MARTINGALE REPRESENTATION, GIRSANOV'S THEOREM, AND A GENERALIZATION OF THE BLACK-SCHOLES MODEL

A Thesis

Presented to

the Faculty of the College of Graduate Studies

Tennessee Technological University

by

Seth Agee

In Partial Fulfillment

of the Requirements of the Degree

MASTER OF SCIENCE

Mathematics

May 2021

ABSTRACT OF A THESIS

MARTINGALE REPRESENTATION, GIRSANOV'S THEOREM, AND A GENERALIZATION OF THE BLACK-SCHOLES MODEL

Seth Agee

Master of Science in Mathematics

In their influential work in options pricing, Black and Scholes derived the theoretical call option premium price and showed that a strategy for eliminating risk to the contract seller exists. The fundamental assumption was that the stock price follows a geometric Brownian motion, thus allowing the premium to be calculated by solving a partial differential equation. We give an alternative derivation, harnessing the power of results in stochastic calculus. Under a change of measure on the underlying probability space, we can transform the discounted stock price into a martingale, then apply the martingale representation theorem to yield a unique representation which gives rise to a replicating strategy. Using properties of Brownian motion, the premium price can then be derived. We make further generalizations on the assumptions of the model and achieve similar results.

TABLE OF CONTENTS

CERTIFICATE OF APPROVAL OF THESIS

MARTINGALE REPRESENTATION, GIRSANOV'S THEOREM, AND A GENERALIZATION OF THE BLACK-SCHOLES MODEL

by

Seth Agee

Graduate Advisory Committee:

____

Motoya Machida, Chair Date

____

Damian Kubiak Date

____

David Smith Date

Approved for the Faculty:

____

Mark Stephens, Dean Date
College of Graduate Studies

CHAPTER 1

STOCHASTIC CALCULUS

In the early 1800's, Robert Brown noticed some properties of particles suspended in water. First, the paths that these particles took seemed to be highly irregular, having no derivative at any given point. Second, the paths of the particles appeared to be independent of one another. These paths would continue to be studied by great minds (including Albert Einstein), and many remarkable properties of this so called "Brownian motion" would be discovered.

From the stock market to rocketry, this strange behavior has had an incredibly large impact on the world of mathematics. Of particular interest in this area of research will be differential equations involving some "white noise," a source of randomness which influences an ordinary differential equation. The source of this randomness will come from Brownian motion.

Before we are able to study these differential equations, much background is needed, and a mathematically rigorous definition of the above concepts is necessary. This chapter will cover first general measure theoretic probability, the stochastic integral, and finally Brownian motion and stochastic differential equations.

1.1 Measure Theoretic Probability Theory

This section gives an outline of the modern language of probability theory. We discuss common definitions, results, and give some basic examples. The first goal we must achieve is determining exactly which events from an experiment we are able to assign probabilities to.

Definition 1.1.1. Let Ω be a set, and $\mathcal{F}$ a collection of subsets of Ω. If $\mathcal{F}$ satisfies

1. $\varnothing, \Omega \in \mathcal{F}$;

2. If $A \in \mathcal{F}$ then $\Omega \setminus A \in \mathcal{F}$;

3. If $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$,

then we call $\mathcal{F}$ a *σ-algebra on* Ω.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*, and sets in $\mathcal{F}$ are called *measurable sets*. We will call Ω the *sample space*, any $\omega \in \Omega$ an *outcome*, and any $E \in \mathcal{F}$ an *event*.

Note that using the definition above, if $\mathcal{F}$ is a σ-algebra, it is also closed under the operation of countable intersections.

From a probabilistic point of view, we think of the sample space as the set of all possible outcomes of an experiment, and we view $\mathcal{F}$ as a set consisting of possible events. As it turns out, measurable sets are exactly those for which we are able to assign probabilities to. We now need a method of assigning probabilities to certain events, and in order to do this we will introduce the measure.

Definition 1.1.2. Let $\mathcal{F}$ be a σ-algebra on Ω. We call a function $\mu : \mathcal{F} \to [0, \infty)$ a *measure on* $\mathcal{F}$ if it satisfies

1. $\mu(\varnothing) = 0$;

2. For any sequence of mutually disjoint sets $A_1, A_2, \ldots \in \mathcal{F}$ we have $\mu(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$.

A triple $(\Omega, \mathcal{F}, \mu)$ consisting of a set Ω, a σ-algebra (on Ω) $\mathcal{F}$, and a measure μ (with respect to $\mathcal{F}$) is called a *measure space*.

Let $(\Omega, \mathcal{F})$ be a measurable space. A measure $P : \mathcal{F} \to [0, 1]$ is called a *probability measure* if it satisfies $P(\Omega) = 1$, and we call $(\Omega, \mathcal{F}, P)$ a *probability space*.

For a simple example, let Ω be the sample space of an experiment where one is flipping a fair coin. Then we could define $\Omega = \{H, T\}$, $\mathcal{F} = \{\varnothing, \{H\}, \{T\}, \Omega\}$, and we can define the measure $P : \mathcal{F} \to [0,1]$ by

$$P(E) = \begin{cases} 0 & \text{if } E = \varnothing \\ 0.5 & \text{if } E = \{H\} \text{ or } E = \{T\} \\ 1 & \text{if } E = \Omega. \end{cases}$$

It is easily checked that $\mathcal{F}$ is a σ-algebra and that P is a probability measure.

We now have a formal way to assign probabilities to events in an experiment, and so questions such as "What is the probability of rolling two 6's when throwing a pair of dice?" can be answered. However, questions such as "What is the chance that the numbers on my pair of dice will add to 7?" do not make sense yet. This latter type of question involves a random variable, which is a mapping from our sample space to some other set of outcomes. We now make this concept formal.

Definition 1.1.3. Let $(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$ be measurable spaces. For a mapping $T : \Omega \to \Omega'$, we say T is *$\mathcal{F}/\mathcal{F}'$-measurable* if $T^{-1}A' = \{\omega \in \Omega : T(\omega) \in A'\} \in \mathcal{F}$ for each $A' \in \mathcal{F}'$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A k-dimensional *random variable* is a function $\mathbf{X} : \Omega \to \mathbb{R}^k$ such that

$$\mathbf{X}^{-1}(U) := \{\omega \in \Omega : \mathbf{X}(\omega) \in U\} \in \mathcal{F}$$

for all open sets $U \subseteq \mathbb{R}^k$ or equivalently all *Borel sets* (those that can be formed from open sets and closed under countable union, countable intersection, and relative complement).

The collection of all Borel sets on a topological space Y forms a σ-algebra called the *Borel σ-algebra* on Y, denoted $\mathcal{B}(Y)$, and is the smallest σ-algebra containing all open sets. Thus, a random variable is an $\mathcal{F}/\mathcal{B}(\mathbb{R}^k)$-measurable function, but often we will simply say that $\mathbf{X}$ is $\mathcal{F}$-measurable

Later, it will be useful to think about the smallest or coarsest σ-algebras that make certain events or random variables measurable. These have a useful interpretation as those containing all relevant information about a set or event.

Definition 1.1.4. Given any family $\mathcal{U}$ of subsets of Ω, there is a smallest σ-algebra $\sigma(\mathcal{U})$ containing $\mathcal{U}$, namely

$$\sigma(\mathcal{U}) = \bigcap\{\mathcal{H} : \mathcal{H} \text{ is a } \sigma\text{-algebra of } \Omega, \mathcal{U} \subset \mathcal{H}\},$$

and we call $\sigma(\mathcal{U})$ the *σ-algebra generated by $\mathcal{U}$*.
If $\mathbf{X} : \Omega \to \mathbb{R}^k$ is any function, the the *σ-algebra generated by* $\mathbf{X}$, denoted $\sigma(\mathbf{X})$, is the smallest σ-algebra on Ω containing the sets $\mathbf{X}^{-1}(U)$ where $U \in \mathcal{B}(\mathbb{R}^k)$. This is the smallest σ-algebra that makes $\mathbf{X}$ measurable.
We say that a measure μ is a *Borel measure* if it is defined on the Borel σ-algebra.

The σ-algebra generated by $\mathbf{X}$ is interpreted as containing all relevant information about $\mathbf{X}$. That is, it gives us the values of $\mathbf{X}(\omega)$ for $\omega \in \Omega$. This property of containing relevant information is given by a theorem of Doob and Dynkin and is presented below. We give necessary lemmas then the result.

Lemma 1.1.5. *If T is $\mathcal{F}/\mathcal{F}'$-measurable and T' is $\mathcal{F}'/\mathcal{F}''$-measurable then $T' \circ T$ is $\mathcal{F}/\mathcal{F}''$-measurable.*

Lemma 1.1.6. *Let $(f_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{F}$-measurable functions on Ω. The ω-set where $(f_n(\omega))_{n=1}^{\infty}$ converges lies in $\mathcal{F}$.*

Lemma 1.1.7. *Let $(f_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{F}$-measurable functions on Ω. If $\lim_{n\to\infty} f_n$ exists everywhere then it is $\mathcal{F}$-measurable.*

For proof of the above lemmas, see Chapter 13 of [2].

Theorem 1.1.8. (The Doob-Dynkin theorem) *Let* $\mathbf{X} = (X_1, \dots, X_k)$ *be a random vector on the measurable space* $(\Omega, \mathcal{F})$*. Then*

1. *The* σ*-algebra* $\sigma(\mathbf{X})$ *consists exactly of sets of the form* $\{\mathbf{X} \in H\}$ *for* $H \in \mathcal{B}(\mathbb{R}^k)$*;*

2. *In order that a random variable* Y *be* $\sigma(\mathbf{X})$*-measurable it is necessary and sufficient that there exist a measurable map* $f : \mathbb{R}^k \to \mathbb{R}$ *such that* $Y(\omega) = f(X_1(\omega), \dots, X_k(\omega))$ *for all* ω*.*

Proof. 1. Let $\mathcal{G}$ be the class of sets of the form $\{\mathbf{X} \in H\}$ for $H \in \mathcal{B}(\mathbb{R}^k)$. Then $\mathcal{G}$ is a σ-algebra since $\{\mathbf{X} \in \varnothing\} = \varnothing$, $\{\mathbf{X} \in \mathbb{R}^k\} = \Omega$, and for arbitrary index I,

$$\bigcup_{i \in I} \{\mathbf{X} \in H_i\} = \{\mathbf{X} \in \cup_{i \in I} H_i\},$$

and

$$\{\mathbf{X} \in H\}^c = \{\mathbf{X} \notin H\} = \{\mathbf{X} \in H^c\}.$$

By definition $\sigma(\mathbf{X})$ is the minimum σ-algebra containing $\mathcal{G}$. Since $\mathcal{G}$ is a σ-algebra, we must have $\mathcal{G} = \sigma(\mathbf{X})$, which proves 1.

2. For sufficiency, note that X is $\sigma(\mathbf{X})/\mathcal{B}(\mathbb{R}^k)$-measurable and f is $\mathcal{B}(\mathbb{R}^k)$-measurable. Then by Lemma 1.1.5 $f(\mathbf{X}) = f(X_1(\omega), \dots, X_k(\omega))$ is $\sigma(\mathbf{X})$-measurable.
For the necessity, we will first suppose that Y is a simple random variable, $\sigma(\mathbf{X})$-measurable, taking values in $\{y_1, \dots, y_m\}$. Then the sets $A_i := \{\omega : Y(\omega) = y_i\}$ are in $\sigma(\mathbf{X})$, and so by the first part must be of the form $\{\omega : \mathbf{X}(\omega) \in H_i\}$ for some $H_i \in \mathcal{B}(\mathbb{R}^k)$. Put $f = \sum_i y_i I_{H_i}$. Then f is simple (or can be rewritten so that it is simple) and is therefore $\mathcal{B}(\mathbb{R}^k)$-measurable. Noting that each A_i is disjoint since Y is simple, if

$$A_i = \{\omega : \mathbf{X}(\omega) \in H_i\}$$

then $\omega \in A_i$ if and only if $\mathbf{X}(\omega) \in H_i$. Thus, $f(\mathbf{X}(\omega)) = y_i = Y(\omega)$.

Now we consider the general case and let Y be a random variable. Then there exists a sequence $(Y_n)_{n=1}^{\infty}$ of simple random variables increasing pointwise to Y. By the work above, for each n there is a measurable function $f_n : \mathbb{R}^k \to \mathbb{R}$ such that $Y_n(\omega) = f_n(X(\omega))$ for all ω. The x-set M for which $(f_n(x))$ converges lies in $\mathcal{B}(\mathbb{R}^k)$ by Lemma 1.1.6. Let

$$f(x) = \begin{cases} \lim_{n\to\infty} f_n(x) & \text{for } x \in M \\ 0 & \text{otherwise.} \end{cases}$$

Then $f = \lim_{n\to\infty} f_n I_M$, and since $f_n I_M$ is measurable, Lemma 1.1.7 gives us that f is measurable. For each ω,

$$f_n(\mathbf{X}(\omega)) = Y_n(\omega) \to Y(\omega).$$

Then $\mathbf{X}(\omega)$ lies in M and so

$$f_n(\mathbf{X}(\omega)) \to f(\mathbf{X}(\omega))$$

for all ω. Thus, by uniqueness of limits, we indeed have that $Y(\omega) = f(\mathbf{X}(\omega))$ for all ω. $\square$

So, the possible values of any $\sigma(\mathbf{X})$-measurable random variable Y can be determined from the values of $\mathbf{X}$.

As an example, let $(\Omega, \mathcal{F}, P)$ be a probability space, and $E \in \mathcal{F}$. Define the mapping $X : \Omega \to \mathbb{R}$ by

$$X(\omega) = \begin{cases} 0 & \text{if } \omega \notin E \\ 1 & \text{if } \omega \in E. \end{cases}$$

X is a random variable which we call the *indicator function of* E, denoted by I_E. If $U \in \mathcal{B}(\mathbb{R})$, we have

$$X^{-1}(U) = \begin{cases} \varnothing & \text{if } 1 \notin U,\, 0 \notin U \\ E & \text{if } 1 \in U,\, 0 \notin U \\ \Omega \setminus E & \text{if } 1 \notin U,\, 0 \in U \\ \Omega & \text{if } 1 \in U,\, 0 \in U. \end{cases}$$

Then $\sigma(X) = \{\varnothing, E, \Omega \setminus E, \Omega\}$, and is easily seen to be the smallest σ-algebra that makes X measurable. In particular, $\sigma(X)$ tells us all the information about the value of $X(\omega)$.

The next definition will be extremely important for the sections to come. We have the concept of random variable, and now we examine collections of random variables. This will give us a way of describing a system, evolving over time, subject to random fluctuations.

Definition 1.1.9. A *stochastic process* is a parametrized collection of random variables

$$(\mathbf{X}(t,\omega))_{t\in T} = (\mathbf{X}_t(\omega))_{t\in T}$$

defined on $(\Omega, \mathcal{F}, P)$ and assuming values in $\mathbb{R}^k$. For each $\omega \in \Omega$, the mapping $t \mapsto \mathbf{X}(t,\omega)$ is the corresponding *sample path.*
The σ-algebra generated by $\mathbf{X}$ up to t is given by $\sigma(\mathbf{X}_s, 0 \le s \le t)$, generally denoted by $\mathcal{F}_t$.

A stochastic process is both a function of T and Ω, though the latter argument is generally omitted. We will then take the convention of writing $(\mathbf{X}_t(\omega))_{t\in T} = (\mathbf{X}_t)_{t\in T}$, and for a fixed t we generally write $\mathbf{X}_t$ or $\mathbf{X}(t)$, omitting reference to ω.

We usually view t as a specific point in time, and for our purposes set $T = [0, \infty)$. Using our interpretation from before, $\mathcal{F}_t$ is the σ-algebra which contains all relevant information about the process up to the time t.

First, a simple example where T is discrete. Let $X_1, X_2, \ldots$ be independent (see Definition 1.1.16) random variables taking the value 1 or -1 with equal probability $\frac{1}{2}$. Then we can define $Y_n = \sum_{i=1}^{n} X_i$ for $n = 0, 1, 2, \ldots$. Note then that $Y_0 = 0$. The sequence $(Y_n)_{n=0}^{\infty}$ is called a random walk, and is one of the simplest examples of a stochastic process.

Below are simulated sample paths for $n = 10$ and $n = 100$.

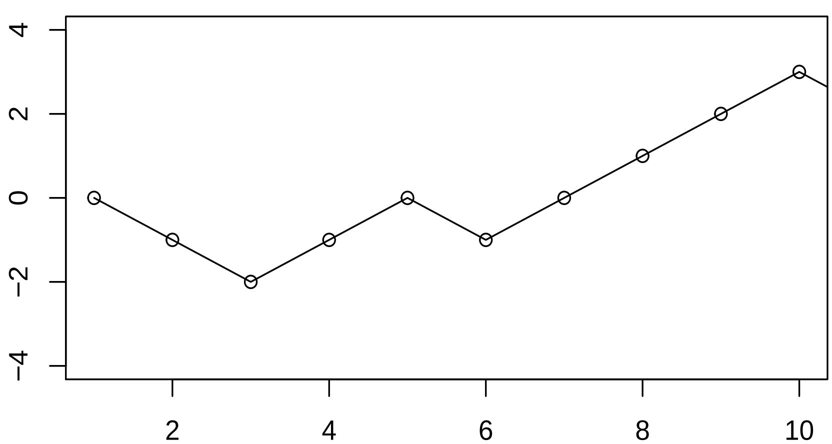

Figure 1.1: Sample path for a random walk with $n = 10$.

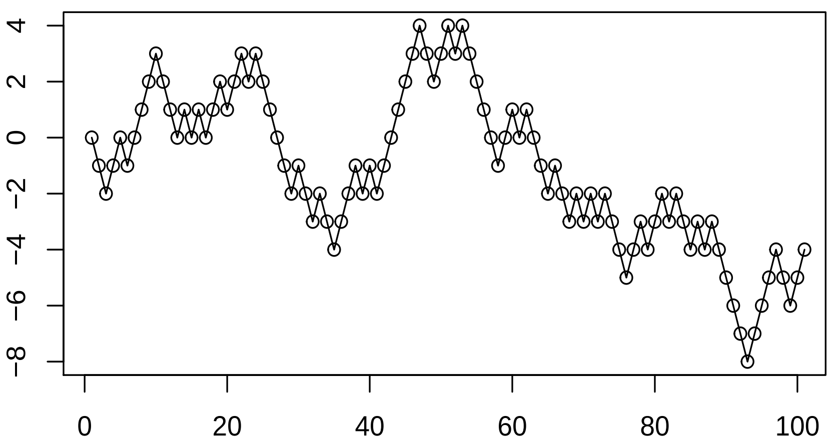

Figure 1.2: Sample path for a random walk with $n = 100$.

In probability theory, we commonly want to know the average or expected value of a random variable. In the discrete case, we would multiply each value by its probability of occuring, then take the sum of these products. We will generalize this idea in order to find the expected value of a much larger class of random variables. First, we need to define the integral of a random variable with respect to a measure, and so below we outline the common construction of the Lebesgue integral.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We say that a random variable $\varphi : \Omega \to [0, \infty)$ is *simple* if it has the representation $\varphi = \sum_{i=1}^{k} a_i I_{E_i}$, where I is the indicator function, $E_i \in \mathcal{F}$ are disjoint measurable sets, and a_i are distinct constants. The integral of the simple function with the above representation is defined as $\int_\Omega \varphi dP = \sum_{i=1}^{k} a_i P(E_i)$.

We then extend the definition to nonnegative random variables. For $X : \Omega \to [0, \infty)$, we define the (Lebesgue) integral of X with respect to P by $\int_\Omega X dP = \sup\{\int_\Omega \varphi dP : 0 \leq \varphi \leq X, \varphi \text{ simple}\}$.

For random variables $X : \Omega \to \mathbb{R}$, define $X^+(\omega) = \max(X(\omega), 0)$ and $X^-(\omega) = \max(-X(\omega), 0)$. Then we define the (Lebesgue) integral of X with respect to P by $\int_\Omega X dP = \int_\Omega X^+ dP - \int_\Omega X^- dP$, provided that either $\int_\Omega X^+ dP < \infty$ or $\int_\Omega X^- dP < \infty$.

Whenever $\mathbf{X} : \Omega \to \mathbb{R}^k$ is a vector-valued random variable with $\mathbf{X} = (X_1, X_2, \ldots, X_k)$, we define $\int_\Omega \mathbf{X} dP = (\int_\Omega X_1 dP, \int_\Omega X_2 dP, \ldots, \int_\Omega X_k dP)$.

We now look at another useful type of integration.

Definition 1.1.10. Let F be a real valued function on $\mathbb{R}$. We say that F is *right continuous* if

$$F(a) = F(a+) = \lim_{x \to a+} F(x)$$

for all $a \in \mathbb{R}$. For an increasing, right-continuous function, let μ_F be the Borel measure determined by the relation $\mu_F((a, b]) = F(b) - F(a)$. In general, we say that $F : \mathbb{R} \to \mathbb{R}$ is of *bounded variation* if

$$\lim_{x \to \infty} \sup\{\sum_{i=1}^{n} |F(x_j) - F(x_{j-1})| : -\infty < x_0 < x_1 < \ldots < x_n = x, n \in \mathbb{N}\} < \infty.$$

We denote the space of all such functions of bounded variation by BV. If $F \in BV$ satisfies $\lim_{x \to -\infty} F(x) = 0$ and F is right continuous, we say that F is of *normalized bounded variation*, and we denote the space of all such functions by NBV.
If $F \in NBV$, we denote the Lebesgue integral of a measurable function g on $\mathbb{R}$ with respect to the signed measure μ_F by $\int g dF$ if g is integrable with respect to μ_F^+ and μ_F^-. We omit definitions of signed measure as well as the Hahn decomposition, but refer to Chapter 3 of [3].

It is important to note that any increasing function is of bounded variation. This will be useful later in our construction of the stochastic integral.

With knowledge of integration, we can now define the expectation of a random variable.

Definition 1.1.11. We define the *expectation* (with respect to P) of a random variable $\mathbf{X}$ by

$$E[\mathbf{X}] := \int_\Omega \mathbf{X}(\omega) dP(\omega).$$

We generally omit the argument ω and simply write $\int_\Omega \mathbf{X} dP$.
The *distribution* $\mu_\mathbf{X}$ of $\mathbf{X}$ is given by $\mu_\mathbf{X}(B) = P(\mathbf{X}^{-1}(B)) = P(\mathbf{X} \in B)$, where $B \in \mathcal{B}(\mathbb{R}^k)$.

We assume that knowledge of the basic properties and results on the Lebesgue integral are known by the reader. Now, a fundamental class of spaces is introduced.

Definition 1.1.12. Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. We define the *p-norm* for $1 \le p < \infty$ by

$$\|X\|_p := \left(\int_\Omega |X|^p dP \right)^{1/p}.$$

We define the space $L^p(\Omega, \mathcal{F}, P)$ as the space of random variables which are p-integrable. That is,

$$L^p(\Omega, \mathcal{F}, P) = \{X : X \text{ is a random variable and } \|X\|_p < \infty\}.$$

We will sometimes shorten the notation if there is not confusion. For example, we may shorten $L^p(\Omega, \mathcal{F}, P)$ to just L^p provided it is obvious which space we are working on.

It is assumed that the reader has knowledge of basic topology and real analysis, including open and closed sets, normed and inner product spaces, completeness, compactness, denseness, linear operators, sequences, modes of convergence, and so on. Working on this assumption, we introduce the following results.

Theorem 1.1.13. *Let* $1 \leq p < \infty$. *Then*

1. L^p *is a Banach space;*

2. L^2 *is a Hilbert space with the inner product*

$$\langle X, Y \rangle = \int_\Omega XY \, dP = E[XY].$$

3. *Simple functions are dense in* L^p*;*

4. *Over probability spaces,* $L^q \subseteq L^p$ *for* $p < q < \infty$.

Proof. See [3] or [8]. $\square$

Now we focus our attention towards one method of estimating random variables. As we will shortly see, the best estimate of a random variable given some prior information is given by the conditional expectation, which we define below.

Definition 1.1.14. The *conditional expectation* of an integrable random variable X given a σ-algebra $\mathcal{G}$ is the random variable $E(X \mid \mathcal{G})$ that satisfies the properties

1. $E(X \mid \mathcal{G})$ is $\mathcal{G}$-measurable;

2. $E(X \mid \mathcal{G})$ satisfies

$$\int_G E(X \mid \mathcal{G}) dP = \int_G X dP$$

   for every $G \in \mathcal{G}$.

Recall our interpretation of σ-algebra as containing information. Then if we are given some σ-algebra $\mathcal{G}$, we think of this as gaining prior knowledge about X.

Given this $\mathcal{G}$, we need to ensure that an estimate is made using the information from $\mathcal{G}$ (condition 1), and that the estimate has the same expectation as X over the sets in $\mathcal{G}$ (condition 2). There is, however, a different interpretation using the projection.

Consider the space $L^2(\Omega, \mathcal{G}, P)$ where $\mathcal{G} \subseteq \mathcal{F}$. This is a closed vector subspace of $L^2(\Omega, \mathcal{F}, P)$. A well known result on Hilbert spaces tells us that, given $X \in L^2(\mathcal{F})$, there is a unique closest element to X in $L^2(\mathcal{G})$. That is, for $X \in L^2(\mathcal{F})$, there is a unique $Y \in L^2(\mathcal{G})$ satisfying

$$\|X - Y\|_2 = \inf\{\|X - Z\|_2 : Z \in L^2(\mathcal{G})\}.$$

This element Y is called the *projection of X onto $L^2(\mathcal{G})$*, and we can see from the definition of conditional expectation that $Y = E(X \mid \mathcal{G})$. This tells us that the conditional expecatation is the best estimate for X with the given information $\mathcal{G}$.

The above result and some other very important properties of conditional expectation are given below.

Theorem 1.1.15. *Conditional expectation has the following properties:*
Let $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ be σ-algebras, $X, Y \in L^2(\Omega, \mathcal{F}, P)$, and a, b constants. Then

1. *$E(aX + bY \mid \mathcal{G}) = aE(X \mid \mathcal{G}) + bE(Y \mid \mathcal{G})$;*
2. *Since X is $\mathcal{F}$-measurable, for a function f, we have $E(f(X) \mid \mathcal{F}) = f(X)$;*
3. *Whenever $\mathcal{F} = \{\varnothing, \Omega\}$, we have $E(X \mid \mathcal{F}) = E[X]$;*
4. *$E(E(X \mid \mathcal{G}) \mid \mathcal{H}) = E(X \mid \mathcal{H})$;*
5. *$E(X \mid \mathcal{G})$ is the projection (in L^2) of X onto the linear subspace $L^2(\Omega, \mathcal{G}, P)$. That is, $E(X \mid \mathcal{G})$ is the best estimate of X given $\mathcal{G}$.*

Proof. See Chapter 8 of [8]. $\square$

The distribution of a random variable tells us all information about the probabilistic behavior of the variable. It is therefore useful to have different ways to characterize the distribution of a given random variable. Also of central importance is whether two variables have some influence over one another, which is the concept of independence. Both of these can be related to what is called the moment generating function.

Definition 1.1.16. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_i \subseteq \mathcal{F}$, $i = 1, 2, \ldots$. We say that $(\mathcal{F}_i)_{i=1}^{\infty}$ are *independent* if for all choices of $1 \leq k_1 < \ldots < k_m$ and of events $A_{k_i} \in \mathcal{F}_{k_i}$, we have

$$P(A_{k_1} \cap A_{k_2} \cap \ldots \cap A_{k_m}) = P(A_{k_1})P(A_{k_2}) \ldots P(A_{k_m}).$$

For random variables $\mathbf{X}_i : \Omega \to \mathbb{R}^k$ (i=1,2,...), we say $(\mathbf{X}_i)_{i=1}^{\infty}$ are *independent* if for all integers $j \geq 2$ and all $B_i \in \mathcal{B}(\mathbb{R}^k)$, $i = 1, 2, \ldots, j$, we have

$$P(\mathbf{X}_1 \in B_1, \mathbf{X}_2 \in B_2, \ldots, \mathbf{X}_j \in B_j) = P(\mathbf{X}_1 \in B_1)P(\mathbf{X}_2 \in B_2) \ldots P(\mathbf{X}_j \in B_j).$$

Definition 1.1.17. We define the *moment generating function* $M_X : \mathbb{R} \to \mathbb{R}$ of a random variable X by

$$M_X(t) = E[e^{tX}].$$

For a random vector $\mathbf{X} = (X_1, \ldots, X_k)$, we can define the *joint moment generating function* $M_{\mathbf{X}} : \mathbb{R}^k \to \mathbb{R}$ by

$$M_{\mathbf{X}}(t) = E[e^{t^T \mathbf{X}}] = E[e^{t_1 X_1 + t_2 X_2 + \ldots + t_k X_k}].$$

It should be noted that the moment generating function may not exist, but in our applications it will. To ensure existence, the characteristic function is often used. Below, we give some results about characterization of the distribution of a random variable using the moment generating function.

Theorem 1.1.18. *The moment generating function has the following properties:*

1. *If* $X_1, X_2, \ldots, X_k$ *are random variables, then* $M_{X_1+X_2+\ldots+X_k}(t) = M_{X_1}(t)M_{X_2}(t)\ldots M_{X_k}(t)$ *if* $X_1, X_2, \ldots, X_k$ *are independent. The joint moment generating function of the vector* $\mathbf{X} = (X_1, X_2, \ldots, X_k)$ *is* $M_{\mathbf{X}}(t) = M_{X_1}(t_1)M_{X_2}(t_2)\ldots M_{X_k}(t_k)$ *if and only if* $X_1, X_2, \ldots, X_k$ *are independent;*

2. *If* X, Y *are random variables or vectors satisfying* $M_X(t) = M_Y(t)$*, then* X *and* Y *have the same distribution;*

3. *If* X *is normally distributed with mean 0 and variance* σ^2*, then* $M_X(t) = e^{\frac{1}{2}(t\sigma)^2}$.

Proof. See Chapter 6 of [8] and Chapter 2 of [4]. $\square$

1.2 The Stochastic Integral

With the foundation we have laid out, we will now begin the construction of another type of integration, which will be discussed in depth. In particular, we define $\int X dM$, where M is a stochastic process, which we will call the "stochastic integral." First, we must define a very important class of processes which will become appropriate integrators.

Definition 1.2.1. A *filtration* (on $(\Omega, \mathcal{F})$) is a family $\mathcal{M} = (\mathcal{M}_t)_{t\geq 0}$ of σ-algebras $\mathcal{M}_t \subset \mathcal{F}$ such that

$$0 \leq s < t \implies \mathcal{M}_s \subset \mathcal{M}_t.$$

An one-dimensional stochastic process $(M_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is called a *martingale* with respect to a filtration $(\mathcal{M}_t)_{t\geq 0}$ (and with respect to P) if

1. M_t is $\mathcal{M}_t$-measurable for all t;

2. $E[|M_t|] < \infty$ for all t;

3. $E(M_s \mid \mathcal{M}_t) = M_t$ for all $s \geq t$.

Martingales originally came about from a betting strategy where a gambler wins if a coin comes up heads when flipped. If he wins his bet, the game is done. Otherwise, he doubles his previous bet. The reasoning behind this strategy is that upon winning, the gambler would recover all his previous losses as well as win the initial stake.

Upon analyzing this method, one can see that it is not a good betting strategy, and this is due to the fact that regardless of the previous information on bets, one cannot inform their decision on the next bet. This is the concept that martingales encapsulate: a fair betting game where previous information will not inform future decisions.

Now, we must define an appropriate class of integrands. This will turn out to be the predictable processes.

Definition 1.2.2. Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration. A process $(X_t)_{t\geq 0}$ is called $\mathcal{F}_t$*-adapted* if for each t the function X_t is $\mathcal{F}_t$-measurable.
In general, we say that the process is *measurable* if the mapping $(t, \omega) \in [0, \infty) \times \Omega \mapsto X_t(\omega) \in \mathbb{R}$ is $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$-measurable.
By $\mathscr{L}$ we denote the smallest σ-algebra on $[0, \infty) \times \Omega$ such that all left-continuous $\mathcal{F}_t$-adapted processes $Y : [0, \infty) \times \Omega \ni (t, \omega) \mapsto Y_t(\omega) \in \mathbb{R}$ are measurable.
We say that a process is $\mathcal{F}_t$*-predictable* if the mapping $(t, \omega) \mapsto X_t(\omega)$ is $\mathscr{L}$-measurable.
We say that a process $(A_t)_{t\geq 0}$ is an *integrable increasing process* if

1. $\mathcal{F}_t$-adapted;

2. $A_0 = 0$ and $t \mapsto A_t$ is right continuous and increasing almost surely;

3. $E[A_t] < \infty$ for all $t \in [0, \infty)$.

We say that an integrable increasing process $(A_t)_{t\geq 0}$ is *natural* if every bounded martingale M satisfies

$$E\left[\int_0^t M_s dA_s\right] = E\left[\int_0^t M_{s-} dA_s\right]$$

where $s-$ is the limit from the left.

Let M be a square integrable martingale. We define the *quadratic variation process* $\langle M \rangle_t$ to be the unique (see Chapter 2.2 of [6]) natural integrable process such that $M_t^2 - \langle M \rangle_t$ is an $\mathcal{F}_t$-martingale. Since $\langle M \rangle_t$ is a natural integrable increasing process, we can define $\int_0^t M_s d\langle M \rangle_s$ as the Stieltjes integral.

We define the space

$$\mathcal{L}^2(M) = \{(\varphi_t(\omega))_{t\geq 0} : \varphi \text{ is } \mathcal{F}_t\text{-predictable, and for every } T > 0,$$
$$\left(\|\varphi\|_{2,T}^M\right)^2 = E\left[\int_0^T \varphi^2(s) d\langle M \rangle_s\right] < \infty\}.$$

For $\varphi \in \mathcal{L}^2(M)$ we can define $\|\varphi\|_2^M = \sum_{n=1}^\infty 2^{-n}(\|\varphi\|_{2,n}^M \wedge 1)$. We can define the set of square integrable processes $\mathcal{L}^0$, consisting of those φ such that there exists a sequence $0 = t_0 < t_1 < \ldots < t_n < \ldots \to \infty$ and a sequence of random variables $(\psi_i(\omega))_{i=0}^\infty$, where each ψ_i is $\mathcal{F}_{t_i}$-measurable, $\sup \|\psi_i\|_\infty < \infty$, and

$$\varphi(t, \omega) = \begin{cases} \psi_0(\omega) & \text{if } t = 0 \\ \psi_i(\omega) & \text{if } t \in (t_i, t_{i+1}],\ i = 0, 1, \ldots \end{cases}$$

Such φ has the expression

$$\varphi(t, \omega) = \psi_0(\omega) I_{\{t=0\}}(t) + \sum_{i=1}^\infty \psi_i(\omega) I_{\{t\in(t_i,t_{i+1}]\}}(t).$$

It can be shown that $\mathcal{L}^0$ is dense in $\mathcal{L}^2(M)$ under the norm $\|\cdot\|_2^M$ (see Chapter 2.2 of [6]), and so for $\varphi \in \mathcal{L}^0$ having the form above, we set

$$\int_0^t \varphi(s)dM_s = \sum_{i=0}^{n-1} \psi_i(\omega)(M_{t_{i+1}} - M_{t_i}) + \psi_n(\omega)(M_t - M_{t_n})$$

for $t_n \leq t \leq t_{n+1}$, $n = 0, 1, \ldots$. It can be shown that $\int_0^t \varphi(s)dM_s$ is a square integrable martingale, and that the isometry

$$E\left[\left(\int_0^t \varphi(s)dM_s\right)^2\right] = E\left[\int_0^t \varphi^2(s)d\left\langle M\right\rangle_s\right]$$

holds true. Then if $\varphi \in \mathcal{L}^2(M)$, there exists a sequence $(\varphi_n)_{n=1}^{\infty} \subseteq \mathcal{L}^0$ such that $\|\varphi - \varphi_n\|_2^M \to 0$.

It is a fact that under the norm $|||M.||| = \sum_{T=1}^{\infty} 2^{-T}(E[M_T^2]^{1/2} \wedge 1)$ the space of all square integrable martingales is a complete metric space. By the isometry given above, have that $||| \int_0^{\cdot} \varphi_s dM_s ||| = \|\varphi\|_2^M$. It follows that

$$\|\varphi_m - \varphi_n\|_2^M = ||| \int_0^{\cdot} \varphi_m(s)dM_s - \int_0^{\cdot} \varphi_n(s)dM_s ||| \to 0 \text{ as } n \to \infty.$$

Then $\left(\int_0^t \varphi_n(s)dM_s\right)_{n=1}^{\infty}$ forms a Cauchy sequence in the space of square integrable martingales, and is therefore convergent to some unique element of that space.

Using this unique element, we can define the stochastic integral.

Definition 1.2.3. The *stochastic integral* of $\varphi \in \mathcal{L}^2(M)$ with respect to the square integrable martingale M is defined to be

$$\int_0^t \varphi(s)dM_s = \lim_{n\to\infty} \int_0^t \varphi_n(s)dM_s$$

where $(\varphi_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{L}^0$ converging to φ in $\mathcal{L}^2(M)$.

The following useful properties hold for the stochastic integral.

Theorem 1.2.4. (Properties of the stochastic integral) *Let M be a square integrable martingale, $f, g \in \mathcal{L}^2(M)$, and $0 \leq S < U < T$. Then*

1. $\int_S^T f dM_t = \int_S^U f dM_t + \int_U^T f dM_t$ *for almost all* $\omega \in \Omega$*;*

2. $\int_S^T (cf + g) dM_t = c \cdot \int_S^T f dM_t + \int_S^T g dM_t$ *(c constant) for almost all* $\omega \in \Omega$*;*

3. $E\left[\int_S^T f dM_t\right] = 0$*;*

4. $\int_S^T f dM_t$ *is* $\mathcal{F}_T$*-measurable;*

5. $\int_0^t f(s,\omega) dM_s$ *is an* $\mathcal{F}_t$*-martingale.*

Proof. See Chapter 2.2 of [6]. □

The evolution over time of a system can commonly be described by a differential equation. In many cases, it may be necessary to also include some randomness in addition to the deterministic differential equation. This included randomness can be described using stochastic processes. With this in mind, we give the definition which will be the model for evolution over time of a system which includes a source of randomness or noise.

Definition 1.2.5. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{t\geq 0}$ a filtration. A stochastic process $(X_t)_{t\geq 0}$ defined on this space is called a *semi-martingale* if it has the form

$$X_t = X_0 + M_t + A_t$$

where X_0 is an $\mathcal{F}_0$-measurable random variable, $(M_t)_{t\geq 0}$ is an $\mathcal{F}_t$-martingale, and $(A_t)_{t\geq 0}$ is a continuous $\mathcal{F}_t$-adapted process such that $A_0 = 0$ almost surely and the map $t \mapsto A(t)$ is of bounded variation on each finite interval almost surely.

Now we give what is arguably the most important result in stochastic calculus, which will give us a way to simplify finding solutions to stochastic differential equations. We give a basic form of the result, but a generalized version is given in [6].

Theorem 1.2.6. (Itô's formula for semi-martingales) *Let* $Y_t = g(t, X_t)$ *be twice continuously differentiable (that is,* $g \in C^2([0,\infty) \times \mathbb{R})$*) on* $\mathbb{R}$ *and* $(X_t)_{t\geq 0}$ *a semi-martingale. Then the process* $(Y_t)_{t\geq 0}$ *is also a semi-martingale (with respect to* $(\mathcal{F}_t)_{t\geq 0}$*) and*

$$dY_t = \frac{\partial}{\partial t} g(t, X_t)dt + \frac{\partial}{\partial x} g(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} g(t, X_t)d\left\langle X \right\rangle_t$$

holds true, where $\left\langle X \right\rangle_t = \left\langle M \right\rangle_t$ *as* M_t *was given in Definition 1.2.5.*

Proof. See Chapter 2.5 of [6]. □

In the formula above, by $\frac{\partial}{\partial x} g(t, X_t)$, we mean take the partial derivative with respect to x of the function $g(t, x)$, then substitute $x = X_t$, and similarly for the other terms.

1.3 Brownian Motion and Itô Calculus

As noted at the beginning of the chapter, the main topic we wish to explore is Brownian motion. We now have sufficient background to give a mathematical definition, and study its properties.

Definition 1.3.1. A *Brownian motion* or *Wiener process* is a stochastic process $(B_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ with the following properties:

1. The process starts at 0:

$$P(B_0 = 0) = 1;$$

2. The increments are independent: If

$$0 \leq t_0 < t_1 < \ldots < t_k,$$

then

$$P(B_{t_i} - B_{t_{i-1}} \in H_i, i \leq k) = \prod_{i \leq k} P(B_{t_i} - B_{t_{i-1}} \in H_i);$$

3. For $0 \leq s < t$ the increment $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$:

$$P(B_t - B_s \in H) = \frac{1}{\sqrt{2\pi(t-s)}} \int_H e^{-x^2/2(t-s)} dx;$$

4. For each ω, $B(t, \omega)$ is continuous in t and $B(0, \omega) = 0$ (note that this is not assumed on B_t, but is simply a property. See for example [2], [10]).

Without proof, note that Brownian motion exists and is $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$-measurable (Again, see [2], [10]).

Brownian motion possesses many fascinating properties. As noted above, the sample paths are continuous, but it turns out that these paths are nowhere differentiable. Due to this fact, a study of differential equations might seem impossible. However, we will see that this gives way to a rich theory. The first step is defining integration with respect to Brownian motion, and so we first give a result.

Theorem 1.3.2. *Let $(B_t)_{t \geq 0}$ be a Brownian motion on $(\Omega, \mathcal{F}, P)$. Then B_t is an $\mathcal{F}_t$-martingale, where $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$.*

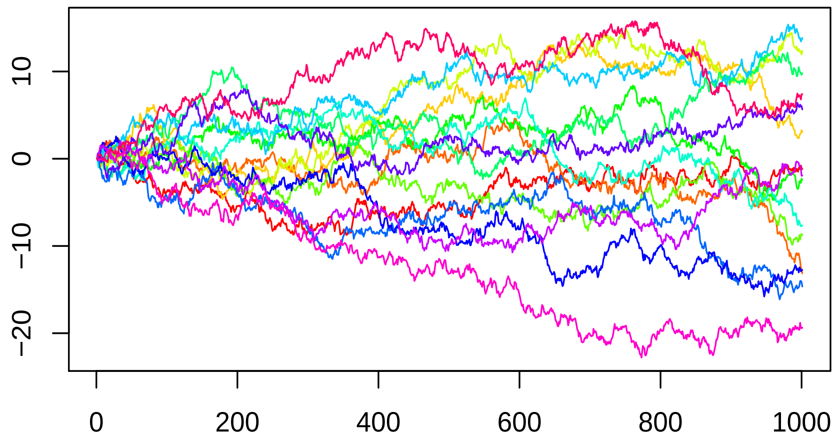

Figure 1.3: Fifteen simulated sample paths for Brownian motion.

Proof. We know B_t has mean 0 for all t, and so the integrability condition is met. We also have for $s \leq t$,

$$E(B_t \mid \mathcal{F}_s) = E(B_t - B_s + B_s \mid \mathcal{F}_s) = E(B_t - B_s \mid \mathcal{F}_s) + E(B_s \mid \mathcal{F}_s) = E(B_t - B_s \mid \mathcal{F}_s) + B_s.$$

Since $B_t - B_s$ is independent of $\mathcal{F}_s$,

$$E(B_t - B_s \mid \mathcal{F}_s) = E[B_t - B_s] = E[B_t] - E[B_s] = 0.$$

Therefore, $E(B_t \mid \mathcal{F}_s) = B_s$, and so B_t is an $\mathcal{F}_t$-martingale. $\square$

Definition 1.3.3. $\mathcal{V}(S,T) = \mathcal{V}$ is the class of functions

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$$

such that

1. $(t,\omega) \to f(t,\omega)$ is $\mathcal{B}([0,\infty)) \otimes \mathcal{F}$-measurable;

2. $f(t,\omega)$ is $\mathcal{F}_t$-adapted;

3. $E[\int_S^T f(t,\omega)^2 dt] < \infty$.

Those functions satisfying 1 and 2 are the *progressively measurable* functions.

It is useful to think of the above functions as only depending upon information from the past, and being appropriately jointly measurable in the variables t and ω together. This class of functions is less general than those defined in Definition 1.2.2 due to the fact that predictable is a stronger condition than adapted. For our purposes, however, adapted will be sufficiently strong and more convenient.

Definition 1.3.4. For $f \in \mathcal{V}(S,T)$ we define the *Itô integral of f with respect to B* as the stochastic integral with respect to the martingale B.

The following isometry will be of great importance when studying integrals of this type.

Theorem 1.3.5. (The Itô isometry) *For all $f \in \mathcal{V}(S,T)$ we have*

$$E\left[\left(\int_S^T f(t,\omega)dB_t\right)^2\right] = E\left[\int_S^T f^2(t,\omega)dt\right]$$

Proof. See Lemma 3.1.5 and Corollary 3.1.7 of [10]. $\square$

Since the Itô integral is in fact a special case of the stochastic integral from the previous section, the following properties hold true.

Theorem 1.3.6. (Properties of the Itô integral) *Let $f, g \in \mathcal{V}(0,T)$ and let $0 \leq S < U < T$. Then*

1. *$\int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t$ for almost all $\omega \in \Omega$;*

2. $\int_S^T (cf+g)dB_t = c\cdot\int_S^T f dB_t + \int_S^T g dB_t$ *(c constant) for almost all* $\omega\in\Omega$*;*

3. $E\left[\int_S^T f dB_t\right] = 0$*;*

4. $\int_S^T f dB_t$ *is* $\mathcal{F}_T$*-measurable;*

5. $\int_0^t f(s,\omega)dB_s$ *is an* $\mathcal{F}_t$*-martingale.*

Proof. See Section 3.2 of [10]. □

With integration and its properties laid out, we will begin our study of stochastic differential equations. We will do this by studying a specific type of semi-martingale, called an Itô process.

Definition 1.3.7. Let $(B_t)_{t\geq 0}$ be 1-dimensional Brownian motion on $(\Omega,\mathcal{F},P)$. We define $\mathcal{W}(S,T)$ to be the class of processes $f(t,\omega)\in\mathbb{R}$ satisfying

1. $(t,\omega)\to f(t,\omega)$ is $\mathcal{B}([0,\infty))\otimes\mathcal{F}$-measurable;

2. There exists an increasing family of σ-algebras $(\mathcal{F}_t)_{t\geq 0}$ such that B_t is an $\mathcal{F}_t$-martingale and f_t is $\mathcal{F}_t$-adapted;

3. $P\left(\int_S^T f(s,\omega)^2 ds < \infty\right) = 1$.

A (1-dimensional) *Itô process* is a stochastic process $(X_t)_{t\geq 0}$ on $(\Omega,\mathcal{F},P)$ of the form

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s,$$

where $v\in\mathcal{W}$, so that

$$P\left(\int_0^t v(s,\omega)^2 ds < \infty \text{ for all } t\geq 0\right) = 1.$$

We also assume that u is $\mathcal{F}_t$-adapted and

$$P\left(\int_0^t |u(s,\omega)|ds < \infty \text{ for all } t\geq 0\right) = 1.$$

If X_t is an Itô process of the above form we sometimes write it as

$$dX_t = u_t dt + v_t dB_t.$$

Generally, a stochastic differential equation is any differential equation which allows one of the terms to be a stochastic process. Itô processes are a type of stochastic differential equation, and the only one we will study at length.

One of the most powerful methods for computing solutions to these Itô processes is given below.

Theorem 1.3.8. (Itô's formula) *Let $(X_t)_{t\geq 0}$ be an Itô process given by*

$$dX_t = u_t dt + v_t dB_t$$

Let $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$. Then

$$Y_t = g(t, X_t)$$

is again an Itô process and

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t) \cdot d\left\langle X \right\rangle_t$$

where $d\left\langle X \right\rangle_t = dX_t \cdot dX_t$ is computed according to the rules

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \;\; dB_t \cdot dB_t = dt.$$

Proof. See Theorem 4.1.2 of [10]. $\square$

As noted in Theorem 1.2.6, by $\frac{\partial}{\partial x} g(t, X_t)$ we mean take the partial derivative with respect to x of the function $g(t,x)$, then substitute $x = X_t$, and similarly for the other terms.

The above formula coincides with our previous Itô formula for martingales, this time in differential form. Notice that for Brownian motion, $\left\langle B \right\rangle_t = t$. As it turns out, any martingale whose quadratic variation is t is a Brownian motion. A version of this characterization is given in Chapter 3.

Now we give another important result on products of Itô processes, including a proof leveraging the power of Itô's formula.

Corollary 1.3.9. (Itô product rule) *Let* $(X_t)_{t\geq 0}, (Y_t)t \geq 0$ *be processes given by*

$$dX_t = \sigma_t dB_t + \mu_t dt$$

$$dY_t = \rho_t dB_t + \nu_t dt$$

for Brownian motion $(B_t)_{t\geq 0}$*. Then*

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + \sigma_t\rho_t dt = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

Proof. Let $Z_t = X_t^2$. Apply Itô's formula to get

$$dZ_t = d(X_t^2) = 2X_t dX_t + (dX_t)^2. \tag{1.1}$$

Since $X_tY_t = \frac{1}{2}((X_t + Y_t)^2 - X_t^2 - Y_t^2)$, use linearity to get

$$d(X_tY_t) = d(\frac{1}{2}((X_t + Y_t)^2 - X_t^2 - Y_t^2)) = \frac{1}{2}d((X_t + Y_t)^2) - \frac{1}{2}d(X_t^2) - \frac{1}{2}d(Y_t^2),$$

then apply (1.1) to each term to get

$$d(X_tY_t) = (X_t + Y_t)d(X_t + Y_t) + \frac{1}{2}(d(X_t + Y_t))^2 - X_t dX_t - \frac{1}{2}(dX_t)^2 - Y_t dY_t - \frac{1}{2}(dY_t)^2. \tag{1.2}$$

Apply linearity again to reduce the terms

$$(X_t + Y_t)d(X_t + Y_t) = (X_t + Y_t)(dX_t + dY_t) = X_t dX_t + X_t dY_t + Y_t dX_t + Y_t dY_t$$

and

$$\frac{1}{2}(d(X_t + Y_t))^2 = \frac{1}{2}(dX_t)^2 + dX_t dY_t + \frac{1}{2}(dY_t)^2.$$

Cancel terms in (1.2) to get

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

We also have

$$dX_t dY_t = \sigma_t \rho_t (dB_t)^2 + \sigma_t \nu_t dB_t \cdot dt + \rho_t \mu_t dB_t \cdot dt + \mu_t \nu_t (dt)^2 = \sigma_t \rho_t dt$$

according to the rules in Theorem 1.3.8. $\square$

CHAPTER 2

REPRESENTATION OF MARTINGALES

In the first chapter, we found that the Itô process $X_t = X_0 + \int_0^t v(s,\omega)dB_s$ is an $\mathcal{F}_t$-martingale. As we will show in this chapter, the converse is also true. That is, given any $\mathcal{F}_t$-martingale M_t, M_t can be represented as an Itô process.

Mathematical analysis is often considered to be a pure topic, studied with no thought of applications. The results in this chapter draw from almost every area of analysis, and at first may seem to be purely abstract. However, the representation theorem presented here will be of extreme importance in deriving one of the most important modern concepts in mathematical finance, the Black-Scholes equation for options pricing.

2.1 Preliminary Results

From now on, we will assume we are working on the probability space $(\Omega, \mathcal{F}, P)$. When discussing a Brownian motion $(B_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$, $\mathcal{F}_T$ is defined to be the smallest σ-algebra containing all sets of the form

$$\{\omega : B_{t_1} \in F_1, \ldots, B_{t_k} \in F_k\}$$

where the points $t_j \leq t$ are dense in $[0,T]$ and $F_j \subseteq \mathbb{R}$ are Borel sets, $j \leq k$, and $k = 1, 2, \ldots$. It is important to note that $\mathcal{F}_T = \sigma(B_s, 0 \leq s \leq T)$. Certainly, $\mathcal{F}_T \subseteq \sigma(B_s, 0 \leq s \leq T)$. Since $(t_j)_{j=1}^{\infty}$ is dense in $[0,T]$, for any $t \in [0,T]$, we can find $t_{j_k} \to t$. Since B is continuous, we have $B_{t_{j_k}} \to B_t$. Since limits of measurable functions are measurable, $B_t \in \mathcal{F}_T$, and since t was arbitrary, we have $\mathcal{F}_T \supseteq \sigma(B_s, 0 \leq s \leq T)$.

The main result of this section shows that a certain class of random variables is dense in the space $L^2(\Omega, \mathcal{F}_T, P)$. We will sometimes refer to this space as $L^2(\mathcal{F}_T, P)$. Before this main result, we need to show that the set of continuously differentiable functions with compact support on $\mathbb{R}^k$ is dense in the space $L^p(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mu)$ for a Borel measure μ.

Lemma 2.1.1. *Let $X : \Omega \to \mathbb{R}$ be a random variable, Y be $\sigma(X)$-measurable, and $p \leq 1 < \infty$. Then $Y = f(X) \in L^p(\Omega, \sigma(X), P)$ if and only if $f \in L^p(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mu_X)$, where μ_X is the distribution of X.*

Proof. We claim that

$$\int_\Omega |f(X)|^p dP = \int_{\mathbb{R}^k} |f|^p d\mu_X,$$

which would prove the result.

Note first that if f is measurable then $|f|^p$ is also measurable, and put $g = |f|^p$ for simplicity. If $g = I_A$ for $A \in \mathcal{B}(\mathbb{R}^k)$, then $g(X) = I_{X^{-1}A}$, and so we get the equality

$$P(X^{-1}A) = \mu_X(A)$$

by definition of μ_X. Then if we take g to be a nonnegative simple function, the equality holds by linearity using the same argument. There is a sequence of nonnegative simple functions $(g_n)_{n=1}^\infty$ such that $g_n \uparrow g$, and so we also have that $g_n(X) \uparrow g(X)$, where each $g_n(X)$ is nonnegative. For each n, we have

$$\int_{\mathbb{R}^k} g_n d\mu_X = \int_\Omega g_n(X) dP$$

and so

$$\lim_{n\to\infty} \int_{\mathbb{R}^k} g_n d\mu_X = \lim_{n\to\infty} \int_\Omega g_n(X) dP.$$

Then by the monotone convergence theorem

$$\lim_{n\to\infty} \int_{\mathbb{R}^k} g_n d\mu_X = \int_{\mathbb{R}^k} g d\mu_X$$

and

$$\lim_{n\to\infty} \int_\Omega g_n(X) dP = \int_\Omega g(X) dP$$

which completes the proof. $\square$

Lemma 2.1.2. (Leúy's zero-one law) *Let $X \in L^p(\Omega, \mathcal{F}, P)$, $p \geq 1$ and $(\mathcal{N}_k)_{k=1}^{\infty}$ be an increasing family of σ-algebras each contained in $\mathcal{F}$. Define $\mathcal{N}_\infty$ to be the σ-algebra generated by $(\mathcal{N}_k)$. Then*

$$E(X \mid \mathcal{N}_k) \to E(X \mid \mathcal{N}_\infty)$$

as $k \to \infty$ almost everywhere P and in $L^p(\Omega, \mathcal{F}, P)$.

Proof. We omit the proof for brevity. See Appendix C of [10]. $\square$

Definition 2.1.3. A Borel measure μ is called a *Radon measure* if it is finite on compact sets, and satisfies (outer regular)

$$\mu(B) = \inf\{\mu(U) : B \subset U, U \text{ open}\}$$

for all Borel sets B, and (inner regular)

$$\mu(O) = \sup\{\mu(K) : K \subset O, K \text{ compact}\}$$

for all open sets O.

The *support* of a function $f : \mathbb{R}^k \to \mathbb{R}$, denoted $\operatorname{supp}(f)$, is the closure of the set $\{x \in X : f(x) \neq 0\}$. We define $C_0^\infty(\mathbb{R}^k)$ to be the space of infinitely differentiable functions on $\mathbb{R}^k$ whose support is compact and contained in $\mathbb{R}^k$.

Lemma 2.1.4. *If μ is a Radon measure on $\mathcal{B}(\mathbb{R}^k)$, then μ is inner regular on all of its measurable sets with finite measure.*

Proof. Suppose $\mu(E) < \infty$. Given $\varepsilon > 0$, we can choose an open $U \supset E$ such that $\mu(U) < \mu(E) + \varepsilon$ and a compact $F \subset U$ such that $\mu(F) > \mu(U) - \varepsilon$. Since $\mu(U \setminus E) < \varepsilon$ we can also choose an open $V \supset U \setminus E$ such that $\mu(V) < \varepsilon$. Let

$K = F \setminus V$. Then K is compact, $K \subset E$, and

$$\mu(K) = \mu(F) - \mu(F \cap V) > \mu(E) - \varepsilon - \mu(V) > \mu(E) - 2\varepsilon.$$

Thus μ is inner regular on E. $\square$

Lemma 2.1.5. (Urysohn's lemma) *If $K \subset \mathbb{R}^k$ is compact and U is an open set containing K, there exists $f \in C_0^\infty(\mathbb{R}^k)$ such that $0 \leq f \leq 1$, $f = 1$ on K, and supp$(f) \subset U$.*

Proof. The proof is omitted for the sake of brevity. See Chapter 8 of [3], where the result is proven in (8.18). $\square$

Lemma 2.1.6. *If $1 \leq p < \infty$ and μ is a Radon measure on $\mathbb{R}^k$, then $C_0^\infty(\mathbb{R}^k)$ is dense in $L^p(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mu)$.*

Proof. Since the L^p simple functions are dense in L^p (see [3] Proposition 6.7), it suffices to show that for any $E \in \mathcal{B}(\mathbb{R}^k)$ with $\mu(E) < \infty$, I_E can be approximated in L^p by elements of C_0^∞. Given $\varepsilon > 0$, by Lemma 2.1.4, we can choose a compact $K \subset E$ and an open $E \subset U$ such that $\mu(U \setminus K) < \varepsilon$, and by Urysohn's lemma (Lemma 2.1.5) we can choose an $f \in C_0^\infty(\mathbb{R}^k)$ such that $I_K \leq f \leq I_U$. Then

$$\|I_E - f\|_p \leq \mu(U \setminus K)^{1/p} < \varepsilon^{1/p}.$$

$\square$

With the results we have established so far, we can now prove that two specific sets of random variables are dense in the space $L^2(\mathcal{F}_T, P)$. We will use the fact that any Borel probability measure defined on $\mathcal{B}(\mathbb{R}^k)$ is a Radon measure (see Theorem 7.8 of [3]) along with Lemma 2.1.6 to prove Lemma 2.1.7 below.

Lemma 2.1.7. *Fix $T > 0$. Then the set of random variables*

$$\{\phi(B_{t_1}, \dots, B_{t_n}) : t_i \in [0, T], \phi \in C_0^\infty(\mathbb{R}^k), n = 1, 2, \dots\}$$

is dense in $L^2(\mathcal{F}_T, P)$.

Proof. Let $(t_i)_{i=1}^\infty$ be a dense subset of $[0, T]$ and for each $n = 1, 2, \dots$ let $\mathcal{H}_n = \sigma(B_{t_1}, \dots, B_{t_n})$. Then

$$\mathcal{H}_n \subseteq \mathcal{H}_{n+1} \subseteq \dots$$

and $\mathcal{F}_T$ is the smallest σ-algebra containing each $\mathcal{H}_n$, and so it is the σ-algebra generated by $(\mathcal{H}_n)_{n=1}^\infty$.

Let $g \in L^2(\mathcal{F}_T, P)$. By Lemma 2.1.2 we have that

$$E(g \mid \mathcal{H}_n) \to E(g \mid \mathcal{F}_T) = g$$

as $n \to \infty$ almost everywhere P and in $L^2(\mathcal{F}_T, P)$. For each n, $E(g \mid \mathcal{H}_n)$ is $\mathcal{H}_n$ measurable, and by Theorem 1.1.8 this is true if and only if there exists a Borel measurable map $g_n : \mathbb{R}^k \to \mathbb{R}$ such that

$$E(g \mid \mathcal{H}_n) = g_n(B_{t_1}, \dots, B_{t_n})$$

for each n. Then by Lemma 2.1.1, since each $g_n(B_{t_1}, \dots, B_{t_n}) \in L^2(\mathcal{F}_T, P)$, $g_n \in L^2(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mu_B)$ and so by Lemma 2.1.6 each g_n can be approximated in $L^2(\mathcal{F}_T, P)$ by some $\phi_n \in C_0^\infty$ where

$$\|g_n(B_{t_1}, \dots, B_{t_n}) - \phi_n(B_{t_1}, \dots, B_{t_n})\|_2 < \varepsilon$$

for arbitrary $\varepsilon > 0$. Then we have that every $g \in L^2(\mathcal{F}_T, P)$ can be approximated by some $\phi \in C_0^\infty$, proving the result. $\square$

Theorem 2.1.8. (Fubini's theorem) *Suppose* $(X, \mathcal{M}, \mu)$ *and* $(Y, \mathcal{N}, \nu)$ *are probability spaces. If* $f \in L^1(\mu \otimes \nu)$*, then*

$$\int \left[\int f(x,y) d\nu(y) \right] d\mu(x) = \int \left[\int f(x,y) d\mu(x) \right] d\nu(y)$$

Proof. See Theorem 2.37 of [3]. $\square$

Lemma 2.1.9. *Let S be a subset of a Hilbert space H. Then span(S) is dense in H if and only if the orthogonal complement* $S^\perp = \{u \in H : \langle u, v \rangle = 0 \text{ for all } v \in S\} = 0$.

Proof. See Section 16.1 of [11], Corollary 4. $\square$

The proof of the result below relies upon results from complex analysis. Some reasoning will be given, but we will not elaborate upon these points.

Lemma 2.1.10. *Let $\mathcal{S}$ be the set of functions of the form*

$$\exp\left\{ \int_0^T h(t) dB_t(\omega) - \frac{1}{2} \int_0^T h^2(t) dt \right\}$$

where $h \in L^2[0,T] = L^2([0,T], \mathcal{B}[0,T], m)$ *(with standard Lebesgue measure m) is a deterministic function. Then span$(\mathcal{S})$ is dense in* $L^2(\mathcal{F}_T, P)$.

Proof. Recall that $L^2(\mathcal{F}_T, P)$ is a Hilbert space, and the inner product of two elements $X, Y \in L^2(\mathcal{F}_T, P)$ is given by

$$\langle X, Y \rangle = \int_\Omega XY dP.$$

Let $0 = t_0 < t_1 < t_2 < \ldots < t_k \leq T$ and define the function

$$h(t) = \begin{cases} \lambda_i + \ldots + \lambda_k & \text{if } t_{i-1} \leq t < t_i \\ 0 & \text{if } t_k \leq t < T. \end{cases}$$

Then $h \in L^2[0,T]$ and is deterministic, and

$$\begin{aligned}\int_0^T h(t)dB_t &= \int_0^{t_1}(\lambda_1+\ldots+\lambda_k)dB_t + \int_{t_1}^{t_2}(\lambda_1+\ldots+\lambda_k)dB_t + \ldots + \int_{t_{k-1}}^{t_k}\lambda_k dB_t \\ &= (\lambda_1+\ldots+\lambda_k)B_{t_1} + (\lambda_2+\ldots+\lambda_k)(B_{t_2}-B_{t_1}) + \ldots + \lambda_k(B_{t_k}-B_{t_{k-1}}) \\ &= \lambda_1 B_{t_1} + \lambda_2 B_{t_2} + \ldots + \lambda_k B_{t_k}.\end{aligned}$$

Furthermore, $\frac{1}{2}\int_0^T h^2(t)dt$ is a constant, and so we can multiply appropriately to get

$$e^{\sum_{i=1}^k \lambda_i B_{t_i}} \in \text{ span}(\mathcal{S}).$$

Let $g \in \mathcal{S}^\perp$. Then we have

$$G(\lambda) := \int_\Omega e^{\sum_{i=1}^k \lambda_i B_{t_i}} g(\omega)dP(\omega) = 0$$

for all $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$ and all $t_1, \ldots, t_k \in [0,T]$. Then $G(\lambda)$ is real analytic in $\mathbb{R}^k$ and so G has an analytic extension to $\mathbb{C}^k$ given by

$$G(z) = \int_\Omega e^{\sum_{i=1}^k z_i B_{t_i}} g(\omega)dP(\omega) = 0$$

for all $z = (z_1, \ldots, z_k) \in \mathbb{C}^k$ since $G(z) = G(\lambda)$ on $\mathbb{R}^k$ and $G(z)$ is infinitely differentiable. Since $G = 0$ on $\mathbb{R}^k$ and it is analytic, $G = 0$ on $\mathbb{C}^k$, and in particular

$$G(iy_1, \ldots, iy_k) = 0 \tag{2.1}$$

for all $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$. Then for $\phi \in C_0^\infty(\mathbb{R}^k)$ we have

$$\int_\Omega \phi(B_{t_1}\ldots, B_{t_k})g(\omega)dP(\omega) = \int_\Omega (2\pi)^{-k/2}\left(\int_{\mathbb{R}^k}\hat{\phi}(y)e^{i\sum_{j=1}^k y_j B_{t_j}}dy\right)g(\omega)dP(\omega) \tag{2.2}$$

$$= (2\pi)^{-k/2}\int_{\mathbb{R}^k}\hat{\phi}(y)\left(\int_\Omega e^{i\sum_{j=1}^k y_j B_{t_j}}g(\omega)dP(\omega)\right)dy \tag{2.3}$$

$$= (2\pi)^{-k/2}\int_{\mathbb{R}^k}\hat{\phi}(y)G(iy)dy = 0 \tag{2.4}$$

where (2.2) is using the inverse Fourier transform

$$\hat{\phi}(y) = (2\pi)^{-k/2} \int_{\mathbb{R}^k} \hat{\phi}(y) e^{-ix\cdot y} dx,$$

(2.3) is using Theorem 2.1.8, and (2.4) is by (2.1).

By the above work and Lemma 2.1.6, we get that g is orthogonal to a dense subset of $L^2(\mathcal{F}_T, P)$, and so we have that $g = 0$ by Lemma 2.1.9. Then $\mathcal{S}^\perp = \{0\}$, and so the result follows by Lemma 2.1.9. $\square$

Lemma 2.1.11. *If $f(t,\omega) \in \mathcal{V}(S,T)$ and $f_n(t,\omega) \in \mathcal{V}(S,T)$ for $n = 1, 2, \ldots$ and*

$$E\left[\int_S^T (f_n(t,\omega) - f(t,\omega))^2 dt\right] \to 0$$

as $n \to \infty$, then

$$\int_S^T f_n(t,\omega) dB_t(\omega) \to \int_S^T f(t,\omega) dB_t(\omega)$$

in $L^2(\Omega, \mathcal{F}, P)$ as $n \to \infty$.

Proof. By the Itô isometry (Theorem 1.3.5),

$$E\left[\int_S^T (f_n(t,\omega) - f(t,\omega))^2 dt\right] = E\left[\left(\int_S^T (f_n(t,\omega) - f(t,\omega)) dB_t\right)^2\right] \to 0$$

as $n \to \infty$. Then $\int_S^T f_n dB \to \int_S^T f dB$ in $L^2(\Omega, \mathcal{F}, P)$. $\square$

2.2 The Itô and Martingale Representation Theorems

With the preliminary results established, we give two representation theorems. The first result gives a representation of a general square integrable random variable as a stochastic process. Once this representation is known, we will use it to prove the main theorem of the section.

Theorem 2.2.1. (The Itô representation theorem) *Let $F \in L^2(\mathcal{F}_T, P)$. Then there exists a unique stochastic process $f(t,\omega) \in \mathcal{V}(0,T)$ such that*

$$F(\omega) = E[F] + \int_0^T f(t,\omega) dB_t. \tag{2.5}$$

Proof. Suppose $F \in \mathcal{S}$ from Lemma 2.1.10. That is, for some deterministic $h(t) \in L^2[0,T]$,

$$F(\omega) = \exp\left\{\int_0^T h(t)dB_t(\omega) - \frac{1}{2}\int_0^T h^2(t)dt\right\}$$

Define

$$Y_t(\omega) = \exp\left\{\int_0^t h(s)dB_s(\omega) - \frac{1}{2}\int_0^t h^2(s)ds\right\}$$

for $0 \le t \le T$. Setting $dX_t = h(t)dB_t$ and $g(t,x) = \exp\left(x - \frac{1}{2}\int_0^t h^2(s)ds\right)$, we have

$$\frac{\partial}{\partial t} g(t,x) = -\frac{1}{2} h^2(t) g(t,x)$$

and

$$\frac{\partial}{\partial x} g(t,x) = \frac{\partial^2}{\partial x^2} g(t,x) = g(t,x).$$

Then by Theorem 1.3.8,

$$\begin{aligned} dY_t &= -Y_t \frac{1}{2} h^2(t)dt + Y_t dX_t + \frac{1}{2} Y_t (dB_t)^2 \\ &= Y_t(h(t)dB_t - \frac{1}{2}h^2(t)dt) + \frac{1}{2} Y_t (h(t)dB_t)^2 \\ &= Y_t h(t) dB_t \end{aligned}$$

so that, since $Y_0 = 1$, we get

$$Y_t = 1 + \int_0^t Y_s h(s) dB_s$$

where $t \in [0,T]$. Therefore

$$F = Y_T = 1 + \int_0^T Y_s h(s) dB_s$$

and hence $E[F] = 1$ since $Y_t h(t) \in \mathcal{V}$ and so $E[\int_0^T Y_s h(s) dB_s] = 0$ by Theorem 1.3.6. Then (2.5) holds in this case, and in fact holds for any function in span($\mathcal{S}$) (in the sense of Lemma 2.1.10) by linearity. Thus, by Lemma 2.1.10, for arbitrary $F \in L^2(\mathcal{F}_T, P)$ there is a sequence $(F_n)_{n=1}^{\infty} \subset \text{span}(\mathcal{S})$ so that $F_n \to F$ in $L^2(\mathcal{F}_T, P)$. Then for each n we have

$$F_n(\omega) = E[F_n] + \int_0^T f_n(s, \omega) dB_s(\omega)$$

where $f_n \in \mathcal{V}(0, T)$.

If $F_n, F_m \in \text{span}(\mathcal{S})$ with corresponding functions $f_n, f_m \in \mathcal{V}(0, T)$, then

$$F_n - F_m = E[F_n - F_m] + \int_0^T (f_n - f_m) dB$$

by linearity. Then we have

$$E[(F_n - F_m)^2] = E\left[\left(E[F_n - F_m] + \int_0^T (f_n - f_m) dB\right)^2\right] \tag{2.6}$$

$$= E\left[(E[F_n - F_m])^2 - 2E[F_n - F_m]\int_0^T (f_n - f_m) dB + \left(\int_0^T (f_n - f_m) dB\right)^2\right] \tag{2.7}$$

$$= (E[F_n - F_m])^2 - 2E[F_n - F_m]E\left[\int_0^T (f_n - f_m) dB\right] + E\left[\left(\int_0^T (f_n - f_m) dB\right)^2\right] \tag{2.8}$$

$$= (E[F_n - F_m])^2 + E\left[\int_0^T (f_n - f_m)^2 dt\right] \tag{2.9}$$

$$= (E[F_n - F_m])^2 + \int_0^T E[(f_n - f_m)^2] dt \tag{2.10}$$

where (2.7) is by expanding (2.6), (2.8) is by linearity and the fact that $E[F_n - F_m]$ is constant, (2.9) is using Theorem 1.3.6 (since $f_n - f_m \in \mathcal{V}$) and Theorem 1.3.5, and (2.10) is using Theorem 2.1.8. As $n, m \to \infty$, we have $E[(F_n - F_m)^2] \to 0$, and by the Cauchy-Schwarz inequality this implies that $(E[F_n - F_m])^2 \to 0$ as $n, m \to \infty$. Then

$$\int_0^T E[(f_n - f_m)^2] dt \to 0$$

as $n, m \to \infty$, and so (f_n) is a Cauchy sequence in $L^2([0,T] \times \Omega)$, and therefore converges to some $f \in L^2([0,T] \times \Omega)$ (and so the first and third conditions in the definition of functions in $\mathcal{V}$ are met). Then there is some subsequence $(f_{n_k}(t,\omega))_{k=1}^{\infty}$ converging to $f(t,\omega)$ for almost all $(t,\omega) \in [0,T] \times \Omega$, and so $f(t,\cdot)$ is $\mathcal{F}_t$-measurable for almost all t. Thus, we can modify $f(t,\omega)$ on a set of t-measure 0 so that it becomes $\mathcal{F}_t$-adapted. Therefore we get that $f \in \mathcal{V}(0,T)$.

Since $F_n \to F$ in $L^2(\mathcal{F}_T, P)$, we also have $F_n \to F$ in $L^1(\mathcal{F}_T, P)$, and Lemma 2.1.11 gives us that $\int_0^T f_n dB \to \int_0^T f dB$ in $L^2(\mathcal{F}_T, P)$, we have

$$F = \lim_{n\to\infty} F_n = \lim_{n\to\infty} (E[F_n] + \int_0^T f_n dB) = E[F] + \int_0^T f dB$$

where the limit is being taken in $L^2(\mathcal{F}_T, P)$. This gives us that (2.5) holds for all $F \in L^2(\mathcal{F}_T, P)$.

For uniqueness, suppose we have $f_1, f_2 \in \mathcal{V}(0,T)$ such that

$$F(\omega) = E[F] + \int_0^T f_1(t,\omega) dB_t(\omega) = E[F] + \int_0^T f_2(t,\omega) dB_t(\omega)$$

and so

$$\int_0^T f_1(t,\omega) - f_2(t,\omega) dB_t(\omega) = 0.$$

Then

$$0 = E\left[\left(\int_0^T f_1(t,\omega) - f_2(t,\omega) dB_t(\omega)\right)^2\right] = \int_0^T E[(f_1(t,\omega) - f_2(t,\omega))^2] dt$$

by Theorem 2.2.1 and Theorem 2.1.8, and so $f_1(t,\omega) = f_2(t,\omega)$ for almost all $(t,\omega) \in [0,T] \times \Omega$. $\square$

We now replace the general random variable by a martingale. The theorem below shows that it can be represented as an Itô process.

Theorem 2.2.2. (The martingale representation theorem) *Suppose M_t is an $\mathcal{F}_t$-martingale (with respect to P) and that $M_t \in L^2(\mathcal{F}_t, P)$ for all $t \geq 0$. Then there exists a unique stochastic process $g(s, \omega)$ such that $g \in \mathcal{V}(0, t)$ for all $t \geq 0$ and*

$$M_t(\omega) = E[M_0] + \int_0^t g(s, \omega) dB_s$$

almost surely for all $t \geq 0$.

Proof. Letting $T = t$ and $F = M_t$ from Theorem 2.2.1, we have that there exists a unique $f^{(t)}(s, \omega) \in L^2(\mathcal{F}_t, P)$ such that

$$M_t(\omega) = E[M_t] + \int_0^t f^{(t)}(s, \omega) dB_s(\omega) = E[M_0] + \int_0^t f^{(t)}(s, \omega) dB_s(\omega)$$

where the second equality is due to the fact that for all t,

$$E[M_t] = E[E(M_t \mid \mathcal{F}_0)] = E[M_0].$$

Assume $0 \leq t_1 < t_2$. Then

$$\begin{aligned} M_{t_1} &= E(M_{t_2} \mid \mathcal{F}_{t_1}) = E\left(E[M_0] + \int_0^{t_2} f^{(t_2)}(s, \omega) dB_s(\omega) \mid \mathcal{F}_{t_1}\right) \\ &= E[M_0] + E\left(\int_0^{t_2} f^{(t_2)}(s, \omega) dB_s(\omega) \mid \mathcal{F}_{t_1}\right) \\ &= E[M_0] + \int_0^{t_1} f^{(t_2)}(s, \omega) dB_s(\omega) \end{aligned}$$

where we have used the fact that

$$E\left(\int_0^{t_2} f^{(t_2)}(s, \omega) dB_s(\omega) \mid \mathcal{F}_{t_1}\right) = \int_0^{t_1} f^{(t_2)}(s, \omega) dB_s(\omega)$$

by part 5 of Theorem 1.3.6. We also have that

$$M_{t_1} = E[M_0] + \int_0^{t_1} f^{(t_1)}(s, \omega) dB_s(\omega)$$

and so

$$0 = E\left[\left(\int_0^{t_1} (f^{(t_2)} - f^{(t_1)}) dB\right)^2\right] = \int_0^{t_1} E[(f^{(t_2)} - f^{(t_1)})^2] ds$$

by Theorem 1.3.5. Thus, we have for almost all $(s, \omega) \in [0, t_1] \times \Omega$ that $f^{(t_1)}(s, \omega) = f^{(t_2)}(s, \omega)$ by uniqueness. Therefore, we can define $f(s, \omega)$ for almost all $s \in [0, \infty) \times \Omega$ by

$$f(s, \omega) = f^{(N)}(s, \omega)$$

where $s \in [0, N]$. Then we have that

$$M_t = E[M_0] + \int_0^t f^{(t)}(s, \omega) dB_s(\omega) = E[M_0] + \int_0^t f(s, \omega) dB_s(\omega)$$

for all $t \geq 0$. $\square$

CHAPTER 3

CHANGE OF MEASURE FOR BROWNIAN MOTION

Let $(X_t)_{t\geq 0}$ be an Itô process of the form

$$dX_t = \mu_t dt + \sigma_t dB_t$$

where $(B_t)_{t\geq 0}$ is a Brownian motion. As we saw in Chapter 1, $\int_0^t \sigma(s)dB_s$ is a martingale, but in general X_t is only a semi-martingale. This is due to the "drift term" $\mu_t dt$, which is undesirable in many cases. Due to this, we are interested in having some way to make this drift term disappear, thus making X_t a martingale.

As we will see in this chapter, we are able to do this by introducing an equivalent measure that eliminates drift and leaves us with a Brownian motion. The power of this theorem will be seen in Chapter 4 in several applications to mathematical finance.

3.1 Levy's Characterization of Brownian Motion

First, we give a characterization of Brownian motion using the moment generating function.

Lemma 3.1.1. *Let $(B_t)_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$, and assume $B_0 = 0$. Then the following are equivalent:*

1. *B_t is a Brownian motion;*

2. *For real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ and for $0 = t_0 < t_1 < \ldots < t_n \leq T$ we have the equality*

$$E[\exp[(\lambda_1 - \lambda_2)B(t_1) + (\lambda_2 - \lambda_3)B(t_2) + \ldots + (\lambda_{n-1} - \lambda_n)B(t_{n-1}) + \lambda_n B(t_n)]] = \prod_{k=1}^{n} \exp\left[\frac{1}{2}(t_k - t_{k-1})\lambda_k^2\right].$$

Proof. Notice that

$$E[\exp[(\lambda_1-\lambda_2)B(t_1)+(\lambda_2-\lambda_3)B(t_2)+\ldots+(\lambda_{n-1}-\lambda_n)B(t_{n-1})+\lambda_n B(t_n)]]$$

$$= E[\exp[(\lambda_1 B(t_1)+(\lambda_2(B(t_2)-B(t_1))+\ldots+\lambda_n(B(t_n)-B(t_{n-1}))]] \tag{3.1}$$

and so (3.1) is the (joint) moment generating function of

$$\mathbf{X}=(B(t_1),B(t_2)-B(t_1),\ldots,B(t_n)-B(t_{n-1})).$$

Then if we assume B is a Brownian motion, the increments $B(t_k)-B(t_{k-1})$ are independent, and so Theorem 1.1.18 part 1 gives us the result.
Now if we suppose 2, we can use Theorem 1.1.18 again to give us that B is a Brownian motion. □

We now give the Levy characterization of Brownian motion. It should be noted that the assumption that X_t and X_t^2-t are $\mathcal{F}_t$-martingales is equivalent to the assumption that the quadratic variation process $\langle X_t\rangle = t$.

Theorem 3.1.2. (Levy's characterization of Brownian motion) *Let $X_0=0$, and suppose X_t and X_t^2-t are $\mathcal{F}_t$-martingales. Then X_t is a Brownian motion.*

Proof. Put $Y_t=g(t,X_t)=\exp\left(\lambda X_t-\frac{1}{2}\lambda^2 t\right)$ and apply Theorem 1.2.6 to get

$$dY_t=\left(-\frac{\lambda^2}{2}dt+\lambda dX_t+\frac{\lambda^2}{2}dt\right)\exp\left(\lambda X_t-\frac{\lambda^2}{2}t\right)=\lambda\exp\left(\lambda X_t-\frac{\lambda^2}{2}t\right)dX_t.$$

Let $Y_0=g(0,X_0)=1$. Then

$$E(Y_t\mid\mathcal{F}_s)=E\left(Y_s+\int_s^t\lambda\exp\left(\lambda X_u-\frac{\lambda^2}{2}u\right)dX_u\mid\mathcal{F}_s\right)=Y_s$$

by Theorem 1.2.4 item 3, and so Y_t is a martingale. Now if we let λ_2 be real and $t_1 < t_2 < T$, we have

$$\begin{aligned} E(\exp(\lambda_2 X_{t_2}) \mid \mathcal{F}_{t_1}) &= E\left(\exp\left(\frac{\lambda_2^2}{2}t_2 + \lambda_2 X_{t_2} - \frac{\lambda_2^2}{2}t_2\right) \mid \mathcal{F}_{t_1}\right) \\ &= \exp\left(\frac{\lambda_2^2}{2}t_2\right) E\left(\exp\left(\lambda_2 X_{t_2} - \frac{\lambda_2^2}{2}t_2\right) \mid \mathcal{F}_{t_1}\right) \\ &= \exp\left(\frac{\lambda_2^2}{2}t_2\right) E\left(Y_{t_2} \mid \mathcal{F}_{t_1}\right) = \exp\left(\frac{\lambda_2^2}{2}t_2\right) \exp\left(\lambda_2 X_{t_1} - \frac{1}{2}\lambda_2^2 t_1\right). \end{aligned}$$

Moreover, for real values λ_1, λ_2 and $t_1 < t_2 < T$, we obtain

$$\begin{aligned} E(\exp((\lambda_1 - \lambda_2)X_{t_1} + \lambda_2 X_{t_2}) \mid \mathcal{F}_{t_1}) &= \exp((\lambda_1 - \lambda_2)X_{t_1}) E(\lambda_2 X_{t_2} \mid \mathcal{F}_{t_1}) \\ &= \exp((\lambda_1 - \lambda_2)X_{t_1}) \exp\left(\frac{\lambda_2^2}{2}t_2\right) \exp\left(\lambda_2 X_{t_1} - \frac{1}{2}\lambda_2^2 t_1\right) \\ &= \exp\left((\lambda_1 - \lambda_2)X_{t_1} + \frac{\lambda_2^2}{2}t_2 + \lambda_2 X_{t_1} - \frac{1}{2}\lambda_2^2 t_1\right) \\ &= \exp\left(\lambda_1 X_{t_1} + \frac{\lambda_2^2}{2}(t_2 - t_1)\right). \end{aligned}$$

As a result of the above, we have the base of induction

$$\begin{aligned} E\left[\exp((\lambda_1 - \lambda_2)X_{t_1} + \lambda_2 X_{t_2})\right] &= E\left[E(\exp((\lambda_1 - \lambda_2)X_{t_1} + \lambda_2 X_{t_2}) \mid \mathcal{F}_{t_1})\right] \\ &= E\left[\exp\left(\lambda_1 X_{t_1} + \frac{\lambda_2^2}{2}(t_2 - t_1)\right)\right] \\ &= \exp\left(\frac{\lambda_2^2}{2}(t_2 - t_1)\right) E(\lambda_1 X_{t_1} \mid \mathcal{F}_0) \\ &= \exp\left(\frac{\lambda_2^2}{2}(t_2 - t_1) + \frac{\lambda_1^2}{2}t_1\right). \end{aligned}$$

Now, suppose for $0 < t_1 < t_2 < \ldots < t_{n-1} < T$ and real values $\lambda_1, \ldots, \lambda_{n-1}$ we have

$$E[\exp((\lambda_1 - \lambda_2)X(t_1) + (\lambda_2 - \lambda_3)X(t_2) + \ldots + (\lambda_{n-2} - \lambda_{n-1})X(t_{n-2}) + \lambda_{n-1}X(t_{n-1}))]$$

$$= E\left[\exp\left(\sum_{k=1}^{n-2}(\lambda_k - \lambda_{k+1})X(t_k) + \lambda_{n-1}X_{t_{n-1}}\right)\right] = \prod_{k=1}^{n-1} \exp\left[\frac{1}{2}(t_k - t_{k-1})\lambda_k^2\right].$$

Then we can apply the above to find

$$
\begin{aligned}
&E\left[\exp\left(\sum_{k=1}^{n-1}(\lambda_k-\lambda_{k+1})X(t_k)+\lambda_n X(t_n)\right)\right]\\
&=E\left[E\left(\exp\left(\sum_{k=1}^{n-1}(\lambda_k-\lambda_{k+1})X(t_k)+\lambda_n X(t_n)\right)\Big|\mathcal{F}_{t_{n-1}}\right)\right]\\
&=E\left[\exp\left(\sum_{k=1}^{n-2}(\lambda_k-\lambda_{k+1})X(t_k)\right)E\left(\exp((\lambda_{n-1}-\lambda_n)X(t_{n-1})+\lambda_n X(t_n))\Big|\mathcal{F}_{t_{n-1}}\right)\right]\\
&=E\left[\exp\left(\sum_{k=1}^{n-2}(\lambda_k-\lambda_{k+1})X(t_k)\right)\exp\left(\lambda_{n-1}X_{t_{n-1}}+\frac{\lambda_n^2}{2}(t_n-t_{n-1})\right)\right]\\
&=E\left[\exp\left(\sum_{k=1}^{n-2}(\lambda_k-\lambda_{k+1})X(t_k)+\lambda_{n-1}X_{t_{n-1}}\right)\right]\exp\left(\frac{\lambda_n^2}{2}(t_n-t_{n-1})\right)\\
&=\prod_{k=1}^{n-1}\exp\left[\frac{1}{2}(t_k-t_{k-1})\lambda_k^2\right]\exp(\frac{\lambda_n^2}{2}(t_n-t_{n-1}))=\prod_{k=1}^{n}\exp\left[\frac{1}{2}(t_k-t_{k-1})\lambda_k^2\right].
\end{aligned}
$$

By Lemma 3.1.1, we have that X_t is a Brownian motion. □

3.2 Girsanov's Theorem

The following lemma will be necessary in proving theorem 3.2.2.

Lemma 3.2.1. *Let μ and ν be probability measures on the measurable space $(\Omega,\mathcal{F})$ such that $d\nu(\omega)=f(\omega)d\mu(\omega)$ for some $f\in L^1(\Omega,\mathcal{F},\mu)$. Let X be a random variable on $(\Omega,\mathcal{F})$ such that*

$$E_\nu[|X|]=\int_\Omega |X| f\,d\mu<\infty,$$

and let $\mathcal{H}\subseteq\mathcal{F}$ be a σ-algebra. Then

$$E_\nu(X\mid\mathcal{H})\cdot E_\mu(f\mid\mathcal{H})=E_\mu(fX\mid\mathcal{H})$$

almost surely.

Proof. First, note that we have for some $H \in \mathcal{H}$

$$\begin{aligned}\int_H E_\nu(X \mid \mathcal{H}) f d\mu &= \int_H E_\nu(X \mid \mathcal{H}) d\nu \\ &= \int_H X d\nu = \int_H X f d\mu \\ &= \int_H E_\nu(fX \mid \mathcal{H}) d\mu.\end{aligned}$$

Also, we have

$$\begin{aligned}\int_H E_\nu(X \mid \mathcal{H}) f d\mu &= E_\mu[E_\nu(X \mid \mathcal{H}) \cdot f \cdot I_H] \\ &= E_\mu[E_\mu(E_\nu(X \mid \mathcal{H}) \cdot f \cdot I_H \mid \mathcal{H})] \\ &= E_\mu[E_\nu(X \mid \mathcal{H}) \cdot I_H E_\mu(f \mid \mathcal{H})] \\ &= \int_H E_\nu(X \mid \mathcal{H}) \cdot E_\mu(f \mid \mathcal{H}) d\mu.\end{aligned}$$

Combine the above results to get $E_\nu(fX \mid \mathcal{H}) = E_\nu(X \mid \mathcal{H}) \cdot E_\mu(f \mid \mathcal{H})$ almost surely.
□

We now give the main result. As stated in the introduction to the chapter, this will give us a way to remove any drift in a given Itô process, leaving us with a Brownian motion.

Theorem 3.2.2. (Girsanov's theorem) *Let $(Y_t)_{t\geq 0}$ be an Itô process of the form*

$$dY_t = a(t,\omega)dt + dB_t; \quad t \leq T, Y_0 = 0$$

where $T \leq \infty$ is a given constant and $(B_t)_{t\geq 0}$ is a Brownian motion on $(\Omega, \mathcal{F}_T, P)$. Put

$$M_t = \exp\left(-\int_0^t a(s,\omega)dB_s - \frac{1}{2}\int_0^t a^2(s,\omega)ds\right); \quad t \leq T.$$

Assume that $a(s,\omega)$ satisfies Novikov's condition

$$E\left[\exp\left(\frac{1}{2}\int_0^T a^2(s,\omega)ds\right)\right] < \infty$$

where $E = E_P$ is the expectation with respect to P. Define the measure Q on $(\Omega, \mathcal{F}_T)$ by

$$dQ(\omega) = M_T(\omega)dP(\omega).$$

Then Y_t is a Brownian motion with respect to Q for $t \leq T$.

Proof. First, we find dM_t and show that M_t is a martingale. To that end, let $X_t = \ln M_t$ and apply Itô's formula to get

$$\begin{aligned} dX_t &= \left(-\frac{1}{2}a_t^2 - a_t dB_t\right) dt - a_t dB_t \\ &= -a_t dB_t - \frac{1}{2}a_t^2 dt. \end{aligned}$$

Then we can find

$$\begin{aligned} dM_t &= \exp(X_t)dX_t + \frac{1}{2}\exp(X_t)(dX_t)^2 \\ &= M_t\left(-a_t dB_t - \frac{1}{2}a_t^2 dt\right) + \frac{1}{2}M_t\left(-a_t dB_t - \frac{1}{2}a_t^2 dt\right)^2 \\ &= M_t\left(-a_t dB_t - \frac{1}{2}a_t^2 dt\right) + \frac{1}{2}M_t a_t^2 dt = -M_t a_t dB_t. \end{aligned}$$

By the above, we now have that $M_t = M_0 + \int_0^t M_u a_u dB_u$, where $M_0 = 1$. In general M_t is known to be locally square integrable, and the Itô integral must be appropriately extended, in which M_t is only guaranteed to be a local martingale. However, assuming Novikov's condition, it has been proven that M_t becomes a martingale under the extended Ito integral (see [6]).

Now, by the Levy characterization proven above (Theorem 3.1.2), it suffices to show that Y and $Y^2 - t$ are Q-martingales. We will show the former first:

Put $K_t = M_t Y_t$. Then by the Itô product rule (Corollary 1.3.9) we have

$$\begin{aligned} dK_t &= M_t dY_t + Y_t dM_t + dY_t dM_t \\ &= M_t(a_t dt + dB_t) + Y_t(-M_t a_t dB_t) + (a_t dt + dB_t)(-M_t a_t dB_t) \\ &= M_t(a_t dt + dB_t - Y_t a_t dB_t - a_t dt) \\ &= M_t(1 - Y_t a_t) dB_t \end{aligned}$$

and so $K_t = \int_0^t M_u - M_u Y_u a_u dB_u$. Then for $0 \leq s < t$

$$E(K_t \mid \mathcal{F}_s) = E\left(\int_0^t M_u - M_u Y_u a_u dB_u \mid \mathcal{F}_s\right) = \int_0^s M_u - M_u Y_u a_u dB_u = K_s$$

where Novikov's condition again guarantees that $\int_0^t M_u - M_u Y_u a_u dB_u$ (the extended Itô integral) is a martingale. Thus, we have that K_t is a martingale. Now by Lemma 3.2.1, we have

$$E_Q(Y_t \mid \mathcal{F}_s) = \frac{E(M_t Y_t \mid \mathcal{F}_s)}{E(M_t \mid \mathcal{F}_s)} = \frac{M_s Y_s}{M_s} = Y_s,$$

and so Y is indeed a martingale.

Now, we prove the latter of Levy's conditions, in a similar manner to the above. Put $Z_t = Y_t^2 - t$, and define $J_t = M_t Z_t$. Notice that

$$dZ_t = -dt + 2Y_t dY_t + \frac{1}{2} 2dt = 2Y_t dY_t$$

by the Itô formula. Then

$$\begin{aligned} dJ_t &= M_t dZ_t + Z_t dM_t + dM_t dZ_t \\ &= M_t(2Y_t dY_t) + Z_t(-M_t a_t dB_t) + (2Y_t dY_t)(-M_t a_t dB_t) \\ &= 2M_t Y_t(a_t dt + dB_t) + Z_t(-M_t a_t dB_t) + (2Y_t a_t dt + 2Y_t dB_t)(-M_t a_t dB_t) \\ &= 2Y_t M_t a_t dt + 2M_t Y_t dB_t - Z_t M_t a_t dB_t - 2Y_t M_t a_t dt \\ &= 2M_t Y_t dB_t - Z_t M_t a_t dB_t \\ &= M_t(2Y_t - Z_t a_t) dB_t. \end{aligned}$$

Thus, J_t is a martingale. Then we can follow the same logic as in the first part to get

$$E_Q(Z_t \mid \mathcal{F}_s) = \frac{E(M_t Z_t \mid \mathcal{F}_s)}{E(M_t \mid \mathcal{F}_s)} = \frac{M_s Z_s}{M_s} = Z_s$$

and so $Z_t = Y_t^2 - t$ is indeed a martingale. $\square$

CHAPTER 4
BROWNIAN MOTION IN FINANCE

Before we continue to the content, we give a simple example. Suppose a corporation wishes to buy 100 bushels of corn in one year's time, and due to budget constraints they will not be willing to pay any more than \$5 per bushel. They may then go to a farmer and sign an agreement, ensuring that they will be able to buy their corn for at most \$5 per bushel, regardless of the current market price of corn. Of course, the farmer does not wish to lose any money on this deal, and so he takes some amount of money initially.

The corporation has nothing more to think about; they will either pay less than their alloted budget or at their alloted budget. The farmer, however, has to ensure that he does not lose money in this process. Then there are two questions he has to answer:

1. How much money should I charge for the contract?

2. With this money, how can I invest it so that I can cover any losses involved in selling corn below market price?

This is the same set of questions which plagued derivative sellers on the stock market for many years. Eventually, these questions were answered due to the work of Fischer Black and Myron Scholes in the late 60's and early 70's.

Their solution to the problem involved solving explicitly a partial differential equation. In this chapter we give an alternative solution using the results established in the previous chapters.

4.1 The Standard Black-Scholes Model

First, we give the necessary background for setting up the options pricing problem. We also give our first model of a stock and bond using Brownian motion.

A *bond* is an interest bearing security (financial instrument) which can make regular interest payments and/or a lump sum payment at maturity. A bond usually has no risk associated with it. A *stock* is a security representing partial ownership of a company. A stock usually has risk associated with it.

The prices of the stock and bond can be modelled using differential equations. Since the stock price has some risk or randomness associated with it, it will involve a stochastic component. In particular, we will give a specific Itô process that behaves very similarly to real stock prices.

An *option* is a contract that gives one the right to buy (call option) or sell (put option) a stock at/by a future date for a specified fee. The *strike price* is the price at which an asset may be bought or sold under an option. The *premium* of an option is the fee paid by the buyer for the right to buy/sell an option at a certain strike price. In our example from before, the option contract was the agreement of the farmer to sell corn at the strike price of $5 per bushel. The premium would be whatever value he collected initially from the corportation.

We now give some foundational definitions from mathematical finance.

Definition 4.1.1. A *portfolio* is a pair of processes $(\varphi_t)_{t\geq 0}$ and $(\psi_t)_{t\geq 0}$ which describe respectively the number of units of stock and of bond which we hold at time t. The processes can be negative or positive. The portfolio $(\varphi_t, \psi_t)_{t\geq 0}$ must be adapted to the filtration of the underlying model associated with the stock price.

The pair of processes in the portfolio essentially gives exactly how much stock and how much bond to hold at any given time. This means that given the portfolio, stock price, and bond price, we can determine the value of your portfolio at any given time. We formalize this notion below.

Definition 4.1.2. We say a portfolio (φ_t, ψ_t) with stock price X_t, bond price B_t, and value at time t, $V_t = \varphi_t X_t + \psi_t B_t$, is *self-financing* if and only if

$$dV_t = \varphi_t dX_t + \psi_t dB_t.$$

A self-financing portfolio is one which requires no additional money after the initial formation of the portfolio. All trades are financed by buying or selling assets within the portfolio. Note that we use B to denote the bond price, not Brownian motion as before. We will generally denote Brownian motion by U or W in this chapter.

Definition 4.1.3. Given a riskless bond B, a risky stock S with volatility σ, and a claim C on events up to time T, a *replicating strategy* for C is a self-financing portfolio (φ_t, ψ_t) such that $\int_0^T \sigma_t^2 \varphi_t^2 dt < \infty$ almost surely and

$$V_T = \varphi_T S_T + \psi_T B_T = C.$$

If C gives the amount an option contract seller needs to pay off at the time T, the existence of a replicating strategy implies that the price of C at time t is given by V_t. Specifically, it allows us to determine the theoretical price of an option contract premium by calculating V_0.

The following will be our model for a given stock and bond:

Consider the following differential equations for risk-free bond price B and stock price

X:

$$dB_t = rB_t dt; \quad B_0 = 1,$$
$$dX_t = \mu X_t dt + \sigma X_t dU_t; \quad X_0 = P_0 > 0,$$

where $(U_t)_{t\geq 0}$ is a Brownian motion on $(\Omega, \mathcal{F}, P)$, r is the fixed interest rate of the bond, μ is the constant drift term, and σ is the constant nonzero volatility term. Note that we can solve dB_t to get $B_t = e^{rt}$, and so $dB_t^{-1} = -rB_t^{-1}dt$.

In the model above, X_t is said to follow a geometric Brownian motion. Many consider this to be the best model for a stock price, although it does have various shortcomings. One example of this is that the volatility is assumed to be constant, but this is generally not how volatility behaves in the real market. In the sections to come, we will explore models which do not assume constant volatility term.

We want to find a replicating strategy (φ_t, ψ_t) so that, given a strike price k, the value V of the portfolio at time T is

$$V_T = [X_T - k]_+ = \begin{cases} X_T - k & \text{if } X_T - k \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

If we can find such a portfolio, we can pay off the difference between the value of the stock price and the value of the option price (that is, the seller of the contract loses no money in the process), and determine what the premium of the option contract should be. To do this, we will follow a three-step process:

1. Find a measure Q_T on $(\Omega, \mathcal{F})$ under which Y_t is a martingale, where $Y_t = B_t^{-1}X_t = e^{-rt}X_t$ is the discounted stock price;

2. Form the best estimate of the payoff Z and find an adapted process φ_t such that $dZ_t = \varphi_t dY_t$;

3. Form a portfolio for $[X_T - k]_+$ using the above steps and verify that it is a replicating strategy.

Theorem 4.1.4. *There exists a replicating strategy (φ_t, ψ_t) for the claim $[X_T - k]_+$.*

Proof. Step 1:

First, define the stochastic differential equation

$$dW_t := \frac{\mu - r}{\sigma} dt + dU_t.$$

μ, r, and σ are constant and therefore satisfy Novikov's condition. Then by Theorem 3.2.2 (Girsanov's Theorem), W_t becomes a Brownian motion under some new measure Q_T on $(\Omega, \mathcal{F})$. Put $Y_t = B_t^{-1} X_t$. Then we have the following by Corollary 1.3.9 (Ito's product rule):

$$\begin{aligned} dY_t = d(B_t^{-1} X_t) &= X_t dB_t^{-1} + B_t^{-1} dX_t \\ &= X_t dB_t^{-1} + B_t^{-1}(\mu X_t dt + \sigma_t X_t dU_t) \\ &= -X_t r B_t^{-1} dt + X_t B_t^{-1}(\mu dt + \sigma dU_t) \\ &= X_t B_t^{-1}(-r dt + \mu dt + \sigma dU_t) \\ &= \sigma X_t B_t^{-1}\left(\frac{-r}{\sigma} dt + \frac{\mu}{\sigma} dt + dU_t\right) \\ &= \sigma Y_t dW_t. \end{aligned}$$

Thus, $Y_t = e^{-rt} X_t$ satisfies $dY_t = \sigma Y_t dW_t$ with $Y_0 = P_0$.

Step 2: Let $\mathcal{G}_t$ be the filtration of W_s, $0 \le s \le t$. Then the best estimate of the discounted payoff $e^{-rT}[X_T - k]_+$ given the observation of the stock price up to time t, X_t, and discounted at $t = 0$ is formed by

$$Z_t = E_{Q_T}(e^{-rT}[X_T - k]_+ \mid \mathcal{G}_t)$$

by Theorem 1.1.15. We have

$$E(Z_{t+s} \mid \mathcal{G}_t) = E(E(e^{-rT}[X_T - k]_+ \mid \mathcal{G}_{t+s}) \mid \mathcal{G}_t)$$
$$= E(e^{-rT}[X_T - k]_+ \mid \mathcal{G}_t) = Z_t$$

using Theorem 1.1.15 again. By definition of conditional expectation, we have that Z_t is integrable almost surely. Then Z_t is a $\mathcal{G}_t$-martingale, and so we can apply Theorem 2.2.2 (Martingale Representation Theorem) to get that there exists a unique process $g_t \in \mathcal{V}$ such that

$$Z_t = E[Z_0] + \int_0^t g_s dW_s.$$

Then we can put $\varphi_t = \frac{g_t}{\sigma Y_t}$ and get

$$Z_t = E[Z_0] + \int_0^t \varphi_s dY_s.$$

Then $dZ_t = \varphi_t dY_t$. This φ is adapted since g_t and Y_t are adapted, and φ meets the square integrability requirement since X_t is square integrable by definition of Itô process. Since this implies that X_T is square integrable, we can apply the same logic as in Remark 4.2.2 below, replacing σ_t with σ.

Step 3: Let $\psi_t = Z_t - \varphi_t Y_t$ and

$$V_t = B_t Z_t = e^{-r(T-t)} E_{Q_T}([X_T - k]_+ \mid \mathcal{G}_t).$$

Then V_t satisfies

$$V_t = B_t(\psi_t + \varphi_t Y_t) = \varphi_t X_t + \psi_t B_t.$$

Applying Corollary 1.3.9 (Ito's product rule), we get

$$dV_t = d(B_t Z_t) = Z_t dB_t + B_t dZ_t =$$
$$= (\psi_t + \varphi_t Y_t) dB_t + \varphi_t B_t dY_t$$
$$= \psi_t dB_t + \varphi_t (Y_t dB_t + B_t dY_t).$$

Now, again by Corollary 1.3.9, we have

$$Y_t dB_t + B_t dY_t = Y_t dB_t + B_t dY_t + \sigma \cdot 0dt = d(Y_t B_t) = dX_t,$$

and so $dV_t = \psi_t dB_t + \varphi_t(Y_t dB_t + B_t dY_t) = \psi_t dB_t + \varphi_t dX_t$. This implies that (φ_t, ψ_t) is self-financing. Since $V_T = [X_T - k]_+$, we have proved that (φ_t, ψ_t) is a replicating strategy for $[X_T - k]_+$. $\square$

With this result established, we have shown that for any claim there is a strategy which allows the options contract seller to incur no risk. We now give the formula for the premium, often called the Black-Scholes formula.

Theorem 4.1.5. (The Black-Scholes formula) *The theoretical option premuim price is given by the equation*

$$P_0 \Phi\left(\frac{(r + \frac{\sigma^2}{2})T + \ln \frac{P_0}{k}}{\sigma\sqrt{T}}\right) - ke^{-rT}\Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \ln \frac{P_0}{k}}{\sigma\sqrt{T}}\right).$$

Proof. We have $dX_t = \mu X_t dt + \sigma X_t dU_t$ and $dU_t = dW_t - \frac{\mu - r}{\sigma}dt$. Then

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma X_t dU_t \\ &= \mu X_t dt + \sigma X_t (dW_t - \frac{\mu - r}{\sigma} dt) \\ &= \mu X_t dt + \sigma X_t dW_t - X_t(\mu - r)dt \\ &= \mu X_t dt + \sigma X_t dW_t - \mu X_t dt + r X_t dt \\ &= r X_t dt + \sigma X_t dW_t. \end{aligned}$$

Put $A_t = \ln(X_t)$. Then an application of Itô's formula (Theorem 1.3.8) gives

$$\begin{aligned} dA_t &= \frac{1}{X_t} dX_t - \frac{\sigma^2 X_t^2}{2X_t^2} dt \\ &= \frac{1}{X_t}(rX_t dt + \sigma X_t dW_t) - \frac{\sigma^2}{2} dt \\ &= r dt + \sigma_t dW_t - \frac{\sigma^2}{2} dt \\ &= (r - \frac{\sigma^2}{2}) dt + \sigma dW_t. \end{aligned}$$

Therefore, upon integrating each side from 0 to t, we have

$$A_t - A_0 = A_t - \ln P_0 = \frac{2r - \sigma^2}{2} t + \sigma W_t$$

and so

$$A_t = \ln P_0 + \sigma W_t + (r - \sigma^2/2)t$$

so that we have

$$\exp A_t = X_t = P_0 \exp(\sigma W_t + (r - \sigma^2/2)t).$$

By Theorem 4.1.4, there exists a replicating strategy (φ_t, ψ_t) for $[X_T - k]_+$ with value at time t given by $V_t = \varphi_t X_t + \psi_t B_t$, and so to determine the theoretical option premium price we need only calculate V_0. Note that

$$V_0 = e^{-rT} E_{Q_T}([X_T - k]_+ \mid \mathcal{G}_0) = e^{-rT} E_{Q_T}[[X_T - k]_+].$$

Since W is a Brownian motion, we have that $W_t \sim N(0, t)$, and so $\ln X_T = \sigma W_T + (r - \frac{\sigma^2}{2}) + \ln P_0 \sim N((r - \frac{\sigma^2}{2})T + \ln P_0, \sigma^2 T)$. Let $m = (r - \frac{\sigma^2}{2})T + \ln P_0$. Then, noting that $e^y - k \geq 0$ only when $y \geq \ln k$, we have

$$\begin{aligned} E[[e^{\ln X} - k]_+] &= \frac{1}{\sigma\sqrt{2T\pi}} \int_{-\infty}^{\infty} [e^{\ln y} - k]_+ \exp\left(\frac{(y-m)^2}{2\sigma^2 T}\right) dy \\ &= \frac{1}{\sigma\sqrt{2T\pi}} \int_{\ln k}^{\infty} \exp\left(y - \frac{(y-m)^2}{2\sigma^2 T}\right) dy - \frac{k}{\sigma\sqrt{2T\pi}} \int_{\ln k}^{\infty} \exp\left(-\frac{(y-m)^2}{2\sigma^2 T}\right) dy. \end{aligned} \tag{4.1}$$

The standard normal cumulative distribution function is given by

$$\Phi(u) = \int_{-\infty}^{u} \phi(x)dx = \int_{-u}^{\infty} \phi(z)dz$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and $x = -z$. Then if we let $z = \frac{y-m}{\sigma\sqrt{T}}$, we have that

$$\frac{k}{\sigma\sqrt{2T\pi}} \int_{\ln k}^{\infty} \exp\left(-\frac{(y-m)^2}{2\sigma^2 T}\right) dy = k\Phi\left(\frac{m - \ln k}{\sigma\sqrt{T}}\right)$$

and if we let $z = \frac{y-\sigma^2 T - m}{\sigma\sqrt{T}}$ and use the fact that $y - \frac{(y-m)^2}{2\sigma^2 T} = -\frac{(y-\sigma^2 T - m)^2}{2\sigma^2 T} + \frac{\sigma^2 T}{2} + m$, we get

$$\frac{1}{\sigma\sqrt{2T\pi}} \int_{\ln k}^{\infty} \exp\left(y - \frac{(y-m)^2}{2\sigma^2 T}\right) dy = e^{\frac{\sigma^2 T}{2} + m} \Phi\left(\frac{m + \sigma^2 T - \ln k}{\sigma\sqrt{T}}\right).$$

Then (4.1) becomes

$$e^{\frac{\sigma^2 T}{2} + m} \Phi\left(\frac{m + \sigma^2 T - \ln k}{\sigma\sqrt{T}}\right) - k\Phi\left(\frac{m - \ln k}{\sigma\sqrt{T}}\right)$$

and since $m = (r - \frac{\sigma^2}{2})T + \ln P_0$, we have

$$\begin{aligned}
E_{Q_T}[[X_T - k]_+] &= \frac{1}{\sigma\sqrt{2T\pi}} \int_{-\infty}^{\infty} [e^x - k]_+ \exp\left(\frac{(x + (r - \frac{\sigma^2}{2})T + \ln P_0)^2}{2\sigma^2 T}\right) dx \\
&= e^{\frac{\sigma^2 T}{2} + (r - \frac{\sigma^2}{2}) + \ln P_0} \Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \ln P_0 + \sigma^2 T - \ln k}{\sigma\sqrt{T}}\right) \\
&\quad - k\Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \ln P_0 - \ln k}{\sigma\sqrt{T}}\right) \\
&= P_0 e^{rT} \Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \sigma^2 T + \ln\frac{P_0}{k}}{\sigma\sqrt{T}}\right) - k\Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \ln\frac{P_0}{k}}{\sigma\sqrt{T}}\right) \\
&= P_0 e^{rT} \Phi\left(\frac{(r + \frac{\sigma^2}{2})T + \ln\frac{P_0}{k}}{\sigma\sqrt{T}}\right) - k\Phi\left(\frac{(r - \frac{\sigma^2}{2})T + \ln\frac{P_0}{k}}{\sigma\sqrt{T}}\right).
\end{aligned}$$

Multiplying the above by e^{-rT} then gives us the Black-Scholes formula. $\square$

The above formula is still in widespread use in the stock market today.

4.2 The Generalized Black-Scholes Model

We will now generalize the above result for the case when μ, σ, and r are not constants, but can in fact be arbitrary (with some restrictions) processes themselves. Now we will find some similar results, using a new model for stock and bond prices. Consider the following differential equations for a bond price B and stock price X:

$$dB_t = r(t)B_t dt; \quad B(0) = 1,$$

$$dX_t = \mu(t,\omega)X_t dt + \sigma(t)X_t dU_t; \quad X(0) = P_0 > 0,$$

where U is a Brownian motion on $(\Omega, \mathcal{F}, P)$, μ, σ, and r satisfy those conditions necessary to make the above equations Itô processes (see Definition 1.3.7), and $\frac{\mu(t)-r(t)}{\sigma(t)}$ satisfies Novikov's condition (see Theorem 3.2.2). Again, we want to find a replicating strategy for the claim $[X_T - k]_+$ and determine the value of the portfolio at the time $t = 0$. The former is possible with the model above, but the latter requires some further restrictions on σ and r to get a desirable result.

Note that we can solve dB_t to find that $B_t = \exp(\int_0^t r(t)dt)$, and so $dB_t^{-1} = -r(t)B_t^{-1}dt$.

Theorem 4.2.1. *There exists a replicating strategy (φ, ψ) for the claim $[X_T - k]_+$ under the generalized model.*

Proof. The proof will have a similar structure to that of Theorem 4.1.4, and steps 1-3 have the same goal as before.

Step 1:

Define the stochastic differential equation

$$dW_t := \frac{\mu(t) - r(t)}{\sigma(t)}dt + dU_t.$$

We have assumed that $\frac{\mu(t)-r(t)}{\sigma(t)}$ satisfies Novikov's condition, and so Theorem 3.2.2 (Girsanov's Theorem) tells us that W_t becomes a Brownian motion under some new measure Q_T on $(\Omega, \mathcal{F})$. Put $Y_t = B_t^{-1}X_t$. Just as before apply Corollary 1.3.9 (Ito's product rule) to get

$$\begin{aligned}
dY_t = d(B_t^{-1}X_t) &= X_t dB_t^{-1} + B_t^{-1}dX_t \\
&= X_t dB_t^{-1} + B_t^{-1}(\mu_t X_t dt + \sigma_t X_t dU_t) \\
&= -X_t r_t B_t^{-1} dt + X_t B_t^{-1}(\mu_t dt + \sigma_t dU_t) \\
&= X_t B_t^{-1}(-r_t dt + \mu_t dt + \sigma_t dU_t) \\
&= \sigma_t X_t B_t^{-1}(\frac{-r_t}{\sigma_t}dt + \frac{\mu_t}{\sigma_t}dt + dU_t) \\
&= \sigma_t Y_t dW_t.
\end{aligned}$$

Thus, Y_t satisfies $dY_t = \sigma_t Y_t dW_t$ with $Y_0 = P_0$. That is, $Y_t = P_0 + \int_0^t \sigma_s Y_s dW_s$.

Step 2:

Let $\mathcal{G}_t$ be the filtration of W_s, $0 \le s \le t$. Then the best estimate of the discounted payoff $B_T^{-1}[X_T - k]_+$ given the observation of the stock price up to time t, X_t, and discounted at $t = 0$ is formed by

$$Z_t = E_{Q_T}(B_T^{-1}[X_T - k]_+ \mid \mathcal{G}_t)$$

by Theorem 1.1.15. We have

$$\begin{aligned}
E(Z_{t+s} \mid \mathcal{G}_t) &= E(E(B_T^{-1}[X_T - k]_+ \mid \mathcal{G}_{t+s}) \mid \mathcal{G}_t) \\
&= E(B_T^{-1}[X_T - k]_+ \mid \mathcal{G}_t) = Z_t.
\end{aligned}$$

using Theorem 1.1.15 again. By definition of conditional expectation, Z_t is integrable almost surely. Then Z_t is a $\mathcal{G}_t$-martingale, and so we can apply Theorem 2.2.2 (the martingale representation theorem) to get that there exists a unique process $g_t \in \mathcal{V}$ such that

$$Z_t = E[Z_0] + \int_0^t g_s dW_s.$$

Let $\varphi_t = \frac{g_t}{\sigma_t Y_t}$. Then by Step 1

$$Z_t = E[Z_0] + \int_0^t \varphi_s dY_s$$

and so $dZ_t = \varphi_t dY_t$. This φ is adapted since g_t and Y_t are adapted, and it meets the square integrability requirement if we assume that X_T is square integrable. See Remark 4.2.2 below.

Step 3: Let $\psi_t = Z_t - \varphi_t Y_t$ and

$$V_t = B_t Z_t = B_t E_{Q_T}(B_T^{-1}[X_T - k]_+ \mid \mathcal{G}_t).$$

Then V_t satisfies

$$V_t = B_t(\psi_t + \varphi_t Y_t) = \varphi_t X_t + \psi_t B_t.$$

Applying Corollary 1.3.9 (Ito's product rule), we get

$$\begin{aligned} dV_t = d(B_t Z_t) &= Z_t dB_t + B_t dZ_t = \\ &= (\psi_t + \varphi_t Y_t) dB_t + \varphi_t B_t dY_t \\ &= \psi_t dB_t + \varphi_t (Y_t dB_t + B_t dY_t). \end{aligned}$$

Now, again by Corollary 1.3.9, we have

$$Y_t dB_t + B_t dY_t = Y_t dB_t + B_t dY_t + \sigma_t \cdot 0 dt = d(Y_t B_t) = dX_t,$$

and so $dV_t = \psi_t dB_t + \varphi_t(Y_t dB_t + B_t dY_t) = \psi_t dB_t + \varphi_t dX_t$. This implies that (φ_t, ψ_t) is self-financing. Since $V_T = [X_T - k]_+$, we have proved that (φ_t, ψ_t) is a replicating strategy for $[X_T - k]_+$. $\square$

Remark 4.2.2. In the setting of Section 4.2, the volatility has the form $\sigma_t X_t$, and the self-financing portfolio (φ_t, ψ_t) must satisfy

$$\int_0^T \sigma_t^2 X_t^2 \varphi_t^2 dt < \infty$$

almost surely (Definition 4.1.3). This can be guaranteed by assuming that X_T is square integrable. By Jensen's inequality we can immediately observe that $E[Z_t^2] \leq B_T^{-2} E[(X_T - k)^2] < \infty$, and therefore, that the martingale Z_t is square integrable. In the proof of Theorem 4.2.1 we can find

$$dZ_t = \varphi_t dY_t = \varphi_t \sigma_t X_t B_t^{-1} dW_t,$$

and

$$B_t^{-1} = \exp\left(-\int_0^t r(s)ds\right) \geq B_T^{-1}$$

for $0 \leq t \leq T$. Hence, we obtain by the Itô isometry (Theorem 1.3.5)

$$\begin{aligned} E\left[\int_0^T \sigma_t^2 X_t^2 \varphi_t^2 dt\right] &\leq B_T^2 E\left[\int_0^T (\varphi_t \sigma_t X_t B_t^{-1})^2 dt\right] \\ &= B_T^2 E\left[\left(\int_0^T \varphi_t \sigma_t X_t B_t^{-1} dW_t\right)^2\right] \\ &= B_T^2 E[(Z_T - Z_0)^2] < \infty, \end{aligned}$$

which implies the square-integrability requirement. We remind the readers that $\mathcal{G}_0 = \{\varnothing, \Omega\}$, and therefore, that Z_0 is a constant (which is often denoted by $E[Z_0]$ to emphasize the non-randomness of the initial value).

Under the generalized assumptions of this model, we can also derive the following premium formula.

Theorem 4.2.3. (The generalized Black-Scholes formula) *Suppose, using our most recent model, that σ and r are deterministic functions. Then the theoretical option* *premium price is given by*

$$e^{s^2/2+m+\xi}\Phi\left(\frac{m+s^2-\ln k}{s}\right)-e^{\xi}k\Phi\left(\frac{m-\ln k}{s}\right)$$

where $\xi=-\int_0^T r_s ds$, $m=\int_0^T\left(r_s-\frac{\sigma_s^2}{2}\right)ds+\ln P_0$, *and* $s^2=\int_0^T\sigma_r^2 dr$.

Proof. The proof is similar to Theorem 4.1.5. In the same way as before, we can find that $dX_t=r_tXdt+\sigma_tX_tdW_t$ and that

$$\ln X_T=\int_0^T\left(r_s-\frac{\sigma_s^2}{2}\right)ds+\int_0^T\sigma_s dW_s+\ln P_0.$$

Since W is Brownian motion, $E[\int_0^T\sigma_s dW_s]=0$ and

$$E\left[\left(\int_0^T\sigma_s dW_s\right)^2\right]-E\left[\int_0^T\sigma_s dW_s\right]^2=E\left[\int_0^T\sigma_s^2 ds\right]=\int_0^T\sigma_s^2 ds$$

by Theorem 1.3.5 (Itô isometry). By definition,

$$\int_0^T\sigma_s dW_s=\lim_{n\to\infty}\int_0^T\phi_n(s)dW_s$$

where ϕ_n are elementary functions. Then for each n,

$$\int_0^T\phi_n(s)dW_s=\sum_{j\geq 0}e_j[W_{t_{j+1}}-W_{t_j}]$$

where e_j is constant since σ is deterministic, and so each $\int_0^T\phi_n(s)ds$ is the sum of normal variables. Then $\int_0^T\sigma_s dW_s$ is the limit of sums of normal variables, and is therefore normal itself. Thus, we have $\int_0^T\sigma_s dW_s\sim N\left(0,\int_0^T\sigma_s^2 ds\right)$ and so $\ln X_T\sim N(m,s^2)$ where $m=\int_0^T\left(r_s-\frac{\sigma_s^2}{2}\right)ds+\ln P_0$ and $s^2=\int_0^T\sigma_r^2 dr$. By Theorem 4.2.1 there is a replicating strategy for $[X_T-k]_+$, and so we need only calculate V_0. We have

$$V_0=B_0E_{\mathbb{Q}_T}(B_T^{-1}[X_T-k]_+\mid\mathcal{G}_0)=B_T^{-1}E_{\mathbb{Q}_T}[[X_T-k]_+].$$

Then

$$E_{\mathbb{Q}_T}[[X_T-k]_+]=\frac{1}{s\sqrt{2\pi}}\int_{-\infty}^{\infty}[e^{\ln y}-k]_+\exp\left(\frac{(y-m)^2}{2s^2}\right)dy.$$

We can then follow the same steps as in Theorem 4.1.5 to find that

$$E_{\mathbb{Q}_T}[[X_T - k]_+] = e^{s^2/2+m}\Phi\left(\frac{m + s^2 - \ln k}{s}\right) - k\Phi\left(\frac{m - \ln k}{s}\right)$$

then multiply by $B_T^{-1} = e^{-\int_0^T r_s ds}$ for the desired result. $\square$

Notice that in the above price we need only calculate integrals with deterministic integrands, and has nothing to do with the (possibly) random process μ.

In many applications, it is appropriate and beneficial to assume that the stock price and bond price are jointly log-normally distributed under the measure Q_T used in the theorems above. Suppose X_T and B_T are jointly log-normally distributed under Q_T, $\sigma_1^2 T$ be the variance of $\log X_T$, $\sigma_2^2 T$ be the variance of $\log B_T^{-1}$, and ρ be their correlation. Then the options contract price with strike price k can be calculated according to

$$E_{Q_T}(B_T^{-1})\left(F\Phi\left(\frac{\log\frac{F}{k} + \frac{1}{2}\sigma_1^2 T}{\sigma_1\sqrt{T}}\right) - k\Phi\left(\frac{\log\frac{F}{k} - \frac{1}{2}\sigma_1^2 T}{\sigma_1\sqrt{T}}\right)\right)$$

where $F = \frac{E_{Q_T}(B_T^{-1}X_T)}{E_{Q_T}(B_T^{-1})} = \exp(\rho\sigma_1\sigma_2 T)E_{Q_T}(X_T)$. This form is computationally easier to solve than that given in Theorem 4.2.3. For further discussion see Chapter 6 of [1].

4.3 Diffusion Model

In the previous two sections, we verified that a replicating strategy does exist under the Black-Scholes model for a stock and bond price. However, in proving this existence, we relied upon the martingale representation theorem, and no explicit formula for the replicating strategy is given. Generally, this formula is impossible to express explicitly. In this section, we explore some additional assumptions which allow us to find an explicit formula for the replicating strategy. First, we must introduce diffusion processes.

Definition 4.3.1. A time-homogeneous *Itô diffusion* is a stochastic process $X_t(\omega) = X(t, \omega) : [0, \infty) \times \Omega \to \mathbb{R}$ satisfying a stochastic differential equation of the form

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t; \quad t \geq s, X(s) = x,$$

where B is a Brownian motion, and b, σ satisfy

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq D\,|x - y|$$

for all $x, y \in \mathbb{R}$ and some constant D.

Note that b and σ depend only on x, but not on t.

Definition 4.3.2. Let $f \in C_0^2(\mathbb{R})$. The *generator A of an Itô diffusion* $(X_t)_{t\geq 0}$ is defined as

$$Af(x) = \lim_{t\downarrow 0} \frac{E(f(X_t) \mid X(0) = x) - f(x)}{t}, \quad x \in \mathbb{R}.$$

The set of all functions $f \in C_0^2(\mathbb{R})$ for which the limit exists for all $x \in \mathbb{R}$ is denoted by $\mathcal{D}_A$.

Theorem 4.3.3. (Kolmogorov's backward equation) *Let $f \in C_0^2(\mathbb{R})$ and let $(X_t)_{t\geq 0}$ be an Itô diffusion of the form*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t; \quad t \geq 0, X(0) = x.$$

Define $u(t, x) = E(f(X_t) \mid X(0) = x)$. Then $u(t, \cdot) \in \mathcal{D}_A$ for each t and

$$\frac{\partial u}{\partial t} = Au; \quad t > 0, x \in \mathbb{R};$$
$$u(0, x) = f(x); \quad x \in \mathbb{R}$$

where the right hand side is interpreted as A applied to the function $x \mapsto u(t, x)$. Furthermore, letting $g(x) = u(t, x)$, we have

$$Ag(x) = b(x)\frac{\partial g}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2 g}{\partial x^2}.$$

Proof. For proof, see Theorem 7.3.3 and Theorem 8.1.1 of [10]. $\square$

With the above in mind, we again change our model, this time letting the stock and bond prices be Itô diffusion processes. That is, for bond price B and stock price X,

$$dB_t = r(B_t)dt; \quad B(0) = 1$$

$$dX_t = \mu(X_t)dt + \sigma(X_t)dU_t; \quad X(0) = x$$

where $r(x) = r \cdot x$ for constant r, the processes μ and σ satisfy the conditions in the definition of Itô diffusion, and $\frac{\mu(X_t)-r(X_t)}{\sigma(X_t)}$ satisfies Novikov's condition (see Theorem 3.2.2). We have $B_t = e^{rt}$ and $dB_t^{-1} = -rB_t^{-1}dt$ just as in the first section.

We already know that under this model, there exists a replicating strategy by Theorem 4.2.1. However, the theorem below gives an explicit construction of the functions, and also does not rely upon the martingale representation theorem.

Theorem 4.3.4. *There exists a replicating strategy (φ, ψ) for the claim $[X_T - k]_+$ under the diffusion model. The processes φ and ψ are given by*

$$\varphi(t, X_t) = g_x(t, X_t)B_t$$

and

$$\psi(t, X_t) = Z_t - g_x(t, X_t)X_t$$

where

$$g_x(t, X_t) = \frac{\partial}{\partial x}(E_{Q_T}(B_T^{-1}[X_{T-t} - k]_+ \mid X_0 = x))$$

and

$$Z_t = E_{Q_T}(B_T^{-1}[X_{T-t} - k]_+ \mid X_0 = x).$$

Proof. As usual, define the stochastic differential equation

$$dW_t := \frac{\mu(X_t) - r(X_t)}{\sigma(X_t)} dt + dU_t; \quad W_0 = 0.$$

We have assumed that $\frac{\mu(X_t)-r(X_t)}{\sigma(X_t)}$ satisfies Novikov's condition, and so Theorem 3.2.2 (Girsanov's Theorem) tells us that W_t becomes a Brownian motion under some new measure Q_T on $(\Omega, \mathcal{F})$. Then

$$\begin{aligned} dX_t &= \mu(X_t)dt + \sigma(x_t)dU_t = \mu(X_t)dt + \sigma(X_t)\left(dW_t - \frac{\mu(X_t) - r(X_t)}{\sigma(X_t)}dt\right) \\ &= r(X_t)dt + \sigma(X_t)dW_t. \end{aligned}$$

Note that $X_0 = x$. Now, let $h(x) = B_T^{-1}[x-k]_+$ and introduce

$$\begin{aligned} &w(t,x) = E_{Q_T}(h(X_t) \mid X_0 = x); \quad 0 \le t \le T, \\ &g(t,x) = w(T-t,x). \end{aligned}$$

Then we can apply Theorem 4.3.3 to find that

$$\frac{\partial w}{\partial t}(t,x) = r(x)\frac{\partial w}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2 w}{\partial x^2}$$

We can then set $s(t) = T - t$, and so $g(t,x) = w(s(t),x)$. Then apply the chain rule to get

$$\frac{\partial g}{\partial t} = \frac{\partial w}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} = -\frac{\partial w}{\partial s},$$

and so

$$\frac{\partial g}{\partial t}(t,x) = -r(x)\frac{\partial g}{\partial x} - \frac{1}{2}\sigma^2(x)\frac{\partial^2 g}{\partial x^2}.$$

Let $Z_t = g(t, X_t) = w(T-t, X_t)$, and denote $g_t = \frac{\partial g}{\partial t}(t, X_t)$, $g_x = \frac{\partial g}{\partial x}(t, X_t)$, and $g_{xx} = \frac{\partial^2 g}{\partial x^2}(t, X_t)$. Apply the Itô formula (Theorem 1.3.8) under P to Z to get

$$\begin{aligned} dZ_t &= g_t dt + g_x dX_t + \frac{1}{2} g_{xx}(dX_t)^2 \\ &= g_t dt + g_x(\mu(X_t)dt + \sigma(X_t)dU_t) + \frac{1}{2} g_{xx}\sigma^2(X_t)dt \\ &= (-r(X_t)g_x - \frac{1}{2}\sigma^2(X_t)g_{xx}dt + g_x(\mu(X_t)dt + \sigma(X_t)dU_t) + \frac{1}{2} g_{xx}\sigma^2(X_t)dt \\ &= g_x(\mu(X_t) - \sigma(X_t))dt + g_x\sigma(X_t)dU_t \\ &= g_x\sigma(X_t)dW_t \end{aligned}$$

and so $dZ_t = g_x\sigma(X_t)dW_t$.

We then have

$$Z_0 = g(0, X_0) = w(T, x) = E_{Q_T}(h(X_T) \mid X_0 = x)$$

and

$$Z_T = E_{Q_T}(h(X_0) \mid X_0 = X_T) = h(X_T).$$

Therefore, we conclude that

$$Z_T = h(X_T) = Z_0 + \int_0^T g_x\sigma(X_t)dW_t.$$

Now, let $V_t = B_t Z_t$, $\varphi_t = g_x B_t$, and $\psi_t = Z_t - g_x X_t$. We can see that φ_t is adapted, and it meets the square integrability requirement (see Remark 4.3.5). Then

$$\varphi_t X_t + \psi_t B_t = g_x B_t X_t + (Z_t - g_x X_t)B_t = Z_t B_t = V_t.$$

Introduce $Y_t = B_t^{-1} X_t$. Then under Q_T, we have by the Itô product rule (Corollary 1.3.9)

$$\begin{aligned} dY_t &= d(B_t^{-1} X_t = d(B_t^{-1})X_t + B_t^{-1} dX_t \\ &= -rB_t^{-1} dt X_t + B_t^{-1}(rX_t + \sigma(X_t)dW_t) \\ &= \sigma(X_t)B_t^{-1} dW_t. \end{aligned}$$

Then $dZ_t = g_x\sigma(X_t)dW_t = \varphi_t B_t^{-1}\sigma(X_t)dW_t = \varphi_t dY_t$. We also have $\varphi_t Y_t = g_x B_t B_t^{-1} X_t = g_x X_t$, and so $Z_t = \psi_t + \varphi_t Y_t$. Apply the Itô product rule (Corollary 1.3.9) to get

$$\begin{aligned} dV_t = dB_t Z_t &= Z_t dB_t + B_t dZ_t \\ &= (\psi_t + \varphi_t Y_t)dB_t + B_t(\varphi_t dY_t) \\ &= \psi_t dB_t + \varphi_t(Y_t dB_t + B_t dY_t). \end{aligned}$$

We can then use Corollary 1.3.9 again to get

$$Y_t dB_t + B_t dY_t = d(Y_t B_t) = dX_t,$$

and so

$$dV_t = \psi_t dB_t + \varphi_t dX_t.$$

Therefore, we see that the portfolio (φ, ψ) is self-financing. Now since

$$V_T = Z_T B_T = h(X_T)B_T = [X_T - k]_+,$$

the pair (φ, ψ) is indeed a replicating strategy for the claim, and each was explicitly given during our construction. $\square$

Remark 4.3.5. Just as in Remark 4.2.2, under the assumption that X_T is square integrable, we can see that Z_t is square integrable. In the above proof, we find that

$$Z_T - Z_0 = \int_0^T g_x \sigma(X_t) dW_t$$

and $B_t \leq B_T$ for $0 \leq t \leq T$. Apply the Itô isometry (Theorem 1.3.5) to obtain

$$\begin{aligned}
E\left[\int_0^T (\varphi_t \sigma(X_t))^2 dt\right] &= E\left[\int_0^T (g_x B_t \sigma(X_t))^2 dt\right] \\
&\leq B_T^2 E\left[\int_0^T (g_x \sigma(X_t))^2 dt\right] \\
&= B_T^2 E\left[\left(\int_0^T g_x \sigma(X_t) dW_t\right)^2\right] \\
&= B_T^2 E[(Z_T - Z_0)^2] < \infty,
\end{aligned}$$

and so the square integrability requirement is met.

CHAPTER 5
FUTURE RESEARCH

In this chapter we give possible avenues for future research. We discuss shortcomings of the current models we have explored, possible new models, and applications of our results. We first give an overview of stochastic volatility models with possible applications in cryptocurrency, then discuss possibilities for using explicit form of replicating strategy functions to make decisions in a portfolio.

No topic here is discussed in depth, but potential papers of interest are cited within.

5.1 Further Generalizations of the Black-Scholes Model

In Chapter 4, we explored several models for a stock and bond price. In the first model, the classical Black-Scholes model, it is assumed that the volatility term σ is constant. The second model does not assume this, but does assume that σ is a deterministic function of t. The third model can be viewed as a special case of the second, this time taking σ to be a deterministic function of the stock price.

While each of these cases leads to the highly (theoretically) desirable result that there exists a replicating strategy for the claim $[X_T - k]_+$, research has shown that these assumptions may be unrealistic for the true behavior of volatility in the market [9].

In future research we would like to consider models which allow volatility to take the form of a stochastic process, referred to as a stochastic volatility model. That is, for a stock price X, the price should follow the stochastic differential equation given by

$$dX_t = \mu X_t dt + \sigma(t, \omega) X_t dU_t.$$

Complications arise when we are asked how to define the form of the stochastic volatility.

Lorig and Sircar [9] give an overview of various stochastic volatility models which have appeared in research. One common issue is that there is generally no way to show that a replicating strategy exists in stochastic volatility models. We wish to determine what conditions must be met in order to derive a closed form solution for the option price, despite this complication.

Ways to accurately estimate this closed form are also a possibility for further research. In particular, we may explore Markov Chain Monte Carlo (MCMC) methods for various estimations under the stochastic volatility model. According to [5], MCMC methods appear to be among the most efficient and accurate under the assumptions made in stochastic volatility models. An overview of MCMC methods in options pricing is given in [7].

A particularly interesting model, stochastic volatility with jumps, seems to be well suited to price cryptocurrency options [5]. Cryptocurrency derivatives are of particular interest due to the fact that they have not been available to trade for very long. The underlying assets tend to be highly volatile and experience large swings due in large part to consumer speculation. Possible models and ways to determine option prices is a new and exciting problem which we wish to explore further.

5.2 Implementation of Replicating Strategies

In Theorem 4.1.4 and Theorem 4.2.1, we show the existence of a replicating strategy for the claim $[X_T - k]_+$ under their respective model assumptions. No construction of the processes in the portfolio is given due the our reliance upon Theorem 2.2.2 (the martingale representation theorem).

However, in Theorem 4.3.4, an explicit form is given. We are interested in determining the computational difficulty in finding specific values of the given processes, and also implementing them in a dynamic fashion. That is, given an initial investment, use the replicating portfolio to determine the exact amounts of stock and bond to hold at each instance so that you end at time T with exactly the amount of money needed to pay the difference.

In [1], another potential method for estimating the value of φ_t is given. Assuming you have some method of approximating V_t, the process φ_t can be approximated using

$$\varphi_t \approx \frac{\Delta V_t}{\Delta S_t}$$

where Δ represents a small interval of time $(t, t + \Delta t)$. We recall that, in general, we can find

$$V_t = B_t E_{Q_T}(B_T^{-1}[X_T - k]_+ \mid \mathcal{G}_t),$$

which is again difficult to compute in the case of a complicated model. Therefore, it is of interest to know the best methods for approximating V_t under different assumptions on the way X behaves.

REFERENCES

[1] Baxter, M. W., & Rennie, A. J. O. (1996). *Financial calculus*. Cambridge University Press.

[2] Billingsley, P. (1995). *Probability and measure*. Wiley.

[3] Folland, G. B. (1984). *Real analysis*. Wiley.

[4] Hogg, R. V., McKean, J., & Craig, A. T. (2012). *Introduction to mathematical statistics*. Pearson Education Inc.

[5] Hou, A. J., Wang, W., Chen, C. Y., & Hardle, W. K. (2020). Pricing cryptocurrency options. *Journal of Financial Econometrics 18: 250-279*.

[6] Ikeda, N., & Watanabe, S. (1981). *Stochastic differential equations and diffusion processes*. North-Holland Publishing Company.

[7] Johannes, M., & Polson, N. (2009). Mcmc methods for financial econometrics. *Handbook of Financial Econometrics 2: 1-72*.

[8] Karr, A. F. (1993). *Probability*. Springer-Verlag New York.

[9] Lorig, M., & Sircar, R. (2015). *Stochastic volatility: Modeling and asymptotic approaches to option pricing & portfolio selection* (tech. rep.). Princeton University, ORFE Department.

[10] Øksendal, B. (2003). *Stochastic differential equations*. Springer-Verlag Heidelberg New York.

[11] Royden, H. L. (1988). *Real analysis*. Prentice Hall.

VITA

Seth Agee was born in Lebanon, Tennessee, on May 22, 1996. He attended Gordonsville Elementary School in Smith County, Tennessee, and graduated from Gordonsville High School in 2015 as the valedictorian. The following August, he enrolled as an undergraduate at Tennessee Technological University. In May of 2019, he graduated with a Bachelor of Science in Mathematics and a minor in Computer Science. He returned to Tennessee Technological University the following fall semester and received a Master of Science degree in Mathematics in May of 2021.