# LECTURE NOTES IN MATHEMATICAL FINANCE

X. Sheldon Lin
Department of Statistics & Actuarial Science
University of Iowa
Iowa City, IA 52242

Phone: (319)335-0730
Fax: (319)335-3017
Email: shlin@stat.uiowa.edu

Comments are welcome!

# Contents

# Part I

# Discrete-Time Finance Models

# Chapter 1

# Basic Concepts and One Time-Period Models

## 1.1 The Basic Setup

We consider a security market with the following conditions:

- There are only two consumption dates: the initial date $t = 0$ and the terminal date $t = T$. Trading takes place at $t = 0$ only.

- There are finite number of states of economy

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_J\}$$

  with the probability at state $\omega_j$ being $P(\omega_j)$.

  Hence $(\Omega, \mathcal{F}, P)$ consists of a probability space with the $\sigma$-algebra being all the subsets of $\Omega$.

- There are $N$ primitive securities. The $n$-th security has price $p_n$ at time 0 and terminal payoff

$$d_n = \begin{pmatrix} d_n(\omega_1) \\ d_n(\omega_2) \\ \vdots \\ d_n(\omega_J) \end{pmatrix}$$

  Thus, we have a price system

$$p = (p_1, p_2, \cdots, p_N)',$$

  where $'$ denotes the corresponding transpose, and the payoff matrix

$$D = \begin{pmatrix} d_1(\omega_1) & \cdots & d_N(\omega_1) \\ \vdots & & \vdots \\ d_1(\omega_J) & \cdots & d_N(\omega_J) \end{pmatrix}$$

- Investors are price takers and have the homogeneous belief $P = (P(\omega_1), P(\omega_2), \cdots, P(\omega_J))$.

- There is only one perishable consumption good.

## 1.2 Trading Strategies

If an investor possesses $\theta_n$ shares of security $n$, the portfolio of the securities of the investor has the payoff $\sum_{n=1}^{N} \theta_n d_n$ at time $T$. Let $e(0), e(T)$ be the initial endowment and the terminal endowment for the investor, respectively. Thus, the investor's consumptions are

$$c(0) = e(0) - \sum_{n=1}^{N} \theta_n p_n, \tag{1.1}$$

$$c(T) = e(T) + \sum_{n=1}^{N} \theta_n d_n. \tag{1.2}$$

We call $\theta = (\theta_1, \theta_2, \cdots, \theta_N)'$ a trading strategy. The set $\mathcal{B}(e, p)$ containing all consumption processes $c = (c(0), c(T))$ over all $\theta$ is called the budget set with respect to the endowment process $e = (e(0), e(T))$ and the price system $p$. Mathematically, a budget set is an affine space of $\mathbf{R}^{J+1}$.

A consumption process is said to be attainable if its terminal consumption can be expressed as the payoff of a portfolio, i.e.

$$c(T) = \sum_{n=1}^{N} \theta_n d_n.$$

It is easy to see that a consumption process is attainable if and only if

$$\text{Rank}(D) = \text{Rank}(D, c(T)).$$

It is also easy to see that the terminal consumption of any attainable consumption process is in the image of the matrix $D$, regarded as a linear map. Thus, every consumption process is attainable if and only if $\text{Rank}(D) = J$, therefore, if and only if there are $J$ independent securities. In this case, we say the market is complete. Otherwise, we say the market is incomplete. We will see later that if the market is complete, any consumption process can be priced uniquely.

When the market is not complete, there is a need to create new securities in order to complete the market. One approach is to create derivative securities on the existing securities such as European-type options.

A European call option written on a security gives its holder the right( not obligation) to buy the underlying security at a prespecified price on a prespecified date; whilst a European put option written on a security gives its holder the right( not obligation) to sell the underlying security at a prespecified price on a prespecified date. The prespecified price is called the strike price and the prespecified date is called the expiration or maturity date.

Given a security with terminal payoff $\bar{d} = (\bar{d}(\omega_1), \cdots, \bar{d}(\omega_J))'$, the payoff of a European call option with strike price $K$ then is

$$\max\{\bar{d} - K,\ 0\}.$$

Similarly, the payoff of a European put option with strike price $K$ then is

$$\max\{K - \bar{d},\ 0\}.$$

**Example 1.1** Consider two securities with payoff $d_1 = (1,2,4)'$, $d_2 = (2,0,1)'$, respectively. Since the number of the states is 3 and the number of securities is 2, the market is not complete. Write a European call option on the first security with strike price 1. Then its payoff is $d_3 = (0,1,3)'$. These three securities are algebraically independent and therefore complete the market. $\square$

We now consider no arbitrage strategies. A trading strategy $\theta = (\theta_1, \theta_2, \cdots, \theta_N)'$ is said to admit arbitrage if either

$$\sum_{n=1}^{N} \theta_n p_n = 0, \quad \text{and} \quad \sum_{n=1}^{N} \theta_n d_n \geq 0 \tag{1.3}$$

with $\sum_{n=1}^{N} \theta_n d_n(\omega_j) > 0$ for some $j$,

or

$$\sum_{n=1}^{N} \theta_n p_n < 0, \quad \text{and} \quad \sum_{n=1}^{N} \theta_n d_n \geq 0 \tag{1.4}$$

These conditions imply that with the zero endowment process, we will be able to obtain a nonzero nonnegative consumption process.

## 1.3 Characterisation of No-Arbitrage Strategies

We are now looking for the necessary and sufficient condition under which the price system does not admit arbitrage.

We first recall the Hahn-Banach Theorem which will be used to derive the condition.

**Theorem 1.1**(Hahn-Banach) Let $A$ and $B$ be two disjoint convex sets in a Hilbert space $\mathcal{H}$. Assume that there exist $a \in A$ and $b \in B$ such that $d(A, B) = \|a - b\|$, where

$d(A,B)$ is the distance between $A$ and $B$ defined by

$$d(A,B) = \inf\{\|x-y\|;\ \text{for any } x \in A \text{ and } y \in B\}.$$

Then, there exists a $z \in \mathcal{H}$ and a scalar $h$ such that for any $x \in A,\ x \bullet z > h$, and for any $y \in B,\ y \bullet z < h$. See Appendix B for a proof.

It is easy to see that the price system admits arbitrage if and only if some consumption process with zero endowment process lies in the set $\mathbf{R}_+^{J+1} - \{0\}$. Thus, no arbitrage condition is equivalent to the condition that the sets $\mathcal{B}(0,p)$ and $\mathbf{R}_+^{J+1} - \{0\}$ are separate. Suppose this is the case. Let $A = \{x \in \mathbf{R}_+^{J+1}; x_0 + \cdots + x_J \geq \frac{1}{2}\}$. Then it can be shown (see Appendix B) that there exist $a \in A$ and $b \in \mathcal{B}(0,p)$ such that $d(A, \mathcal{B}(0,p)) = \|a-b\|$. By the Hahn-Banach Theorem, there is a $z = (z_0, z_1, \cdots, z_J)'$ and a scalar $h$ such that for any $x \in A,\ x'z > h$, and for any $y \in \mathcal{B}(0,p),\ y'z < h$. Since $\mathcal{B}(0,p)$ is a linear space, $y'z$ is either 0 or unbounded from above on $\mathcal{B}(0,p)$. Thus, $y'z = 0$ for all $y \in \mathcal{B}(0,p)$. This means that

$$\theta' D' \bar{z} = z_0 \theta' p$$

for any $\theta$, where $\bar{z} = (z_1, \cdots, z_J)'$. That $\theta$ is arbitrary implies

$$D'\alpha = p, \tag{1.5}$$

where $\alpha = (\frac{z_1}{z_0}, \frac{z_2}{z_0}, \cdots, \frac{z_J}{z_0})'$. It is easy to see that $h > 0$. Let $s_j$ be the vector whose $(j+1)$-th component is 1 and others 0. Then $s_j \in A$ and hence $s_j'z > 0$, which implies $z_j > 0,\ j = 0, 1, \cdots, J$. Thus, $\alpha > 0$.

Therefore, the price system does not admit arbitrage implies that there is a vector $\alpha$, all of whose components are positive, such that the equation (1.5) holds.

Conversely, if there is a positive vector $\alpha$ such that the equation (1.5) holds, there will be no arbitrage. Otherwise, let $\hat{\theta}$ be an arbitrage trading strategy. Multiplying $\hat{\theta}'$ both sides of the equation from the left gives an inequality, which is a contradiction.

Summerizing the above arguments, we conclude that

**Theorem 1.2** The price system does not admit arbitrage if and only if there is a positive vector $\alpha$ such that

$$D'\alpha = p. \tag{1.6}$$

Let us now consider the case that one of these securities is a riskless bond, say the first security. Denote $r$ the rate of return of the bond. Thus, $d_1 = (1+r)p_1$. The first equality in equation (1.6) gives $(1+r)(\alpha_1 + \alpha_2 + \cdots + \alpha_J) = 1$. Let $Q(\omega_j) = (1+r)\alpha_j,\ j = 1, \cdots, J$. Then, $Q = (Q(\omega_1), \cdots, Q(\omega_J))'$ is a probability measure on $(\Omega, \mathcal{F})$ and the equation ((1.6) becomes

$$D'Q = (1+r)p. \tag{1.7}$$

The $n$th scalar equation in (1.7) gives

$$\sum_{j=1}^{J} d_n(\omega_j)Q(\omega_j) = (1+r)p_n.$$

Thus,

$$p_n = \frac{E_Q(d_n)}{1+r}$$

and

$$E_Q(R_n) = \frac{1}{p_n}\sum_{j=1}^{K} d_n(\omega_j)Q(\omega_j) - 1 = r,$$

where $R_n = \frac{d_n}{p_n} - 1$.

Hence, the price system does not admit arbitrage if and only if there is a probability measure $Q$ on $(\Omega, \mathcal{F})$ such that under this measure, the price of each security is the discounted value of its expected payoff and all securities have the same expected rate of return. The probability measure $Q$ is often referred to as a risk-neutral probability measure. If the market is complete it uniquely exists under no arbitrage condition. However, it the market is not complete, there are more than one risk-neutral probability measure.

## 1.4 Valuation

We now denote the time-0 price of a consumption process $c = (c(0), c(T))$ by $\phi(c)$. Then, no arbitrage implies that for any attainable comsumption process with $c(T) = \sum_{n=1}^{N} \theta_n d_n$,

$$\phi(c) = c(0) + \sum_{n=1}^{N} \theta_n p_n. \tag{1.8}$$

This formula itself is trivial but it represents a very important principle in pricing securities. That is, if the payoff of a security can be hedged by forming a portfolio of the existing securities, the price should be equal to the initial value of the portfolio. We will see later on that the same principle is applied to many multi-period models.

On the other hand, since the price of each existing security can be written as the discounted expected value of its payoff under a risk-neutral probability measure, we have

$$\phi(c) = c(0) + \frac{\sum_{n=1}^{N} \theta_n E_Q(d_n)}{1+r} = c(0) + \frac{1}{1+r} E_Q\{\sum_{n=1}^{N} \theta_n d_n\}. \tag{1.9}$$

This formula represents another important principle in pricing securities. It says that the price of a security is the discounted expected value of its payoff under a risk-neutral probability measure, discounted at the risk-free rate. This principle is often applied to American type options as well as continuous time financial models.

It is easy to see that the price of an attainable consumption process is uniquely determined no matter which risk-neutral probability measure is used. Thus, when the market is complete, every consumption process is priced uniquely.

Consider now the following securities: for each $j = 1, 2, \cdots, J$, the payoff of the $j$-th security is

$$\chi_j(\omega) = \begin{cases} 1, & \omega = \omega_j \\ 0, & \text{otherwise} \end{cases} \tag{1.10}$$

These securities are usually referred to as the Arrow-Debreu securities which pay one unit at one state and nothing elsewhere. Their prices $\phi_j$, $j = 1, 2, \cdots, J$, called the Arrow-Debreu prices or the state prices, can be easily determined when the market is complete.

For each $j$,

$$\phi_j = \frac{1}{1+r} E_Q(\chi_j) = \frac{1}{1+r} Q(\omega_j). \tag{1.11}$$

In other words, the risk-neutral probability for each state is actually the accumulated value of the corresponding state price at the riskfree rate.

**Example 1.2** Consider an economy with only two states: the upstate and the downstate, respectively. The probability of the upward state is $q$ and the other is $1 - q$. There are two securities, one riskless bond with interest rate $r$ and one stock with initial price $S$ and with return $u$ at the upstate and return $d$ at the downstate, $u > d$(Figure 1.1).

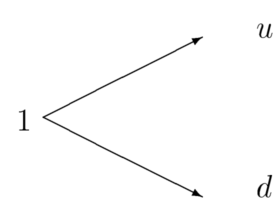


Figure 1.1: **Return of the risky security**

The payoff matrix then is

$$D' = \begin{pmatrix} 1+r & 1+r \\ Su & Sd \end{pmatrix}$$

The no arbitrage condition is equivalent to $u > 1 + r > d$. The risk-neutral probability measure $Q = (q_u, q_d)'$ satisfies

$$q_u = \frac{1+r-d}{u-d}, \quad q_d = \frac{u-1-r}{u-d}.$$

For any given payoff $C = (C_u, C_d)'$, which could be the payoff of a call or put option, we have the price

$$\phi_C = C_u q_u + C_d q_d = \frac{(1+r-d)C_u + (u-1-r)C_d}{u-d}. \tag{1.12}$$

On the other hand, suppose that a portfolio $\Delta S + B$, where $\Delta$ is the number of shares of the stock and $B$ is the bond value, gives the same payoff as $(C_u, C_d)'$. We then have

$$\Delta S u + B(1+r) = C_u \tag{1.13}$$

$$\Delta S d + B(1+r) = C_d \tag{1.14}$$

Thus, $\Delta = \frac{C_u - C_d}{S(u-d)}$, $B = \frac{uC_d - dC_u}{(1+r)(u-d)}$. It is easy to verify that $\phi_C = \Delta S + B$. $\Delta$ is also the derivative of the price of the security with payoff $C$ with respect to the stock price and is often called delta by practitioners. $\square$

We now consider pricing consumption processes in an incomplete market. It suffices to price only the respective terminal consumptions since the price of a consumption process is simply the sum of its initial consumption and the price of its terminal consumption.

Let $\psi$ be a price system on the terminal consumption space $\{c(T) \in \mathbf{R}^J\}$. Then, $\psi(\lambda c(T)) = \lambda\psi(c(T))$ and $\psi(c^1(T) + c^2(T)) = \psi(c^1(T)) + \psi(c^2(T))$. In other words, $\psi$ is a linear functional on $\mathbf{R}^J$. Furthermore, the fact that $\psi$ is a price system and it does not admit arbitrage implies that

$$\psi(\chi_j) > 0, \quad j = 1, \cdots, J, \tag{1.15}$$

and

$$\sum_{j=1}^{J} \psi(\chi_j) = \frac{1}{1+r}. \tag{1.16}$$

If we require the price system $\psi$ to be consistent with the current price system $p = (p_1, \cdots, p_N)'$, i.e. $\psi(d_n) = p_n, n = 1, \cdots, N$, we have

$$\sum_{j=1}^{J} d_n(\omega_j)\psi(\chi_j) = p_n. \tag{1.17}$$

Define $Q_\psi(\omega_j) = (1+r)\psi(\chi_j), j = 1, \cdots, J$. These three conditions (1.15), (1.16) and (1.17) give

$$D'Q_\psi = (1+r)p. \tag{1.18}$$

Thus, they together are also sufficient conditions for a consistent no-arbitrage price system for all consumption processes.

The correspondance $\psi \to Q_\psi$ is an one-to-one correspondance since a linear functional is uniquely determined by its values on a basis which is $\chi_j, j = 1, \cdots, J,$ in our case. Thus, the number of price functionals is equal to the number of the risk-neutral probability measures for the price system $p = (p_1, \cdots, p_N)'$. Moreover, if $\mathrm{Rank}(D) = N$, there are exactly $J - N$ independent price functionals and any other price functional is a linear combination of those.

From the equation (1.18), for any consistent no-arbitrage price system $\psi$, there is a unique risk-neutral probability measure $Q_\psi$ such that

$$\psi(c(T)) = \frac{1}{1+r} E_{Q_\psi}(c(T)). \tag{1.19}$$

## 1.5 Risk Premiums

As in the preceding section, we let

$$R_n = \frac{d_n}{p_n} - 1$$

be the rate of return of security $n$. Denote the expected rate of return under the probability measure $P$ by

$$\mu_n = E_P(R_n) = \frac{E_P(d_n)}{p_n} - 1.$$

This is the expected rate of return based on the investors homogeneous belief. Hence, the difference $\mu_n - r$ between the expected rate of return and the riskfree rate of return is the risk premium for security $n$. If we let $z = \frac{P}{Q} - 1$, independent of the securities, then

$$\mu_n - r = E_P(R_n) - E_Q(R_n) = E_P[(1 - \frac{P}{Q})R_n] = -\mathrm{Cov}_P(z, R_n). \tag{1.20}$$

For any given portfolio $\sum_{n=1}^N \theta_n p_n$ with the rate of return

$$R = \frac{\sum_{n=1}^N \theta_n d_n}{\sum_{n=1}^N \theta_n p_n} - 1,$$

and the expected rate of return $\mu = E_P(R)$,

$$\mu - r = E_P(R) - E_Q(R) = -\mathrm{Cov}_P(z, R), \tag{1.21}$$

since $E_Q(R) = r$.

If $z$ is attainable, i.e. $z = \sum_{n=1}^{N} \xi_n d_n$, then,

$$z = (\sum_{n=1}^{N} \xi_n p_n)(1 + R_z),$$

where $R_z$ is the rate of return of portfolio $z$. We have

$$E_P(R_z) - r = -(\sum_{n=1}^{N} \xi_n p_n)\mathrm{Var}_P(R_z),$$

and

$$\mu - r = -(\sum_{n=1}^{N} \xi_n p_n)\mathrm{Cov}_P(R_z, R).$$

Therefore,

$$\mu - r = \frac{\mathrm{Cov}_P(R_z, R)}{\mathrm{Var}_P(R_z)}(E_P(R_z) - r). \tag{1.22}$$

The equation (1.22) is in the form of the well-known Capital Asset Pricing Model(CAPM). $z$ is referred to as the market portfolio and the quantity $\frac{\mathrm{Cov}_P(R_z,R)}{\mathrm{Var}_P(R_z)}$ is referred to as the market beta. Since the covariance operator and the variance operator are invariant under parallel shifting $R \to R + a$, the above formula also holds when the rates of returns are replaced by the returns per unit. The latter is used in the standard CAPM setting.

# Chapter 2

# Discrete-Time Stochastic Processes and Lattice Models

## 2.1 Discrete-Time Stochastic Processes

let $(\Omega, \mathcal{F}, P)$ be a probability space. $\mathcal{F}$ then is the collection of all possible random events. Thus, $\mathcal{F}$ represents all the information contained in this probability space.

Let $\mathcal{F}_1$ be another $\sigma$-algebra defined on $\Omega$. If $\mathcal{F}_1$ is coarser than $\mathcal{F}$, the probability space $(\Omega, \mathcal{F}_1, P)$ contains less information than $(\Omega, \mathcal{F}, P)$ does.

**Example 2.1** Let $\mathcal{F}_1 = \{\phi,\ \Omega\}$. $\mathcal{F}_1$ is the coarsest $\sigma$-algebra which contains no information at all.

Let $\mathcal{F}_2$ be the set of all subsets of $\Omega$. $\mathcal{F}_2$ is the finest $\sigma$-algebra which contains all the information from the underlying space $\Omega$. ▭

Let $X$ be a random variable defined on $(\Omega, \mathcal{F}, P)$. How much information would we be able to obtain from $X$? Obviously, any random event we could observe through $X$ will be represented by the values of $X$ on the random event. If two events give the same range for $X$, we will be unable to distinguish them. Hence, all possible random events we can observe from $X$ are in the $\sigma$-algebra generated by events $\{X \leq x\}$, for all real numbers $x$. We call the $\sigma$-algebra generated by $\{X \leq x\}, x \in \mathbf{R}$, the Borel $\sigma$-algebra with respect to $X$ and denote it as $\mathcal{B}_X$. Hence, $\mathcal{B}_X$ represents all the information that can be obtained from $X$.

**Example 2.2** Suppose that $X$ is a random variable which only takes a finite number of different values $u_1, u_2, \cdots, u_J$. Let $\omega_j = \{X = u_j\}$, $j = 1, 2, \cdots, J$. Then $\mathcal{B}_X$ is the collection of all subsets of $\{\omega_1, \omega_2, \cdots, \omega_J\}$. $\square$

Consider now all the time-dependent random events in $\mathcal{F}$. Let $\mathcal{F}_t$ be the collection of all possible random events that may happen before or at time $t$. Then, (i) $\mathcal{F}_t$ is a $\sigma$-algebra coarser than $\mathcal{F}$; (ii) if $t < s$, $\mathcal{F}_t \subset \mathcal{F}_s$. Thus, $\{\mathcal{F}_t,\ t \geq 0\}$ define an information structure on $(\Omega, \mathcal{F}, P)$, with $\mathcal{F}_t$ representing the information up to time $t$. In probability theory, any collection of $\sigma$-algebras which satisfies (i) and (ii) is called a filtration on $(\Omega, \mathcal{F}, P)$ and the quadruplet $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is called a filtered space.

At this moment let us consider a discrete-time setting: $t = t_0, t_1, \cdots$. Without loss of generality, assume $t = 0, 1, 2, \cdots$. If we have a sequence of random variables $X(0), X(1), \cdots$, $X(t), \cdots$, such that $X(t)$ is a random variable on $(\Omega, \mathcal{F}_t, P)$, then the sequence $X(0), X(1), \cdots$, $X(t), \cdots$, is called an adapted discrete-time stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. In these notes, we always assume that a stochastic process is adapted and simply call it a stochastic process.

Given a stochastic process $X(t)$, $t = 0, 1, \cdots$, we want to see how much information we will be able to obtain from it. As we have mentioned above, $\mathcal{B}_{X(t)}$ is the information we can obtain from the random variable $X(t)$. Thus, the information up to time $t$ from the stochastic process $X(t)$, $t = 0, 1, \cdots$, is the $\sigma$-algebra generated by the random events in $\mathcal{B}_{X(0)}, \mathcal{B}_{X(1)}, \cdots, \mathcal{B}_{X(t)}$. In other words, it is the smallest $\sigma$-algebra containing $\mathcal{B}_{X(0)}, \mathcal{B}_{X(1)}, \cdots, \mathcal{B}_{X(t)}$. We denote this $\sigma$-algebra as $\mathcal{B}_t$. It is easy to see that

$$\mathcal{B}_0 \subset \mathcal{B}_1 \subset \cdots \subset \mathcal{B}_t \subset \cdots \subset \mathcal{F}$$

Thus, $\mathcal{B}_t,\ t = 0, 1, \cdots$, form a filtration on $(\Omega, \mathcal{F}, P)$, called the Borel or natural filtration with respect to $X(t),\ t = 0, 1, \cdots$. This filtration contains exact information obtained from $X(t),\ t = 0, 1, \cdots$, and $\mathcal{B}_t$ is the exact information obtained from $X(t),\ t = 0, 1, \cdots$, up to time $t$. Since $X(t)$ is $\mathcal{F}_t$-measurable, $\mathcal{B}_t \subset \mathcal{F}_t$. Hence the information contained in the Borel filtration is no more than that in the original filtration $\mathcal{F}_t,\ t = 0, 1, \cdots$.

So far, we assume that we are given an information structure or filtration $\mathcal{F}_t,\ t = 0, 1, \cdots$. Based on this information structure, we define a stochastic process and its Borel information structure. But very often, what we have is a sequence of random variables $X(t),\ t = 0, 1, \cdots$, defined on $(\Omega, \mathcal{F}, P)$ without having the information structure $\mathcal{F}_t,\ t = 0, 1, \cdots$. In other words, the sequence of random variables is the only source we can obtain information from. In this case, we may directly define $\mathcal{B}_t$ from the sequence $X(t),\ t = 0, 1, \cdots$, as above. It is easy to see that $(\Omega, \mathcal{F}, \mathcal{B}_t, P)$ is a filtered space and $X(t),\ t = 0, 1, \cdots$, is a stochastic process on it.

Finally, we extend our discussion to vector-valued stochastic processes. We say a sequence of random vectors

$$(X_1(t), X_2(t), \cdots, X_N(t)),\ t = 0, 1, \cdots,$$

is a stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ if for any $n = 1, 2, \cdots, N, \quad X_n(t),\ t = 0, 1, \cdots$, is a stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. The corresponding Borel $\sigma$-algebra $\mathcal{B}_t$ is defined as the smallest $\sigma$-algebra containing the Borel $\sigma$-algebras generated by $X_n(s),\ n = 1, 2, \cdots, N;\ s = 0, 1, \cdots, t$.

## 2.2 Random Walks

Random walks are one of the simplest discrete-time stochastic processes. Because they are simple, intuitive and have other appealing features, they have been widely used in modeling securities. In this section we show how a random walk is constructed and how it can be used to model securities.

Let $Y_1, Y_2, \cdots, Y_k, \cdots$ be a sequence of independent, identically distributed(iid) Bernoulli random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. First, let us assume that for a given $h > 0$,

$$Y_k = \begin{cases} h \\ -h \end{cases} \tag{2.1}$$

and

$$\Pr(Y_k = h) = \Pr(Y_k = -h) = \frac{1}{2}. \tag{2.2}$$

We now construct a random walk over the time period $[0,\ T]$ as follows:

> An object starts at a position marked 0. It moves once only at a time interval with length $\tau > 0$. We choose $\tau$ such that $T$ is a multiple of $\tau$. During each time interval the object either moves up $h$ units with probability $\frac{1}{2}$ or moves down $h$ units with probability $\frac{1}{2}$.

Let $X(t)$ be the position of the object at time $t$. Then we have

$$X(t) = Y_1 + Y_2 + \cdots + Y_{\bar{t}}, \tag{2.3}$$

where $t = \bar{t}\tau$. Apparently, $X(t)$ is binomially distributed. The Borel $\sigma$-algebra $\mathcal{B}_t$ with respect to this process is the collection of all the subsets of the set of all possible paths up to time $t$. For example, $\mathcal{B}_2$ is the collection of all the subsets of

$$\{(h, h), (h, -h), (-h, h), (-h, -h)\}.$$

Let $P(x, t)$ denote the probability that the object is at position $x$ at time $t$, i.e. $P(x, t) = \Pr(X(t) = x)$. Then when $x$ is reached by moving up $m$ times and moving down $\bar{t} - m$ times, we have

$$P(x, t) = \begin{pmatrix} \bar{t} \\ m \end{pmatrix} (\frac{1}{2})^{\bar{t}}, \quad x = (2m - \bar{t})h, \tag{2.4}$$

$m = 0, 1, \cdots, \bar{t}$.

We can also easily compute its mean and variance.

$$\begin{aligned} E(X(t)) &= \bar{t}E(Y_1) = 0, \quad (2.5) \\ Var(X(t)) &= \bar{t}Var(Y_1) = \bar{t}h^2 = t\frac{h^2}{\tau}. \quad (2.6) \end{aligned}$$

Moreover, there is a recursive relation among $P(x, t), t = 0, \tau, \cdots$. It follows from

$$P(x, t+\tau) = \Pr(X_{\bar{t}} = x - Y_{\bar{t}+1}) = E(\Pr(X_{\bar{t}} = x - y)|Y_{\bar{t}+1} = y)$$

that

$$P(x, t+\tau) = P(h, \tau)P(x-h, t) + P(-h, \tau)P(x+h, t), \quad (2.7)$$

with $P(0, 0) = 1, \quad P(x, 0) = 0, x \neq 0$.

In the above case,

$$P(x, t+\tau) = \frac{1}{2}[P(x-h, t) + P(x+h, t)]. \quad (2.8)$$

Next, we consider random walks with drift. As we have seen in the previous discussion that the mean and variance of the random walk discussed are proportional to the time that has passed. In other words, the average mean and variance over time remain constant. Furthermore, the move at time $t$ only depends on the position of the object at time $t$, not the positions in the past, which is called the Markov property. Those properties are appealing since many risky securities enjoy the same properties. However, in practice, the average mean of a risky security is often nonzero. Thus there is a need to extend the random walk we have considered.

Let us now design a random walk $X(t), \ t = 0, \tau, 2\tau, \ldots,$ with a constant average mean and a constant average variance, namely

$$E(X(t)) = t\mu, \quad Var(X(t)) = t\sigma^2. \quad (2.9)$$

There are two approaches to achieve this goal.

1. Adjust the probabilities of the up movement and the down movement. Let

$$\Pr(Y_k = h) = q, \quad \Pr(Y_k = -h) = 1 - q.$$

To satisfy equations in (2.9), we must have

$$\frac{h(2q-1)}{\tau} = \mu, \quad \frac{4h^2 q(1-q)}{\tau} = \sigma^2,$$

which yields

$$h = \sqrt{\sigma^2\tau + \mu^2\tau^2}, \quad q = \frac{1}{2}\Big[1 + \sqrt{\frac{1}{1+\sigma^2/\mu^2\tau}}\Big]. \tag{2.10}$$

The corresponding recursive formula becomes

$$P(x, t+\tau) = qP(x-h, t) + (1-q)P(x+h, t). \tag{2.11}$$

2. Adjust the magnitude of the up movement and the down movement separately.

$$\Pr(Y_k = h_1) = \frac{1}{2}, \quad \Pr(Y_k = -h_2) = \frac{1}{2}.$$

Since

$$E(Y_k) = \frac{1}{2}(h_1 - h_2), \quad Var(Y_k) = (\frac{h_1 + h_2}{2})^2,$$
$$\frac{h_1 - h_2}{2\tau} = \mu, \quad \frac{(h_1 + h_2)^2}{4\tau} = \sigma^2.$$

This yields

$$h_1 = \mu\tau + \sigma\sqrt{\tau}, \quad h_2 = -\mu\tau + \sigma\sqrt{\tau}. \tag{2.12}$$

The recursive formula is the same as (2.8).

We now illustrate how a random walk with drift can be used to model the price movement of a risky security.

Consider a risky security for the time period $[0,\ T]$. Assume that during the period $[0,\ T]$, there are $\bar{T}$ trading dates, $\bar{t} = 0, 1, \cdots, \bar{T} - 1$, separated in regular intervals, i.e. $\bar{t} = t/\tau$, where $\tau$ is the time between two consecutive trading dates. At each time $t$, there are only two states of economy over the next time interval: the upstate and the downstate. The probablities of the upstate and the downstate are $q$ and $1-q$, respectively. The return of the security over the next time interval is $u$ when the upstate is attained and $d$ when the downstate is attained. Suppose that $S(t)$ is the price of the security at time $t$. Define that $Y_{\bar{t}} = \log S(t) - \log S(t-\tau)$. Then $Y_{\bar{t}}$ is a Bernoulli random variable with

$$\Pr(Y_{\bar{t}} = \log u) = q, \quad \Pr(Y_{\bar{t}} = \log d) = 1 - q. \tag{2.13}$$

Let $X(t) = Y_1 + Y_2 + \cdots + Y_{\bar{t}}$. Then $X(t)$ is a random walk and the price $S(t)$ can be expressed as

$$S(t) = S(0)e^{X(t)}. \tag{2.14}$$

If we further require that the logarithm of $S(t)$ have constant average mean $\mu$ and constant average variance $\sigma^2$, our first approach gives

$$u = e^{\sqrt{\sigma^2\tau+\mu^2\tau^2}}, \quad d = e^{-\sqrt{\sigma^2\tau+\mu^2\tau^2}}. \tag{2.15}$$

and

$$q = \frac{1}{2}\Big[1 + \sqrt{\frac{1}{1+\sigma^2/\mu^2\tau}}\Big]. \tag{2.16}$$

The first-order approximation on $\tau$ in (2.15) and (2.16) yields the well-known binomial model of Cox, Ross and Rubinstein[5].

If we use the second approach, then $q = \frac{1}{2}$ and

$$u = e^{\mu\tau+\sigma\sqrt{\tau}}, \quad d = e^{\mu\tau-\sigma\sqrt{\tau}}. \tag{2.17}$$

This is similar to the model proposed by Hua He[14].

## 2.3 General Lattice Models

We now consider a security market with the following conditions:

- There are $T+1$ consumption dates separated in regular intervals. Without loss of generality, we assume these dates are $t = 0, 1, \cdots, T$. Tradings take place only at $t = 0, 1, \cdots, T-1$.

- There are a finite number of states of economy

  $$\Omega = \{\omega_1, \omega_2, \cdots, \omega_J\}$$

  with the probability at state $\omega_j$ being $P(\omega_j)$.

  Hence the $\sigma$-algebra $\mathcal{F}$ of this probability space $(\Omega, \mathcal{F}, P)$ is the collection of all the subsets of $\Omega$.

- There is an information structure

  $$\{\mathcal{F}_t, \quad t = 0, 1, \cdots, T\} \tag{2.18}$$

  on $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}_0 = \{\phi, \ \Omega\}$ is the trivial $\sigma$-algebra and $\mathcal{F}_T = \mathcal{F}$. Thus, at the beginning of the period, there is no information and at the end of the period, all information is available.

- There are $N$ primitive securities with price process

  $$p(t) = (p_1(t), p_2(t), \cdots, p_N(t))', \quad t = 0, 1, \cdots, T,$$

  where $p_n(t)$ is the price of security $n$ at time $t$. The prices $(p_1(T), p_2(T), \cdots, p_N(T))$ at time $T$ is actually the terminal payoffs of those securities and sometime we denote them by payoff matrix

  $$D = \begin{pmatrix} d_1(\omega_1) & \cdots & d_N(\omega_1) \\ \vdots & & \vdots \\ d_1(\omega_J) & \cdots & d_N(\omega_J) \end{pmatrix}$$

Since at time $t$ the securities are priced based on the information available up to time $t$, the price process

$$(p(0), p(1), \cdots, p(T))$$

is a stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

We further assume that one of these securities, say, the first security, is a riskfree bond with constant interest rate $r$ over each time interval. Thus, $p_1(t) = (1+r)^t p(0)$.

- Investors are price takers. They share the same information represented by $\{\mathcal{F}_t, \ \ t = 0, 1, \cdots, T\}$ and have a homogeneous belief $P$.

- There is only one perishable consumption good.

A trading strategy $\theta(t) = (\theta_1(t), \cdots, \theta_N(t))'$ at the time $t$ is such that after trading the investor owns $\theta_n(t)$ shares of security $n$ at time $t$. A trading strategy for the period $[0, \ T]$ is then

$$\theta(0), \theta(1), \cdots, \theta(T-1).$$

Since $\theta(t)$ is determined at time $t$, it is a random vector on $(\Omega, \mathcal{F}_t, P)$. Thus, $\theta(t), \ t = 0, 1, \cdots, T-1$ is a stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

Let $e(t)$ be an endowment at time $t$ which is a random variable on $(\Omega, \mathcal{F}_t, P)$. Hence, the endowment process $e = (e(0), e(1), \cdots, e(T))$ is a stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

A stochastic process $c = (c(0), c(1), \cdots, c(T))$ is called a consumption process with respect to the endowment process $e$ and the price process $p$ if there is a trading strategy $\theta(t), \ t = -1, 0, 1, \cdots, T$ such that

$$c(t) = e(t) + \sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t), \tag{2.19}$$

for $t = 0, 1, 2, \cdots, T$, where $\theta_n(-1) = 0, \ \theta_n(T) = 0$. Similar to the one time-period case, a consumption process $c = (c(0), c(1), \cdots, c(T))$ is attainable if there is a trading strategy

$\theta(t)$ such that

$$c(t) = \sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t), \tag{2.20}$$

for $t = 1, 2, \cdots, T$. A market is complete if every consumption process is attainable.

A self-financing trading strategy is a trading strategy such that

$$\sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t) = 0, \tag{2.21}$$

for $t = 1, 2, \cdots, T-1$. In other words, under a self-financing trading strategy an investor consumes only his/her endowment, no more and no less, on any intermediate trading date. It is easy to see that a consumption process with no intermediate consumptions is attainable if and only if the corresponding trading strategy is self-financing.

In this section we examine the relation of prices between any two consecutive trading dates.

Let $p(t), \quad t = 0, 1, \cdots, T$ be the price process and $\mathcal{F}_t, \quad t = 0, 1, \cdots, T$ be the information structure we defined in Section 2.3. As we have assumed that $\Omega$ consists of only a finite number of states, it can be shown that each $\mathcal{F}_t$ is generated by a finite partition $\{F_t^1, F_t^2, \cdots, F_t^{m_t}\}$, i.e. $\cup_{i=1}^{m_t} F_t^i = \Omega, \;\; F_t^i \cap F_t^j = \phi, \;\; i \neq j$, and each $F_t^i$ is indivisible in $\mathcal{F}_t$. To see this, given any $\omega \in \Omega$, let $F_t(\omega)$ be the smallest set in $\mathcal{F}_t$ containing $\omega$. All such sets then either coincide or completely separate. Thus all different $F_t(\omega), \;\; \omega \in \Omega$, form a finite partition of $\Omega$. From the indivisibility of each $F_t^i$, every random variable defined on $(\Omega, \mathcal{F}_t, P)$ is constant on $F_t^i$.

Let now the partition $\{F_{t-1}^1, F_{t-1}^2, \cdots, F_{t-1}^{m_{t-1}}\}$ generate $\mathcal{F}_{t-1}$. It follows from $\mathcal{F}_{t-1} \subset \mathcal{F}_t$ that each $F_{t-1}^j$ is the union of a finite number of $F_t^i$'s. Without loss of generality, we may number $\omega$ is such a way that, for each $t$, there are

$$1 \leq i_1 < i_2 < \cdots < i_{m_{t-1}-1} < m_t$$

such that (Figure 2.1)

$$F_{t-1}^1 = \cup_{i=1}^{i_1} F_t^i, \;\; F_{t-1}^2 = \cup_{i=i_1+1}^{i_2} F_t^i, \;\; \cdots, \;\; F_{t-1}^{m_{t-1}} = \cup_{i=i_{m_{t-1}-1}+1}^{m_t} F_t^i. \tag{2.22}$$

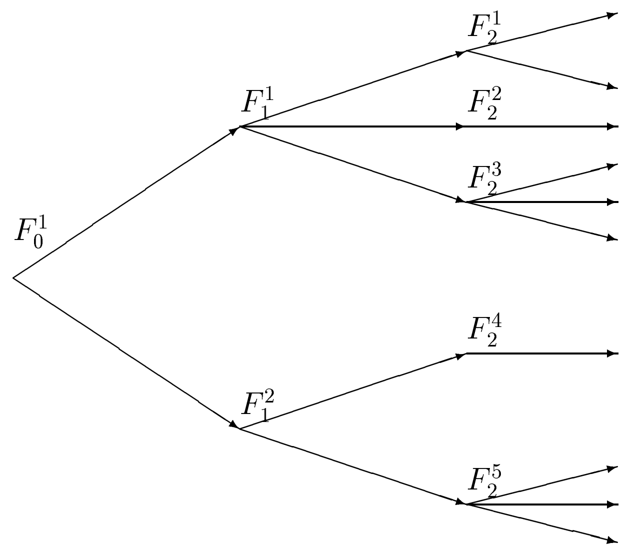


Figure 2.1: **Tree structure of a lattice model**

Thus, $i_j - i_{j-1}$ is the number of sets in $\mathcal{F}_t$ split from $F_{t-1}^j$. Summarizing the discussion above, we see that the model we consider is of a lattice or tree structure. Under this structure,

$$F_{t-1}^j, \ j = 1, 2, \cdots, m_{t-1}$$

are nodes at time $t-1$. The set $F_t^i, \ F_t^i \subset F_{t-1}^j$, is a branch coming from $F_{t-1}^j$.

Recall that $p(t-1)$ is constant on each $F_{t-1}^j$ and $p(t)$ is constant on each $F_t^i$. We may denote these constant vectors as $p(F_{t-1}^j)$ and $p(F_t^i)$, respectively.

For each $F_{t-1}^j$, we now construct a one time-period model as follows: $p(F_{t-1}^j)$ is its price system and

$$D(F_{t-1}^j) = \begin{pmatrix} p_1(F_t^{i_{j-1}+1}) & \cdots & p_N(F_t^{i_{j-1}+1}) \\ \vdots & & \vdots \\ p_1(F_t^{i_j}) & \cdots & p_N(F_t^{i_j}) \end{pmatrix} \tag{2.23}$$

is its payoff matrix.

Thus, the multiple time-period lattice model is decomposed into a collection of the associated one time-period models in which the payoffs of a price system are the prices on the following trading date.

# Chapter 3

# No-Arbitrage Valuation

## 3.1 No-Arbitrage Condition

Similar to the definition of arbitrage for one time-period models, we say a price process $p(t), \quad t = 0, 1, \cdots, T$ admits arbitrage if there exists a trading strategy $\theta(t), \quad t = 0, 1, \cdots, T-1$ such that the associated consumption process

$$c(t) = \sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t), \tag{3.1}$$

for $t = 0, 1, 2, \cdots, T$, is a nonzero, nonnegative consumption process.

We will show below that a price process does not admit arbitrage if and only if every associated one time-period model defined in Section 2.4 does not admit arbitrage.

Suppose that there is an associated one time-period model which admits arbitrage, say the one with price system $p(F_t^i)$ and payoff matrix $D(F_t^i)$. Thus, there is a trading strategy $\theta = (\theta_1, \theta_2, \cdots, \theta_N)'$ such that

$$(-\theta' p(F_t^i), \theta' D'(F_t^i))$$

is nonzero and nonnegative. Let $\theta(t) = \theta$, when the state $F_t^i$ prevails; otherwise, $\theta(t) = 0$. Then $\theta(t), \quad t = 0, 1, \cdots, T-1$ is an arbitrage strategy.

Conversely, Suppose that none of the associated one period models admits arbitrage but the price process $p(t), \;\; t = 0, 1, \cdots, T$ admits arbitrage. Let $\theta(t), \;\; t = 0, 1, \cdots, T-1$ be an arbitrage strategy. Thus,

$$c(t) = \sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t) \geq 0,$$

for $t = 0, 1, \cdots, T$ and $c(s) > 0$ at some node $F_s^i$.

Starting from the last time interval $[T-1, \; T]$, that

$$c(T) = \sum_{n=1}^{N} \theta_n(T-1) p_n(T) \geq 0$$

and every one period model does not admit arbitrage implies

$$\sum_{n=1}^{N} \theta_n(T-1) p_n(T-1) \geq 0.$$

It follows from $c(T-1) \geq 0$ that

$$\sum_{n=1}^{N} \theta_n(T-2) p_n(T-1) \geq 0.$$

Continuing backwards over time in this manner, we have

$$\sum_{n=1}^{N} \theta_n(t) p_n(t) \geq 0,$$

for $t = s, s+1, \cdots, T-1$.

Since $c(s) > 0$ at node $F_s^i$, $\sum_{n=1}^{N} \theta_n(s-1) p_n(s) > 0$ at node $F_s^i$. The same argument as above can show that for $t = 0, 1, \cdots, s-1$, at at least one node at time $t$

$$\sum_{n=1}^{N} \theta_n(t) p_n(t) > 0.$$

In particular, $-c(0) = \sum_{n=1}^{N} \theta_n(0) p_n(0) > 0$, which is contradictory to $c(0) \geq 0$.

Hence, to verify whether a multi-period model admits arbitrage, it is sufficient to verify whether its associated one period models admit arbitrage, which can easily be done as we have shown in Chapter One.

## 3.2 Risk-Neutral Probability Measures

We now always assume that the price process $p(t),\ \ t = 0, 1, \cdots, T$ does not admit arbitrage. We will show in this section that under the no-arbitrage assumption, there is a probability measure on $(\Omega, \mathcal{F})$ such that the present value processes are martingales with respect to the information structure $\mathcal{F}_t$.

Consider the one period model at each node $F^j_{t-1}$. From Theorem 1.2, there is a risk-neutral probability measure, denoted as $\bar{Q}(F^j_{t-1})$ such that

$$D'(F^j_{t-1})\bar{Q}(F^j_{t-1}) = (1+r)p(F^j_{t-1}). \tag{3.2}$$

Define

$$Q(t,\omega) = \bar{Q}(F^j_{t-1})(F^i_t), \quad \text{if} \quad \omega \in F^i_t \subset F^j_{t-1}.$$

Then, $Q(t)$ is measurable on $(\Omega, \mathcal{F}_t)$ and $Q(t)$ is a probability measure on each $F^j_{t-1}$. Let

$$Q = \prod_{t=1}^{T} Q(t). \tag{3.3}$$

We first show that $Q$ is a probability measure on $(\Omega, \mathcal{F})$. The positivity of $Q$ is obvious. Let

$$1 \le i_1 < i_2 < \cdots < i_{m_{T-1}-1} < J$$

be the partition for the last time interval. Then

$$\begin{aligned}
\sum_{j=1}^{J} Q(\omega_j) &= \sum_{j=1}^{J}\prod_{t=1}^{T} Q(t,\omega_j) \\
&= \sum_{k=1}^{m_{T-1}} \sum_{j=i_{k-1}+1}^{i_k} (\prod_{t=1}^{T-1} Q(t,\omega_j))Q(T,\omega_j) \\
&= \sum_{k=1}^{m_{T-1}} (\prod_{t=1}^{T-1} Q(t,F^k_{T-1})) \sum_{j=i_{k-1}+1}^{i_k} Q(T,\omega_j) \\
&= \sum_{k=1}^{m_{T-1}} \prod_{t=1}^{T-1} Q(t,F^k_{T-1}),
\end{aligned} \tag{3.4}$$

where $Q(s, F_t^i) = Q(s, \omega)$ for $\omega \in F_t^i$ and $s \leq t$. It is well defined since $Q(s),\ s = 1, \cdots, t$ all are $\mathcal{F}_t$-measurable. Proceeding in this manner yields

$$\sum_{j=1}^{J} Q(\omega_j) = \cdots = \sum_{k=1}^{m_0} Q(1, F_1^k) = 1. \tag{3.5}$$

Furthermore, for any $F_t^i \in \mathcal{F}_t$,

$$\begin{aligned} Q(F_t^i) &= \sum_{\omega_j \in F_t^i} Q(\omega_j) = \sum_{\omega_j \in F_t^i} \prod_{s=1}^{T} Q(s, \omega_j) \\ &= \prod_{s=1}^{t} Q(s, F_t^i) \sum_{\omega_j \in F_t^i} \prod_{s=t+1}^{T} Q(s, \omega_j). \end{aligned} \tag{3.6}$$

The second factor in (3.6) is equal to one , which can be derived exactly as was done in (3.4) and (3.5). Thus,

$$Q(F_t^i) = \prod_{s=1}^{t} Q(s, F_t^i). \tag{3.7}$$

For any pair $F_t^i \subset F_{t-1}^j$, the conditional probability

$$Q(F_t^i \mid F_{t-1}^j) = \frac{\prod_{s=1}^{t} Q(s, F_t^i)}{\prod_{s=1}^{t-1} Q(s, F_{t-1}^j)} = Q(t, F_t^i) = \bar{Q}(F_{t-1}^j)(F_t^i), \tag{3.8}$$

since $Q(s, F_t^i) = Q(s, F_{t-1}^j)$ for $s = 1, \cdots, t-1$.

We now introduce the present value process

$$a_n(t) = (1 + r)^{-t} p_n(t), \tag{3.9}$$

for security $n,\ n = 1, \cdots, N$. In fact, $a_n(t)$ is the present value at time 0, of the price of security $n$ at time $t$, discounted at the riskfree rate $r$.

Recall that a stochastic process $X(t)$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is called a martingale if

$$E(X(t) \mid \mathcal{F}_{t-1}) = X(t-1). \tag{3.10}$$

We now show that $a_n(t)$ is a martingale under the probabilty measure $Q$.

For any $F_{t-1}^j \in \mathcal{F}_{t-1}$,

$$\begin{aligned} E_Q(a_n(t) \mid F_{t-1}^j) &= (1+r)^{-t} E_Q(p_n(t) \mid F_{t-1}^j) \\ &= (1+r)^{-t} \sum_{F_t^i \subset F_{t-1}^j} p_n(F_t^i) Q(F_t^i \mid F_{t-1}^j) \\ &= (1+r)^{-t} \sum_{F_t^i \subset F_{t-1}^j} p_n(F_t^i) \bar{Q}(F_{t-1}^j)(F_t^i) \end{aligned}$$

by (3.2)

$$= (1+r)^{-t+1} p_n(F_{t-1}^j) = a_n(t-1, F_{t-1}^j).$$

Therefore,

$$E_Q(a_n(t) \mid \mathcal{F}_{t-1}) = a_n(t-1). \tag{3.11}$$

Moreover, for any $s < t$,

$$E_Q(a_n(t) \mid \mathcal{F}_s) = a_n(s). \tag{3.12}$$

This can be easily derived from $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_2)$, where $\mathcal{F}_1$ is finer than $\mathcal{F}_2$.

Hence, if the price process $p(t),\ t = 0, 1, \cdots, T$ does not admit arbitrage, there is a probability measure $Q$ such that the present value processes $a_n(t),\ t = 0, 1, \cdots, T;\ n = 1, \cdots, N$ are martingales on the filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, Q)$. It is proved below that this is also true conversely.

**Theorem 3.1** A price process $p(t),\ t = 0, 1, \cdots, T$ does not admit arbitrage if and only if there is a probability measure $Q$ such that the present value processes $a_n(t),\ t = 0, 1, \cdots, T;\ n = 1, \cdots, N$ are martingales on the filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, Q)$.

Proof: It is enough to prove the sufficient part.

Let $Q$ be such a probability measure. For each node $F_{t-1}^j$, define a probability measure $\bar{Q}(F_{t-1}^j)$ for the associated one period model as follows: for each $F_t^i \subset F_{t-1}^j$,

$$\bar{Q}(F_{t-1}^j)(F_t^i) = Q(F_t^i \mid F_{t-1}^j).$$

From

$$E_Q(a_n(t) \mid \mathcal{F}_{t-1}) = a_n(t-1),$$

we have

$$D'(F_{t-1}^j)\bar{Q}(F_{t-1}^j) = (1+r)p(F_{t-1}^j).$$

Thus, none of the one period models admits arbitrage, neither does the multi-period model.

□

The probability measure $Q$ under which the present value processes are martingales is called the risk-neutral probability measure.

So far, we have assumed that the riskfree rate is a constant and nonrandom throughtout the entire period $[0,\ T]$. More realistically, the interest rate should be assumed to depend on the information available up to a current point in time. Mathematically, this is equivalent to assuming that the interest rate process $r_t,\ t = 1, \cdots, T$, is a predictable process(a stochastic process $X(t)$ is said to be predictable if $X(t+1)$ is a stochastic process). Let us denote the discount function at time $t$ as

$$R_t^{-1} = \frac{1}{(1+r_1)(1+r_2)\cdots(1+r_t)}. \tag{3.13}$$

Then the present value processes are

$$a_n(t) = R_t^{-1} p_n(t),\ t = 1, \cdots, T, \tag{3.14}$$

for $n = 1, \cdots, N$. All the results for multi-period models we have discussed and in the following sections can be extended to hold in the case that the discount function is defined in (3.13).

## 3.3 Valuation

In this section, we assume that the market is complete. The discussion of pricing in an incomplete market will be similar to that in Section 1.4.

Since any consumption process $c = (c(t),\ t = 0, 1, 2, \cdots, T)$, is attainable, we have a trading strategy $\theta(t),\ t = 0, 1, 2, \cdots, T-1$ such that

$$c(t) = \sum_{n=1}^{N} (\theta_n(t-1) - \theta_n(t)) p_n(t), \tag{3.15}$$

for $t = 1, 2, \cdots, T$. No-arbitrage implies that the price of this consumption process is

$$\phi(c) = c(0) + \sum_{n=1}^{N} \theta_n(0) p_n(0) \tag{3.16}$$

This formula suggests that the consumption process $c = (c(t),\ t = 0, 1, 2, \cdots, T)$ could be achieved by rebalancing a portfolio with initial value $\phi(c)$ at each trading date: at time 0, consume $c(0)$ and form a portfolio with $\theta_n(0)$ shares of security $n$; at time 1, adjust the portfolio so that we own $\theta_n(1)$ shares of security $n$ and consume the rest which is exact amount of $c(1)$, and so on. This process is often called replication or dynamic hedging. When a security has no intermediate consumptions, which is the case for many derivative securities, the price of this security is the value of a portfolio which replicates the terminal payoff of the security through a self-financing trading strategy.

Another approach is to use the risk-neutral probability measure. Similar to the one period case, the price of a consumption process can be expressed as the expected present value under the risk-neutral probability measure.

Let

$$a = c(0) + \sum_{t=1}^{T} \frac{c(t)}{(1+r)^t}. \tag{3.17}$$

$a$ is the present value of the consumption process.

$$\begin{aligned} E_Q(a) &= c(0) + \sum_{t=1}^{T} \frac{E_Q(c(t))}{(1+r)^t} \\ &= c(0) + \sum_{t=1}^{T} \sum_{n=1}^{N} \frac{E_Q(\theta_n(t-1)p_n(t)) - E_Q(\theta_n(t)p_n(t))}{(1+r)^t}. \end{aligned}$$

Since

$$E_Q(\theta_n(t-1)p_n(t)) = E_Q\Big[E_Q(\theta_n(t-1)p_n(t)\ |\mathcal{F}_{t-1})\Big] = (1+r)E_Q(\theta_n(t-1)p_n(t-1)),$$

$$
\begin{aligned}
E_Q(a) &= c(0) + \sum_{t=1}^{T}\sum_{n=1}^{N}\Big[\frac{E_Q(\theta_n(t-1)p_n(t-1))}{(1+r)^{t-1}} - \frac{E_Q(\theta_n(t+1)p_n(t))}{(1+r)^t}\Big] \\
&= c(0) + \sum_{n=1}^{N}\theta_n(1)p_n(0) = \phi(c). \qquad (3.18)
\end{aligned}
$$

In particular, if we let

$$
\chi_{F_t^i}(\omega) = \begin{cases} 1, & \omega \in F_t^i \\ 0, & \text{otherwise} \end{cases} \qquad (3.19)
$$

be the Arrow-Debreu security which pays one unit when the state $F_t^i$ prevails and zero otherwise. Then, the corresponding price is

$$
\phi(\chi_{F_t^i}) = \frac{E_Q(\chi_{F_t^i})}{(1+r)^t} = \frac{Q(F_t^i)}{(1+r)^t}. \qquad (3.20)
$$

Although both approaches produce the same price for a financial security under the ideal assumptions we have used, which should be used in practice depends on the type of the security. In many cases, the risk-neutral valuation approach is simpler than the dynamic hedging approach, especially when a security is a European type derivative. This can be seen in the next section where a European call is considered. However, the dynamic hedging approach offers more flexiblities. it allows us not only to deal with non-European derivatives but also to deal with the valuation problem for models under more realistic assumptions. Many features such as dividend payments, transaction costs can be dealt with easily. The valuation of American derivatives is illustrated in the next section. Applications of the dynamic hedging approach to models which incorporates transaction costs can be found in [4] and [3].

## 3.4 Binomial Models of Option Pricing

Consider now a market with only two securities: a riskless bond and a stock. Both are traded over the period $[0,\ T]$. There are $\bar{T}$ trading dates, $\bar{t} = 0, 1, \cdots, \bar{T}-1$, separated in regular intervals. The stock price $S(t)$ follows the random walk model described in Section 2.2. Hence,

$$\begin{aligned} \Pr\{S(t) = uS(t-\tau)\ |S(t-\tau)\} &= q, \\ \Pr\{S(t) = dS(t-\tau)\ |S(t-\tau)\} &= 1-q, \quad 0 < q < 1, \end{aligned} \tag{3.21}$$

for $t = \tau, \cdots, T$. The interest rate of the riskless bond at each trading period is $r$.

It is easy to see that the associated one period models are identical with the one we presents in Example 1.2. Thus, the market is complete and the stock price does not admit arbitrage if and only if $d < 1 + r < u$. Moreover, under the unique risk-neutral probability measure, the conditional probability measure at time $t$, conditional on $t - \tau$ is

$$q_u = \frac{1+r-d}{u-d}, \quad q_d = 1 - q_u = \frac{u-1-r}{u-d}. \tag{3.22}$$

Therefore, the stock price $S(t)$ at time $t$ under the risk-neutral probability measure is binomially distributed and

$$Q(\{S(t) = S(0)u^s d^{\bar{t}-s}\}) = \begin{pmatrix} \bar{t} \\ s \end{pmatrix} q_u^s (1-q_u)^{\bar{t}-s}, \quad s = 0, 1, \cdots, \bar{t}, \tag{3.23}$$

for $t = \tau, \cdots, T$.

We now try to price a European call option written on the stock with the strike price $K$, expired at time $T$. Let

$$C = \max(S(T) - K,\ 0)$$

be the payoff of the call. Then by the result obtained in Section 3.3, the price of the call is

$$\begin{aligned} \phi_c &= (1+r)^{-\bar{T}} E_Q(C) \\ &= (1+r)^{-\bar{T}} \sum_{s=0}^{\bar{T}} \max(S(0)u^s d^{\bar{T}-s} - K,\ 0) \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_u^s (1-q_u)^{\bar{T}-s} \\ &= (1+r)^{-\bar{T}} \sum_{s \geq \frac{\log(K/S(0)) - \bar{T}\log d}{\log(u/d)}} (S(0)u^s d^{\bar{T}-s} - K) \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_u^s (1-q_u)^{\bar{T}-s} \end{aligned}$$

$$
\begin{aligned}
&= (1+r)^{-\bar{T}} S(0) \sum_{s \geq \frac{\log(K/S(0)) - \bar{T}\log d}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} u^s d^{\bar{T}-s} q_u^s (1-q_u)^{\bar{T}-s} \\
&- (1+r)^{-\bar{T}} K \sum_{s \geq \frac{\log(K/S(0)) - \bar{T}\log d}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_u^s (1-q_u)^{\bar{T}-s}.
\end{aligned}
$$

Let

$$
\bar{q}_u = \frac{u q_u}{1+r}. \tag{3.24}
$$

Then, from (3.22)

$$
\bar{q}_d = (1-\bar{q}_u) = \frac{d q_d}{1+r}. \tag{3.25}
$$

The price of the call can then be written as

$$
\begin{aligned}
\phi_c &= S(0) \sum_{s \geq \frac{\log(K/S(0)) - \bar{T}\log d}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} \bar{q}_u^s (1-\bar{q}_u)^{\bar{T}-s} \\
&- (1+r)^{-\bar{T}} K \sum_{s \geq \frac{\log(K/S(0)) - \bar{T}\log d}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_u^s (1-q_u)^{\bar{T}-s} \\
&= S(0) \sum_{s \leq \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} \bar{q}_d^s (1-\bar{q}_d)^{\bar{T}-s} \\
&- (1+r)^{-\bar{T}} K \sum_{s \leq \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_d^s (1-q_d)^{\bar{T}-s}.
\end{aligned} \tag{3.26}
$$

Denote $\bar{d} = \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}$ and $B(x;\ n, p)$ the distribution function of the binomial distribution with parameters $n$ and $p$, i.e.

$$
B(x;\ n, p) = \sum_{s \leq x} \begin{pmatrix} n \\ s \end{pmatrix} p^s (1-p)^{n-s}. \tag{3.27}
$$

We have

$$
\phi_c = S(0) B(\bar{d};\ \bar{T}, \bar{q}_d) - (1+r)^{-\bar{T}} K B(\bar{d};\ \bar{T}, q_d). \tag{3.28}
$$

This formula is the well known option pricing formula of Ross, Cox and Rubinstein[5]. The delta is $B(\bar{d};\ \bar{T}, \bar{q}_d)$. Other Greeks can be calculated easily. It also reveals that to replicate a European call, the strategy is to form a portfolio long in stock and short in bond. Note that the first summation in the formula is the distribution function of the binomial distribution with parameter $\bar{q}_d$ and the second summation is the distribution function of the binomial distribution with parameter $q_d$. Hence both can be evaluated quite easily.

To value a European put option, we can either use the above approach or use the put-call parity.

Let

$$P = \max(K - S(T),\ 0)$$

be the payoff of a European put option with the strike price $K$, expired at time $T$. It is easy to see that

$$\max(S(T) - K,\ 0) - \max(K - S(T),\ 0) = S(T) - K.$$

Hence if $\phi_p$ is the price of the put, we have

$$\phi_c - \phi_p = (1+r)^{-\bar{T}}[E_Q(S(T)) - K] = S(0) - (1+r)^{-\bar{T}}K. \tag{3.29}$$

This identity is called the put-call parity.

We now consider the valuation problem for American options. The payoff structure of an American option is similar to its European counterpart. However, an American option can be exercised at anytime before its expiration date. For example, an American call option and an American put option written on a stock with the price $S(t)$ at time $t$ for the period $[0,\ T]$ can be exercised before time $T$. Their payoffs, if exercised at $t$, will be $\max(S(t) - K,\ 0)$ and $\max(K - S(t),\ 0)$, respectively.

The valuation problem for American options is generally much more difficult than European options. Unlike European options, There are no closed form solutions for American options. This is because the buyer of an American option holds the right to exercise at anytime and the problem becomes how to find the optimal exercise time at which the expected discounted payoff for the buyer is maximized. Since a decision on whether to exercise should be based on the information up to date, an exercise time is a random variable and is described as a stopping time in the probabilistic context. A stopping time $\mathcal{T}$ on a filtered probabilty space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a random variable such that for each $t$, the event $\{\mathcal{T} \leq t\}$ belongs to $\mathcal{F}_t$.

Let $g(S(t), t)$ be the payoff of an American option when it is exercised at time $t$. If the decision to exercise this option is based on a stopping time $\mathcal{T}$, the valuation formula (3.18) gives that the price of this option is

$$E_Q\{(1+r)^{-\mathcal{T}} g(S(\mathcal{T}), \mathcal{T})\}. \tag{3.30}$$

Recalling that the buyer of an American option always wants to maximize the expected discounted payoff, the value of this option then is

$$\phi_g = \max_{\mathcal{T}} E_Q\{(1+r)^{-\mathcal{T}} g(S(\mathcal{T}), \mathcal{T})\}. \tag{3.31}$$

Maximization is taken over all stopping times over the period $[0,\ T]$. It is easy to see that there is no put-call parity for American options since the optimal exercise time for a call is different from the optimal exercise time for the corresponding put.

It is impractical to examine each of these stopping times in (3.31) in order to find the optimal exercise time and the value for the option. However, under the discrete-time framework we have disccussed in this chapter we will be able to find the optimal exercise time and the option value through a backward recursive algorithm.

We begin with the last time interval. For $t = T$, define a random variable $v(T-1, \mathcal{F}_{T-1})$ on $(\Omega, \mathcal{F}_{T-1})$ as

$$v(T-1, \mathcal{F}_{T-1}) = \max\{(1+r)^{-1} E_Q\{g(S(T), T) | \mathcal{F}_{T-1}\},\quad g(S(T-1), T-1)\}. \tag{3.32}$$

For $t = 1, \cdots, T-1$, define a random variable $v(t-1, \mathcal{F}_{t-1})$ on $(\Omega, \mathcal{F}_{t-1})$ as

$$v(t-1, \mathcal{F}_{t-1}) = \max\{(1+r)^{-1} E_Q\{v(t, \mathcal{F}_t) | \mathcal{F}_{t-1}\},\quad g(S(t-1), t-1)\}. \tag{3.33}$$

The value $v(0)$ then is the price of this American option at time 0. Furthermore, $v(t, \mathcal{F}_t)$ is the value of the option at time $t$. In other words, the value of an American option is calculated as the maximum of the expected discounted value of the same option at next trading date and the current payoff. The optimal exercise time of this option then is

$$\mathcal{T}_g = \min\{t; g(S(t), t) > v(t, \mathcal{F}_t)\}, \tag{3.34}$$

where if the set is empty, we define $\mathcal{T}_g = T$.

The rationale behind this algorithm is the following:

We choose $\mathcal{T}_0 = T$ as an initial exercise time which of course is not optimal. If at a node at time $T-1$, say $F^i_{T-1}$,

$$g(S(T-1), T-1) > (1+r)^{-1} E_Q\{g(S(T), T) | F^i_{T-1}\},$$

we define

$$\mathcal{T}_1 = \begin{cases} T-1, & \text{at } F^i_{T-1} \\ \mathcal{T}_0, & \text{otherwise.} \end{cases}$$

Thus $\mathcal{T}_1$ will yield a higher expected discounted payoff than $\mathcal{T}_0$. The same argument applies to intermediate trading dates. After we exhuast all the nodes we obtain the optimal exercise time and the value of the option.

The algorithm we discussed above is quite flexible. It can apply to other types of options. For instance, we may use it to evaluate Bermudan options which allow their buyer to exercise during a given period of time before expiration of the options. In that case, we may use the algorithm for the exercise period and use an option pricing formula for European options for the no exercise period.

Finally, we discuss the valuation of an American call option. We will see in the following that there will never be an early exercise for an American call. Thus, the value of an American call is the same as that of the corresponding European call. To see this it is sufficient to show

$$\max\{S(t-1) - K,\ 0\} \leq (1+r)^{-1} E_Q\{\max\{S(t) - K,\ 0\} | S(t-1)\}. \tag{3.35}$$

The inequality (3.35) is in fact a direct application of the Jensen's inequality which states that for any random variable $X$ and for any convex function $h(x)$, $h(E(X)) \leq E(h(X))$. Now choose $h(S) = \max(S - K,\ 0)$. We have

$$\begin{aligned} & (1+r)^{-1} E_Q\{\max\{S(t) - K,\ 0\} | S(t-1)\} \\ \geq\ & (1+r)^{-1} \max\{E_Q\{S(t)|S(t-1)\} - K,\ 0\} \\ =\ & \max\{S(t-1) - \frac{K}{1+r},\ 0\} \geq \max\{S(t-1) - K,\ 0\}. \end{aligned}$$

## 3.5 Binomial Interest Rate Models

Consider a bond market in which default-free bonds are traded during the period $[0,\ T]$. The trading times are separated in regular intervals(trading periods) of length $\tau$ as we described in Section 2.2. Denoting $\bar{t} = t/\tau$, then,

$$t = \bar{t}\tau,\ \bar{t} = 0, 1, \cdots, \bar{T}$$

are the trading times.

Bonds are uniquely determined by their time of maturity. Thus, at time $t$ there are $\bar{T} - \bar{t}$ different bonds: ones with the maturity times $t = t + \tau, \cdots, T$. Let $p(t, s)$ be the price of a bond at time $t$ which pays one unit matured at time $s$.

Define

$$f(t, s) = -\frac{1}{\tau} \log \Big[\frac{p(t, s+\tau)}{p(t, s)}\Big]. \tag{3.36}$$

$f(t, s)$ is called the forward rate at time $t$ for the time period $[s,\ s+\tau]$. It relates the bond with maturity at $s$ to the bond with maturity at $s + \tau$, because

$$p(t, s+\tau) = p(t, s) e^{-\tau f(t,s)}, \tag{3.37}$$

for $t = 0, \tau, 2\tau, \cdots, T;\ s = t, \cdots, T$.

The formula (3.37) yields

$$p(t,s) = e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} f(t,k\tau)}. \tag{3.38}$$

Hence the bond price structure is uniquely determined by this forward rate structure. To model bond prices over a certain period, it is sufficient to model the respective forward rates over the period.

We now define

$$\delta(t,s) = -\frac{1}{s-t}\log p(t,s). \tag{3.39}$$

Then,

$$p(t,s) = e^{-(s-t)\delta(t,s)}. \tag{3.40}$$

$\delta(t,s)$ is the implied constant force of interest or the continuously compounded interest rate over the period $[t,\ s]$. For each $t$, the sequence

$$\delta(t,s),\ \ s = t+\tau,\cdots,T$$

is called the yield curve at time $t$. The yield curve at time 0 is simply called the yield curve. The structure of the yield curves is called the term structure of interest rates. Comparing (3.38) with (3.40), we have

$$\delta(t,s) = \frac{1}{\bar{s}-\bar{t}}\sum_{k=\bar{t}}^{\bar{s}-1} f(t,k\tau).$$

The short term rate $r_t$ at time $t$ is

$$(1+r_t)^{-1} = e^{-\tau\delta(t,t+\tau)} = p(t,t+\tau).$$

Thus the discount function defined in (3.13) is

$$\begin{aligned} R_t^{-1} &= p(0,\tau)p(\tau,2\tau)\cdots p(t-\tau,t) \\ &= e^{-\tau\sum_{k=0}^{\bar{t}-1} f(k\tau,k\tau)}. \end{aligned} \tag{3.41}$$

It is obviously a predictable process.

The present value process for the $s$-maturity bond is

$$\begin{aligned} a(t,s) &= R_t^{-1} p(t,s) \\ &= e^{-\tau[\sum_{k=0}^{\bar{t}-1} f(k\tau,k\tau)+\sum_{k=\bar{t}}^{\bar{s}-1} f(t,k\tau)]}. \end{aligned} \tag{3.42}$$

We now try to model the forward rates. As we have pointed out that the bond price structure is uniquely determined by the forward rate structure, the simpliest way to model the forward rates is to use random walks.

Let

$$f(0,s), f(\tau,s), \cdots, f(s,s)$$

be the forward rate process. We assume that it follows a random walk:

$$f(t,s) = f(t-\tau,s) + Y(t,s), \tag{3.43}$$

where $Y(t,s)$ is a Bernoulli random variable with

$$\begin{aligned} \Pr\{Y(t,s) &= u(t,s)\} = q(t), \\ \Pr\{Y(t,s) &= d(t,s)\} = 1-q(t), \quad 0 < q(t) < 1. \end{aligned} \tag{3.44}$$

The choice of $u(t,s), d(t,s)$ and $q(t),\ t = 0, \tau, \cdots, T;\ s = t, \cdots, T$ must be such that the bonds are priced to avoid arbitrage. Equivalently, under our choice there is a risk-neutral probability measure $Q$ such that the present value processes for all bonds are martingales under $Q$.

Now,

$$\begin{aligned} a(t,s) &= e^{-\tau[\sum_{k=0}^{\bar{t}-1} f(k\tau,k\tau)+\sum_{k=\bar{t}}^{\bar{s}-1} f(t,k\tau)]} \\ &= e^{-\tau[\sum_{k=0}^{\bar{t}-1} f(k\tau,k\tau)+\sum_{k=\bar{t}}^{\bar{s}-1} f(t-\tau,k\tau)+\sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)]} \\ &= e^{-\tau[\sum_{k=0}^{\bar{t}-2} f(k\tau,k\tau)+\sum_{k=\bar{t}-1}^{\bar{s}-1} f(t-\tau,k\tau)]} e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)} \\ &= a(t-\tau,s) e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)}. \end{aligned} \tag{3.45}$$

Thus, $a(t,s), t = 0, \tau, \cdots, s;\; s = \tau, \cdots, T$ are martingales under $Q$ if and only if

$$E_Q(e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)} \; |\mathcal{B}_{t-\tau}) = 1. \tag{3.46}$$

That is equivalent to the existence of $Q(t)$ such that $0 < Q(t) < 1$, $t = \tau, \cdots, T - \tau$, and

$$Q(t)e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} u(t,k\tau)} + (1 - Q(t))e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} d(t,k\tau)} = 1, \tag{3.47}$$

for all $\bar{s} = \bar{t} + 1, \cdots, \bar{T}$.

In this model at each trading period there are only two states, which implies that there is one factor to determine the price movement. Hence, it is a one-factor model. Even so, this model is not easy to implement. One problem is the determination of its parameters since they are time dependent. Another problem is that the number of the states increases exponentially as the number of trading times increases. This sometime makes the implementation of the model very difficult in practice.

To simplify this binomial model, assume that

1. $q(t) = q$, independent of $t$. Consequently, we look for the risk-neutral probability measure $Q$ such that $Q(t) = \bar{Q}$ is also independent of $t$;

2. The average variance of the forward rate process is constant:

$$Var\Big[f(t,s) - f(t-\tau,s) \; |\mathcal{B}_{t-\tau}\Big] = \sigma^2\tau. \tag{3.48}$$

The second assumption is equivalent to

$$Var(Y(t,s)) = \sigma^2\tau.$$

Since

$$\begin{aligned} Var(Y(t,s)) &= \Big[(u(t,s) - d(t,s))(1-q)\Big]^2 q + \Big[(u(t,s) - d(t,s))q\Big]^2 (1-q) \\ &= \Big[(u(t,s) - d(t,s))\Big]^2 q(1-q) = \sigma^2\tau, \end{aligned}$$

$$u(t,s) = d(t,s) + \sigma\sqrt{\frac{\tau}{q(1-q)}}.$$

Denote

$$\psi = \sigma\sqrt{\frac{\tau}{q(1-q)}}.$$

From (3.47), we have

$$\bar{Q}e^{-\tau\sum_{k=\bar{t}}^{\bar{s}-1}u(t,k\tau)} + (1-\bar{Q})e^{(s-t)\psi}e^{-\tau\sum_{k=\bar{t}}^{\bar{s}-1}u(t,k\tau)} = 1,$$

for all $\bar{s} = \bar{t}+1, \cdots, \bar{T}$. Taking the ratio of

$$e^{-\tau\sum_{k=\bar{t}}^{\bar{s}-1}u(t,k\tau)}\left[\bar{Q} + (1-\bar{Q})e^{(s-t)\psi}\right] = 1$$

to

$$e^{-\tau\sum_{k=\bar{t}}^{\bar{s}}u(t,k\tau)}\left[\bar{Q} + (1-\bar{Q})e^{(s-t+\tau)\psi}\right] = 1$$

yields

$$\begin{aligned} u(t,s) &= \frac{1}{\tau}\log\frac{\bar{Q} + (1-\bar{Q})e^{(s-t+\tau)\psi}}{\bar{Q} + (1-\bar{Q})e^{(s-t)\psi}}, \\ d(t,s) &= \frac{1}{\tau}\log\frac{\bar{Q}e^{\tau\psi} + (1-\bar{Q})e^{(s-t)\psi}}{\bar{Q} + (1-\bar{Q})e^{(s-t)\psi}}. \end{aligned} \tag{3.49}$$

The model we just obtained is the well known Ho-Lee model[16]. This model has a very appealing feature: recombining. In a recombining binomial model, the security prices are determined only by the number of upstates and the number of downstates that have occured in the past and are independent of the order of those states occured. If a model is recombining, the number of states increases linearly instead of exponentially. That enables us to use a computer with limited memory and to greatly reduce computing time when implementing it even with a large number of trading periods.

We now verify that the Ho-Lee model is recombining.

Let us consider the case with two trading periods. The case with more trading periods can be discussed in a similar manner. Since

$$p(t,s) = e^{-\tau\sum_{k=\bar{t}}^{\bar{s}-1}f(t,k\tau)},$$

we have

$$\begin{aligned} p(t+\tau,s) &= p(t,s)e^{\tau f(t,t)-\tau\sum_{k=\bar{t}+1}^{\bar{s}-1} Y(t+\tau,k\tau)} \\ &= p(t-\tau,s)e^{\tau f(t-\tau,t-\tau)+\tau f(t,t)-\tau\sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)-\tau\sum_{k=\bar{t}+1}^{\bar{s}-1} Y(t+\tau,k\tau)} \\ &= p(t-\tau,s)e^{\tau f(t-\tau,t-\tau)+\tau f(t-\tau,t)+\tau Y(t,t)-\tau\sum_{k=\bar{t}}^{\bar{s}-1} Y(t,k\tau)-\tau\sum_{k=\bar{t}+1}^{\bar{s}-1} Y(t+\tau,k\tau)}. \end{aligned}$$

Suppose that at time $t-\tau$, the $s$-maturity bond is $p(t-\tau,s)$. The price $p(t+\tau,s)$ can be attained in two ways: (i) the upstate prevails at time $t$ and then the downstate prevails at time $t+\tau$; (ii) the downstate prevails at time $t$ and then the upstate prevails at time $t+\tau$. Recombining means that the prices obtained from both ways agree. Thus,

$$\begin{aligned} & u(t,t)-\sum_{k=\bar{t}}^{\bar{s}-1} u(t,k\tau)-\sum_{k=\bar{t}+1}^{\bar{s}-1} d(t+\tau,k\tau) \\ = \;& d(t,t)-\sum_{k=\bar{t}}^{\bar{s}-1} d(t,k\tau)-\sum_{k=\bar{t}+1}^{\bar{s}-1} u(t+\tau,k\tau). \end{aligned} \tag{3.50}$$

This equation is automatically satisfied if the difference $u(t,s)-d(t,s)$ is constant and independent of $t$ and $s$, which is the case in the Ho-Lee model as we have seen above.

Binomial models are widely used in practice to value interest rate sensitive securities since they are easy to implement. Other binomial models include the Black-Derman-Toy model[2] and the Pedersen-Shiu-Thorlacius model[22].

## 3.6 Multinomial/Multifactor Interest Rate Models

Binomial models however have some shortcomings. One may be that the yield rates are perfectly correlated. In other words, if one rate moves up, all other rates move up simultaneously and if one rate moves down all other rates move down simultaneously. Although in many real situations interest rate scenarios are of this pattern, there are situations that long term rates and short term rates are moving in an opposite direction. Another possible shortcoming is fitting a binomial model to market data. Since there are only two states in each time step, a binomial model sometimes will not be able to accurately reproduce the current yield curve and the volatility structure. Thus, there is a need to develop more general lattice models to accomodate these situations.

One approach is to extend the binomial model in the previous section to a multinonimial model.

Let $\alpha_j(t),\ j = 1, \cdots, J$ be correlated random variables such that $\alpha_j(t) = 0$ or 1; $\sum_{j=1}^{J} \alpha_j(t) \leq 1$. Denote $P_j(t) = \Pr(\alpha_j(t) = 1)$.

The basic model is as follows: for any $t = \tau, \cdots, T;\ s = t, t + \tau, \cdots, T$,

$$f(t, s) = f(t - \tau, s) + \sum_{j=0}^{J} \alpha_j(t) u_j(t, s), \tag{3.51}$$

where $\alpha_0 = 1 - \alpha_1 - \cdots - \alpha_J$. Hence, $u_j(t, s)$ is the increment of the forward rate process at time $t$ when the state $\{\alpha_j(t) = 1\}$ prevails.

Similar to (3.45), the present value process of the $s$-maturity bond is

$$\begin{aligned} a(t, s) &= e^{-\tau[\sum_{k=0}^{\bar{t}-1} f(k\tau, k\tau) + \sum_{k=\bar{t}}^{\bar{s}-1} f(t, k\tau)]} \\ &= e^{-\tau[\sum_{k=0}^{\bar{t}-1} f(k\tau, k\tau) + \sum_{k=\bar{t}}^{\bar{s}-1} f(t-\tau, k\tau) + \sum_{k=\bar{t}}^{\bar{s}-1} \sum_{j=0}^{J} \alpha_j(t) u_j(t, k\tau)]} \\ &= e^{-\tau[\sum_{k=0}^{\bar{t}-2} f(k\tau, k\tau) + \sum_{k=\bar{t}-1}^{\bar{s}-1} f(t-\tau, k\tau)]} e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} \sum_{j=0}^{J} \alpha_j(t) u_j(t, k\tau)} \\ &= a(t - \tau, s) e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} \sum_{j=0}^{J} \alpha_j(t) u_j(t, k\tau)} \\ &= a(t - \tau, s) e^{-\tau \sum_{j=0}^{J} \alpha_j(t) (\sum_{k=\bar{t}}^{\bar{s}-1} u_j(t, k\tau))}. \end{aligned} \tag{3.52}$$

The no-arbitrage condition implies that for each $t,\ t = \tau, \cdots, T$, there is a probability measure $Q_t = (Q_0(t), Q_1(t), \cdots, Q_J(t))$ such that

$$Q_t(\alpha_j(t) = 1) = Q_j(t),\ j = 0, 1, \cdots, J,$$

and

$$\sum_{j=0}^{J} Q_j(t) e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} u_j(t, k\tau)} = 1, \tag{3.53}$$

for $s = \bar{t} + 1, \cdots, \bar{T}$.

In this model, we do not require that all the values $u_j(t,s)$ be different. This is important since it allows us to incorporate several economical factors into the model. For instance, if we assume that the forward rate process includes two random shocks and they are represented by two correlated binomial processes, we obtain a two factor model. In this case,

$$f(t,s) = f(t-\tau,s) + Y_1(t,s) + Y_2(t,s). \tag{3.54}$$

Let us denote the rate changes at time $t$ for the time period $[s, s+\tau]$ in the upstate and the downstate of the first random shock and the upstate and the downstate of the second random shock as

$$Y_1(t,s) = \begin{pmatrix} V_1(t,s) \\ U_1(t,s) \end{pmatrix}, \quad Y_2(t,s) = \begin{pmatrix} V_2(t,s) \\ U_2(t,s) \end{pmatrix}, \tag{3.55}$$

respectively. The probability distribution of $(Y_1(t,s), Y_2(t,s))$ is denoted as

$$\begin{aligned} \Pr(Y_1 = V_1, Y_2 = V_2) &= q_{00}(t), \ \Pr(Y_1 = V_1, Y_2 = U_2) = q_{01}(t), \\ \Pr(Y_1 = U_1, Y_2 = V_2) &= q_{10}(t), \ \Pr(Y_1 = U_1, Y_2 = U_2) = q_{00}(t). \end{aligned}$$

We obtain the well known discrete version of the Heath-Jarrow-Morton model[15]. It is easy to see that this is a fourth-nomial model with

$$u_0 = V_1 + V_2, \ u_1 = V_1 + U_2, \ u_2 = U_1 + V_2, \ u_3 = U_1 + U_2.$$

Finally, we gives an artificial example to show how to compute a risk-neutral probability measure for a given trinomial model.

**Example 3.3** A Trinomial Model

Let $J = 2$. Assume that

1.

$$\Pr(\alpha_0(t) = 1) = q_1 q_2,$$

$$\begin{aligned}
\Pr(\alpha_1(t) = 1) &= (1 - q_1)q_2, \\
\Pr(\alpha_2(t) = 1) &= 1 - q_2, \\
& \quad 0 < q_1 < 1, \quad 0 < q_2 < 1.
\end{aligned}$$

2. The average variance of the forward rate process is constant:

$$Var\Big[f(t,s) - f(t-\tau,s) \ |\mathcal{B}_{t-\tau}\Big] = \sigma^2\tau.$$

3.

$$u_1(t,s) - u_0(t,s) = u_2(t,s) - u_1(t,s).$$

We are looking for a risk-neutral probabililty measure in the following form:

$$\begin{aligned}
Q_0(t) &= Q_1 Q_2, \\
Q_1(t) &= (1 - Q_1)Q_2, \\
Q_2(t) &= (1 - Q_2), \\
& \quad 0 < Q_1 < 1, \quad 0 < Q_2 < 1.
\end{aligned}$$

From Condition 2,

$$\begin{aligned}
\sigma^2\tau &= Var\Big[\sum_{j=0}^{2} \alpha_j(t) u_j(t,s)\Big] \\
&= [u_1(t,s) - u_0(t,s)]^2 Var(\alpha_2(t) - \alpha_0(t)) \\
&= [u_1(t,s) - u_0(t,s)]^2 q_2[1 + 3q_1 - q_2(1 + q_1)^2].
\end{aligned}$$

Define $\psi$ to be

$$\psi = \sigma\sqrt{\frac{\tau}{q_2[1 + 3q_1 - q_2(1 + q_1)^2]}}.$$

Then,

$$\begin{aligned}
u_0(t,s) &= u_1(t,s) - \psi, \\
u_2(t,s) &= u_1(t,s) + \psi.
\end{aligned}$$

(3.53) implies that

$$\begin{aligned} & Q_1 Q_2 e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} u_1(t,k\tau)+(s-t)\psi} \\ + \quad & (1-Q_1) Q_2 e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} u_1(t,k\tau)} \\ + \quad & (1-Q_2) e^{-\tau \sum_{k=\bar{t}}^{\bar{s}-1} u_1(t,k\tau)-(s-t)\psi} = 1, \end{aligned}$$

for $s > t$.

Thus,

$$u_0(t,s) = \frac{1}{\tau} \log \frac{Q_1 Q_2 e^{(s-t)\psi} + (1-Q_1) Q_2 e^{-\tau\psi} + (1-Q_2) e^{-(s-t+2\tau)\psi}}{Q_1 Q_2 e^{(s-t)\psi} + (1-Q_1) Q_2 + (1-Q_2) e^{-(s-t)\psi}}, \tag{3.56}$$

$$u_1(t,s) = \frac{1}{\tau} \log \frac{Q_1 Q_2 e^{(s-t+\tau)\psi} + (1-Q_1) Q_2 + (1-Q_2) e^{-(s-t+\tau)\psi}}{Q_1 Q_2 e^{(s-t)\psi} + (1-Q_1) Q_2 + (1-Q_2) e^{-(s-t)\psi}}, \tag{3.57}$$

$$u_2(t,s) = \frac{1}{\tau} \log \frac{Q_1 Q_2 e^{(s-t+2\tau)\psi} + (1-Q_1) Q_2 e^{\tau\psi} + (1-Q_2) e^{-(s-t)\psi}}{Q_1 Q_2 e^{(s-t)\psi} + (1-Q_1) Q_2 + (1-Q_2) e^{-(s-t)\psi}}. \tag{3.58}$$

$\square$

# Part II

# Continuous-Time Finance Models

# Chapter 4

# Stochastic Calculus

## 4.1 Characteristic Functions

In this section, we briefly recall the characteristic function of a random variable which will be used to identify the distribution of random variables we consider.

Let $X$ be a random variable on $(\Omega, \mathcal{F}, P)$. The characteristic function of $X$ is defined as follows: for any real $z$,

$$\tilde{f}_X(z) = E(e^{izX}) = \int_\Omega e^{izX} dP, \tag{4.1}$$

where $i = \sqrt{-1}$. Let $F(x) = \Pr(X \le x)$ be the distribution function of $X$ and $f(x) = F'(x)$, if it exists. Then

$$\tilde{f}_X(z) = \int_{-\infty}^{\infty} e^{izx} dF(x) = \int_{-\infty}^{\infty} e^{izx} f(x) dx. \tag{4.2}$$

Remarks. (a) The domain of a characteristic function does not have to be real numbers. It can be complex numbers as long as the corresponding expectations exist as shown in the examples below. (b) If $iz$ is replaced by $-z$ or $z$, we obtain the Laplace transform or the moment generating function respectively.

When $X_1, X_2, \cdots, X_T$ are independent, then the characteristic function of the sum

$X = X_1 + X_2 + \cdots + X_T$ is

$$\tilde{f}_X(z) = \tilde{f}_{X_1}(z)\tilde{f}_{X_2}(z)\cdots\tilde{f}_{X_T}(z). \tag{4.3}$$

It can also be shown that the distribution of a random variable is uniquely determined by its characteristic function.

**Example 4.1** Binomial Distribution

Let $X = X_1 + X_2 + \cdots + X_T$, where $X_t,\ t = 1, \cdots, T$, be iid Bernoulli random variables:

$$\Pr\{X_t = h_1\} = q, \quad \Pr\{X_t = -h_2\} = 1 - q.$$

Then

$$\tilde{f}_X(z) = \left[qe^{ih_1 z} + (1-q)e^{-ih_2 z}\right]^T. \tag{4.4}$$

□

**Example 4.2** Normal Distribution

Let $X$ be a normal random variable with mean $\mu$ and variance $\sigma^2$, i.e.

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

Then,

$$\tilde{f}_X(z) = e^{i\mu z - \frac{1}{2}\sigma^2 z^2}. \tag{4.5}$$

□

## 4.2 Wiener Processes

Recall from Section 2.2 that the price of a risky security can be expressed in terms of a random walk. In that case, the price $S(t)$ at time $t$ is

$$S(t) = S(0)e^{X_\tau(t)}, \quad 0 \le t \le T,$$

where $X_\tau(t)$ is a random walk with length of step $\tau$, average mean $\mu$, and average variance $\sigma^2$.

Imagine that trading becomes more and more frequent and eventually continuous trading is achieved. This is the case when $\tau \to 0$. Thus, if $X_\tau(t)$ approaches a continuous-time stochastic process, say $W(t)$, the price of the security will be expressed as $S(t) = S(0)e^{W(t)}$. Obviously, the limiting stochastic process $W(t)$ will inherit the properties that the random walk $X_\tau(t)$ possesses. Hence, $W(0) = 0,\ E(W(t)) = \mu t$, and $Var(W(t)) = \sigma^2 t$. Since $X_\tau(t)$ is of independent increment, so is $W(t)$. Thus, for any partition $0 < t_1 < t_2 < \cdots < t_j < t$,

$$W(t_1), W(t_2) - W(t_1), \cdots, W(t) - W(t_j)$$

are independent.

We now show that such a limiting stochastic process does exist. we will find its distribution by identifying its characteristic function.

Let $\tilde{f}_\tau(z,t)$ be the characteristic function of $X_\tau(t)$. From Example 4.1,

$$\tilde{f}_\tau(z,t) = \left[qe^{ih_1 z} + (1-q)e^{-ih_2 z}\right]^{\bar{t}},$$

where $\bar{t} = t/\tau$.

Recalling that we may choose

$$h_1 = \mu\tau + \sigma\sqrt{\tau}, \quad h_2 = -\mu\tau + \sigma\sqrt{\tau}, \quad q = \frac{1}{2},$$

we have

$$\tilde{f}_\tau(z,t) = e^{i\mu t z}\Big[\frac{e^{i\sigma\sqrt{\tau}z} + e^{-i\sigma\sqrt{\tau}z}}{2}\Big]^{t/\tau}.$$

Using the Taylor expansion, it is easy to see

$$\lim_{\tau\to 0}\tilde{f}_\tau(z,t) = e^{i\mu t z - \frac{1}{2}\sigma^2 t z^2}.$$

Since the limit is a continuous function, $X_\tau(t)$ converges weakly to a random variable $W(t)$.

Comparing it with the characteristic function of a normal random variable in Example 4.2, we see that $W(t)$ is a normal random variable with mean $\mu t$ and variance $\sigma^2 t$.

Remark: Weak convergence is defined as follows: a sequence of random variable $X_n$ converges weakly to a random variable $X$ if for any bounded continuous function $h(x)$, $\lim_{n\to\infty} E(h(X_n)) = E(h(X))$. It can be shown (see Appendix A) that weak convergence is equivalent to one of the following:

(i) The distribution function of $X_n$ converges to the distribution function of $X$ at any continuous point;

(ii) The characteristic function of $X_n$ converges to the characteristic function of $X$ as long as the characteristic function of $X$ is continuous at $z = 0$.

Summarizing what we have derived above plus the fact that any linear combination of normal random variables is still a normal random variable, we can make the following conclusions.

1. The stochastic process $W(t)$ is of independent increment. Moreover, for any partition $0 < t_1 < t_2 < \cdots < t_j < t$,

$$W(t_1), W(t_2) - W(t_1), \cdots, W(t) - W(t_j)$$

are independent normal random variables. An implication of this property is that $W(t)$ is a Markovian process whose future position depends only on the current position but not on the positions in the past. A very useful corollary is that for any function $h(w)$,

$$E(h(W(s)) \,|W(u),\ 0 \le u \le t) = E(h(W(s)) \,|W(t)).$$

2. For any $s > t$,

$$E(W(s) - W(t)) = \mu(s-t), \tag{4.6}$$

$$Var(W(s) - W(t)) = \sigma^2(s-t). \tag{4.7}$$

The stochastic process $W(t)$ is called a Wiener process( or Brownian motion) with drift $\mu$ and infinitesimal variance $\sigma^2$. Especially, a Wiener process with drift $\mu = 0$ and infinitesimal variance $\sigma^2 = 1$ is called a standard Wiener process.

The next property characterises Wiener processes.

3. A stochastic process $W(t)$ is a Wiener process if and only if for any real $\lambda$, the stochastic process

$$Z_\lambda(t) = e^{\lambda W(t) - \lambda\mu t - \frac{1}{2}\lambda^2\sigma^2 t} \tag{4.8}$$

is a martingale( with respect to the Borel filtration generated by $W(t)$).

The necessary part follows from

$$E(e^{\lambda[W(s)-W(t)]} \,|\mathcal{B}_t) = e^{\lambda\mu(s-t)+\frac{1}{2}\lambda^2\sigma^2(s-t)},$$

which can be obtained by letting $z = -i\lambda$ in Example 4.2.

For the sufficient part, we proceed as follows.

$$E(Z_\lambda(t)) = E(Z_\lambda(0)) = 1.$$

Thus, $E(e^{\lambda W(t)}) = e^{\lambda\mu t + \frac{1}{2}\lambda^2\sigma^2 t}$. $W(t)$ is normal with mean $\mu t$ and variance $\sigma^2 t$. For any $s > t$ and real numbers $\lambda_1,\ \lambda_2$,

$$\begin{aligned}
& E(e^{\lambda_1[W(s)-W(t)]+\lambda_2 W(t)}) \\
= \ & E\Big\{E(e^{\lambda_1[W(s)-W(t)]+\lambda_2 W(t)}) \ |\mathcal{B}_t\Big\} \\
= \ & e^{\lambda_1\mu(s-t)+\frac{1}{2}\lambda_1^2\sigma^2(s-t)} E(e^{\lambda_2 W(t)}) \\
= \ & e^{\lambda_1\mu(s-t)+\frac{1}{2}\lambda_1^2\sigma^2(s-t)+\lambda_2\mu t+\frac{1}{2}\lambda_2^2\sigma^2 t}.
\end{aligned} \tag{4.9}$$

Differentiating (4.9) with respect to $\lambda_1,\ \lambda_2$ at $\lambda_1 = 0,\ \lambda_2 = 0$ yields

$$E([W(s) - W(t)]W(t)) = E(W(s) - W(t))E(W(t)).$$

Hence $W(s)-W(t)$ and $W(t)$ are independent since two normal random variables are independent if and only if their covariance is zero. Thus, $W(t)$ is a Wiener process.

Many useful martingales related to $W(t)$ can then be derived from $Z_\lambda(t)$. Noting the fact that the derivative of a martingale is still a martingale, if it exists,

$$W(t) - \mu t = \frac{\partial Z_\lambda(t)}{\partial \lambda}|_{\lambda=0}$$

is a martingale.

$$\left(W(t) - \mu t\right)^2 - \sigma^2 t = \frac{\partial^2 Z_\lambda(t)}{\partial \lambda^2}|_{\lambda=0}$$

is also a martingale.

Finally, we state without a proof that

4. All the paths (with probability one) of $W(t)$ are continuous.

It is easy to see that for any Wiener process $W(t)$, $\frac{1}{\sigma}(W(t) - \mu t)$ is a standard Wiener process. Thus, from now on we always denote a standard Wiener process as $W(t)$ and alternative Wiener processes are written in the form $\mu t + \sigma W(t)$.

Sometime we need to deal with a Wiener process starting at a point, say $x$, away from zero. A standard Wiener process in this case is $x + W(t)$. We call it the standard Wiener process starting at $x$ and denote it as $W(t),\ W(0) = x$.

As an application of Wiener processes, we consider a market with a riskfree bond and a risky security over the period $[0,\ T]$. Denote $\delta$ as the force of interest or the continuously compounded interest rate for the riskfree bond. We assume that the price of the risky security at time $t$ is

$$S(t) = S(0)e^{\mu t + \sigma W(t)}, \quad 0 \le t \le T. \tag{4.10}$$

We call the exponential of a Wiener process a geometric Wiener process, Geometric Brownian motion or Lognormal process.

We now wish to price a European call option maturing at time $T$ on the risky security with strike price $K$. As we mentioned earlier, $S(t)$ is actually the limit of the price process of the binomial model we considered in Section 3.2. Let $\phi_c$ be the price under the above continuous-time model and $\phi_\tau$ be the price the call under the binomial model, respectively. Then $\phi_c = \lim_{\tau\to 0}\phi_\tau$.

Recall from Section 3.4 that

$$\begin{aligned}\phi_\tau &= S(0) \sum_{s \le \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} \bar{q}_d^s (1-\bar{q}_d)^{\bar{T}-s} \\ &- (1+r)^{-\bar{T}} K \sum_{s \le \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_d^s (1-q_d)^{\bar{T}-s},\end{aligned}$$

with $u = e^{\mu\tau+\sigma\sqrt{\tau}}$, $d = e^{\mu\tau-\sigma\sqrt{\tau}}$. Let

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}dy,\ n(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \tag{4.11}$$

be the distribution function and the densilty function of the standard normal random variable. By the Central Limit Theorem(Appendix A),

$$\lim_{\tau\to 0} \sum_{s \le \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} \bar{q}_d^s (1-\bar{q}_d)^{\bar{T}-s} = N(d_1)$$

where

$$d_1 = \lim_{\tau\to 0} \frac{\frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)} - \bar{T}\bar{q}_d}{\sqrt{\bar{T}\bar{q}_d(1-\bar{q}_d)}},$$

and

$$\lim_{\tau\to 0} \sum_{s \le \frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)}} \begin{pmatrix} \bar{T} \\ s \end{pmatrix} q_d^s (1-q_d)^{\bar{T}-s} = N(d_2),$$

where

$$d_2 = \lim_{\tau\to 0} \frac{\frac{\bar{T}\log u + \log(S(0)/K)}{\log(u/d)} - \bar{T}q_d}{\sqrt{\bar{T}q_d(1-q_d)}}.$$

Since $1 + r = e^{\delta\tau}$, (3.22 ), (3.24) and (3.25) yield

$$\begin{aligned}
q_d &= \frac{e^{\sigma\sqrt{\tau}} - e^{(\delta-\mu)\tau}}{e^{\sigma\sqrt{\tau}} - e^{-\sigma\sqrt{\tau}}} = \frac{\sigma\sqrt{\tau} + (\mu - \delta + \sigma^2/2)\tau + O(\tau^{3/2})}{2\sigma\sqrt{\tau} + O(\tau^{3/2})}, \\
&= \frac{\sigma\sqrt{\tau} + (\mu - \delta + \sigma^2/2)\tau + O(\tau^{3/2})}{2\sigma\sqrt{\tau}}, \\
1 - q_d &= \frac{e^{(\delta-\mu)\tau} - e^{-\sigma\sqrt{\tau}}}{e^{\sigma\sqrt{\tau}} - e^{-\sigma\sqrt{\tau}}} = \frac{\sigma\sqrt{\tau} - (\mu - \delta + \sigma^2/2)\tau + O(\tau^{3/2})}{2\sigma\sqrt{\tau}}, \\
\bar{q}_d &= \frac{e^{(\mu-\delta)\tau} - e^{-\sigma\sqrt{\tau}}}{e^{\sigma\sqrt{\tau}} - e^{-\sigma\sqrt{\tau}}} = \frac{\sigma\sqrt{\tau} + (\mu - \delta - \sigma^2/2)\tau + O(\tau^{3/2})}{2\sigma\sqrt{\tau}}, \\
1 - \bar{q}_d &= \frac{e^{\sigma\sqrt{\tau}} - e^{(\mu-\delta)\tau}}{e^{\sigma\sqrt{\tau}} - e^{-\sigma\sqrt{\tau}}} = \frac{\sigma\sqrt{\tau} - (\mu - \delta - \sigma^2/2)\tau + O(\tau^{3/2})}{2\sigma\sqrt{\tau}}.
\end{aligned}$$

Thus,

$$\begin{aligned}
d_1 &= \lim_{\tau\to 0} \frac{\frac{(\mu T + \log(S(0)/K))\sqrt{\tau} + \sigma T}{2\sigma\tau} - \frac{\sigma T + (\mu - \delta - \sigma^2/2)T\sqrt{\tau} + O(\tau^2)}{2\sigma\tau}}{\sqrt{\frac{(\sigma^2\tau + O(\tau^2))T}{2\sigma\tau}}} \\
&= \frac{\log(S(0)/K) + (\delta + \sigma^2/2)T}{\sigma\sqrt{T}}. \qquad (4.12)
\end{aligned}$$

Similarly,

$$\begin{aligned}
d_2 &= \lim_{\tau\to 0} \frac{\frac{(\mu T + \log(S(0)/K))\sqrt{\tau} + \sigma T}{2\sigma\tau} - \frac{\sigma T + (\mu - \delta + \sigma^2/2)T\sqrt{\tau} + O(\tau^2)}{2\sigma\tau}}{\sqrt{\frac{(\sigma^2\tau + O(\tau^2))T}{2\sigma\tau}}} \\
&= \frac{\log(S(0)/K) + (\delta - \sigma^2/2)T}{\sigma\sqrt{T}} \qquad (4.13)
\end{aligned}$$

Therefore, the price of the European call option

$$\phi_c = S(0)N(d_1) - e^{-\delta T}KN(d_2), \qquad (4.14)$$

where $d_1$ and $d_2$ are given in (4.12) and (4.13). This formula is the well known Black-Scholes option pricing formula.

## 4.3 Reflection Principle

**Theorem 4.1**(Reflection Principle) Let $W(t)$ be a standard Wiener process and $a,\ h$ be two nonnegative real numbers. Then,

$$\Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\ge a+h\Big\} = \Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\le a-h\Big\}. \tag{4.15}$$

**Corollary 4.2** Let $[a+h_1,\ a+h_2],\ h_2\ge h_1\ge 0$, be any interval above the horizontal line $x=a$. Its symmetric interval about $x=a$ is $[a-h_2,\ a-h_1]$. Then we have

$$\Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\in[a+h_1,\ a+h_2]\Big\} = \Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\in[a-h_2,\ a-h_1]\Big\}. \tag{4.16}$$

*Proof:* Obviously,

$$\begin{aligned} & \Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\in[a+h_1,\ a+h_2]\Big\} \\ = \ & \Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\ge a+h_1\Big\} - \Pr\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\ge a+h_2\Big\}. \end{aligned}$$

The right hand side can be written similarly. Applying Theorem 4.1 immediately obtains the identity. ▭

*Interpretation:* $\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\in[a+h_1,\ a+h_2]\Big\}$ is the set of all paths which hit the horizontal line(barrier) $x=a$ before reaching the interval $[a+h_1,\ a+h_2]$. Similarly, $\Big\{\max_{0<s\le t} W(s)\ge a, W(t)\in[a-h_2,\ a-h_1]\Big\}$ is the set of all paths which hit the barrier $x=a$ before reaching the interval $[a-h_2,\ a-h_1]$. Since $h_1,\ h_2$ are arbitrary, the reflection principle roughly says that if a path hits the barrier $x=a$ at some time $s<t$, there is another path which is identical to the first path before and at time $s$ and is the mirror image of the first path about the barrier $x=a$ after time $s$. With the symmetric property of the standard Wiener process, the reflection principle also implies that we may

replace the path hitting the barrier $x = a$ at some time $s < t$ by a path which is the the mirror image of the first path about the barrier $x = a$ before time $s$ and identical to the first path after time $s$.

A very important application of the reflection principle is to compute various barrier hitting probabilities of the standard Wiener process. We will see in a later chapter these hitting probabilities are very useful in valuation of barrier options.

In this section, we consider three cases.

*First Passage Time*

Define the following random variable

$$\tau_a = \inf\{t > 0;\ W(t) = a\}. \tag{4.17}$$

Then $\tau_a$ is the time $W(t)$ first hits the barrier $x = a$. We call it the first passage time of $W(t)$.

To find the distribution of $\tau_a$, consider the event $\{\tau_a \leq t\}$ for any $t$. Since

$$\{\tau_a \leq t\} = \{\max_{0<s\leq t} W(s) \geq a\},$$

we have

$$\begin{aligned}
\Pr\{\tau_a \leq t\} &= \Pr\{\max_{0<s\leq t} W(s) \geq a\} \\
&= \Pr\{\max_{0<s\leq t} W(s) \geq a, W(t) \geq a\} + \Pr\{\max_{0<s\leq t} W(s) \geq a, W(t) \leq a\} \\
&\quad (\textit{Since}\ \Pr\{W(t) = a\} = 0) \\
&= 2\Pr\{\max_{0<s\leq t} W(s) \geq a, W(t) \geq a\} \quad (\textit{Reflection Principle}) \\
&= 2\Pr\{W(t) \geq a\} \quad (\textit{Since}\ \{W(t) \geq a\} \subset \{\max_{0<s\leq t} W(s) \geq a\}) \\
&= \sqrt{\frac{2}{\pi t}} \int_a^\infty e^{-\frac{1}{2t}x^2} dx = \sqrt{\frac{2}{\pi}} \int_{a/\sqrt{t}}^\infty e^{-\frac{1}{2}x^2} dx.
\end{aligned}$$

If $a < 0$, we have

$$\Pr\{\tau_a \le t\} = \sqrt{\frac{2}{\pi}} \int_{-a/\sqrt{t}}^{\infty} e^{-\frac{1}{2}x^2} dx.$$

This can be obtained by considering $-W(t)$ instead of $W(t)$.

It is easy to see from the above that $\tau_a$ is finite since

$$\Pr\{\tau_a < \infty\} = \lim_{t\to\infty} \Pr\{\tau_a \le t\} = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} e^{-\frac{1}{2}x^2} dx = 1.$$

Thus, for any horizontal line, every path of a standard Wiener process will hit the line sooner or later.

Now, let $f_a(t)$ be the density function of $\tau_a$. We have

$$f_a(t) = \frac{d}{dt} \Pr\{\tau_a \le t\} = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}, \quad t > 0. \tag{4.18}$$

This distribution is called the one-sided stable distribution of index $\frac{1}{2}$, which can be obtained as a limit of Inverse Guassian distributions[9]. As we will see later on, this result can be used to find the first passage time of a goemetric Wiener process.

*Single Barrier*

Let $g_a(x)$, $a > 0$ be the density function of $W(T)$, $W(t) = a$, for some $0 < t \le T$. Thus $g_a(x)dx$ is the probability that a path hits the barrier $x = a$ and then reaches point $x$ at time $T$. It is easy to see that $g_a(x)$ is a defective density function( a density function is called defective if its integral is less than one). We will show below

$$g_a(x) = \begin{cases} \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-2a)^2}{2T}}, & x < a \\ \frac{1}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}}, & x \ge a \end{cases} \tag{4.19}$$

Obviously, for $x \ge a$, any path to reach $x$ at $T$ will hit the barrier $x = a$ before or at $T$. Hence,

$$g_a(x) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}}, \; x \ge a.$$

To see the rest, we use the reflection principle. The probability of a path that hits the barrier $a$ and then reaches $x$ at $T$ is equal to the probability of a path that starts at $2a$, hits the barrier $a$, and then reaches $x$ at $T$. Since the latter is just the probability that a path starts at $2a$ and then reaches $x$ at $T$. Hence it is equal to

$$\frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-2a)^2}{2T}}.$$

We derived the density.

*Double Barriers*

We now consider the case there are two barriers: one upper barrier and one lower barrier. We will derive the distribution of paths which never hit these barriers before time $T$.

Let $f(x;\ a_l, a_u),\ a_l < 0 < a_u$ be the density function of $W(T),\ a_l < W(t) < a_u$ for all $0 < t \leq T$. Then

$$f(x;\ a_l, a_u) = \begin{cases} \frac{1}{\sqrt{2\pi T}} \sum_{n=-\infty}^{n=\infty} \{ e^{-\frac{[x+2n(a_u-a_l)]^2}{2T}} - e^{-\frac{[x-2a_u+2n(a_u-a_l)]^2}{2T}} \}, & a_l < x < a_u \\ 0, & x \leq a_l \text{ or } x \geq a_u. \end{cases} \tag{4.20}$$

The reflection principle will repeatly be used in the derivation.

Let $g(x;\ a_l, a_u)$ be the density function of $W(T),\ W(t) = a_u$ or $a_l$ for some $0 < t \leq T$. Thus

$$f(x;\ a_l, a_u) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}} - g(x;\ a_l, a_u),\ a_l < x < a_u$$

Let

$A_n = \{$there exist $0 < t_1 < t_2 < \cdots < t_n \leq T$ such that $W(t_{2k+1}) = a_u, W(t_{2k}) = a_l$ and $W(T) \in dx\}$,

$B_n = \{$there exist $0 < t_1 < t_2 < \cdots < t_n \leq T$ such that $W(t_{2k}) = a_u, W(t_{2k+1}) = a_l$ and $W(T) \in dx\}$,

Then, $A_{n-1} \cap B_{n-1} = A_n \cup B_n$. Recalling that

$$\Pr(A_{n-1} \cup B_{n-1}) = \Pr(A_{n-1}) + \Pr(B_{n-1}) - \Pr(A_{n-1} \cap B_{n-1}),$$

we have

$$\begin{aligned} g(x;\ a_l, a_u) &= \Pr(A_1 \cup B_1) = \Pr(A_1) + \Pr(B_1) - \Pr(A_1 \cap B_1) \\ &= \Pr(A_1) + \Pr(B_1) - \Pr(A_2 \cup B_2) \\ &\cdots \\ &= \sum_{n=1}^{\infty} (-1)^{n-1} [\Pr(A_n) + \Pr(B_n)]. \end{aligned}$$

The problem then becomes how to compute $\Pr(A_n)$ and $\Pr(B_n)$. It is easy to see

$$\Pr(A_1) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-2a_u)^2}{2T}} dx.$$

By applying the reflection principle twice, we have

$$\Pr(A_2) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{[x+2(a_u-a_l)]^2}{2T}} dx.$$

In general, we have

$$\Pr(A_{2n+1}) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{[x-2a_u-2n(a_u-a_l)]^2}{2T}} dx,$$

and

$$\Pr(A_{2n}) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{[x+2n(a_u-a_l)]^2}{2T}} dx.$$

Exchanging $a_u$ and $a_l$ in the above, we have

$$\Pr(B_{2n+1}) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{[x-2a_u-2(n+1)(a_u-a_l)]^2}{2T}} dx,$$

and

$$\Pr(B_{2n}) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{[x-2n(a_u-a_l)]^2}{2T}} dx.$$

With some tedious algebra, we obtain the density function (4.20).

## 4.4 Stochastic(Ito) Integral

In this section, we will deal with the integration of a stochastic process with respect to a standard Wiener process. For simplicity, we always assume that stochastic processes considered are continuous, namely, all of their paths are continuous.

Let $X(t)$ be a continuous stochastic process on $(\Omega, \mathcal{F}, \mathcal{B}_t, P)$, $0 \le t \le T$, where $\mathcal{B}_t$ is the Borel filtration generated by the standard Wiener process.

Define the integral of $X(t)$ on $[a,\ b]$, $0 \le a < b \le T$, as

$$\int_a^b X(t)dW(t) = \lim_{\max|t_j - t_{j-1}| \to 0} \sum_{j=1}^{J} X(t_{j-1})[W(t_j) - W(t_{j-1})], \tag{4.21}$$

where $a = t_0, t_1 < \cdots < t_{J-1} < t_J = b$ is a partition on $[a,\ b]$ and the limit is taken in the sense of uniform convergence in probability. It can be shown that the above limit always exists and is independent of the choice of partitions. We call this limit the Ito integral of $X(t)$ on $[a,\ b]$.

Remark: The definition of an Ito integral only requires that the process $X(t)$ satisfy

$$\int_0^T X^2(t)dt < \infty,$$

which is automatically satisfied by continuous stochastic processes.

Many properties of the usual Riemannian integration are carried over to Ito integration. For instance,

1. $$\int_a^b [X_1(t) + X_2(t)]dW(t) = \int_a^b X_1(t)dW(t) + \int_a^b X_2(t)dW(t);$$

2. $$\int_a^c X(t)dW(t) = \int_a^b X(t)dW(t) + \int_b^c X(t)dW(t);$$

However, there are fundamental differences between these integrations. Firstly, $\int_a^b X(t)dW(t)$ is a random variable on $(\Omega, \mathcal{B}_b, P)$ and if

$$E\Big(\int_a^b X^2(t)dt\Big) < \infty, \tag{4.22}$$

then

$$E\Big(\int_a^b X(t)dW(t)\Big) = 0, \tag{4.23}$$

and

$$Var\Big\{\Big(\int_a^b X(t)dW(t)\Big)^2\Big\} = E\Big\{\Big(\int_a^b X(t)dW(t)\Big)^2\Big\} = E\Big(\int_a^b X^2(t)dt\Big). \tag{4.24}$$

The first implication is obvious from the definition. The second implication follows from

$$E\Big(X(t_{j-1})[W(t_j) - W(t_{j-1})]\Big) = E\Big(E\Big(X(t_{j-1})[W(t_j) - W(t_{j-1})]\Big)|\ W(t_{j-1})\Big) = 0,$$

and

$$\begin{aligned} & E\Big(X(t_{j-1})X(t_{i-1})[W(t_j) - W(t_{j-1})][W(t_i) - W(t_{i-1})]\Big) \\ &= \begin{cases} E\Big(X^2(t_{j-1})[W(t_j) - W(t_{j-1})]^2\Big), & i = j \\ 0, & i < j \end{cases} \\ &= \begin{cases} E\Big(X^2(t_{j-1})(t_j - t_{j-1})\Big), & i = j \\ 0, & i < j. \end{cases} \end{aligned}$$

Second and more importantly, the point at which the value of $X(t)$ is taken in each subinterval $[t_{j-1},\ t_j]$ is critical. Under Ito integration, it is always the value at the left endpoint. Unlike the usual Riemannian integration, different choice of points will lead to different stochastic integration. For example, if we choose the midpoint of each subinterval, the limit obtained will be a Stratonovich integral. It can also be seen from the following example.

**Example 4.3** Consider

$$\begin{aligned} \int_a^b W(t)dW(t) &= \lim_{\max|t_j - t_{j-1}|\to 0} \sum_{j=1}^{J} W(t_{j-1})[W(t_j) - W(t_{j-1})] \\ &= \frac{1}{2} \lim_{\max|t_j - t_{j-1}|\to 0} \sum_{j=1}^{J} \{W^2(t_j) - W^2(t_{j-1}) - [W(t_j) - W(t_{j-1})]^2\} \\ &= \frac{1}{2}[W^2(b) - W^2(a)] - \frac{1}{2} \lim_{\max|t_j - t_{j-1}|\to 0} \sum_{j=1}^{J} [W(t_j) - W(t_{j-1})]^2. \end{aligned}$$

The summation in the second term is the sum of squares of independent normal random variables. Moreover,

$$E\Big(\sum_{j=1}^{J}[W(t_j)-W(t_{j-1})]^2\Big)=\sum_{j=1}^{J}(t_j-t_{j-1})=b-a,$$

and

$$Var\Big(\sum_{j=1}^{J}[W(t_j)-W(t_{j-1})]^2\Big)=2\sum_{j=1}^{J}(t_j-t_{j-1})^2\to 0,\quad \text{as}\quad \max|t_j-t_{j-1}|\to 0.$$

Thus,

$$\lim_{\max|t_j-t_{j-1}|\to 0}\sum_{j=1}^{J}[W(t_j)-W(t_{j-1})]^2=b-a.$$

We then have

$$\int_a^b W(t)dW(t)=\frac{W^2(b)-W^2(a)-(b-a)}{2}.$$

Now if we use the right endpoint of each subinterval instead of the left endpoint, the similar argument will give

$$\begin{aligned}
\text{(R)}\int_a^b W(t)dW(t) &= \lim_{\max|t_j-t_{j-1}|\to 0}\sum_{j=1}^{J}W(t_j)[W(t_j)-W(t_{j-1})]\\
&= \frac{1}{2}\lim_{\max|t_j-t_{j-1}|\to 0}\sum_{j=1}^{J}\{W^2(t_j)-W^2(t_{j-1})+[W(t_j)-W(t_{j-1})]^2\}\\
&= \frac{W^2(b)-W^2(a)+(b-a)}{2},
\end{aligned}$$

which is different from what we have from Ito integration. □

## 4.5 Stochastic Differential Equations and Ito's Lemma

Let $\alpha(t,x)$ and $\sigma(t,x)$ be two continuous functions in their domain. If there exists a continuous stochastic process $X(t),\ 0\le t\le T$ on $(\Omega,\mathcal{F},\mathcal{B}_t,P)$ such that

$$X(t)=X(0)+\int_0^t\alpha(s,X(s))ds+\int_0^t\sigma(s,X(s))dW(s),\ 0\le t\le T, \tag{4.25}$$

we say that $X(t)$ is a solution of the stochastic differential equation(SDE)

$$dX(t) = \alpha(t, X(t))dt + \sigma(t, X(t))dW(t), \tag{4.26}$$

with initial condition $X(0)$, or simply

$$dX = \alpha(t, X)dt + \sigma(t, X)dW. \tag{4.27}$$

The function $\alpha(t, X)$ and $\sigma(t, X)$ are often referred to as the drift and the infinitesimal deviation of the SDE. In finance, $\sigma(t, X(t))$ is also called the volatility of the stochastic process $X(t)$. The solution $X(t)$ is also called an Ito process.

It can be proved that if there is $L > 0$, independent of $t$ and $x$, such that

1. (Lipschitz condition) for any $x_1$ and $x_2$,

$$|\alpha(t, x_1) - \alpha(t, x_2)| + |\sigma(t, x_1) - \sigma(t, x_2)| \leq L|x_1 - x_2|; \tag{4.28}$$

2. (Linear Growth condition) for any $x$,

$$|\alpha(t, x)| + |\sigma(t, x)| \leq L(1 + |x|), \tag{4.29}$$

there exists an unique solution on $[0,\ T]$ for the SDE (4.24) with initial condition $X(0)$.

**Example 4.4** Consider

$$dX = \alpha dt + \sigma dW, \tag{4.30}$$

where $\alpha$ and $\sigma$ are constants. Then its solution is

$$X(t) = X(0) + \alpha t + \sigma W(t). \tag{4.31}$$

□

We now state without a proof two of the most important results in stochastic calculus. Not only are they the building blocks for many other important results in stochastic calculus but also they provide a methodology in solving stochastic differential equations.

The first result is called Ito's Lemma. It is the stochastic version of the chain rule.

**Theorem 4.3**(Ito's Lemma)

Let $X(t)$ be a solution of the stochastic differential equation (4.25) and $g(t, x)$ be a function which is continuously differentiable in $t$ and continuously twice differentiable in $x$. Then $g(t, X(t))$ is a solution of the following SDE

$$\begin{aligned} dg(t, X) &= [\frac{\partial}{\partial t} g(t, X) + \alpha(t, X) \frac{\partial}{\partial x} g(t, X) + \frac{1}{2} \sigma^2(t, X) \frac{\partial^2}{\partial x^2} g(t, X)] dt \\ &+ \sigma(t, X) \frac{\partial}{\partial x} g(t, X) dW. \end{aligned} \tag{4.32}$$

**Example 4.5** Consider the following SDE

$$dX = \alpha X dt + \sigma X dW, \tag{4.33}$$

where $\alpha$ and $\sigma$ are constants.

Let $g(t, x) = \log x$. Ito's Lemma yields

$$d \log X = (\alpha - \frac{1}{2} \sigma^2) dt + \sigma dW.$$

Thus,

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

Therefore, the solution of the equation is

$$X(t) = X(0) e^{(\alpha - \frac{1}{2} \sigma^2) t + \sigma W(t)}, \tag{4.34}$$

a geometric Wiener process. $\square$

**Example 4.6** Consider the following SDE

$$dX = (\alpha X + \beta) dt + \gamma dW, \tag{4.35}$$

where $\alpha$, $\beta$ and $\gamma$ are constants.

For the special case $\beta = \gamma = 0$, it becomes an ordinary differential equation(ODE) and one of its solutions is $e^{\alpha t}$. Thus we are seeking a solution of (4.35) in the form of

$$X(t) = e^{\alpha t} Y(t),$$

for some stochastic process $Y(t)$.

Since $Y(t) = e^{-\alpha t} X(t)$, we let $g(t,x) = e^{-\alpha t} x$. Applying Ito's Lemma, we obtain

$$dY = \beta e^{-\alpha t} dt + \gamma e^{-\alpha t} dW.$$

Thus,

$$Y(t) = Y(0) + \beta \int_0^t e^{-\alpha s} ds + \gamma \int_0^t e^{-\alpha s} dW(s).$$

Therefore,

$$X(t) = X(0) e^{\alpha t} + \beta \int_0^t e^{\alpha(t-s)} ds + \gamma \int_0^t e^{\alpha(t-s)} dW(s). \tag{4.36}$$

It is easy to see that for each $t$, $X(t)$ is a normal random variable with mean $X(0)e^{\alpha t} + \beta \int_0^t e^{\alpha(t-s)} ds$ and variance $\gamma^2 \int_0^t e^{2\alpha(t-s)} ds$. $\square$

Next result is the stochastic version of integration by parts.

**Theorem 4.4** Let $X_1(t)$ and $X_2(t)$ be solutions of the following SDEs:

$$dX_1 = \alpha_1(t, X_1) dt + \sigma_1(t, X_1) dW, \tag{4.37}$$

and

$$dX_2 = \alpha_2(t, X_2) dt + \sigma_2(t, X_2) dW, \tag{4.38}$$

respectively.

Then,

$$\begin{aligned} dX_1X_2 &= X_2 dX_1 + X_1 dX_2 + \sigma_1(t, X_1)\sigma_2(t, X_2) dt & (4.39) \\ &= [\alpha_1(t, X_1) X_2 + \alpha_2(t, X_2) X_1 + \sigma_1(t, X_1)\sigma_2(t, X_2)] dt \\ &+ [\sigma_1(t, X_1) X_2 + \sigma_2(t, X_2) X_1] dW. & (4.40) \end{aligned}$$

Ito's Lemma and integration by parts are very powerful tools, especially when they are used together as we show in the next example and in the next section.

**Example 4.7** Consider the following SDE

$$dX = (\alpha X + \beta)dt + (\gamma + \sigma X)dW, \quad \sigma \neq 0. \tag{4.41}$$

where $\alpha$, $\beta$, $\gamma$ and $\sigma$ are constants. We will show how to apply the integration by parts to solve this equation.

First we assume $\gamma = 0$.

When $\beta = 0$, we obtain the SDE in Example 4.5, whose solution, denoted by $\Phi(t)$, is a geometric Wiener process

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

Similar to Example 4.6, we are seeking a solution in the form of $X(t) = \Phi(t)Y(t)$, for some process $Y(t)$.

Since $Y(t) = \Phi^{-1}(t)X(t)$, we may apply integration by parts as long as we can express $\Phi^{-1}(t)$ as a solution of a SDE. From

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

It satisfies

$$d\Phi^{-1} = (-\alpha + \sigma^2)\Phi^{-1}dt - \sigma\Phi^{-1}dW,$$

(simply compare it with the SDE in Example 4.5).

Hence Theorem 4.4 yields

$$dY = \beta\Phi^{-1}dt.$$

$$Y(t) = Y(0) + \beta\int_0^t \Phi^{-1}(s)ds = Y(0) + \beta\int_0^t e^{-(\alpha - \frac{1}{2}\sigma^2)s - \sigma W(s)}ds.$$

Therefore,

$$\begin{aligned} X(t) &= \Phi(t)[Y(0) + \beta\int_0^t e^{-(\alpha - \frac{1}{2}\sigma^2)s - \sigma W(s)}ds] \\ &= X(0)e^{(\alpha - \frac{1}{2}\sigma^2)t + \sigma W(t)} + \beta\int_0^t e^{(\alpha - \frac{1}{2}\sigma^2)(t-s) + \sigma[W(t) - W(s)]}ds. \end{aligned} \tag{4.42}$$

When $\gamma \neq 0$, we let $\tilde{X}(t) = X(t) + \frac{\gamma}{\sigma}$. Then

$$d\tilde{X} = (\alpha\tilde{X} + \beta - \frac{\alpha\gamma}{\sigma})dt + \sigma\tilde{X}dW.$$

We have

$$\begin{aligned} X(t) &= \tilde{X}(t) - \frac{\gamma}{\sigma} \\ &= -\frac{\gamma}{\sigma} + (X(0) + \frac{\gamma}{\sigma})e^{(\alpha-\frac{1}{2}\sigma^2)t+\sigma W(t)} \\ &+ (\beta - \frac{\alpha\gamma}{\sigma})\int_0^t e^{(\alpha-\frac{1}{2}\sigma^2)(t-s)+\sigma[W(t)-W(s)]}ds. \end{aligned} \tag{4.43}$$

□

## 4.6 Feynman-Kac Formula and Other Applications

The Feynman-Kac formula provides solutions for a particular class of partial differential equations(PDEs). Since the prices of many European-type derivatives are often a solution of some PDE as we will see in the next section, the Feynman-Kac formula plays an important role in option pricing.

**Theorem 4.5**(Feynman-Kac formula)

Let $u(t,x)$ be the solution of the following PDE

$$\frac{\partial}{\partial t}u(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x) + \gamma(x)u(t,x) = 0, \quad 0 \leq t \leq T, \tag{4.44}$$

with terminal condition $u(T,x) = \psi(x)$, where $\gamma(x)$ is a continuous function bounded from above and $\psi(x)$ is a continuous function satisfying the linear growth condition (4.29). Further assume that $|u(t,x)| \leq Le^{x^k}$, for some $L > 0$ and $k < 2$.

Then

$$u(t,x) = E\Big\{e^{\int_t^T \gamma(W(s))ds}\psi(W(T)) \; |W(t) = x\Big\}, \tag{4.45}$$

where $W(t)$ is a standard Wiener process.

*Proof:* Define

$$X_1(t) = e^{\int_0^t \gamma(W(s))ds}, \quad \text{and} \quad X_2(t) = u(t, W(t)).$$

Then

$$dX_1 = \gamma(W)X_1 dt$$

and

$$dX_2 = [\frac{\partial}{\partial t}u(t,W) + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,W)]dt + \frac{\partial}{\partial x}u(t,W)dW.$$

The first equation is obtained by ordinary differentiation and the second equation is obtained by Ito's Lemma.

Theorem 4.4 then yields

$$\begin{aligned} dX_1X_2 &= X_2dX_1 + X_1dX_2 \\ &= e^{\int_0^t \gamma(W(s))ds}\frac{\partial}{\partial x}u(t,W)dW. \end{aligned}$$

Thus

$$X_1(t)X_2(t) = X_1(0)X_2(0) + \int_0^t e^{\int_0^r \gamma(W(s))ds}\frac{\partial}{\partial x}u(r, W(r))dW(r).$$

Using the property that an Ito integral is a local martingale(this is beyond the scope of these notes) and that the tail distribution of the hitting time of a standard Wiener process to the boundary of the interval $[-a, \ a]$ is of order $e^{-ca^2}$, we can show that indeed

$$E\Big\{\int_{t'}^t e^{\int_0^r \gamma(W(s))ds}\frac{\partial}{\partial x}u(r, W(r))dW(r)\Big\} = 0.$$

Hence, $X_1(t)X_2(t)$ is a martingale. Thus,

$$E\Big\{e^{\int_0^T \gamma(W(s))ds}u(T, W(T)) \ |W(t) = x\Big\} = e^{\int_0^t \gamma(W(s))ds}u(t,x).$$

Dividing both sides by $e^{\int_0^t \gamma(W(s))ds}$ yields the result, since $u(T, W(T)) = \psi(W(T))$. □

As we have seen in Section 4.2, the stochastic process

$$Z_\lambda(t) = e^{\lambda W(t) - \lambda\mu t - \frac{1}{2}\lambda^2\sigma^2 t}$$

is a martingale. Using this martingale, we have derived other martingales in terms of the Wiener process. Applying Ito's Lemma and integration by parts, we are able to extend this result to a solution of a SDE.

**Theorem 4.6** Let $X(t)$ be a solution of the following SDE

$$dX = \alpha(t, X)dt + \sigma(t, X)dW, \tag{4.46}$$

where $\sigma(t, x)$ is bounded.

For any bounded stochastic process $b(t)$, define a stochastic process

$$Z_b(t) = e^{\int_0^t b(s)dX(s) - \int_0^t b(s)\alpha(s,X(s))ds - \frac{1}{2}\int_0^t b^2(s)\sigma^2(s,X(s))ds}. \tag{4.47}$$

Then $Z_b(t)$ is a martingale.

*Proof:* Define

$$X_1(t) = e^{-\int_0^t b(s)\alpha(s,X(s))ds - \frac{1}{2}\int_0^t b^2(s)\sigma^2(s,X(s))ds}$$

and

$$X_2(t) = e^{\int_0^t b(s)dX(s)}.$$

Then $Z_b(t) = X_1(t)X_2(t)$. Since

$$dX_1 = -[b(t)\alpha(t, X) + \frac{1}{2}b^2(t)\sigma^2(t, X)]X_2 dt$$

and by Ito's Lemma

$$dX_2 = [b(t)\alpha(t, X) + \frac{1}{2}b^2(t)\sigma^2(t, X)]X_1 dt + b(t)\sigma(t, X)]X_1 dW,$$

$$dZ_b = X_2 dX_1 + X_1 dX_2 = b(t)\sigma(t, X)Z_b dW.$$

Since $b$ and $\sigma$ are bounded, if $E(Z_b^2(t))$ is bounded, then $Z_b(t)$ is a martingale. The boundedness of $E(Z_b^2(t))$ can be proved in two steps: first, assume $b(t)$ is a step function;

then assume $b(t)$ is the limit of a sequence of step functions. The boundedness of the expectation in the second step is ensured by the Fatou's lemma[25]. $\square$

**Corollary 4.7** For any real $\lambda$,

$$Z_\lambda(t) = e^{\lambda X(t) - \lambda \int_0^t \alpha(s,X(s))ds - \frac{1}{2}\lambda^2 \int_0^t \sigma^2(s,X(s))ds} \tag{4.48}$$

is a martingale.

*Proof:* Simply let $b(t) = \lambda$. $\square$

**Corollary 4.8** the stochastic processes

$$X(t) - \int_0^t \alpha(s, X(s))ds \tag{4.49}$$

and

$$[X(t) - \int_0^t \alpha(s, X(s))ds]^2 - \int_0^t \sigma^2(s, X(s))ds \tag{4.50}$$

are martingales.

*Proof:* Notice that (4.49) and (4.50) are the first and second derivative of $Z_\lambda(t)$ with respect to $\lambda$ at $\lambda = 0$. The corollary follows from the fact that the derivative of a martingale is a martingale. $\square$

## 4.7 Option Pricing: Dynamic Hedging Approach

In this section, we apply the Feynman-Kac formula to option pricing.

Suppose that there is a market with a riskfree bond and a risky security. The prices of the bond and the security at time $t$ are denoted as $B(t)$ and $S(t)$, $0 \le t \le T$, respectively.

Assume that

$$dB = \delta B dt, \tag{4.51}$$

$$dS = \alpha S dt + \sigma S dW. \tag{4.52}$$

This is the case we discussed in the end of Section 4.2 with $\mu = \alpha - \frac{1}{2}\sigma^2$.

Consider now a European-type derivative which pays $\Phi(S(T))$ at time $T$, where $\Phi(x)$ is a continuous function with linear growth. Let $\phi(t,S)$ be the price of the derivative at time $t$ when the price of the underlying security at $t$ is $S$. Suppose that there is a self-financing strategy under which we are able to construct a portfolio from the bond and the risky security such that the value of this portfolio at time $T$ will be exactly equal to the payoff of the derivative $\Phi(S(T))$. Thus, if the market does not admit arbitrage, the value of the portfolio at time $t$ must be equal to the price of the derivative at time $t$. Let $\Theta(t,S)$ be the value of the risky security in the portfolio at time $t$. The value of the bond at time $t$ then is $\phi(t,S) - \Theta(t,S)$. The self-financing strategy implies that

$$d\phi(t,S) = [\phi(t,S) - \Theta(t,S)]\frac{dB}{B} + \Theta(t,S)\frac{dS}{S},$$

because that the right-hand side is the capital gain during the period $[t,\ t+dt]$, which is fully reinvested under this equation. Thus,

$$d\phi(t,S) = [\delta\phi(t,S) + (\alpha-\delta)\Theta(t,S)]dt + \sigma\Theta(t,S)dW. \tag{4.53}$$

On the other hand, if we assume that $\phi(t,S)$ satisfies the conditions of Ito's Lemma, we have

$$d\phi(t,S) = [\frac{\partial}{\partial t}\phi(t,S) + \alpha S\frac{\partial}{\partial S}\phi(t,S) + \frac{1}{2}\sigma^2 S^2\frac{\partial^2}{\partial S^2}\phi(t,S)]dt + \sigma S\frac{\partial}{\partial S}\phi(t,S)dW. \tag{4.54}$$

Equating the respective coefficients in these two equations, we have

$$\Theta(t,S) = S\frac{\partial}{\partial S}\phi(t,S), \tag{4.55}$$

$$\delta\phi(t,S) + (\alpha-\delta)\Theta(t,S) = \frac{\partial}{\partial t}\phi(t,S) + \alpha S\frac{\partial}{\partial S}\phi(t,S) + \frac{1}{2}\sigma^2 S^2\frac{\partial^2}{\partial S^2}\phi(t,S). \tag{4.56}$$

Eliminating $\Theta(t,S)$ from (4.56) yields

$$\frac{\partial}{\partial t}\phi(t,S) + \frac{1}{2}\sigma^2 S^2\frac{\partial^2}{\partial S^2}\phi(t,S) + \delta S\frac{\partial}{\partial S}\phi(t,S) - \delta\phi(t,S) = 0, \tag{4.57}$$

for any $0 \le t \le T$ and $S \ge 0$. Thus, $\phi(t, S)$ is the solution of the PDE (4.57) with terminal condition $\phi(T, S) = \Phi(S)$.

Now our problem becomes how to solve the above PDE. It is obvious we can not apply the Feynman-Kac formula directly. However, we are able to convert it into a PDE to which the Feynman-Kac formula can apply. We will do it in two steps.

1. Eliminating $S$ and $S^2$ in the coefficients.

Noting that the PDE is similar to an Euler-type ODE, we let $S = e^{\sigma x}$ and $v(t, x) = \phi(t, e^x)$. Hence,

$$\begin{aligned}
\frac{\partial}{\partial x} v(t, x) &= \sigma e^{\sigma x} \frac{\partial}{\partial x} \phi(t, e^x) = \sigma S \frac{\partial}{\partial S} \phi(t, S), \\
\frac{\partial^2}{\partial x^2} v(t, x) &= \sigma^2 S^2 \frac{\partial^2}{\partial S^2} \phi(t, S) + \sigma^2 S \frac{\partial}{\partial S} \phi(t, S).
\end{aligned}$$

Thus, $v(t, x)$ is the solution of the PDE

$$\frac{\partial}{\partial t} v(t, x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} v(t, x) + \frac{1}{\sigma}(\delta - \frac{1}{2}\sigma^2) \frac{\partial}{\partial x} v(t, x) - \delta v(t, x) = 0, \tag{4.58}$$

with $v(T, x) = \Phi(e^{\sigma x})$.

2. Eliminating the first-order term $\frac{\partial}{\partial x} v(t, x)$.

Since (4.58) is similar to a second-order linear ODE, the technique used there can apply. Let $v(t, x) = e^{\kappa x} u(t, x)$, where $\kappa$ will be determined later.

$$\begin{aligned}
\frac{\partial}{\partial x} v(t, x) &= e^{\kappa x} \frac{\partial}{\partial x} u(t, x) + \kappa e^{\kappa x} u(t, x), \\
\frac{\partial^2}{\partial x^2} v(t, x) &= e^{\kappa x} \frac{\partial^2}{\partial x^2} u(t, x) + 2\kappa e^{\kappa x} \frac{\partial}{\partial x} u(t, x) + \kappa^2 e^{\kappa x} u(t, x).
\end{aligned}$$

Thus,

$$\frac{\partial}{\partial t} u(t, x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} u(t, x) + [\kappa + \frac{1}{\sigma}(\delta - \frac{1}{2}\sigma^2)] \frac{\partial}{\partial x} u(t, x) + [\frac{1}{2}\kappa^2 + \frac{\kappa}{\sigma}(\delta - \frac{1}{2}\sigma^2) - \delta] u(t, x) = 0. \tag{4.59}$$

Let $\kappa = -\frac{1}{\sigma}(\delta - \frac{1}{2}\sigma^2)$. Then, $u(t,x)$ is the solution of

$$\frac{\partial}{\partial t}u(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x) - [\frac{1}{2\sigma^2}(\delta - \frac{1}{2}\sigma^2)^2 + \delta]u(t,x) = 0, \tag{4.60}$$

with $u(T,x) = e^{\frac{1}{\sigma}(\delta - \frac{1}{2}\sigma^2)x}\Phi(e^{\sigma x})$.

We now can apply the Faynman-Kac formula to (4.60). Hence,

$$u(t,x) = e^{-[\frac{1}{2\sigma^2}(\delta-\frac{1}{2}\sigma^2)^2+\delta](T-t)}E\Big\{e^{\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)W(T)}\Phi(e^{\sigma W(T)})\ |W(t)=x\Big\}. \tag{4.61}$$

Noting that $W(T)$, condition on $W(t) = x$, where $x = \frac{1}{\sigma}\log S$, is a normal random variable with mean $x$ and variance $T - t$.

$$\begin{aligned}
\phi(t,S) &= e^{-[\frac{1}{2\sigma^2}(\delta-\frac{1}{2}\sigma^2)^2+\delta](T-t)-\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)x} \\
&\times\ E\Big\{e^{\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)W(T)}\Phi(e^{\sigma W(T)})\ |W(t)=x\Big\} \\
&= \frac{1}{\sqrt{2\pi(T-t)}}e^{-[\frac{1}{2\sigma^2}(\delta-\frac{1}{2}\sigma^2)^2+\delta](T-t)-\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)x} \\
&\times \int_{-\infty}^{\infty} e^{\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)y}\Phi(e^{\sigma y})e^{-\frac{1}{2(T-t)}(y-x)^2}dy \\
&= \frac{e^{-\delta(T-t)}}{\sqrt{2\pi(T-t)}}\int_{-\infty}^{\infty}\Phi(Se^{\sigma y})e^{-\frac{1}{2(T-t)}[y-\frac{1}{\sigma}(\delta-\frac{1}{2}\sigma^2)(T-t)]^2}dy \\
&= \frac{e^{-\delta(T-t)}}{\sqrt{2\pi(T-t)}}\int_{-\infty}^{\infty}\Phi(Se^{(\delta-\frac{1}{2}\sigma^2)(T-t)+\sigma y})e^{-\frac{1}{2(T-t)}y^2}dy.
\end{aligned}$$

The above formula can be interpreted as follows: We imagine that there is a phantom risky security in the same period. Its price $\tilde{S}(t)$ also follows a geometric Wiener process but its expected return is exactly the same as the return of the riskfree bond, that is,

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

Then the derivative price we obtained can be written as

$$\phi(t,S) = e^{-\delta(T-t)}E\Big\{\Phi(\tilde{S}(T)\ |\tilde{S}(t)=S\Big\}. \tag{4.62}$$

It says that the price of the derivative at time $t$ is the discounted value, at the riskfree rate $\delta$, at time $t$ of the expected payoff, when the underlying security is the phantom security we define above.

We now use this formula to reproduce the Black-Scholes formula for a European call option we have derived in Section 4.2. Let $\phi_c(t,S)$ be the price. Then,

$$\begin{aligned}
\phi_c(t,S) &= \frac{e^{-\delta(T-t)}}{\sqrt{2\pi(T-t)}}\int_{-\infty}^{\infty}\max(Se^{(\delta-\frac{1}{2}\sigma^2)(T-t)+\sigma y}-K,\ 0)e^{-\frac{1}{2(T-t)}y^2}dy \\
&= \frac{e^{-\delta(T-t)}}{\sqrt{2\pi(T-t)}}\int_{y\le\frac{\log(S/K)+(\delta-\frac{1}{2}\sigma^2)(T-t)}{\sigma}}(Se^{(\delta-\frac{1}{2}\sigma^2)(T-t)-\sigma y}-K)e^{-\frac{1}{2(T-t)}y^2}dy \\
&= \frac{S}{\sqrt{2\pi(T-t)}}\int_{y\le\frac{\log(S/K)+(\delta-\frac{1}{2}\sigma^2)(T-t)}{\sigma}} e^{-\frac{1}{2(T-t)}[y+\sigma(T-t)]^2}dy \\
&- \frac{e^{-\delta(T-t)}K}{\sqrt{2\pi(T-t)}}\int_{y\le\frac{\log(S/K)+(\delta-\frac{1}{2}\sigma^2)(T-t)}{\sigma}} e^{-\frac{1}{2(T-t)}y^2}dy \\
&= \frac{S}{\sqrt{2\pi}}\int_{y\le\frac{\log(S/K)+(\delta+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}} e^{-\frac{1}{2}y^2}dy \\
&- \frac{e^{-\delta(T-t)}K}{\sqrt{2\pi}}\int_{y\le\frac{\log(S/K)+(\delta-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}} e^{-\frac{1}{2}y^2}dy.
\end{aligned}$$

Thus,

$$\phi_c(t,S) = SN(d_1) - Ke^{-\delta(T-t)}N(d_2), \tag{4.63}$$

where

$$d_1 = \frac{\log(S/K)+(\delta+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \tag{4.64}$$

and

$$d_2 = \frac{\log(S/K)+(\delta-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}. \tag{4.65}$$

We obtain the Black-Scholes formula.

## 4.8 Girsanov Theorem

**Theorem 4.9**(Girsanov) Let $W(t)$, $0 \leq t \leq T$, be a standard Wiener process on a probability space $(\Omega, \mathcal{F}, P)$ and $b(t)$ a bounded stochastic process. Define a function on all the events in $\mathcal{F}$ in the following way: for any $F \in \mathcal{F}$,

$$Q(F) = \int_F e^{\int_0^T b(t)dW(t) - \frac{1}{2}\int_0^T b^2(t)dt} dP. \tag{4.66}$$

Then,

1. $Q$ is a probability measure on $(\Omega, \mathcal{F})$ and it is equivalent to $P$;
2. the stochastic process
$$\tilde{W}(t) = -\int_0^t b(s)ds + W(t) \tag{4.67}$$
is a standard Wiener process under the probability measure $Q$.

*Proof:* Define a stochastic process

$$z(t) = e^{\int_0^t b(s)dW(s) - \frac{1}{2}\int_0^t b^2(s)ds}. \tag{4.68}$$

Since $b(t)$ is bounded, by Theorem 4.6 $z(t)$ is a martingale under $P$( $\alpha = 0$ and $\sigma = 1$ in this case). Thus,

$$\begin{aligned} Q(\Omega) &= \int_\Omega e^{\int_0^T b(t)dW(t) - \frac{1}{2}\int_0^T b^2(t)dt} dP \\ &= E(z(T)) = z(0) = 1. \end{aligned}$$

It is easy to see from the definition that $Q$ is nonnegative and additive on $\mathcal{F}$. Hence $Q$ is a probability measure. Since $z(t)$ is the Radon-Nikodym derivative of $Q$ with respect to $P$ and it is always finite and positive, $Q$ and $P$ are equivalent.

Next, we show that $\tilde{W}(t)$ is a standard Wiener process under $Q$.

Recall (Section 4.2) that $\tilde{W}(t)$ is a standard Wiener process if and only if for any real $\lambda$,

$$Z_\lambda(t) = e^{\lambda \tilde{W}(t) - \frac{1}{2}\lambda^2 t}$$

is a martingale. Denote that $E_Q$ the expectation under $Q$. We need to show that for any $s > t$,

$$E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} = Z_\lambda(t). \tag{4.69}$$

Given any $F \in \mathcal{B}_t$, we have

$$\begin{aligned} & \int_F E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} dQ = \int_F E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} z(T) dP \\ = & \int_F E\Big\{E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} z(T) \ |\mathcal{B}_t\Big\} dP = \int_F E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} z(t) dP. \end{aligned}$$

On the other hand,

$$\begin{aligned} & \int_F E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} dQ = \int_F Z_\lambda(s) dQ = \int_F Z_\lambda(s) z(T) dP \\ = & \int_F E\Big\{Z_\lambda(s) z(T) \ |\mathcal{B}_t\Big\} dP = \int_F E\Big\{E\Big\{Z_\lambda(s) z(T) \ |\mathcal{B}_s\Big\} \ |\mathcal{B}_t\Big\} dP \\ = & \int_F E\Big\{Z_\lambda(s) z(s) \ |\mathcal{B}_t\Big\} dP. \end{aligned}$$

Hence

$$E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} z(t) = E\Big\{Z_\lambda(s) z(s) \ |\mathcal{B}_t\Big\}. \tag{4.70}$$

We now show that $Z_\lambda(t) z(t)$ is a martingale under $P$. Define

$$X_1(t) = e^{\lambda W(t) + \int_0^t b(s) dW(s)}$$

and

$$X_2(t) = e^{-\frac{1}{2}\lambda^2 t - \lambda \int_0^t b(s) ds - \frac{1}{2}\int_0^t b^2(s) ds}.$$

Then, Applying Ito's Lemma, we obtain

$$dX_1 = \frac{1}{2}[\lambda + b(t)]^2 X_1 dt + [\lambda + b(t)] X_1 dW$$

and

$$dX_2 = -\frac{1}{2}[\lambda + b(t)]^2 X_2 dt.$$

The integration by parts yields

$$d[Z_\lambda z] = d[X_1 X_2] = [\lambda + b(t)] Z_\lambda z dW.$$

Similar to the proof of Theorem 4.6, $Z_\lambda(t)z(t)$ is a martingale under $P$. Therefore,

$$E\Big\{Z_\lambda(s)z(s) \ |\mathcal{B}_t\Big\} = Z_\lambda(t)z(t).$$

Since $z(t)$ is positive, we have from (4.70)

$$E_Q\Big\{Z_\lambda(s) \ |\mathcal{B}_t\Big\} = Z_\lambda(t).$$

The proof is complete. $\square$

It is important to note that $z^{-1}(t)$ can be written as

$$z^{-1}(t) = e^{-\int_0^t b(s)d\tilde{W}(s) - \frac{1}{2}\int_0^t b^2(s)ds}. \tag{4.71}$$

Thus, $z^{-1}(t)$ is a martingale under the probability measure $Q$. This property is often used when we need to change form the probability measure $Q$ to the probability measure $P$.

We now examine the SDE

$$dX = \alpha(t, X)dt + \sigma(t, X)dW \tag{4.72}$$

under the new probability measure $Q$.

Replacing $W(t)$ by $\int_0^t b(s)ds + \tilde{W}(t)$, we have

$$dX = [\alpha(t, X) + \sigma(t, X)b(t)]dt + \sigma(t, X)d\tilde{W}. \tag{4.73}$$

Thus, under the probability measure $Q$, the SDE has new drift $\alpha(t, X) + \sigma(t, X)b(t)$ but the same volatility $\sigma(t, X)$. This leads to the following corollary

**Corollary 4.10** Let $X(t)$ be a solution of the SDE (4.72), where $\sigma(t,x)$ is a positive function. Also let $\beta(t,x)$ be a continuous function such that

$$\frac{\beta(t,x)-\alpha(t,x)}{\sigma(t,x)} \tag{4.74}$$

is bounded.

Then, there exists a probability measure $Q$ such that under this measure $X(t)$ is a solution of the following SDE

$$dX = \beta(t,X)dt + \sigma(t,X)d\tilde{W}, \tag{4.75}$$

where $\tilde{W}(t)$ is a standard Wiener process under $Q$. Moreover, the Radon-Nikodym derivative of $Q$ with respect to $P$ is

$$\frac{dQ}{dP} = e^{\int_0^T b(t)dW(t) - \frac{1}{2}\int_0^T b^2(t)dt}, \tag{4.76}$$

where

$$b(t) = \frac{\beta(t,X(t))-\alpha(t,X(t))}{\sigma(t,X(t))}. \tag{4.77}$$

Corollary 4.10 is a very important result for continuous-time finance models. As we will see in the next chapter that no arbitrage in a price system will imply that there is a unique probability measure such that all present value processes under this new measure are martingales. It is equivalent to the condition that the SDEs representing price processes will have the same drift. Thus, Corollary 4.10 provides a methodology to find a such measure. We will study it further in the next chapter.

Another application of the Girsanov Theorem is to compute barrier hitting probabilities for Wiener processes with drift.

Consider a Wiener process with drift $\mu$: $W_\mu(t) = \mu t + W(t)$, where $W(t)$ is the standard Wiener process. Let $x = a + kt$ be a straight line(horizontal or nonhorizontal). Define

$$\tau = \inf\{t;\quad W_\mu(t) = a + kt\}. \tag{4.78}$$

Then $\tau$ is the first passage time for the barrier $x = a + kt$. It is easy to see that $\tau = \inf\{t;\ \ W(t) - (k-\mu)t = a\}$. Let

$$b = k - \mu \text{ and } z(t) = e^{bW(t) - \frac{1}{2}b^2 t}. \tag{4.79}$$

Then $\tilde{W}(t) = W(t) - (k-\mu)t$ is a standard Wiener process under $\tilde{Q}$ generated by $z(T)$ in (4.79) and $\tau$ is the first passage time of $\tilde{W}(t)$ under $\tilde{Q}$. Thus its density under $\tilde{Q}$ is given in (4.18). Let $f(t; a, k)$ be the density function of $\tau$. Then we have

$$\begin{aligned} f(t; a, k)dt &= E\{\chi_{\{\tau \in dt\}}\} \\ &= E_{\tilde{Q}}\{\chi_{\{\tau \in dt\}} z^{-1}(t)\} \\ &= z^{-1}(t) f_a(t) dt = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}(t + a/b)^2} dt. \end{aligned}$$

Thus,

$$f(t; a, k) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}(t + a/b)^2}, \tag{4.80}$$

where $b$ is given in (4.79).

To identify the distribution of $\tau$, we introduce the inverse Guassian(IG) distribution which has density

$$f_{IG}(x) = \frac{\alpha}{\sqrt{2\pi\beta x^3}} e^{-\frac{1}{2\beta x}(x-\alpha)^2}, \quad x > 0, \tag{4.81}$$

where $\alpha > 0,\ \beta > 0$. $\alpha$ and $\beta$ are called the location parameter and dispersion parameter respectively. Its distribution function can be expressed in terms of the distribution function of the standard normal distribution

$$F_{IG}(x) = N\Big(\frac{x - \alpha}{\sqrt{\beta x}}\Big) + e^{2\alpha/\beta} N\Big(-\frac{x + \alpha}{\sqrt{\beta x}}\Big). \tag{4.82}$$

The Laplace transform is

$$\tilde{f}_{IG}(z) = e^{\frac{\alpha}{\beta}(1 - \sqrt{1 + 2\beta z})}. \tag{4.83}$$

If we allow $\alpha$ to be negative($\alpha$ in the coefficient is replaced by $|\alpha|$ accordingly), we obtain a defective inverse Guassian distribution since

$$f_{IG}(x) = e^{2\alpha/\beta} \hat{f}_{IG}(x), \tag{4.84}$$

where $\hat{f}_{IG}(x)$ is the nondefective density of an IG distribution with location parameter $-\alpha$ and dispersion parameter $\beta$.

By comparing the density of the IG distribution and the function form of $f(t;a,k)$, we have the following Corollary.

**Corollary 4.11** The distribution of the first passage time $\tau$ defined in (4.78) for a Wiener process with drift $\mu$ is an IG distribution with $\alpha = -a/b$ and $\beta = 1/b^2$. If $-a/b > 0$, the distribution is nondefective hence $\tau$ is a finite random variable. In this case, the Wiener process $\mu t + W(t)$ will hit the barrier $x = a + kt$ sooner or later. If $-a/b < 0$, the distribution is defective and $\Pr\{\tau < \infty\} = e^{-2ab}$.

Next, we consider the distribution of $W(T), W(t) + \mu t = a + kt$ for some $0 \le t \le T$. Let $g(x;a,k)$ be its density function. Then,

$$\begin{aligned} g(x;a,k)dt &= E\{\chi_{\{W(T)\in dx, W_\mu(t)=a+kt, 0\le t\le T\}}\} \\ &= E_{\tilde{Q}}\{\chi_{\{\tilde{W}(T)\in dy, \tilde{W}(t)=a, 0\le t\le T\}} z^{-1}(T)\} \\ &\quad \text{(where } y = x + (\mu - k)T = x - bT\text{)} \\ &= e^{-by-\frac{1}{2}b^2T} g_a(y)dy \quad \left(\text{since } z^{-1}(T) = e^{-b\tilde{W}(T)-\frac{1}{2}b^2T} = e^{-by-\frac{1}{2}b^2T}\right) \\ &= e^{-bx+\frac{1}{2}b^2T} g_a(x - bT)dx, \end{aligned}$$

where $g_a(x)$ is given in (4.19). We thus have

$$g(x;a,k) = \begin{cases} \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-2a)^2}{2T} - 2ab}, & x < a + bT \\ \frac{1}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}}, & x \ge a + bT \end{cases} \tag{4.85}$$

Similarly for $a < 0$, we have

$$g(x;a,k) = \begin{cases} \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-2a)^2}{2T} - 2ab}, & x > a + bT \\ \frac{1}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}}, & x \le a + bT \end{cases} \tag{4.86}$$

For the double-barrier case, the density function of $W(T)$, $a_l+kt < W(t)+\mu t < a_u+kt$ for all $0 \le t \le T$, is

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi T}} \sum_{n=-\infty}^{n=\infty} e^{2nb(a_u-a_l)}\{e^{-\frac{[x+2n(a_u-a_l)]^2}{2T}} - e^{-\frac{[x-2a_u+2n(a_u-a_l)]^2}{2T}-2ba_u}\}, \\ \qquad a_l + bT < x < a_u + bT \\ 0, \\ \qquad x \le a_l + bT \text{ or } x \ge a_u + bT. \end{cases} \tag{4.87}$$

The derivation is very similar to the single barrier case and is omitted.

## 4.9 Multi-Dimensional Ito Processes

In this section, we extend the results in previous sections to multi-dimensional Ito processes.

A stochastic process $\mathbf{W}(t) = (W_1(t), W_2(t), \cdots, W_I(t))$ is said to be an $I$-dimensional standard Wiener process if $W_i(t), i = 1, \cdots, I$ are independent standard Wiener processes. In general, an $I$-dimensional stochastic process is a Wiener process if each of its components is a Wiener process. However, the components do not need to be independent of each other.

Let $\lambda = (\lambda_1, \cdots, \lambda_I)'$ be a real-valued vector, where $'$ denotes the corresponding transpose. Similar to (4.8), $\mathbf{W}(t)$ is a standard Wiener process if and only if

$$Z_\lambda(t) = e^{\lambda' \mathbf{W}(t) - \frac{1}{2}\lambda' \lambda t} \tag{4.88}$$

is a martingale (with respect to the Borel filtration generated by $\mathbf{W}(t)$).

A stochastic process $\mathbf{X}(t) = (X_1(t), X_2(t), \cdots, X_N(t))$ is a solution of an $N$-dimensional SDE

$$dX_n = \alpha_n(t, \mathbf{X})dt + \sum_{i=1}^{I} \sigma_{ni}(t, \mathbf{X})dW_i, \tag{4.89}$$

$n = 1, \cdots, N$, if

$$\begin{aligned} X_n(t) = X_n(0) \quad &+ \quad \int_0^t \alpha_n(s, \mathbf{X}(s))ds \\ &+ \quad \sum_{i=1}^{I} \int_0^t \sigma_{ni}(s, \mathbf{X}(s))dW_i(s). \end{aligned} \tag{4.90}$$

Easy to see that every term in (4.90) is well-defined.

The conditions which guarantee the existence and uniqueness of a solution remain the same as that in the one-dimensional case, i.e. all $\alpha_n(t,x)$ and $\sigma_{ni}(t,x)$ are Lipschitz in $x$ and grow linearly in $x$.

The following theorems are the analogy of Theorems 4.3, 4.6 and Corollary 4.10.

**Theorem 4.11**(Ito's Lemma)

Let $\mathbf{X}(t)$ be a solution of the SDE (4.89) and $g(t,\mathbf{x}) = g(t, x_1, \cdots, x_N)$ a function which is continuously differentiable in $t$ and continuously twice differentiable in each $x_n$. Then $g(t, \mathbf{X}(t))$ is a solution of the following SDE

$$\begin{aligned} dg(t,\mathbf{X}) &= [\frac{\partial}{\partial t} g(t,\mathbf{X}) + \sum_{n=1}^{N} \alpha_n(t,\mathbf{X}) \frac{\partial}{\partial x_n} g(t,\mathbf{X}) \\ &+ \frac{1}{2} \sum_{i=1}^{I} \sum_{m=1}^{N} \sum_{n=1}^{N} \sigma_{mi}(t,\mathbf{X}) \sigma_{ni}(t,\mathbf{X}) \frac{\partial^2}{\partial x_m \partial x_n} g(t,\mathbf{X})] dt \\ &+ \sum_{i=1}^{I} \sum_{n=1}^{N} \sigma_{ni}(t,\mathbf{X}) \frac{\partial}{\partial x_n} g(t,\mathbf{X}) dW_i. \end{aligned} \tag{4.91}$$

It can be seen that Theorem 4.4 is an immediate consequence of this theorem with $g(t, x_1, x_2) = x_1 x_2$.

**Theorem 4.12** Denote that $\alpha(t,\mathbf{x}) = (\alpha_1(t,\mathbf{x}), \cdots, \alpha_N(t,\mathbf{x}))'$ and $\sigma(t,\mathbf{x}) = \left[\sigma_{ni}(t,\mathbf{x})\right]$, a $N \times I$ matrix. If $\sigma(t,\mathbf{x})$ is bounded, then for any bounded stochastic process $\mathbf{b}(t) = (b_1(t), \cdots, b_N(t))'$,

$$Z_b(t) = e^{\int_0^t \mathbf{b}'(s) d\mathbf{X}(s) - \int_0^t \mathbf{b}'(s)\alpha(s,\mathbf{X}(s))ds - \frac{1}{2}\int_0^t \mathbf{b}'(s)\sigma'(s,\mathbf{X}(s))\sigma(s,\mathbf{X}(s))\mathbf{b}(s)ds} \tag{4.92}$$

is a martingale.

**Theorem 4.13**(Girsanov Theorem) Let $\mathbf{X}(t)$ be a solution of the SDE (4.89). For a given vector-valued function $\beta(t,\mathbf{x}) = (\beta_1(t,\mathbf{x}), \cdots, \beta_N(t,\mathbf{x}))'$, suppose that there is a bounded

$I$-dimensional function $\gamma(t, \mathbf{x}) = (\gamma_1(t, \mathbf{x}), \cdots, \gamma_I(t, \mathbf{x}))'$ such that

$$\beta(t, \mathbf{x}) = \alpha(t, \mathbf{x}) + \sigma(t, \mathbf{x})\gamma(t, \mathbf{x}). \tag{4.93}$$

Then, there exists an equivalent probability measure $Q$ such that under $Q$, $\mathbf{X}(t)$ is a solution of the following SDE

$$dX_n = \beta_n(t, \mathbf{X})dt + \sum_{i=1}^{I} \sigma_{ni}(t, \mathbf{X})d\tilde{W}_i, \tag{4.94}$$

$n = 1, \cdots, N$, where $\tilde{\mathbf{W}}(t) = (\tilde{W}_1(t), \cdots, \tilde{W}_I(t))$ is a standard Wiener process under $Q$.

# Chapter 5

# Continuous-Time Finance Models

## 5.1 Security Markets and Valuation

Consider an economy characterised by a probability space $(\Omega, \mathcal{F}, P)$, where $\Omega$ is the state space, $\mathcal{F}$ the collection of all possible events and $P$ the probability measure which may be interpreted as the homogeneous belief among investors. In this economy, there is a continuous trading security market with trading period $[0,\ T]$. $\mathcal{F}_t,\ 0 \leq t \leq T$ is the corresponding information structure with $\mathcal{F}_0 = \{\phi,\ \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. The precise definition of an information structure and its interpretation are given in the sections 2.1 and 2.3. Assume that there are $N+1$ traded securities. Their prices at time $t$ are $p_0(t), p_1(t), \cdots, p_N(t)$, respectively. Among them, one, say the first security $p_0$, is the money market account. The spot rate for the money market account is defined as

$$r(t) = \lim_{\tau \to 0^+} \frac{1 - p(t, t+\tau)}{\tau}, \tag{5.1}$$

$p(t, t+\tau)$ is the price at time $t$ of a default-free bond which pays one unit at time $t+\tau$(see Section 3.5), i.e. $r(t)$ is the instantaneous interest rate at time $t$. Hence we have

$$p_0(t) = e^{\int_0^t r(s)ds}, \tag{5.2}$$

or

$$dp_0 = r(t)p_0dt. \tag{5.3}$$

Others are described by a system of stochastic differential equations:

$$dp_n = \alpha_n(t, \mathbf{p})dt + \sum_{i=1}^{N} \sigma_{ni}(t, \mathbf{p})dW_i, \tag{5.4}$$

$n = 1, \cdots, N$, where $\mathbf{p} = (p_0, p_1, \cdots, p_N)$. Moreover, we assume that the matrix $\sigma(t, \mathbf{p}) = [\sigma_{ni}(t, \mathbf{p})]$ is nonsingular for all $\mathbf{p} > 0$. The difference between the money market account and other securities is that the former does not have a diffusion term hence is of finite variation.

A trading strategy $\theta(t) = (\theta_1(t), \cdots, \theta_N(t))'$ of an investor is a stochastic process on $[0,\ T]$ such that the investor owns $\theta_n(t)$ shares of security $n$ at time $t$. In these notes, we restrict ourselves to those trading strategies which satisfy

$$E\Big(\int_0^T \theta_n^2(t)\alpha_n^2(t, \mathbf{p}(t))dt\Big) < \infty, \text{ and } E\Big(\int_0^T \theta_n^2(t)\sigma_{ni}^2(t, \mathbf{p}(t))dt\Big) < \infty. \tag{5.5}$$

A trading strategy $\theta(t)$ is self-financing if for any $t$,

$$\sum_{n=0}^{N} \theta_n(t)p_n(t) = \sum_{n=0}^{N} \theta_n(0)p_n(0) + \sum_{n=0}^{N} \int_0^t \theta_n(s)dp_n(s). \tag{5.6}$$

The second term in the right-hand side represents the capital gain achieved by $\theta(t)$ over $[0,\ t)$. We may write equation (5.6) in matrix form

$$\theta'(t)\mathbf{p}(t) = \theta'(0)\mathbf{p}(0) + \int_0^t \theta'(s)d\mathbf{p}(s), \tag{5.7}$$

or simply

$$d[\theta'(t)\mathbf{p}(t)] = \theta'(t)d\mathbf{p}(t). \tag{5.8}$$

**Example 5.1** Simple Self-Financing Strategies

Let $0 = t_0 < t_1 < \cdots < t_M = T$ be a sequence of preset times. A trading strategy $\theta(t)$ is said to be simple if

$$\theta(t) = \theta(t_m), \quad t_m \leq t < t_{m+1}, \quad m = 0, 1, \cdots, M-1. \tag{5.9}$$

A simple trading strategy is self-financing if and only if

$$\sum_{n=0}^{N} \theta_n(t_m) p_n(t_{m+1}) = \sum_{n=0}^{N} \theta_n(t_{m+1}) p_n(t_{m+1}). \tag{5.10}$$

Thus, under a simple trading strategy, trading occurs only at a finite number of preset times. The value of the portfolio held by the investor before each trading is equal to the value of the portfolio after the trading. □

The price system $\mathbf{p}(t)$ does not admit arbitrage if for any self-financing trading strategy $\theta(t)$, neither

$$\theta'(0)\mathbf{p}(0) = 0 \ \text{ and } \ \theta'(T)\mathbf{p}(T) > 0 \tag{5.11}$$

nor

$$\theta'(0)\mathbf{p}(0) < 0 \ \text{ and } \ \theta'(T)\mathbf{p}(T) \geq 0 \tag{5.12}$$

can happen.

A contingent claim is such that its payoff, made at time $T$, depends on the prevailing state at time $T$. Thus any contingent claim is represented by a random variable $X$ on $(\Omega, \mathcal{F}, P)$, where $X$ is the value of the payoff at time $T$. A simple example is a European option expired at time $T$. The above concept can be generalised to contingent claims paid at other times. However, if we assume that any payoff will be deposited right away into the money market account and there is no withdraw until time $T$, all situations are equivalent. In these notes, we only consider the contingent claims whose second moments exist. Hence, all contingent claims form a Hilbert space(Appendix B). A contingent claim $X$ is said to be attainable if there is a self-financing trading strategy $\theta(t)$ such that

$$X = \theta'(0)\mathbf{p}(0) + \int_0^T \theta'(t) d\mathbf{p}(t). \tag{5.13}$$

Similar to the discrete case, a market is called complete if every contingent claim is attainable. In general, a market might not be complete. However, if $\sigma(t, \mathbf{p})$ is nonsingular for all $\mathbf{p} > 0$, every contingent claim is attainable and the market is complete. This can be proved by the Martingale Representation Theorem[8].

We now look into the valuation of contingent claims. We always assume that the price system $\mathbf{p}(t)$ does not admit arbitrage. Let $X$ be a contingent claim. If it is attainable, then

$$X = \theta'(0)\mathbf{p}(0) + \int_0^T \theta'(t) d\mathbf{p}(t),$$

for some self-financing trading strategy $\theta(t)$. Naturally, we define the price $\phi(X)$ at time 0 of $X$ as $\theta'(0)\mathbf{p}(0)$. Under the no-arbitrage condition, $\phi(X) = \theta'(0)\mathbf{p}(0)$ is uniquely defined even if there is more than one self-financing trading strategy. It is also easy to see that $\phi(X)$ is a strictly positive, continuous, linear functional on the space of all contingent claims, since the market is complete.

We are now able to price all contingent claims. However, the method we employ is not very practical. The corresponding self-financing trading strategy is unsovable except for a few cases. Below we are looking for a different approach to define the price functional which will provide a practical method for pricing contingent claims.

Let us first define the present value process of each security:

$$a_n(t) = e^{-\int_0^t r(s)ds} p_n(t), \ \ n = 0, 1, \cdots, N. \tag{5.14}$$

Thus, $a_n(t)$ is the time-0 value of $p_n(t)$, discounted at the short rate process. Thus, $p_0(t)$ serves as a benchmarking security.

**Lemma 5.1** Let $\theta(t)$ be a self-financing trading strategy for $\mathbf{p}(t)$. Then it is also a self-financing trading strategy for $\mathbf{a}(t) = (a_0(t), a_1(t), \cdots, a_N(t))'$. i.e.

$$\theta'(t)\mathbf{a}(t) = \theta'(0)\mathbf{a}(0) + \int_0^t \theta'(s) d\mathbf{a}(s). \tag{5.15}$$

*Proof:*

$$\begin{aligned} d[\theta'(t)\mathbf{a}(t)] &= d[e^{-\int_0^t r(s)ds}\theta'(t)\mathbf{p}(t)] \\ &= e^{-\int_0^t r(s)ds}d[\theta'(t)\mathbf{p}(t)] - e^{-\int_0^t r(s)ds}r(t)\theta'(t)\mathbf{p}(t)dt \\ &= e^{-\int_0^t r(s)ds}\theta'(t)d\mathbf{p}(t) - e^{-\int_0^t r(s)ds}r(t)\theta'(t)\mathbf{p}(t)dt \\ &= \theta'(t)d\mathbf{a}(t). \end{aligned}$$

□

**Theorem 5.2** The price system $\mathbf{p}(t)$ does not admit arbitrage if and only if there is a unique probability measure $Q$ equivalent to $P$ such that each present value process $a_n(t)$ is a martingale under $Q$. Furthermore, for any contingent claim $X$,

$$\phi(X) = E_Q\Big(e^{-\int_0^T r(t)dt}X\Big). \tag{5.16}$$

The probability measure $Q$ is referred to as the risk-neutral probability measure.

*Proof:* Sufficiency. Let $Q$ be a probability measure such that each present value process $a_n(t)$ is a martingale under $Q$. For any self-financing trading strategy $\theta(t)$, (5.15) yields

$$\theta'(T)\mathbf{a}(T) = \theta'(0)\mathbf{a}(0) + \int_0^T \theta'(s)d\mathbf{a}(s). \tag{5.17}$$

Since $a_n(t)$ is a martingale under $Q$, we have

$$E_Q\Big(\int_0^T \theta_n(s)da_n(s)\Big) = 0.$$

Thus,

$$E_Q\Big(\theta'(T)\mathbf{a}(T)\Big) = \theta'(0)\mathbf{a}(0).$$

It is easy to see that neither (5.11) nor (5.12) can happen under this identity.

Necessity. By the Riesz Representation Theorem(Appendix B), there is a unique random variable $Z$ such that

$$\phi(X) = E\Big(e^{-\int_0^T r(t)dt}XZ\Big). \tag{5.18}$$

Since $\phi$ is strictly positive, so is $Z$. Furthermore,

$$E(Z) = \phi(e^{\int_0^T r(t)dt}) = 1.$$

This is because the claim $e^{\int_0^T r(t)dt}$ can be replicated by depositing one unit into the money market account. Thus, we may define a probability measure $Q$ as follows: for any $F \in \mathcal{F}$,

$$Q(F) = E\Big(\chi_F Z\Big), \tag{5.19}$$

where $\chi_F$ is the indicator of $F$. It is easy to see that $Q$ is equivalent to $P$ since $Z$ is the Radon-Nikodym derivative of $Q$ with respect to $P$ and it is strictly positive.

We now show that each $a_n(t)$ is a martingale under $Q$. For any $t < s$ and $F \in \mathcal{F}_t$, construct a self-financing simple trading strategy $\theta(t)$ as follows:

1. $\theta_{n'}(t') = 0,\ n' \neq 0, n; 0 \leq t' \leq T.$

2. $\theta_0(t') = \theta_n(t') = 0,\ 0 \leq t' < t.$

3. $\theta_n(t') = \chi_F$ and $\theta_0(t') = -a_n(t)\chi_F$, for $t \leq t' < s$.

4. $\theta_n(t') = 0$ and $\theta_0(t') = [a_n(s) - a_n(t)]\chi_F$, for $s \leq t'$.

This strategy says that we buy one share of security $n$ at time $t$ and then sell it at time $s$. Since this strategy costs nothing the price of the terminal payoff

$$X = e^{\int_0^T r(u)du}[a_n(s) - a_n(t)]\chi_F$$

is zero by the no-arbitrage condition, i.e.

$$0 = \phi(X) = E\Big([a_n(s) - a_n(t)]\chi_F Z\Big) = E_Q\Big([a_n(s) - a_n(t)]\chi_F\Big).$$

Since $F$ is arbitrary, $a_n(t)$ is a martingale under $Q$. $\square$

The advantage of this result is that instead of finding a self-financing strategy we only need to find an equivalent probability measure such that under the new measure all present value processes are martingale. In many cases this can be done quite easily thanks to the Girsanov theorem.

To illustrate how to use this technique, we reconsider the Black-Scholes model. Using the same notations as in (4.51) and (4.52), the present value process

$$a(t) = e^{-\delta t} S(t).$$

Thus,

$$da = (\alpha - \delta) a dt + \sigma a dW.$$

Since $a(t)$ is a martingale if and only if its drift term is zero. By Corollary 4.10, we may choose

$$b(t) = \frac{\delta - \alpha}{\sigma} \tag{5.20}$$

and

$$Z(t) = e^{\frac{\delta - \alpha}{\sigma} W(t) - \frac{1}{2}(\frac{\delta - \alpha}{\sigma})^2 t}. \tag{5.21}$$

Under the probability measure

$$Q(F) = \int_F Z(T) dP, \tag{5.22}$$

$a(t)$ is a martingale. Also, by Theorem 4.9,

$$\tilde{W}(t) = -\frac{\delta - \alpha}{\sigma} t + W(t) \tag{5.23}$$

is a standard Wiener process under $Q$. Thus,

$$dS = \alpha S dt + \sigma S dW = \delta S dt + \sigma S d\tilde{W}, \tag{5.24}$$

i.e. $S(t)$ is a geometric Wiener process with parameters $\delta$ and $\sigma$ under $Q$.

For any European option which pays $\Phi(S)$ at time $T$ when the security value at $T$ is $S$, its price is

$$\phi(\Phi(S)) = e^{-\delta T} E_Q(\Phi(S)),$$

which is exactly (4.62) and we obtain the Black-Scholes formula again.

We now consider some variations of the Black-Scholes pricing formula.

First, let us assume that in addition to the prices of a bond and a risky security subject to Equations (4.51) and (4.52), the risky security pays dividends continuously at a fixed rate $\zeta$. This model was discussed in [27] and [28]. In this case the present value process is

$$a(t) = \zeta \int_0^t e^{-\delta u} S(u) du + e^{-\delta t} S(t).$$

and

$$da = [a - \int_0^t e^{-\delta u} S(u) du][(\alpha + \zeta - \delta)dt + \sigma dW].$$

Thus,

$$b(t) = \frac{\delta - \alpha - \zeta}{\sigma}$$

and

$$Z(t) = e^{\frac{\delta-\alpha-\zeta}{\sigma} W(t) - \frac{1}{2}(\frac{\delta-\alpha-\zeta}{\sigma})^2 t}.$$

Under the probability measure

$$Q(F) = \int_F Z(T) dP,$$

$a(t)$ is a martingale and

$$\tilde{W}(t) = -\frac{\delta - \alpha - \zeta}{\sigma} t + W(t)$$

is a standard Wiener process under $Q$. Thus,

$$dS = \alpha S dt + \sigma S dW = (\delta - \zeta) S dt + \sigma S d\tilde{W}, \tag{5.25}$$

i.e. $S(t)$ is a geometric Wiener process with parameters $\delta - \zeta$ and $\sigma$ under $Q$. Let $\phi_d$ be the price at $t$ of a European call whose payoff at $T$ is $\max\{S(T) - K,\ 0\}$. Then,

$$\phi_d = e^{-\zeta(T-t)} S(t) N(\hat{d}_1) - K e^{-\delta(T-t)} N(\hat{d}_2), \tag{5.26}$$

where

$$\hat{d}_1 = \frac{\log(S(t)/K) + (\delta - \zeta + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$\hat{d}_2 = \frac{\log(S(t)/K) + (\delta - \zeta - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

Next, we derive the price of a futures option. A futures contract on an underlying security promises its holder to purchase the security at a certain time at a certain price. Its buyer and seller both have obligation to honour the contract. Hence the price of a futures contract is stochastic during the lifetime of the contract. Let $F(t)$ be the futures price at time $t$ of a risky security $S(t)$ to be delivered at time $T$. Under the Black-Scholes framework, it is easy to see that

$$F(t) = e^{\delta(T-t)} E_Q(S(T)) = e^{\delta(T-t)} S(t),$$

where $Q$ is defined in (5.22).

Suppose that a futures option maturing at time $T_1 < T$ with strike price $K$. Hence its payoff is $\max\{F(T_1) - K,\ 0\}$, delivered at $T$. The price of this option is

$$\phi_f = e^{-\delta T} E_Q\{\max\{F(T_1) - K,\ 0\}\},$$

which is equivalent to

$$\begin{aligned} \phi_f &= e^{-\delta T_1} E_Q\{\max\{S(T_1) - K e^{-\delta(T-T_1)},\ 0\}\} \\ &= e^{-\delta T}[F(0)N(\tilde{d}_1) - KN(\tilde{d}_2)], \end{aligned} \tag{5.27}$$

where

$$\tilde{d}_1 = \frac{\log(F(0)/K) + \frac{1}{2}\sigma^2 T_1}{\sigma\sqrt{T_1}}$$

and

$$\tilde{d}_2 = \frac{\log(F(0)/K) - \frac{1}{2}\sigma^2 T_1}{\sigma\sqrt{T_1}}.$$

The above formula is the well known Black futures formula.

The final variation we consider in this section is the compound option discussed in [10]. This is the case of an option on an option. We illustrate this type of options by a call on a call. We have seen that the payoff of a call maturing at time $T$ is $\max\{S(T)-K,\ 0\}$. The price of such a call at time $T_1 < T$ then is $\phi_c(T_1, S(T_1))$ given in (4.63). Then a compound call with strike price $K_1$ maturing at $T_1$ on a call with strike price $K$ and maturing at $T$, has payoff $\max\{\phi_c(T_1, S(T_1)) - K_1,\ 0\}$. The price of this compound call is

$$\phi_{co} = e^{-\delta T_1} E_Q\Big\{ \max\{\phi_c(T_1, S(T_1)) - K_1,\ 0\}\Big\}.$$

Let $x_0$ be the solution of $S(0)e^{(\delta-\frac{1}{2}\sigma^2)T_1+\sigma\sqrt{T_1}x_0}N(d_1(x_0)) - Ke^{-\delta(T-T_1)}N(d_2(x_0)) = K_1$, where $d_1(x_0) = d_2(x_0) + \sigma\sqrt{T-T_1}$ and $d_2(x_0) = \frac{\log(S(0)/K)+(\delta-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T_1}x_0}{\sigma\sqrt{T-T_1}}$.

Thus,

$$\phi_{co} = S(0)\int_{x_0}^{\infty} N(d_1(x+\sigma\sqrt{T_1}))n(x)dx - e^{-\delta T}K\int_{x_0}^{\infty} N(d_2(x))n(x)dx - e^{-\delta T_1}K_1 N(-x_0). \tag{5.28}$$

It is important to point out that we choose the money market account as the benchmarking security in this section because it is commonly used in practice and because it results in simple derivation of the results. However, any positive valued security can serve as the benchmarking security and there is a need to use other securities in some valuation problems. The choice of a benchmarking security largely depends on what security market we deal with. The money market account is often used for an equity market. For a bond market we may use either the money market account or a long term zero-coupon bond. In a swap market, a long term swap sometime is more appropriate.

## 5.2 Digital and Barrier Options

In this section, we apply the results in the previous section to exotic options. We focus on some digital(binary) options and European-type barrier options under the Black-Scholes framework. A digital option has a step payoff function, contingent on several random events. A simple example is the cash-or-nothing option which pays a fixed amount if its underlying security reaches a predetermined level at a preset time, otherwise nothing. Barrier options as their names suggested are those whose payoffs depend upon whether their underlying security hits a predetermined barrier or not. Typically, there is a predetermined value $H$ called barrier. When the value of the underlying security hits the barrier the status of a predesignated option changes. Their definition will become clear later. Barrier options are classified as IN options and OUT options. Under an IN option, the predesignated option starts when the value of the underlying security hits the barrier whilst under an OUT option, the predesignated option starts immediately but will expire when the value of the underlying security hits the barrier. In this section, we describe a down-and-in European call option and a up-and-out European call option in details. Other barrier options such as down-and-in put, down-and-out call, down-and-out put, up-and-out put, up-and-in call and up-and-in put are similar and can be understood easily from their names.

*Down-And-In European Call* Like a usual European call, this option has strike price $K$ and expiration time $T$. A barrier $H,\ H < S(0)$, is predetermined, where $S(t)$ is the value of the underlying security at time $t$. If $S(t)$ hits the barrier $H$ before $T$, the usual European call starts and the terminal payoff is $\max\{S(T) - K,\ 0\}$. Otherwise the value of terminal payoff is zero.

*Up-And-Out European Call* In this case, barrier $H > S(0)$, If $S(t)$ hits the barrier $H$ before $T$, the value of the terminal payoff is zero, Otherwise the value of the terminal payoff is $\max\{S(T) - K,\ 0\}$.

From the above description, the terminal payoff of a barrier option depends not only on the terminal value of the underlying security but its value in the past. Thus barrier options are path-dependent options.

We now turn to the valuation problem for digital and barrier options. Again, we assume a Black-Scholes economy characterised by (4.51) and (4.52). Under the unique risk-neutral probability measure $Q$ defined by (5.20)-(5.22), the value of the risky security $S(t)$ is a geometric Wiener process satisfying

$$dS = \delta S dt + \sigma S dW,$$

or

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

The price of an option is then the discounted value of its terminal payoff under the risk-neutral probability measure $Q$. We assume that $\delta - \frac{\sigma^2}{2} > 0$. The case $\delta - \frac{\sigma^2}{2} < 0$ can be dealt with accordingly. The case $\delta - \frac{\sigma^2}{2} = 0$ is the limiting case of either above case.

Three probability densities in (4.80) and (4.85) will play a very important role in valuation.

Let

1. $f_H(t)$, the density that represents the probability when the value of the security $S(t)$ hits the barrier $H$ at time $t$ for the first time.

2. $f_D(x)$, the density of $W(T)$ on the event that the value of the security $S(t)$ hits the barrier $H$, $H < S(0)$, before time $T$.

3. $f_U(x)$, the density of $W(T)$ on the event that the value of the security $S(t)$ hits the barrier $H$, $H > S(0)$, before time $T$.

Remark. The first density might be defective depending on the value of $H$. The last two densities are defective.

Define the first hitting time $\tau_H$ of $S(t)$ as

$$\tau_H = \inf\{t; \quad S(t) = H\}. \tag{5.29}$$

Then,

$$\tau_H = \inf\{t; \quad W(t) - bt = a\},$$

where

$$a = \frac{1}{\sigma}\log(H/S(0)), \; b = -\frac{1}{\sigma}(\delta - \frac{\sigma^2}{2}). \tag{5.30}$$

Then the density of $\tau_H$ is just a restatement of (4.80):

$$f_H(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}(t+a/b)^2}. \tag{5.31}$$

Next, we derive the densities $f_D(x)$ and $f_U(x)$.

Let us first consider $H < S(0)$. In this case, $a < 0$. (4.86) yields

**Lemma 5.3** The defective density function of $W(T)$ on the event that the path $S(t)$ hits the barrier $H, H < S(0)$, is

$$f_D(x) = \begin{cases} \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2T}x^2}, & x \le \frac{1}{\sigma}\log(H/S(0)) - \frac{\delta - \frac{\sigma^2}{2}}{\sigma}T \\ \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2T}[x - \frac{2}{\sigma}\log(H/S(0))]^2 + 2\frac{\delta - \frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))}, & x > \frac{1}{\sigma}\log(H/S(0)) - \frac{\delta - \frac{\sigma^2}{2}}{\sigma}T. \end{cases} \tag{5.32}$$

The derivation of $f_U(x)$ is very similar. In this case, $a > 0$. Thus, (4.85) yields

**Lemma 5.4** The defective density function of $W(T)$ on the event that the path $S(t)$ hits the barrier $H, H > S(0)$, is

$$f_U(x) = \begin{cases} \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2T}[x - \frac{2}{\sigma}\log(H/S(0))]^2 + 2\frac{\delta - \frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))}, & x < \frac{1}{\sigma}\log(H/S(0)) - \frac{\delta - \frac{\sigma^2}{2}}{\sigma}T \\ \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2T}x^2}, & x \ge \frac{1}{\sigma}\log(H/S(0)) - \frac{\delta - \frac{\sigma^2}{2}}{\sigma}T \end{cases} \tag{5.33}$$

We are now in the position to value some digital and barrier options. We begin with a digital option which pays a lump sum amount when the value of the underlying security hits a predetermined barrier. Then we evaluate the down-and-in call option and the up-and-out call option we have discussed. Other single barrier options can be evaluated in a similar manner( also see [23, 26]).

Suppose that a digital option pays amount $K$ at the time when the value of its underlying security $S(t)$ hits barrier $H > S(0)$ before time $T$. The present value of its payoff is $e^{-\delta\tau_H}K$. Thus, its price at time 0 is

$$\begin{aligned}
\phi_d &= E_Q\{e^{-\delta\tau_H}K\chi_{\{\tau_H\le T\}}\} = K\int_0^T e^{-\delta t}f_H(t)dt \\
&= \frac{aK}{\sqrt{2\pi}}\int_0^T t^{-\frac{3}{2}}e^{-\delta t}e^{-\frac{b^2}{2t}(t+a/b)^2}dt \\
&= \frac{aK}{\sqrt{2\pi}}e^{-a(b+\sqrt{b^2+2\delta})}\int_0^T t^{-\frac{3}{2}}e^{-\frac{b^2+2\delta}{2t}(t-\frac{a}{\sqrt{b^2+2\delta}})^2}dt \\
&= Ke^{-a(b+\sqrt{b^2+2\delta})}N(\frac{\sqrt{b^2+2\delta}T-a}{\sqrt{T}}) + Ke^{-a(b-\sqrt{b^2+2\delta})}N(-\frac{\sqrt{b^2+2\delta}T+a}{\sqrt{T}}).
\end{aligned}$$

Noting that $a = \frac{1}{\sigma}\log(H/S(0))$, $b = -\frac{1}{\sigma}(\delta - \frac{\sigma^2}{2})$, and $\sqrt{b^2+2\delta} = \frac{1}{\sigma}(\delta + \frac{\sigma^2}{2})$, we have

$$\phi_d = K[\frac{H}{S(0)}]^{-1}N(\tilde{d}_1) + K[\frac{H}{S(0)}]^{2\delta/\sigma^2}N(\tilde{d}_2), \tag{5.34}$$

where

$$\tilde{d}_1 = \frac{\log(S(0)/H) + (\delta + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}},\ \tilde{d}_2 = \frac{\log(S(0)/H) - (\delta + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

We now consider barrier options. For the down-and-in call, the terminal payoff is

$$X_{DI} = \begin{cases} \max\{S(T)-K, 0\}, & \text{if } \max_{0<t\le T} S(t) \le H \\ 0, & \text{if } \max_{0<t\le T} S(t) > H \end{cases} \tag{5.35}$$

Hence,

$$\begin{aligned}
E_Q\Big(X_{DI}\Big) &= \int_{-\infty}^{\infty} \max\{S(0)e^{(\delta-\frac{\sigma^2}{2})T+\sigma x} - K, 0\}f_D(x)dx \\
&= \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} \Big(S(0)e^{(\delta-\frac{\sigma^2}{2})T+\sigma x} - K\Big)f_D(x)dx
\end{aligned}$$

If $K \ge H$,

$$E_Q\Big(X_{DI}\Big) = S(0)\int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} e^{(\delta-\frac{\sigma^2}{2})T+\sigma x}\cdot\frac{1}{\sqrt{2\pi T}}e^{-\frac{1}{2T}[x-\frac{2}{\sigma}\log(H/S(0))]^2 + 2\frac{\delta-\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))}dx$$

$$
\begin{aligned}
&- \quad K \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2T}[x - \frac{2}{\sigma}\log(H/S(0))]^2 + 2\frac{\delta-\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} dx \\
&= \quad S(0)e^{\delta T + 2\frac{\delta+\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} \frac{1}{\sqrt{2\pi T}} \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} e^{-\frac{1}{2T}[x - (\frac{2}{\sigma}\log(H/S(0)) + \sigma T)]^2} dx \\
&- \quad K e^{2\frac{\delta-\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} \frac{1}{\sqrt{2\pi T}} \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} e^{-\frac{1}{2T}[x - \frac{2}{\sigma}\log(H/S(0))]^2} dx.
\end{aligned}
$$

A change of variables easily yields

$$
E_Q\Big(X_{DI}\Big) = S(0)e^{\delta T + 2\frac{\delta+\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} N(d_3) - K e^{2\frac{\delta-\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} N(d_4), \tag{5.36}
$$

where

$$
d_3 = \frac{\log\frac{H^2}{KS(0)} + (\delta + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \text{ and } d_4 = \frac{\log\frac{H^2}{KS(0)} + (\delta - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}. \tag{5.37}
$$

If $K < H$, the density $f_D(x)$ is piecewise. Write

$$
\int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty} = \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\frac{1}{\sigma}\log(H/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T} + \int_{\frac{1}{\sigma}\log(H/S(0)) - \frac{\delta-\frac{\sigma^2}{2}}{\sigma}T}^{\infty}.
$$

It is easy to see from above and the Black-Scholes formula that

$$
\begin{aligned}
E_Q\Big(X_{DI}\Big) \quad &= \quad S(0)e^{\delta T}\Big\{e^{2\frac{\delta+\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} N(d_3(H)) + N(d_1(K)) - N(d_1(H))\Big\} \\
&- \quad K\Big\{e^{2\frac{\delta-\frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))} N(d_4(H)) + N(d_2(K)) - N(d_2(H))\Big\},
\end{aligned} \tag{5.38}
$$

where $d_1(H)$ and $d_2(H)$ are given in the Black-Scholes formula with strike price $H$, and $d_3(H)$ and $d_4(H)$ are given in (5.38) with $K = H$. In summary, we have the following theorem

**Theorem 5.6** Let $\phi(X_{DI})$ be the price of the down-and-in call with stike price $K$, barrier $H$ and expiration time $T$.

If $K \geq H$, then

$$
\phi(X_{DI}) = S(0)[H/S(0)]^{2\delta/\sigma^2+1} N(d_3) - K e^{-\delta T}[H/S(0)]^{2\delta/\sigma^2-1} N(d_4), \tag{5.39}
$$

where $d_3,\ d_4$ are given in (5.37).

If $K < H$, then

$$\begin{aligned} \phi(X_{DI}) &= S(0)\Big\{[H/S(0)]^{2\delta/\sigma^2+1}N(d_3(H)) + N(d_1(K)) - N(d_1(H))\Big\} \\ - & Ke^{-\delta T}\Big\{[H/S(0)]^{2\delta/\sigma^2-1}N(d_4(H)) + N(d_2(K)) - N(d_2(H))\Big\}, \end{aligned} \tag{5.40}$$

where $d_1(H),\ d_2(H),\ d_3(H)$ and $d_4(H)$ are given in (5.38).

*Proof:* It follows immediately from

$$\phi(X_{DI}) = e^{-\delta T} E_Q\Big(X_{DI}\Big).$$

$\square$

To value the up-and-out option, we may use the put-call parity

$$\phi(X) = \phi(X_{UI}) + \phi(X_{UO}), \tag{5.41}$$

where $X_{UI}$ and $X_{UO}$ denote the terminal payoff of a up-and-in option and a up-and-out option, respectively, or calculate it directly using Lemma 5.5. We here show how to calculate it directly. We suppose that $K < H$, otherwise the value is null. To keep the notation simple, we use $a$ and $b$ instead for the moment. The defective density for the up-and-out option is $f_{UO}(x) = f(x) - f_U(x)$, where $f(x)$ is the density of a normal distribution with mean 0 and variance $T$. Thus,

$$f_{UO}(x) = \begin{cases} \frac{1}{\sqrt{2\pi T}}[e^{-\frac{1}{2T}x^2} - e^{-\frac{1}{2T}(x-2a)^2-2ab}], & x \le a + bT \\ 0, & x > a + bT. \end{cases} \tag{5.42}$$

$$\begin{aligned} E_Q\Big(X_{UO}\Big) &= \int_{-\infty}^{a+bT} \max\{S(T) - K, 0\} f_{UO}(x)dx \\ = & \int_{\frac{1}{\sigma}\log(K/S(0)) - \frac{\delta - \frac{\sigma^2}{2}}{\sigma}T}^{a+bT} [S(T) - K][e^{-\frac{1}{2T}x^2} - e^{-\frac{1}{2T}(x-2a)^2-2ab}]dx \\ = & \{S(0)e^{\delta T}[N(d_1(K)) - N(d_1(H))] - K[N(d_2(K)) - N(d_2(H))]\} \\ - & \{S(0)e^{\delta T + 2\frac{\delta + \frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))}[N(d_3) - N(d_3(H))] - Ke^{2\frac{\delta - \frac{\sigma^2}{2}}{\sigma^2}\log(H/S(0))}[N(d_4) - N(d_4(H))]\}. \end{aligned}$$

We then have

**Theorem 5.7** Let $\phi(X_{UO})$ be the price of the down-and-in call with stike price $K$, barrier $H,\ H > K$, and expiration time $T$. Then,

$$\begin{aligned} \phi(X_{UO}) &= S(0)\Big\{[N(d_1(K)) - N(d_1(H))] - [H/S(0)]^{2\delta/\sigma^2+1}[N(d_3) - N(d_3(H))]\Big\} \\ - \quad & Ke^{-\delta T}\Big\{[N(d_2(K)) - N(d_2(H))] - [H/S(0)]^{2\delta/\sigma^2-1}[N(d_4) - N(d_4(H))]\Big\}. \qquad (5.43) \end{aligned}$$

## 5.3 Interest Rate Models

## 5.4 Swaps and Swaptions

# Appendix A

# Probability Theory

Let $\Omega$ be a space and $\mathcal{F}$ is a collection of subsets of $\Omega$. $\mathcal{F}$ is said to be a $\sigma$-algebra on $\Omega$ if it satisfies

1. The empty set $\phi$ and the whole space $\Omega$ are in $\mathcal{F}$;
2. If $F_1, F_2, \cdots$ are in $\mathcal{F}$, then $\cup_{n=1}^{\infty} F_n$ is in $\mathcal{F}$;
3. If $F$ is in $\mathcal{F}$, its complement $F^c$ is in $\mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Let $X$ be a real-value function defined on $\Omega$. If for any $x \in \mathbf{R}$, the set $\{X \leq x\} \in \mathcal{F}$, $X$ is said to be measurable on $(\Omega, \mathcal{F})$.

Let $P : \mathcal{F} \rightarrow \mathbf{R}_+$ satisfy

1. $P(\phi) = 0, \quad P(\Omega) = 1$;
2. For any $F_1$ and $F_2$ with $F_1 \cap F_2 = \phi$, $P(F_1 \cup F_2) = P(F_1) + P(F_2)$;

$P$ is called a probability measure on the space $(\Omega, \mathcal{F})$ and the triplet $(\Omega, \mathcal{F}, P)$ is called a probability space.

A real-value function $X$ defined on $\Omega$ is a random variable if it is measurable on $(\Omega, \mathcal{F}, P)$. The expectation of a random variable $X$ is defined as

$$E(X) = \int_{\Omega} X dP.$$

Given two $\sigma$-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ on $\Omega$, we says $\mathcal{F}_1$ is finer than $\mathcal{F}_2$ if $\mathcal{F}_2 \subseteq \mathcal{F}_1$

Suppose that $\mathcal{F}_1$ is coarser than $\mathcal{F}$, the expectation of $X$ conditional on $\mathcal{F}_1$, $E(X|\mathcal{F}_1)$ is a random variable on $(\Omega, \mathcal{F}_1, P)$ such that for any $F \in \mathcal{F}_1$,

$$\int_F E(X|\mathcal{F}_1) dP = \int_F X dP.$$

Hence if $\mathcal{F}_1$ is finer than $\mathcal{F}_2$, $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_2)$.

Let $\mathcal{F}_t, 0 \leq t \leq T$ be a collection of increasing (finer and finer) $\sigma$-algebras which are coarser that $\mathcal{F}$. Then $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is called a filtered space and $\mathcal{F}_t, 0 \leq t \leq T$ is the filtration on the probability space $(\Omega, \mathcal{F}, P)$.

A collection of random variables $X(t)$, $0 \leq t \leq T$ is called a (adapted) stochastic process if each $X(t)$ is a random variable on $(\Omega, \mathcal{F}_t, P)$. A (adapted) stochastic process $X(t)$ is called a martingale with respect to $\mathcal{F}_t, 0 \leq t \leq T$ if for any $s > t$,

$$E\Big\{X(s) \ |\mathcal{F}_t\Big\} = X(t).$$

**Levy's Convergence Theorem** Let $F_n(x)$ be a sequence of distribution functions and $\tilde{f}_n(z)$ be the corresponding characteristic functions. If there is a distribution function $F(x)$ with its characteristic function $\tilde{f}(z)$ such that

1. $\tilde{f}(z)$ is continuous at $z = 0$;

2. $\lim_{n \to \infty} \tilde{f}_n(z) = \tilde{f}(z)$.

Then, $F_n(x)$ is weakly convergent to $F(x)$. Therefore,

$$\lim_{n \to \infty} \Pr\{a < X_n \leq b\} = \Pr\{a < X \leq b\},$$

where $X_n$ and $X$ has df $F_n(x)$ and $F(x)$.

**Central Limit Theorem** Let $X_n$ be a sequence of iid random variables with mean $\mu$ and standard deviation $\sigma$. Let

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}\sigma},$$

the normalised sum of $X_1, \cdots, X_n$. Then,

$$\lim_{n\to\infty} \Pr\{a < Y_n \leq b\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

# Appendix B

# Functional Analysis

A linear space $\mathcal{H}$ is a Hilbert space if (i) there is a symmetric bilinear map $(x, y) \to x \bullet y$ from $\mathcal{H} \times \mathcal{H}$ to $\mathbf{R}$ such that $x \bullet x \geq 0; x \bullet x = 0$ only if $x = 0$. The norm of each $x$ is defined as $\|x\| = \sqrt{x \bullet x}$ and the distance between $x$ and $y$ is $d(x, y) = \|x - y\|$; (ii) the space is complete under this distance( i.e. the limit of any Cauchy sequence in $\mathcal{H}$ still belongs to $\mathcal{H}$).

A linear functional on $\mathcal{H}$ is a linear map from $\mathcal{H}$ to $\mathbf{R}$. A linear functional $f(x)$ is said to be continuous if there is $L > 0$ such that $\|f(x)\| \leq L\|x\|$. A set $N$ in $\mathcal{H}$ is called a hyperplane if there exist a continuous functional $f(x)$ and a real number $h$ such that $N = \{x;\ f(x) = h\}$.

**Riesz Representation Theorem** Let $f(x)$ be a continuous linear functional. Then there is a unique $z \in \mathcal{H}$ such that

$$f(x) = x \bullet z, \text{ for any } x \in \mathcal{H}.$$

*An Application:* Let $\mathcal{H}$ be the set of all random variables whose second moment exist on a probablity space. Define $X \bullet Y = E(XY)$. Then $\mathcal{H}$ is a Hilbert space. Hence for any continuous linear functional $f(X)$, there is a unique random variable $Z$ whose second moment exists such that $f(X) = E(XZ)$.

**Hahn-Banach Theorem** Let $A$ and $B$ be two disjoint convex sets in a Hilbert space $\mathcal{H}$. Assume that there exist $a \in A$ and $b \in B$ such that $d(A, B) = ||a - b||$, where $d(A, B)$ is the distance between $A$ and $B$ defined by $d(A, B) = \inf\{||x - y||;\text{ forany } x \in A \text{ and } y \in B\}$. Then, there exists a $z \in \mathcal{H}$ and a scalar $h$ such that for any $x \in A,\ x \bullet z > h$, and for any $y \in B,\ y \bullet z < h$. In other words, the sets $A$ and $B$ are separated by a hyperplane.

**Proof**: We first show that for any $x \in A,\ \ (x - a) \bullet (b - a) \leq 0$, and for any $y \in B,\ \ (y - b) \bullet (a - b) \leq 0$. Let $0 < \lambda < 1$. then $x_\lambda = (1 - \lambda)a + \lambda x$ is in $A$ from the convexity of $A$. Thus,

$$||b - a||^2 \leq ||b - x_\lambda||^2 = ||b - a - \lambda(x - a)||^2 = ||b - a||^2 - 2\lambda(b - a) \bullet (x - a) + \lambda^2||x - a||^2.$$

This gives

$$0 \leq -2(b - a) \bullet (x - a) + \lambda||x - a||^2.$$

Let $\lambda \to 0$. We obtain the first assertion. The second assertion can be obtained similarly.

Now, let $z = a - b$. Then we have

$$x \bullet z \geq a \bullet z, \quad y \bullet z \leq b \bullet z,$$

for any $x \in A$ and $y \in B$. Since $||a - b|| > 0, a \bullet z > b \bullet z$. We may choose $h$ such that $a \bullet z > h > b \bullet z$, we prove the theorem. ▭

Remark. If A is compact(equivalently closed and bounded in a finite dimensional space) and B is closed, the distance condition in the above theorem is satisfied automatically.

To show that in Theorem 1.2 of Chapter 1, there exist $a \in A$ and $b \in \mathcal{B}(0, p)$ satisfying the condition in the Hahn-Banach Theorem, we define the set

$$A_1 = \{x \in A;\ x_0 + \cdots + x_J = \frac{1}{2}\}.$$

From the above remark, there are $a \in A_1$ and $b \in \mathcal{B}(0,p)$ such that $d(A_1, \mathcal{B}(0,p)) = \|a - b\|$. We will show it is true for $A$ as well. Actually, it suffices to show $d(A_1, \mathcal{B}(0,p)) \leq d(A, \mathcal{B}(0,p))$.

For any $x \in A$, let $b_x \in \mathcal{B}(0,p)$ such that

$$\|x - b_x\| = d(x, \mathcal{B}(0,p)).$$

Then for any $y \in \mathcal{B}(0,p)$, $(x - b_x)'(y - b_x) \leq 0$. Thus, $(x - b_x)'(y - b_x) = 0$, which implies $(x - b_x)'y = 0$, for any $y$. Choose $0 < \lambda \leq 1$ such that $\lambda x \in A_1$. We then have from the same argument that

$$\|\lambda x - b_\lambda\| = d(\lambda x, \mathcal{B}(0,p)),$$

and for any $y \in \mathcal{B}(0,p)$, $(\lambda x - b_\lambda)'y = 0$, for any $y$. Hence, $b_\lambda = \lambda b_x$. This implies that

$$d(\lambda x, \mathcal{B}(0,p)) = \lambda d(x, \mathcal{B}(0,p)) \leq d(x, \mathcal{B}(0,p)).$$

We complete our proof. □

# Bibliography

[1] Black F. and Scholes M.J. (1973). The pricing of options and corporate li abilities. *Journal of Political Economy,* 81, 637-654.

[2] Black F., Derman, E. and Toy, W. (1990). A one factor model of interest rates and its application to treasury bond options, *Financial Analysts Journal,* 46, 33-39.

[3] Boyle, P. and Lin, X. (1996). "Valuation of options on several risky assets when there are transactions costs," *Advances in Futures and Options Research*. (to appear)

[4] Boyle P.P., and Vorst, T. (1992). Option replication in discrete time with transaction costs. *Journal of Finance*, 47, 271-293.

[5] Cox J., Ross S.A. and Rubinstein M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7, 229-263.

[6] Dothan W.U. (1990). *Prices in Financial Markets*, Oxford University Press, Oxford.

[7] Duffie, D. (1992). *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton.

[8] Durrett, R. (1984). *Brownian Motion and Martingale in Analysis*, Wadsworth, Inc.

[9] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, John Wiley, New York.

[10] Geske, R. (1979). The valuation of compound options, *Journal of Financial Economics,* 7, 63-81.

[11] Gihman, I.I. and Skorohod, A.V. (1979). *The Theory of Stochastc processes III*, Springer-Verlag, New York.

[12] Harrison J.M. and Krep D.M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory,* 20, 381-408.

[13] Harrison J.M. and Pliska, S. (1981). Martingales and Stochastic integrals in the theory of continuous trading, *Stochastic Processes and Their Applications*, 11, 215-260.

[14] He, H. (1990). Convergence from discrete to continuous time contingent claim prices. *Review of Financial Studies,* 3, 523-546.

[15] Heath, D., Jarrow, R. and Morton, A. (1990). Bond pricing and the term structure of interest rates: A discrete time approximation, *Journal of Financial and Quantitative Analysis,* 25, 419-440.

[16] Ho, T.S.Y. and Lee, S.B. (1986). Term structure movements and pricing interest rate contingent claims, *Journal of Finance,* 41, 1011-1029.

[17] Hull J. (1993). *Options, Futures, and Other Derivative Securities*, 2nd ed. Pratice-Hall, Englewood Cliffs, N.J.

[18] Hull, J. and White, A. (1990). Pricing interest rate derivative securities, *Review of Financial Studies,* 3, 573-592.

[19] McKean, H.P. (1969). *Stochastc Integrals*, Academic Press.

[20] Merton R.C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science,* 4, 141-183.

[21] Oksendal, B. (1985). *Stochastc Differential Equations*, Springer-Verlag.

[22] Pedersen H.W., Shiu E.S.W. and Thorlacius A.E. (1989). Arbitrage-free pricing of interest-rate contingent claims. *Transactions of the Society of Actuaries,* 41, 231-265; Discussion 267-279.

[23] Rich, D.R. (1995). The mathematical foundations of barrier option pricing theory, Preprint.

[24] Ross S.A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory,* 13, 341-360.

[25] Royden, H. (1968). *Real Analysis*, Macmillan, New York.

[26] Rubinstein, M. and Reiner, E. (1991). Barrier options, Preprint.

[27] Samuelson, P.A. (1965). Rational theory of warrant prices, *Industrial Management Review,* 6, 13-31.

[28] Samuelson, P.A. and Merton R.C. (1969). A complete model of warrant pricing that maximizes utility, *Industrial Management Review,* 10, 17-46.

[29] Varadhan, S.R.S. (1980). *Diffusion Problems and Partial Differential Equations*, Springer-Verlag, New York.

[30] Vasicek, O.A. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics,* 5, 177-188.

[31] Webster, R. (1994). *Convexity*, Oxford University Press.

[32] Williams, D. (1994). *Probability with Martingales*, Cambridge University Press.