Stochastic Calculus with Applications to Finance

Bubacarr Kandeh

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Applied Mathematics and Computer Science

Eastern Mediterranean University
July 2020
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Prof. Dr. Nazim Mahmudov
Chair, Department of Mathematics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Asst. Prof. Dr. Arran Fernandez
Co-Supervisor

Prof. Dr. Aghamirza Bashirov
Supervisor

Examining Committee

1. Prof. Dr. Aghamirza Bashirov
2. Asst. Prof. Dr. Arran Fernandez
3. Asst. Prof. Dr. Mustafa Kara
4. Asst. Prof. Dr. Mohammad Momenzadeh
5. Asst. Prof. Dr. Yücel Tandoğdu

ABSTRACT

The two most fundamental aspects of mathematical finance are; portfolio optimization and portfolio pricing. Portfolio optimization uses concepts from linear algebra and ordinary multi-variable calculus. On the other hand, portfolio pricing is modelled by stochastic calculus. In this work we will focus our interest in the development of stochastic calculus and how it is applied to finance in Portfolio Pricing.

Keywords: Stochastic and Wiener Processes, Itō calculus, Instantaneous Interest rate models, Continuous time pricing of the European call option, Black Scholes formula.

ÖZ

Matematiksel finansın en temel iki yönü; portföy optimizasyonu ve portföy fiyatlandırmasıdır. Portföy optimizasyonu, doğrusal cebirden ve sıradan çok değişkenli hesaplardan kavramları kullanır. Öte yandan, portföy fiyatlaması stokastik hesapla modellenmiştir. Bu çalışmada stokastik analizin gelişimine ve Portföy Fiyatlandırmasında finansmana nasıl uygulandığına odaklanacağız.

Anahtar Kelimeler: Stokastik ve Wiener Süreçleri, Itō hesabı, Anlık Faiz oranları modelleri, Avrupa çağrı seçeneğinin sürekli zaman fiyatlandırması, Black Scholes formülü.

ACKNOWLEDGMENT

I would like to thank Allah the Almighty for giving me the strength and wisdom to successfully complete this thesis. I would like to express my most sincere appreciations to my supervisors Prof. Dr. Agamirza Bashirov and Asst. Prof. Arran Fernandez for their invaluable contributions and guidance during the preparation of my thesis. I would like to thank all my instructors at the department of mathematics. My warmest thanks and gratitude goes to my friends and colleagues for their support and encouragements. Finally I would like to extend my very special thanks to my family for their continued support, encouragement and understanding throughout my masters programme.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

CDF	Cumulative Distribution Function
MMSEE	Minimum Mean Squared Error Estimator
PDF	Probability Density Function
PMF	Probability Mass Function
RV	Random Variable
SDE	Stochastic Differential Equation

Chapter 1

INTRODUCTION

The functions we encounter in ordinary (Newtonian) calculus like Polynomials, Exponentials, Logarithms, Radicals, etc do not capture the zig-zag motion followed by the graph modelling the prices of a stock in a stock market. To model and be able to describe the changes and behaviours of a stock in a stock market, there became a necessity to extend ordinary calculus to be able to describe the derivatives and integrals of functions that are continuous but have no slope (not differentiable) at any point. In this work we have introduced the basic concepts from probability and measure theory which are needed to build up a stochastic process. This part was written based on Brzeźniak and Zastawniak, *Basic Stochastic Processes*, with other references. This is covered in chapter 2 and 3. In chapter 4 we dealt with Stochastic Calculus and how to use the Itō formula, this chapter was written based on Thomas Mikosch, *Elementary stochastic calculus*, with some other reference materials. An important aspect of stochastic calculus is their differential equations, this has been taken up in this chapter as well based on Bernt Θksendal, *"Stochastic Differential Equation with Application in Finance"*. Finally at the summit of the thesis, we used all these tools to model stock-prices, interest rates and pricing the European call options and the derivations of Black Scholes formula.This is also given in the differential equation form whose explicit solution is the Black Scholes formula. I used Karatzas-Shreve, *Methods of Mathematical Finance*, and several other references to write this part of the thesis and they are mentioned in the bibliography.

Chapter 2

PRELIMINARIES

To begin with, here are some necessary definitions and concepts to build up to the subject matter in these thesis.

2.1 Random Experiment

Definition 2.1.1. A mapping that assigns all elements in set D a unique element in C is a function, this we denote $f : D \to C$. D is the domain and C is co-domain of f. The subset of C which is mapped to elements in D is the range R which are images of D under f.

For a function f, the following may happen;

(i) **Injective** (one-to-one) if every element in D has a unique image in C.

(ii) **Surjective** (onto) if $C = R$.

(iii) **Bijective** if (i) and (ii) are both satisfied.

Remark 2.1.1. Condition (iii) is necessary for f to posses an inverse.

Definition 2.1.2. Assuming D is finite, the number of elements in D is its cardinality which we denote $|D|$. Two sets D and C are of equi-cardinal if there is a bijection from D to C. All finite or sets having a bijection with $\mathbf{N}$ (natural numbers) are countable. Sets having a bijection with $\mathbf{R}$ (real numbers), are uncountable they are termed continuum. Note there are other forms of uncountable sets.

Definition 2.1.3 (Random experiment). Is an experiment where certainty of the outcomes cannot be predicted: for example we cannot say with certainty that we are going to have a head as an outcome before tossing a fair coin. An outcome of this experiment is termed an event. The set where all events are contained is termed the sample space and we will denote this by Θ.

2.1.1 Sigma Fields

Definition 2.1.4 (field). A group $\mathcal{F}_0$ of subsets of Θ is termed a field if it fulfils the conditions below:

(i) $\varnothing \in \mathcal{F}_0$

(ii) $D \in \mathcal{F}_0$, then $\mathrm{D}^{\mathrm{c}} \in \mathcal{F}_0$,

(iii) $D_1, D_2, \ldots, D_n \in \mathcal{F}_0$, then $\bigcup_{i=1}^{n} D_i \in \mathcal{F}_0$.

From (i) we see that $\Theta \in \mathcal{F}_0$ since $\varnothing^c = \Theta$. Also from (ii) and (iii) and De-Morgan principle $\bigcap_{i=1}^{n} D_i \in \mathcal{F}_0$. Since an field excludes events generated by infinite unions and intersections, so it is not enough to describe a complete probability theory. To achieve this, we need a strong field.

Definition 2.1.5 (σ-field). A σ-field $\mathcal{F}$ is a field with the additional property,

$$D_1, D_2, D_3, \ldots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} D_i \in \mathcal{F}.$$

We can similarly show that for

$$D_1, D_2, D_3, \ldots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} D_i \in \mathcal{F}.$$

Given a σ-field $\mathcal{F}$, the sets found in $\mathcal{F}$ are referred to as $\mathcal{F}$-measurable.

Theorem 2.1.1. An intersection of a group of σ-fields of Θ, is also σ-field of Θ.

This can be easily shown by verifying the conditions needed for a σ-field .

Theorem 2.1.2 (Generated σ-field). Let $\mathcal{F}_0$ be an field , the least σ-field which contain $\mathcal{F}_0$ is a generated σ-field. We denote it as $\sigma(\mathcal{F}_0)$.

Note: The existence of $\sigma(\mathcal{F}_0)$ is by construction that is adding infinite unions and intersections. Define:

$$\sigma(\mathcal{F}_0) = \bigcap_{\alpha} \mathcal{F}_\alpha,$$

where $\{\mathcal{F}_\alpha, \alpha \in I\}$, hence $\sigma(\mathcal{F}_0)$, is the least σ-field containing $\mathcal{F}_0$.

If Θ is with a σ-field say $\mathcal{F}$, the pair $(\Theta, \mathcal{F})$ is termed measurable space. When Θ is countable, we take $\mathcal{F} = 2^\Theta$. When Θ is uncountable, 2^Θ becomes too large and thus we have to settle for a smaller σ-field , hence $\mathcal{F} \subset 2^\Theta$.

2.1.2 Measure Space

Definition 2.1.6. A measure $\vartheta : \mathcal{F} \to [0, \infty]$ on $\mathcal{F}$, satisfies;

1. $\vartheta(\varnothing) = 0$,
2. (Countable additivity) Let $D_1, D_2, D_3, \ldots$ is a family of countable and disjoint $\mathcal{F}-$measurable sets,

$$\vartheta\left(\bigcup_{i=1}^{\infty}\right) D_i = \sum_{i=1}^{\infty} \vartheta(D_i)$$

The triple $(\Theta, \mathcal{F}, \vartheta)$, is termed measure space. Since $\Theta \in \mathcal{F}$ so $\vartheta(\Theta)$ is well defined and therefore in $(\Theta, \mathcal{F}, \vartheta)$, following may happen.

1. $\vartheta(\Theta) < \infty$, ϑ is finite,
2. $\vartheta(\Theta) = \infty$, ϑ is infinite,

3. $\vartheta(\Theta) = 1$, ϑ is a probability.

Note that $\mathbf{P}$, denotes probability measures. Clearly $\mathbf{P}$ is finite, $(\Theta, \mathcal{F}, \mathbf{P})$, is termed probability space. We will be working with this kinds of spaces for the rest of the thesis.

Definition 2.1.7 (Borel Sets). Consider $(\Theta, \mathcal{F}, \mathbf{P})$ where $\Theta = \mathbf{R}$. Let $\mathcal{C}_0$ be a group of the open intervals of $\mathbf{R}$. Then $\sigma(\mathcal{C}_0)$ generated by $\mathcal{C}_0$ is termed Borel σ-field on $\mathbf{R}$, and this we denote as $\mathcal{B}(\mathbf{R})$.

All set in $\mathcal{B}(\mathbf{R})$ are Borel measurable or simply Borel sets. Since

$$\{b\} = \bigcap_{m=1}^{\infty} \left(b - \frac{1}{m}, b + \frac{1}{m} \right) \cap \Theta.$$

The singleton $\{b\}$ is Borel, therefore all intervals of $\mathbf{R}$, are Borel. $\mathcal{B}(\mathbf{R})$ is a proper subset of $2^{\mathbf{R}}$. So not all elements of $2^{\mathbf{R}}$ are Borel. Non Borel sets are very weird and rare hence they are only of academic interest. So the Borel σ-field contains all sets that are of probabilistic interest and even more such as the Cantor set. The Vitali sets is an example of a non Borel set.

Let $\Theta = (0,1)$, for any $(a,b) \subset \Theta$, define the length of (a,b) as $\mathbf{P}$, This is termed the Lebesque measure λ and it is a probability.

Definition 2.1.8. Given $(\Theta, \mathcal{F}, \mathbf{P})$, all sets in $\mathcal{F}$ are events, hence an event is any $\mathcal{F}$-measurable set.

2.2 Conditioning and Independence

Definition 2.2.1. Given $(\Theta, \mathcal{F}, \mathbf{P})$, let Γ, Λ be events and $\mathbf{P}(\Lambda) > 0$. Since $\Gamma, \Lambda \in \mathcal{F}$ then we define the probability of Γ given Λ as

$$\mathbf{P}(\Gamma|\Lambda) = \frac{\mathbf{P}(\Gamma \cap \Lambda)}{\mathbf{P}(\Lambda)}.$$

The equation above is termed, the conditional probabilityof Γ given Λ.

Remark 2.2.1. We cannot condition on an event of zero measure.

Theorem 2.2.1. If $\Lambda \in \mathcal{F}$ with $\mathbf{P}(\Lambda) > 0$, then $\mathbf{P}_\Lambda$ is a probability measure, where $\mathbf{P}_\Lambda(\cdot) = \mathbf{P}(\cdot|Y)$.

Proof is by checking all axioms of measure.

Theorem 2.2.2. Let Λ_k for $k = 1, 2, 3, \dots$ form a partition of Θ in $(\Theta, \mathcal{F}, \mathbf{P})$, that is, $\bigcup_{k \geq 1} \Lambda_k = \Theta$ and $\Lambda_k \cap \Lambda_i = \varnothing$, $\forall k \neq i$. Let $\Gamma \in \mathcal{F}$ and suppose $\forall k$, $\mathbf{P}(\Lambda_k) > 0$. Then,

$$\mathbf{P}(\Gamma) = \sum_{k \geq 1} \mathbf{P}(\Gamma|\Lambda_k)\mathbf{P}(\Lambda_i) \tag{2.1}$$

Proof. Consider

$$\begin{aligned} \sum_{k \geq 1} \mathbf{P}(\Gamma \cap \Lambda_k) &= \mathbf{P}\left(\bigcup_{k \geq 1} (\Gamma \cap \Lambda_k)\right) \\ &= \mathbf{P}\left(\Gamma \cap \left(\bigcup_{k \geq 1} \Lambda_k\right)\right) \\ &= \mathbf{P}(\Gamma \cap \Theta) = \mathbf{P}(\Gamma). \end{aligned}$$

This completes the proof. $\square$

In particular if $\Lambda \in \mathcal{F}$ and $0 < \mathbf{P}(\Lambda) < 1$. Then for any $\Gamma \in \mathcal{F}$,

$$\mathbf{P}(\Gamma) = \mathbf{P}(\Gamma|\Lambda)\mathbf{P}(\Lambda) + \mathbf{P}(\Gamma|\Lambda^c)\mathbf{P}(\Lambda^c).$$

This is an import of the above theorem, since Λ and Λ^c form a partition of Θ, hence

$$\mathbf{P}(\Gamma|\Lambda)\mathbf{P}(\Lambda) = \mathbf{P}(\Gamma \cap \Lambda).$$

Theorem 2.2.3 (Bayes Rule). Let $A \in \mathcal{F}$ with $\mathbf{P}(\Gamma) > 0$, and let Λ_k's be as in theorem 2.2.2. For $k \in \mathbf{N}$,

$$\mathbf{P}(\Lambda_k|\Gamma) = \frac{\mathbf{P}(\Gamma|\Lambda_k)\mathbf{P}(\Lambda_i)}{\sum\limits_{k\geq 1} \mathbf{P}(\Gamma|\Lambda_k) \cdot \mathbf{P}(\Lambda_k)}.$$

Proof. Notice that

$$\mathbf{P}(\Gamma|\Lambda_k)\mathbf{P}(\Lambda_k) = \mathbf{P}(\Gamma \cap \Lambda_k),$$

using the results of theorem 2.2.2,

$$\sum_{k\geq 1} \mathbf{P}(\Gamma|\Lambda_k)\mathbf{P}(\Lambda_k) = \mathbf{P}(\Gamma).$$

So Bayes Rule is a consequence of definition 2.2.1 and theorem 2.2.2. □

2.2.1 Independence

Definition 2.2.2. Given $(\Theta, \mathcal{F}, \mathbf{P})$, Let $D, \Lambda \in \mathcal{F}$. The events Γ and Λ are independent under $\mathbf{P}$ if

$$\mathbf{P}(\Gamma \cap \Lambda) = \mathbf{P}(\Gamma) \cdot \mathbf{P}(\Lambda).$$

If $\mathbf{P}(\Lambda) > 0$, we have $\mathbf{P}(\Gamma|\Lambda) = \mathbf{P}(\Gamma)$. If two events are independent, that does not necessarily mean they have nothing to do with each other. The solution of

$$\mathbf{P}(Z) = (\mathbf{P}(Z))^2,$$

is either $\mathbf{P}(Z) = 0$ or $\mathbf{P}(Z) = 1$. So self independent events are the impossible and the sure event. Z is not necessarily $\varnothing$ or Θ.

2.2.2 Independence of Sigma Fields

Definition 2.2.3. A family $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_n \in \mathcal{F}$, of events are independent if $I \subseteq \{1, 2, 3, \ldots, n\}$ and for all $I \neq \varnothing$, then

$$\mathbf{P}\left(\bigcap_{i \in I} \Gamma_i\right) = \prod_{i \in I} \mathbf{P}(\Gamma_i).$$

Definition 2.2.4. Let $\{\Gamma_i, i \in I\}$ be an arbitrary group of events; these are independent if for every finite group $\Gamma_{i_1}, \Gamma_{i_2}, \ldots, \Gamma_{i_n}$ of them,

$$\mathbf{P}\left(\bigcap_{1 \leq k \leq n} \Gamma_{i_k}\right) = \prod_{1 \leq k \leq n} \mathbf{P}(\Gamma_{i_k}).$$

Definition 2.2.5. Let $\mathcal{F}$ and $\mathcal{G}$ be σ-fields of Θ, They are independent if for every $\Gamma \in \mathcal{F}$ and $\Lambda \in \mathcal{G}$, Γ and Λ are independent.

2.3 Random Variables

Let $(\Theta, \mathcal{F}, \mathbf{P})$ be a probability space.

Definition 2.3.1. A mapping $Z : \Theta \to \mathbf{R}$ is $\mathcal{F}$-measurable when each $D \in \mathcal{B}(\mathbf{R})$, the image of the inverse $Z^{-1}(D)$ is an event in $\mathcal{F}$, where

$$Z^{-1}(D) = \{\omega \in \Theta : Z(\cdot) \in D\}.$$

Definition 2.3.2. A random variable Z is a $\mathcal{F}$-measurable function, where $Z : \Theta \to \mathbf{R}$; here the domain of Z, is Θ and the range is in $\mathbf{R}$.

Definition 2.3.3. Let Z be a random variable, the probability law $\mathbf{P}_Z : \mathcal{B}(\mathbf{R}) \to [0,1]$ of Z for all Borel sets is:

$$\mathbf{P}_Z(A) = \mathbf{P}(Z^{-1}(A)) = \mathbf{P}\left(\{\omega \in \Theta : Z(\cdot) \in A\}\right)$$

Remark 2.3.1. In other words the probability law is $\mathbf{P}_Z = \mathbf{P} \circ Z^{-1}$, which is the composite of $\mathbf{P}$ and Z^{-1}.

We generate $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \mathbf{P}_Z)$, as follows:

$$\sigma((-\infty, z] : z \in \mathbf{R}) = \mathcal{B}(\mathbf{R}).$$

So $\mathbf{P}_Z((-\infty, z])$ is well defined for every $z \in \mathbf{R}$ since $(-\infty, z] \in \mathcal{B}(\mathbf{R})$, then for $z \in \mathbf{R}$,

$$\mathbf{P}_Z((-\infty, z]) = \mathbf{P}(\{\omega \in \Theta | Z(\cdot) \leq z\}).$$

This is called the cumulative distribution functions for short CDF, and this we denote

$$F_Z(z) = \mathbf{P}(\{\omega \in \Theta | Z(\cdot) \leq z\}).$$

Of the 3 types of random variables, singular random variables are only of academic interest since they have no real applications. In this thesis we will deal only with discrete and continuous random variables.

2.3.1 Discrete and Continuous Random Variables

Definition 2.3.4. A random variable Z is discrete when its domain is countable. in other words Z assigns values in a countable set almost surely. $\mathbf{P}_Z = \mathbf{P}(Z = z)$, is termed probability mass function for short PMF of Z and it completely describes its probability.

Definition 2.3.5. Let ϑ_1 and ϑ_2 be measures defined on (Θ, $\mathcal{B}$), ϑ_1 is with respect to ϑ_2, absolutely continuous if for each $B \in \mathcal{B}$ with $\vartheta_1(B) = 0$, then $\vartheta_2(B) = 0$.

Definition 2.3.6. A random variable Z is continuous if with respect to λ (Lebesque), $\mathbf{P}_Z$ is absolutely continuous. In other words, if for each $B \in \mathcal{B}(\mathbf{R})$ of $\lambda(B) = 0$, then $\mathbf{P}_Z(B) = 0$.

Theorem 2.3.1 (Radon Nikodym special case). Let Z be a continuous random variable, there exist $f_Z : \mathbf{R} \to [0,\infty)$, a non negative measurable function where $\forall B \in \mathcal{B}(\mathbf{R})$ with

$$\mathbf{P}_Z(B) = \int_B f_Z(z)\,dz.$$

Putting $B = (-\infty, z]$, in the theorem above

$$\mathbf{P}_Z((-\infty,z]) = \mathrm{F}_Z(z) = \int_{-\infty}^{z} f_Z(z)\,dz, \quad \forall z \in \mathbf{R}.$$

Here $f_Z(z)$, is termed the probabilitydensity function for short PDF of Z.

2.3.2 Multi-Random Variables

Let Z, Y be random variable's on Θ where the point $(X(\cdot), Y(\cdot)) \in \mathbf{R}^2$, the Borel σ-field on $\mathbf{R}^2$, is

$$\mathcal{B}(\mathbf{R}^2) = \sigma((-\infty, z] \times (-\infty, y] : z, y \in \mathbf{R}).$$

So for $B \in \mathcal{B}(\mathbf{R}^2)$ we have,

$$\mathbf{P}_{ZY}(B) = \mathbf{P}(\{\omega \in \Theta : (Z(\cdot), Y(\cdot)) \in B\}).$$

This is termed joint probability law of the random variable's. In particular,

$$\mathbf{P}_{Z,Y}((-\infty, z] \times (-\infty, y]) = \mathbf{P}(\{\omega \in \Theta : Z \leq z\} \cap \{Y \leq y\}) = F_{ZY}(z, y).$$

This is the joint cumulative distribution function (CDF) of the random variables. In short-hand we write,

$$F_{ZY}(z, y) = \mathbf{P}(Z \leq z, Y \leq y).$$

When $F_{ZY}(z, y)$, is given, we can find the marginal CDFs, that is $F_Z(z)$ and $F_Y(y)$ but the converse is not necessarily true.

Definition 2.3.7. Let $\sigma(Y)$, $\sigma(Z)$ be independent σ-fields generated by Y, Z, then Y and Z are independent.

In particular if Y, Z are independent then their joint CDF factors to their individual CDFs. That is

$$F_{ZY}(z,y) = F_Z(z)F_Y(y).$$

Definition 2.3.8. Given Y and Z are discrete random variable's with PMF $\mathbf{P}_{YZ}(y,z)$ on $(\Theta, \mathcal{F}, \mathbf{P})$. We define the conditional PMF as,

$$\mathbf{P}_{Y|Z}(y,z) = \mathbf{P}(Y = y|Z = z) = \frac{\mathbf{P}_{YZ}(y,z)}{\mathbf{P}_Z(z)} \quad \text{where } \mathbf{P}_Z(z) > 0.$$

Definition 2.3.9. Given Y and Z are continuous random variable's in $(\Theta, \mathcal{F}, \mathbf{P})$, with joint PDF $f_{Y,Z}(y,z)$; then conditional PDF of Y given Z is defined as,

$$f_{Y|Z}(y,z) = \frac{f_{YZ}(y,z)}{f_Z(z)} \quad \text{where } f_Z(z) > 0.$$

Hence for an event $D \in Y$,

$$\mathbf{P}(D \in Y|Z = z) = \int\limits_D f_{Y|Z}(y|z)\,dy = \int_{-\infty}^{\infty} I_D(Y)\,\frac{f_{YZ}(y,z)}{f_Y(z)},$$

where $f_Z(z) > 0$, and I_D (indicator function) is

$$I_D(\omega) = \begin{cases} 1 & ,\omega \in D, \\ 0 & ,\omega \notin D. \end{cases}$$

2.3.3 The Gaussian

[1] A Gaussian or normal random variable X, has a distribution of the form

$$\mathbf{P}_X((a,b]) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx,\ -\infty < a < b < \infty,$$

where $m \in \mathbf{R}$ and $\sigma^2 > 0$ are parameters, this is denoted by $X \sim \mathcal{N}(m, \sigma^2)$. A constant random variable $X = m$, is Gaussian and is a degenerate. A Gaussian random variable X, belongs to $L_2(\Theta, \mathcal{F}, \mathbf{P})$ space and m, σ^2 denotes expectation and variance respectively, that is $\mathbf{E}[X] = m$ and $var(X) = \sigma^2$.

Theorem 2.3.2. If X, Y are independent and Gaussian with (m_1, σ_1^2), (m_2, σ_2^2) respectively: then

$$aX + bY \sim \mathcal{N}(am_1 + bm_2, a^2\sigma_1^2 + b^2\sigma_2^2),$$

where $a, b \in \mathbf{R}$.

This makes the Gaussian random variable very useful in many applications.

Definition 2.3.10. A family $\mathcal{N}(X_\alpha : \alpha \in A)$ of Gaussian random variables is a Gaussian system provided that any linear combinations of random variables from $\mathcal{N}$ are themselves Gaussian.

Theorem 2.3.3. Suppose $X, Z \in L_2(\Theta, \mathcal{F}, \mathbf{P})$, defines a Gaussian system, X and Z are independent iff $cov(X, Z) = 0$.

An example of a Gaussian system is, the Wiener process.

2.4 Conditional Expectation

Definition 2.4.1. Given Y and Z are discrete random variables with PMF $\mathbf{P}_{YZ}(y, z)$. Suppose the conditional PMF of Y, Z are,

$$\mathbf{P}_{Y|Z}(y|z) = \frac{\mathbf{P}_{Y,Z}(y, z)}{P_Z(z)} \quad \text{where } \mathbf{P}_Z(z) > 0.$$

We define,

$$\mathbf{E}(Y|Z=z) = \sum_{\Theta} y\mathbf{P}_{Y|Z}(y|z),$$

This depends on z, so $\sum\limits_{\Theta} y\mathbf{P}_{Y|Z}(y,z)$ is a mapping of z which we may denote $\Psi(z)$ and therefore

$$\Psi(z) = \mathbf{E}(Y|Z=z).$$

Note: $\Psi(Z)$ is a random variable and its termed conditional expectation of Y given Z.

Definition 2.4.2. Let Y, Z be jointly continuous with $f_{Z|Y}(z|y)$. We define,

$$\mathbf{E}(Y|Z=z) = \int_{\Theta} y f_{Y|Z}(y|z)dy.$$

Remark 2.4.1. Here $\mathbf{E}(Y|Z)$ is not a number it is a random variable.

In a more general situation we can show that for a measurable function g with $\mathbf{E}(|g|) < \infty$,

$$\mathbf{E}(Y - \Psi(Z)g(Z)) = 0. \tag{2.2}$$

When $g = 1$ we have the law of iterated expectation. From equation (2.2),

$$\mathbf{E}(Yg(\cdot)) = \mathbf{E}(\Psi(\cdot)g(\cdot)),$$

so $Y - \Psi(\cdot)$ and $g(\cdot)$ are uncorrelated. This equation forms the basic idea of conditional expectation in a general measure space. To prove equation (2.2) recall the Correlation coefficient

$$\rho_{y,z} = \frac{cov(Y,Z)}{\sigma_Z \sigma_Y}.$$

When both Y, Z are of 0 mean, $cov(Z,Y)$ forms an inner product where σ_Z and σ_Y play the role of the norms thus it forms a Hilbert space. Since $Y - \Psi(Z)$ and $g(Z)$ are uncorrelated so they are independent thus orthogonal on the Hilbert space. That is

$$\mathbf{E}[(Y - \Psi(Z))g(Z)] = 0.$$

Hence $\mathbf{E}(Y|Z)$ is an estimate of Y when Z is known and $Y - \mathbf{E}(Y|Z)$ is the Error of estimation and it is uncorrelated with any function of Z. Since $Y - \Psi(Z)$ and $g(Z)$ are orthogonal on the Hilbert space, this makes $\mathbf{E}[Y|Z]$ the "minimum mean-squared error estimator" in short (MMSE) of Y. Meaning for any measurable function φ,

$$E[(Y - E(Y|Z))^2] \leq \mathbf{E}[(Y - \gamma(Z)^2].$$

We can verify this by expanding the equation

$$\mathbf{E}[(Y - \Psi(Z) + \Psi(X) - \varphi(Z))^2].$$

If we consider

$$\inf_{z \in \sigma(Z)} \mathbf{E}[(Y - Z)^2],$$

this infimum exist since Hilbert spaces are complete so this is the conditional expectation in general a.s.

2.4.1 Conditioning on Event and Sigma Fields

Definition 2.4.3. Given $(\Theta, \mathcal{F}, \mathbf{P})$, let $B \in \mathcal{F}$ and Z be an integrable random variable with $\mathbf{P}(B) > 0$. Then

$$\mathbf{E}[Z|B] = \frac{1}{\mathbf{P}(B)} \int_A Z d\mathbf{P}.$$

is conditional expectation of Z given B.

For all events $A \in \sigma(Y)$,

$$\int_A \mathbf{E}[Z|Y] d\mathbf{P} = \int_A Z d\mathbf{P}.$$

Since $\mathbf{E}[Z|Y]$, depends only on $\sigma(Y)$ and not on the values of Z. So

$$\mathbf{E}[Z|\sigma(Y)] = \mathbf{E}[Z|Y].$$

Definition 2.4.4. Given $(\Theta, \mathcal{F}, \mathbf{P})$, let Z be an integrable random variable and $\mathcal{H}$ be a σ-field with $\mathcal{H} \subset \mathcal{F}$, then the random variable $\mathbf{E}[Z|\mathcal{H}]$ is conditional expectation of Z given $\mathcal{H}$ and $\forall A \in \mathcal{H}$,

$$\int_A \mathbf{E}[Z|\mathcal{H}] d\mathbf{P} = \int_A Z d\mathbf{P}.$$

Remark 2.4.2. Given an event A,

$$\mathbf{P}(A|\mathcal{H}) = \mathbf{E}[I_A(\cdot)|\mathcal{H}].$$

Theorem 2.4.1 (Law of iterated expectation). Let $\Psi(Z) = \mathbf{E}(Y|Z)$, then $\mathbf{E}(\Psi(Z)) =$

$\mathbf{E}(Y)$.

Proof. From

$$\mathbf{E}(\Psi(Z)) = \sum_{Z} \mathbf{P}_Z(z)\mathbf{E}(Y(\cdot)|Z(\cdot) = z).$$

Consider

$$\begin{aligned}
\sum_{Z} \mathbf{P}_Z(z) \sum_{Y} y \mathbf{P}_{\mathrm{y|z}}(y|z) &= \sum_{Z} \mathbf{P}_Z(z) \sum_{Y} y \frac{\mathbf{P}_{ZY}(z,y)}{\mathbf{P}_Z(z)}, \\
&= \sum_{ZY} y \mathbf{P}_{ZY}(z,y) = \mathbf{E}(Y).
\end{aligned}$$

This completes the proof [7] [2]. $\square$

Chapter 3

STOCHASTIC PROCESSES AND BROWNIAN MOTION

3.1 Filtrations and Martingales

Definition 3.1.1. Given $(\Theta, \mathcal{F}, \mathbf{P})$, a group $\{\mathcal{F}_\mu\}$ of sub-σ-fields is a filtration of $\mathcal{F}$, if it satisfies

$$\mathcal{F}_r \subset \mathcal{F}_\mu, 0 \leq r \leq t.$$

A Filtration $\{\mathcal{F}_\mu\}$ is complete if each event of zero probability is in $\mathcal{F}_0$. Since filtrations are increasing sequences of σ-fields, then

$$\bigcup_{r<\mu} \mathcal{F}_r \subset \mathcal{F}_\mu \subset \bigcap_{r>\mu} \mathcal{F}_r.$$

For all $\mu \geq 0$, we denote

$$\mathcal{F}_\mu^- = \bigcap_{r<\mu} \mathcal{F}_r, \quad \text{and} \quad \mathcal{F}_0^- = \mathcal{F}_0.$$

Similarly

$$\mathcal{F}_\mu^+ = \bigcap_{r>\mu} \mathcal{F}_r.$$

So $\{\mathcal{F}_\mu\}$ is;

1. Left continuous if $\mathcal{F}_\mu = \mathcal{F}_\mu^-$.
2. Right continuous if $\mathcal{F}_\mu = \mathcal{F}_\mu^+$.

3. Continuous if $\mathcal{F}_\mu = \mathcal{F}_\mu^+ = \mathcal{F}_\mu^-$.

The natural filtration for some random process Z, is

$$\mathcal{F}_\mu^Z = \sigma(Z_\mu; 0 \leq \mu \leq T),$$

where $T \geq 0$. If Z is $\mathcal{F}_\mu-$measurable for all $\mu \geq 0$, then Z is $\mathcal{F}_\mu$-adapted. Clearly Z is $\mathcal{F}_\mu^Z-$adapted.

Definition 3.1.2. The process $Z_\mu, \mu \geq 0$, is termed with respect to the filtration $\mathcal{F}_\mu$, a continuous time martingale if

1. $\mathbf{E}[Z_\mu] < \infty, \forall \mu \geq 0$,
2. Z is $\mathcal{F}_\mu^Z-$adapted,
3. $\mathbf{E}[Z_\mu|\mathcal{F}_s] = Z_s$ i.e Z_s is best estimation of Z_μ knowing $\mathcal{F}_s$.

For $k = 0, 1, \ldots$, with respect to $\{\mathcal{F}_k\}$, $\{Z_k\}$ is a discrete martingale, if

1. $\mathbf{E}[Z_k] < \infty$,
2. Z_k is $\mathcal{F}_k-$adapted,
3. $\mathbf{E}[Z_{k+1}|\mathcal{F}_k] = Z_k$.

A constant expectation function, is a remarkable property for martingales. I.e, for $r < \mu$ since $\mathbf{E}[Z_\mu|\mathcal{F}_r] = Z_r$, we obtain

$$\mathbf{E}[Z_r] = \mathbf{E}(\mathbf{E}[Z_\mu|\mathcal{F}_r]) = \mathbf{E}[Z_\mu], 0 \leq r \leq \mu.$$

3.2 Stochastic Processes

Definition 3.2.1. Given $(\Theta, \mathcal{F}, \mathbf{P})$, a group $\{Z_\mu\}_{\mu \geq 0}$ of parametrized random variables is termed stochastic process. For each $\omega \in \Theta$ and $\mu \in [0, \infty)$, Z_μ is a random variable and $Z_\mu(\omega)$ is a path.

A process Z_μ is a Gaussian if Z_μ, $\mu \in [0, T]$ forms a Gaussian system. The process Z_μ is termed a second order process if $\mathbf{E}[Z_\mu^2] < \infty$, $\forall \mu \in [0, T]$.

Definition 3.2.2 (Stationary increments). A random process Z, have stationery increments of wide sense provided for each μ, s, where $\mu, \mu + s \in [0, T]$, we have

1. $\mathbf{E}[Z_{\mu+s} - Z_s] = 0$,
2. $\mathbf{E}[Z_{\mu+s} - Z_s]^2 = \mathbf{E}[Z_\mu - X_0]^2$.

Definition 3.2.3 (independent increments). A random process Z have independent increments, provided we have for each finite number of $\mu_0, \dots, \mu_k \in [0, T]$ satisfying $\mu_0 < \cdots < \mu_k$. Then $Z_{\mu_1} - Z_{\mu_0}, \dots, Z_{\mu_k} - Z_{\mu_{k-1}}$ become independent random variable's.

3.3 Brownian Motion and Wiener Process

In 1827, R. Brown watched the development of dust grains as they connect with the atoms of fluids under a magnifying instrument. This development was weird and exceptionally unpredictable thus it was named Brownian motion. In 1905 physical thinking for Brownian motion was given by A. Einstain as a movement of a tiny molecule dictated by its impact with fluid atoms.

To decide the Brownian path of a molecule, we need full data of the atoms of a given fluid at some fixed moment, their position, speed, and so on. This is certifiably not a deterministic methodology. However, a stochastic method was recommended by N. Wiener to deal with any motion of Brownian nature.

Definition 3.3.1 (Wiener process). [6] Consider a microscopic particle on a liquid and let Θ be the group of all sample cases leading to unique paths of the particle. Then its position can be modelled by $W_{\mu,\omega}, \mu \geq 0$, $\omega \in \Theta$, a function of two variables. Since the particle has continuous paths , one can select $\Theta = C(0,\infty;\mathbf{R})$ and presume that the chosen path $\omega \in \Theta$, leads to $W_{\mu,\omega} = \omega_\mu$, $\mu \geq 0$. Clearly, different family of paths can exist at different possibilities. Hence, we may presume a probabilitydistribution $\mathbf{P}_A$ on Θ is specifically defined and thus $\mathbf{P}_A(B)$, manifests the probabilityof the existence of the particle's followed path in $B \subset \Theta$. Hence from this the following import cab be made, there is a Θ with a distribution on Θ, for which the particle's motion is a random process $W : [0,\infty) \times \Theta \to \mathbf{R}$.

To further understand the properties of this process, we need a microscopic particle on a liquid with the conditions:

1. Homogeneous with a fixed viscosity,
2. No outside force interference.

Lemma 3.3.1. Let the random process Z be stationery in wide sense and has independent increments. If $\psi(\mu) = \mathbf{E}[Z_\mu - Z_0]^2, \mu \in [0,T]$, is continuous then $\psi(\mu) = \lambda\mu$ for some constant $\lambda \geq 0$.

Proof. Consider

$$\psi(\mu) - \psi(r) = \mathbf{E}[Z_\mu - Z_0]^2 - \mathbf{E}[Z_r - Z_0]^2$$

$$
\begin{aligned}
&= \mathbf{E}[Z_\mu - Z_r + Z_r - Z_0]^2 - \mathbf{E}[Z_r - Z_0]^2 \\
&= \mathbf{E}[Z_\mu - Z_r]^2 + \mathbf{E}[(Z_\mu - Z_r)(Z_r - Z_0)] \\
&= \mathbf{E}[Z_\mu - Z_r]^2 \geq 0.
\end{aligned}
$$

We have used the fact that $Z_\mu - Z_r$ and $Z_r - Z_0$ are independent random variables so they are uncorrelated thus

$$\mathbf{E}[(Z_\mu - Z_r)(Z_r - Z_0)] = 0.$$

Since $\psi(\mu)$ is a non-decreasing function; substituting $\mu = r+s$ we have

$$
\begin{aligned}
\psi(r+s) &= \psi(r) + \mathbf{E}[Z_{r+s} - Z_r]^2 \\
&= \psi(r) + \mathbf{E}[Z_s - Z_0]^2 \\
&= \psi(r) + \psi(s),\ \psi(0) = 0.
\end{aligned}
$$

So a unique continuous and non-decreasing solution for ψ is a solution of Cauchy's functional equation. Therefore $\psi(\mu) = \lambda\mu$, where λ is a constant. □

Thus as a result, for $\lambda \geq 0$, we have

$$\mathbf{E}[W_\mu - W_r]^2 = \lambda(\mu - r),\ 0 \leq r \leq \mu \leq T.$$

When $\lambda = 0$, then W_μ is constant. This is when there is so much cohesion in the liquid's molecules that it behaves like a solid. Hence we can always assume $\lambda > 0$.

Here are the fantastic properties of W:

1. As a consequence of condition (i) and (ii) W has stationary increments in the

wide sense. So we have $\mathbf{E}(W_r - W_s) = 0$, $\mathbf{E}(W_r - W_s)^2$ depends only on $|r-s|$.

2. W has independent increments, this is because the liquid contains a lot of molecules and so in different intervals $(\beta,\mu]$, $(r,\sigma]$, the particle comes into contact with different molecules. Hence, increment $W_\mu - W_\beta$, $W_r - W_\sigma$ are formed by independent collisions.

3. $W_\mu - W_\beta$ is a Gaussian random variable, since $W_\mu - W_\beta$, can be considered as a sum of small increments which are independent and equidistributed. From lemma (3.3.1), we have $\mathbf{E}(W_\mu - W_\beta) = 0$ and $\mathbf{E}(W_\mu - W_\beta)^2 = \lambda(\mu-\beta)$ so $W_\mu - W_\beta \sim \mathcal{N}(0, \lambda(\mu-\beta))$.

4. W has continuous paths.

5. At $\mu = 0$ the position of the particle is selected to be the origin $W_0 = 0$ and the viscosity of the liquid is such that $\lambda = 1$. These two are just normalising conditions.

A process $W : [0,\infty) \times \Theta \to \mathbf{R}$ satisfying all the above properties is termed a standard Wiener process, the existence of such W was given by N. Wiener in 1923.

Theorem 3.3.1 (Wiener). Given Borel σ-field of $C(0,\infty);\mathbf{R}$. There is a precise probability measure $\mathbf{P}_B$, where the coodinate proceşs $W_{\mu,\omega} = \omega_\mu$, $\omega \in C(0,\infty);\mathbf{R}$, with respect to $\mathbf{P}_B$, is a standard Wiener process with probability 1.

The paths of a Wiener process have the following interesting properties:

1. They are not monotone.
2. They are continuous.
3. They are nowhere differentiable.

4. They have infinite variation even on the smallest possible interandom variableal, that is across every partitions $\sigma : 0 = s_0 < \cdots < s_k = S$ of $[0,S]$. we have

$$\sup_{\sigma} \sum_{1 \leq i \leq k} |W_{\mu_i,\omega} - W_{\mu_{i-1},\omega}| = \infty.$$

Property 3 and 4, seems to be against the genuine nature of Brownian motion, since a microscopic particles have limited speed at every instance and goes over a limited distance for all limited intervals. However the Wiener process is known to be a decent estimate of the Brownian motion. with respect to its natural filtration, the Wiener process is martingale and some of its transformations are themselves martingales.

Theorem 3.3.2 (Levy's). Suppose $\{W_\mu\}_{\mu \geq 1}$ is a stochastic process where W_μ is $\mathcal{F}_\mu$-adapted and martingale a.s with respect to $\mathcal{F}_\mu$, if $[W,W](\mu) = \mu$, then $\{W_\mu\}_{\mu \geq 1}$ is Brownian [1] [6] [7].

Chapter 4

ITō CALCULUS

The failure of ordinary integration methods with regards to the paths described in Wiener's theorem, is an import of their nowhere differentiability and their unbounded variation and hence the need for stochastic calculus with a totally different approach.

4.1 Riemann Integral

Given a function f on $[0,T]$ where it is real-valued and

$$\sigma_m : 0 = k_0 < s_1 < \cdots < k_{m-1} < k_m = T,$$

a partition of $[0,T]$: and for $k = 1, \ldots, m$, define

$$\Delta_k = \mu_k - \mu_{k-1}.$$

An intermediate partition φ_m of σ_m is given by y_k where $\mu_{k-1} \leq y_k \leq \mu_k$, gives. now for Riemann integral:

$$S_m = \sum_{1 \leq k \leq m} f(y_k)(\mu_k - \mu_{k-1}) = \sum_{1 \leq k \leq m} f(y_k)\Delta_k.$$

Here, S_m is an estimate of the area under the curve f, where f takes only non–ve numbers. Let

$$\text{mesh}(\varphi_m) = \max_{k=1,\ldots,m} \Delta_k \to 0.$$

As

$$\max_{j=k,\ldots,m} \Delta_k \to 0,$$

if the limit

$$S = \lim_{m\to\infty} S_m,$$

exist and S is not dependent on choices of φ_m and σ_m, thus S is termed the Riemann integral of f in $[0,T]$. We write

$$S = \int_0^T f(\mu)d\mu.$$

Note: $\int_0^T f(\mu)d\mu$ is known to exist for f continuous or piecewise continuous.

4.2 Riemann Stieltjes Integral

Consider partitions φ_m and σ_m similar to section 4.1, on $[0,T]$. Given f,h on $[0,T]$ and are real-valued, for $i = 1,\ldots,m$, define

$$\Delta_k h = h(\mu_k) - h(\mu_{k-1}).$$

Corresponding to σ_m and φ_m, we obtain

$$S_m = \sum_{1\le k\le m} f(y_k)[h(\mu_k) - h(\mu_{k-1})] = \sum_{1\le k\le m} f(y_k)\Delta_k h,$$

which is referred to as Riemann-Stieltjes sum. We can see that when $h(\nu) = \nu$ we have a Riemann sum. Thus weighing $f(y_k)$, with $\Delta_k h$, in $[\mu_{k-1},\mu_k]$ of h, gives The Riemann-Stieltjes sum.

As mesh $(\sigma_m) \to 0$, if the limit

$$S = \lim_{m \to \infty} S_m$$

exist and S is not dependent on choices of σ_m and φ_m, then with respect to g on $[0,T]$, S is termed, Riemann-Stieltjes integral of f. i.e

$$S = \int_0^T f(\mu) dh(\mu).$$

If h has bounded-variation, that is

$$\sup_{\sigma} \sum_{1 \le k \le m} |h(\mu_k) - h(\mu_{k-1})| < \infty,$$

and f is continuous where supremum is across all partitions σ of $[0,T]$, then $\int_0^T f(\mu) dh(\mu)$ is known to exist.

Since the Wiener process W_t has an unbounded variation so the above result is not applicable. Nonetheless, for the existence of $\int_0^T f(\mu) dg(\mu)$, it is not necessary for g to be of bounded variation. Finding relaxed conditions for this existence is an open question. However, we can state these:

Definition 4.2.1. A function h on $[0,T]$ with

$$\sup_{\sigma} \sum_{k=1}^{m} |h(\mu_k) - h(\mu_{k-1})|^q < \infty,$$

for some $q > 0$, is $q-$variation bounded. Here h is real-valued and supremum is across every partition σ of $[0,T]$.

So this holds for h when $q = 1$. If f and g agrees with the conditions below

1. All discontinuities are at different points,
2. For some $p, q > 0$ and $p^{-1} + q^{-1} > 1$, they are bounded $p, q-$variation respectively.

Hence $\int_0^T f(\mu) dh(\mu)$ exists. For a particular interval, a Brownian path W_μ, is of $q-$variation with the provision that $q > 2$, and for $q \leq 2$, its unbounded.

Consider a sample path $f(\mu, \cdot)$, where f is a stochastic process . With respect to the sample path W_μ, we can define

$$\int_0^T f(\mu) dW_\mu,$$

provided W_μ is of $q-$variation with $q < 2$.

Given f is differentiable, let $f'(\mu)$ be bounded then by the mean-value theorem it follows:

$$|f(\mu) - f(\nu)| \leq \Lambda(\mu - \nu), \nu < \mu,$$

where $\Lambda > 0$, is some constant. So we have

$$\sup_\sigma \sum_{j=1}^{k} |f(\mu_j) - f(\mu_{j-1})| \leq \Lambda \sum_{j=1}^{k} |\mu_j - \mu_{j-1}| = \Lambda < \infty,$$

hence bounded variation holds for f. As and import the following statement also holds for f on $[0, T]$;

$$\int_0^T f(\mu)dW_\mu,$$

exist for every Brownian sample path W_μ.

Since Wiener processes are nowhere differentiable so the above results are not applicable so the integral

$$\int_0^T W_\mu dW_\mu, \tag{4.1}$$

does not exist. So with respect to a Brownian motion, path-wise integration does not give a vast group of integrable functions, hence the need for an integral which is not define based on a path. The mean-square limit approach is used define Itō stochastic integral.

4.3 Itō Stochastic Integral

Definition 4.3.1. The stochastic process $\{C(\mu)\}_{\mu=0}^T$ is simple if it agrees with the condition below: there exist a partition φ_m and a family $\{Z_k\}_{k=1}^n$ of random variables with

$$C(\mu) = \begin{cases} Z_n, & \mu = T, \\ \\ Z_k, & \mu \in [\mu_{k-1}, \mu_k), \end{cases}$$

where

$$\varphi_m : 0 = t_0 < t_1 < \cdots < tn - 1 < t_n = T,$$

here $\{Z_k\}$ is $\{\mathcal{F}_\mu$-adapted and agrees with $\mathbf{E}[Z_k^2] < \infty$ for all $k = 1, \ldots, n$.

Let $\{W_\mu\}_{k=1}^{T}$ be a Wiener process adapted to its natural filtration

$$\mathcal{F}_\mu = \sigma(W_s, s < \mu), \mu > 0.$$

For a simple process $C(\mu)$ on $[0,T]$, define

$$\int_0^T C(\mu) dW_\mu = \sum_{1 \le k \le n} C(\mu_{k-1})(W_{\mu_k} - W_{\mu_{k-1}}) = \sum_{j=1}^{k} Z_k \Delta_k W, \tag{4.2}$$

is Itō integral. So clearly given simple process $C(\mu)$, then expression in 4.2 has a representation of the Riemann-Steiltjes sum.

To generalise the Itō integral, we assume $C(t)$ agrees with the following additional conditions:

1. $C(\mu)$ is W_μ-adapted and its a function of $W_s, 0 \le s \le \mu$. where W_μ is Brownian.
2. $\int_0^T \mathbf{E}[C(\mu)^2] d\mu < \infty$.

Then we can find $\{C^{(n)}\}$, a family of simple processes with

$$\int_0^T \mathbf{E}[C(\mu) - C^{(n)(\mu)}]^2 d\mu \to 0.$$

Now we define

$$I\left(C^{(n)(\mu)}\right) = \int_0^T C^{(n)}(\mu) dW_\mu.$$

One can prove that this limit of a family of random variable's exist, which gives the Itō stochastic integral of $C(\mu)$, we write this as

$$I(C) = \int_0^T C(\mu) dW_\mu, \, \mu \in [0,T].$$

4.3.1 Properties Of Itō Integral

1. The process $I_\mu(C)$ has continuous paths.
2. $I_\mu(C)$ is $\{\mathcal{F}_\mu\}$-measurable and has zero expectation.
3. with respect to $\{\mathcal{F}_\mu\}, \mu \in [0,T]$, the natural Brownian filtration, $I_\mu(C)$ is martingale.
4. $\int_0^T [aC_1(\mu) + bC_2(\mu)] \, dW_\mu = \int_0^T aC_1(\mu) dW_\mu + \int_0^T bC_2(\mu) dW_\mu$. Where a, b are constants which is the linearity of $I_\mu(C)$.
5. $\int_0^T C(\mu) dW_\mu = \int_0^T C(\mu) dW_\mu + \int_\mu^T C(\mu) dW_\mu$. Linearity on adjacent intervals.
6. $\mathbf{E}[I^2(C)] = \int_0^T C^2(\mu) d\mu$, Isometry property.

4.4 Itō Formula

The chain rule from classical calculus is

$$[\Gamma(\Lambda(\mu))]' = \Gamma'(\Lambda(\mu))\Lambda'(\mu).$$

If Γ and Λ are differentiable functions. In integral form

$$\Gamma(\Lambda(\mu)) - \Gamma(\Lambda(0)) = \int_0^t \Gamma'(\Lambda(\mu))\Lambda'(\mu) d\mu = \int_0^t \Gamma'(\Lambda(\mu)) d\Lambda(\mu).$$

The Itō formula's import is, extending the chain rule to stochastic differentials. Given $\Gamma(\mu, X_\mu)$, assume the differentials $\Gamma_\mu(\mu, X_\mu)$, $\Gamma_\nu(\mu, X_\mu)$ and $\Gamma_{\nu\nu}(\mu, X_\mu)$ exist and are continuous. If W_μ, is Wiener process, then

$$\Gamma(T, W_T) - \Gamma(0, W_0) = \int_0^T \Gamma_\mu(\mu, W_\mu) dt + \int_0^T \Gamma_\nu(\mu, W_\mu) +$$

$$\frac{1}{2}\int_0^T \Gamma_{\nu\nu}(\mu, W_\mu)dW_\mu dW_\mu.$$

We know $dW_\mu dW_\mu = d\mu$, so we obtain

$$\Gamma(T, W_T) - \Gamma(0, W_0) = \int_0^T \Gamma_\mu(\mu, W_\mu)dt + \int_0^T \Gamma_\nu(\mu, W_\mu) + \frac{1}{2}\int_0^T \Gamma_{\nu\nu}(\mu, W_\mu)d\mu.$$

In differential form, the Itō formula is

$$d\Gamma(\mu, W_\mu) = \Gamma_\mu(\mu, W_\mu)d\mu + \Gamma_\nu(\mu, W_\mu)dW_\mu + \frac{1}{2}\Gamma_\mu(\mu, W_\mu)d\mu.$$

Theorem 4.4.1. Given $\Gamma(x)$ is twice differentiable and W_t is a Wiener process. Applying Itō formula

$$\Gamma(W_t) - \Gamma(W_s) = \int_s^t \Gamma'(W_z)dW_z + \frac{1}{2}\int_s^t \Gamma''(W_z)dz.$$

As an import of the theorem above, we can find

$$\int_0^t W_z dW_z.$$

For this choose $\Gamma(\mu) = \mu^2$, so that $\Gamma'(\mu) = 2\mu$ and $\Gamma''(\mu) = 2$. Hence by the above theorem,

$$W_t^2 - W_s^2 = 2\int_s^t W_z dW_z + \int_s^t dz.$$

Putting $s = 0$,

$$\int_0^t W_z dW_z = \frac{1}{2}(W_t^2 - t).$$

So the integral in (3.1), exist in the Itō sense.

Using integration by parts from ordinary calculus, the Riemann Stieltjes integral yields

$$\int_0^t \Lambda(\nu)d\Lambda(\nu) = \int_0^t \Lambda(\nu)\Lambda'(\nu)d\nu = \frac{1}{2}\Lambda^2(t).$$

We have already seen in the Itō integral, we have an additional term $\frac{1}{2}t$, which comes from the quadratic variation of the Wiener process W_t.

4.4.1 Itō Process

Definition 4.4.1. Defined the process;

$$X_s = X_0 + \int_0^s \Delta(\mu)dW_\mu + \int_0^s \varphi(\mu)d\mu$$

Here $\Delta(s)$, $\varphi(s)$ are $\mathcal{F}_\mu$-adapted stochastic processes associated to $\{W_\mu\}$. In short-hand we write this as a stochastic differential equation (SDE).

$$dX_\mu = \Delta(\mu)dW_\mu + \varphi(\mu)d\mu.$$

This SDE is sometimes called Itō's lemma, and it has the following quadratic variation

$$dX_\mu dX_\mu = [\Delta(\mu)dX_\mu + \varphi(\mu)d\mu]^2 = \Delta^2(\mu)d\mu,$$

thus

$$[X,X](s) = \int_0^s \Delta^2(\mu)d\mu.$$

This is nothing but the isometry of the Itō integral $I(t)$. Under normal conditions we assume

$$\mathbf{E}\left[\int_0^T \Delta^2(\mu)d\mu\right] < +\infty,$$

the variance of $I(\mu)$ is

$$Var(I(\mu)) = \mathbf{E}[I^2(\mu)] - \mathbf{E}[I(\mu)].$$

Since $I(\mu)$ is a martingale, $\mathbf{E}[I(\mu)] = 0$; hence

$$Var(I(\mu)) = \mathbf{E}[I^2(\mu)] = \int_0^T \Delta^2(\mu)d\mu.$$

Let $\Gamma(X_\mu, Y_\mu) = X_\mu, Y_\mu$ the product rule from Itō's formula is

$$d\Gamma(X_\mu, Y_\mu) = Y_\mu dX_\mu + X_\mu dY_\mu + dX_\mu dY_\mu.$$

If X_μ, Y_μ is not a Wiener process and is a function of classical calculus, the cross quadratic variation $dX_\mu dY_\mu$ becomes zero. So in classical calculus the quadratic variation $dX_\mu dY_\mu$ does not appear.

4.4.2 Itō Formula for the Itō Process

Given $\Gamma(\mu, X_\mu)$, applying Itō formula to get

$$\begin{aligned} d\Gamma(\mu, X_\mu) &= \Gamma_\mu(\mu, X_\mu)d\mu + \Gamma_\nu(\mu, X_\mu)dX_\mu + \frac{1}{2}\Gamma_{\nu\nu}(\mu, X_\mu)dX_\mu dX_\mu, \\ &= \Gamma_\mu(\mu, X_\mu)d\mu + \Gamma_\nu(\mu, X_\mu)dX_\mu + \frac{1}{2}\Gamma_{\nu\nu}(\mu, X_\mu)\Delta^2(t)d\mu, \end{aligned} \tag{4.3}$$

Substituting the SDE

$$dX_\mu = \Delta(\mu)dW_\mu + \varphi(\mu)d\mu,$$

expression 4.3, becomes

$$d\Gamma(\mu, X_\mu) = \Gamma_\mu(\mu, X_\mu)d\mu + \Gamma_\nu(\mu, X_\mu)[\Delta(\mu)dW_\mu + \varphi(\mu)d\mu] + \frac{1}{2}\Gamma_{\nu\nu}(\mu, X_\mu)dX_\mu dX_\mu$$
$$= [\Gamma_\mu + \varphi(\mu)\Gamma_\nu + \frac{1}{2}\Delta^2(\mu)\Gamma_{\nu\nu}]d\mu + \Delta(\mu)\Gamma_\nu dW_\mu.$$

This SDE is the Itō formula for $\{X_\mu\}$.

4.4.3 Itō Formula for Higher Dimensions

Let

$$\mathbf{W}_t = W_{1t}, W_{2t}, \ldots, W_{nt}$$

be a Brownian vector where $\{W_{kt}\}$ is Brownian $\forall k,\ 1 \leq k \leq n$. Then $\mathbf{W}_t$ has the following quadratic variation.

$$d\mathbf{W}_{kt}d\mathbf{W}_{jt} = \begin{cases} dt & k = j, \\ 0 & k \neq j. \end{cases}$$

Consider the case where $n = 2$:

$$X_\mu = X_0 + \int_0^\mu \varphi_1(s)ds + \int_0^\mu \sigma_{11}(s)dW_1(s) + \int_0^\mu \sigma_{12}(s)dW_2(s).$$
$$Y_\mu = Y_0 + \int_0^\mu \varphi_2(s)ds + \int_0^\mu \sigma_{21}(s)dW_1(s) + \int_0^\mu \sigma_{22}(s)dW_2(s).$$

These give the following SDEs:

$$dX_\mu = \varphi_1(\mu)d\mu + \sigma_{11}(\mu)dW_1(\mu) + \sigma_{12}(\mu)dW_2(\mu).$$

$$dY_\mu = \varphi_2(\mu)dt + \sigma_{21}(\mu)dW_1(\mu) + \sigma_{22}(\mu)dW_2(\mu).$$

From these, we obtain the quadratic variations below.

$$[X,X](\mu) = \int_0^{\mu} (\sigma_{11}^2(s) + \sigma_{12}^2(s))ds \tag{4.4}$$

$$[Y,Y](\mu) = \int_0^{\mu} (\sigma_{21}^2(s) + \sigma_{22}^2(s))ds \tag{4.5}$$

$$[X,Y](\mu) = \int_0^{\mu} (\sigma_{11}(s)\sigma_{21}(s) + \sigma_{12}(s)\sigma_{22}(s))ds \tag{4.6}$$

For 2 dimensions, let $\Gamma(\mu, X_\mu, Y_\mu)$ be such that Γ_μ, Γ_ν, Γ_τ, $\Gamma_{\nu\nu}$, $\Gamma_{\nu\tau}$ and $\Gamma_{\tau\tau}$ exist and are continuous, then we have;

$$d\Gamma(\mu, X_\mu, Y_\mu) = \Gamma_\mu(\mu, X_\mu, Y_\mu)d\mu + \Gamma_{nu}(\mu, X_\mu, Y_\mu)dX_\mu + \Gamma_\tau(\mu, X_\mu, Y_\mu)dY_\mu +$$
$$\frac{1}{2}\Gamma_{\nu\nu}(\mu, X_\mu, Y_\mu)dX_\mu dX_\mu + \Gamma_{\nu\tau}(\mu, X_\mu, Y_\mu)dX_\mu dY_\mu + \frac{1}{2}\Gamma_{\tau\tau}(\mu, X_\mu, Y_\mu)dY_\mu dY_\mu.$$

We can now use this formula and the quadratic variations in (3.1), (3.2) and (3.3) with their SDEs to find a formula for the Itō process in 2 dimensions.

Remark 4.4.1. So in essence the quadratic variation $[W,W](t)$ determines if a stochastic process is Brownian [5] [7] [8].

4.5 Stochastic Differential Equations

Given W_μ, is a Wiener process; A stochastic differential equation (SDE) is an equation of the form

$$dX_\mu = \alpha(\mu, X_\mu)d\mu + \sigma(\mu, X_\mu)dW_\mu, \tag{4.7}$$

Our goal here is to find a stochastic process X_μ so that, X_μ agrees with the SDE in (4.7), that is

$$X_\mu = X_0 + \int_0^{\mu} \alpha(s, X_s)dt + \int_0^{\mu} \sigma(s, X_s)dW_\mu.$$

Here X_μ is solution to (4.7), where $\alpha(\mu, X_\mu), \sigma(\mu, X_\mu) \in \mathbf{R}$ and $X_\mu(0) = X_0$.

4.6 Existence and Uniqueness

Given the SDE

$$dX_\mu = \alpha(\mu, X_\mu)d\mu + \sigma(\mu, X_\mu)dW_\mu,$$

where $\varphi(,) : [0,S] \times \mathbf{R}^k \to \mathbf{R}^k$ and $\sigma(,) : [0,S] \times \mathbf{R}^k \to \mathbf{R}^{k\times m}$ are measurable for $S \geq 0$, satisfying:

1. $|\varphi(r,u)| + |\sigma(r,u)| \leq \lambda(1+|u|)$ and $|\sigma|^2 = \Sigma|\sigma_{ij}|^2$.
2. $|\varphi(r,u) - \varphi(r,v)| + |\sigma(r,u) - \sigma(r,v)| \leq \Upsilon|u-v|$.

Where λ, Υ are constants, while $u, v \in \mathbf{R}^k$ and $t \in [0,T]$.

Let Z be an $\mathcal{F}_\infty^{(m)}$ independent random variable with $\mathbf{E}[|Z|^2] < \infty$ generated by the Wiener process $W_s(\cdot)$. Then (4.7), possesses unique and continuous solution $X_t(\cdot)$, for $s \leq t$ and $X_0 = Z$, where

$$X_\mu = X_0 + \int_0^\mu \varphi(r, X_r)d\mu + \int_0^\mu \sigma(r, X_r)dW_\mu.$$

The process X_μ is $\mathcal{F}_s^Z$- adapted where $W_r(\cdot); r \leq s$ and

$$E\left[\int_0^\mu |X_s|^2\right] < \infty.$$

4.7 Solutions to Special Cases

Finding a solution to stochastic differential equations depends mostly on the application of the Itō formula and our basic knowledge of ordinary differential equations. These solutions are sometimes called closed form solution, not all SDEs have a closed form solution, Here are examples.

4.7.1 Geometric Brownian

This has the following stochastic differential equation;

$$X_t = \mu X_t dt + \sigma X_t dB_t.$$

Let $X_t = f(t, B_t)$, so that

$$dX_t = df(t, B_t) = (f_t + \frac{1}{2} f_{xx})dt + f_x dB_t.$$

By comparison we have

$$f_t + \frac{1}{2} f_{xx} = \mu X_t = \mu f \quad \text{and}$$

$$f_x = \sigma f \implies f(t, B_t) = e^{\sigma x + g(t)}$$

From the latter form, we see that $f_t = g'(t) f$ and $f_{xx} = \sigma^2 f$, so we have

$$\gamma'(\cdot) = \mu - \frac{1}{2}\sigma^2 \implies \gamma(\cdot) = (\mu - \frac{1}{2}\sigma^2)t + k,$$

so that

$$f(t, B_t) = e^{\sigma x + (\mu - \frac{1}{2}\sigma^2)t + k},$$

where $f(0,0) = e^k = X_0$. So we have

$$X_t = X_0 e^{\sigma x + (\mu - \frac{1}{2}\sigma^2)t}.$$

This process is very useful in modelling stock prices since it is a strictly positive Brownian motion.

4.7.2 Ornstein Uhlenbeck Process

This process is used in the study of the behaviour of certain gasses and it has the following SDE:

$$dX_t = -\alpha X_t dt + \sigma dB_t.$$

We have already seen

$$df(X_t e^{\alpha t}) = \alpha X_t e^{\alpha t} dt + e^{\alpha t} dX_t$$

substituting dX_t we get

$$df(X_t e^{\alpha t}) = \alpha X_t e^{\alpha t} dt + e^{\alpha t}[-\alpha X_t dt + \sigma dB_t] = e^{\alpha t}\sigma dB_t.$$

So by integrating from 0 to t:

$$X_t e^{\alpha t} - X_0 = \sigma \int_0^t e^{\alpha s} dB_s \implies X_t = X_0 e^{-\alpha t} + \int_0^t e^{\alpha(s-t)} dB_s.$$

This process has a variation, termed the *mean-reverting* Ornstein Uhlenbeck process whose SDE is:

$$dX_t = (c - X_t)dt + \sigma dB_t.$$

Here we make the following guess

$$X_t = c + \vartheta(t)\left[X_0 + \tau c + \int_0^t h(s) dB_s\right].$$

Differentiating we get

$$dX_t = \frac{\vartheta'(t)}{\vartheta(t)}(X_t - c)dt + \vartheta(t)h(t)dB_t.$$

Hence

$$-\vartheta'(t) = \vartheta(t) \implies \vartheta(t) = e^{-t} \quad \text{and}$$

$$\vartheta(t)h(t) = \sigma \implies h(t) = \sigma e^t,$$

$$d(e^{-t}(X_0 + \tau c)) = (c - X_0)e^{-t}dt \implies \tau = -1,$$

so we have

$$X_t = c + (X_0 - c)e^{-t} + \int_0^t e^{s-t}dB_s.$$

So we see that

$$E[X_t] = c + (X_0 - c)e^{-t},$$

and since

$$Var(X_t) := E[(X_t - E[X_t])^2],$$

hence as a result:

$$Var(X_t) = \frac{\sigma^2}{2}[1 - e^{-2t}].$$

4.7.3 Numerical Solutions

[3] For SDEs without close form solutions, we can make numerical approximations for the solution. The method is similarly to that of ordinary differential equations. In ordinary differential equations we use the finite difference from the Taylor series. Given $[a,b]$, we can make n equal partitions with a step size h. Then following approximation can be made: for $Z'(x), x \in [a,b]$:

$$Z((i+1)h) = Z(ih) + hZ'(ih).$$

This can also be done in two variables.To adapt this method for a SDE, we need to take a sample Brownian motion or a path, this path can be fixed according to B_t. If we have a fixed path for B_t then for

$$dX_t = \mu dt + \sigma dB_t$$

we can make the following approximation:

$$\begin{aligned} Z((i+1)h) &= Z(ih) + hdX(ih) \\ &= Z(ih) + h[\mu(ih, Z(ih))dt + \sigma(ih, Z(ih))dB(ih)] \end{aligned}$$

From the fixed sample path,

$$dB(ih) = B((i+1)h) - B(ih)\, dt = h$$

To get this fixed path, we use the **Monte Carlo simulation**.

When numerically solving SDEs by hand where the solution is path dependent, the tree method can be used. Some SDEs can be transformed into the well known heat PDE,

$$\frac{\partial \varphi}{\partial t} = \frac{\partial^2 \varphi}{\partial x^2}.$$

In a perfectly insulated infinitely long conductor where at $t = 0$, the distribution of heat is known and heat can only travel in the x-axis, after some time t

$$\frac{\partial \varphi}{\partial t} = \frac{\partial^2 \varphi}{\partial x^2},$$

tells us how the distribution of heat will be along the substance. With initial conditions $\varphi(0,x) = \vartheta(x)$. The solution to the above PDE is:

$$\varphi(t,x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{-(x-s)^2}{4t}} \vartheta(s) ds.$$

The Black Scholes equation can be further studied using this PDE which is well understood [1] [3] [9].

Chapter 5

APPLICATIONS OF ITō CALCULUS IN FINANCE

The modelling of instantaneous interest rates and the pricing of the European call option are made possible by applying Itō calculus, since both of these have a Brownian behaviour and thus can be represented or approximated by a stochastic process.

5.1 Stock Prices as Stochastic Processes

[8] Given a stock in a stock market, we want to write its price as a stochastic process at time t. Let $\varphi(t)$ be the drift term, while $\Delta(t)$ is the stock volatility, so that we obtain

$$dX_t = \Delta(t)dW_t + \varphi(t)dt.$$

Since prices are always non-negative, we need a strictly positive stochastic process. To satisfy this restriction we use a geometric Brownian process.

5.1.1 Geometric Brownian Process

Let S_t be the price for a stock, define an Itō process

$$X_t = \int_0^t \sigma dW_t + \int_0^t (\alpha - \frac{1}{2}\sigma^2)dt,$$

where $X_0 = 0$, so we have

$$dX_t = \sigma dW_t + (\alpha - \frac{1}{2}\sigma^2)dt.$$

So we see that

$$[dX_t]^2 = \sigma^2 dt,$$

is the quadratic variation. Since σ and α are positive constants, we have

$$X_t = \sigma W_t + (\alpha - \frac{1}{2}\sigma^2)t.$$

Now we can define S_t to be

$$S_t = S_0 e^{X_t} = S_0 e^{\sigma W_t + (\alpha - \frac{1}{2}\sigma^2)t}.$$

This is called the asset price, W_t makes S_t random.

Appling the Itō formula to $\varphi(X_t) = S_0 e^{X_t}$, we have

$$d\varphi(X_t) = dS_t = \varphi'(X_t)dX_t + \frac{1}{2}\varphi''(X_t)[dX_t]^2.$$

Substituting dX_t and the quadratic variation, we have

$$\begin{aligned} dS_t &= S_0 e^{X_t}(\sigma dW_t + (\alpha - \frac{1}{2}\sigma^2)dt) + \frac{1}{2}S_0 e^{X_t}\sigma^2 dt \\ &= \sigma S_t dW_t + \alpha S_t dt = S_t(\sigma dW_t + \alpha dt). \end{aligned}$$

This SDE is very useful in describing the change in price of an asset in the future to some degree of accuracy since S_t is random.

5.1.2 Analysing Two Stocks

[4] Suppose we have two stocks in a market with the following asset prices $S_1(\mu)$ and $S_2(\mu)$ such that:

$$dS_1(\mu) = \alpha_1 S_1(\mu)d\mu + \sigma_1 S_1(\mu)dW_1(\mu) \tag{5.1}$$

$$dS_2(\mu) = \alpha_2 S_2(\mu)d\mu + \sigma_2 S_2(\mu)[\rho dW_1(\mu) + \sqrt{1-\rho^2}dW_2(\mu)]. \tag{5.2}$$

Where $\rho \in [0,1]$ is $cov(W_1(\mu), W_2(\mu))$. For $\rho = 0, 1$ $S_2(\mu)$ depends only on $W_1(\mu)$ or $W_2(\mu)$ respectively. When $W_1(\mu)$ and $W_2(\mu)$ are independent we can define

$$dW_3(\mu) = \rho dW_1(\mu) + \sqrt{1-\rho^2}dW_2(\mu).$$

So we see that,

$$dW_3(\mu)dW_3(\mu) = d\mu \implies [W_3, W_3](\mu) = \mu.$$

So $W_3(\mu)$ is Brownian as result of Levy's condition and so we can adapt it to the same filtration as $W_1(\mu)$ and $W_2(\mu)$. Hence from

$$dS_2(\mu) = \alpha_2 S_2(\mu)d\mu + \sigma_2 S_2(\mu)dW_3(\mu),$$

we see that the correlation coefficient of $W_1(\mu)$ and $W_3(\mu)$ is not zero, in fact

$$[W_1, W_3](\mu) = \rho \implies E[W_1(\mu)W_3(\mu)] = \int_0^\mu \rho d\mu = \rho\mu.$$

So there is a relation between $S_1(\mu)$ and $S_2(t)$ so in essence a change in the price of asset $S_1(\mu)$ will affect the price of asset $S_2(\mu)$.

5.2 Interest Rate

[5] Suppose we invest some capital say P and after some time t we are paid and added sum

$$P + Pr = P(r+1),$$

here r is our rate of interest. In banks, r keeps changing with time, we want to know the instantaneous rate. Let $R(s,t)$ be the rate from s to t, with the condition $0 \leq s \leq t$. Then

$$R_t = \lim_{s \to t} R(s,t),$$

gives the instantaneous rate of interest. R_t is a random process but we also want it to be an Itō process.

5.2.1 Vasicek's Model of Interest Rate

Let σ, α and τ be positive constants, define

$$dR_t = (\alpha - \tau R_t)dt + \sigma dW_t.$$

This model has a closed form solution

$$R_t = R_0 e^{-\tau t} + \frac{\alpha}{\tau}(1 - e^{-\tau t}) + \sigma e^{-\tau t} \int_0^t e^{\tau s} dw(s).$$

We will verify this using Itō formula, let

$$X_t = \int_0^t e^{\tau t} dW_u \implies dX_t = e^{-\tau t} dW_t,$$

where $X_0 = 0$ and define:

$$\Gamma(t, X_t) = R_0 e^{-\tau t} + \frac{\alpha}{\sigma}(1 - e^{-\tau t}) + \sigma e^{-\tau t} X_t$$

$$\Gamma_t(t, X_t) = -\tau R_0 e^{-\tau t} + \alpha e^{-\tau t} - \sigma e^{-\tau t} X_t,$$

$$\Gamma_{\nu}(t, X_t) = \sigma e^{-\tau t}, \Gamma_{\nu\nu}(t, X_t) = 0.$$

Substituting these in

$$dR_t = d\Gamma(t, X_t) = \Gamma_t(t, X_t)dt + \Gamma_{\nu}(t, X_t)dX_t + \frac{1}{2}\Gamma_{\nu\nu}(t, X_t)dX_t dX_t,$$

we have

$$\begin{aligned} dR_t &= [-\tau R_0 e^{-\tau t} + \alpha e^{-\tau t} - \sigma e^{-\tau t} X_t]dt + \sigma e^{-\tau t} dX_t \\ &= [\alpha - \tau(R_0 e^{-\tau t} + \frac{\alpha}{\tau}(1 - e^{-\tau t}) + \sigma e^{-\tau t} X_t)]dt + \sigma dW_t \\ &= (\alpha - \tau R_t)dt + \sigma dW_t. \end{aligned}$$

From

$$X_t = \int_0^t e^{\tau t} dW_u \implies dX_t = e^{-\tau t} dW_t,$$

it follows $E[X_t] = 0$ and

$$Var(X_t) = [X, X](t) \implies Var(X_t) = \int_0^t e^{2\tau} ds = \frac{1}{2\tau}(e^{2\tau} - 1).$$

So that $X_t \sim \mathcal{N}\left(0, \frac{1}{2\tau}(e^{2\tau} - 1)\right)$.

5.2.2 Cox Ingersoll Ross Model

This model has the representation below,

$$dR_t = \alpha(\beta - R_t)dt + \sigma\sqrt{R_t}dW_t.$$

Where $\alpha(\beta - R_t)$ is the drift to the mean and $\sigma\sqrt{R_t}$ is the diffusion which presents the market volatility. The Cox-Ingersoll-Ross model's properties,

1. Is a non-negative interest rate model.

2. It is a Wiener process with drift.

3. Mean reverting the long term mean.

This model has no close form solution however an approximation using the Monte Carlo Simulation can generate a numerical solution.

5.2.3 Compound Interest

Suppose we invest 1 dollar in money market with rate r, for $t_i = \frac{i}{n}T \in [0,T]$ so that the instantaneous interest rate is:

$$R_T = \lim_{n\to\infty}\left(1+\frac{rT}{n}\right)^n = e^{rT}$$

so for a P dollar investment, we get Pe^{rT} as our total sum after time T. Now let B_0 be the money to be invested and B_t is the amount after time t then we have.

$$B_t = B_0 e^{rt} \implies B_0 = B_t e^{-rt}.$$

So by knowing how much we want to make in a period t for r, we can calculate how much to invest at 0. This is called the discounted price of the fixed deposit for time 0 at time t. When $t = 0$, $S_0 = e^{-rt}S_t$. Since

$$dS_t = \alpha S_t dt + \sigma S_t dW_t,$$

and thus

$$dS_t = S_0 e^{rt}[\alpha dt + \sigma dW_t].$$

This is the Geometric Brownian process and we have already seen in the last part of the previous chapter its solution. The **Black Scholes equation** gives the premium $X_t = \gamma(\mu, S_\mu)$ [4].

5.3 Portfolio Pricing

In the study of mathematical finance the two most fundamental aspects are **portfolio pricing** and **portfolio optimization**. In optimizing the portfolio, the mathematical tools used the most is multi-variable calculus, linear field and the Lagrange multipliers and the most fundamental results is the Karush-Kuhn-Turker (KKT) condition. The portfolio pricing on the other hand depends almost entirely on the Itō calculus. For this write up we will focus more on the Portfolio pricing for a European call option which directly applies the Itō calculus in the continuous time model.

5.3.1 Pricing in Discrete Time

Definition of some terms used in financial markets:

(i) **Call option**: a stock is a call option when the holder is ready to sell the stock.

(ii) **Put option**: is when a buyer is ready to buy a call option.

(iii) **Strike price**: is the agreed price to sell the option denoted by k.

Suppose I hold an option in a market and let S_t be the asset price where $t \in [0, T]$ such that the expiration time for the option is T. Supposed I enter in a contract with a buyer of my asset S_t with strike price k. In this contract you are not obligated to buy from me if the market price is less than k and I am oblige selling you the asset at k, even if the market price is higher than k. For all scenarios, I am at a disadvantage to avoid this, the contract for the option will involve a premium price sometimes called the **option price**. **Black Scholes Equation** is used for finding the option price for an asset necessitated the.

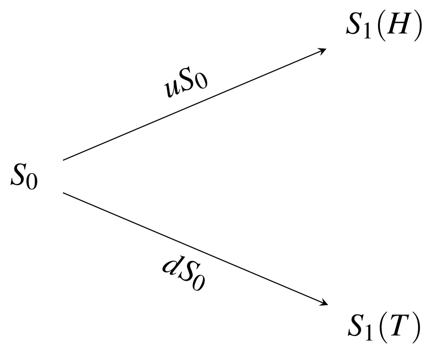

Figure 5.1: One Period Binomial Model

For a stock S_t with strike price k, the option value denoted by $V(t)$ is:

$$V(t) = \max\{S_t(\cdot) - k, 0\}.$$

For T the option value is

$$V(T) = \max\{S_T(\cdot) - k, 0\}.$$

This is the only time to exercise in European call options. Here finding the value of V_0, gives the option price or premium. So we can use V_0 to Hedge the option. A market where hedging is always possible is called a complete market but this is not always the case in most markets around the world.

5.3.2 The Binomial (Cox Loss Rubenstein) Model

At $t = 0$, let S_0 stock price. Now toss a fair coin with probabilities $\mathbf{P}(H) = p$ and $\mathbf{P}(T) = q$. Suppose the prices goes up by u when H shows up and decrease with d when T shows up, we illustrate this as in Figure 1.

Suppose we have a simple market, let Z_0 be the amount charged for a call option. We can use some portion of Z_0 to buy $\Lambda_0 S_0$ amount of stock and invest the rest which is $Z_0 - \Lambda_0 S_0$ in banks at rate r, so that for $t = 1$, With the no arbitrage condition $0 < d < r+1 < u$. We obtain

$$Z_1 = \Lambda_0 S_1 + (r+1)[Z_0 - \Lambda_0 S_0] \implies Z_1 = (r+1)Z_0 + \Lambda_0[S_1 - (r+1)S_0].$$

With the no arbitrage condition $0 < d < i+r < u$ and let $Z_0 = V_0$ hence $Z_1(\cdot) = V_1(\cdot)$ so we have the following from the binomial model presented in Figure 1.

$$\frac{V_1(H)}{(r+1)} = Z_0 + \Lambda_0\left[\frac{V_1(H)}{(r+1)} - S_0\right] \tag{5.3}$$

and

$$\frac{V_1(T)}{(r+1)} = Z_0 + \Lambda_0\left[\frac{V_1(T)}{(r+1)} - S_0\right]. \tag{5.4}$$

Let $\hat{p} + \hat{q} = 1$ multiply equation (5.3) with $\hat{p}$ and (5.4) with $\hat{q}$, summing the results to obtain:

$$\frac{1}{r+1}(\hat{p}V_1(H) + \hat{q}V_1(T)) = Z_0 + \Lambda_0\left[\frac{1}{r+1}(\hat{p}S_1(H) + \hat{q}S_1(T)) - S_0\right].$$

Let choose $\hat{p}$ and $\hat{q}$ such that

$$\frac{1}{r+1}(\hat{p}S_1(H) + \hat{q}S_1(T)) = S_0. \tag{5.5}$$

Hence it follows

$$Z_0 = \frac{1}{r+1}(\hat{p}V_1(H) + \hat{q}V_1(T)).$$

from equation (5.5), we have

$$S_0 = \frac{1}{r+1}(\hat{p}uS_0 + (1-\hat{p})dS_0) \implies 1 = \frac{1}{r+1}[\hat{p}u + (1-\hat{p})d],$$

hence,

$$\hat{p} = \frac{r+1-d}{u-d}, \hat{q} = \frac{u-r-1}{u-d}.$$

Since $0 < d < r+1 < u$, therefore $\hat{q}, \hat{p} > 0$. Thus $\hat{p}, \hat{q}$ are probabilitymeasures but they are not the market probabilities. If we subtract equation (5.3) and (5.4) we have

$$\Lambda_0 =: \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)}.$$

This is called the delta of the hedging. From (5.5) we see that

$$S_0 = \frac{1}{r+1}\hat{\mathbf{E}}[S_1] = \hat{\mathbf{E}}\left[\frac{S_1}{r+1}\right].$$

So S_t is a martingale Under $\hat{p}, \hat{q}$, this means the stock market's growth is the same as that of the money market. So we can invest only in the money market which is risk-free. $\hat{p}, \hat{q}$ are called the risk-neutral probabilities.

5.3.3 Two Period Binomial Model

Assuming we have similar conditions as in the one period binomial model.

Assume at $t = 2$

$$V_2(\cdot) = \max\{S_2(\cdot) - k, 0\},$$

at $t = 1$ we buy $\Lambda_1(\cdot)S_1(\cdot)$ shares and we invest the rest in the money market, so the

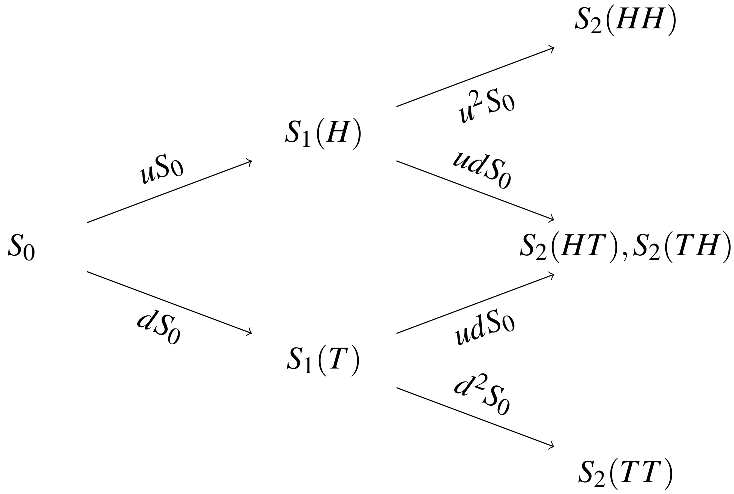

Figure 5.2: Two Period Binomial Model

wealth equation becomes

$$V_2(\cdot) = \Lambda_1(\cdot)S_1(\cdot) + (r+1)[V_1(\cdot) - \Lambda_1(\cdot)S_1(\cdot)].$$

From the figure below, we get the following equations for each possibility.

$$V_1(H) = \Lambda_0 S_1(H) + (r+1)[V_0 - \Lambda_0 S_0].$$

$$V_1(T) = \Lambda_0 S_1(T) + (r+1)[V_0 - \Lambda_0 S_0].$$

$$V_2(HH) = \Lambda_1(H)S_(HH) + (r+1)[V_1(H) - \Lambda_1 S_1(H)].$$

$$V_2(HT) = \Lambda_1(H)S_(HT) + (r+1)[V_1(H) - \Lambda_1 S_1(H)].$$

$$V_2(TH) = \Lambda_1(T)S_(TH) + (r+1)[V_1(T) - \Lambda_1 S_1(T)].$$

$$V_2(TT) = \Lambda_1(T)S_(TT) + (r+1)[V_1(T) - \Lambda_1 S_1(T)].$$

Under $\hat{p}, \hat{q}$,

$$\Lambda_1(H) =: \frac{V_2(HH) - V_2(HT)}{S_2(HH) - S_2(HT)},$$

and

$$V_1(H) = \frac{1}{r+1}[\hat{p}V_2(HH) + \hat{q}V_2(HT)] = \hat{\mathbf{E}}\left[\frac{V_2}{r+1}\right].$$

For the K-period model with no arbitrage condition under $\hat{p}, \hat{q}$, let V_K be the pay-off of the option at $t = K$. Then for $\omega_1, \ldots, \omega_K$ we can backtrack $V_K \to V_{K-1} \to, \ldots, \to V_0$ so for any $0 < k < K$, we have

$$\begin{aligned}
\Lambda_k(\omega_1, \ldots, \omega_k) &= \frac{V_{n+1}(\omega_1, \ldots, \omega_k H) - V_{k+1}(\omega_1, \ldots, \omega_k T)}{S_{k+1}(\omega_1, \ldots, \omega_k H) - S_{k+1}(\omega_1, \ldots, \omega_k T)}. \\
V_k(\omega_1, \ldots, \omega_k) &= \frac{1}{r+1}[\hat{p}V_{k+1}(\omega_1, \ldots, \omega_k H) + \hat{q}V_{k+1}(\omega_1, \ldots, \omega_k T)], \\
&= \frac{1}{r+1}\hat{\mathbf{E}}[V_{k+1}](\omega_1, \ldots, \omega_k).
\end{aligned}$$

The discounted price under $\hat{p}, \hat{q}$ is a martingale that is,

$$\hat{\mathbf{E}}\left[\frac{S_{k+1}}{(r+1)^{k+1}}\right] = \frac{1}{(r+1)^k}\hat{\mathbf{E}}\left[\frac{S_{k+1}}{r+1}\right] = \frac{S_k}{(r+1)^k}.$$

5.3.4 Portfolio Wealth Process

Assume $\mathcal{F}$ is generated by coin tosses and stochastic process $\{\Lambda_0, \Lambda_1, \ldots, \Lambda_k, \ldots, \Lambda_{K-1}\}$ is $\mathcal{F}-$adapted. Let $\{Z_1, Z_2, \ldots, Z_k, \ldots, Z_{K-1}\}$ be the wealth process which is also $\mathcal{F}-$adapted then we have following wealth equation.

$$Z_{k+1} = \Lambda_k S_{k+1} + (r+1)[Z_k - \Lambda_k S_k]. \tag{5.6}$$

Theorem 5.3.1. Under $\hat{p}, \hat{q}$, the discounted wealth process $\frac{Z_k}{(r+1)^k}$ is a martingale for each k. That is,

$$\frac{Z_k}{(r+1)^k} = \hat{\mathbf{E}}\left[\frac{Z_{k+1}}{(r+1)^{k+1}}\right].$$

Proof. We will use equation (5.6).

$$\begin{aligned}
\hat{\mathbf{E}}\left[\frac{Z_{k+1}}{(r+1)^{1+k}}\right] &= \hat{\mathbf{E}}\left[\frac{\Lambda_k S_{k+1}}{(r+1)^{1+k}} + \frac{Z_k - \Lambda_k S_k}{(r+1)^k}\right] \\
&= \Lambda_k \hat{\mathbf{E}}\left[\frac{S_{k+1}}{(r+1)^{1+k}}\right] + \left[\frac{Z_k - \Lambda_k S_k}{(r+1)^k}\right] \\
&= \frac{\Lambda_k S_k}{(r+1)^k} + \left[\frac{Z_k - \Lambda_k S_k}{(r+1)^k}\right] = \frac{Z_k}{(r+1)^k}.
\end{aligned}$$

□

Since $Z_k = V_k$ so we have,

$$\frac{V_k}{(r+1)^k} = \hat{\mathbf{E}}\left[\frac{V_{k+1}}{(r+1)^{k+1}}\right].$$

This is the discounted portfolio value and its analogue in continuous time is the Black Scholes formula [4] [7] [8].

5.4 Pricing in Continuous Time

In the last section of Chapter 1 we discussed about conditional expectation, we will used some of those ideas here.

Given $(\Theta, \mathcal{H}, \mathbf{P})$, assume $Z > 0$ is a random variable with such that $\forall A \in \mathcal{H}$ and for $\hat{\mathbf{P}}$,

$$\hat{\mathbf{P}}(A) = \int_A Z(\cdot) d\mathbf{P}(\cdot).$$

If $\mathbf{E}[Z] = 1$ then

$$\hat{\mathbf{P}}(\Theta) = \int_\Theta Z(\cdot) d\mathbf{P}(\cdot).$$

Here, Z is termed Radon Nikodym derivative and $Z = \dfrac{d\hat{\mathbf{P}}}{d\mathbf{P}}$. Suppose X is a random variable then $\hat{\mathbf{E}}[X] = [XZ]$ so for $\mathbf{P}\{Z > 0\} = 1$,

$$\mathbf{E}[X] = \hat{\mathbf{E}}\left[\frac{1}{Z}X\right].$$

If $X \sim \mathcal{N}(0,1)$ and $Z = e^{-(\theta x + \frac{1}{2}\theta^2)}$ where θ is a constant, suppose W is a random variable in $\hat{\mathbf{P}}$ with $W = \theta + X$, then $\mathbf{E}[W] = \theta$ and $\hat{\mathbf{E}}[W] = E[WZ] = 0$. Let $\mathcal{G}$ be contained in $\mathcal{F}$ and $\mathbf{E}[X|\mathcal{G}]$. Then using partial averaging we have

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

For $A = \Theta$,

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

So this makes $\mathcal{G}$ an unbiased estimator of X.

For $t, s \in [0, T]$ with $s < t$, from tower property we discussed in conditional expectation, we can proof Z being a martingale that is

$$\mathbf{E}[Z_t|\mathcal{H}_s] = Z_s, \quad \text{and} \quad \mathbf{E}[Z|\mathcal{H}_t] = Z_t,$$

and this is the Radon Nikodym process.

Lemma 5.4.1. Given $W(\cdot)$ is $\mathcal{H}_t$-measurable an $0 < t < T$ we have

$$\hat{\mathbf{E}}[W(\cdot)] = E[W(\cdot)Z_t].$$

Proof. Consider

$$\begin{aligned}\hat{\mathbf{E}}[W(\cdot)] = \mathbf{E}[W(\cdot)Z] &= \mathbf{E}[\mathbf{E}[W(\cdot)Z|\mathcal{H}_t]] \\ &= \mathbf{E}[W(\cdot)\mathbf{E}[Z|\mathcal{H}_t]] \\ &= \mathbf{E}[W(\cdot)Z_t], \forall \mathcal{H}_t \subset \mathcal{F}\end{aligned}$$

□

Lemma 5.4.2. Let r and t be such that $0 < r < t < T$ let Y be $\mathcal{F}_t$-measurable random variable then

$$\hat{\mathbf{E}}[Y(\cdot)|\mathcal{H}_s] = \frac{1}{Z_s} E[Y(\cdot)Z_t|\mathcal{H}_s].$$

Proof. Consider

$$\int_A \hat{\mathbf{E}}[Y(\cdot)|\mathcal{H}_s] d\hat{\mathbf{P}} = \int_A Y(\cdot) d\hat{\mathbf{P}}, \forall A \in \mathcal{H}_s.$$

So we need to show that

$$\int_A \frac{1}{Z(s)} \mathbf{E}[Y(\cdot)Z(t)|\mathcal{H}_s] d\hat{\mathbf{P}} = \int_A Y d\hat{\mathbf{P}}.$$

Since

$$\frac{1}{Z_s} \mathbf{E}[Y(\cdot)Z(t)|\mathcal{H}_s] d\hat{\mathbf{P}},$$

is $\mathcal{H}_s$-measurable so

$$\begin{aligned}\int_A \frac{1}{Z_s} E[Y(\cdot)Z_t|\mathcal{H}_s] d\hat{\mathbf{P}} &= \int_\Theta I_A \frac{1}{Z_s} \mathbf{E}[Y(\cdot)Z_t|\mathcal{H}_s] d\hat{\mathbf{P}}, \\ &= \hat{\mathbf{E}}\left[I_A \frac{1}{Z_s} \mathbf{E}[Y(\cdot)Z_t|\mathcal{H}_s]\right], \\ &= E[I_A \mathbf{E}[Y(\cdot)Z_t|\mathcal{H}_s]].\end{aligned}$$

Where I_A is $\mathcal{H}_s$-measurable and an unbiased estimator, so

$$\mathbf{E}[\mathbf{E}[I_A Y(\cdot) Z_t | \mathcal{H}_s]] = \mathbf{E}[I_A Y(\cdot) Z_t | \mathcal{H}_s].$$

By Lemma (5.4.1)

$$\mathbf{E}[I_A Y(\cdot) Z_t | \mathcal{H}_s] = \mathbf{E}[I_A Y(\cdot) Z_t] = \hat{\mathbf{E}}[I_A Y(\cdot)] = \int_A Y(\cdot) d\hat{\mathbf{P}}.$$

□

Theorem 5.4.1 (Girsanov). Given $(\Theta, \mathcal{F}, \mathbf{P})$, suppose W_μ is a Brownian and $\mathcal{F}_\mu$ be associated with it. Assume $\theta(\cdot)$ is adapted to $\mathcal{F}_\mu$, define:

$$Z_\mu = \exp\left\{ -\int_0^\mu \theta(\mu) dW_\mu - \frac{1}{2}\int_0^\mu \theta^2(\mu) d\mu \right\}.$$

and

$$\hat{W}_\mu = W_\mu + \int_0^\mu \theta(u) du.$$

Assuming

$$E\left[\int_0^\mu \theta^2(u) du\right] < +\infty,$$

where $o < \mu < T$, set $Z = Z(T)$. Then $E[Z] = 1$ and if $\hat{\mathbf{P}}$ is the new probability measure given by Z then under $\hat{\mathbf{P}}$, $\hat{W}_\mu$ is Brownian.

We will prove this, using Levy's condition.

Proof. $\hat{W}_\mu$ must meet the conditions below:

1. $\hat{W}(0) = 0$,
2. $[\hat{W}_\mu, \hat{W}_\mu] = \mu$ quadratic variation,

3. Under $\hat{\mathbf{P}}$, $\hat{W}_\mu$ is a martingale.

From the definition of $\hat{W}_\mu$ we see that (i) and (ii) are trivial. For (iii), first we need to prove the martingale property of Z_μ. Lets define,

$$X_\mu = -\int_0^\mu \theta(\nu)dW_\nu - \frac{1}{2}\int_0^\mu \theta^2(\nu)d\nu \implies dX_\mu = -\theta(\mu)dW_\mu - \frac{1}{2}\theta^2(\mu)d\mu$$

so that

$$dX^2(\mu) = \theta^2(\mu)D_\mu.$$

Let $\Lambda(x) = e^x$ then $\Lambda'(x) = \Lambda''(x) = \Lambda(x) = e^x$, using Itō formula to get

$$dZ_\mu = d\Lambda(X_\mu) = \Lambda'(X_\mu)dX_\mu + \frac{1}{2}\Lambda''(X_\mu)dX_\mu^2.$$

Hence

$$\begin{aligned} dZ_\mu &= e^{X_\mu}[-\theta(\mu)dW_\mu - \frac{1}{2}\theta^2(\mu)dt] + \frac{1}{2}\theta^2(\mu)D_\mu e^{X_\mu} \\ &= -Z_\mu\theta(\mu)dW_\mu. \end{aligned}$$

Hence it follows

$$Z_\mu = Z_0 - \int_0^\mu Z_\nu\theta(\nu)dW_\nu.$$

Z_μ is a martingale since it has an Itō representation. Hence

$$\mathbf{E}[Z(\cdot)] = \mathbf{E}[Z_T] = \mathbf{E}[Z(0)] = 1,$$

and

$$Z_\mu = \mathbf{E}[Z_T|\mathcal{F}_\mu] = \mathbf{E}[Z(\cdot)|\mathcal{F}_\mu].$$

So Z_μ is a Radon Nikodym process. Now we will show the martingale property for $\{\hat{W}_\mu Z_\mu\}$.

$$\begin{aligned}
d(\hat{W}_\mu Z_\mu) &= \hat{W}_\mu Z_\mu + Z(t)d\hat{W}_t + d\hat{W}_\mu Z_\mu \\
&= -\hat{W}_\mu Z_\mu \theta(\mu)dW_\mu + Z_\mu[dW_\mu + \theta(\mu)d\mu] - \theta(\mu)Z(\mu)d\mu \\
&= -\hat{W}_\mu Z_\mu \theta(t)dW_\mu + Z_\mu dW_\mu \\
&= [1 - \hat{W}_\mu \theta(\mu)]Z(\mu)dW_\mu.
\end{aligned}$$

So we have

$$\hat{W}_\mu Z_\mu = \int_0^\mu [1 - \hat{W}_\nu \theta(\nu)]Z_\nu dW_\nu.$$

So this is an Itō integral, hence $\{\hat{W}_\mu Z_\mu\}$ is a martingale. Now we need to show that with respect to $\mathcal{F}_\nu$, for $0 \leq \nu \leq \mu \leq T$, $\{\hat{W}_\mu\}$ is a martingale .

$$\hat{\mathbf{E}}[\hat{W}_\mu|\mathcal{F}_\nu] = \frac{1}{Z_\nu}\mathbf{E}[\hat{W}_\mu Z_\mu|\mathcal{F}_\nu],$$

by Lemma (5.4.2),

$$= \frac{1}{Z_\nu}\hat{W}_\nu Z_\nu = \hat{W}_\nu,$$

hence under $\hat{P}$, $\hat{W}_\mu$ is a martingale, thus levy's condition is meet. □

This is all true in the Almost surely sense. $\hat{W}_t$ is used for pricing in continuous time.

5.4.1 Risk-Neutral Pricing.

Given $(\Theta, \mathcal{F}, \mathbf{P})$, let $\{W_t\}$, be a Wiener process for $\mu \in [0,T]$ and $\{\mathcal{F}_\mu\}$ is a filtration associated with $\{W_\mu\}$. Recall the following SDE

$$dS_\mu = \alpha(t)S_\mu dt + \sigma(t)S_\mu dW_t,$$

where $\alpha(t)$ and $\sigma(t)$ denotes stock mean return and volatility respectively. This has the following solution

$$S_\mu = S_0 \exp\left\{\int_0^t \sigma(s)w(s) - \int_0^t \left[\alpha(s) - \frac{1}{2}\sigma^2(s)ds\right]\right\}.$$

Assume $\{R_\mu\}$ is the stochastic process denoting the rate of interest adapted to $\left\{\mathcal{F}_\mu\right\}_{t=0}^{T}$. The discounted process $\{D_\mu\}$ is define as:

$$D_\mu = e^{-\int_0^\mu R_\nu d\nu}.$$

Let $\Lambda(\nu) = e^{-\nu}$ then we have $\Lambda'(\nu) = -e^{-\nu}$ and $\Lambda''(\nu) = \Lambda(\nu)$. Now let

$$I(\mu) = \int_0^\mu R_\nu d\nu,$$

So that

$$dI(\mu) = R_\mu d\mu \implies dD_\mu = -D_\mu R_\mu d\mu.$$

Appling the product rule

$$\begin{aligned} d(D_\mu S_\mu) &= D_\mu dS_\mu + S_\mu dD_\mu + dR_\mu dS_\mu, \\ &= D_\mu S_\mu[\alpha(\mu) - R_\mu]d\mu + \sigma(\mu)D_\mu S_\mu dW_\mu. \end{aligned}$$

Set

$$\theta(\mu) = \frac{\alpha(\mu) - R_\mu}{\sigma(\mu)},$$

so that

$$d(D_\mu S_\mu) = \sigma(\mu) D_\mu S_\mu [\theta(\mu) d\mu + dW_\mu],$$
$$= \sigma(\mu) D_\mu S_\mu d\hat{W}_\mu.$$

Hence

$$D_\mu S_\mu = S_0 + \int_0^\mu D_\nu S_\nu \sigma(\nu) d\hat{W}_\nu.$$

So under $\hat{P}$, $D_\mu S_\mu$ is martingale. Here

$$\theta(\mu) = \frac{\alpha(\mu) - R_\mu}{\sigma(\mu)},$$

is the risk-price.

$$dS_t = \alpha(t) S_\mu d|mu + \sigma(\mu) S_\mu dW_\mu,$$
$$= [\theta(\mu)\sigma(\mu) + R_\mu] S_\mu d\mu + \sigma(\mu) S_\mu dW_\mu,$$
$$= R_\mu S_\mu d\mu + \sigma(\mu) S_\mu [\theta(\mu) d\mu + dW_\mu],$$
$$= R_\mu S_\mu d\mu + \sigma(\mu) S_\mu d\hat{W}_\mu.$$

This is the risk-neutral SDE for S_μ.

5.4.2 Pricing the European Call Option

Let $\{S_\mu\}_{\mu=0}^T$ be the stock price for call option so at $\mu = 0$ we have a contract so that we can sell shares of S_μ at some price k, hence the option value is;

$$X_\mu = \max\{S_\mu - k, 0\},$$

and $X_\mu = V(\mu)$, $X_0 = V(0)$. Recall the portfolio wealth equation:

$$dX_\mu = \Lambda(\mu)dS_\mu + R_\mu[X_\mu - \Lambda(\mu)S_\mu]d\mu.$$

Substituting the Geometric Brownian, to get:

$$dX_\mu = \Delta(t)[\alpha(t)S_\mu dt + \sigma(t)S_\mu dW_\mu] + R_\mu[X_\mu - \Delta(t)S_\mu]dt$$

$$= \Delta(t)S_\mu\sigma(t)d\hat{W}_\mu + R_\mu X_\mu dt.$$

By Itō product rule we have

$$d(D_\mu X_\mu) = D_\mu dX_\mu + X_\mu dD_\mu + dD_\mu dX_\mu, \tag{5.7}$$

since D_μ is not Brownian so $dD_\mu dX_\mu = 0$. Substituting dX_μ in (5.7) gives,

$$d(D_\mu X_\mu) = D_\mu\Lambda(\mu)[[\alpha(\mu)S_\mu d\mu + \sigma(\mu)S_\mu dW_\mu] +$$

$$R_\mu[X_\mu - S_\mu]d\mu] - D_\mu R_\mu X_\mu d\mu$$

$$= D_\mu\Lambda(\mu)[\Lambda(\mu)dS_\mu - R_\mu S_\mu d\mu]$$

$$= D_\mu\Lambda(\mu)[\Lambda(\mu)[\alpha(\mu)S_\mu d\mu + \sigma(t)S_\mu dW_\mu] - R_\mu S_\mu d\mu]$$

$$= D_\mu\Lambda(\mu)[S_\mu[\alpha(\mu) - R_\mu]d\mu + \sigma(\mu)S_\mu dW_\mu]$$

$$= D_\mu\Lambda(\mu)[\sigma(\mu)\theta(\mu)S_\mu d\mu + \sigma(\mu)S_\mu dW_\mu]$$

$$= \Lambda(\mu)D_\mu\sigma(\mu)S_\mu d\hat{W}_\mu.$$

Hence

$$D_\mu X_\mu = X_0 + \int_0^\mu D_\nu \Lambda(\nu)\sigma(\nu) S_\nu d\hat{W}_\nu.$$

So under $\hat{P}$:

1. $D_\mu X_\mu$ is martingale,
2. The Black Scholes is a continuous time model for pricing European call-options [4] [5].

5.4.3 Black Scholes

Let T be the expiration time so that the option value at T is

$$V(T) = \max\{S_T - k, 0\},$$

is a $\mathcal{F}_\mu$-measurable random variable and at $\mu_0 = 0$ we have the stating price $V(\mu_0) = V(0)$. Suppose we have a market with the following conditions:

1. The market has only one stock with a money market (Bank or Bond)
2. Complete market (always possible to hedge option)
3. Free Arbitrage (Same price for same stock in different markets)
4. Risk-neutral pricing (does not matter weather you invest in stock or bond)

This is called a simple market.

The following gives the discounted price:

$$D_\mu V(\mu) = \hat{\mathbf{E}}[V(T) D_T | \mathcal{F}_\mu],$$

so that

$$V(\mu) = \hat{\mathbf{E}}\left[V(T)\frac{D_T}{D_\mu} | \mathcal{F}_\mu\right],$$

$$= \hat{\mathbf{E}}\left[V(T)e^{-\int_\mu^T R_\nu d\nu}|\mathcal{F}_\mu\right].$$

This is a risk-neutral pricing. Suppose the money market has an interest rate $R_\mu = r \implies D_\mu = e^{-r\mu}$ and the stock volatility σ, is constant. Then the risk-neutral SDE becomes,

$$dS_\mu = rS_\mu d\mu + \sigma S_\mu d\hat{W}_\mu.$$

Define $V(\mu) = \gamma(\mu, S_\mu)$, then it follows

$$\gamma(\mu, S_\mu) = \hat{\mathbf{E}}[e^{-(T-\mu)}\max\{S_T - k, 0\}|\mathcal{F}_\mu].$$

$$S_\mu = S_0 exp\{\sigma\hat{W}_\mu + (r - \frac{1}{2}\sigma^2)\mu\}. \tag{5.8}$$

$$S_T = S_0 exp\{\sigma\hat{W}_T + (r - \frac{1}{2}\sigma^2)T\}. \tag{5.9}$$

Dividing equation (5.9) by (5.8), gives

$$S_T = S_\mu \exp\{\sigma(\hat{W}_T - \hat{W}_\mu) + (r - \frac{1}{2}\sigma^2)(T - \mu)\}.$$

Let $\lambda = T - \mu$, and $\Theta = \frac{\hat{W}_T - \hat{W}_\mu}{\sqrt{\lambda}}$, so that $\hat{W}_T - \hat{W}_\mu \sim \mathcal{N}(0, \lambda)$ and therefore $\Theta \sim \mathcal{N}(0, 1)$. So we have,

$$S_T = S_\mu \exp\{-\sigma\sqrt{\lambda}\Theta + (r - \frac{1}{2}\sigma^2)\lambda\}.$$

This is independent of $\mathcal{F}_\mu$ since λ is beyond μ, hence

$$\gamma(\mu, z) = \hat{\mathbf{E}}\left[e^{-r\lambda}\max\{x\exp[-\sigma\sqrt{\lambda}\Theta + (r - \frac{1}{2}\sigma^2)\lambda] - k, 0\}|\mathcal{F}_\mu\right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-r\lambda} \max\{z \exp[-\sigma\sqrt{\lambda}\Theta + (r - \frac{1}{2}\sigma^2)\lambda] - k, 0\} e^{-\frac{1}{2}\theta^2} d\theta.$$

We need

$$\gamma(\mu, z) \geq 0 \implies z \exp[-\sigma\sqrt{\lambda}\theta + (r - \frac{1}{2}\sigma^2)\lambda] \geq k,$$
$$ln\left(\frac{z}{k}\right) \geq \sigma\sqrt{\lambda}\theta - (r - \frac{1}{2}\sigma^2)\lambda \implies, \theta \leq \frac{1}{\sigma\sqrt{\lambda}}\left[ln\left(\frac{z}{k}\right) + (r - \frac{1}{2}\sigma^2)\lambda\right].$$

Let

$$\varphi_-(\lambda, z) = \frac{1}{\sigma\sqrt{\lambda}}\left[ln\left(\frac{z}{k}\right) + (r - \frac{1}{2}\sigma^2)\lambda\right],$$

so that $\theta \leq \varphi_-(\lambda, z)$ so this gives,

$$\gamma(\mu, z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varphi_-(\lambda,z)} e^{-r\tau}[z \exp\{-\sigma\sqrt{\tau}\Theta + (r - \frac{1}{2}\sigma^2)\lambda\} - k] e^{-\frac{1}{2}\theta^2} d\theta,$$
$$= \frac{z}{\sqrt{2\pi}} \int_{-\infty}^{\varphi_-(\lambda,z)} \exp\left[-\frac{1}{2}\theta^2 - \sigma\sqrt{\lambda}\Theta + \frac{1}{2}\sigma^2\lambda\right] d\theta,$$
$$- \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varphi_-(\lambda,z)} e^{-r\lambda} k e^{-\frac{1}{2}\theta^2} d\theta,$$
$$= \frac{z}{\sqrt{2\pi}} \int_{-\infty}^{\varphi_-(\lambda,z)} \exp\left[-\frac{1}{2}(\theta + \sigma)^2\right] dy - e^{-r\lambda} k \Phi(\varphi_-(\lambda, z)).$$

Let $y = \theta + \sigma\sqrt{\lambda}$, then $Y \leq \varphi_+(\lambda, z)$, where

$$\varphi_+(\lambda, z) = \varphi_-(\lambda, z) + \sigma\sqrt{\lambda}.$$

So we have

$$\gamma(\mu, z) = \frac{z}{\sigma\sqrt{\tau}} \int_{-\infty}^{\varphi_+(\tau,x)} e^{-\frac{1}{2}y^2} dy - k e^{-r\lambda} \Phi(\varphi_-(\lambda, z)),$$
$$= z\Phi(\varphi_+(\lambda, z)) - k e^{-r\lambda} \Phi(\varphi_-(\lambda, z)).$$

Replacing z with S_μ we have

$$\gamma(\mu,z) = S_\mu \Phi(\varphi_+(\lambda, S_\mu)) - ke^{-r\lambda}\Phi(\varphi_-(\lambda, S_\mu)). \tag{5.10}$$

Since $\lambda = T - \mu$, (5.10) becomes

$$\gamma(\mu,z) = S_\mu \Phi(\varphi_+(T-\mu, S_\mu)) - ke^{-r(T-\mu)}\Phi(\varphi_-(T-\mu, S_\mu)).$$

This pricing formula is termed the Black Scholes.

Putting $\mu = 0$ then $\gamma(0, S_0) = V_0$, where S_0 is the initial stock price, hence

$$V_0 = S_0 \Phi(\varphi_+(T, S_0)) - ke^{-rT}\Phi(\varphi_-(T, S_0)).$$

This gives the option price V_0 (or premium) for the stock.

To find $\Lambda(\mu)$ we use the representation theorem which says.

Assume $\{W_\mu\}_{\mu=0}^{T}$ is Wiener process and $\{\mathcal{F}_\mu\}_{\mu=0}^{T}$ be a filtration. If $\{M(\mu)\}_{\mu=0}^{T}$ is an $\mathcal{F}_\mu$ -adapted Brownian and its a martingale then, there exist $\Gamma(\nu)$ an $\mathcal{F}_\mu$-adapted process with

$$M(\mu) = M(0) + \int_0^t \Gamma(\nu)dW_\nu.$$

In essence every martingale can be represented as an Itō integral. We know

$$D_\mu V(\mu) = \hat{E}[D_\mu V(T)|\mathcal{F}_\mu],$$

is a martingale so it must have an Itō integral.

$$D_{\mu}V(\mu) = V(0) + \int_{0}^{\mu} \Gamma(\nu) d\hat{W}_{\nu} \tag{5.11}$$

Since $X_{\mu} = V(\mu)$ and $X_0 = V(0)$, hence

$$D_{\mu}X_{\mu} = X_0 + \int_{0}^{\mu} \Lambda(\nu)\sigma(\nu)S_{\nu} d\hat{W}_{\nu}, \tag{5.12}$$

and thus by comparing equation 5.11 and 5.12, we obtain

$$\Lambda(\mu) = \frac{\Gamma(\mu)}{\sigma(\mu)S_{\mu}}.$$

So this gives the delta for the hedging.

5.5 Black Scholes Stochastic Differential Equation

Let a portfolio in a market have the expiration time T. Then

$$\gamma(T, S_T) = max\{S_T - k, 0\},$$

gives option value. For $\mu = 0$, the option price is $\gamma(0, S_0)$, to avoid losses, the holder's option price must match the option's Expiry value $X_T = \gamma(T, S_T)$. Suppose we have a simple market. At $\mu = 0$, the option price is $X_0 = \gamma(0, S_0)$ we can use this money to buy more stock and invest the rest in banks with rate r. Continue this process such that at time T we have $X_T = \gamma(T, S_T)$. X_T is the option price (premium) while $\gamma(T, S_T)$ is the portfolio value. This process is called **hedging** the option, where ever this is possible its risk free but in world markets this is not always the case. What is guaranteed is that, in the money markets investment will always generate interest. Examples of this are fixed deposits and bonds.

5.5.1 Black Scholes derivation

Suppose $\Lambda(\mu)$ is the amount of stock held at μ then $\Lambda(\mu)S_\mu$ is the total worth of the stock. Invest $X_\mu - \Lambda(\mu)S_\mu$ in money market with interest rate r to get $r(X_\mu - \Lambda(\mu)S_\mu)$ amount. Then the wealth equation is

$$X_\mu = \Lambda(\mu)S_\mu + r(X_\mu - \Lambda(\mu)S_\mu),$$

and therefore

$$dX_\mu = \Lambda(\mu)dS_\mu + r(X_\mu - \Lambda(\mu)S_\mu)d\mu.$$

Substituting

$$dS_\mu = \alpha S_\mu d\mu + \sigma S_\mu dW_\mu,$$

we have

$$\begin{aligned} dX_\mu &= \Lambda(\mu)[\alpha S_t d\mu + \sigma S_\mu dW_\mu]d\mu + r(X_\mu - \Lambda(\mu)S_\mu)d\mu \\ &= [rX_\mu + (\alpha - r)\Lambda(\mu)S_\mu]d\mu + \sigma\Lambda(\mu)S_\mu dW_\mu. \end{aligned}$$

Consider $\Gamma(\mu, Y) = e^{-r\mu}Y$. Then $\Gamma_{yy} = 0$, applying Itō formula gives

$$d\Gamma(\mu, Y) = (e^{-r\mu}Y)' = \Gamma_\mu(\mu, Y)d\mu + \Gamma_y(\mu, Y)dY,$$

substituting S_μ for Y, yields

$$\begin{aligned} (e^{-r\mu}S_\mu)' &= -re^{-r\mu}S_\mu d\mu + e^{-r\mu}dS_\mu = -re^{-r\mu}S_\mu d\mu + e^{-r\mu}[\alpha S_\mu d\mu + \sigma S_\mu dW_\mu], \\ &= (\alpha - r)e^{-r\mu}S_\mu d\mu + \sigma e^{-r\mu}S_\mu dW_\mu. \end{aligned}$$

Similarly

$$d\Gamma(\mu, X_\mu) = d(e^{-r\mu}X_\mu) = -re^{-r\mu}X_\mu d\mu + e^{-r\mu}dX_\mu,$$

we can now substitute

$$dX_\mu = [rX_\mu + (\alpha - r)\Lambda(\mu)S_\mu]d\mu + \sigma\Lambda(\mu)S_\mu dW_\mu,$$

to get

$$\begin{aligned}
dX_\mu &= -re^{-r\mu}X_\mu d\mu + e^{-r\mu}[rX_\mu d\mu + (\alpha - r)\Lambda(\mu)S_\mu d\mu + \sigma\Lambda(\mu)S_\mu dW_\mu],\\
&= e^{-r\mu}[(\alpha - r)\Lambda(\mu)S_\mu d\mu + \sigma\Lambda(\mu)S_\mu dW_\mu],\\
&= \Lambda(\mu)[(\alpha - r)e^{-r\mu}S_\mu d\mu + \sigma e^{-r\mu}S_\mu dW_\mu],\\
&= \Lambda(\mu)d(e^{-r\mu}S_\mu).
\end{aligned}$$

The above gives the evolution of our portfolio value X_μ. Now we need the evolution of the option-value (or premium) $\gamma(\mu, S_\mu)$. let $\gamma(\mu, S_\mu) \to \gamma(\mu, z)$ in continuous time.

Apply Itō formula on $\gamma(\mu, S_\mu)$, we have

$$d\gamma(\mu, S_\mu) = [\gamma_\mu(\mu, S_\mu) + \alpha S_\mu \gamma_\nu(\mu, S_\mu) + \frac{1}{2}\sigma^2 S_\mu^2 \gamma_{\nu\nu}(\mu, S_\mu)]d\mu + \sigma S_\mu \gamma_\nu(\mu, S_\mu)dW_\mu.$$

Consider $\Gamma(\mu, Z) = e^{-r\mu}Z$, so that $\Gamma_{zz}(\mu, Z) = 0$, hence

$$\begin{aligned}
d(e^{-r\mu}\gamma(\mu, S_\mu)) &= df(\gamma(\mu, S_\mu)) = \Gamma_\mu(\gamma(\mu, S_\mu))d\mu + \Gamma_z(\gamma(\mu, S_\mu))d\gamma(\mu, S_\mu),\\
&= -re^{-r\mu}\gamma(\mu, S_\mu)d\mu + e^{-r\mu}d\gamma(\mu, S_\mu).
\end{aligned}$$

So we have,

$$
\begin{aligned}
d(e^{-rt}\gamma(\mu,S_\mu)) &= e^{-rt}[[C_\mu(t,S_\mu)+\alpha S_\mu C_x(t,S_\mu)+\frac{1}{2}\sigma^2 S_\mu^2 C_{xx}(t,S_\mu)]dt,\\
&+\sigma S_\mu \gamma_z(\mu,S_\mu)dW_\mu]-re^{-rt}\gamma(\mu,S_\mu)dt,\\
&= e^{-rt}[-r\gamma(\mu,S_\mu)+C_\mu(t,S_\mu)+\alpha S_\mu C_x(t,S_\mu)+\frac{1}{2}\sigma^2 S_\mu^2 C_{xx}(t,S_\mu)]dt,\\
&+\sigma S_\mu C_x(t,S_\mu)dW_\mu].
\end{aligned}
$$

Now we need

$$
d(e^{-rt}X_\mu)=d(e^{-rt}\gamma(\mu,S_\mu)),
$$

where $X_0=C(0,S_0)$ and $X_T=\gamma(T,S_T)=\max\{S_T-k,0\}$.

$$
d(e^{-rt}X_\mu)=e^{-rt}\Delta(t)(\alpha-r)S_t dt+\sigma e^{-rt}\Delta(t)S_\mu dW_\mu. \tag{5.13}
$$

$$
\begin{aligned}
d(e^{-r\mu}\gamma(\mu,S_\mu)) = e^{-r\mu}[-r\gamma(\mu,S_\mu)+\gamma_\mu(\mu,S_\mu)+\alpha S_\mu \gamma_z(\mu,S_\mu)+\frac{1}{2}\sigma^2 S_\mu^2 \gamma_{zz}(\mu,S_\mu)]d\mu\\
+\sigma e^{-r\mu}S_\mu \gamma_z(\mu,S_\mu)dW_\mu. \quad (5.14)
\end{aligned}
$$

Comparing the coefficients of dW_μ, in (5.13) and (5.14) to get

$$
\Lambda(\mu)=\gamma_z(\mu,S_\mu),\forall\mu\in[0,T].
$$

This gives a formula for the hedging and $\gamma_z(\mu,S_\mu)$, is the delta of the option, since it helps us calculate how much stock to buy given we know the option price S_μ.

Equating the coefficients of $d\mu$,

$$
\Lambda(\mu)(\alpha-r)S_\mu=-\gamma(\mu,S_\mu)r+\gamma_\mu(\mu,S_\mu)+\alpha S_\mu\gamma_z(\mu,S_\mu)+\frac{1}{2}\sigma^2 S_\mu^2\gamma_{zz}(\mu,S_\mu).
$$

Rearranging and substituting the delta of the hedging, gives

$$\gamma(\mu, S_\mu)r = \gamma_\mu(\mu, S_\mu) + \gamma_z(\mu, S_\mu)S_\mu + \frac{1}{2}\sigma^2 S_\mu^2 \gamma_{zz}(\mu, S_\mu) \quad \forall \mu \in [0, T].$$

In continuous time μ we need $\gamma(\mu, z)$, which satisfies the PDE

$$\gamma(\mu, z)r = \gamma_\mu(\mu, z) + \gamma_z(\mu, z)S_\mu + \frac{1}{2}\sigma^2 S_\mu^2 \gamma_{zz}(\mu, z). \tag{5.15}$$

Where $z \geq 0$ and $\forall \mu \in [0, T]$, under the terminal condition

$$X_T = \gamma(T, z) = \max\{z - k, 0\},$$

where $z = S_T$. The equation in (5.15) is called the Black Scholes equation. This SDE has no closed form solution so to approximate its solution, we can perform a Monte Carlo simulation. However, we have already seen an explicit solution for (5.15), i.e

$$\gamma(\mu, S_\mu) = S_\mu \Phi(\varphi_+(T - \mu, S_\mu)) - ke^{-r(T-\mu)}\Phi(\varphi_-(T - \mu, S_\mu)),$$

and for this, a Nobel price was won in economics [4] [8].

REFERENCES

[1] Bashirov, A. E. (99-05). Partially Observable Linear Systems Under Dependent Noises. *Print, AS*.

[2] Brzeźniak Zdzislaw, & Zastawniak Tomasz (1999). Basic Stochastic Processes. *Springer Verlag*.

[3] Giham I. I., & Skorohod A. V. (1977). Stochastic Differential Equations. *Springer Verlag, New York*.

[4] Karatzas, & I, Shreve, S. E. (1998). Methods of Mathematical Finance. *Springer, Heidelberg*.

[5] Klebaner, F. C. (1998). Introduction to Stochastic Calculus with Applications. *Imperial College Press*.

[6] Lipster, R. S., & Shiryaev, A. N. (1977). Statistics of Random Processes I. General Theory. *Springer Verlag, New York*.

[7] Mikosch Thomas (1998). Elementary Stochastic Calculus with Finance in view. *World Scientific, Singapore New Jersey London*.

[8] Neftci, S. N. (2000). Introduction to the Mathematics of Financial Derivatives. *Academic press*.

[9] Øksendal Bernt (1985). Stochastic Differential Equations an Introduction with Applications *Springer, Berlin*.