STOCHASTIC CALCULUS APPLIED TO ARBITRAGE-FREE OPTIONS PRICING

EDWIN SURESH

Abstract. We provide a thorough explanation of stochastic (random) processes, specifically Brownian motion, before rigorously introducing Itô Calculus. This allows us to evaluate integrals with respect to Brownian motion and solve stochastic differential equations (SDE). We then develop the SDE and measure-theory machinery to derive the famous Black-Scholes model for pricing options. A detailed overview of the Girsanov Theorem, the Feynman-Kac Formula, and the concept of arbitrage is included to tie together intuition, theory, and application.

Contents

Introduction

Ordinary calculus is characterized by differentiable functions; however, many random processes are necessarily non-differentiable, having "rough" paths. If they were in fact differentiable, that would allow one to predict how the process would change over time before observing it. This renders ordinary calculus ineffective for stochastic processes, which appear frequently in real-world applications. However, we can still describe how a stochastic process changes or accumulates over time by extending our usual notions of calculus to Stochastic Calculus.

One framework for doing so is Itô Calculus, named after Kiyosi Itô who developed much of its fundamental theory and techniques. At the heart of Itô Calculus is the Itô Integral and Itô's Formula. The former gives precise meaning to integrating with respect to Brownian motion, which is a continuous but non-differentiable stochastic process. The latter is the equivalent of the chain rule for Stochastic Calculus, and has wide-ranging applications.

Most notably, Itô Calculus is essential in the derivation of the extensively-used Black-Scholes Equation and Black-Scholes Formula, which comprise the Black-Scholes Model for pricing financial derivatives. The partial differential equation describes how prices changes over time assuming that arbitrage – risk-free profit – is impossible. The formula allows one to calculate the price of an option given certain information. This gave much more credibility and popularity to options markets, accelerating the fields of financial economics and quantitative finance.

In this paper, we first explain in detail the fundamentals of probability theory and measure theory, which may be review to some but nonetheless provides the necessary background for Itô Calculus. We introduce Brownian motion, which models continuous random movement, and proceed to develop the calculus theory. We examine common stochastic differential equations in finance and devise intuitive methods for changing probability measures, developing useful machinery for more-involved financial applications. We are then able to derive the Feynman-Kac Formula, Black-Scholes Equation, and Black-Scholes Formula while ensuring that we have clear intuition for how and why we apply various Itô Calculus and measure theory results.

1. Probability Spaces and Random Variables

We introduce the basics of theoretical probability before delving into more advanced applications in later sections. This section mainly consists of essential definitions that many readers may be familiar with.

We use a *probability space* to observe and analyze some procedure (i.e., experiment, observation, process) that has various possible outcomes. The realized outcome of the procedure is in some way chosen "randomly" rather than deterministically.

Definition 1.1 (Sample Space)**.** A *sample space* Ω is a nonempty set of all outcomes of a procedure.

After a procedure has taken place, exactly one outcome must have taken place. In many cases, the probability of an outcome ω may be zero but ω is still a possible outcome. As an example, when picking a random number from $[-1, 1]$, the probability of picking any particular number is zero. Oftentimes, *events*, or sets of zero or more outcomes, are more natural for defining probabilities. In the same example, the probability of picking a positive number is one half. The use of *σ-algebras* helps us avoid such issues. In probability theory, the σ-algebra is the set of relevant events or "event space," where each event can be assigned a probability.

Definition 1.2 (σ-Algebra)**.** A *σ-algebra* $\mathcal{F}$ on a set Ω is a collection of subsets of Ω satisfying the following properties:

- $\Omega \in \mathcal{F}$.
- If $A \in \mathcal{F}$ then $A^C := \Omega \setminus A \in \mathcal{F}$. (Closed under Complement)
- If $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. (Closed under Countable Unions)

As a consequence of these properties, we know that a σ-algebra contains the empty set and is closed under countable intersections. Also, we know that the power set 2^Ω is a σ-algebra on Ω.

We often use a function σ, where for a collection F of subsets of Ω, $\sigma(F)$ gives the smallest σ-algebra on Ω such that $F \subset \sigma(F)$. It is defined as the intersection of all σ-algebras on Ω that contain F; the result is called the σ-algebra *generated* by F.

The *Borel σ-algebra* $\mathcal{R}$ is the σ-algebra generated by all open sets of $\mathbb{R}$. Elements of $\mathcal{R}$ are called Borel sets. Open sets, closed sets, countable unions of open or closed sets, and countable intersections of open or closed sets are all Borel sets, although this is not an exhaustive description.

Oftentimes, $\mathcal{F}$ will not be explicitly defined. We can think of the σ-algebra $\mathcal{F}$ as the "information" available to us. Event E is an element of $\mathcal{F}$ if and only if we can determine whether or not any given outcome ω belongs to E. In other words, given any ω, we can determine whether or not any event $E \in \mathcal{F}$ has occurred. So, $\mathcal{F}$ represents what questions we can answer from the procedure.

Definition 1.3 (Probability Measure)**.** A *probability measure* $\mathbb{P}$ is a function $\mathbb{P}\colon \mathcal{F} \to [0, 1]$ that satisfies the following properties:

- $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$.
- If $E_1, E_2, \ldots \in \mathcal{F}$ are pairwise disjoint, then $\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$.

Together, these make up a *probability space*, denoted $(\Omega, \mathcal{F}, \mathbb{P})$ which is used to model a procedure. Note that we can define different probability measures on the same sample space to define two different scenarios. For example, a fair coin flip and an unfair coin flip will have the same sample space and σ-algebra, with different probability measures.

An event E is *$\mathcal{F}$-measurable* if $E \in \mathcal{F}$. So, the event is measurable if we can tell whether or not the event occurred based on the information available, that is, the σ-algebra. This allows us to measure the probability of the event occurring. A function X defined on Ω is $\mathcal{F}$-measurable if for any Borel set $B \in \mathcal{R}$, the pre-image $X^{-1}(B) = \{X \in B\}$ is $\mathcal{F}$-measurable. This allows us to *measure*, or evaluate, $\mathbb{P}(X \in B)$ which is the probability that an outcome ω occurred such that $X(\omega) \in B$. We call such a function a *random variable*.

Note that we write $\{X \in B\}$ as short-hand for $\{\omega \in \Omega : X(\omega) \in B\}$. Also, we write $\mathbb{P}(X \in B)$ as short-hand for $\mathbb{P}(\{X \in B\})$.

Definition 1.4 (Random Variable)**.** A function $X\colon \Omega \to \mathbb{R}$ is a *random variable* on $(\Omega, \mathcal{F}, \mathbb{P})$ if it is $\mathcal{F}$-measurable. Explicitly, X is a random variable if

$$\{X \in B\} \in \mathcal{F}$$

for all $B \in \mathcal{R}$.

We can extend the function σ to random variables: $\sigma(X)$ is the smallest σ-algebra such that X is measurable. Explicitly,

$$\sigma(X) := \{\{X \in B\}\colon B \in \mathcal{R}\}.$$

Given an event $E \in \mathcal{F}$, one commonly used random variable is the *indicator random variable* 1_E, given by

$$1_E(\omega) := \begin{cases} 1 & \omega \in E, \\ 0 & \omega \notin E. \end{cases}$$

Each random variable has a distribution function, a distribution (i.e. induced measure), and oftentimes a density.

Definition 1.5 (Distribution Function)**.** The *distribution function* $F\colon \mathbb{R} \to [0,1]$ of a random variable X on Ω is defined by

$$F(x) := \mathbb{P}(X \leq x).$$

If F is continuous, then X is a *continuous random variable.*

Definition 1.6 (Density and Distribution)**.** Let X be a continuous random variable. If there exists a function f such that

$$F(x) = \int_{-\infty}^{x} f(y)dy,$$

then f is the *density* of X. The *distribution* of X is the *induced measure* $\mathbb{P}_X$ on $\mathbb{R}$ where $\mathbb{P}_X$ is defined by $\mathbb{P}_X(A) = \mathbb{P}(X \in A)$ where A is Borel. If f is the density of X, then

$$\mathbb{P}_X(A) = \int_A f(x)dx.$$

Definition 1.7 (Independence)**.** Two events E_1 and E_2 are *independent* if

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2).$$

Two random variables X and Y are independent if for all Borel $A, B \in \mathcal{R}$, the events $E_1 = \{X \in A\}$ and $E_2 = \{Y \in B\}$ are independent. Two σ-algebras $\mathcal{F}$ and $\mathcal{G}$ are independent if for all $E_1 \in \mathcal{F}$, $E_2 \in \mathcal{G}$, events E_1 and E_2 are independent.

It is very useful to consider the *expectation* (i.e., expected value) of a random variable. It can be thought of as a weighted average of all possible outcomes, with each being weighted by its likelihood.

Definition 1.8 (Expectation)**.** The *expectation* of a continuous random variable X is given by

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

If X has density f, then

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

In contrast, a discrete random variable X that only takes on values $x_1, x_2, \ldots, x_n$ has no density function. Its expectation is given by

$$\mathbb{E}[X] := \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

Definition 1.9 (Moment-Generating Function)**.** The *moment-generating function* of a random variable X is given by

$$m(t) := \mathbb{E}[e^{tX}].$$

If two random variables have the same moment-generating function, they are said to be identically distributed.

Definition 1.10 (Conditional Expectation)**.** If X is a random variable on probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{F}_0 \subset \mathcal{F}$ is a sub-σ-algebra, then the *conditional expectation* $E[X \mid \mathcal{F}_0]$ is the unique random variable satisfying the following properties:

- $E[X \mid \mathcal{F}_0]$ is $\mathcal{F}_0$-measurable.
- If $E \in \mathcal{F}_0$, then $\mathbb{E}[E[X|\mathcal{F}_0]1_E] = \mathbb{E}[X1_E]$.

If Y is another random variable, then $E[X|Y] := E[X|\sigma(Y)]$.

Definition 1.11 (Variance)**.** The *variance* of X is given by

$$\mathrm{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Definition 1.12 (Standard Normal Distribution)**.** A random variable with *standard normal distribution* has distribution function

$$\Phi(b) := \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

From this we can see that the standard normal distribution has density

$$\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

If X has a normal distribution with mean μ and variance σ^2, we write $X \sim N(\mu, \sigma^2)$. In this case, X has moment-generating function $m(t) = e^{\mu t} e^{\sigma^2 t^2/2}$.

Proposition 1.13. *If X, Y are independent $N(0,1)$ random variables, then*

$$Z = \frac{X+Y}{\sqrt{2}},\ W = \frac{X-Y}{\sqrt{2}}$$

are independent $N(0,1)$ random variables.

For the proof, see Proposition 2.2.1 on page 39 of [6].

Theorem 1.14 (Central Limit Theorem)**.** *Let $X_1, X_2, \dots$ be independent, identically distributed random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$. Let*

$$Z_n := \frac{(X_1 + \cdots + X_n) - n\mu}{\sigma\sqrt{n}}.$$

Then as $n \to \infty$, the distribution of Z_n approaches a standard normal distribution. I.e., if $a < b$, then

$$\lim_{n\to\infty} \mathbb{P}(a \le Z_n \le b) = \Phi(b) - \Phi(a).$$

For the proof, see Section 3 of [9].

2. Stochastic Processes

The rest of this paper will heavily involve *stochastic process.* We define this and important concepts to analyze them, focusing on martingales.

Definition 2.1 (Stochastic Process)**.** A *stochastic process* X_t is a collection of random variables indexed by time $t \in T \subset \mathbb{R}$. If T is an interval in $\mathbb{R}$ then time is considered *continuous*; if T is a countable set in $\mathbb{R}$ then time is considered *discrete.*

We may think of this explicitly as a collection $\{X_t\}_{t\in T}$ of random variables, or as a random variable whose value is a function $X : T \to \mathbb{R}$. We can also consider stochastic processes $Y(t, X_t)$ that depend on both $t \in T$ and a random variable X_t from another stochastic process.

Definition 2.2 (Filtration)**.** Let X_t be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. For all $t \in T$, let $\mathcal{F}_t \subset \mathcal{F}$ be a σ-algebra on Ω such that if $r < s$, then $\mathcal{F}_r \subset \mathcal{F}_s$. Then the collection of increasing σ-algebras $\{\mathcal{F}_t\}$ is a filtration.

For stochastic process X_t and filtration $\{\mathcal{F}_t\}$, we think of the σ-algebra $\mathcal{F}_t$ as the "information" from X_s for $0 \leq s \leq t$. The condition that if $r < s$, then $\mathcal{F}_r \subset \mathcal{F}_s$ means that no information is lost over time; as time progresses, we can answer the same questions as before, plus additional ones. Note from the definition of filtration that there is some flexibility in defining the σ-algebra for some $t \in T$. In this paper however, we will be using the *natural filtration* unless otherwise specified. The natural filtration holds all information available from the process so far, and nothing more. It is constructed so that each σ-algebra $\mathcal{F}_t$ is the smallest one such that X_t is $\mathcal{F}_t$-measurable.

Definition 2.3 (Natural Filtration)**.** Let X_t be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. The natural filtration $\{\mathcal{F}_t\}$ is the filtration where

$$\mathcal{F}_t := \sigma(\{X_s\}_{s\leq t}).$$

If we have multiple stochastic processes X_t and Y_t, then

$$\mathcal{F}_t := \sigma(\{X_s\}_{s\leq t}, \{Y_s\}_{s\leq t}).$$

If X_t is a stochastic process such that X_t is $\mathcal{F}_t$-measurable for all t, then we say X_t is *adapted* to filtration $\{\mathcal{F}_t\}$. It may seem like this is true for any process. When we use a natural filtration, as we do in this paper, this in fact is the case; more generally, however, it is not.

Proposition 2.4. *Let X_t be a stochastic process with respect to filtration $\{\mathcal{F}_t\}_{t\in T}$ and let Y be a random variable. Then the following properties hold:*

- *If Y is $\mathcal{F}_t$-measurable, then $E[Y|\mathcal{F}_t] = Y$.*
- *For any $\mathcal{F}_t$-measurable event E, $\mathbb{E}[E[Y|\mathcal{F}_t]1_E] = \mathbb{E}[Y 1_E]$. Letting $E = \Omega$, we get that $\mathbb{E}[E[Y|\mathcal{F}_t]] = \mathbb{E}[Y]$.*
- *If $\{X_s\}_{s\leq t}$ random variables are independent of Y, then $\mathcal{F}_t$ has no information about Y, so $E[Y|\mathcal{F}_t] = \mathbb{E}[Y]$.*
- *$E[aY + bZ|\mathcal{F}_t] = aE[Y|\mathcal{F}_t] + bE[z|\mathcal{F}_t]$ for random variables Y, Z and constants a, b. (Linearity)*
- *If $s < t$, then $E[E[Y|\mathcal{F}_t]|\mathcal{F}_s] = E[Y|\mathcal{F}_s]$. (Tower Property)*
- *If Z is an $\mathcal{F}_t$-measurable random variable, then $E[YZ|\mathcal{F}_t] = ZE[Y|\mathcal{F}_t]$.*

For the proof, see Proposition 1.1.1 on page 6 of [6].

Definition 2.5 (Martingale)**.** A stochastic process M_t is a *martingale* with respect to filtration $\{\mathcal{F}_t\}$ if:

- For all t, M_t is an $\mathcal{F}_t$-measurable random variable with $\mathbb{E}[|M_t|] < \infty$.
- If $s < t$, then

$$E[M_t|\mathcal{F}_s] = M_s.$$

The second condition can be equivalently expressed as

$$E[M_t - M_s | \mathcal{F}_s] = 0.$$

When M_t is a martingale, we know

$$\mathbb{E}[M_t] = \mathbb{E}\left[E[M_t|\mathcal{F}_0]\right] = \mathbb{E}[M_0] = M_0.$$

A martingale is typically meant to model a "fair game." We can think of M_t as the price of an asset or the winnings in a game, where regardless of past prices or past winnings before time t, the expected change from time t to any time $s > t$ is 0.

3. Example: Simple Random Walk

In this section, we explore the simple random walk to see concrete examples of concepts defined in the previous sections. The simple random walk is closely related to Brownian motion, which we will examine in great detail in the next section. In particular, we aim to make σ-algebras and filtrations more understandable using concrete examples.

Consider an infinite stochastic process of identically distributed random variables $X_1, X_2, \ldots$ where $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Then

$$\Omega = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_i = -1 \text{ or } \omega_i = 1 \text{ for all } i\}.$$

It is helpful to consider the finite process with $X_1, X_2, \ldots, X_n$ and

$$\Omega_n = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) : \omega_i = -1 \text{ or } \omega_i = 1 \text{ for all } i\}.$$

The sample space Ω_n has 2^n outcomes, each equally likely. On this finite sample space, we use the power set 2^{Ω_n} as our σ-algebra . The probability measure function on $(\Omega_n, 2^{\Omega_n})$, which we denote $\mathbb{P}_n \colon 2^{\Omega_n} \to [0,1]$, is given by

$$\mathbb{P}_n(E) := \sum_{\omega \in E} p_n(\omega)$$

where

$$p_n(\omega) := \frac{1}{2^n}, \omega \in \Omega_n.$$

We let $\mathcal{F}_n$ be the collection of subsets E of Ω such that there exists $E' \in 2^{\Omega_n}$ satisfying $E = \{\omega : (\omega_1, ..., \omega_n) \in E'\}$. In fact, each $\mathcal{F}_n$ is the σ-algebra generated by the random variables $X_1, X_2, \ldots$; we are actually forming the natural filtration explicitly. Note that the σ-algebra $\mathcal{F}_n$ and the σ-algebra 2^{Ω_n} are on different sample spaces. Each $\mathcal{F}_n$ is a σ-algebra on Ω, though the σ-algebra $\mathcal{F}$ of $(\Omega, \mathcal{F}, \mathbb{P})$ will not be any one of these σ-algebras, as we will see.

For $n = 1$, we have:

$$\Omega_1 = \{(1), (-1)\},$$

$$\begin{aligned}\mathcal{F}_1 &= \{\emptyset, \{\omega : (\omega_1) = (1)\}, \{\omega : (\omega_1) = (-1)\}, \Omega\} \\ &= \{E \subset \Omega : E = \{\omega : (\omega_1) \in E' \in 2^{\Omega_1}\}\}.\end{aligned}$$

For $n = 2$, we have:

$$\Omega_2 = \{(1,1), (1,-1), (-1,1), (-1,-1)\},$$

$$\mathcal{F}_2 = \{E \subset \Omega : E = \{\omega : (\omega_1, \omega_2) \in E' \in 2^{\Omega_2}\}\}.$$

Note then that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \dots$. For example, $\{\omega : (\omega_1) = (1)\} \in \mathcal{F}_1$ as seen above, but where does it appear in $\mathcal{F}_2$? It is the event

$$E = \{\omega\colon (\omega_1, \omega_2) \in \{(1,1),(1,-1)\} \in 2^{\Omega_2}\}.$$

These increasing σ-algebras give the natural filtration $\{\mathcal{F}_n\}$ for this process. Overall, the σ-algebra we use for the infinite process is

$$\mathcal{F} := \sigma\left(\bigcup_{i=1}^{\infty} \mathcal{F}_i\right).$$

The probability measure $\mathbb{P}$ for the infinite process is defined for any event E belonging to some $\mathcal{F}_n$, and is accordingly given by

$$\mathbb{P}(E) := \mathbb{P}_n(E')$$

where $E \subset \Omega$ and $E' \in 2^{\Omega_n}$ as described earlier.

Consider random variable $Y := \frac{X_1}{X_2}$. This is an example of a random variable that is $\mathcal{F}_2$- measurable but not $\mathcal{F}_1$-measurable.

Note that $\mathbb{E}[X_i] = 1\cdot\frac{1}{2}+(-1)\cdot\frac{1}{2} = 0$, and $\mathrm{Var}[X_i] = \mathbb{E}[X_i^2]-\mathbb{E}[X_i]^2 = 1\cdot1-0 = 1$. Consider the stochastic process S_n given by $S_n = X_1 + X_2 + \cdots + X_n$. S_n is called a *simple random walk*. As we can see, it is adapted to the natural filtration $\{\mathcal{F}_n\}$, as expected. By Central Limit Theorem, we see that the distribution of $Z_n = \frac{S_n}{\sqrt{n}}$ approaches a standard normal distribution. Then as $n \to \infty$, $\mathbb{E}\left[\frac{S_n}{\sqrt{n}}\right] = 0$, so $\mathbb{E}[S_n] = \sqrt{n}\cdot 0 = 0$. Also, as $n \to \infty$,

$$\mathrm{Var}\left[\frac{S_n}{\sqrt{n}}\right] = \mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^2\right] - \mathbb{E}\left[\frac{S_n}{\sqrt{n}}\right]^2 = \mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^2\right] = 1.$$

Then $\mathbb{E}[S_n^2] = \mathrm{Var}[S_n] = n$, so the distribution of S_n approaches $N(0, n)$. We will further explore the relationship between random walks and normally distributed random variables when we study Brownian Motion in the next section.

Proposition 3.1. *Let $X_1, \dots, X_n$, S_n, and $\mathcal{F}_n$ be defined as above. Then S_n is a martingale with respect to $\mathcal{F}_n$.*

Proof. Let $m < n$. We want to show that

$$E[S_n|\mathcal{F}_m] = S_m.$$

We know that X_j is independent of $\mathcal{F}_m$ if $j > m$. Then

$$\begin{aligned}
E[S_n|\mathcal{F}_m] &= E[S_m + (S_n - S_m)|\mathcal{F}_m] \\
&= E[S_m|\mathcal{F}_m] + E[S_n - S_m|\mathcal{F}_m] \\
&= S_m + E[X_{m+1} + \cdots + X_n|\mathcal{F}_m] \\
&= S_m + E[X_{m+1}|\mathcal{F}_m] + \cdots + E[X_n|\mathcal{F}_m] \\
&= S_m + 0\cdot(n-m) \\
&= S_m.
\end{aligned}$$

□

4. Brownian Motion

In this section, we introduce *Brownian Motion*, one of the most important stochastic processes. It also goes by the name *Wiener Process*.

In the last section, the simple random walk was assumed to have time increment $\Delta t = 1$ and space increment $\Delta x = 1$. We can think of Brownian Motion as the limit of a random walk as each increment approaches zero, while preserving certain normalization. Note that in the previous example, $\mathrm{Var}[S_1] = 1$. We want to preserve this as we take our limit.

Let N be a large positive integer. We denote our process as $W_t^{(N)}$. Let time increment $\Delta t := \frac{1}{N}$. As before, we observe the process $X_1, X_2, \dots$ where $\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$, and each X_i corresponds to the i-th jump of length Δx up or down. Now, however, the i-th step is completed at time $i\Delta t$ rather than at time i.

Then at time $1 = N\Delta t$, N steps have been completed, giving

$$W_1^{(N)} = \Delta x(X_1 + \cdots + X_N).$$

As mentioned earlier, we want $\mathrm{Var}[W_1^{(N)}] = 1$. Note that $\mathbb{E}[W_1^{(N)}] = 0$. Then

$$\begin{aligned}\mathrm{Var}[W_1^{(N)}] &= \mathrm{Var}[\Delta x(X_1 + \cdots + X_N)] \\ &= (\Delta x)^2(\mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_N]) \\ &= (\Delta x)^2 N \\ &= 1,\end{aligned}$$

so $\Delta x = \sqrt{\dfrac{1}{N}} = \sqrt{\Delta t}$. This result is important; in Brownian Motion, we will show that $dB_t^2 = dt$ which is analogous to the square of this equation.

From Central Limit Theorem, we see that the distribution of $Z_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}} = \Delta x(X_1 + \cdots + X_N) = W_1^{(N)}$ approaches $N(0, 1)$.

This gives an intuitive idea of what Brownian Motion is. Brownian Motion is a stochastic process modeling continuous random motion. $B_t = B(t)$ is the value of the Brownian Motion at time t. We can think of the process either as a collection of random variables B_t defined for each $t \geq 0$, or as a random function (a function-valued random variable) $t \mapsto B_t$. When we say that it is continuous, we mean that any such function $t \mapsto B_t$ is continuous in the standard sense. We often say that $B_0 = 0$ but other starting conditions are also valid. In this paper, we only consider the one-dimensional case, and $B_0 = 0$ always.

Definition 4.1 (Brownian Motion)**.** A continuous stochastic process B_t is a one-dimensional *Brownian motion* with *drift* m and *variance* σ^2 starting at the origin if it satisfies the following properties:

- $B_0 = 0$.
- For $s \leq t$, the distribution of $B_t - B_s$ is $N(m(t-s), \sigma^2(t-s))$.
- If $s \leq t$, then the random variable $B_t - B_s$ is independent of values B_r for $r \leq s$.
- With probability one, the function $t \mapsto B_t$ is a continuous function of t.

When $m = 0, \sigma^2 = 1$, we call B_t a *standard Brownian motion*.

As a result of these properties, we see that the distribution of $B_t = B_t - B_0$ is $N(mt, \sigma^2 t)$. In particular, $B_1 \sim N(m, \sigma^2)$, and so for standard Brownian motion, B_1 has standard normal distribution. Also, we know that if $X \sim N(0,1)$ and $Y = \sigma Z + m$, then $Y \sim N(m, \sigma^2)$. Using this, we can show that for standard Brownian motion B_t, $W_t = \sigma B_t + mt$ is Brownian motion with drift m and variance σ^2.

Note that we have merely described a standard Brownian motion B_t without ever proving that there exists such a process satisfying all necessary properties. Once this existence is proven, existence of non-standard Brownian motion naturally follows. For the complete proof, see Section 2.5 of [6]. The overall proof layout is as follows:

- Define B_t on dyadic rationals t.
- Prove that with probability one, the function $t \mapsto B_t$ is continuous on the dyadics.
- Extend B_t to all other t by continuity.

We will carry out the first step, defining B_t for dyadic rationals t on $[0,1]$ to give intuition for how the standard Brownian motion is constructed. Note that even carrying out the first step is a lengthy task, and is not necessary for understanding and using Brownian Motion. The reader may skip over it, but due to its importance in Itô Calculus, we provide the proof for curious readers. Once Brownian motion is constructed for all $t \in [0,1]$, we can "connect" countably many such processes to define any B_t for $t \geq 0$.

Proposition 4.2. *There exists a stochastic process B_q defined on the dyadic rationals in $[0,1]$ that satisfies the first three properties of Brownian motion.*

Proof. Let

$$\mathcal{D}_n := \left\{ \frac{k}{2^n} : k = 0, 1, \dots, 2^n \right\}$$

denote the dyadic rationals in $[0,1]$ that are multiples of 2^{-n}. Let $\mathcal{D} := \bigcup_{n=0}^{\infty} \mathcal{D}_n$ denote all dyadic rationals in $[0,1]$.

We proceed by defining B_t on $\mathcal{D}_0 = \{\frac{0}{1}, \frac{1}{1}\}$, then on $\mathcal{D}_1 \setminus \mathcal{D}_0 = \{\frac{1}{2}\}$, then on $\mathcal{D}_2 \setminus \mathcal{D}_1 = \{\frac{1}{4}, \frac{3}{4}\}$, and so on recursively such that the properties of Brownian motion are satisfied. $\mathcal{D}$ is countable, and we use a corresponding $N(0,1)$ random variable Z_q from the countable set

$$\{Z_q \sim N(0,1) \colon q \in \mathcal{D}\}$$

to help define each B_q, $q \in \mathcal{D}$ besides B_0.

We define $B_0 = 0$ as our standard Brownian motion initial condition, and $B_1 = Z_1$ since it is $N(0,1)$ for standard Brownian motion. Then

$$B_{1/2} = \frac{B_1}{2} + \frac{Z_{1/2}}{2}.$$

We can think of this as $E[B_{1/2}|B_0, B_1]$ plus some independent randomness so that it has the appropriate variance. We see that

$$B_1 - B_{1/2} = \frac{B_1}{2} - \frac{Z_{1/2}}{2},$$

so by Proposition 1.13, $B_{1/2}$ and $B_1 - B_{1/2}$ are independent $N(0,1/2)$ variables. Continuing, we note that if $q \in \mathcal{D}_{n+1} \setminus \mathcal{D}_n$ then $q = (2k+1)/2^{n+1}$ for some $k = 0, 1, \ldots, 2^n - 1$. Then we define

$$B_q := B_{k/2^n} + \frac{B_{(k+1)/2^n} - B_{k/2^n}}{2} + \frac{Z_q}{2^{(n+2)/2}}.$$

Again, we think of this as

$$B_q = E[B_q | B_{k/2^n}, B_{(k+1)/2^n}] + \text{independent randomness}$$

where

$$E[B_q | B_{k/2^n}, B_{(k+1)/2^n}] = \frac{B_{k/2^n} + B_{(k+1)/2^n}}{2}.$$

In other words, it is an average of its $\mathcal{D}_n$ neighbors' values, plus randomness.

Now we examine the randomness. We know from earlier that for $n = 1$, the random variables $\{B_{k/2^n} - B_{(k-1)/2^n} \colon k = 1, \ldots, 2^n\}$ are independent with $N(0, 2^{-n})$ distribution. Suppose this is true for $n = m$. We define B_q for any $q \in \mathcal{D}_{m+1} \setminus \mathcal{D}_m$ as above. Then

$$B_q - B_{k/2^m} = \frac{B_{(k+1)/2^m} - B_{k/2^m}}{2} + \frac{Z_q}{2^{(m+2)/2}},$$

$$B_{(k+1)/2^m} - B_q = \frac{B_{(k+1)/2^m} - B_{k/2^m}}{2} - \frac{Z_q}{2^{(m+2)/2}}.$$

Note that $B_{(k+1)/2^m} - B_{k/2^m}$ is a random variable with distribution $N(0, 2^{-m})$ or equivalently $2^{-m/2} \cdot N(0,1)$. Let this $N(0,1)$ variable be denoted X. Then

$$B_q - B_{k/2^m} = \frac{2^{-m/2} X}{2} + \frac{Z_q}{2^{(m+2)/2}} = 2^{-(m+1)/2} \left(\frac{X + Z_q}{\sqrt{2}} \right),$$

$$B_{(k+1)/2^m} - B_q = \frac{2^{-m/2} X}{2} - \frac{Z_q}{2^{(m+2)/2}} = 2^{-(m+1)/2} \left(\frac{X - Z_q}{\sqrt{2}} \right).$$

Then by Proposition 1.13, $B_q - B_{k/2^m}$ and $B_{(k+1)/2^m} - B_q$ are independent with $2^{-(m+1)/2} N(0,1)$ distributions, or equivalently $N(0, 2^{-(m+1)})$ distributions. Since this is true for all k, we have shown that for $n = m+1$, the random variables $\{B_{k/2^n} - B_{(k-1)/2^n} \colon k = 1, \ldots, 2^n\}$ are independent with $N(0, 2^{-n})$ distribution. So, by induction, we have shown this is true for all n.

Then, it is not hard to show that $\{B_q \colon q \in \mathcal{D}\}$ satisfies all properties of Brownian motion except continuity of paths. □

Although Brownian motion is continuous, it is quite "rough." In fact, the paths are differentiable nowhere with probability 1. To see this, imagine discretizing time by choosing a small Δt such that you sample

$$B_0, B_{\Delta t}, B_{2\Delta t}, \ldots$$

Then for any k, $B_{(k+1)\Delta t} - B_{k\Delta t}$ is a random variable with distribution $N(0, \Delta t)$ or equivalently $\sqrt{\Delta t} N(0,1)$. Let $N_0, N_1, \ldots$ be independent $N(0,1)$ random variables. Then

$$\Delta B_{k\Delta t} := B_{(k+1)\Delta t} - B_{k\Delta t} = \sqrt{\Delta t} N_k.$$

Then $\mathbb{E}[|B_{(k+1)\Delta t} - B_{k\Delta t}|] = \sqrt{\Delta t}\mathbb{E}[|N_k|]$ which turns out to be $\sqrt{\Delta t} \cdot \sqrt{\frac{2}{\pi}} \approx \sqrt{\Delta t}$.

The derivative at time t, if it existed, would be

$$\lim_{\Delta t \to 0} \frac{B_{t+\Delta t} - B_t}{\Delta t}.$$

However, the abolute value of the numerator is of order $\sqrt{\Delta t}$, which is much larger than denominator Δt for small Δt. Intuitively then, we can say that this limit does not exist.

Theorem 4.3. *With probability one, the function $t \mapsto B_t$ is nowhere differentiable.*

For the full proof, see Theorem 2.6.1 on pages 48-51 of [6]. The fact that Brownian motion is not differentiable partially explains why we need a new form of Calculus to address continuous stochastic processes.

Proposition 4.4. *A standard Brownian motion B_t is a continuous martingale with respect to filtration $\{\mathcal{F}_t\}$.*

Proof. Let $s < t$. Then

$$E[B_t|\mathcal{F}_s] = E[B_s|\mathcal{F}_s] + E[B_t - B_s|\mathcal{F}_s] = B_s + E[B_t - B_s|\mathcal{F}_s] = B_s.$$

□

Proposition 4.5. *Suppose B_t is a standard Brownian motion and $a > 0$. Then*

$$W_t := \frac{B_{at}}{\sqrt{a}}$$

is a standard Brownian motion.

Proof. $W_0 = \dfrac{B_0}{\sqrt{a}} = 0$.

Let $s < t$. Then the distribution of $W_t - W_s = \dfrac{B_{at} - B_{as}}{\sqrt{a}}$ is $a^{-1/2}N(0, a(t-s))$, or equivalently $N(0, t-s)$.

The random variable $W_t - W_s = \dfrac{B_{at} - B_{as}}{\sqrt{a}}$ is independent of values $W_r = \dfrac{B_{ar}}{\sqrt{a}}$ for $r \leq s$ since $B_{at} - B_{as}$ is independent of values B_{ar} for $ar \leq as$. Dividing both quantities by the constant $\sqrt{a}$ does not change independence.

Continuity is implied by continuity of B_{at} which is in turn implied by continuity of B_t. □

Definition 4.6 (Quadratic Variation)**.** If X_t is a process, the *quadratic variation* of the process is given by

$$\langle X \rangle_t := \lim_{n \to \infty} \sum_{j \leq tn} \left[X\left(\frac{j}{n}\right) - X\left(\frac{j-1}{n}\right) \right]^2.$$

Quadratic variation will be a very important property in stochastic calculus. To determine the quadratic variation of Brownian motion, we first consider the similar sum for some fixed n and $t = 1$:

$$Q_1^{(n)} := \sum_{j \leq n} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right]^2 = \sum_{j=1}^{n} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right]^2.$$

We can rewrite each term:

$$\left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right)\right]^2 = \frac{1}{n}\left[\frac{B(\frac{j}{n}) - B(\frac{j-1}{n})}{1/\sqrt{n}}\right]^2.$$

Note that by Proposition 4.5, each $\dfrac{B(\frac{j}{n}) - B(\frac{j-1}{n})}{1/\sqrt{n}}$ is an $N(0,1)$ random variable, which motivates us to write

$$Q_1^{(n)} = \sum_{j=1}^{n} \frac{1}{n} Y_j = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

where each Y_j has distribution Z^2 where Z is standard normal. Through integration by parts, we see that $\mathbb{E}[Z^2] = 1$ and $\mathbb{E}[Z^4] = 3$. Then $\mathrm{Var}[Y_j] = \mathbb{E}[Y_j^2] - \mathbb{E}[Y_j]^2 = 3 - 1 = 2$. Hence,

$$\mathbb{E}[Q_1^{(n)}] = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[Y_j] = \frac{n}{n} = 1,$$

$$\mathrm{Var}[Q_1^{(n)}] = \frac{1}{n^2}\sum_{j=1}^{n} \mathrm{Var}[Y_j] = \frac{2n}{n^2} = \frac{2}{n}.$$

Then as $n \to \infty$, the random variable $Q_1^{(n)}$ tends to the constant (zero variance) random variable of 1. If we consider

$$Q_t^{(n)} := \sum_{j \le tn} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right)\right]^2$$

we see that

$$\mathbb{E}[Q_t^{(n)}] = \frac{1}{n}\sum_{j \le tn} \mathbb{E}[Y_j] = \frac{n\lfloor t\rfloor}{n} = \lfloor t\rfloor,$$

$$\mathrm{Var}[Q_t^{(n)}] = \frac{1}{n^2}\sum_{j \le tn} \mathrm{Var}[Y_j] = \frac{2n\lfloor t\rfloor}{n^2} = \frac{2\lfloor t\rfloor}{n}.$$

As $n \to \infty$, the random variable $Q_t^{(n)}$ tends to the constant (zero variance) random variable t. Then $\langle B\rangle_t = t$ for standard Brownian motion B_t. We now generalize.

Theorem 4.7. *Suppose W_t is a Brownian motion with drift m and variance σ^2. Then $\langle W\rangle_t = \sigma^2 t$.*

Proof. Brownian motion with drift m and variance σ^2 can be written $W_t = \sigma B_t + mt$. Let

$$\begin{aligned}
Q_t^{(n)} &:= \sum_{j \le tn} \left[W\left(\frac{j}{n}\right) - W\left(\frac{j-1}{n}\right) \right]^2 \\
&= \sum_{j \le tn} \left[\sigma B\left(\frac{j}{n}\right) + m\left(\frac{j}{n}\right) - \sigma B\left(\frac{j-1}{n}\right) - m\left(\frac{j-1}{n}\right) \right]^2 \\
&= \sum_{j \le tn} \left[\sigma \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right] + \frac{m}{n} \right]^2 \\
&= \sigma^2 \sum_{j \le tn} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right]^2 + \frac{2\sigma m}{n} \sum_{j \le tn} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right] + \sum_{j \le tn} \frac{m^2}{n^2}.
\end{aligned}$$

As $n \to \infty$,

$$\sigma^2 \sum_{j \le tn} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right]^2 \to \sigma^2 \langle B \rangle_t = \sigma^2 t,$$

$$\frac{2\sigma m}{n} \sum_{j \le tn} \left[B\left(\frac{j}{n}\right) - B\left(\frac{j-1}{n}\right) \right] = \frac{2\sigma m}{n} B_t \to 0,$$

$$\sum_{j \le tn} \frac{m^2}{n^2} = \frac{m^2 \lfloor t \rfloor}{n} \to 0.$$

Then

$$\langle W \rangle_t = \lim_{n \to \infty} Q_t^{(n)} = \sigma^2 t.$$

□

5. Itô Calculus

In regular calculus, we examine differential equations of the form

$$df(t) = C(t, f(t))dt, \text{ or equivalently, } \frac{df}{dt} = f'(t) = C(t, f(t)).$$

A solution to such an equation satisfying initial condition $f(0) = x_0$ would be

$$f(t) = x_0 + \int_0^t C(s, f(s))ds.$$

With stochastic calculus, we aim to examine *stochastic differential equations* or SDE's of the form

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t.$$

A solution to such an equation satisfying initial condition $X_0 = x_0$ would be

$$X_t = x_0 + \int_0^t m(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

The *Itô integral* gives precise meaning to this last term. The reader will likely find it similar to the Riemann integral.

Definition 5.1 (Simple Process)**.** A stochastic process A_t is a *simple process* if there exist times $0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = \infty$ and $\mathcal{F}_{t_j}$-measurable random variables Y_j for $j = 0, 1, \ldots, n$ such that

$$A_t = Y_j,\ t_j \le t < t_{j+1}.$$

We think of A_t as a step function. Note that A_t is $\mathcal{F}_{t_j}$-measurable and is therefore $\mathcal{F}_t$-measurable. Then A_t is adapted to filtration $\{\mathcal{F}_t\}$.

Definition 5.2 (Itô Integral for Simple Processes)**.** Let A_t be a simple process as defined in Definition 5.1, with the additional condition that $\mathbb{E}[Y_j^2] < \infty$. Then we define $Z_t := \int_0^t A_s dB_s$ by:

$$Z_{t_j} := \sum_{i=0}^{j-1} Y_i[B_{t_{i+1}} - B_{t_i}],$$

$$Z_t := Z_{t_j} + Y_j[B_t - B_{t_j}] \text{ if } t_j \le t < t_{j+1},$$

$$\int_r^t A_s dB_s := Z_t - Z_r.$$

Proposition 5.3. *Let A_t, C_t be simple processes with $Z_t = \int_0^t A_s dB_s$, and let B_t be a standard Brownian motion.*

- *Let a, c be constants. Then $aA_t + cC_t$ is a simple process and*
$$\int_0^t (aA_s + cC_s)dB_s = a\int_0^t A_s dB_s + c\int_0^t C_s dB_s. \text{ (Linearity)}$$
- *Z_t is a martingale with respect to $\{\mathcal{F}_t\}$. (Martingale Property)*
- $\mathrm{Var}[Z_t] = \mathbb{E}[Z_t^2] = \int_0^t \mathbb{E}[A_s^2]ds$. *(Variance Rule)*
- *With probability one, the function $t \mapsto Z_t$ is continuous. (Continuity)*

Proof. Linearity follows from the definition of the integral, and continuity follows from the continuity of $t \mapsto B_t$.

Let $s < t$. To prove the Martingale Property, it suffices to show that

$$E(Z_t - Z_s|\mathcal{F}_s) = 0.$$

Let $t_j \le s < t_{j+1}$ and $t_k \le t < t_{k+1}$ where $j \le k$. Then $Z_s = Z_{t_j} + Y_j[B_s - B_{t_j}]$ and $Z_t = Z_{t_k} + Y_k[B_t - B_{t_k}]$. Then

$$\begin{aligned} E(Z_t - Z_s|\mathcal{F}_s) =& E\left(Y_j[B_{t_{j+1}} - B_s] + \sum_{i=j+1}^{k-1} Y_i[B_{t_{i+1}} - B_{t_i}] + Y_k[B_t - B_{t_k}]\middle|\mathcal{F}_s\right) \\ =& E(Y_j[B_{t_{j+1}} - B_s]|\mathcal{F}_s) \\ &+ \sum_{i=j+1}^{k-1} E(Y_i[B_{t_{i+1}} - B_{t_i}]|\mathcal{F}_s) \\ &+ E(Y_k[B_t - B_{t_k}]|\mathcal{F}_s). \end{aligned}$$

Each term in this summation is of the form $E(Y_a[B_{t_c} - B_{t_b}]|\mathcal{F}_s)$ where $s \le t_a \le t_b \le t_c$. Then

$$\begin{aligned} E(Y_a[B_{t_c} - B_{t_b}]|\mathcal{F}_s) &= E(E(Y_a[B_{t_c} - B_{t_b}]|\mathcal{F}_{t_a})|\mathcal{F}_s) \\ &= E(Y_a E(B_{t_c} - B_{t_b}|\mathcal{F}_{t_a})|\mathcal{F}_s) \\ &= E(Y_a \cdot 0|\mathcal{F}_s) \\ &= 0. \end{aligned}$$

Then $E(Z_t - Z_s|\mathcal{F}_s) = 0$, so Z_t is a martingale.

To prove the variance rule for $t = t_j$, we must show that

$$\mathrm{Var}[Z_t] = \mathbb{E}[Z_t^2] = \sum_{i=0}^{j-1}\sum_{k=0}^{j-1} \mathbb{E}[Y_i[B_{t_{i+1}} - B_{t_i}]Y_k[B_{t_{k+1}} - B_{t_k}]] = \int_0^t \mathbb{E}[A_s^2]ds.$$

When $i < k$,

$$\begin{aligned} \mathbb{E}[Y_i[B_{t_{i+1}} - B_{t_i}]Y_k[B_{t_{k+1}} - B_{t_k}]] &= \mathbb{E}[E(Y_i[B_{t_{i+1}} - B_{t_i}]Y_k[B_{t_{k+1}} - B_{t_k}]|\mathcal{F}_k] \\ &= \mathbb{E}[Y_i[B_{t_{i+1}} - B_{t_i}]Y_k E(B_{t_{k+1}} - B_{t_k}|\mathcal{F}_k)] \\ &= \mathbb{E}[Y_i[B_{t_{i+1}} - B_{t_i}]Y_k \cdot 0] \\ &= 0 \end{aligned}$$

Similarly, when $i > k$,

$$\mathbb{E}[Y_i[B_{t_{i+1}} - B_{t_i}]Y_k[B_{t_{k+1}} - B_{t_k}]] = \mathbb{E}[E(Y_i[B_{t_{i+1}} - B_{t_i}]Y_k[B_{t_{k+1}} - B_{t_k}]|\mathcal{F}_k] = 0.$$

Then we want to show that

$$\mathrm{Var}[Z_t] = \mathbb{E}[Z_t^2] = \sum_{i=0}^{j-1} \mathbb{E}[Y_i^2[B_{t_{i+1}} - B_{t_i}]^2] = \int_0^t \mathbb{E}[A_s^2]ds.$$

We proceed:

$$\begin{aligned} \mathbb{E}[Y_i^2[B_{t_{i+1}} - B_{t_i}]^2] &= \mathbb{E}[E(Y_i^2[B_{t_{i+1}} - B_{t_i}]^2|\mathcal{F}_i)] \\ &= \mathbb{E}[Y_i^2 E([B_{t_{i+1}} - B_{t_i}]^2|\mathcal{F}_i)] \\ &= \mathbb{E}[Y_i^2(t_{i+1} - t_i)] \\ &= \mathbb{E}[Y_i^2](t_{i+1} - t_i). \end{aligned}$$

The function $s \mapsto \mathbb{E}[A_s^2]$ is a step function taking value $\mathbb{E}[Y_i^2]$ when $t_i \le s < t_{i+1}$. Then

$$\mathrm{Var}[Z_t] = \mathbb{E}[Z_t^2] = \sum_{i=0}^{j-1} \mathbb{E}[Y_i^2](t_{i+1} - t_i) = \int_0^t \mathbb{E}[A_s^2]ds.$$

□

We can generalize the integral to process A_t adapted to filtration $\{\mathcal{F}_t\}$ having piece-wise continuous paths. We first generalize to A_t bounded, continuous paths by approximating A_t with simple processes. We then account for unbounded paths, and finally piece-wise continuous paths.

Lemma 5.4. *Let A_t be a stochastic process adapted to filtration $\{\mathcal{F}_t\}$ with continuous paths. Suppose there exists $C < \infty$ such that with probability one $|A_t| \leq C$ for all t. Then there exists a sequence of simple processes $A_t^{(n)}$ such that for all t,*

$$\lim_{n\to\infty} \int_0^t \mathbb{E}\left[\left|A_s - A_s^{(n)}\right|^2\right] ds = 0$$

and for all n, t,

$$|A_t^n| \leq C.$$

For the proof, see Lemma 3.2.2 of [6].

Definition 5.5 (Itô Integral for Bounded Process with Continuous Path)**.** Let A_t be a bounded process adapted to filtration $\{\mathcal{F}_t\}$ having continuous paths. Then there exists a sequence of simple processes $A_t^{(n)}$ such that

$$\lim_{n\to\infty} \int_0^t \mathbb{E}\left[\left|A_s - A_s^{(n)}\right|^2\right] ds = 0.$$

Then we define

$$Z_t := \lim_{n\to\infty} \int_0^t A_s^{(n)} dB_s.$$

Definition 5.6 (Itô Integral for Unbounded Process with Continuous Paths)**.** Let A_t be a possibly unbounded process adapted to filtration $\{\mathcal{F}\}$ with continuous paths. Let $T_n = \inf\{t : |A_t| = n\}$ for all $n = 0, 1, \ldots < \infty$. Then $A_t^{(n)} = A_{\min(t,T_n)}$ is a sequence of continuous, bounded processes with corresponding well-defined Itô integrals $Z_t^{(n)} = \int_0^t A_s^{(n)} dB_s$. Then we define

$$Z_t := \lim_{n\to\infty} Z_t^{(n)}.$$

Continuity and Linearity still hold for unbounded processes, and the Variance Rule holds although it is possible that $\mathrm{Var}[Z_t] = \infty$. However, the Martingale Property might not hold since A_s can grow to infinity. We still know that the integral is a *local martingale* though. Local martingales involve stopping times; T is a *stopping time* if it is a positive integer random variable with respect to $\{\mathcal{F}_n\}$ such that for each n the event $\{T = n\}$ is $\mathcal{F}_n$-measurable.

Definition 5.7 (Local Martingale)**.** A continuous process M_t adapted to filtration $\{\mathcal{F}_t\}$ is a *local martingale* on $[0, T)$ if there exist stopping times

$$\tau_1 \leq \tau_2 \leq \tau_3 \cdots$$

such that

$$\lim_{j\to\infty} \tau_j = T$$

and for each j, $M_t^{(j)} = M_{\min(t,\tau_j)}$ is a martingale.

For the proof of how Proposition 5.3 extends to unbounded processes with continuous paths, see Theorem 3.4 in [2].

Returning to our previous discussion, we can now make sense of stochastic differential equations of the form

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t,$$

as their integral form

$$X_t = X_0 + \int_0^t m(s, X_s)dt + \int_0^t \sigma(s, X_s)dB_s$$

is now well defined.

We can now derive Itô's Formula, which is the foundation of Itô Calculus. Itô's Formula is the analog of the chain rule in ordinary calculus. Due to the non-differentiability and the non-zero quadratic variation of Brownian motion, we must include more terms in the Taylor expansion for chain rule. Note that the proof is lengthy and not necessary for understanding its applications. The reader may skip over it, but due to its importance in Itô Calculus, we provide the proof for curious readers.

We say that a function $f(x)$ is C^k in x if it has k continuous x-derivatives; we use it to denote that a function is sufficiently differentiable.

Theorem 5.8 (Itô's Formula)**.** *Let $f(t, x)$ be C^1 in t and C^2 in x. Let B_t be a standard Brownian motion. Then*

$$f(t, B_t) = f(0, B_0) + \int_0^t \left[\partial_s f(s, B_s) + \frac{1}{2}\partial_{xx} f(s, B_s)\right] ds + \int_0^t \partial_x f(s, B_s)dB_s$$

and in differential form

$$df(t, B_t) = [\partial_t f(t, B_t) + \partial_{xx} f(t, B_t)]\, dt + \partial_x f(t, B_t)dB_t.$$

Proof. We write the Taylor expansion:

$$\begin{aligned} f(t + \Delta t, x + \Delta x) - f(t, x) =& \partial_t f(t, x)\Delta t + o(\Delta t) \\ &+ \partial_x f(t, x)\Delta x + \frac{1}{2}\partial_{xx} f(t, x)(\Delta x)^2 + o\left((\Delta x)^2\right). \end{aligned}$$

Let $\Delta t := \frac{1}{n}$. We write the telescoping sum for $f(t, B_t)$:

$$f(t, B_t) - f(0, B_0) = \sum_{j \le tn} \left[f\left(\frac{j}{n}, B_{j/n}\right) - f\left(\frac{j-1}{n}, B_{(j-1)/n}\right)\right].$$

Let $\Delta_{j,n} := B_{j/n} - B_{(j-1)/n}$. Applying the Taylor expansion, we get

$$\begin{aligned} & f\left(\frac{j}{n}, B_{j/n}\right) - f\left(\frac{j-1}{n}, B_{(j-1)/n}\right) \\ =& \partial_t f\left(\frac{j}{n}, B_{j/n}\right)\frac{1}{n} + o\left(\frac{1}{n}\right) \\ &+ \partial_x f\left(\frac{j}{n}, B_{j/n}\right)\Delta_{j/n} + \frac{1}{2}\partial_{xx} f\left(\frac{j}{n}, B_{j/n}\right)(\Delta_{j,n})^2 + o\left((\Delta_{j,n})^2\right). \end{aligned}$$

Then $f(t, B_t) - f(0, B_0)$ is equal to

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{j\le tn} \partial_t f\left(\frac{j}{n}, B_{j/n}\right)\frac{1}{n} \\
&+ \lim_{n\to\infty} \sum_{j\le tn} o\left(\frac{1}{n}\right) \\
&+ \lim_{n\to\infty} \sum_{j\le tn} \partial_x f\left(\frac{j}{n}, B_{j/n}\right)(\Delta_{j,n}) \\
&+\frac{1}{2}\lim_{n\to\infty} \sum_{j\le tn} \partial_{xx} f\left(\frac{j}{n}, B_{j/n}\right)(\Delta_{j,n})^2 \\
&+ \lim_{n\to\infty} \sum_{j\le tn} o\left((\Delta_{j,n})^2\right).
\end{aligned}$$

The first term sum is the Riemann integral approximation of $\partial_t f(t, B_t)$, so taking the limit, the first term is equal to $\int_0^t \partial_s f(s, B_s)ds$. The second term sum is proportional to

$$nt \cdot o\left(\frac{1}{n}\right) = t \cdot \frac{o\left(\frac{1}{n}\right)}{\frac{1}{n}},$$

so taking the limit, the second term is equal to 0 by definition of $o\left(\frac{1}{n}\right)$. The third term sum is a simple process approximation of the Itô integral of $\partial_x f(t, B_t)$, so taking the limit, the third term is equal to $\int_0^t \partial_x f(s, B_s)ds$.

The fourth term is less straightforward. Suppose that $\partial_{xx} f\,(t, B_t))$ is constant with value a on some interval from t_1 to t_2. Then we know that

$$\begin{aligned}
\frac{1}{2}\lim_{n\to\infty} \sum_{t_1 n\le j\le t_2 n} \partial_{xx} f\left(\frac{j}{n}, B_{j/n}\right)(\Delta_{j,n})^2 &= \frac{a}{2}\lim_{n\to\infty} \sum_{t_1 n\le j\le t_2 n} [B_{j/n} - B_{(j-1)/n}]^2 \\
&= \frac{a}{2}[\langle B_{t_2}\rangle - \langle B_{t_1}\rangle] \\
&= \frac{a}{2}(t_2 - t_1).
\end{aligned}$$

For any $\varepsilon > 0$, we can approximate $\partial_{xx} f\,(t, x)$ by a simple process $g_\varepsilon(t, x)$ such that for all t, x, $|g_\varepsilon - f| < \varepsilon$. Since the sum can be broken up into several sums on which g is constant, we see that for fixed ε,

$$\frac{1}{2}\lim_{n\to\infty} \sum_{j\le tn} g_\varepsilon\left(\frac{j}{n}, B_{j/n}\right)[B_{j/n} - B_{(j-1)/n}]^2 = \frac{1}{2}\int_0^t g_\varepsilon(s, B_s)dt.$$

Also,

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\left|\sum_{j\le tn}\left[g_\varepsilon\left(\frac{j}{n}, B_{j/n}\right) - \partial_{xx} f\left(\frac{j}{n}, B_{j/n}\right)\right][B_{j/n} - B_{(j-1)/n}]^2\right| \\
\le&\lim_{\varepsilon\to 0}\varepsilon\sum_{j\le tn}[B_{j/n} - B_{(j-1)/n}]^2 \\
=&0.
\end{aligned}$$

Then the fourth term is equal to

$$\frac{1}{2}\lim_{\varepsilon\to 0} g_\varepsilon\left(\frac{j}{n}, B_{j/n}\right)[B_{j/n}-B_{(j-1)/n}]^2 = \frac{1}{2}\lim_{\varepsilon\to 0}\int_0^t g_\varepsilon(t,B_t)dt = \frac{1}{2}\int_0^t \partial_{xx}f(t,B_t)dt.$$

Finally, for the fifth term, let $Y_{j,n} \sim N(0,1)$. Then $(\Delta_{j,n})^2 = [B_{j/n} - B_{(j-1)/n}]^2 = \left[\sqrt{\frac{1}{n}}Y_{j,n}\right]^2 = \frac{1}{n}Y_{j,n}^2$. Since $E[(\Delta_{j,n})^2] = \frac{1}{n}E\left[Y_{j,n}^2\right] = \frac{1}{n}$, it follows that $o\left((\Delta_{j,n})^2\right) \approx o\left(\frac{1}{n}\right)$, so after taking the limit, the fifth term is equal to 0 just like the second term. Then

$$f(t,B_t) = f(0,B_0) + \int_0^t [\partial_s f(s,B_s) + \partial_{xx}f(s,B_s)]ds + \int_0^t \partial_x f(s,B_s)dB_s.$$

□

Definition 5.9 (Itô Process)**.** If $dX_t = A_t dt + C_t dB_t$, where A_t and C_t are adapted, continuous processes, then X_t is an *Itô process.*

All processes we mention hereafter will be Itô processes; it may not always be immediately apparent, but after simplifying and grouping dt and dB_t terms, the SDE can be expressed in the form above. We can think of this Itô process as a Brownian motion with variance C_t^2 and drift A_t at time t. Note that if $A_t \neq 0$ for all t, then X_t cannot be a martingale. This will be important, as we may know that a given process is a martingale and can thus set all dt terms to 0. For example, we will see this in the Feynman-Kac Formula derivation in Section 8.

Theorem 5.10 (Itô's Formula, Generalized)**.** *Let $f(t,x)$ be C^1 in t and C^2 in x, and let X_t and Y_t be Itô processes where $dX_t = A_t dt + C_t dB_t$ and $dY_t = J_t dt + K_t dB_t$. Then*

$$\begin{aligned} df(X_t,Y_t) =& \partial_x f(X_t,Y_t)dX_t + \partial_y f(X_t,Y_t)dY_t \\ &+ \frac{1}{2}C_t^2\partial_{xx}f(X_t,Y_t)dt + \frac{1}{2}K_t^2\partial_{yy}f(X_t,Y_t)dt \\ &+ C_tK_t\partial_{xy}f(X_t,Y_t)dt \\ =& [A_t\partial_x f(X_t,Y_t) + J_t\partial_y f(X_t,Y_t) \\ &+ \frac{1}{2}C_t^2\partial_{xx}f(X_t,Y_t) + \frac{1}{2}K_t^2\partial_{yy}f(X_t,Y_t) \\ &+ C_tK_t\partial_{xy}f(X_t,Y_t)]dt \\ &+ [C_t\partial_x f(X_t,Y_t) + K_t\partial_y f(X_t,Y_t)]dB_t. \end{aligned}$$

The proof of this theorem is similar to the original Itô's Formula proof.

Theorem 5.11. *Let $f(t,x)$ be C^1 in t and C^2 in x, and let X_t be an Itô process where $dX_t = A_t dt + C_t dB_t$. Then*

$$\begin{aligned} df(t,X_t) =& \partial_x f(t,X_t)dX_t + \partial_t f(t,X_t)dt + \frac{1}{2}C_t^2\partial_{xx}f(t,X_t)dt \\ =& [A_t\partial_x f(t,X_t) + \partial_t f(t,X_t) + \frac{1}{2}C_t^2\partial_{xx}f(t,X_t)]dt + C_t\partial_x f(t,X_t)dB_t. \end{aligned}$$

Proof. Let $Y_t = J_t dt + K_t dB_t = t$. Then $J_t = 1$ and $K_t = 0$. Applying Theorem 5.10 yields the result, as all terms with K_t vanish. □

Recall that for standard Brownian motion B_t,

$$\langle B\rangle_t = \lim_{n\to\infty} \sum_{j\le tn} \left[X\left(\frac{j}{n}\right) - X\left(\frac{j-1}{n}\right)\right]^2 = t.$$

We can rethink these expressions as an Itô integral and a Riemann integral:

$$\int_0^t 1(dB_s)^2 = \int_0^t 1(ds).$$

From this, we get the formal rule that $dB_t^2 = dt$. We will soon show that $dB_t dt = 0$ as well. This motivates the following definition.

Definition 5.12 (Covariation)**.** Given two processes X_t and Y_t, their *Covariation* is given by

$$\langle X, Y\rangle_t := \lim_{j\le tn} \left[X\left(\frac{j}{n}\right) - X\left(\frac{j-1}{n}\right)\right]\left[Y\left(\frac{j}{n}\right) - Y\left(\frac{j-1}{n}\right)\right].$$

Note that the quadratic variation of a process is the same as its covariation with itself.

Proposition 5.13. *If X_t is continuous and Y_t has finite variation, then $\langle X, Y\rangle_t = 0$.*

For the proof, see Theorem 3.12 of [2].

Since Brownian motion B_t is continuous and t has finite variation, from this we get the additional formal rules $dB_t dt = 0$ and $dt^2 = 0$.

Theorem 5.14. *Let X_t and Y_t be Itô processes where $dX_t = A_t dt + C_t dB_t$ and $dY_t = J_t dt + K_t dB_t$. Then $\langle X, Y\rangle_t = \int_0^t C_s K_s ds$ or in differential form $d\langle X, Y\rangle_t = C_t K_t dt$.*

Proof. We give a formal proof using differential form:

$$\begin{aligned} d\langle X, Y\rangle_t &= (dX_t)(dY_t) \\ &= [A_t dt + C_t dB_t][J_t dt + K_t dB_t] \\ &= C_t K_t dB_t^2 + [A_t K_t + C_t J_t] dB_t dt + A_t J_t dt^2 \\ &= C_t K_t dt. \end{aligned}$$

□

From this, we see that for adapted process A_t with continuous paths, the quadratic variation of the Itô integral $Z_t = \int_0^t A_s dB_s$ is given by

$$\langle Z\rangle_t = \int_0^t A_s^2 ds.$$

We now look to derive the stochastic product rule. Before, however, it will be helpful to derive the usual product rule formally.

$$\begin{aligned}
d(f(t)\cdot g(t)) &= f(t+dt)\cdot g(t+dt) - f(t)\cdot g(t)\\
&= f(t+dt)\cdot g(t+dt) - f(t)\cdot g(t) + [g(t+dt)\cdot f(t) - g(t+dt)\cdot f(t)]\\
&= [f(t+dt)-f(t)]g(t+dt) + f(t)[g(t+dt)-g(t)]\\
&= (df)(g+dg) + f(dg)\\
&= (df)g + f(dg) + (df)(dg).
\end{aligned}$$

In ordinary calculus, due to the differentiability of f and g, the final term $(df)(dg)$ can be ignored. This is related to the quadratic variation and covariation of functions that are differentiable being 0. In the Stochastic Product Rule, however, we must include the covariation, giving the following theorem.

Theorem 5.15 (Stochastic Product Rule)**.** *Let X_t and Y_t be Itô processes where $dX_t = A_t dt + C_t dB_t$ and $dY_t = J_t dt + K_t dB_t$. Then $d(X_t \cdot Y_t) = X_t dY_t + Y_y dX_t + d\langle X, Y\rangle_t$. Equivalently,*

$$\begin{aligned}
X_tY_t =& X_0\cdot Y_0 \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X\cdot Y\rangle_t\\
=& X_0\cdot Y_0 + \int_0^t [X_sJ_s + Y_sA_s + C_sK_s]ds\\
&+ \int_0^t [X_sC_s + Y_sA_s]dB_s.
\end{aligned}$$

6. Stochastic Differential Equations in Finance

In this section, we apply what we know from stochastic calculus to stochastic differential equations that we will encounter in finance.

Example 6.1. Let $f(t,x) = x_0 e^{at+bx}$, and let $X_t = f(t, B_t) = e^{at+bB_t}$ Then

$$\begin{aligned}
dX_t &= df(t,B_t)\\
&= [\partial_t f(t,B_t) + \frac{1}{2}\partial_{xx} f(t,B_t)]dt + \partial_x f(t,B_t)dB_t\\
&= \left(a + \frac{b^2}{2}x_t\right)dt + bX_t dB_t.
\end{aligned}$$

Suppose we know that $dX_t = mX_t dt + \sigma X_t dB_t$. Then by equating terms, we get $a = m - \frac{\sigma^2}{2}$ and $b = \sigma$. Then

$$X_t = x_0 e^{\left(m - \frac{\sigma^2}{2}\right)t + \sigma B_t}.$$

X_t is an example of *geometric Brownian motion*. Geometric Brownian motion is used commonly in finance because it models how asset prices change as a percentage rather than as a difference, which is oftentimes more accurate and intuitive.

Definition 6.2 (Geometric Brownian Motion)**.** Let B_t be a standard Brownian motion. A process X_t is called a *geometric Brownian motion* with drift m and volatility σ if it satisfies

$$dX_t = mX_t dt + \sigma X_t dB_t = X_t[mdt + \sigma dB_t].$$

Definition 6.3 (Exponential SDE)**.** The *exponential SDE* is

$$dX_t = x_0 A_t X_t dB_t, \text{ where } X_0 = x_0.$$

Example 6.4. We look to prove that $X_t = x_0 \exp\left\{\int_0^t A_s dB_s - \frac{1}{2}\int_0^t A_s^2 ds\right\}$ solves the exponential SDE.

Let

$$Y_t = \int_0^t A_s dB_s - \frac{1}{2}\int_0^t A_s^2 ds.$$

Then

$$dY_t = -\frac{A_t^2}{2}dt + A_t dB_t.$$

Let $f(t,x) = x_0 e^x$. Then $f(x) = f'(x) = f''(x)$. Then

$$\begin{aligned} df(Y_t) =& f'(Y_t)dY_t + \frac{1}{2}A_t^2 f''(Y_t)dt \\ =& f(Y_t)[-\frac{A_t^2}{2}dt + A_t dB_t] + \frac{A_t^2}{2}f(Y_t)dt \\ =& f(Y_t)A_t dB_t. \end{aligned}$$

Then $f(Y_t) = x_0 \exp\left\{\int_0^t A_s dB_s - \frac{1}{2}\int_0^t A_s^2 ds\right\}$ solves the exponential SDE.

7. Change of Measure and Girsanov Theorem

In this section, we explore how and why we change our probability measure. In particular, we want to use a different measure to change the drift of a Brownian motion.

Definition 7.1 (Absolutely Continuous Measures)**.** Let $\mathbb{P}$, $\mathbb{Q}$ be probability measures on $(\Omega, \mathcal{F})$. $\mathbb{Q}$ is *absolutely continuous* with respect to $\mathbb{P}$ if for every $E \in \mathcal{F}$, if $\mathbb{P}(E) = 0$ then $\mathbb{Q}(E) = 0$. In other words, if $\mathbb{Q}(E) > 0$, then $\mathbb{P}(E) > 0$. This is denoted $\mathbb{Q} \ll \mathbb{P}$. If $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$, then $\mathbb{Q}$ and $\mathbb{P}$ are *mutually absolutely continuous* or *equivalent* measures.

Definition 7.2 (Singular Measures)**.** Let $\mathbb{P}$, $\mathbb{Q}$ be probability measures on $(\Omega, \mathcal{F})$. If there exists an event $E \in \mathcal{F}$ such that $\mathbb{P}(E) = 0$ and $\mathbb{P}(\Omega \setminus E) = 0$, then $\mathbb{P}$ and $\mathbb{Q}$ are called *singular measures.* This is denoted $\mathbb{P} \perp \mathbb{Q}$.

Example 7.3. Let Ω be all continuous functions from $[0,1]$ to $\mathbb{R}$. This set contains all possible values that any Brownian motion's function-valued random variable $t \mapsto B_t$ could potentially take, regardless of drift or variance. For any Brownian motion with drift 0 and variance σ^2, there is a probability measure P_σ which describes the distribution of this random variable $t \mapsto B_t$. We call this the *Wiener measure.* Consider P_σ and $P_{\sigma'}$ where $\sigma \neq \sigma'$. Let E_v denote the set of functions f that have $\langle f \rangle_1 = v^2$. We know that Brownian motion with variance σ will have quadratic variation σ^2 at $t = 1$ and Brownian motion with variance σ' will have quadratic variation $(\sigma')^2$ at $t = 1$. Then $P_\sigma(E_\sigma) = 1$, $P_{\sigma'}(E_{\sigma'}) = 1$, and $E_{\sigma'} \cap E_\sigma = \emptyset$. Then $P_\sigma(\Omega \setminus E_\sigma) = 0$ and $P_{\sigma'}(E_\sigma) = 0$, so $P_\sigma \perp P_{\sigma'}$.

This also implies that for a given measure, two Itô processes with equal dB_t terms have the same events of probability 1 and the same events of probability 0. This will be important later as we see that we can change the drift term alone by using a different but equivalent probability measure.

Example 7.4. Lebesgue measure μ is a measure on $(\mathbb{R}, \mathcal{R})$ such that the measure of any interval is its length. Then for a continuous random variable X with density f, the distribution P_X can be denoted

$$P_x(A) = \mathbb{P}(X \in A) = \int_A f(x)dx = \int_A f d\mu.$$

Note that $P_X(A) \ll \mu$. Then we say that $\frac{dP_X}{d\mu} = f$, so then $P_X(A) = \int_A \frac{dP_X}{d\mu} d\mu$.

Let Y be another continuous random variable with density g. If $P_Y \ll P_X$, then we say that $\frac{dP_Y}{dP_X} := \frac{g}{f}$. Then

$$P_Y(A) = \mathbb{P}(Y \in A) = \int_A g d\mu = \int_A \frac{g}{f} f d\mu = \int_A \frac{dP_Y}{dP_X} dP_X.$$

Definition 7.5 (σ-finite Measure)**.** A measure μ is *σ-finite* if there exist $A_1, A_2, \ldots$ such that $\mu(A_i) < \infty$ and $\Omega = \bigcup_{i=1}^{\infty} A_i$.

Theorem 7.6 (Radon-Nikodym Theorem)**.** *Let $\mathbb{P}$ and $\mathbb{Q}$ be σ-finite measures on $(\Omega, \mathcal{F})$ with $\mathbb{Q} \ll \mathbb{P}$. Then there exists f such that for every $E \in \mathcal{F}$,*

$$\mathbb{Q}(E) = \int_E f d\mathbb{P}.$$

This f is the Radon-Nikodym derivative *of $\mathbb{Q}$ with respect to $\mathbb{P}$ and is denoted*

$$f = \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

We think of this as the ratio of the measures for any value; note that if x is a point, it is possible for $\mathbb{P}(x) = 0$, $\mathbb{Q}(x) = 0$, but $\frac{d\mathbb{Q}}{d\mathbb{P}} > 0$. We use notation $\mathbb{E}_{\mathbb{P}}$ to denote expectation with respect to measure $\mathbb{P}$, and likewise for $\mathbb{Q}$.

Note that

$$\mathbb{Q}(E) = \int_E d\mathbb{Q} = \int_E \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \int_\Omega \frac{d\mathbb{Q}}{d\mathbb{P}} 1_E d\mathbb{P} = \mathbb{E}_{\mathbb{P}} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} 1_E \right].$$

Also,

$$\mathbb{E}_{\mathbb{Q}}[X] = \int_\Omega X d\mathbb{Q} = \int_\Omega X \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \mathbb{E}_{\mathbb{P}} \left[X \frac{d\mathbb{Q}}{d\mathbb{P}} \right].$$

We now explore how changing measure will allow us to change the drift of a Brownian motion with drift m and variance σ^2:

$$dX_t = mdt + \sigma dB_t.$$

To illustrate this, we discretize the process as a random walk, similar to our discussion in Sections 4 and 5. We wish to discretize in such a way that the path remains on a lattice of points (multiples of $\sigma \Delta x = \sigma\sqrt{\Delta t}$). Then at any time, the process has two options: $X(t + \Delta t) - X(t) = \pm\sigma\Delta t$. We also must maintain that $\mathbb{E}[X(t + \Delta t) - X_t|\mathcal{F}_t] = m(t, X_t)\Delta t$, so we cannot simply let the probability of either outcome be the same, which would always result in expected change of 0. We arrive at the equation $m(t, X_t) = p(\sigma\Delta t) + (1 - p)(-\sigma\Delta t)$ where $p = \mathbb{P}\{X(t + \Delta t) - X(t) = \sigma\Delta t\}$. Solving yields

$$\mathbb{P}\{X(t + \Delta t) - X(t) = \sigma\Delta t\} = \frac{1}{2} \left[1 + \frac{m(t, X_t)\Delta t}{\sigma} \right],$$

$$\mathbb{P}\{X(t + \Delta t) - X(t) = -\sigma\Delta t\} = \frac{1}{2} \left[1 - \frac{m(t, X_t)\Delta t}{\sigma} \right].$$

We fix $\sigma = 1$, and we look to sample from X_t using values from standard Brownian motion B_t. However, a set of paths that might be likely for B_t might be possible but less likely for X_t; as we will see, a change of measure is required. So, we need a way to describe the ratio between probabilities of a path for X_t and B_t.

We observe a discretized standard Brownian motion B_t. Let N be a very large positive integer, and let $\Delta t := \frac{1}{N}$, so $\Delta x = \sqrt{\Delta t} = \frac{1}{\sqrt{N}}$. After n steps, there are 2^n equally probable paths that B_t could have taken. We denote these paths as

$$\omega = (\omega_1, \omega_2, \dots, \omega_n)$$

where ω_i is 1 or -1 if the i-th step is up or down, respectively. Let J be the number of steps up in the path ω that we observe. Let $r := \frac{2J-n}{2\sqrt{N}}$. The position at time $t = n\Delta t = \frac{n}{N}$ is

$$\begin{aligned} B(t) &= B(n\Delta t) \\ &= \Delta x(\omega_1 + \cdots + \omega_n) \\ &= \frac{1}{\sqrt{N}}(J - (n - J)) \\ &= \frac{2J - n}{\sqrt{N}} \\ &= 2r. \end{aligned}$$

For each possible ω, the probability of it occurring as B_t is $(\frac{1}{2})^n$. The probability of it occurring as X_t however is

$$\mathbb{Q}(\omega) = \left(\frac{1}{2}\right)^n \left[1 + m\sqrt{t}\right]^J \left[1 - m\sqrt{t}\right]^{n-J}.$$

Note that $J = (n/2) + r\sqrt{N}$. Then the ratio of probabilities from X_t to B_t is

$$\begin{aligned} [1 + m\sqrt{t}]^J[1 - m\sqrt{t}]^{n-J} &= \left[1 + \frac{m}{\sqrt{N}}\right]^J \left[1 - \frac{m}{\sqrt{N}}\right]^{n-J} \\ &= \left[1 + \frac{m}{\sqrt{N}}\right]^{(n/2)+r\sqrt{N}} \left[1 - \frac{m}{\sqrt{N}}\right]^{(n/2)-r\sqrt{N}} \\ &= \left[\left[1 + \frac{m}{\sqrt{N}}\right]^{n/2} \left[1 - \frac{m}{\sqrt{N}}\right]^{n/2}\right] \left[1 + \frac{m}{\sqrt{N}}\right]^{r\sqrt{N}} \left[1 - \frac{m}{\sqrt{N}}\right]^{r\sqrt{N}} \\ &= \left[1 - \frac{m^2}{N}\right]^{n/2} \left[1 + \frac{m}{\sqrt{N}}\right]^{r\sqrt{N}} \left[1 - \frac{m}{\sqrt{N}}\right]^{r\sqrt{N}} \\ &= \left(\left[1 + \frac{-m^2}{N}\right]^N\right)^{t/2} \left(\left[1 + \frac{m}{\sqrt{N}}\right]^{\sqrt{N}}\right)^r \left(\left[1 + \frac{-m}{\sqrt{N}}\right]^{\sqrt{N}}\right)^{-r}. \end{aligned}$$

We use $(1 + \frac{a}{N})^N \sim e^a$ and find that the limit as $N \to \infty$ is

$$(e^{-m^2})^{t/2}(e^m)^r(e^{-m})^{-r} = e^{\frac{-m^2 t}{2}} e^{rm} e^{rm} = e^{\frac{-m^2 t}{2} + 2rm} = e^{mB_t - \frac{m^2 t}{2}}.$$

This gives us the ratio of probabilities between X_t and B_t, meaning that we can sample from B_t and weight the samples by

$$e^{mB_t - \frac{m^2 t}{2}}.$$

In fact, this gives the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = M_t$$

where $\mathbb{Q}$ is the probability measure for X_t and $\mathbb{P}$ is the probability measure for B_t. In other words, for any $V \in \mathcal{F}_t$,

$$Q(V) = \mathbb{E}[M_t 1_V].$$

Note from Example 6.4 that M_t solves the exponential SDE

$$dM_t = e^{mB_t - \frac{m^2}{t}},\ M_0 = 1.$$

Proposition 7.7. *Let B_t be a standard Brownian motion. Then $M_t = e^{mB_t - \frac{m^2 t}{2}}$ is a martingale.*

Proof. Let $s < t$. Recall that $B_t - B_s$ is independent of B_s and has distribution $N(0, t-s)$, meaning that it has moment generating function $f(m) = e^{\mu \cdot t} e^{\sigma^2 m^2/2} = e^{0 \cdot t} e^{(t-s)m^2/2} = e^{(t-s)m^2/2}$ Then

$$\begin{aligned} E[M_t|M_s] &= E[e^{mB_t - \frac{m^2 t}{2}} | e^{mB_s - \frac{m^2 s}{2}}] \\ &= e^{\frac{-m^2 t}{2}} E[e^{m(B_t - B_s) + mB_s} | B_s] \\ &= e^{\frac{-m^2 t}{2} + mB_s} \mathbb{E}[e^{m(B_t - B_s)}] \\ &= e^{\frac{-m^2 t}{2} + mB_s} e^{(t-s)m^2/2} \\ &= e^{\frac{-m^2 s}{2} + mB_s} \\ &= M_s. \end{aligned}$$

□

Example 7.8 (Risk-Neutral Measure)**.** Let B_t be a standard Brownian motion under measure $\mathbb{P}$. Let X_t be a geometric Brownian motion satisfying

$$dX_t = X_t[mdt + \sigma dB_t].$$

We aim to find an equivalent measure $\mathbb{Q}$ such that X_t has drift r under measure $\mathbb{Q}$. Then if W_t is a standard Brownian motion with respect to $\mathbb{Q}$, we want $mdt + \sigma dB_t = rdt + \sigma dW_t$, so $dB_t = \frac{r-m}{\sigma} dt + dW_t$. Since only the drift term changed, we know that $\mathbb{Q}$ is an equivalent probability measure. Oftentimes, in finance, if X_t models stock price and R_t satisfying $dR_t = rR_t dt$ models bond price, we want to consider a *risk-neutral measure* such that X_t/R_t is a martingale. As long as certain conditions are met to avoid local martingales, we see that changing the measure so that the drift term is rdt is a useful step. This will be applied in the next section, when deriving the Black-Scholes Formula.

So far, we have changed measure so that a standard Brownian motion gains a constant drift m. However, with the Girsanov Theorem, we can give Brownian motion drift A_t.

Theorem 7.9 (Girsanov Theorem)**.** *Let B_t be a standard Brownian motion under measure $\mathbb{P}$. Let M_t satisfy*

$$dM_t = A_t M_t dB_t,\ M_0 = 1.$$

That is, $M_t = e^{Y_t}$ *where* $Y_t = \int_0^t A_s dB_s - \frac{1}{2}\int_0^t A_s^2 ds$*. Let*

$$T_n = \inf\{t : M_t + \langle Y \rangle_t = n\},\ T := T_\infty := \lim_{n\to\infty} T_n.$$

Note that if M_t *is a non-negative martingale, then* $T = \infty$*. Let* $\mathbb{Q}$ *be the equivalent probability measure such that for all* $V \in \mathcal{F}_t$*,*

$$\mathbb{Q}(V) := \mathbb{E}_{\mathbb{P}}[1_V M_t].$$

Let

$$W_t = B_t - \int_0^t A_s ds,\ t < T.$$

Then with respect to measure $\mathbb{Q}$*, the process* W_t *for* $t < T$ *is a standard Brownian motion, so*

$$dBt = A_t dt + dWt,\ t < T.$$

If any of the following conditions hold, then M_s *is a martingale for* $s \le t$*.*

- $\mathbb{Q}(M_t + \langle Y \rangle_t < \infty) = 1.$
- $\mathbb{E}_{\mathbb{P}}[M_t] = 1.$
- $\mathbb{E}_{\mathbb{P}}\left[\exp\left\{\frac{\langle Y \rangle_t}{2}\right\}\right] < \infty.$

For the proof, see pages 153-155 of [6]. An intuitive explanation can be found through binomial approximation as shown earlier.

We can think of this as "tilting" the Brownian motion's measure by the specified local martingale, and weighting samples accordingly as we described earlier. This will be used in deriving the Black-Scholes Formula.

8. Application: Options Pricing

One of the most significant uses of Itô Calculus is in the derivation of the *Black-Scholes Equation.* This partial differential equation describes the price of various financial assets, most notably European options. We derive the related Feynman-Kac partial differential equation before deriving the Black-Scholes Equation, and finally we solve the Black-Scholes Equation for pricing European options, arriving at the *Black-Scholes Formula.* Note that the Black-Scholes Model works under certain assumptions about the market, many of which are not completely accurate. In modern-day applications, it is often altered to take into account these simplifications, but it is still extremely useful.

The owner of a *European call option* has the right to buy an underlying stock at a specified *strike price* K, at a specified expiry time T. It is important to realize that this is a *right* instead of an obligation, so naturally the option is only exercised (i.e., the stock is bought at K dollars) if the stock price is above K dollars; otherwise the option owner simply does nothing. It is also important to realize that unlike an American option, the European option cannot be exercised before expiry time T, even if the owner finds the price suitable. A *put option* gives the right to sell the underlying stock; mathematically this is no more complicated, and any calculations or analysis we perform for call options could be easily modified to address put options. From now on, when we say *option* without further specification, we assume the option to be a European call option.

Suppose that the underlying stock's price is described by

$$dS_t = m(t, S_t)dt + \sigma(t, S_t)dB_t.$$

For example, it could be the case that $m(t, S_t) = mS_t$ and $\sigma(t, S_t) = \sigma S_t$, which would give geometric Brownian motion. Note that S_t is a *Markov* process, which means that if $r > t$, then $E[S_r|\mathcal{F}_t] = E[S_r|S_t]$.

It only makes sense to exercise this right if $S_T > K$, so the payoff of this option at T is

$$F(S_T) = (S_T - K)_+ = \begin{cases} (S_T - K) & , S_T > K, \\ 0 & , S_T \leq K. \end{cases}$$

We also suppose an inflation rate of $r(t, x)$ such that R_0 dollars at time 0 is worth R_t dollars at time t, where

$$dR_t = r(t, S_t)R_t dt.$$

Then

$$R_t = R_0 \exp\left\{\int_0^t r(s, S_s)ds\right\}.$$

To illustrate how we use this, say we have X dollars at time b and we want to find the value of this amount at time $a < b$. Due to inflation, we know that our answer should be less than X. We know that R_a dollars at time a are worth R_b dollars at time b, so $X\frac{R_a}{R_b} = X \exp\left\{-\int_a^b r(t, S_t)dt\right\}$ dollars at time a are worth X dollars at time b.

We aim to find the expectation of the payoff at expiry time T, and find the value of that payoff amount in time $t \leq T$ dollars. Mathematically, we are looking for

$$f(t, x) = E\left[\frac{R_t}{R_T}F(S_T)|S_t = x\right] = E\left[\exp\left\{-\int_t^T r(s, X_s)ds\right\}F(S_T)|S_t = x\right].$$

We assume that f is C^1 in t and C^2 in x.

Consider the process

$$M_t = E[R_T^{-1}F(S_T)|\mathcal{F}_t].$$

Since R_t is $\mathcal{F}_t$-measurable and S_t is a Markov process, we can also write

$$M_t = R_t^{-1}E[\exp\left\{\int_t^T r(s, S_s)ds\right\}F(S_T)|S_t = x] = R_t^{-1}f(t, X_t).$$

Note that $M_T = E[R_T^{-1}F(S_T)|\mathcal{F}_T] = R_T^{-1}F(S_T)$, so we plug in to get $M_t = E[M_T|\mathcal{F}_t]$. We then note that M_t is a martingale, since if $s < t$, then

$$E[M_t|\mathcal{F}_s] = E[E[M_T|\mathcal{F}_t]|\mathcal{F}_s] = E[M_T|\mathcal{F}_s] = M_s$$

by Tower Property. Since it is a martingale, we can apply Stochastic Product Rule and Itô's Formula and set all dt terms to 0. Note that R_t has finite variation so the resulting covariation is 0. Also, we can see from normal calculus that $d[R_t^{-1}] = R_t^{-1}(-r(t, S_t))dt$. Then

$$
\begin{aligned}
dM_t =& d[R_t^{-1} f(t,x)] \\
=& f(t,x) d[R_t^{-1}] + R_t^{-1} df(t,x) + 0 \\
=& f(t,x)[R_t^{-1}(-r(t,S_t)dt] \\
&+ R_t^{-1}[\partial_x f(t,S_t) dS_t + \partial_t f(t,S_t)dt + \frac{1}{2}\sigma(t,S_t)^2 \partial_{xx} f(t,S_t)dt] \\
=& f(t,x)[R_t^{-1}(-r(t,S_t)dt] \\
&+ R_t^{-1}[\partial_x f(t,S_t)[m(t,S_t)dt + \sigma(t,S_t)dB_t] + \partial_t f(t,S_t)dt + \frac{1}{2}\sigma(t,S_t)^2 \partial_{xx} f(t,S_t)dt] \\
=& R_t^{-1}[(-r(t,S_t) + m(t,S_t)\partial_x f(t,S_t) + \partial_t f(t,S_t) + \frac{1}{2}\sigma(t,S_t)^2 \partial_{xx} f(t,S_t)]dt \\
&+ [\sigma(t,S_t)\partial_x f(t,S_t)]dB_t.
\end{aligned}
$$

We set the dt term to 0 and isolate $\partial_t f(t,x)$ on the left hand side to get the Feynman-Kac PDE:

$$\partial_t f(t,x) = -m(t,x)\partial_x f(t,x) - \frac{1}{2}\sigma(t,x)^2 \partial_{xx} f(t,x) + r(t,x) f(t,x).$$

We have proven the following theorem.

Theorem 8.1 (Feynman-Kac Formula)**.** *Suppose the price S_t of a stock is described by*

$$dS_t = m(t,S_t)dt + \sigma(t,S_t)dB_t$$

and the value of R_0 dollars at time t is R_t satisfying

$$dR_t = r(t,S_t)R_t dt$$

where $r(t,S_t)$ is the rate of inflation. Suppose there exists a payoff $F(S_T)$ at time T based on the stock price, satisfying $\mathbb{E}[|F(S_t)|] < \infty$. Then if

$$f(t,x) = E[R_T/R_t F(S_T)|S_t = x]$$

is C^1 in t and C^2 in x, then it satisfies

$$\partial_t f(t,x) = -m(t,x)\partial_x f(t,x) - \frac{1}{2}\sigma(t,x)^2 \partial_{xx} f(t,x) + r(t,x) f(t,x)$$

for $0 \le t \le T$ with $f(T,x) = F(x)$.

The Feynman-Kac PDE is very similar to but still different from the Black-Scholes PDE, which describes how options are priced. It turns out that pricing an option according to its expected value in time t dollars, though an intuitively sensible idea, can lead to arbitrage opportunities. *Arbitrage* occurs when a trading strategy has a positive probability of making money and a zero probability of losing money. In other words, there is risk-less profit; we naturally expect this to be impossible when things are priced correctly. Someone seeking arbitrage might *hedge*, or reduce risks, using a *self-financing* portfolio. Changes in a self-financing portfolio's value are only caused by changes in its assets; there is no inflow or outflow of funds to the portfolio.

Suppose we have a call option on a stock priced at $f(t,S_t)$ with payoff $F(S_T)$ at time T. Note that at time T, there is no uncertainty about the payoff, so the option price would be equal to the payoff. If $f(T,S_T) < F(S_T)$ then one could arbitrage by buying the option and collecting the payoff; in the other case, one would sell the option knowing that the payoff is less. Even if $f(T, S_T) = F(S_T)$, arbitrage for this option could still occur if one could construct a self-financing portfolio (a_t, b_t) with the same payoff $F(S_T)$ as the option, having a_t shares of the underlying stock at price S_t and b_t risk-free bonds at price R_t, but the portfolio can be obtained for a price different from the option. In other words, arbitrage is possible if the portfolio value $V_t = a_t S_t + b_t R_t$ satisfies $V_T = F(S_T) = f(T, S_T)$ with probability one but $V_t \neq f(t, S_t)$ at some $t < T$. If $V_t < f(t, S_t)$, then an investor could sell an option for $f(t, S_t)$ dollars, then invest V_t dollars in the portfolio and the remaining $f(t, S_t) - V(t)$ dollars in risk-free bonds; the portfolio will have the same outcome as the option, so $f(t, S_t) - V(t)$ would be instant risk-less profit. Conversely, if the option is under-priced, one could buy the derivative and sell the portfolio accordingly.

We now aim to derive the Black-Scholes Equation by assuming no-arbitrage pricing of an option and evaluating a self-financing portfolio that replicates its price. Suppose that the stock price S_t satisfies

$$dS_t = S_t[m(t, S_t)dt + \sigma(t, S_t)dB_t].$$

This time, we let $r(t, S_t)$ represent the risk-free rate so that R_t represents the risk-free bond price, satisfying the same SDE

$$dR_t = r(t, S_t)R_t dt.$$

This is a similar notion to before, as an R_0 dollar bond bought at time 0 should be worth R_t dollars at time t. Let V_t be the value of a portfolio of a_t stocks and b_t bonds that is constructed to guarantee a value of $V_T = F(S_T)$ at expiry. We note that constructing such a portfolio is not only possible but straightforward: the initial value of the portfolio is simply the option price

$$V_0 = f(0, S_0)$$

and we then switch between stocks and bonds such that the overall value is always equal to the option price

$$V_t = f(t, S_t)$$

thereby guaranteeing that $V_T = f(T, S_T) = F(S_T)$. We can solve for the exact manner in which we would have to switch between stocks and bonds in order to maintain the relationship $V_t = f(t, S_t)$. In fact, this is the exact portfolio we will examine, as $V_T = F(S_T)$ with probability one but there is no possibility for the arbitrage process we described earlier, since $V_t = f(t, S_t)$.

By Stochastic Product Rule and Proposition 5.13, we know that

$$dV_t = a_t dS_t + S_t da_t + \langle a, S\rangle_t + b_t dR_t + R_t db_t.$$

However, the mathematical consequence of being self-financing is that

$$dV_t = a_t dS_t + b_t dR_t,$$

so after placing this self-financing condition, we can proceed by substituting dS_t and dR_t:

$$\begin{aligned} dV_t &= a_t[S_t[m(t, S_t)dt + \sigma(t, S_t)dB_t]] + b_t[r(t, S_t)R_t dt] \\ &= a_t[S_t[m(t, S_t)dt + \sigma(t, S_t)dB_t]] + r(t, S_t)[V_t - a_t S_t]dt \\ &= [m(t, S_t)a_t S_t + r(t, S_t)(V_t - a_t S_t)]\, dt + \sigma(t, S_t)a_t S_t dB_t. \end{aligned}$$

Alternatively, we can apply Itô's Formula:

$$
\begin{aligned}
df(t,x) =& \partial_t f(t,S_t)dt + \partial_x f(t,S_t)dS_t + \frac{1}{2}S_t^2\sigma(t,S_t)^2\partial_{xx}f(t,S_t)dt \\
=& \partial_t f(t,S_t)dt + \partial_x f(t,S_t)[S_t[m(t,S_t)dt + \sigma(t,S_t)dB_t]] \\
&+ \frac{S_t^2\sigma(t,S_t)^2}{2}\partial_{xx}f(t,S_t)dt \\
=& \left[\partial_t f(t,S_t) + m(t,S_t)S_t\partial_x f(t,S_t) + \frac{\sigma(t,S_t)^2S_t^2}{2}\partial_{xx}f(t,S_t)\right]dt \\
&+ \sigma(t,S_t)S_t\partial_x f(t,S_t)dB_t.
\end{aligned}
$$

Since $V_t = f(t,S_t)$, we now equate the dB_t terms, and then the dt terms. The former tells us how exactly to manage the portfolio, as it gives an equation for a_t which in turn gives a formula for b_t:

$$a_t = \partial_x f(t,S_t),\ b_t = \frac{f(t,x) - a_tS_t}{R_t}.$$

After substituting for a_t, the latter gives the Black-Scholes Equation:

$$\partial_t f(t,x) = r(t,x)f(t,x) - r(t,x)x\partial_x f(t,x) - \frac{\sigma(t,x)^2x^2}{2}\partial_{xx}f(t,x).$$

Note that $m(t,S_t)$ does not appear in this PDE. This is because Itô processes with the same $\sigma(t,S_t)$ have the same events of probability one; this fact was shown in Example 7.3. So, this relationship describing how the price changes over time should be independent of the drift term of S_t. Also, if we apply the Feynman-Kac Formula to this scenario, we get

$$\partial_t f(t,x) = r(t,x)f(t,x) - m(t,x)x\partial_x f(t,x) - \frac{\sigma(t,x)^2x^2}{2}\partial_{xx}f(t,x).$$

The only difference is that the $m(t,x)$ in the Feynman-Kac PDE is replaced by $r(t,x)$ in the Black-Scholes PDE. We have proven our earlier statement that pricing according to Feynman-Kac PDE may lead to arbitrage opportunities: arbitrage is only eliminated when pricing follows the Black-Scholes PDE, and it cannot simultaneously follow both PDE's when $m(t,x) \neq r(t,x)$, so in this case arbitrage must be possible under Feynman-Kac PDE pricing. It is worth considering how one would find arbitrage in this case.

Example 8.2 (Arbitrage with Feynman-Kac PDE Pricing)**.** We construct a self-financing portfolio with value $V_t = a_tS_t + b_tR_t$ as before with $V_0 = f(0,S_0)$, although V and f will not always be equal. We simplify the SDE for dV_t under the self-financing condition, and simplify the SDE for $df(t,x)$ under Itô's formula. This time, however, we do not equate V_t and $f(t,x)$. We equate the dB_t terms to get the same result $a_t = \partial_x f(t,x)$. This tells us how to manage the self-financing portfolio. We can then subtract $df(t,x)$ from dV_t, and since the dB_t terms are eliminated, we get a deterministic PDE. Solving this, we get an expression for $V_T - f(T,S_T)$. We can then arbitrage as described earlier based on the sign of this expression.

Nevertheless, if we were to have $m(t,x) = r(t,x)$, then the Feynman-Kac PDE would hold true; the option price at time t would be described by the expected value of the payoff in time t dollars. We use this fact to solve the Black-Scholes equation and get the *Black-Scholes Formula* which gives a closed form equation for the option price. We want to use the Feynman-Kac Formula to price the option, since the Feynman-Kac formula gives a closed form expression for $f(t,x)$ using expected value. Since changing the drift term does not affect the Black-Scholes PDE but can allow us to use the Feynman-Kac Formula, we must find a different probability measure $\mathbb{Q}$ such that if W_t is a standard Brownian motion in $\mathbb{Q}$, then

$$dS_t = S_t[r(t,S_t)dt + \sigma(t,S_t)dW_t].$$

This is in fact the *risk-neutral measure* referred to in Example 7.8. Since we know that $dS_t = S_t[m(t,S_t)dt + \sigma(t,S_t)dB_t]$, we can set the inner expressions equal to each other and solve to get

$$dB_t = \frac{r(t,S_t) - m(t,S_t)}{\sigma(t,S_t)}dt + dW_t.$$

Note that since these Brownian motions only differ in their drift term, the corresponding measures are equivalent; this was shown in Example 7.3.

We must "tilt" the measure by the local martingale M_t that satisfies

$$dM_t = M_t\frac{r(t,S_t) - m(t,S_t)}{\sigma(t,S_t)}dB_t.$$

We know how to solve this SDE, as shown in Example 6.4. Under certain conditions, such as when $\frac{r(t,S_t)-m(t,S_t)}{\sigma(t,S_t)}$ is uniformly bounded, M_t is in fact a martingale. This allows us to apply Girsanov Theorem for all $t \geq 0$.

So, we know that we can find an equivalent measure $\mathbb{Q}$ by Girsanov's Theorem such that $dS_t = S_t[r(t,S_t)dt + \sigma(t,S_t)dW_t]$, so then the portfolio value and option price satisfy the Feynman-Kac PDE using expectation with respect to $\mathbb{Q}$:

$$V_t = f(t,S_t) = E_{\mathbb{Q}}(R_t/R_T F(S_T)|S_t) = E_{\mathbb{Q}}(R_t/R_T F(S_T)|\mathcal{F}_t).$$

The processes $\widetilde{S}_t = S_t/R_t$ and $\widetilde{V}_t = V_t/R_t$ are the stock price and portfolio value discounted by the bond rate. Applying Stochastic Product Rule, we see that

$$\begin{aligned}
d\widetilde{S}_t &= d[S_t R_t^{-1}] \\
&= S_t d[R_t^{-1}] + R_t^{-1} dS_t \\
&= S_t[R_t^{-1}(-r(t,S_t)dt] + R_t^{-1}[S_t[r(t,S_t)dt + \sigma(t,S_t)dW_t]] \\
&= S_t/R_t[(-r(t,S_t)dt + r(t,S_t)dt + \sigma(t,S_t)dW_t] \\
&= \widetilde{S}_t\sigma(t,S_t)dW_t.
\end{aligned}$$

Then, as expected under risk-neutral measure $\mathbb{Q}$, $\widetilde{S}_t$ is a martingale given certain conditions on $\sigma(t,S_t)$, for example it being uniformly bounded. Also,

$$\widetilde{V}_t = V_t/R_t = R_t^{-1}E_{\mathbb{Q}}(R_t/R_T F(S_T)|\mathcal{F}_t) = E_{\mathbb{Q}}(R_T^{-1}F(S_T)|\mathcal{F}_t) = E_{\mathbb{Q}}(\widetilde{V}_T|\mathcal{F}_t).$$

By multiplying by R_t, we get the following theorem.

Theorem 8.3. *Let B_t be a standard Brownian motion with respect to probability measure $\mathbb{P}$. Let stock price S_t satisfy*

$$dS_t = S_t[m(t,S_t)dt + \sigma(t,S_t)dB_t]$$

and risk-free bond value R_t with risk-free rate $r(t,S_t)$ satisfy

$$dR_t = r(t,S_t)R_t dt.$$

Suppose that $\frac{r(t,S_t)-m(t,S_t)}{\sigma(t,S_t)}$ is uniformly bounded and $\sigma(t,S_t) > 0$ is uniformly bounded. Then there exists a probability measure $\mathbb{Q}$ that is equivalent to $\mathbb{P}$ such that *the discounted stock price $\widetilde{S}_t = S_t/R_t$ is a martingale under $\mathbb{Q}$. Suppose there is an option for the stock with payoff $F(S_T)$ at time T, satisfying $E_{\mathbb{Q}}[R_T^{-1}|F(S_T)|] < \infty$. Then the arbitrage-free price of the option at time t is*

$$V_t = R_t E_{\mathbb{Q}}(R_T^{-1} F(S_T)|\mathcal{F}_t).$$

One necessary assumption for the Black-Scholes Formula is that $r(t, S_t)$ and $\sigma(t, S_t)$ are constants r and σ. Then we let $\widetilde{S}_t = e^{-rt}S_t$, $\widetilde{V}_t = e^{-rt}V_t$, $\widetilde{K} = e^{-rT}K$, so that

$$\widetilde{V}_T = e^{-rt}V_T = e^{-rt}F(S_T) = e^{-rt}(S_T - K)_+ = (\widetilde{S}_T - \widetilde{K})_+.$$

Then if Z is the conditional distribution of $\widetilde{S}_T$ given $\mathcal{F}_t$ and g is the density of Z, we get

$$\widetilde{V}_t = E_{\mathbb{Q}}(\widetilde{V}_T|S_t) = E_{\mathbb{Q}}((\widetilde{S}_T - \widetilde{K})_+|S_t) = \int_{-\infty}^{\infty} (z - \widetilde{K})_+ g(z)dz = \int_{\widetilde{K}}^{\infty} (z - \widetilde{K})g(z)dz.$$

Now we must find Z and $g(z)$, and then compute. $\widetilde{S}_t$ satisfies the exponential SDE

$$d\widetilde{S}_t = \sigma \widetilde{S}_t dW_t,$$

so we know from Example 6.4 that

$$\begin{aligned} \widetilde{S}_t &= \widetilde{S}_0 \exp\left\{\int_0^t \sigma dW_s - \frac{1}{2}\int_0^t \sigma^2 ds\right\} \\ &= \widetilde{S}_0 \exp\left\{\sigma W_t - \frac{\sigma^2 t}{2}\right\} \end{aligned}$$

and hence

$$\widetilde{S}_T = \widetilde{S}_t \exp\left\{\sigma(W_T - W_t) - \frac{\sigma^2(T-t)}{2}\right\}.$$

Given $\mathcal{F}_t$, the distribution of $W_T - W_t$ is $\sqrt{T-t}D$ where D has a standard normal distribution. The rest of the expression is known. Then

$$Z = \widetilde{S}_t \exp\left\{\sigma\sqrt{T-t}D - \frac{\sigma^2(T-t)}{2}\right\} = \exp\left\{\sigma\sqrt{T-t}D - \frac{\sigma^2(T-t)}{2} + \log(S_t)\right\}.$$

Let $a = \sigma\sqrt{T-t}$ and $y = \log(S_t) - \frac{a^2}{2}$ be constants. Then

$$Z = \exp\{aD + y\}.$$

Note that $aD + y \sim N(y, a^2)$. Then since $\log(Z)$ has a normal distribution $N(y, a^2)$, we say that Z has a *log-normal distribution* with density $g(z) = \frac{1}{az}\phi(\frac{\log z - y}{a})$. We omit this density calculation; we refer the reader to page 10 of [8].

Lemma 8.4. *If Z has log-normal distribution with density function $g(z)$ under measure $\mathbb{Q}$, and the variance of $\log(Z)$ is a^2, then*

$$\int_K^{\infty} (z - K)g(z)dz = \mathbb{E}_{\mathbb{Q}}[Z]\Phi(d_1) - K\Phi(d_2)$$

where

$$d_1 = \frac{\log(\mathbb{E}_{\mathbb{Q}}[Z]/K) + \frac{a^2}{2}}{a},\ d_2 = \frac{\log(\mathbb{E}_{\mathbb{Q}}[Z]/K) - \frac{a^2}{2}}{a}$$

For the proof, see pages 380-381 of [7].

Then

$$\widetilde{V}_t = \int_{\widetilde{K}}^{\infty} (z - \widetilde{K}) g(z) dz = \mathbb{E}_{\mathbb{Q}}[Z]\Phi(d_1) - \widetilde{K}\Phi(d_2)$$

where

$$d_1 = \frac{\log(\mathbb{E}_{\mathbb{Q}}[Z]/\widetilde{K}) + \frac{a^2}{2}}{a}, \; d_2 = \frac{\log(\mathbb{E}_{\mathbb{Q}}[Z]/\widetilde{K}) - \frac{a^2}{2}}{a}.$$

We know that $\mathbb{E}_{\mathbb{Q}}[Z] = E_{\mathbb{Q}}(\widetilde{S}_T|\widetilde{S}_t) = \widetilde{S}_t$ since $\widetilde{S}_t$ is a martingale. We substitute this for $\mathbb{E}_{\mathbb{Q}}[Z]$ and substitute $a = \sigma\sqrt{T-t}$, multiply by e^{rt}, and then simplify in terms of the original (not discounted) prices.

$$\begin{aligned} V_t =& e^{rt}\widetilde{S}_t\Phi\left(\frac{\log(\widetilde{S}_t/\widetilde{K}) + \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right) - e^{rt}\widetilde{K}\Phi\left(\frac{\log(\widetilde{S}_t/\widetilde{K}) - \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right) \\ =& S_t\Phi\left(\frac{\log\left(\frac{S_t e^{-rt}}{Ke^{-rT}}\right) + \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right) - e^{rt}e^{-rT}K\Phi\left(\frac{\log\left(\frac{S_t e^{-rt}}{Ke^{-rT}}\right) - \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right) \\ =& S_t\Phi\left(\frac{\log(S_t/K) + r(T-t) + \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right) \\ &- e^{-r(T-t)}K\Phi\left(\frac{\log(S_t/K) + r(T-t) - \frac{\sigma(T-t)}{2}}{\sigma\sqrt{T-t}}\right). \end{aligned}$$

This gives the Black-Scholes Formula, which says that the no-arbitrage price of this option is

$$V_t = S_t\Phi(d_1) - e^{-r(T-t)}K\Phi(d_2),$$

where

$$d_1 = \frac{\log(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \; d_2 = \frac{\log(S_t/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

Acknowledgments

I express my sincere gratitude to my mentors Ronno Das and Oliver Wang for their insightful guidance over the course of the past two months. I am also pleased to thank professor Peter May for arranging this eye-opening Math Research Experience for Undergraduates program and for reviewing this paper.

References

[1] Bernt Oksendal. Stochastic Differential Equations: An Introduction with Applications. Springer-Verlag. 2003.

[2] Brian Chen. Stochastic Calculus and an Applications to the Black-Scholes Model. http://math.uchicago.edu/ may/REU2013/REUPapers/ChenB.pdf.

[3] Fabrice D. Rouah. Four Derivations of the Black-Scholes PDE. https://www.frouah.com/finance%20notes/Black%20Scholes%20PDE.pdf.

[4] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. https://www.cs.princeton.edu/courses/archive/fall09/cos323/papers/black_scholes73.pdf.

[5] Gregory F. Lawler. Notes On Probability. http://www.math.uchicago.edu/ lawler/probnotes.pdf.

[6] Gregory F. Lawler. Stochastic Calculus: An Introduction with Applications. http://www.math.uchicago.edu/ lawler/finbook2.pdf.

[7] John C. Hull. Options, Futures, and Other Derivatives. Pearson. 1989.
[8] Ovidiu Calin. An Introduction to Stochastic Calculus with Applications to Finance. https://people.emich.edu/ocalin/Teaching_files/D18N.pdf
[9] Rick Durrett. Probability: Theory and Examples. Cambridge University Press. 2010.