## FINANCIAL MODELS

Thesis submitted by Konstantin Kvatch in partial fulfillment of the requirements for the Degree of

## MASTER OF SCIENCE

 inStochastic Processes in Finance and Insurance VICTORIA UNIVERSITY OF WELLINGTON

## CONTENTS

Foreword ..... 3
1 Introduction ..... 4
2 Brownian Motion ..... 7
3 Stochastic Calculus ..... 10
3.1 INTRODUCTION ..... 10
3.2 QUADRATIC VARIATION OF BROWNIAN MOTION ..... 13
3.3 ITÔ'S LEMMA AND THE ITÔ PROCESS ..... 17
3.4 EXAMPLES ..... 18
4 Risk, Arbitrage, and Pricing ..... 23
4.1 IDEAS BEHIND ARBITRAGE PRICING THEORY ..... 23
4.2 GIRSANOV THEOREM ..... 32
4.3 FUNDAMENTAL THEOREM OF ASSET PRICING ..... 36
4.4 BLACK-SCHOLES EQUATION ..... 39
4.5 BLACK-SCHOLES FORMULA ..... 41
4.6 EXAMPLES ..... 45
5 Simple Diffusion Models: Conclusion ..... 51
6 So That's It? ..... 53
7 Shortcomings of Geometric Brownian Motion ..... 54
7.1 NON-CONSTANT VOLATILITY ..... 54
7.2 EVIDENCE OF JUMPS ..... 55
7.3 HEAVY TAILS ..... 59
7.4 MARKET PRICES AGAINST THE BLACK-SCHOLES FOR- MULA ..... 61
8 Addressing the Problem: Stochastic Volatility ..... 66
8.1 INTRODUCTION ..... 66
8.2 BLACK-SCHOLES EQUATION ..... 67
8.3 RISK NEUTRAL PRICING ..... 70
8.4 CONSEQUENCES OF USING STOCHASTIC VOLATILITY ..... 76
9 Addressing the Problem: Introducing Jumps ..... 77
9.1 THE COMPOUND POISSON PROCESS ..... 77
9.2 JUMP DIFFUSION ..... 79
9.3 PRICING ..... 80
9.4 CONSEQUENCES OF USING JUMP MODELS ..... 86
10 Conclusion ..... 88
References ..... 90

This is by no means a technical or formal paper. Instead of pushing the frontiers of mathematical finance, my goal here is to look at what has been done and display my understanding of material that I believe is most relevant to contemporary mathematical finance. My aim is not to blindly treat existing theory as fact but rather, step back with a critical eye and look at the reasoning that put the current models in place. By maintaining this outlook I hope to gain and demonstrate good understanding of the financial modeling of today.

In addition I wish to show that the material should not just be left to students of popular "elite" schools, but in fact with some interest and determination this content is really beautiful and reachable to all.

I would also like to thank the teachers that have guided me through my studies and my loved ones for their support.

## Konstantin Kvatch

## 1 INTRODUCTION

The thesis will have two main parts. First, let us start with an example.

In finance, the standard version of the Black-Scholes formula is a beautiful closed form solution used to price European options. This famous formula is ingenious, but has a flaw that relegates it to something that should be admired, and perhaps not be used in the real world. It relies on the assumption that prices of shares evolve according to geometric Brownian motion. This means that we are willing to accept that extreme shocks to prices are almost impossible. Is this a realistic assumption? Of course not. The stock market crashes of 1929, 1987 are great examples to show that extreme events do happen. More recently, the 1997 Asian crisis and 2000 crash of the NASDAQ show that in addition, such events are not so rare. These jumps occur even more frequently and are larger in magnitude for share prices of individual companies.

This problem is by no means new, and a plethora of models and pricing techniques have been developed. The standard Black-Scholes formula is just one example, but this is simply illustration of the matter at hand. The process that we use to model a financial time series is of paramount importance, whether we do it for forecasting purposes or for pricing financial derivatives. If we choose to use a model that does not capture the key empirical aspects of the data, then any subsequent inference may be very unfavourably biased.

It is because of this problem that we should investigate the more standard modeling that assumes continuity and normal or log-normal distribution of financial time series. We will begin from the very basics and we will see that this is a wonderful piece of theory, deserving of the reputation it has in being simple, groundbreaking and extremely useful. This work should bring us to a position where we can evaluate a second goal.

Stochastic processes with jumps and "heavy-tails" have existed for some time, but have begun to filter through to the financial industry only recently. This lag is due to the perceived added conceptual difficulty in the introduction of such models, although we will see that this should not be the case. There is plenty of real evidence that financial time series exhibit discontinuous behaviour and that these series are far from normally or log-normally distributed. Rather than looking at standard models as correct, and jump or stochastic volatility models as complicated, we should look upon standard models as educational but not sufficient for the real world. Stochastic volatility or jump models should instead be viewed as natural.

The theme of the thesis is the importance of choosing a correct model for the underlying process. Although we may speak of the implications of some models to hedging, we will not actually look at specific hedging techniques. The particular aspect of pricing is also not considered in full scope although we will see the Black-Scholes pricing formula. We will consider that the main problem is to specify the model correctly where the method of pricing is a subsequent technicality. In examples we may take pricing tools like MonteCarlo simulation as a given.

We will not strive for full generality or formality, but rather take a physical approach and aim for clarity and understanding. Let us now move on to the beginning, with the introduction of our primary source of randomness.

## Simple Diffusion Models

## 2 BROWNIAN MOTION

The reason that mathematical finance exists is that there is uncertainty about the future. We attempt to model what really goes on in the financial world. We need to begin by providing some sort of fundamental engine to drive the stochasticity in these models. To do this we will use Brownian motion.

The discovery of this process is usually accredited to Robert Brown who in 1827 observed pollen particles under a microscope. He noticed that the particles had restless motion that seemed to be completely random. Repeating this experiment with dust particles he ruled out the movement being due to some process related to life. The idea is that the dust is continuously colliding with gas molecules which cause it to drift around unpredictably.

Several notable mathematicians went on to describe the process independently, but for our purpose the most notable is Louis Bachelier. He is regarded as the first person to use the concept of Brownian motion with application in finance. In his 1900 paper La Théorie de la Spéculation he used Brownian motion to describe the evolution of stock prices.

The idea is fully justifiable. In this case, the price of the stock is the dust particle, the traders are the gas molecules, and the acts of buying and selling are the collisions.

Now we should look at the properties of Brownian motion. In financial circles this process may also be called a Wiener process. Here is a rough definition:

A continuous time process $W_{t}$ is called a Brownian motion or Wiener process if it is characterised by the following:

1. $W_{0}=0$
2. Non-overlapping increments $W_{t_{2}}-W_{t_{1}}$ and $W_{t_{4}}-W_{t_{3}}$ are independent.
3. For $t_{2}>t_{1}$, the increment $W_{t_{2}}-W_{t_{1}} \sim N\left(0, t_{2}-t_{1}\right)$


Figure 1. A path of Brownian motion
At first glance this graph does indeed look like a figure displaying a stock price or exchange rate, but there is an important element that needs to be noticed. The above process moves below zero, and it is obvious that stock prices or exchange rates do not do this. Therefore we need to go further and see how this problem has been addressed. We cannot use this particular model, but the Wiener process is nevertheless very crucial, as for us it will be an engine to drive the randomness in the models we see later. In particular the property of independent increments is of value because it
agrees with the efficient market hypothesis seen in finance. This hypothesis states that markets are informationally efficient and prices reflect all known information, thus a steady profit cannot be made unless by luck. If we had non-independent increments in our engine for randomness, it would imply that prices do not reflect all known information and the market could be outperformed. This would disagree with the efficient market hypothesis. Let us now move on to see how more useful models have been made by using Brownian motion. We need to investigate Stochastic Calculus.

## 3 STOCHASTIC CALCULUS

### 3.1 INTRODUCTION

At the end of the last section we may have said, why not take the exponential of the Wiener process? We obtain something that is always positive and closely resembles share prices. Furthermore we know the distribution of this transformed variable and it is not difficult to work with it. We don't need any differential equations or integrals.

This may be enough to model a share price, but this is just one process and one model. Let us look at some other examples.

Perhaps we want to model an interest rate. A process like this is usually regarded as stationary so let us begin with a simple stationary model, say the $\operatorname{AR}(1)$ process in discrete-time form. However, since we want to use this model for financial data, where the step sizes are tiny, let us derive the continuous time version.

Example (Continuous analogy to the $A R(1)$ process) Let us define our $\operatorname{AR}(1)$ process as $X_{n+1}=a+\mu X_{n}+Z_{n+1}$ where the $Z_{i}$ are standard normal and $|\mu|<1$. In addition we know this process has a long-run mean of $a /(1-\mu)$. To make this model in continuous time we must consider what happens when we take smaller time steps. Instead of having time steps of width 1 , let us make them of width $1 / m$ and $m \rightarrow \infty$. To make this step we need to keep in mind that the process must retain its characteristics. That
means for example $X_{n+1}=X_{n+\frac{1}{2}+\frac{1}{2}}=X_{n+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}} \cdots$

$$
\begin{aligned}
X_{n+1} & =a+\mu X_{n}+Z_{n+1} \\
X_{n+\frac{1}{2}} & =\frac{a}{\mu^{1 / 2}+1}+\mu^{1 / 2} X_{n}+Z_{n+\frac{1}{2}}^{\prime} \\
X_{n+\frac{1}{3}} & =\frac{a}{\mu^{2 / 3}+\mu^{1 / 3}+1}+\mu^{1 / 3} X_{n}+Z_{n+\frac{1}{3}}^{\prime \prime} \\
X_{n+\frac{1}{m}} & =\frac{a}{\sum_{i=0}^{m-1}\left(\mu^{1 / m}\right)^{i}}+\mu^{1 / m} X_{n}+Z_{n+\frac{1}{m}}^{\prime \prime \prime} \\
X_{n+\frac{1}{m}} & =\frac{a\left(1-\mu^{1 / m}\right)}{1-\mu}+\mu^{1 / m} X_{n}+Z_{n+\frac{1}{m}}^{\prime \prime \prime}
\end{aligned}
$$

Notice that the random component $Z$ changes as well. After all, the process must retain its characteristics over the time step of width 1 . Then for smaller time steps the random component must also be "smaller" (in the sense that these smaller components should add up to the original $Z_{n+1}$ ). This amounts to a reduction in variance, and in the continuous case, it turns out that $Z_{n+1}$ becomes $d W_{t}$. Back to the construction, let us now consider the increment of $X$.

$$
\begin{aligned}
X_{n+\frac{1}{m}}-X_{n} & =\frac{a\left(1-\mu^{1 / m}\right)}{1-\mu}+\mu^{1 / m} X_{n}-X_{n}+Z_{n+\frac{1}{m}}^{\prime \prime \prime} \\
\Delta X_{n} & =\frac{a\left(1-\mu^{1 / m}\right)}{1-\mu}-\left(1-\mu^{1 / m}\right) X_{n}+Z_{n+\frac{1}{m}}^{\prime \prime \prime} \\
\Delta X_{n} & =\left(\frac{a}{1-\mu}-X_{n}\right)\left(1-\mu^{1 / m}\right)+Z_{n+\frac{1}{m}}^{\prime \prime \prime}
\end{aligned}
$$

We notice that the long-run mean of the $\operatorname{AR}(1)$ process appears in the equation, call this level c. We also notice the expression $\left(1-\mu^{1 / m}\right)$ which tends to 0 and retains dependence on $\mu$. This determines how fast the process returns to its mean. We can say $\left(1-\mu^{1 / m}\right)=b d t$ and since $d t$ goes to 0 arbitrarily, we can specify it as $d t=1 / m$ and $m \rightarrow \infty$. The normal random variable changes only in variance as we take smaller time steps, and in fact if we take increments of a standard Brownian motion, in the continuous time case we preserve the characteristics of the $\mathrm{AR}(1)$ process. Finally we can write the expression using continuous time notation. This is often called the Ornstein-Uhlenbeck process.

$$
\begin{aligned}
& d X_{t}=b\left(c-X_{t}\right) d t+d W_{t} \text { in continuous time } \\
& \text { with } c=\frac{a}{1-\mu}
\end{aligned}
$$

We still need to find the limit of $m\left(1-\mu^{1 / m}\right)$. This requires a couple clever steps as well. Notice especially, the change of variable $1 / m=x$.

$$
\text { and } \begin{aligned}
b & =\lim _{m \rightarrow \infty} m\left(1-\mu^{1 / m}\right) \\
& =\lim _{x \rightarrow 0} \frac{\left(1-\mu^{x}\right)}{x} \\
& =\lim _{x \rightarrow 0} \frac{d\left(1-\mu^{x}\right) / d x}{d(x) / d x} \text { using L'Hôpital's rule } \\
& =\lim _{x \rightarrow 0}-\log \mu e^{x \log \mu} \\
& =-\log \mu
\end{aligned}
$$

We discover that a differential equation appears even after the introduction of such an elementary process. Furthermore, not only is it a differential equation, but it has a random component $\left(d W_{t}\right)$. Equations of this type are called stochastic differential equations. We will discover later that to solve this type of equation, stochastic calculus is paramount.

Example (Stock holdings process) This is a very natural example. Let us say we have an asset price that evolves according to some random process $\left\{X_{t}\right\}$. We also have another process, the amount of the asset that we hold, $Y_{t}$. So, it time $t_{0}$ we buy $Y_{t_{0}}$ of the asset. Then, at time $t_{1}$ the asset price has moved and we have a change in wealth $Y_{t_{0}} X_{t_{1}}-Y_{t_{0}} X_{t_{0}}=Y_{t_{0}}\left(X_{t_{1}}-X_{t_{0}}\right)$. Then at time $t_{1}$ we want to change our position to $Y_{t_{1}}$. After another time step we again have a change in wealth $Y_{t_{1}} X_{t_{2}}-Y_{t_{1}} X_{t_{1}}=Y_{t_{1}}\left(X_{t_{2}}-X_{t_{1}}\right)$. If we continue this, our total wealth change at $t_{i+1}$ is:
$Y_{t_{0}}\left(X_{t_{1}}-X_{t_{0}}\right)+Y_{t_{1}}\left(X_{t_{2}}-X_{t_{1}}\right)+\ldots+Y_{t_{i}}\left(X_{t_{i+1}}-X_{t_{i}}\right)=\sum_{k=0}^{t_{i}} Y_{t_{k}}\left(X_{t_{k+1}}-X_{t_{k}}\right)$
Once again, we want to take this to continuous time:

$$
\lim _{\Delta t \rightarrow 0} \sum_{k=0}^{t_{i}} Y_{t_{k}}\left(X_{t_{k+1}}-X_{t_{k}}\right)=\int_{0}^{t} Y_{s} d X_{s}
$$

Now we have an integral, and it is extremely important to note that the integrator is random. This is a stochastic integral.

These two examples should demonstrate that differentials and integrals are unavoidable. We need to have some tools to operate with such equations. A calculus is necessary and we will see that it is not only necessary but also very desirable because we open ourselves up to a myriad of representations. There is a mass of processes that cannot be described explicitly and need to be written in differential or integral form. Now someone might argue that this is no problem since we have all learnt calculus at school, but there is an extremely crucial difference between ordinary calculus and stochastic calculus. We should investigate:

### 3.2 QUADRATIC VARIATION OF BROWNIAN MOTION

Let us say we want to find the infinitesimal difference of a simple, smooth function, for example $f(t)=t^{3}$. To do this we can use the Taylor expansion.

$$
\begin{aligned}
\Delta f(t) & =\frac{d f(t)}{d t} d t+\frac{d^{2} f(t)}{2 d t^{2}}(d t)^{2}+\frac{d^{3} f(t)}{3!d t^{3}}(d t)^{3}+\ldots \\
\Delta t^{3} & =3 t^{2} d t+3 t(d t)^{2}+(d t)^{3}+\ldots \\
\Delta t^{3} & =3 t^{2} d t
\end{aligned}
$$

Terms of order $(d t)^{2}$ and higher are negligible so we can say $d t^{3}=3 t^{2} d t$. To demonstrate, we know that dt is an infinitesimal increment. If we take for example $d t=0.01$ and decrease it by a factor of 10 to $0.001,(d t)^{2}$ will decrease by a factor of 100 to 0.00001 , so in comparison to $d t$, $(d t)^{2}$ will become less and less significant as $d t$ goes to 0 . Obviously all terms of higher order are even less significant. Let us now consider the same function, but with a different input. Instead of taking a function of time, we will take a
function of the Wiener process: $f\left(W_{t}\right)=W_{t}^{3}$.

$$
\begin{align*}
\Delta f\left(W_{t}\right) & =\frac{d f\left(W_{t}\right)}{d W_{t}} d W_{t}+\frac{d^{2} f\left(W_{t}\right)}{2!d W_{t}^{2}}\left(d W_{t}\right)^{2}+\frac{d^{3} f\left(W_{t}\right)}{3!d W_{t}^{3}}\left(d W_{t}\right)^{3}+\ldots \\
\Delta W_{t}^{3} & =3 W_{t}^{2} d W_{t}+3 W_{t}\left(d W_{t}\right)^{2}+\left(d W_{t}\right)^{3}+\ldots \\
\Delta W_{t}^{3} & =3 W_{t}^{2} d W_{t}
\end{align*}
$$

Is this correct? It turns out that we cannot apply the same reasoning as before since $\left(d W_{t}\right)^{2}$ is not negligible. This may be already obvious since $E\left[\left(d W_{t}\right)^{2}\right]$ is the variance of the Wiener process increments and in our definition of the Wiener process we stated that $\operatorname{Var}\left[d W_{t}\right]=d t$. It can be shown that while $\int_{0}^{t} d W_{s}=W_{t}$ is random, $\int_{0}^{t}\left(d W_{s}\right)^{2}$ is not.

For a heuristic understanding it is enough to consider that the variance of Brownian motion is equal to $d t$, but this is the cornerstone of Stochastic Calculus and we should examine the proof.

Theorem 3.2.1. (Quadratic Variation of Brownian Motion) Consider a partition $0<\ldots<t_{k}<\ldots<t_{n}<T$ with $n \rightarrow \infty$ and $t_{k+1}-t_{k}=\Delta t_{k} \rightarrow 0$. Then:

$$
\begin{array}{r}
\lim _{\Delta t_{k} \rightarrow 0} E\left[\sum_{k}\left(W_{t_{k+1}}-W_{t_{k}}\right)^{2}-T\right]^{2}=0 \\
\sum_{k}\left(W_{t_{k+1}}-W_{t_{k}}\right)^{2} \xrightarrow{\text { a.s }} T \tag{2}
\end{array}
$$

Proof:

$$
\begin{gathered}
\text { Set } \quad V_{n}=\sum_{k}\left(W_{t_{k+1}}-W_{t_{k}}\right)^{2}=\sum_{k}\left(\Delta W_{t_{k}}\right)^{2} \\
V_{n}-T=\sum_{k}\left[\left(\Delta W_{t_{k}}\right)^{2}-\Delta t_{k}\right]
\end{gathered}
$$

Taking the second moment of this quantity

$$
E\left[V_{n}-T\right]^{2}=E\left[\sum_{k} \sum_{j}\left[\left(\Delta W_{t_{k}}\right)^{2}-\Delta t_{k}\right]\left[\left(\Delta W_{t_{j}}\right)^{2}-\Delta t_{j}\right]\right]
$$

For non-overlapping increments (where $k \neq j$ ) we have

$$
E\left[\left(\Delta W_{t_{k}}\right)^{2}-\Delta t_{k}\right] E\left[\left(\Delta W_{t_{j}}\right)^{2}-\Delta t_{j}\right]=0
$$

This is simply due to the properties of Brownian motion. Non-overlapping increments are independent so the expectation can be taken inside, and $E\left[\Delta W_{t_{k}}\right]^{2}=\Delta t_{k}$. Hence

$$
\begin{aligned}
E\left[V_{n}-T\right]^{2} & =E \sum_{k}\left[\left(\Delta W_{t_{k}}\right)^{2}-\Delta t_{k}\right]^{2} \\
& =\sum_{k} E\left[\Delta t_{k}\left(\frac{\left(\Delta W_{t_{k}}\right)^{2}}{\Delta t_{k}}-1\right)\right]^{2} \\
& =\sum_{k} E\left[\Delta t_{k}\left(Z^{2}-1\right)\right]^{2} \text { where } Z \sim N(0,1) \\
& =\sum_{k}\left(\Delta t_{k}\right)^{2} E\left[Z^{2}-1\right]^{2}
\end{aligned}
$$

But $E\left[Z^{2}-1\right]^{2}$ is some constant, call it $C$. Hence

$$
\begin{aligned}
E\left[V_{n}-t_{n}\right]^{2} & =C \sum_{k}\left(\Delta t_{k}\right)^{2} \\
& \leq C \max _{k}\left(\Delta t_{k}\right) \sum_{k} \Delta t_{k} \\
& \leq C \max _{k}\left(\Delta t_{k}\right) T \rightarrow 0 \text { if } \Delta t_{k} \rightarrow 0
\end{aligned}
$$

This proves part (1) of the theorem. To prove part (2) we need to specify our partition a little bit. Instead of saying it is arbitrary, let us say the increments are now all of equal width and $\Delta t_{k}=\epsilon_{i} / i^{2}$ with $i \rightarrow \infty$, remembering that $\Delta t_{k}$ still goes to 0 (this means that $\Delta t_{k} \rightarrow 0$ faster than $1 / i^{2} \rightarrow 0$ ).

$$
\mathbb{P}\left(\left(V_{n}-T\right)^{2}>2 \epsilon_{i}\right) \leq \frac{E\left[V_{n}-T\right]^{2}}{2 \epsilon_{i}} \leq \frac{C \max _{k}\left(\Delta t_{k}\right) T}{2 \epsilon_{i}} \leq \frac{C \epsilon_{i} T}{2 \epsilon_{i} i^{2}}=\frac{C^{\prime}}{i^{2}}
$$

Since $\sum_{i} \frac{C^{\prime}}{i^{2}}<\infty$, the sum of probabilities in the above expression is finite. Applying the Borel-Cantelli lemma, the probability that $\left(V_{n}-T\right)^{2}>2 \epsilon_{i}$ for infinitely many $i$ is zero. Since $\epsilon_{i} \rightarrow 0$, this means that almost surely $(V-T) \rightarrow 0$. In other words

$$
\sum_{k}\left(W_{t_{k+1}}-W_{t_{k}}\right)^{2} \xrightarrow{\text { a.s. }} T
$$



Figure 2. This plot shows 3 approximations of $\int_{O}^{1}\left(d W_{s}\right)^{2}$ for different time grids. The lightest line has $\Delta t=1 / 50$, the darker line has $\Delta t=1 / 200$, and the darkest line has $\Delta t=1 / 6400$. Notice that the darkest line is almost identical to $f(t)=t$.

Returning to our previous example, the difference $(\dagger)$ should be as below:

$$
\Delta W_{t}^{3}=3 W_{t}^{2} d W_{t}+3 W_{t} d t
$$

The fundamental difference between ordinary calculus and stochastic calculus is this second order term. In ordinary calculus it is negligible, whereas in stochastic calculus it is not. It is also worth noting that it can be shown $\int_{o}^{t} d W_{s} d s=0$. With this knowledge we can move on to a very famous and useful formula in continuous time finance.

### 3.3 ITÔ'S LEMMA AND THE ITÔ PROCESS

We reasoned earlier that Brownian motion itself will not likely be a sufficient model for financial time series. By taking functions of it, we can however use it as an engine of randomness. Since it is most convenient to represent financial time series in differential form, we need a tool to obtain the differential of a given function of Brownian motion. So let us generalise what we discovered in the last section.

$$
\begin{aligned}
d f\left(W_{t}\right) & =\frac{d f\left(W_{t}\right)}{d W_{t}} d W_{t}+\frac{1}{2} \frac{d^{2} f\left(W_{t}\right)}{d W_{t}^{2}}\left(d W_{t}\right)^{2} \\
& =\frac{d f\left(W_{t}\right)}{d W_{t}} d W_{t}+\frac{1}{2} \frac{d^{2} f\left(W_{t}\right)}{d W_{t}^{2}} d t
\end{aligned}
$$

At this stage we should also consider that we may not want to begin with a standard Brownian motion, but perhaps alter the variance and drift. We introduce the Itô process which is simply a generalised Brownian motion, where we have time and/or state dependent parameters.

$$
d X_{t}=\mu\left(X_{t}, t\right) d t+\sigma\left(X_{t}, t\right) d W_{t}
$$

We may also write this using shorter notation:

$$
d X_{t}=\mu_{X} d t+\sigma_{X} d W_{t}
$$

So our standard Brownian motion is an Itô process with $\mu\left(X_{t}, t\right)=0$ and $\sigma\left(X_{t}, t\right)=1$. If we have $\mu\left(X_{t}, t\right)=a$ where $a$ is a constant, we simply add drift to our standard Brownian motion. Changing $\sigma\left(X_{t}, t\right)$ alters the variance of the process. The next step is to consider what happens when we take the function of this process. As is often useful in finance, we can go one step further and consider a function of the process, and time. Finally here is the lemma:

Lemma 3.3.1. (Itô's Lemma) For an Itô process $d X_{t}=\mu_{X} d t+\sigma_{X} d W_{t}$, and any twice continuously differentiable function $f$, the process $f\left(X_{t}, t\right)=Y_{t}$ is also an Itô process satisfying $d Y_{t}=\mu_{Y} d t+\sigma_{Y} d W_{t}$.

We make an informal derivation by once again using the Taylor expansion.

$$
\begin{aligned}
d f\left(X_{t}, t\right) & =\frac{\partial Y_{t}}{\partial X_{t}} d X_{t}+\frac{\partial Y_{t}}{\partial t} d t+\frac{1}{2} \frac{\partial^{2} Y_{t}}{\partial X_{t}^{2}}\left(d X_{t}\right)^{2} \\
& =\left(\mu_{X} \frac{\partial Y_{t}}{\partial X_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma_{X}^{2} \frac{\partial^{2} Y_{t}}{\partial X_{t}^{2}}\right) d t+\sigma_{X} \frac{\partial Y_{t}}{\partial X_{t}} d W_{t} \\
& =\mu_{Y} d t+\sigma_{Y} d W_{t}
\end{aligned}
$$

In principle, we can take a function of more variables, and we do not need to consider a function of $X_{t}$ and $t$. We can just as well take a function of two Itô processes. To adjust, all we need to do is use the Taylor expansion with the appropriate number of variables and investigate what happens to the infinitesimal terms. We will see examples in the next section and later.

### 3.4 EXAMPLES

Example (Geometric Brownian motion) At the end of the last chapter we argued, why not simply take the exponential of Brownian motion to get a process that is always positive. This we thought, might be a good model of stock price dynamics. With this in mind let us consider a certain Itô process.

$$
d S_{t}=\mu S_{t} d t+\sigma S_{t} d W_{t}
$$

As it stands, the $S_{t}$ in the drift and volatility terms make it difficult to integrate. We use Itô's Lemma to help us solve it. Let us see what happens if we take $Y_{t}=\log S_{t}$.

$$
\begin{aligned}
d \log S_{t}=d Y_{t} & =\frac{d \log \left(S_{t}\right)}{d S_{t}} d S_{t}+\frac{d^{2} \log \left(S_{t}\right)}{2 d S_{t}^{2}}\left(d S_{t}\right)^{2} \\
& =\frac{1}{S_{t}}\left(\mu S_{t} d t+\sigma S_{t} d W_{t}\right)-\frac{1}{S_{t}^{2}} \sigma^{2} S_{t}^{2} d t \\
& =\left(\mu-\frac{\sigma^{2}}{2}\right) d t+\sigma d W_{t} \\
& =a d t+\sigma d W_{t}
\end{aligned}
$$

But this is now simply a Brownian motion with drift. Since taking the logarithm of $S_{t}$ gives us a Brownian motion, taking the exponential of a Brownian motion must give us $S_{t}$. This is exactly what we were considering for a stock price model. It turns out that this is indeed the first and most prominent model for a stock encountered in mathematical finance and it is called geometric Brownian motion. We will use this as the model of choice to explain stock price dynamics for the remainder of this chapter. Before continuing let us find the solution to this equation.

$$
\begin{aligned}
\int_{0}^{t} d \log S_{u} & =\int_{0}^{t}\left(\left(\mu-\frac{\sigma^{2}}{2}\right) d u+\sigma d W_{u}\right) \\
\log S_{t}-\log S_{0} & =\left(\left(\mu-\frac{\sigma^{2}}{2}\right) t+\sigma W_{t}\right) \\
\log S_{t} & =\log S_{0}+\left(\left(\mu-\frac{\sigma^{2}}{2}\right) t+\sigma W_{t}\right) \\
S_{t} & =S_{0} e^{\left(\mu-\frac{1}{2} \sigma^{2}\right) t+\sigma W_{t}}
\end{aligned}
$$

Of course it is worth noting that $S_{t}$ is a random variable (log-normal to be exact). We can also consider the logarithm of this, which is normally distributed. Finding the mean and variance is elementary:

$$
\log S_{t} \sim N\left(\log S_{0}+\left(\mu-\frac{\sigma^{2}}{2}\right) t, \sigma^{2} t\right)
$$

Example (Ornstein-Uhlenbeck process solution) In the introduction to this section we discovered that this is the continuous-time analogy of the $\operatorname{AR}(1)$ process. To find the solution to this differential equation we must again use Itô's Lemma. We must first eliminate the state variable in the equation to allow easy integration, so we should try to find some transformation that would address this problem. It turns out that if we set $f\left(X_{t}, Y_{t}\right)=f\left(X_{t}, t\right)=$ $Z_{t}=X_{t} e^{b t}$ we get the desired result.

$$
\begin{aligned}
d X_{t} & =b\left(\mu-X_{t}\right) d t+\sigma d W_{t} \\
d Z_{t} & =\frac{\partial f\left(X_{t}, t\right)}{\partial X_{t}} d X_{t}+\frac{\partial f\left(X_{t}, t\right)}{\partial t} d t+\ldots \\
& =e^{b t}\left(b\left(\mu-X_{t}\right) d t+\sigma d W_{t}\right)+b X_{t} e^{b t} d t \\
& =e^{b t}\left(b \mu d t+\sigma d W_{t}\right)
\end{aligned}
$$



Figure 3. Left is a path of standard Brownian motion with single standard deviation bands. Right is a path of geometric Brownian motion generated by the process on the left, with $\mu=0.05$ and $\sigma=0.2$, and the corresponding bands.

$$
\begin{aligned}
\int_{0}^{t} d Z_{s} & =\int_{0}^{t} e^{b s}\left(b \mu d s+\sigma d W_{s}\right) \\
\int_{0}^{t} d Z_{s} & =b \mu \int_{0}^{t} e^{b s} d s+\sigma \int_{0}^{t} d W_{s} \\
Z_{t}-Z_{0} & =b \mu\left(\frac{e^{b t}}{b}-\frac{1}{b}\right)+\sigma \int_{0}^{t} e^{b s} d W_{s} \\
X_{t} e^{b t}-X_{0} & =\mu e^{b t}-\mu+\sigma \int_{0}^{t} e^{b s} d W_{s} \\
X_{t} & =e^{-b t} X_{0}+\mu\left(1-e^{-b t}\right)+\sigma \int_{0}^{t} e^{b(s-t)} d W_{s}
\end{aligned}
$$

Given the normality of the driving Brownian motion, this is also a normal random variable. The mean is simple to find. To find the variance we have to remember that non-overlapping increments of Brownian motion are independent and hence $E\left[\sigma \int_{0}^{t} e^{b(s-t)} d W_{s}\right]^{2}=\sigma \int_{0}^{t} e^{2 b(s-t)} E\left[d W_{s}\right]^{2}$.

$$
X_{t} \sim N\left(e^{-b t} X_{0}+\mu\left(1-e^{-b t}\right), \sigma^{2} \frac{1-e^{-2 b t}}{2 b}\right)
$$

We showed that this is a continuous-time analogy of the $\operatorname{AR}(1)$ process, so it
is also stationary. We can take $t \rightarrow \infty$ to obtain the stationary distribution:

$$
X_{t} \sim N\left(\mu, \frac{\sigma^{2}}{2 b}\right)
$$



Figure 4. An Ornstein-Uhlenbeck process generated by the same process as in figure 5 with $\mu=1, b=0.4, \sigma=\sqrt{2 b}$ and an initial value of 0 . Single standard deviation bands are also shown

Example ( $A$ Numéraire) In French this word means money, or face value, and it's meaning in mathematics is exactly that. It is a unit of measure for the value of something. For example the US dollar is used as numéraire for goods in the USA and the Euro for goods in Europe. Let us assume that the exchange rate between the two currencies is $\$ 1=€ 0.64$. A burger in the USA costs $\$ 2.50$. Taking the Euro as numéraire the price becomes $2.40 \times 0.64=€ 1.54$. The only condition we require is that the numéraire is always positive.

We can look at another useful change of numéraire. We will find out later why it is useful, but let us just calculate it for now. Consider a stock price evolving according to geometric Brownian motion and a risk-less bond.

$$
\begin{aligned}
d S_{t} & =\mu S_{t} d t+\sigma S_{t} d W_{t} \\
d \beta_{t} & =r \beta_{t} d t \quad \text { notice the lack of random component (no risk) }
\end{aligned}
$$

We will use the bond as numéraire, so in effect we will be calculating how
many bonds the stock is worth rather than how many dollars. That is $f\left(S_{t}, \beta_{t}\right)=Y_{t}=S_{t} / \beta_{t}$, and we are looking for $d\left(S_{t} / \beta_{t}\right)$. To find it we must again use Itô's Lemma.

$$
\begin{aligned}
d\left(Y_{t}\right) & =\frac{\partial f\left(S_{t}, \beta_{t}\right)}{\partial S_{t}} d S_{t}+\frac{\partial f\left(S_{t}, \beta_{t}\right)}{\partial \beta_{t}} d \beta_{t}+\ldots \\
& =\frac{1}{\beta_{t}} d S_{t}-\frac{S_{t}}{\beta_{t}^{2}} d \beta_{t} \\
& =\frac{1}{\beta_{t}}\left(\mu S_{t} d t+\sigma S_{t} d W_{t}\right)-\frac{S_{t}}{\beta_{t}^{2}} r \beta_{t} d t \\
& =\frac{S_{t}}{\beta_{t}}\left((\mu-r) d t+\sigma d W_{t}\right) \\
d Y_{t} & =(\mu-r) Y_{t} d t+\sigma Y_{t} d W_{t}
\end{aligned}
$$

This is again a geometric Brownian motion. In fact if we take $\beta_{0}=1$ this gives us the stock price dynamics discounted by the risk free rate. The drift term corresponds to the risk premium of the stock.

## 4 RISK, ARBITRAGE, AND PRICING

### 4.1 IDEAS BEHIND ARBITRAGE PRICING THEORY

This is perhaps the most important sub-section in the first part of the thesis as it explains the reasoning behind arbitrage-free pricing in very simple terms. The real theory should be much easier to understand after reading this section.

Let us first recall some simple fundamentals of finance. The first is the timevalue of money. It is better to have $\$ 1$ now than $\$ 1$ in the future because $\$ 1$ now can be invested and a profit can be made. A government bond is usually used as a tool to measure this value (of course not any government bond, but one issued by some roughly speaking 'developed' and stable government. Bonds are usually given a rating to signify their reliability). Those with the best rating are compounded and discounted at the risk-free rate. Basically we assume that there is no chance of default and we are guaranteed the future return. So if we consider one time step, simple geometric compounding, and a risk-free rate of $2.5 \%$ then $\$ 10$ invested now returns $\$ 10(1+0.025)=\$ 10.25$ at the end of the time period. This is illustrated by the blue line in the figure 5 below.

Of course there are assets that are not risk-free. Take for example a stock. Let us say that right now it is worth $\$ 10$, but in one period's time we are not sure what it will be worth (hence the risk involved). It will either go up to $\$ 11.50$ or down to $\$ 9.50$ with equal probability. The expected value is however $\$ 10.50$ (red dotted line in the figure 5 ). We notice that the expected returns from the two assets differ by $\mathbf{\Delta}-\boldsymbol{\square}$. Since the initial price is the same, the rates of return are different. The discount factor for the risky asset is called the risk-adjusted discount rate. For this particular asset it is
$(10.50-10) / 10=5 \%$. The difference between the risk-adjusted discount rate and the risk-free rate is called the risk premium. This can be interpreted as the reward for bearing risk.

- (risk-adjusted discount rate) $-($ risk-free rate $)=$ risk premium

All assets have this risk-adjusted discount rate and we use this rate to discount the expected future return of an asset. This discounted value gives a fair current price of the asset.


Figure 5. Risky and risk-less assets.

So, what does this have to do with derivative pricing?

Let us consider the most simple derivative, a European call option on the stock in the current example. That is, the option gives us the right to buy one unit of stock at the end of the time period for some given strike price $(K)$, let us say $K=\$ 10$. At the end of the time period one of two things can happen:

- $S_{1}=\$ 11.50:$ We exercise the option, and buy one unit of stock for $\$ 10$. We can immediately sell it for the market price at $\$ 11.50$ and make $\$ 1.50$.
- $S_{1}=\$ 9.50$ : We let the option expire since to buy one unit of stock at $\$ 10$ would be useless considering that we could do so for less on the market.

We can also calculate the expected payoff from holding the option since we know the probabilities with which the stock price will change.

$$
\begin{aligned}
E\left[C_{1}\right] & =E\left[\max \left(S_{1}-K, 0\right)\right] \\
& =0.5(11.50-10)+0.5(0) \\
& =\$ 0.75
\end{aligned}
$$

This is where we hit a wall. After all, what we are after is the current option price. To calculate it, we need to discount $E\left[C_{1}\right]$ at the appropriate risk-adjusted discount rate, call it $\rho$.

$$
C_{0}=\frac{E\left[C_{1}\right]}{1+\rho}
$$

It is important to note that although the payoff of the option depends completely on the stock, their risk-adjusted discount rates are very different since the payoffs and initial prices are different. We also cannot simply use the risk-free rate. So how do we value this option?

There are two approaches to solving this problem:

## 1. Absence of arbitrage:

Arbitrage is a free lunch. It's a cash inflow with no risk involved.

We argue that if we can make some asset that has identical cash flows to the option, then the price of the option should be the same as the price of this new asset. Otherwise we exploit the difference in prices to make a risk-free profit i.e. arbitrage. So how can we make this asset with identical cash flows? Consider a portfolio that consists of stocks and bonds:

$$
\begin{aligned}
\text { Portfolio } & =a S_{0}+b \beta_{0} \\
& =a 10.00+b 10.00
\end{aligned}
$$

At the end of the period, we set the portfolio value equal to the option pay-off. We can do this because the option payoff is completely dependent on the stock price.

$$
\begin{aligned}
a S_{1}+b \beta_{1} & =C_{1} \\
a 11.50+b 10.25 & =1.50 \\
a 9.50+b 10.25 & =0.00
\end{aligned}
$$

We can solve for the unknowns so $a=0.75$ and $b=-0.695122$. Now we know the value of the portfolio because we know what the stock and the bond are worth at the beginning of the time period. But since this portfolio has the same cash flows as the option, if there is no arbitrage the option price must be equal to the portfolio price at the start. The option price is therefore:

$$
\begin{aligned}
C_{0} & =a S_{0}+b \beta_{0} \\
C_{0} & =a 10.00+b 10.00 \\
& =\$ 0.5488
\end{aligned}
$$

2. The second approach uses risk-neutrality:

Consider a world where all assets have risk premia equal to 0 (a risk neutral world). That is, all expected returns of assets are discounted at the risk-free rate. For our particular example, this would mean that the probabilities with which the stock price evolves change. This is illustrated in the figure below. $E\left[\hat{S}_{1}\right]=\$ 10.25$


Figure 6. A risk neutral world.

We can do the same as we did in the arbitrage-free approach, and create a portfolio to price the option.

$$
\begin{aligned}
a 11.50+b 10.25 & =1.50 \\
a 9.50+b 10.25 & =0.00
\end{aligned}
$$

So $a=0.75$ and $b=-0.695122$, as before and $C_{0}=\$ 0.5488$ as before. This has not changed because none of the cash flows have changed, only the probabilities with which they occur. The price of the option is the same.

So is there any benefit in considering this risk-neutral world? Well, our first attempt to price the option consisted of taking an expectation, but we could not proceed because we did not know the appropriate discount factor. But in this example all assets are discounted at the risk-free rate. Our portfolio consists of two assets that are discounted at the risk-free rate, so any combination of these two assets is also discounted at the risk-free rate. Therefore, under risk-neutrality the option payoff is also discounted at the risk-free rate.

$$
C_{0}=\frac{E\left[\max \left(\hat{S}_{1}-K, 0\right)\right]}{1+r}=\frac{E\left[\max \left(S_{1}-K, 0\right)\right]}{1+\rho}
$$

The no arbitrage approach gives the same price regardless of riskneutrality but under risk neutrality it is much easier to calculate the price of the option since we can take the expected value and discount by the risk-free rate. Notice that the only assumption we make is that of no arbitrage. We do not assume that the world is risk neutral. Risk neutrality is a tool that we use to calculate the no-arbitrage price of an option.

$$
\begin{aligned}
C_{0} & =\frac{E\left[\max \left(\hat{S}_{1}-K, 0\right)\right]}{1+r} \\
& =\frac{0.375 \times 1.50+0.625 \times 0}{1+0.025} \\
& =\$ 0.5488
\end{aligned}
$$

There is another, slightly more philosophical point. The assumption of no arbitrage is used to calculate the fair price of a derivative. We do not actually assume that the world really is arbitrage-free as this would be far from the truth. Knowing the fair price of a derivative gives us a benchmark against which we can set market prices and determine mis-pricings that we can later exploit.

Before we move on to the real pricing we should examine a few more points using this really simple example.

1. Change of measure and the Girsanov theorem:

We notice that after assuming risk neutrality, the probabilities with which the share price evolves change. This is to allow for the change in expected value without altering the cash flows. In this extremely simple example it is easy to calculate these probabilities. However, when we assume that a stock price evolves according to GBM, we will need to find these probabilities. Since under GBM, the stock price at $t=1$ is a continuous random variable (where the probability of any single outcome is 0 ) we will need to consider something called change of measure.

The idea is really quite simple. We can measure the height of a man in metres or in feet. Going from one to the other is called changing measure. Just in the same way that we can assign a height to a man, we can assign a probability to an event. In fact we have a change
of measure in the example above. Under the real-world measure the probability of the stock price going to $\$ 11.50$ is 0.5 , whereas under the risk-neutral measure it is 0.375 .

The tool that we will use to find this risk neutral measure is the Girsanov theorem. Simply put, it states that change of measure is the same as change of drift. To understand this using our example we can see that there are actually two ways to make the stock RADR equal to the risk-free rate. The first is changing the probabilities of the outcomes (as we have done) but the second way is keeping the probabilities the same, and instead changing the price outcomes.


Figure 7. Illustration of two ways of changing the expected stock price at $t=1$.

## 2. Equivalent measure:

This example, and in particular the shortcoming of it illustrates another very important point, equivalent measure.

A few pages back we saw that risk-neutrality does not change the price of the option because the cash flows do not change, only the probability with which they occur. Our reason to use risk-neutrality was to make the expected pay-off from the stock equal to that of the bond (in order to obtain a discount rate for the expected option pay-off).

Now we are saying that we can make the expected values equal by
changing the price outcomes for the stock instead of the probabilities. But does this still give us the same option price?

$$
\begin{aligned}
a 11.25+b 10.25 & =1.25 \\
a 9.25+b 10.25 & =0.00, \quad a=0.625, b=-0.564024 \\
C_{0} & =0.6098
\end{aligned}
$$

The answer is no, and it is obviously because the cash flows have changed. So it seems that for this example we cannot say that change of measure is the same as change of drift. This is exactly where equivalent measure comes in. If all events possible under $\mathbb{P}$ are the same as the events possible under $\mathbb{Q}$ (the measures have the same support) then the two measures $\mathbb{P}$ and $\mathbb{Q}$ are quivalent.

If $\mathbb{P}(A)>0 \Leftrightarrow \mathbb{Q}(A)>0$ then $\mathbb{P}$ and $\mathbb{Q}$ are quivalent.

The two graphs in figure 7 above demonstrate clearly that the two measures are not equivalent. This is precisely why the option price changes. So when changing measure, to avoid changing the option price, we must always go to an equivalent measure. This way the outcomes of the stock price do not change, only the probabilities. This in turn does not change the cash flows, and therefore the option price remains the same.

This is where the Girsanov theorem and change in drift is so useful. If at $t=1$ the stock price is log-normally distributed as in GBM, then it can take all positive values. To go to the risk-neutral measure we can again shift the probability mass. In this case however, we can use the theorem and instead change the drift since doing so does not change the possible values of the stock price, thus preserving all cash flows. In other words, we can change the drift of a log-normal process instead of changing the probability mass because all log-normally distributed variables have equivalent measures (all have $(0, \infty)$ as support).

## 3. Statement of the Fundamental Theorem of Asset Pricing:

We began this sub-section with the idea that assuming absence of arbitrage allows us to calculate the price of a derivative. Furthermore we discovered that risk-neutrality can be used as a very effective tool to help us in calculating this price.

There is a theorem tying these two ideas together, and we should first observe how it is stated using our current example. Later on we will examine it in more detail. There is a type of process that we need to know before continuing, that is we need to know what is a martingale. A martingale is a process in which the expectation of the next value given the current value is equal to the current value. In other words the process has no drift. For example, Brownian motion is a martingale.

We call a process a martingale if $E[|X|]<\infty$ and

$$
E\left[X_{t+s} \mid X_{t}\right]=X_{t}
$$

Now let us state the Fundamental Theorem of Asset Pricing and observe and illustrate what the statement means. Later we will prove it.

Fundamental Theorem of Asset Pricing: Given a market with assets $Y_{t}$, and a numéraire $N_{t}$, the market is arbitrage free if and only if there exists an equivalent martingale measure for $Y_{t} / N_{t}$.

Remembering that any asset that is always positive can serve as numéraire, in this example we take the bond. The statement is well illustrated by graph 8 below. The first step is looking at the discounted process $Y_{t} / N_{t}$. If we can find a measure that makes this new process a martingale, then no-arbitrage is admitted. Working the other way, if there is no-arbitrage, then there is this equivalent martingale measure.

Observe the probabilities given to make the discounted stock process a martingale. They are the same as in the risk neutral stock process. In fact, taking the bond as numéraire gives the same probabilities as those under risk neutrality.


Figure 8. This is no longer simply the stock price process, but we have taken the bond as numéraire. In the second plot the measure has been changed to the equivalent martingale measure.

We will look at this theorem in more depth a little later, but now let us examine a tool that we will use to find equivalent martingale measures.

### 4.2 GIRSANOV THEOREM

As we discussed previously, the Girsanov theorem will help us greatly in finding and specifying new measures. We will look at this theorem for the special case of Brownian motion, although it can be extended to more general processes. The underlying point is that change of measure is the same as change of drift.

Theorem 4.2.1. (Change of Measure (Girsanov)) Let $W_{t}$ be a standard Brownian motion with a corresponding probability measure $\mathbb{P}$. Now consider a process $\theta_{t}$ such that $\int_{0}^{t} \theta_{s} d W_{s}$ is well defined and define a process $M_{t}$ where:

$$
M_{t}=e^{\int_{0}^{t} \theta_{s} d W_{s}-\frac{1}{2} \int_{0}^{t} \theta_{s}^{2} d s}
$$

A new measure is given by $d \mathbb{Q}=M d \mathbb{P}$ and the process $\tilde{W}_{t}=-\int_{0}^{t} \theta_{s} d s+W_{t}$ is $a$ standard Brownian motion under the new measure $\mathbb{Q}$.

Proof:

We will illustrate why our choice of $M_{t}$ helps a little later but now we will prove only that $\tilde{W}_{t}$ is indeed a standard Brownian motion under $\mathbb{Q}$. To do this we simply need to show that $\tilde{W}_{t}$ has independent, normally distributed increments with the right variances, and this is done by showing that the joint moment generating function of $n$ increments is the same as the moment generating function of n independent normal random variables.

Recall that the MGF for a standard normal random variable is $e^{(\tau \sigma)^{2} / 2}$, so we want to show:

$$
E_{\mathbb{Q}}\left[e^{\sum_{1}^{n} \tau_{i}\left(\tilde{W}_{t_{i+1}}-\tilde{W}_{t_{i}}\right)}\right]=E_{\mathbb{P}}\left[\prod_{1}^{n} e^{\tau_{i}^{2}\left(t_{i+1}-t_{i}\right) / 2}\right]
$$

Let us just consider one such increment:

$$
\begin{aligned}
E_{\mathbb{Q}}\left[e^{\tau_{i}\left(\int_{t_{i}}^{t_{i+1}} d \tilde{W}_{s}\right)}\right] & =E_{\mathbb{Q}}\left[e^{\tau_{i}\left(-\int_{t_{i}}^{t_{i+1}} \theta_{s} d s+\int_{t_{i}}^{t_{i+1}} d W_{s}\right.}\right] \\
& =E_{\mathbb{P}}\left[e^{\tau_{i}\left(-\int_{t_{i}}^{t_{i+1}} \theta_{s} d s+\int_{t_{i}}^{t_{i+1}} d W_{s}\right)}\left(M_{t_{i+1}}-M_{t_{i}}\right)\right] \\
& =E_{\mathbb{P}}\left[e^{\tau_{i}\left(-\int_{t_{i}}^{t_{i+1}} \theta_{s} d s+\int_{t_{i}}^{t_{i+1}} d W_{s}\right)} e^{\int_{t_{i}}^{t_{i+1}} \theta_{s} d W_{s}-\frac{1}{2} \int_{t_{i}}^{t_{i+1}} \theta_{s}^{2} d s}\right] \\
& =E_{\mathbb{P}}\left[e^{-\frac{1}{2} \int_{t_{i}}^{t_{i+1}}\left(2 \tau_{i} \theta_{s}+\theta_{s}^{2}\right) d s+\int_{t_{i}}^{t_{i+1}}\left(\tau_{i}+\theta_{s}\right) d W_{s}}\right] \\
& =e^{\tau_{i}^{2}\left(t_{i+1}-t_{i}\right) / 2} E_{\mathbb{P}}\left[e^{\left.-\frac{1}{2} \int_{t_{i}}^{t_{i+1}}\left(\tau_{i}+\theta_{s}\right)^{2} d s+\int_{t_{i}}^{t_{i+1}\left(\tau_{i}+\theta_{s}\right) d W_{s}}\right]}\right. \\
& =e^{\tau_{i}^{2}\left(t_{i+1}-t_{i}\right) / 2} \quad \text { as required }
\end{aligned}
$$

Going to line 5 we simply completed the square, and for the last step we used the fact that the expectation is equal to 1 (see below).

The proof of the theorem is not so difficult, what is more interesting is our choice of $M_{t}$. It is called a Radon-Nikodym process, and the idea behind it is simple. We use this value to change the measure of random variables and it must be such that $E_{\mathbb{Q}}\left[W_{t}\right]=E_{\mathbb{P}}\left[M_{t} W_{t}\right]$ and $E_{\mathbb{P}}\left[M_{t}\right]=1$. This ensures
that $\mathbb{Q}$ is still a probability measure. For some events $M_{t}>1$, so more mass is shifted on those events. Likewise for some outcomes, $M_{t}<1$ so these become less likely. The total "amount" of probability does not change, only where it is placed.

We will show that that $M_{t}$ is a martingale and $E\left[M_{t}\right]=1$ (for non-random $\theta_{t}$ ). This is best done by finding $d M_{t}$ using Itô's lemma:

$$
\begin{aligned}
d M_{t} & =\frac{\partial f\left(W_{t}, t\right)}{\partial W_{t}} d W_{t}+\frac{\partial f\left(W_{t}, t\right)}{\partial t} d t+\frac{\partial^{2} f\left(W_{t}, t\right)}{2 \partial W_{t}^{2}}\left(d W_{t}\right)^{2}+\ldots \\
& =\theta_{t} M_{t} d W_{t}-\frac{1}{2} \theta_{t}^{2} M_{t} d t+\frac{1}{2} \theta_{t}^{2} M_{t}\left(d W_{t}\right)^{2} \\
& =\theta_{t} M_{t} d W_{t} \\
E\left[d M_{t} \mid M_{t}\right] & =E\left[\theta_{t} M_{t} d W_{t} \mid M_{t}\right] \\
& =\theta_{t} M_{t} E\left[d W_{t} \mid W_{t}\right] \quad M_{t} \text { given implies } \theta_{t} \text { and } W_{t} \text { are given } \\
& =0
\end{aligned}
$$

This shows that $M_{t}$ is a martingale. The fact that $E\left[M_{t}\right]=1$ is easily deduced since $M_{0}=1$. Let us now look at an example and observe how the theorem works with some numbers.

Example (Change of measure) Let us now consider $\theta_{t}$ as in the theorem above, equal to 2 . Then by the theorem:

$$
\begin{aligned}
\tilde{W}_{t} & =-\int_{0}^{t} \theta_{s} d s+W_{t} \\
& =-2 t+W_{t}
\end{aligned}
$$

and $\tilde{W}_{t}$ is a standard Brownian motion under $\mathbb{Q}$. Let us have a look at $M_{t}$ :

$$
\begin{aligned}
M_{t} & =e^{\int_{0}^{t} \theta_{s} d W_{s}-\frac{1}{2} \int_{0}^{t} \theta_{s}^{2} d s} \\
& =e^{2 W_{t}-2 t}
\end{aligned}
$$

for $W_{t}>t, M_{t}>1$ so more weight is placed on upward trajectories of $W_{t}$. This means that the density is shifted upwards. We can also observe this
by computing the new density function by $\phi_{\mathbb{Q}}=M \phi_{\mathbb{P}}$ where $\phi$ represents the normal density function under the respective measure. Let us do this for $t=1:$


Figure 9. The densities under $\mathbb{P}$ (blue) and $\mathbb{Q}($ red $)$.

And the next figure perhaps best illustrates the theorem in action:


Figure 10. The two plots illustrate change of measure. The brightness of the colour represents the density associated with the particular path. The plot on the right has identical paths to the plot on the left, but different measure. The densities under $\mathbb{P}$ and $\mathbb{Q}$ have been over-laid to clarify the relationship between change of measure and change of drift.

Let us now move on to a very important section that will finally provide us
with the tools to price financial derivatives. The Girsanov theorem will be crucial in the development of this section.

### 4.3 FUNDAMENTAL THEOREM OF ASSET PRICING

The first theorem in this sub-section is the first fundamental theorem of asset pricing. We have almost all the terminology we need. To continue we only need to have a quick look over the self-financing portfolio strategy.

The example used throughout 4.1 contained a portfolio strategy $\left(p_{t}\right)$ in a stock and a bond $\left(Y_{t}\right)$. The portfolio value would then be denoted $V_{t}$.

$$
V_{t}=p_{t} Y_{t}^{\prime}=\left[a_{t}, b_{t}\right]\left[S_{t}, \beta_{t}\right]^{\prime}=a_{t} S_{t}+b_{t} \beta_{t}
$$

A self-financing portfolio strategy is a strategy such that there are no funds added or withdrawn and $V_{0}=0$. This means that $d V_{t}=p_{t} d Y_{t}$ should hold.

We can make such a strategy using the same example:

$$
\begin{aligned}
{\left[a_{t}, b_{t}, c_{t}\right]\left[S_{t}, \beta_{t}, C_{t}\right]^{\prime} } & =a_{t} S_{t}+b_{t} \beta_{t}+c_{t} C_{t} \\
{[-0.75,0.6951,1]\left[S_{t}, \beta_{t}, C_{t}\right]^{\prime} } & =-0.75 S_{t}+0.6951 \beta_{t}+C_{t}
\end{aligned}
$$

Such a portfolio has $V_{0}=0$ at $t=0$, and there are no funds added. This type of portfolio should give us insight about the relationship between arbitrage and portfolios. Considering the example in 4.1, if the price of the option is fair then $V_{0}=0$ and $V_{1}=0$. But what happens if the option is mis-priced? Let us say that it is undervalued and costs $\$ 0.50$ instead of the no-arbitrage price of $\$ 0.5488$. We can create a portfolio to exploit this mis-pricing and earn a risk-free profit, in other words we can exploit this arbitrage opportunity.

The easiest way to do this, is to buy the under-priced option, and sell the replicating portfolio. Basically we have the same asset, since the future cash flows are the same. So we buy the option for $\$ 0.50$ and then sell essentially
the same thing (the replicating portfolio) for $\$ 0.5488$, making an instantaneous profit of $\$ 0.0488$. We have no risk of losing money at $t=1$ since the cash flows of the assets we hold are the same and will cancel each other out. However, for the proof of the fundamental theorem of asset pricing we need to fulfill the condition that our portfolio has $V_{0}=0$ but for our current construction $V_{0}=0.0488$.

We can exploit this opportunity in a slightly different way, by receiving a guaranteed future cash flow. Consider a portfolio strategy $p_{t}=[-0.75,0.7,1]$ (never mind for now how we selected these particular numbers). Then at $t=0$ :

$$
V_{0}=[-0.75,0.7,1][10,10,0.5]^{\prime}=0
$$

And let us have a look at what happens at $t=1$. Remembering that there are two possible outcomes:

$$
\begin{aligned}
& V_{1}
\end{aligned}=[-0.75,0.7,1][11.5,10.25,1.5]^{\prime}=0.59 \text { or } \quad V_{1}=[-0.75,0.7,1][9.5,10.25,0]^{\prime}=0.58
$$

Either way, we are sure to make a profit of $\$ 0.50$ without having injected any of our own money at the start. This is also a way of exploiting this arbitrage opportunity. Notice that these two approaches in fact give the same payoff. When we discount the arbitrage payoff using the risk-free rate, we obtain the same value as when using the first approach $(0.05 / 1.025=0.0488)$.

Now that we have a feeling for these ideas we can see how no-arbitrage and equivalent martingale measure are tied together in the theorem itself:

Theorem 4.3.1. (Fundamental Theorem of Asset Pricing (I)) Given a market with assets $Y_{t}$, and a numéraire $N_{t}$, the market is arbitrage free if and only if there exists an equivalent martingale measure for $Y_{t} / N_{t}$

Proof:
We will prove only the sufficiency of this theorem, it will be enough for us to continue.

Consider a self financing portfolio $p_{t}$ in the assets $Y_{t}$. The portfolio value is then $V_{t}=p_{t} Y_{t}$ and $V_{0}=0$. Let us assume that there is an arbitrage resulting from holding this portfolio. That is $\mathbb{P}\left(V_{t} \leq 0\right)=0$.

$$
\begin{align*}
\mathbb{P}\left(V_{t} \leq 0\right)=0, \mathbb{P}\left(V_{t}>0\right)>0 & \text { this is an arbitrage } \\
\mathbb{P}\left(V_{t} / N_{t} \leq 0\right)=0, \mathbb{P}\left(V_{t} / N_{t}>0\right)>0 & \text { since } N_{t} \text { is always positive }
\end{align*}
$$

But we have an equivalent martingale measure, so this means that:

$$
\begin{aligned}
E_{\mathbb{Q}}\left(V_{t} / N_{t}\right) & =V_{0} / N_{0}+E_{\mathbb{Q}} \int_{0}^{t} d\left(V_{s} / N_{s}\right) \\
E_{\mathbb{Q}}\left(V_{t} / N_{t}\right) & =V_{0} / N_{0}+E_{\mathbb{Q}} \int_{0}^{t} p_{s} d\left(Y_{s} / N_{s}\right) \\
& =0 \quad \text { since } Y_{s} / N_{s} \text { is a martingale under } \mathbb{Q}
\end{aligned}
$$

But because the expectation is 0 , it must be that $\mathbb{Q}\left(V_{t} / N_{t} \leq 0\right)>0$ and therefore $\mathbb{Q}\left(V_{t} \leq 0\right)>0$. Since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, it must be that $\mathbb{P}\left(V_{t} \leq 0\right)>0$. This is a contradiction to ( $\dagger$ ). It is in fact possible for $V_{t} \leq 0$ and therefore this is not arbitrage.

The next theorem follows on very naturally from above and also uses the Girsanov theorem. It tells us how the drift and volatility coefficients of a discounted asset price should be related to allow no-arbitrage. It will be crucial in the development of the Black-Scholes equation.

Theorem 4.3.2. Let there be a market with assets $Y_{t}$ and a numéraire $N_{t}$. Let us say the process $Y_{t} / N_{t}$ has the form:

$$
d Y_{t} / N_{t}=\mu_{Y / N} d t+\sigma_{Y / N} d W_{t}
$$

No-arbitrage is admitted if and only if there exists a function $\lambda_{N}$ (the $N$ subscript denotes that the function depends on choice of numéraire) such that:

$$
\mu_{Y / N}=\sigma_{Y / N} \lambda_{N}
$$

## Proof:

We saw from the theorem above that to have absence of arbitrage we need an equivalent martingale measure for the discounted process $Y_{t} / N_{t}$. But by the Girsanov theorem, change of measure is the same as change of drift. So if there exists an equivalent martingale measure then there exists a $\lambda_{N}$ such that:

$$
d \tilde{W}_{t}=\lambda_{N} d t+d W_{t}
$$

and

$$
d\left(Y_{t} / N_{t}\right)=\sigma_{Y / N} d \tilde{W}_{t} \quad \text { (the discounted process is a martingale) }
$$

Substituting $d W_{t}=d \tilde{W}_{t}-\lambda_{N} d t$ into $d Y_{t} / N_{t}=\mu_{Y / N} d t+\sigma_{Y / N} d W_{t}$ we see that:

$$
\begin{aligned}
d Y_{t} / N_{t} & =\mu_{Y / N} d t+\sigma_{Y / N}\left(d \tilde{W}_{t}-\lambda_{N} d t\right) \\
& =\left(\mu_{Y / N}-\sigma_{Y / N} \lambda_{N}\right) d t+\sigma_{Y / N} d \tilde{W}_{t}
\end{aligned}
$$

The drift term vanishes (the process is a martingale) only if $\mu_{Y / N}=\sigma_{Y / N} \lambda_{N}$, as stated in the theorem. Note that $\lambda_{N}$ has the form $-\int_{0}^{t} \theta_{s} d s$ in the Girsanov theorem sub-section (4.2).

### 4.4 BLACK-SCHOLES EQUATION

In this sub-section we will use what we have learned so far to arrive at the famous Black-Scholes ${ }^{1}$ equation. This will be a very general form, and may seem slightly abstract at first. We will however review it using a couple of nice examples in the examples sub-section.

Theorem 4.4.1. (Black-Scholes equation) Let there be a market with assets $Y_{t}$ and a numéraire $N_{t}$. Let us say the discounted process $Y_{t} / N_{t}$ has the form:

$$
d Y_{t} / N_{t}=\mu_{Y / N} d t+\sigma_{Y / N} d W_{t}
$$

[^0]No-arbitrage is admitted if and only if there exist functions $\Lambda$ and $R$ such that the equation (Black-Scholes) below is satisfied:

$$
\mu_{Y}-R Y_{t}=\sigma_{Y} \Lambda
$$

Proof:
We will need to actually compute $d\left(Y_{t} / N_{t}\right)$ and use theorem 4.3.2 in order to arrive at the equation. We will use Itô's lemma to do this, remembering of course that $d Y_{t}=\mu_{Y} d t+\sigma_{Y} d W_{t}$ and $d N_{t}=\mu_{N} d t+\sigma_{N} d W_{t}$ (both are Itô processes).

$$
\begin{aligned}
d \frac{Y_{t}}{N_{t}} & =\frac{1}{N_{t}} d Y_{t}-\frac{Y_{t}}{N_{t}^{2}} d N_{t}+\frac{1}{2}\left(0-\frac{1}{N_{t}^{2}}\left(d Y_{t}\right)\left(d N_{t}\right)-\frac{1}{N_{t}^{2}}\left(d N_{t}\right)\left(d Y_{t}\right)+\frac{2 Y}{N_{t}^{3}}\left(d N_{t}\right)^{2}\right) \\
& =\frac{1}{N_{t}} d Y_{t}-\frac{Y_{t}}{N_{t}^{2}} d N_{t}-\frac{1}{N_{t}^{2}}\left(d N_{t}\right)\left(d Y_{t}\right)+\frac{Y}{N_{t}^{3}}\left(d N_{t}\right)^{2} \\
& =\frac{1}{N_{t}}\left(\mu_{Y} d t+\sigma_{Y} d W_{t}\right)-\frac{Y_{t}}{N_{t}^{2}}\left(\mu_{N} d t+\sigma_{N} d W_{t}\right)-\frac{1}{N_{t}^{2}} \sigma_{N} \sigma_{Y} d t+\frac{Y}{N_{t}^{3}} \sigma_{N}^{2} d t
\end{aligned}
$$

From above we then obtain:

$$
\sigma_{Y / N}=\frac{1}{N_{t}^{2}}\left(\sigma_{Y} N_{t}-Y_{t} \sigma_{N}\right)
$$

and

$$
\mu_{Y / N}=\frac{1}{N_{t}^{2}}\left(\mu_{Y} N_{t}-Y_{t} \mu_{N}\right)-\sigma_{Y / N} \frac{\sigma_{N}}{N_{t}}
$$

In the previous theorem (4.3.2) we saw how $\sigma_{Y / N}$ and $\mu_{Y / N}$ are related if there is no-arbitrage. The theorem stated that $\mu_{Y / N}=\sigma_{Y / N} \lambda$ in such a case. Using this, we obtain:

$$
\begin{aligned}
\left(\mu_{Y} N_{t}-Y_{t} \mu_{N}\right) & =\left(\sigma_{Y} N_{t}-Y_{t} \sigma_{N}\right)\left(\lambda+\frac{\sigma_{N}}{N_{t}}\right) \\
& =\left(\sigma_{Y} N_{t}-Y_{t} \sigma_{N}\right) \Lambda
\end{aligned}
$$

We can now re-arrange this to:

$$
\mu_{Y}-\frac{\mu_{N}-\sigma_{N} \Lambda}{N_{t}} Y_{t}=\sigma_{Y} \Lambda
$$

This is exactly the expression $\mu_{Y}-R Y_{t}=\sigma_{Y} \Lambda$ as stated in the theorem with:

$$
R=\frac{\mu_{N}-\sigma_{N} \Lambda}{N_{t}} \quad \text { and } \quad \Lambda=\lambda+\frac{\sigma_{N}}{N_{t}}
$$

Clearly, if there is no-arbitrage, then the Black-Scholes equation is satisfied. Conversely, if $\mu_{Y}-R Y_{t}=\sigma_{Y} \Lambda$ is given first, then $\mu_{N}-R N_{t}=\sigma_{N} \Lambda$ must also hold. Rearranging this we obtain the same expression for $R$ as above. Consequently there is no arbitrage.

It is very interesting to examine this equation. $R$ in the Black-Scholes equation can be interpreted as the risk-free rate on a risk-free asset $Y$ (an asset with $\sigma_{Y}=0$ ). This $R$ does not have to be constant and can depend on time. The expression $\sigma_{Y} \Lambda$ is then viewed as the excess return on $Y_{t}$ if it is not risk-free. If we rearrange the equation to obtain $\Lambda=\left(\mu_{Y}-R Y_{t}\right) / \sigma_{Y}$, we see that it is the excess return per unit of volatility. A common name for this is the market price of risk (also commonly called the Sharpe ratio).

### 4.5 BLACK-SCHOLES FORMULA

The example in 4.1 showed us that the price of a European call option can be evaluated in two ways. The first way was to solve simultaneous equations and in fact this is parallel to the Black-Scholes equation when we introduce a continuous-time framework. We will see in the examples sub-section that the same arguments are used and we will arrive at a partial differential equation, the solution to which gives the option price. However, we also saw that we can use risk-neutral pricing as a tool, and it is widely regarded that this is an easier method. By simply taking the expectation of the payoff under risk-neutrality then discounting this expected value by the risk-free rate we arrive at the same price.

So let us consider the model with the stock dynamics explained by geometric Brownian motion, a European call option on the stock (with strike price K) and a risk-free bond (note that we can use the technique in the sub-section to price any option with a payoff at a non-random time. It does not have to
be a European call).

$$
\begin{aligned}
d S_{t} & =\mu S_{t} d t+\sigma S_{t} d W_{t} \\
d \beta_{t} & =r \beta_{t} d t
\end{aligned}
$$

In a risk-neutral world all assets are expected to grow at the risk free rate, so the drift term for the stock price should be $r \tilde{S}_{t} d t$ instead of $\mu S_{t} d t$. Let us just confirm this fact by using the theory that we have learned so far. We know that arbitrage is ruled out if we can find an equivalent martingale measure. The first step is to choose a numéraire and we know that taking the bond results in finding the risk-neutral measure. As we have done before, we use Itô's lemma to find how the discounted price evolves.

$$
d \frac{S_{t}}{\beta_{t}}=(\mu-r) \frac{S_{t}}{\beta_{t}} d t+\sigma \frac{S_{t}}{\beta_{t}} d W_{t}
$$

Now we need to find a measure that will make this process a martingale, but by the Girsanov theorem, change of measure is the same as change of drift. So, all we need to do is find a $\lambda$ in $d \tilde{W}_{t}=\lambda d t+d W_{t}$ such that the above process is a martingale (drift is 0 ).

$$
\begin{aligned}
d \frac{S_{t}}{\beta_{t}} & =(\mu-r) \frac{S_{t}}{\beta_{t}} d t+\sigma \frac{S_{t}}{\beta_{t}}\left(d \tilde{W}_{t}-\lambda d t\right) \\
& =((\mu-r)-\sigma \lambda) \frac{S_{t}}{\beta_{t}} d t+\sigma \frac{S_{t}}{\beta_{t}} d \tilde{W}_{t}
\end{aligned}
$$

So $\lambda=(\mu-r) / \sigma$. Using this change of measure in our model for the stock price gives the same outcome:

$$
\begin{aligned}
d \tilde{S}_{t} & =\mu \tilde{S}_{t} d t+\sigma \tilde{S}_{t}\left(d \tilde{W}_{t}-\lambda d t\right) \\
& =\mu \tilde{S}_{t} d t+\sigma \tilde{S}_{t}\left(d \tilde{W}_{t}-\frac{\mu-r}{\sigma} d t\right) \\
& =r \tilde{S}_{t} d t+\sigma \tilde{S}_{t} d \tilde{W}_{t}
\end{aligned}
$$

The payoff of the option is $\max \left(S_{t}-K, 0\right)$. The value of the option is then:

$$
C_{0}=e^{-\rho t} E\left[\max \left(S_{t}-K, 0\right)\right]
$$

However the difficulty in estimating $\rho$ means that we use risk-neutrality as
a tool. Risk-neutrality changes the drift of the underlying process, but it will give us the correct price only if we remember that the initial values of the actual and the risk-neutral processes must be the same. Under risk neutrality the price is given by:

$$
C_{0}=e^{-r t} E\left[\max \left(\tilde{S}_{t}-K, 0\right) \mid \tilde{S}_{0}=S_{0}\right]
$$

Where $\tilde{S}_{t}$ is as above. All that remains for us is to find $E\left[\max \left(\tilde{S}_{t}-K, 0\right) \mid \tilde{S}_{0}=\right.$ $\left.S_{0}\right]$. We know the distribution of $\tilde{S}_{t}$ so we can proceed.

$$
\begin{aligned}
X=\log \tilde{S}_{t} & \sim N\left(\nu, \varphi^{2}\right) \\
\nu & =\log \tilde{S}_{0}+\left(r-\sigma^{2} / 2\right) t \\
\varphi^{2} & =\sigma^{2} t
\end{aligned}
$$

To find $E\left[\max \left(e^{X}-K, 0\right)\right]$ we can consider the moment generating function of the normal random variable evaluated at $s=1 . \phi$ and $\Phi$ represent the normal density and distribution function respectively.

$$
\begin{aligned}
&\left.E\left[e^{X s}\right]\right|_{s=1}=E\left[e^{X}\right] \\
& E\left[e^{X s}\right]=\int_{-\infty}^{\infty} e^{x s} \phi_{\left(\nu, \varphi^{2}\right)} d x \\
& E\left[\max \left(e^{X s}-K, 0\right)\right]= \int_{\log K}^{\infty}\left(e^{x s}-K\right) \phi_{\left(\nu, \varphi^{2}\right)} d x \\
&= \frac{1}{\sqrt{2 \pi \varphi^{2}}} \int_{\log K}^{\infty} e^{x s} e^{-\frac{(x-\nu)^{2}}{2 \varphi^{2}}} d x-K \int_{\log K}^{\infty} \phi_{\left(\nu, \varphi^{2}\right)} d x \\
&= a-b
\end{aligned}
$$

$$
\begin{array}{ll}
a=\frac{1}{\sqrt{2 \pi \varphi^{2}}} e^{\nu s} \int_{\log K-\nu}^{\infty} e^{z s} e^{-\frac{z^{2}}{2 \varphi^{2}}} d z & b=K\left(1-\Phi_{\left(\nu, \varphi^{2}\right)}[\log K]\right) \\
a=\frac{1}{\sqrt{2 \pi \varphi^{2}}} e^{\nu s} \int_{\log K-\nu}^{\infty} e^{\left(-\frac{z^{2}}{2 \varphi^{2}}+z s-\frac{s^{2} \varphi^{2}}{2}+\frac{s^{2} \varphi^{2}}{2}\right)} d z & b=K \Phi_{(0,1)}[(\nu-\log K) / \varphi] \\
a=\frac{1}{\sqrt{2 \pi \varphi^{2}}} e^{\nu s} \int_{\log K-\nu}^{\infty} e^{\left(\frac{\left(\left(z-s \varphi^{2}\right)^{2}\right.}{2 \varphi^{2}}+\frac{s^{2} \varphi^{2}}{2}\right)} d z & \\
a=\frac{1}{\sqrt{2 \pi \varphi^{2}}} e^{\left(\nu s+\frac{s^{2} \varphi^{2}}{2}\right)} \int_{\log K-\nu-s \varphi^{2}}^{\infty} e^{\frac{-y^{2}}{2 \varphi^{2}}} d y & \sqrt{2 \pi \varphi^{2}} \phi_{\left(0, \varphi^{2}\right)}^{\infty} d y \\
a=\frac{1}{\sqrt{2 \pi \varphi^{2}}} e^{\left(\nu s+\frac{s^{2} \varphi^{2}}{2}\right)} \int_{\log K-\nu-s \varphi^{2}}^{\infty} \\
a=e^{\left(\nu s+\frac{s^{2} \varphi^{2}}{2}\right)}\left(1-\Phi_{\left(0, \varphi^{2}\right)}\left[\log K-\nu-s \varphi^{2}\right]\right) \\
a=e^{\left(\nu s+\frac{s^{2} \varphi^{2}}{2}\right)} \Phi_{(0,1)}[(\nu-\log K) / \varphi+s \varphi]
\end{array}
$$

$$
E\left[\max \left(e^{X}-K, 0\right)\right]=e^{\left(\nu+\frac{\varphi^{2}}{2}\right)} \Phi_{(0,1)}\left[\frac{\nu-\log K}{\varphi}+\varphi\right]-K \Phi_{(0,1)}[(\nu-\log K) / \varphi]
$$

Remembering that $X=\log \tilde{S}_{t}$ we can finally write down the formula:

$$
\begin{aligned}
C_{0}= & e^{-r t} \\
m_{1}= & \left.\frac{\log \left(e^{\left(\log \tilde{S}_{0}+\left(r-\sigma^{2} / 2\right) t+\frac{\sigma^{2} t}{2}\right)} \Phi_{(0,1)}\left[m_{1}\right]-K \Phi_{(0,1)}\left[m_{2}\right]\right.}{\sigma \sqrt{t}}\right) \\
m_{2}= & \frac{\log \left(\tilde{S}_{0} / K\right)+\left(r-\sigma^{2} / 2\right) t}{\sigma \sqrt{t}} \\
& C_{0}=\tilde{S}_{0} \Phi_{(0,1)}\left[m_{1}\right]-e^{-r t} K \Phi_{(0,1)}\left[m_{2}\right]
\end{aligned}
$$

There remains only one thing for us to remember. As it stands, the formula contains $\tilde{S}$, the risk neutral stock price. This differs only in drift from $S$ and at $t=0$ these two have to be equal for us to obtain the correct price.

Changing this gives us the famous Black-Scholes option pricing formula:


Figure 11. Black-Scholes formula option prices for parameters $\sigma=0.2, r=0.05$ and $S_{0}=10$. The blue lines correspond to different maturity times. The solid line is for $\tau=1 / 24$, dashed line for $\tau=2 / 24$, and dotted line for $\tau=3 / 24$.

### 4.6 EXAMPLES

In this sub-section we look at two assets, a European option, and a bond. We compare the differential equations given by general Black-Scholes and the differential equations given by simple derivation using portfolios. We will find that the equations we obtain are identical.

Example (Black-Scholes equation for a European option) We can use the Black-Scholes equation as it was derived in theorem 4.4.1, but as stated before this may seem slightly abstract. We can actually arrive at the same equation by using a very simple approach. The benefit in using 4.4.1 is that it is very general, while the derivation is specific to this example.

1. Using the general Black-Scholes equation directly:

$$
\mu_{Y}-R Y_{t}=\sigma_{Y} \Lambda
$$

where

$$
R=\frac{\mu_{N}-\sigma_{N} \Lambda}{N_{t}} \quad \text { and } \quad \Lambda=\lambda+\frac{\sigma_{N}}{N_{t}}
$$

The asset we want to price is a European call option, so $Y_{t}$ is the price of the option. The payoff of this option is $\max \left(S_{t}-K, 0\right)$ where $K$ is the strike price for some given expiry date. $Y_{t}$ must then depend on the current stock price $S_{t}$ and time $t$. Therefore $Y_{t}=f\left(S_{t}, t\right)$. If the stock price follows geometric Brownian motion:

$$
\begin{aligned}
d S_{t}= & \mu S_{t} d t+\sigma S_{t} d W_{t} \\
d \beta_{t}= & r \beta_{t} d t \\
d Y_{t}= & \frac{\partial f\left(S_{t}, t\right)}{\partial S_{t}} d S_{t}+\frac{\partial f\left(S_{t}, t\right)}{\partial t} d t+\frac{1}{2} \frac{\partial^{2} f\left(S_{t}, t\right)}{\partial S_{t}^{2}}\left(d S_{t}\right)^{2} \\
= & \left(\mu S_{t} \frac{\partial f\left(S_{t}, t\right)}{\partial S_{t}}+\frac{\partial f\left(S_{t}, t\right)}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} f\left(S_{t}, t\right)}{\partial S_{t}^{2}}\right) d t \\
& +\sigma S_{t} \frac{\partial f\left(S_{t}, t\right)}{\partial S_{t}} d W_{t}
\end{aligned}
$$

Taking the bond as numéraire, $R=r$ and $\Lambda=\lambda$. $\lambda$ is the price of risk associated with the driving Brownian motion. If we apply the BlackScholes formula to the underlying stock price, we obtain $\lambda=(r-\mu) / \sigma$. The same Brownian motion drives the option price, so the $\lambda$ associated with it must be the same. Note that this is identical to $\lambda$, as calculated in the previous sub-section.

$$
\mu S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}-r Y_{t}=\sigma S_{t} \frac{\partial Y_{t}}{\partial S_{t}} \lambda
$$

$$
\begin{gathered}
(\mu-\sigma \lambda) S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}-r Y_{t}=0 \\
r S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}-r Y_{t}=0
\end{gathered}
$$

The option price $Y_{t}$ must satisfy the above partial differential equation.
2. By creating a risk-less portfolio:

Let us consider now this slightly different approach, creating a portfolio comprising of one option, the amount $a$ in stocks and an investment in the money market (risk-less bond).

$$
V_{t}=a S_{t}+Y_{t}+\beta_{t} \quad \text { and } \quad d V_{t}=a d S_{t}+d Y_{t}+r \beta_{t} d t
$$

Writing out the full expression for $d V_{t}$ :
$d V_{t}=\left(a \mu S_{t}+\left(\mu S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}\right)+r \beta\right) d t+\left(a \sigma S_{t}+\sigma S_{t} \frac{\partial Y_{t}}{\partial S_{t}}\right) d W_{t}$
To make the portfolio riskless, we must eliminate $\sigma_{V}=a \sigma S_{t}+\sigma S_{t} \frac{\partial Y_{t}}{\partial S_{t}}$. This is done when $a=-\partial Y_{t} / \partial d S_{t}$ (remembering that $a$ is the amount of shares we buy).

$$
d V_{t}=\left(-\frac{\partial Y_{t}}{\partial S_{t}} \mu S_{t}+\left(\mu S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}\right)+r \beta\right) d t
$$

If the portfolio is risk-less, it must grow at the same rate as a risk-less money market account. Therefore $d V_{t}=r V_{t} d t$

$$
\left(\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}+r \beta\right) d t=r\left(\frac{\partial Y_{t}}{\partial S_{t}} S_{t}+Y_{t}+\beta_{t}\right) d t
$$

Simplifying this expression we obtain the partial differential equation that the option price must satisfy.

$$
r S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}-r Y_{t}=0
$$

This is identical to the result given by the general Black-Scholes equation.

Example (Black-Scholes equation for a bond in the Vasicek model) Here we consider the valuation of a bond. In this case however, the interest rate is not constant, or even deterministic. The rate evolves according to an Ornstein-Uhlenbeck process.

$$
\begin{aligned}
d r_{t} & =b\left(\mu-r_{t}\right) d t+\sigma d W_{t} \\
& =\mu_{r} d t+\sigma d W_{t}
\end{aligned}
$$

The bond pays us $\$ 1$ some pre-determined time in the future, so the bond price is a function of the interest rate and time.

$$
\beta_{t}=f\left(r_{t}, t\right)
$$

Our goal is to find the bond's current price (or at least the differential equation that is satisfied by this price).

1. Using the general Black-Scholes equation directly (note that $r_{t}$ and $R$ are quite different here. $r_{t}$ is the random short rate, whereas $R$ is some risk-free rate):

$$
\mu_{\beta}-R \beta_{t}=\sigma_{\beta} \Lambda
$$

So we must find how the bond price evolves, and as usual we use Itô's lemma:

$$
\begin{aligned}
d \beta_{t} & =\left(\mu_{r} \frac{\partial \beta_{t}}{\partial r_{t}}+\frac{\partial \beta_{t}}{\partial t}+\frac{1}{2} \sigma^{2} \frac{\partial^{2} \beta_{t}}{\partial r_{t}^{2}}\right) d t+\sigma \frac{\partial \beta_{t}}{\partial r_{t}} d W_{t} \\
& =\mu_{\beta} d t+\sigma_{\beta} d W_{t}
\end{aligned}
$$

Using this, we can write the equation to be satisfied by the bond price:

$$
\begin{gathered}
\mu_{\beta}-R \beta_{t}=\sigma_{\beta} \Lambda \\
(\mu-\sigma \Lambda) \frac{\partial \beta_{t}}{\partial r_{t}}+\frac{\partial \beta_{t}}{\partial t}+\frac{1}{2} \sigma^{2} \frac{\partial^{2} \beta_{t}}{\partial r_{t}^{2}}-R \beta_{t}=0
\end{gathered}
$$

2. We can also price this bond using a similar approach to the previous example. When pricing the option however, we had a perfectly correlated asset (the stock), which we could use to completely remove the risk. But what do we hedge the bond against? The trick here is
to consider arbitrage occurring over different time frames. The idea is that the yield from buying two single period bonds consecutively should be the same as form holding a two period bond over the same time period. Otherwise there would be an arbitrage opportunity. The trick is to consider two bonds of different maturities.

$$
\beta_{t}=f\left(r_{t}, t\right) \quad \text { and } \quad \alpha_{t}=g\left(r_{t}, t\right)
$$

$$
\begin{aligned}
d \beta_{t} & =\left(\mu_{r} \frac{\partial \beta_{t}}{\partial r_{t}}+\frac{\partial \beta_{t}}{\partial t}+\frac{1}{2} \sigma^{2} \frac{\partial^{2} \beta_{t}}{\partial r_{t}^{2}}\right) d t+\sigma \frac{\partial \beta_{t}}{\partial r_{t}} d W_{t} \\
d \alpha_{t} & =\left(\mu_{r} \frac{\partial \alpha_{t}}{\partial r_{t}}+\frac{\partial \alpha_{t}}{\partial t}+\frac{1}{2} \sigma^{2} \frac{\partial^{2} \alpha_{t}}{\partial r_{t}^{2}}\right) d t+\sigma \frac{\partial \alpha_{t}}{\partial r_{t}} d W_{t}
\end{aligned}
$$

$$
\begin{aligned}
& d \beta_{t}=\mu_{\beta} d t+\sigma_{\beta} d W_{t} \\
& d \alpha_{t}=\mu_{\alpha} d t+\sigma_{\alpha} d W_{t}
\end{aligned}
$$

Let us consider a portfolio $V_{t}=\beta_{t}+c \alpha_{t}$ where $d V_{t}=d \beta_{t}+c d \alpha_{t}$. Then to remove the risk in this portfolio much in the same way as we did for the previous example, we must make $\left(\sigma_{\beta}+c \sigma_{\alpha}\right) d W_{t}$ equal to 0 . That is, $c=-\frac{\sigma_{\beta}}{\sigma_{\alpha}}$. Then since the portfolio has no more risk, it must grow at some riskless rate.

$$
\begin{aligned}
d V_{t} & =R V_{t} d t \\
\mu_{\beta} d t-\frac{\sigma_{\beta}}{\sigma_{\alpha}} \mu_{\alpha} d t & =R V_{t} d t \\
\sigma_{\alpha} \mu_{\beta}-\sigma_{\beta} \mu_{\alpha} & =R\left(\sigma_{\alpha} \beta_{t}-\sigma_{\beta} \alpha_{t}\right)
\end{aligned}
$$

Rearranging, we obtain:

$$
\frac{\mu_{\beta}-R \beta_{t}}{\sigma_{\beta}}=\frac{\mu_{\alpha}-R \alpha_{t}}{\sigma_{\alpha}}=\Lambda
$$

But this is "excess return per volatility" and our other name for this is the price of risk. And since this is the same for both bonds, it must
not depend on the maturity time. Let us say we want to price the first bond, we can write:

$$
\frac{\mu_{\beta}-R \beta_{t}}{\sigma_{\beta}}=\Lambda
$$

Rearranging, and making substitutions for $\mu_{\beta}$ and $\sigma_{\beta}$ we obtain:

$$
\begin{gathered}
\mu_{\beta}-R \beta_{t}=\sigma_{\beta} \Lambda \\
(\mu-\sigma \Lambda) \frac{\partial \beta_{t}}{\partial r_{t}}+\frac{\partial \beta_{t}}{\partial t}+\frac{1}{2} \sigma^{2} \frac{\partial^{2} \beta_{t}}{\partial r_{t}^{2}}-R \beta_{t}=0
\end{gathered}
$$

Again, this is the same as given by the general Black-Scholes equation.

## 5 SIMPLE DIFFUSION MODELS: CONCLUSION

This brings us to the end of the first part. Not wishing to derive solutions, the point of interest here is the reasoning behind our pricing constructions. There are many methods to solve differential equations analytically, numerically and by simulation however, the most important part is the construction. One can always find a solution thereafter even if not an analytical one. While we did obtain the Black-Scholes formula which gives us the price of a European option, more importantly we have the Black-Scholes equation. Having this means that we can always price an asset that pays a payoff at some fixed time. What we must move on to now is our investigation of the underlying processes that govern the state variables, for we will discover that incorrect choice of such models can make pricing misleading.

## 6 SO THAT'S IT?

We have arrived at the second part of the paper. We have discovered that using Brownian motion as the engine of stochasticity provides us with a self contained theoretical framework that allows modeling underlying processes and pricing of financial instruments. The elegance and ease of working in such a framework undoubtedly makes Brownian motion the star of mathematical finance.

Graphs like the figure below are usually used to motivate the use of geometric Brownian motion as the appropriate process to model a stock price. The two processes do indeed look indistinguishable. However, this graph conceals several crucial components that we must investigate.


Figure 12. The graph shows the Dow Jones ${ }^{2}$ index and a simulated path of geometric Brownian motion with the same daily return and volatility. But which is which?

[^1]
## 7 SHORTCOMINGS OF GEOMETRIC BROWNIAN MOTION

### 7.1 NON-CONSTANT VOLATILITY

The first thing we can do to investigate is look at the returns of this process ( $d X_{t} / X_{t}$ ) which should be normally distributed if the process really is geometric Brownian motion. The returns are illustrated in the figure below. At one glance it is immediately evident that the geometric Brownian motion fit to the index and the actual index have very different structures. The index returns (blue) exhibit volatility clustering. The returns for 1992-1997 and 2004-2007 look to have relatively lower variance than for 1997-2004. The simulated returns (red) on the other hand show constant variance throughout. It seems then, that the characteristic of constant volatility is not exhibited by the Dow Jones index. This is the first clue suggesting that geometric Brownian motion is not a suitable model.


Figure 13. Daily returns of the Dow Jones and a simulated geometric Brownian motion with the same mean and volatility.

We have looked at the Dow Jones index, a composite of several stock prices. If we look at individual stocks, we will arrive at the same conclusion. Figure 14 displays actual and simulated returns of the Microsoft Corporation stock. Once again it is evident that the actual returns exhibit non-constant volatility. Although we will not be overly pedantic by checking the returns for all existing stocks, it has been shown the vast majority of stocks display this behaviour.


Figure 14. Daily returns of the Microsoft Corporation stock and a simulated geometric Brownian motion with the same mean and volatility.

There have been several suggestions to deal with the problem of non-constant volatility, which we will have a look at later. Next we must investigate a far more serious drawback of geometric Brownian motion.

### 7.2 EVIDENCE OF JUMPS

A very important property of geometric Brownian motion is continuity and scale invariance. Continuity means that all paths are continuous and exhibit no jumps at all, on any scale. Scale invariance means that no matter
how much we "zoom in", the process shows the same behaviour. Looking at a figure such as figure 12 (which is usually used to motivate the choice of geometric Brownian motion) we may agree that both actual and simulated paths are continuous.

Recall that Robert Brown observed pollen particles and reasoned that molecules would collide with, and cause the pollen to move at random. We reasoned that a market place should act the same. A large number of traders would buy and sell a stock at random, causing the price to move around unpredictably. This is true, but it is not the picture in its entirety. For the case of the pollen on the petri dish, with the assumption that there is no outside influence the model is sufficient. However, this model is far more simple than a market. We rule out outside influence. But what if we have someone bump into the table? That is, the table that our microscope and our petri dish is on. This would surely cause all the particles on the dish to make an uncharacteristically large movement in some direction.

This is of course unlikely under experimental conditions and indeed; one would hope that Mr. Brown was careful while undertaking his famous observations. But as we said, the marketplace is a predominantly more complex object than a petri dish. There are governments, natural disasters, and large competing firms just to name a few of the highly influential items affecting traders' behaviour. Indeed if a hurricane sinks an Exxon Mobil oil tanker, traders will not naively keep trading at random, causing Mobil's share price to move around unpredictably. These traders have brains, and undoubtedly they will recognise the importance and consequences of such an event. They would surely sell their Exxon Mobil stocks in anticipation of the problems arising for the company. In such an event the price of the stock would surely fall, and not by a negligible distance but possibly by an amount that the Geometric Brownian motion model would classify as nearly impossible. An event like this would be our jump. For a recent example we may take the sub-prime mortgage crisis in the USA. Numerous financial institutions have seen tremendous drops in their share price in a matter of minutes.

For one particular case we can look at the stock price of National City Corporation, one of the largest commercial banks in the United States.


Figure 15. NCC stock price and a simulated geometric Brownian motion, 19982008.


Figure 16. NCC stock price and a simulated geometric Brownian motion, 20072008.

Figure 15 displays the NCC stock price since 1998 and a path of simulated geometric Brownian motion with the same daily return and volatility. It is difficult to distinguish between the two. Figure 16 displays the same paths, but for a much smaller time scale, beginning at 2007. Although it is not immediately obvious which path is simulated and which is the actual process, a more experienced individual will notice the very large deviation in the blue path, towards the end. This is a revealing sign that the blue path is the actual process, since we know that such high deviations are extremely unlikely under geometric Brownian motion. Let us zoom in even further and look at this deviation:

NCC Stock Price Mar 13 - Mar 19, 2008


Figure 17. NCC stock price and a simulated geometric Brownian motion, March 13 - March 19, 2008.

Figure 17 shows the NCC stock price for March 13 to March 19, 2008 with the stock price recorded at 5 minute intervals. This plot should remove all doubt as to which process is the real stock price and which is the simulated Brownian motion. It is the ideal instance exemplifying the lack of scale
invariance and continuity in financial time series. As we can see, the red path still behaves in a similar fashion as the previous two plots. It wanders around unpredictably, but always in small steps. The blue path exhibits very different behaviour. Firstly, we can say that at this scale, the path moves very differently to the blue path in the two previous figures. Although the majority of steps are indeed small, the total variation is dominated by several huge jumps. Secondly, the process is obviously not continuous. We see that the price actually moves by jumps. The purist may begin a discussion about our inability to ever record a truly continuous process because we can only observe values at given instances, but without going into such arguments it is still evident that geometric Brownian motion does not have the flexibility to account for such movement. Even if we disregard the fact that the price took jumps in excess of $30 \%$ of it's entire value over time steps of only 5 minutes and consider daily data, we will still arrive at the same conclusion. At the end of March 17, the NCC stock closed at $\$ 7.49$ compared to the previous trading day's close of $\$ 13.15$. This is a loss of $43.0 \%$ when the standard deviation of daily returns for the NCC stock is only $2.1 \%$. If the price really did evolve according to geometric Brownian motion, an event like this would occur every $3.6 \times 10^{88}$ years! We can rule out this jump happening by chance because other stocks such as Citigroup Inc. and the Bear Stearns Comp. have also recently displayed sudden movements of similar magnitude. In general, financial time series of other types also exhibit discontinuity. Although share prices are most active in this light, interest rates, exchange rates and commodities also show evidence of jumps.

### 7.3 HEAVY TAILS

The two previous sub-sections can be summarised here. Both processes with non-constant volatility and jumps have distributions which display heavier tails. The term refers to the behaviour of the distribution function for large values of the random variable (by large we mean limit to infinity). We say that a random variable has a heavy right tail of the distribution if the tail
is not exponentially bounded, that is:

$$
\lim _{x \rightarrow \infty} e^{\kappa x} \mathbb{P}(X>x)=\infty \text { for all } \kappa>0
$$

The following figure illustrates heavy tailed behaviour observed in the market.


Figure 18. Distribution of daily returns of the Dow Jones and a fitted normal distribution.

In figure 18 both returns have the same mean and the same variance, so if they are both normally distributed, the two distribution curves should be the same. It seems however, that Dow Jones index returns are not normally distributed. Instead we see this heavy tail phenomenon. This means that while the two distributions have the same variance, the heavy tailed version gets more of its variability from more extreme events. This coincides with both concepts of jumps and non-constant volatility. In figure 19 are several graphs depicting empirical and normal distributions with the same daily mean and volatility for several NYSE stocks.


Figure 19. Distribution of daily returns for several NYSE companies. Blue represents the empirical distribution and red, a fitted normal distribution..

### 7.4 MARKET PRICES AGAINST THE BLACK-SCHOLES FORMULA

Another piece of evidence against the use of geometric Brownian motion is the difference in market option prices and the prices given by the standard Black-Scholes formula. Initially we may say that this is not a problem, or even a good thing since we are very confident in our model, and this difference in prices means that we can even exploit it for a profit. In our derivation
of the Black-Scholes equation and formula we argued that the assumption of no-arbitrage should give us the fair price of the option. If the market price does not coincide with this, then it is the market price that is wrong, not the model. Before we get caught up in over-confidence we should look at figure 20 which shows some interesting data.

The data is for European call option prices of three Dow Jones, NYSE listed companies. All three graphs show that the Black-Scholes formula overprices options quite considerably. While we have shown three graphs, the case is very similar for the vast majority of options. This must place a seed of doubt in our confidence regarding the Black-Scholes formula. Now it must be said that in this case we have used a model that does not take into account dividends or transaction costs which both contribute in reducing the option price. We will speak about this later but now let us consider error from parameter estimation. The current underlying stock price is observable, the time to expiry is observable, and we definitely know the strike prices.

The two variables that cause us difficulty are the interest rate and the volatility of the underlying stock. Although volatility is not observable, if we assume the stock price evolves according to geometric Brownian motion, then the volatility coefficient is just a constant ( $\sigma$ ) and it does not pose difficulty in estimation. We can simply take the standard deviation of historical returns $\left(d S_{t} / S_{t}=\mu d t+\sigma d W_{t}\right)$. For a large data set we can get very accurate estimates so under geometric Brownian motion this should not be a problem. Our other problem is the interest rate. It is observable in the market, but the obstacle here is that it is not constant as assumed in the Black-Scholes formula. Although there are pricing equations that consider non-constant interest rates, our purpose here is to investigate the performance of geometric Brownian motion. For the plots in figure 18 an approximate rate was used, taken from US T-Bill ${ }^{3}$ yield rates. We haven't accounted for dividends or transaction costs but since under geometric Brownian motion we can get very a accurate estimate for $\sigma$, we might say that the mis-pricing is due to a poor choice of the interest rate. We would be saying; since the option price and the interest rate have a positive relationship $(\partial C / \partial r>0)$, overpricing

[^2]would be due to an overestimated interest rate. However, taking values of the interest rate up to five times smaller than the observed market rate still leads to overpricing and more importantly a poor fit of the pricing curve. This is again evidence that the problem must lie in the choice of underlying process.


In light of these figures we may choose to acknowledge that the standard Black-Scholes formula is not a reliable tool to determine option prices. Indeed, the truth is that in the real world the formula is not used to price derivatives, but instead it is used to determine something called implied volatility.

The idea here is that we agree that the formula gives erroneous answers.

For vanilla ${ }^{4}$ options, markets are very large, providing high liquidity. We therefore assume that market forces determine the fair price better than our formula. Since $\sigma$ is the only unobservable variable in the formula, perhaps our estimate of $\sigma$ is incorrect. So instead of treating the parameter as input and obtaining the price, we use the price to obtain the parameter thus finding the underlying stock volatility implied by the option. This technique is used to say something about the underlying process. Of course, this reasoning automatically admits the rigidity of the standard Black-Scholes model. As stated in [13], this is the "wrong number which, plugged into the wrong formula, gives the right answer."

Figure 21 displays implied volatility surfaces for European call options on the same three company stocks as in figure 20 . These surfaces are sometimes called volatility smiles or volatility skews. None of the plots show a flat profile for the surface as would be expected if the underlying stock process really was geometric Brownian motion. The most disturbing and perhaps obvious aspect of all is the multiple values of $\sigma$ for one price! We must stress that the input parameters for the plots are strike price and time, where each plot is for a single given price of the underlying asset. Different values of $\sigma$ for one value of $S_{t}$ ? Is this not completely overwhelming evidence that the assumptions about the underlying asset price are flawed?

We stated earlier that we have not accounted for dividends or transaction costs. Both these inputs can be approximated by incorporating constants into the drift term of the underlying asset when treating the option price as an expected value under risk neutrality. In our current model, using an underestimated interest rate actually amounts to the same thing. As we saw in figure 20 this explains the over-pricing to an extent but, although we have used a very vulgar approximation by simply taking a smaller value for the rate, we still cannot obtain the correct shape of the pricing curve. Even models that accurately take into account other parameters like dividends and transaction costs simply cannot attain the correct curve. This wrong curvature results in the smile phenomenon and thus implies misplaced as-

[^3]sumptions about the underlying asset process.

We will take steps and investigate the attempts at solving the problems that have been highlighted in this section.


Implied Volatility surface for ExxonMobil options


Implied Volatility surface for Microsoft Corp. options


Figure 21. Graphs showing European call option implied volatility surfaces using closing prices from April 24 - May 16 (2008) for options expiring on May 16 (2008). The z-axis displays the volatility ( $\sigma$ for daily returns), the x-axis displaying strike price, and y-axis displaying trading days elapsed from April 24 (at 15, we are 0 trading days from expiry). Note that the empty values in-the-money are due to the Black-Scholes overpricing and our exclusion of dividends and transaction costs. Infrequent trading of in-the-money options is also a cause. It is common that even for the smallest values of $\sigma$, the pricing function cannot attain observed market prices. When this is not the case, we obtain a "smile".

## 8 ADDRESSING THE PROBLEM: STOCHASTIC VOLATILITY

### 8.1 INTRODUCTION

The first option we will explore is modeling with stochastic volatility. Our set up here is almost identical to that of the standard Black-Scholes, but the volatility coefficient itself is now a random process $\left(\sigma=\sigma_{t}\right)$.

$$
d S_{t}=\mu S_{t} d t+\sigma_{t} S_{t} d W_{t}
$$

As seen above, we do not change the underlying asset process and each increment remains log-normally distributed conditional on $\sigma_{t}$, but we need to agree on our choice of process for the volatility. It is reasonable to assume that the volatility will not attain unrealistically high values and with the evidence from figures of asset returns, we can safely say that volatility varies with time, but around some level. We can then say $\sigma_{t}$ evolves according to some mean-reverting process. In addition, it is obvious that volatility cannot be negative, so we must add the positivity restraint. In general then, $\sigma_{t}$ should be a mean reverting, positive Itô process.

$$
d \sigma_{t}=\mu_{\sigma} d t+\sigma_{\sigma} d B_{t}
$$

Another important point to make here is the relationship between $B_{t}$ and $W_{t}$. We may assume that these two are independent however, market data shows that this is not always the case. It is commonly observed that when prices take downward movements, they also become more volatile. When prices grow on the other hand, they tend to exhibit more calm behaviour. This is also summarised well from another perspective, as seen in asymmetry of returns. Losses are usually accounted for by few large movements as opposed to gains, where large gains are not as common. We will see this in
the example at the end of this sub-section. These observations are a good indicator that $B_{t}$ and $W_{t}$ should be correlated. We will say $E\left[d W_{t} d B_{t}\right]=\rho d t$, where $\rho<0$. The construction of $d W_{t}$ and $d B_{t}$ is such that:

$$
\left[d W_{t}, d B_{t}\right]^{\prime}=M\left[d \hat{W}_{t}, d \hat{B}_{t}\right]^{\prime}
$$

Where M is a $2 \times 2$ matrix and $d \hat{W}_{t}$ and $d \hat{B}_{t}$ have no correlation. $M M^{\prime}=\Sigma$ is a Cholesky decomposition and $\Sigma$ is the covariance matrix of $d W_{t}$ and $d B_{t}$.

Before we move on, here are some examples of popular processes used in volatility modeling:

Square root process: $\quad d \sigma_{t}=a\left(b-\sigma_{t}\right) d t+c \sqrt{\sigma_{t}} d B_{t}$ Log-Ornstein-Uhlenbeck: $\quad d \log \sigma_{t}=a\left(b-\log \sigma_{t}\right) d t+c d B_{t}$

Heston model (here we do change $d \sigma_{t}=a\left(b-\sigma_{t}\right) d t+c \sqrt{\sigma_{t}} d B_{t}$ the underlying asset process): $\quad d S_{t}=\mu S_{t} d t+\sqrt{\sigma_{t}} S_{t} d W_{t}$

The choice of model is a question of the modeler's beliefs about the underlying processes. We can forgo this debate and move on to pricing. In the next sub-section we will use the Black-Scholes equation to derive a general pricing equation.

### 8.2 BLACK-SCHOLES EQUATION

Using our general Black-Scholes equation, the derivation here is surprisingly simple. To move on we must nonetheless agree on what has changed from the standard Black-Scholes setup.

Standard

$$
\begin{aligned}
& d S_{t}=\mu S_{t} d t+\sigma S_{t} d W_{t} \\
& \sigma_{t}=\sigma
\end{aligned}
$$

$$
d \beta_{t}=r \beta_{t} d t
$$

$$
Y_{t}=f\left(S_{t}, t\right)
$$

Stochastic volatility
$d S_{t}=\mu S_{t} d t+\sigma_{t} S_{t} d W_{t}$
$d \sigma_{t}=\mu_{\sigma} d t+\sigma_{\sigma} d B_{t}$
$E\left[d W_{t} d B_{t}\right]=\rho d t$
$d \beta_{t}=r \beta_{t} d t$
$Y_{t}=f\left(S_{t}, \sigma_{t}, t\right)$

We have already discussed the nature of the volatility process, what remains to be noticed is the change of the derivative pricing function. It would have been a mistake on our part to continue without realising that $Y_{t}$ has changed. Since we have introduced a second random component, the pricing function now has dependence on this new variable. This may seem obvious but we can consider the standard setup for another interpretation. We know that a derivative is more expensive if the underlying asset has higher volatility. If volatility is now allowed to vary over time, obviously the derivative will have lower prices over periods when $\sigma_{t}$ is low, and higher prices over periods when $\sigma_{t}$ is high. Hence the dependence on $\sigma_{t}$ in the pricing function and the modification of $Y_{t}$.

Recall that the general Black-Scholes equation is as given below:

$$
\mu_{Y}-R Y_{t}=\sigma_{Y} \Lambda
$$

To continue we must have expressions for $\mu_{Y}$ and $\sigma_{Y}$ so we must use Itô's lemma to compute $d Y_{t}$.

$$
\begin{aligned}
d Y_{t}= & \frac{\partial f\left(S_{t}, \sigma_{t}, t\right)}{\partial S_{t}} d S_{t}+\frac{\partial f\left(S_{t}, \sigma_{t}, t\right)}{\partial t} d t+\frac{\partial f\left(S_{t}, \sigma_{t}, t\right)}{\partial \sigma_{t}} d \sigma_{t} \\
& +\frac{1}{2}\left(\frac{\partial^{2} f\left(S_{t}, \sigma_{t}, t\right)}{\partial S_{t}^{2}}\left(d S_{t}\right)^{2}+\frac{\partial^{2} f\left(S_{t}, \sigma_{t}, t\right)}{\partial \sigma_{t}^{2}}\left(d \sigma_{t}\right)^{2}+2 \frac{\partial^{2} f\left(S_{t}, \sigma_{t}, t\right)}{\partial S_{t} \partial \sigma_{t}}\left(d \sigma_{t}\right)\left(d S_{t}\right)\right) \\
= & \left(\mu S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\mu_{\sigma} \frac{\partial Y_{t}}{\partial \sigma_{t}}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}+\frac{1}{2} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial \sigma_{t}^{2}}+\rho \sigma_{t} S_{t} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial S_{t} \partial \sigma_{t}}\right) d t \\
& +\left[\sigma S_{t} \frac{\partial Y_{t}}{\partial S_{t}}, \sigma_{\sigma} \frac{\partial Y_{t}}{\partial \sigma_{t}}\right]\left[d W_{t}, d B_{t}\right]^{\prime}
\end{aligned}
$$

The price of the derivative must then satisfy the following differential equation:

$$
\begin{aligned}
\mu S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\mu_{\sigma} \frac{\partial Y_{t}}{\partial \sigma_{t}} & +\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}+\frac{1}{2} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial \sigma_{t}^{2}}+\rho \sigma_{t} S_{t} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial S_{t} \partial \sigma_{t}}-R Y_{t}= \\
& =\left[\sigma S_{t} \frac{\partial Y_{t}}{\partial S_{t}}, \sigma_{\sigma} \frac{\partial Y_{t}}{\partial \sigma_{t}}\right]\left[\Lambda_{W}, \Lambda_{B}\right]^{\prime}
\end{aligned}
$$

It is very interesting to observe the nature of $\Lambda$, the price of risk. In the standard Black-Scholes setup there is only one source of uncertainty, the Brownian motion driving the underlying asset price. Here on the other hand, we have two Brownian motions, two sources of uncertainty, and therefore a price of risk associated with each. Under our current model assumptions, this equation can be simplified a little further. Considering that investment in the money market grows at a constant risk-free rate, taking the bond as numéraire means that $R$ is actually the specified risk-free rate $r$. We can also say something about $\Lambda_{W}$. This is the price of risk associated with the Brownian motion driving the underlying asset price. From using the Black-Scholes formula for the underlying asset we notice that the price of risk is $(\mu-r) / \sigma_{t}$. This of course is the price of risk associated with $W_{t}$. Hence $\Lambda_{W}=(\mu-r) / \sigma_{t}$.

Finally, the derivative price must satisfy the following differential equation:

$$
r S_{t} \frac{\partial Y_{t}}{\partial S_{t}}+\frac{\partial Y_{t}}{\partial t}+\left(\mu_{\sigma}-\sigma_{\sigma} \Lambda_{B}\right) \frac{\partial Y_{t}}{\partial \sigma_{t}}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} Y_{t}}{\partial S_{t}^{2}}+\frac{1}{2} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial \sigma_{t}^{2}}+\rho \sigma_{t} S_{t} \sigma_{\sigma} \frac{\partial^{2} Y_{t}}{\partial S_{t} \partial \sigma_{t}}-r Y_{t}=0
$$

This is a very general pricing equation, and in fact we can see that if $\sigma_{t}$ is constant, all quotients with dependence on $\sigma_{t}$ obviously become 0 . The equation then reduces to that which is obtained under a constant volatility model. Although we have made an assumption about the underlying asset process, that the increments are still geometric Brownian motion conditional on $\sigma_{t}$, it would not be difficult to generalise this to other processes well.

As we have seen, the derivation of this equation is not difficult with our current tools in hand. Even finding solutions to this is not such a difficult problem. Without even delving into finding explicit solutions we may use a numerical approach like the finite differences method. As we will see in the section below, we may also use risk-neutral pricing with the Monte-Carlo method as a tool. What causes more problems for us is the parameter estimation, in particular with respect to $\Lambda_{B}$. Assuming only absence of arbitrage
it is not generally possible to specify $\Lambda_{B}$. This can be done by developing a full general equilibrium model of asset prices. There have also been recent developments in nonparametric estimation techniques using observed derivative prices [6]. We will leave this topic and move onto to the next sub-section, where we will compare prices given by the standard setup and the setup allowing for stochastic volatility.

### 8.3 RISK NEUTRAL PRICING

We can also use risk-neutrality as a tool to price derivatives when volatility is stochastic. As with the previous sub-section, to work through we simply need to keep in mind that the volatility is now random and adjust accordingly. Recall that the price of a derivative paying some amount at time time has the current price as follows:

$$
C_{0}=e^{-\delta t} E\left[f\left(S_{t}, \sigma_{t}, t\right) \mid S_{0}, \sigma_{0}\right]
$$

However, difficulties in estimating the risk-adjusted discount rate $\delta$ means that we must assume no-arbitrage and use risk-neutrality as a tool.

$$
C_{0}=e^{-r t} E\left[f\left(\tilde{S}_{t}, \tilde{\sigma}_{t}, t\right) \mid \tilde{S}_{0}=S_{0}, \tilde{\sigma}_{0}=\sigma_{0}\right]
$$

Finding explicit expressions of this expectation may be difficult but we can always use Monte-Carlo simulation. We can use the following example to compare stochastic volatility modeling and the standard Black-Scholes model.

Example (Modeling volatility by the log-Ornstein-Uhlenbeck process) In this example we will price a European call option. We will compare the prices given by the standard Black-Scholes formula with prices given assuming stochastic volatility. To calculate the latter, we will use Monte-Carlo simulation. Our model is as follows:

$$
\begin{aligned}
& d S_{t}=\mu S_{t} d t+\sigma_{t} S_{t} d W_{t} \\
& d \log \sigma_{t}=a\left(b-\log \sigma_{t}\right) d t+c d B_{t} \\
& d \beta_{t}=r \beta_{t} d t \\
& E\left[d W_{t} d B_{t}\right]=\rho d t \\
& \mu=0.15, \quad a=1, \quad b=-1.6, \quad c=0.8, \quad r=0.05, \quad \rho=-0.75
\end{aligned}
$$

We will be looking at several strike prices and exercise times.


Figure 22. This figure illustrates the relationship between the underlying volatility and the asset price process. The left plot shows the volatility with the stationary average $\left(e^{b+c^{2} / 4 a}\right)$. Notice the increase in variability of the asset price process towards the end, corresponding with the increase of the volatility coefficient. Notice also the sharp rise in volatility over $t \approx(0.7,0.9)$. This is also the time when the asset price shows a slight downward trend. This illustrates the correlation between the two driving Brownian motions.

Figures 22 and 23 illustrate some properties of our model. Notice the resemblance of the empirical distribution in figure 22 to that of the distributions seen in sub-section 7.3.

In comparing the outcomes, the first thing we want to do is obtain prices using the standard Black-Scholes formula. To do this we assume that we do not know that the process has stochastic volatility and simply treat it as


Figure 23. Here we see the empirical distribution and density functions of daily returns in blue, made by generating 200000 paths. The red curves are for a fitted normal distribution.
geometric Brownian motion.

$$
d S_{t}=\nu S_{t} d t+\theta S_{t} d W_{t}
$$

This means that the volatility coefficient is constant. We can look at returns of this process to obtain an estimate for $\theta$. Having all other parameters, we then use the formula to obtain our prices.

Our next step is to look at what happens if we now know that the process has random volatility. We will use risk neutral pricing to obtain the answer and to do this we must first find out how the process evolves under risk neutrality. This is the same as finding an equivalent martingale measure with the bond taken as numéraire. As done before, we use Itô's lemma to find how the discounted price evolves. We must then find $\lambda$ in $d \tilde{W}_{t}=\lambda_{t} d t+d W_{t}$ such that the following process is a martingale:

$$
d \frac{S_{t}}{\beta_{t}}=(\mu-r) \frac{S_{t}}{\beta_{t}} d t+\sigma_{t} \frac{S_{t}}{\beta_{t}} d W_{t}
$$

This yields $\lambda_{t}=(\mu-r) / \sigma_{t}$. Notice that the only difference here to the derivation in sub-section 4.5 is that $\sigma$ is random, otherwise the structure is the same. With this knowledge we can state that the underlying asset price
evolves according to:

$$
d \hat{S}_{t}=r \tilde{S}_{t} d t+\sigma_{t} \tilde{S}_{t} d \tilde{W}_{t}
$$

This is not the final step in finding the risk-neutral process. As we saw in the previous section, there is also a price of risk and hence a risk premium associated with the Brownian motion driving the volatility. We need to find $\Lambda_{2}$ in the expression below. However, considering the difficulties associated with estimating $\Lambda_{2}$, we will assume that this risk premium is zero in this example.

$$
d \log \tilde{\sigma}_{t}=a\left(b-\log \tilde{\sigma}_{t}\right) d t+c\left(-\Lambda_{2} d t+d \tilde{B}_{t}\right)
$$

In other words we assume that $d \sigma_{t}=d \tilde{\sigma}_{t}$. With this we have our risk neutral process:

$$
d \hat{S}_{t}=r \tilde{S}_{t} d t+\tilde{\sigma}_{t} \tilde{S}_{t} d \tilde{W}_{t}
$$

Below is an illustration showing comparison in valuation of a European call option with one month to expiry using our two different setups. The initial stock price is $\$ 10$. The initial $\sigma_{t}$ value was specified as having the stationary distribution of the volatility process. This results in obtaining an "average" price over all possible initial $\sigma_{t}$ values. As we can see, there is a significant difference in these curves. The stochastic volatility price of the option has a higher degree of curvature at the money, which agrees with prices seen in the market. Another characteristic seen here is the slightly higher value given by the stochastic volatility approach well in and out of the money. Both these phenomena can be explained by the heavier tailed distribution of the asset returns under stochastic volatility.


Figure 24. Comparing option prices given by models with and without stochastic volatility. The blue line gives the risk-neutral price obtained by using Monte-Carlo simulation, assuming stochastic volatility. The dotted line gives the Black-Scholes price assuming that the underlying process is a geometric Brownian motion. 100000 draws were used.

## Implied Volatility surface



Figure 25. Implied volatility surface obtained by using the standard Black-Scholes formula on the prices given by the stochastic volatility model. At 30 days, we are 0 days from option expiry.

The above plot also shows how stochastic volatility can help in explaining the smile phenomenon. This is also a result of the heavier tails of the distribution of asset returns.

### 8.4 CONSEQUENCES OF USING STOCHASTIC VOLATILITY

The example in the previous sub-section helps to show the magnitude of improvement seen in introducing stochastic volatility. However there is also added difficulty in implication. For example we have not spoken about parameter estimation in stochastic volatility models. The fact that the volatility coefficient cannot be observed is an obvious barrier to tractability. There is also the question of the risk premium associated with the underlying volatility. We have stated that estimation of this parameter is also troublesome. Nevertheless, there are developments in these fields and there are certain models where parameters can be estimated with a degree of accuracy. Overall, the introduction of a stochastic volatility framework is a vast improvement on the standard Black-Scholes model.

There is the obvious benefit in more accurate fitting of the underlying asset returns. In addition to producing heavier tails, we can also account for the asymmetry of returns. We can even account for the correlation of returns. There is the great step in explaining the volatility smile, although it has been shown that this model does not fare so well over time horizons. They cannot yield a realistic term structure [13]. We can also produce perfect hedging strategies due to the retained continuity of the underlying asset price.

Retained continuity is also the point that is a sizable drawback for this model. After all, discontinuity of returns is something it cannot explain. Reproducing the jumps seen in the market is possible only at the expense of ridiculously high and sudden values of volatility of volatility. Jump models however manage to capture this phenomenon as a generic property. Jump models also manage to produce a variety of volatility smiles and perform much better over maturities [1].

Finally in the next section we introduce jumps.

## 9 ADDRESSING THE PROBLEM: INTRODUCING JUMPS

### 9.1 THE COMPOUND POISSON PROCESS

To begin we need to introduce something to drive the jumps in our model. This is done by taking a rather elementary tool.

As we reasoned before, these jumps that we see are a result of arrivals of important pieces of information (about the underlying asset). We may assume that these arrivals are independent. We can describe the probability of an event occurring over the time interval $(t, t+d t)$ in the following way:
$\mathbb{P}(J u m p$ does not occur in the time interval $(t, t+d t))=1-\lambda d t+o(d t)$
$\mathbb{P}(J u m p$ occurs once in the time interval $(t, t+d t))=\lambda d t+o(d t)$
$\mathbb{P}($ Jump occurs more than once in the time interval $(t, t+d t))=o(d t)$
Where $o(d t)$ is negligible with respect to $d t$.

If the jump sizes are all of the same magnitude, this construction results in the very well know Poisson process where $\mathbb{P}\left(\Delta N_{t}=k\right)=e^{-\lambda \Delta t}(\lambda \Delta t)^{k} / k$ !. The next step is to consider what happens when one of these important pieces of information arrives. Information of different kinds has different bearings on the asset price. In other words the effect of each arriving piece of information is random, i.e. the magnitude of the jump is random. This added feature generates a compound Poisson process.

A compound Poisson process with intensity $\lambda>0$ and jump size distribution
$f$ is a stochastic process $X_{t}$ defined as:

$$
X_{t}=\sum_{i=1}^{N_{t}} Y_{i}
$$

Where the $Y_{i}$ are independently, identically distributed and $N_{t}$ is a Poisson process with intensity $\lambda$, independent from $Y_{i}$.


Figure 26. On the left we see a Poisson process with intensity $\lambda=5$ and jumps of size 1. On the right is a compound Poisson process generated by the process on the left, with standard normal distribution of jump sizes.

Using conditional expectation and the law of total variance we can also write down the expressions for expected value and variance of the compound Poisson process:

$$
\begin{aligned}
& E\left[X_{t}\right]=E\left[E\left[X_{t} \mid N_{t}\right]\right]=E\left[N_{t} E[Y]\right]=\lambda t E[Y] \\
& \begin{aligned}
\operatorname{Var}\left[X_{t}\right] & =E\left[\operatorname{Var}\left[X_{t} \mid N_{t}\right]\right]+\operatorname{Var}\left[E\left[X_{t} \mid N_{t}\right]\right] \\
& =E\left[N_{t} \operatorname{Var}[Y]\right]+\operatorname{Var}\left[N_{t} E[Y]\right] \\
& =\lambda t \operatorname{Var}[Y]+E[Y]^{2} \operatorname{Var}\left[N_{t}\right] \\
& =\lambda t\left(\operatorname{Var}[Y]+E[Y]^{2}\right) \\
& =\lambda t\left(E\left[Y^{2}\right]-E[Y]^{2}+E[Y]^{2}\right) \\
& =\lambda t E\left[Y^{2}\right]
\end{aligned}
\end{aligned}
$$

In the next sub-section we will incorporate this process into our underlying asset price evolution.

### 9.2 JUMP DIFFUSION

As we saw with stochastic volatility, the introduction of the idea was not difficult at all. It is the same here. All we do is take our ordinary diffusion process and add a compound Poisson process.

$$
d S_{t}=\mu_{S} S_{t} d t+\sigma_{S} S_{t} d W_{t}+S_{t} d X_{t}
$$

Where $X_{t}=\sum_{i=1}^{N_{t}} Y_{i}$ as in the previous sub-section.

The change in the asset price is presented as a composition of two types of changes. The standard "vibrations" in price due to short-term changes in supply and demand, changes in traders' outlooks and other marginal effects. This part is modeled as before, by using a Brownian motion to drive the stochasticity. The second component consists of shocks to the system pointed out earlier. These are due to the arrival of important new information with specific relevance to the asset in question. This is the kind of "rare" information that we model by events at discrete points in time. These points in time are exactly when jumps occur, reflecting the non-marginal effect of the information. This second type of change is modeled by a compound Poisson process.

What is left to us is the specification of the drift term and $Y$. Ideally the drift term should be specified in such a way that $\mu_{S}=\mu-\eta$ where $\eta$ compensates the compound Poisson process. This is not a strict requirement, but makes things much easier when tempering with the drift of the process since it would be determined only by $\mu$.

We are also free in our specification of $Y$. This depends on our beliefs about the jumps in the underlying process. The choice is important since it determines to a large extent the tails of the distribution of returns. There are
two popular models; the original jump model proposed by Merton [12] and a double exponential jump model proposed by Kou [8].

$$
\begin{array}{ll}
\text { Merton } & \text { Kou } \\
Y \sim N\left(\nu, \theta^{2}\right) & \mathbb{P}(\text { positive jump })=p \\
& \mathbb{P}(\text { negative jump })=1-p \\
& \frac{1}{p} Y_{+} \sim \exp \left(\lambda_{1}\right)_{+} \\
& \frac{1}{1-p} Y_{-} \sim \exp \left(\lambda_{2}\right)_{-}
\end{array}
$$

Both models are an improvement but it has been shown that the Kou model is a better fit to market data [8]. The use of the exponential distribution means that the tails of returns decay exponentially (much slower) than the tails of Merton model returns. This is a better reflection of market data and the better fit subsequently results in more accurate pricing and implied volatility modeling. Due it's simplicity however, we will investigate the Merton model in an example.

### 9.3 PRICING

In stochastic volatility modeling we can still use risk-neutral pricing without alteration to our fundamental assumptions. Although finding this measure may prove more difficult due to the obvious increased complexity of the model, there is still just one risk-neutral measure.

Pricing in jump models introduces a problem called incompleteness. Two sources of randomness driving the evolution of the price means that we have many choices for measure change. To obtain a process that is expected to grow at the risk-free rate we may change the measure of the diffusion, of the jump, or of both components. To make this choice we usually need to make additional assumptions about the market. In the next example, we will see that it is assumed that jumps can be "diversified" away, by holding many assets. This means that only the drift of the diffusion is changed, attaching no risk premium to the jump component. We have an additional
assumption that the losses and gains incurred as a result of jumps will even themselves out over many assets and time. This may not be the best choice of measure change, because it is only useful if jumps across assets and time are completely independent. However, markets show that when these jumps occur, they are not independent (during a market crash many stocks may exhibit downward jumps at the same time). Attaching no risk premium to the jumps is just one possible choice and an illustration of the matter at hand. This problem actually adds realism to our model. If there were no jumps in reality, all market participants could perfectly price and hedge away risk. This is not the case in and better or worse choices can be made. Ultimately, it is up to the user to make assumptions depending on their beliefs about the market. Once we have a choice of measure change, we can take the expectation of the option payoff and then discount to obtain the current option price. The following example shows how we can price a European call option via a jump model.

Example (Option pricing via the Merton model) Let us investigate the setup:

$$
\begin{aligned}
d S_{t}= & (\mu-\lambda \kappa) S_{t} d t+\sigma S_{t} d W_{t}+S_{t} d X_{t} \\
X_{t}= & \sum_{i=1}^{N_{t}} Y_{i} \\
& Y \text { i.i.d. } N\left(\nu, \theta^{2}\right) \text { and } N_{t} \text { is a Poisson process with intensity } \lambda \\
\kappa= & E\left[e^{Y}-1\right] \\
d \beta_{t}= & r \beta_{t} d t
\end{aligned}
$$

$$
\mu=0.10, \quad \sigma=0.20, \quad \lambda=10, \quad \nu=0, \quad \theta=0.10, \quad r=0.05
$$

The first thing that stands out is the choice of drift term. Why this $\lambda \kappa$ ? As stated above, from the onset it would be preferable to specify just one parameter that determines the growth rate. We must consider that the inclusion of the compound process can have a contribution towards the growth if the $Y_{i}$ have non-zero mean. Let us look at what happens if we have one jump, $Y$. Ignoring the Brownian motion, it's contribution to the stock price is a change $S_{t+d t}=S_{t} e^{Y}$. The percentage change contribution is then $\left(e^{Y}-e^{0}\right) / e^{0}$. We


Figure 27. A figure showing one path of underlying asset price under Merton model dynamics. The parameters are $\mu=0.10, \sigma=0.20, \lambda=10$, $\nu=0, \theta=0.10$. The plot on the right shows how we would observe the process without knowledge of the jump times.
can specify the overall drift component in such a way that $\mu$ is the parameter determining growth. To do this we should offset any contribution by the compound Poisson process. The key in recognising that we need to subtract the percentage change contribution of the process per unit of time. This is done by subtracting $E\left[N_{1}\left(e^{Y}-e^{0}\right) / e^{0}\right]=\lambda E\left[e^{Y}-1\right]$. The next figure gives an example of the process.

Figure 27 illustrates a problem resulting from introduction of jumps. If we do not know when the jumps occur it is very difficult for us to decompose the process and hence there are difficulties in parameter estimation, though we will not go into this here.

In valuing a European call option on this underlying asset we can use riskneutral pricing.

The notion of incompleteness, a result of two sources of randomness contributing to the drift means that we have many choices for risk-neutral measure. Merton proposed changing only the drift of the Brownian motion (as we have done previously). The choice was justified by arguing that "jump risk" is diversifiable and there is therefore no risk premium attached to it


Figure 28. Comparison of empirical distribution and density functions of daily returns for the Merton model and a model assuming normal returns.
(though there is evidence that this is far from the truth). We will take this approach for this example. Assuming there is no risk premium associated with the jumps, we find the risk neutral measure of the Brownian motion in the same way as we have done many times before. This yields the following risk neutral process.

$$
d \tilde{S}_{t}=(r-\lambda \kappa) \tilde{S}_{t} d t+\sigma \tilde{S}_{t} d \tilde{W}_{t}+\tilde{S}_{t} d X_{t}
$$

We can also write down the solution (the steps taken to find it are almost identical to finding the solution of geometric Brownian motion).

$$
\tilde{S}_{t}=\tilde{S}_{0} e^{\left(r-\sigma^{2} / 2-\lambda \kappa\right) t+\sigma \tilde{W}_{t}+X_{t}}
$$

The European call option price is then given by:

$$
C_{0}=e^{-r t} E\left[\max \left(\tilde{S}_{0} e^{\left(r-\sigma^{2} / 2-\lambda \kappa\right) t+\sigma \tilde{W}_{t}+X_{t}}-K, 0\right) \mid \tilde{S}_{0}=S_{0}\right]
$$

$X_{t}=\sum_{i=1}^{N_{t}} Y_{i}$ with $X_{t}$ distributed as $N\left(n \nu, n \theta^{2}\right)$ when $N_{t}=n$ is given. By conditioning on the number of jumps, we can express the price as a weighted sum of Black-Scholes prices:

$$
C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\max \left(\tilde{S}_{0} e^{\left(r-\sigma^{2} / 2-\lambda \kappa\right) t+\sigma \tilde{W}_{t}+i Y}-K, 0\right) \mid \tilde{S}_{0}=S_{0}\right]
$$

We can make a clever step here. As it stands, we take expectation over two normal random variables $i Y$ and $\tilde{W}_{t}$. However, if we can "eliminate" $i Y$, we can take the expected value over just $\tilde{W}_{t}$. We make this elimination by adding/subtracting a transformation of $\tilde{W}_{t}$ that has the same variance as $i Y$. Consider $\pm i \nu \pm \theta \sqrt{\frac{i}{t}} \tilde{W}_{t}$. We are not saying that $i Y=i \nu+\theta \sqrt{\frac{i}{t}} \tilde{W}_{t}$, we are saying that taking the expected value under this substitution will not change the answer.

$$
\begin{aligned}
& C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\left.\max \left(\tilde{S}_{0} e^{\left(r-\sigma^{2} / 2-\lambda \kappa\right) t+\sigma \tilde{W}_{t}+i Y \pm i \nu \pm \theta \sqrt{\frac{i}{t}} \tilde{W}_{t}}-K, 0\right) \right\rvert\, \tilde{S}_{0}=S_{0}\right] \\
& C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\left.\max \left(\tilde{S}_{0} e^{\left(r-\sigma^{2} / 2-\lambda \kappa\right) t+\sigma \tilde{W}_{t}+i \nu-\theta \sqrt{\frac{i}{t}} \tilde{W}_{t}}-K, 0\right) \right\rvert\, \tilde{S}_{0}=S_{0}\right]
\end{aligned}
$$

Notice that we have taken $-\theta \sqrt{\frac{i}{t}} \tilde{W}_{t}$ and not $+\theta \sqrt{\frac{i}{t}} \tilde{W}_{t}$. Once again, we may do this because the expectation will be the same. Our choice depends on the form we desire in our answer. We also take $\pm i \theta^{2} / 2$.

$$
\begin{gathered}
C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\left.\max \left(\tilde{S}_{0} e^{i \nu-\lambda \kappa t} e^{\left(r-\sigma^{2} / 2\right) t+\sqrt{\sigma^{2}+\frac{\theta^{2}}{t}} \tilde{W}_{t} \pm i \theta^{2} / 2}-K, 0\right) \right\rvert\, \tilde{S}_{0}=S_{0}\right] \\
C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\left.\max \left(\tilde{S}_{0} e^{i \nu+i \theta^{2} / 2-\lambda \kappa t} e^{\left(r-\left(\sigma^{2}+\theta^{2} / t\right) / 2\right) t+\sqrt{\sigma^{2}+\frac{\theta^{2} i}{t}} \tilde{W}_{t}}-K, 0\right) \right\rvert\, \tilde{S}_{0}=S_{0}\right] \\
C_{0}=e^{-r t} \sum_{i=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{i}}{i!} E\left[\max \left(\tilde{S}_{i} e^{\left(r-\sigma_{i}^{2} / 2\right) t+\sigma_{i} \tilde{W}_{t}}-K, 0\right) \mid \tilde{S}_{0}=S_{0}\right]
\end{gathered}
$$

$$
\text { where } \quad \tilde{S}_{i}=\tilde{S}_{0} e^{i \nu+i \theta^{2} / 2-\lambda \kappa t} \quad \text { and } \quad \sigma_{i}^{2}=\sigma^{2}+\frac{\theta^{2} i}{t}
$$

As we can see, the expectation in the framed expression is simply the BlackScholes price of a European call option but with new given values for volatility and initial underlying asset price. The Merton model price is a weighted sum of standard Black-Scholes prices.


Figure 29. Comparing option prices given by the standard Black-Scholes and Merton models. The blue line gives the Merton model price. The dotted blue line gives the Black Scholes price assuming that the underlying process is a geometric Brownian motion.

Figure 29 shows the comparison in prices given by the standard Black-Scholes and the Merton model for a European option with expiry in one month. The initial underlying asset price is $\$ 10$. Notice that the Merton price has a higher degree of convexity at the money, as seen in market data. This again goes towards explaining the volatility smile phenomenon. Note that to obtain the standard Black-Scholes price we assume that the underlying process is geometric Brownian motion and by using the variance of asset returns obtain $\sigma_{G B M}^{2}=\sigma^{2}+\lambda \theta^{2}$.

Figure 30 illustrates the implied volatility surface generated by using the Merton model. The introduction of jumps and resulting heavier tails of the distribution of asset returns results in the observed pattern.

## Implied Volatility surface



Figure 30. Implied volatility surface obtained by using the standard Black-Scholes formula on the prices given by the Merton model. At 30 days, we are 0 days from option expiry.

### 9.4 CONSEQUENCES OF USING JUMP MODELS

Introducing jumps attains two important goals in the aim of correctly modeling observed financial phenomena. This first is capturing the underlying asset price dynamics. As we have discussed, jumps are commonly seen due to the arrivals of important, non-negligible information at discrete times. This
is especially evident in illiquid markets, where such models are of particular importance. With the introduction of jumps, we simultaneously explain the volatility smile phenomenon, and it has been shown that pricing with jumps provides a better fit than stochastic volatility models for the surface.

A further important outcome of using jumps is the loss of perfect hedging strategies. Perfect hedging is something that both geometric Brownian motion and stochastic volatility models admit. This means that under the assumptions of these continuous models, options are made redundant since they can be perfectly replicated by existing assets. In the financial world, it is indeed impossible to attain a perfect hedging strategy, and there are actually multiple techniques available depending on the preference of the user. They are determined according to which risks a user most wants to avoid. Some of these techniques are superhedging, utility maximization, and quadratic hedging, where all include some kind of optimization. Jump models on the other hand help in explaining the existence of option markets. Options allow better allocation of risk among market participants, which was the original reason behind their invention.

Although using jump models gives a great degree of flexibility in modeling the tail behaviour of returns, there is a drawback. They do not help in explaining the correlation of returns. As specified in our model, the jumps have independent times and ignoring the jumps, we still have a geometric Brownian motion with independent returns. This does not help in explaining the commonly observed effect of volatility clustering.

There is the added difficulty in parameter estimation due to our inability to distinguish jumps and the "normal" diffusion movements, but this is an inevitable consequence of introducing a more complicated model. Overall, like stochastic volatility models, jump diffusion models are a vast improvement above the standard Black-Scholes model as they help explain several market phenomena and provide a better fit to market prices.

## 10 CONCLUSION

Mathematical finance has beginnings from very simple constructions, with great properties. The entire theory is built with only one assumption - that there is no arbitrage. This is extremely useful because the assumption is not about how the market actually behaves, but an assumption made to determine fair prices. Crucially, we have not assumed anything about the market participants' beliefs or preferences. It has been pointed out that this is a very sharp and nimble piece of theory and it deserves the full scope of credit that it receives.

It has also been pointed out that while the theory is beautiful, and the assumption of no-arbitrage is good, the choices we make in modeling the underlying market processes have a significant influence on creating realistic physical descriptions. The world beyond standard Black-Scholes should not be concealed at any cost. Ideally the theory should be introduced along with the evidence that it is a simplification and the path to overcoming the difficulties should always be shown ahead.

It was a goal to illustrate that it is not at all difficult to show this path. As intended, the second part of the thesis is presented in relatively simple terms, even more simple then the first. This should show that the aura surrounding modeling beyond geometric Brownian motion is very artificial. Introducing stochastic volatility or jumps is extremely natural and should not be looked upon with fear.

The second part is indeed only an introduction to modeling beyond geometric Brownian motion. To include the full scope of current models we would have needed more than several volumes. Stochastic volatility or jump models on their own are not the best available to capture financial process dynamics. There are subordination models where the time steps taken are themselves random, so the data is viewed in "financial time". There are infinite activity
models where the process moves only by jumps. The class of so called Lévy process provides an extremely rich framework for financial modeling. We can even go beyond Lévy processes to additive models without stationary increments. There are models such as the Bates, where we assume both jumps and stochastic volatility. They are an even better fit to market data.

The above topics are left for the future, but I hope that I have attained my goal - to show my understanding of the topics and issues covered and illustrate that they are not as complicated as is sometimes perceived.

## References

[1] Cont, R. \& Tankov, P. Financial Modeling with Jump Processes. Chapman \& Hall/CRC, 2004.
[2] Evans, L. C. "Introduction to Stochastic Differential Equations" Course Notes. University of California, Berkley, California, USA
[3] Feller, W. An Introduction to Probability Theory and Its Applications, Volume 2. John Wiley \& Sons, 1966.
[4] Gamarnik, D. "Advanced Stochastic Processes" Course Notes. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, USA 2005.
[5] Guthrie, G. "Derivatives" Course Notes. Victoria University of Wellington, Wellington, New Zealand, 2007.
[6] Guthrie, G. "Options" Course Notes. Victoria University of Wellington, Wellington, New Zealand, 2006.
[7] Kohn, R. V. "Continuous Time Finance" Course Notes. Section 10, Courant Institute of Mathematical Sciences, New York University, New York, USA, 2004
[8] Kou, S. G. A jump diffusion model for option pricing. Management Science, Vol. 48, 1086-1101, 2002.
[9] Kou, S. G. \& Wang, H. Option pricing under a double exponential jump diffusion model. Management Science, Vol. 50, 1178-1192, 2004.
[10] Lalley, S. "Mathematical Finance" Course Notes. Section 10, University of Chicago, Illinois, USA, 2001
[11] Merton, R. C. Continuous-Time Finance. Blackwell Publishing Ltd, 1990.
[12] Merton, R. C. Option Pricing When Underlying Stock Returns are Discontinuous. Journal of Financial Economics 3, January-February 1976.
[13] Rebonato, R. Volatility and Correlation in the Pricing of Equity, FX and Interest Rate Options. John Wiley \& Sons, 1999.
[14] Schumacher, J. M. "Financial Models" Course Notes. Tilburg University, Netherlands, 2007.
[15] Shreve, S. Stochastic Calculus for Finance II: Continuous-Time Models. Springer Finance, 2004.
[16] Taylor, H. M. \& Karlin, S. An introduction to Stochastic Modeling, Third Edition. Academic Press, 1998
-Financial data obtained from www.finance.yahoo.com and www.scottrade.com


[^0]:    ${ }^{1}$ Fischer Black (1938-1995), Myron Scholes, Nobel Laureate 1997

[^1]:    ${ }^{2}$ The Dow Jones Industrial Average is a weighted average of stocks from 30 of the largest companies in the United States.

[^2]:    ${ }^{3}$ US Treasury Bills, a risk-free short term (less than a year, usually at 4 weeks, 13 weeks, 26 weeks) debt obligation backed by the US government.

[^3]:    ${ }^{4}$ A vanilla option is any "common", simple option without any special features. There is no definitive distinction between these and exotic options. The latter tends to be specialised. European, American, Bermudan are all examples of vanilla options.

