MATHEMATICAL FINANCE III

Course Outline

This course is an introduction to modern mathematical finance. Topics include

1. single period portfolio optimization based on the mean-variance analysis, capital asset pricing model, factor models and arbitrage pricing theory.

2. pricing and hedging derivative securities based on a fundamental state model, the well-received Cox-Ross-Rubinstein’s binary lattice model, and the celebrated Black-Scholes continuum model;

3. discrete-time and continuous-time optimal portfolio growth theory, in particular the universal log-optimal pricing formula;

4. necessary mathematical tools for finance, such as theories of measure, probability, statistics, and stochastic process.

Prerequisites

Calculus, Knowledge on Excel Spreadsheet, or Matlab, or Mathematica, or Maple.

Textbooks

Financial Calculus, Martin Baxter & Andrew Rennie.
Xinfu Chen, Lecture Notes, available online www.math.pitt.edu/~xfc.

Recommended References


Grading Scheme

Homework 40%  Take Home Midterms 40%  Final 40%
# Contents

## MATHEMATICAL FINANCE III

1 Stochastic Process  
1.1 Certain Mathematical Tools .............................................. 1  
1.2 Random Walk .............................................................. 4  
1.2.1 Description .............................................................. 4  
1.2.2 Characteristic Properties of a Random Walk ...................... 5  
1.3 Brownian Motion .......................................................... 7  
1.3.1 Brownian Motion as the Limit of Random Walks .................. 7  
1.3.2 A Fourier Series Representation of the Brownian Motion ......... 8  
1.3.3 Brownian Motion on Space of Continuous Functions ............. 10  
1.3.4 Filtration ............................................................... 11  
1.3.5 Vector Valued Brownian Motion ..................................... 12  
1.4 Filtration and Martingale ................................................ 13  
1.4.1 Filtration ............................................................... 13  
1.4.2 Filtration, Partition, and Information .............................. 14  
1.4.3 Conditional Probability .............................................. 17  
1.4.4 Martingale ............................................................. 19  
1.5 Itô Integrals .................................................................... 22  
1.5.1 Stochastic (Itô) Integration ........................................... 23  
1.5.2 Itô's Lemma .............................................................. 24  
1.5.3 Stochastic Differential Equation (SDE) ............................. 25  
1.6 Diffusion Process ........................................................... 27  
1.6.1 Itô diffusion ............................................................. 27  
1.6.2 Semigroup Generator .................................................. 28  
1.6.3 Kolmogorov’s Backward Equation .................................. 28  
1.6.4 Kolmogorov’s Forward or Fokker-Planck equation ............. 30  
1.6.5 The Feynman-Kac Formula .......................................... 30  

2 The Black–Scholes Theory .................................................. 33  
2.1 Replicating Portfolio ....................................................... 33  
2.1.1 Modelling Unit Share Prices of Securities ......................... 33  
2.1.2 Portfolios ............................................................... 34  
2.1.3 Claim and Replicating Strategy ..................................... 35  
2.1.4 Expectation Pricing and Arbitrage-free Pricing ................. 35  
2.2 The Black–Scholes Equation .............................................. 37  
2.3 Solutions to the Black–Scholes Equation .............................. 40
2.3.1 A PDE Method ....................................................... 40  
2.3.2 Probabilistic Method ............................................. 42  
2.3.3 The Risk-Neutral Measure ....................................... 43  
2.4 Examples .............................................................. 44  
  2.4.1 European Options .................................................. 44  
  2.4.2 Foreign Exchange ................................................. 45  
  2.4.3 Equities and Dividends ......................................... 47  
  2.4.4 Quantos ............................................................ 48  
2.5 Binomial Tree Model .................................................. 52  
2.6 Pricing Options ....................................................... 57  
2.7 Replicating Portfolio for Derivative Security ...................... 61  
2.8 A Model for Stock Prices .......................................... 64  
2.9 Continuous Model As Limit of Discrete Model ...................... 67  
2.10 The Black-Scholes Equation ......................................... 71

References 75

Index 76
Chapter 1

Stochastic Process

Stochastic process has been used in mathematical finance to model various kinds of indeterministic quantities such as stock prices, interest rates, etc. It has become a necessary tool for modern theories on finance. In this chapter, we introduce certain elementary theories on the stochastic process for our basic needs in this course.

1.1 Certain Mathematical Tools

Here we introduce necessary mathematical tools from probability that are needed for the study of stochastic process.

A **probability space** is a triple \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\Omega\) is a set, \(\mathcal{F}\) is a \(\sigma\)-algebra on \(\Omega\), and \(\mathbb{P}\) is a probability measure on \((\Omega, \mathcal{F})\).

A **random variable** is a measurable function on a probability space.

A **stochastic process** is a collection \(\{S_t\}_{t \in T}\) of random variables on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\); here \(T\) is a set for time such as \(T = [0, T], T = [0, \infty), T = \{0, 1, 2, \cdots\}\), or \(T = \{t_0, t_1, t_2, \cdots, t_n\}\), etc.

Here by \(\mathcal{F}\) being a \(\sigma\)-algebra on \(\Omega\) it means that \(\mathcal{F}\) is a non-empty collection of subsets of \(\Omega\) that is closed under the operation of compliment and countable union.

Also, by a probability measure, it means that \(\mathbb{P} : \mathcal{F} \rightarrow [0, 1]\) is a non-negative function satisfying \(\mathbb{P}(\Omega) = 1\) and for every countable disjoint sets \(A_1, A_2, \cdots\) in \(\mathcal{F}\),

\[
\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).
\]

Finally, a function \(f : \Omega \rightarrow \mathbb{R}\) is called \(\mathcal{F}\) measurable if

\[
\{\omega \in \Omega \mid f(\omega) < r\} \in \mathcal{F}\quad \forall r \in \mathbb{R}.
\]

The simplest measurable function is the **characteristic function** \(1_A\) of a measurable set \(A \in \mathcal{F}\) defined by

\[
1_A(\omega) = \begin{cases} 
1 & \text{if } \omega \in A, \\
0 & \text{if } \omega \notin A.
\end{cases}
\]
A simple function is a linear combination of finitely many characteristic functions. For a simple function \( \sum_{i=1}^{n} c_i 1_{A_i} \), its integral is defined as

\[
\int_{\Omega} \left( \sum_{i=1}^{n} c_i 1_{A_i}(\omega) \right) \, d\mathbb{P}(\omega) := \sum_{i=1}^{n} c_i \mathbb{P}(A_i).
\]

The integral of a general measurable function is defined as the limit (if it exists) of integrals of an approximation sequence of simple functions.

Two random variables \( X \) and \( Y \) are called equal and write \( X = Y \) if \( \mathbb{P}\{ \omega \in \Omega; \ | \ X(\omega) \neq Y(\omega) \} = 0 \).

In a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), each \( \omega \in \Omega \) is called a sample event, and each \( A \in \mathcal{F} \) is called an observable event with observable probability \( \mathbb{P}(A) \). Similarly, if \( A \) is not measurable, then \( A \) is called a non-observable event.

Let \( \Omega \) be a set and \( S \) be a collection of subsets of \( \Omega \). We denote by \( \sigma(S) \) the smallest \( \sigma \)-algebra that contains \( S \). \( \sigma(S) \) is called the \( \sigma \)-algebra generated by \( S \).

In \( \mathbb{R} \), the algebra \( \mathcal{B} \) generated by all open intervals is called the Borel \( \sigma \)-algebra and each set in \( \mathcal{B} \) is called a Borel set.

If \( X \) is a random variable on \( (\Omega, \mathcal{F}, \mathbb{P}) \) and \( B \) is a Borel set of \( \mathbb{R} \), then

\[
X^{-1}(B) := \{ \omega \in \Omega \mid X(\omega) \in B \}
\]

is a measurable set. In the sequel, we use notation, for every Borel set \( B \) in \( \mathbb{R} \),

\[
\mathbb{P}(X \in B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{ \omega \in \Omega \mid X(\omega) \in B \}).
\]

Note that the mapping:

\[
\mathbb{P}X^{-1} : B \in \mathcal{B} \to \mathbb{P}(X^{-1}(B))
\]

defines a probability measure on \( (\mathbb{R}, \mathcal{B}) \); namely, \( (\mathbb{R}, \mathcal{B}, \mathbb{P}X^{-1}) \) is a probability space.

In the study of a single random variable \( X \), all the properties of the random variable are observed through the probability space \( (\mathbb{R}, \mathcal{B}, \mathbb{P}X^{-1}) \), since \( X \) is experimentally observed by measuring the probability of the outcome \( X^{-1}(B) \) for every \( B \in \mathcal{B} \). Given a random variable \( X \), its distribution function is defined by

\[
F(x) := \mathbb{P}(X \leq x) := \mathbb{P}(X \in (-\infty, x]) := \mathbb{P}(\{ \omega \in \Omega \mid X(\omega) \leq x \}) \quad \forall x \in \mathbb{R}.
\]

The distribution density is defined as the derivative of \( F \):

\[
\rho(x) = \frac{dF(x)}{dx} \quad \forall x \in \mathbb{R}.
\]

A random variable is called \( N(\mu, \sigma^2) \) \( (\mu \in \mathbb{R}, \sigma > 0) \) distributed, or normally distributed with mean \( \mu \) and standard deviation \( \sigma \) (i.e. variance \( \sigma^2 \)) if it has probability density

\[
\rho(\mu, \sigma; x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).
\]
1.1. CERTAIN MATHEMATICAL TOOLS

When $\sigma = 0$, a $N(\mu, 0)$ random variable $X$ becomes a deterministic constant function $X(\omega) \equiv \mu$ for (almost) all $\omega \in \Omega$.

In the sequel, we use $\mathbb{E}$ for expectation and $\text{Var}$ for the variance

\[
\mathbb{E}[X] := \int_{\Omega} X(\omega) \mathbb{P}(d\omega), \\
\text{Var}[X] := \int_{\Omega} \left( X(\omega) - \mathbb{E}(X) \right)^2 \mathbb{P}(d\omega) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.
\]

The Law of the unconscious statistician

Given a random variable $X$ with probability density function $\rho$ and an integrable real function $f$ on $(\mathbb{R}, \mathcal{B})$, the expectation of $f(X)$ is

\[
\mathbb{E}[f(X)] := \int_{\Omega} f(X(\omega)) \mathbb{P}(d\omega) = \int_{-\infty}^{\infty} f(x) \rho(x) \, dx.
\]

As far as a single random variable $X$ is concerned, in certain sense all relevant information in $\mathbb{P}$ is contained in the measure $\mathbb{P}_{X^{-1}}$, i.e. the distribution function $F$. Of course in such a study we lost track of the model underlying the random variable. Nothing is lost just so long as we are interested in one random variable.

If however we have two random variables $X$ and $Y$ on $(\Omega, \mathcal{F}, \mathbb{P})$, then two distribution functions for $\mathbb{P}_{X^{-1}}$ and $\mathbb{P}_{Y^{-1}}$ are not by themselves sufficient to say all there is about $X, Y$ and their interrelation. We need to consider $Z = (X, Y)$ as a map of $\Omega$ into $\mathbb{R}^2$ and define a measure by $\mathbb{P}_{Z^{-1}}$. This measure on the Borel sets of the plane defines what is usually called the joint distribution of $X$ and $Y$, and is sufficient for a complete study of $X, Y$ and their interrelations. This idea of course extends to any finite number of random variables.

The concept of a stochastic process is now a straightforward generalization of these ideas.

Let $\mathbf{T}$ be any index set of points $t$. A stochastic process on $\mathbf{T}$ is a collection of random variables $\{\xi_t\}_{t \in \mathbf{T}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

For each $\omega \in \Omega$, the map $\xi(\omega) : \mathbf{T} \to \mathbb{R}$ defined by $t \to \xi_t(\omega)$ is a function from $\mathbf{T}$ to $\mathbb{R}$, which we call a sample path. Let's denote by $\text{Map}(\mathbf{T}; \mathbb{R})$ the collection of all functions from $\mathbf{T}$ to $\mathbb{R}$. Thus, a stochastic process can be viewed as a function $\xi : \Omega \to \text{Map}(\mathbf{T}; \mathbb{R})$ by the realization $\xi(\omega)(t) = \xi_t(\omega)$ for all $t \in \mathbf{T}$. In many applications, one simply take $\Omega$ as a subset of $\text{Map}(\mathbf{T}; \mathbb{R})$. In such a case, any $\omega \in \Omega$ is a function in $\text{Map}(\omega; \mathbb{R})$ and hence, the function from $\Omega \to \text{Map}(\mathbf{T}; \mathbb{R})$ is realized through the default

\[
\pi(\omega) := \omega(\cdot) \quad \forall t \in \mathbf{T}, \omega \in \Omega \subset \text{Map}(\mathbf{T}; \mathbb{R}).
\]

Note that the probability is built upon $\text{Map}(\mathbf{T}; \mathbb{R})$.

Exercise 1.1. Assume that $f$ is a simple function. Prove the law of unconscious statistician.

Exercise 1.2. Let $\Omega = \{1, 2, 3, 4\}$. Let $\mathcal{F}$ be the smallest $\sigma$-algebra that contains $\{1\}$ and $\{2\}$.

(i) List all the elements in $\mathcal{F}$;
(ii) Is the event $\{1, 2, 3\}$ observable?
Exercise 1.3. For every random variable $X$, show that
\[
\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.
\]

Exercise 1.4. Suppose $X$ is $N(\mu, \sigma^2)$ distributed. Find the following:
\[
\mathbb{E}[X], \quad \text{Var}[X], \quad \mathbb{E}[(X - \mathbb{E}[X])^2], \quad \mathbb{E}[(X - \mathbb{E}[X])^4], \quad \mathbb{E}[e^{i\lambda X}] \quad (\lambda \in \mathbb{C}).
\]

Exercise 1.5. Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions. Show that there exists a smallest $\sigma$-algebra $\mathcal{F}$ on $\mathbb{R}$ such that both $f$ and $g$ are $\mathcal{F}$ measurable. Also show that each element in $\mathcal{F}$ is a Borel set.

Suppose $f(x) = 1$ and $g(x) = 2$ for all $x \in \mathbb{R}$. Describe $\mathcal{F}$.

1.2 Random Walk

Brownian motion is one of the most important building block for stochastic process. To study Brownian motion, we begin with a random walk, a discretized version of the Brownian motion. Roughly speaking,

A random walk is the motion of a particle on a line, which walks in each unit time step a unit space step in one direction or the opposite with probability 1/2 each.

1.2.1 Description

Mathematically, we describe a random walk on the real line by the following steps.

1. Let $X_1, X_2, X_3, \ldots$ be a sequence of independent binomial random variables taking values +1 and −1 with equal probability:
\[
\text{Probability}(X_i = 1) = \frac{1}{2}, \quad \text{Probability}(X_i = -1) = \frac{1}{2}.
\]
Such a sequence can be obtained, for example, by tossing a fair coin, head for 1 and tail for −1.

2. Let $\Delta t$ be the unit time step and $\Delta x$ be the unit space step. Let $W_{i\Delta t}$ be the position of the particle at time $t_i := i\Delta t$. It is a random variable, and can be defined as
\[
W_0 := 0,
\]
\[
W_{i\Delta t} := \sum_{j=1}^{i} X_j \Delta x = W_{[i-1]\Delta t} + X_i \Delta x, \quad i = 1, 2, \ldots.
\]
Here the last equation $W_{i\Delta t} = W_{[i-1]\Delta t} + X_i \Delta x$ means that the particle moves from the position $W_{[i-1]\Delta t}$ at time $t_{i-1} = [i-1]\Delta t$ to a new position $W_{i\Delta t}$ at time $t_i = i\Delta t$ by walking one unit space step $\Delta x$, either to the left or to the right, depending on the choice of $X_i$ being −1 or +1.

3. Though not necessary, it is sometimes convenient to define the position $W_t$ of the particle at an arbitrary time $t \geq 0$. There are jump version and continuous versions. Here we use a constant speed version by a linear interpolation:
\[
W_t = \left( [i + 1] - \frac{t}{\Delta t} \right) W_{i\Delta t} + \left( \frac{t}{\Delta t} - i \right) W_{[i+1]\Delta t} \quad \forall t \in (i\Delta t, [i+1]\Delta t), \quad i = 0, 1, \ldots.
\]
(1.1)
We call $\{W_t\}_{t \geq 0}$ the process of a random walk with time step $\Delta t$ and space step $\Delta x$. 
1.2.2 Characteristic Properties of a Random Walk

The random walk is described by the stochastic process \( \{W_t\}_{t \geq 0} \) defined earlier. Here the probability space for the process is determined by the probability space associated with the random variables \( \{X_i\}_{i=1}^{\infty} \).

A standard probability space \((\Omega, \mathcal{F}, \mathbb{P})\) for a sequence \( \{X_i\}_{i=1}^{\infty} \) of binary random variables can be obtained as follows.

1. First we set
   \[
   \Omega = \{-1, 1\}^\mathbb{N} = \{(x_1, x_2, \cdots) \mid x_i \in \{-1, 1\} \forall i \in \mathbb{N}\}.
   \]
   Each \( \omega = (x_1, x_2, \cdots) \in \Omega \) can be regarded as the record of a sequences of coin tossing where \( x_i \) records the outcome of the \( i \)th toss, \( x_i = +1 \) for head and \( x_i = -1 \) for tail.

2. We define random variables \( X_i \) for each \( i \in \mathbb{N} \) by
   \[
   X_i(\omega) = x_i \quad \forall \omega = (x_1, x_2, \cdots) \in \Omega.
   \]
   Thus, \( X_i(\omega) \) is the \( i \)-th outcome of the event \( \omega \in \Omega \). Note that \( X_i \) takes only two values, \(+1\) and \(-1\).

   We denote
   \[
   \Omega_i^{+1} := \{\omega \in \Omega \mid X_i(\omega) = +1\} = \{(x_1, x_2, \cdots) \in \Omega \mid x_i = +1\},
   \]
   \[
   \Omega_i^{-1} := \{\omega \in \Omega \mid X_i(\omega) = -1\} = \{(x_1, x_2, \cdots) \in \Omega \mid x_i = -1\}.
   \]

3. The \( \sigma \)-algebra \( \mathcal{F} \) on \( \Omega \) is the smallest \( \sigma \)-algebra on \( \Omega \) such that all \( X_1, X_2, \cdots \) are measurable. Clearly, it is necessary and sufficient to define \( \mathcal{F} \) as the \( \sigma \)-algebra generated by the the countable boxes
   \[
   \Omega_i^{+1}, \Omega_i^{-1}, \Omega_2^{+1}, \Omega_2^{-1}, \cdots.
   \]
   Under this \( \sigma \)-algebra, each \( X_i \) is \((\Omega, \mathcal{F})\) measurable.

   For each \( (x_1, \cdots, x_n) \in \{-1, 1\}^n \), the cylindrical box \( c(x_1, \cdots, x_n) \) is defined by
   \[
   c(x_1, \cdots, x_n) := (x_1, \cdots, x_n) \times \prod_{j=n+1}^{\infty} \{-1, 1\} = \{(x_1, \cdots, x_n, y_{n+1}, y_{n+2}, \cdots) \mid y_j \in \{-1, 1\} \forall j \geq n + 1\}.
   \]
   Note that each cylindrical box belongs to \( \mathcal{F} \) since
   \[
   c(x_1, \cdots, x_n) = \bigcap_{i=1}^{n} \{\omega \in \Omega \mid X_i(\omega) = x_i\}.
   \]

4. To define \( \mathbb{P} \) such that both \( X_i = 1 \) and \( X_i = -1 \) has probability \( 1/2 \), we first define \( \mathbb{P} \) on the each cylindrical box by
   \[
   \mathbb{P}(c(x_1, \cdots, x_n)) = 2^{-n} \quad \forall n \in \mathbb{N}, (x_1, \cdots, x_n) \in \{-1, 1\}^n.
   \]
   One can show that such defined \( \mathbb{P} \) on all cylindrical boxes can be extended onto \( \mathcal{F} \) to become a probability measure. Hence, we have a standard probability space \((\Omega, \mathcal{F}, \mathbb{P})\).

5. One can show that under the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), \( \{X_i\}_{i=1}^{\infty} \) is a sequence of i.i.d. random variables having the property that the probability of \( X_i = \pm 1 \) is \( 1/2 \).
The random walk, modelled by the stochastic process \( \{W_t\}_{t \geq 0} \) on \((\Omega, \mathcal{F}, P)\), has the following properties:

1. \( W_0(\omega) \equiv 0 \) for every \( \omega \in \Omega \);
2. \( \mathbb{E}[W(t)] = 0 \) and \( \text{Var}[W(t)] = \sigma t \) for every \( t \in T \), where
\[
T = \{i \Delta t\}_{i=0}^\infty, \quad \sigma = \frac{(\Delta x)^2}{\Delta t}.
\]
3. For every \( t_0, t_1, t_2, \ldots, t_n \in T \) with \( 0 = t_0 < t_1 < \cdots < t_n \), the following random variables are independent:
\[
W(t_n) - W(t_{n-1}), \; W(t_{n-1} - t_{n-2}), \; \cdots, \; W(t_1) - W(t_0).
\]
4. For every \( \omega \in \Omega \), the function \( t \in [0, \infty) \rightarrow W_t(\omega) \in \mathbb{R} \) is a continuous function.

Exercise 1.6. From the definition, show that
\[
P(\{\omega \in \Omega \mid X_i(\omega) = 1\}) = \frac{1}{2},
\]
\[
\text{Prob}(X_i \in A, X_j \in B) = \text{Prob}(X_i \in A) \text{Prob}(X_j \in B) \quad \forall i \neq j.
\]

Exercise 1.7. Show that the random walk has the above four listed properties.

Exercise 1.8. Show that for every positive integer \( n \), \( W_n \Delta x \) has the range \( \{k \Delta x\}_{k=-n}^n \) and
\[
\text{Prob}(W_n \Delta t = k \Delta x) = \frac{C^k_n}{2^n}, \quad C^k_n := \frac{k!(n-k)!}{n!}.
\]

Exercise 1.9. Let \( n \) be a positive integer and set \( \Delta t = 1/n \) and \( \Delta x = 1/\sqrt{n} \). Let \( \Omega_n = \{-1,1\}^n \), \( \mathcal{F}_n = 2^{\Omega_n} \) (the collection of all subsets of \( \Omega_n \)) and
\[
P(\omega) = \frac{1}{2^n} \quad \forall \omega \in \Omega_n.
\]

1. Show that \( (\Omega_n, \mathcal{F}_n, P_n) \) is a probability space.
2. In time interval \([0,1]\), show that there are a total of \( 2^n \) samples random walks. We denote by \( W^1, \cdots, W^{2^n} \) all the samples random walks.
3. Denote by \( C([0,1]) = C([0,1]; \mathbb{R}) \) the space of all continuous functions from \([0,1]\) to \( \mathbb{R} \). Denote by \( \mathcal{B} \) the Borel algebra on \( C([0,1]) \) generated by all open sets under the norm
\[
\|x\| = \sup_{t \in [0,1]} |x(t)| \quad \forall x \in C([0,1]).
\]

On \( C([0,1]) \) we define
\[
P(\{W^i\}) = \frac{1}{2^n} \quad \forall i = 1, \cdots, 2^n,
\]
\[
P(C([0,1]) \setminus \{W^1, \cdots, W^{2^n}\}) = 0.
\]

Show that there exists an extension of \( P \) on \( \mathcal{B} \) such that \( (C([0,1]); \mathcal{B}, P) \) is a probability space.
[For each \( B \in \mathcal{B} \), one can define \( P(B) \) as the number of random walks contained in \( B \).]
4. Show that the map \( \pi \) from \( \omega \) to the function \( t \rightarrow W_t(\omega) \) lift the probability space \( (\Omega_n, \mathcal{F}_n, \mathbb{P}_n) \) to the probability space \( (C([0, 1]), \mathcal{B}, \mathbb{P}) \); that is, prove the following:

(a) for every \( A \in \mathcal{F}_n \), \( \pi(A) \in \mathcal{B} \) and \( \mathbb{P}_n(A) = \mathbb{P}(\pi(A)) \);

[Notice that \( A \) is a finite set so \( \pi(A) \) is also a finite set, and hence is a closed set in \( C([0, 1]) \) which is of course a Borel set.]

(b) for every \( B \in \mathcal{B} \), \( \pi^{-1}(B) \in \mathcal{F}_n \) and \( \mathbb{P}(B) = \mathbb{P}_n(\pi^{-1}(B)) \).

[Count how many random walks are in the set \( B \). Be careful about the definition of \( \mathbb{P} \) on \( \mathcal{B} \).]

5. When \( n = 6 \), calculate the probability

\[
\text{Prob}(W_1 = 0, W_{1/2} = 0).
\]

1.3 Brownian Motion

The Brownian motion was first described as “continuous swarming motion” by the English botanist Robert Brown in studying the motion of small pollen grains immersed in a liquid medium in 1827. Albert Einstein in 1905 showed that the swarming motion, now called Brownian motion, could be the consequence of the continual bombardment of particles by the molecules of liquid. A formal mathematical description of the Brownian motion and its properties was first given by Norbert Wiener in 1918. It is especially interesting to note that the Brownian motion was used by the French mathematician Bachelier to model stock prices, for his doctoral dissertation in 1900!

It was nearly a century after Brown first observed microscopic particles zigzagging that the mathematical model for their movement was properly developed. There are many mathematical ways to model the Brownian motion. Here, we provide three mathematical models.

1.3.1 Brownian Motion as the Limit of Random Walks

Fix a positive integer \( n \). We denote by \( \{W^n_t\}_{t \geq 0} \) the random walk obtained by taking

\[
\Delta t = \frac{1}{n}, \quad \Delta x = \frac{1}{\sqrt{n}} = \sqrt{\Delta t}.
\]

Fix \( t > 0 \). Consider the sequence of random variables \( \{W^n_t\}_{t=1}^{\infty} \). Denote by \( [nt] \) the largest integer no bigger than \( nt \). We have, by (1.1) with \( i = [nt] \),

\[
W^n_t = \left( [nt] + 1 - nt \right) W^n_{[nt]} + \left( nt - [nt] \right) W^n_{[nt]+1/n} + \frac{1}{\sqrt{n}} \sqrt{\Delta t} \sum_{i=1}^{[nt]} X_i + \frac{(nt - [nt])}{\sqrt{n}} X_{[nt]+1}
\]

\[
= \sqrt{n} \frac{[nt]}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i + \frac{(nt - [nt])}{\sqrt{n}} X_{[nt]+1}.
\]

To consider the limit as \( n \to \infty \), we recall the celebrated central limit theorem.
If \( \{X_i\}_{i=1}^{\infty} \) is a sequence of i.i.d (independent identically distributed) random variables with mean \( \mu \) and variance \( \sigma^2 \), then in probability, the random variable
\[
\frac{\sum_{i=1}^{m} X_i - m\mu}{\sqrt{m\sigma^2}}
\]
approaches, as \( m \to \infty \), a random variable that is \( N(0,1) \) distributed.

Using the central limit theorem, we see that there exists a random variable \( B_t \) such that
\[
\lim_{n \to \infty} W^n_t = B_t \quad \text{in probability.}
\]
Here limit in probability means
\[
\lim_{n \to \infty} \Pr(W^n_t \in (a,b)) = \Pr(B_t \in (a,b)) \quad \forall (a,b) \subset \mathbb{R}.
\]
In addition, \( B_t \) is \( N(0, t) \) distributed.

The stochastic process \( \{B_t\}_{t \geq 0} \) has the following properties:

1. \( B_0 = 0 \);
2. For each \( t \geq 0 \), \( B_t \) is a \( N(0, t) \) distributed random variable;
3. For each positive integer \( k \) and each \( 0 = t_0 < t_1 < t_2 < \cdots < t_k \), the following random variables are independent:
\[
B_{t_0} - B_{t_1}, \quad B_{t_2} - B_{t_1}, \quad \cdots, \quad B_{t_{k-1}} - B_{t_k}.
\]

We shall henceforth call a stochastic process \( \{B_t\}_{t \geq 0} \) a Brownian motion if it has the above stated properties.

We remark that the limit \( W^n_t \to B_t \) is only in probability, not pointwise. Hence, it is a mathematical nightmare to see what the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) that the Brownian motion is defined on. So far we shall be content with what we have obtained.

### 1.3.2 A Fourier Series Representation of the Brownian Motion

Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space on which there exists a sequence \( \{\xi_n\}_{n=0}^{\infty} \) of i.i.d random variables which are \( N(0,1) \) distributed. Consider the stochastic process \( \{B_t\}_{t \in [0,\pi/2]} \) defined by
\[
B_t(\omega) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \xi_n(\omega) \frac{\sin((2n+1)t)}{2n+1} \quad \forall \omega \in \Omega, \quad \forall t \in [0,\pi/2).
\]
One can show that the series is convergent for \( \mathbb{P} \)-almost all \( \omega \in \Omega \). We claim that \( \{B_t\}_{t \in [0,\pi/2]} \) is a Brownian motion on \( (\Omega, \mathcal{F}, \mathbb{P}) \).

For this purpose, let \( 0 < t_1 < \cdots < t_m < \pi/2 \) be arbitrary times in \( (0,\pi/2) \). We consider the joint distribution of the \( m \)-dimensional random variable \( Z := (B_{t_1}, B_{t_2}, \cdots, B_{t_m}) \). We want to show that \( Z \) has a Gaussian distribution with mean \( (0, \cdots, 0) \) and covariance matrix
\[
\text{Cov}(B_{t_i}, B_{t_j}) := \mathbb{E}[(B_{t_i} - 0)(B_{t_j} - 0)] = \min\{t_i, t_j\}.
\]
Suppose this is proven. Then

(i) $\text{Var}(B_t) = t$ for all $t \in (0, \pi/2)$ so as $t \searrow 0$, $B_t \to 0 = B_0$ in measure;

(ii) $B_t$ is $N(0, t)$ distributed;

(iii) $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_m} - B_{t_{m-1}}$ are independent.

Thus, $\{B_t\}_{t \in [0, \pi/2]}$ is a Brownian motion.

Indeed, an alternative definition for a stochastic process $\{B_t\}_{t \geq 0}$ to be a Brownian motion is

for every $0 < t_1 < \cdots < t_m$, the joint distribution of $(B_{t_1}, \ldots, B_{t_m})$ is Gaussian with mean vector $(0, \ldots, 0)$ and covariance matrix $(\text{min}\{t_i, t_j\})_{m \times m}$.

To show that $Z = (B_{t_1}, \ldots, B_{t_m})$ is Gaussian distributed, we need only show that the characteristic function $\psi(\lambda_1, \ldots, \lambda_m)$ of $Z$ is the characteristic function of the required Gaussian distribution. Here the characteristic function of a random variable is defined as the Fourier transform of the density function of the Random variable. Thus if we let $\rho(x_1, \ldots, x_m)$ be the joint distribution of $Z$, then the characteristic function $\psi(\lambda_1, \ldots, \lambda_m)$ of $Z$ is

$$
\psi(\lambda_1, \ldots, \lambda_m) := \int_{\mathbb{R}^m} e^{i \sum_{j=1}^m \lambda_j x_j} \rho(x_1, \ldots, x_m) \, dx_1 \cdots dx_m
$$

$$
= \mathbb{E}\left[ e^{i \sum_{j=1}^m \lambda_j B_{t_j}} \right]
$$

$$
= \prod_{n=0}^{\infty} \exp \left( i \xi_n \sum_{j=1}^m \frac{2 \lambda_j \sin([2n + 1] t_j)}{\sqrt{\pi}(2n + 1)} \right)
$$

$$
= \prod_{n=0}^{\infty} \mathbb{E} \left[ \exp \left( i \xi_n \sum_{j=1}^m \frac{2 \lambda_j \sin([2n + 1] t_j)}{\sqrt{\pi}(2n + 1)} \right) \right]
$$

since all $\xi_0, \xi_1, \ldots$ are independent random variables.

Since $\xi_n$ is $N(0,1)$ distributed, we have

$$
\mathbb{E}[e^{i \lambda \xi_n}] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i \lambda x - x^2/2} \, dx = e^{-\lambda^2/2} \quad \forall \lambda \in \mathbb{C}.
$$

It then follows that

$$
\psi(\lambda_1, \ldots, \lambda_m) = \prod_{n=1}^{\infty} \exp \left( - \frac{1}{2} \sum_{j=1}^m \frac{2 \lambda_j \sin([2n + 1] t_j)}{\sqrt{\pi}(2n + 1)} \right)
$$

$$
= \prod_{n=1}^{\infty} \exp \left( - \frac{1}{2} \sum_{i,j=1}^m \frac{4 \lambda_i \lambda_j \sin([2n + 1] t_j) \sin([2n + 1] t_i)}{\pi(2n + 1)^2} \right)
$$

$$
= \exp \left( - \frac{1}{2} \sum_{i,j=1}^m \sigma_{ij} \lambda_i \lambda_j \right)
$$

where

$$
\sigma_{ij} := \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin([2n + 1] t_j) \sin([2n + 1] t_i)}{(2n + 1)^2}
$$

$$
= \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1 - \cos([2n + 1] t_i + t_j)}{(2n + 1)^2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\cos([2n + 1] t_i - t_j) - 1}{(2n + 1)^2}
$$

$$
= \frac{1}{2} |t_i + t_j| - \frac{1}{2} |t_i - t_j| = \min\{t_i, t_j\}.
$$
Here in the last equation, we have used the identity
\[ |t| = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1 - \cos([2n+1]t)}{(2n+1)^2} \quad \forall t \in (-\pi, \pi). \]

To see this, we first use the Fourier sine series expansion for the constant function 1 to obtain
\[ 1 = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin([2n+1]s)}{(2n+1)} \quad \forall s \in (0, \pi). \]

Integrating both sides on \( s \in (0, t) \) then gives required identity.

### 1.3.3 Brownian Motion on Space of Continuous Functions

Quite often people use Wiener process as a synonym for the Brownian motion. Though Wiener process and Brownian motion are the same thing, deep mathematical theories are associated with the Wiener process. Here we introduce two important aspects of the Wiener process: continuous sample path and filtration.

For each fixed \( \omega \in \Omega \), the function \( t \in [0, \infty) \rightarrow B_t(\omega) \) is called a sample path. Ideally, one would like to build a stochastic process upon sample paths. Hence, we would like, if possible, to relate our probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with the collection of sample paths. It is said that Brownian motion has continuous sample paths; this means that for almost all \( \omega \in \Omega \), the function \( t \in [0, \infty) \rightarrow B_t(\omega) \) is continuous in \( t \). If we delete that measure zero set from \( \Omega \), then every sample path is continuous.

In our definition, \( \mathbb{E}[\cdot] \) is understood as the expectation derived from the measure space \( (\Omega, \mathcal{F}, \mathbb{P}) \) on which each individual random variable \( B_t(\cdot) : \Omega \rightarrow \mathbb{R} \) is defined. One fundamental question is the very existence of such a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) for which each sample path \( t \rightarrow B_t(\omega) \) is continuous. Clearly, a natural probability space is to take directly \( \Omega \) as the space of continuous functions. Here we briefly describe how a Wiener process can be built upon such a space.

Let \( T > 0 \) be a fixed time. The Brownian motion can be defined on the space of continuous functions
\[ \Omega_T := C([0, T]). \]

Here \( C([0, T]) = C([0, T]; \mathbb{R}) \) denotes the set of all continuous functions from \( [0, T] \) to \( \mathbb{R} \).

For each \( \omega \in \Omega \), the value \( B_t(\omega) \) is defined as
\[ B_t(\omega) = \omega(t) \quad \forall \omega \in \Omega_T \quad \forall t \in [0, T]. \quad (1.2) \]

One can show that there is a (minimum) \( \sigma \)-algebra \( F_T \) on \( \Omega_T \) and a probability measure \( \mathbb{P}_T \) on \( (\Omega_T, F_T) \) such that the so defined \( \{ B_t \}_{t \in [0, T]} \) is a Brownian motion on \( (\Omega_T, F_T, \mathbb{P}_T) \).

In this definition, we see that for each \( \omega \in \Omega \), the function \( t \in [0, T] \rightarrow B_t(\omega) \) is indeed exactly the function \( t \in [0, T] \rightarrow \omega(t) \). Since \( \omega \in \Omega_T \) is continuous, we see that every Brownian motion sample path \( \{(t, B_t(\omega))\}_{0 \leq t \leq T} = \{(t, \omega(t))\}_{0 \leq t \leq T} \) is a continuous curve on the time-space coordinate system.

We remark that \( C([0, T]) \) is a Banach space under the norm
\[ ||\omega|| = \max_{t \in [0, T]} |\omega(t)| \quad \forall \omega \in \Omega_T = C([0, T]). \]
We know a Banach space admits a default $\sigma$-algebra—the Borel $\sigma$-algebra, being the $\sigma$-algebra generated by all open sets. The $\sigma$-algebra $F_T$ for the Brownian motion is indeed the default Borel algebra. Thus, under the above norm, every open set is in $F_T$.

Consider the set

$$A := \{ \omega \in C([0, T]) \mid \omega(0) = 0 \};$$

$$B := \{ \omega \in C([0, T]) \mid \omega(0) \neq 0 \}.$$

Then $A$ is a closed set and $B$ is an open set. From the property of Brownian motion, we see that $P_T(A) = 1$, $P_T(B) = 0$.

One can show more regularity (smoothness) of Brownian motion paths. For this, we introduce the Hölder space. For each $\alpha \in (0, 1]$, we define the norm

$$\|\omega\|_{C^\alpha([0, T])} = \max_{t \in [0, T]} |\omega(t)| + \sup_{0 \leq s < t \leq T} \frac{|\omega(t) - \omega(s)|}{|t - s|^{\alpha}} \quad \forall \omega \in C([0, T]).$$

We set

$$C^\alpha([0, T]) := \{ \omega \in C([0, T]) \mid \|\omega\|_{C^\alpha([0, T])} < \infty \}.$$

One can show that $C^\alpha([0, 1])$ is a closed set in $C([0, T])$ and is $F_T$ measurable. More importantly, we have the following zero-one law.

Almost every sample path of the Brownian motion is Hölder continuous with exponent $\alpha \in [0, 1/2]$; namely,

$$P_T(C^\alpha([0, T])) = 1 \quad \forall \alpha \in [0, 1/2).$$

For every $\alpha \in (1/2, 1]$, almost none of the Brownian motion sample path is Hölder continuous with exponent $\alpha \in (1/2, 1]$; namely,

$$P_T(C^\alpha([0, T])) = 0 \quad \forall \alpha \in (1/2, 1].$$

Since $C^1([0, 1])$ is traditionally used to denote the space of contnuously differentiable functions, when $\alpha = 1$, the Hölder space with exponent $\alpha = 1$ is denoted by $C^{0,1}([0, T])$, which is indeed the space of all Lipschitz continuous functions. Clearly, a differentiable function is also Hölder continuous for every exponent $\alpha \in [0, 1]$. The above statement shows that no Brownian motion sample path is everywhere differentiable. Indeed, one can shows that, except a set of measure zero, every Brownian motion sample path is nowhere differentiable.

### 1.3.4 Filtration

Now we set

$$\Omega = C([0, \infty)).$$

For each $\omega \in \Omega$, let $\omega_{[0,T]}$ be the restriction of $\omega$ on $[0, T]$; then $\omega \rightarrow \omega_{[0,T]}$ is a projection from $\Omega$ to $\Omega_T$. This projection allows us to lift $F_T$ to a $\sigma$-algebra $\mathcal{F}_T$ on $\Omega$ and lift $P_T$ to a probability measure $P$ on $(\Omega, \mathcal{F}_T)$ as follows.
For each set $A$ in $F_T$, we define a cylinder $c_T(A)$ in $\Omega$ by

$$c_T(A) = \{ \omega \in \Omega \mid \omega_{[0,T]} \in A \} \quad \forall A \in F_T.$$ 

Also, we define

$$\mathcal{F}_T := \{ c_T(A) \mid A \in F \} \quad \forall T \geq 0,$$

$$P(c_T(A)) := P_T(A) \quad \forall A \in F_T, T \geq 0.$$ 

Then $\mathcal{F}_T$ is a $\sigma$-algebra on $\Omega$ and $P$ is a probability measure on $(\Omega, \mathcal{F}_T)$. It is easy to see that

(i) $F_t \subset F_s$ if $0 < s < t$;

(ii) $B_t$ is $\mathcal{F}_t$ measurable.

(iii) $P$ is a probability measure on $(\Omega, \mathcal{F}_t)$ for every $t \geq 0$.

Finally, we define

$$\mathcal{F}_\infty = \bigcup_{t \geq 0} \mathcal{F}_t.$$ 

We then have a Wiener process $\{B_t\}_{t \geq 0}$ defined on $(\Omega, \mathcal{F}_\infty, P)$. For this we introduce the following important concept.

A filtration on a set $\Omega$ is family of $\sigma$-algebra $\{\mathcal{F}_t\}_{t \geq 0}$ on $\Omega$ such that

$$\mathcal{F}_s \subset \mathcal{F}_t \quad \forall 0 \leq s < t.$$ 

A stochastic process $\{S_t\}_{t \geq 0}$ on $(\Omega, \mathcal{F}_\infty, P)$ is called adapted to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if $\mathcal{F}_\infty = \cup_{t \geq 0} \mathcal{F}_t$ and for each $t \geq 0$, $S_t$ is a random variable on the probability space $(\Omega, \mathcal{F}_t, P)$.

### 1.3.5 Vector Valued Brownian Motion

If $\{B_t\}$ is a standard Brownian motion and $a \in \mathbb{R}$, $\{a + B_t\}$ is often called Brownian motion started at $a$. If $B^1, \ldots, B^n$ are $n$ independent standard Brownian motions, then the process $B_t = (B^1_t, \ldots, B^n_t)$ is called an $n$-dimensional Brownian motion, which can be realized on $C(T, \mathbb{R}^n)$. If $A = (a_{ki})_{n \times m}$ is a matrix, then the process $\hat{B}_t = B_t A = (\hat{B}^1_t, \ldots, \hat{B}^m_t)$ is a general $m$ dimensional Brownian motion where

$$\hat{B}^i_t = \sum_{k=1}^{m} B^k_t a_{ki} \quad \forall i = 1, \ldots, m.$$ 

Note that the covariance

$$\sigma_{ij} := \frac{1}{t} \text{Cov}(\hat{B}^i_t, \hat{B}^j_t) = \frac{1}{t} \text{Cov}(B^i_t, B^j_t) = \frac{1}{t} \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ki} a_{lj} \text{Cov}(B^k_t, B^l_t)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ki} a_{lj} \delta_{kl} = \sum_{k=1}^{n} a_{ki} a_{kj} = (A^T A)_{ij}$$

where $A^T$ is the transpose of $A$ and $(A^T A)_{ij}$ represents the entry on $i$th row $j$th column. Hence, we say $\{\hat{B}_t\}$ is an $m$-dimensional Brownian motion with covariance matrix $C = (\sigma_{ij})_{m \times m} = A^T A$. Every vector valued Brownian motion is characterized by its mean vector and covariance matrix.

Many books and countless research papers have been written about Brownian motion. A few introductory references are (in increasing order of difficulty) [8], [22], and [14].
Exercise 1.10. Suppose $0 < s < t$, show that $Z = (B_t, B_s)$ has distribution density

$$
\rho(x, t; y, s) = \frac{e^{-\frac{(x-y)^2}{2\sigma^2}}}{\sqrt{2\pi(t-s)}} \times \frac{e^{-\frac{y^2}{2\sigma^2}}}{\sqrt{2\pi s}}.
$$

Exercise 1.11. From the basic property of the Brownian motion listed in subsection 1.3.1 show that for every $0 < t_1 < \cdots < t_n$, the random variable $Z := (B_{t_1}, \cdots, B_{t_n})$ is Gaussian distributed with mean $(0, \cdots, 0)$ and covariance matrix $(\min\{t_i, t_j\})_{m \times m}$.

Hint: It is equivalent to show that the characteristic function $\psi(\lambda) := \mathbb{E}(e^{i\lambda Z})$ is given by

$$
\psi(\lambda_1, \cdots, \lambda_m) = e^{-\frac{1}{2} \sum_{i,j=1}^m \sigma_{ij} \lambda_i \lambda_j}, \quad \sigma_{ij} = \min\{t_i, t_j\}.
$$

Exercise 1.12. Suppose $Z$ is a normally distributed random variable. Consider the process $\{X_t\}_{t \geq 0}$ defined by $X_t = tZ$. Show that for each $t \geq 0$, $X_t$ is $N(0, t)$ distributed. Explain that $\{X_t\}_{t \geq 0}$ is not a Brownian motion.

Exercise 1.13. Suppose $\{B_t\}_{t \geq 0}$ is Brownian motion.

(i) Show that for each $t_2 > t_1 \geq 0$, $B_{t_2} - B_{t_1}$ is $N(0, t_2 - t_1)$ distributed. Also, show that for each $t_0 \geq 0$, the stochastic process $\{B_{t+t_0} - B_{t_0}\}_{t \geq 0}$ is also a Brownian motion.

(ii) Show that for each $a > 0$ and $t > 0$, $\frac{1}{\sqrt{a}}B_{at}$ is $N(0, t)$ distributed. Defining $\tilde{B}_t := \frac{1}{\sqrt{a}}B_{at}$, show that $\{\tilde{B}_t\}_{t \geq 0}$ is a Brownian motion.

(iii) Are $\{B_t\}$ and $\{\tilde{B}_t\}$ independent?

Exercise 1.14. Suppose $\{B^1_t\}_{t \geq 0}$ and $\{B^2(t)\}_{t \geq 0}$ are two independent Brownian motions and $\theta \in [0, 2\pi]$ is a constant. Show that $\{B^1_t \cos \theta + B^2_t \sin \theta\}_{t \geq 0}$ is a Brownian motion.

Exercise 1.15. For the Brownian motion process $\{B_t\}$, show that $(dB)^2 = dt$ in the sense that

$$
\lim_{h \searrow 0} \mathbb{P}\left( \left| \frac{(B_{t+h} - B_t)^2}{h} - 1 \right| > \varepsilon \right) = 0 \quad \forall \varepsilon > 0.
$$

[Calculate $\mathbb{E}((B_{t+h} - B_t)^2 - h)$ where $B_{t+h} - B_t$ has distribution density $(2\pi h)^{-1}e^{-x^2/(2h^2)}$.]

Exercise 1.16. Show that $\lim_{t \searrow 0} B_t = 0$ in measure; that is, show that

$$
\lim_{t \searrow 0} \mathbb{P}(|B_t| > \varepsilon) = 0 \quad \forall \varepsilon > 0.
$$

[ Use the inequality $\mathbb{E}[|B_t|] \geq \varepsilon \mathbb{P}(|B_t| > \varepsilon)$ and calculate $\mathbb{E}[|B_t|]$ using the fact that $B_t$ is $N(0, t)$ distributed.]

Exercise 1.17. Show that for every $\beta > 0$, there exist $\alpha > 0$ and $C > 0$ such that

$$
\sup_{t > s \geq 0} \frac{\mathbb{E}[|B_t - B_s|^\beta]}{|t - s|^\alpha} \leq C.
$$

1.4 Filtration and Martingale

1.4.1 Filtration

We recall the following
A filtration on a state space $\Omega$ is a collection $\{F_t\}_{t \geq 0}$ of $\sigma$-algebras on $\Omega$ that is indexed by time and satisfies $F_s \subset F_t$ for all $s < t$.

A stochastic process $\{X_t\}_{t \geq 0}$ defined on $(\Omega, F_\infty, \mathbb{P})$ is called adapted to a filtration $\{F_t\}_{t \geq 0}$ if $F_\infty = \bigcup_{t \geq 0} F_t$ and for each $t \geq 0$, $X_t(\cdot)$ is a random variable on the probability space $(\Omega, F_t, \mathbb{P})$.

Note that if $\mathbb{P}$ is a probability measure on each $F_t = \bigcup_{0 \leq \tau \leq t} F_\tau$, then by setting $\mathcal{F}_\infty := \bigcup_{t \geq 0} F_t$ every needed probability (involving finite time) can be calculated. Namely, it is not necessary to extend $\mathbb{P}$ from $\mathcal{F}_\infty$ to $\sigma(\mathcal{F}_\infty)$ (the smallest $\sigma$-algebra containing $\mathcal{F}_\infty$), though it can be done (since $\mathcal{F}_\infty$ is an algebra).

Given a stochastic process $\{S_t\}_{t \in [0,T]}$ ($T \in (0, \infty]$) on probability space $(\Omega, F, \mathbb{P})$, we can define a filtration $\mathcal{F}_t$ as follows:

For each $t \geq 0$, $\mathcal{F}_t$ is defined as the smallest $\sigma$-algebra on $\Omega$ such that each $S_s$, $s \in [0,t]$, is $\mathcal{F}_t$ measurable.

It is easy to see that $\{\mathcal{F}_t\}_{t \in [0,T]}$ is a filtration and $\{S_t\}_{t \in [0,T]}$ is adapted to $\{\mathcal{F}_t\}_{t \in [0,T]}$. Such a filtration is called the natural filtration of the process.

In the sequel, all Wiener process are assumed to be adapted to its natural filtration.

### 1.4.2 Filtration, Partition, and Information

A filtration is usually taken to represent the flow of information; that is $\mathcal{F}_t$ is used to accommodate the degree of detail on information that can be revealed up to time $t$. In application $\mathcal{F}_t$ is designed so it cannot resolve any more detailed information revealed after time $t$.

Consider the situation that the state space $\Omega$ is finite. Then an $\sigma$-algebra $\mathcal{F}$ can be characterized by its atoms; here a set $A \in \mathcal{F}$ is called an atom if $A$ does not contain any proper subset in $\mathcal{F}$, i.e. if $B \subset A$ and $B \in \mathcal{F}$, then either $B = A$ or $B = \emptyset$.

Thus on a finite state space, there is a one to one correspondence between $\sigma$-algebras and partitions. The finer the partition, the smaller the atoms, and the larger the $\sigma$-algebra. Hence, if $\mathcal{F}_s \subset \mathcal{F}_t$, then $\mathcal{F}_s$ has coarser partition than $\mathcal{F}_t$, in other words $\mathcal{F}_t$ has finer partition than $\mathcal{F}_s$.

Taking an example of locating an address of a person, a coarse partition only resolve the detail up to, say city, whereas a finer partition many have a resolution up to street, or even street numbers.

We work on a few examples.

1. Consider the game of tossing coins. We let $N$ be the total number of coin toss. Denote by $+1$ for head and $-1$ for tail. We set

\[
\Omega := \{1, -1\}^N, \quad \Omega_i := \{1, -1\}^i \quad \forall \ i = 1, \cdots, N, \\
F_i = 2^{\Omega_i} \times (-1, 1)^{N-i} = \{\omega \times \{1, -1\}^{N-i} : \omega \in \Omega_i\},
\]

\[
c(x_1, \cdots, x_i) := (x_1, \cdots, x_i) \times \{1, -1\}^{N-i} \\
= \{(x_1, \cdots, x_i, y_{i+1}, \cdots, y_N) : y_j \in \{1, -1\} \forall j \geq i+1 \} \quad \forall (x_1, \cdots, x_i) \in \Omega_i.
\]

Here we use convention $\emptyset \times \{-1, 1\}^{N-i} = \emptyset$ so $\mathcal{F}_i$ is a $\sigma$-algebra. One can show that $\{F_i\}_{i=1}^N$ is a filtration on $\Omega$ for $T := \{1, \cdots, N\}$.
Note that for each \((x_1, \cdots, x_i) \in \Omega_i\), the cylinder \(c(x_1, \cdots, x_i)\) is an atom of \(\mathcal{F}_i\); that is to say, if \(B\) is a proper subset of \(c(x_1, \cdots, x_i)\) and \(B \in \mathcal{F}_i\), then \(B = \emptyset\).

Now suppose a function \(X\) is \(\mathcal{F}_i\) measurable. Then, \(X(x_1, x_2, \cdots, x_i, y_{i+1}, \cdots, y_n)\) is independent of \(y_j\) for every \(j \geq i + 1\). Indeed, since \(b(x_1, \cdots, x_i)\) is an atom of \(\mathcal{F}_i\), and \(X\) is \(\mathcal{F}_i\) measurable, \(X\) has to be a constant function on the set \(c(x_1, \cdots, x_i)\). Hence, if \(X\) is an \(\mathcal{F}_i\) measurable function, then there exists a function \(\tilde{X}\) defined on \(\Omega_i\) such that

\[
X(x_1, \cdots, x_N) = \tilde{X}(x_1, \cdots, x_i) \quad \forall (x_1, \cdots, x_N) \in \Omega.
\]

Now suppose \(X\) is a random variable on \((\Omega, \mathcal{F}_N)\) and for each sequence of outcome \((x_1, \cdots, x_N) \in \Omega\), \(X(x_1, \cdots, x_N)\) represents the award that once can collect from a particular gambling contract. If \(X\) is \(\mathcal{F}_i\) measurable, then one knows the award after the \(i\)th tossing of the coin; namely, there is no need to wait for the results of all \(N\) coin tossing to find out the award. This is so since \(X(x_1, \cdots, x_N) = \tilde{X}(x_1, \cdots, x_i)\); as long as the results \(x_1, \cdots, x_i\) of the first \(i\)th coin tossing are known, the value \(X(x_1, \cdots, x_N)\) is also known. For example, suppose \(X\) represent the award of $8.00 for the first three consecutive heads, and lost $1.00 otherwise. Then as long as the result of the first three coin toss are revealed, the game with agreement \(X\) can be considered as finished, since the payment is clear.

On the other hand, if \(X\) is not \(\mathcal{F}_{N-1}\) measurable, for example,

\[
X(x_1, \cdots, x_N) = \sum_{i=1}^{N} (2x_i - 1),
\]

then one has to find the results of the final tossing before knows the exact amount of award.

In summary, for this example, by saying that \(X: \Omega \to \mathbb{R}\) is \(\mathcal{F}_i\) measurable, it means that there exists a function \(\tilde{X}: \Omega_i \to \mathbb{R}\) such that \(X(x_1, \cdots, x_N) = \tilde{X}(x_1, \cdots, x_i)\) for all \((x_1, \cdots, x_N) \in \Omega\); that is, \(X\) does not dependent on the variables \((x_{i+1}, \cdots, x_N)\).

2. Consider the Brownian motion \(\{B_t\}_{t \geq 0}\) adapted to the filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) that we defined in the earlier section. Given an event \(z \in \Omega = C([0, \infty); \mathbb{R})\), we have \(B_t(z) = z(t)\); thus no matter what value \(z(t + h)\) is, as long as \(z(t) = x\), we always have \(B_t(z) = x\). On the other hand, if we divide the set \(A := \{z \in \Omega \mid z(t) = x\}\) into two sets: \(A_1 := \{z \in \Omega \mid z(t) = x, z(t + h) > 0\}\) and \(A_2 := \{z \in \Omega \mid z(t) = x, z(t + h) \leq 0\}\). Then this information \(A_1\) or \(A_2\) cannot be resolved by \(\mathcal{F}_t\). We have to use a finer \(\sigma\)-algebra \(\mathcal{F}_{t+h}\) to categorize such information, which can only be known for sure on or after time \(t + h\).

Here that \(A_1\) is not \(\mathcal{F}_t\) measurable means that it is an non-observable event; that is, at time \(t\), it is impossible to observe an event which tells that \(B_t = x\) and \(B_{t+h} > 0\).

Now consider the stochastic process \(\{B^*_t\}_{t \geq 0}\) defined by

\[
B^*_t(\omega) = \max_{0 \leq s \leq t} B_s(\omega) \quad \forall \omega \in \Omega.
\]

This is the running high of the Brownian motion. It is a stochastic process adapted to the same filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) as that of \(\{B_t\}\). Clearly, for every \(t > 0\) and every \(\omega \in \Omega = C([0, \infty))\), we can always know the value \(B^*(t)\) from the restriction \(\omega|_{[0,t]}\).

Next consider the random variable

\[
X(\omega) = \int_0^1 B_s(\omega)ds \quad \forall \omega \in \Omega.
\]
This random variable is $\mathcal{F}_1$ measurable, but not $\mathcal{F}_s$ measurable for every $s < 1$. That is to say, the outcome of $X$ can be calculated only after time $t \geq 1$. For example, at time $t < 1$, it is impossible to observe an event which tells that $X < 0$.

Of course, as a random variable, the expectation, variance, etc of $X$ can be evaluated by using the finest filtration $\mathcal{F}_\infty$.

The filtration is vital in studying stopping times. For example, consider the first time that a Brownian motion has a running high equal to 1:

$$\tau(\omega) := \inf\{t > 0 \mid B_t(\omega) \geq 1\} = \sup\{t > 0 \mid B_s(\omega) < 1 \forall s \in [0,t]\} \quad \forall \omega \in \Omega.$$ 

The measurability of the function $\tau$ deeply relies on the filtration. For a natural filtration, $\tau$ is a stopping time in the sense that at each time $t$, one knows whether $\tau \leq t$ has happened or not, based on the knowledge of $\mathcal{F}_t$.

3. Consider another example, the stock price of a company. We use a space $\Omega$ sufficiently large so as to contain all information (past, current, and future) that is needed to determine the stock price (past, current, and future). That is, for each $\omega \in \Omega$, all financial conditions specified by $\omega$ provide a unique stock price $S_t(\omega)$ for all $t \in \mathbb{T} := [0, \infty)$. Suppose current tome is $t = 1$ and the stock price is $S = 10$. Then, in the language of probability, we say we have observed the event $E_1 = \{\omega \in \Omega \mid S_1(\omega) = 10\}$. Suppose we also know that at time $t = 1/2$ the stock price is $S = 9$. Then we say we have observed the event $E_2 = \{\omega \in \Omega \mid S_1(\omega) = 10, S_{1/2}(\omega) = 9\}$ which could be a fairly large set in $\Omega$. As we see, the more information we have, the smaller the set $E$. Nevertheless, usually it is impossible to use all past information to pin down a future event. That is, even if we know the value of $S$ for all $t \leq 1$, it is impossible to know the value $S$ at $t = 1.1$, since, for example, the set $E := \{\omega \in \Omega \mid S_{1.1}(\omega) = 11\}$ is not in $\cup_{t \leq 1} \mathcal{F}_t$ so it is impossible to observe such an event at time $t$ (e.g. predict with sure that $S = 11$ at time $t = 1.1$).

The following definition may clarify some of our thought about filtration as information.

**Box: If current time is $t$ and an event $E$ is in $\mathcal{F}_{t+h} \setminus \mathcal{F}_t$, then $E$ is called a future event.**

4. Let’s go back to the consider coin tossing game. Suppose one starts from a total of capital $V_0$ at tome $t = 0$ and will bet the outcome of the tossing of a coin at each integer time $t_j = j$. Suppose current time is $t = t_{j-1}$ and one has $V_{j-1}$ amount of capital. One makes a bet $b_{j-1}V_{j-1}$ on a head outcome and $c_{j-1}V_{j-1}$ on tail outcome where $b_{j-1} > 0, c_{j-1} > 0$ and $b_{j-1} + c_{j-1} \leq 1$. At time $t_j = j$, a coin is tossed and if the outcome is a head (denoted by $X_{j+1} = 1$) one collects award $b_{j-1}V_{j-1}$ from the bet on the head outcome and lost a penalty of $c_{j-1}V_{j-1}$ from the bet on the tail, so at time $t = t_j$, one has capital

$$V_j = V_{j-1}\left\{1 + (b_{j-1} - c_{j-1})X_j\right\} = V_0 \prod_{i=1}^{j}(1 + [b_{i-1} - c_{i-1}]X_i).$$

Similarly, if the outcome is a tail (denoted by $X_j = -1$), one collects award $c_{j-1}V_{j-1}$ from the bet on tail and lost $b_{j-1}V_{j-1}$ from the bet on head. Still one have same formula for $V_j$.

Now the question is how to maximize the final capital $V_N$, say at $N = 10$. Here the capital $V_{10}$ depends on various kinds of betting strategies and on the outcomes of the coin tossing.

Here for a fair game, a central restriction is that both $b_{j-1}$ and $c_{j-1}$ can depend only on a $V_0, V_1, \cdots, V_{j-1}$ and on $X_1, \cdots, X_{j-1}$, i.e., known information. It would be regarded as cheating if
In the form above such that the expected value $V$ also depends on $\{1, 4, \text{filtration and martingale}\}$.

In the language of filtration, the restriction means that both $b_{j-1}$ and $c_{j-1}$ have to be $F_{j-1}$ measurable. In terms of information tree, it means that one can only use the past and current information $\cup_{i\leq j-1}F_i = F_{j-1}$ to make a bet (investment) at time $t_{j-1} = j - 1$.

Since $V_1, \cdots, V_{j-1}$ are functions of $V_0, b_0, c_0, \cdots, b_{j-2}, c_{j-1}, X_0, \cdots, X_{j-1}$, a game strategy is represented a set of functions

$$b_{j-1} = b_{j-1}(V_0, x_1, \cdots, x_{j-1}), \quad c_{j-1} = c_{j-1}(V_0, x_1, \cdots, x_{j-1}), \quad j = 1, \cdots, N$$

such that at the out comes of $X_0 = x_0, \cdots, X_{j-1} = x_{j-1}$, one bets the portion $b_{j-1}(V_0, x_0, \cdots, x_{j-1})$ of total capital $V_{j-1}$ on heads and $c_{j-1}(V_0, x_0, \cdots, x_{j-1})$ of total capital on tails. Here although $b_{j-1}$ can also depend on $\{b_0, \cdots, b_{j-2}, c_0, \cdots, c_{j-2}\}$, the above strategy covers such case since by induction, each $b_i, c_i$, for each $i \leq j - 2$, are functions of $V_0, x_0, \cdots, x_i$.

Given a target, say maximize the expectation of $V_N$, the optimal strategy is to find a game strategy in the form above such that the expected value $V_N$ is maximized.

### 1.4.3 Conditional Probability

1. We recall that if $\mathcal{A}$ and $\mathcal{B}$ are two measurable sets of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then the conditional probability that

   the outcome is in $\mathcal{A}$ knowing that the outcome is in $\mathcal{B}$

   is defined as

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

   In the sequel, for a Borel set $A$ in $\mathbb{R}$ and random variable $X$ on $(\Omega, \mathcal{F}, \mathbb{P})$, the set $\mathcal{A} := \{\omega \in \Omega | X(\omega) \in A\}$ will be simply written as $\{X \in A\}$ or simply $X \in A$. Hence, given Borel sets $A$ and $B$ of $\mathbb{R}$, we can define

   $$\mathbb{P}(X \in A | X \in B) = \mathbb{P}(A | B), \quad \text{where } A := \{\omega \in \Omega | X(\omega) \in A\}, B := \{\omega | X(\omega) \in B\}.$$ 

   In particular, if $\rho(x)$ is the density function of $X$, then

   $$\mathbb{P}(A \cap B) = \mathbb{P}(X \in A \cap B) = \int_{A \cap B} \rho(x) \, dx, \quad \mathbb{P}(B) = \mathbb{P}(X \in B) = \int_B \rho(x) \, dx$$

   so

   $$\mathbb{P}(A | B) = \mathbb{P}(X \in A | X \in B) = \frac{\int_{A \cap B} \rho(x) \, dx}{\int_B \rho(x) \, dx}.$$ 

2. Let’s consider an application of the conditional probability for the Brownian motion $\{B_t\}_{t \geq 0}$. Let $0 \leq s < t$, $A$ be a Borel set of $\mathbb{R}$ and $y \in \mathbb{R}$. We want to define the conditional probability $\mathbb{P}(B_t \in A | B_s = y)$, i.e. the probability $B_t \in A$ under the condition that $B_s = y$. Since $\mathbb{P}(B_s = y) = 0$, we cannot directly define such a conditional probability by our definition. Nevertheless, for Brownian motion, we can define it through a limiting process:

   $$\mathbb{P}(B_t \in A | B_s = y) := \lim_{\Delta y \downarrow 0} \mathbb{P}(B_t \in A | B_s \in (y - \Delta y, y + \Delta y)).$$
Since the density of the joint distribution of \((B_t, B_s)\) is \(\rho(x, t; y, s) = e^{-(x-y)^2/[2(t-s)]} - y^2/[2s] / (2\pi\sqrt{(t-s)s})\), we find that

\[
P(B_t \in A \mid B_s = y) = \lim_{\Delta y \to 0} \frac{\int_A dx \frac{\int_{y-\Delta y}^{y+\Delta y} e^{-(x-y)^2/[2(t-s)]} - y^2/[2s] \sqrt{2\pi s}}{2\pi\sqrt{(t-s)s}} \, dz}{\int_{y-\Delta y}^{y+\Delta y} e^{-(x-y)^2/[2(t-s)]} - y^2/[2s] \sqrt{2\pi s}} \, dz
\]

\[
= \int_A \rho^{(y,s)}(x,t)dx,
\]

where

\[
\rho^{(y,s)}(x,t) := e^{-(x-y)^2/[2(t-s)]} / \sqrt{2\pi |t-s|}.
\]

If in particular \(A = (x, x + dx)\), then

\[
P(B_t \in (x, x + dx) \mid B_s = y) = \rho^{(y,s)}(x,t)dx.
\]

We call \(P(B_t \in A \mid B_s = y)\) the transition probability measure and \(\rho^{(s,y)}(x, t)\) the transition probability density.

3. Suppose \(X\) is a random variable on \((\Omega, \mathcal{F}, P)\) and \(A\) is a measurable set. We denote by \(E[X \mid A]\) the conditional expectation of \(X\) under \(A\) defined by

\[
E[X \mid A] := \int_A X(\omega)P(d\omega \mid A) = \int_A X(\omega) \frac{P(d\omega)}{P(A)} = \frac{\int_A X(\omega)P(d\omega)}{P(A)}.
\]

When \(A\) has measure zero, we need to take a limit process.

For example, for the Brownian motion, when \(t > s\),

\[
E[B_t \mid B_s = y] = \lim_{\Delta y \to 0} E[B_t \mid B_s \in (y - \Delta y, y + \Delta y)] = \int_{\mathbb{R}} dx \frac{\int_{y-\Delta y}^{y+\Delta y} e^{-(x-y)^2/[2(t-s)]} - y^2/[2s] \sqrt{2\pi s}}{2\pi\sqrt{(t-s)s}} \, dz
\]

\[
= \int_{\mathbb{R}} x \rho^{(y,s)}(x,t)dx = y \quad \forall y \in \mathbb{R}.
\]

4. Suppose \(X\) is a random variable on \((\Omega, \mathcal{F}, P)\). Assume \(\mathcal{G}\) is an \(\sigma\)-algebra on \(\Omega\) satisfying \(\mathcal{G} \subset \mathcal{F}\).

We say that that \(Y\) is the expectation of \(X\) under \(\mathcal{G} \subset \mathcal{F}\) and written as \(Y = \mathbb{E}[X \mid \mathcal{G}]\) if \(Y\) is \(\mathcal{G}\) measurable and

\[
\mathbb{E}[X - Y \mid A] = 0 \quad \forall A \in \mathcal{G}.
\]

Note that

\[
\mathbb{E}[Y - X \mid \mathcal{G}] = 0 \iff \int_A \left\{X(\omega) - Y(\omega)\right\}P(d\omega) = 0 \quad \forall A \in \mathcal{G}.
\]

To illustrate the idea, we consider the following example:

\[
\Omega = \{-1,1\}^2, \quad \mathcal{F} = 2^\Omega, \quad P(\omega) = \frac{1}{4} \quad \forall \omega \in \Omega,
\]

\[
\mathcal{G} = \{\emptyset, \{1\} \times \{-1,1\}, \{-1\} \times \{-1,1\}, \Omega\}.
\]
1.4. FILTRATION AND MARTINGALE

It is easy to see that \((\Omega, \mathcal{F}, \mathbb{P})\) is a probability space. Also, \(\mathcal{G}\) is a sub-\(\sigma\)-algebra of \(\mathcal{F}\) on \(\Omega\).

Now consider the functions
\[
X(x_1, x_2) := \sin(x_1) + x_1\{x_2 + (x_2)^2\}, \\
Y(x_1, x_2) := \tilde{Y}(x_1) := \sin(x_1) + x_1.
\]

It is easy to see that \(X\) and \(Y\) are random variables on \((\Omega, \mathcal{F})\). Moreover, \(Y\) is \(\mathcal{G}\) measurable. It is easy to verify that \(\mathbb{P}(X | A) = \mathbb{P}(Y | A)\) for every \(A \in \mathcal{G}\). For example, for each \(x_1 \in \{-1, 1\}\), \(A = \{x_1\} \times \{-1, 1\} \in \mathcal{G}\) is an atom of \(\mathcal{G}\), so
\[
\mathbb{E}[Y | A] = \tilde{Y}(x_1).
\]

On the other hand,
\[
\mathbb{E}[X | A] = \int_A Y(\omega) \frac{\mathbb{P}(d\omega)}{\mathbb{P}(A)} = \frac{1}{4}\{\sin(x_1) + x_1[1 + 1^2]\} + \frac{1}{4}\{\sin(x_1) + x_1[-1 + (-1)^2]\} = \sin(x_1) + x_1 = \tilde{Y}(x_1) = Y(x_1, x_2) \quad \forall x_2 \in \{-1, 1\}.
\]

In general, for the coin tossing example, a random variable \(X\) is a function of \((x_1, \ldots, x_N)\). The condition expectation \(\mathbb{E}[X | \mathcal{F}_i]\) is a function of \((x_1, \ldots, x_i)\).

Note that by definition, if \(Y\) is \(\mathcal{G}\) measurable, then
\[
Y = \mathbb{E}[Y | \mathcal{G}].
\]

For the Brownian motion we have a simple expression for conditional expectation. From the identity (1.3), we see that
\[
\mathbb{E}[B_t | \mathcal{F}_s] = B_s \quad \forall s \geq 0, t \geq s.
\]

Finally, we derive the following equivalent conditions to illustrate the idea. Assume that \(X\) is random variable on \((\Omega, \mathcal{F}, \mathbb{P})\) and \(Y\) is \(\mathcal{G}\) measurable where \(\mathcal{G}\) is a sub-\(\sigma\)-algebra of \(\mathcal{F}\). Assume that both \(X\) and \(Y\) have nice density functions. Then
\[
Y = \mathbb{E}[X | \mathcal{G}] \iff \int_A \{Y(\omega) - X(\omega)\} \mathbb{P}(d\omega) = 0 \quad \forall A \in \mathcal{G}
\iff \mathbb{E}[X - Y | Y = y] = 0 \quad \forall y \in \mathbb{R},
\iff \mathbb{E}[X | Y = y] = y \quad \forall y \in \mathbb{R}.
\]

Here we point out that for every \(y \in \mathbb{R}\), \(Y^{-1}((y - dy, y + dy))\) is a measurable set in \(\mathcal{G}\).

1.4.4 Martingale

Brownian motion is the canonical example of many classical stochastic processes. Here we introduce a few of them.

A stochastic process \(\{X_t\}_{t \geq 0}\) on \((\Omega, \mathcal{F}_\infty, \mathbb{P})\) is a martingale with respect to an adapted filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) if for each \(t \geq 0\), \(\mathbb{E}(|X_t|) < \infty\) and
\[
\mathbb{E}[X_{t+h} | \mathcal{F}_t] = X_t \quad \text{(i.e. } \mathbb{E}[X_{t+h} | X_t = x] = x \text{ a.s.)} \quad \forall h > 0, t \geq 0.
\]
A process \( \{S_t\}_{t \geq 0} \) is said to have **stationary increment** if for every \( h > 0 \), the distribution of \( S_{t+h} - S_t \) is independent of \( t \geq 0 \).

**A Markov process** is a stochastic process \( \{X_t\}_{t \geq 0} \) on certain probability space \((\Omega, \mathcal{F}, \mathbb{P})\) such that

\[
\mathbb{P}\left(X_t \in A \mid X_{s_1}, \ldots, X_{s_n} \in A_n\right) = \mathbb{P}\left(X_t \in A \mid X_s \in A \right)
\forall n \in \mathbb{N}, \ s_1 < \cdots < s_n < t, \ A, A_1, \ldots, A_n \in \mathcal{B}.
\]

From (1.3), we can see that Brownian motion is a martingale. Also, for the Brownian motion, \( B_{t+h} - B_t \) has \( N(0, h) \) distribution, which is independent of \( t > 0 \), so Brownian motion has stationary increment.

The condition for Markov process means that \( X_t \) depend only on the most recent history. Namely, the probability distribution of \( X_t \) under condition \( X_{s_1}, \ldots, X_{s_n} \in A_n \) is the same as the probability distribution of \( X_t \) under the sole condition condition \( X_{s_n} \in A_n \), provided that \( s_1, \ldots, s_{n-1} < s_n \).

A random walk, for example, is a (discrete) Markov process, since the position \( W \) only on the most recent history. Hence, Brownian motion is a Markov process.

For the Brownian motion, one can show the following:

\[
\mathbb{P}(B_t \in A \mid B_{s_1} = x_1, \ldots, B_{s_n} = x_n) = \mathbb{P}(B_t \in A \mid B_{s_n} = x_n)
\forall 0 \leq s_1 < \cdots < s_n < t, x_1, \ldots, x_n \in \mathbb{R}, A \in \mathcal{B}.
\]

Hence, Brownian motion is a Markov process.

A Markov process is determined essentially by the **transition probabilities**

\[
p^{(y,s)}(A, t) := \mathbb{P}\left(X_t \in A \mid X_s = y\right), \quad y \in \mathbb{R}, s < t, A \in \mathcal{B}.
\]

Let’s regard \( X_t \) as the position of a particle at time \( t \). Knowing the particle at position \( y \) at time \( s \), the probability that the particle is in \( A \) at time \( t > s \) is \( p^{(y,s)}(A, t) \). It has nothing to do with where the particle was at any time before \( s \). From here, we see that it is history independent.

The transition probability must satisfy the **Chapman-Kolmogorov equation**

\[
p^{(y,s)}(A; t) = \int_\mathbb{R} p^{(y,s)}((z, z + dz), \tau))p^{(z, \tau)}(A; t) \forall s < \tau < t.
\]

This equation says that to enter \( A \) at time \( t \) from a position \( y \) at time \( s \), the particle has to appear at some interval \((z, z + dz)\) at time \( \tau \) with probability \( p^{(y,s)}((z, z + dz), \tau) \) and then from there enter \( A \) at time \( t \), with probability \( p^{(z, \tau)}(A, t) \).

Conversely, let \( \{p^{(y,s)}(\cdot; t)\}_{y \in \mathbb{R}, t > s > 0} \) be a collection of probabilities satisfies the Chapman-Kolmogorov equation (1.5), and in addition assume that \( X_0 \) is given. Then there is a unique Markov process \( \{X_t\}_{t \geq 0} \) having the given initial distribution and the given transition probabilities.

In [22], the Brownian motion is indeed constructed from the Markov process by using the transition probability density

\[
p^{(y,s)}(x, t) = \frac{e^{-(x-y)^2/(2(t-s))}}{\sqrt{2\pi(t-s)}}.
\]
In using transition probability density and Markov process to define a stochastic process, it is a priori unknown if the resulting process has continuous paths. The following theorem characterizes a large class of Markov processes.

A Markov process \( \{X_t\}_{t \geq 0} \) with transition probabilities \( p^{(\cdot,\cdot)}(\cdot,\cdot) \) can be realized in the space of continuous functions if for every \( \varepsilon > 0 \),

\[
\lim_{\delta \searrow 0} \frac{1}{\delta} \sup_{y \in \mathbb{R}, 0 < t - s < \delta} p^{(y,s)} \left( \{ z \mid |z - y| > \varepsilon \} , t \right) = 0.
\]

Exercise 1.18. Suppose \( X \) is a random variable on \( (\Omega, \mathcal{F}) \) and \( A \) is an atom of \( \mathcal{F} \). Show that \( X|_A \) is a constant function; namely, for every \( \omega_1, \omega_2 \in A \), \( X(\omega_1) = X(\omega_2) \).

Exercise 1.19. Assume that \( 0 \leq s_1 < s_2 < t \). Let \( A = (x,y), A_1 = (x_1,y_1), A_2 = (x_2,y_2) \) be non-empty intervals. For the Brownian motion show that

\[
P(B_t \in A \mid B_{s_1} \in A_1, B_{s_2} \in A_2) = P(B_t \in A \mid B_{s_2} \in A_2).
\]

Exercise 1.20. Let \( n \gg 1 \) be an integer and set \( T = \{0,1,\ldots,n\} \) so we are considering the discrete version of a stochastic process. Suppose we are entering a game of tossing a fair coin. We start with \( X_0 = 1 \). If we have \( X_i \) capital after \( i \)th game, we bet \( X_i \) for the next game. If the tossing result is a head, we win half of the bet and otherwise we lose half of the bet, so that \( X_{i+1} = \frac{1}{2}X_i \) for head and \( X_{i+1} = \frac{1}{2}X_i \) for tail. Show that \( \{X_i\}_{i \in T} \) is a Markov process as well as a martingale, after building up appropriately a filtration (information tree).

Suppose \( Y_0 = 1, Y_1 = 1 \), and for \( i \geq 2 \), \( Y_{i+1} = \frac{1}{2}Y_i \) if both the \( i \)th and \((i+1)\)th outcomes of tossing are the same (both heads or both tails) and \( Y_{i+1} = \frac{1}{2}Y_i \) if the outcomes of the \( i \)th tossing and the \((i+1)\)th tossing are different. Show that \( \{Y_i\}_{i \in T} \) is a martingale, but not a Markov process.

Exercise 1.21. Let \( \{B_t\}_{t \geq 0} \) be the standard Brownian motion process adapted to a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \). Show that \( \{B_t^2\}_{t \geq 0} \) is a stochastic process that is not a martingale; that is, for some \( s < t \) and \( x \geq 0 \),

\[
E[B_t^2 \mid B_s^2 = x] \neq x.
\]

Also show that \( \{B_t^2\}_{t \geq 0} \) is a Markov process; that is, for every \( s_1 < s_2 < \cdots < s_n < t \) and \( a_1, \cdots, a_n \in (0, \infty) \),

\[
P(B_t^2 \mid B_{s_i}^2 = a_i, i = 1, \cdots, n) = P(B_t^2 \mid B_{s_n}^2 = a_n);
\]

namely,

\[
P(B_t^2 \mid B_{s_i}^2 \in (a_i, a_i + dx), i = 1, \cdots, n) = P(B_t^2 \mid B_{s_n}^2 \in (a_n, a_n + dx)).
\]

Exercise 1.22. Let \( \{B_t\}_{t \geq 0} \) be a Brownian motion adapted to a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \). Show that \( \{B_t^2 - t\}_{t \geq 0} \) is a martingale; that is, show that for every \( h > 0, t \geq 0 \), and \( x \in \mathbb{R} \),

\[
E[B_{t+h}^2 - (t+h) \mid B_t^2 - t = x] = x.
\]
Exercise 1.23. Show that the following two functions are transition probability densities:

\[ \rho_1(x,y,t,s) = \frac{1}{\pi} \frac{t-s}{(x-y)^2 + (t-s)^2}, \quad \rho_2(x,y,t,s) = \frac{e^{-(x-y)^2/[2(t-s)]}}{\sqrt{2\pi(t-s)}}, \quad \forall x,y \in \mathbb{R}, s < t. \]

Also, show that the resulting Markov process has stationary increment.

Exercise 1.24. Suppose \( \Omega = C([0, \infty), \mathbb{C}) \), \( B_t(\omega) = \omega(t) \) for all \( \omega \in \Omega \) and \( t \in [0, \infty) \). Assume that \( \{B_t\}_{t \geq 0} \) is a Brownian motion on \( (\Omega, \mathcal{F}_\infty, \mathbb{P}) \) and is adapted to a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \). Consider the first hitting time

\[ \tau(\omega) = \sup\{s > 0 \mid B_t(\omega) < 1 \forall t \in [0,s]\}. \]

Show that \( \tau \) is a stopping time; that is, show that for every \( t \geq 0 \), the set \{\( \omega \mid \tau(\omega) \leq t \}\} is \( \mathcal{F}_t \) measurable set.

Hint: Let \( r_0 = t \) and \( \{r_i\}_{i=1}^\infty \) be the set of all rational numbers in \( [0,t] \). Show that

\[ \{\omega \mid \tau(\omega) > t\} = \bigcap_{i=0}^\infty \{\omega \mid B_{r_i}(\omega) < 1\}. \]

Exercise 1.25. Assume that \( \mathcal{G} \subset \mathcal{F} \subset \mathcal{F}_\infty \). Assume that \( X \) is \( \mathcal{F}_\infty \) measurable. Show that

\[ \mathbb{E}\left[ \mathbb{E}[X \mid \mathcal{F}] \mid \mathcal{G} \right] = \mathbb{E}[X \mid \mathcal{G}]. \]

1.5 Itô Integrals

An ordinary differential equation (ODE) takes the form of

\[ \frac{dx}{dt} = f(x,t) \quad \text{or} \quad dx = f(x,t)dt \quad \text{or} \quad dX_t = f(X_t,t)dt. \]

Many systems in natural and social sciences are modelled by equations as above, subject to some random noise. The ubiquity of Brownian motion makes it natural to consider equations of the form:

\[ dX_t = f(X_t,t)dt + g(X_t,t)dB_t \]

where \( \{B_t\} \) is the Winner Process or Brownian motion. The fact that Brownian paths are nowhere differentiable makes this formulation problematic at first sight. To get around this problem, we could try to interpret the differential equation in its integrated form:

\[ X_t(\omega) = X_0(\omega) + \int_0^t f(X_s(\omega),s)ds + \int_0^t g(X_s(\omega),s)dB_s(\omega) \quad \forall \omega \in \Omega. \]

The fact that Brownian motion paths are not of finite variation on any interval means that the standard analytic techniques for defining the second integral in the equation do not apply (see, next subsection or [28]). So we have to find a new way to interpret integrals with respect to \( dB_s \) and for this we go back to the construction of the integral via Riemann sums.

In the sequel, we assume that the Brownian motion \( \{B_t\}_{t \geq 0} \) defined on \( (\Omega, \mathcal{F}_\infty, \mathbb{P}) \) is adapted to a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \) where \( \mathcal{F}_\infty = \cup_{t \geq 0} \mathcal{F}_t \).
1.5.1 Stochastic (Itô) Integration

The simplest way to define an integral is by Riemann sums. Thus, for a Riemann integrable function on \([a, b]\) one takes partitions \(\Pi = \{t_0, t_1, \ldots, t_N\}\) with \(a = t_0 < t_1 < t_2 \cdots < t_n = b\), and the integral is defined as

\[
\int_a^b f(t)\,dt = \lim_{n \to \infty} \sum_{i=1}^{n} f(t_i^*) (t_i - t_{i-1})
\]

where the limit is as the mesh of the partition (size of the largest interval) tends to zero, and each \(t_i^* \in [t_{i-1}, t_i]\). **Riemann integrable** functions are those functions for which the limit exists and is independent of choices of the \(t_i^*\) (left endpoint, right endpoint, midpoint, etc.). This can be generalized to

\[
\int_a^b f(t)\,dg(t) = \lim_{n \to \infty} \sum_{i=1}^{n} f(t_i^*) (g(t_i) - g(t_{i-1}))
\]

which is called the **Stieltjes integral** (see [28]). Again the limit (for \(f\) integrable) will be independent of the way of choosing the \(t_i^*\) when the function \(g\) is of **finite variation**, i.e.

\[
\sup_{a=t_0 < t_1 < \cdots < t_n = b} \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| < \infty.
\]

Unfortunately, this is not true for sample paths of Brownian motion, as shown below.

Let’s try to define \(\int_0^1 B_t\,dB_t\) using Riemann sums. For this let’s denote by \(S_R\) and \(S_L\) the Riemann sum resulting from the corresponding right-endpoint rule and left-endpoint rule respectively:

\[
S_R := \sum_{i=1}^{n} W_i(B_t - B_{t_{i-1}}), \quad S_L := \sum_{i=1}^{n} B_{t_{i-1}}(B_t - B_{t_{i-1}}).
\]

Unfortunately, we find the following

\[
\mathbb{E}[S_R - S_L] = \mathbb{E} \left[ \sum_{i=1}^{n} B_{t_i}(B_t - B_{t_{i-1}}) - \sum_{i=1}^{n} B_{t_{i-1}}(B_t - B_{t_{i-1}}) \right]
= \sum_{i=1}^{n} \mathbb{E} \left( B_{t_i} - B_{t_{i-1}} \right)^2 = \sum_{i=1}^{n} (t_i - t_{i-1}) = t_n - t_0 = b - a.
\]

Thus, we cannot define \(\int_a^b B_t\,dB_t\) as the Riemann integral.

The Itô stochastic integral is defined by **taking the left-endpoint rule**. It has the advantage of preserving the martingale property, but changes the rules of calculus, as exemplified by Itô’s lemma. (For the record, other choices are possible, although we will never use them. For example, the Stratonovich integral follows by taking the midpoint, preserves the ordinary rules of calculus, and is popular with people studying stochastic differential geometry).

The idea is then to approximate any stochastic process with “simple” stochastic processes. There are different ways of doing this in different books, but they are all essentially equivalent. The remainder of this section follows [8]. For simplicity, we use \(1_I\) to denote the characteristic function of the set \(I \subset \mathbb{R}:

\[
1_I(t) = 1 \quad \text{if} \quad t \in I, \quad 1_I(t) = 0 \quad \text{if} \quad t \not\in I.
\]
Definition 1.1. With respect to a filtration \( \{F_t\}_{t \geq 0} \), a (non-anticipating) simple function \( f \) is one for which there is a partition \( 0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = \infty \) such that

\[
f(\omega, t) = \sum_{i=0}^{n} a_i(\omega) 1_{[t_i, t_{i+1})}(t) \quad \forall t \geq 0, \omega \in \Omega,
\]

where for each \( i = 0, \ldots, n \), \( a_i : \Omega \to \mathbb{R} \) must of course be \( F_{t_i} \)-measurable.

For a simple function as above, its stochastic integral is defined by, for every \( \omega \in \Omega \) and \( t \geq 0 \),

\[
\int_0^t f(\omega, s) dB_s(\omega) := \lim_{f_n \to f, f_n \text{ simple}} \int_0^t f_n(\omega, s) dB_s(\omega) \quad \forall \omega \in \Omega, t \in [0, T].
\]

To define stochastic integral for general \( f \), the idea would then be to approximate \( f \) by simple functions and take limit, just as one does in classical integration theory. Although this doesn’t work for all \( f \), but it works for a large class of functions, in particular the class of non-anticipating functions.

The following theorem is the foundation for stochastic integrals; see [8] for a proof.

**Theorem 1.1** Assume that \( f(\omega, t) \) is \( F_\infty \times \mathcal{B} \) measurable (\( \mathcal{B} \) is the Borel \( \sigma \)-algebra on \([0, \infty)\)), \( \{F_t\} \) adapted, and \( \int_0^T E(f^2(\cdot, t)) dt < \infty \). Then the (Itô) stochastic integral

\[
\int_0^t f(\omega, s) dB_s(\omega) := \lim_{f_n \to f, f_n \text{ simple}} \int_0^t f_n(\omega, s) dB_s(\omega) \quad \forall \omega \in \Omega, t \in [0, T]
\]

exists, is \( \{F_t\} \) adapted, and satisfies the Itô’s Isometry

\[
E \left[ \left( \int_0^t f(\cdot, s) dB_s(\cdot) \right)^2 \right] = \int_0^t E[f^2(\cdot, s)] ds = E \left[ \int_0^t f^2(\cdot, s) ds \right] \quad \forall t \in [0, T].
\]

1.5.2 Itô’s Lemma

An Itô process is a stochastic process \( X_t \) of the form

\[
X_t(\omega) = X_0(\omega) + \int_0^t a(s, \omega) ds + \int_0^t b(s, \omega) dB_s
\]

(1.6)

where \( b \) is (stochastically) integrable and \( a(\cdot, \omega) \) is integrable for a.e. \( \omega \). It is a convention that one writes (1.6) in the shorthand differential form

\[
dX_t = a \, dt + b \, dB_t.
\]

(1.7)

Hence, (1.7) is understood in the sense of (1.6) with the Itô’s stochastic integral.
**Theorem 1.2 (Itô’s Lemma)** Let $X_t$ be an Itô process given by (1.7) in the sense of the stochastic integral (1.6). Let $g(x, t) \in C^{2,1}(\mathbb{R} \times [0, \infty))$ (i.e. $g$ is twice differentiable in $x$ and differentiable in $t$ on $\mathbb{R} \times [0, \infty)$). Then $g(X_t, t)$ is again an Itô process, and
\[
 dg(X_t, t) = \frac{\partial g(x, t)}{\partial t} \, dt + \frac{\partial g(x, t)}{\partial x} \, dX_t + \frac{1}{2} \frac{\partial^2 g(x, t)}{\partial x^2} (dX_t)^2 |_{x=X_t},
\]
where $(dX_t)^2 = b^2 dt$ which is computed according to the rules
\[
dt \cdot dt = dt \, dB_t = dB_t \, dt = 0, \quad (dB_t)^2 = dt.
\]

The proof of this theorem is essentially an application of Taylor’s theorem. The fact that $(dB_t)^2 = dt$ can be seen from Exercise 1.15. For a full proof, see [22].

The stochastic integral and Itô’s Lemma can be easily extended to multi-dimensions. We omit the details.

### 1.5.3 Stochastic Differential Equation (SDE)

Quite often a stochastic differential equation (SDE) is an ordinary differential equation with an addition of a random noise modelled by the Brownian motion process. Thus, denoting by $\{B_t\}_{t \geq 0}$ the Winners process (standard Brownian motion process), described in certain probability space $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}_{t \geq 0}$, a **classical** one-dimensional stochastic differential equation takes the form

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dB_t \quad \forall \, t > 0, \quad X_t \bigg|_{t=0} = X_0
\]

where $\sigma(x, t)$ and $\mu(x, t)$ are given (smooth) functions on $\mathbb{R} \times [0, \infty)$, $X_0$ is a given random variable measurable on $\mathcal{F}_0$, and $\{X_t\}_{t \geq 0}$ is an unknown process to be solved from the SDE. The solution is required to be adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. As mentioned earlier, the SDE has to be understood in the **stochastic integral formulation**

\[
X_t(\omega) = X_0(\omega) + \int_0^t \mu(X_s(\omega), s) \, ds + \int_0^t \sigma(X_s(\omega), s) \, dB_s(\omega) \quad \forall \, \omega \in \Omega.
\]

The following is the basic existence and uniqueness theorem for stochastic differential equations.
Theorem 1.3 (Existence and Uniqueness for SDEs) Let $T > 0$, $\mathcal{B} = \{B_t\}_{0 \leq t \leq T}$ be an $m$-dimensional (column vector) standard Brownian motion, and $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$, $A : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ be measurable functions satisfying, for some constant $C > 0$,

\[
\|b(0, t)\| + \|A(0, t)\| \leq C \quad \forall t \in [0, T]
\]
\[
\|b(x, t) - b(y, t)\| + \|A(x, t) - A(y, t)\| \leq C\|x - y\| \quad \forall x, y \in \mathbb{R}^n, \ t \in [0, T].
\]

Then for any initial value $X_{t|t=0} = X_0$ which is square $\mathcal{F}_0$-integrable and independent of $W$, the stochastic differential equation

\[
dX_t = b(X, t)dt + A(X, t) dB_t \quad 0 < t \leq T
\]

has a unique, adapted, square-integrable solution $\{X_t\}_{0 \leq t \leq T}$.

The theorem was originally proved by Itô using Picard iteration, much like the corresponding result for ordinary differential equations; see for example, [14].

Exercise 1.26. Let $f$ be a simple functional and set $g(\omega, t) = \int_0^t f(\omega, s)dB_s(\omega)$. Show the following:

1. $g(\omega, \cdot)$ is continuous for almost all $\omega$;
2. $\{g(\cdot, t)\}_{t \geq 0}$ is a martingale, i.e. $\mathbb{E}[g(\cdot, t + h) | g(\cdot, t) = x] = x$ for every $t \geq 0, h > 0, x \in \mathbb{R}$;
3. $\{g(\cdot, t)\}_{t \geq 0}$ is adapted;
4. for every $t \geq 0$, $\mathbb{E}(g^2(\cdot, t)) = \mathbb{E}\left(\int_0^t f^2(\omega, s) ds\right) = \int_0^t \mathbb{E}(f^2(\cdot, s))ds$.

Exercise 1.27. Let $\{X_t\}$ be an Ito process and $f(\cdot, \cdot), g(\cdot, \cdot)$ be smooth functions from $\mathbb{R} \times [0, \infty)$ to $\mathbb{R}$. Show that for the processes $f = \{f(X_t, t)\}_{t \geq 0}$ and $g = \{g(X_t, t)\}_{t \geq 0},$

\[
d(fg) = fdg + gdf + dfdg.
\]

Exercise 1.28. If $X_t = e^{B_t}$ then what is $dX_t$?

Exercise 1.29. Assume that $\mu$ and $\sigma$ are constants. Show that $X_t := e^{B_t - vt}$ where $\nu := \mu - \frac{1}{2} \sigma^2$ is the solution to $dX_t = X_t(\mu dt + \sigma dB_t)$.

Exercise 1.30. Suppose $dS_t = S_t(\mu dt + \sigma dB_t)$ where $\mu$ and $\sigma$ are constants. Show that

\[
S_t = S_0 e^{\nu t + \sigma B_t} \quad \nu := \mu - \frac{1}{2} \sigma^2.
\]

Exercise 1.31. Is $X_t = B_t^2$ the solution to $\frac{dX_t}{2\sqrt{X_t}} = dB_t$?

Exercise 1.32. Let $\{B_t\}$ be the Wiener process. Consider the integral

\[
Y_t = \int_0^t B_s ds.
\]
1. Let $0 = t_0 < t_1 < \cdots < t_n = t$ be a partition of $[0, t]$ and $s_i, \tau_i \in [t_{i-1}, t_i]$ for every $i$. Consider the Riemann sums

$$S_1 := \sum_{i=0}^{n-1} B_{s_i}(t_{i+1} - t_i), \quad S_2 := \sum_{i=0}^{n-1} B_{\tau_i}(t_{i+1} - t_i).$$

Show that

$$\mathbb{E}[|S_1 - S_2|^2] \leq t \Delta t$$

where $\Delta t = \max_i |t_{i+1} - t_i|$.

2. Show that $Y_t$ is a Riemann integral, i.e.,

$$Y_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} B_{s_i}(t_{i+1} - t_i)$$

where $\lim$ is taken as mesh size approaches zero.

3. Show that $\mathbb{E}[Y_t] = 0$ for every $t \geq 0$.

4. Show that $\{Y_t\}_{t \geq 0}$ is not a martingale. In particular, show that for every $h > 0$ and $t \geq 0$,

$$\mathbb{E}[Y_{t+h} | \mathcal{F}_t] = Y_t + hB_t.$$ That is, show that $\mathbb{E}[Y_{t+h} - Y_t - hB_t | \mathcal{F}_t] = 0$, or more generally,

$$\mathbb{E}[Y_{t+h} - Y_t | B_t = x] = hx.$$

5. Show that $\{Y_t - tB_t\}_{t \geq 0}$ is a martingale.

1.6 Diffusion Process

1.6.1 Itô diffusion

**Definition 1.2.** An Itô diffusion, or a diffusion process is a stochastic process $X_t$ satisfying a stochastic differential equation of the form

$$dX_t = b(X_t)dt + A(X_t)dB_t, \quad t \geq s \quad X_s = x$$

where $B_t$ is an $m$-dimensional standard Brownian motion and $b : \mathbb{R}^n \to \mathbb{R}^n$ and $A : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ satisfy the hypotheses of the existence and uniqueness theorem.

The following theorem gives the Markov Property for Itô diffusions, which defines them as a sort of stochastic analogue of dynamical systems. Heuristically, it says that the future evolution of the process is independent of the past, given its current position.

**Theorem 1.4 (Markov Property)** Let $f$ be a bounded (and measurable) function from $\mathbb{R}^n \to \mathbb{R}$, and $X_t$ a diffusion process. Then, for $t, h \geq 0$, $y \in \mathbb{R}^n$,

$$\mathbb{E}[f(X_{t+h}) | X_t = y] = \mathbb{E}[f(X_h) | X_0 = y]$$
1.6.2 Semigroup Generator

We denote by $\mathcal{BC}(\mathbb{R}^n; \mathbb{R})$ the space of all bounded continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$.

Let $\{X_t\}$ be a diffusion process. Consider the operators $T_t : f \in \mathcal{BC}(\mathbb{R}^n; \mathbb{R}) \to T_tf \in \mathcal{BC}(\mathbb{R}^n; \mathbb{R})$ defined by

$$T_tf(x) = \mathbb{E}[f(X_t) \mid X_0 = x] \quad \forall x \in \mathbb{R}^n, t \geq 0.$$  

The Markov property implies that the collection of operators $\{T_t\}_{t \geq 0}$ is a semigroup on $\mathcal{BC}(\mathbb{R}^n; \mathbb{R})$, i.e.

(i) $T_0f = f \quad \forall f,$

(ii) $T_tT_s = T_{t+s} \quad \forall t, s \geq 0.$

A deep analysis shows that in a dense subspace $D$ of $\mathcal{BC}(\mathbb{R}^n, \mathbb{R})$, the following limit exists:

$$A f := \lim_{h \to 0} \frac{T_h f - f}{h} \quad \forall f \in D. \quad (1.8)$$

Note that $A$ can be regarded as the derivative of $T_t$ at $t = 0$. The operator $A$, together with its definition domain $D$, is called the semigroup generator. There is a one-to-one correspondence between (nice) semigroups and generators, by the (symbolic) relation

$$\frac{d}{dt} T_t = A T_t = T_t A, \quad T_t = e^{tA} \quad \forall t \geq 0.$$  

One can easily prove this if $A$ were a number or a matrix; here this is not the case, but things can still be worked out by using functional analysis and stochastic calculus. Performing a calculation with Itô’s lemma on equation (1.8) one can find that (at least for smooth enough $f$)

$$A f = \sum_{i=1}^n (b)_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\mathbf{A} \mathbf{A}^T)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j},$$

which is the differential operator that appears in all the PDEs below.

1.6.3 Kolmogorov’s Backward Equation

There are many important connections between stochastic differential equations and partial differential equations. Here we review some of the ones that are most vital for mathematical finance.

Suppose $f \in D$. Consider the function $u$ defined by

$$u(x, t) = \mathbb{E}[f(X_t) \mid X_0 = x] = T_t f(x).$$

At least formally, we can differentiating with respect to $t$ to obtain

$$\frac{\partial u}{\partial t} = \frac{d}{dt} T_t f = AT_t f = Au.$$  

The following can be proved rigorously using stochastic methods (see [22]).

**Theorem 1.5 (Kolmogorov’s Backward Equation)** Assume that $dX_t = b(X_t) + \mathbf{A}(X_t)d\mathbf{B}_t$. Let $f \in \mathcal{BC}(\mathbb{R}^n; \mathbb{R})$. Then $u(x, t) := \mathbb{E}[f(X_t) \mid X_0 = x]$ solves

$$\frac{\partial u}{\partial t} = Au \quad \forall t > 0, x \in \mathbb{R}^n, \quad u(x, 0) = f(x) \quad \forall x \in \mathbb{R}^n.$$
Let $\rho(t, x, y)$ be the probability density function of $X_t$ given that $X_0 = x$. That is,
\[
\mathbb{P}(X_t \in A \mid X_0 = x) = \int_A \rho(t, x, y) dy \quad \forall A \in \mathcal{B}^n
\]
Here $\mathcal{B}^n$ is the Borel $\sigma$-sigma of $\mathbb{R}^n$. The function $\rho$ is often called the transition density of the process. Assume that such a density exists and is as smooth as it needs to be. Then by the definition of expectation,
\[
u(x, t) := \mathbb{E}[f(X_t) \mid X_0 = x] = \int_{\mathbb{R}^n} f(y) \rho(t, x, y) dy
\]
Then the Kolmogorov’s backward equation $\frac{\partial \nu}{\partial t} = \mathcal{A} \nu$ becomes
\[
\int_{\mathbb{R}^n} f(y) \frac{\partial \rho(t, x, y)}{\partial t} dy = \int_{\mathbb{R}^n} f(y) \mathcal{A} \rho(t, x, y) dy
\]
where $\mathcal{A}$ is supposed to operate on the $x$ variable. Since this is true for all $f$ we have
\[
\frac{\partial \rho}{\partial t} = \mathcal{A} \rho := \sum_{i=1}^n (b_i) \frac{\partial \rho}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (A A^T)_{ij} \frac{\partial^2 \rho}{\partial x_i \partial x_j}.
\]
The initial condition for the above equation is simply $\rho(0, x, y) = \delta(x - y)$ where $\delta$ is the Dirac’s delta “function”; it is either probabilistically obvious to us, or may be derived from the initial condition for the equation with $f$ in it. This gives us a PDE satisfied by the transition probability $\rho(t, x, y)$ of a diffusion process, where $(x, t)$ are variables, and $y$ is considered as a parameter.

**Example 1.1.** Consider the one space dimensional case and for simplicity assume that $b = \nu \in \mathbb{R}, A = \sigma > 0$ are constants. Then the equation for the density function becomes
\[
\frac{\partial \rho}{\partial t} = \frac{\nu^2}{2} \frac{\partial^2 \rho}{\partial x^2} + \nu \frac{\partial \rho}{\partial x} \quad \forall x \in \mathbb{R}, t > 0, \quad \rho(0, x, y) = \delta(x - y).
\]
One can verify that the solution is given by
\[
\rho(t, x, y) = \frac{e^{-(x+\nu t-y)^2/(2\sigma^2 t)}}{\sqrt{2\pi t \sigma^2}} \quad \forall t > 0, x, y \in \mathbb{R}.
\]
Consequently,
(i) For the corresponding diffusion process,
\[
\mathbb{P}(X_t \in A \mid X_s = x) = \int_A \frac{e^{-(x+\nu (t-s)-y)^2/(2\sigma^2 (t-s))}}{\sqrt{2\pi (t-s) \sigma^2}} dy \quad \forall t > s \geq 0, x \in \mathbb{R}, A \in \mathcal{B}.
\]
(ii) For $f \in \mathcal{BC}(\mathbb{R}; \mathbb{R})$, the function $u(x, t) = \mathbb{E}(f(X_t) \mid X_0 = x)$, which is the solution to
\[
\frac{\partial u}{\partial t} = \frac{\nu^2}{2} \frac{\partial^2 u}{\partial x^2} + \nu \frac{\partial u}{\partial x} \quad \forall x \in \mathbb{R}, t > 0, \quad u(x, 0) = f(x) \quad \forall x \in \mathbb{R}
\]
is given by
\[
u(x, t) = \int_{\mathbb{R}} \frac{e^{-(x+\nu t-y)^2/(2\sigma^2 t)}}{\sqrt{2\pi t \sigma^2}} f(y) dy.
1.6.4 Kolmogorov’s Forward or Fokker-Planck equation

For a smooth function $f(x)$, Itô’s stochastic integral and lemma provide

$$f(X_t) - f(X_0) = \int_0^t df(X_s) = \int_0^t \left\{ \frac{\partial f}{\partial x_i} dX_s + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} dX_s^i dX_s^j \right\}$$

by using $dX_s = b ds + A dB_s$ and the definition of $A$. Taking expectation, conditioned on starting at $x$, gives

$$E[f(X_t) \mid X_0 = x] - f(x) = E \left[ \int_0^t A f(X_s) ds \mid X_0 = x \right]$$

(remember that the Itô integral is a martingale having expectation zero). In terms of the transition density it is equivalent to

$$\int_{\mathbb{R}^n} f(y) \rho(t, x, y) dy - f(x) = \int_0^t \int_{\mathbb{R}^n} A_y f(y) \rho(s, x, y) dy ds$$

Here $A_y$ means $A$ acting on $y$ variables. Differentiating with respect to $t$ and integrating by parts gives

$$\int_{\mathbb{R}^n} f(y) \frac{\partial \rho(t, x, y)}{\partial t} dy = \int_{\mathbb{R}^n} A_y f(y) \rho(t, x, y) dy = \int f(y) A^\ast \rho(t, x, y) dy$$

where $A^\ast$, acting on the $y$ variables, is the “adjoint” of the operator $A$, given by

$$A^\ast g(y) = \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} \left( (A^T A)_{ij} g \right) - \sum_i \frac{\partial}{\partial y_i} \left( (b)_i g \right)$$

Since this holds for all $f$ we have the Kolmogorov forward equation,

$$\frac{\partial \rho(t, x, y)}{\partial t} = A^\ast \rho := \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial y_i \partial y_j} \left\{ (A A^T (y))_{ij} \rho(t, x, y) \right\} - \sum_{i=1}^n \frac{\partial}{\partial y_i} \left\{ (b(y))_i \rho(t, x, y) \right\}.$$

Again, the initial condition is $\rho(0, x, y) = \delta(x - y)$ (the Dirac delta). In this equation for $\rho(t, x, y)$, $(t, y)$ are the variables, whereas $x$ is considered as a parameter.

1.6.5 The Feynman-Kac Formula

The Feynman-Kac formula was inspired by Feynman’s path integral formulation of quantum mechanics, which was not mathematically rigorous and where there would be a $\sqrt{-1}$ in the PDE, but was proved rigorously by Kac in the form below; see [14] for more details.

Fix $T > 0$, and let $f(x) : \mathbb{R}^n \to \mathbb{R}$, $g(x, t) : \mathbb{R}^n \times [0, T] \to \mathbb{R}$ and $k(x, t) : \mathbb{R}^n \times [0, T] \to \mathbb{R}$ be “nice” functions (e.g continuous and bounded).
Theorem 1.6 (Feynman-Kac Formula) Suppose \( dX_t = b dt + A dB_t \) where the coefficients \( b, A \) satisfy the hypotheses of the existence and uniqueness theorem. Let \( v(x, t) : \mathbb{R}^n \times [0, T] \to \mathbb{R} \) be the solution to the pde problem
\[
\frac{\partial v(x, t)}{\partial t} + A v + k v = g \quad \text{in } \mathbb{R}^n \times [0, T), \quad v(x, T) = f(x) \quad \forall x \in \mathbb{R}^n.
\]
Then \( v(t, x) \) admits the stochastic representation, for every \( x \in \mathbb{R}^n \) and \( t \in [0, T] \),
\[
v(x, t) = E \left[ f(X_T) e^{\int_t^T k(X_s) ds} - \int_t^T g(X_s, \tau) e^{\int_t^\tau k(X_s) ds} d\tau \mid X_t = x \right].
\]

One important application of the Feynman-Kac formula is the Monte-Carlo simulation. Here a typical Monte-Carlo simulation involves first generating a group \( \{ X(s, \omega_i) \}_{i=1}^N \) of sample paths from the SDE \( dX = b dt + A dB \) with “initial condition” \( X_t = x \) and then calculating the average (approximation of the expectation) of the integral inside the expectation bracket. Usually the Monte-carlo simulation is very easy to program and quite stable, and most importantly, simulating the reality. The drawback is that quite often, the method needs sufficient amount of sample paths and is not so easy to obtain highly accurate results.

Exercise 1.33. Assume that \( \{X_t\}_{t \geq 0} \) is an Ito process, i.e. \( dX_t = b(X_t) dt + A(X_t) dB_t \). Let \( f \in \mathcal{B}C(\mathbb{R}^n; \mathbb{R}) \) be a smooth function. Using \( dE[f(X_t)]|X_0 = x] = E[df(X_t)|X_0 = x] \) and Itô’s lemma formally derive that
\[
A f := \left. \frac{d}{dt} f \right|_{t=0} = \sum_{i=1}^n (b(x))_i \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (A(x)A^T(x))_{ij} \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.
\]

Exercise 1.34. Let \( \{B_t\} \) be the Brownian motion and \( X_t = x + 2t + W_t \) so \( dX_t = 2 dt + dB_t \) and \( X_0 = x \). Using the fact that \( B_t \) is \( N(0, t) \) distributed, show that the function \( u(x, t) := E[f(X_t) \mid X_0 = x] = E[f(x + 2t + B_t)] \) is given by
\[
u(x, t) = \int_{\mathbb{R}} f(x + 2t + y) \rho(y, t) dy, \quad \rho(y, t) := \frac{e^{-y^2/(2t)}}{\sqrt{2\pi t}}.
\]
Also show that \( u(x, t) \) satisfies the Kolmogorov backwards equation.

Hint: Write \( u(x, t) = \int_{\mathbb{R}} f(z) \rho(z-x-2t, t) dz \).

Exercise 1.35. Integrating the Kolmogorov forward equation show that
\[
\frac{d}{dt} \int_{\mathbb{R}^n} \rho(t, x, y) dy = 0 \quad \forall t > 0, \quad \int_{\mathbb{R}^n} \rho(t, x, y) dy = 1 \quad \forall t \geq 0.
\]
Explain the probabilistic meaning of the identity.

Exercise 1.36. In case of one space dimension solve the Kolmogorov forward equation for the transition probabilities for the standard Brownian motion.

Exercise 1.37. In one space dimension, simplifying the Feynman-Kac Formula in the case \( k, b = \nu, A = \sigma \) are all constants.
Exercise 1.38. (Monte-Carlo simulation) Consider the pde problem for $V = V(x,t)$:

$$
\frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial x^2} + V = 1 \quad \text{in} \quad \mathbb{R} \times [0,T], \quad V(x,T) = f(x) := \cos x \quad \forall x \in \mathbb{R}.
$$

(i) Show that the exact solution is $V(x,t) = 1 - e^{T-t} + e^{(T-t)/2} \cos x$.

(ii) Let $\{B_t\}$ be standard the Brownian motion. Using the Feynman-Kac formula show that

$$
V(x,t) = 1 - e^{T-t} + e^{T-t} \mathbb{E}\left[ f(x + B_T) \right]
$$

(iii) Consider the case $x = 0$ and $T = 1$. Let $N$ be a large integer, say $N = 1024$. Use a random generator generating $N$ sample Brownian motions paths, $\omega_1(\cdot), \cdots, \omega_N(\cdot) \in C([0,1])$. Here random walks can be regarded as Brownian motions in numerical simulations. Assuming each path has probability $1/N$, we can approximate the Feynman-Kac formula by

$$
V_N(0,1) = 1 - e + \frac{e}{N} \sum_{i=1}^{N} \cos(\omega_i(1)).
$$

Find the approximation and estimate the error $V_N(0,1) - V(0,1)$. 
Chapter 2

The Black–Scholes Theory

In this chapter we study the well-celebrated theory in mathematical finance—the theory of Black and Scholes. Their paper [1] initiated the modern approach to option’s evaluation. They discovered the rational option price derived from risk-neutral probability, an astonishing consequence of the seemingly trivial no-arbitrage assumption. Another earlier significant contribution was Merton [19, 18]. The simplified approach using binomial lattice (or random walks) was first presented by Sharpe in [29] and later developed by Cox, Ross, Robinstein [4] and also Rendleman and Bartter [24].

2.1 Replicating Portfolio

2.1.1 Modelling Unit Share Prices of Securities

Suppose we have $m$ securities (assets) $a := (a_1, \cdots, a_m)$. In modern theory of mathematical finance, we use a $m$ dimensional vector valued stochastic process $\{S_t\}_{t \in [0,T]}$ to denote the unit share prices of the securities in time interval $T = [0, T]$; here $S_t := (S^1_t, \cdots, S^m_t)$ where for each $i = 1, \cdots, m$, $S^i_t$ is the unit share price of the asset $a_i$ at time $t \in T$.

By stochastic process, of course we have a probability space $(\Omega, F, P)$. Hence, each $S^i_t$ is a real (indeed non-negative) valued measurable function with definition domain $\Omega$ whereas $S_t$ is a $m$ dimensional vector valued measurable function with definition domain $\Omega$.

In theory, by introducing random variables, any reasonable model has to incorporate the very fact that future prices of risky assets are indeterministic, whereas history prices of any assets are known. For this reason, filtration is introduced to model the flow of information.

Mathematically, a filtration is a collection $\{F_t\}_{t \in T}$ of $\sigma$-algebras on $\Omega$ such that $F_s \subset F_t$ for every $s, t \in T$ with $s < t$. Cooperating flow of information into filtration is done by requiring that $\{S_t\}_{t \in T}$ is $\{F_t\}_{t \in T}$ adapted. That is, $S_t$ is $F_t$ measurable for every $t \in T$. Of course, with filtration, $F$ is taken to be the biggest $\sigma$-algebra, $F = \cup_{t \in T} F_t = F_T$. Quite often, one uses the natural filtration.

Now how filtration models the information. We take an example of a single risky asset whose unit share price is modelled by a stochastic process $\{S_t\}_{t \in T}$, adapted to (a natural) filtration $\{F_t\}_{t \in T}$ in the probability space $(\Omega, F_T, P)$. To illustrate the idea, let’s assume that $T = \{t_1, \cdots, t_k = T\}$ is finite and $\Omega$ is also finite. Suppose we are at time $T$ so we have observed the stock prices $f_1 = S_{t_1}, \cdots, f_k = S_{t_k}$. Then by observing the whole history, we know that a particular event $\omega^* \in \Omega$ actually happened. This event has the property that $S_{t_i}(\omega^*) = f_i$ for all $i = 1, \cdots, k$. 

33
Now let’s move back to an earlier times. At \( t = t_1 \), we only know one price \( S_{t_1} = f_1 \). This information tells that

\[
\omega^* \in A_1 = \{ \omega \in \Omega \mid S_{t_1}(\omega) = f_1 \}.
\]

What we know at time \( t = t_1 \) is the set \( A_1 \); namely, we narrowed down the true event \( \omega^* \) from being in the biggest set \( \Omega \) to being in the smaller set \( A_1 \).

Similarly, at time \( t = t_2 \), we know two pieces of information, so we know that

\[
\omega^* \in A_2 = \{ \omega \in \Omega \mid S_{t_1}(\omega) = f_1, S_{t_2}(\omega) = f_2 \}.
\]

As we are getting more and more information, the set \( A_i \) gets smaller and smaller, therefore, we are getting closer and closer to the final revelation of \( \omega^* \).

Now what happens if at a time \( t = t_j \) we find that \( A_j \) is a single element. Well, in this case, we know that this single element has to be \( \omega^* \). Consequently, we know at time \( t = t_j \) all future stock price \( S_{t_l}, l = j + 1, \ldots, k \). Clearly, this is against our intuition that it is impossible to pin down future prices.

Hence, to exclude such a possibility, we require that \( A_j \) have to be \( F_j \) measurable. If \( j < k \), then except set of measure zero, the smallest set in \( F_j \) is not a singleton! Thus, \( A_j \) is not a singleton, and at time \( t = t_j \), we only know that \( \omega^* \in A_j \). That is all.

The finesse of the filtration \( F_j \) is just small enough so that on each atom \( A \) of \( F_j \), the price \( S_{t_i} \) on \( A \) is a constant function for each \( i = 1, \ldots, j \), but \( S_{t_i} \) on \( A \) for \( l \geq j + 1 \) is not a constant function. Hence, a good model uses a good filtration so that an atom in \( F_j \) can pin down all historical stock prices up to \( t \) with 100% certainty, but for any future information, one only has a conditional probability distribution \( \mathbb{P}(\cdot \mid A) \) for every \( A \in F_j \). Since \( A \in F_j \) can be as small as the atom of \( F_j \), we often write \( \mathbb{P}(\cdot \mid F_j) \).

### 2.1.2 Portfolios

A **portfolio** in a time interval \([0, T]\) is a collection \( \{n_t := (n^1_t, \ldots, n^m_t)\}_{t \in [0,T]} \) describing the number \( n^i_t \) of unit shares of security \( a_i \), held at time \( t \) (or more precisely, in time interval \([t, t + dt]\)), \( i = 1, \ldots, m \).

When unit share prices are \( S_t = (S^1_t, \ldots, S^m_t) \), the total value \( V_t \) of the portfolio at time \( t \) is

\[
V_t = n_t \cdot S_t = \sum_{i=1}^{m} n^i_t S^i_t.
\]

A portfolio \( \{n_t := (n^1_t, \ldots, n^m_t)\}_{t \in [0,T]} \) is called a **trading strategy** if it is adapted to \( \{F_t\}_{t \in [0,T]} \), the same filtration modelling stock prices.

A trading strategy \( \{n_t\}_{t \in [0,T]} \) is called **self-financing** if for every \( t \in [0, T] \),

\[
dV_t = n_t \cdot dS_t = \sum_{i=1}^{m} n^i_t dS^i_t.
\]

Here by requiring \( n_t \) to be adapted to \( F_t \), it means that \( n_t \) is **non-anticipating**, i.e., one does not need future information on stock prices to determine the current strategy; nevertheless, one does need all information of the past and the current to make a decision. In some textbook (e.g.,[3]), a trading
strategy is required to be pre-visible; that is, $n_t$ depends only on information up to time $t$ but not $t$ itself. Mathematically, this means that $n_t$ is $\bigcup_{s<t} \mathcal{F}_s$ measurable. Since in this chapter we are dealing only with continuous processes and continuous trading strategies, there is no distinction between $n_t$ being $\bigcup_{s<t} \mathcal{F}_s$ measurable and $n_t$ being $\mathcal{F}_t = \bigcup_{s\leq t} \mathcal{F}_s$ measurable.

Note that by self-financing it means no external money is put in or taken out, except the initial investment. Hence, the value change is totally due the unit price changes of the securities.

### 2.1.3 Claim and Replicating Strategy

While trading strategies can be used for designing optimal growth of wealth, in the Black–Scholes theory, it is used to design strategies to hedge, i.e., to pay prearranged payments in contracts.

A (contingent) **claim** at time $T$ is an $\mathcal{F}_T$ measurable random variable $X$ where $X(\omega)$ represents prearranged payment at time $T$ when the event $\omega \in \Omega$ occurs.

A **replicating strategy** for a claim $X$ at time $T$ is a self-financing strategy $\{n_t\}_{t \in [0,T]}$ such that $V_T(\omega) = X(\omega)$ for all $\omega \in \Omega$, i.e.,

$$X(\omega) = n_T(\omega) \cdot S_T(\omega) = \sum_{i=1}^{m} n_T^i(\omega) S_T^i(\omega) \quad \forall \omega \in \Omega.$$  

If a claim $X$ has a replicating strategy $\{n_t\}_{t \in [0,T]}$, then for each $t \in [0,T]$, the value $V_t$ of the portfolio $\{n_t\}$ is called the **arbitrage-free price** of the claim at time $t$; in particular, its initial arbitrage-free price is

$$V_0 = n_0 \cdot S_0 = \sum_{i=1}^{m} n_0^i S_0^i.$$  

We remark that it is required that $\mathcal{F}_0$ is the coarsest $\sigma$-algebra, i.e., $\mathcal{F}_0 = \{, \Omega\}$. Hence, $n_0, S_0$ being $\mathcal{F}_0$ measurable means that both $n_0$ and $S_0$ are constant vectors. It follows that $V_0$ is a deterministic value. Similarly, if $g \in C([0,t]; \mathbb{R}^n)$ is a given function, then under the observable event

$$A := \{\omega \in \Omega \mid S_s(\omega) = g(s) \text{ for all } s \in [0,t]\},$$  

the function $n_t$ is known constant vector on $A$, so $V_t$ is a known quantity.

### 2.1.4 Expectation Pricing and Arbitrage-free Pricing

Given a claim $X$ for time $T$, under a probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$, its **expectation price** is

$$\mathbb{E}(X) := \int_\Omega X(\omega) \mathbb{P}(d\omega).$$  

Until 1973, the financial community had been using expectation to price future contracts. However, Black and Scholes, in their 1973 revolutionary paper [1], pointed out how wrong in using expectation to price claims.

Here we illustrate Black–Scholes idea by a simple example.
Problem: Suppose there is a future contract which states that party A is to deliver 100 unit shares of ABC company’s stock one year from now to party B, and party B is to give party A at the time of receiving stocks an $P$ amount of money. The question here is what price $P$ should be written in the contract? Assume that unbearable penalty is imposed on either party should their obligation are not fulfilled.

There are abundant reasons for such kind of contract to come into play. The first party who will deliver the stock want to secure a fixed income one year from now, whereas the second party who will buy the stock want to acquire the stock in the future with a secured price. This scenario arises, for example, when the second party has a certificate of deposit which is to be matured one year from now so the cash is only available at that time. The first party, meanwhile, may have stocks available only one year from now, e.g., an employee of the ABC company who promised to give free stocks to its employees as bonus at the specific time.

Suppose the current ABC company’s stock price is $S_0 = $40.00. Detailed painstaking analysis indicates that the annual (continuously compounded) growth rate is $\nu = 15\%$ with volatility $\sigma = 20\%$. That is, in the language of probability, $S_t = S_0 e^{r_t+\sigma \sqrt{T}X}$ where $X$ is $N(0,1)$ distributed. Then in probability, the expected price is

$$E[S_T] = \int_{\Omega} S_T(\omega)P(d\omega) = \int_{\mathbb{R}} S_0 e^{\nu T + \sigma \sqrt{T}x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = S_0 e^{(\nu + \sigma^2/2)T}.$$ 

Here we call $\mu = \nu + \sigma^2/2 = 17\%$ the (instantaneous) return rate. Setting $T = 1$, $\nu = 15\%$, $\sigma = 20\%$, $S_0 = $40.00, we see that one year from now, 100 unit shares of ABC’s company stock is expected to worth about

$$100 \times \$40.00 \times e^{(0.15 + 0.2^2/2) \times 1} = \$4741.22.$$ 

Thus, it is reasonable to write in the future contract $\$4741.22$ as the delivery price for 100 units of ABC company’s stock one year from now. We call $E[S_T] = \$4741.22$ the expected price. Intuitively, this should be the price that both parties would agree on. Well, this is the case until Black–Scholes paper which twisted people’s logic.

Now suppose that on the market there is a risk-free investment opportunity (typically modelled by government bonds) which can be traded, long or short, freely without any transition cost or whatsoever. Mathematically, this is a reasonable assumption since we are dealing with fair games: no one should be in the advantage or disadvantage of borrowing or lending money; namely, the interest rate of lending and borrowing should be exactly the same.

Assume that the annual risk-free interest rate $r = 10\%$. Then according to Black–Scholes, the future price for such a contract should be a multiple $e^{rT}$ of the current price; that is, the contract price to deliver 100 unit shares of the stock one year from now should be

$$100 \times \$40.00 \times e^{0.10 \times 1} = \$4420.68.$$ 

We call $\$4420.68$ the arbitrage-free price, which is determined by the risk-free interest rate and current cost of the commodity.

Black and Scholes argue that the arbitrage-free price should be the fair price entered into the contract. This price has nothing to do with the probabilistic expected price of the stock. Here is the reasoning.

Suppose the price in the contract is $P$ and $P$ is smaller than the arbitrage price. Then one enters the contract in the stock buyer’s side, short sells 100 share of stock obtaining $\$40 \times 100 = \$4,000$ and
Consider a one period model. There is a stock whose current unit share price is \( X \) out of money, and to obtain a guaranteed future payment exactly matches probabilistic event \( \omega \). It is expected that at the end of period \( T \), the portfolio will exactly offset the option claim. From the portfolio, find the price of the call option.

Namely, fine an initial investment of \( n^S \) share of stocks and \( n^B \) share of bounds, so that at the end of the period, the portfolio exactly offset the option claim. From the portfolio, find the price of the call option.

Also, find the price from expectation.

Exercise 2.1. Consider a one period model. There is a stock whose current unit share price is \( S_0 = 50 \). It is expected that at the end of period \( T \) there is a 2/3 chance that the price will become \( S_T = \frac{5}{6} 60 \) and a 1/3 chance that the price will be \( S_T = 45 \). Also there is a risk-free bond which gives a guaranteed 10% return per period.

Consider a European Call option which gives option buyer right, but not obligation to buy a unit share of stock at \$50 at the end of first period. Find a replication portfolio that duplicate the claim. Namely, fine an initial investment of \( n^S \) share of stocks and \( n^B \) share of bounds, so that at the end of the period, the portfolio exactly offset the option claim. From the portfolio, find the price of the call option.

Also, find the price from expectation.

Exercise 2.2. Assume that \( \Omega = \{ \omega_1, \ldots, \omega_8 \} \), \( \mathcal{F} = 2^\Omega \), and \( \mathbb{P}(\omega_i) = 1/8 \) for each \( i = 1, \ldots, 8 \). Let \( \mathcal{F}_1 \) be the smallest \( \sigma \)-algebra generated by \( \{ \omega_1, \{ \omega_2, \omega_3 \}, \{ \omega_4, \omega_5, \omega_6, \omega_7, \omega_8 \} \} \). Assume that \( S \) is a random variable given by \( S(\omega_i) = 1 + i + i^2 \) for all \( i = 1, \ldots, 8 \). Find \( \mathbb{E}[S \mid \mathcal{F}_1] \).

## 2.2 The Black–Scholes Equation

In the sequel, we shall use \( \{ W_t \}_{t \in [0, T]} \) to denote a standard Wiener process on \( (\Omega, \mathcal{F}_T, \mathbb{P}) \) adapted to a filtration \( \{ \mathcal{F}_t \}_{t \geq 0} \). The notation \( \{ B_t \} \) will have nothing to do with Brownian motion.

Now we consider a simple situation that there are two securities, one is a stock and the other is a risk free bound. Let \( S_t \) and \( B_t \) be the unit prices of the securities at time \( t \). We assume that \( \{ S_t \}_{t \geq 0} \) is
a stochastic process satisfying the sde
\[ dS_t = S_t (\mu dt + \sigma dW_t); \]

Here \( \mu \) is the **instant return rate** and \( \sigma > 0 \) is the **volatility**. Meanwhile, we assume that the bond has a **risk-free interest rate** \( r \) so we can write
\[ dB_t = r B_t dt \quad \text{(or } B_t = B_0 e^{-\int_0^t r(s) ds}). \]

Consider a contract which has a claim \( X \) at a specific future maturity date \( T \). In the original Black–Scholes paper, one of such claim is a **European call option**—the right, not obligation, to buy a unit share of stock at time \( T \) with price \( k \). In such a case, at time \( t = T \), if \( S_T > k \), the option holder has a net profit of \( X = S_T - k \); if \( S_T \leq k \), the option is automatically voided, and the option holder gets not profit, \( X = 0 \). Combining both situations, we see that the call option has a claim \( X = \max(S_T - k, 0) \), upon the option writer. The option writer will receive a payment at the time of selling the call option from the option buyer in exchange for the future obligation of paying \( \max(S_T - k, 0) \) to the option buyer. To fulfill the obligation, the option writer seeks a replicating strategy so that no matter where the stock price lands at the maturity date, the replication portfolio always ends up with an exact payment for claim \( X \); that is, if the stock price \( S_T \) is bigger than \( k \), then there is one share of stock and \( -k \) cash in the replicating portfolio; if the stock price \( S_T \) is less than \( k \), then the amount of stock available in the portfolio worths exactly the amount of cash own to the risk-free account, so the net worth of the portfolio is exactly zero. The initial cost for the replicating portfolio will then be the base price charged to the buyer.

In the sequel, we consider the situation that \( X = f(S_T) \), where \( f : (0, \infty) \to \mathbb{R} \) is a given function. For example, for the above European call option with strike price \( k = 10 \), \( f(x) = \max(x - k, 0) = \max(x - 10, 0) \).

Let’s use \( \{(n^S_t, n^B_t)\}_{t \in [0,T]} \) for a portfolio where \( n^S_t \) and \( n^B_t \) are the numbers of units of share of stock and bond respectively. Denote the total value of the portfolio by \( V_t \). Then
\[ V_t = n^B_t B_t + n^S_t S_t. \]

For the portfolio to be self-financing, we need
\[ dV_t = n^B_t dB_t + n^S_t dS_t = \sigma n^S_t S_t dW_t + (\mu n^S_t S_t + rn^B_t B_t) dt. \]

Without a clue how to construct a replicating strategy, we seek the possibility of such a portfolio that there is a function \( V(s, t) : (0, \infty) \times [0, T] \to \mathbb{R} \) such that
\[ V_t = V(S_t, t). \]

Let’s see if it is possible. By Itô lemma, we know that
\[
\begin{align*}
    dV(S_t, t) &= \frac{\partial V}{\partial s} dS_t + \frac{\partial V}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V}{\partial s^2} (dS_t)^2 \\
    &= \sigma S_t \frac{\partial V(S_t, t)}{\partial s} dW_t + \left( \frac{\partial V(S_t, t)}{\partial s} \mu S_t + \frac{\partial V(S_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 V(S_t, t)}{\partial s^2} \sigma^2 S_t^2 \right) dt.
\end{align*}
\]

From here, we see that a self-financing trading strategy can be constructed out of the given function \( V \) if and only if
\[
\begin{align*}
    V(S_t, t) &= n^S_t S_t + n^B_t B_t, \\
    \sigma n^S_t S_t &= \sigma S_t \frac{\partial V(S_t, t)}{\partial s}, \\
    \mu n^S_t S_t + rn^B_t B_t &= \frac{\partial V(S_t, t)}{\partial s} \mu S_t + \frac{\partial V(S_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 V(S_t, t)}{\partial s^2} \sigma^2 S_t^2
\end{align*}
\]
This is equivalent to design the portfolio by taking

\[ n_i^S = \frac{\partial V(S_t, t)}{\partial s}, \quad n_i^B = \frac{1}{r B_t} \left\{ \frac{\partial V(S_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 V(S_t, t)}{\partial s^2} \sigma^2 S_t^2 \right\}. \]  

(2.1)

This portfolio has the value \( V_t = V(S_t, t) \) if and only if the function \( V(\cdot, \cdot) \) is a priori chosen to satisfy

\[ V(s, t) = s \frac{\partial V(s, t)}{\partial s} + \frac{1}{r} \left\{ \frac{\partial V(s, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 V(s, t)}{\partial s^2} \sigma^2 s^2 \right\} \quad \forall s > 0, t \in [0, T). \]

After simplification, the above equation reads

\[ \frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial V(s, t)}{\partial s^2} + r s \frac{\partial V(s, t)}{\partial s} - r V(s, t) = 0 \quad \forall s > 0, t \in [0, T) \]

(2.2)

This is the famous Black-Scholes equation.

Now let’s see what we have obtained.

Suppose \( V \) is continuous and satisfies the Black–Scholes equation. Consider a portfolio \( \{ (n_i^S, n_i^B) \}_{t \in [0,T]} \) where \( n_i^S \) and \( n_i^B \) are given by the formula (2.1). This portfolio has the following properties:

1. \( \{ (n_i^S, n_i^B) \} \) is a trading strategy since it depends only on \( S_t \) and \( B_t \); namely, as long as one knows the stock price and bound price, one knows how to organize the portfolio. (Here we are dealing with continuous processes, so non-anticipating and pre-visible are the same).

2. The total value of the portfolio at time \( t \) is

\[ V_t = n_i^S S_t + n_i^B B_t = s \frac{\partial V(s, t)}{\partial s} + \frac{1}{r} \left\{ \frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V(s, t)}{\partial s^2} \right\} \bigg|_{s=S_t} = V(S_t, t) \]

by the differential equation satisfied by \( V \).

3. Finally, we find that the infinitesimal change of the value \( dV_t \) of the portfolio is, by Itô’s formula,

\[ dV_t = dV(S_t, t) = \frac{\partial V(s, t)}{\partial s} dS_t + \left( \frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V(s, t)}{\partial s^2} \right) dt \bigg|_{s=S_t} = n_i^S dS_t + n_i^B dB_t \]

by the definition of \( n_i^B \) and \( n_i^S \). Thus, we have a self-financing trading strategy.

If \( V \) is a continuous function satisfying (2.2), then the portfolio \( \{ (n_i^S, n_i^B) \}_{t \in [0,T]} \) given by (2.1) is a self-financing trading strategy, whose value at time \( t \) is \( V(S_t, t) \) for any \( t \in [0, T] \).

Equation (2.6) has infinitely many solutions so there are infinitely many self-financing trading strategies. What we are looking for is a strategy that replicates the claim \( X = f(S_T) \). That is, we want a particular strategy such that \( V_T = X \), equivalently,

\[ X(\omega) = f(S_T(\omega)) = V(S_T(\omega), T) \quad \forall \omega \in \Omega. \]

Since \( \{ S_T(\omega) \mid \omega \in \Omega \} \) is typically the set \( (0, \infty) \), to guarantee the validity of the above relation for every possible event \( \omega \in \Omega \), it is necessary and sufficient to require \( V \) to satisfies the terminal condition

\[ V(s, T) = f(s) \quad \forall s > 0. \]  

(2.3)
One can show that (2.2) subject to the terminal condition (2.3) admits a unique solution. Thus we have the following.

Any claim of the form \( X = f(S_T) \) admits a unique replicating strategy, obtained by setting a portfolio of \( n_S^i \) unit shares of stock and \( n_B^i \) unit shares of bond at time \( t \) where \((n_S^i, n_B^i)\) are determined from formula (2.1) with \( V \) the solution to (2.2)–(2.3). Consequently, the arbitrage-free price of the claim at time \( t = 0 \) is \( V(S_0, 0) \). At any time \( t < T \) if the stock price is \( S_t \), then the arbitrage-free value of the claim is \( V(S_t, t) \).

One observes that the return rate \( \mu \) does not appear the Black–Scholes equation, and therefore also not in the final price formula.

Exercise 2.3. Consider a discrete model of three steps: \( T = \{0, 1, 2, 3\} \). Let \( \{W_t\}_{t \in T} \) be the random walk of time step \( \Delta t = 1 \) and space step \( \Delta x = 1 \). There are a total of 8 different sample paths, which we denote by \( \Omega = \{\omega_1, \cdots, \omega_8\} \).

Assume that \( S_t = 10e^{\nu t + \sigma W_t} \) where \( \nu = 0.1 \) and \( \sigma = 0.2 \). Also assume that \( B_t = e^{rt} \) with \( r = 0.05 \). Find a replication portfolio for a claim \( X \) defined by \( X = \max\{S_3 - 11, 0\} \).

Exercise 2.4. Consider a Black–Scholes model in which there are two risky securities and one risk free bond. At each time \( t \), the unit share prices of risky securities are \( S_t^1 \) and \( S_t^2 \) and the unit price of the bond is \( B_t \). They obey

\[
\frac{dB_t}{B_t} = r, \quad \frac{dS_t^1}{S_t^1} = \mu_1 dt + \sigma_{11} dW_t^1 + \sigma_{12} dW_t^2, \quad \frac{dS_t^2}{S_t^2} = \mu_2 dt + \sigma_{21} dW_t^1 + \sigma_{22} dW_t^2
\]

where \( \{(W_t^1, W_t^2)\} \) is the standard Wiener process satisfying \( dW_t^i dW_t^j = dt \) if \( i = j \) and \( = 0 \) if \( i \neq j \).

Suppose \( V(x_1, x_2, t) \) is a smooth function of three variables. The Itô formula says that

\[
dV(S_t^1, S_t^2, t) = \sum_{i=1}^{2} \frac{\partial V}{\partial x_i} dS_t^i + \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} (\sigma_{i1} \sigma_{j1} + \sigma_{i2} \sigma_{j2}) S_t^i S_t^j \right) dt \bigg|_{x_1=S_t^1, x_2=S_t^2}.
\]

Derive the Black–Scholes partial differential equation for \( V \) such that \( V(S_t^1, S_t^2, t) \) is the value at time \( t \) of a self-financing trading strategy.

### 2.3 Solutions to the Black–Scholes Equation

Now we solve the Black–Scholes equations (2.2)–(2.3). There are at least two ways to solve it, one is a pde method, the other is a probabilistic method. To obtain closed formula, we assume that \( r \) and \( \sigma > 0 \) are constants.

#### 2.3.1 A PDE Method

First we use the pde method. As we are dealing with geometric motion, it is convenient to use variables \((x, \tau)\) instead of \((s, t)\) where

\[ s = e^x, \quad t = T - \tau \quad \Leftrightarrow \quad x = \ln s, \quad \tau = T - t. \]

Note that \( \tau \geq 0 \) is the amount of time left to the expiration data \( T \), simple called time to expiration. Write

\[ V(s, t) = v(x, \tau)|_{x=\ln s, \tau=T-t}. \]
Using the chain rule, we find that
\[
\frac{\partial V}{\partial s} = \frac{\partial v}{\partial x} \frac{\partial x}{\partial s}, \\
\frac{\partial^2 V}{\partial s^2} = -\left(\frac{1}{s} \frac{\partial v}{\partial x}\right) + \frac{1}{s^2} \frac{\partial v}{\partial x} \frac{\partial x}{\partial s}, \\
\frac{\partial V}{\partial t} = \frac{\partial v}{\partial \tau} \frac{\partial \tau}{\partial t} = -\frac{\partial v}{\partial \tau}.
\]

Thus, the problem for \(V(s,t)\) is equivalent to the problem for \(v(x,\tau)\):
\[
\frac{\partial v(x,\tau)}{\partial \tau} = \frac{\sigma^2}{2} \frac{\partial^2 v(x,\tau)}{\partial x^2} + \left(r - \frac{\sigma^2}{2}\right) \frac{\partial v(x,\tau)}{\partial x} - rv(x,\tau) \quad \forall x \in \mathbb{R}, \tau > 0,
\]
\[
v(x,0) = f(e^x) \quad \forall x \in \mathbb{R}.
\]

Note that this is a linear equation with constant coefficients! We can further simplify the problem by using new variables
\[
y = x + \alpha \tau, \quad v(x,\tau) = u(y,\tau)e^{\beta \tau}
\]
where \(\alpha\) and \(\beta\) are constant to be determined. By the chain rule, the equation for \(v\) is equivalent to the following equation for \(u\), using subscripts for partial differentiation,
\[
(u_\tau + u_y \alpha)e^{\beta \tau} + \beta ue^{\beta \tau} = \frac{1}{2} \sigma^2 u_{yy} e^{\beta \tau} + (r - \frac{1}{2} \sigma^2) u_y e^{\beta \tau} - ru e^{\beta \tau}.
\]
Hence, taking
\[
\alpha = r - \frac{1}{2} \sigma^2, \quad \beta = -r,
\]
we see that \(u(y,t)\) solves the heat equation
\[
\begin{cases}
  u_\tau(y,\tau) = \frac{1}{2} \sigma^2 u_{yy}(y,\tau) & \forall y \in \mathbb{R}, t > 0, \\
  u(y,0) = f(e^y) & \forall y \in \mathbb{R}.
\end{cases}
\tag{2.4}
\]
Since we know that the pde for \(u\) is the same as that for the density function of the Brownian motion of standard-deviation \(\sigma\), we obtain the formula
\[
u(x,\tau) = e^{-r\tau} u(y,\tau) \big|_{y=x+[r-\sigma^2/2] \tau} = e^{-r\tau} \int \frac{f(e^{y+z\sqrt{\tau}})}{\sqrt{2\pi}} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz.
\]
Hence,
\[
v(x,\tau) = e^{-r\tau} u(y,\tau) \big|_{y=x+[r-\sigma^2/2] \tau} = e^{-r\tau} \int f(e^{x+z\sqrt{\tau}}) e^{-z^2/2} \sqrt{2\pi} dz.
\]
Finally, we obtain
\[
V(s,t) = e^{-r(T-t)} \int f(se^{[r-\sigma^2/2][T-t]+\sigma z \sqrt{T-t}}) e^{-z^2/2} \sqrt{2\pi} dz.
\]
We then obtain the following:
Thus, setting \( t \in [0, T) \) with unit share stock price \( S_t = s \), the value of the a claim \( X = f(S_T) \) at \( T \) is
\[
V_t = V(s, t) = e^{-r(T-t)} \int_{\mathbb{R}} f(s e^{(r-\sigma^2/2)(T-t)+\sigma z\sqrt{T-t}}) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz
\]
In particular, its initial price is
\[
V_0 = V(S_0, 0) = e^{-rT} \int_{\mathbb{R}} f(s_0 e^{(r-\sigma^2/2)T+\sigma z\sqrt{T}}) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz
\]

### 2.3.2 Probabilistic Method

Next we use a probabilistic method. Fix an arbitrary \( t_1 < T \). Consider the discounted stock price
\[
\tilde{S}_t := e^{(r-\mu)(t-t_1)} S_t.
\]
We find that
\[
d\tilde{S}_t = \tilde{S}_t \{ rdtd + \sigma dW_t \}.
\]
Now using the Feynman-Kac Formula with \( k = -r, g = 0 \) we find that
\[
V(s, t_1) = \mathbb{E}[f(\tilde{S}_T) e^{-r(T-t_1)} | \tilde{S}_{t_1} = s] = e^{-r(T-t_1)} \mathbb{E}[f(S_T e^{(r-\mu)(T-t_1)} | S_{t_1} = s)] \quad \forall s > 0, t_1 \in [0, T).
\]
Thus, setting \( t_1 = t \in [0, T) \) and \( s = S_t \) we obtain
\[
V_t \bigg|_{S_t = s} = V(s, t) = e^{-r(T-t_1)} \mathbb{E} \left[ f(S_T e^{(r-\mu)(T-t)} | S_t = s) \right].
\]
This formula can also be written as
\[
V_t = V(S_t, t) = e^{-r(T-t)} \mathbb{E} [ f(S_T e^{(r-\mu)(T-t)} ) | \mathcal{F}_t ].
\]
In the sequel we explain that the probabilistic solution and the pde solution are indeed the same.

We recall that the solution to \( dS_t = S_t(\mu dt + \sigma dW_t) \) is given by
\[
S_t = S_0 e^{[\mu-\sigma^2/2]t+\sigma W_t}, \quad S_T = S_t e^{[\mu-\sigma^2/2](T-t)+\sigma(W_T-W_t)}.
\]
The probabilistic solution then gives
\[
V(S_t, t) = e^{-r(T-t)} \mathbb{E} [ f(S_t e^{[r-\sigma^2/2](T-t)+\sigma(W_T-W_t)} ) | \mathcal{F}_t ].
\]
Since \( W_T - W_t \) is \( N(0, T-t) \) distributed, we then obtain
\[
V(S_t, t) = e^{-r(T-t)} \int_{\mathbb{R}} f(S_t e^{[r-\sigma^2/2](T-t)+\sigma z}) \frac{e^{-z^2/(2(T-t))}}{\sqrt{2(T-t)}} dz.
\]
This formula is the same as that derived from the PDE solution.
2.3.3 The Risk-Neutral Measure

Now we investigate the initial price of the claim. We have

\[ V_0 = e^{-rT}E[f(S_0e^{(r-\sigma^2/2)T+\sigma W_T})] = e^{-rT} \int_{\mathbb{R}} f(S_0e^{(r-\sigma^2/2)T+\sigma z\sqrt{T}}) e^{-z^2/2} \frac{1}{\sqrt{2\pi}} dz. \]

Using the change of variable \( z = \tilde{z} + (\mu - r)\sqrt{T}/\sigma \) we have

\[ V_0 = e^{-rT} \int_{\mathbb{R}} f(S_0e^{(\mu-\sigma^2/2)T+\tilde{z}\sqrt{T}}) e^{-(\tilde{z}+|\mu-r|\sqrt{T}/\sigma)^2/2} \frac{1}{\sqrt{2\pi}} d\tilde{z} \]
\[ = e^{-rT} \int_{\mathbb{R}} f(S_0e^{(\mu-\sigma^2/2)T+\tilde{z}\sqrt{T}}) e^{-(\tilde{z}-(\mu-r)\sqrt{T}/\sigma)^2/2} \frac{1}{\sqrt{2\pi}} d\tilde{z} \]
\[ = e^{-rT} \int_{\Omega} f(S_0e^{(\mu-\sigma^2/2)T+\sigma W_T}) e^{-(\mu-r)W_T(\omega)}/\sigma-(\mu-r)^2T/(2\sigma^2) e^{-z^2/2} \frac{1}{\sqrt{2\pi}} dz \]
\[ = e^{-rT} \int_{\Omega} X(\omega)e^{-(\mu-r)W_T(\omega)/\sigma-(\mu-r)^2T/(2\sigma^2)} d\omega \]

since \( W_T(\cdot)/\sqrt{T} \) under the measure \( \mathbb{P} \) is \( N(0,1) \) distributed and \( X = f(S_T) = f(S_0e^{\mu-\sigma^2/2}T+\sigma W_T) \).

Now we introduce another measure \( \mathbb{Q} \) on the Wiener process \( \{W_t\} \) on \( (\Omega, \{\mathcal{F}_t\}, \mathbb{P}) \) via

\[ \frac{\mathbb{Q}(d\omega)}{\mathbb{P}(d\omega)} := e^{-(\mu-r)W_T(\omega)/\sigma-(\mu-r)^2T/(2\sigma^2)} \]

It is easy to see that \( \mathbb{Q} \) is a probability measure. Then we have

\[ V_0 = e^{-rT} \int_{\Omega} X(\omega)\mathbb{Q}(d\omega) = \mathbb{E}_\mathbb{Q}[e^{-rT}X]. \]

where \( \mathbb{E}_\mathbb{Q} \) means the expectation under the measure \( \mathbb{Q} \).

We note that \( e^{-rT}X \) is the discounted value of the claim. In the modern approach of the Black–Scholes theory, the probability measure \( \mathbb{Q} \) is called the risk-neutral measure. Hence, we have the following simple explanation of the Black–Scholes theory:

The initial price of a claim \( X = f(S_T) \) is the expectation, under the risk-neutral probability measure, of the discounted claim \( e^{-rT}X \). At any time \( t < T \), the value of the claim is \( \mathcal{F}_t \) conditional expectation of the discounted claim under the risk-neutral measure:

\[ V_t = V(S_t, t) = \mathbb{E}_\mathbb{Q}[e^{-r(T-t)}X|\mathcal{F}_t]. \]

In the typical approach, the risk-neutral measure is obtained via the change of measure from \( \mathbb{P} \) to \( \mathbb{Q} \) for the Wiener process so that the discounted stock price \( S_t e^{-rt} \) is a martingale. Then using the unique martingale representation, one derives the formula for the arbitrage price.

Risk-neutral probability measure, also called the martingale measure, is an important concept in the modern mathematical layout of the Black–Scholes theory.

In the case when \( \sigma \) and \( r \) are constants, the risk-neutral measure \( \mathbb{Q} \) can be calculated, so the representation of the martingale can be obtained explicitly. Nevertheless, when \( \sigma \) and/or \( r \) are not constants, one only knows the existence of a replicating portfolio, i.e., the existence of a martingale representation of the discounted claim. To calculate the exact value, one has to solves directly the Black-Scholes PDE problem, for which one does not need martingales or their measures.

Hence, martingale measure can explain arbitrage–free price in a straight forward way, the Black–Scholes equation can be used to solve the problem numerically.
CHAPTER 2. THE BLACK–SCHOLES THEORY

Exercise 2.5. Assume that \( f \in C(\mathbb{R}; \mathbb{R}) \) is uniformly bounded. Show that the solution to
\[
\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2} \quad \forall x \in \mathbb{R}, t > 0, \quad u(x, 0) = f(x) \quad \forall x \in \mathbb{R}
\]
is given by
\[
u(x, t) = \int_{\mathbb{R}} f(z) e^{-\frac{(x-z)^2}{4t}} \sqrt{\frac{4}{\pi}} \, dz = \int_{\mathbb{R}} f(x - 2\sqrt{t}z) e^{-\frac{-z^2}{\pi}} \, dz.
\]

Exercise 2.6. Complete the proof that solution \( V_t \) from the pde method agrees with that from the probabilistic method.

Exercise 2.7. Show that \( Q \) is a probability measure.

Exercise 2.8. Consider a discrete model of three steps: \( T = \{0, 1, 2, 3\} \). Let \( \{W_t\}_{t \in T} \) be the random walk with \( \Delta t = 1 \) and \( \Delta x = 1 \). Denote by \( \omega_1, \cdots, \omega_8 \) the eight-total samples paths. Set \( \Omega = \{\omega_1, \cdots, \omega_8\} \). Assume that \( S_t = S_0 e^{rT + \sigma W_t} \) and \( B_t = e^{rt} \) for all \( t \in T \). Find the risk-neutral probability measure \( Q \) on \( \Omega \). That is, find \( Q(\omega_i) \) for \( i = 1, \cdots, 8 \) such that every contingent claim \( X \) at \( T = 3 \) has initial price given by
\[
V_0 = e^{-rT} E_Q[X] = e^{-rT} \sum_{i=1}^{8} X(\omega_i) Q(\omega_i).
\]

2.4 Examples

2.4.1 European Options

A European call option is a right, but not obligation to buy a unit share of a stock at a predetermined price, say \( k \), and at a particular exercise date, say \( T \).

A European put option is a right, but not obligation to sell a unit share of a stock at a predetermined price \( k \), and at a particular exercise date \( T \).

In the Black–Scholes model presented in the earlier sections, we are dealing with contingent claims
\[
X_{EC} = \max\{S_t - k\} =: (S_t - k)^+, \quad X_{EP} = \max\{k - S_T, 0\} =: (k - S_T)^+,
\]
where \( EC \) stands for European call and \( EP \) stands for European put; also, \(( \cdot )^+\) stands the positive part of the argument.

Using the formula derived earlier, we find that
\[
V(s, t) = e^{-r(T-t)} \int_{\ln(k/s) - (r-\sigma^2/2)(T-t)/[\sigma \sqrt{T-t}]}^{\infty} \left( se^{(r-\sigma^2/2)(T-t)+\sigma \sqrt{T-t}}} - k \right) e^{-z^2/2} \sqrt{2\pi} \, dz.
\]

We use \( \Phi \) to denote the cumulative normal integral
\[
\Phi(x) := \int_{-\infty}^{x} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz.
\]

Then the above gives the Black–Scholes formula for the price of the European call option.
The value of the European call option at time $t$, stock price $S_t = s$ is

$$V(s, t) = s \Phi\left( \frac{\ln \frac{se^{r(T-t)}}{k} + \frac{1}{2} \sigma^2 (T-t)}{\sigma \sqrt{T-t}} \right) - k e^{-r(T-t)} \Phi\left( \frac{\ln \frac{se^{r(T-t)}}{k} - \frac{1}{2} \sigma^2 (T-t)}{\sigma \sqrt{T-t}} \right).$$

To hedge the option, the portfolio consists of $n^S_t$ unit share of stock and $n^B_t$ unit share of bond where

$$n^S_t = \frac{\partial V}{\partial s} \bigg|_{s=S_t} = \Phi\left( \frac{\ln \frac{S_te^{r(T-t)}}{k} + \frac{1}{2} \sigma^2 (T-t)}{\sigma \sqrt{T-t}} \right),$$
$$n^B_t = -\frac{e^{-rT}}{B_0} \Phi\left( \frac{\ln \frac{S_te^{r(T-t)}}{k} - \frac{1}{2} \sigma^2 (T-t)}{\sigma \sqrt{T-t}} \right).$$

### 2.4.2 Foreign Exchange

In the foreign exchange market, like the stock market, holding the basic asset, the currency, is a risky business. The dollar value of, say one UK pound, varies from moment to moment just as a US stock does. And with this risk comes demand for derivatives: claims based on the future value of one unit of currency in terms of another.

Let’s consider the following forward contract problem:

At a future date $t = T > 0$, one party is to exchange one pound sterling to $F_0$ US dollars.
At the current time $t = 0$, how to settle a reasonable value $F_0$?

This situation occurs very often in companies having international business. In the above example, the income is to be received at time $T$ in UK pound and has to be changed to US dollars to pay domestic investors. To reduced the risk of foreign exchange rate, one considers a froward contract as above.

Let’s assume that the dollar has an constant interest rate $r$ and pound has interest rate $u$. Hence, an initial one dollar cash bond will become $D_t = e^{rt} ($) at time $t$ and an initial one pound bond will become $e^{ut}$ at time $t$. We assume that the exchange rate is a stochastic process $C_t = C_0 e^{\sigma W_t + \nu t}$ where $\{W_t\}$ is a Brownian motion, $\sigma > 0$ is a constant, and $\nu$ is a constant, positive of negative.

#### Black–Scholes Currency Model

<table>
<thead>
<tr>
<th></th>
<th>$D_t = e^{rt} ($)</th>
<th>$P_t = e^{ut}$ (\pounds),</th>
<th>$C_t = C_0 e^{\sigma W_t + \nu t}$ ($ \pounds)\bigg(\frac{$}{\pounds}\bigg)$.</th>
</tr>
</thead>
</table>

We consider a portfolio $\{(n^D_t, n^P_t)\}_{t \in [0,T]}$ consists of $n^D_t$ share of dollar bond and $n^P_t$ share of sterling bond at time $t \in [0,T]$.

Let’s use dollar as our numeraire. The dollar value at time $t$ of the portfolio is

$$V_t = n^D_tD_t + n^P_tP_tC_t.$$
Assume that this is a self-financing trading strategy. Then we need

\[ dV_t = n_t^D dD_t + n_t^P d(P_t C_t). \]

If we compare with the option model, we see that we can regard \( S_t := P_t C_t \) as the stock price, and regard \( D_t \) as the risk-free bond. Hence,

\[
dV_t = n_t^D dD_t + n_t^P dS_t = r n_t^D D_t dt + n_t^P dS_t.
\]

Note that, since \( P_t \) is deterministic,

\[
ds_t = d(P_t C_t) = dP_t C_t + P_t dC_t = P_t C_t \{[\mu + u] dt + \sigma dW_t\} = S_t \{[\mu + u] dt + \sigma dW_t\}
\]

where \( \mu = \nu + \sigma^2/2 \).

As before, we seek a smooth function \( V(s, t) \) such that \( V_t = V(S_t, t) \). If this is true, we have

\[
dV_t = dV(S_t, t) = \frac{\partial V(s, t)}{\partial s} dS_t + \left\{ \frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \partial^2 V(s, t) \right\} dt \bigg|_{s=S_t}.
\]

This gives us the same sort of equation for the value \( V_t \) of the portfolio and the trading strategy \( \{(n_t^D, n_t^P)\} \):

\[
V(s,t) = n_t^D(s,t) D_t + n_t^P(s,t) s,
\]

\[
n_t^P(s,t) = \frac{\partial V(s,t)}{\partial s},
\]

\[
n_t^D(s,t) = \frac{1}{r D_t} \left\{ \frac{\partial V(s,t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V(s,t)}{\partial s^2} \right\}
\]

This gives us the familiar Black-Scholes equation

\[
\frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V}{\partial s^2} + r s \frac{\partial V}{\partial s} = r V \quad \forall s \in (0, \infty), t < T.
\]

At time \( T \), if we agree to buy one UK pound with \( F_0 \) US dollar, the payoff at time \( T \) is

\[ X = C_T - F_0 = \frac{S_T}{P_T} - F_0. \]

Hence, in order for the self-financing portfolio to replicating the payment, we need

\[ V(s,T) = \frac{s}{P_T} - k = e^{-uT} s - F_0, \quad \forall s > 0. \]

It is easy to verify that the exactly solution for \( V \) is

\[ V(s,t) = se^{-uT} - F_0 e^{r(t-T)} \quad \forall s > 0, t \leq T. \]

To be fair, we need \( V(S_0, 0) = 0 \). Recalling that \( S_t = C_t P_t \), we hence have

\[ F_0 = \frac{S_0 e^{-uT}}{e^{-rT}} = C_0 e^{(r-u)T}. \]

Thus, the at time \( t = 0 \), the exchange rate from pound to dollar at time \( T \) is \( k = C_0 e^{(r-u)T} \) (\$/\£); this is the current price for sterling discounted by a factor depending on the difference between interest rates of the two currencies.

Similarly, we can derive the following:
At any time $t < T$ and currency exchange rate $C_t$ ($\$/\£$), the rate $F_t$ to be determined at time $t$ for pound–dollar exchanged at time $T$ is

$$F_t = C_t e^{(r - \mu)(T - t)}.$$

We call $F_t$ the **forward exchange rate**; it is an exchange rate for time $T$ that is determined at time $t < T$.

Now we consider a European call option on foreign exchange:

What is the price of selling an option having the right, but not obligation, to buy, at time $T$, one UK pound at $k$ US dollars?

In this example, we know the payoff at time $T$ is

$$X = \max\{C_T - k, 0\} = \max\{S_T e^{-uT} - k, 0\}.$$

Thus, we seek a solution to the Black-Scholes equation with terminal value

$$V(s, t) = \max\{se^{-uT} - k, 0\}.$$ 

Using the method as before, we find that the price of the option is

$$V_0 = V(s_0, 0) = e^{-rT} \left\{ F_0 \Phi \left( \frac{\log F_0 - \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}} \right) - k \left( \frac{\log F_0 - \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}} \right) \right\}.$$ 

The hedging for the portfolio is

$$n_t^D = -ke^{-rT} \Phi \left( \frac{\log F_t - \frac{1}{2} \sigma^2 (T - t)}{\sigma \sqrt{T - t}} \right),$$

$$n_t^P = e^{-uT} \Phi \left( \frac{\log F_t - \frac{1}{2} \sigma^2 (T - t)}{\sigma \sqrt{T - t}} \right)$$

where $F_t = e^{(r - t)(T - t)} C_t$ is the forward exchange rate.

### 2.4.3 Equities and Dividends

An equity is a stock makes periodic cash payment to the current holder. Our previous models treated a stock as a pure asset, but they can be modified to handle dividend payments.

**Equity model with continuous dividends**

The bond price $B_t$ is given by $B_t = e^{rt}$ and the stock price $S_t$ is given by $S_t = S_0 e^{r + \sigma W_t}$.

The dividend payment made in time interval of length $dt$ starting at time $t$ is

$$dQ_t = \delta S_t dt$$

where $\delta$ is constant of proportionality.
We now study self-financing trading strategies. Assume that at time $t$, the portfolio has $n_t^B$ unit share of bond and $n_t^S$ unit share of stock, so

$$V_t = n_t^B B_t + n_t^S S_t.$$ 

By self-financing, it means that the value change of the portfolio is totally due to the price change of stock and bond, plus the dividend. Hence, we need

$$dV_t = n_t^B dB_t + n_t^S (dS_t + dQ_t) = \left\{ r n_t^B B_t + \delta n_t^S S_t \right\} dt + n_t^B dS_t.$$ 

Again as before, we want $V_t = V(S_t, t)$ for some smooth function, we need

$$dV_t = dV(S_t, t) = \left. \frac{\partial V}{\partial s} dS_t + \left\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V}{\partial s^2} \right\} dt \right|_{s=S_t}. $$

This gives us the weights of the portfolio

$$n_t^S(s, t) = \frac{\partial V(s, t)}{\partial s},$$

$$n_t^B(s, t) = \frac{1}{r B_t} \left\{ \frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V(s, t)}{\partial s^2} - \delta s \frac{\partial V(s, t)}{\partial s} \right\}$$

and the Black–Scholes pde

$$\frac{\partial V(s, t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial^2 V(s, t)}{\partial s^2} + (r - \delta) s \frac{\partial V(s, t)}{\partial s} = r V(s, t).$$

In the case of European call option, the claim is $X = (S_T - k)^+$, so we have the terminal condition

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

Solve the pde we find that

$$V_0 = e^{-rT} \left\{ F_0 \Phi \left( \frac{\log F_0 + \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}} \right) - k \Phi \left( \frac{\log F_0 - \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}} \right) \right\}, \quad F_0 := e^{(r - \delta)T} S_0.$$ 

One can show that $F_0 = S_0 e^{(r - \delta)T}$ is the forward price of the stock, that is, the price, determined now, to buy one unit share of stock at time $T$.

### 2.4.4 Quantos

BP, short hand for British Petroleum, is a UK company having a sterling denominated stock price. But instead of thinking that stock price just in pounds, we could also consider it as a pure number which could be denominated in any currency. Contracts like this which pay off in the “wrong” currency are quantos.

Quantos are best described with examples. Here are three:

- A forward contract, namely receiving the BP stock price at time $T$ as if it were in dollars in exchange for a pre-agreed dollar amount;
- A digital contract which pays one dollar at time $T$ if the BP stock price is larger that some pre-agreed strike;
- An option to receive the BP stock price less a strike price, in dollars.
In each case, a simple derivative is given the added twist of paying off in a currency other than in which the underlying security is denominated.

### The Quanto model

- **sterling bond**: \( P_t = e^{u_t}(\mathcal{L}) \),
- **dollar bond**: \( D_t = e^{r_t}(\$) \),
- **exchange rate**: \( C_t = C_0 e^{\sigma_{11} W_{1t} + \sigma_{12} W_{2t} + \nu_1 t} (\$) \),
- **stock price**: \( S_t = S_0 e^{\sigma_{21} W_{1t} + \sigma_{22} W_{2t} + \nu_2 t} (\$) \).

where \( \{(W_{1t}, W_{2t})\}_{t \geq 0} \) is the standard Brownian motion, \( \text{Cov}(dW_{1t}, dW_{2t}) = \delta_{ij} dt \).

We remark that the stock price in sterling is
\[
S_t / C_t = S_0 / C_0 e^{(\nu_2 - \nu_1) + [\sigma_{21} - \sigma_{11}] W_{1t} + [\sigma_{22} - \sigma_{12}] W_{2t}} (\mathcal{L}) .
\]

Now we consider a portfolio \( \{n_t^S, n_t^P, n_t^D\} \) representing the number of units of stocks, sterling bond, and dollar bond in the portfolio at time \( t \), respectively.

Using the dollar denominator, the value of the portfolio is
\[
V_t = n_t^D D_t + n_t^S S_t + n_t^P C_t .
\]

For the portfolio to be self-financing, we need
\[
dV_t = n_t^D dD_t + n_t^S dS_t + n_t^P d(C_t P_t) .
\]

Set \( S_t = C_t P_t \) and \( S_t^2 = S_t \). We seek a smooth function \( (s_1, s_2, t) \) such that \( V_t = V(S_t^1, S_t^2, t) \).

We find
\[
dS_t^i = S_t^i \left\{ \mu_i dt + \sigma_{i1} dW_{1t} + \sigma_{i2} dW_{2t} \right\} .
\]

where
\[
\mu_1 = u + \nu_1 + \sigma_{11}^2 / 2 + \sigma_{12}^2 / 2, \quad \mu_2 = \nu_2 + \sigma_{21}^2 / 2 + \sigma_{22}^2 / 2 .
\]

Hence, we have
\[
dV(S_t^1, S_t^2, t) = \sum_{i=1}^{2} \frac{\partial V(s_1, s_2, t)}{\partial s_i} dS_t^i + \left\{ \frac{\partial V(s_1, s_2, t)}{\partial t} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial^2 V(s_1, s_2, t)}{\partial s_i \partial s_j} b_{ij} s_i s_j \right\} dt \bigg|_{s_1=S_1^t, s_2=S_2^t}
\]

where
\[
b_{ij} = \sigma_{i1} \sigma_{j1} + \sigma_{i2} \sigma_{j2} .
\]

Thus, what we need is to let
\[
n_t^S = n^S(S_t^1, S_t^2, t), \quad n_t^P = n^P(S_t^1, S_t^2, t), \quad n_t^D = n^D(S_t^1, S_t^2, t)
\]
where
\[ n(s_1, s_2, t) = \frac{\partial V(s_1, s_2, t)}{\partial s_2}, \]
\[ n^P(s_1, s_2, t) = \frac{\partial V(s_1, s_2, t)}{\partial s_1}, \]
\[ n^D(s_1, s_2, t) = \frac{1}{r} \mathbb{E} \left\{ \frac{\partial V(s_1, s_2, t)}{\partial t} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial^2 V(s_1, s_2, t)}{\partial s_i \partial s_j} b_{ij} s_i s_j \right\}. \]

The portfolio is a self-financing trading strategy is \( V \) solved the Black–Scholes equation
\[ \frac{\partial V(s_1, s_2, t)}{\partial t} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial^2 V(s_1, s_2, t)}{\partial s_i \partial s_j} b_{ij} s_i s_j + r s_1 \frac{\partial V}{\partial s_1} + r s_2 \frac{\partial V}{\partial s_2} = r V \quad \forall s_1 > 0, s_2 > 0, t < T. \]

Given a claim \( X = f(S^1_T, S^2_T) = f(C_T e^{-uT}, S_T) \) at time \( T \), the current dollar price of the claim is \( V(C_0, S_0, 0) \) where \( V(s_1, s_2, t) \) solves the Black–Scholes equation, together with the terminal condition
\[ V(s_1, s_2, T) = f(s_1, s_2) \quad \forall s_1 > 0, s_2 > 0. \]
Given a bounded \( f \), one can show that the BS equation together with the terminal condition admits a unique smooth solution.

When all parameters are constants, explicit formulas are available. We shall not derive the formula here.

Exercise 2.9. Find the change of variables thus verify the Black–Scholes formula for European put option. Also, verify the formula for the hedge.

Exercise 2.10. Find explicit formulas for the value of the European call option and the replicating portfolios.

Exercise 2.11. (Discrete Hedging) Consider a European call option with \( T = 3/12 \) (year), \( S_0 = \$40, k = \$41, r = 0.10 \) (1/year), \( \sigma = 0.20 \) (1/\( \sqrt{\text{year}} \)), and \( \mu = 0.15 \) (1/year). Then \( B_t = e^{rt} \) for all \( t \in [0, T] \).

1. Generate a sample path of the stock price \( \{S_t\}_{0 \leq t \leq T} \) as follows.
   Divide \([0, T]\) into \( N = 90 \) equal length intervals \([t_i, t_{i+1}]\) for \( i = 0, \ldots, N-1 \). Use a Random number generator generating a sample \( \{x_0, \ldots, x_{N-1}\} \) for i.i.d Binomial random variables \( \{X_0, X_1, \ldots, X_{N-1}\} \) satisfying \( \text{Prob}(X_i = \pm 1) = 1/2 \). Then set
   \[ s_i = s_{i-1} \left\{ \mu [t_i - t_{i-1}] + \sigma \sqrt{t_i - t_{i-1}} x_i \right\}, \quad \forall i = 1, \ldots, N. \]

   We regard \( \{s_i\}_{i=0}^N \) as a sample path of \( \{S_t\}_{t=0}^N \).

2. Use the Black-Scholes formula find the value \( V_0 \) and the unit of shares of stock \( n_i^S = n^S(s_i, t_i) \) to be held in the hedging portfolio at time \( t = t_i, i = 0, \ldots, N-1 \).

3. Find \( n_i^B \) such that \( V_0 = n_0^B B_{t_0} + n_0^S S_0 \).

4. For each \( i = 1, \ldots, N-1 \), first find the value of the portfolio at \( t = t_i \): \( V_i = n_{i-1}^B B_{t_i} + n_{i-1}^S s_i \). Then find \( n_i^B \) by using the formula \( V_i = n_i^B B_{t_i} + n_i^S s_i \), so no money is pumped in or taken out of the portfolio. Notice that \( n_i^B \) may differ from the Black–Scholes formula since we are dealing with discrete hedging strategy.
5. Calculate \( V_n = n_{N-1}^R B_T + n_{N-1}^S s_N \). This is the amount of money in the portfolio at the final call expiration date \( T \).

6. Calculate the difference \( V_N - X \) where \( X = \max\{s_N - k\} \).

7. Use a spreadsheet explaining the growth of the portfolio and the final payment against the claim \( X = \max\{S_T - k\} \).

Exercise 2.12. (Experiment on Option)


2. Use the quoted stock prices calculate the daily return

\[
R_i = \frac{s_{i+1} - s_i}{s_i}.
\]

Assume that there are 2500 trading days from Jan 1st 1996 to Jan 1st 2006. Using these 2500 daily returns find the mean return \( \mu \) and variance \( \sigma \) by

\[
\mu_{\text{day}} := \frac{1}{2500} \sum_{i=1}^{2500} R_i, \quad \sigma_{\text{day}} := \sqrt{\frac{\sum_{i=1}^{2500} (R_i - \mu_{\text{day}})^2}{N - 1}},
\]

\[
\mu = \frac{\mu_{\text{day}}}{250}, \quad \sigma = \frac{\sigma_{\text{day}}}{\sqrt{250}}.
\]

3. Denote by \( S_0 \) the stock price at Jan 1st 2006. Consider a European call option with strike price \( k = S_0 + \$1 \), \( T = 3/12 \) year, and risk-free interest rate \( r = 0.06 \) (1/year). Using the Black-Scholes formula calculate the price \( V_0 \) of the option.

4. Construct a hedging portfolio by using the actual stock daily prices from Jan 1st 2006 to March 31 2006. First use the Black-Scholes formula find the number \( n_S^i \) unit share of stocks at each trading day. Then use spreadsheet calculate the number of shares of bond that can be purchased from the money left from the portfolio, after the purchasing the required numbers of share of stocks. Ignoring any transaction costs.

5. At the last date, check the balance of the hedging portfolio against the claim from the call.

Exercise 2.13. Derive the price and hedging formulas for the call option on foreign exchanges.

Exercise 2.14. For the equity model with continuous dividends, show that at a current time \( t < T \), the forward stock price to be written in the forward contract to exchange a unit share of stock with cash \( F_t \) at time \( T \) is \( F_t = S_t e^{(r - \delta)(T - t)} \).

Also, derive the Black–Scholes formula for European call option, as well as the hedging strategy for the option.

Exercise 2.15. In the quanto model, first derive the Black–Scholes equation when sterling is used as the denominate for the value of the replication portfolio. Then show that the resulting replication portfolio are the same in using dollar or sterling as numeraire.
True multi-period investments fluctuate in values, distribute random dividends, exist in an environment of variable interest rates, and are subject to a continuing variety of uncertainties. By asset dynamics it means the change of values of assets with time. Two primary model types are used to represent asset dynamics: trees and stochastic processes.

Binomial tree models are finite state models based on the assumption that there are only two possible outcomes between each single period. Consider a binomial tree model for a stock price. If we use $U$ for up and $D$ for down, then in a $n$-period binary tree model, each state can be represented by a word of $n$ letters of only two symbols “U” and “D”. At the final time moment $T$, there are $2^n$ possible states. Quite often tree branches can be combined to form lattices. For example, if each U and D represent the increment of stock price by a factor of $u > 1$ and $d < 1$, respectively. Then after $k$ ups and $n - k$ downs, the stock price at time $T$ is $u^kd^{n-k}$ fold of its initial price. While there are $C_n^k$ many ways to reach this price, if only the stock prices are relevant to the problem, then we can combine all those nodes which give the same price. In this ways, the $2^n$ binary states can be replaced by an equivalent $n + 1$ states, resulting the commonly used binomial lattice model.

Stochastic process, or continuum model, on the other hand, are more realistic than binomial tree models in the sense that they have a continuum of possible stock prices at each period, not just two. The continuum models allow sophisticated analytical tools get involved so some problems can be solved analytically, as well as computationally. They also provide the foundation for constructing binomial lattice models in a clear and consistent manner (once the necessary mathematical tools are assessed). Stochastic process, particularly the Ito process, models are fundamental to dynamic problems.

Binomial tree models are conceptual easier to understand and analytically simpler than the Ito process. They provide an excellent basis for computational work associated with investment problems. We shall present a binary tree model known as the Cox-Ross-Rubinstein (CRR) Model [4] which is a specific example of finite state model present in the previous chapter.

From the CRR model, we shall take a limit to derive the famous Black–Scholes model [7]. Historically the paper of Black–Scholes came first and initial the modern approach to option’s evaluation. They discovered the rational option price derived from risk-neutral probability, an astonishing consequence of the seemingly trivial no-arbitrage assumption. Another earlier significant contribution was Merton [19, 18]. The simplified approach using binomial lattice was first presented by Sharpe in [29] and later developed by Cox, Ross, Robinstein [4] and also Rendleman and Bartter [24].

### 2.5 Binomial Tree Model

To define a binomial tree model, a basic period length of time is established (such as one week or one day). According to the traditional Cox–Ross–Robinstein (CRR) model, if the price is known at the beginning of a period, the price at the beginning of the next period is only one of two possible values—a multiple of $u$ for up and a multiple of $d$ for down, here $u > 1 > d > 0$. The probabilities of these possibilities are $q$ and $1 – q$ respectively. Therefore, if the initial price is $S$, at the beginning of nth period, there are only $n + 1$ possible stock prices, $Su^kd^{n-k}$, $k = 0, 1, \ldots, n$ with a binomial probability distribution of parameter $q \in (0, 1)$. The state model based on a tree of size $\sum_{i=0}^{n} 2^i = 2^{n+1} – 1$ can be collapsed to a binary lattice \{(t_k, u^i d^{k-i}) \mid k = 0, 1, \ldots, i = 0, \ldots, k\} of size $\sum_{i=0}^{n} (i+1) = (n+1)(n+2)/2$. From each node, there are two outgoing arrows, one for up and one for down. For a tree, each node (except root) has only one incoming arrow; for lattice, each node (except the root and the first level) has two incoming arrows. Hence, for a tree structure history is unique, whereas for lattice, one cannot trace the history. In applications, if history is not needed, using lattice saves computation time. For most cases, a tree structure is much clearer and much more versatile than a lattice structure and therefore is strongly
Statistically one can gather historical data to estimate two of the most important parameters: the expected return $E(R)$ and variance $\text{Var}(R)$. Hence, in applying the theoretical model to reality, the model parameter $(u, d, q)$ has to satisfy the matching condition

$$qu + (1 - q)d = 1 + E(R), \quad q(1 - q)(u - d)^2 = \text{Var}(R).$$

There are three parameters for two equations, so one of the parameter is free. One can show that if the single period is sufficiently short, then the choice of the free parameter is not very much relevant. In application, one adds in one of the following equations to fix the parameter:

(i) $q = 1/2$;  
(ii) $ud = 1$;  
(iii) $p = 1/2$

where $p$ is the risk-neutral probability (to be explained later). Each choice has its own advantages.

Here we shall take a simpler but more fundamental and theoretical approach than the general binomial tree or lattice approach described above. We shall simulate the stock price by using the digitized Brownian motion: in each time period, the position of a particle either move to the left or to the right by exactly one unit length, both with probability $1/2$. Such digitized Brownian motion seems very special, nevertheless, in its limit, it can form most of the known stochastic processes. This phenomena has its root from the central limit theorem which asserts that almost all averages of i.i.d random variables are empirically normally distributed.

1. Trading Dates.

The time interval in our consideration is $[0, T]$ which is divided into $n$ subinterval of equal length. Times of trading are at the end points of these subintervals. Hence, we set

$$T = \{t_0, t_1, \ldots, t_n\}, \quad t_i = i\Delta t \quad \forall \ i = 0, 1, \ldots, n, \quad \Delta t = \frac{T}{n}.$$

In most applications, people use Excel spreadsheet, taking $n$ ranging from 5 to 100; that is, $\Delta t$ can be one day, one week, one month, or even one year. One keeps in mind that without accurate input (e.g.
evaluation of important parameters), taking large $n$ but not the limit of $n \to \infty$, may not help much. Nevertheless, for mathematical beauty and deep theoretical analysis, we would like to take the limit as $\Delta t \to 0$ to obtain continuous models, for which, calculus will be very useful.

2. Assets.

In our consideration there are two assets:

$a_0$ : a risk-free asset with constant continuous compounded interest rate $\nu_0$;

$a_1$ : a security (e.g. stock) whose unit share price is $S^t$, a random variable satisfying

$$S^0 = S, \quad S^t_i = S^{t_{i-1}} e^{\nu \Delta t + \sigma \sqrt{\Delta t} \epsilon_i} \quad \forall i = 1, 2, \ldots, n$$

where $\epsilon_i$ is a random variable with probability distribution

$$\text{Prob}(\epsilon_i = 1) = \text{Prob}(\epsilon_i = -1) = \frac{1}{2} \quad \forall i = 1, 2, \ldots, n.$$

It is assumed that $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are independent, identically distributed random variables.

The random variable $\nu \Delta t + \sigma \sqrt{\Delta t} \epsilon_i$ is the continuously compounded return rate of the stock in a single period. The conventional one period return rate $R_i$ introduced in the mean-variance theory is given by $R_i := e^{\nu \Delta t + \sigma \sqrt{\Delta t}} - 1$. Hence, in a single period, the expected return rate $\mathbb{E}(R_i)$ and variance $\text{Var}(R_i)$ are given by

$$\mathbb{E}(R_i) = \frac{1}{2} e^{\nu \Delta t + \sqrt{\Delta t}} + \frac{1}{2} e^{\nu \Delta t - \sqrt{\Delta t}} - 1 = \left(\nu + \frac{\sigma^2}{2}\right) \Delta t + O(\Delta t^{3/2}),$$

$$\text{Var}(R_i) = \frac{1}{4} (e^{\nu \Delta t + \sigma \sqrt{\Delta t}} - e^{\nu \Delta t - \sigma \sqrt{\Delta t}})^2 = \sigma^2 \Delta t + O(\Delta t^{3/2}).$$

It is very important to notice that

*when interests are compounded continuously, the expected growth rate $\nu$ differ from the expected return rate $\mu$ by approximately half of the variance.*

3. State Space

Base on the behavior of the stock price, we build a state space as follows. We use

$$\Omega = B^n := \{(z_1, \ldots, z_n) \mid |z_i| = 1 \quad \forall i = 1, \ldots, n\}, \quad B = \{-1, 1\}.$$ 

The information tree $\{P^t\}_{t \in T}$ is then defined by

$$P^t_i = \{(z_1, \ldots, z_i) \times B^{n-i} \mid |z_k| = 1 \quad \text{for all} \quad k = 1, \ldots, i\}, \quad i = 0, 1, \ldots, n.$$ 

4. State Economy.

For the risk-free asset, its unit share price $S^t_0$ at time $t$ is easily calculated to be

$$S^t_0(\omega) = S^0_0 e^{\nu_0 t} \quad \forall \omega \in \Omega, \ t \in T.$$
The stock price $S^t$ can be calculated by, for every $\omega = (z_1, \cdots, z_n) \in \Omega$,

$$S^{t_k}(\omega) = S^{t_{k-1}} e^{\nu \Delta t + \sigma \sqrt{\Delta t} z_k} = \cdots = S e^{\nu t_k + \sigma \sqrt{\Delta t} \sum_{i=1}^{k} z_i} \quad \forall k = 1, \cdots, n.$$ 

We remark that according to our construction, the natural probability is

$$\text{Prob}(\{\omega\}) = 2^{-n} = \frac{1}{|\Omega|} \quad \forall \omega \in \Omega.$$ 

Consequently, using combinatorial, for any $k = 0, 1, \cdots, n$,

$$q_k := \text{Prob}(S^T = S e^{\nu T + \sigma \sqrt{\Delta t} (2k-n)}) = 2^{-n} \binom{n}{k} = 2^{-n} C_n^k = \frac{n!}{2^n k!(n-k)!}.$$ 

That is to say,

$$\frac{\ln S^T - \nu T}{\sigma \sqrt{\Delta t}} \text{ is binomially distributed.}$$

We remark that the binary tree can be collapsed to binary lattice by combing the nodes with same sum $\sum_{i=1}^{k} z_i$, for every $k = 1, \cdots, n$. This is allowed as far as only spot stock prices are concerned.

Here we remark that the possibility of collapsing of a binomial tree model to a binary lattice model relies on the constancy of $\mu$ and $\sigma$. In sophisticated models, both $\mu$ and $\sigma$ are functions of $S^t$ and $t$ and one realizes it to be very hard to collapse a binary tree model to a binary lattice. Nevertheless, one can use a trinomial lattice model.


First we calculate the risk-neutral transition probability. For any node $(z_1, \cdots, z_k) \times B^{n-k} \in \mathcal{P}^{t_k}$, there are two immediate successors, $(z_1, \cdots, z_k, 1) \times B^{n-k-1}$ and $(z_1, \cdots, z_k, -1) \times B^{n-k-1}$. Denote by $p$ the probability for up and by $1-p$ the probability for down. It is easy to derive the equation for $p$:

$$e^{\nu \Delta t} = p e^{\nu \Delta t + \sigma \sqrt{\Delta t}} + (1-p) e^{\nu \Delta t - \sigma \sqrt{\Delta t}} \quad \implies p = \frac{e^{(\nu - \nu_0) \Delta t} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}}.$$ 

For the model to be good, i.e., arbitrage free, we need $0 < p < 1$. This is equivalent to

$$\sigma > |\nu - \nu_0| \sqrt{\Delta t}.$$ 

Under this condition, we have unique risk-neutral probability. That is, the state model is arbitrage free and complete.

From the transition probabilities, we can derive the risk-neutral probability

$$\mathbb{P}(\{\omega\}) = p^ \frac{1}{2} \sum_{z_1=1}^{z_1}(1+z_1) (1-p)^ \frac{1}{2} \sum_{z_1=1}^{n}(1-z_1) \quad \forall \omega = (z_1, \cdots, z_n) \in \Omega.$$ 

From which, we can also calculate the risk-neutral probability distribution of $S^T$:

$$p_k := \mathbb{P}(\{\omega \mid S^T(\omega) = S e^{\nu T + \sigma \sqrt{\Delta t} (2k-n)}\}) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \quad k = 0, \cdots, n.$$

As we know from the previous chapter, in pricing contingent claims, one has to use the risk-neutral probability distribution. This is a revolutionary idea of Black and Scholes. In old days, people used natural probabilities pricing derivative securities, resulting mismatch with reality.
5. Pricing Formula

Suppose we have a derivative security whose payoff is given by

$$X = f(S^T)$$ only at time $T$

where $f : \mathbb{R} \to \mathbb{R}$ is a given smooth function. Then by our pricing formula, the price for the derivative security is

$$P(X) = \mathbb{E} \left( \frac{S^T_0}{S^T_T} \right) = e^{-\nu_0 T} \mathbb{E}(X) = e^{-\nu_0 T} \mathbb{E}(f(S^T)) = e^{-\nu_0 T} \sum_{\omega \in \Omega} f(S^T(\omega)) \mathbb{P}(\{\omega\})$$

$$= \sum_{k=0}^{n} p_k e^{-\nu_0 T} f(Se^{\nu T + \sigma \sqrt{T} (2k-n)})$$

$$= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} e^{-\nu_0 T} f(Se^{\nu T + \sigma \sqrt{T} (2k-n)/\sqrt{n}}).$$

We can summarize our calculation as follows:

**Theorem 2.1 (Contingent Claim Price Formula)** Assume that the risk-free continuously compounded interest rate is $\nu_0$ and the log of the underlying security price obeys the digitized Brownian Motion with mean $\nu \Delta t$ and variance $\sigma^2 \Delta t$:

$$\ln S_{t_k} - \ln S_{t_{k-1}} = \nu \Delta t + \sigma \sqrt{\Delta t} \epsilon_k, \quad t_k = k \Delta t \quad \forall k \in \mathbb{N},$$

where $\epsilon_1, \epsilon_2, \cdots$ are independent, identically distributed random variables satisfying $\text{Prob}(\epsilon_k = 1) = \text{Prob}(\epsilon_k = -1) = 1/2$ for all $k \in \mathbb{N}$. Also assume that $\sigma > |\nu - \nu_0| \sqrt{\Delta t}$.

Then for any derivative security $X$ with payoff $f(S^T)$ at time $T = n \Delta t$, its price at initial time is

$$P(X) = P_{\Delta t}(S, T) := \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} e^{-\nu_0 T} f(S e^{\nu T + \sigma \sqrt{T} (2k-n)/\sqrt{n}}).$$

where $S$ is the current price of the underlying security and $p = \frac{e^{(\nu_0 - \nu) \Delta t} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}}$.  

Exercise 2.16. Under the assumption of Theorem 2.1, show that if the current time is $t = T - kn \Delta t$ and spot security price is $S$, then the price of a contingent claim with payoff $f(S^T)$ at time $T$ is

$$P_{\Delta t}(S, T - t) := \sum_{i=0}^{m} \frac{m!}{i!(m-i)!} p^i (1-p)^{m-i} e^{-\nu_0 (T-t)} f(S e^{\nu (T-t) + \sigma (2i-m) \sqrt{T-t}}).$$  

(2.5)
2.6 Pricing Options

An option is the right, but not obligation, to buy or sell an asset under specific terms.

A **call option** is the one that gives the right to purchase something.

An **put option** is the one that gives the right to sell something.

To **exercise** an option means the actually buying or selling asset according to option terms.

An option **buyer** or **holder** has the right to exercise an option according to the option terms.

An option **writer** or **seller** has the obligation to fulfil buyer’s right.

The specifications of an option include, but may not limited to the following:

1. A clear description of what can be bought (for a call) or sold (for a put).
   For options on stock, each option is usually for 100 shares of a specified stock. Mathematically, one option means for one share of stock.

2. The **exercise price** or **strike price** for the underlying asset to be sold or bought at.

3. The period of time that the option is valid. This is typically defined as **expiration date**.

There are quite a number of options types:

1. **European option** In this option, the right of the option can be exercise only on the expiration date. Strike price is fixed.

2. **American option** The option right can be exercised any time on or before the expiration data. Strike price is fixed.

3. **Bermudan option** The exercise dates are restricted, in some case to specific dates, in other cases to specific periods within the lifetime of the option.

4. **Asian option** The payoff depend on the average price $S_{avg}$ of the underlying asset during the period of the option. There are basically two ways that the average can be used.
   (i) $S_{avg}$ is served as strike price; the payoff for a call is $\max\{S_T - S_{avg}, 0\}$.
   (ii) $S_{avg}$ is served as the final asset price; the payoff for a call is $\max\{S_{avg} - K, 0\}$.

5. **Look-back option** The effective strike price is determined by the minimum (in the case of call) or maximum (in the case of put) of the price of the underlying asset during the period of the option. For example, a European look-back call has payoff = $\max\{S_T - S_{min}\}$ where $S_{min}$ is the minimum value of the price $S$ over the period from initiation to termination $T$.

6. **Cross-ratio option** These are foreign-currency options denominated in another foreign currency; for example, a call for 100 US dollars with an exercise price of 95 euros.

7. **Exchange option** Such option gives one the right to exchange one specified security for another.

8. **Compound option** A compound is an option on an option.

9. **Forward start option** These are options paid for one date, but do not begin, until a later date.
"As you like it" option The holder can, at a specific time, declare the option to be either a put or a call.

PS: The words European, American, Bermudan, Asian are words for the structure of the option, no matter where they are issued.

Option prices can be calculated by using tree structures, if the dynamics of the price of underlying asset can be described by a finite state model. In the sequel, we provide a few examples demonstrating how to calculate prices of options by using a binomial tree structure. The key here is we have to use risk-neutral probability, not natural probability!

Assume that risk-free interest rate is \( \nu_0 = 0.10 \). Let’s consider of a stock of initial price \( S^0 = 100 \), growth rate \( \nu = 0.12 \)/year, and volatility \( \sigma = 0.20/\sqrt{\Delta t} \). We shall consider options of duration three months, so we construct a three month tree structure. For simplicity, de take \( \Delta t = 1 \) (month) = 1/12 (year). We find

\[
\begin{align*}
    u &= e^{\nu \Delta t + \sigma \sqrt{\Delta t}} = 1.0701, \\
    d &= e^{\nu \Delta t - \sigma \sqrt{\Delta t}} = 0.9534, \\
    p &= \frac{e^{(\nu_0 - \nu)\Delta t} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}} = 0.4712
\end{align*}
\]

The tree is displayed as follows:

<table>
<thead>
<tr>
<th>states</th>
<th></th>
<th>ddd</th>
<th>ddu</th>
<th></th>
<th>dud</th>
<th>duu</th>
<th></th>
<th>udd</th>
<th>udu</th>
<th></th>
<th>uud</th>
<th>uuU</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_k )</td>
<td>.1479</td>
<td>.1318</td>
<td>.1318</td>
<td>.01174</td>
<td>.1318</td>
<td>.1174</td>
<td>.1174</td>
<td>.01046</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( S^3(k) )</td>
<td>86.66</td>
<td>97.26</td>
<td>97.26</td>
<td>109.17</td>
<td>97.26</td>
<td>109.17</td>
<td>109.17</td>
<td>122.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( S^2(k) )</td>
<td>90.89</td>
<td>102.02</td>
<td>102.02</td>
<td>114.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( S^1(k) )</td>
<td>95.33</td>
<td>107.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( S^0 )</td>
<td>100.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here the second line displays all possible states, and the top line provides the risk-neutral probability of these states, being \( p_k = p^k(1 - p)^{3-k}, k = 0, 1, 2, 3 \). The lines after are possible prices of the stock. Since we have constant \( \nu, \nu_0, \sigma \), the risk-neutral probabilities are constants in the sense that probability \( p \) for up and \( 1 - p \) for down.

We now consider the following options. Their durations are all three months.

1. A **European call option** with strike price \( K = 100 \). The caller can buy a stock from seller at price 100 and sell it on the market at price \( S^T \), so the payoff is \( S^T - 100 \). Of course, if \( S^T \leq 100 \), the caller just let the option void. Hence the payoff is \( X = \max\{S^3 - 100, 0\} \). From this, we can use the risk-neutral probability to determine its price

\[
EC = e^{-\nu_0 T} \sum_{k=1}^{8} p_k \cdot X(k) = e^{-\nu_0 T} \sum_{i=0}^{8} p_k \cdot \max\{S^3(k) - 100, 0\}
\]

\[
= \frac{e^{(\nu_0 - \nu)\Delta t} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}} \left\{ 0 + 3 \ast 0 + 3 \ast 9.17 \ast 0.1174 + 22.53 \ast 0.1046 \right\} = \$5.44.
\]

Thus, one European call option worths \$5.44.

Note that for European call options, we need only a lattice structure.

2. A **European put option** with strike price \( K = 100 \). The option holder can by one share of stock at price \( S^T \) and sell to the writer at \( K = 100 \), with profit \( 100 - S^T \). Surely, if \( S^T \geq 100 \), the
2.6. PRICING OPTIONS

option expires quietly. Hence the payoff of is \( X = \max\{100 - S^3\} \). Using risk-neutral probability, we calculate the option’s price as

\[
EP = e^{-\nu_0 T} \sum_{i=0}^{8} p_k \max\{100 - S^3(k), \ 0\}
\]

\[
= e^{-0.1 \times 3/12} \left\{ 13.34 \times 0.1479 + 3 \times 2.74 \times 0.1318 + 3 \times 0 + 0 \right\} = 2.97.
\]

We check the **put-call option parity formula** \( EC - EP = S^0 - Ke^{-\nu_0 T} \):

\[
EC - EP = 5.44 - 2.97 = 2.47, \quad S - Ke^{-\nu_0 T} = 100 - 100e^{-0.1 \times 3/12} = 2.47.
\]

3. **An Asian call option** with strike price \( K = (\sum_{i=1}^{3} S^i)/3 \) (The average of past three month). Then the payoff is \( X = \max\{0, \ S^3 - (\sum_{i=1}^{3} S^i)/3\} \). Tracking the history, we find value’s of \( X \) and the corresponding probability as follows:

<table>
<thead>
<tr>
<th>history</th>
<th>uuu</th>
<th>uud</th>
<th>udu</th>
<th>ddu</th>
<th>ddd</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>114.68</td>
<td>110.23</td>
<td>106.07</td>
<td>102.18</td>
<td>102.10</td>
</tr>
<tr>
<td>S^3</td>
<td>122.53</td>
<td>109.17</td>
<td>109.17</td>
<td>109.17</td>
<td>97.26</td>
</tr>
<tr>
<td>p</td>
<td>0.1046</td>
<td>0.1174</td>
<td>.1174</td>
<td>0.1174</td>
<td>0.1317</td>
</tr>
</tbody>
</table>

Here \( K \) is calculated by taking the average of three stock prices in the last three month. For example,

\[
K(udu) = \left\{ S^0 u + S^0 ud + S^0 udu \right\}/3 = 106.07.
\]

Thus, the price of the Asian call option is

\[
AC = e^{-0.1 \times 3/12} \left\{ (122.53 - 114.68) \times 0.1046 + (109.17 - 106.07) \times 0.1174 \\
+ (109.17 - 102.18) \times 0.1174 + (97.26 - 94.50) \times 0.1317 \right\} = $2.31 .
\]

4. **An Asian Put Option** with strike price being the average of past three month. Then the payoff is \( X = \max\{(\sum_{i=1}^{3} S^i)/3 - S^3, 0\} \). Tracking the history, we find value’s of \( X \) and calculate its price by

\[
AP = e^{-0.1 \times 3/12} \left\{ (110.23 - 109.17) \times 0.1174 + (102.10 - 97.26) \times 0.1317 \\
+ (98.21 - 97.26) \times 0.1174 + (90.96 - 86.66) \times 0.1479 \right\} = $1.41 .
\]

Thus, the Asian put option should have a price of $1.41.

5. **An American Call Option** with strike price \( K = 100 \). It can be argued that the best strategy for American call is to exercise the right at the last day. There is no advantage to exercise earlier! We leave the detailed calculation as an exercise.

6. **American Put.** This is a much hard problem. We have to work backwards to obtain its solution.

<table>
<thead>
<tr>
<th>( S^3 )</th>
<th>122.53</th>
<th>109.17</th>
<th>97.26</th>
<th>86.66</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P^3 = (K - S^3)^+ := \max{K - S^3, 0} )</td>
<td>0</td>
<td>0</td>
<td>2.74</td>
<td>13.34</td>
</tr>
</tbody>
</table>
At $t = 2$, we first use the risk-neutral probability calculate the price of the claim. It’s value is

$$\hat{P}^2(\omega) = e^{-r\Delta t} \{ p P^3(\omega u) + (1 - p) P^3(\omega d) \}$$

One can show that a replicating portfolio at time $t = 2$ can be prepared to pay the $P^3$ exactly at time $t = 3$, and the cost of the portfolio is $\hat{P}^2$.

Since we have option to exercise the right valued at max{$K - S^2$}. If we know the value $\hat{P}^2$ is smaller this, then we should exercise the option. Hence, we take the maximum of $(K - S^2)^+$ and $\hat{P}^2$. The calculation is as follows:

<table>
<thead>
<tr>
<th>$S^2$</th>
<th>114.51</th>
<th>102.02</th>
<th>90.89</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(K - S^2)^+$</td>
<td>0</td>
<td>0</td>
<td>9.11</td>
</tr>
<tr>
<td>$e^{-r\Delta t}[P^3(i\omega u) + (1 - p)P^3(\omega d)]$</td>
<td>0</td>
<td>1.43</td>
<td>8.28</td>
</tr>
<tr>
<td>$P^2$</td>
<td>0</td>
<td>1.43</td>
<td><strong>9.11</strong></td>
</tr>
</tbody>
</table>

In the next step, we perform the same calculation

$$P^1(\omega) = \max \{ (K - S^1(\omega))^+, e^{-r\Delta t}(pP^2(\omega d) + (1 - p)P^2(\omega d)) \}$$

<table>
<thead>
<tr>
<th>$S^1$</th>
<th>107.01</th>
<th>95.34</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(K - S^1)^+$</td>
<td>0</td>
<td>4.66</td>
</tr>
<tr>
<td>$e^{-r\Delta t}[P^2(i) + (1 - p)P^2(i + 1)]$</td>
<td>0.75</td>
<td>5.46</td>
</tr>
<tr>
<td>$P^1$</td>
<td>0.75</td>
<td>5.46</td>
</tr>
</tbody>
</table>

Finally, $P^0 = \max\{ (K - S^0), e^{-r\Delta t}(pP^1(u) + (1 - p)P^1(d)) \} = 3.21$. Thus, the American put has value $3.21$.

7. Summary. From these example, we can summarize the method as follows:

1. If the claim has only a final payment $X$. Then the value at earlier time $t_k$ can be evaluated by

$$P^{t_k}(\omega) = p P^{t_{k+1}}(\omega u) + (1 - p) P^{t_{k+1}}(\omega d).$$

2. If there are options to exercise a right at time $t_k$ with value $X^{t_k}(S^{t_k})$ after the revelation of stock price $S^{t_k}$ at time $t_k$, then

$$P^{t_k}(\omega) = \max \{ X^{t_k}(S^{t_k}(\omega))^+, p P^{t_{k+1}}(\omega u) + (1 - p) P^{t_{k+1}}(\omega d) \}$$

Finally, we have to emphasize that the value we calculated is obtained by taking $\Delta t = 1$ month. It is too big for the value to be accurate. Theoretically we should take $\Delta t = dt$, an infinitesimal.

Exercise 2.17. For the American call option, follow the process as the American put option, go backward step by step find the value of the American call option. Show that at each time, the value of the call option is always bigger than $S^1 - K$, so option holder should not exercise the option right. Consequently, the value of the American option is the same as that of the European put option.

Exercise 2.18. A stock has an initial price $S^0 = 100$ and we are considering a four month security. Take $\Delta t$ = one month, $\nu_0 = 0.1$/year, $\nu = 0.15$/year and $\sigma^2 = 0.04$/year, construct both a binomial tree and binomial lattice. Compute the risk-neutral probabilities.

1. Use the lattice to calculate the prices of a European put option, a European call option, an American call option, and an American put option. Here strike prices are the same as the current stock price and duration is fourth month.

2. Use the tree to calculate the price for an Asian call option to be called at the beginning of the fifth month with strike price is $K = (S^2 + S^3 + S^4)/3$ where $S^t$ is the price at the first day of $(t + 1)$th month.
2.7 Replicating Portfolio for Derivative Security

As a special example of the general finite model, here we see how a portfolio is managed to provide the exact payment of a contingent claim, thereby providing the price of the claim. We continue to use the notations in the previous section.

Suppose $X$ is a contingent claim with payoff $f(S^T)$. Here we shall use the binary tree model. Let $\{(n^0_t(\cdot),n^1_t(\cdot))\}_{t\in T}$ be the trading strategy, where $n^0_t$ is the number of shares of risk-free asset whose unit price is $e^{r_t}$ and $n^1_t(\cdot)$ is the number of shares of the underlying security whose unit price is $S^t(\cdot)$.

We denote by $V^t(\omega)$ the value of the portfolio at time $t$ and event $\omega \in \Omega$. Hence,

$$V^t(\cdot) = n^0_t(\cdot)e^{r_t} + n^1_t(\cdot)S^t(\cdot).$$

Clearly, at time $T$, we require

$$V^T(\omega) = f(S^T(\omega)) \quad \forall \omega \in \Omega.$$

We now investigate how a portfolio can be established at time $t_{n-1}$ so that regardless of the outcome at time $T$, the updated value generated by the portfolio matches, with 100% certainty, with the needs, e.g. pay the claim off at time $T$.

Hence, assume that we are at time $t = t_{n-1}$ and at a state $(z_1, \cdots, z_{n-1}) \times B$. For convenience we use the following notations

$$z = (z_1, \cdots, z_{n-1}, ?) \in \mathcal{P}^{n-1},$$

$$z_U = (z_1, \cdots, z_{n-1}, 1), \quad z_D = (z_1, \cdots, z_{n-1}, -1).$$

$$s = S^{t_{n-1}}(z), \quad s_U = S^T(z_U) = su, \quad s_D = S^T(z_D) = sd$$

where

$$u = e^{r\Delta t + \sigma \sqrt{\Delta t}}, \quad d = e^{r\Delta t - \sigma \sqrt{\Delta t}}.$$

At current time $t = t_{n-1}$, we know event $z$ happened and we have $n^0_t(z)$ shares of risk-free asset whose unit price is $e^{r_t}$ and $n^1_t(z)$ shares of security whose unit price is $s = S^t(z)$. At the end period (e.g. at time $T$), the outcome is either $z_U$ or $z_D$. For the value of the portfolio to prepare for the payment of the claim $X$ in either outcome, it is necessary and sufficient to have

$$f(s_U) = n^0_t(z)e^{r_t} + n^1_t(z)s_U \quad t = t_{n-1}, s = S^t(z)$$

$$f(s_D) = n^0_t(z)e^{r_t} + n^1_t(z)s_D \quad s_U = su, \quad s_D = sd$$

This system has a unique solution given by, for $t = t_{n-1},$

$$n^0_t(z) = e^{-r\Delta t} \frac{f(s_d)u - f(s_u)d}{u - d}, \quad n^1_t(z) = \frac{f(s_u) - f(s_d)}{su - sd}.$$

Notice that the value of the portfolio at time $t = t_{n-1}$ is

$$V^t(z) = n^0_t(z)e^{r_t} + n^1_t(z)S^t(z) = e^{r_t(t-T)}\{pf(s_u) + (1-p)f(s_d)\}.$$  

Thus, the portfolio at time $t = t_{n-1}$ is completely determined by the claim.

Now we consider the general time period from $t = t_k$ to $t_{k+1} = t + \Delta t$. Similar to the above discussion, we denote a general block in $\mathcal{P}^k$ by

$$z = (z_1, \cdots, z_k, ?, \cdots, ?) \in \mathcal{P}^k,$$

$$z_U = (z_1, \cdots, z_k, 1, ?, \cdots, ?) \in \mathcal{P}^{k+1},$$

$$z_D = (z_1, \cdots, z_k, -1, ?, \cdots, ?) \in \mathcal{P}^{k+1},$$

$$s = S^t(z), \quad s_U = S^{t+\Delta t}(z_U) = su, \quad s_D = S^{t+\Delta t}(z_D) = sd.$$
Suppose we know the value of $V^{t+\Delta t}(z_U) = V^{t+\Delta t}(su)$ and $V^{t+\Delta t}(z_D) = V^{t+\Delta t}(sd)$. Then for the portfolio constructed at time $t = t_{n-1}$ to match exactly with the needed portfolio at time $t + \Delta t$, we need

$$V^{t+\Delta t}(z_U) = e^{v_0(t+\Delta t)} n_0^0(z) + s u n_1^1(z), \quad V^{t+\Delta t}(z_D) = e^{v_0(t+\Delta t)} n_0^1(z) + s d n_1^1(z).$$

It follows that there are a unique portfolio $(n_0^0, n_1^1)$ that provides the need for payment at at time $t_{k+1}$:

$$n_0^0(z) = e^{-v_0(t+\Delta t)} \frac{V^{t+\Delta t}(z_D) u - V^{t+\Delta t}(z_U) d}{u - d},$$

$$n_1^1(z) = \frac{V^{t+\Delta t}(z_U) - V^{t+\Delta t}(z_D)}{su - sd}.$$

The value of the portfolio at $t = t_k$ is

$$V^t(z) = e^{-r_0 t} \left( p V^{t+\Delta t}(z_U) + (1 - p) V^{t+\Delta t}(z_D) \right).$$

With this reduction formula, we then use an induction to derive the following formula, at $t = t_{n-m} = T - m \Delta t$ and spot security price $s = S^t(z)$,

$$V^t(z) = \sum_{i=0}^{m} m! \frac{p^i (1 - p)^{m-i}}{i! (m-i)!} e^{v_0(t-T)} f \left( s e^{\nu(T-t)+\sigma \sqrt{T t}} (2i-m) \right).$$

It is important to observe that the right-hand sides depends only on $s$, i.e., it depends only on $\sum_{i=1}^{m} z_i$; namely, the binomial tree model can be replaced by a binomial lattice model, whose states are

$$\{(t_k, su^i d^{n-i}) \mid i = 0, \cdots, k, k = 0, \cdots, n\}.$$ 

Hence, we can summarize our result in terms of the lattice model setting as follows.

**Theorem 2.2 (Portfolio Replication Theorem)** Assume the condition of Theorem 2.1. Then at any time $t \in T$ and spot underlying security price $s$, the value $V = V(s, t)$ of the contingent claim is

$$V(s, t) = \sum_{i=0}^{m} m! \frac{p^i (1 - p)^{m-i}}{i! (m-i)!} e^{v_0(t-T)} f \left( s e^{\nu(T-t)+\sigma \sqrt{T t}} (2i-m) \right), \quad m = \frac{T - t}{\Delta t}.$$ 

The portfolio replicating the contingent claim is unique. At any time $t - \Delta t$ and spot price security $s$, it consists of $n_r(s, t - \Delta t)$ shares of risk-free asset and $n_S(s, t - \Delta t)$ shares of security, where

$$n_r(s, t - \Delta t) = e^{-r_0 \Delta t} \frac{V(s, t) u - V(su, t) d}{u - d},$$

$$n_S(s, t - \Delta t) = \frac{V(su, t) - V(s, t)}{su - sd},$$

$$u = e^{\nu t + \sigma \sqrt{T t}}, \quad d = e^{\nu t - \sigma \sqrt{T t}}, \quad p = \frac{e^{(v_0 - \nu) \Delta t} - e^{-\sigma \sqrt{T t}}}{e^{\sigma \sqrt{T t}} - e^{-\sigma \sqrt{T t}}} = \frac{e^{\nu_0 \Delta t} - d}{u - d}.$$
Example 2.1. Consider a European call option with duration 3 month and strike price 19. The risk-free interest rate is \( r_0 = 0.1 \)\% year, current stock price is 20 with variance \( \sigma^2 = 0.02 \)\% year and return rate
\[ \nu + \frac{\sigma^2}{2} = 0.22 \]\% year. Set \( \Delta t = 1/12 \) we have \( n = 3 \) and \( u = 1.079, \quad d = 0.961, \quad p = 0.399. \)

The stock price, value of the call, and number of shares are given as follows:

<table>
<thead>
<tr>
<th>time</th>
<th>stock price</th>
<th>time</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20.00</td>
<td>0</td>
<td>1.73</td>
</tr>
<tr>
<td>1</td>
<td>19.23</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td>2</td>
<td>18.48 20.75</td>
<td>2</td>
<td>0.37</td>
</tr>
<tr>
<td>3</td>
<td>17.77 19.95</td>
<td>3</td>
<td>0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>time</th>
<th>shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(-14.56, 0.815)</td>
</tr>
<tr>
<td>1</td>
<td>(-11.93, 0.676)</td>
</tr>
<tr>
<td>2</td>
<td>(-7.53, 0.435)</td>
</tr>
<tr>
<td>3</td>
<td>(-7.53, 0.435)</td>
</tr>
</tbody>
</table>

Here the table for value is constructed as follows: (i) Last row is the payoff \( \max\{S - 19, 0\} \). (ii) From each row up, \( v(i, j) = [0.399 \times v(i + 1, j + 1) + (1 - 0.399) \times v(i + 1, j)] \times e^{-r \Delta t} \). For example, third row last number: \( 6.13 \times p + 3.39 \times (1 - p) \times e^{-r \Delta t} = 4.44 \). The initial price of the call is 1.73.

Now we see how portfolio is constructed, i.e. how a writer of the call prepares for the call.

At time \( t = 0 \): the writer sells an option for 1.73 and sets up a portfolio consisting of -14.56 cash (share of risk-free asset) and 0.815 share of stock. We can check the balance \(-14.56 + 0.815 \times 20.00 = 1.73\).

At time \( t = 1 \) month, if stock price drops to 19.23, then the portfolio is worth 0.98. The writer responds by shorting 11.93 \times e^{0.1/12} \) cash and longing 0.676 share of stock, total \(-11.93 \times e^{0.1/12} + 0.676 \times 19.23 = 0.98\). If stock price rises up to 21.58, then the portfolio is worth 2.89. The writer responds by shorting 18.53 share of risk-free asset and longing 1.000 share of stock, totaling \(-18.53 \times e^{0.1/12} + 1.000 \times 21.58 = 2.89\).

At time \( t = 2 \) month, if stock price is 20.75 or 23.28, the writer rolls over the portfolio. If price drops to 18.48, the writer shorts 7.53 \times e^{0.1 \times 2/12} \) cash and buys 0.435 share of stock, net value \(-7.53 \times e^{0.1 \times 2/12} + 0.435 \times 18.46 = 0.37\).

At \( t=3 \), in each situation, the portfolio pays the call exactly. Note that if stock price drop from 18.48 to 17.77, the portfolio worths \(-7.53 \times e^{0.1 \times 3/12} + 0.435 \times 17.77 = 0\), no need to answer call. If it moves up from 18.48 to 19.95, then the portfolio worths \(-7.50 \times e^{0.1 \times 3/12} + 0.435 \times 19.95 = 0.95\) which, adding caller’s 19, is just enough to buy one share of stock at 19.95/share to give it to the caller. In other situations, -18.53 share of risk-free becomes -19.00 cash. Hence, the portfolio has -19.00 cash balance and one share of stock, just enough to pay the caller.

We emphasize again that we take \( \Delta t = 1 \) (month) is for illustration only. If possible, smaller \( \Delta t \) is preferred.

Exercise 2.19. Using the parameters the example 2.1, calculate the price and corresponding replicating portfolio for (i) six month European call, strike price 19. (ii) Six month European put with strike price 19.

Also, study the American call and put options with duration 3 months and strike price 20.
2.8 A Model for Stock Prices

In this section, we present a model for stock prices, by taking the limit of discretized model.

1. Random Walk and Brownian Motion

Let \( \{\epsilon_i\}_{i=0}^{\infty} \) be a sequence of independent, identically distributed real valued random variables with mean zero and variance one. For simplicity, we assume that

\[
\text{Prob}(\epsilon_i = 1) = \text{Prob}(\epsilon_i = -1) = \frac{1}{2} \quad \forall i = 0, 1, \ldots.
\]

Fix any \( \Delta t > 0 \). Consider the following random variables \( \{z^{\Delta t}(t)\}_{t \geq 0} \) defined by

\[
z^{\Delta t}(0) = 0, \quad z^{\Delta t}(t) = \sqrt{\Delta t} \sum_{0 \leq i < t} \epsilon_i \quad \forall t > 0.
\]

A random walk is the motion of a particle whose position at \( t \) is given by \( z^{\Delta t}(t) \) for all \( t \geq 0 \).

Now fix a positive time \( t \). Let \( n = n(\Delta t, t) \) be an integer such that \( t \in ((n-1)\Delta t, n\Delta t] \). Then

\[
z^{\Delta t}(t) = \sqrt{\Delta t} \sum_{i=0}^{n-1} \epsilon_i = \sqrt{\frac{n\Delta t}{t}} \sum_{i=0}^{n-1} \epsilon_i.
\]

According to the central limit theorem, as \( \Delta t \to 0 \), the right-hand side has a limit in distribution, and the limit is normally distributed. Let’s denote this limit by \( B_t \). We know \( B_t \) is a random variable in some measure space \((\Omega, \mathcal{F}, \mathbb{P})\). The measure space \((\Omega, \mathcal{F}, \mathbb{P})\) is extremely hard to construct so we shall not worry about it. What we care is the distribution of \( B_t \), which we can derive without much difficulty.

(A) First of all, from the central limit theorem, \( B^t \) is normally distributed with variance \( t \):

\[
\text{Prob}(B^t \in A) := N(0, t; A) \quad \forall A \in \mathcal{B}
\]

where \( \mathcal{B} \) is the \( \sigma \)-algebra of Borel set on \( \mathbb{R} \) and \( N(\mu, \sigma; A) \) represents the normal distribution

\[
N(\mu, \sigma; A) := \frac{1}{\sqrt{2\pi} \sigma} \int_{A} e^{-x^2/(2\sigma^2)} dx \quad \forall A \in \mathcal{B}.
\]

(B) Next, for any fixed \( t \) and \( \tau \) satisfying \( t > \tau \geq 0 \), we investigate the random variable \( B^t - B^\tau \). Assume that \( t \in ((n-1)\Delta t, n\Delta] \) and \( \tau \in ((k-1)\Delta t, k\Delta t) \) we have

\[
\frac{z^{\Delta t}(t) - z^{\Delta t}(\tau)}{t - \tau} = \sqrt{\frac{(n - k)\Delta t}{t - \tau}} \sum_{i=k-1}^{n-1} \epsilon_i.
\]

Again, sending \( \Delta t \to 0 \) we see that \( B_t - B_\tau \) is normally distributed, with mean zero and variance \( t - \tau \).

(C) Finally, take any \( \{t_i, \tau_i\}_{i=1}^{m} \) that satisfies \( 0 \leq \tau_1 < t_1 < \tau_2 < t_2 < \cdots < \tau_m < t_m \). When \( \Delta t \) is sufficiently small, we know that \( z^{\Delta t}(t_i) - z^{\Delta t}(\tau_i), \ i = 1, \cdots, m, \) are independent. Hence in the limit, we know that \( B_{t_i} - B_{\tau_i}, i = 1, \cdots, m, \) are also independent, i.e.

\[
\text{Prob}(B_{t_i} - B_{\tau_i} \in A_i \ \forall i = 1, \cdots, m) = \prod_{i=1}^{m} \text{Prob}(B_{t_i} - B_{\tau_i} \in A_i) \quad \forall A_1, \cdots, A_m \in \mathcal{B}.
\]

These three properties characterize all needed properties of the Brownian motion, also known as Winner Process. We formalize it as follows.

A stochastic process is a collection \( \{\xi_t\}_{t \geq 0} \) of random variables in certain measure space \((\Omega, \mathcal{F}, \mathbb{P})\).

A Brownian motion or Wiener process is a stochastic process \( \{B_t\}_{t \geq 0} \) satisfying the following:
2.8. A MODEL FOR STOCK PRICES

1. \( B_0 = 0 \);
2. for any \( t > \tau \geq 0 \), \( B_t - B_\tau \) are normally distributed with mean zero and variance \( t - \tau \);
3. for any \( 0 \leq t_1 < t_2 < \cdots < t_m \), the following increment are independent:
   \[
   B_{t_m} - B_{t_{m-1}}, B_{t_{m-1}} - B_{t_{m-2}}, \ldots, B_{t_2} - B_{t_1}.
   \]

We know that to model \( n \) independent real valued random variables, we need to use a sample space as big as \((\mathbb{R}^n, \mathcal{B}^n)\). If we are going to describe a sequence of i.i.d. random variables, we need a space something like \((\mathbb{R}^N, \mathcal{B}^N)\). Now to model the Brownian motion, we could use the space \( \mathbb{R}^{[0, \infty)} \).

However, this space is enormously big and we can hardly define a \( \sigma \)-algebra and meaningful measure on it. Deep mathematical analysis shows that Brownian motion can be realized on the space of all continuous functions from \([0, \infty) \to \mathbb{R}\). That is, one can take \( \Omega = C([0, \infty); \mathbb{R}) \), on it build a \( \sigma \)-algebra and define a measure. As a result, for every \( \omega \in \Omega \), \( B_t(\omega) \) is a continuous function from \( t \to B_t(\omega) \).

2. Generalized Winner Process and Ito Process

Note that in discretized approximation of the Brownian motion, we have
\[
z^{\Delta t}(t + \Delta t) - z^{\Delta t} = \epsilon(t)\sqrt{\Delta t}
\]
where \( \epsilon(t) \) is a random variable with mean zero and variance one. Hence, symbolically we can write
\[
\mathrm{d}B_t = \epsilon(t)\sqrt{\mathrm{d}t}
\]
where \( \epsilon(t) \) is normally distributed having mean zero and variance 1. We have to say that the Brownian process is nowhere differentiable since
\[
\mathbb{E}\left( \frac{(B_t - B_\tau)^2}{t - \tau} \right) = \frac{t - \tau}{(t - \tau)^2} = \frac{1}{t - \tau} \to \infty \quad \text{as} \quad \tau \searrow t.
\]
In engineering field, people use \( \mathrm{d}B_t/\mathrm{d}t \) symbolically to describe \textbf{white noise}.

The general Wiener process can be written as
\[
W_t = \nu t + \sigma B_t \quad \text{or} \quad \mathrm{d}W_t = \nu \mathrm{d}t + \sigma \mathrm{d}B_t.
\]

An \textbf{Ito process} is a solution to the stochastic differential equation
\[
\mathrm{d}x_t = a(x_t, t)\mathrm{d}t + b(x_t, t)\mathrm{d}B_t
\]

Since with probability one the sample path \( t \to B_t(\omega) \) is not differentiable, spacial tools, e.g. stochastic calculus are need. The following Ito’s lemma \cite{13} is no doubt one of the most important results.

\textbf{Lemma 2.1 (Ito Lemma).} Suppose \( \{x_t\} \) is a stochastic process satisfying \( \mathrm{d}x = a(x, t)\mathrm{d}t + b(x, t)\mathrm{d}B_t \). Let \( f \) be a smooth function : \( \mathbb{R} \times [0, \infty) \to \mathbb{R} \). Then
\[
\mathrm{d}f(x_t, t) = \frac{\partial f}{\partial x} \mathrm{d}x + \frac{\partial f}{\partial t} \mathrm{d}t + \frac{b^2}{2} \frac{\partial^2 f}{\partial x^2} \mathrm{d}t.
\]
If we use the ordinary differentials, the rule can be memorized as
\[(dt)^2 = 0, \quad dtdB_t = 0, \quad (dB_t)^2 = dt.\]

3. The Lognormal Process for Stock Prices.

A basic assumption in the Black–Scholes model is the **geometric Brownian motion** for the stock price \(S_t\). In the form of stochastic differential equation, it reads
\[
\frac{dS_t}{S_t} = \left(\nu + \frac{\sigma^2}{2}\right) dt + \sigma dB_t.
\]
Here \(B_t\) is the standard Brownian motion also called Wiener process.

In the discretized case, we can write it as
\[
\frac{S_{t+\Delta t} - S_t}{S_t} = \left(\nu + \frac{\sigma^2}{2}\right) \Delta t + \sigma \sqrt{\Delta t} \epsilon(t)
\]
where \(\text{Prob}(\epsilon(t) = 1) = \text{Prob}(\epsilon(t) = -1) = 1/2\). It then follows that
\[
\ln S_{t+\Delta t} - \ln S_t = \ln \left\{ 1 + \frac{S_{t+\Delta t} - S_t}{S_t} \right\} = \frac{S_{t+\Delta t} - S_t}{S_t} - \frac{1}{2} \left( \frac{S_{t+\Delta t} - S_t}{S_t} \right)^3 + O(1) \left( \frac{S_{t+\Delta t} - S_t}{S_t} \right)^3
\]
\[
= \left(\nu + \frac{\sigma^2}{2}\right) \Delta t + \sigma \sqrt{\Delta t} \epsilon(t) - \frac{1}{2} \sigma^2 \Delta t \epsilon(t)^2 + O(\Delta t^{3/2})
\]
where we use the assumption \(\text{Prob}(\epsilon^2(t) = 1) = 1\). Hence, in the limit, we should have, symbolically
\[
d\ln S_t = \nu dt + \sigma dB_t.
\]

Hence, a geometric Brownian motion is also called a **lognormal process**. Here lognormal means log is normal not log of normal. To make everything rigorous, one needs first define stochastic calculus. For this we omit here.

Finally, since \(d(\ln S_t - \mu t - B_t) = 0\), the solution is given by
\[
\ln S_t = \ln S_0 + \nu t + B_t \quad \text{or} \quad S_t = S_0 e^{\nu t + \sigma B_t}.
\]

Exercise 2.20. A stock price is governed by \(S_0 \equiv 1\) and \(d\ln S_t = \nu dt + \sigma dB_t\). Find the following:
\[
E(\ln S_t), \quad \text{Var}(\ln S_t), \quad \ln E(S_t), \quad \text{Var}(S_t).
\]

Exercise 2.21. If \(R_1, R_2, \ldots, R_n\) are return rates of a stock in each of \(n\) periods. The arithmetic mean \(R_A\) and geometric mean \(R_G\) return rates are defined by
\[
R_A = \frac{1}{n} \sum_{i=1}^n R_i, \quad R_G = \left( \prod_{i=1}^n (1 + R_i) \right)^{\frac{1}{n}} - 1.
\]

Suppose \$40 is invested. During the first it increases to \$60 and the second year it decreases to \$48. What is the arithmetic mean and geometric mean?

When is it appropriate to use these means to describe investment performance?
Exercise 2.22. The following is a list of stock price in 12 weeks:

10.00, 10.08, 10.01, 9.59, 9.89, 10.55, 10.96, 11.25, 10.86, 11.01, 11.79, 11.74.

(1) From these data, find appropriate \( \nu \) and \( \sigma^2 \) in the geometric Brownian motion process for the stock price, take unit by year.

(2) Suppose \( R \) is the annual rate of return of the stock. Find approximation for \( \mathbb{E}(R) \) and \( \text{Var}(R) \).

2.9 Continuous Model As Limit of Discrete Model

In this section, we shall derive, in an indirect way, the Black-Scholes equation for pricing derivative securities. In the sequel, we shall assume that \( \nu_0, \nu, \sigma, T \) are all fixed constants, where \( T > 0, \sigma > 0 \).

We start with the price formula from the finite state model. We wish to derive the limit of the price, as \( \Delta t \to 0 \), so that we can obtain the continuous limit.

We shall use Taylor’s expansion

\[
e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + O(x^4), \quad \ln(1 + x) = x - \frac{x^2}{2} + O(x^3) \quad \text{as} \quad x \to 0.
\]

We can expand the risk neutral probability \( p \) by

\[
p = \frac{\exp(\nu_0 - \nu)\Delta t - \exp(-\sigma \sqrt{\Delta t})}{\exp(\sigma \sqrt{\Delta t}) - \exp(-\sigma \sqrt{\Delta t})}
\]

\[
= \frac{\sigma \sqrt{\Delta t} + (\nu_0 - \nu - \sigma^2/2)\Delta t + \sigma^3 \Delta t^{3/2}/6 + O(\Delta t^3)}{2\sigma \sqrt{\Delta t} + 2\sigma^3 (\Delta t)^{3/2} / 6 + O(\Delta t^{5/2})}
\]

\[
= \frac{1}{2} \left[ 1 + \frac{\nu_0 - \nu - \frac{\sigma^2}{2}}{\sigma} \right] \sqrt{\Delta t} + O(\Delta t^{3/2})
\]

It then follows that for any integer \( k \in [0, n] \),

\[
\ln \left[ \begin{array}{c}
k \ln p + (n - k) \ln(1 - p) \\

\end{array} \right] = k \ln p + (n - k) \ln(1 - p)
\]

\[
= k \ln \left( \frac{1}{2} \left[ 1 + \frac{\nu_0 - \nu - \frac{\sigma^2}{2}}{\sigma} \sqrt{\Delta t} + O(\Delta t^{3/2}) \right] \right) + (n - k) \ln \left( \frac{1}{2} \left[ 1 - \frac{\nu_0 - \nu - \frac{\sigma^2}{2}}{\sigma} \sqrt{\Delta t} + O(\Delta t^{3/2}) \right] \right)
\]

\[
= n \ln \frac{1}{2} + (2k - n) \frac{(\nu_0 - \nu - \frac{\sigma^2}{2})}{\sigma} \sqrt{\Delta t} - \frac{n\Delta t}{2} \left( \frac{\nu_0 - \nu - \frac{\sigma^2}{2}}{\sigma} \right)^2 + nO(\Delta t^{3/2})
\]

\[
= -n \ln 2 + \frac{(2 - n)}{\sqrt{n}} \frac{(\nu_0 - \nu - \frac{\sigma^2}{2})}{\sigma} \sqrt{T} - \frac{1}{2} \left( \frac{(\nu + \frac{\sigma^2}{2} - \nu_0)}{\sigma} \sqrt{T} \right)^2 + O\left( \frac{1}{\sqrt{n}} \right).
\]

Now we define

\[
x_k = \frac{2k - n}{\sqrt{n}} \quad \forall k \in \mathbb{Z}, \quad \Delta x = x_{k+1} - x_k = \frac{2}{\sqrt{n}}.
\]

Then

\[
k = \frac{n}{2} \left( 1 + \frac{x_k}{\sqrt{n}} \right), \quad n - k = \frac{n}{2} \left( 1 - \frac{x_k}{\sqrt{n}} \right).
\]

For the factories, we use the Stirling’s formula

\[
k! = \sqrt{2\pi} \exp \left( (k + 1/2) \ln k - k + \frac{\theta_k}{12k} \right) \quad \forall k \geq 1, 0 < \theta_k < 1.
\]
Therefore, substituting $k = 1, \cdots, n - 1$,
$$\frac{1}{\Delta x} \frac{n!}{k!(n-k)!} = \frac{1}{\sqrt{2\pi}} \exp \left(\frac{n+1}{2} \ln n - \ln 2 - \frac{1}{2} \ln 2 - \frac{1}{2} \left(\ln n - \frac{1}{2}\right) + \frac{1}{2} \left(\ln n - \frac{1}{2}\right)^2\right)$$
where
$$\xi_k = \frac{\theta_n}{12n} - \frac{\theta_k}{12k} - \frac{\theta_{n-k}}{12(n-k)} \quad 0 < \theta_k, \theta_{n-k}, \theta_n < 1.$$  
Therefore, substituting $k = n[1 + x_k/\sqrt{n}] / 2$ and $n-k = n[1 - x_k/\sqrt{n}] / 2$ we have
$$\ln \left(\frac{\sqrt{2\pi}}{\Delta x} \frac{n!}{k!(n-k)!}\right) - \xi_k = \left[n + 1\right] \ln n - \ln 2 - \frac{1}{2} \ln 2 - \frac{1}{2} \left(\ln n - \frac{1}{2}\right) + \frac{1}{2} \left(\ln n - \frac{1}{2}\right)^2$$
When $|x_k| < 2^{1/4}$, we have
$$\frac{n+1}{2} \ln \left(1 - \frac{x_k^2}{n}\right) + \frac{\sqrt{n}x_k}{2} - \frac{1}{2} \left(\ln n + \frac{1}{n}\right) - \frac{1}{2} \left(\ln n + \frac{1}{n}\right)^2.$$  
Combining all these together, we then obtain when $|x_k| \leq 2^{1/4}$,
$$\frac{1}{\Delta x} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{1}{\sqrt{2\pi}} \exp \left\{-\frac{1}{2} \left(x_k + \frac{(\nu_0 - \nu - \sigma^2/2)\sqrt{T}}{\sigma}\right)^2 - \frac{1}{2} \left(\frac{\nu + \sigma^2/2 - \nu_0}{\sqrt{T}}\right)^2 + O(1)x_k^4\right\}$$
When $|x_k| \geq 2^{1/4}$, one can verify that
$$\rho_k := \frac{1}{\Delta x} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \leq O(1)e^{-\sqrt{n}}$$
since when $2^{1/4} \leq |x_k| \leq 2^{1/4} + 1, \rho_k < e^{-\sqrt{n}}$ and
$$\frac{\rho_{k+1}}{\rho_k} = \frac{(n-k)p}{(k+1)(1-q)} > 1 \quad \text{when} \quad x_k < -2^{1/4},$$
$$\frac{\rho_{k+1}}{\rho_k} = \frac{(n-k)p}{(k+1)(1-q)} < 1 \quad \text{when} \quad x_k > 2^{1/4}.$$  
Hence, assume that $f$ is continuous and bounded, we have
$$P_{\Delta t}(S, T) := \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} e^{-\nu_0 T} f(S e^{\nu T + \sigma \sqrt{T}(2k-n)})$$
$$= \sum_{|x_k| < 2^{1/4}} \frac{e^{-\nu_0 T}}{\sqrt{2\pi}} e^{-\frac{1}{2} \left|x_k + \frac{(\nu + \sigma^2/2 - \nu_0)\sqrt{T}}{\sigma}\right|^2 + O(1)x_k^4 / n} f(S e^{\nu T + \sigma \sqrt{T} x_k}) \Delta x + O(1)ne^{-\sqrt{n}}.$$
2.9. CONTINUOUS MODEL AS LIMIT OF DISCRETE MODEL

Sending $\Delta t \to \infty$ (i.e. $n \to \infty$) we then obtain

\[
\lim_{\Delta t \to 0} P_{\Delta t}(S,T) = \int_{\mathbb{R}} e^{-\nu_0 T} e^{-\frac{1}{2}(x+\nu_0 T + \frac{\sigma^2}{2} T^2)} e^{-r T} f(S e^{\nu T + \sigma \sqrt{T} x}) dx
\]

\[
= \frac{e^{-\nu_0 T}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-z^2/2} f(e^{\ln S + \sigma z \sqrt{T} + (\nu_0 - \sigma^2/2)T}) dz.
\]

We summarize our calculation as follows.

**Theorem 2.3** Suppose the risk-free interest rate is a constant $\nu_0$ and the unit price of the underlying stock is a geometric Brownian (or lognormal) process

\[
\log S^t = \ln S + \nu t + \sigma B_t \quad \forall t \geq 0
\]

where $B_t$ is the standard Brownian motion process. Then a contingent claim at time $T > 0$ with payoff $f(S^T)$ has price given by the Black-Scholes’ pricing formula

\[
P(S, T) = \frac{e^{-\nu_0 T}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-z^2/2} f(e^{\ln S + \sigma z \sqrt{T} + (\nu_0 - \sigma^2/2)T}) dz.
\]

In addition, at any time $t \in (0, T)$ and spot price $s$, the value of the contingent claim is

\[
V(s, t) = \frac{e^{-\nu_0 (T-t)}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-z^2/2} f(e^{\ln s + \sigma z \sqrt{T-t} + (\nu_0 - \sigma^2/2)(T-t)}) dz.
\]

Furthermore, at any time $t \in (0, T)$ and spot price $s$, the portfolio replicating the contingent claim is given by $n_S(s, t)$ shares of stock and $n_{rf}(s, t)$ shares of risk-free asset (whose unit share price is $e^{\nu_0 t}$) where

\[
n_S(s, t) = \frac{\partial V(s, t)}{\partial s}, \quad n_{rf}(s, t) = e^{-\nu_0 t} \left\{ V(s, t) - s n_S(s, t) \right\}.
\]

We leave the derivation of formula for $n_S$ and $n_{rf}$ as an exercise.

Here we make a few observations:

(i) The parameter $\nu$ does not appear in the formula. Namely, the mean expected return of the stock is irrelevant to the price. This sounds very strange, but it explains the importance of Black-Scholes’ work.

(ii) One notices that

\[
V(s, t) = P(s, T - t).
\]

That is, if the current stock price is $s$ and there is $T - t$ time remaining toward to final time $T$, then the price of the contingent claim is $P(s, T - t)$, so is the value $V(s, t)$ of the portfolio.

(iii) Denote by

\[
\Gamma(x, \tau) := \frac{1}{\sqrt{2\pi \tau}} e^{-x^2/(2\tau)} \quad \forall x \in \mathbb{R}, \tau > 0.
\]
Then a change of variable

\[ y = \ln S + \sigma z \sqrt{T} + (\nu_0 - \sigma^2/2)T \]

we have

\[
P(S,T) = e^{-rT} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{(y - \ln S - [\nu_0 - \sigma^2/2]T)^2}{2\sigma^2 T}} f(e^y) dy.
\]

Direct differentiation gives

\[
\frac{\partial P(S,T)}{\partial S} = -e^{-\nu_0 T} S \int_{\mathbb{R}} \Gamma_x f dy,
\]

\[
\frac{\partial^2 P(S,T)}{\partial S^2} = -\frac{1}{S} \frac{\partial P}{\partial S} + e^{-\nu_0 T} \int_{\mathbb{R}} \Gamma_{xx} f dy,
\]

\[
\frac{\partial P(S,T)}{\partial T} = -\nu_0 P - (\nu_0 - \sigma^2/2) e^{-\nu_0 T} \int_{\mathbb{R}} \Gamma_x f dy + \sigma^2 \int_{\mathbb{R}} \Gamma f dy.
\]

Finally using \( \Gamma_\tau = \frac{1}{2} \Gamma_{xx} \) we then obtain

\[
\frac{\partial P}{\partial T} = -\nu_0 P + \left(\nu_0 - \frac{\sigma^2}{2}\right) S \frac{\partial P}{\partial S} + \frac{\sigma^2 S^2}{2} \left\{ \frac{\partial^2 P}{\partial S^2} + \frac{1}{S} \frac{\partial P}{\partial S} \right\} = \frac{\sigma^2 S^2}{2} \frac{\partial^2 P}{\partial S^2} + \nu_0 S \frac{\partial P}{\partial S} - \nu_0 P.
\]

Using \( V(s,t) = P(s,T-t) \) we can derive an equation for \( V \). We summarize our result as follows.

**Theorem 2.4** Suppose the risk-free interest rate is \( \nu_0 \) and the price \( S_t \) of a security satisfies

\[
\ln S_t = \ln S + \nu t + \sigma B_t \quad \forall t \geq 0
\]

where \( B_t \) is the Brownian motion process. Consider a derivative security whose payoff occurs only at \( t = T \) and equals \( f(S^T) \). Then its price at \( t = 0 \) is \( P(S,T) \) which, as function of \( S > 0, T > 0 \), satisfies the following **Black-Scholes’ equation**

\[
\frac{\partial P(S,T)}{\partial T} = \frac{\sigma^2 S^2}{2} \frac{\partial^2 P}{\partial S^2} + \nu_0 S \frac{\partial P}{\partial S} - \nu_0 P \quad \forall S > 0, T > 0. \tag{2.6}
\]

Analogously, at any time \( t \in [0, T] \) and spot price \( s \) of the security, the value \( V(s,t) \) of the derivative security satisfies

\[
\frac{\partial V(s,t)}{\partial t} + \frac{\sigma^2 s^2}{2} \frac{\partial V}{\partial s^2} + \nu_0 s \frac{\partial V}{\partial s} = \nu_0 V \quad \forall t < T, s > 0. \tag{2.7}
\]

Both \( P(S,T) \) and \( V(s,t) \) can be solved by supplying the respective initial conditions

\[
P(S,0) = f(S) \quad \forall S > 0, \quad V(s,T) = f(s) \quad \forall s > 0.
\]

It is very important to know that \( \nu \) plays no rule here. This is one of the Black-Scholes’ most significant contribution towards the investment science.

That \( \nu \) is irrelevant is due to the fact that only risk-neutral probability play roles here.
2.10 The Black–Scholes Equation

Here we provide a direct derivation for the Black-Scholes equation and a proof for the pricing formula.

Considered in the problem is a market system consisting of a risk-free bond and a risky stock. The price $B_t$ of the bond and the spot price $S_t$ of the stock obey the stochastic differential equations

$$\frac{dB_t}{B_t} = r dt, \quad \frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

where $\{W_t\}_{t \geq 0}$ is the standard Wiener (Brownian motion) process. To be more general, we shall not assume that $r, \mu, \sigma$ are constants. Instead, we assume that $r = r(S, t)$, $\mu = \mu(S, t)$ and $\sigma = \sigma(S, t)$ are given functions $S$ and $t$. Of course, for the Black-Scholes equation to be well-posed (existence of a unique solution) we do need to assume that $\sigma$ is positive and all $r, \mu, \sigma$ are bounded and continuous.

The problem here is to price, at time $t < T$ and spot stock price $S_t$, a derivative security (legal document) which will pay $f(S_T)$ at time $T$.

We shall carry out the task in two steps. In the first step, we show that if the derivative security can be replicated by a portfolio of stocks and bonds, then the value $V(S,t)$ of the portfolio at time $t$ and spot price $S$ must satisfy the Black-Scholes equation.

In the second step, we construct explicitly a self-financing portfolio whose value $V(S,t)$ is exactly the solution to the Black-Scholes equation. Thus, the price of the derivative is equal to $V(S,t)$; that is to say, the solution to the black-Scholes equation provides the price to the derivative security.

We have to say that step 1 is only a derivation of the equation. It is not part of the proof. If only a proof is needed, then step 1 is totally unnecessary. That is, only step 2 is the real proof that the price of derivative security satisfies the Black-Scholes equation. We present step 1 here is to let the reader see how Black-Scholes equation is first formally derived, and then shown to be the right one.

1. Assume that there is a replicating portfolio for the security. We denote by $n_s(S,t)$ the number of shares of stock and by $n_b(S,t)$ the number of shares of bond in the portfolio at time $t$ and spot stock price $S$.

   At time $t$ and spot stock price $S_t$, the portfolio’s value is
   $$V_t = n_s(S_t, t)S_t + n_b(S_t, t)B_t.$$  

   At time $t + dt$ and spot stock price $S_t + dt$, the portfolio’s value is
   $$V_{t+dt} = n_s(S_t, t)S_t + n_b(S_t, t)B_t,$$

   Thus, the change $dV_t$ of the value of the portfolio due to change of prices of the stock and bond is

   $$dV_t = V_{t+dt} - V_t = n_s(S_t, t)(S_t^{t+dt} - S_t) + n_b(S_t, t)(B_t^{t+dt} - B_t)$$  

   After we plug in the assumed dynamics for prices of the stock and bond. Note that $dV_t$ is a random variable, normally distributed.

   Now assume that $V_t$ can be written as $V(S_t, t)$ where $V(\cdot, \cdot)$ is a certain known function. Let’s see how can we do this. Given a function $V(s, t)$ on $\mathbb{R} \times (-\infty, T]$, when we replace $s$ by $S_t$, we obtain a
random variable defined on the same space as that of the Brownian motion. By Ito’s lemma, we know that \( V(s, t)|_{s=S^t} \) relates the Brownian motion according to

\[
dV(S^t, t) = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} (dS)^2
\]

\[
= \left\{ \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \right\} dt + \sigma S \frac{\partial V}{\partial S} dW. \tag{2.10}
\]

Here we have used the fact that \((dt)^2 = 0, dt dW = 0, (dW)^2 = dt\) so \((dS)^2 = \sigma^2 S^2 dt\).

Hence, to have \( V^t = V(s, t)|_{s=S^t} \), it is necessary and sufficient for coefficients of \(dt\) and \(dW\) in (2.9) and (2.10) to be exactly equal. Thus, we must have

\[
\begin{align*}
\mu n_s S + n_b r B &= \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2}, \\
\sigma n_s S &= \sigma S \frac{\partial V}{\partial S}.
\end{align*}
\]

This is equivalent to require

\[
n_s = \frac{\partial V}{\partial S}, \quad n_b = \frac{1}{r B} \left\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \right\}. \tag{2.11}
\]

Therefore, the portfolio and the only portfolio that can replicate the derivative security is \( n_s \) shares of stock and \( n_b \) shares of bound, where \( n_s \) and \( n_b \) are as above. This is the only way that we can get rid of the randomness caused by Brownian motion process \( dW \) in the price change of stock.

We repeat a few more words about the randomness. Here \( S^{t+dt} \) is a random variable with normal distribution. The function \( V(s, t + dt) \) itself is not a random variable, it becomes a random variable only when we replace \( s \) by \( S^{t+dt} \). That the random variable \( n_s(S^t, t)S^{t+dt} + n_b(S^t, t)B^{t+dt} \) is exactly the same as \( V(S^{t+dt}, t + dt) \) is a very strong requirement. In the discrete model, we have learned of how to construct a replicating portfolio that matches exactly the required payment for contingent claim, regardless of which state the stock price lands on. Here is the same situation. The randomness is get rid of by matching the coefficients of two \( dV \)'s. The former from the actual behavior of the stock price change, the other from Ito's lemma and our hypothesis that \( V^{t+dt} = V(s, t + dt)|_{s=S^{t+dt}} \), where \( V(s, t + dt) \) is a function to be constructed without knowledge of the outcomes of actual price.

The value of the portfolio is \( V = n_s S + n_b B \). Hence we need, in view of (2.11),

\[
V = n_s S + n_b B = S \frac{\partial V}{\partial S} + B \frac{1}{r B} \left\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \right\}.
\]

After simplification, this becomes

\[
\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} + r S \frac{\partial V}{\partial S} = r V,
\]

which is exactly the famous Black-Scholes equation.

So far we have derived the Black-Scholes equation. We know that if a derivative security can be replicated by a portfolio, then its value or price must satisfy the Black-Scholes equation.

2. Now let \( V \) be the solution to the Black-Scholes equation with “initial” condition \( V(s, T) = f(s) \) for all \( s > 0 \). Let \( n_s \) and \( n_b \) be defined as in (2.11). Consider the portfolio consisting of \( n_s(S^t, t) \) shares of stock and \( n_b(S^t, t) \) shares of bound at time \( t + 0 \) and spot stock price \( S^t \).

First of all the value of the portfolio is

\[
n_s S + n_b B = V(s, t)
\]
by the definition of \( n_s, n_b \) and the differential equation for \( V \).

Now we show that the portfolio is self-financing. For this, we calculate the capital needed to maintain such a portfolio.

At time \( t = 0 \), we have a portfolio of \( n_s(S^t, t) \) shares of stock and \( n_b(S^t, t) \) shares of bond. At time \( t + dt \) before rebalancing, its value is \( n_s(S^{t+dt}, t + dt) S^{t+dt} + n_b(S^{t+dt}, t + dt) B^{t+dt} \) which we wish to rebalance to \( n_s(S^{t+dt}, t + dt) \) shares of stock and \( n_b(S^{t+dt}, t + dt) \) shares of bond. The capital \( \delta \) needed to perform such a trade is

\[
\delta := \{ n_b(S^{t+dt}, t + dt) S^{t+dt} + n_s(S^{t+dt}, t + dt) B^{t+dt} \} - \{ n_s(S^t, t) S^{t+dt} + n_b(S^t, t) B^{t+dt} \} (2.12)
\[
= \left\{ n_b(S^{t+dt}, t + dt) S^{t+dt} + n_b(S^{t+dt}, t + dt) B^{t+dt} \right\} - \left\{ n_s(S^t, t) S^t + n_b(S^t, t) B^t \right\}
\]

\[
+ n_s(S^t, t) \left[ S^t - S^{t+dt} \right] + n_b(S^t, t) \left[ B^t - B^{t+dt} \right]
\]

\[
= dV - n_s dS - n_b dB
\]

\[
= dV - \frac{\partial V}{\partial S} \left\{ \mu S dt + \sigma S dW \right\} - \frac{1}{2} \left\{ \frac{\partial^2 V}{\partial S^2} \right\} \left\{ \sigma^2 S^2 dt \right\} - r B dt
\]

\[
= dV - \left\{ \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \right\} dt - \frac{\partial V}{\partial S} dW - r B dt
\]

\[
= 0.
\]

Thus, the portfolio is self-financing. Since the outcome of the portfolio at time \( T \) is \( V(S^T, t) \) which is equal exactly \( f(S^T) \), regardless what \( S^T \) is, we see that the value of the contingent claim at any time \( t \) and spot stock price \( S^t \) had to be \( V(S^t, t) \). Since if the the claim is sold for more, say at \( V(S^t, t) > V(S^t, t) \) at this time, the value of the portfolio exactly pays the claim. Thus there is no future payoff and we have a profit at time \( t \). Similarly, if \( V(S^t, t) < V(S^t, t) \) one can do other way around. Since such arbitrage is excluded from the mathematical perfection, we conclude that the value of the contingent claim has to be \( V(S^t, t) \).

Therefore, the price of the derivative security must be equal to \( V \) which is the unique solution to the Black-Scholes equation.

We summarize our result as follows.

**Theorem 2.5** Consider a system consisting of a risk-free asset and a risky asset whose price obeys a geometric Brownian motion process. Then any contingent claim with only a fixed one time payment can be uniquely replicated and therefore be priced, and the price can be calculated from the solution to the Black-Scholes equation.

---

**Exercise 2.23.** Complete the argument that if the price \( \hat{V}(S^t, t) \) of the contingent claim is smaller than \( V(S^t, t) \) at some time \( t < T \) and some spot stock price \( S^t \), then there is an arbitrage.

**Exercise 2.24.** Note that \( \delta \) in (2.12) can be expresses as

\[
\delta = S^{t+dt}[n_s(S^{t+dt}, t + dt) - n_s(S^t, t)] + B^{t+dt}[n_b(S^{t+dt}, t + dt) - n_b(S^t, t)]
\]

\[
= S^{t+dt}dn_s + B^{t+dt}dn_b = \{S + dS\}dn_s + (B + dB)dn_b.
\]
Using Ito’s lemma, the expression of \( n_s \) and \( n_b \) in (2.11), and the Black-Scholes equation for \( V \) show directly that \( \delta = 0 \).

Exercise 2.25. Assume that \( r \) and \( \sigma \) are constants and \( \sigma > 0 \). Make the change of variables from \((S,t,V)\) to \((x,\tau,v)\) by

\[
x = \ln S + \left(r - \frac{\sigma^2}{2}\right)(T - \tau), \quad \tau = \frac{\sigma^2}{2}(T - t), \quad V(s,t) = v(x,\tau)e^{-r(T-t)}
\]

Show that the Black-Scholes equation for \( V(S,t) \) becomes the following linear equation for \( v \):

\[
\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} \quad \forall x \in \mathbb{R}, \tau > 0, \quad v(x,0) = f(e^x) \quad \forall x \in \mathbb{R}.
\]

Also show that the solution for \( v \) is given by the following formula:

\[
v(x,t) = \frac{1}{\sqrt{4\pi\tau}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4\tau}} f(e^y)dy \quad \forall x \in \mathbb{R}, \tau > 0.
\]

Exercise 2.26. Assume risk-free rate is \( r \) and volatility of a stock is \( \sigma > 0 \). Both \( r \) and \( \sigma \) are constants. Find the price for European put and call options, with duration time \( T \) and strike price \( K \).
Bibliography


