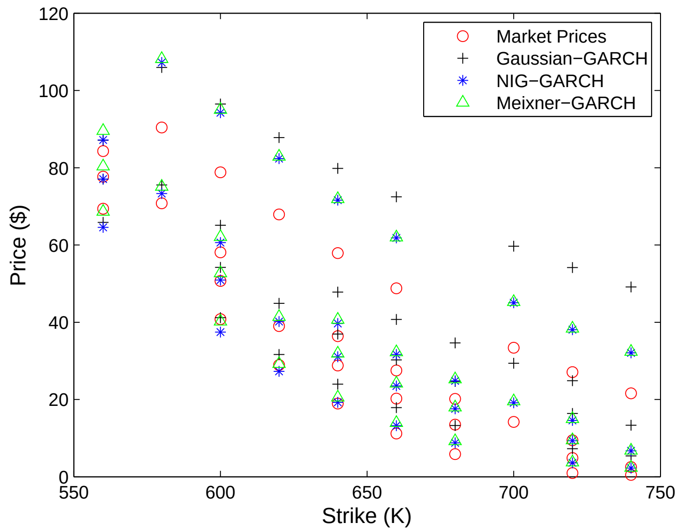

Figure 7.15: S&P 100 Comparison of out-of-sample prices.

Figure 7.15 plots the out-of-sample predictions for the Gaussian-GARCH, NIG-GARCH and Meixner-GARCH models against the market prices. Here we see, like the S&P 500 options, that the Lévy models outperform the Gaussian-GARCH model. The NIG-GARCH and Meixner-GARCH provide very similar predictions for the market option prices.

7.5.1 Impact of the restriction $g(\boldsymbol{\theta})$

In the empirical analysis, not a single instance occurred where the simulated squared volatility process in the NIG-GARCH and Meixner-GARCH models reached the cap value, $g(\boldsymbol{\theta})$, for both the S&P 500's and S&P 100's maximum likelihood estimates and the calibrated parameters. Hence, the additional restriction had no impact in the results obtained for these two data sets. Whether or not this restriction will influence the results for other data sets is not known.

7.A Appendix

7.A.1 S&P 500 Data

Strike	May 2002	June 2002	Sep. 2002	Dec. 2002	March 2003	June 2003	Dec. 2003
975			161.60	173.30			
995			144.80	157.00		182.10	
1025			120.10	133.10	146.50		
1050		84.50	100.70	114.80		143.00	171.40
1075		64.30	82.50	97.60			
1090	43.10						
1100	35.60		65.50	81.20	96.20	111.30	140.40
1110		39.50					
1120	22.90	33.50					
1125	20.20	30.70	51.00	66.90	81.70	97.00	
1130		28.00					
1135		25.60	45.50				
1140	13.30	23.20		58.90			
1150		19.10	38.20	53.90	68.30	83.30	112.80
1160		15.30					
1170		12.10					
1175		10.90	27.70	42.50	56.60		99.80
1200			19.60	33.00	46.10	60.90	
1225			13.20	24.90	36.90	49.80	
1250				18.30	29.30	41.20	66.90
1275				13.20	22.50		

Table 7.8: S&P 500 Option Data.

7.A.2 S&P 100 Data

Strike	22 March 2008	19 April 2008	17 May 2008	19 June 2008	20 Sep. 2008	20 Dec. 2008	19 Dec. 2009
400						217.80	
420						199.40	
440						181.30	189.60
460						163.60	
480		135.80	136.70	138.40	142.70	146.00	158.40
485	130.20	130.90					
490	125.20	126.00	127.20				
495	120.20						
500	115.20	116.40	117.80	120.20	125.40		143.50
505	110.00	111.50					
510		106.80	108.60				
515	100.20	102.00					
520	95.30	97.30	99.50	102.40	108.70	114.00	129.30
525	90.30	92.60					
530	85.30	88.00					
535	80.40						
540		78.80	81.90	85.40	92.90	98.70	115.00
545	70.60	74.30					
550	65.80	69.90	73.40				
555	61.00	65.50					
560	56.30	61.30	65.10	69.40	77.70	84.30	
565	51.60						
570		52.90	57.20				
575	42.60	48.90					
580	38.20	45.00	49.60			70.80	90.40
585	34.00	41.20					
590	30.00	37.50					
595	26.10	33.90					
600	22.50	30.50	35.70	40.80	50.70	58.10	78.80
605	19.05	27.20					

Table 7.9: S&P 100 Option Data (Part 1).

Strike	22 March 2008	19 April 2008	17 May 2008	19 June 2008	20 Sep. 2008	20 Dec. 2008	19 Dec. 2009
610	15.90		29.40				
615		21.05					
620	10.15	18.35	23.60	28.90	39.00		67.90
625	7.90	15.75					
630	5.90	13.30	18.45				
635	4.35	11.20					
640	3.00		14.00	18.95	28.80	36.40	57.90
645		7.45					
650	1.30	5.95					
655	0.88	4.65					
660	0.57	3.65	7.15	11.20	20.25	27.50	48.80
665	0.38	2.75					
670	0.25		4.75				
675		1.40					
680	0.13	0.88	2.95	5.85	13.50	20.15	
685	0.08	0.65					
690	0.15	0.45	1.70				
695	0.15	0.38					
700	0.05	0.35	0.95			14.20	33.40
705	0.05	0.50					
710	0.30	0.28	0.55				
715	0.30						
720			0.35	0.95	4.85	9.45	27.10
725	0.30						
730	0.10	0.50					
740	0.30	0.50	0.28	0.50	2.50		21.60
750	0.30						
760	0.30	0.50		0.50	1.15		17.05
780	0.30	0.50	0.50	0.50	0.55	2.10	13.20
800		0.50		0.50		1.15	

Table 7.10: S&P 100 Option Data (Part 2).

Summary and Conclusion

Duan [21] proposed the first option pricing model which modeled volatility using a GARCH process. In this dissertation we have explored an extension of Duan's [21] GARCH option pricing model. This extension allowed us to drop the assumption of lognormal returns. We replaced the assumption of normality by proposing Lévy process innovations. These models were aptly named Lévy-GARCH option pricing models. More specifically, the normal inverse Gaussian and Meixner distributions were proposed. The main advantage of the models we proposed was that they allowed the daily conditional logarithmic returns to be heavy tailed and skewed. These GARCH models (both Duan's and the Lévy models) are able to capture another stylized fact of financial data, namely volatility clustering. A disadvantage of the Lévy-GARCH models was the necessity of incorporating an additional restriction on the evolution of the squared volatility. For the NIG and Meixner distributions, the additional restriction required us to cap the squared volatility process.

The normal inverse Gaussian and Meixner distributions were shown to be acceptable candidates to govern the innovation process. Both distributions can be skewed, either positively or negatively, and have kurtosis greater than that of the normal distribution. These distributions still allow for desirable properties of Duan's GARCH option pricing model to persist. The normal inverse Gaussian and Meixner innovations were standardized so that the innovation process had an equivalent mean and variance to the Gaussian-GARCH's innovation process. This allowed for direct comparisons to Duan's Gaussian-

GARCH model.

In addition, we were able to show that an equivalent martingale measure for the NIG or Meixner distribution exists and furthermore that the equivalent martingale measure generates respectively a NIG or Meixner distribution.

We proposed a new method to generate random numbers from a Meixner distribution. This method is based on the rejection method discussed in Ross [43]. Using the NIG distribution as a basis, we matched the first three moments of the Meixner and NIG distributions so that the distributions were as 'close' as possible. However, matching the moments is not always possible. In the cases where the moments are not matched, we used a transform of the Meixner parameters to get the NIG parameters. A brief justification was given on the choice of transformation used. We showed, in a simulation study, that the rejection method was significantly more efficient, irrespective of whether the moments can or cannot be matched, than the inverse transform method.

In our empirical analysis we used data on the S&P 500 and S&P 100 index series as well as a set of European options on each index. Empirical tests on the ability of the GARCH models to price options were provided.

Modeling the innovations with the NIG and Meixner distributions significantly improved the goodness of fit statistics. Furthermore, the χ^2 test rejected the hypothesis that the innovations are normally distributed but did not reject the hypothesis that the innovations are normal inverse Gaussian or Meixner distributed.

The Lévy GARCH models outperformed the Gaussian-GARCH and Black-Scholes models in both in-sample calibrations and out-of-sample predictions. Overall, we conclude that the Lévy GARCH models outperform the Gaussian-GARCH model in both fitting the stock price process and in the calibrating and predicting of option prices. The skewness and semi-heavy tails of the NIG and Meixner innovations seem to generate significant improvements in the empirical results.

The main drawbacks to our model was the necessity of incorporating a restriction on the squared volatility process and that there was no leverage effect incorporated into our model. Looking ahead we would try to relieve the Lévy GARCH models of these drawbacks.

The first drawback is a difficult one to overcome. Due to the complex nature of the characteristic functions for the distributions we have considered and many other more general heavy tailed distributions, this restriction must be incorporated. However, our empirical results were not affected by this restriction as the squared volatility process never reached the cap value. This could be due to the data sets we used or there is a remote possibility that this will always be the case irrespective of the data used. This will be left for further investigation.

As for the leverage effect, it can be incorporated into the model with a modification in the squared volatility process. We can replace the GARCH(1,1) process by one of two modified GARCH processes. The first being the nonlinear asymmetric GARCH or NGARCH model. The NGARCH(1,1) model is given by:

$$h_t = \alpha_0 + \alpha_1 h_t(\epsilon_t - \rho)^2 + \beta_1 h_t. \tag{7.22}$$

The other option is to incorporate positive innovations and negative innovations separately through the use of indicator functions. The innovations are then given different weights in the squared volatility process. This is done as follows:

$$h_t = \alpha_0 + \alpha_1 h_t \epsilon_t^2 \mathbb{I}(\epsilon_t > 0) + \alpha_2 h_t \epsilon_t^2 \mathbb{I}(\epsilon_t < 0) + \beta h_t. \tag{7.23}$$

These two variations of the GARCH process add an additional parameter to the model. The increase in pricing performance of the GARCH models due to the inclusion of this additional parameter is also left for further investigation.

Bibliography

[1] Abramowitz, M. and Stegun, I.A. (1968) *Handbook of mathematical functions.* New York: Dover Publications.

[2] Anderson, T.W. and Darling, D.A. (1954) *Asymptotic theory of certain goodness of fit criteria based on stochastic processes.* The annals of mathematical statistics, Vol. 23, No. 2, pp. 193-212.

[3] Anderson, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001) *The distribution of realized stock return volatility.* Journal of financial economics, Vol. 61, pp. 43-76.

[4] Barndorff-Nielsen, O.E. (1995) *Normal inverse Gaussian distributions and the modeling of stock returns.* Research report No. 300, Department of statistics, Aarhus university.

[5] Barndorff-Nielsen, O.E. (1997) *Normal inverse Gaussian distributions and stochastic volatility models.* Scandinavian journal of statistics, Vol. 24, pp. 1-13.

[6] Barndorff-Nielsen, O.E. (2001) *Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics.* Journal of the royal statistical society B, Vol. 63, pp. 167-241.

[7] Benth, F.E., Groth, M. and Kettler, P.C. (2005) *A Quasi-Monte Carlo algorithm for the normal inverse Gaussian distribution and valuation of fi-*

nancial derivatives. University of Oslo, Department of Mathematics, No. 14. (www.math.uio.no/eprint/pure_math/2005/14-05.pdf)

[8] Black, F. and Scholes, M. (1973) *The pricing of options and corporate liabilities.* The journal of political economy, Vol. 81, No. 3, pp. 637-654.

[9] Bollen, B. and Inder, B. (2002) *Estimating daily volatility in financial markets utilizing intraday data.* Journal of empirical finance, Vol. 9, pp. 551-562.

[10] Bollerslev, T. (1986) *Generalized autoregressive conditional heteroscedasticity.* Journal of Econometrics, Vol. 31, pp. 307-327.

[11] Bollerslev, T. (1987) *A conditionally heteroskedastic time series model for speculative prices and rates of return.* Review of economics and statistics, Vol. 69, pp. 542-547.

[12] Brennan, M. (1979) *The pricing of contingent claims in discrete time models.* Journal of finance, Vol. 34, pp. 53-68.

[13] Carr, P., Geman, H., Madan, D.B. and Yor, M. (2002) *The fine structure of asset returns: An empirical investigation.* Journal of business, Vol. 75, pp. 305-332.

[14] Chung, K.L. (1968) *A course in probability theory.* New York: Harcourt, Brace and World Inc.

[15] Clark, P. (1973) *A subordinated stochastic process model with finite variance for speculative prices.* Econometrica, Vol. 41, pp. 135-156.

[16] Cont, R. (2001) *Empirical properties of asset returns: stylized facts and statistical issues.* Quantitative finance, Vol. 1, pp. 223-236.

[17] Cont, R. and Tankov, P. (2003) *Financial modeling with jump processes.* New York: Chapman & Hall/CRC Press.

[18] Copson, E.T. (1935) *An introduction to the theory of functions of a complex variable.* Oxford: Oxford University Press.

[19] Curto, J.D., Pinto, J.C. and Tavares, G.N. (2007) *Modeling stock markets' volatility using GARCH models with Normal, Student's t and stable Paretian distributions.* Springer: Statistical papers, Regular Article (DOI 10.1007/s00362-007-0080-5).

[20] Delbaen, F. and Schachermayer, W. (1994) *A general version of the fundamental theorem of asset pricing.* Mathematische annalen, Vol. 300, pp. 463-520.

[21] Duan, J.C. (1995) *The GARCH option pricing model.* Mathematical finance, Vol 5, No. 1, pp. 13-32.

[22] Duan, J.C., Ritchken, P. and Sun, Z. (2004) *Jump starting GARCH: Pricing and hedging options with jumps in returns and volatilities.* FRB of Cleveland, Working paper no. 06-19. (http://ssrn.com/abstract=479438)

[23] Engle, R.F. (1982) *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.* Econometrica, Vol. 50, No. 4, pp. 987-1008.

[24] Engle, R.F. and Ng, V.K. (1993) *Measuring and testing the impact of news on volatility.* Journal of econometrics, Vol. 48, No. 5, pp. 1749-1778.

[25] Geman, H., Madan, D. and Yor, M. (2001) *Time changes for Lévy processes.* Mathematical Finance, Vol. 11, pp. 79-96.

[26] Griebenow, G. (2006) *GARCH models based on Brownian inverse Gaussian innovation processes.* Doctoral thesis, North-West University.

[27] Grigelionis, B. (1999) *Processes of Meixner type.* Lithuanian mathematics journal, Vol. 39, No. 1, pp. 33-41.

[28] Hansen, P.R. and Lunde, A. (2005) *A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?.* Journal of applied econometrics, Vol. 20, No. 7, pp. 873-889.

[29] Heston, S.L. (1993) *A closed-form solution for options with stochasic volatility woth applications to bond and currency options.* The review of financial studies, Vol. 6, No. 2, pp. 327-343.

[30] Heston, S.L. and Nandi, S. (2000) *A closed-form GARCH option valuation model.* The review of financial studies, Vol. 13, No. 3, pp. 585-625.

[31] Hsieh, K.C. and Ritchken, P. (2005) *An empirical comparison of GARCH option pricing models.* Review of derivatives research, Vol. 8, No. 3, pp. 129-150.

[32] Hull, J. and White, A. (1987) *The pricing of options on assets with stochastic volatilities.* The journal of finance, Vol. 42, No. 2, pp. 281-300.

[33] Kim, S.K., Rachev, S.T. and Chung, D.M. (2006) *The modified tempered stable distribution, GARCH models and option pricing.* Technical report.

[34] Loève, M. (1973) *Paul Lévy, 1886 - 1971.* The annals of probability, Vol. 1, No. 1, pp. 1-18.

[35] Madan, D.B. and Seneta, E. (1990) *The VG model for share market returns.* Journal of business, Vol. 63, No. 4, pp. 511-524.

[36] Marshall, J.P. (2008) *Stochastic volatility modeling of the Ornstein Uhlenbeck type: Pricing and calibration.* Masters Dissertation, University of Johannesburg.

[37] Menn, C. and Rachev, S.T. (2005a) *A GARCH option pricing model with α-stable innovations.* European journal of operational research, Vol. 163, pp. 201-209.

[38] Menn, C. and Rachev, S.T. (2005b) *Smoothly truncated stable distributions, GARCH-models and option pricing.* Technical report. (http://www.statistik.uni-karlsruhe.de/technical_reports/sts-option.pdf).

[39] Merton, R.C. (1973) *Theory of rational option pricing.* The Bell journal of economics and management science, Vol. 4, No. 1, pp. 141-183.

[40] Merton, R.C. (1976) *Option pricing when underlying stock returns are discontinuous.* Journal of financial economics, Vol. 3, pp. 125-144.

[41] Michael, J.R., Schucany, W.R. and Haas, R.W. (1976) *Generating random variates using transformations with multiple roots.* The American statistician, Vol. 30, No. 2, pp. 88-90.

[42] Robbertse, J.L. (2007) *On the modeling of asset returns and the calibration of European option pricing models.* Masters Dissertation, University of Johannesburg.

[43] Ross, S.M. (2001) *Simulation.* 3^{rd} Edition. California: Academic press.

[44] Rubinstein, M. (1976) *The valuation of uncertain income streams and the pricing of options.* Bell journal of economics, Vol. 7, No. 2, pp. 407-425.

[45] Rydberg, T. (1997) *The normal inverse Gaussian Lévy process: simulations and approximation.* Communications in statistics: Stochastic models, Vol. 13, pp. 887-910.

[46] Sato, K-I. (2005) *Lévy processes and infinitely divisible distributions.* Cambridge: Cambridge university press.

[47] Schoutens, W. (2000) *Stochastic processes and orthogonal polynomials.* Lecture notes in statistics, Vol. 146. Springer.

[48] Schoutens, W. (2001) *The Meixner process in finance.* EURANDOM Report 2001-002. EURANDOM, Eindhoven.

[49] Schoutens, W. (2002) *Meixner processes: theory and applications in finance.* EURANDOM Report 2002-004. EURANDOM, Eindhoven.

[50] Schoutens, W. (2003) *Lévy Processes in Finance: Pricing financial derivatives.* New York: John Wiley & sons.

[51] Schoutens, W. and Teugels, J.L. (1998) *Lévy processes, polynomials and martingales.* Communications in statistics: Stochastic models, Vol. 14, pp. 335-349.

[52] Schuster, J. (1968) *On the inverse Gaussian distribution.* Journal of the American statistical association, Vol. 63, No. 324, pp. 1514-1516.