Utrecht University

Bachelor Thesis Mathematics

Solving the Black-Scholes equation using Martingales

Author:
Tim Waroux
5971241

Supervisor:
Dr. Karma Dajani

June 14, 2019

Utrecht University

Contents

Acknowledgements **2**

Introduction **3**

1 Expectation and Conditional Expectation **4**
1.1 Expectation 4
1.2 Conditional Expectation 5

2 Martingales and Brownian Motion **9**
2.1 Stochastic Processes 9
2.2 Martingales 9
2.3 Brownian Motion 11

3 Itô Integral **16**
3.1 Itô integral for Simple Processes 16
3.2 Itô's Integral in the General Case 19

4 The Black–Scholes Equation **21**
4.1 Itô Process 21
4.2 The Itô-Doeblin Formula 21
4.3 The Black-Scholes Equation 23
4.4 Girsanov's Theorem 26
4.5 Derivation of the Black-Scholes Formula 28

5 Parameter Estimation **30**
5.1 Maximum Likelihood 30
5.2 Estimating the Volatility 31
5.3 Example for AEX 32

6 Conclusion **35**

7 References **36**

Appendix **37**

Acknowledgements

I would like to thank my supervisor Dr. Karma Dajani for coming up with the idea to write about solving the Black-Scholes equation using Martingales. Also I would like to thank her for all the effort and time she put in helping me while writing the thesis itself.

Introduction

Martingales are very important when we talk about stochastic processes. Martingales have a constant expectation regardless of time. In Financial Mathematics Martingales are crucial, when we model a financial asset as a random process we want to use a different measure under which this process is a Martingale. This particular measure is often called the risk-neutral measure and we will be using this to solve the Black-Scholes partial stochastic differntial equation. The Black-Scholes equation is one of the most well known equations in Financial Mathematics. The solution to this equation, the Black-Scholes formula, calculates the price of European call options. European call options give the buyer the right to buy stock at time T for a fixed price K, which is often called the strike price of the option. We also must note that the buyer is not obligated to buy the stock at time T, he can also decide to not buy it at all.

Section 1 will introduce two important concepts in the probability theory, namely expectation and conditional expectation. An exposure to Measure theory is recommended while reading this section as some concepts will be assumed to be known (Such as the $\mathcal{L}^2$-space and the Lebesgue integral). For section 1 we used Chapter 1-2 from [1] as a guideline while Chapters 1-4 in [2] are referred to for more background information and a more abstract approach. In Section 2 we treat the concept of Martingales and Brownian motion, which will be used frequently in later Sections. Guidelines for this section were Chapter 3 from [1] and Chapters 6-7 from [5], while the introduction for Brownian motion was partly inspired by [8]. In Section 3 we talk about the Itô Integral where we use the Brownian motion to introduce the stochastic integral and a whole different form of Calculus, namely Stochastic Calculus. Chapter 10 from [5] and Chapter 4 [1] were used as guidelines for this section. Section 4 introduces the Black-Scholes equation and derives the solution to this equation making use of Martingales to move to a risk-free measure. The ideas proposed in this section are based upon Chapter 4 in [1]. Section 5 is the last section of this thesis. We talk about how to estimate the volatility of the model, i.e. Geometric Brownian motion, used to determine the stock prices in the Black-Scholes Equation and Formula. In this section [3],[4] and [6] were consulted to derive a proper estimation.

1 Expectation and Conditional Expectation

In this chapter we will cover the basis of expectation of a so called *random variable*. We will use this chapter as the setup for the concepts introduced in later chapters. We start with the concept of expectation in general measure spaces.

1.1 Expectation

The average value of a random variable is very important while we also would like to account for the probabilities that one value might occur. We do this by defining the expectation in this section.

Definition 1.1.1 (Random Variable and Probability Measure)

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an arbitrary measurable function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a *random variable*. Now we define the *probability measure* $\mu_X : \mathcal{B}(\mathbb{R}) \to [0, 1]$ by

$$\mu_X(A) = \mathbb{P}\left(X^{-1}(A)\right),$$

where A is a Borel measurable set of $\mathbb{R}$. Often μ_X is also called the distribution measure of X. Note that this is equivalent to

$$\mu_X(A) = \mathbb{P}\{\omega \in \Omega; X(\omega) \in A\} = \mathbb{P}\{X \in A\}.$$

Remark 1.1.1:

We claim that μ_X is indeed a probability measure, we can check this by checking the definition of a (probability) measure. The definition of a measure is as follows, let $\mathcal{F}$ be a sigma algebra then $\mathbb{P} : \mathcal{F} \to [0, 1]$ is called a probability measure when

(i) $\mathbb{P}(\emptyset) = 0$

(ii) Let $\mathcal{A}$ be a countable collection where all the $A_i \in \mathcal{A}$ are disjoint then there must hold that

$$\mathbb{P}(\mathcal{A}) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

It is easy to see that (i) holds for μ_X, for (ii) consider a countable collection $\mathcal{A}$ where all the $A_i \in \mathcal{A}$ are disjoint we have that

$$\mu_X(\mathcal{A}) = \mathbb{P}\left(X^{-1}(\mathcal{A})\right) = \mathbb{P}\left(X^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} X^{-1}(A_i)\right) = \sum_{i=1}^{\infty} \mathbb{P}\left(X^{-1}(A_i)\right) = \sum_{i=1}^{\infty} \mu_X(A_i)$$

Definition 1.1.2 (Probability Density Function)

Now if there exists a measurable function $f_X \geq 0$ on $\mathbb{R}$ such that

$$\int_A f_X(x) d\lambda(x) = \mu_X(A),$$

where λ is the Lebesgue measure, then we call f_X the *probability densitity function* of the random variable X. Note that when f_X is Riemann-integrable this is equal to

$$\int_A f_X(x) dx = \mu_X(A).$$

Definition 1.1.3 (Expectation)

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable function $X \in \mathcal{L}^1$, we define the *expectation* of X as

$$\mathbb{E}[X] := \int_\Omega X d\mathbb{P}.$$

The expectation of a random variable X can be seen as an average value where we account for different probabilites of $\omega \in \Omega$. We require $X \in \mathcal{L}^1$ because it implies that $\mathbb{E}[|X|] < \infty$.

Remark 1.1.2

Suppose we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider the expectation of $X \in \mathcal{L}^1$ as defined in Definition 1.1.3. Now consider that we have a different measure on this measurable space $(\Omega, \mathcal{F})$, say $\widetilde{\mathbb{P}}$, then we can calculate the expectation under this measure $\widetilde{\mathbb{P}}$ which will be denoted by

$$\widetilde{\mathbb{E}}[X] = \int_\Omega X d\widetilde{\mathbb{P}}.$$

1.2 Conditional Expectation

From an elementary probabality course one may know that for two given events A and B with $\mathbb{P}(B) > 0$, we have that the conditional probability of A given B is defined by

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In a similar fashion we can define the conditional expectation. Consider first a specific situation. Given a random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$. Then we define the *conditional expectation of X given A* by

$$\mathbb{E}[X|A] := \frac{\int_A X d\mathbb{P}}{\mathbb{P}(A)} = \frac{\mathbb{E}[\mathbb{I}_A X]}{\mathbb{P}(A)}.$$

We can extend this notion of conditional expectation. Consider a sample space Ω for which we can write

$$\Omega = \bigcup_{i=1}^{\infty} A_i.$$

where A_i are disjoint for all i and we have that $\mathbb{P}(A_i) > 0$. We define $\mathcal{G} := \sigma(A_1, A_2, \ldots)$ as the σ-algebra generated by the A_i. We then define

$$\mathbb{E}[X|\mathcal{G}](\omega) = \mathbb{E}[X|A_i], \qquad \omega \in A_i$$

and thus follows

$$\mathbb{E}[X|\mathcal{G}] = \sum_{i=1}^{\infty} \mathbb{E}[X|A_i]\mathbb{I}_{A_i}.$$

Note that we have that $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measureable, i.e. the preimage of every Borel-measurable set is in $\mathcal{F}$. Now we will give a more general definition of the conditional expectation.

Definition 1.2.1 (Conditional Expectation)

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub-σ-algebra on $\mathcal{F}$. If $X \in \mathcal{L}^1$ is a random variable and if we have a random varibale Y which is $\mathcal{G}$-measurable such that

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P},$$

for every $A \in \mathcal{G}$. We say that Y is the *conditional expectation* given a σ-algebra of X and we write $Y = \mathbb{E}[X|\mathcal{G}]$. The Y in Definition 1.2.4 always exists and is unique almost surely. Hence we prove the following theorem

Theorem 1.2.1 (Existence and Uniqueness of Conditional Expectation)

The conditional expectation defined in Defenition 1.2.1, always exists and is unique almost everywhere.

Proof
Consider a random variable X. We assume that $X \in \mathcal{L}^1$, hence it meets the condition in Definition 1.2.4. We denote $X^+ = \max\{X, 0\}$ and $X^- = \max\{-X, 0\}$. Then we know that $X = X^+ - X^-$ and we have written X as a sum of postive elements. Consider a sub-σ-algebra $\mathcal{G}$, we will define $\mathbb{Q}^+ : \mathcal{G} \to [0, \infty)$ as

$$\mathbb{Q}^+(A) := \int_A X^+ d\mathbb{P}.$$

By the Theorem A.1 presented in Appendix (A.1) we know that $\mathbb{Q}^+(A)$ is a measure. Because $\mathbb{Q}^+$ is absolutely continous with respect to $\mathbb{P}|_{\mathcal{G}}$ we know by the Radon-Nikodym Theorem, see Theorem 4.2.2 in [1], that there must exist a $\mathcal{G}$-measurable function Y_1 such that

$$\mathbb{Q}^+(A) = \int_A Y_1 d\mathbb{P},$$

where A is arbitrary in $\mathcal{G}$. Similary since X^- is postive, there exists $\mathcal{G}$-measurable function Y_2 such that

$$\mathbb{Q}^-(A) = \int_A Y_2 d\mathbb{P}.$$

Now if we write that $Y := Y_1 - Y_2$, then Y is a $\mathcal{G}$-measurable function and for any $A \in \mathcal{G}$ we have

$$\begin{aligned}\int_A X d\mathbb{P} &= \int_A X^+ - X^- d\mathbb{P} \\ &= \int_A X^+ d\mathbb{P} - \int_A X^- d\mathbb{P} \\ &= \int_A Y_1 d\mathbb{P} - \int_A Y_2 d\mathbb{P} \\ &= \int_A (Y_1 - Y_2)\, d\mathbb{P} = \int_A Y d\mathbb{P}\end{aligned}.$$

Thus we have proven the existence of such a Y. Suppose we have Y and Z both $\mathcal{G}$-measurable such that

$$\int_A X d\mathbb{P} = \int_A Y d\mathbb{P} = \int_A Z d\mathbb{P}.$$

Then we know that

$$\int_A (Y - Z) d\mathbb{P} = 0,$$

for any $A \in \mathcal{G}$. Hence $Y = Z$ a.e. Thus we have that the conditional expectation always exists and is unique a.e.

Theorem 1.2.2 (Properties of the Conditional Expectation)

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ then the following hold

1. $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[X|\mathcal{G}], \quad \text{with } a, b \in \mathbb{R}$
2. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$
3. If X is $\mathcal{G}$-measurable and XY is integrable, we have that $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$
4. Let $\mathcal{H} \subset \mathcal{G}$ then we have that $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$

Proof

1. We know that $\forall A \in \mathcal{G}$ there holds

$$\begin{aligned} \int_A \mathbb{E}[aX + bY|\mathcal{G}] d\mathbb{P} &= \int_A aX + bY d\mathbb{P} \\ &= a \int_A X d\mathbb{P} + b \int_A Y d\mathbb{P} \\ &= a \int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} + b \int_A \mathbb{E}[Y|\mathcal{G}] d\mathbb{P} \\ &= \int_A a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}] d\mathbb{P} \end{aligned}.$$

   Thus we see that $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[X|\mathcal{G}]$.

2. If we take $A = \Omega$ we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \int_\Omega \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_\Omega X d\mathbb{P} = \mathbb{E}[X].$$

   Thus there follows that $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.

3. If we have that $X = \mathbb{I}_A$ with $A \in \mathcal{G}$ then for $B \in \mathcal{G}$ we have

$$\begin{aligned} \int_B \mathbb{E}[\mathbb{I}_A Y|\mathcal{G}] &= \int_B \mathbb{I}_A Y d\mathbb{P} \\ &= \int_{A \cap B} Y d\mathbb{P} \\ &= \int_{A \cap B} \mathbb{E}[Y|\mathcal{G}] d\mathbb{P} \\ &= \int_B \mathbb{I}_A \mathbb{E}[Y|\mathcal{G}] d\mathbb{P} \end{aligned}.$$

   Thus the result hold whenever X is a simple function, we know that we can write a non-negative measurable function X as a convergent sequence of increasing simple functions, then with the Montone Convergence Theorem one can conclude this holds for all non-negative measurable X. To show that this hold for all measurable X consider $X = X^+ - X^-$, where we define X^+ and X^- in the same way as the proof of Theorem 1.2.1. Since both X^+ and X^- are non-negative $\mathcal{G}$-measurable functions then we can conclude that this holds for all $\mathcal{G}$-measurable functions.

4. For $A \in \mathcal{H} \subset \mathcal{G}$ we have that

$$\int_A \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{H}] d\mathbb{P}.$$

   Thus we see that indeed $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.

   Note that for more properties one could view Theorem 2.3.2 in [1].

Theorem 1.2.3

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $Z \geq 0$ a.s. and $\mathbb{E}[Z] = 1$. We define $\widetilde{\mathbb{P}}(A) = \int_A Zd\mathbb{P}$. Suppose X is measurable function then

$$\widetilde{\mathbb{E}}[X] = \mathbb{E}[XZ].$$

Proof
First of all note that $\widetilde{\mathbb{P}}$ is indeed a measure, see Appendix A.1. Now suppose $X = \mathbb{I}_A$ then

$$\widetilde{\mathbb{E}}[X] = \widetilde{\mathbb{E}}\left[\mathbb{I}_A\right] = \int_\Omega \mathbb{I}_A ZdP = \mathbb{E}\left[\mathbb{I}_A Z\right] = \mathbb{E}\left[XZ\right].$$

Then we can show that this holds for simple functions, i.e. a sum of indicator functions. This works because of the linearity of the expectation. Then in a similar way as for the proof of Theorem 1.2.2(3) we can show this for arbitrary measurable functions. Note that we cannot have that

$$\mathbb{E}[X^+Z] = \mathbb{E}[X^-Z] = \infty.$$

This is because then

$$\mathbb{E}[XZ] = \mathbb{E}\left[\left(X^+ - X^-\right)Z\right] = \mathbb{E}[X^+Z] - \mathbb{E}[X^-Z] = \infty - \infty,$$

hence we must have that one of the two expectations is finite. Note that if $X \in \mathcal{L}^1(\widetilde{\mathbb{P}})$ this problem wil not occur. So Theorem 1.2.3 holds for all $X \in \mathcal{L}^1(\widetilde{\mathbb{P}})$

2 Martingales and Brownian Motion

In this chapter we will talk about the Martingale property of an adpated process. We will also introduce the concept of Brownian motion. We will see that given the filtration $\{\mathcal{F}(t) : t \geq 0\}$ that a Brownian motion is a Martingale. This is an important result as it will be used often in later chapters.

2.1 Stochastic Processes

We would like to make sense of two important classes of stochastic processes, namely Martingales and Brownian Motions. Since these are stochastic processes we should first define what we mean with a stochastic process.

Definition 2.1.1 (Filtration)

Let Ω be a non empty set. Let T be a fixed positive number, which mostly denotes time in the thesis. We assume that for every $t \in [0, T]$ there is a $\sigma-$algebra $\mathcal{F}(t)$. If $\forall s \leq t$ we have that $\mathcal{F}(s) \subset \mathcal{F}(t)$, we call the collection $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, a *filtration*.

Definition 2.1.2 (Random Variable Adapted to a Filtration)

A stochastic process is a sequence of random variables $X(t) : \Omega \to \mathbb{R}$ parametrized by $t \in [0, T]$. The stochastic process $X : [0, T] \times \Omega \to \mathbb{R}$ is a measurable mapping. In this thesis the we will most likely talk about stochastic processes indexed by an interval $[0, T]$. Since this interval is uncountable we say that in this case we have a continuous stochastic variable. If we were to have a stochastic process indexed by a countable set, we say that the process is discrete. The filtration $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, generated by a process with $\mathcal{F}(t) = \sigma(\{X(s) : 0 \leq s \leq t\})$ is called a natural filtration for $\{X(t) : 0 \leq t \leq T\}$. We say that $\{X(t) : 0 \leq t \leq T\}$ is *adapted to the filtration* $\{\mathcal{F}(t)\colon 0 \leq t \leq T\}$, if $X(t)$ is measurable with respect to $\mathcal{F}(t)$ for all t.

2.2 Martingales

In essence a Martingale, is a process for which based on what happened until time s, the expected change in the future is 0. Hence our best guess for what happens at time t where $t > s$ is just to expect it having the same value as at time s. Now remember when we say that $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ is the natural filtration of a process $\{X(t) : t \geq 0\}$. It can be thought of as saying that $\mathcal{F}(t)$ contains all the information up until t. Now we will define the Martingale property of a stochastic process.

Definition 2.2.1 (Martingale)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose we have a fixed number $T \geq 0$. Let $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted process $\{X(t) : 0 \leq t \leq T\}$. Then if for any $0 \leq s \leq t \leq T$ we have that

$$\mathbb{E}[X(t)|\mathcal{F}(s)] = X(s),$$

$\{X(t),\, 0 \leq t \leq T\}$, is called a *Martingale*. On the otherhand if

$$\mathbb{E}[X(t)|\mathcal{F}(s)] \geq X(s),$$

for $0 \leq s \leq t \leq T$ then $\{X(t),\, 0 \leq t \leq T\}$, is called a *submartingale*, if we have that

$$\mathbb{E}[X(t)|\mathcal{F}(s)] \leq X(s),$$

for $0 \leq s \leq t \leq T$ we say that $\{X(t),\, 0 \leq t \leq T\}$ is called a *supermartingale*.

Theorem 2.2.1

If $\{X(t) : 0 \leq t \leq T\}$, is a martingale with respect to the filtration $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ for which $\mathcal{F}(0) = \{\emptyset, \Omega\}$, then $\mathbb{E}[X(t)] = \mathbb{E}[X(0)]$ for all $0 \leq t \leq T$.

Proof
We have that for $0 \leq s \leq t$

$$\mathbb{E}[X(t)] = \mathbb{E}[\mathbb{E}[X(t)|\mathcal{F}(s)]] = \mathbb{E}[X(s)].$$

If we let $s = 0$ we get the desired result. This means that the expectations of all $X(t)$ are equal.

Definition 2.2.2 (Radon-Nikodým Derivative Process)

Consider $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{F}(t)$ a filtration for $0 \leq t \leq T$, let $Z \geq 0$ almost surely and let $\mathbb{E}[Z] = 1$, we define the measure

$$\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P}.$$

Since $\mathbb{P}(Z > 0) = 1$, Z is called the Radon-Nikodým derivative of $\widetilde{\mathbb{P}}$ and

$$Z(t) = \mathbb{E}[Z|\mathcal{F}(t)] \quad 0 \leq t \leq T,$$

is called a *Radon-Nikodým derivative process.*

Theorem 2.2.2

Suppose we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ be filtration. Let $Z(t)$ be a Radon-Nikodým derivative process as defined in definition 2.2.2. Define $\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P}$ and let Y be a $\mathcal{F}(t)$-measurable function.

Then $\{Z(t) : 0 \leq t \leq T\}$ is a martingale and

$$\widetilde{\mathbb{E}}[Y] = \mathbb{E}\left[YZ(t)\right], \quad \widetilde{\mathbb{E}}\left[Y|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right].$$

Proof
Let $0 \leq s \leq t \leq T$, then

$$\mathbb{E}\left[Z(t)|\mathcal{F}(s)\right] = \mathbb{E}\left[\mathbb{E}\left[Z|\mathcal{F}(t)\right]|\mathcal{F}(s)\right] = \mathbb{E}\left[Z|\mathcal{F}(s)\right] = Z(s),$$

hence $Z(t)$ is a martingale. By Theorem 1.2.2 and since Y is $\mathcal{F}(t)$-measurable

$$\widetilde{\mathbb{E}}[Y] = \mathbb{E}\left[YZ\right] = \mathbb{E}\left[\mathbb{E}\left[YZ|\mathcal{F}(t)\right]\right] = \mathbb{E}\left[Y\mathbb{E}\left[Z|\mathcal{F}(t)\right]\right] = \mathbb{E}\left[YZ(t)\right].$$

It is clear that $\frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right]$ is $\mathcal{F}(s)$-measurable hence we show that

$$\int_A \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right] d\widetilde{\mathbb{P}} = \int_A Y d\widetilde{\mathbb{P}}, \quad \forall A \in \mathcal{F}.$$

Thus with what we just proved

$$\begin{aligned}\int_A \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right] d\widetilde{\mathbb{P}} &= \int_\Omega \mathbb{I}_A \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right] d\widetilde{\mathbb{P}} \\ &= \widetilde{\mathbb{E}}\left[\mathbb{I}_A \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right]\right] \\ &= \mathbb{E}\left[\mathbb{I}_A \mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right]\right] \\ &= \mathbb{E}\left[\mathbb{I}_A YZ(t)\right] \\ &= \widetilde{\mathbb{E}}\left[\mathbb{I}_A Y\right] \\ &= \int_A Y d\widetilde{\mathbb{P}}\end{aligned},$$

hence

$$\widetilde{\mathbb{E}}\left[Y|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\mathbb{E}\left[YZ(t)|\mathcal{F}(s)\right].$$

2.3 Brownian Motion

Biologist Robert Brown discovered Brownian motion, hence the name, back in 1827. Brown was not able to fully explain what the origin of Brownian motion was. Then Louis Bachelier provided a theory of Brownian motion in his PhD thesis in the year 1900. While Albert Einstein, independently provided a probablisitc model that adequatly explained Brownian motion. Later Norbert Wiener had a great contribution in the field of Brownian motion, hence Brownian motion is sometimes called a Wiener process.
Brownian motion are often used in Financial Mathematics. One could assume that, for example, stock prices are random walks. A Brownian motion is essentially a random walk where the change in value is unrelated to future or past changes, for more information about the ideas behind Brownian motion see Chapter 3 Section 1 and 2 of [1]. Hence this is why Brownian motion can be used in modelling the financial markets. Brownian motion have useful mathematical properties, for example they are easy to do calculations with and with the right filtration they are Martingales, hence we can apply the theory we derived for Martingales. For this reason it is used often when dealing with processes from an unkown origin, i.e. the Stock market. Later on we assume that the stock price is modelled by a geometric Brownian motion. Hence the concept of Brownian motion is explained.

Defintion 2.3.1 (Brownian Motion)

A stochastic process $\{W(t) : t \geq 0\}$ is called a *Brownian motion* if it satisfies the following properties:

(1) $W(0) = 0$ and $t \mapsto W(t)$, $t \geq 0$, is continious with probability 1.

(2) For $0 \leq s \leq t$ we have that $W(t) - W(s) \sim \mathcal{N}(0, t-s)$, i.e. Normally distributed with mean 0 and variance $t - s$.

(3) $\{W(t) : t \geq 0\}$ has stationary and independent increments.

Remark 2.3.1

Consider the Brownian motion $\{W(t) : t \geq 0\}$. Let $t \geq s$, since $\mathbb{E}\left[W(t) - W(s)\right] = 0$, then

$$\begin{aligned}\mathbb{E}\left[(W(t) - W(s))^2\right] &= \mathrm{Var}\left(W(t) - W(s)\right) + \mathbb{E}\left[W(t) - W(s)\right]^2 \\ &= t - s\end{aligned}.$$

Definition 2.3.2 (Filtration for the Brownian Motion)

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{W(t) : t \geq 0\}$ be a Brownian motion defined on this space. A *filtration for the Brownian motion* $\{\mathcal{F}(t) : t \geq 0\}$ satisfies the following:

(1) For $0 \leq s < t$, we have that $\mathcal{F}(s) \subset \mathcal{F}(t)$.

(2) For $t \geq 0$, we have that $W(t)$ is $\mathcal{F}(t)$-measurable.

(3) For $0 \leq u < t$, we have that $W(t) - W(u)$ is independent of $\mathcal{F}(u)$.

When we talk about a Brownian motion $\{W(t) : t \geq 0\}$ and a filtration $\{\mathcal{F}(t) : t \geq 0\}$, we suppose that the properties in Definition 2.3.2 hold unless otherwise specified.

Theorem 2.3.1

Let $\{W(t) : t \geq 0\}$ be a Brownian motion with the filtration $\{\mathcal{F}(t) : t \geq 0\}$ satisfying the properties of Definition 2.3.2. Then $\{W(t) : t \geq 0\}$ is a Martingale with respect to the filtration $\{\mathcal{F}(t) : t \geq 0\}$.

Proof Let $W(t)$ be a Brownian motion adapted to the filtration $\{\mathcal{F}(t) : t \geq 0\}$. Then we have for $0 \leq s \leq t$ that

$$\begin{aligned}\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[W(t) + W(s) - W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] + \mathbb{E}[W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)] + W(s) \\ &= W(s)\end{aligned}.$$

Example 2.3.1

Let $\{W(t) : t \geq 0\}$, be a Brownian motion, and let $\{\mathcal{F}(t) : t \geq 0\}$ be a filtration for this Brownian motion. Let $a \in \mathbb{R}$ then

$$X(t) = e^{aW(t) - \frac{1}{2}a^2 t},$$

is a Martingale.

We will show this by using a standard method that is used a lot when dealing with showing that functions of Brownian motions are Martingales. Suppose $0 \leq s \leq t$. Define $Y(t) = aW(t) - \frac{1}{2}a^2 t$, then we can write

$$\begin{aligned}X(t) &= e^{Y(t)} \\ &= e^{Y(t) - Y(s) + Y(s)} \\ &= e^{Y(t) - Y(s)} e^{Y(s)} \\ &= X(s) e^{a(W(t) - W(s)) - \frac{1}{2}a^2(t-s)}\end{aligned}.$$

Then since $X(s)$ is $\mathcal{F}(s)$-measurable and $W(t) - W(s)$ is independent of $\mathcal{F}(s)$

$$\begin{aligned}\mathbb{E}\left[X(t)|\mathcal{F}(s)\right] &= \mathbb{E}\left[X(s) e^{a(W(t)-W(s)) - \frac{1}{2}a^2(t-s)} \middle| \mathcal{F}(s)\right] \\ &= X(s)\mathbb{E}\left[e^{a(W(t)-W(s)) - \frac{1}{2}a^2(t-s)}\right] \\ &= X(s) e^{-\frac{1}{2}a^2(t-s)} \mathbb{E}\left[e^{a(W(t)-W(s))}\right] \\ &= X(s) e^{-\frac{1}{2}a^2(t-s)} \mathbb{E}\left[e^{a\sqrt{t-s}Z}\right]\end{aligned},$$

where $Z = \dfrac{W(t) - W(s)}{\sqrt{(t-s)}} \sim \mathcal{N}(0,1)$, since $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ per definition. By calculating the Moment-generating function of a standard normal random variable one can show that $\mathbb{E}\left[e^{tX}\right] = e^{\frac{1}{2}t^2}$ when $X \sim \mathcal{N}(0,1)$. But we are assuming that this is already known. Hence

$$\mathbb{E}\left[e^{a\sqrt{t-s}Z}\right] = e^{\frac{1}{2}a^2(t-s)}.$$

Thus we conclude

$$X(s)e^{-\frac{1}{2}a^2(t-s)}\mathbb{E}\left[e^{a\sqrt{t-s}Z}\right] = X(s),$$

and we now have shown that $X(t)$ is a Martingale.

Definition 2.3.3 (Quadratic Variation of the Brownian Motion)

Let $\{W(t) : t \geq 0\}$ a Brownian motion, let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0,T]$ such that $0 < t_1 < \cdots < t_n = T$. Consider the mesh of the partition $\|\Pi\| := \max_{0 \leq i \leq (n-1)} |t_{i+1} - t_i|$, then *the quadratic variation of the Brownian motion* is defined by

$$[W,W](T) := \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} |W(t_{i+1}) - W(t_i)|^2,$$

with convergence in the $\mathcal{L}^2$-sense, where $\mathcal{L}^2$ is induced with the $\mathcal{L}^2$-norm.

Theorem 2.3.2

Let $\{W(t) : t \geq 0\}$ a Brownian motion, let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0,T]$ such that $0 < t_1 < \cdots < t_n = T$. Then the Quadratic Variation of the Brownian motion is given by and equal to

$$[W,W](T) := \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} |W(t_{i+1}) - W(t_i)|^2 = T,$$

with convergence in the $\mathcal{L}^2$ sense. Where $\mathcal{L}^2$ is equipped with the semi-norm

$$\|X\|_2 := \sqrt{\mathbb{E}[X^2]}.$$

We will proof this after we have proven the following 2 lemmas.

Lemma 2.3.1

Let $Z \sim \mathcal{N}(0,1)$ then

$$\mathbb{E}[Z^4] = 3.$$

Proof
The moment generating function of $Z \sim \mathcal{N}(0,1)$ is given by

$$M_Z(t) = e^{\frac{1}{2}t^2}.$$

Therefore

$$\mathbb{E}[Z^4] = \left[\frac{d^4}{dt^4} e^{\frac{1}{2}t^2}\right]_{t=0}.$$

It can be show by elementary calculations that

$$\frac{d^4}{dt^4} e^{\frac{1}{2}t^2} = (3 + 3t^2)e^{\frac{1}{2}t^2} + (3t^2 + t^4)e^{\frac{1}{2}t^2}.$$

Thus

$$\mathbb{E}[Z^4] = 3.$$

Lemma 2.3.2

Consider the random variables $(I_i)_{i=1}^n$ where we define $I_i := (W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)$. Then

$$\mathbb{E}\left[I_i^2\right] = 2(t_{i+1} - t_i)^2.$$

Proof
Define $W_i := W(t_{i+1}) - W(t_i)$. Then we find that

$$\begin{aligned}\mathbb{E}\left[I_i^2\right] &= \mathbb{E}\left[(W_i^2 - (t_{i+1} - t_i))^2\right] \\ &= \mathbb{E}\left[W_i^4 - 2W_i^2(t_{i+1} - t_i) + (t_{i+1} - t_i)^2\right] \\ &= \mathbb{E}\left[W_i^4\right] - 2(t_{i+1} - t_i)\mathbb{E}\left[W_i^2\right] + (t_{i+1} - t_i)^2\end{aligned}.$$

By definiton of the Brownian motion we know that $W_i \sim \mathcal{N}(0, t_{i+1} - t_i)$. Let $Z \sim \mathcal{N}(0,1)$, then we know that $\sqrt{t_{i+1} - t_i}Z \sim \mathcal{N}(0, t_{i+1} - t_i)$. Using this, $\mathbb{E}[Z^2] = 1$ and Lemma 2.3.1 we conclude:

$$\mathbb{E}[I_i^2] = 3(t_{i+1} - t_i)^2 - 2(t_{i+1} - t_i)^2 + (t_{i+1} - t_i)^2 = 2(t_{i+1} - t_i)^2.$$

Proof of Theorem 2.3.2
Define $M(\Pi) := \sum_{i=0}^{n-1} \left|W_{t_{i+1}} - W_{t_i}\right|^2$ and note that $\sum_{i=0}^{n-1} I_i = M(\Pi) - T$, thus

$$\left\|\sum_{i=0}^{n-1} I_i\right\|_2 = \sqrt{\mathbb{E}\left[\left(\sum_{i=0}^{n-1} I_i\right)^2\right]} = \sqrt{\mathbb{E}\left[\sum_{i=0}^{n-1} I_i^2 + \sum_{i=0}^{n-1}\sum_{j\neq i} I_i I_j\right]}. \tag{1}$$

We know

$$\mathbb{E}\left[\sum_{i=0}^{n-1} I_i^2 + \sum_{i=0}^{n-1}\sum_{j\neq i} I_i I_j\right] = \sum_{i=0}^{n-1}\mathbb{E}\left[I_i^2\right] + \sum_{i=0}^{n-1}\sum_{j\neq i}\mathbb{E}[I_i I_j]. \tag{2}$$

Then follows by using the same notation for W_i as in Lemma 2.3.2

$$\begin{aligned}\mathbb{E}[I_i I_j] &= \mathbb{E}\left[(W_i^2 - (t_{i+1} - t_i))(W_j^2 - (t_{j+1} - t_j))\right] \\ &= \mathbb{E}[W_i^2 W_j^2 - W_i^2(t_{j+1} - t_j) - W_j^2(t_{i+1} - t_i) + (t_{i+1} - t_i)(t_{j+1} - t_j)]\end{aligned}.$$

Thus

$$\mathbb{E}[I_i I_j] = \mathbb{E}[W_i^2 W_j^2] - (t_{j+1} - t_j)\mathbb{E}[W_i^2] - (t_{i+1} - t_i)\mathbb{E}[W_j^2] + (t_{i+1} - t_i)(t_{j+1} - t_j).$$

By the fact that the t_i, $0 \leq i \leq n$, form a partition we know that $W(t_{i+1}) - W(t_i)$ is independent of $W(t_{j+1}) - W(t_j) \quad \forall j \neq i$ (The independent increments property of the Brownian motion). Thus W_i is independent of $W_j \quad \forall j \neq i$, and so follows W_i^2 is independent of $W_j^2 \quad \forall i \neq j$. Then we know by Remark 2.3.1 that $\mathbb{E}[W_i^2] = (t_{i+1} - t_i)$. Thus follows

$$\mathbb{E}[I_i I_j] = \mathbb{E}[W_i^2]\mathbb{E}[W_j^2] - (t_{j+1} - t_j)\mathbb{E}[W_i^2] - (t_{i+1} - t_i)\mathbb{E}[W_j^2] + (t_{i+1} - t_i)(t_{j+1} - t_j)$$

Thus

$$\mathbb{E}[I_i I_j] = (t_{i+1} - t_i)(t_{j+1} - t_j) - (t_{j+1} - t_j)(t_{i+1} - t_i) - (t_{i+1} - t_i)(t_{j+1} - t_j) + (t_{i+1} - t_i)(t_{j+1} - t_j) = 0.$$

Note that this only holds for all $i \neq j$. But this means we can use it in (2). So plugging in gives us

$$\mathbb{E}\left[\sum_{i=0}^{n-1} I_i^2 + \sum_{i=0}^{n-1}\sum_{j\neq i} I_i I_j\right] = \sum_{i=0}^{n-1}\mathbb{E}\left[I_i^2\right].$$

If we use this information in (1) we find

$$\left\|\sum_{i=0}^{n-1} I_i\right\|_2 = \sqrt{\sum_{i=0}^{n-1} \mathbb{E}[I_i^2]}.$$

If we now use Lemma 2.3.2 we find

$$\left\|\sum_{i=0}^{n-1} I_i\right\|_2 = \sqrt{\sum_{i=0}^{n-1} 2(t_{i+1}-t_i)^2} \leq \sqrt{2\|\Pi\| \sum_{i=0}^{n-1} t_{i+1}-t_i} = \sqrt{2\|\Pi\|T}. \tag{3}$$

If we now want to show that $M(\Pi)$ converges to T we want to show that

$$\lim_{||\Pi||\to 0} \|M(\Pi)-T\|_2 = \lim_{\|\Pi\|\to 0} \left\|\sum_{i=0}^{n-1} I_i\right\|_2 = 0.$$

By (3)

$$0 \leq \lim_{\|\Pi\|\to 0} \|M(\Pi)-T\|_2 = \lim_{\|\Pi\|\to 0} \left\|\sum_{i=0}^{n-1} I_i\right\|_2 \leq \lim_{\|\Pi\|\to 0} \sqrt{2\|\Pi\|T} = 0.$$

Then by the Squeeze Theorem there must hold that

$$\lim_{\|\Pi||\to 0} \|M(\Pi)-T\|_2 = 0,$$

and thus we have shown that

$$\lim_{\|\Pi\|\to 0} M(\Pi) = T,$$

in the $\mathcal{L}^2$-sense.

Informally we will write

$$dW(t)dW(t) = dt.$$

Theorem 2.3.3 (Lévy's Characterization of Brownian Motion)

Let $\{W(t) : t \geq 0\}$, be a martingale relative to a filtration $\{\mathcal{F}(t) : t \geq 0\}$. If $W(0) = 0$, $W(t)$ has continous paths and $[W, W](t) = t$, $\forall t \geq 0$. Then $W(t)$ is a Brownian motion.

Proof See Theorem 4.6.4 in [2].

Definiton 2.3.4 (Brownian Motion with Drift)

Let $\{W(t) : t \geq 0\}$ denote a $\mathbb{P}$-Brownian motion, i.e. $\{W(t) : t \geq 0\}$ is a Brownian motion with respect to the measure $\mathbb{P}$. Let θ be a given constant then we define

$$X(t) := W(t) + \theta,$$

we call $\{X(t) : t \geq 0\}$ a *Brownian motion with drift.*

Definition 2.3.5 (Geometric Brownian Motion)

Let $\alpha > 0$ and σ be constant, let $W(t)$ be a Brownian motion. We define the *geometric Brownian motion* by

$$S(t) = S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t}.$$

Note that this is closely related to the stochastic process shown in Example 2.3.1.

3 Itô Integral

In this section we will define the Itô integral, i.e. an integral with respect to a Brownian motion. First we will define the integral for a simple adapted stochastic process $\Delta(t)$, and prove some useful properties of the Itô integral defined this way. Then we will extend this notion of integral for general stochastic processes.

3.1 Itô integral for Simple Processes

Suppose we have a Brownian motion $\{W(t) : t \geq 0\}$ with a filtration $\{\mathcal{F}(t) : t \geq 0\}$, we want to make sense of the integral of an adapted process where we integrate with respect to the Brownian motion $\{W(t) : t \geq 0\}$, which is called the Itô integral. Hence we first define the Itô integral for the more simpler case, the simple processes.

Definition 3.1.1 (Simple Process)

Let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0, T]$. Assume that $\{\Delta(t) : 0 \leq t \leq T\}$ is constant in t on the interval $[t_i, t_{i+1})$, for all i. We call $\{\Delta(t) : 0 \leq t \leq T\}$ a *simple process*.

Definition 3.1.2 (Itô Integral for Simple Processes)

Let $\{W(t) : t \geq 0\}$ be a Brownian motion with the filtration $\{\mathcal{F}(t) : t \geq 0\}$ and let $\{\Delta(t) : 0 \leq t \leq T\}$ be a simple adapted process. Then if $t_k \leq t \leq t_{k+1}$ define the *Itô Integral* by

$$I(t) = \sum_{i=0}^{k-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right] + \Delta(t_k) \left[W(t) - W(t_k)\right] = \int_0^t \Delta(t) dW(t).$$

Theorem 3.1.1

The Itô Integral for simple processes is a Martingale with respect to the filtration $\{\mathcal{F}(t) : t \geq 0\}$.

Proof
Let $0 \leq s \leq t \leq T$, we assume that s and t are in different intervals of Π. Thus $\exists l \leq k$ such that $s \in [t_l, t_{l+1})$ and $t \in [t_k, t_{k+1}]$. We must show that

$$\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s).$$

Note that

$$I(t) = \sum_{i=0}^{l-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right] + \Delta(t_l) \left[W(t_{l+1}) - W(t_l)\right] + \sum_{i=l+1}^{k-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right] + \Delta(t_k) \left[W(t) - W(t_k)\right].$$

By the linearity of the expectation we consider the conditional expectation of each term. So for the first term

$$\mathbb{E}\left[\sum_{i=0}^{l-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right] \middle| \mathcal{F}(s)\right].$$

Because $t_l \leq s$ we have that $\sum_{i=0}^{l-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right]$ is $\mathcal{F}(s)$-measurable then by Theorem 1.2.2(3) we have that

$$\mathbb{E}\left[\sum_{i=0}^{l-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right] \middle| \mathcal{F}(s)\right] = \sum_{i=0}^{l-1} \Delta(t_i) \left[W(t_{i+1}) - W(t_i)\right].$$

For the second term

$$\mathbb{E}\left[\Delta(t_l) \left[W(t_{l+1}) - W(t_l)\right] \middle| \mathcal{F}(s)\right].$$

Because again $t_l \leq s$ and because of Theorem 2.3.1 we have that

$$\begin{aligned}\mathbb{E}\left[\Delta(t_l)\left[W(t_{l+1}) - W(t_l)\right]\middle|\mathcal{F}(s)\right] &= \Delta(t_l)\mathbb{E}\left[W(t_{l+1}) - W(t_l)\middle|\mathcal{F}(s)\right] \\ &= \Delta(t_l)\left(\mathbb{E}\left[W(t_{l+1})\middle|\mathcal{F}(s)\right] - W(t_l)\right). \\ &= \Delta(t_l)\left(W(t_s) - W(t_l)\right)\end{aligned}$$

We see that the conditional expectation of the first two terms combined gives us

$$\sum_{i=0}^{l-1}\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right] + \Delta(t_l)\left(W(t_s) - W(t_l)\right) = I(s).$$

For the third term

$$\mathbb{E}\left[\sum_{i=l+1}^{k-1}\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right]\middle|\mathcal{F}(s)\right].$$

Because $t_i \geq s$ for $i \in [l+1, \cdots, k-1]$ we know that $\mathcal{F}_s \subset \mathcal{F}_{t_i}$ then by linearity, Theorem 1.2.2 (4) and Theorem 2.3.1

$$\begin{aligned}\mathbb{E}\left[\sum_{i=l+1}^{k-1}\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right]\middle|\mathcal{F}(s)\right] &= \sum_{i=l+1}^{k-1}\mathbb{E}\left[\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right]\middle|\mathcal{F}(s)\right] \\ &= \sum_{i=l+1}^{k-1}\mathbb{E}\left[\mathbb{E}\left[\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right]|\mathcal{F}_{t_i}\right]\middle|\mathcal{F}(s)\right] \\ &= \sum_{i=l+1}^{k-1}\mathbb{E}\left[\Delta(t_i)\left(\mathbb{E}\left[W(t_{i+1})|\mathcal{F}(t_i)\right] - W(t_i)\right)\middle|\mathcal{F}(s)\right]. \\ &= \sum_{i=l+1}^{k-1}\mathbb{E}\left[\Delta(t_i)\left(W(t_i) - W(t_i)\right)\middle|\mathcal{F}(s)\right] = 0\end{aligned}$$

For the last term follows in the same way that as for the third term that

$$\mathbb{E}\left[\Delta(t_k)\left[W(t) - W(t_k)\right]\right] = 0.$$

Theorem 3.1.2 (Itô Isometry)

The Itô integral for simple processes statisfies

$$\mathbb{E}\left[I^2(t)\right] = \mathbb{E}\left[\int_0^t \Delta^2(u)du\right].$$

Proof
We know that the Itô Integral is given by

$$I(t) = \sum_{i=0}^{k-1}\Delta(t_i)\left[W(t_{i+1}) - W(t_i)\right] + \Delta(t_k)\left[W(t) - W(t_k)\right].$$

We can write this as one summation using $\Delta W_i = W(t_{i+1}) - W(t_i)$ and $\Delta W_k = W(t) - W(t_k)$. Hence

$$I(t) = \sum_{i=0}^{k}\Delta(t_i)\Delta W_i.$$

Then follows

$$\begin{aligned} I^2(t) &= \left(\sum_{i=0}^{k} \Delta(t_i)\Delta W_i\right)^2 \\ &= \sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2 + \sum_{i\neq j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j \quad . \\ &= \sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2 + 2\sum_{i<j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}\left[I^2(t)\right] &= \mathbb{E}\left[\sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2 + 2\sum_{i<j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2\right] + \mathbb{E}\left[2\sum_{i<j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right]. \\ &= \sum_{i=0}^{k} \mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] + 2\sum_{i<j} \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right] \end{aligned}$$

Then because ΔW_j is independent of $\mathcal{F}(t_j)$ and $\Delta(t_i)\Delta(t_j)\Delta W_i$ is $\mathcal{F}(t_j)$-measurable for $i < j$

$$\begin{aligned} \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right] &= \mathbb{E}\left[\mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j | \mathcal{F}(t_j)\right]\right] \\ &= \mathbb{E}\left[\Delta W_j \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i | \mathcal{F}(t_j)\right]\right] \\ &= \mathbb{E}\left[\Delta W_j\right] \mathbb{E}\left[\mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i | \mathcal{F}(t_j)\right]\right]. \\ &= \mathbb{E}\left[\Delta W_j\right] \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i\right] \end{aligned}$$

Then because ΔW_j is the increment of a Brownian motion

$$\mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right] = \mathbb{E}\left[\Delta W_j\right] \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i\right] = 0.$$

On the otherhand since $\Delta^2(t_j)$ is $\mathcal{F}(t_j)$-measurable and ΔW_j^2 is independent of $\mathcal{F}(t_j)$

$$\mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] = \mathbb{E}\left[\Delta^2(t_i)\right] \mathbb{E}\left[\Delta W_i^2\right].$$

Then using Remark 2.3.1 we find that $\mathbb{E}[\Delta W_i^2] = t_{i+1} - t_i$ and $\mathbb{E}[\Delta W_k^2] = t - t_k$, hence

$$\begin{aligned} \sum_{i=0}^{k} \mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] &= \sum_{i=0}^{k-1} \mathbb{E}\left[\Delta^2(t_i)\right](t_{i+1} - t_i) + \mathbb{E}\left[\Delta^2(t_k)\right](t - t_k) \\ &= \mathbb{E}\left[\Delta^2(t_i)\left(t_{i+1} - t_i\right)\right] + \mathbb{E}\left[\Delta^2(t_k)\left(t - t_k\right)\right] \quad . \\ &= \mathbb{E}\left[\Delta^2(t_i)\left(t_{i+1} - t_i\right) + \Delta^2(t_k)\left(t - t_k\right)\right] = \int_0^t \Delta^2(u)du \end{aligned}$$

Hence we have shown that

$$\mathbb{E}\left[I^2(t)\right] = \mathbb{E}\left[\int_0^t \Delta^2(u)du\right].$$

Theorem 3.1.3

The quadratic variation of the Itô integral for a simple process

$$[I, I](t) = \int_0^t \mathit{\Delta}^2(u)du.$$

Proof We will proof this in increments. Let $[t_i, t_{i+1}]$ be a subinterval on which $\{\mathit{\Delta}(t) : 0 \le t \le T\}$ is constant. Now consider the partition of the interval given by

$$t_i = s_0 < s_1 < \cdots < s_n = t_{i+1}.$$

Define $\|\Pi\| := \max_{0 \le i \le (n-1)} |t_{i+1} - t_i|$. Then the quadratic variation on this interval is given by

$$\begin{aligned} \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [I(s_{j+1} - I(s_j))]^2 &= \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [\mathit{\Delta}(t_i)(W(s_{j+1}) - W(s_j))]^2 \\ &= \mathit{\Delta}^2(t_i) \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [W(s_{j+1}) - W(s_j)]^2 \end{aligned}.$$

Note that the limit given in this expression is the quadratic variation of a Brownian motion between t_{i+1} and t_i. Then in a similar way as we did in Theorem 2.3.2 we can show that the quadratic variation is in this case equal to $t_{i+1} - t_i$. Then because $\mathit{\Delta}(u) = \mathit{\Delta}(t_j)$ for $u \in [t_j, t_{j+1}]$

$$\mathit{\Delta}^2(t_i) \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [W(s_{j+1}) - W(s_j)]^2 = \mathit{\Delta}^2(t_i)(t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} \mathit{\Delta}(u)^2 du.$$

Note that the quadratic variation between times t_k and t is evenso given by

$$\int_{t_k}^t \mathit{\Delta}^2(u)du.$$

Then by summing the quadratic variations for the subintervals we find that indeed

$$[I, I](t) = \int_0^t \mathit{\Delta}^2(u)du.$$

3.2 Itô's Integral in the General Case

In the previous sections we assumed that $\mathit{\Delta}(t)$ was constant on each subinterval of some fixed partition of $[0, T]$. Now we will look at the general case where $\mathit{\Delta}(t)$ need not have piecewise constant paths on the interval $[0, T]$. In order to define the Itô Integral for such a process we assume that $\{\mathit{\Delta}(t) : t \ge 0\}$ is adapted to the filtration $\{\mathcal{F}(t) : t \ge 0\}$. We will also assume that

$$\mathbb{E}\left[\int_0^T \mathit{\Delta}^2(t)dt\right] < \infty. \tag{4}$$

We need this assumption to prevent convergence issues. The idea is to find a sequence $\mathit{\Delta}_n(t)$ such that $\mathit{\Delta}_n(t)$ is a simple process for each n, and approaches $\mathit{\Delta}(t)$ as $n \to \infty$. We can do this by choosing a partition of $[0, T]$, i.e. $\Pi_1 = \{t_0, t_1, \cdots, t_n\}$. Then we define $\mathit{\Delta}_1(t)$ to take the value at the left end part, i.e.

$$\mathit{\Delta}_1(t) = \mathit{\Delta}(t_i) \quad t_i < t < t_{i+1}.$$

We then proceed to create an increasing sequence of partitions $(\Pi_n)_{n \in \mathbb{N}}$, i.e. $\Pi_n \subset \Pi_{n+1}$ for all n, with $\lim_{n \to \infty} \|\Pi\| = 0$. Remember that $\|\Pi\| := \max_{0 \le i \le (n-1)} |t_{i+1} - t_i|$. We let $\mathit{\Delta}_n$ the simple process associated with Π_n, defined the same way as we did for Δ_1 and Π_1. Then, we can show that the sequence $\Delta_n(t)$ converges to $\Delta(t)$ in $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. That is

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\Delta_n(t) - \Delta(t)|^2 \, dt\right] = 0.$$

Let $I_n(t) = \int_0^t \Delta_n(u)dW(u)$, that is the Itô integral we defined for simple processes. The idea is to show that $(I_n(t))_{n\in\mathbb{N}}$ is a cauchy sequence. We consider the same norm as in Theorem 2.3.2. By Theorem 3.1.2

$$\mathbb{E}\left[(I_n(t) - I_m(t))^2\right] = \mathbb{E}\left[\int_0^t |\Delta_n(u) - \Delta_m(u)|^2 \, du\right].$$

Then as n and m approach infinity

$$\mathbb{E}\left[\int_0^t |\Delta_n(u) - \Delta_m(u)|^2 \, du\right] \to 0.$$

This is because of assumption (4). Hence we conclude that $(I_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence. Since $\mathcal{L}^2$ induced with the norm we considered is a complete space, we must have that the sequence converges in $\mathcal{L}^2$. In particular the sequence converges to the Itô integral of the adapted process $\Delta(t)$. We define the Itô integral of $\Delta(t)$ to be

$$\int_0^t \Delta(u)dW(u) = \lim_{n\to\infty} \int_0^t \Delta_n(u)dW(u).$$

Note that all the properties proven for the Itô integral for simple processes also hold in the general case. This is very important for theorems that appear later on.

4 The Black–Scholes Equation

In this chapter we will derive the Black-Scholes equation. The Black-Scholes equation is a partial stocahstic differential equation and it describes the price changes of a European call option. An European option gives the buyer the option to buy a stock for a certain price, called the strike price, regardless of the price of the stock at that time. The buyer decides at time T whether or not to buy the stock, he is not obligated to do so. In order to do this we will need to introduce the Itô-Doeblin Formula. Then when we have derived the Black-Scholes stochastic differntial equation, we will derive its solution, The Black-Scholes Formula, using Girsanov's Theorem and using properties of Martingales.

4.1 Itô Process

We can extend the notion of the Itô integral for a more general case. So we can integrate with respect to a Itô process.

Definition 4.1.1 (Itô Process)

Let $\{W(t) : t \geq 0\}$ be a Brownian Motion and $\{\mathcal{F}(t) : t \geq 0\}$ a filtration for $\{W(t) : t \geq 0\}$, an *Itô process* is a stochastic process of the form

$$X(t) = X(0) + \int_0^t \mathit{\Delta}(u)dW(u) + \int_0^t \Theta(u)du,$$

where $X(0)$ is non-random and $\{\mathit{\Delta}(u) : u \geq 0\}$, $\{\Theta(u) : u \geq 0\}$ adapted stochastic processes. We assume that the integrals exist and in particular that $\mathbb{E}\left[\int_0^T \mathit{\Delta}^2(t)dt\right] < \infty$. Hence assumption (4) is satisfied. Note that a Brownian motion is a Itô process.

Definition 4.1.2 (Integral with respect to an Itô Process)

Let $\{X(t) : t \geq 0\}$ be an Itô process, let $\{\Gamma(t) : t \geq 0\}$ be an adapted process. Then the *integral with respect to the Itô process* is defined by

$$\int_0^t \Gamma(u)dX(u) = \int_0^t \Gamma(u)\mathit{\Delta}(u)dW(u) + \int_0^t \Gamma(u)\Theta(u)du.$$

We assume that the integrals above exist, and in particular that $\mathbb{E}\left[\int_0^T \mathit{\Delta}^2(t)dt\right] < \infty$. Hence it satisfies assumption (4)

4.2 The Itô-Doeblin Formula

The Itô-Doeblin Formula will be used regularly in the remaining part of this chapter. Hence consider the following theorem

Theorem 4.2.1 (Itô-Doeblin Formula)

Let $f(t, x)$ be a function which has continous partial derivatives $\frac{\partial}{\partial t}f(t, x)$, $\frac{\partial}{\partial x}f(t, x)$ and $\frac{\partial^2}{\partial x^2}f(t, x)$. Let $\{X(t) : t \geq 0\}$ be an Itô process as in Definition 4.1.1 then

$$f(T, X(T)) = f(0, X(0)) + \int_0^T \frac{\partial}{\partial t}f(t, X(t))dt + \int_0^T \frac{\partial}{\partial x}f(t, X(t))dX(t) + \frac{1}{2}\int_0^T \frac{\partial^2}{\partial x^2}f(t, X(t))d[X, X](t).$$

Proof
Consider the interval $[0,T]$ and let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of this interval. Recall that we used $\|\Pi\| := \max_{0\le i\le (n-1)} |t_{i+1} - t_i|$. Then we write

$$f(T, X(T)) - f(0, X(0)) = \sum_{i=0}^{n-1} f\left(t_{i+1}, X\left(t_{i+1}\right)\right) - f\left(t_i, X\left(t_i\right)\right).$$

Then with the use of a 2-dimensional Taylor expansion we write,

$$\begin{aligned}
\sum_{i=0}^{n-1} f\left(t_{i+1}, X\left(t_{i+1}\right)\right) - f\left(t_i, X\left(t_i\right)\right) &= \sum_{i=0}^{n-1} \frac{\partial}{\partial t} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right) \\
&+ \sum_{i=0}^{n-1} \frac{\partial}{\partial x} f\left(t_i, X\left(t_i\right)\right)\left(X(t_{i+1}) - X(t_i)\right) \\
&+ \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial x^2} f\left(t_i, X\left(t_i\right)\right)\left(X(t_{i+1}) - X(t_i)\right)^2 \\
&+ \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t \partial x} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right)\left(X(t_{i+1}) - X(t_i)\right) \\
&+ \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right)^2 + \text{H.O.T. (Higher order terms)}.
\end{aligned}$$

Then if $\|\Pi\| \to 0$ we have that

$$\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \frac{\partial}{\partial t} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right) = \int_0^T \frac{\partial}{\partial t} f(t, X(t)) dt,$$

and

$$\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \frac{\partial}{\partial x} f\left(t_i, X\left(t_i\right)\right)\left(X(t_{i+1}) - X(t_i)\right) = \int_0^T \frac{\partial}{\partial x} f\left(t, X\left(t\right)\right) dX(t),$$

$$\lim_{\|\Pi\|\to 0} \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial x^2} f\left(t_i, X\left(t_i\right)\right)\left(X(t_{i+1}) - X(t_i)\right)^2 = \frac{1}{2} \int_0^T \frac{\partial^2}{\partial x^2} f(t, X(t)) d[X, X](t).$$

While for the other terms

$$\begin{aligned}
&\lim_{\|\Pi\|\to 0} \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t \partial x} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right)\left(X(t_{i+1}) - X(t_i)\right) \\
&\le \lim_{\|\Pi\|\to 0} \max_{0\le i\le (n-1)} \left(X(t_{i+1}) - X(t_i)\right) \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t \partial x} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right) \cdot \\
&= 0 \cdot \int_0^T \frac{\partial^2}{\partial t \partial x} f\left(t, X\left(t\right)\right) dX(t) = 0
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{\|\Pi\|\to 0} \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right)^2 \\
&\le \lim_{\|\Pi\|\to 0} \|\Pi\| \frac{1}{2} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f\left(t_i, X\left(t_i\right)\right)\left(t_{i+1} - t_i\right) \cdot \\
&= 0 \cdot \int_0^T \frac{\partial^2}{\partial t^2} f\left(t, X\left(t\right)\right) dt
\end{aligned}$$

Hence

$$f(T,X(T)) - f(0,X(0)) = \int_0^T \frac{\partial}{\partial t} f(t, X(t))\, dt + \int_0^T \frac{\partial}{\partial x} f(t, X(t))\, dX(t) + \frac{1}{2}\int_0^T \frac{\partial^2}{\partial x^2} f(t,X(t)) d[X,X](t).$$

and thus yields the desired equality which is often written in differential form

$$df(t,X(t)) = \frac{\partial}{\partial t} f(t, X(t))\, dt + \frac{\partial}{\partial x} f(t, X(t))\, dX(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t,X(t))dX(t)dX(t)$$

Remark 4.2.1

Note that in Theorem 4.2.1 we can informally write

$$dtdt = 0$$

and

$$dtdX(t) = dX(t)dt = 0$$

Theorem 4.2.2 (Product Rule)

Let $X(t)$ and $Y(t)$ be Itô processes, then

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)Y(t)$$

Proof

In similar fashion as for Theorem 4.2.1, i.e. using a 3-dimensional Taylor expansion, one could show that given $f(t,x,y)$ a function which has continious partial derivatives $\frac{\partial}{\partial t}f(t,x,y)$, $\frac{\partial}{\partial x}f(t,x,y)$, $\frac{\partial}{\partial y}f(t,x,y)$, $\frac{\partial^2}{\partial x^2}f(t,x,y)$, $\frac{\partial^2}{\partial y^2}f(t,x,y)$, $\frac{\partial^2}{\partial x \partial y}f(t,x,y)$ and $\frac{\partial}{\partial y \partial x}f(t,x,y)$ that

$$\begin{aligned} df(t,X(t),Y(t)) = &\frac{\partial}{\partial t} f(t, X(t), Y(t))\, dt + \frac{\partial}{\partial x} f(t, X(t), Y(t))\, dX(t) + \frac{\partial}{\partial y} f(t, X(t), Y(t))\, dY(t) \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t, X(t), Y(t))\, dX(t)X(t) + \frac{1}{2}\frac{\partial^2}{\partial y^2} f(t, X(t), Y(t))\, dY(t)Y(t) \\ &+ \frac{\partial^2}{\partial x \partial y} f(t, X(t), Y(t))\, dX(t)Y(t) \end{aligned} \quad . \tag{5}$$

Now let $f(t,x,y) = xy$, then by (5)

$$\begin{aligned} d(X(t)Y(t)) &= df(t,X(t),Y(t)) = 0 + Y(t)dX(t) + X(t)dY(t) + 0 + 0 + dX(t)dY(t) \\ &= X(t)dY(t) + Y(t)dX(t) + dX(t)Y(t) \end{aligned} .$$

4.3 The Black-Scholes Equation

We will derive the Black-Scholes equation for the price of an European option based on an asset modeled as a geometric Brownian motion. Suppose we have a portfolio, where the value of the portfolio is given by $X(t)$ for each time t. The portfolio invests in a money market account which yields a constant rate of interest r. It also invests in a stock which is modeled by a geometric Brownian motion. Recall from Chapter 2 that a geometric Brownian motion is defined by

$$S(t) = S(0)e^{\sigma W(t) + (\alpha - \frac{1}{2}\sigma^2)t},$$

where $\alpha > 0$, σ are constants and $\{W(t) : t \geq 0\}$ is a Brownian motion. Using the Itô Doeblin formula derived in Theorem 4.2.1 we can write this in its differential form. Let $f(t,x) = S(0)e^{\sigma x + \left(\alpha - \frac{1}{2}\sigma^2\right)t}$ then

$$\begin{aligned} dS(t) = df(t, W(t)) &= \left(\alpha - \frac{1}{2}\sigma^2\right) S(t)dt + \sigma S(t)dW(t) + \frac{1}{2}\sigma^2 S(t)dt \\ &= \alpha S(t)dt + \sigma S(t)dW(t) \end{aligned}$$

we will use this later on. Suppose that at each time $t \geq 0$ we have $\Delta(t)$ shares of the stock. The value of $\Delta(t)$ can be random, but must at all times be adapted to the filtration induced by the Brownian motion $\{W(t) : t \geq 0\}$. The remaining value of the portfolio $X(t) - \Delta(t)S(t)$ is invested in the money market account. Then its clear that the change in portfolio value $dX(t)$ is described by the change in capital gain $\Delta(t)dS(t)$ and the earnings by interest $r\left(X(t) - \Delta(t)S(t)\right)dt$, hence

$$dX(t) = \Delta(t)dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt.$$

Using the differntial form of the geometric Brownian motion we can write

$$\begin{aligned} \Delta(t)dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt &= \Delta(t)\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) + r\left(X(t) - \Delta(t)S(t)\right)dt \\ &= \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt \end{aligned}.$$

We will often consider the discounted stock price $e^{-rt}S(t)$ and the discounted value of our portfolio $e^{-rt}X(t)$. Using the Itô-Doeblin formula we can derive the differentials of these functions. Hence

$$\begin{aligned} d\left(e^{-rt}S(t)\right) &= \frac{\partial}{\partial t}e^{-rt}x\bigg|_{x=S(t)} dt + \frac{\partial}{\partial x}e^{-rt}x\bigg|_{x=S(t)} dS(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}e^{-rt}x\bigg|_{x=S(t)} dS(t)dS(t) \\ &= -re^{-rt}S(t)dt + e^{-rt}dS(t) \\ &= -re^{-rt}S(t)dt + e^{-rt}\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) \\ &= (\alpha - r)\,e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t) \end{aligned},$$

and analogously

$$\begin{aligned} d\left(e^{-rt}X(t)\right) &= \frac{\partial}{\partial t}e^{-rt}x\bigg|_{x=X(t)} dt + \frac{\partial}{\partial x}e^{-rt}x\bigg|_{x=X(t)} dX(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}e^{-rt}x\bigg|_{x=X(t)} dX(t)dX(t) \\ &= -re^{-rt}X(t)dt + e^{-rt}dX(t) \\ &= -re^{-rt}X(t)dt + e^{-rt}\left(\Delta(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt\right) \\ &= \Delta(t)\,(\alpha - r)\,e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t) \\ &= \Delta(t)d\left(e^{-rt}S(t)\right) \end{aligned}.$$

Suppose we have a European call option that pays us $(S(T) - K)^+$ at time T. Where

$$(S(T) - K)^+ = \begin{cases} S(T) - K & \text{if } S(T) > K \\ 0 \quad \text{if } S(T) \leq K \end{cases},$$

and K is a contractual determined value, note that K usually is called the strike price. Black, Scholes and Merton assumed that the value of this call at any time t, should depend upon the price of the stock and the time, which is a reasonable assumption. It is true that the value of the call also depends upon the values r, σ and K, but these can be considered constants. Hence we define $c(t,x)$ to be the function which denotes the value of the call at a given time t when $S(t) = x$, at that particular time. Hence $c(t, S(t))$ is a stochastic process. Our goal is to find the function $c(t,x)$ so we have information about the future call values for future stock prices. Using the Itô Doeblin formula we find that

$$dc(t, S(t)) = \frac{\partial}{\partial t}c(t,x)dt\bigg|_{x=S(t)} + \frac{\partial}{\partial x}c(t,x)dS(t)\bigg|_{x=S(t)} + \frac{1}{2}\frac{\partial^2}{\partial x^2}c(t,x)dS(t)dS(t)\bigg|_{x=S(t)}.$$

Using Remark 4.2.1

$$\begin{aligned} dS(t)dS(t) &= \alpha^2 S^2(t)dtdt + 2\sigma\alpha S(t)dW(t)dt + \sigma^2 S^2(t)dW(t)dW(t) \\ &= \sigma^2 S^2(t)dt \end{aligned}.$$

Hence

$$\begin{aligned} dc(t,S(t)) &= \frac{\partial}{\partial t}c(t,x)dt\Big|_{x=S(t)} + \frac{\partial}{\partial x}c(t,x)dS(t)\Big|_{x=S(t)} + \frac{1}{2}\frac{\partial^2}{\partial x^2}c(t,x)\sigma^2 S^2(t)dt\Big|_{x=S(t)} \\ &= \frac{\partial}{\partial t}c(t,x)dt\Big|_{x=S(t)} + \frac{\partial}{\partial x}c(t,x)\left(\alpha S(t)dt + \sigma S(t)dW(t)\right)\Big|_{x=S(t)} + \frac{1}{2}\frac{\partial^2}{\partial x^2}c(t,x)\sigma^2 S^2(t)dt\Big|_{x=S(t)}, \\ &= \left(\frac{\partial}{\partial t}c(t,x) + \alpha S(t)\frac{\partial}{\partial x}c(t,x) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,x)\right)dt\Big|_{x=S(t)} + \sigma S(t)\frac{\partial}{\partial x}c(t,x)dW(t)\Big|_{x=S(t)} \end{aligned}$$

where for convenience we will now write

$$dc(t,S(t)) = \left(\frac{\partial}{\partial t}c(t,S(t)) + \alpha S(t)\frac{\partial}{\partial x}c(t,S(t)) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,S(t))\right)dt + \sigma S(t)\frac{\partial}{\partial x}c(t,S(t))dW(t).$$

Analogously, we derive for the discounted option price,

$$\begin{aligned} d\left(e^{-rt}c(t,S(t))\right) = e^{-rt}&\left(-rc(t,S(t)) + \frac{\partial}{\partial t}c(t,S(t)) + \alpha S(t)\frac{\partial}{\partial x}c(t,S(t)) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,S(t))\right)dt \\ &+ e^{-rt}\sigma S(t)\frac{\partial}{\partial x}c(t,S(t))dW(t) \end{aligned}.$$

For a short hedging portfolio, with initial captial $X(0)$, we must have that the portfolio value $X(t)$ for each time $0 \leq t \leq T$ agrees with the value of the option $c(t,S(t))$. This is only true when

$$e^{-rt}X(t) = e^{-rt}c(t,S(t)), \quad \forall t,$$

This can be ensured by having

$$\begin{cases} d\left(e^{-rt}X(t)\right) = d\left(e^{-rt}c(t,S(t))\right) \\ X(0) = c(0,S(0)) \end{cases},$$

then integration yields

$$\begin{aligned} e^{-rt}X(t) - X(0) &= e^{-rt}c(t,S(t)) - c(0,S(0)), \quad \forall t, \\ e^{-rt}X(t) &= e^{-rt}c(t,S(t)), \quad \forall t, \end{aligned}$$

Hence the desired equality is satisfied. Because $d\left(e^{-rt}X(t)\right) = d\left(e^{-rt}c(t,S(t))\right)$

$$\begin{aligned} \Delta(t)d\left(e^{-rt}S(t)\right) &= \Delta(t)\left((\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)\right) \\ &= e^{-rt}\left(-rc(t,S(t)) + \frac{\partial}{\partial t}c(t,S(t)) + \alpha S(t)\frac{\partial}{\partial x}c(t,S(t)) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,S(t))\right)dt. \\ &+ e^{-rt}\sigma S(t)\frac{\partial}{\partial x}c(t,S(t))dW(t) \end{aligned}$$

Since the left hand side and the right hand side must be equal we have that

$$\Delta(t)\sigma e^{-rt}S(t)dW(t) = e^{-rt}\sigma S(t)\frac{\partial}{\partial x}c(t,S(t))dW(t).$$

Hence

$$\Delta(t) = \frac{\partial}{\partial x} c(t, S(t)) = \frac{\partial}{\partial x} c(t, x)\bigg|_{x=S(t)},$$

for $0 \leq t < T$. On the otherhand

$$\Delta(t)\,(\alpha - r)\,e^{-rt} S(t) dt = e^{-rt}\left(-rc(t, S(t)) + \frac{\partial}{\partial t} c(t, S(t)) + \alpha S(t) \frac{\partial}{\partial x} c(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) \frac{\partial^2}{\partial x^2} c(t, S(t))\right) dt,$$

hence

$$\frac{\partial}{\partial t} c(t, S(t)) + rS(t) \frac{\partial}{\partial x} c(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) \frac{\partial^2}{\partial x^2} c(t, S(t)) = rc(t, S(t)),$$

for $0 \leq t \leq T$. Hence we should find a twice differentiable function that is a solution to

$$\begin{cases} \frac{\partial}{\partial t} c(t, x) + rx \frac{\partial}{\partial x} c(t, x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} c(t, x) = rc(t, x) \\ c(T, x) = (x - K)^+ \end{cases},$$

where

$$\frac{\partial}{\partial t} c(t, x) + rx \frac{\partial}{\partial x} c(t, x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} c(t, x) = rc(t, x),$$

is called the Black-Scholes-Merton equation.

4.4 Girsanov's Theorem

In this section we will prove Girsanov's Theorem. Hence consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{W(t) : 0 \leq t \leq T\}$ be a Brownian motion, let $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ be a filtration for $\{W(t) : 0 \leq t \leq T\}$, and let $\{\Theta(t) : 0 \leq t \leq T\}$ be an adapted process. We define

$$\begin{aligned} Z(t) &= e^{-\int_0^t \Theta(u) dW(u) - \frac{1}{2}\int_0^t \Theta^2(u) du} \\ \widetilde{W}(t) &= W(t) + \int_0^t \Theta(u) du \end{aligned}.$$

Let $Z = Z(T)$, we define

$$\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P}.$$

Then $\mathbb{E}[Z] = 1$ and $\widetilde{W}(t)$ is a $\widetilde{\mathbb{P}}$-Brownian motion. In order for the Itô integral to converge properly we assume that

$$\mathbb{E}\left[\int_0^T (\Theta(t) Z(t))^2 \, dt\right] < \infty.$$

Proof Let

$$X(t) = -\int_0^t \Theta(u) dW(u) - \frac{1}{2}\int_0^t \Theta^2(u) du,$$

in differntial form

$$dX(t) = -\Theta(t) dW(t) - \frac{1}{2}\Theta^2(t) dt,$$

and

$$dX(t) dX(t) = \Theta^2(t) dt,$$

then by the Itô Doeblin formula

$$\begin{aligned}
dZ(t) &= \frac{d}{dx}e^x dX(t) + \frac{1}{2}\frac{d^2}{dx^2}e^x dX(t)dX(t)\Big|_{x=X(t)} \\
&= e^{X(t)}\left(-\Theta(t)dW(t) - \frac{1}{2}\Theta^2(t)dt\right) + \frac{1}{2}e^{X(t)}\Theta^2(t)dt \\
&= -\Theta(t)Z(t)dW(t)
\end{aligned}.$$

Hence we have that

$$\begin{aligned}
Z(t) - Z(0) &= -\int_0^t \Theta(u)Z(u)dW(u) \\
Z(t) &= Z(0) - \int_0^t \Theta(u)Z(u)dW(u)
\end{aligned}.$$

Since the Itô integral is a Martingale we have that

$$\begin{aligned}
\mathbb{E}\left[Z(t)|\mathcal{F}(s)\right] &= \mathbb{E}\left[Z(0) - \int_0^t \Theta(u)Z(u)dW(u)\Big|\mathcal{F}(s)\right] \\
&= Z(0) - \mathbb{E}\left[\int_0^t \Theta(u)Z(u)dW(u)\Big|\mathcal{F}(s)\right] \\
&= Z(0) - \int_0^s \Theta(u)Z(u)dW(u) = Z(s)
\end{aligned}.$$

Thus $Z(t)$ is a $\mathbb{P}$-Martingale. Then by Theorem 2.2.1 we have that

$$\mathbb{E}[Z] = \mathbb{E}[Z(T)] = \mathbb{E}[Z(0)] = \mathbb{E}[e^0] = 1.$$

and because of the Martingale property

$$Z(t) = \mathbb{E}\left[Z(T)|\mathcal{F}(t)\right] = \mathbb{E}\left[Z|\mathcal{F}(t)\right],$$

so $Z(t)$ is also a Radon-Nikodým derivative process as defined in Definition 2.2.1. Note that

$$d\widetilde{W}(t) = dW(t) + \Theta(t)dt.$$

Then by using the product rule from Theorem 4.2.2 we find that

$$\begin{aligned}
d\left(\widetilde{W}(t)Z(t)\right) &= \widetilde{W}(t)dZ(t) + Z(t)d\widetilde{W}(t) + d\widetilde{W}(t)dZ(t) \\
&= -\widetilde{W}\Theta(t)Z(t)dW(t) + Z(t)\left(dW(t) + \Theta(t)dt\right) + \left(dW(t) + \Theta(t)dt\right)\left(-\Theta(t)Z(t)dW(t)\right), \\
&= \left(-\widetilde{W}\Theta(t) + 1\right)Z(t)dW(t)
\end{aligned}$$

and thus we can show in a similar way as for $Z(t)$ that $\widetilde{W}(t)Z(t)$ is a $\mathbb{P}$-Martingale. Then with Theorem 2.2.2

$$\widetilde{E}\left[\widetilde{W}(t)|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\mathbb{E}\left[\widetilde{W}(t)Z(t)|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\widetilde{W}(s)Z(s) = \widetilde{W}(s),$$

hence $\widetilde{W}(s)$ is a $\widetilde{\mathbb{P}}$-Martingale. Since $\widetilde{W}(0) = W(0) = 0$ and $d\widetilde{W}(t)d\widetilde{W}(t) = dW(t)dW(t) = dt$, then by Theorem 2.3.4 we have that $\widetilde{W}(t)$ is a $\widetilde{\mathbb{P}}$-Brownian motion.

4.5 Derivation of the Black-Scholes Formula

Now we have all the necessities to derive the solution to the Black-Scholes equation. In Section 4.3 we showed that

$$\begin{aligned} S(t) &= S(0)e^{\sigma W(t)+\left(\alpha-\frac{1}{2}\sigma^2\right)t} \\ d\left(e^{-rt}S(t)\right) &= (\alpha-r)\,e^{-rt}S(t)dt+\sigma e^{-rt}S(t)dW(t) \\ &= \sigma e^{-rt}S(t)\left(\frac{\alpha-r}{\sigma}dt+dW(t)\right) \end{aligned}.$$

Define $\Theta = \dfrac{\alpha-r}{\sigma}$ then by Girsanov's Theorem we have that

$$d\left(e^{-rt}S(t)\right) = \sigma e^{-rt}S(t)d\widetilde{W}(t).$$

Using $\sigma W(t) = \sigma\widetilde{W}(t) - (\alpha-r)t$, we have

$$S(t) = S(0)e^{\sigma\widetilde{W}(t)+\left(r-\frac{1}{2}\sigma^2\right)t}.$$

If we let $X(t)$ denote the portfolio value as we did in Section 4.3 we have

$$d\left(e^{-rt}X(t)\right) = \mathit{\Delta}(t)d\left(e^{-rt}S(t)\right) = \mathit{\Delta}(t)\sigma e^{-rt}S(t)d\widetilde{W}(t),$$

Thus we see that $e^{-rt}X(t)$ is a $\widetilde{\mathbb{P}}$-Martingale. In Section 4.3 we derived the Black-Sholes equation for a European call, we wanted to know $X(0)$ and what portfolio process was needed to hedge a short position in the call, that is

$$X(T) = (S(T)-K)^+ \quad \text{a.s.}$$

We say that $V(T) = (S(T)-K)^+$ is called the derivative security payoff. Since we know that $e^{-rt}X(t)$ is a $\widetilde{\mathbb{P}}$-Martingale

$$e^{-rt}X(t) = \widetilde{\mathbb{E}}\left[e^{-rT}X(T)|\mathcal{F}(t)\right] = \widetilde{\mathbb{E}}\left[e^{-rT}V(T)|\mathcal{F}(t)\right].$$

If we now call $X(t)$ the price $V(t)$ of the derivative security at time t we have

$$\begin{aligned} e^{-rt}V(t) &= \widetilde{\mathbb{E}}\left[e^{-rT}V(T)|\mathcal{F}(t)\right] \\ V(t) &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)|\mathcal{F}(t)\right]. \end{aligned}$$

Where $V(t)$ is often called the risk-neutral pricing formula. Note that this is specific to the assumptions made in section 4.3. In general σ, r, α need not to be constants. Because the geometric Brownian motion depends on the stock price $S(t)$ and on time t at which the expectation is computed, but not on the stock price prior to t, we can conclude it is a markov process, see Definition 2.3.6 and Remark 2.3.7 from [2]. Hence there exists a function $c(t,x)$ such that

$$\begin{aligned} c(t,S(t)) &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)|\mathcal{F}(t)\right] \\ &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}\left(S(T)-K\right)^+|\mathcal{F}(t)\right] \end{aligned}.$$

Let $Y(t) = \sigma\widetilde{W}(t) + \left(r-\frac{1}{2}\sigma^2\right)t$ then

$$\begin{aligned} S(T) &= S(0)e^{Y(T)} \\ &= S(0)e^{Y(T)+Y(t)-Y(t)} \\ &= S(0)e^{Y(t)}e^{Y(T)-Y(t)} \\ &= S(t)e^{Y(T)-Y(t)}) \\ &= S(t)e^{\sigma\left(\widetilde{W}(T)-\widetilde{W}(t)\right)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)} \end{aligned}.$$

For convience we write $\tau = T - t$, with $Z = \dfrac{\widetilde{W}(T) - \widetilde{W}(t)}{\sqrt{\tau}}$ hence

$$S(T) = S(t)e^{\sigma\sqrt{\tau}Z + \left(r - \frac{1}{2}\sigma^2\right)\tau},$$

where $Z \sim \mathcal{N}(0,1)$. Then since $S(t)$ is $\mathcal{F}(t)$-measurable and $e^{\sigma\sqrt{\tau}Z + \left(r - \frac{1}{2}\sigma^2\right)\tau}$ is independent of $\mathcal{F}(t)$

$$\begin{aligned} c(t, S(t)) &= \widetilde{\mathbb{E}}\left[e^{-r\tau}\left(S(t)e^{\sigma\sqrt{\tau}Z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right)^+\right] \\ c(t, x) &= \widetilde{\mathbb{E}}\left[e^{-r\tau}\left(xe^{\sigma\sqrt{\tau}Z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right)^+\right]. \end{aligned}$$

Then because $Z \sim \mathcal{N}(0,1)$

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(xe^{\sigma\sqrt{\tau}Z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right)^+\right] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-r\tau}\left(xe^{\sigma\sqrt{\tau}z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right)^+ e^{-\frac{1}{2}z^2} dz$$

Then

$$\left(xe^{\sigma\sqrt{\tau}z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right)^+ > 0,$$

if and only if

$$z > \frac{\log(\frac{K}{x}) - \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} = \alpha.$$

Hence we change the set over which we integrate, which yields

$$c(t, x) = \int_{\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-r\tau}\left(xe^{\sigma\sqrt{\tau}z + \left(r - \frac{1}{2}\sigma^2\right)\tau} - K\right) e^{-\frac{1}{2}z^2} dz.$$

Let $u = -z$, then $du = -dz$ and we have that $u < \frac{\log(\frac{x}{K}) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} = \beta_1$ hence the integral becomes

$$\begin{aligned} c(t, x) &= \int_{-\infty}^{\beta_1} \frac{1}{\sqrt{2\pi}} e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}u + \left(r - \frac{1}{2}\sigma^2\right)\tau}\right) du \\ &= x\int_{-\infty}^{\beta_1} \frac{1}{\sqrt{2\pi}} e^{-\sigma\sqrt{\tau}u - \frac{1}{2}\sigma^2\tau} e^{-\frac{1}{2}u^2} du - e^{-r\tau}K\int_{-\infty}^{\beta_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du. \\ &= x\int_{-\infty}^{\beta_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(u + \sigma\sqrt{\tau}\right)^2} du - e^{-r\tau}K\int_{-\infty}^{\beta_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ &= x\Phi(\beta_1 + \sigma\sqrt{\tau}) - e^{-r\tau}K\Phi(\beta_1) \end{aligned}$$

Hence the price $c(t, x)$ is given by

$$c(t, x) = x\Phi(\beta_2) - e^{-r\tau}K\Phi(\beta_1),$$

where $\Phi(z) = \mathbb{P}(Z < z)$, $Z \sim \mathcal{N}(0,1)$ and $\beta_2 = \beta_1 + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K}) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K}) + \left(r + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$.

5 Parameter Estimation

In this chapter we will take a look at how to estimate the parameters in the Black-Scholes formula.

5.1 Maximum Likelihood

The Maximum likelihood method of estimating parameters is a method based on observations. We deal with stock prices and want to derive estimates for the geometric Brownian motion used to model these prices. That means that we have plenty of observations, hence the maximum likelihood method is a good choice.

Definition 5.1.1 (Maximum Likelihood Function)

Consider $X_1, X_2, \cdots, X_n$ random variables with realisations $X_1 = x_1, X_2 = x_2, \cdots X_n = x_n$. Suppose these random variables follow a distribution characterized by a parameter θ. Suppose that these random variables have a joint density or mass function $f_\theta(x_1, x_2, \cdots, x_n)$. Then we define the *Maximum likelihood function* by

$$L(\theta) = f_\theta(x_1, x_2, \cdots, x_n),$$

Which is precisely the chance of these realisations occuring. The maximum of this function is called the Maximum Likelihood Estimator (M.L.E). Hence $\hat{\theta}$ for which $L(\theta)$ is maximal, is precisely the value for which this realisation has the biggest chance of occuring.

Remark 5.1.1

Suppose that random variables $X_1, \cdots, X_n$ in Definition 6.1.1 are identically independtly distributed (i.i.d), then we can write

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i),$$

where $f_\theta(x)$ is the density function or mass function of X. Sometimes this function could be hard to differentiate. Hence often we refer to the log-likelihood function

$$l(\theta) = \log\left(\prod_{i=1}^{n} f_\theta(x_i)\right) = \sum_{i=1}^{n} \log(f_\theta\,(x_i)).$$

This definition is perfectly fine because the natural logarithm is a monotonically increasing function, hence if we calculate the maximum of $l(\theta)$ this gives us the same value as if we were to do it for $L(\theta)$.

Example 5.1.1

Consider $X_1, \cdots, X_n$ i.i.d $\sim \mathcal{N}(\mu, \sigma^2)$, each with realisation $x_1, \cdots, x_n$. Then we can calculate the Maximum Likelihood Estimates for μ and σ. The density function is given by

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

Hence the log-likelihood function is given by

$$\begin{aligned} l(\mu, \sigma^2) &= \sum_{i=1}^{n} \log\left(f_{\mu,\sigma^2}(x_i)\right) \\ &= \sum_{i=1}^{n} -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left(\sigma^2\right) - \frac{(x_i-\mu)^2}{2\sigma^2}. \\ &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned}$$

If we now calculate its maximum by differentiation we find

$$\begin{aligned}\frac{\partial l}{\partial \mu} &= \frac{\partial}{\partial \mu}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\right) \\ &= \sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2}\end{aligned}.$$

If we set this equal to zero

$$\begin{aligned}\sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2} &= 0, \\ \sum_{i=1}^{n} x_i - n\mu &= 0.\end{aligned}$$

Hence

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Since $\hat{\mu}$ is the only extremum and the second derivative with respect to μ is negative this should be a maximum. On the otherhand

$$\begin{aligned}\frac{\partial l}{\partial \sigma} &= \frac{\partial}{\partial \sigma}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\right) \\ &= -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^3}\end{aligned}.$$

Hence we solve

$$-n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 = 0.$$

Which results in

$$\hat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n}.$$

We see that $\hat{\mu}$ and $\hat{\sigma}$ indeed maximize $l(\mu, \sigma^2)$, then since the pair $(\hat{\mu}, \hat{\sigma}^2)$ result in a maximum of $l(\mu, \sigma^2)$ we conclude that the Maximum Likelihood Estimates are given by

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \hat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n}.$$

5.2 Estimating the Volatility

Recall from Section 4.5 that the solution to the Black-Scholes-Merton Equation is given by

$$c(t,x) = x\Phi(\beta_2) - e^{-r\tau}K\Phi(\beta_1),$$

where $\Phi(z) = \mathbb{P}(Z < z)$, $Z \sim \mathcal{N}(0,1)$ and $\beta_2 = \beta_1 + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K}) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K}) + \left(r + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$. The parameters in the solution are t, x, σ, K, r. We consider the risk free interest rate r to be known, hence the parameter that we need to estimate is the volatility σ. Recall that we assume that the stock prices are modelled by a Geometric Brownian Motion, and that the stock price is given by

$$S(t) = S(0)e^{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t}.$$

Hence we can write

$$\log(S(t)) = \log(S(0)) + \sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right) t,$$

where we must have that $S(0) \neq 0$. Hence for $u < t$ we can write that

$$\log(S(t)) - \log(S(u)) = \sigma\left(W(t) - W(u)\right) + \left(\alpha - \frac{1}{2}\sigma^2\right)(t-u).$$

Since $\{W(t) : t \geq 0\}$ is a Brownian motion, we know that $(W(t) - W(u)) \sim \mathcal{N}(0, (t-u))$. Hence

$$\log(S(t)) - \log(S(u)) \sim \mathcal{N}\left(\left(\alpha - \frac{1}{2}\sigma^2\right)(t-u), \sigma^2(t-u)\right).$$

Now suppose we know the values $S(t_i)$ for $1 \leq i \leq n$. For example if we look at the history of some stock. Take the value of AEX-INDEX at Yahoo! Finance[1], where we can consider the weekly closing price for the year 2018. Hence we have 53 different stock prices recorded every week for the year 2018. We define

$$X(t_i) := \log(S(t_i)) - \log(S(t_{i-1})).$$

If we let $t_i - t_{i-1} = \Delta t$, $\forall i$. We have that $X(t_i) \sim \mathcal{N}\left(\left(\alpha - \frac{1}{2}\sigma^2\right)\Delta t, \sigma^2 \Delta t\right)$. If we assume all the $X(t_i)$ are independent. Let $x(t_i)$ be the realisation of $X(t_i)$ at time t_i, then we know by Example 5.1.1

$$\left(\hat{\alpha} - \frac{1}{2}\hat{\sigma}^2\right)\Delta t = \frac{1}{n}\sum_{i=1}^{n} x(t_i), \quad \hat{\sigma}^2 \Delta t = \frac{\sum_{i=1}^{n}\left(x(t_i) - \frac{1}{n}\sum_{i=1}^{n} x(t_i)\right)^2}{n}.$$

Hence we can solve this to find the values of $\hat{\sigma}^2$ and $\hat{\alpha}$. Note that we must have that $X(t_i)$ are all independent, else we cannot use Example 5.1.1. We cannot know this for sure, but we can calculate the autocorrelation, and if it does not exhibit significant lags the assumption of independence is justifiable as shown in [4].

5.3 Example for AEX

Consider the weekly closing prices of the AEX, they can be found at, for example, Yahoo! Finance. We get data similar to Table 1 shown in Appendix A.2. Then using this data we can calculate the log differences of every week. So we calculate all the $x(t_i)$ where $1 \leq i \leq 52$. We use MATLAB to plot the Autocorrelation and Partial Autocorrelation Function of the $x(t_i)$, this can be done by using their MATLAB functions. This results in the following figure

[1] https://finance.yahoo.com/quote/%5EAEX/history?p=%5EAEX

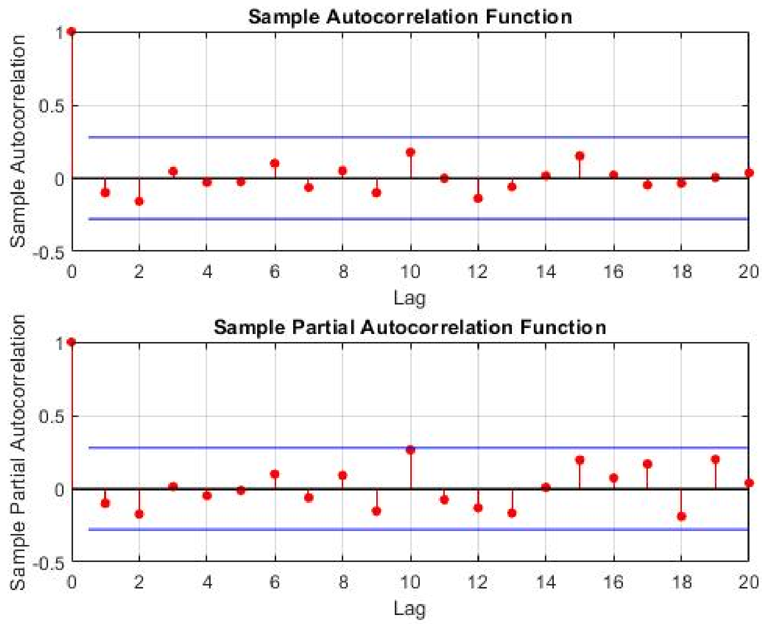

Figure 1: Autocorrelation Functions in MATLAB

If the value does not exceed the blue horizontal line we say that the value is not significant from 0. Hence we see that the there are no significant lags hence the assumption that the log differences are Normally distributed is justifiable. We have 1 observation every week one year long. Hence $\Delta t = \frac{1}{52}$ and

$$\hat{\sigma}^2 = \sum_{i=1}^{52}(x(t_i) - \frac{1}{52}\sum_{i=1}^{52} x(t_i))^2,$$

where $x(t_i) = \log(S(t_i)) - \log(S(t_{i-1}))$ and $S(t_i)$ denotes the closed stock price at time t_i which can be found in the table where we referred to earlier. Then $\hat{\sigma}$ can be found by taking the square root. Since we now know $\hat{\sigma}^2$ we can use this to calculate $\hat{\alpha}$. Hence

$$\hat{\alpha} = \sum_{i=1}^{52} x(t_i) + \frac{1}{2}\hat{\sigma}^2.$$

We use the Python code in Appendix A.3 and run the function Parameter_estimation(). This function has 2 input values, 'data' and 'invdt'. The data variable is just the .csv file we download from Yahoo! Finance containing our closing prices. The 'invdt' variable is $\frac{1}{\Delta t}$. If we run this program on the dataset given in Appendix A.2 we find[2]

$$\hat{\alpha} = -0.12418951308832828, \quad \hat{\sigma} = 0.144129205548125.$$

Hence we have estimated the parameters in the Geometeric Brownian Motion used to model the stock prices. To see if these estimates make sense we can generate 20000 data sets containing 52 normally distributed random variables with mean and standard deviation as given by our python program, i.e. the mean and standard deviation estimated by the dataset. Then we proceed to calculate the mean and standard deviation for each of these data sets. Afterwards we use these to calculate $\hat{\alpha}$ and $\hat{\sigma}$. Hence this will result in 20000 values for $\hat{\alpha}$ and $\hat{\sigma}$, we can do this by running the code in Appendix A.4 in MATLAB[3]. If we run

[2]If the data file is called 'AEX.csv' the line 'Parameter_estimation('AEX.csv', 52)' executes the code.
[3]The code used can also be found in [3] Chapter 7

```
GBMSIM(-0.0025880024429672135,0.01998712464983283,20000,52)
```

We get the following figure

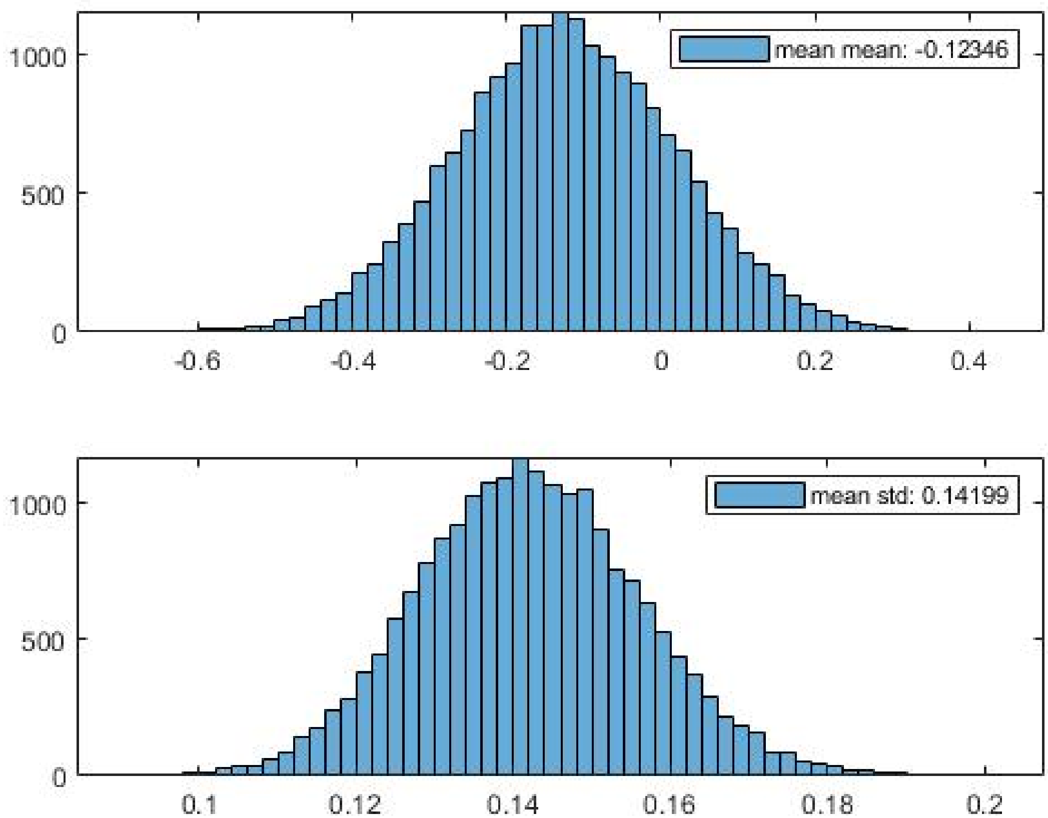

Figure 2: Distribution of the mean (top) and standard deviation (bottom) of 20000 datasets

We conclude that these values which we found by simulation, are rather close to the theoretical values we found using the maximum likelihood method.

6 Conclusion

In Section 4.3 we assumed that the price of a stock

$$S(t) = S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t},$$

or in differential form

$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t).$$

We then formulated the portfolio value in terms of an investment in a money market account and the capital formed by the amount of stock held and the stock price, which resulted in

$$dX(t) = \Delta(t)dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt.$$

We then derived a partial differential equation

$$\begin{cases} \frac{\partial}{\partial t}c(t,x) + rS(t)\frac{\partial}{\partial x}c(t,x) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,x) = rc(t,x) \\ c(T,x) = (x-K)^+ \end{cases},$$

where

$$\frac{\partial}{\partial t}c(t,x) + rS(t)\frac{\partial}{\partial x}c(t,x) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,x) = rc(t,x),$$

is called the Black-Scholes-Merton equation. Then using Girsanov's Theorem we created a risk-neutral measure. This measure was used to derive a solution to the Black-Scholes equation and we found

$$c(t,x) = x\Phi(\beta_2) - e^{-r\tau}K\Phi(\beta_1),$$

where $\Phi(z) = \mathbb{P}(Z < z)$, $Z \sim \mathcal{N}(0,1)$ and $\beta_2 = \beta_1 + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K})+\left(r-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K})+\left(r+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$. Since we assume the risk-free interest rate to be constant in $c(t,x)$ we noticed that σ is the only parameter for which we do not immediatly have a value. Hence we used the Maximum likelihood method on the log differences from the stock prices to derive an estimate for σ. This was possible because we noticed that

$$\log(S(t)) - \log(S(u)) \sim \mathcal{N}\left(\left(\alpha - \frac{1}{2}\sigma^2\right)(t-s), \sigma^2(t-u)\right).$$

Hence we found that

$$\hat{\sigma}^2\Delta t = \frac{\sum_{i=1}^n (x(t_i) - \frac{1}{n}\sum_{i=1}^n x(t_i))^2}{n},$$

and thus found an estimate to for the value of σ which can be used in the Black-Scholes formula. We also used the AEX closing prices to construct an example to see this estimation in practice. Afterwards we used a simulation to see if our estimate was plausible, and we found that our estimate was close to the simulated answer.

7 References

[1] Cohn, Donald L. Cohn *Measure Theory.* Birkhäuser; 2013, 978-1-4614-6955-1

[2] Shreve, Steven E. Shreve *Stochastic Calculus for Finance II Continous-Time Models*, Springer, 2008, 978-0-387-40101-0

[3] Kubrom Hisho Teka *Parameter Estimation Of The Black-Scholes-Merton Model*, `https://krex.k-state.edu/dspace/bitstream/handle/2097/15669/KubromTeka2013rev.pdf;sequence=3`

[4] Brigo, D., Dalessandro, A., Neugebauer, M., Triki, F. *A Stochastic Processes Toolkit for Risk Management*, `https://arxiv.org/pdf/0812.4210.pdf`

[5] Choe, G.H. *Stochastic Analysis for Finance with Simulations*, Springer, 2016, 978-3-319-25589-7

[6] Dekking, F.M., Kraaikamp, C., Lopuhaä, H.P., Meester, L.E.*A Modern Introduction to Probability and Statistics*, Springer, 2010, 978-1-84996-952-9

[7] Rice, J.A.*Mathematical Statistics and Data Analysis*, Brooks/Cole, 2007, 978-0-495-11868-8

[8] Ermogenous, A.*Brownian Motion and its Applications in the Stock Market*, `https://ecommons.udayton.edu/mth_epumd/15/`

Appendix A

A.1 Measure Defined by the Integral of a Postive Function

Suppose we have a measure space $(\Omega, \mathcal{F}, \mu)$ and that $f : \Omega \to [0, \infty)$, if we define $\mathbb{Q}(A) := \int_A f d\mu$, then $\mathbb{Q}(A) : \mathcal{F} \to [0, \infty)$ is a measure.

Proof
We are going to prove that $\mathbb{Q}(A)$ statisfies the conditions of a measure. We know that

$$\mathbb{Q}(\emptyset) = \int_\emptyset f d\mu = \int_\Omega f \mathbb{I}_\emptyset d\mu = \int_\Omega 0 d\mu = 0.$$

Suppose $\mathcal{B}$ is a countable collection of disjoint subsets of $\mathcal{A}$ then we know that $\cup\mathcal{B} = \cup_i B_i$ with B_i disjoint for all i. Define

$$f_n := f\mathbb{I}_{\cup_{i=1}^n B_i}.$$

Because all B_i are disjoint we know that $\mathbb{I}_{\cup_{i=1}^n B_i} = \sum_{i=1}^n \mathbb{I}_{B_i}$. So there follows

$$\int_\Omega f_n d\mu = \int_\Omega f\mathbb{I}_{\cup_{i=1}^n B_i} d\mu = \int_\Omega f(\sum_{i=1}^n \mathbb{I}_{B_i}) d\mu = \sum_{i=1}^n \int_\Omega f\mathbb{I}_{B_i} d\mu = \sum_{i=1}^n \int_{B_i} f d\mu.$$

We also see that $\{f_n\}$ is a monotone increasing sequence of functions because $f : \Omega \to [0, \infty]$ and

$$f_{n+1} = f\mathbb{I}_{\cup_{i=1}^{n+1} B_i} = f\mathbb{I}_{\cup_{i=1}^n B_i} + f\mathbb{I}_{B_{n+1}} = f_n + f\mathbb{I}_{B_{n+1}}.$$

Note that also holds that $\lim_{n\to\infty} f_n = f\mathbf{1}_{\cup_{i=1}^\infty B_i} = f\mathbf{1}_{\bigcup\mathcal{B}}$. Then by monotone convergence we know that

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \sum_{i=1}^\infty \int_{B_i} f d\mu = \int_\Omega \lim_{n\to\infty} f_n d\mu = \int_\Omega f\mathbb{I}_{\bigcup\mathcal{B}} d\mu = \int_{\bigcup\mathcal{B}} f d\mu.$$

Thus follows that

$$\mathbb{Q}(\bigcup\mathcal{B}) = \sum_{i=1}^\infty \mathbb{Q}(B_i).$$

Hence

$$\mathbb{Q}(A) := \int_A f d\mu,$$

is a measure.

A.2 Table With Data

Table 1: AEX closing value

t_i	Date	Close
0	2017-12-31	558.159973
1	2018-01-07	561.099976
2	2018-01-14	569.299988
3	2018-01-21	566.789978
4	2018-01-28	550.080017
5	2018-02-04	518.330017
6	2018-02-11	532.270020
7	2018-02-18	534.090027

8	2018-02-25	518.719971
9	2018-03-04	537.140015
10	2018-03-11	536.919983
11	2018-03-18	521.450012
12	2018-03-25	529.520020
13	2018-04-01	539.289978
14	2018-04-08	548.049988
15	2018-04-15	550.380005
16	2018-04-22	554.940002
17	2018-04-29	555.700012
18	2018-05-06	562.270020
19	2018-05-13	567.030029
20	2018-05-20	562.770020
21	2018-05-27	559.179993
22	2018-06-03	560.030029
23	2018-06-10	561.710022
24	2018-06-17	560.340027
25	2018-06-24	551.679993
26	2018-07-01	553.619995
27	2018-07-08	560.119995
28	2018-07-15	572.200012
29	2018-07-22	576.239990
30	2018-07-29	572.289978
31	2018-08-05	562.979980
32	2018-08-12	552.950012
33	2018-08-19	560.289978
34	2018-08-26	558.419983
35	2018-09-02	538.510010
36	2018-09-09	540.530029
37	2018-09-16	549.789978
38	2018-09-23	549.619995
39	2018-09-30	539.510010
40	2018-10-07	516.289978
41	2018-10-14	525.169983
42	2018-10-21	507.519989
43	2018-10-28	521.799988
44	2018-11-04	529.549988
45	2018-11-11	522.429993
46	2018-11-18	513.849976
47	2018-11-25	519.369995
48	2018-12-02	503.980011
49	2018-12-09	506.529999
50	2018-12-16	484.809998
51	2018-12-23	484.170013
52	2018-12-30	487.880005

A.3 Python Code

```
def Parameter_estimation(data, invdt):
```

```
import csv
import math
with open(data, mode='r') as csv_file:
    csv_reader = csv.DictReader(csv_file)
    dataset = []
    for row in csv_reader:
        dataset.append(row["Close"])
def mean(A):
    dummy = 0
    for i in range(len(A)-1):
        dummy = dummy + math.log(float(A[i+1]))-math.log(float(A[i]))
    dummy = dummy/(len(A)-1)
    return dummy
def var(A):
    dummy = 0
    for i in range(len(A)-1):
        dummy = dummy + (math.log(float(A[i+1]))-math.log(float(A[i]))-m)**2
    print(dummy)
    dummy = dummy/(len(A)-1)
    return dummy
m = mean(dataset)
v = var(dataset)
sigma = math.sqrt(v*invdt)
std = math.sqrt(v)
alpha = 1/2*v*invdt+invdt*m
print('mean= '+ str(m),'sigma= ' + str(sigma),
'std= ' + str(samplestd),
'variance= ' +str(v),'alpha= '+ str(alpha))
```

A.4 MATLAB Code

```
function GBMSIM(samplemean,samplestd,datasetsize,deltainv)
Z = normrnd(samplemean , samplestd , datasetsize, deltainv);
R = Z';
m = sum(R)/(deltainv);
p = m';
for k = 1:datasetsize;
        for j = 1:(deltainv);
                D(k,j) = (Z(k,j)-p(k,1)).^2;
        end;
end;
V = sum(D')/(deltainv);
V=V';
t = 1/(deltainv);
S = sqrt(V/t);
A = 0.5*S.^2+(1/t)*p;
mymean = mean(A);
mystd = mean(S);
subplot(2,1,1);
histogram(A)
hold on
legend("mean mean: " + string(mymean))
```

```
hold off
subplot(2,1,2);
histogram(S)
hold on
legend("mean std: " + string(mystd))
hold off
end
```