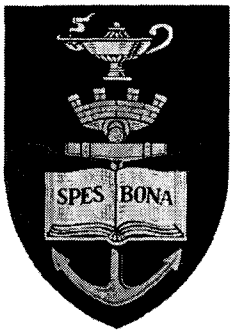

Pricing Path Dependent Options under Variance Gamma Dynamics

Craig Anderson

Submitted to the Department of Mathematics and Applied Mathematics
in partial fulfillment of the requirements for the degree of

Master of Science in Mathematics of Finance
at the
UNIVERSITY OF CAPE TOWN

May 29, 2007

Supervisor: Dr. P. Ouwehand

Abstract

The pricing of derivative securities in finance has been dominated by modelling the underlying security as an Itô diffusion process. A major problem with diffusion processes is that their sample paths are continuous almost surely and operate in a complete market model. Whilst local (and stochastic) volatility models in a diffusion setting do well in alleviating the volatility smile, they ignore one fundamental qualitative aspect of risky asset price dynamics: prices jump. Thus price dynamics driven by general Lévy processes may provide a more realistic framework.

This dissertation explores the variance gamma process, which consists entirely of jumps, as a model for pricing derivatives. General properties are discussed including a method of calibrating the model to the implied risk-neutral distribution. The pricing of an up-and-out call, a lookback put and a double barrier option are discussed assuming Geometric Brownian motion dynamics, followed by variance gamma dynamics. An efficient Monte Carlo method known as the 'truncated difference-of-gammas sampling' is discussed and implemented. A comparison of prices concludes the dissertation.

Plagiarism Declaration

1. This dissertation is my own work. It has not been submitted before for any degree or examination to any other University.

2. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

3. Each significant contribution to, and quotation in, this dissertation from the work of other people has been cited and referenced.

_____

Signature

_____

Date

Acknowledgements

I would like to thank my supervisor, Dr Peter Ouwehand, for his help and commitment.

I would also like to thank my parents, Stuart and Denise Anderson for their unconditional moral support.

Contents

List of Figures

Chapter 1

Introduction

Pricing and hedging of financial derivatives involves choosing a stochastic process for the underlying security. Of these, Itô diffusion processes have received the most interest. Loosely speaking, Itô diffusions are processes that are driven by a drift component and a continuous martingale component, given by a Brownian motion. Itô diffusions have continuous sample paths (because Brownian motion does) which means that the underlying process that you're trying to model, moves at each instant in time. Pure jump processes, on the other hand, move by jumps (possibly an infinite number of jumps in any finite time period) breaking down the continuity of sample paths.

The underlying stochastic variables of interest in this dissertation will be stock prices. However, all pricing and calibration methods discussed in this dissertation apply equally well to indices, futures prices and various other traded instruments. The question of deciding whether the above mentioned variables belong to the class of Itô diffusion processes or pure jump processes will not be considered. Motivation for the use of pure jump processes will be expressed in the light of empirical research and qualitative aspects. Indeed, the modelling of stock prices is an art (although highly scientific in its methodology) and the choice of a particular underlying process should be based on a set of reasonable and justifiable assumptions. We give a brief discussion on Itô diffusion processes and pure jump processes and provide motivation for making use of the latter.

1.1 Itô Diffusion Processes

Definition 1.1.1 *Suppose that $(S_t)_t$ is an N-vector Itô diffusion process. The round brackets around S indicate that S is an adapted process. The dynamics of $(S_t)_t$ are given by the following stochastic differential equation:*

$$dS = \mu(t, S)dt + \sigma(t, S)dW \tag{1.1}$$

where $\mu(t,S)$ is an $N \times 1$ vector of drifts and $\sigma(t,S)$ is an $N \times K$ matrix such that $\sigma\sigma^T$ is an $N \times N$ matrix of volatilities and co-volatilities. W is a K-dimensional Brownian motion. We interpret the stochastic integral in the Itô sense.

□

Note that in the one-dimensional case ($N = 1$), if we set $\mu(t,S) = \mu S$ and $\sigma(t,S) = \sigma S$, where $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, the dynamics of $(S_t)_t$ are given by

$$dS = \mu S dt + \sigma S dW \tag{1.2}$$

which is the familiar geometric Brownian motion.

Itô diffusion processes (of type (1.2)) have dominated the financial literature in the pricing of univariate and especially multivariate[1] options for a number of reasons. Perhaps the most important result in Itô diffusion theory is that of Girsanov's theorem which guarantees (provided that the dimension of the Brownian motion equals the number of underlying assets, i.e. $N = K$) the existence of a unique risk-neutral measure implying that our market model is arbitrage-free and complete. See Karatzas & Shreve (1998) for a proof. In addition, the volatility matrix $\sigma\sigma^T$ is invariant under change of measure which means that the volatility matrix can be estimated from the 'real world' or historical measure. In fact, the maximum likelihood estimate of $\sigma\sigma^T$ is the sample covariance matrix which is observable under the 'real world' measure.

Simulation of multidimensional Brownian motion is an elementary task (e.g. using Cholesky's decomposition of the covariance matrix) and hence prices of multivariate (path-dependent) options can always be computed with relative ease by Monte Carlo methods. See Glasserman (2004) for details.

Many univariate options (e.g. vanilla calls, puts, barrier options) and some multivariate options (e.g. rainbow options) can be computed in closed form; meaning that they can be written in terms of a integral involving a standard (multi)normal density. This integral, in low dimensions, can be efficiently computed. Otherwise, efficient Monte Carlo methods can always be employed since multidimensional Brownian motion is easy to simulate.

Stochastic volatility models also belong to the class of Itô diffusion based models where the volatility itself is driven by a Itô diffusion process.

[1] We define a multivariate option as an option which has a payoff connected to more than one underlying asset

1.2 Pure Jump Processes

Pure jump processes attempt to model underlying variables with a process that consists purely of jumps. The 'jumps' in the process produce discontinuous sample paths. Pure jump processes can broadly be categorised into finite activity and infinite activity processes. The former consists of sample paths that possess a finite number of jumps in any finite time period (e.g. the compound Poisson process) whereas the latter consists of an infinite number of jumps in any finite time period. This dissertation will solely be interested in infinite activity pure jump processes. Finite activity processes do not possess enough 'movement' (unless coupled with a diffusion process) and lack modelling flexibility. There are also processes which consist of a mixture of a diffusion process and a pure jump process. They are usually termed 'Jump-diffusion' processes. These processes consist of a continuous diffusion component (usually given by a Brownian motion) punctuated by jumps at random times.

1.3 Rationale for using Pure Jump Processes

Stock prices, indices and futures prices are essentially determined by supply and demand. The Johannesburg Stock Exchange lists a price at every point in time which is the average of the highest bid and lowest offer price at which a buyer and seller are willing to pay/receive. Once the bid and offer price coincide, then a sale is conducted in which case the highest bid and lowest offer price change. This results in a change in the price of the stock in question. This price change clearly corresponds to a discontinuity or jump in the price process. Price changes clearly move in a discrete fashion in time and would seem unrealistic to model a stock price using a stochastic process with continuous sample paths which assumes that the price changes at every instant in time. Having said this, we will encounter pure jump processes which possess a countably infinite number of jumps. Whilst these processes have discontinuous sample paths, the extent to which they add to the realism of stock price movements is debatable.

Empirical research indicates that return distributions of financial variables (e.g. stock prices) exhibit asymmetry and fat tails. This applies to both historical return distributions and implied risk-neutral distributions. Existence of the volatility skew (and smile in exchange rate options) in the Black-Scholes model provides evidence that geometric Brownian motion is inadequate in capturing the return distributions of financial variables. One can alleviate this problem to a certain degree by introducing stochastic volatility. Stochastic volatility models do well in flattening the skew (more so with longer term options than short-dated options) but have several problems. Firstly, stochastic volatility models do not explain

the phenomenon of the skew. They are merely in place to provide a better fit of observed option prices to theoretical prices. The skew is present in option prices because of investor sentiment that financial variables are far more volatile (exhibit fat-tails) than what the normal distribution suggests. Whilst stochastic volatility models do well in this regard, they come at the cost of unreasonably high and nonstationary values of the diffusion coefficient driving the volatility process, which are required to capture the (implied) fat-tails of return distributions. This poses estimation and calibration problems.

Derivatives that trade on an exchange or OTC cannot be perfectly replicated. If this were the case, then one could synthetically create any such payoff using just the underlying variables. Why does a derivatives market exist? The answer is because derivatives cannot be perfectly replicated. Markets are not complete. Itô diffusion processes (provided that the dimension of the Brownian motion equals the number of underlying assets) give rise to models that are complete (as a consequence of the martingale representation theorem) which is clearly unrealistic. Processes consisting of only jumps give rise to models that are generally incomplete. This is far more realistic and lends itself to the study of hedging strategies (dynamic and static).

1.4 The way ahead....

It should be clear that we are interested in an underlying process which does not suffer from the drawbacks of Itô diffusion processes. We seek processes with jumps. Lévy processes are processes which incorporate jumps. The purpose of this dissertation is to give a comparison of prices of options obtained in a Black-Scholes model and that of a model involving a pure jump process. The breadth and scope of Lévy processes is vast and we have chosen one popular type known as the variance gamma process (Madan & Seneta, 1990). Studying the properties of the variance gamma process as a model for option prices does require some general knowledge of Lévy processes. This is discussed in Chapter 2. Chapter 3 moves on to the specifics of the variance gamma process discussing all aspects (quantitative and qualitative) required in pricing options. Barrier options have been chosen to make the various comparisons. These options seem to pose the greatest difficulty in a pure jump setting. An overview of the literature involving the pricing of barrier options in a Black-Scholes world is given in the beginning of Chapter 4. In particular, we focus on three types of path-dependent options namely

- Up-and-Out Call option

- Lookback Put
- Knock-Out Double Barrier Call option

The chapter also discusses a very efficient method of calculating prices of double barrier options in a Black-Scholes world. We move on to pricing the same three options, but this time using a Lévy process, in particular, the variance gamma process. We briefly discuss the methods available in pricing path dependent options in a Lévy world and focus on Monte Carlo methods. We start off by discussing available methods of simulation of the variance gamma process including a method known as the 'truncated difference-of-gammas bridge sampling' (TDGBS) algorithm developed by Avramidis & L'Ecuyer (2004), which is very efficient. The TDGBS algorithm applied to barrier options virtually eliminates all bias induced by discretization of sample paths. The final chapter concludes with a brief comparison of option prices obtained using the Black-Scholes and the variance gamma model. Conclusions and recommendations wrap up the dissertation.

Chapter 2

Lévy Processes

This chapter explores the definitions and general properties of Lévy processes. The majority of the theorems and definitions given are fairly general and only special cases of these results will be used in this dissertation. This dissertation will only provide proofs for theorems which are of interest to us whereas references to proofs of more general results will be stated. Good references include Sato (1999), Schoutens (2003) and Bertoin (1998).

2.1 Definitions and properties of Lévy Processes

Definition 2.1.1 *Let $(\Omega, \Im, \mathbb{P})$ be a probability space equipped with a filtration $\Im_t$ which we assume satisfies the usual conditions. A d-dimensional Cádlág stochastic process $(X_t)_t$ with $X_0 = 0$ is called a Lévy process if it possesses the following properties:*

- ***Independent increments**: if $t_0, ..., t_n$ is a finite increasing sequence of times then the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.*

- ***Stationarity**: the distribution of $X_{t+h} - X_t \stackrel{d}{=} X_h$ for every $h > 0$.*

- ***Stochastic continuity**: $\forall \epsilon > 0, [\lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0]$*

□

The purpose of this dissertation is to model underlying financial variables (e.g. stock prices or indices) using a particular type of Lévy process and to price various contingent claims thereon. Let us spend some time discussing why a Lévy process would be appropriate for our underlying financial variables.

Firstly, independent increments of the process would imply that given a σ-algebra $\Im_t$ of available information at time t, the change in the underlying variable $X_{t+h} - X_t$ is independent of what is known in our filtration. Here we assume the filtration is generated by $(X_t)_t$. Suppose we discard the independent increments property; then changes in the underlying variable would be dependent on past information (non-Markovian) implying that the weak-form of market efficiency breaks down. Hence prices and index levels do not fully reflect all information contained in them and there's room for statisticians to make profits.
Secondly, stationarity of increments would imply that changes in our underlying variable $X_{t+h} - X_t$ (of length h) have the same distribution for all times t. Suppose we discard the stationarity property; then changes in the distribution (of time length h) would depend on time. This essentially means that we would have some knowledge about how the distribution of our underlying variables is going to change at some time in the future. This is an unrealistic assumption.
Lastly, the stochastic continuity of X_t by no means implies that the sample paths are continuous. Cont & Tankov (2002) provide a good discussion. It is merely a condition which excludes processes which have predictable jumps (which are not useful for our applications in finance).

Intuitively, a Lévy process is the continuous analogue of a random walk. Consider sampling a Lévy process $(X_t)_t$ at a set of evenly spaced discrete times given by $t_0, \ldots, t_n$. The process $X_{t_n} = \sum_{i=1}^{n} Y_i$ where $Y_i = X_{t_i} - X_{t_{i-1}}$ is a random walk. Indeed, X_{t_n} it is a sum of independent, identically distributed (i.i.d.) random variables. This property of dividing X_t into n i.i.d. parts is not true for all measures on random variables which leads us to our next definition.

Definition 2.1.2 *A distribution function F on $\mathbb{R}^d$ is said to be **infinitely divisible** if for any integer $n \geq 2$, there exists an i.i.d. sequence of random variables $Y_1, Y_2, ..., Y_n$ such that $\sum_{i=1}^{n} Y_i$ has distribution F.*

It follows that any Lévy process is infinitely divisible. Conversely, it can be shown that given an infinitely divisible distribution F then there exists a Lévy Process $(X_t)_t$ such that X_1 has distribution F. Sato (1999) provides a detailed proof.

□

Proposition 2.1.1 *Let $(X_t)_t$ be a Lévy process on $\mathbb{R}^d$. The characteristic function of X_t has the form*

$$\mathbb{E}[e^{izX_t}] = e^{t\varphi(z)}$$

where $z \in \mathbb{R}^d$ and $\varphi : \mathbb{R}^d \to \mathbb{C}$.

Proof: Define $\Psi(t) = \mathbb{E}[e^{izX_t}]$. Now X_{t+s} can be written as $X_s + X_{t+s} - X_s$. Since $X_{t+s} - X_s$ is independent of X_s

$$\begin{aligned}\Psi(t+s) = \mathbb{E}[e^{izX_{t+s}}] &= \mathbb{E}[e^{izX_s}]\mathbb{E}[e^{iz(X_{t+s}-X_s)}] \\ &= \mathbb{E}[e^{izX_s}]\mathbb{E}[e^{izX_t}] \text{ (stationarity property)} \\ &= \Psi(s)\Psi(t).\end{aligned}$$

i.e. $\Psi(t)$ is a multiplicative function of t. Note that the stochastic continuity of $(X_t)_t$ implies that $X_t \to X_s$ in distribution as $t \to s$. Hence $\Psi(t) \to \Psi(s)$ as $t \to s$ (Lévy Continuity Theorem). Therefore, Ψ is a continuous and multiplicative function of t. It follows that there exists a function φ such that $\mathbb{E}[e^{izX_t}] = e^{t\varphi(z)}$. A proof of this result can be found in Feller (1968).

□

There are two very important types of Lévy processes. Firstly, it's easy to verify that Brownian motion satisfies the conditions of a Lévy process. The second important type of Lévy process is known as the compound Poisson process to be defined below. We first need to define a counting process.

Definition 2.1.3 *Let $\{T_n, n \geq 1\}$ represent an increasing sequence of random times with the property that $\mathbb{P}(T_n \to \infty) = 1$. Define X_t by*

$$X_t = \sum_{n \geq 1} 1_{t \geq T_n}$$

is known as a counting process.

Definition 2.1.4 *Let N_t be a Poisson process with intensity $\lambda > 0$[1]. Recall that N_t is a counting process with independent and stationary increments with probability mass function given by*

$$\mathbb{P}(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

Let Y_i be an i.i.d. sequence of random variables with characteristic function F and distribution μ. A compound Poisson process, X_t, is defined by

$$X_t = \sum_{i=1}^{N_t} Y_i$$

□

[1] λ can be interpreted as the average number of jumps or arrivals per unit time

It's easy to verify that X_t is a Lévy process. See Cont & Tankov (2002) for a proof. The characteristic function of X_t can be computed as follows:

$$\begin{aligned}\mathbb{E}[e^{izX_t}] &= \mathbb{E}[\mathbb{E}[e^{izX_t}|N_t]] = \mathbb{E}[F^{N_t}] \\ &= \sum_{i=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^i (F)^i}{i!} \\ &= e^{t\lambda(F-1)} \\ &= e^{t\lambda \int_{\mathbb{R}}(e^{izx}-1)d\mu(x)} \end{aligned} \tag{2.1}$$

It should be clear that a compound Poisson process has discontinuous sample paths. A typical sample path is a step function.
Define a new measure ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by

$$\nu(A) = \lambda \int_A d\mu(x) \tag{2.2}$$

where $A \in \mathcal{B}(\mathbb{R})$. $\nu(A)$ can be interpreted as the expected number of jumps in A per unit time. We will call ν as the Lévy measure of the compound Poisson process. The Lévy measure is not specific to compound Poisson processes and can be generalized to all Lévy processes. This leads us to our next definition.

Definition 2.1.5 *Let $(X_t)_t$ be a Lévy process on $\mathbb{R}^d$. The Lévy measure of $(X_t)_t$ is defined by*

$$\nu(A) = \mathbb{E}[\#\{t \in [0,1]\} : \Delta X_t \neq 0, \Delta X_t \in A] \text{ where } \Delta X_t = X_t - \lim_{s \uparrow t} X_s$$

$\nu(A)$ is the expected number of jumps in A per unit time. The Lévy measure $\nu(A)$ as defined above is a mapping from $\mathcal{B}(\mathbb{R}^d) \to \mathbb{R}^+$. The Lévy processes considered in this dissertation will have Lévy measures which are absolutely continuous with respect to Lebesgue measure, i.e. they have densities. This is certainly not always the case. For our purposes, it suffices to interpret the Lévy measure as follows

$$\nu(A) = \int_A \nu^d(x_1, x_2, \ldots, x_d) dx_1 dx_2 \ldots dx_d$$

where $\nu^d(x_1, x_2, \ldots, x_d)$ is known as the Lévy density of the Lévy measure.

□

The compound Poisson process on $\mathbb{R}$ has a particularly simple form of Lévy measure given by equation (2.2). See Cont & Tankov (2002) for details. Now suppose that μ has a density f. Defining a new density $\nu^d(x) = \lambda f(x)$, we have $\nu(A) = \int_A \nu^d(x)dx$ where $\nu^d(x)$ is the density of the Lévy measure ν. Note that

$\nu^d(x)$ is not a probability density since it integrates to λ and not 1.

Example. Suppose $(X_t)_t$ is a compound Poisson process on $\mathbb{R}$ with intensity λ and jump size density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The jumps of $(X_t)_t$ arrive according to a Poisson process with a average rate of λ per unit time and the size of the jumps are given by a normally distributed random variable with mean μ and variance σ^2. Let A denote the set of jumps which have a size of less than 1, then $(X_t)_t$ has Lévy measure

$$\nu(A) = \lambda \int_A f(x)dx = \frac{\lambda}{\sqrt{2\pi}\sigma} \int_{-\infty}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \lambda N\left(\frac{1-\mu}{\sigma}\right)$$

where $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}p^2} dp$. Note that when $A = \mathbb{R}$ (jumps of all sizes) then $\nu(A) = \lambda < \infty$. The expected number of jumps per unit time on any bounded interval is finite, i.e. the compound Poisson process has a finite number of jumps in any time period. The compound Poisson process is an example of a pure jump process of finite activity to be defined below.

Definition 2.1.6 *Let $(X_t)_t$ be a Lévy process on $\mathbb{R}^d$ and let ν be its Lévy measure. $(X_t)_t$ is said to be of finite activity if $\nu(\mathbb{R}^d) < \infty$. $(X_t)_t$ is said to be of infinite activity if $\nu(\mathbb{R}^d) = \infty$.*

The compound Poisson process is of finite activity. This dissertation is mainly concerned with infinite activity processes since finite activity processes, unless coupled with arithmetic Brownian motion to form a jump-diffusion process, do not adequately explain market movements of stock returns. Highly liquid stocks or indices have hundreds of trades per day which would require a very large λ. Infinite activity models would provide a much more accurate approximation of how stock prices and index levels evolve.

2.2 Jump Measures and Poisson Random Measures

The construction of Lévy processes requires knowledge of random measures and in particular, Poisson random measures. This section starts off with some definitions and ends off by showing that a compound Poisson process can be written in terms of a Poisson random measure. The Lévy-Itô decomposition, which is probably the most fundamental result in Lévy process theory, makes use of this representation of a compound Poisson process.

Definition 2.2.1 *Let $(X_t)_t$ be a stochastic process on $(\Omega, \Im, \mathbb{P})$ with values in $\mathbb{R}^d$. Let $H = (0,\infty) \times \mathbb{R}^d \backslash \{0\}$. A random measure is a mapping $J_X : \mathcal{B}(H) \times \Omega \to \mathbb{N}$ defined by $J_X(A,\omega) = \#\{t : (t, \Delta X_t(\omega)) \in A\}$ for each Borel set $A \subseteq H$ and $\omega \in \Omega$. If $(X_t)_t$ is a Lévy process then we will call the random variable $J_X(A)$ the jump measure of $(X_t)_t$.*

$J_X(A)$ counts the number of jumps ΔX_t such that $(t, \Delta X_t) \in A$. It is clearly a non-negative integer-valued random variable since the number of jumps of $(X_t)_t$ depends on each $\omega \in \Omega$. The jump measure of $(X_t)_t$ in any bounded time interval has a mass function which leads us to define a particularly important class of jump measures.

□

Definition 2.2.2 *Let N_X be the jump measure of a Lévy Process $(X_t)_t$. N_X is a Poisson random measure with intensity measure μ if it satisfies the following criteria:*

- $\mathbb{P}(N_X(A) = k) = \frac{e^{-\mu(A)}(\mu(A))^k}{k!}$ *where $A \subseteq H$ i.e. $N_X(A)$ is a Poisson random variable with intensity $\mu(A)$.*

- *if $A \cap B = \emptyset$, the random variables $N_X(A)$ and $N_X(B)$ are independent.*

We'll reserve the notation J_X for a general jump measure of $(X_t)_t$ and N_X for the special case where the jump measure is also a Poisson random measure.

□

Proposition 2.2.1 *Let $X_t = \sum_{i=1}^{N_t} Y_i$ be a compound Poisson process with intensity λ and jump size distribution μ_p. Its jump measure J_X is a Poisson random measure with intensity $\mu = \lambda^{Leb} \otimes \nu$ where $\nu = \lambda\mu_p$ is the Levy measure of the Poisson process and λ^{Leb} is the Lesbesgue measure.*

In different notation, the proposition states that if $(X_t)_t$ is a compound Poisson process then $J_X \equiv N_X$. We refer the reader to Cont & Tankov (2002) for a proof. In fact, the above proposition can be generalized to any Lévy process.

Proposition 2.2.2 *Let $(X_t)_t$ be an arbitrary Lévy process. Its jump measure J_X is a Poisson random measure with intensity $\lambda \otimes \nu$ where λ is the Lesbesgue measure and ν is the Lévy measure of $(X_t)_t$.*

□

A proof can be found in Sato (1999). This proposition allows us to characterize compound Poisson processes in terms of its jump measure. Let $(X_t)_t$ be a compound Poisson process on $\mathbb{R}^+$. We can then write

$$X_t = \int_{[0,t]\times A} x N_X(d(s,x)) = \sum_{s\in[0,t]} \Delta X_s 1_{\Delta X_s \in A}$$

as a well defined compound Poisson process. Since the compound Poisson process has a finite number of jumps in any time interval, the stochastic integral has no convergence problems. Note that one could generalize the above construction to any pure jump Lévy process provided that the process is of finite activity[2], i.e. let $(X_t)_t$ be a pure jump Lévy process on $\mathbb{R}^d$ with jump measure J_X. Then X_t can be written as

$$X_t = \int_{(0,t]\times\mathbb{R}^d\setminus\{0\}} x\ J_X(d(s,x)) \tag{2.3}$$

Note that X_t is d-dimensional which requires us to interpret the above integral component-wise. Now consider a jump-diffusion process (a Lévy process) given by

$$X_t = \gamma t + \sigma W_t + Z_t \tag{2.4}$$

where γ is the deterministic drift, W_t is a d-dimensional Brownian motion and Z_t is a pure jump process (independent of W_t) with jump measure J_Z. Note that $J_X \equiv J_Z$ since $\gamma t + \sigma W_t$ is continuous a.s. (W_t has no jumps). We can equivalently write

$$X_t = \gamma t + \sigma W_t + \int_{(0,t]\times\mathbb{R}^d\setminus\{0\}} x\ J_Z(d(s,x)) \tag{2.5}$$

using the same argument as above. Every pure jump Lévy process can be written in the form given by equation (2.5) provided that J_Z has a finite number of jumps in any time interval or equivalently, the Lévy measure is of finite activity. Not every Lévy process can be written in the form given by equation (2.5) since $\nu(A)$ could be infinite (as in infinite activity processes), whereas the number of jumps of size ≥ 1 can be shown to be finite on each bounded set, a.s. The stochastic integral appearing in equation (2.5) will not necessarily converge. This is because Z_t may have an infinite number of small jumps and, when summed up, could be infinite. This will impose certain conditions on the Lévy measure which alters the decomposition given in (2.5).

[2]Else the stochastic integral may have convergence issues

2.3 Lévy-Itô decomposition

The Lévy-Itô decomposition allows one to decompose every Lévy process into four independent processes. This decomposition generalizes equation (2.5) to any Lévy Process (finite and infinite activity). Here is the proposition in the case where the Lévy measure has a density:

Proposition 2.3.1 *Let $(X_t)_t$ be a Lévy process on $\mathbb{R}^d$. Then we may write*

$$X_t = \gamma t + B_t + X_t^l + \lim_{\epsilon \to 0} \tilde{X}_t^\epsilon \tag{2.6}$$

where the four processes on the right are independent and

- *ν, the Lévy measure of X_t, is a Radon measure (i.e. finite on compact sets) with $\int_{\mathbb{R}^d} |x|^2 \wedge 1 \nu(x)dx < \infty$ and with $\nu(\{0\}) = 0$.*
- *γ is the deterministic drift*
- *B_t is a $d-$dimensional Brownian motion with covariance matrix Σ.*
- *$X_t^l = \int_{(0,t]\times\{|x|\geq 1\}} x N_X(d(s,x))$*
- *$\tilde{X}_t^\epsilon = \int_{(0,t]\times\{\epsilon \leq |x| < 1\}} x[N_X(d(s,x)) - \nu(x)ds]$*

The processes $X_t^l, \tilde{X}_t^\epsilon$ are compound Poisson processes.

□

The four terms are independent and convergence of $\tilde{X}_t^\epsilon$ is almost sure. The proof of this proposition is difficult and readers may refer to Sato (1999). The Lévy-Itô decomposition tells us that every Lévy Process can be factored into four independent components. The decomposition is useful since it allows one to compute the characteristic function of any Lévy process with ease. On the practical side, the decomposition allows one to approximate the simulation of any Lévy process by truncating jumps of size smaller than ϵ. This approximation is useful when dealing with intractable Lévy processes. Note that the three parameters (Σ, ν, γ) uniquely identify the above decomposition which we will now refer to as the *characteristic triplet* of the Lévy process $(X_t)_t$.

The first two components $\gamma t + B_t$ (arithmetic Brownian motion with drift) are the only components which are continuous (almost surely). The d-dimensional Brownian motion does allow for arbitrary covariance matrix Σ. The third component, X_t^l, is a compound Poisson process as discussed in the previous section.

There are no convergence issues since X_t^l has truncated all jumps of size[3] less than 1. $\nu(\{|x| \geq 1\})$ can be shown to be finite, implying that the number of jumps of size greater than 1 must be finite. As stated before, $\nu(\{|x| \leq 1\})$ may be infinite which means in this case that there is an infinite number of small jumps per unit time and when summed up, could be infinite. The compound Poisson process appearing in the fourth term subtracts the average number of jumps at each point in time. Once each jump has been centered, the stochastic integral $\tilde{X}_t^\epsilon$, converges almost surely as $\epsilon \to 0$. Again, we refer the reader to Sato (1999) for details.

□

The Lévy-Itô decomposition can be simplified under certain conditions. One condition which will be of importance to us is the case where the Lévy process is of finite variation.

Proposition 2.3.2 *A Lévy process with characteristic triplet (Σ, ν, γ) is of finite variation if and only if $\Sigma = 0$ and $\int_{|x|\leq 1} |x|\nu(x)dx < \infty$.*

The proof of this can be found in Cont & Tankov (2002). An important corollary to Proposition 2.3.1 is given below.

Corollary 2.3.1 *Let $(X_t)_t$ be a Lévy process with triplet (Σ, ν, γ) on $\mathbb{R}^d$ which is of finite variation. X_t can then be uniquely expressed as follows*

$$X_t = bt + \int_{(0,t]\times\mathbb{R}^d\setminus\{0\}} x\, N_X(d(s,x)) \tag{2.7}$$

where $b = \gamma - \int_{|x|\leq 1} x\nu(x)dx$. This decomposition should make intuitive sense since there is no Brownian component and there is no need to truncate jumps of size smaller than ϵ since $(X_t)_t$ is of finite variation. This allows us to directly set $\epsilon = 0$ in equation (2.6) and hence obtain equation (2.7).

2.4 Lévy-Khinchin representation

The Lévy-Khinchin representation allows us to construct the characteristic function of any Lévy process. Once we have the result on the Lévy-Itô decomposition, it is easy to derive the characteristic function of any general Lévy process.

[3]Using the usual Euclidean norm, i.e. if $\mathbf{x}$ is a d-dimensional vector then $|\mathbf{x}| = \sqrt{\sum_{i=1}^d x_i^2}$

Proposition 2.4.1 *Let $(X_t)_t$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (Σ, ν, γ). The characteristic function is given by*

$$\mathbb{E}[e^{i<z,X_t>}] = e^{t(-\frac{1}{2}z^{tr}\Sigma z + i<\gamma,z> + \int_{\mathbb{R}^d}(e^{i<z,x>} - 1 - i<z,x>1_{|x|\leq 1})\nu(x)dx)} \quad (2.8)$$

Proof: The Lévy-Itô decomposition tells us that any Lévy process X_t can be written as the almost sure limit of $\gamma t + B_t + X_t^l + \tilde{X}_t^\epsilon$ as $\epsilon \to 0$, where B_t has covariance matrix Σ. Since almost sure convergence implies convergence in characteristic functions, by independence we have

$$\begin{aligned}
\mathbb{E}[e^{i<z,X_t>}] &= \mathbb{E}[e^{i<z,(\gamma t + B_t)>}]\mathbb{E}[e^{i<z,X_t^l>}]\mathbb{E}[e^{i<z,\tilde{X}_t^\epsilon>}] \\
&= [e^{-\frac{1}{2}tz^{tr}\Sigma z + i<z,\gamma>t}][e^{t\int_{|x|\geq 1}(e^{i<z,x>}-1)v(x)dx}][e^{t\int_{\epsilon\leq|x|<1}(e^{i<z,x>}-1-i<z,x>)v(x)dx}] \\
&= e^{t(-\frac{1}{2}z^{tr}\Sigma z + i<z,\gamma> + \int_{|x|\geq\epsilon}(e^{i<z,x>}-1-i<z,x>1_{|x|\leq 1})\nu(x)dx}
\end{aligned}$$

□

The above term converges to equation (2.8) as $\epsilon \to 0$. The cut-off of jumps in $\tilde{X}_t^\epsilon$ of size smaller than one is arbitrary and one may use any $\epsilon > 0$[4]. The representation given in equation (2.8) assumes that the Lévy measure ν has a density and is in fact not the most general form[5]. It should also be stressed that the γ term is not necessarily the drift of the Lévy process $(X_t)_t$. This constant allows one to re-write the characteristic function of any Lévy process in the form given by equation (2.8) and hence easily read off from the characteristic triplet. For example, consider a compound Poisson process $(X_t)_t$ on $\mathbb{R}$ given by $X_t = \sum_{i=1}^{N_t} Y_i$ where Y_i is an i.i.d. sequence of random variables with common density $f(x)$. It is possible to re-write its characteristic function given by equation (2.1) in the form given by equation (2.8).

$$\begin{aligned}
\mathbb{E}[e^{izX_t}] &= e^{t\lambda\int_{\mathbb{R}}(e^{izx}-1)f(x)dx} \\
&= e^{t(\int_{\mathbb{R}}(e^{izx}-1-izx1_{|x|\leq 1})\lambda f(x)dx + \int_{|x|\leq 1} izx\lambda f(x)dx)} \\
&= e^{t(iz\int_{|x|\leq 1} x\lambda f(x)dx + \int_{\mathbb{R}}(e^{izx}-1-izx1_{|x|\leq 1})\lambda f(x)dx)} \\
&= e^{t(iz\gamma + \int_{\mathbb{R}}(e^{izx}-1-izx1_{|x|\leq 1})\nu(x)dx)} \quad (2.9)
\end{aligned}$$

The characteristic triplet can easily be read off as $(0, \nu, \gamma)$ where $\nu(x) = \lambda f(x)$ and $\gamma = \int_{|x|\leq 1} x\lambda f(x)dx$ which is clearly not the drift since the compound Poisson process has no continuous part. The Lévy-Khinchin representation given by equation (2.8) allows us to work with triplets rather than the complicated characteristic function since the triplet uniquely identifies the characteristic function. The Lévy-Khinchin representation has a number of simplifications. One important simplification is given below.

[4]The number 1 seems standard in the literature

[5]This general form is not necessary for our purposes

Proposition 2.4.2 *Let $(X_t)_t$ be a Lévy process with triplet $(0, \nu, \gamma)$ on $\mathbb{R}^d$ which is of finite variation. Its characteristic function is given by*

$$\mathbb{E}[e^{i<z,X_t>}] = e^{t(i<z,b>+\int_{\mathbb{R}^d}(e^{i<z,x>}-1)\nu(x)dx)} \tag{2.10}$$

where $b = \gamma - \int_{|x|\leq 1} x\nu(x)dx$. This representation allows one to interpret b as the drift of $(X_t)_t$. This can easily be proved by making use of the decomposition given by equation (2.7). The Lévy triplet of $(X_t)_t$ is given by $(0, \nu, \gamma)$ and not $(0, \nu, b)$. The next section deals with a particular type of Lévy process which will be extremely relevant when introducing the variance gamma process.

2.5 Subordinators

Definition 2.5.1 *Let $(S_t)_t$ be a Lévy process on $\mathbb{R}$ defined on a probability space $(\Omega, \Im, \mathbb{P})$ with characteristic triplet (Σ, ϕ, γ). $(S_t)_t$ is said to be a subordinator if $\mathbb{P}(S_t \geq S_s) = 1$ for every pair $s < t$.*

□

A subordinator is an (almost surely) increasing Lévy process. Subordinators have a number of properties which are deduced from the definition. Note that the following statements are equivalent.

- The Lévy-Itô decomposition of $(S_t)_t$ cannot consist of a Brownian motion since $(S_t)_t$ is increasing and therefore of finite variation. This implies that $\Sigma = 0$.

- $\phi(A) = 0$ for $A = \{(-\infty, a) : a \leq 0\}$. There cannot be any negative jumps since $(S_t)_t$ is increasing.

- $S_t \geq 0$ for every t since $S_0 = 0$ and $(S_t)_t$ is increasing.

- S_t can be decomposed as $bt + \int_{(0,t]\times\mathbb{R}^d\setminus\{0\}} x\ N_S(d(s,x))$ where $b = \gamma - \int_{|x|\leq 1} x\phi(x)dx$ and N_S is the jump measure of $(S_t)_t$. This follows from Corollary 2.3.1 since $(S_t)_t$ is of finite variation.

The Lévy-Khinchin representation reduces to

$$\mathbb{E}[e^{izS_t}] = e^{t(izb+\int_0^\infty(e^{izx}-1)\phi(x)dx)} \tag{2.11}$$

Since S_t is a positive random variable, it is more customary to express the distribution of S_t in terms of its moment generating function (m.g.f.). The m.g.f. (for $u \leq 0$) can be obtained by setting $z = -iu$ in equation (2.11) and is given by[6]

$$\mathbb{E}[e^{uS_t}] = e^{t(ub + \int_0^\infty (e^{ux}-1)\phi(x)dx)} \tag{2.12}$$

The moment generating function of any subordinator can be written as $\mathbb{E}[e^{uS_t}] = e^{tl(u)}$. We will call $l(u)$ the Laplace exponent of S_t.

□

2.6 Brownian Subordination

There are various ways of specifying a Lévy process and one such way is via Brownian subordination. Brownian subordination involves specifying Brownian motion with drift and time-changing the process by a subordinator, i.e. if $X(t,\omega) = \vartheta t + \sigma B(t,\omega)$ is a Brownian motion with drift with ϑ,σ constant and $S(t,\omega)$ is a subordinator then $X(S(t,\omega),\omega)$ is said to be subordinated by $S(t,\omega)$ with the resulting process given by $X(S(t,\omega),\omega) = \vartheta S(t,\omega) + \sigma B(S(t,\omega),\omega)$. The next proposition is fundamental in constructing Brownian subordinated processes.

Proposition 2.6.1 *Let $X(t,\omega) = \vartheta t + \sigma B(t,\omega)$ be a Brownian motion with drift (a Lévy Process on $\mathbb{R}^d$) with characteristic triplet[7] $(\Sigma, 0, \vartheta)$. Let $S(t,\omega)$ be a subordinator independent of $X(t,\omega)$ on $\mathbb{R}$ with Laplace exponent $l(u)$ and characteristic triplet $(0,\phi,b)$. The process Y_t defined pathwise is given by $Y(t,\omega) = X(S(t,\omega),\omega) = \vartheta S(t,\omega) + \sigma B(S(t,\omega),\omega)$ is a Lévy Process with characteristic triplet $(\Sigma^S, \nu^S, \vartheta^S)$ defined below. The characteristic function of Y_t is given by*

$$\mathbb{E}[e^{izY_t}] = e^{tl(iz\vartheta - \frac{z^2\sigma^2}{2})} \tag{2.13}$$

where $l(u)$ is the Laplace exponent of $S(t,\omega)$. The characteristic triplet of Y_t is given by

- $\Sigma^S = b\Sigma$
- $\nu^S(x) = \int_0^\infty f(x,s)\phi(s)ds$

[6]Recall that if $\Phi_X(z)$ is the characteristic function of a random variable X then its m.g.f. is given by $\Phi_X(-iu)$ where $i \in \mathbb{C}$ and $u \in \mathbb{R}$ provided the m.g.f. exists

[7]Note that $\sigma\sigma^{tr} = \Sigma$

- $\vartheta^S = b\vartheta + \int_0^\infty \int_{|x|\leq 1} xf(x,s)\phi(s)dxds$

where $f(x,s)$ has a multivariate normal density with mean vector ϑs and covariance matrix $s\Sigma$.

□

We give a proof of the characteristic function and refer the reader to Cont & Tankov (2002) for the completion of the proof.

$$\begin{aligned}
\mathbb{E}[e^{izY_t}] &= \mathbb{E}[e^{iz(\vartheta S(t,\omega)+\sigma B(S(t,\omega),\omega))}] \\
&= \mathbb{E}[\mathbb{E}[e^{iz(\vartheta S(t,\omega)+\sigma B(S(t,\omega),\omega))}|S(t,\omega)]] \\
&= \mathbb{E}[e^{iz\vartheta S(t,\omega)}\mathbb{E}[e^{iz\sigma B(S(t,\omega),\omega))}|S(t,\omega)]] \\
&= \mathbb{E}[e^{iz\vartheta S(t,\omega)}e^{-\frac{z^2\sigma^2 S(t,\omega)}{2}}] \\
&= \mathbb{E}[e^{(iz\vartheta-\frac{z^2\sigma^2}{2})S(t,\omega)}] \\
&= e^{tl(iz\vartheta-\frac{z^2\sigma^2}{2})} \qquad (2.14)
\end{aligned}$$

The process $Y_t = X(S(t,\omega),\omega)$ is no longer a Brownian motion with drift. It is a pure jump process provided the subordinator has no drift. One can interpret the subordinator $S(t,\omega)$ as a stochastic clock. For each $\omega \in \Omega$ determines whether the clock runs faster or slower than calendar time t. Note that conditional on a sample path of $S(t,\omega)$, the process $X(S(t,\omega),\omega)$ is a Brownian motion with drift, but now $S(t,\omega)$ is the time index and not calendar time t.

The interpretation is highly intuitive when applied to modelling underlying financial instruments such as stock prices or indices. Information comes in dribs and drabs and most certainly not in a continuous fashion like Brownian motion suggests. As discussed in Chapter 1, stock price movements are driven by new information in the market at discrete points in time. The subordinator models this information flow. Geman *et al.* (1998) argue that asset prices are Brownian motions; but only in business time. The business time is modelled by the subordinator.

Brownian subordinated models have become very popular in pricing financial derivatives because of the intuitive concept of the stochastic clock (given by the subordinator) and their analytical tractability. In order to specify a Brownian subordinated model, one needs to decide on an appropriate subordinator which must have some desirable properties. Properties would include

- the ability to simulate the subordinator.
- provide a realistic view of the randomness involved in information arrival.

- the expected time of information arrival at calendar time t should be equal to t, i.e. $\mathbb{E}[S(t,\omega)] = t$. On average, we desire our stochastic clock to coincide with calendar time.

- contain sufficient parameters to maintain control of skewness and excess kurtosis of the underlying security.

The univariate variance gamma model (Madan & Seneta, 1990) is one particular choice of a Brownian subordinated model which uses a gamma process as its subordinator. Another popular choice is the normal inverse Gaussian model which uses the normal inverse Gaussian process as its subordinator. This dissertation focuses on the variance gamma model.

Chapter 3

Univariate Variance Gamma Model

The variance gamma model was introduced by Madan & Seneta (1990). Their objective was to provide an alternate model to using Brownian motion as the martingale component. This model needed to be both practical and empirically relevant. They argue that the long tailedness of the variance gamma distribution over the normal distribution provides a good empirical fit to stock returns. Madan & Milne (1991) argue that relative to the Black-Scholes model, the variance gamma option prices are higher. This is particularly so when the options are out-of-the-money with long maturities. Carr *et al.* (2002) extended the variance gamma process to include a diffusion component. They also allowed the process to be either of finite or infinite activity as well as a Lévy measure which may have finite or infinite variation. This is known as the CGMY model. Their models were calibrated to both historical time-series and option price data. They conclude that whilst individual stocks may contain a diffusion component, indices do not. They also report a significant increase in skewness and kurtosis in the risk-neutral process when compared to the historical process. They conclude that the risk-neutral process is mainly of infinite activity but of finite variation.

□

3.1 The Gamma and Variance Gamma Process

We start by giving an account of the gamma process followed by the variance gamma process.

Definition 3.1.1 *The gamma process denoted by $G_t(\alpha, \beta)$ is a strictly increasing (driftless) process with independent and stationary increments hence satisfying*

the conditions for a subordinator[1]*. The stationarity of increments implies that* $G_{t+h}(\alpha,\beta) - G_t(\alpha,\beta) \stackrel{d}{=} G_h(\alpha,\beta)$ *for* $h > 0$*.* $G_h(\alpha,\beta)$ *has a gamma density given by*

$$f_{G_h(\alpha,\beta)}(x) = \frac{\beta^{-\alpha h} x^{\alpha h - 1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha h)} 1_{x \geq 0} \tag{3.1}$$

□

It is easy to show $\mathbb{E}[G_h(\alpha,\beta)] = \alpha\beta h$ and $Var[G_h(\alpha,\beta)] = \alpha\beta^2 h$. The Laplace exponent of $G_t(\alpha,\beta)$ is given by

$$l(u) = -\alpha \ln(1 - u\beta) \quad \text{for } u < \frac{1}{\beta} \tag{3.2}$$

Since the Laplace exponent of a general subordinator is given by equation (2.12) we must have that $ub + \int_0^\infty (e^{ux} - 1)\phi(x)dx = -\alpha \ln(1 - u\beta)$. Noting that $b = 0$ (the gamma process has no drift) we can solve for ϕ, the Lévy density of the gamma process. This can be done as follows:

$$\begin{aligned}
-\alpha \ln(1 - u\beta) &= -\alpha \int_0^u \frac{1}{\beta^{-1} - y} dy \quad \text{for } u \leq 0 \\
&= -\alpha \int_0^u \int_0^\infty e^{-x\beta^{-1} + yx} dx dy \\
&= -\alpha \int_0^\infty e^{-x\beta^{-1}} \int_0^u e^{yx} dy dx \\
&= \alpha \int_0^\infty e^{-x\beta^{-1}} \frac{1}{x}(e^{ux} - 1) dx \quad \text{since } u \leq 0 \\
&= \int_0^\infty (e^{ux} - 1) \frac{\alpha e^{-\frac{x}{\beta}}}{x} dx
\end{aligned}$$

Hence the solution for $\phi(x)$ is given by

$$\phi(x) = \frac{\alpha e^{-\frac{x}{\beta}}}{x} \tag{3.3}$$

The derivation above can be found in Sato (1999). Note that $\int_0^\infty (x \wedge 1)\phi(x)dx < \infty$ satisfying the condition of a Lévy measure of a subordinator. The Lévy density $\phi(x)$ describes how frequently various jump sizes arrive. Since $\int_0^\infty \phi(x)dx = \infty$, the gamma process has an infinite number of jumps arriving per unit time (i.e. infinite activity process). The majority of these jumps are small since the number of jump arrivals (per unit time) of size greater than 1 is finite. Equation

[1]Existence of a Lévy process is guaranteed since the $Gamma(\alpha,\beta)$ distribution is infinitely divisible. Therefore, there exists a Lévy process $(X_t)_t$ such that X_1 has a gamma distribution

(3.3) verifies this since an exponential damping factor of β curtails the arrival of large jumps. The γ parameter[2] must be chosen such that $b = 0$. Therefore, $\gamma = \int_0^1 \alpha e^{-\frac{x}{\beta}} dx = \alpha\beta(1 - e^{-\frac{1}{\beta}})$. In summary, the gamma process $G_t(\alpha, \beta)$, is a subordinator with characteristic triplet $(0, \frac{\alpha e^{-\frac{x}{\beta}}}{x}, \alpha\beta(1 - e^{-\frac{1}{\beta}}))$.

□

Definition 3.1.2 *Suppose $X(t,\omega)$ is a one-dimensional arithmetic Brownian motion with constant drift ϑ and volatility σ given by $X(t,\omega) = \vartheta t + \sigma B(t,\omega)$ where $B(t,\omega)$ is a one-dimensional standard Brownian motion, i.e. $X(t,\omega)$ is a Lévy process with characteristic triplet $(\sigma^2, 0, \vartheta)$. Let $G(t,\omega)$ be a gamma[3] subordinator, i.e. $G(t,\omega)$ is a driftless gamma process. The process $Y(t,\omega) = X(G(t,\omega),\omega) = \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega)$ is known as a variance gamma process.*

The name originates from the fact that the variance of the Brownian motion has a gamma distribution. The univariate variance gamma process is simply a one-dimensional Brownian motion with drift subordinated by a gamma process. Let us discuss why a gamma process would be a suitable choice.

Suppose we are dealing with a dynamic random experiment which models the arrivals of a certain object[4]. Let N_t represent the number of arrivals by time t, i.e. N_t is a counting process which is clearly increasing in t. Assume that the process $(N_t)_t$ has independent and stationary increments. Define a random variable T which represents the time till the first arrival. Now the set $\{N_t = 0\}$ occurs iff $\{T > t\}$ which implies they have equal measure, i.e. $\mathbb{P}(N_t = 0) = \mathbb{P}(T > t)$. Letting $\beta(t) = \mathbb{P}(N_t = 0)$ we can deduce from the independent and stationary increments of N_t that

$$\begin{aligned}
\beta(t+s) &= \mathbb{P}(N_{t+s} = 0) \\
&= \mathbb{P}(N_{t+s} - N_t = 0 \cap N_t = 0) \\
&= \mathbb{P}(N_{t+s} - N_t = 0)\mathbb{P}(N_t = 0) \\
&= \mathbb{P}(N_s = 0)\mathbb{P}(N_t = 0) \\
&= \beta(s)\beta(t)
\end{aligned}$$

i.e. β is a multiplicative continuous decreasing function of t. Therefore, there exists a $\lambda > 0$ such that $\mathbb{P}(N_t = 0) = e^{-\lambda t}$. Since $\mathbb{P}(N_t = 0) = \mathbb{P}(T > t)$ we can deduce that $\mathbb{P}(T > t) = e^{-\lambda t}$. We have just shown that T (the time till arrival of the first object) has an exponential distribution with parameter

[2]From the characteristic triplet

[3]We use a slight change in notation. $G(t,\omega) = G_t(\alpha,\beta)$ for some α and β. For each fixed t, $G(t,\omega)$ has density given by equation (3.1) with $h = t$

[4]In our case, information

λ. Now letting T_i represent the time till the i^{th} arrival with $T_0 = 0$, we can write $T_N = \sum_{i=1}^{N}(T_i - T_{i-1})$ as a telescoping series. Our objective is to find the distribution of T_N. Observe that

$$\begin{aligned}
\mathbb{P}(T_i - T_{i-1} > t) &= \mathbb{P}(N_{s+t} - N_s = 0 | T_{i-1} = s) \\
&= \mathbb{P}(N_{s+t} - N_s = 0 | N_s = i-1) \\
&= \mathbb{P}(N_{s+t} - N_s = 0) \\
&= \mathbb{P}(N_t = 0) \\
&= \mathbb{P}(T_1 > t) \qquad (3.4)
\end{aligned}$$

We have just shown that the sequence of random variables $T_i - T_{i-1}$ for $i = 1, ..., N$ are identically distributed. Each $T_i - T_{i-1}$ has the same distribution as T_1 which is exponential with parameter λ. To show independence of $(T_i - T_{i-1})_{i=1,...,N}$ note that

$$\begin{aligned}
\mathbb{P}(T_i - T_{i-1} > t | T_{i-1} - T_{i-2} = p) &= \mathbb{P}(N_{s+t} - N_s = 0 | T_{i-1} = s, T_{i-1} - T_{i-2} = p) \\
&= \mathbb{P}(N_{s+t} - N_s = 0 | T_{i-2} = s - p) \\
&= \mathbb{P}(N_{s+t} - N_s = 0 | N_{s-p} = i-2) \\
&= \mathbb{P}(N_{s+t} - N_s = 0) \\
&= \mathbb{P}(T_1 > t) \qquad (3.5)
\end{aligned}$$

i.e. $\mathbb{P}(T_i - T_{i-1} > t | T_{i-1} - T_{i-2} = p) = \mathbb{P}(T_i - T_{i-1} > t)$ implies that the random variable $T_i - T_{i-1}$ is independent of $T_{i-1} - T_{i-2}$ for $i = 2, \ldots, N$. We have shown that $T_N = \sum_{i=1}^{N} T_i - T_{i-1}$ is a sum of independent and identically distributed exponential random variables. It is easy to show that if $X_1, \ldots, X_N$ form an i.i.d. sequence of exponential random variables with common parameter λ then $\sum_{i=1}^{N} X_i$ follows a gamma distribution[5] with parameters N and $\frac{1}{\lambda}$. We have found the distribution of T_N which is gamma with parameters N and $\frac{1}{\lambda}$ with density function $f_{T_N}(x) = \frac{\lambda^{-N} x^{N-1} e^{-\frac{x}{\lambda}}}{\Gamma(N)} 1_{x \geq 0}$. Since T_N consists of a sum of i.i.d. random variables it should be clear that these properties translate to the process $(T_N)_N$. i.e. $(T_N)_N$ has independent and stationary increments.

The subordinator used in our applications is to model information arrival. If G_t represents information arrivals and if the time between arrivals has independent and stationary increments then G_t should be modelled using a gamma distribution. The assumption of independent increments of information arrivals is realistic since if time dependence existed between information arrivals then patterns of trade times could be detected which is hardly realistic especially if our underlying security is highly liquid. There should be no reason why we expect our daily/weekly distribution of information arrivals to change at some time

[5]This can be done via characteristic functions

in the future[6] since it's unreasonable to assume that we have knowledge of the future liquidity of the underlying security. The assumptions of independent and stationary increments are reasonable and justifies the usage of the gamma process in modelling information arrivals.

□

The variance gamma process uses a gamma process to model information arrival. The gamma process has an infinite arrival rate. Most of these arrivals are very small. This can be verified by noting that the Lévy measure of the gamma process is infinite on a set $A \subseteq [0, 1]$. The gamma process is of infinite activity but of finite variation. See Applebaum (2004) for details.

3.2 Properties of the Univariate Variance Gamma Process

We will henceforth denote the process

$$Y(t, \omega) = X(G(t, \omega), \omega) = \vartheta G(t, \omega) + \sigma B(G(t, \omega), \omega)$$

where $G(t, \omega) = G_t(\alpha, \beta)$ is a gamma process (a subordinator) as a variance gamma process. Recall that $X(t, \omega)$ is an arithmetic Brownian motion. We need to choose α and β such that $\mathbb{E}[G(t, \omega)] = t$. This is because we would like, on average, our stochastic clock $G(t, \omega)$ to coincide with calendar time t. Since $\mathbb{E}[G(t, \omega)] = \alpha\beta t$, we must choose $\alpha = \frac{1}{\beta}$. We now have a reduction in one parameter. Defining $\beta = \nu$ (not to be confused with the Lévy measure notation) we now have a gamma process $G(t, \omega) = G_t(\frac{1}{\nu}, \nu)$. Note that in this parameterization, we have the $\mathbb{E}[G(t, \omega)] = t$ and $Var[G(t, \omega)] = \nu t$. This parameterization for $G(t, \omega)$ will be used throughout the rest of the chapter.

3.2.1 Characteristic triplet of $Y(t, \omega)$

Let the characteristic triplet of $Y(t, \omega)$ be given by $(\Sigma^Y, \nu^Y, \gamma^Y)$. The gamma subordinator is discontinuous since it consists purely of jumps and has no drift. Since $Y(t, \omega)$ jumps when $G(t, \omega)$ does, the process $Y(t, \omega)$ has no points of continuity and hence cannot consist of a Brownian component. This would imply that the covariance matrix modelling $Y(t, \omega)$ is 0, i.e. $\Sigma^Y = 0$. This fact can be confirmed by making use of Proposition 2.6.1.
The Lévy measure of the gamma process has density given by $\phi(x) = \frac{e^{-\frac{x}{\nu}}}{\nu x}$ using

[6] Implying non-stationarity

the result given by equation (3.3). We know that $\phi(x)$ has an infinite integral over $\mathbb{R}^+$ implying that there are an infinite number of jumps in any time interval. Making use of Proposition 2.6.1, ν^Y has density

$$\begin{aligned}
\nu^Y(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 s}} e^{-\frac{(x-\vartheta s)^2}{2\sigma^2 s}} \phi(s) ds \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 s}} e^{-\frac{(x-\vartheta s)^2}{2\sigma^2 s}} \frac{e^{-\frac{s}{\nu}}}{\nu s} ds \\
&= \frac{e^{\frac{x\vartheta}{\sigma^2}}}{\nu} \frac{|x|}{|x|} \int_0^\infty \frac{(2\pi\sigma^2)^{-\frac{1}{2}}}{s^{\frac{3}{2}}} e^{-\frac{x^2}{2\sigma^2 s} - \frac{\vartheta^2 s}{2\sigma^2} - \frac{s}{\nu}} ds \\
&= \frac{e^{-\frac{|x|\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}}}{\sigma^2}}}{e^{-\frac{|x|\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}}}{\sigma^2}}} \frac{e^{\frac{x\vartheta}{\sigma^2}}}{\nu|x|} \int_0^\infty \frac{(2\pi\sigma^2 x^{-2})^{-\frac{1}{2}}}{s^{\frac{3}{2}}} e^{-\left(\frac{\vartheta^2}{2\sigma^2} + \frac{1}{\nu}\right)s - \left(\frac{x^2}{2\sigma^2}\right)\frac{1}{s}} ds \\
&= \frac{e^{\frac{x\vartheta}{\sigma^2} - |x|\frac{\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}}}{\sigma^2}}}{\nu|x|} \int_0^\infty \frac{\frac{x}{\sqrt{2\pi}\sigma}}{s^{\frac{3}{2}}} e^{-\left(\frac{\vartheta^2}{2\sigma^2} + \frac{1}{\nu}\right)s - \left(\frac{x^2}{2\sigma^2}\right)\frac{1}{s} + x\frac{\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}}}{\sigma^2}} ds
\end{aligned}$$

The integrand evaluates to one since it can be recognised as a normal inverse Gaussian density given by

$$f(s) = \frac{ct}{s^{\frac{3}{2}}} e^{-\lambda s - (\pi c^2 t^2)\frac{1}{s} + 2ct\sqrt{\pi\lambda}} 1_{s>0}$$

The above is a special case with $c = \frac{x}{\sqrt{2\pi}\sigma}$, $t = 1$ and $\lambda = \left(\frac{\vartheta^2}{2\sigma^2} + \frac{1}{\nu}\right)$. Therefore, the Lévy density of Y_t is given by

$$\nu^Y(x) = \frac{e^{\frac{x\vartheta}{\sigma^2} - |x|\frac{\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}}}{\sigma^2}}}{\nu|x|} \tag{3.6}$$

The 'drift' or γ^Y parameter of $Y(t,\omega)$ can again be obtained by making use of Proposition 2.6.1 and is given by

$$\begin{aligned}
\gamma^Y &= \int_0^\infty \frac{e^{-\frac{s}{\nu}}}{\nu s} \int_{|x|\leq 1} x \frac{1}{\sqrt{2\pi\sigma^2 s}} e^{-\frac{(x-\vartheta s)^2}{2\sigma^2 s}} dx ds \\
&= \int_{|x|\leq 1} x \int_0^\infty \frac{e^{-\frac{s}{\nu}}}{\nu s} \frac{1}{\sqrt{2\pi\sigma^2 s}} e^{-\frac{(x-\vartheta s)^2}{2\sigma^2 s}} ds dx \\
&= \int_{|x|\leq 1} x\nu^Y(x) dx
\end{aligned} \tag{3.7}$$

The characteristic triplet of $Y(t,\omega)$ is $(0, \nu^Y, \gamma^Y)$ with ν^Y and γ^Y given by equations (3.6) and (3.7) respectively.

□

3.2.2 Characteristic Function and Moments of $Y(t,\omega)$

The characteristic function of $Y(t,\omega)$ can be obtained via the Lévy-Khinchin representation since we know the characteristic triplet. This can be done in principle, but requires some difficult integrals. An alternate and far simpler derivation using conditional expectation is given below.

$$\begin{aligned}
\mathbb{E}[e^{iuY_t}] &= \mathbb{E}[\mathbb{E}[e^{iuY_t}|G(t,\omega)]] \\
&= \mathbb{E}[\mathbb{E}[e^{iu(\vartheta G(t,\omega)+\sigma B_{G(t,\omega)})}|G(t,\omega)]] \\
&= \mathbb{E}[e^{G(t,\omega)(i\vartheta u-\frac{1}{2}\sigma^2u^2)}] \\
&= (1-\nu(i\vartheta u-\frac{1}{2}\sigma^2u^2))^{-\frac{t}{\nu}} \qquad (3.8)
\end{aligned}$$

The last step is obtained by making use of the characteristic function of a gamma distribution (Applebaum, 2004). Once the characteristic function of a random variable is known, moments can be found by successive differentiation (provided the moments exist). Letting μ_i denote the i^{th} central moment[7], the first four central moments (Seneta, 2004) are given by

$$\begin{aligned}
\mu_1 &= \vartheta t \\
\mu_2 &= t(\vartheta^2\nu+\sigma^2) \\
\mu_3 &= t\nu\vartheta(3\sigma^2+2\vartheta^2\nu) \\
\mu_4 &= t(3\sigma^4\nu+6\vartheta^4\nu^3+12\sigma^2\vartheta^2\nu^2)+t^2(3\sigma^4+6\sigma^2\vartheta^2\nu+3\vartheta^4\nu^2)
\end{aligned}$$

It should be clear that ϑ primarily controls the skewness of $Y(t,\omega)$ and ν controls the kurtosis. Positive values of ϑ lead to positive skewness and negative values of ϑ lead to negative skewness. Note that when $\vartheta=0$, the skewness is zero and the coefficient of kurtosis[8] at time period 1 is $3(1+\nu)$, i.e. ν is the percentage excess kurtosis over the normal distribution. These additional parameters ϑ and ν give us control over skewness and kurtosis in return distributions with the intent of flattening the volatility skew.

□

3.2.3 Probability Density Function of $Y(t,\omega)$

The probability density function can be expressed in terms of special functions in mathematics as shown by Madan *et al.* (1998). Since the variance gamma process is conditionally Gaussian, one can use conditional probability to obtain

[7] $\mu_1=\mathbb{E}[Y_t]$ and $\mu_i=\mathbb{E}[(Y_t-E[Y_t])^i]$ for $i=2,3,\ldots$

[8] This is obtained by taking μ_4 and dividing by the square of μ_2

the density function. Suppose that $G(t,\omega) = g$ then the conditional density of $Y(t,\omega)$ is

$$f_{Y_t|G_t=g}(x) = \frac{1}{\sigma\sqrt{2\pi g}} e^{-\frac{(x-\vartheta g)^2}{2\sigma^2 g}}$$

Making use of the standard result in statistics which states that if $f_X(x)$, $f_Y(y)$ and $f_{X|Y}(x)$ are the respective p.d.f.'s of X,Y and $X|Y$ then $f_X(x) = \int_C f_{X|Y}(y) f_Y(y) dy$ where C is the range of Y. Applying this result, we obtain

$$\begin{aligned} f_{Y_t}(x) &= \int_0^\infty \frac{1}{\sigma\sqrt{2\pi g}} e^{-\frac{(x-\vartheta g)^2}{2\sigma^2 g}} \frac{g^{\frac{t}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})} dg \\ &= \frac{2}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}} \frac{1}{2} \int_0^\infty g^{(\frac{t}{\nu}-\frac{3}{2})} e^{-\frac{(x^2-2x\vartheta g+\vartheta^2 g^2)}{2\sigma^2 g} - \frac{g}{\nu}} dg \\ &= \frac{2e^{(\frac{x\vartheta}{\sigma^2})}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}} \frac{1}{2} \int_0^\infty g^{(\frac{t}{\nu}-\frac{1}{2})-1} e^{-\frac{1}{2}(\frac{x^2}{\sigma^2 g} + g(\frac{\vartheta^2}{\sigma^2}+\frac{2}{\nu}))} dg \end{aligned}$$

Letting $p = \frac{\sqrt{2\frac{\sigma^2}{\nu}+\vartheta^2}}{x} g$, we obtain

$$\begin{aligned} f_{Y_t}(x) &= \frac{2e^{(\frac{x\vartheta}{\sigma^2})}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}} \left(\frac{x^2}{2\frac{\sigma^2}{\nu}+\vartheta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} \frac{1}{2} \int_0^\infty p^{(\frac{t}{\nu}-\frac{1}{2})-1} e^{-\frac{\sqrt{x^2\frac{\sigma^2}{\nu}+\vartheta^2}}{2\sigma^2}(\frac{1}{p}+p)} dp \\ &= \frac{2e^{(\frac{x\vartheta}{\sigma^2})}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}} \left(\frac{x^2}{\frac{2\sigma^2}{\nu}+\vartheta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{(\frac{t}{\nu}-\frac{1}{2})}\left(\frac{\sqrt{x^2\left(\frac{2\sigma^2}{\nu}+\vartheta^2\right)}}{\sigma^2}\right) \end{aligned} \quad (3.9)$$

where $K_\eta(\omega)$ is known as the modified Bessel function of the third kind. It has the following integral representation:

$$K_\eta(\omega) = \frac{1}{2} \int_0^\infty \nu^{\eta-1} e^{(-\frac{\omega}{2}(\frac{1}{\nu}+\nu))} d\nu$$

We have thus simplified the p.d.f. of $Y(t,\omega)$ into a term multiplied by a special function in mathematics. The reason why we have bothered to make such a simplification is that there exist many efficient numerical algorithms for computing modified Bessel functions[9]. The following plot illustrates the shape of the variance gamma p.d.f. relative to the normal p.d.f.

[9]Matlab permits the use of the modified Bessel function of the third kind

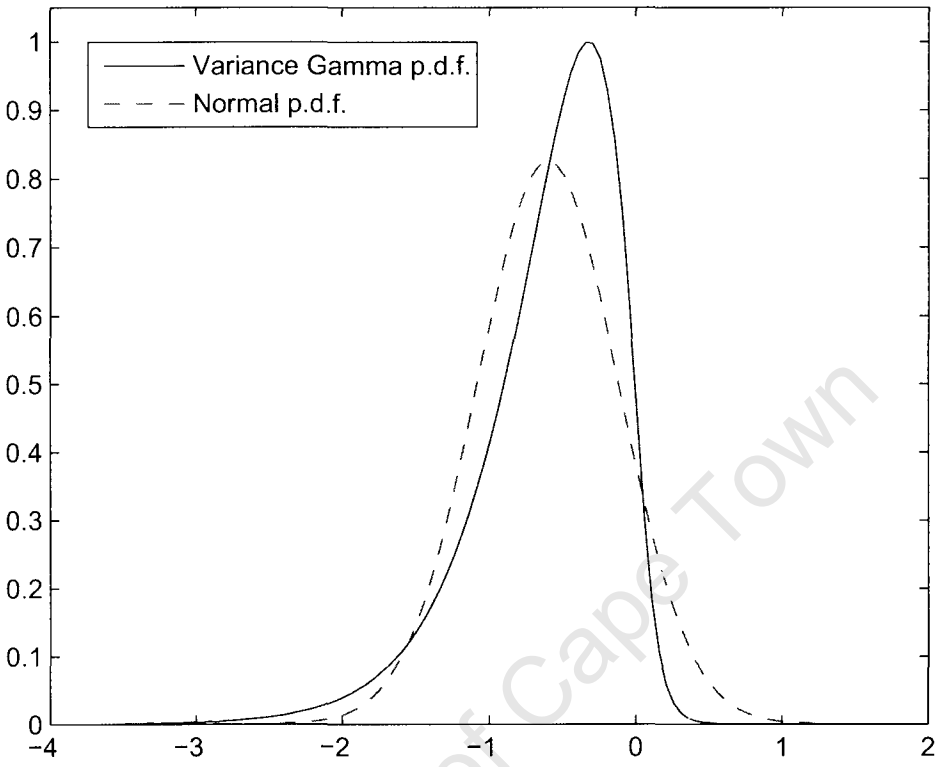

Figure 3.1: Variance gamma p.d.f. with matched mean and variance

Figure 3.1 is a plot of the Variance Gamma p.d.f. given by equation (3.9). The parameters of the variance gamma p.d.f. were $\vartheta = -0.6$, $\sigma = 0.3$, $\nu = 0.4$ and $t = 1$. The normal density is plotted on the same set of axes to illustrate the non-zero skewness and excess kurtosis of the variance gamma density. The mean and variance parameters of the normal density were chosen to match the mean and variance of the variance gamma density. The variance gamma density clearly displays negative skewness (given by ϑ) and is more leptokurtotic (given by ν) than the normal density.

□

3.2.4 Representation of a Variance Gamma Process

The variance gamma process has a representation that becomes particularly useful when simulating. In fact, this representation is fundamental in applying the 'truncated difference-of-gammas sampling' discussed in later chapters.

Proposition 3.2.1 *Let* $X_t = \vartheta G_t + \sigma W_{G_t}$ *be a variance gamma process with*

$G_{t+h} - G_t \sim G\left(\frac{h}{\nu}, \nu\right)$, i.e. $G_{t+h} - G_t$ has mean h and variance νh. Then

$$Y_t = \Gamma^+(t) - \Gamma^-(t) \stackrel{d}{=} X_t$$

where $\Gamma^+(t) \sim G\left(\frac{t}{\nu}, \frac{\nu_p}{\mu_p}\right)$ and $\Gamma^-(t) \sim G\left(\frac{t}{\nu}, \frac{\nu_n}{\mu_n}\right)$ are independent with

$$\mu_p = \frac{1}{2}(\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}} + \vartheta) \qquad \nu_p = \mu_p^2 \nu$$
$$\mu_n = \frac{1}{2}(\sqrt{\vartheta^2 + \frac{2\sigma^2}{\nu}} - \vartheta) \qquad \nu_n = \mu_n^2 \nu$$

The proof is straightforward and we refer you to Madan *et al.* (1998) for details. The above proposition provides an alternative method of simulating the variance gamma process.

□

Chapter 4

Option Pricing

This chapter explores the pricing of various options under both Black-Scholes and variance gamma dynamics. We begin with a short literature review on option pricing under variance gamma dynamics.

The paper written by Madan *et al.* (1998) provide a method of finding a risk-neutral measure. They provide parameter estimates of both the historical and risk-neutral variance gamma process. They conclude that the statistical distribution is symmetric while the risk-neutral distribution, as implied by option prices quoted, is asymmetric. They conclude that the variance gamma model is superior in pricing options when compared to Black-Scholes. This was conducted via likelihood ratio tests.

Throughout the rest of the chapter, the technique of risk-neutral valuation will be used to price options. The technique of risk-neutral valuation is briefly discussed below.

Let $(S_t)_t$ be an adapted stochastic process. Let $f(S_T)$ be a European derivative on S. Assuming that we can find a risk-neutral measure $\mathbb{Q}$ (which is equivalent to stating that our market model is arbitrage-free) then a no-arbitrage price V_t of this derivative is given by

$$V_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[f(S_T)] \tag{4.1}$$

where r is the constant risk-free rate of interest. Recall that a risk-neutral measure is a measure which is equivalent[1] to the 'real world' measure under which the discounted (by the constant risk-free rate) securities are martingales, i.e. $\mathbb{Q}$ is a risk-neutral measure iff $\mathbb{Q}$ satisfies $(u < t)$

$$\mathbb{E}^{\mathbb{Q}}[S_t e^{-rt}|\Im_u] = S_u e^{-ru} \tag{4.2}$$

[1] $\mathbb{P}$ is equivalent to $\mathbb{Q}$ if $\mathbb{P}(A) = 0$ whenever $\mathbb{Q}(A) = 0$ and conversely

It should be stressed that equation (4.1) will give you a no-arbitrage price but certainly does not guarantee this price to be unique. The uniqueness of V_t requires the additional assumption that every derivative $f(S_T)$ can be replicated. The replication depends on what model you choose for the underlying security S_t.

□

The variance gamma market, where $\ln \frac{S_t}{S_0}$ has a 'real world' density given by

$$f_{Y_t}(x) = \frac{2e^{\left(\frac{x\vartheta}{\sigma^2}\right)}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}}\left(\frac{x^2}{\frac{2\sigma^2}{\nu}+\vartheta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{(\frac{t}{\nu}-\frac{1}{2})}\left(\frac{\sqrt{x^2\left(\frac{2\sigma^2}{\nu}+\vartheta^2\right)}}{\sigma^2}\right)$$

belongs to the class of incomplete models. The variance gamma model assumes that $S_t = S_0 e^{\vartheta G(t,\omega)+\sigma B(G(t,\omega),\omega)}$ and their are many ways to manipulate S_t in order to satisfy equation (4.2) whilst remaining in a variance gamma market.

Since the variance gamma model is incomplete, different risk-neutral measures will result in different derivative prices. We make an assumption that we can find a risk-neutral measure (which excludes arbitrage opportunities in our model). However, for these prices to be meaningful, approximate hedges need to be constructed. Prices in an incomplete model are not enough. These need to be compared to the cost of the approximate hedge. Subsequent chapters will only be concerned with pricing and further work is necessary in constructing hedges. Even dynamic hedging (e.g. delta hedging) in a Black-Scholes world becomes infeasible in the case of barrier options. This is because the option's delta becomes infinite as the underlying approaches the barrier. Alternate hedging strategies such as static hedging needs to be investigated. See Bosman (2003) for a discussion of static hedging of barrier options in a Black-Scholes world.

4.1 Construction of a Risk-Neutral Measure

Since the variance gamma model belongs to the class of incomplete markets, one needs to artificially construct a risk-neutral measure such that equation (4.2) is satisfied. One such method is to solve for a constant η such that $\mathbb{E}[S_t] = \mathbb{E}[S_0 e^{\eta t+\vartheta G(t,\omega)+\sigma B(G(t,\omega),\omega)}] = S_0 e^{(r-q)t}$ where q is the constant continuous dividend yield. This method is adapted from Madan *et al.* (1998). Note that this is but one particular method of artificially satisfying equation (4.2). This parameterization suggests that the risk-neutral log price of S_t is given by

$$\ln \frac{S_t}{S_0} = \eta t + \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega). \tag{4.3}$$

ηt is a mean-correcting term which varies linearly with time which seems plausible. We need to solve for η such that equation (4.2) is satisfied so

$$\begin{aligned}
\mathbb{E}[S_t] = \mathbb{E}[S_0 e^{\eta t + \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega)}] &= S_0 e^{(r-q)t} \\
\mathbb{E}[\mathbb{E}[e^{\eta t + \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega)} | G(t,\omega)]] &= e^{(r-q)t} \\
\mathbb{E}[e^{\frac{\sigma^2 G(t,\omega)}{2} + \vartheta G(t,\omega) + \eta t}] &= e^{(r-q)t} \\
e^{\eta t}\mathbb{E}[e^{G(t,\omega)(\vartheta + \frac{\sigma^2}{2})}] &= e^{(r-q)t} \\
e^{\eta t}(1 - (\vartheta + \frac{\sigma^2}{2})\nu)^{-\frac{t}{\nu}} &= e^{(r-q)t}
\end{aligned}$$

Rearranging and solving for η yields

$$\eta = r - q + \nu^{-1}\ln(1 - (\vartheta + \frac{\sigma^2}{2})\nu) \tag{4.4}$$

Choosing η as above satisfies equation (4.2). We have found one particular risk-neutral measure which we'll call $\mathbb{Q}$. Under $\mathbb{Q}$, the log price of S_t is given by

$$\ln\frac{S_t}{S_0} = t(r - q + \nu^{-1}\ln(1 - (\vartheta + \frac{\sigma^2}{2})\nu)) + \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega) \tag{4.5}$$

The density of $Y_t = \ln\frac{S_t}{S_0}$ under $\mathbb{Q}$ which we denote by $f^{\mathbb{Q}}_{Y_t}(x)$ can be obtained using equation (3.9). Note that

$$\begin{aligned}
f^{\mathbb{Q}}_{Y_t}(x) &= \frac{d}{dx}\mathbb{P}(\eta t + \vartheta G_t + \sigma B_{G_t} \leq x) \\
&= \frac{d}{dx}\mathbb{P}(\vartheta G_t + \sigma B_{G_t} \leq x - \eta t) \\
&= f_{Y_t}(x - \eta t)
\end{aligned}$$

where $f_{Y_t}(x)$ is given by equation (3.9). Hence, Y_t has density under $\mathbb{Q}$ given by

$$f^{\mathbb{Q}}_{Y_t}(x) = \frac{2e^{\left(\frac{(x-\eta t)\vartheta}{\sigma^2}\right)}}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu})\sigma\sqrt{2\pi}}\left(\frac{(x-\eta t)^2}{\frac{2\sigma^2}{\nu} + \vartheta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{(\frac{t}{\nu}-\frac{1}{2})}\left(\frac{\sqrt{(x-\eta t)^2\left(\frac{2\sigma^2}{\nu} + \vartheta^2\right)}}{\sigma^2}\right) \tag{4.6}$$

4.2 Pricing European Options

The previous section constructed a risk-neutral measure $\mathbb{Q}$ under which all discounted securities are martingales. We can now use this risk-neutral measure to price contingent claims on these securities. Suppose we wish to price an option with payoff $f(S_T)$ at expiry T where $\ln\frac{S_T}{S_0}$ has 'real world' density given by

equation (3.9). Using the principle of risk-neutral valuation, the price V_0 of this option at time 0 is given by

$$\begin{aligned} V_0 &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[f(S_T)] \\ &= e^{-rT}\int_0^{\infty} g(w)f(w)dw \end{aligned}$$

where r is the constant risk-free rate and g is the risk-neutral density of S_T. Letting $x = \ln\frac{w}{S_0}$ or $w = S_0e^x$, we obtain

$$V_0 = e^{-rT}\int_{-\infty}^{\infty} S_0e^x g(S_0e^x)f(S_0e^x)dx$$

Note that $S_0e^x g(S_0e^x) = f^{\mathbb{Q}}_{Y_T}(x)$, i.e. the risk-neutral density of $\ln\frac{S_T}{S_0}$ given by equation (4.6). To see this, note that

$$\begin{aligned} f^{\mathbb{Q}}_{Y_T}(x) = \frac{d}{dx}\mathbb{Q}(Y_T \leq x) &= \frac{d}{dx}\mathbb{Q}(\ln\frac{S_T}{S_0} \leq x) \\ &= \frac{d}{dx}\mathbb{Q}(S_T \leq S_0e^x) \\ &= \frac{d}{dx}\int_0^{S_0e^x} g(w)dw \\ &= S_0e^x g(S_0e^x) \end{aligned}$$

We can therefore write V_0 in terms of the risk-neutral log density of S_T as

$$V_0 = e^{-rT}\int_{-\infty}^{\infty} f(S_0e^x)f^{\mathbb{Q}}_{Y_t}(x)dx \tag{4.7}$$

Suppose we would like to price a European call option with strike K which has payoff given by $f(S_T) = \max(S_T - K, 0)$. The price of this option at time period 0 is given by

$$\begin{aligned} V_0 &= e^{-rT}\int_{-\infty}^{\infty} \max(S_0e^x - K, 0)f^{\mathbb{Q}}_{Y_t}(x)dx \\ &= e^{-rT}\int_{\ln\frac{K}{S_0}}^{\infty} (S_0e^x - K)f^{\mathbb{Q}}_{Y_t}(x)dx \end{aligned} \tag{4.8}$$

The above expression for a European call option has been given a 'closed form' solution in terms of degenerate hypergeometric functions. See Madan *et al.* (1998) for details. The author believes that there are some numerical issues and care should be taken when implementing the 'closed form' solution. The integral representation of a European option with payoff $f(S_T)$ given by equation (4.7) can be approximated numerically using various quadrature techniques. This will be discussed below.

□

4.3 Pricing Path Dependent Options

The pricing of path dependent options almost always requires Monte Carlo simulation. There are very few cases (e.g. pricing barrier options in a Black-Scholes market) where 'closed form' or even integral representations of option prices exist. Pricing path dependent options requires simulation of entire paths and not just the value at expiry of the option. It now becomes necessary to discuss simulation of variance gamma sample paths with the aid of an example. A far more comprehensive discussion of simulating the variance gamma process is given later in the chapter.

Suppose we would like to price a European[2] Asian option on an underlying security S with strike K and payoff at expiry T given by $f(S_T) = \max(\frac{1}{N}\sum_{i=1}^{N} S_{t_i} - K, 0)$ with $t_N = T$. The sequence of times t_i is increasing but $t_i - t_{i-1}$ is not necessarily equal to $t_j - t_{j-1}$ for $i \neq j$. This particular option requires knowledge of the process at times $(t_i)_{i=1,\ldots N}$. Recall that in a variance gamma market, S_{t_i} evolves in the following manner

$$S_{t_i} = S_0 e^{\eta t_i + \vartheta G(t_i,\omega) + \sigma B(G(t_i,\omega),\omega)}$$

Note that the times t_i are measured in years[3] and t_i represents the number of years since the current time 0. The independent increments of the variance gamma process allow us to write G_{t_i} as a sum of independent gamma variates, i.e.

$$G_{t_i} = \sum_{j=1}^{i} G_{t_j} - G_{t_{j-1}}$$

and similarly for Brownian motion. This representation is used in the simulation algorithm described below.

1. Simulate N independent gamma random variates $P_i \sim G(\frac{t_i - t_{i-1}}{\nu}, \nu)$ for $i = 1, \ldots, N$ where $t_0 = 0$
2. Simulate N independent standard normal random variates N_i for $i = 1, \ldots, N$
3. Compute $W_i = S_0 e^{\eta t_i + \vartheta \sum_{j=1}^{i}(P_j - P_{j-1}) + \sigma \sum_{j=1}^{i} \sqrt{(P_j - P_{j-1})} N_j}$ for $i = 1, \ldots, N$ with $P_0 = 0$
4. Compute $\mathbb{D}_j = \max(\frac{1}{N}\sum_{i=1}^{N} W_i - K, 0)$ for $j = 1, \ldots, M$

[2]This often causes confusion. The 'European' part of the name implies that the option can only be exercised at expiry

[3]This is only a convention since volatility is quoted in years and our parameters ϑ, σ, ν are assumed to be annualized

5. Repeat the first 4 steps M times (usually around 50000 times)

The resulting price of this Asian option is given by $e^{-rT}\frac{1}{M}\sum_{j=1}^{M}\mathbb{D}_j$. Note that each payoff $\mathbb{D}_j$ is exact. There is no discretization error. The only source of approximation error is the intrinsic error induced by Monte Carlo methods. This is an example of a path dependent option which can be priced exactly. A barrier option is an example of a path dependent option which cannot be priced exactly. More sophisticated simulation algorithms need to be constructed to eliminate discretization error. This is discussed further on in detail.
There are many existing algorithms for generating gamma and normal random variates. See William H. Press & Flannery (1992) for efficient algorithms.

□

4.4 Pricing Barrier Options under Black-Scholes Dynamics

Barrier options are path dependent options whose payoff depends on whether the underlying has breached a barrier[4] or not. They provide an attractive alternative to vanilla options since they are cheaper and provide the holder the opportunity to incorporate some of his personal views as to how the underlying will move. Barrier options can be broadly categorized into *knock-out* and *knock-in* options. The former option has zero payoff as soon as the underlying breaches the barrier whereas the latter has zero payoff until the barrier is hit. Note that even if a *knock-in* option breaks the barrier then the option can still have zero payoff. This will occur when the option expires out-of-the-money. The pricing of barrier options requires knowledge of maximal and minimal processes which leads us to our next definition.

Definition 4.4.1 *Let $(X_t)_t$ be a stochastic process defined on a probability space. The maximum and minimum processes associated with $(X_t)_t$ are defined as*

$$\begin{aligned} X_t^{Max} &= \sup_{s\leq t} X_s \\ X_t^{Min} &= \inf_{s\leq t} X_s \end{aligned}$$

Figure (4.1) illustrates a sample path of Brownian motion with its associated maximum and minimum processes. It should be clear from the definition that both X_t^{Max} and X_t^{Min} are of finite variation since the former is increasing and the latter is decreasing in t.

[4]There may be two barriers as in a double barrier option

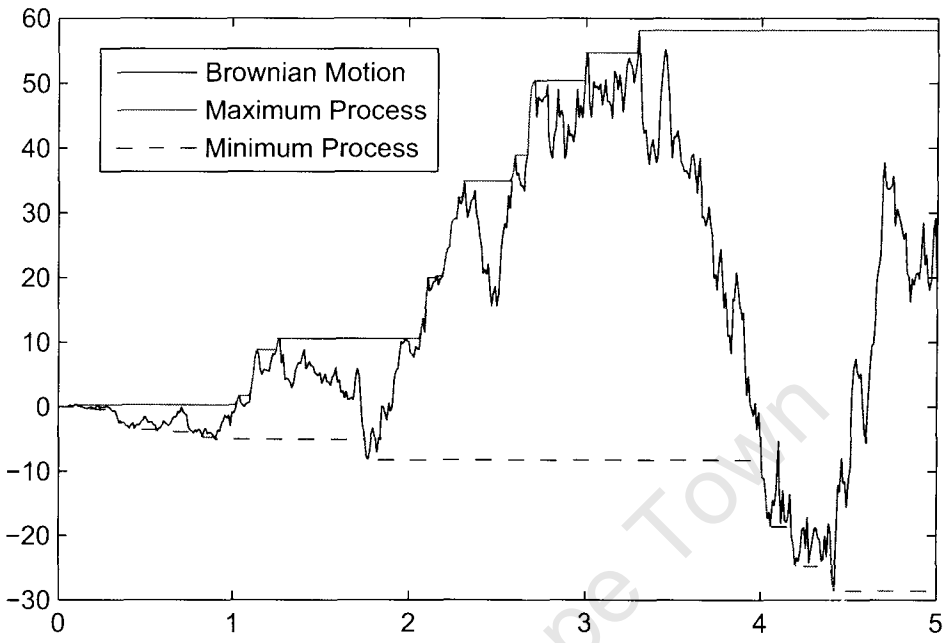

Figure 4.1: Sample path of Brownian motion with its maximum and minimum processes

We will encounter various types of barrier options. Consider a *knock-out* barrier option on an underlying variable S with barrier $L > S_0$ and payoff $\Phi(S_T)$. We will assume that the above option is European and expires at time T. This option has the same payoff $\Phi(S_T)$ as a European option provided S_t for all $t \in [0, T]$ has not breached the barrier. If we assume we can find a risk-neutral measure $\mathbb{Q}$ such the $S_t e^{rt}$ is a $\mathbb{Q}$-martingale then a no-arbitrage price is given by

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)1_{S_T^{Max}<L}] \tag{4.9}$$

It should be clear that the underlying S_t has not breached the barrier L through its entire life from $t = 0$ to $t = T$ if and only if $\{\omega \in \Omega : S_T^{Max} < L\}$ has occurred. The joint distribution of S_T and S_T^{Max} is required in order to compute the expectation. We actually only require to find the joint distribution $\ln S_t, \ln S_t^{Max}$ since it is possible to make a logarithmic transformation which simplifies the computation of the integral. This joint distribution can be obtained in closed form in certain circumstances which is, of course, model dependent. This section assumes that we live in a Black-Scholes world where assets are driven by geometric Brownian motion.

We attempt to find closed form solutions for European *knock-in* and *knock-out* barrier options for arbitrary payoff $\Phi(S_T)$. We will henceforth assume risk-neutral

dynamics of $(S_t)_t$ as follows

$$dS = rSdt + \sigma SdW$$

which we'll label $\mathbb{Q}$-dynamics[5]. $(W_t)_t$ is a standard $\mathbb{Q}$-Brownian motion and r is the constant continuously compounded risk-free rate of interest. This is the familiar equation of geometric Brownian motion and the distribution of S_T is easily solved using Itô's formula, i.e. $\ln S_T \sim N(\ln S_0 + (r - \frac{1}{2}\sigma^2)T, \sigma^2 T)$ under $\mathbb{Q}$. In order to price barrier options we need to find the joint distribution of $Y_t = \ln S_0 + \ln S_t$ and its maximum/minimum processes Y_t^{Max},Y_t^{Min} under $\mathbb{Q}$. Y_t is of course, arithmetic Brownian motion. This is the subject of the next section.

4.4.1 Derivation of several useful joint distributions

We first need to state and prove the reflection principle of Brownian motion.

Proposition 4.4.1 *Let W_t be a standard Brownian motion defined on a probability space $(\Omega, \Im, \mathbb{P})$ and let W_t^{Max} be its maximum process. Now for each fixed t*

$$\mathbb{P}(W_t \leq x, W_t^{Max} \geq y) = \mathbb{P}(W_t \geq 2y - x) = N\left(\frac{x - 2y}{\sqrt{t}}\right) \tag{4.10}$$

where $N(x)$ denotes the standard cumulative normal distribution function, i.e.

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

Proof: Let $\tau = \inf\{t : W_t \geq y\}$. τ is the first time that W_t hits some barrier y. Now $\{\omega \in \Omega : \tau \leq t\}$ is clearly known at time t and hence $\Im_t$-measurable. Thus, τ is a $\Im_t$-stopping time. Consider a new process $\widehat{W}_t$ defined by

$$\widehat{W}_t = W_t 1_{\{\tau < t\}} + (2y - W_t) 1_{\{\tau \geq t\}}$$

$\widehat{W}_t$ is known as the reflected Brownian motion. For each sample path $\omega \in \Omega$, $\widehat{W}_t$ is identical to W_t if or until W_t hits a barrier y which occurs at a random time τ. Thereafter, $\widehat{W}_t$ reflects in the horizontal line $W_\tau = y$, i.e. $\widehat{W}_t$ (for $\tau \geq t$) is the symmetric image of W_t around the line $W_\tau = y$. One would expect $\widehat{W}_t$ to be a $\mathbb{P}$-Brownian motion. To see this mathematically, consider the process $\hat{W}_t = W_{\tau+t} - W_\tau$. Now, $\hat{W}_t \stackrel{d}{=} W_t$ since Brownian motion is stationary and strong Markov. It follows that $\hat{W}_t$ is a $\mathbb{P}$-Brownian motion. Note that W_t and $\widehat{W}_t$ can be written as

$$\begin{aligned} W_t &= W_t 1_{\{\tau < t\}} + (y + \hat{W}_{t-\tau}) 1_{\{\tau \geq t\}} \\ \widehat{W}_t &= W_t 1_{\{\tau < t\}} + (y - \hat{W}_{t-\tau}) 1_{\{\tau \geq t\}} \end{aligned}$$

[5]There should be no ambiguity about which risk-neutral measure we have chosen since $\mathbb{Q}$ is unique in a Black-Scholes world

Now since $\hat{W}_{t-\tau}$ is a Brownian motion then so is $-\hat{W}_{t-\tau}$. It follows that they must have the same distributions, i.e. $W_t \stackrel{d}{=} \widehat{W}_t$. Since they both have continuous a.s. sample paths it follows that $\widehat{W}_t$ is a $\mathbb{P}$-Brownian motion. Now, let $x < y$. A little thought shows you that the following two sets are equal.

$$\{\omega \in \Omega : W_t \leq x, W_t^{Max} \geq y\} = \{\omega \in \Omega : \widehat{W}_t \geq 2y - x, \widehat{W}_t^{Max} \geq y\}$$

Let $A = \{\omega \in \Omega : \widehat{W}_t \geq 2y - x\}$ and let $B = \{\omega \in \Omega : \widehat{W}_t^{Max} \geq y\}$. Now $A \subset B$, which implies $A \cap B = A$. Therefore,

$$\{\omega \in \Omega : \widehat{W}_t \geq 2y - x, \widehat{W}_t^{Max} \geq y\} = \{\omega \in \Omega : \widehat{W}_t \geq 2y - x\}$$

Now since $W_t \stackrel{d}{=} \widehat{W}_t$ it follows that

$$\begin{aligned} \mathbb{P}(W_t \leq x, W_t^{Max} \geq y) = \mathbb{P}(\widehat{W}_t \geq 2y - x) &= \mathbb{P}(W_t \geq 2y - x) \\ &= 1 - N\left(\frac{2y - x}{\sqrt{t}}\right) \\ &= N\left(\frac{x - 2y}{\sqrt{t}}\right) \end{aligned}$$

using the fact that $1 - N(x) = N(-x)$ and the result is proved.

□

A useful corollary to the above proposition is required.

Corollary 4.4.1 *Let W_t be a standard Brownian motion under $\mathbb{P}$ and let W_t^{Max} be its maximum process. Let $F_t(x, y)$ denote the joint distribution function of W_t and W_t^{Max} for a fixed t. The joint distribution function (for $x \leq y$) is given by*

$$F_t(x, y) = \mathbb{P}(W_t \leq x, W_t^{max} \leq y) = N\left(\frac{x}{\sqrt{t}}\right) - N\left(\frac{x - 2y}{\sqrt{t}}\right) \tag{4.11}$$

Proof: We only consider the case where $x \leq y$. This is because the set $\{\omega \in \Omega : W_t \in [y, y + \Delta y], W_t^{max} \leq y\}$ has $\mathbb{P}$-measure zero. Hence $F_t(s, y) = F_t(y, y)$ for all $s \geq y$. We also must have that $y \geq 0$ since W_t is a standard Brownian motion which starts at zero and therefore cannot have a maximum process which takes on negative values. Recall that $\mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap \overline{B})$ is true for any probability measure $\mathbb{P}$ and measurable sets A and B. $\overline{B}$ denotes the complement of B. It follows that

$$\begin{aligned} F_t(x, y) = \mathbb{P}(W_t \leq x, W_t^{max} \leq y) &= \mathbb{P}(W_t \leq x) - \mathbb{P}(W_t \leq x, W_t^{Max} \geq y) \\ &= N\left(\frac{x}{\sqrt{t}}\right) - N\left(\frac{x - 2y}{\sqrt{t}}\right) \end{aligned}$$

by making use of Proposition 4.4.1 and the result is proved.

□

We have now found the joint distribution for standard Brownian motion and its maximum process. We would like to extend this result to arithmetic Brownian motion which starts at some point α.

Recall some facts about changes of measure in a Itô diffusion setting. See Bjork (2004) for a more comprehensive text. Recall that if $X_t = \mu t + \sigma W_t$ is an arithmetic $\mathbb{P}$-Brownian motion then for all $\lambda \in \mathbb{R}$, $\tilde{W}_t = W_t - \lambda t$ is also a Brownian motion but under a different measure, say $\tilde{\mathbb{P}}$. For the special case of Girsanov's theorem applied to a one-dimensional arithmetic Brownian motion then λ always exists and is given by $\lambda = \sigma^{-1}(\nu - \mu)$ where $\nu \in \mathbb{R}$. λ is known as the Girsanov kernel. The resulting arithmetic $\tilde{\mathbb{P}}$-Brownian motion will now have drift ν but the same volatility σ, i.e. $\tilde{X}_t = \nu t + \sigma \tilde{W}_t$ is an arithmetic $\tilde{\mathbb{P}}$-Brownian motion. Moreover, the Doleans exponential $\xi(\lambda W)_t$ at a fixed time t is nothing other than the Radon-Nikodym derivative of the change of measure from $\mathbb{P}$ to $\tilde{\mathbb{P}}$. i.e.

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \xi(\lambda W)_t = e^{\lambda W_t - \frac{1}{2}\lambda^2 t}$$

Then clearly the process $\left(\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right)_t$ is a $\mathbb{P}$-martingale[6]. Changing measures from $\tilde{\mathbb{P}}$ to $\mathbb{P}$ (with Girsanov kernel $-\lambda$) induces the following Radon-Nikodym derivative

$$\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} = \xi(-\lambda \tilde{W})_t = e^{-\lambda \tilde{W}_t - \frac{1}{2}\lambda^2 t}$$

which is now a $\tilde{\mathbb{P}}$-martingale. The above result immediately implies that if $(X_t)_t$ is an Itô diffusion diven by a $\mathbb{P}$-Brownian motion then for a fixed time t

$$\mathbb{E}^{\mathbb{P}}[X_t] = \int X_t d\mathbb{P} = \int X_t \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} d\tilde{\mathbb{P}} = \int X_t e^{-\lambda \tilde{W}_t - \frac{1}{2}\lambda^2 t} d\tilde{\mathbb{P}} = \mathbb{E}^{\tilde{\mathbb{P}}}[X_t e^{-\lambda \tilde{W}_t - \frac{1}{2}\lambda^2 t}]$$

This terminates our short diversion and we continue with our exposition of finding the joint distribution of arithmetic Brownian motion and its maximum process.

Proposition 4.4.2 *Let $X_t = \alpha + \mu t + \sigma W_t$ be an arithmetic $\mathbb{P}$-Brownian motion starting at α with constant drift μ and volatility σ. Let X_t^{Max} be its maximum process. The joint distribution of X_t and X_t^{Max} is given by*

$$\mathbb{P}(X_t \leq x, X_t^{Max} \leq y) = N\left(\frac{x - \alpha - \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}} N\left(\frac{x + \alpha - 2y - \mu t}{\sigma\sqrt{t}}\right) \quad (4.12)$$

[6]Itô's formula can be used to prove this

Proof: We only require the distribution for $x \leq y$ and for $y > \alpha$. We begin by finding $\mathbb{P}(X_t \leq x, X_t^{Max} \geq y)$ and use the same probability trick to find the joint distribution function. We would first like to transform the process $X_t - \alpha = \mu t + \sigma W_t$ into a martingale under some measure. Lets call this new measure $\tilde{\mathbb{P}}$ and solve for its kernel. We can eliminate the drift of $X_t - \alpha$ by setting ν equal to zero. Hence λ (the Girsanov kernel) has solution $-\frac{\mu}{\sigma}$. Its easy to see that $\frac{X_t - \alpha}{\sigma}$ is not only a $\tilde{\mathbb{P}}$-martingale but also a $\tilde{\mathbb{P}}$-Brownian motion.

$$\frac{X_t - \alpha}{\sigma} = \frac{\mu}{\sigma}t + (W_t) = \frac{\mu}{\sigma}t + (\tilde{W}_t - \frac{\mu}{\sigma}t) = \tilde{W}_t$$

We know that to change measure from $\mathbb{P}$ to $\tilde{\mathbb{P}}$ and to change from $\tilde{\mathbb{P}}$ to $\mathbb{P}$ we need to know their Radon-Nikodym derivatives. They are given by

$$\begin{aligned} \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} &= e^{-\frac{\mu}{\sigma}W_t - \frac{1}{2}(-\frac{\mu}{\sigma})^2 t} \\ \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} &= e^{\frac{\mu}{\sigma}\tilde{W}_t - \frac{1}{2}(\frac{\mu}{\sigma})^2 t} \end{aligned}$$

Now since the process $\left(\frac{X_t - \alpha}{\sigma}\right)_t$ is a $\tilde{\mathbb{P}}$-Brownian motion, we can make use of Proposition 4.4.1. We first set $\alpha = 0$ and $\sigma = 1$. Then $X_t = \tilde{W}_t$ is $\tilde{\mathbb{P}}$-Brownian motion. Let $A = \{\omega \in \Omega : \tilde{W}_t \leq x, \tilde{W}_t^{Max} \geq y\}$. We need to determine $\mathbb{P}(A)$ and not $\tilde{\mathbb{P}}(A)$. So let $\tau = \{\omega \in \Omega : \tilde{W}_t \geq y\}$ be a stopping time. Define a new process

$$\widehat{W}_t = \tilde{W}_t 1_{\{\tau < t\}} + (2y - \tilde{W}_t) 1_{\{\tau \geq t\}}$$

which is the reflected Brownian motion discussed previously. We have already shown that $\widehat{W}_t$ is a $\tilde{\mathbb{P}}$-Brownian motion. Let $B = \{\omega \in \Omega : \widehat{W}_t \leq x, \widehat{W}_t^{Max} \geq y\}$. Since $\widehat{W}_t$ is the reflected Brownian motion of $\tilde{W}_t$ then B occurs iff $\{\omega \in \Omega : \tilde{W}_t \geq 2y - x\}$ occurs. It follows that

$$\mathbb{E}^{\tilde{\mathbb{P}}}[1_A e^{\mu\tilde{W}_t - \frac{1}{2}\mu^2 t}] = \mathbb{E}^{\tilde{\mathbb{P}}}[1_B e^{\mu\widehat{W}_t - \frac{1}{2}\mu^2 t}] \tag{4.13}$$

This is because both $\tilde{W}_t$ and $\widehat{W}_t$ are $\tilde{\mathbb{P}}$-Brownian motions and hence identically distributed. We can know find $\mathbb{P}(A)$ as follows:

$$\mathbb{P}(A) = \int_A d\mathbb{P} = \int_A \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} d\tilde{\mathbb{P}} = \int_A e^{\mu\tilde{W}_t - \frac{1}{2}\mu^2 t} d\tilde{\mathbb{P}} = \mathbb{E}^{\tilde{\mathbb{P}}}[1_A e^{\mu\tilde{W}_t - \frac{1}{2}\mu^2 t}]$$

It follows by making use of equation (4.13) and using the definition of $\widehat{W}_t$ that

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{E}^{\tilde{\mathbb{P}}}[1_A e^{\mu\tilde{W}_t - \frac{1}{2}\mu^2 t}] = \mathbb{E}^{\tilde{\mathbb{P}}}[1_B e^{\mu\widehat{W}_t - \frac{1}{2}\mu^2 t}] = \mathbb{E}^{\tilde{\mathbb{P}}}[1_B e^{\mu(2y - \tilde{W}_t) - \frac{1}{2}\mu^2 t}] \\ &= e^{2\mu y}\mathbb{E}^{\tilde{\mathbb{P}}}[1_{\{\tilde{W}_t \geq 2y - x\}} e^{-\mu\tilde{W}_t - \frac{1}{2}\mu^2 t}] \end{aligned}$$

Now, the factor of $e^{-\mu\tilde{W}_t-\frac{1}{2}\mu^2 t}$ in the expectation above is the Radon-Nikodym derivative of some measure, lets call it $\overline{\mathbb{P}}$, i.e. $\frac{d\overline{\mathbb{P}}}{d\tilde{\mathbb{P}}} = e^{-\mu\tilde{W}_t-\frac{1}{2}\mu^2 t}$ from which it's easy to read off its Girsanov kernel given by $-\mu$. It follows that

$$\overline{W}_t = \tilde{W}_t + \mu t$$

i.e. $\overline{W}_t$ has drift of μ under the $\overline{\mathbb{P}}$ measure. It follows that

$$\begin{aligned}\mathbb{P}(A) = e^{2\mu y}\tilde{\mathbb{E}}^{\mathbb{P}}[1_{\{\tilde{W}_t\geq 2y-x\}}e^{-\mu\tilde{W}_t-\frac{1}{2}\mu^2 t}] &= e^{2\mu y}\overline{\mathbb{E}}^{\overline{\mathbb{P}}}[1_{\overline{W}_t-\mu t\geq 2y-x}] \\ &= e^{2\mu y}\overline{\mathbb{P}}(\overline{W}_t \geq 2y-x+\mu t) \\ &= e^{2\mu y}N\left(\frac{x-2y-\mu t}{\sqrt{t}}\right)\end{aligned}$$

i.e. If $X_t = \mu t + W_t$ then

$$\mathbb{P}(X_t \leq x, X_t^{Max} \geq y) = e^{2\mu y}N\left(\frac{x-2y-\mu t}{\sqrt{t}}\right) \tag{4.14}$$

All that's left is to generalize the above to arbitrary α, σ so that $X_t = \alpha+\mu t+\sigma W_t$ and lastly, transform into a distribution function. We first suppose that $\sigma \neq 1$ and $\alpha = 0$ so that $\frac{X_t}{\sigma}$ is an arithmetic $\mathbb{P}$-Brownian motion with drift $\frac{\mu}{\sigma}$ and volatility 1. Using equation (4.14) we have

$$\begin{aligned}\mathbb{P}(X_t \leq x, X_t^{Max} \geq y) &= \mathbb{P}\left(\frac{X_t}{\sigma} \leq \frac{x}{\sigma}, \frac{X_t^{Max}}{\sigma} \geq \frac{y}{\sigma}\right) \\ &= e^{\frac{2\mu y}{\sigma^2}}N\left(\frac{x-2y-\mu t}{\sigma\sqrt{t}}\right)\end{aligned}$$

Now let $\alpha \neq 0$. Note that $X_t - \alpha$ is an arithmetic $\mathbb{P}$-Brownian motion starting at zero and therefore

$$\begin{aligned}\mathbb{P}(X_t \leq x, X_t^{Max} \geq y) &= \mathbb{P}(X_t-\alpha \leq x-\alpha, (X-\alpha)_t^{Max} \geq (y-\alpha)) \\ &= e^{\frac{2\mu(y-\alpha)}{\sigma^2}}N\left(\frac{x+\alpha-2y-\mu t}{\sigma\sqrt{t}}\right)\end{aligned}$$

And finally, using the fact that $\mathbb{P}(A\cap B) = \mathbb{P}(A) - \mathbb{P}(A\cap\overline{B})$ we have proved Proposition 4.4.2, i.e.

$$\mathbb{P}(X_t \leq x, X_t^{Max} \leq y) = N\left(\frac{x-\alpha-\mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}}N\left(\frac{x+\alpha-2y-\mu t}{\sigma\sqrt{t}}\right)$$

As a corollary to the above result we know that $\{\omega\in\Omega : X_t^{Max} \leq x\} \subseteq \{\omega\in\Omega : X_t \leq x\}$. It follows that

Corollary 4.4.2

$$\begin{aligned}\mathbb{P}(X_t^{Max} \leq x) &= \mathbb{P}(X_t \leq x, X_t^{Max} \leq x) \\ &= N\left(\frac{x-\alpha-\mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu(x-\alpha)}{\sigma^2}} N\left(\frac{-x+\alpha-\mu t}{\sigma\sqrt{t}}\right) \quad (4.15)\end{aligned}$$

which is the distribution function of the maximum process of an arithmetic Brownian motion which starts at a point α.

□

As we shall see in the sequel, the pricing of barrier options can be simplified to taking risk-neutral expectations of stopped processes. Recall that if τ is a stopping time and X_t is a stochastic process then the process $X_t^\tau = X_{\tau\wedge t}$ is known as the stopped process of X_t, i.e. $X_{\tau\wedge t} = X_t$ on the set $\{\omega \in \Omega : \tau > t\}$ and $X_{\tau\wedge t} = \tau$ on the set $\{\omega \in \Omega : \tau \leq t\}$. We attempt to find the distribution of $X_t^{\tau^y}$ and $X_t^{\tau_y}$ where X_t is an arithmetic Brownian motion starting at α with $\tau^y = \inf\{t : X_t \geq y\}$ and $\tau_y = \inf\{t : X_t \leq y\}$. We are actually more concerned with the densities as opposed to the distribution functions. We consider the two cases separately.

Proposition 4.4.3 *Let $X_t = \alpha + \mu t + \sigma W_t$ be an arithmetic $\mathbb{P}$-Brownian motion starting at α with constant drift μ and volatility σ. Let $\tau^y = \inf\{t : X_t \geq y\}$ then for $y > \alpha$ the density $f_{X_t^{\tau^y}}(x)$ of $X_t^{\tau^y}$ evaluated at x is given by*

$$\begin{aligned} f_{X_t^{\tau^y}}(x) = \phi(x, \alpha+\mu t, \sigma^2 t) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}} \phi(x, 2y-\alpha+\mu t, \sigma^2 t) \quad &for \quad x < y \\ 0 \quad &for \quad x \geq y \end{aligned}$$

where $\phi(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Proof: We first find the distribution functions and then differentiate to find the density. Noting that for $x < y$, $\{\omega \in \Omega : X_t^{\tau^y} \leq x\} = \{\omega \in \Omega : X_t \leq x, X_t^{Max} < y\}$ and for $x \geq y$, $\{\omega \in \Omega : X_t^{\tau^y} \leq x\} = \Omega$. It follows that for $x < y$

$$\begin{aligned}\mathbb{P}(X_t^{\tau^y} \leq x) &= \mathbb{P}(X_t \leq x, X_t^{Max} < y) \\ &= N\left(\frac{x-\alpha-\mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}} N\left(\frac{x+\alpha-2y-\mu t}{\sigma\sqrt{t}}\right)\end{aligned}$$

The density function can be obtained by simply taking the first derivative with respect to x.

□

Proposition 4.4.4 *Let $X_t = \alpha + \mu t + \sigma W_t$ be an arithmetic $\mathbb{P}$-Brownian motion starting at α with constant drift μ and volatility σ. Let $\tau_y = \inf\{t : X_t \leq y\}$ then for $y < \alpha$ the density $f_{X_t^{\tau_y}}(x)$ of $X_t^{\tau_y}$ evaluated at x is given by*

$$f_{X_t^{\tau_y}}(x) = \begin{cases} \phi(x, \alpha + \mu t, \sigma^2 t) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}} \phi(x, 2y - \alpha + \mu t, \sigma^2 t) & \text{for } x > y \\ 0 & \text{for } x \leq y \end{cases}$$

where $\phi(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Proof: First noting that for $x > y$, $\{\omega \in \Omega : X_t^{\tau_y} \leq x\} = \{\omega \in \Omega : X_t \leq x, X_t^{Min} > y\}$ and for $x < y$, $\{\omega \in \Omega : X_t^{\tau_y} \leq x\} = \emptyset$. Note that the minimum process of X_t can also be written in terms of its maximum process. It's easy to see that $X_t^{Min} = -(-X)_t^{Max}$. For $x > y$ we have,

$$\begin{aligned} \mathbb{P}(X_t^{\tau_y} \leq x) = \mathbb{P}(X_t \leq x, X_t^{Min} > y) &= \mathbb{P}(-X_t \geq -x, -(-X)_t^{Max} > y) \\ &= \mathbb{P}(Y_t \geq -x, Y_t^{Max} < -y) \end{aligned}$$

where $Y_t = (-X)_t$. Using the fact that $\mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(\overline{A} \cap B)$ and the fact that the density is given by the first derivative of the distribution function with respect to x, we have

$$\begin{aligned} f_{X_t^{\tau_y}}(x) = \frac{d}{dx}\mathbb{P}(X_t^{\tau_y} \leq x) &= \frac{d}{dx}\mathbb{P}(Y_t \geq -x, Y_t^{Max} < -y) \\ &= \frac{d}{dx}\mathbb{P}(Y_t^{Max} < -y) - \frac{d}{dx}\mathbb{P}(Y_t < -x, Y_t^{Max} < -y) \\ &= -\frac{d}{dx}\mathbb{P}(Y_t < -x, Y_t^{Max} < -y) \end{aligned}$$

We can now make use of Proposition 4.4.2 to solve the above. The process Y_t is an arithmetic $\mathbb{P}$-Brownian motion starting at $-\alpha$ with drift $-\mu$ and volatility $-\sigma$. It follows that

$$\begin{aligned} f_{X_t^{\tau_y}}(x) &= -\frac{d}{dx}\mathbb{P}(Y_t < -x, Y_t^{Max} < -y) \\ &= -\frac{d}{dx}\left\{N\left(-\frac{x - \alpha - \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu(y-\alpha)}{\sigma^2}} N\left(-\frac{x + \alpha - 2y - \mu t}{\sigma\sqrt{t}}\right)\right\} \\ &= f_{X_t^{\tau^y}}(x) \end{aligned}$$

and the proposition is proved.

□

Both $X_t^{\tau^y}$ and $X_t^{\tau_y}$ have the same densities but with different support on the real line. The former having a non-zero density on $(-\infty, y)$ and the latter on (y, ∞). Note that both $f_{X_t^{\tau^y}}(x)$ and $f_{X_t^{\tau_y}}(x)$ are not probability densities and hence do not integrate to 1. This is because both $X_t^{\tau^y}$ and $X_t^{\tau^y}$ have atoms at y.

4.4.2 Reducing Barrier Options to European-Style Derivatives

The objective of this section is to show that if we can price a particular European option (which only concerns the payoff at expiry) then it's possible to price any barrier version of that option. These results are due to Bjork (2004). For example, consider an up-and-out call option with strike K on an underlying S_t with expiry T. Let the barrier level be $L > K > S_0$. This option pays out $S_T - K$ if S_T is above K and if S_t has never breached the barrier L, $\forall t \in [0, T]$. Otherwise it has zero payoff, i.e.

$$\Phi(S_T) = \max(S_T - K, 0) 1_{S_T^{Max} < L}$$

Our intention is to show that it's only necessary to price the option with payoff

$$\Phi^\star(S_T) = \max(S_T - K, 0) 1_{S_T < L}$$

which is considerably simpler since $\Phi^\star(S_T)$ is path independent. Even if no closed form solution exists for the price of the option with payoff $\Phi^\star(S_T)$ then the efficiency of a Monte Carlo approach is considerably enhanced since its now only necessary to simulate a large number of log-normal random variables as opposed to simulating a large number of Brownian paths. Monte Carlo pricing of barrier options always involves discretization error which can otherwise be avoided. For instance, certain discretized paths may indicate that a barrier has not been breached whereas the continuous version may have breached the barrier. This discretization error leads to over-pricing of barrier options.

This simplification is certainly true in a Black-Scholes world where S_t is driven by geometric Brownian motion. It may cease to exist if alternate dynamics are specified such as a pure jump process.

Definition 4.4.2 *Let $\Phi(S_T)$ denote the payoff function on some underlying S. Define the truncated payoff of Φ as follows:*

$$\Phi^L(S_T) = \Phi(S_T) 1_{S_T < L} \qquad \Phi_L(S_T) = \Phi(S_T) 1_{S_T > L}$$

We begin by distinguishing between four types of barrier options. We denote the time zero prices of an up-and-out, up-and-in, down-and-out and down-and-in option with barrier L by F^{LO}, F^{LI}, F_{LO} and F_{LI} respectively. The O and I indicate whether it's an *out* or *in* contract. The superscript and subscript indicate whether the barrier L is above or below the time zero price of the underlying.

Definition 4.4.3 *Let $\Phi(S_T)$ denote the payoff function of an underlying security S with a time zero price of S_0. Let $\Phi^L(S_T)$ and $\Phi_L(S_T)$ denote their truncated payoffs. Let the time zero prices of these derivatives be respectively given by*

$$F(S_0, T, \Phi^L(S_T)) \qquad F(S_0, T, \Phi_L(S_T))$$

Proposition 4.4.5 *Let $\Phi(S_T)$ be the payoff function on an underlying security S. Let $L > S_0$ be the barrier level. The price of an up-and-out barrier option is given by*

$$F^{LO}(S_0, T, \Phi(S_T)) = F(S_0, T, \Phi^L(S_T)) - \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2}-1} F\left(\frac{L^2}{S_0}, T, \Phi^L(S_T)\right)$$

where r is the constant risk-free rate of interest.

Proof: By risk-neutral valuation

$$F^{LO}(S_0, T, \Phi(S_T)) = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)1_{S_T^{Max}<L}] \tag{4.16}$$

where $\mathbb{Q}$ is the unique risk-neutral measure which transforms $S_t e^{rt}$ into a $\mathbb{Q}$-martingale, i.e. S has $\mathbb{Q}$-dynamics

$$dS = rSdt + \sigma SdW \tag{4.17}$$

Now let $\tau^L = \inf\{t : S_t \geq L\}$ and define $S_t^{\tau^L}$ to be its associated stopped process. It's easy to verify that $\Phi^L(S_T^{\tau^L}) = \Phi(S_T^{\tau^L})1_{S_T^{\tau^L}<L}$ has the same payoff as $\Phi(S_T)1_{S_T^{Max}<L}$. By the Law of One Price, they must have the same time zero prices. We can therefore rewrite equation (4.16) as

$$F^{LO}(S_0, T, \Phi(S_T)) = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi^L(S_T^{\tau^L})]$$

Now let $X_t = \ln S_t = \ln S_0 + (r - \frac{1}{2}\sigma^2)t + \sigma W_t$ be an arithmetic $\mathbb{Q}$-Brownian motion. This can easily be verified by letting $X_t = \ln S_t$ and applying Itô's formula. Now since $\{\omega \in \Omega : S_t > L\} = \{\omega \in \Omega : \ln S_t > \ln L\}$ we can rewrite

$$\tau^L = \inf\{t : S_t \geq L\} = \inf\{t : \ln S_t \geq \ln L\} = \inf\{t : X_t \geq \ln L\}$$

Now X_t is an arithmetic $\mathbb{Q}$-Brownian motion and so we know the density of $X_T^{\tau^L}$ using Proposition 4.4.3. Letting $f_{S_T^{\tau^L}}(x)$ and $f_{X_T^{\tau^L}}(x)$ denote the densities evaluated at x of the random variables $S_T^{\tau^L}$ and $X_T^{\tau^L}$ we have

$$\begin{aligned} F^{LO}(S_0, T, \Phi(S_T)) = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi^L(S_T^{\tau^L})] &= e^{-rT}\int_0^\infty \Phi^L(x) f_{S_T^{\tau^L}}(x)dx \\ &= e^{-rT}\int_{-\infty}^\infty \Phi^L(e^y) e^y f_{S_T^{\tau^L}}(e^y)dy \\ &= e^{-rT}\int_{-\infty}^\infty \Phi^L(e^y) f_{X_T^{\tau^L}}(y)dy \end{aligned}$$

The last line is obtained using the result that $e^x f_{S_T^{\tau^L}}(e^x) = f_{X_T^{\tau^L}}(x)$. This can be verified by noting that

$$e^x f_{S_T^{\tau^L}}(e^x) = \frac{d}{dx}\mathbb{Q}(S_T^{\tau^L} \leq e^x) = \frac{d}{dx}\mathbb{Q}(\ln S_T^{\tau^L} \leq x) = f_{X_T^{\tau^L}}(x)$$

By making use of Proposition 4.4.3 with $\alpha = \ln S_0$, $t = T$, $\mu = r - \frac{1}{2}\sigma^2$ and $y = \ln L$, we can write the density of $X_T^{\tau^L}$ evaluated at x as

$$f_{X_T^{\tau^L}}(x) = \phi(x, \ln S_0 + (r - \frac{1}{2}\sigma^2)T, \sigma^2 T) - \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2}-1} \phi(x, \ln \frac{L^2}{S_0} + (r - \frac{1}{2}\sigma^2)T, \sigma^2 T)$$

for $x < \ln L$ otherwise $f_{X_T^{\tau^L}}(x) = 0$. The density decomposes into two terms. The first ϕ, call it ϕ_1, is the density of $\ln S_T$ with initial price of $\ln S_0$ and the second ϕ, call it ϕ_2, is the density of $\ln S_T$ with initial price $\ln \frac{L^2}{S_0}$. Letting $p = \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2}-1}$ we have

$$\begin{aligned} F^{LO}(S_0, T, \Phi(S_T)) &= e^{-rT} \int_{-\infty}^{\infty} \Phi^L(e^y)\phi_1(y)dy - e^{-rT} p \int_{-\infty}^{\infty} \Phi(e^y)\phi_2(y)dy \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi^L(S_T)|S_0 = S_0] - e^{-rT} p \mathbb{E}^{\mathbb{Q}}\left[\Phi^L(S_T)|S_0 = \frac{L^2}{S_0}\right] \\ &= F(S_0, T, \Phi^L(S_T)) - \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2}-1} F\left(\frac{L^2}{S_0}, T, \Phi^L(S_T)\right) \end{aligned}$$

and the proposition is proved.

□

Having priced the up-and-out option we automatically get the up-and-in for free. This follows from in-out parity. Holding an up-and-out option and an up-and-in option is equivalent to holding an ordinary European option with payoff $\Phi(S_T)$. Applying the Law of One Price,

$$F^{LO}(S_0, T, \Phi(S_T)) + F^{LI}(S_0, T, \Phi(S_T)) = F(S_0, T, \Phi(S_T))$$

The only barrier options we have left to price are the down-and-out and down-and-in options.

Proposition 4.4.6 *Let $\Phi(S_T)$ be the payoff function on an underlying security S. Let $L < S_0$ be the barrier level. The price of a down-and-out option is given by*

$$F_{LO}(S_0, T, \Phi(S_T)) = F(S_0, T, \Phi_L(S_T)) - \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2}-1} F\left(\frac{L^2}{S_0}, T, \Phi_L(S_T)\right)$$

Proof: The proof is virtually identical to the previous one. The crux is identifying that $\Phi(S_T)1_{S_T^{Min} > L}$ has the same payoff as $\Phi_L(S_T^{\tau_L})$ and proceed as before. Again, we get the down-and-in price for free because of in-out parity, i.e.

$$F_{LO}(S_0, T, \Phi(S_T)) + F_{LI}(S_0, T, \Phi(S_T)) = F(S_0, T, \Phi(S_T))$$

4.4.3 Examples of Barrier Options

In this section we derive analytical solutions to three types of Barrier options. We first consider an up-and-out call, a lookback put[7] and a knock-out double barrier call option. The final chapter investigates the comparison between these 'Black-Scholes' prices and prices obtained using a pure jump process, namely the variance gamma process.

Up-and-out Call

An up-and-out call is identical to a usual call option except that it has a zero payoff if the underlying process breaches a barrier $L > S_0$ during the life of the option. Sticking with same notation, its payoff is given by

$$\Phi(S_T) = \max(S_T - K, 0)1_{S_T^{Max} < L}$$

We only consider the case where $L > K$ else the call option never expires in-the-money and hence has zero value. Now for $L > K$, the truncated payoff is given by

$$\Phi^L(S_T) = \max(S_T - K, 0)1_{S_T < L}$$

Once we have valued the above payoff, we can make use of Proposition 4.4.5 to price the up-and-out call. The truncated payoff can easily be replicated by constructing the following portfolio:

- Long a vanilla call with strike K, expiry T
- Short a vanilla call with strike L, expiry T
- Short $L - K$ digital calls with strike L, expiry T

It's a trivial exercise to show that this portfolio perfectly replicates the truncated payoff above and by the Law of One Price, they must have the same time zero prices. Using standard results from option pricing theory in a Black-Scholes world, the price of this option with the truncated payoff is given by

$$F(S_0, T, \Phi^L(S_T)) = S_0(N(d_1^K) - N(d_1^L)) + e^{-rT}K(N(d_2^L) - N(d_2^K))$$

where $d_i^j = \frac{\ln S_0 - \ln j + (r - (-1)^i 0.5\sigma^2)T}{\sigma\sqrt{T}}$ for $j = K, L$. One can now easily solve for $F(\frac{L^2}{S_0}, T, \Phi^L(S_T))$ by making a direct substitution of $\frac{L^2}{S_0}$ for S_0. The price of an up-and-out call $F^{L0}(S_0, T, \Phi(S_T))$ is given by Proposition 4.4.5.

Lookback Put

A lookback put allows the holder at expiry to obtain the difference between the

[7]Strictly speaking, this is not a barrier option but may be referred to as one

realized maximum of the underlying S and its price at expiry. Its payoff is given by

$$\Phi(S_T) = \max_{t\in[0,T]} S_t - S_T = \sup_{t\leq T} S_t - S_T$$

It's clear that $\Phi(S_T) \geq 0$. By risk-neutral valuation its price, denoted by $F^{LookPut}(S_0, T)$, is given by

$$F^{LookPut}(S_0,T) = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\sup_{t\leq T} S_t - S_T\right] = e^{-rT}\mathbb{E}^{\mathbb{Q}}[S_T^{Max}] - e^{-rT}\mathbb{E}^{\mathbb{Q}}[S_T]$$

The second expectation is easy to compute since it is, by construction of $\mathbb{Q}$, equal to S_0. Now equation (4.15) gives us the distribution function of X_T^{Max} where X is an arithmetic Brownian motion which starts at α. The density can easily be found by straightforward differentiation. Its price is given by

$$\begin{aligned} F^{LookPut}(S_0,T) &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[S_T^{Max}] - S_0 \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[e^{X_T^{Max}}] - S_0 \\ &= e^{-rT}\int_{\ln S_0}^{\infty} e^x f_{X_T^{Max}}(x)dx - S_0 \end{aligned}$$

where $f_{X_T^{Max}}$ is the density of X_T^{Max} with $X_T = \ln S_0 + (r - 0.5\sigma^2)T + \sigma W_T$. $(W_t)_t$ is a $\mathbb{Q}$-Brownian motion. One can now make use of equation (4.15) by finding the density and setting $\alpha = \ln S_0$, $\mu = r - 0.5\sigma^2$ and $t = T$. Perform the integration (by parts) and you'll end up with the following price for a lookback put.

$$F^{LookPut}(S_0,T) = -S_0N(-d_1) + S_0e^{-rT}N(-d_2) + S_0\frac{\sigma^2}{2r}N(d_1) - S_0e^{-rT}\frac{\sigma^2}{2r}N(-d_2)$$

where $d_1 = \frac{(r+0.5\sigma^2)T}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \sigma\sqrt{T}$

Double Barrier Options

Double barrier options, like the name suggests, are options which either get *knocked-in* or *knocked-out* as soon as one of the two barriers is breached. This is in contrast to previous barrier options we've encounted which had either an upper barrier or lower barrier but not both. Let the upper barrier be denoted by A and the lower barrier by B where $B < S_0 < A$. Suppose we wish to price a knock-out double barrier option with payoff $\Phi(S_T)$. By risk-neutral valuation, this option has a time zero price of

$$\begin{aligned} V_0 &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)1_{(S_T^{Max}<A)\cap(S_T^{Min}>B)}] \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(e^{\ln S_T})1_{((\ln S)_T^{Max}<\ln A)\cap((\ln S)_T^{Min}>\ln B)}] \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T})1_{(X_T^{Max}<\ln A)\cap(X_T^{Min}>\ln B)}] \end{aligned}$$

Using a slightly different notation,

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T}); X_T^{Max} < A^* \cap X_T^{Min} > B^*] \tag{4.18}$$

where $A^* = \ln A$, $B^* = \ln B$ and $X_T = \ln S_T = \ln S_0 + (r - 0.5\sigma^2)T + \sigma W_T$ under $\mathbb{Q}$. The semi-colon notation indicating that the expectation is taken over the set which appears on the right-hand side of the semi-colon. Note that $\ln S_T^{Max} = (\ln S)_T^{Max}$. In order to evaluate equation (4.18) we need to be very precise since we are evaluating an expectation involving three dependent random variables. We first need to define three stopping times. They are as follows:

$$\begin{aligned} \tau^{A^*} &= \inf\{t : X_t \geq A^*\} \\ \tau_{B^*} &= \inf\{t : X_t \leq B^*\} \\ \tau &= \tau^{A^*} \wedge \tau_{B^*} \wedge T \end{aligned}$$

We now define $X_T^\tau = X_{\tau^{A^*} \wedge \tau_{B^*} \wedge T} = A^* 1_{\tau = \tau^{A^*}} + B^* 1_{\tau = \tau_{B^*}} + X_T 1_{\tau = T}$. Now X_T^τ under $\mathbb{Q}$[8], is neither a discrete nor a continuous random variable. It has an absolutely continuous component (with respect to Lebesgue measure) and two atoms. Letting μ denote the distribution of X_T^τ, then by the Lebesgue decomposition theorem, we can write $\mu = \mu^{ac} + \mu^{nd}$ where μ^{ac} is the absolutely continuous part of μ and μ^{nd} is the non-diffuse part of μ. Since μ^{ac} is absolutely continuous with respect to Lebesgue measure, it has a density. Let us denote its density by $p(x, T)$. Now, assuming that $\Phi(e^{X_T^\tau})$ is integrable, we can write

$$\mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau})] = \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau}); \tau = T] + \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau}); \tau = \tau^{A^*}] + \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau}); \tau = \tau_{B^*}]$$

Rearranging the subject of the formula and noting that $\{\omega \in \Omega : \tau = T\} = \{\omega \in \Omega : X_T^{Max} < A^* \cap X_T^{Min} > B^*\}$ we have

$$\begin{aligned} V_0 e^{rT} &= \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau})] - \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau}); \tau = \tau^{A^*}] - \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau}); \tau = \tau_{B^*}] \\ &= \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau})] - \Phi(e^{A^*})\mathbb{Q}(X_T^\tau = A^*) - \Phi(e^{B^*})\mathbb{Q}(X_T^\tau = B^*) \end{aligned}$$

Since X_T^τ has distribution $\mu = \mu^{ac} + \mu^{nd}$ we can write the first term as

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\Phi(e^{X_T^\tau})] &= \int_{-\infty}^{\infty} \Phi(e^x) d\mu(x) \\ &= \int_{B^*}^{A^*} \Phi(e^x) d\mu^{ac}(x) + \Phi(e^{A^*})\mathbb{Q}(X_T^\tau = A^*) + \Phi(e^{B^*})\mathbb{Q}(X_T^\tau = B^*) \\ &= \int_{B^*}^{A^*} \Phi(e^x) p(x, T) dx + \Phi(e^{A^*})\mathbb{Q}(X_T^\tau = A^*) + \Phi(e^{B^*})\mathbb{Q}(X_T^\tau = B^*) \end{aligned}$$

We can therefore price a double barrier option using this representation

$$V_0 = e^{-rT} \int_{B^*}^{A^*} \Phi(e^x) p(x, T) dx \tag{4.19}$$

[8]The same measure under which X_t is an arithmetic Brownian motion

Our objective is to find this truncated density $p(x,T)$. I use the word 'truncated' since this density does not integrate to 1. This can be done using a probabilistic argument where the density is not known in closed form but given by a rapidly decreasing infinite sum. Instead of using a probabilistic approach we derive the density using an analytical approach which gives the density in terms of a fast converging Fourier series.

To formalise, let $X_t = X_0 + \mu t + \sigma W_t$ be an arithmetic $\mathbb{P}$-Brownian motion starting at X_0 with absorbing barriers A and B where $B < X_0 < A$. For simplicity, we begin by setting $X_0 = 0$ and adjust the barriers further on in order to obtain a non-zero starting point. Note that $B < 0$ in this particular case. We will henceforth rewrite $p(x,t)$ as $p_0^{A,B}(x,t)$ to remind us that we finding this density on the interval $[B,A]$. We have also included a subscript of 0 to remind us that $X_0 = 0$. We follow the works of Cox & Miller (1965). Since X_t satisfies an Itô diffusion equation, it must satisfy Kolmogorov's forward equation. See Bjork (2004) for a statement and easy to follow proof of Kolmogorov's forward and backward equations. Since μ and σ are constant, Kolmogorov's forward equation simplifies to

$$\frac{1}{2}\sigma^2\frac{\partial^2 p}{\partial x^2} - \mu\frac{\partial p}{\partial x} = \frac{\partial p}{\partial t}$$

Since we are only finding its density over $[-B,A]$ we must have that $x \in [-B,A]$ and $t \in [0,T]$. Now $p_0^{A,B}(x,t)$ is a truncated density which we require to be zero whenever $x = -B$ or whenever $x = A$ for any $t \in [0,T]$. These form our boundary conditions. We also have an initial condition which states that the density at the current time zero must be equal to the Dirac Delta function located at zero, i.e.

$$p_0^{A,B}(x,0) = \delta_0$$

We therefore need to solve the following boundary value problem:

$$\frac{1}{2}\sigma^2\frac{\partial^2 p}{\partial x^2} - \mu\frac{\partial p}{\partial x} = \frac{\partial p}{\partial t} \tag{4.20}$$

for $x \in [B,A]$ subject to

$$p_0^{A,B}(x,0) = \delta_0 \qquad p_0^{A,B}(A,t) = 0 \qquad p_0^{A,B}(B,t) = 0 \quad \forall t \in [0,T]$$

Cox & Miller (1965) propose a solution of the form

$$p_0^{A,B}(x,t) = e^{\kappa x - \lambda t}\sin\left\{\frac{n\pi(x+B)}{A+B}\right\} \tag{4.21}$$

where $n \in \mathbb{N}$ and κ, λ are, as yet, undetermined constants. The sine function was chosen since $p_0^{A,B}(x,t)$ vanishes if $x = A, -B$ and hence our boundary conditions

are satisfied. The exponential function always seems to crop up in solutions to differential equations since it has some desirable differentiability properties. Now we will choose κ and λ such that equation (4.21) satisfies equation (4.20). This can easily be solved by straightforward differentiation and solutions to κ and λ are given by

$$\kappa = \frac{\mu}{\sigma^2} \qquad \lambda = \lambda_n = \frac{1}{2}\left(\frac{\mu^2}{\sigma^2} + \frac{n^2\pi^2\sigma^2}{(A+B)^2}\right) \tag{4.22}$$

Our last concern is to manipulate equation (4.21) so that it satisfies our initial condition. Since equation (4.21) is a solution to equation (4.20) then so is any linear combination. We have,

$$p_0^{A,B}(x,t) = e^{\frac{\mu x}{\sigma^2}} \sum_{n=1}^{\infty} a_n e^{-\lambda_n t} \sin\left(\frac{n\pi(x+B)}{A+B}\right) \tag{4.23}$$

for any constants a_n. These constants give us an extra degree of freedom which allows us to solve for them in order to satisfy our initial condition, $p_0^{A,B}(x,0) = \delta_0$. Substituting $t = 0$ in equation (4.23) we have

$$\delta_0 e^{-\frac{\mu^2 x}{\sigma^2}} = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi(x+B)}{A+B}\right) \tag{4.24}$$

where the constants a_n can be solved using Fourier series. It's an easy exercise in integration to show that $\sin(\frac{k\pi(x+B)}{A+B})_{k\in\mathbb{N}}$ forms an orthogonal basis on the interval $[-B, A]$, i.e.

$$\int_{-B}^{A} \sin\left(\frac{n\pi(x+B)}{A+B}\right) \sin\left(\frac{k\pi(x+B)}{A+B}\right) dx = \frac{A+B}{2} 1_{k=n}$$

For ease of notation, we let $g(x) = \sin\left(\frac{k\pi(x+B)}{A+B}\right)$. Multiplying both sides of equation (4.24) by $g(x)$ and integrating over $[-B, A]$ we have

$$\begin{aligned} \int_{-B}^{A} \delta_0 e^{-\frac{\mu^2 x}{\sigma^2}} g(x)dx &= \int_{-B}^{A} \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi(x+B)}{A+B}\right) g(x)dx \\ g(0) &= a_k \frac{A+B}{2} \\ a_k &= \frac{2}{A+B} \sin\left(\frac{k\pi B}{A+B}\right) \end{aligned}$$

We have thus solved for the k^{th} coefficient a_k. Simply substitute this result into equation (4.23) and we have found $p_0^{A,B}(x,t)$ which satisfies the Kolmogorov

forward equation, the two boundary conditions and the initial condition. The solution is given by

$$p_0^{A,B}(x,t) = \frac{2e^{\frac{\mu x}{\sigma^2}}}{A+B}\sum_{n=1}^{\infty} e^{-\lambda_n t}\sin\left(\frac{n\pi B}{A+B}\right)\sin\left(\frac{n\pi(x+B)}{A+B}\right) \tag{4.25}$$

where the constants λ_n are given by equation (4.22). The final step is to generalize the above transition density to an arbitrary starting point $X_0 = x_0$ where $x_0 > 0$. This is can be done by simply shifting the upper and lower barriers. It's easy to verify (using a diagram) that $p_{x_0}^{A,B}(x,t) = p_0^{A-x_0,x_0-B}(x-x_0,t)$. Recall, by definition, that $p_{x_0}^{A,B}(x,t)$ is the density we require. It's given by

$$\begin{aligned} p_{x_0}^{A,B}(x,t) &= \frac{2e^{\frac{\mu(x-x_0)}{\sigma^2}}}{A-B}\sum_{n=1}^{\infty} e^{-\lambda_n t}\sin\left(\frac{n\pi(x_0-B)}{A-B}\right)\sin\left(\frac{n\pi(x-B)}{A-B}\right) \\ \lambda_n &= \frac{1}{2}\left(\frac{\mu^2}{\sigma^2}+\frac{n^2\pi^2\sigma^2}{(A-B)^2}\right) \end{aligned} \tag{4.26}$$

The above density converges very rapidly since λ_n goes to ∞ at a quadratic rate and hence $e^{-\lambda_n}$ goes to zero at an even faster rate. As few as ten terms are needed in the sum in order to obtain double precision.

We can use equation (4.26) to price a European double barrier option. Suppose we have a European option with payoff function $\Phi(S_T)$ on an underlying S_t which follows geometric $\mathbb{P}$-Brownian motion with drift μ and volatility σ. We consider a double barrier version of this option which has zero payoff if S_t has breached either A or B during its entire life from 0 to T. Clearly, we must have $B < S_0 < A$ otherwise the option will always have zero value. Making use of equation (4.19), this option has value

$$V_0 = e^{-rT}\int_{B^*}^{A^*} \Phi(e^x) p_{\ln S_0}^{A^*,B^*}(x,T)dx$$

where $p_{\ln S_0}^{A^*,B^*}(x,T)$ is given by

$$\begin{aligned} p_{\ln S_0}^{A^*,B^*}(x,T) &= \frac{2e^{\frac{(r-0.5\sigma^2)(x-\ln S_0)}{\sigma^2}}}{A^*-B^*}\sum_{n=1}^{\infty} e^{-\lambda_n T}\sin\left(\frac{n\pi(\ln S_0 - B^*)}{A^*-B^*}\right)\sin\left(\frac{n\pi(x-B^*)}{A^*-B^*}\right) \\ \lambda_n &= \frac{1}{2}\left(\frac{(r-0.5\sigma^2)^2}{\sigma^2}+\frac{n^2\pi^2\sigma^2}{(A^*-B^*)^2}\right) \end{aligned}$$

We will only consider pricing a *knock-out* double barrier call option so in this case, $\Phi(x) = \max(x-K,0)$, where K is the agreed upon strike price. Its price,

in a Black-Scholes world, is given by

$$\begin{aligned} V_0 &= e^{-rT}\int_{B^*}^{A^*}\Phi(e^x)p^{A^*,B^*}_{\ln S_0}(x,T)dx \\ &= e^{-rT}\int_{B^*}^{A^*}(\max(e^x-K,0)p^{A^*,B^*}_{\ln S_0}(x,T)dx \\ &= e^{-rT}\int_{\ln K}^{A^*}(e^x-K)p^{A^*,B^*}_{\ln S_0}(x,T)dx \end{aligned} \tag{4.27}$$

The above is easy and very efficient to calculate. In fact, it's possible to find an antiderivative and hence a closed form exists for a desired level of precision.

We have given derivations and closed form expressions for the prices of an up-and-out call, a lookback put and a knock-out double barrier call option. All of which assumed that the underlying was driven by geometric Brownian motion. The next section is devoted to pricing the same three options but assuming dynamics given by a variance gamma process. A comparison is discussed in the final chapter.

□

4.5 Pricing Barrier Options under Variance Gamma Dynamics

The pricing of barrier options in a Lévy world is considerably more complex than pricing in a Black-Scholes world for a number of reasons. The reflection principle of Brownian motion does not exist for an asymmetric Lévy process with a non-trivial jump component. In addition, since most Lévy processes have discontinuous sample paths, it's possible for the process to cross the barrier without hitting it.

4.5.1 Methods of Pricing Barrier Options

There are essentially three methods for pricing barrier options (more generally, path dependent options) in a Lévy world. They are listed as follows:

- Wiener-Hopf factorization identities
- Partial integro-differential equations
- Monte Carlo methods

of computing time. PC's have limited capabilities in this regard and may take several hours to run say 10 million iterations. Secondly, and more importantly, software packages have a finite cycle length of the number of pseudo-random variates it can produce[9]. Once a length of a cycle is complete, the pseudo-random variates simply repeat themselves! For these reasons, so-called variance reduction techniques have been created to reduce the standard error of a simulation by reducing $\sigma_{g(X)}$ as opposed to simply increasing n. The following is a list of the most commonly used variance reduction techniques:

- Control Variates
- Stratification
- Antithetic Variates
- Importance Sampling

The methods can be used simultaneously provided the specific problem at hand lends itself to its use. The type of simulations performed in our analysis lends itself to the first two techniques. The method of Control variates is briefly discussed below.

Control Variates

This is potentially the most powerful variance reduction technique and the idea is fairly simple. Suppose we have a random variable X with known distribution function F and we require to find $\mathbb{E}[g(X)]$. We assume that we know how to simulate $g(X)$. Suppose there exists another random variable Y with **known** mean $\mathbb{E}[Y]$. Set

$$Z = g(X) - b(Y - \mathbb{E}[Y])$$

for some $b \in \mathbb{R}$. It's easy to show that $\mathbb{E}[Z] = \mathbb{E}[g(X)]$ which implies that if (X_i, Y_i), $i = 1, \ldots, n$, is an i.i.d. sample from (X, Y) then for arbitrary dependence

$$\hat{\mathbb{E}}[g(X)] = \frac{1}{n}\sum_{i=1}^{n} g(X_i) - b(Y_i - \mathbb{E}[Y])$$

is an unbiased estimator of $\mathbb{E}[g(X)]$. Most importantly,

$$Var(Z) = Var(g(X) - b(Y - \mathbb{E}[Y])) = \sigma^2_{g(X)} - 2b\sigma_{g(X)}\sigma_Y\rho_{g(X)Y} + b^2\sigma^2_Y$$

where $\rho_{g(X)Y}$ denotes the correlation between $g(X)$ and Y. Now choosing $b = \frac{\sigma_{g(X)}}{\sigma_Y}\rho_{g(X)Y} = b^*$ produces the smallest possible $Var(Z)$ for a given set of inputs.

[9]This information is sometimes difficult to obtain

Choosing this value of b we have

$$\begin{aligned} Var(Z) &= \sigma^2_{g(X)} - 2b^*\sigma_{g(X)}\sigma_Y\rho_{g(X)Y} + b^{*2}\sigma^2_Y \\ &= \sigma^2_{g(X)}(1-\rho^2_{g(X)Y}) \\ &\leq \sigma^2_{g(X)} \end{aligned}$$

Clearly, the greater $|\rho_{g(X)Y}|$, the greater the reduction in the standard error of a Monte Carlo simulation for a fixed n.

We are interested in pricing three types of path dependent options using the variance gamma process. We price these options using Monte Carlo techniques incorporating two variance reduction methods, namely control variates and stratification. We discuss two types of Monte Carlo approaches. The first type, which we'll refer to as sequential sampling is discussed in the next section followed by the second type, which we'll refer to as bridge sampling. We show that bridge sampling has efficiency gains and lends itself to a very efficient method known as the 'Truncated Difference-of-Gammas Bridge Sampling' (TDGBS) developed by Avramidis & L'Ecuyer (2004).

4.5.3 Monte Carlo: Sequential Sampling

In this section we discuss the pricing of the three options using sequential sampling in a Monte Carlo simulation. Sequential sampling generates a path of a stochastic process X_t in a chronological order, i.e. Discretize $[0,T]$ into $p+1$ points $t_0 = 0, t_1, \ldots, t_p = T$. Now simulate $X_{t_1}, X_{t_2}, \ldots, X_{t_p}$ in that order. Compute the associated payoff and repeat a large number of times. The stochastic process chosen in this dissertation is the variance gamma process. To formalise, let $S_t = S_0 e^{X_t + \eta t}$ where

$$X_t = \vartheta G_t + \sigma W_{G_t}$$

is a $\mathbb{Q}$-variance gamma process and η (given by equation (4.4)) is a constant which transforms $S_t e^{-rt}$ into a $\mathbb{Q}$-martingale. Recall that ϑ is a constant and $G_{t+h} - G_t$ has a gamma distribution with mean h and variance νh. Below is an algorithm of how to simulate a discretized path of S_t for $t \in [0,T]$ using sequential sampling.

- Divide the interval $[0,T]$ into $p+1$ equally spaced points. Let the points be given by $t_0 = 0, t_1, t_2, \ldots, t_p = T$. Let $c = t_i - t_{i-1}$ be the constant difference.
- Generate p independent gamma variates G_i for $i = 1, \ldots, p$ with mean c and variance νc.
- Generate p independent normal variates N_i for $i = 1, \ldots, p$ with mean θG_i and variance $\sigma^2 G_i$

- The discretized variance gamma path is given by $X_{t_i} = \sum_{j=1}^{i} N_j$
- The discretized path for S_{t_i} is given by $S_{t_i} = S_0 e^{X_{t_i} + wt_i}$

It remains to compute the payoff for the three types of options. The payoff of the j^{th} path for an up-and-out call option (with barrier L and strike K) is given by

$$\Phi(S^j) = (S_T - K)1_{(S_T > K) \cap (\max_i S_{t_i} < L)}$$

Compute the above payoff a large number of times n. The Monte Carlo price is given by

$$\hat{V}_0 = e^{-rT} \frac{1}{n} \sum_{j=1}^{n} \Phi(S^j) \tag{4.28}$$

One would need to employ variance reduction techniques to reduce the standard error. These techniques are discussed in the following section. The lookback put and the knock-out double barrier call option (lower barrier = B, upper barrier = A) can be priced in a similar vein. Their payoffs of the j^{th} path are respectively given by

$$\begin{aligned} \Phi(S^j) &= \max_i S_{t_i} - S_T \\ \Phi(S^j) &= (S_T - K)1_{(S_T > K) \cap (\max_i S_{t_i} < A) \cap (\min_i S_{t_i} > B)} \end{aligned}$$

The associated Monte Carlo prices are then given by equation (4.28). It should be clear that for all three options, $\hat{V}_0$ is a biased estimate for all n. This is because there exist paths where $(\max_i S_{t_i} < L)$ but $(S_T^{Max} > L)$ and similarly for the minimum process. Stated mathematically, $\{\omega \in \Omega : S_T^{Max} < L\} \subset \{\omega \in \Omega : \max_i S_{t_i} < L\}$. It follows that $\hat{V}_0$ is always an overestimate[10] of the 'continuous time' price. It should be pointed out that sequential sampling of the variance gamma process requires p gamma variates and p normal variates for each path, which is fairly costly[11].

However, `Matlab` was the chosen language in writing the above algorithms. `Matlab` is a very efficient vector language and sequential sampling lends itself to fully vectorized code. The code does not contain any loops which would otherwise significantly increase computing time.

Sequential sampling does not lend itself to stratification of terminal values which is the primary source of variance in the option payoff. We seek a more efficient method which does not suffer from these disadvantages.

[10] Underestimate for the lookback put option

[11] Gamma random variates take especially long to generate

4.5.4 Monte Carlo: Bridge Sampling

The Brownian bridge construction was one of the first applications of what is known as bridge sampling. The algorithm works as follows: Let $t_0 = 0, t_1, \ldots, t_m = T$ be an increasing sequence of times where $m = 2^k$ for $k \in \mathbb{N}$.

- Generate $B_T = N(0, T)$
- Generate $B_{\frac{T}{2}}$ which is a normal random variable conditional on B_T
- Generate $B_{\frac{T}{4}}$ which is a normal random variable conditional on $B_{\frac{T}{2}}$
- Generate $B_{\frac{3T}{4}}$ which is a normal random variable conditional on $B_{\frac{T}{2}}$ and B_T

etc. Repeat until you obtain a sufficiently fine path. The Brownian bridge above is constructed in such a way that the resulting path is a discretized Brownian motion. This bridge construction exists for other stochastic processes apart from Brownian motion. In particular, it exists for the variance gamma process (Ribeiro & Webber, 2004). Here's the proposition.

Proposition 4.5.1 *Let G_t be a gamma process on the interval $[0, T]$ with $G_t \sim G\left(\frac{t\mu^2}{\nu}, \frac{\nu}{\mu}\right)$. Let $0 < \tau_1 < t < \tau_2 < T$. The conditional distribution of G_t given G_{τ_1} and G_{τ_2} is equal in distribution to*

$$G_{\tau_1} + (G_{\tau_2} - G_{\tau_1})Y$$

where $Y \sim B\left(\frac{(t-\tau_1)\mu^2}{\nu}, \frac{(\tau_2-t)\mu^2}{\nu}\right)$, i.e. Y has a beta distribution.

We henceforth refer to the following representation of the beta distribution. If $X \sim B(\alpha, \beta)$ then X has density function

$$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy}$$

Note that X has support on the interval $(0, 1)$. The beta distribution has a rather unusual shape for certain parameters. The p.d.f. can even have a bimodal shape. The familiar uniform distribution is a special case of the above with both α and β set to 1.
The proof heavily relies on quite a well known result in statistics. We state and prove the result. The proof is a modified version of the one given is Stewart (2003).

Proposition 4.5.2 *Let $X \sim G(a,\beta), Y \sim G(b,\beta)$ be two independent gamma variates. The ratio*

$$U = \frac{X}{X+Y}$$

has a beta distribution with parameters a, b, i.e. $U \sim B(a,b)$

Proof: Since X, Y are independent, their joint p.d.f. $f_{XY}(x,y)$ is merely the product of their densities, i.e.

$$f_{XY}(x,y) = \frac{x^{a-1}y^{b-1}e^{-\beta^{-1}(x+y)}}{\beta^{a+b}\Gamma(a)\Gamma(b)}$$

Define $U = \frac{X}{X+Y}$ and $V = X + Y$ with inverse transformations $x = uv$ and $y = v(1-u)$. This bounds U between $(0,1)$. The Jacobian $\mathbb{J}$ of this transformation is given by v. Using a standard result in statistics, U and V have joint density $f_{UV}(u,v) = f_{XY}(uv, v(1-u))\mathbb{J}$, i.e.

$$\begin{aligned} f_{UV}(u,v) &= \frac{vu^{a-1}v^{a-1}v^{b-1}(1-u)^{b-1}e^{-\frac{v}{\beta}}}{\beta^{a+b}\Gamma(a)\Gamma(b)} \\ &= \left(\frac{u^{a-1}(1-u)^{b-1}}{\Gamma(a)\Gamma(b)}\right)\left(\frac{v^{a+b-1}e^{-\frac{v}{\beta}}}{\beta^{a+b}}\right) \end{aligned}$$

Multiplying the numerator and denominator by $\Gamma(a+b)$ we see that the joint density factorizes into two marginal density functions. The first of which can be recognized as a beta distribution with parameters a and b. The second term can be recognized as a gamma distribution with parameters $a+b$ and β. Since $f_{UV}(u,v)$ can be factorized, it implies that U and V are independent. Thus, $U \sim B(a,b)$ which completes the proof.

□

We make use of this result in proving Proposition 4.5.1.

Proof of Proposition 4.5.1. Note that since G_t has independent increments,

$$G_t - G_{\tau_1} \sim G\left(\frac{(t-\tau_1)\mu^2}{\nu}, \frac{\nu}{\mu}\right) \qquad G_{\tau_2} - G_t \sim G\left(\frac{(\tau_2 - t)\mu^2}{\nu}, \frac{\nu}{\mu}\right)$$

are independent gamma variates. It follows that the ratio

$$Y = \frac{G_t - G_{\tau_1}}{G_t - G_{\tau_1} + G_{\tau_2} - G_t} = \frac{G_t - G_{\tau_1}}{G_{\tau_2} - G_{\tau_1}} \tag{4.29}$$

has a beta distribution with parameters $\frac{(t-\tau_1)\mu^2}{\nu}$ and $\frac{(\tau_2-t)\mu^2}{\nu}$. Since we interested in the distribution G_t conditional on G_{τ_1} and G_{τ_2} (i.e. G_{τ_1} and G_{τ_2} are constants)

we can simply change the subject of the formula in equation (4.29) to G_t, i.e. The conditional distribution of G_t is equal in distribution to

$$G_{\tau_1} + (G_{\tau_2} - G_{\tau_1})Y$$

and the proposition is proved.

□

Since the variance gamma process X_t can be written as a difference between two independent gamma processes, we can use this result to construct a variance gamma bridge. The algorithm works as follows:

- Generate two random variates from $\Gamma^+(T)$ and $\Gamma^-(T)$ independently of one another
- Set $X_T = \Gamma^+(T) - \Gamma^-(T)$
- Generate a draw from $Y \sim B\left(\frac{T-0.5T}{\nu}, \frac{0.5T}{\nu}\right)$
- Set $\Gamma^+(0.5T) = \Gamma^+(0) + (\Gamma^+(T) - \Gamma^+(0))Y$
- Generate another independent draw from $Y \sim B\left(\frac{T-0.5T}{\nu}, \frac{0.5T}{\nu}\right)$
- Set $\Gamma^-(0.5T) = \Gamma^-(0) + (\Gamma^-(T) - \Gamma^-(0))Y$
- Set $X_{0.5T} = \Gamma^+(0.5T) - \Gamma^-(0.5T)$
- Generate two independent draws from $Y^i \sim B\left(\frac{0.5T-0.25T}{\nu}, \frac{0.25T}{\nu}\right)$ for $i = 1, 2$
- Set $\Gamma^+(0.25T) = \Gamma^+(0) + (\Gamma^+(0.5T) - \Gamma^+(0))Y^1$
- Set $\Gamma^-(0.25T) = \Gamma^-(0) + (\Gamma^-(0.5T) - \Gamma^-(0))Y^2$
- Set $X_{0.25T} = \Gamma^+(0.25T) - \Gamma^-(0.25T)$

Subsequent steps are not listed but the pattern should be clear. Repeat until a desired level of discretization is achieved. This, by construction, gives a discretized path of a variance gamma process.

□

Bridge sampling offers a number of advantages over sequential sampling. Firstly, since the variance in most option payoffs is concentrated in the terminal value of the underlying process, bridge sampling allows one to stratify the terminal values. This is not possible using sequential sampling. Stratification of terminal values reduces the variance in the option payoff thereby reducing standard errors.
Secondly, for a given number of discretized points per path, bridge sampling requires as many calls to a beta inverse function as does sequential sampling to a

gamma and normal inverse function. The calls to the beta inverse function are in fact calls to a symmetric beta distribution, i.e. $B(\alpha, \alpha)$. Efficient algorithms have been developed which exploit this symmetry which significantly reduce computing time. In this regard, we adopt the method of L'Ecuyer & Simard (2006) for generating random variates from a symmetrical beta distribution. Finally, bridge sampling lends itself to the method of TDGBS to be discussed further on.

It remains to price the three types of options using this bridge construction. The pricing of the three options remains exactly the same as discussed in the previous section on sequential sampling. It's only the generation of the paths that differs. However, since this bridge sampling method was implemented in pricing the three options, we discuss two variance reduction techniques that were employed, namely a control variate and stratification of terminal values.

Control Variates

We start with the up-and-out call. We seek a random variable with known expectation which correlates well with the payoff of an up-and-out call. We used the following path independent option with payoff

$$\Phi(S_T) = \max(S_T - K, 0)1_{S_T<L} \tag{4.30}$$

where $K < L$ is the strike of the up-and-out call and $L > S_0$ is the barrier. One would expect this payoff to correlate fairly well with the up-and-out call payoff. Its expectation is known since it is merely the 'future value' of the price of the option with the above payoff. We therefore need to price this option and multiply by e^{rT}. As discussed in the previous chapter, the payoff can be replicated with a long call with strike K, short a call with strike L and short $L - K$ digital calls with strike L. We know how to price the call options in 'closed form' using a variance gamma underlying[12]. It remains to price the digital call. Its price is given by

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}[1_{S_T>L}] = e^{-rT}\mathbb{Q}(S_T > L) = e^{-rT}(1 - F(L))$$

where $F(L)$ is the distribution function of S_T evaluated at L under the risk-neutral measure $\mathbb{Q}$. This distribution function is known in 'closed form'. We can price both the call and the digital option and thereby price the option with payoff given by equation (4.30). The control variate's expectation is known. The optimal coefficient b^* was estimated using sample correlation and sample variances. Sample correlations of up to 80% were achieved for $L >> K$. Standard errors of simulations were significantly reduced.

The control variate chosen for the lookback put is simply the terminal underlying value S_T. The control variate chosen for the knock-out double barrier call option

[12]This is discussed in Chapter 3

was the following payoff:

$$\Phi(S_T) = \max(S_T - K, 0)1_{S_T<A \cap S_T>B}$$

where A and B denote the upper and lower barriers respectively. We only consider the case where $K \in [B, A]$. In this case, the payoff simplifies to equation (4.30) with $L = A$ which coincides with the same control variate as for the up-and-out call. We do not expect this control variate to be as effective as with the up-and-out call but nevertheless does its job in reducing the standard error.

Stratification of Terminal Values

Since the variability in the three option's payoffs is primarily concentrated on the underlying's terminal values, it would make sense to stratify the uniform random variates that generate these values. Since we have used the representation of a variance gamma process as the difference between two independent gamma processes, we require two uniform variates in order to generate one terminal value. We therefore require to stratify the unit square as opposed to one-dimensional stratification of uniform random variates. Stratification forces a certain number of uniform draws into what's known as strata. This has the effect of covering the $[0, 1]$ interval more uniformly thereby reducing the variance of the uniform draws. We divide each one-dimensional $[0, 1]$ interval into K evenly spaced strata, i.e.

$$\left(0, \frac{1}{K}\right), \left(\frac{1}{K}, \frac{2}{K}\right) \cdots \left(\frac{K-1}{K}, 1\right)$$

These form one-dimensional stratum. In two-dimensions, we seek the Cartesian product of intervals, i.e. the i^{th} interval of the first uniform variate and the j^{th} interval of the second uniform variate have an interval of the form

$$\left(\frac{i-1}{K}, \frac{i}{K}\right) \times \left(\frac{j-1}{K}, \frac{j}{K}\right)$$

In total, there will be K^2 such intervals. This requires us to produce at least K^2 pairs of independent uniform draws. Given K^2 uniform pairs, we need to force that one and only one pair gets allocated to each of the K^2 strata. Given the j^{th} pair of independent uniform draws $(U_1, U_2)_j$ for $j = 1, \ldots, K^2$, we simply define the stratified pair as follows:

$$V_1^j = \frac{j - 1 + U_1}{K} \quad \text{for } j = 1, 2, \ldots, K$$
$$V_2^i = \frac{i - 1 + U_2}{K} \quad \text{for } i = 1, 2, \ldots, K$$

The desired K^2 pairs are (V_1^j, V_2^i). We found that sufficient stratification took place when $K = 10$.

The bridge sampling technique discussed above has a number of advantages over the sequential sampling technique. However, they are both inefficient. As an example, we choose the up-and-out call option. Consider a final draw from $X_T = \Gamma^+(T) - \Gamma^-(T)$ which results $S_T = S_0 e^{X_T + wT} > L$. The payoff of this particular path is already known at this point which equals 0. The bridge algorithm still continues to produce a variance gamma path for this particular terminal draw even though its payoff is already known with certainty. This creates inefficiencies since valuable computing time is wasted. In addition, it's not possible to vectorize one's code in `Matlab` using bridge sampling which means that one has to resort to loops which significantly increase the computing time. The TDGBS approach addresses this issue not just at terminal times but at subsequent points along the path. These Monte Carlo simulations are costly and there is a trade-off between convergence of the continuous time price (i.e. increasing the number of discretized points in a path) and computing time.

4.5.5 Truncated Difference-of-Gammas Bridge Sampling

The TDGBS is an algorithm specifically developed for pricing path dependent options using a variance gamma process. We follow the paper written by Avramidis & L'Ecuyer (2004) although we make minor modifications when deemed necessary. The essential ingredients of TDGBS are firstly, the representation of the variance gamma process as a difference between two independent gamma processes, and secondly, the bridge algorithm discussed above. This representation allows one to compute bounds (for any $t \in [0, T]$) on the underlying process. These bounds translate into bounds for the 'continuous time' option payoff. For certain types of options, for example barriers options, it's possible to sample until the bounds coincide which gives the exact 'continuous time' payoff thereby eliminating all bias. In addition, the bounds on the underlying process dictate the option payoff after a varying number of points in a path. This allows one to terminate or truncate a particular path since it becomes redundant to continue the construction of a path if the option payoff is already known. This leads to significant savings in terms of computing time. We start with a discussion on how to construct these bounds on the (exponential) variance gamma process.

Bounds on the Exponential Variance Gamma Process

Let $X_t = \Gamma^+(t) - \Gamma^-(t)$ denote a variance gamma process with the same parameters as previously discussed. Suppose we sample a path of the variance gamma process in this order:

$$T, \frac{T}{2}, \frac{T}{4}, \frac{3T}{4}, \frac{T}{8}, \ldots$$

by making use of Proposition 4.5.1. Suppose we have sampled m points per path where $m = 2^k$, $k \in \mathbb{N}$. Let $t_{m,0} = 0, t_{m,1}, \ldots, t_{m,m-1}, t_{m,m} = T$ denote the above times arranged in increasing order. For example, $t_{4,3} = \frac{3T}{4}$. Here's the proposition:

Proposition 4.5.3 *Let $t \in (t_{m,i-1}, t_{m,i})$ for $i = 1, \ldots, m$. Define three processes as follows:*

$$\begin{aligned} L_{m,i} &= S_0 e^{\eta t_{m,i-1} - \Gamma^-(t_{m,i}) + \Gamma^+(t_{m,i-1})} \\ S_t &= S_0 e^{\eta t + \Gamma^+(t) - \Gamma^-(t)} \\ U_{m,i} &= S_0 e^{\eta t_{m,i} + \Gamma^+(t_{m,i}) - \Gamma^-(t_{m,i-1})} \end{aligned}$$

then S_t (the exponential variance gamma process) is contained between $L_{m,i}$ and $U_{m,i}$ for each $i = 1, \ldots, m$, i.e. $L_{m,i} \leq S_t \leq U_{m,i}$.

Proof: Fix $t \in (t_{m,i-1}, t_{m,i})$. Note that $\Gamma^+(t) \leq \Gamma^+(t_{m,i})$ since Γ^+ is a surbordinator (an increasing Lévy process). Similarly, $\Gamma^-(t) \geq \Gamma^-(t_{m,i-1})$. It should be clear that this implies $S_t \leq U_{m,i}$. A symmetrical argument proves that $S_t \geq L_{m,i}$.

□

One would expect that these bounds decrease monotonously as the number of points per path (i.e. m) increases. This is indeed the case, i.e. $L_{m,i} \leq L_{m+1,i} \leq S_t \leq U_{m+1,i} \leq U_{m,i}$. The proof is straightforward and is given in Avramidis & L'Ecuyer (2004).
The truncation of paths is best explained via an example. Suppose we are interested in pricing an up-and-out call option by Monte Carlo with S_t having exponential variance gamma dynamics. Let the barrier be L and the strike be K. Suppose we have already sampled m points in a particular path. Denote these values by $S_{t_{m,1}}, S_{t_{m,2}}, \ldots, S_{t_{m,m}}$ where $S_{t_{m,0}} = S_0$. We also suppose that none of the $S_{t_{m,i}}$ have breached L, i.e. $\max_i S_{t_{m,i}} < L$. Truncating early at this point (after m steps) and calculating the payoff of $\max(S_{t_{m,m}} - K, 0)$ introduces bias since it's possible that S_t for $t \in (t_{m,i-1}, t_{m,i})$ for $i = 1, \ldots, m$ to be greater than L. However, using the above proposition, it's possible to tell whether $S_t, t \in [0, T]$, could be greater than L or not. $U_{m,1}$ provides an upper bound for S_t over the interval $[0, t_{m,1})$ and more generally, $U_{m,i}$, provides an upper bound for S_t over the interval $(t_{m,i-1}, t_{m,i})$. We continue as follows:

Calculate $U_{m,i}$ for $i = 1, \ldots, m$. If $\max_i U_{m,i} < L$ then truncate the sampling procedure. The payoff of the option in this path is $\max(S_{t_{m,m}} - K, 0)$. This is because it's impossible for S_t to breach the barrier L for all $t \in [0, T]$.
If $\max_i U_{m,i} > L$ then it's still possible for S_t to breach the barrier L. In this case, we would continue to sample more points in this particular path. You will need to impose an upper bound M on the number of points per path as it is possible

that $\max_i U_{m,i} > L$ and $\max_i S_{t_{m,i}} < L$ will occur for very large m. It should be clear that this TDGBS clearly reduces computing time. Moreover, assuming that $m < M$, the discounted payoff is an unbiased estimate of the continuous time price of the option. Choosing M sufficiently large creates a virtually[13] unbiased estimate of the continuous time price.

Barrier options provide a special case where it's possible to eliminate virtually all bias[14]. Lookbacks and Asian options are examples where the bias can never be completely alleviated. However, the bounds on the underlying process translate into bounds on the continuous time option payoff. This is assuming that the option payoff is monotonic in S_t given already sampled points $S_{t_{m,1}}, S_{t_{m,2}}, \ldots, S_{t_{m,m}}$. For suppose we price a lookback put with continuous time payoff

$$\Phi(S) = S_T^{Max} - S_T$$

Define two estimators of the payoff by

$$\begin{aligned} C_{L,m} &= \max_i L_{m,i} - S_T \\ C_{U,m} &= \max_i U_{m,i} - S_T \end{aligned}$$

Using straight forward monotonicity arguments, it should be clear that $C_{L,m} \leq \Phi(S) \leq C_{U,m}$, i.e. We have bounds on the continuous time payoff which narrow monotonously with increasing m. In fact, we can improve on our lower bound since $\max_i S_{m,i} \geq \max_i L_{m,i}$ which allows us to write $C_{L,m} = \max_i S_{m,i} - S_T$. It remains to give an account of how to employ the TDGBS algorithm applied to pricing the three options.

Barrier Options

We only discuss the up-and-out call option since the knock-out double barrier call option is very similar. We discuss the sampling of one simulation. The algorithm works as follows:

Generate two independent gamma random variates, $\Gamma^+(T)$ and $\Gamma^-(T)$. We employed stratification (on the unit square) of the pseudo-random draws. Check immediately whether $S_T = S_0 e^{\eta T + \Gamma^+(T) - \Gamma^-(T)} > L$ or whether $S_T < K$. If so, terminate sampling, set the payoff equal to zero, and move on to the next path. If $K < S_T < L$ then generate two symmetrical beta random variates Y^1 and Y^2. Compute

$$S_{0.5T} = S_0 e^{0.5\eta T + \Gamma^+(T)Y^1 - \Gamma^-(T)Y^2}$$

[13] It is possible that $m = M$

[14] i.e. bias induced by using a discretized variance gamma path

If $S_{0.5T} > L$ then terminate sampling and set the payoff equal to zero. Otherwise compute $U_{2,1}$ and $U_{2,2}$. If $\max(U_{2,1}, U_{2,2}) < L$ then set the payoff equal to $S_T - K$ and terminate sampling. Note that for the double barrier case, you would also need to compute $L_{2,1}$ and $L_{2,2}$ and check that $\min(L_{2,1}, L_{2,2}) > B$ where B is your lower barrier. If $\max(U_{2,1}, U_{2,2}) > L$ then it's still possible for S_t to cross the barrier L. Continue sampling. This time you are required to generate four beta variates. Compute $S_{0.25T}$ and $S_{0.75T}$. Check if $\max(S_{0.25T}, S_{0.75T}) > L$ and if so, terminate sampling and set the payoff equal to zero. If not, compute $U_{4,1}, U_{4,2}, U_{4,3}, U_{4,4}$. Check if $\max(U_{4,1}, U_{4,2}, U_{4,3}, U_{4,4}) < L$ and if so, set the payoff equal to $S_T - K$. If not, continue sampling etc. Continue sampling until the algorithm terminates the path or until some upper bound M is reached

Repeat the above a large number of times incorporating a control variate. The same control variate was used (i.e. the truncated call option). Estimate the optimal b^* coefficient using sample variances and sample correlations. Compute the associated summation and remembering to discount, you have obtained a virtually unbiased estimate of the continuous time payoff whilst making significant savings in computing time.

Lookback Put

Unlike barrier options, it is not possible to eliminate all bias when pricing a lookback option. However, as discussed above, one may obtain bounds which contain the continuous time payoff. Simulate $S_{m,i}$ for $i = 1, \ldots, m$ using the variance gamma bridge approach incorporating stratification of terminal uniform draws. Compute

$$\begin{aligned} C_{L,m} &= \max_i S_{m,i} - S_T \\ C_{U,m} &= \max_i U_{m,i} - S_T \end{aligned}$$

These bounds narrow monotonously for increasing m. Note that the continuous time payoff is contained between these two bounds. In order to obtain one payoff, extrapolation techniques are required. One method is to simply average $C_{L,m}$ and $C_{U,m}$. This was the method adopted in this dissertation. The terminal exponential variance gamma value (which is stratified) was used as a control variate. Repeat a large number of times, compute the associated summation and you have obtained a biased estimate of the continuous time payoff. Avramidis & L'Ecuyer (2004) show that if your bias is of a particular asymptotic form, then the TDGBS approach reduces the order of the bias quite substantially.

Chapter 5

Calibration of Variance Gamma Model

This section discusses how to make use of equation (4.7) to price European options with payoff $f(S_T)$. Equation (4.7) requires us to make use of the risk-neutral density $f^{\mathbb{Q}}_{Y_t}(x)$ given by equation (4.6). This will require us to obtain parameter estimates of ϑ,ν and σ since r and q are observed in the market. Estimation of the three parameters can be implemented in several ways (e.g. Maximum likelihood estimation or moment matching techniques). We would like to estimate our parameters so that our model prices options in a manner consistent with the market. This is known as the market calibration problem.

The conventional method of calibrating a model which prices options in a consistent manner is that of least squares. The choice of the type of option which needs to be calibrated depends on the liquidity of the option. In most exchanges, European call options appear to be the most liquid, which is indeed the case in the South African market. The calibration problem can be formulated as follows: Given quoted market prices of European call options $C^*(K_i)$ for fixed maturity and various strikes $K_i, i = 1, \ldots, N$, model prices of European call options denoted by $C^{VG}(K_i, \vartheta, \sigma, \nu)$ where $C^{VG}(K_i, \vartheta, \sigma, \nu)$ are given by equation (4.8), we need to solve the following least squares problem:

$$\min_{\vartheta, \nu, \sigma} \sum_{i=1}^{N} (C^*(K_i) - C^{VG}(K_i, \vartheta, \sigma, \nu))^2 \tag{5.1}$$

subject to $\vartheta \in \mathbb{R}\backslash\{0\}$, $\nu > 0$ and $\sigma > 0$.
It should be noted that several problems arise with the formulation of equation (5.1). The most important problem is that the formulation is ill-posed. The parameters $\hat{\vartheta}, \hat{\nu}, \hat{\sigma}$ which minimize equation (5.1), are highly sensitive to changes in $C^*(K_i)$. This is clearly an undesirable feature of a calibration engine. Secondly, as pointed out by Tankov (2004), the least squares formulation does not guarantee a solution. Tankov (2004) give an example of a least squares formulation

which does not admit a solution whereas a solution does in fact exist.

Note that the computation of $C^{VG}(K_i, \vartheta, \sigma, \nu)$ involves some lengthy and expensive calculations since numerical integration techniques are required to compute equation (4.8). The integrand itself is written in terms of an integral (the modified Bessel function) although this does not pose a serious problem since efficient `C++` and `Matlab` source code is available. Functions were written for $C^{VG}(K_i, \vartheta, \sigma, \nu)$ in `C++` and `Matlab` which coincided with Monte Carlo integration. The technique of Romberg integration was implemented in `C++`. Romberg integration is an efficient algorithm which reduces the number of partitions required in using the Trapezoidal quadrature technique without loss of precision. See William H. Press & Flannery (1992) for a discussion which includes source code. In `Matlab`, we made use of the function `quad`. The function uses an adaptive recursive Simpson's rule.

The above minimization problem was solved in `C++` by making use of the Nelder and Mead method. The Nelder and Mead method minimizes functions of the form $f : \mathbb{R}^d \to \mathbb{R}^+$ by changing the d parameters. The function used in our case would be $f(\vartheta, \sigma, \nu) = \sum_{i=1}^{N}(C^{VG}(K_i, \vartheta, \sigma, \nu) - C^*(K_i))^2$. Run times took several minutes given the complexity of just one call to f. The method is not guaranteed to find a global minimum (should it exist) but will find local minima depending on how you initialize your starting points.

The same least squares formulation was carried out in `Matlab` using the `lsqnonlin` command. The `lsqnonlin` requires a vector-valued function f as input and a vector of parameters. The command implicitly constructs sums of squared differences. Denoting f^i as the i^{th} component of f, then in our case, f would take the following form: $f^i(K_i, \vartheta, \sigma, \nu) = C^{VG}(K_i, \vartheta, \sigma, \nu) - C^*(K_i)$ for $i = 1, \ldots N$. The command then minimizes $\sum_{i=1}^{N}(f^i)^2$ by changing ϑ, σ, ν. `lsqnonlin` uses the interior-reflective Newton method. The run times for the algorithm were lengthy and we again encountered several problems. Different starting points were chosen and all produced local minima. The problem with this method is that the various estimated parameters that produced local minima were quite distinct. The decision of which triplet of parameters $(\hat{\vartheta}, \hat{\sigma}, \hat{\nu})$ to use is not clear. The algorithm is essentially finding a 'valley' of local minima with different values of estimated parameters. This is a problem specific to least squares calibration as pointed out in Tankov (2004).
The least squares formulation is not only ill-posed but very dependent on how you initialize ϑ, σ, ν. This confirmed that the problem at hand was not the minimization algorithm but a problem specific to the least squares formulation.

One must bear in mind that the choice of parameters are very important for the pricing and hedging of more exotic options. Suppose that $\acute{\vartheta}, \acute{\sigma}, \acute{\nu}$ and $\tilde{\vartheta}, \tilde{\sigma}, \tilde{\nu}$

The first method is probably the least well known and a brief discussion is presented in Cont & Tankov (2002). The second approach, abbreviated by PIDE, can be set up if one knows the Lévy measure of the underlying Lévy process. The PIDE looks very similar to a PDE that one obtains using Kolmogorov's equations although there is an extra term. This term is a stochastic integral involving the jump measure of the Lévy process. PIDE's are considerably more complex to firstly set up and secondly to solve. The author believes that PIDE's would be a plausible method to apply if one uses a jump-diffusion process or at least a Lévy process which has a finite Lévy measure. This is far too restrictive in our setting, since we require the variance gamma process to price options which has an infinite Lévy measure. Methods do exist to solve PIDE's for processes with infinite Lévy measure but involves some sophisticated mathematics worth exploring in a separate topic. See Cont & Tankov (2002) for a lengthy discussion. PIDE's will not be pursued any further.

The Monte Carlo method has been chosen in this dissertation as the method of pricing path dependent options in a Lévy world. We will focus on pricing the same three options as we did in the previous chapter. They are:

- Up-and-out Call option
- Lookback Put option
- Knock-out double barrier Call option

We have chosen the variance gamma process (discussed in detail in Chapter 3) to price the above options. A comparison will be made between prices obtained for the above three options priced in a Black-Scholes world and a variance gamma world.

4.5.2 Monte Carlo Methods

We give a brief discussion of applying Monte Carlo methods, alluding to the advantages and disadvantages specific to pricing path dependent options. For an extremely well written and comprehensive text on Monte Carlo methods in finance, we refer you to Glasserman (2004).

Suppose we have a random variable X with known distribution function F and associated parameter(s). Suppose that we require $\mathbb{E}[g(X)]$ but this is not known in closed form. Monte Carlo methods can provide an estimate of the above expectation by simulating observations with distribution F, transforming them by g and computing the sample mean. Since F is known then so is its inverse F^{-1} since one can also obtain F^{-1} either in closed form or by numerical methods such as Newton's method or acceptance/rejection methods. The method works as follows:

- Simulate n uniform random variables U_i for $i = 1, .., n$
- Compute $X_i = F^{-1}(U_i)$
- Compute $g(X_i)$
- Compute $\frac{1}{n}\sum_{i=1}^{n} g(X_i) = \hat{\mathbb{E}}[g(X)]$ is an estimator of $\mathbb{E}[g(X)]$

See Glasserman (2004) for details. This is the most simplistic approach to estimating $\mathbb{E}[g(X)]$ and is usually referred to as a brute force Monte Carlo simulation since no variance reduction techniques have been employed. An estimator, $\hat{\theta} = \hat{\mathbb{E}}[g(X)]$ of $\theta = \mathbb{E}[g(X)]$, is said to be unbiased if $\mathbb{E}[\hat{\theta}] = \theta$ and consistent if $\mathbb{E}[\hat{\theta}] \to \theta$ in probability as $n \to \infty$. An estimator (even if it's unbiased) is meaningless unless we can quantify the error is in our approximation. The most commonly used measure is referred to as the mean square error, defined as follows:

$$\begin{aligned} MSE(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \theta)^2] \\ &= \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2] + (\mathbb{E}[\hat{\theta}] - \theta)^2 \\ &= Var(\hat{\theta}) + Bias^2(\hat{\theta}) \end{aligned}$$

The formula can be further simplified since θ is an expectation of the form $\theta = \mathbb{E}[g(X)]$. It follows that

$$Var(\hat{\theta}) = Var(\hat{\mathbb{E}}[g(X)]) = Var\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(g(X_i)) = \frac{\sigma^2_{g(X)}}{n}$$

where $\sigma^2_{g(X)}$ is the variance of $g(X)$. We will henceforth refer to the ratio

$$\sigma_e = \frac{\sigma_{g(X)}}{\sqrt{n}}$$

as the standard error of a Monte Carlo simulation. This becomes the important measure of error when one deals with unbiased estimators. However, in practical situations, $\sigma_{g(X)}$ is unknown. This is because one requires knowledge of $\mathbb{E}[g(X)]$ in order to compute $\sigma_{g(X)}$ which is precisely what we trying to estimate in the first place. However, $\sigma_{g(X)}$, can be consistently estimated by

$$s_{g(X)} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(g(X_i) - \hat{\mathbb{E}}[g(X)])^2}$$

which is known, albeit a random variable since it changes from simulation to simulation. Clearly, the smaller the standard error, the more confidence one would have in one's estimate of $\mathbb{E}[g(X)]$. A naive answer is to simply increase n but this has two serious drawbacks. Firstly, the larger the n, the greater the length

are such that they produce the same local minima for some calibration technique (e.g. least squares). Suppose we'd like to price a European barrier option or any type of exotic option. These two models will produce wildly different prices if $||(\acute{\vartheta}, \acute{\sigma}, \acute{\nu}) - (\tilde{\vartheta}, \tilde{\sigma}, \tilde{\nu})||$ is large. Perhaps more importantly, is that hedging strategies (e.g. delta hedging) will be very different. We hence need to find a calibration technique which satisfies the following criteria:

1. Produces theoretical prices which are 'closest' to market prices. The 'closeness' measure will depend on the calibration engine you choose.

2. Continuity of solution. Perturbations of market prices leading to a small and smooth change in the estimated parameters.

3. A solution vector of parameters which does not depend on how one initializes them.

The least squares formulation does not satisfy these criteria. Another approach, which will be adopted in this dissertation, is that instead of calibrating the observed call option prices to equation (4.8), one may calibrate the observed risk-neutral density of the underlying security to that of the theoretical risk-neutral density given by equation (4.6). This approach has the advantage of satisfying the above criteria but has drawbacks in that the solution vector of parameters may not produce a local minima in the least squares setting.

Our task involves calibrating the risk-neutral distribution of our underlying security and not the 'real world' density given by equation (3.9). Estimation techniques[1] can be used to calibrate the 'real world' density using past log returns. Estimates of the three parameters can be obtained but there is no reason why these are the same three parameters as in equation (4.6). In a Black-Scholes market we have the remarkable result that the volatility parameter σ, is invariant under change of equivalent measure (Girsanov's Theorem), i.e. $\sigma^{RealWorld} = \sigma^{RNworld}$. No such results exist in a variance gamma world and we have no reason to believe that $\vartheta^{RealWorld} = \vartheta^{RNworld}$ etc. The risk-neutral world is, unfortunately, not directly observable. We do, however, have the extremely useful result given by Breeden & Litzenberger (1978), coupled with some interpolating techniques, which allows us to back out an implied risk-neutral distribution of our underlying security.

5.1 Data Used in Market Calibration

SAFEX (The South African Futures Exchange) is an organization which allows for exchange traded futures contracts and options on futures contracts. SAFEX

[1] maximum likelihood seems to dominate the literature

offers futures contracts and options on these contracts on various indices expiring every 3 months on the 3^{rd} Thursday of that particular month. The option prices used in our market calibration will be European call options on futures contracts on the ALSI (The All Share Index). These option contracts are the most liquid on SAFEX and have a wide range of strikes available. These option contracts are traded on volatilities and the SAFEX Black formula[2] is used to recover the price in Rands.

The data used in our calibration was extracted from volatilities quoted on 24 March 2005. Note that these volatilities are not mark-to-market volatilities[3] but volatilities on the very last trade on 24 March 2005. These European call options on futures contracts on the ALSI expire on the 16 June 2005 at which time the futures contract is also closed out. Since they both expire on the same date, the futures option price must coincide with that of the underlying index option price to exclude arbitrage. This allows us to price these futures option contracts as if they are index options. We used the 3-month JIBAR rate as a proxy for the constant risk-free rate over the period. The 3-month JIBAR is a quoted simple rate which converted to 7.48% as a continuously compounded rate. The futures price on the ALSI on 24 March 2005 was 12050 which means that the implied index level must be $12050e^{-0.0748(0.23)} = 11843$ to exclude arbitrage[4]. We can henceforth treat the observed option prices as options on an index with spot price $S_0 = 11843$, i.e. we can assume index dynamics which would be an exponential variance gamma process. There are 46 different strikes available with their associated volatilities. The strikes range from 8555.5 to 13978 and all contracts have a time to expiry of 84 days (0.23 years).

5.2 Construction of the Implied Risk-Neutral Distribution

Breeden & Litzenberger (1978) showed that the implied risk-neutral probability density function of an underlying security is given by the second derivative of the call price with respect to strike. To be precise, let $F_{t,T}$ be the time t-futures price of an underlying security S such that $F_{T,T} = S_T$. Letting $C_t(K,T)$ be the time

[2]This is very similar to Black's formula for futures options but is modified since options are fully margined at SAFEX and not paid for upfront. See West (2006) for an excellent explanation

[3]this is what SAFEX publishes (for margining purposes) but is incorrect for analysis since it is a weighted average of the last couple of hours trading and not the last day's trade

[4]we have assumed a constant risk-free rate which implies that futures and forward prices coincide

t-price of a futures call option then

$$\frac{\partial C_0(K,T)}{\partial K} = \mathbb{Q}(S_T \leq K) - 1 \tag{5.2}$$

$$\frac{\partial^2 C_0(K,T)}{\partial K^2} = f^{\mathbb{Q}}_{S_T}(K) \tag{5.3}$$

where $\mathbb{Q}$ is a risk-neutral measure and $f^{\mathbb{Q}}_{S_T}$ is the risk-neutral density of S_T. See Knox & Ouwehand (2006) for an easy to understand one line proof. This is an incredibly powerful theoretical result and only useful for practical purposes if we have a continuum of strikes available (else the partial derivatives are not defined). It's obvious that only a discrete set of strikes are available and one would need to construct a differentiable curve fitting all, in our case, 46 pairs of data points through the option price, strike space. This can be made possible via the cubic spline technique of interpolation which fits a third degree piecewise defined polynomial to all data points. To be specific, let K_i and C_i denote the 46 strikes and observed call option prices for $i = 1, \ldots, 46$. We attempt to fit the following cubic equation:

$$C(K) = a_i + b_i(K - K_i) + c_i(K - K_i)^2 + d_i(K - K_i)^3 \text{ for } K_i \leq K \leq K_{i+1}$$

The C_i and K_i are known and we attempt to fit a cubic curve to solve for $C(K)$ for $K \neq K_i$. Note that we have to solve for a total of 180 (45×4) coefficients since there are only 45 splines connecting all data points. We must have the constraint that $C(K_i) = C_i$ (the spline must pass through all points) and that the curve is continuous and differentiable in K. This leaves us with a further 46 constraints on the coefficients in order to obtain a unique solution for the 180 coefficients. See Hagan & West (2006) for details. We adopt the natural cubic spline which requires the entire curve to be twice differentiable and that both end points have a second derivative of zero. This was implemented in `Matlab` using the `scape` command in the `Spline Toolbox`.

We now have a differentiable curve $C(K)$ for all $K \in [8555.5, 13978]$. The c.d.f. (cumulative distribution function) can be obtained by differentiating the above curve with respect to strike and shifting the curve up by 1. This can be achieved in `Matlab` using the `fnder` command in the `Spline Toolbox`. Let us define the implied risk-neutral c.d.f. obtained by $F_{imp}(K)$. The graph below depicts the implied risk-neutral distribution function of the index level (for expiry 16-June 2005).

It should be pointed out that the implied risk-neutral distribution function obtained by differentiation is defined for $Index \in [8555.5, 13978]$ and not for $Index \in [0, \infty)$. The above curve has a minimum of $F_{imp}(8555.5) = 0.02065$ and a maximum $F_{imp}(13978) = 0.99925$. It becomes necessary to include two extra points which have the property that $F_{imp}(0) = 0$ and $F_{imp}(\infty) = 1$. In

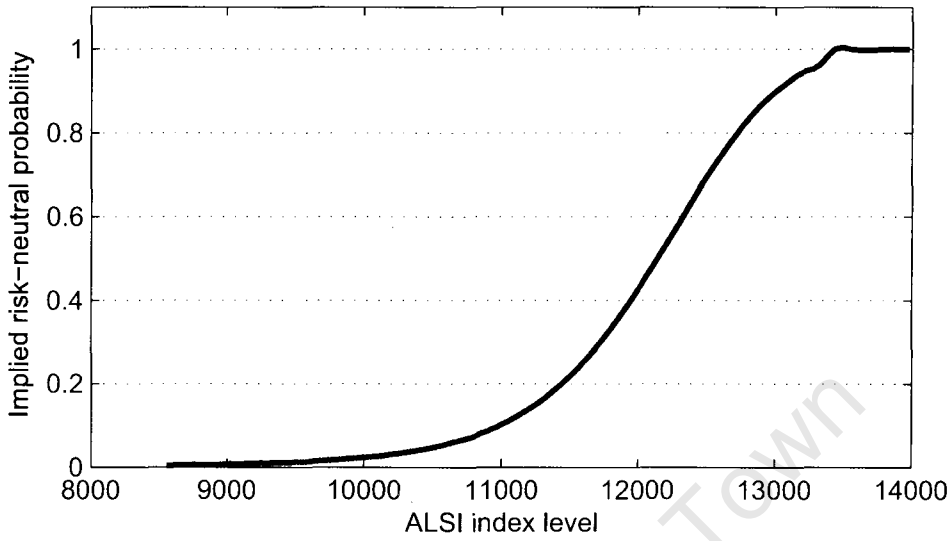

Figure 5.1: Implied risk-neutral c.d.f.

addition, we require two smooth curves joining the newly constructed points. The reason for this is because we require to simulate from this c.d.f. $F^{-1}_{imp}(x)$ for $x \in [0, 0.02065] \cup [0.99925, 1]$ which is not defined unless we add two extra points along with their associated curves joining the points. The upper tail did not pose a problem since a random number drawn in $[0.99925, 1]$ occurs with probability less than 0.001 but random numbers drawn in $[0, 0.02065]$ occurred far more regularly. It was decided to fit log-normal[5] tails. The log-normal distribution has two parameters and we have two constraints in that the curve must pass through the points $F_{imp}(0) = 0$ and $F_{imp}(8555.5) = 0.02065$. The two parameters for the log-normal density were (uniquely) solved using a two by two system of non-linear equations and $F^{-1}_{imp}(x)$ is now defined for $x \in [0, 0.99925]$.

We will not bother with the probability density function (p.d.f.) since the second derivative of a cubic function is a linear function which means that the implied p.d.f. obtained will consist of kinked straight lines. This would then require simulation from the density and apply kernel smoothing techniques to obtain a smooth p.d.f.

[5]this is equivalent to fitting Gaussian tails for the log density

5.3 Market Calibration of the Variance Gamma Model

This section provides a method for calibrating the risk-neutral variance gamma p.d.f. given by equation (4.6) to the implied risk-neutral density given by $\frac{\partial F_{imp}(K)}{\partial K}$ with $F_{imp}(K)$ obtained in the previous section. We will adopt the method of matching the first four moments of the risk-neutral variance gamma p.d.f. to that of the implied risk-neutral distribution. Recall that $S_t = S_0 e^{\eta t + \vartheta G(t,\omega) + \sigma B(G(t,\omega),\omega)}$ under the risk-neutral measure $\mathbb{Q}$ with η given by equation (4.4). Defining $R(n)$ to be the n^{th} risk-neutral raw moment, it is easy to show that

$$R(n) = \mathbb{E}^{\mathbb{Q}}[S_t^n] = S_0^n e^{n\eta t}\left(1 - n\nu\left(\vartheta + \frac{n\sigma^2}{2}\right)\right)^{-\frac{t}{\nu}} \text{ for } n \in \mathbb{N} \tag{5.4}$$

This can be verified by conditioning the expectation on $G(t,\omega) = z$ and proceeding in a similar fashion as we did in obtaining the characteristic function of $Y(t,\omega) = \ln \frac{S_t}{S_0}$. It follows that the first four central moments μ_i of S_t are given by:

$$\begin{aligned}
\mu_1 &= R(1) \\
\mu_2 &= R(2) - R(1)^2 \\
\mu_3 &= R(3) - 3R(2)R(1) + 2R(1)^3 \\
\mu_4 &= R(4) - 4R(3)R(1) + 6R(2)R(1)^2 - 3R(1)^4
\end{aligned}$$

The μ_i given above are all functions of ϑ,ν and σ which we require to estimate. We need to match up these four theoretic moments to the implied risk-neutral moments and solve for the three parameters.

The implied risk-neutral moments can be accurately estimated by simulating observations with c.d.f. $F_{imp}(K)$ and obtain sample estimates of the first four central moments. Recall that if U has a uniform distribution on $[0,1]$ then $F_{imp}^{-1}(U)$ has distribution F_{imp}. This boils down to finding zeros of $F_{imp}(u) - u$ where u is a drawing from a uniform random variable on $[0,1]$. Newton's method is probably the most efficient method to find zeros but may not converge for certain values of u. The Bisection method was chosen which is less efficient than Newton's method but guaranteed to find a root should it exist. We now have a method to simulate observations with c.d.f. $F_{imp}(K)$. We repeat this a large number of times (10000 in our case) and obtain sample estimates[6] of the first four central moments. We will denote our sample estimates of the implied risk-neutral central moments by $\hat{\mu}_i$. The following sample estimates of the moments of $F_{imp}(K)$ were obtained

[6] i.e. if x_i is an i.i.d sequence of N drawings from X then $\hat{\mu}_1 = \sum_{i=1}^N x_i$
$\hat{\mu}_2 = \frac{1}{N}\sum_{i=1}^N (x_i - \bar{x})^2$, $\hat{\mu}_3 = \frac{1}{N}\sum_{i=1}^N (x_i - \bar{x})^3$, $\hat{\mu}_4 = \frac{1}{N}\sum_{i=1}^N (x_i - \bar{x})^4$

from 10000 simulations:

Stdev	Coef Skewness	Coef Kurtosis
967.89	-1.5478	6.6982

Note that these measures displayed in the table above are normalized versions of $\hat{\mu}_i$ which more useful for interpretation.

We are now in a position to solve for ϑ,ν and σ by equating $\mu_i = \hat{\mu}_i$ for $i = 2, 3, 4$. It is not necessary to equate $\mu_1 = \hat{\mu}_1$ since this is independent of the three parameters and by construction of $\mathbb{Q}$ yields $S_0 e^{t(r-q)} = \mu_1$. We have a system of three non-linear equations in three unknowns ϑ,ν and σ. It's not clear a priori whether a unique solution exists but a least squares formulation produced a minimum of very close to zero which suggested that a unique solution may exist. We instead formulated the problem as a three by three system of non-linear equations.

$$\text{Solve for } \vartheta, \nu, \sigma \text{ such that } \mu_i(\vartheta, \nu, \sigma) - \hat{\mu}_i = 0 \text{ for } i = 2, 3, 4 \tag{5.5}$$

This can be achieved through the command `fsolve` in the `Optimization Toolbox` in `MATLAB`. `fsolve` uses the Gauss-Newton method and convergence is fast. The solution vector is given by

$$\vartheta = -0.24065 \quad \nu = 0.32634 \quad \sigma = 0.13489 \tag{5.6}$$

These parameters translate into the following coefficients of skewness and kurtosis of the risk-neutral log density of the underlying ALSI index:

Coef Skewness	Coef Kurtosis
-2.1176	34.8816

The estimates are consistent with the theory that risk-neutral return distributions (or log densities) display negative skewness and are leptokurtic.

Chapter 6

Numerical Results and Conclusions

Chapter 4 has investigated the pricing of three path dependent options assuming underlying dynamics given by geometric Brownian motion and dynamics given by an exponential variance gamma process. Apart from the desirable qualitative characteristics of the variance gamma process, we would expect the associated model to price observed vanilla options more consistently than the Black-Scholes model. The variance gamma process has three parameters which provide control over skewness and kurtosis whereas Brownian motion assumes zero skewness and a constant coefficient of kurtosis equaling 3. The volatility parameter σ in the Black-Scholes model is the only parameter we have available for calibrating the model. In the absence of hedging strategies, we can tentatively conclude that we would have more confidence in pricing path dependent options under variance gamma dynamics.

6.1 Comparison of Prices of Path Dependent Options

In this section we attempt to make comparisons of prices obtained for the three options of interest. It's not immediately obvious as to how a comparison can be constructed since there are many ways of performing this. We have selected one such method described below. Chapter 5 provided a discussion on how to calibrate the variance gamma model. The method worked as follows:

- Collect data on vanilla option prices in the market with different strikes. We chose call options.
- Fit a continuous curve in (C, K)-space where C is the call option price with strike K.

- Differentiate once and add 1 to obtain an implied risk-neutral distribution function of S_T at a fixed time T.
- Simulate from the distribution function and obtain sample estimates of the 2^{nd},3^{rd} and 4^{th} central moments.
- Match these sample moments to the $2^{nd}, 3^{rd}, 4^{th}$ theoretical central moments. Solve for the triplet $(\hat{\vartheta}, \hat{\nu}, \hat{\sigma})$.

The triplet $(\hat{\vartheta}, \hat{\nu}, \hat{\sigma})$ best fits the implied risk-neutral distribution given our method of calibration. Our data set obtained from $SAFEX$ consisted of 46 pairs of call option prices (on the ALSI futures index) and strikes with a fixed maturity of 84 days. We obtained the following triplet:

$$\hat{\vartheta} = -0.24065 \quad \hat{\nu} = 0.32634 \quad \hat{\sigma} = 0.13489 \tag{6.1}$$

This triplet will be used in subsequent analysis. In order to make a meaningful comparison, we adopt the same method for the Black-Scholes model. Since the Black-Scholes model has only one parameter σ, we match this up to the second sample central moment (the sample variance) of S_T. The second sample moment obtained was 967.89^2. The theoretical variance of S_T in a Black-Scholes world[1] is given by

$$V(S_T) = S_0^2 e^{2rT}(e^{\sigma^2 T} - 1)$$

where $r = 0.0748$, $S_0 = 11843$ and $T = 0.23$. This yields a unique solution for σ given by 0.167. This is the parameter used as an input in our various Black-Scholes pricing calculators.

We first examine how the two models match up with observed prices of options. We use the same data. Figure (6.1) displays a comparison of the prices of vanilla options using the variance gamma model and the Black-Scholes model with 'optimal' parameters given above. We only plot the option interpolated curves for strikes greater than 10500. Both models priced deep in-the-money call options (for strikes < 10500) extremely well.

It's clear that the Black-Scholes model calibrated well for call options struck in-the-money whilst over pricing deep out-of-the-money options . One striking feature is that the variance gamma model consistently overpriced options struck at or close to the money. The model did very well in pricing deep out-of-the-money options since the process is capturing the kurtosis in the implied risk-neutral distribution. One such quantitative measure of pricing accuracy is the sum of squared pricing errors where the sum is taken over all 46 strikes. The variance gamma model faired better in this regard although not too much cognizance

[1] under the risk-neutral measure $\mathbb{Q}$

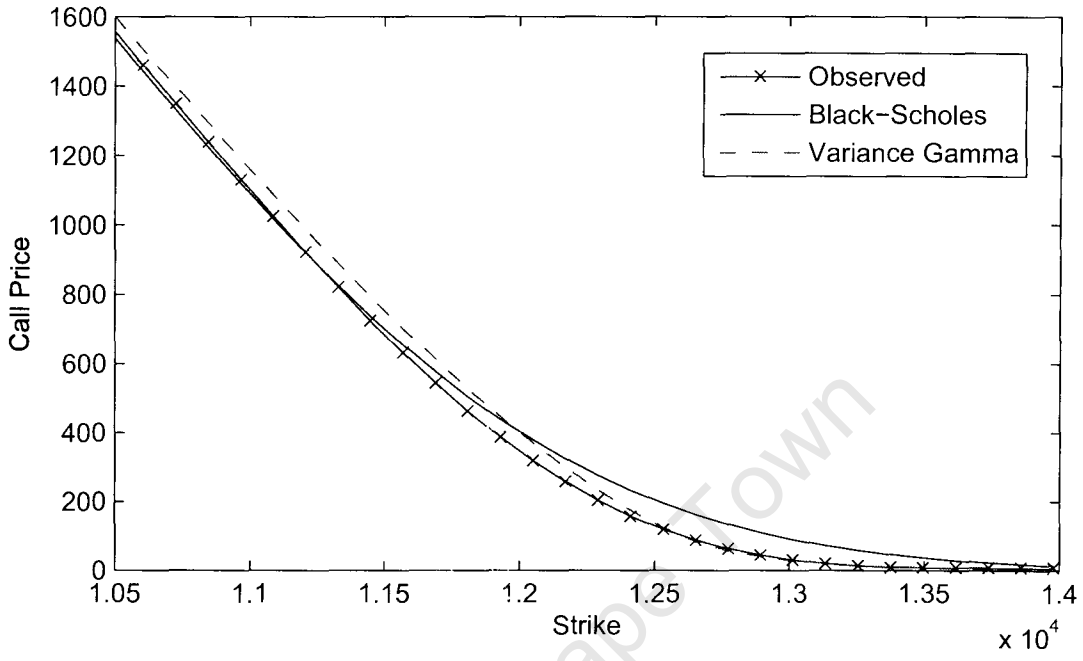

Figure 6.1: Comparison of Calibrated Models

should be taken from this. We now make a comparison of how the models faired in pricing the three path dependent options.

Up-and-Out Call Option

We implemented the TDGBS algorithm in pricing the up-and-out call option under variance gamma dynamics. In fact, the algorithm was used in pricing all three path dependent options. The algorithm was programmed in `Matlab`. Unfortunately, the TDBGS algorithm does not lend itself to vectorized code in `Matlab` and loops had to be used instead. This slowed down the simulations quite significantly when compared to sequential sampling[2] which was implemented with vectorized code. However, since the TDGBS algorithm terminates the sampling of a path if the payoff is already known, it proved to be far quicker than sequential sampling. In addition, our symmetric beta generator was more than two and a half times the speed of the `betainv` function in `Matlab`. Most importantly, the payoff is a virtually unbiased estimate of the 'continuous time' payoff.

We used the closed form solution given by Proposition 4.4.5 to price the barrier option in a Black-Scholes world. We use the same parameters as discussed above. We decided to fix the barrier level at 12500 and vary the strike ranging from 11300 to 12400. The time to maturity for these options was less than 3

[2]i.e. for a given path consisting of the same number of discretized points

months (0.23 years). On the same set of axes, we also priced the same set of options but with a longer maturity (1 year). Since the expected the value of S_1 under $\mathbb{Q}$ is larger than the barrier, it was decided to shift the barrier (for $T = 1$) to 13000. Figure (6.2) plots all four price curves for various strikes. It's clear

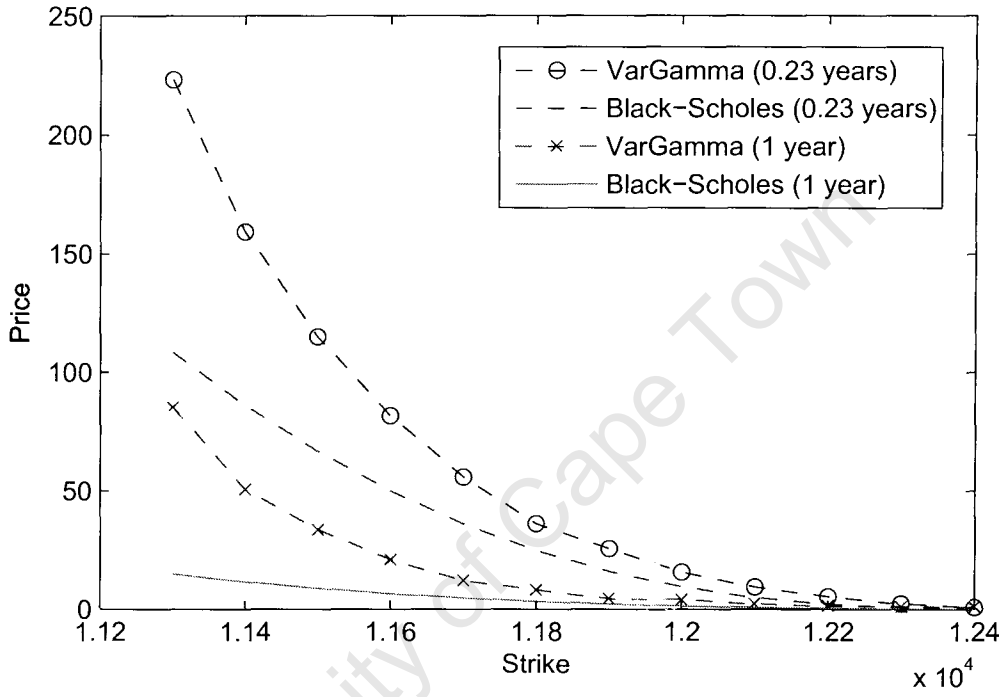

Figure 6.2: Up-and-out Call option

that the variance gamma prices are far greater than Black-Scholes for all strikes. The gap widens as the option becomes deeper in-the-money.

Lookback Put

Lookback puts do not have any strikes or barriers so plots cannot be produced. We priced the short-dated option with expiry 0.23 years and found the Black-Scholes price of 674.12 far greater than the variance gamma price of 460.53. Pricing the longer term option with expiry 1 year had a similiar result. The variance gamma price of 1089.4 is significantly less than the Black-Scholes price of 1225.5.

Double barrier call option

We implemented the TDGBS algorithm and priced the double barrier option. We chose the same set of strikes as we did with the up-and-out call option. We fixed the lower and upper barriers at 11000 and 12500 respectively. Figure (6.3) plots the set of prices for various strikes and a fixed maturity of 0.23 years. The corresponding Black-Scholes prices obtained using equation (4.27) is super-imposed

on the same set of axes. We also made comparisons for longer term options. Here we chose a maturity of 1 year. Here we pushed the lower and upper barriers out to 10500 and 13500 respectively to avoid obtaining prices that are too small and hence difficult to compare.

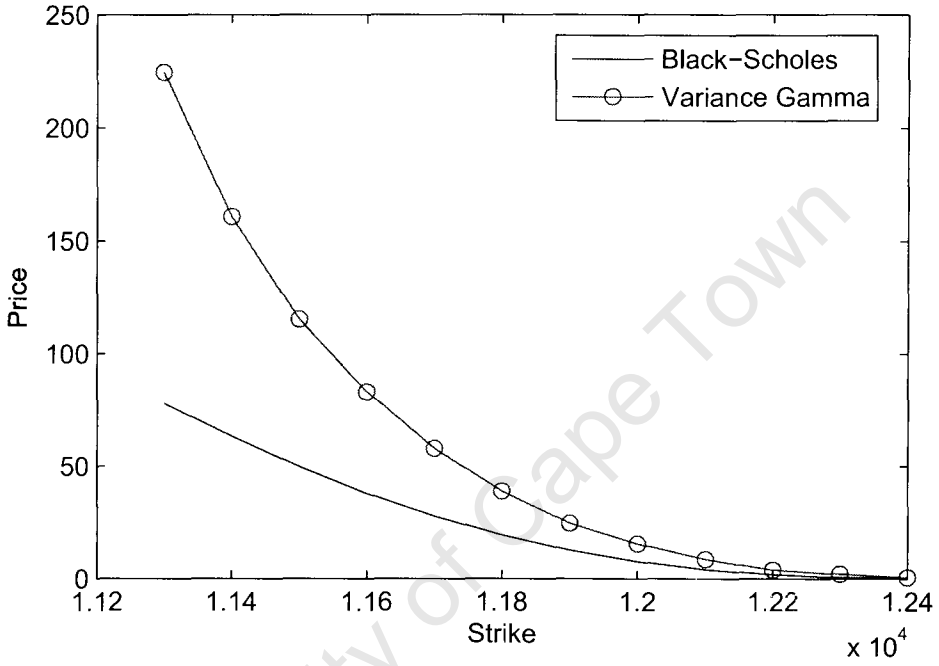

Figure 6.3: Double barrier knock-out call option. $T = 0.23$

6.2 Conclusions and Recommendations

This dissertation has investigated an alternate method of pricing options. The variance gamma process forms part of the very general class of Lévy processes. The process is pure jump which contrasts to popular diffusion based processes. We advocate the variance gamma model for its qualitative properties rather than statistical issues such as

- How well the model fits market prices of options.
- Sufficient parameters to maintain control over skewness and kurtosis in implied return distributions

Since stock prices, futures prices and indices in reality move by jumps, it seems natural to choose a model which incorporates jumps. Pure jump processes are far more difficult to analyze than diffusion processes and the relevant theory is discussed in Chapters 2 and 3. The variance gamma model has been well

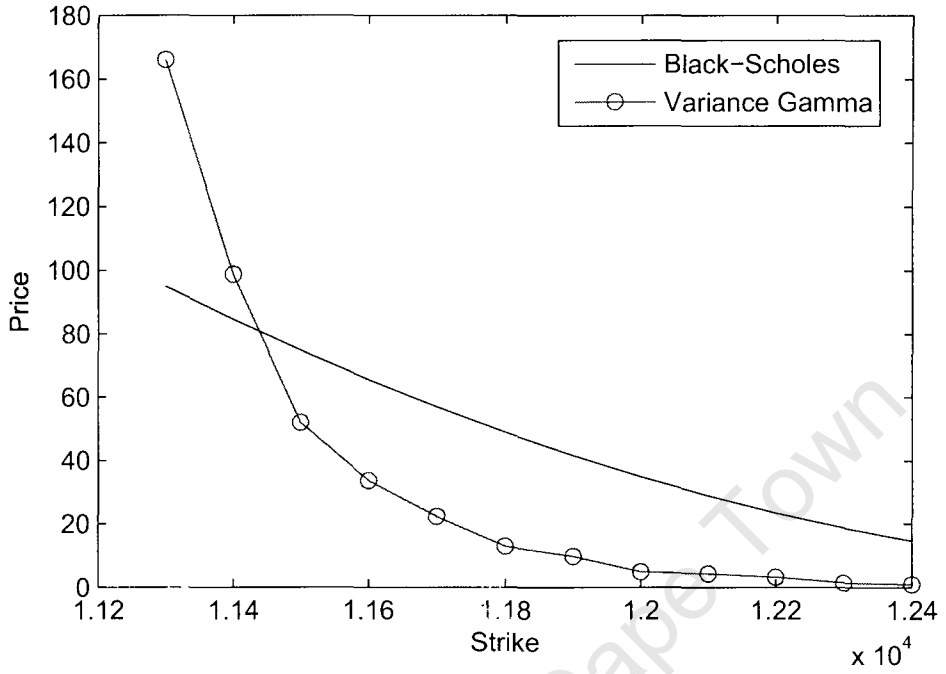

Figure 6.4: Double barrier knock-out call option. $T = 1$

studied in pricing European options where payoffs are path independent. This dissertation concerns the pricing of more exotic type options whose payoff depends on the path. Since the variance gamma process is pure jump, the pricing of barrier options is considerably more complex. We have resorted to Monte Carlo methods and implement the very efficient TDGBS algorithm developed by Avramidis & L'Ecuyer (2004). The algorithm not only dramatically improves CPU times but it becomes possible to virtually eliminate all bias in pricing barrier options. We chose three path dependent options to make a comparison with Black-Scholes prices. The results show dramatic differences in prices. It's quite remarkable how the two models calibrated equally well with the implied risk-neutral distributions yet yield such vastly different prices when applied to exotic type options. This illustrates just how important your decision is in choosing underlying dynamics. Whilst vanilla options may yield similiar prices to the Black-Scholes model, prices of exotic options are wildly different. The choice of which price is 'correct' or more rational cannot be answered without constructing approximate hedges. Recall that the variance gamma model is incomplete and we recommend future research into static and dynamic hedging of path dependent options under variance gamma dynamics.

Bibliography

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge University Press.

Avramidis, A.N. & L'Ecuyer, P. (2004). Efficient Monte Carlo and Quasi-Monte Carlo Option Pricing Under the Variance Gamma Model. Working Paper, Department of Computer Science and Operations Research, Université de Montréal.

Bertoin, J. (1998). *Lévy Processes*. Cambridge University Press.

Bjork, T. (2004). *Arbitrage Theory in Continuous Time*. Oxford University Press Inc., 2nd edn.

Bosman, P. (2003). *Static Hedging of Barrier Options: A Review of Four Methods*. Masters Dissertation, University of Cape Town.

Breeden, D. & Litzenberger, R. (1978). Prices of State-Contingent Claims implicit in Option Prices. *Journal of Business 51, 4, 621-651*.

Carr, P., Geman, H., Madan, D. & Yor, M. (2002). The Fine Structure of Asset Returns: An Empirical Investigation. *The Journal of Business, volume 75 (2002)*.

Cont, R. & Tankov, P. (2002). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, 1st edn.

Cox, D. & Miller, H. (1965). *Theory of Stochastic Processes*. Chapman and Hall, 1st edn.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Wiley, volume 1, 1st edn.

Geman, H., Madan, D. & Yor, M. (1998). Asset Prices are Brownian motion: only in Business Time, working Paper. http://citeseer.ist.psu.edu/geman98asset.html.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. New York: Springer-Verlag, 1st edn.

Hagan, P.S. & West, G. (2006). Interpolation Methods for Curve Construction. *Applied Mathematical Finance, 2006, vol. 13, issue 2, 89-129.*

Karatzas, Y. & Shreve, S. (1998). *Methods of Mathematical Finance*. Springer.

Knox, S. & Ouwehand, P. (2006). Pricing Rainbow Options: Nonparametric Methods Using Copulas. *The Investment Analysts Journal, no.64, pp 35-42.*

L'Ecuyer, P. & Simard, R. (2006). Inverting the Symmetrical Beta Distribution. *ACM Transactions on Mathematical Software, 2006, vol. 32, issue 4, 509-520.*

Madan, D. & Milne, F. (1991). Option pricing with VG Martingale Components. *Mathematical Finance. Vol. 1 , No. 4 (October 1991). 39-55.*

Madan, D. & Seneta, E. (1990). The Variance Gamma (V.G.) Model for Share Market Returns. *Journal of Business, 63, pp. 511 - 524.*

Madan, D., Carr, P. & Chang, E. (1998). The Variance Gamma Process and Option Pricing. *European Finance Review, 2, 79-105.*

Ribeiro, C. & Webber, N. (2004). Valuing Path Dependent Options in the Variance-Gamma Model by Monte Carlo with a Gamma Bridge. *The Journal of Computational Finance 7 (2),81-100.*

Sato, K.I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1st edn.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley.

Seneta, E. (2004). Fitting the Variance-Gamma Model to Financial Data. *Applied Probability Trust, Speical Vol 41A, 177-187.*

Stewart, T.J. (2003). Topics in Probability Distribution Theory, (Lecture Notes, UCT).

Tankov, P. (2004). *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*. Ph.D. thesis, Centre de Mathematiques Appliquees (Ecole Polytechnique).

West, G. (2006). South African Financial Markets, (Working Paper), http://www.finmod.co.za/safm.pdf.

William H. Press, W.T.V., Saul A. Teukolsky & Flannery, B.P. (1992). *Numerical recipes in C*. Irwin, 2nd edn.