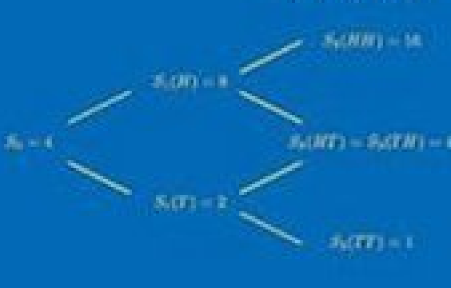
Springer Finance
Textbook

Steven E. Shreve

# Stochastic Calculus for Finance I
## The Binomial Asset Pricing Model




Springer

# Steven Shreve: Stochastic Calculus and Finance

PRASAD CHALASANI
Carnegie Mellon University
chal@cs.cmu.edu

SOMESH JHA
Carnegie Mellon University
sjha@cs.cmu.edu



July 25, 1997

# Contents

# Chapter 1

# Introduction to Probability Theory

## 1.1 The Binomial Asset Pricing Model

The *binomial asset pricing model* provides a powerful tool to understand arbitrage pricing theory and probability theory. In this course, we shall use it for both these purposes.

In the binomial asset pricing model, we model stock prices in discrete time, assuming that at each step, the stock price will change to one of two possible values. Let us begin with an initial positive stock price $S_0$. There are two positive numbers, $d$ and $u$, with

$$0 < d < u, \tag{1.1}$$

such that at the next period, the stock price will be either $dS_0$ or $uS_0$. Typically, we take $d$ and $u$ to satisfy $0 < d < 1 < u$, so change of the stock price from $S_0$ to $dS_0$ represents a *downward* movement, and change of the stock price from $S_0$ to $uS_0$ represents an *upward* movement. It is common to also have $d = \frac{1}{u}$, and this will be the case in many of our examples. However, strictly speaking, for what we are about to do we need to assume only (1.1) and (1.2) below.

Of course, stock price movements are much more complicated than indicated by the binomial asset pricing model. We consider this simple model for three reasons. First of all, within this model the concept of arbitrage pricing and its relation to risk-neutral pricing is clearly illuminated. Secondly, the model is used in practice because with a sufficient number of steps, it provides a good, computationally tractable approximation to continuous-time models. Thirdly, within the binomial model we can develop the theory of conditional expectations and martingales which lies at the heart of continuous-time models.

With this third motivation in mind, we develop notation for the binomial model which is a bit different from that normally found in practice. Let us imagine that we are tossing a coin, and when we get a "Head," the stock price moves up, but when we get a "Tail," the price moves down. We denote the price at time $1$ by $S_1(H) = uS_0$ if the toss results in head (H), and by $S_1(T) = dS_0$ if it

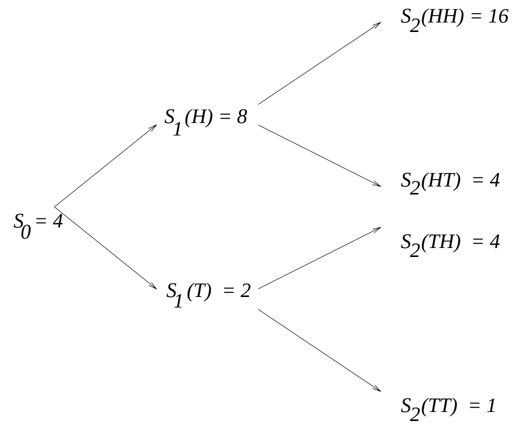



Figure 1.1: *Binomial tree of stock prices with* $S_0 = 4$, $u = 1/d = 2$.

results in tail (T). After the second toss, the price will be one of:

$$S_2(HH) = uS_1(H) = u^2 S_0, \quad S_2(HT) = dS_1(H) = duS_0,$$

$$S_2(TH) = uS_1(T) = udS_0, \quad S_2(TT) = dS_1(T) = d^2 S_0.$$

After three tosses, there are eight possible coin sequences, although not all of them result in different stock prices at time $3$.

For the moment, let us assume that the third toss is the last one and denote by

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

the set of all possible outcomes of the three tosses. The set $\Omega$ of all possible outcomes of a random experiment is called the *sample space* for the experiment, and the elements $\omega$ of $\Omega$ are called *sample points*. In this case, each sample point $\omega$ is a sequence of length three. We denote the $k$-th component of $\omega$ by $\omega_k$. For example, when $\omega = HTH$, we have $\omega_1 = H$, $\omega_2 = T$ and $\omega_3 = H$.

The stock price $S_k$ at time $k$ depends on the coin tosses. To emphasize this, we often write $S_k(\omega)$. Actually, this notation does not quite tell the whole story, for while $S_3$ depends on all of $\omega$, $S_2$ depends on only the first two components of $\omega$, $S_1$ depends on only the first component of $\omega$, and $S_0$ does not depend on $\omega$ at all. Sometimes we will use notation such $S_2(\omega_1, \omega_2)$ just to record more explicitly how $S_2$ depends on $\omega = (\omega_1, \omega_2, \omega_3)$.

**Example 1.1** Set $S_0 = 4$, $u = 2$ and $d = \frac{1}{2}$. We have then the binomial "tree" of possible stock prices shown in Fig. 1.1. Each sample point $\omega = (\omega_1, \omega_2, \omega_3)$ represents a path through the tree. Thus, we can think of the sample space $\Omega$ as either the set of all possible outcomes from three coin tosses or as the set of all possible paths through the tree.

To complete our binomial asset pricing model, we introduce a *money market* with *interest rate* $r$; \$1 invested in the money market becomes $\$(1+r)$ in the next period. We take $r$ to be the interest rate for both *borrowing* and *lending*. (This is not as ridiculous as it first seems, because in a many applications of the model, an agent is either borrowing or lending (not both) and knows in advance which she will be doing; in such an application, she should take $r$ to be the rate of interest for her activity.) We assume that

$$d < 1 + r < u. \tag{1.2}$$

The model would not make sense if we did not have this condition. For example, if $1+r \geq u$, then the rate of return on the money market is always at least as great as and sometimes greater than the return on the stock, and no one would invest in the stock. The inequality $d \geq 1+r$ cannot happen unless either $r$ is negative (which never happens, except maybe once upon a time in Switzerland) or $d \geq 1$. In the latter case, the stock does not really go "down" if we get a tail; it just goes up less than if we had gotten a head. One should borrow money at interest rate $r$ and invest in the stock, since even in the worst case, the stock price rises at least as fast as the debt used to buy it.

With the stock as the underlying asset, let us consider a *European call option* with strike price $K > 0$ and expiration time $1$. This option confers the right to buy the stock at time $1$ for $K$ dollars, and so is worth $S_1 - K$ at time $1$ if $S_1 - K$ is positive and is otherwise worth zero. We denote by

$$V_1(\omega) = (S_1(\omega) - K)^+ \stackrel{\Delta}{=} \max\{S_1(\omega) - K, 0\}$$

the value (payoff) of this option at expiration. Of course, $V_1(\omega)$ actually depends only on $\omega_1$, and we can and do sometimes write $V_1(\omega_1)$ rather than $V_1(\omega)$. Our first task is to compute the *arbitrage price* of this option at time zero.

Suppose at time zero you sell the call for $V_0$ dollars, where $V_0$ is still to be determined. You now have an obligation to pay off $(uS_0 - K)^+$ if $\omega_1 = H$ and to pay off $(dS_0 - K)^+$ if $\omega_1 = T$. At the time you sell the option, you don't yet know which value $\omega_1$ will take. You *hedge* your short position in the option by buying $\Delta_0$ shares of stock, where $\Delta_0$ is still to be determined. You can use the proceeds $V_0$ of the sale of the option for this purpose, and then borrow if necessary at interest rate $r$ to complete the purchase. If $V_0$ is more than necessary to buy the $\Delta_0$ shares of stock, you invest the residual money at interest rate $r$. In either case, you will have $V_0 - \Delta_0 S_0$ dollars invested in the money market, where this quantity might be negative. You will also own $\Delta_0$ shares of stock.

If the stock goes up, the value of your portfolio (excluding the short position in the option) is

$$\Delta_0 S_1(H) + (1+r)(V_0 - \Delta_0 S_0),$$

and you need to have $V_1(H)$. Thus, you want to choose $V_0$ and $\Delta_0$ so that

$$V_1(H) = \Delta_0 S_1(H) + (1+r)(V_0 - \Delta_0 S_0). \tag{1.3}$$

If the stock goes down, the value of your portfolio is

$$\Delta_0 S_1(T) + (1+r)(V_0 - \Delta_0 S_0),$$

and you need to have $V_1(T)$. Thus, you want to choose $V_0$ and $\Delta_0$ to also have

$$V_1(T) = \Delta_0 S_1(T) + (1+r)(V_0 - \Delta_0 S_0). \tag{1.4}$$

These are two equations in two unknowns, and we solve them below

Subtracting (1.4) from (1.3), we obtain

$$V_1(H) - V_1(T) = \Delta_0(S_1(H) - S_1(T)), \tag{1.5}$$

so that

$$\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)}. \tag{1.6}$$

This is a discrete-time version of the famous "delta-hedging" formula for derivative securities, according to which the number of shares of an underlying asset a hedge should hold is the derivative (in the sense of calculus) of the value of the derivative security with respect to the price of the underlying asset. This formula is so pervasive the when a practitioner says "delta", she means the derivative (in the sense of calculus) just described. Note, however, that my *definition* of $\Delta_0$ is the number of shares of stock one holds at time zero, and (1.6) is a consequence of this definition, not the definition of $\Delta_0$ itself. Depending on how uncertainty enters the model, there can be cases in which the number of shares of stock a hedge should hold is not the (calculus) derivative of the derivative security with respect to the price of the underlying asset.

To complete the solution of (1.3) and (1.4), we substitute (1.6) into either (1.3) or (1.4) and solve for $V_0$. After some simplification, this leads to the formula

$$V_0 = \frac{1}{1+r}\left[\frac{1+r-d}{u-d}V_1(H) + \frac{u-(1+r)}{u-d}V_1(T)\right]. \tag{1.7}$$

This is the *arbitrage price* for the European call option with payoff $V_1$ at time $1$. To simplify this formula, we define

$$\tilde{p} \triangleq \frac{1+r-d}{u-d}, \quad \tilde{q} \triangleq \frac{u-(1+r)}{u-d} = 1 - \tilde{p}, \tag{1.8}$$

so that (1.7) becomes

$$V_0 = \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)]. \tag{1.9}$$

Because we have taken $d < u$, both $\tilde{p}$ and $\tilde{q}$ are defined,i.e., the denominator in (1.8) is not zero. Because of (1.2), both $\tilde{p}$ and $\tilde{q}$ are in the interval $(0,1)$, and because they sum to $1$, we can regard them as probabilities of $H$ and $T$, respectively. They are the *risk-neutral* probabilites. They appeared when we solved the two equations (1.3) and (1.4), and have nothing to do with the actual probabilities of getting $H$ or $T$ on the coin tosses. In fact, at this point, they are nothing more than a convenient tool for writing (1.7) as (1.9).

We now consider a European call which pays off $K$ dollars at time $2$. At expiration, the payoff of this option is $V_2 \triangleq (S_2 - K)^+$, where $V_2$ and $S_2$ depend on $\omega_1$ and $\omega_2$, the first and second coin tosses. We want to determine the arbitrage price for this option at time zero. Suppose an agent sells the option at time zero for $V_0$ dollars, where $V_0$ is still to be determined. She then buys $\Delta_0$ shares of stock, investing $V_0 - \Delta_0 S_0$ dollars in the money market to finance this. At time $1$, the agent has a portfolio (excluding the short position in the option) valued at

$$X_1 \stackrel{\Delta}{=} \Delta_0 S_1 + (1+r)(V_0 - \Delta_0 S_0). \tag{1.10}$$

Although we do not indicate it in the notation, $S_1$ and therefore $X_1$ depend on $\omega_1$, the outcome of the first coin toss. Thus, there are really two equations implicit in (1.10):

$$\begin{aligned} X_1(H) &\stackrel{\Delta}{=} \Delta_0 S_1(H) + (1+r)(V_0 - \Delta_0 S_0), \\ X_1(T) &\stackrel{\Delta}{=} \Delta_0 S_1(T) + (1+r)(V_0 - \Delta_0 S_0). \end{aligned}$$

After the first coin toss, the agent has $X_1$ dollars and can readjust her hedge. Suppose she decides to now hold $\Delta_1$ shares of stock, where $\Delta_1$ is allowed to depend on $\omega_1$ because the agent knows what value $\omega_1$ has taken. She invests the remainder of her wealth, $X_1 - \Delta_1 S_1$ in the money market. In the next period, her wealth will be given by the right-hand side of the following equation, and she wants it to be $V_2$. Therefore, she wants to have

$$V_2 = \Delta_1 S_2 + (1+r)(X_1 - \Delta_1 S_1). \tag{1.11}$$

Although we do not indicate it in the notation, $S_2$ and $V_2$ depend on $\omega_1$ and $\omega_2$, the outcomes of the first two coin tosses. Considering all four possible outcomes, we can write (1.11) as four equations:

$$\begin{aligned} V_2(HH) &= \Delta_1(H)S_2(HH) + (1+r)(X_1(H) - \Delta_1(H)S_1(H)), \\ V_2(HT) &= \Delta_1(H)S_2(HT) + (1+r)(X_1(H) - \Delta_1(H)S_1(H)), \\ V_2(TH) &= \Delta_1(T)S_2(TH) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)), \\ V_2(TT) &= \Delta_1(T)S_2(TT) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)). \end{aligned}$$

We now have six equations, the two represented by (1.10) and the four represented by (1.11), in the six unknowns $V_0$, $\Delta_0$, $\Delta_1(H)$, $\Delta_1(T)$, $X_1(H)$, and $X_1(T)$.

To solve these equations, and thereby determine the arbitrage price $V_0$ at time zero of the option and the hedging portfolio $\Delta_0$, $\Delta_1(H)$ and $\Delta_1(T)$, we begin with the last two

$$\begin{aligned} V_2(TH) &= \Delta_1(T)S_2(TH) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)), \\ V_2(TT) &= \Delta_1(T)S_2(TT) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)). \end{aligned}$$

Subtracting one of these from the other and solving for $\Delta_1(T)$, we obtain the "delta-hedging formula"

$$\Delta_1(T) = \frac{V_2(TH) - V_2(TT)}{S_2(TH) - S_2(TT)}, \tag{1.12}$$

and substituting this into either equation, we can solve for

$$X_1(T) = \frac{1}{1+r}[\tilde{p}V_2(TH) + \tilde{q}V_2(TT)]. \tag{1.13}$$

Equation (1.13), gives the value the hedging portfolio should have at time $1$ if the stock goes down between times $0$ and $1$. We define this quantity to be the *arbitrage value of the option at time* $1$ *if* $\omega_1 = T$, and we denote it by $V_1(T)$. We have just shown that

$$V_1(T) \triangleq \frac{1}{1+r}[\tilde{p}V_2(TH) + \tilde{q}V_2(TT)]. \tag{1.14}$$

The hedger should choose her portfolio so that her wealth $X_1(T)$ if $\omega_1 = T$ agrees with $V_1(T)$ defined by (1.14). This formula is analgous to formula (1.9), but postponed by one step. The first two equations implicit in (1.11) lead in a similar way to the formulas

$$\Delta_1(H) = \frac{V_2(HH) - V_2(HT)}{S_2(HH) - S_2(HT)} \tag{1.15}$$

and $X_1(H) = V_1(H)$, where $V_1(H)$ is the value of the option at time $1$ if $\omega_1 = H$, defined by

$$V_1(H) \triangleq \frac{1}{1+r}[\tilde{p}V_2(HH) + \tilde{q}V_2(HT)]. \tag{1.16}$$

This is again analgous to formula (1.9), postponed by one step. Finally, we plug the values $X_1(H) = V_1(H)$ and $X_1(T) = V_1(T)$ into the two equations implicit in (1.10). The solution of these equations for $\Delta_0$ and $V_0$ is the same as the solution of (1.3) and (1.4), and results again in (1.6) and (1.9).

The pattern emerging here persists, regardless of the number of periods. If $V_k$ denotes the value at time $k$ of a derivative security, and this depends on the first $k$ coin tosses $\omega_1, \ldots, \omega_k$, then at time $k-1$, after the first $k-1$ tosses $\omega_1, \ldots, \omega_{k-1}$ are known, the portfolio to hedge a short position should hold $\Delta_{k-1}(\omega_1, \ldots, \omega_{k-1})$ shares of stock, where

$$\Delta_{k-1}(\omega_1, \ldots, \omega_{k-1}) = \frac{V_k(\omega_1, \ldots, \omega_{k-1}, H) - V_k(\omega_1, \ldots, \omega_{k-1}, T)}{S_k(\omega_1, \ldots, \omega_{k-1}, H) - S_k(\omega_1, \ldots, \omega_{k-1}, T)}, \tag{1.17}$$

and the value at time $k-1$ of the derivative security, when the first $k-1$ coin tosses result in the outcomes $\omega_1, \ldots, \omega_{k-1}$, is given by

$$V_{k-1}(\omega_1, \ldots, \omega_{k-1}) = \frac{1}{1+r}[\tilde{p}V_k(\omega_1, \ldots, \omega_{k-1}, H) + \tilde{q}V_k(\omega_1, \ldots, \omega_{k-1}, T)] \tag{1.18}$$

## 1.2 Finite Probability Spaces

Let $\Omega$ be a set with finitely many elements. An example to keep in mind is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \tag{2.1}$$

of all possible outcomes of three coin tosses. Let $\mathcal{F}$ be the set of all subsets of $\Omega$. Some sets in $\mathcal{F}$ are $\emptyset$, $\{HHH, HHT, HTH, HTT\}$, $\{TTT\}$, and $\Omega$ itself. How many sets are there in $\mathcal{F}$?

**Definition 1.1** A *probability measure* $I\!P$ is a function mapping $\mathcal{F}$ into $[0,1]$ with the following properties:

**(i)** $I\!P(\Omega) = 1$,

**(ii)** If $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then

$$I\!P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} I\!P(A_k).$$

Probability measures have the following interpretation. Let $A$ be a subset of $\mathcal{F}$. Imagine that $\Omega$ is the set of all possible outcomes of some random experiment. There is a certain probability, between $0$ and $1$, that when that experiment is performed, the outcome will lie in the set $A$. We think of $I\!P(A)$ as this probability.

**Example 1.2** Suppose a coin has probability $\frac{1}{3}$ for $H$ and $\frac{2}{3}$ for $T$. For the individual elements of $\Omega$ in (2.1), define

$$\begin{array}{ll}
I\!P\{HHH\} = \left(\frac{1}{3}\right)^3, & I\!P\{HHT\} = \left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right), \\
I\!P\{HTH\} = \left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right), & I\!P\{HTT\} = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2, \\
I\!P\{THH\} = \left(\frac{1}{3}\right)^2\left(\frac{1}{3}\right), & I\!P\{THT\} = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2, \\
I\!P\{TTH\} = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2, & I\!P\{TTT\} = \left(\frac{2}{3}\right)^3.
\end{array}$$

For $A \in \mathcal{F}$, we define

$$I\!P(A) = \sum_{\omega \in A} I\!P\{\omega\}. \tag{2.2}$$

For example,

$$I\!P\{HHH, HHT, HTH, HTT\} = \left(\frac{1}{3}\right)^3 + 2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{1}{3},$$

which is another way of saying that the probability of $H$ on the first toss is $\frac{1}{3}$.

As in the above example, it is generally the case that we specify a probability measure on only some of the subsets of $\Omega$ and then use property (ii) of Definition 1.1 to determine $I\!P(A)$ for the remaining sets $A \in \mathcal{F}$. In the above example, we specified the probability measure only for the sets containing a single element, and then used Definition 1.1(ii) in the form (2.2) (see Problem 1.4(ii)) to determine $I\!P$ for all the other sets in $\mathcal{F}$.

**Definition 1.2** Let $\Omega$ be a nonempty set. A $\sigma$-algebra is a collection $\mathcal{G}$ of subsets of $\Omega$ with the following three properties:

**(i)** $\emptyset \in \mathcal{G}$,

**(ii)** If $A \in \mathcal{G}$, then its complement $A^c \in \mathcal{G}$,

**(iii)** If $A_1, A_2, A_3, \ldots$ is a sequence of sets in $\mathcal{G}$, then $\cup_{k=1}^{\infty} A_k$ is also in $\mathcal{G}$.

Here are some important $\sigma$-algebras of subsets of the set $\Omega$ in Example 1.2:

$$\begin{aligned}
\mathcal{F}_0 &= \Big\{\emptyset, \Omega\Big\}, \\
\mathcal{F}_1 &= \Big\{\emptyset, \Omega, \{HHH, HHT, HTH, HTT\}, \{THH, THT, TTH, TTT\}\Big\}, \\
\mathcal{F}_2 &= \Big\{\emptyset, \Omega, \{HHH, HHT\}, \{HTH, HTT\}, \{THH, THT\}, \{TTH, TTT\}, \\
&\qquad\qquad \text{and all sets which can be built by taking unions of these}\Big\}, \\
\mathcal{F}_3 &= \mathcal{F} = \text{The set of all subsets of } \Omega.
\end{aligned}$$

To simplify notation a bit, let us define

$$\begin{aligned}
A_H &\triangleq \{HHH, HHT, HTH, HTT\} = \{H \text{ on the first toss}\}, \\
A_T &\triangleq \{THH, THT, TTH, TTT\} = \{T \text{ on the first toss}\},
\end{aligned}$$

so that

$$\mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_T\},$$

and let us define

$$\begin{aligned}
A_{HH} &\triangleq \{HHH, HHT\} = \{HH \text{ on the first two tosses}\}, \\
A_{HT} &\triangleq \{HTH, HTT\} = \{HT \text{ on the first two tosses}\}, \\
A_{TH} &\triangleq \{THH, THT\} = \{TH \text{ on the first two tosses}\}, \\
A_{TT} &\triangleq \{TTH, TTT\} = \{TT \text{ on the first two tosses}\},
\end{aligned}$$

so that

$$\begin{aligned}
\mathcal{F}_2 = \;& \{\emptyset, \Omega, A_{HH}, A_{HT}, A_{TH}, A_{TT}, \\
& A_H, A_T, A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT}, \\
& A_{HH}^c, A_{HT}^c, A_{TH}^c, A_{TT}^c\}.
\end{aligned}$$

We interpret $\sigma$-algebras as a record of information. Suppose the coin is tossed three times, and you are not told the outcome, but you are told, for every set in $\mathcal{F}_1$ whether or not the outcome is in that set. For example, you would be told that the outcome is not in $\emptyset$ and is in $\Omega$. Moreover, you might be told that the outcome is not in $A_H$ but is in $A_T$. In effect, you have been told that the first toss was a $T$, and nothing more. The $\sigma$-algebra $\mathcal{F}_1$ is said to contain the "information of the first toss", which is usually called the "information up to time $1$". Similarly, $\mathcal{F}_2$ contains the "information of the first two tosses," which is the "information up to time 2." The $\sigma$-algebra $\mathcal{F}_3 = \mathcal{F}$ contains "full information" about the outcome of all three tosses. The so-called "trivial" $\sigma$-algebra $\mathcal{F}_0$ contains no information. Knowing whether the outcome $\omega$ of the three tosses is in $\emptyset$ (it is not) and whether it is in $\Omega$ (it is) tells you nothing about $\omega$

**Definition 1.3** Let $\Omega$ be a nonempty finite set. A *filtration* is a sequence of $\sigma$-algebras $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ such that each $\sigma$-algebra in the sequence contains all the sets contained by the previous $\sigma$-algebra.

**Definition 1.4** Let $\Omega$ be a nonempty finite set and let $\mathcal{F}$ be the $\sigma$-algebra of all subsets of $\Omega$. A random variable is a function mapping $\Omega$ into $I\!R$.

**Example 1.3** Let $\Omega$ be given by (2.1) and consider the binomial asset pricing Example 1.1, where $S_0 = 4$, $u = 2$ and $d = \frac{1}{2}$. Then $S_0$, $S_1$, $S_2$ and $S_3$ are all random variables. For example, $S_2(HHT) = u^2 S_0 = 16$. The "random variable" $S_0$ is really not random, since $S_0(\omega) = 4$ for all $\omega \in \Omega$. Nonetheless, it is a function mapping $\Omega$ into $I\!R$, and thus technically a random variable, albeit a degenerate one.

A random variable maps $\Omega$ into $I\!R$, and we can look at the preimage under the random variable of sets in $I\!R$. Consider, for example, the random variable $S_2$ of Example 1.1. We have

$$\begin{aligned}
&S_2(HHH) = S_2(HHT) = 16,\\
&S_2(HTH) = S_2(HTT) = S_2(THH) = S_2(THT) = 4,\\
&S_2(TTH) = S_2(TTT) = 1.
\end{aligned}$$

Let us consider the interval $[4, 27]$. The preimage under $S_2$ of this interval is defined to be

$$\{\omega \in \Omega; S_2(\omega) \in [4, 27]\} = \{\omega \in \Omega; 4 \le S_2 \le 27\} = A^c_{TT}.$$

The complete list of subsets of $\Omega$ we can get as preimages of sets in $I\!R$ is:

$$\emptyset, \Omega, A_{HH}, A_{HT} \cup A_{TH}, A_{TT},$$

and sets which can be built by taking unions of these. This collection of sets is a $\sigma$-algebra, called the *$\sigma$-algebra generated by the random variable* $S_2$, and is denoted by $\sigma(S_2)$. The information content of this $\sigma$-algebra is exactly the information learned by observing $S_2$. More specifically, suppose the coin is tossed three times and you do not know the outcome $\omega$, but someone is willing to tell you, for each set in $\sigma(S_2)$, whether $\omega$ is in the set. You might be told, for example, that $\omega$ is not in $A_{HH}$, is in $A_{HT} \cup A_{TH}$, and is not in $A_{TT}$. Then you know that in the first two tosses, there was a head and a tail, and you know nothing more. This information is the same you would have gotten by being told that the value of $S_2(\omega)$ is $4$.

Note that $\mathcal{F}_2$ defined earlier contains all the sets which are in $\sigma(S_2)$, and even more. This means that the information in the first two tosses is greater than the information in $S_2$. In particular, if you see the first two tosses, you can distinguish $A_{HT}$ from $A_{TH}$, but you cannot make this distinction from knowing the value of $S_2$ alone.

**Definition 1.5** Let $\Omega$ be a nonemtpy finite set and let $\mathcal{F}$ be the $\sigma$-algebra of all subsets of $\Omega$. Let $X$ be a random variable on $(\Omega, \mathcal{F})$. The *$\sigma$-algebra $\sigma(X)$ generated by $X$* is defined to be the collection of all sets of the form $\{\omega \in \Omega; X(\omega) \in A\}$, where $A$ is a subset of $I\!R$. Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. We say that $X$ *is $\mathcal{G}$-measurable* if every set in $\sigma(X)$ is also in $\mathcal{G}$.

Note: We normally write simply $\{X \in A\}$ rather than $\{\omega \in \Omega; X(\omega) \in A\}$.

**Definition 1.6** Let $\Omega$ be a nonempty, finite set, let $\mathcal{F}$ be the $\sigma$-algebra of all subsets of $\Omega$, let $I\!P$ be a probabilty measure on $(\Omega, \mathcal{F})$, and let $X$ be a random variable on $\Omega$. Given any set $A \subseteq I\!R$, we define the *induced measure* of $A$ to be

$$\mathcal{L}_X(A) \triangleq I\!P\{X \in A\}.$$

In other words, the induced measure of a set $A$ tells us the probability that $X$ takes a value in $A$. In the case of $S_2$ above with the probability measure of Example 1.2, some sets in $I\!R$ and their induced measures are:

$$\begin{aligned}
&\mathcal{L}_{S_2}(\emptyset) = I\!P(\emptyset) = 0,\\
&\mathcal{L}_{S_2}(I\!R) = I\!P(\Omega) = 1,\\
&\mathcal{L}_{S_2}[0, \infty) = I\!P(\Omega) = 1,\\
&\mathcal{L}_{S_2}[0, 3] = I\!P\{S_2 = 1\} = I\!P(A_{TT}) = \left(\frac{2}{3}\right)^2.
\end{aligned}$$

In fact, the induced measure of $S_2$ places a mass of size $\left(\frac{1}{3}\right)^2 = \frac{1}{9}$ at the number $16$, a mass of size $\frac{4}{9}$ at the number $4$, and a mass of size $\left(\frac{2}{3}\right)^2 = \frac{4}{9}$ at the number $1$. A common way to record this information is to give the *cumulative distribution function* $F_{S_2}(x)$ of $S_2$, defined by

$$F_{S_2}(x) \triangleq I\!P(S_2 \leq x) = \begin{cases} 0, & \text{if } x < 1, \\ \frac{4}{9}, & \text{if } 1 \leq x < 4, \\ \frac{8}{9}, & \text{if } 4 \leq x < 16, \\ 1, & \text{if } 16 \leq x. \end{cases} \tag{2.3}$$

By the *distribution* of a random variable $X$, we mean any of the several ways of characterizing $\mathcal{L}_X$. If $X$ is discrete, as in the case of $S_2$ above, we can either tell where the masses are and how large they are, or tell what the cumulative distribution function is. (Later we will consider random variables $X$ which have densities, in which case the induced measure of a set $A \subseteq I\!R$ is the integral of the density over the set $A$.)

**Important Note.** In order to work through the concept of a risk-neutral measure, we set up the definitions to make a clear distinction between random variables and their distributions.

A *random variable* is a mapping from $\Omega$ to $I\!R$, nothing more. It has an existence quite apart from discussion of probabilities. For example, in the discussion above, $S_2(TTH) = S_2(TTT) = 1$, regardless of whether the probability for $H$ is $\frac{1}{3}$ or $\frac{1}{2}$.

The *distribution* of a random variable is a measure $\mathcal{L}_X$ on $I\!R$, i.e., a way of assigning probabilities to sets in $I\!R$. It depends on the random variable $X$ and the probability measure $I\!P$ we use in $\Omega$. If we set the probability of $H$ to be $\frac{1}{3}$, then $\mathcal{L}_{S_2}$ assigns mass $\frac{1}{9}$ to the number $16$. If we set the probability of $H$ to be $\frac{1}{2}$, then $\mathcal{L}_{S_2}$ assigns mass $\frac{1}{4}$ to the number $16$. The distribution of $S_2$ has changed, but the random variable has not. It is still defined by

$$\begin{aligned}
&S_2(HHH) = S_2(HHT) = 16,\\
&S_2(HTH) = S_2(HTT) = S_2(THH) = S_2(THT) = 4,\\
&S_2(TTH) = S_2(TTT) = 1.
\end{aligned}$$

Thus, a random variable can have more than one distribution (a "market" or "objective" distribution, and a "risk-neutral" distribution).

In a similar vein, two *different random variables* can have the *same distribution*. Suppose in the binomial model of Example 1.1, the probability of $H$ and the probability of $T$ is $\frac{1}{2}$. Consider a European call with strike price $14$ expiring at time $2$. The payoff of the call at time $2$ is the random variable $(S_2 - 14)^+$, which takes the value $2$ if $\omega = HHH$ or $\omega = HHT$, and takes the value $0$ in every other case. The probability the payoff is $2$ is $\frac{1}{4}$, and the probability it is zero is $\frac{3}{4}$. Consider also a European put with strike price $3$ expiring at time $2$. The payoff of the put at time $2$ is $(3 - S_2)^+$, which takes the value $2$ if $\omega = TTH$ or $\omega = TTT$. Like the payoff of the call, the payoff of the put is $2$ with probability $\frac{1}{4}$ and $0$ with probability $\frac{3}{4}$. The payoffs of the call and the put are different random variables having the same distribution.

**Definition 1.7** Let $\Omega$ be a nonempty, finite set, let $\mathcal{F}$ be the $\sigma$-algebra of all subsets of $\Omega$, let $I\!P$ be a probabilty measure on $(\Omega, \mathcal{F})$, and let $X$ be a random variable on $\Omega$. The *expected value* of $X$ is defined to be

$$I\!EX \stackrel{\Delta}{=} \sum_{\omega\in\Omega} X(\omega) I\!P\{\omega\}. \tag{2.4}$$

Notice that the expected value in (2.4) is defined to be a sum *over the sample space* $\Omega$. Since $\Omega$ is a finite set, $X$ can take only finitely many values, which we label $x_1, \ldots, x_n$. We can partition $\Omega$ into the subsets $\{X_1 = x_1\}, \ldots, \{X_n = x_n\}$, and then rewrite (2.4) as

$$\begin{aligned}
I\!EX &\stackrel{\Delta}{=} \sum_{\omega\in\Omega} X(\omega) I\!P\{\omega\}\\
&= \sum_{k=1}^{n} \sum_{\omega\in\{X_k = x_k\}} X(\omega) I\!P\{\omega\}\\
&= \sum_{k=1}^{n} x_k \sum_{\omega\in\{X_k = x_k\}} I\!P\{\omega\}\\
&= \sum_{k=1}^{n} x_k I\!P\{X_k = x_k\}\\
&= \sum_{k=1}^{n} x_k \mathcal{L}_X\{x_k\}.
\end{aligned}$$

Thus, although the expected value is defined as a sum over the sample space $\Omega$, we can also write it as a sum over $I\!R$.

To make the above set of equations absolutely clear, we consider $S_2$ with the distribution given by (2.3). The definition of $I\!\!ES_2$ is

$$\begin{aligned} I\!\!ES_2 &= S_2(HHH)I\!\!P\{HHH\} + S_2(HHT)I\!\!P\{HHT\} \\ &\quad + S_2(HTH)I\!\!P\{HTH\} + S_2(HTT)I\!\!P\{HTT\} \\ &\quad + S_2(THH)I\!\!P\{THH\} + S_2(THT)I\!\!P\{THT\} \\ &\quad + S_2(TTH)I\!\!P\{TTH\} + S_2(TTT)I\!\!P\{TTT\} \\ &= 16 \cdot I\!\!P(A_{HH}) + 4 \cdot I\!\!P(A_{HT} \cup A_{TH}) + 1 \cdot I\!\!P(A_{TT}) \\ &= 16 \cdot I\!\!P\{S_2 = 16\} + 4 \cdot I\!\!P\{S_2 = 4\} + 1 \cdot I\!\!P\{S_2 = 1\} \\ &= 16 \cdot \mathcal{L}_{S_2}\{16\} + 4 \cdot \mathcal{L}_{S_2}\{4\} + 1 \cdot \mathcal{L}_{S_2}\{1\} \\ &= 16 \cdot \frac{1}{9} + 4 \cdot \frac{4}{9} + 4 \cdot \frac{4}{9} \\ &= \frac{48}{9}. \end{aligned}$$

**Definition 1.8** Let $\Omega$ be a nonempty, finite set, let $\mathcal{F}$ be the $\sigma$-algebra of all subsets of $\Omega$, let $I\!\!P$ be a probabilty measure on $(\Omega, \mathcal{F})$, and let $X$ be a random variable on $\Omega$. The *variance* of $X$ is defined to be the expected value of $(X - I\!\!EX)^2$, i.e.,

$$\text{Var}(X) \triangleq \sum_{\omega \in \Omega} (X(\omega) - I\!\!EX)^2 I\!\!P\{\omega\}. \tag{2.5}$$

One again, we can rewrite (2.5) as a sum over $I\!R$ rather than over $\Omega$. Indeed, if $X$ takes the values $x_1, \ldots, x_n$, then

$$\text{Var}(X) = \sum_{k=1}^{n} (x_k - I\!\!EX)^2 I\!\!P\{X = x_k\} = \sum_{k=1}^{n} (x_k - I\!\!EX)^2 \mathcal{L}_X(x_k).$$

## 1.3 Lebesgue Measure and the Lebesgue Integral

In this section, we consider the set of real numbers $I\!R$, which is uncountably infinite. We define the *Lebesgue measure* of intervals in $I\!R$ to be their length. This definition and the properties of measure determine the Lebesgue measure of many, but not all, subsets of $I\!R$. The collection of subsets of $I\!R$ we consider, and for which Lebesgue measure is defined, is the collection of *Borel sets* defined below.

We use Lebesgue measure to construct the *Lebesgue integral*, a generalization of the Riemann integral. We need this integral because, unlike the Riemann integral, it can be defined on abstract spaces, such as the space of infinite sequences of coin tosses or the space of paths of Brownian motion. This section concerns the Lebesgue integral on the space $I\!R$ only; the generalization to other spaces will be given later.

**Definition 1.9** The *Borel $\sigma$-algebra*, denoted $\mathcal{B}(\mathbb{R})$, is the smallest $\sigma$-algebra containing all open intervals in $\mathbb{R}$. The sets in $\mathcal{B}(\mathbb{R})$ are called *Borel sets*.

Every set which can be written down and just about every set imaginable is in $\mathcal{B}(\mathbb{R})$. The following discussion of this fact uses the $\sigma$-algebra properties developed in Problem 1.3.

By definition, every open interval $(a, b)$ is in $\mathcal{B}(\mathbb{R})$, where $a$ and $b$ are real numbers. Since $\mathcal{B}(\mathbb{R})$ is a $\sigma$-algebra, every union of open intervals is also in $\mathcal{B}(\mathbb{R})$. For example, for every real number $a$, the *open half-line*

$$(a, \infty) = \bigcup_{n=1}^{\infty} (a, a + n)$$

is a Borel set, as is

$$(-\infty, a) = \bigcup_{n=1}^{\infty} (a - n, a).$$

For real numbers $a$ and $b$, the union

$$(-\infty, a) \cup (b, \infty)$$

is Borel. Since $\mathcal{B}(\mathbb{R})$ is a $\sigma$-algebra, every complement of a Borel set is Borel, so $\mathcal{B}(\mathbb{R})$ contains

$$[a, b] = \Big((-\infty, a) \cup (b, \infty)\Big)^c.$$

This shows that every closed interval is Borel. In addition, the *closed half-lines*

$$[a, \infty) = \bigcup_{n=1}^{\infty} [a, a + n]$$

and

$$(-\infty, a] = \bigcup_{n=1}^{\infty} [a - n, a]$$

are Borel. Half-open and half-closed intervals are also Borel, since they can be written as intersections of open half-lines and closed half-lines. For example,

$$(a, b] = (-\infty, b] \cap (a, \infty).$$

Every set which contains only one real number is Borel. Indeed, if $a$ is a real number, then

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n}\right).$$

This means that every set containing finitely many real numbers is Borel; if $A = \{a_1, a_2, \ldots, a_n\}$, then

$$A = \bigcup_{k=1}^{n} \{a_k\}.$$

In fact, every set containing countably infinitely many numbers is Borel; if $A = \{a_1, a_2, \ldots\}$, then

$$A = \bigcup_{k=1}^{n} \{a_k\}.$$

This means that the set of rational numbers is Borel, as is its complement, the set of irrational numbers.

There are, however, sets which are not Borel. We have just seen that any non-Borel set must have uncountably many points.

**Example 1.4** (The Cantor set.) *This example gives a hint of how complicated a Borel set can be. We use it later when we discuss the sample space for an infinite sequence of coin tosses.*

*Consider the unit interval* $[0,1]$*, and remove the middle half, i.e., remove the open interval*

$$A_1 \triangleq \left(\frac{1}{4}, \frac{3}{4}\right).$$

*The remaining set*

$$C_1 = \left[0, \frac{1}{4}\right] \cup \left[\frac{3}{4}, 1\right]$$

*has two pieces. From each of these pieces, remove the middle half, i.e., remove the open set*

$$A_2 \triangleq \left(\frac{1}{16}, \frac{3}{16}\right) \bigcup \left(\frac{13}{16}, \frac{15}{16}\right).$$

*The remaining set*

$$C_2 = \left[0, \frac{1}{16}\right] \bigcup \left[\frac{3}{16}, \frac{1}{4}\right] \bigcup \left[\frac{3}{4}, \frac{13}{16}\right] \bigcup \left[\frac{15}{16}, 1\right].$$

*has four pieces. Continue this process, so at stage* $k$*, the set* $C_k$ *has* $2^k$ *pieces, and each piece has length* $\frac{1}{4^k}$*. The* Cantor set

$$C \triangleq \bigcap_{k=1}^{\infty} C_k$$

*is defined to be the set of points not removed at any stage of this nonterminating process.*

*Note that the length of* $A_1$*, the first set removed, is* $\frac{1}{2}$*. The "length" of* $A_2$*, the second set removed, is* $\frac{1}{8} + \frac{1}{8} = \frac{1}{4}$*. The "length" of the next set removed is* $4 \cdot \frac{1}{32} = \frac{1}{8}$*, and in general, the length of the* $k$*-th set removed is* $2^{-k}$*. Thus, the total length removed is*

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = 1,$$

*and so the Cantor set, the set of points not removed, has zero "length."*

*Despite the fact that the Cantor set has no "length," there are lots of points in this set. In particular, none of the endpoints of the pieces of the sets* $C_1, C_2, \ldots$ *is ever removed. Thus, the points*

$$0, \frac{1}{4}, \frac{3}{4}, 1, \frac{1}{16}, \frac{3}{16}, \frac{13}{16}, \frac{15}{16}, \frac{1}{64}, \ldots$$

*are all in* $C$*. This is a countably infinite set of points. We shall see eventually that the Cantor set has uncountably many points.* $\diamond$

**Definition 1.10** Let $\mathcal{B}(\mathbb{R})$ be the $\sigma$-algebra of Borel subsets of $\mathbb{R}$. A *measure on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a function $\mu$ mapping $\mathcal{B}$ into $[0, \infty]$ with the following properties:

**(i)** $\mu(\emptyset) = 0$,

**(ii)** If $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{B}(\mathbb{R})$, then

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

*Lebesgue measure* is defined to be the measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ which assigns the measure of each interval to be its length. Following Williams's book, we denote Lebesgue measure by $\mu_0$.

A measure has all the properties of a probability measure given in Problem 1.4, except that the total measure of the space is not necessarily $1$ (in fact, $\mu_0(\mathbb{R}) = \infty$), one no longer has the equation

$$\mu(A^c) = 1 - \mu(A)$$

in Problem 1.4(iii), and property (v) in Problem 1.4 needs to be modified to say:

**(v)** If $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{B}(\mathbb{R})$ with $A_1 \supseteq A_2 \supseteq \cdots$ *and* $\mu(A_1) < \infty$, then

$$\mu\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} \mu(A_n).$$

To see that the additional requirment $\mu(A_1) < \infty$ is needed in (v), consider

$$A_1 = [1, \infty), A_2 = [2, \infty), A_3 = [3, \infty), \ldots.$$

Then $\cap_{k=1}^{\infty} A_k = \emptyset$, so $\mu_0(\cap_{k=1}^{\infty} A_k) = 0$, but $\lim_{n\to\infty} \mu_0(A_n) = \infty$.

We specify that the Lebesgue measure of each interval is its length, and that determines the Lebesgue measure of all other Borel sets. For example, the Lebesgue measure of the Cantor set in Example 1.4 must be zero, because of the "length" computation given at the end of that example.

The Lebesgue measure of a set containing only one point must be zero. In fact, since

$$\{a\} \subseteq \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$$

for every positive integer $n$, we must have

$$0 \le \mu_0\{a\} \le \mu_0\left(a - \frac{1}{n}, a + \frac{1}{n}\right) = \frac{2}{n}.$$

Letting $n \to \infty$, we obtain

$$\mu_0\{a\} = 0.$$

The Lebesgue measure of a set containing countably many points must also be zero. Indeed, if $A = \{a_1, a_2, \ldots\}$, then

$$\mu_0(A) = \sum_{k=1}^{\infty} \mu_0\{a_k\} = \sum_{k=1}^{\infty} 0 = 0.$$

The Lebesgue measure of a set containing uncountably many points can be either zero, positive and finite, or infinite. We may not compute the Lebesgue measure of an uncountable set by adding up the Lebesgue measure of its individual members, because there is no way to add up uncountably many numbers. The integral was invented to get around this problem.

In order to think about Lebesgue integrals, we must first consider the functions to be integrated.

**Definition 1.11** Let $f$ be a function from $\mathbb{R}$ to $\mathbb{R}$. We say that $f$ is *Borel-measurable* if the set $\{x \in \mathbb{R}; f(x) \in A\}$ is in $\mathcal{B}(\mathbb{R})$ whenever $A \in \mathcal{B}(\mathbb{R})$. In the language of Section 2, we want the *$\sigma$-algebra generated by* $f$ to be contained in $\mathcal{B}(\mathbb{R})$.

Definition 3.4 is purely technical and has nothing to do with keeping track of information. It is difficult to conceive of a function which is not Borel-measurable, and we shall pretend such functions don't exist. Hencefore, "function mapping $\mathbb{R}$ to $\mathbb{R}$" will mean "Borel-measurable function mapping $\mathbb{R}$ to $\mathbb{R}$" and "subset of $\mathbb{R}$" will mean "Borel subset of $\mathbb{R}$".

**Definition 1.12** An *indicator function* $g$ from $\mathbb{R}$ to $\mathbb{R}$ is a function which takes only the values $0$ and $1$. We call

$$A \triangleq \{x \in \mathbb{R}; g(x) = 1\}$$

the set *indicated* by $g$. We define the *Lebesgue integral* of $g$ to be

$$\int_{\mathbb{R}} g \, d\mu_0 \triangleq \mu_0(A).$$

A *simple function* $h$ from $\mathbb{R}$ to $\mathbb{R}$ is a linear combination of indicators, i.e., a function of the form

$$h(x) = \sum_{k=1}^{n} c_k g_k(x),$$

where each $g_k$ is of the form

$$g_k(x) = \begin{cases} 1, & \text{if } x \in A_k, \\ 0, & \text{if } x \notin A_k, \end{cases}$$

and each $c_k$ is a real number. We define the *Lebesgue integral* of $h$ to be

$$\int_R h \, d\mu_0 \triangleq \sum_{k=1}^{n} c_k \int_{\mathbb{R}} g_k d\mu_0 = \sum_{k=1}^{n} c_k \mu_0(A_k).$$

Let $f$ be a nonnegative function defined on $\mathbb{R}$, possibly taking the value $\infty$ at some points. We define the *Lebesgue integral* of $f$ to be

$$\int_{\mathbb{R}} f \, d\mu_0 \triangleq \sup \left\{ \int_{\mathbb{R}} h \, d\mu_0; h \text{ is simple and } h(x) \leq f(x) \text{ for every } x \in \mathbb{R} \right\}.$$

It is possible that this integral is infinite. If it is finite, we say that $f$ *is integrable*.

Finally, let $f$ be a function defined on $\mathbb{R}$, possibly taking the value $\infty$ at some points and the value $-\infty$ at other points. We define the *positive* and *negative parts* of $f$ to be

$$f^+(x) \triangleq \max\{f(x), 0\}, \ f^-(x) \triangleq \max\{-f(x), 0\},$$

respectively, and we define the *Lebesgue integral* of $f$ to be

$$\int_{\mathbb{R}} f\, d\mu_0 \triangleq \int_{\mathbb{R}} f^+\, d\mu_0 - - \int_{\mathbb{R}} f^-\, d\mu_0,$$

provided the right-hand side is not of the form $\infty - \infty$. If both $\int_{\mathbb{R}} f^+\, d\mu_0$ and $\int_{\mathbb{R}} f^-\, d\mu_0$ are finite (or equivalently, $\int_{\mathbb{R}} |f|\, d\mu_0 < \infty$, since $|f| = f^+ + f^-$), we say that $f$ is *integrable*.

Let $f$ be a function defined on $\mathbb{R}$, possibly taking the value $\infty$ at some points and the value $-\infty$ at other points. Let $A$ be a subset of $\mathbb{R}$. We define

$$\int_A f\, d\mu_0 \triangleq \int_{\mathbb{R}} \mathbb{I}_A f\, d\mu_0,$$

where

$$\mathbb{I}_A(x) \triangleq \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A, \end{cases}$$

is the *indicator function of* $A$.

The Lebesgue integral just defined is related to the Riemann integral in one very important way: if the Riemann integral $\int_a^b f(x)dx$ is defined, then the Lebesgue integral $\int_{[a,b]} f\, d\mu_0$ agrees with the Riemann integral. The Lebesgue integral has two important advantages over the Riemann integral. The first is that the Lebesgue integral is defined for more functions, as we show in the following examples.

**Example 1.5** Let $Q$ be the set of rational numbers in $[0,1]$, and consider $f \triangleq \mathbb{I}_Q$. Being a countable set, $Q$ has Lebesgue measure zero, and so the Lebesgue integral of $f$ over $[0,1]$ is

$$\int_{[0,1]} f\, d\mu_0 = 0.$$

To compute the Riemann integral $\int_0^1 f(x)dx$, we choose partition points $0 = x_0 < x_1 < \cdots < x_n = 1$ and divide the interval $[0,1]$ into subintervals $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$. In each subinterval $[x_{k-1}, x_k]$ there is a rational point $q_k$, where $f(q_k) = 1$, and there is also an irrational point $r_k$, where $f(r_k) = 0$. We approximate the Riemann integral from above by the *upper sum*

$$\sum_{k=1}^n f(q_k)(x_k - x_{k-1}) = \sum_{k=1}^n 1 \cdot (x_k - x_{k-1}) = 1,$$

and we also approximate it from below by the *lower sum*

$$\sum_{k=1}^n f(r_k)(x_k - x_{k-1}) = \sum_{k=1}^n 0 \cdot (x_k - x_{k-1}) = 0.$$

No matter how fine we take the partition of $[0, 1]$, the upper sum is always $1$ and the lower sum is always $0$. Since these two do not converge to a common value as the partition becomes finer, the Riemann integral is not defined. $\diamond$

**Example 1.6** Consider the function

$$f(x) \triangleq \begin{cases} \infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

This is not a simple function because simple function cannot take the value $\infty$. Every simple function which lies between $0$ and $f$ is of the form

$$h(x) \triangleq \begin{cases} y, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0, \end{cases}$$

for some $y \in [0, \infty)$, and thus has Lebesgue integral

$$\int_{\mathbb{R}} h \, d\mu_0 = y\mu_0\{0\} = 0.$$

It follows that

$$\int_{\mathbb{R}} f \, d\mu_0 = \sup \left\{ \int_{\mathbb{R}} h \, d\mu_0; h \text{ is simple and } h(x) \leq f(x) \text{ for every } x \in \mathbb{R} \right\} = 0.$$

Now consider the Riemann integral $\int_{-\infty}^{\infty} f(x) \, dx$, which for this function $f$ is the same as the Riemann integral $\int_{-1}^{1} f(x) \, dx$. When we partition $[-1, 1]$ into subintervals, one of these will contain the point $0$, and when we compute the upper approximating sum for $\int_{-1}^{1} f(x) \, dx$, this point will contribute $\infty$ times the length of the subinterval containing it. Thus the upper approximating sum is $\infty$. On the other hand, the lower approximating sum is $0$, and again the Riemann integral does not exist. $\diamond$

The Lebesgue integral has all *linearity* and *comparison* properties one would expect of an integral. In particular, for any two functions $f$ and $g$ and any real constant $c$,

$$\begin{aligned} \int_{\mathbb{R}} (f + g) \, d\mu_0 &= \int_{\mathbb{R}} f \, d\mu_0 + \int_{\mathbb{R}} g \, d\mu_0, \\ \int_{\mathbb{R}} cf \, d\mu_0 &= c \int_{\mathbb{R}} f \, d\mu_0, \end{aligned}$$

and whenever $f(x) \leq g(x)$ for all $x \in \mathbb{R}$, we have

$$\int_{\mathbb{R}} f \, d\mu_0 \leq \int_{\mathbb{R}} gd \, d\mu_0.$$

Finally, if $A$ and $B$ are disjoint sets, then

$$\int_{A \cup B} f \, d\mu_0 = \int_A f \, d\mu_0 + \int_B f \, d\mu_0.$$

There are three *convergence theorems* satisfied by the Lebesgue integral. In each of these the situation is that there is a sequence of functions $f_n, n = 1, 2, \ldots$ converging *pointwise* to a limiting function $f$. *Pointwise convergence* just means that

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ for every } x \in \mathbb{R}.$$

There are no such theorems for the Riemann integral, because the Riemann integral of the limiting function $f$ is too often not defined. Before we state the theorems, we given two examples of pointwise convergence which arise in probability theory.

**Example 1.7** Consider a sequence of normal densities, each with variance $1$ and the $n$-th having mean $n$:

$$f_n(x) \triangleq \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-n)^2}{2}}.$$

These converge pointwise to the function

$$f(x) = 0 \text{ for every } x \in \mathbb{R}.$$

We have $\int_{\mathbb{R}} f_n d\mu_0 = 1$ for every $n$, so $\lim_{n\to\infty} \int_{\mathbb{R}} f_n d\mu_0 = 1$, but $\int_{\mathbb{R}} f\, d\mu_0 = 0$. $\diamond$

**Example 1.8** Consider a sequence of normal densities, each with mean $0$ and the $n$-th having variance $\frac{1}{n}$:

$$f_n(x) \triangleq \sqrt{\frac{n}{2\pi}}\, e^{-\frac{x^2}{2n}}.$$

These converge pointwise to the function

$$f(x) \triangleq \begin{cases} \infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

We have again $\int_{\mathbb{R}} f_n d\mu_0 = 1$ for every $n$, so $\lim_{n\to\infty} \int_{\mathbb{R}} f_n d\mu_0 = 1$, but $\int_{\mathbb{R}} f\, d\mu_0 = 0$. The function $f$ is not the Dirac delta; the Lebesgue integral of this function was already seen in Example 1.6 to be zero. $\diamond$

**Theorem 3.1** (Fatou's Lemma) *Let $f_n, n = 1, 2, \ldots$ be a sequence of nonnegative functions converging pointwise to a function $f$. Then*

$$\int_{\mathbb{R}} f\, d\mu_0 \le \liminf_{n\to\infty} \int_{\mathbb{R}} f_n\, d\mu_0.$$

If $\lim_{n\to\infty} \int_{\mathbb{R}} f_n\, d\mu_0$ is defined, then Fatou's Lemma has the simpler conclusion

$$\int_{\mathbb{R}} f\, d\mu_0 \le \lim_{n\to\infty} \int_{\mathbb{R}} f_n\, d\mu_0.$$

This is the case in Examples 1.7 and 1.8, where

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n\, d\mu_0 = 1,$$

while $\int_{\mathbb{R}} f \, d\mu_0 = 0$. We could modify either Example 1.7 or 1.8 by setting $g_n = f_n$ if $n$ is even, but $g_n = 2f_n$ if $n$ is odd. Now $\int_{\mathbb{R}} g_n \, d\mu_0 = 1$ if $n$ is even, but $\int_{\mathbb{R}} g_n \, d\mu_0 = 2$ if $n$ is odd. The sequence $\{\int_{\mathbb{R}} g_n \, d\mu_0\}_{n=1}^{\infty}$ has two cluster points, $1$ and $2$. By definition, the smaller one, $1$, is $\liminf_{n\to\infty} \int_{\mathbb{R}} g_n \, d\mu_0$ and the larger one, $2$, is $\limsup_{n\to\infty} \int_{\mathbb{R}} g_n \, d\mu_0$. Fatou's Lemma guarantees that even the smaller cluster point will be greater than or equal to the integral of the limiting function.

The key assumption in Fatou's Lemma is that all the functions take only nonnegative values. Fatou's Lemma does not assume much but it is is not very satisfying because it does not conclude that

$$\int_{\mathbb{R}} f \, d\mu_0 = \lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu_0.$$

There are two sets of assumptions which permit this stronger conclusion.

**Theorem 3.2** (Monotone Convergence Theorem) *Let* $f_n, n = 1, 2, \ldots$ *be a sequence of functions converging pointwise to a function* $f$*. Assume that*

$$0 \leq f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots \text{ for every } x \in \mathbb{R}.$$

*Then*

$$\int_{\mathbb{R}} f \, d\mu_0 = \lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu_0,$$

*where both sides are allowed to be* $\infty$*.*

**Theorem 3.3** (Dominated Convergence Theorem) *Let* $f_n, n = 1, 2, \ldots$ *be a sequence of functions, which may take either positive or negative values, converging pointwise to a function* $f$*. Assume that there is a nonnegative integrable function* $g$ *(i.e.,* $\int_{\mathbb{R}} g \, d\mu_0 < \infty$*) such that*

$$|f_n(x)| \leq g(x) \text{ for every } x \in \mathbb{R} \text{ for every } n.$$

*Then*

$$\int_{\mathbb{R}} f \, d\mu_0 = \lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu_0,$$

*and both sides will be finite.*

## 1.4 General Probability Spaces

**Definition 1.13** A *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ consists of three objects:

**(i)** $\Omega$, a nonempty set, called the *sample space*, which contains all possible outcomes of some random experiment;

**(ii)** $\mathcal{F}$, a $\sigma$-algebra of subsets of $\Omega$;

**(iii)** $\mathbb{P}$, a probability measure on $(\Omega, \mathcal{F})$, i.e., a function which assigns to each set $A \in \mathcal{F}$ a number $\mathbb{P}(A) \in [0, 1]$, which represents the probability that the outcome of the random experiment lies in the set $A$.

**Remark 1.1** We recall from Homework Problem 1.4 that a probability measure $I\!P$ has the following properties:

**(a)** $I\!P(\emptyset) = 0$.

**(b)** (Countable additivity) If $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then

$$I\!P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} I\!P(A_k).$$

**(c)** (Finite additivity) If $n$ is a positive integer and $A_1, \ldots, A_n$ are disjoint sets in $\mathcal{F}$, then

$$I\!P(A_1 \cup \cdots \cup A_n) = I\!P(A_1) + \cdots + I\!P(A_n).$$

**(d)** If $A$ and $B$ are sets in $\mathcal{F}$ and $A \subseteq B$, then

$$I\!P(B) = I\!P(A) + I\!P(B \setminus A).$$

In particular,

$$I\!P(B) \geq I\!P(A).$$

**(d)** (Continuity from below.) If $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{F}$ with $A_1 \subseteq A_2 \subseteq \cdots$, then

$$I\!P\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} I\!P(A_n).$$

**(d)** (Continuity from above.) If $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{F}$ with $A_1 \supseteq A_2 \supseteq \cdots$, then

$$I\!P\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} I\!P(A_n).$$

We have already seen some examples of finite probability spaces. We repeat these and give some examples of infinite probability spaces as well.

**Example 1.9** Finite coin toss space.
Toss a coin $n$ times, so that $\Omega$ is the set of all sequences of $H$ and $T$ which have $n$ components. We will use this space quite a bit, and so give it a name: $\Omega_n$. Let $\mathcal{F}$ be the collection of all subsets of $\Omega_n$. Suppose the probability of $H$ on each toss is $p$, a number between zero and one. Then the probability of $T$ is $q \stackrel{\Delta}{=} 1 - p$. For each $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ in $\Omega_n$, we define

$$I\!P\{\omega\} \stackrel{\Delta}{=} p^{Number\ of\ H\ in\ \omega} \cdot q^{Number\ of\ T\ in\ \omega}.$$

For each $A \in \mathcal{F}$, we define

$$I\!P(A) \stackrel{\Delta}{=} \sum_{\omega\in A} I\!P\{\omega\}. \tag{4.1}$$

We can define $I\!P(A)$ this way because $A$ has only finitely many elements, and so only finitely many terms appear in the sum on the right-hand side of (4.1). $\diamond$

**Example 1.10** Infinite coin toss space.
Toss a coin repeatedly without stopping, so that $\Omega$ is the set of all nonterminating sequences of $H$ and $T$. We call this space $\Omega_\infty$. This is an uncountably infinite space, and we need to exercise some care in the construction of the $\sigma$-algebra we will use here.

For each positive integer $n$, we define $\mathcal{F}_n$ to be the $\sigma$-algebra determined by the first $n$ tosses. For example, $\mathcal{F}_2$ contains four basic sets,

$$\begin{aligned}
A_{HH} &\stackrel{\Delta}{=} \{\omega = (\omega_1, \omega_2, \omega_3, \ldots); \omega_1 = H, \omega_2 = H\} \\
&= \text{The set of all sequences which begin with } HH, \\
A_{HT} &\stackrel{\Delta}{=} \{\omega = (\omega_1, \omega_2, \omega_3, \ldots); \omega_1 = H, \omega_2 = T\} \\
&= \text{The set of all sequences which begin with } HT, \\
A_{TH} &\stackrel{\Delta}{=} \{\omega = (\omega_1, \omega_2, \omega_3, \ldots); \omega_1 = T, \omega_2 = H\} \\
&= \text{The set of all sequences which begin with } TH, \\
A_{TT} &\stackrel{\Delta}{=} \{\omega = (\omega_1, \omega_2, \omega_3, \ldots); \omega_1 = T, \omega_2 = T\} \\
&= \text{The set of all sequences which begin with } TT.
\end{aligned}$$

Because $\mathcal{F}_2$ is a $\sigma$-algebra, we must also put into it the sets $\emptyset$, $\Omega$, and all unions of the four basic sets.

In the $\sigma$-algebra $\mathcal{F}$, we put every set in every $\sigma$-algebra $\mathcal{F}_n$, where $n$ ranges over the positive integers. We also put in every other set which is required to make $\mathcal{F}$ be a $\sigma$-algebra. For example, the set containing the single sequence

$$\{HHHHH\cdots\} = \{H \text{ on every toss}\}$$

is not in any of the $\mathcal{F}_n$ $\sigma$-algebras, because it depends on all the components of the sequence and not just the first $n$ components. However, for each positive integer $n$, the set

$$\{H \text{ on the first } n \text{ tosses}\}$$

is in $\mathcal{F}_n$ and hence in $\mathcal{F}$. Therefore,

$$\{H \text{ on every toss}\} = \bigcap_{n=1}^{\infty} \{H \text{ on the first } n \text{ tosses}\}$$

is also in $\mathcal{F}$.

We next construct the probability measure $I\!P$ on $(\Omega_\infty, \mathcal{F})$ which corresponds to probability $p \in [0, 1]$ for $H$ and probability $q = 1 - p$ for $T$. Let $A \in \mathcal{F}$ be given. If there is a positive integer $n$ such that $A \in \mathcal{F}_n$, then the description of $A$ depends on only the first $n$ tosses, and it is clear how to define $I\!P(A)$. For example, suppose $A = A_{HH} \cup A_{TH}$, where these sets were defined earlier. Then $A$ is in $\mathcal{F}_2$. We set $I\!P(A_{HH}) = p^2$ and $I\!P(A_{TH}) = qp$, and then we have

$$I\!P(A) = I\!P(A_{HH} \cup A_{TH}) = p^2 + qp = (p + q)p = p.$$

In other words, the probability of a $H$ on the second toss is $p$.

Let us now consider a set $A \in \mathcal{F}$ for which there is no positive integer $n$ such that $A \in \mathcal{F}$. Such is the case for the set $\{H$ on every toss$\}$. To determine the probability of these sets, we write them in terms of sets which are in $\mathcal{F}_n$ for positive integers $n$, and then use the properties of probability measures listed in Remark 1.1. For example,

$$\begin{aligned}\{H \text{ on the first toss}\} &\supseteq \{H \text{ on the first two tosses}\}\\ &\supseteq \{H \text{ on the first three tosses}\}\\ &\supseteq \cdots,\end{aligned}$$

and

$$\bigcap_{n=1}^{\infty}\{H \text{ on the first } n \text{ tosses}\} = \{H \text{ on every toss}\}.$$

According to Remark 1.1(d) (continuity from above),

$$I\!\!P\{H \text{ on every toss}\} = \lim_{n\to\infty} I\!\!P\{H \text{ on the first } n \text{ tosses}\} = \lim_{n\to\infty} p^n.$$

If $p = 1$, then $I\!\!P\{H$ on every toss$\} = 1$; otherwise, $I\!\!P\{H$ on every toss$\} = 0$.

A similar argument shows that if $0 < p < 1$ so that $0 < q < 1$, then every set in $\Omega_\infty$ which contains only one element (nonterminating sequence of $H$ and $T$) has probability zero, and hence very set which contains countably many elements also has probabiliy zero. We are in a case very similar to Lebesgue measure: every point has measure zero, but sets can have positive measure. Of course, the only sets which can have positive probabilty in $\Omega_\infty$ are those which contain uncountably many elements.

In the infinite coin toss space, we define a sequence of random variables $Y_1, Y_2, \ldots$ by

$$Y_k(\omega) \triangleq \begin{cases} 1 & \text{if } \omega_k = H,\\ 0 & \text{if } \omega_k = T,\end{cases}$$

and we also define the random variable

$$X(\omega) = \sum_{k=1}^{n} \frac{Y_k(\omega)}{2^k}.$$

Since each $Y_k$ is either zero or one, $X$ takes values in the interval $[0,1]$. Indeed, $X(TTTT\cdots) = 0$, $X(HHHH\cdots) = 1$ and the other values of $X$ lie in between. We define a "dyadic rational number" to be a number of the form $\frac{m}{2^k}$, where $k$ and $m$ are integers. For example, $\frac{3}{4}$ is a dyadic rational. Every dyadic rational in (0,1) corresponds to two sequences $\omega \in \Omega_\infty$. For example,

$$X(HHTTTTT\cdots) = X(HTHHHHH\cdots) = \frac{3}{4}.$$

The numbers in (0,1) which are not dyadic rationals correspond to a single $\omega \in \Omega_\infty$; these numbers have a unique binary expansion.

Whenever we place a probability measure $I\!P$ on $(\Omega, \mathcal{F})$, we have a corresponding induced measure $\mathcal{L}_X$ on $[0,1]$. For example, if we set $p = q = \frac{1}{2}$ in the construction of this example, then we have

$$\mathcal{L}_X\left[0, \frac{1}{2}\right] = I\!P\{\text{First toss is } T\} = \frac{1}{2},$$
$$\mathcal{L}_X\left[\frac{1}{2}, 1\right] = I\!P\{\text{First toss is } H\} = \frac{1}{2},$$
$$\mathcal{L}_X\left[0, \frac{1}{4}\right] = I\!P\{\text{First two tosses are } TT\} = \frac{1}{4},$$
$$\mathcal{L}_X\left[\frac{1}{4}, \frac{1}{2}\right] = I\!P\{\text{First two tosses are } TH\} = \frac{1}{4},$$
$$\mathcal{L}_X\left[\frac{1}{2}, \frac{3}{4}\right] = I\!P\{\text{First two tosses are } HT\} = \frac{1}{4},$$
$$\mathcal{L}_X\left[\frac{3}{4}, 1\right] = I\!P\{\text{First two tosses are } HH\} = \frac{1}{4}.$$

Continuing this process, we can verify that for any positive integers $k$ and $m$ satisfying

$$0 \leq \frac{m-1}{2^k} < \frac{m}{2^k} \leq 1,$$

we have

$$\mathcal{L}_X\left[\frac{m-1}{2^k}, \frac{m}{2^k}\right] = \frac{1}{2^k}.$$

In other words, the $\mathcal{L}_X$-measure of all intervals in $[0,1]$ whose endpoints are dyadic rationals is the same as the Lebesgue measure of these intervals. The only way this can be is for $\mathcal{L}_X$ to be Lebesgue measure.

It is interesing to consider what $\mathcal{L}_X$ would look like if we take a value of $p$ other than $\frac{1}{2}$ when we construct the probability measure $I\!P$ on $\Omega$.

We conclude this example with another look at the Cantor set of Example 3.2. Let $\Omega_{pairs}$ be the subset of $\Omega$ in which every even-numbered toss is the same as the odd-numbered toss immediately preceding it. For example, $HHTTTTHH$ is the beginning of a sequence in $\Omega_{pairs}$, but $HT$ is not. Consider now the set of real numbers

$$C' \triangleq \{X(\omega); \omega \in \Omega_{pairs}\}.$$

The numbers between $(\frac{1}{4}, \frac{1}{2})$ can be written as $X(\omega)$, but the sequence $\omega$ must begin with either $TH$ or $HT$. Therefore, none of these numbers is in $C'$. Similarly, the numbers between $(\frac{1}{16}, \frac{3}{16})$ can be written as $X(\omega)$, but the sequence $\omega$ must begin with $TTTH$ or $TTHT$, so none of these numbers is in $C'$. Continuing this process, we see that $C'$ will not contain any of the numbers which were removed in the construction of the Cantor set $C$ in Example 3.2. In other words, $C' \subseteq C$. With a bit more work, one can convince onself that in fact $C' = C$, i.e., by requiring consecutive coin tosses to be paired, we are removing exactly those points in $[0,1]$ which were removed in the Cantor set construction of Example 3.2. $\diamond$

In addition to tossing a coin, another common random experiment is to pick a number, perhaps using a random number generator. Here are some probability spaces which correspond to different ways of picking a number at random.

**Example 1.11**
Suppose we choose a number from $I\!\!R$ in such a way that we are sure to get either $1$, $4$ or $16$. Furthermore, we construct the experiment so that the probability of getting $1$ is $\frac{4}{9}$, the probability of getting $4$ is $\frac{4}{9}$ and the probability of getting $16$ is $\frac{1}{9}$. We describe this random experiment by taking $\Omega$ to be $I\!\!R$, $\mathcal{F}$ to be $\mathcal{B}(I\!\!R)$, and setting up the probability measure so that

$$I\!\!P\{1\} = \frac{4}{9},\ I\!\!P\{4\} = \frac{4}{9},\ I\!\!P\{16\} = \frac{1}{9}.$$

This determines $I\!\!P(A)$ for every set $A \in \mathcal{B}(I\!\!R)$. For example, the probability of the interval $(0,5]$ is $\frac{8}{9}$, because this interval contains the numbers $1$ and $4$, but not the number $16$.

The probability measure described in this example is $\mathcal{L}_{S_2}$, the measure induced by the stock price $S_2$, when the initial stock price $S_0 = 4$ and the probability of $H$ is $\frac{1}{3}$. This distribution was discussed immediately following Definition 2.8. $\diamond$

**Example 1.12** Uniform distribution on $[0,1]$.
Let $\Omega = [0,1]$ and let $\mathcal{F} = \mathcal{B}([0,1])$, the collection of all Borel subsets containined in $[0,1]$. For each Borel set $A \subseteq [0,1]$, we define $I\!\!P(A) = \mu_0(A)$ to be the Lebesgue measure of the set. Because $\mu_0[0,1] = 1$, this gives us a probability measure.

This probability space corresponds to the random experiment of choosing a number from $[0,1]$ so that every number is “equally likely” to be chosen. Since there are infinitely mean numbers in $[0,1]$, this requires that every number have probabilty zero of being chosen. Nonetheless, we can speak of the probability that the number chosen lies in a particular set, and if the set has uncountably many points, then this probability can be positive. $\diamond$

I know of no way to design a physical experiment which corresponds to choosing a number at random from $[0,1]$ so that each number is equally likely to be chosen, just as I know of no way to toss a coin infinitely many times. Nonetheless, both Examples 1.10 and 1.12 provide probability spaces which are often useful approximations to reality.

**Example 1.13** Standard normal distribution.
Define the standard normal density

$$\varphi(x) \triangleq \frac{1}{\sqrt{2\pi}}\, e^{-\dfrac{x^2}{2}}.$$

Let $\Omega = I\!\!R$, $\mathcal{F} = \mathcal{B}(I\!\!R)$ and for every Borel set $A \subseteq I\!\!R$, define

$$I\!\!P(A) \triangleq \int_A \varphi\, d\mu_0. \tag{4.2}$$

If $A$ in (4.2) is an interval $[a,b]$, then we can write (4.2) as the less mysterious Riemann integral:

$$\mathbb{P}[a,b] \triangleq \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx.$$

This corresponds to choosing a point at random on the real line, and every single point has probability zero of being chosen, but if a set $A$ is given, then the probability the point is in that set is given by (4.2). $\diamond$

The construction of the integral in a general probability space follows the same steps as the construction of Lebesgue integral. We repeat this construction below.

**Definition 1.14** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X$ be a random variable on this space, i.e., a mapping from $\Omega$ to $\mathbb{R}$, possibly also taking the values $\pm\infty$.

- If $X$ is an *indicator*, i.e,

  $$X(\omega) = \mathbb{I}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \in A^c, \end{cases}$$

  for some set $A \in \mathcal{F}$, we define

  $$\int_\Omega X\, d\mathbb{P} \triangleq \mathbb{P}(A).$$

- If $X$ is a *simple function*, i.e,

  $$X(\omega) = \sum_{k=1}^n c_k \mathbb{I}_{A_k}(\omega),$$

  where each $c_k$ is a real number and each $A_k$ is a set in $\mathcal{F}$, we define

  $$\int_\Omega X\, d\mathbb{P} \triangleq \sum_{k=1}^n c_k \int_\Omega \mathbb{I}_{A_k}\, d\mathbb{P} = \sum_{k=1}^n c_k \mathbb{P}(A_k).$$

- If $X$ is *nonnegative* but otherwise general, we define

  $$\int_\Omega X\, d\mathbb{P} \triangleq \sup\left\{ \int_\Omega Y\, d\mathbb{P};\, Y \text{ is simple and } Y(\omega) \le X(\omega) \text{ for every } \omega \in \Omega \right\}.$$

  In fact, we can always construct a sequence of simple functions $Y_n,\ n = 1, 2, \ldots$ such that

  $$0 \le Y_1(\omega) \le Y_2(\omega) \le Y_3(\omega) \le \ldots \text{ for every } \omega \in \Omega,$$

  and $Y(\omega) = \lim_{n\to\infty} Y_n(\omega)$ for every $\omega \in \Omega$. With this sequence, we can define

  $$\int_\Omega X\, d\mathbb{P} \triangleq \lim_{n\to\infty} \int_\Omega Y_n\, d\mathbb{P}.$$

- If $X$ is *integrable*, i.e,
$$\int_\Omega X^+\,d\mathbb{P} < \infty, \quad \int_\Omega X^-\,d\mathbb{P} < \infty,$$
where
$$X^+(\omega) \triangleq \max\{X(\omega), 0\}, \quad X^-(\omega) \triangleq \max\{-X(\omega), 0\},$$
then we define
$$\int_\Omega X\,d\mathbb{P} \triangleq \int_\Omega X^+\,d\mathbb{P} - - \int_\Omega X^-\,d\mathbb{P}.$$

If $A$ is a set in $\mathcal{F}$ and $X$ is a random variable, we define
$$\int_A X\,d\mathbb{P} \triangleq \int_\Omega \mathbb{I}_A \cdot X\,d\mathbb{P}.$$
The *expectation* of a random variable $X$ is defined to be
$$\mathbb{E}X \triangleq \int_\Omega X\,d\mathbb{P}.$$

The above integral has all the linearity and comparison properties one would expect. In particular, if $X$ and $Y$ are random variables and $c$ is a real constant, then
$$\begin{aligned}\int_\Omega (X+Y)\,d\mathbb{P} &= \int_\Omega X\,d\mathbb{P} + \int_\Omega Y\,d\mathbb{P},\\ \int_\Omega cX\,d\mathbb{P} &= c\int_\Omega X\,dP,\end{aligned}$$
If $X(\omega) \le Y(\omega)$ for every $\omega \in \Omega$, then
$$\int_\Omega X\,d\mathbb{P} \le \int_\Omega Y\,d\mathbb{P}.$$
In fact, we don't need to have $X(\omega) \le Y(\omega)$ for *every* $\omega \in \Omega$ in order to reach this conclusion; it is enough if the set of $\omega$ for which $X(\omega) \le Y(\omega)$ has probability one. When a condition holds with probability one, we say it holds *almost surely*. Finally, if $A$ and $B$ are disjoint subsets of $\Omega$ and $X$ is a random variable, then
$$\int_{A\cup B} X\,d\mathbb{P} = \int_A X\,d\mathbb{P} + \int_B X\,d\mathbb{P}.$$

We restate the Lebesgue integral convergence theorem in this more general context. We acknowledge in these statements that conditions don't need to hold for every $\omega$; almost surely is enough.

**Theorem 4.4** (Fatou's Lemma) *Let $X_n, n = 1, 2, \ldots$ be a sequence of almost surely nonnegative random variables converging almost surely to a random variable $X$. Then*
$$\int_\Omega X\,d\mathbb{P} \le \liminf_{n\to\infty} \int_\Omega X_n\,d\mathbb{P},$$
*or equivalently,*
$$\mathbb{E}X \le \liminf_{n\to\infty} \mathbb{E}X_n.$$

**Theorem 4.5** (Monotone Convergence Theorem) *Let $X_n, n = 1, 2, \ldots$ be a sequence of random variables converging almost surely to a random variable $X$. Assume that*

$$0 \le X_1 \le X_2 \le X_3 \le \cdots \text{ almost surely.}$$

*Then*

$$\int_\Omega X \, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega X_n d\mathbb{P},$$

*or equivalently,*

$$\mathbb{E}X = \lim_{n\to\infty} \mathbb{E}X_n.$$

**Theorem 4.6** (Dominated Convergence Theorem) *Let $X_n, n = 1, 2, \ldots$ be a sequence of random variables, converging almost surely to a random variable $X$. Assume that there exists a random variable $Y$ such that*

$$|X_n| \le Y \text{ almost surely for every } n.$$

*Then*

$$\int_\Omega X \, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega X_n \, d\mathbb{P},$$

*or equivalently,*

$$\mathbb{E}X = \lim_{n\to\infty} \mathbb{E}X_n.$$

In Example 1.13, we constructed a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by integrating the standard normal density. In fact, whenever $\varphi$ is a nonnegative function defined on $R$ satisfying $\int_{\mathbb{R}} \varphi \, d\mu_0 = 1$, we call $\varphi$ a *density* and we can define an associated probability measure by

$$\mathbb{P}(A) \triangleq \int_A \varphi \, d\mu_0 \text{ for every } A \in \mathcal{B}(\mathbb{R}). \tag{4.3}$$

We shall often have a situation in which two measure are related by an equation like (4.3). In fact, the market measure and the risk-neutral measures in financial markets are related this way. We say that $\varphi$ in (4.3) is the *Radon-Nikodym derivative* of $d\mathbb{P}$ with respect to $\mu_0$, and we write

$$\varphi = \frac{d\mathbb{P}}{d\mu_0}. \tag{4.4}$$

The probability measure $\mathbb{P}$ weights different parts of the real line according to the density $\varphi$. Now suppose $f$ is a function on $(R, \mathcal{B}(\mathbb{R}), \mathbb{P})$. Definition 1.14 gives us a value for the abstract integral

$$\int_{\mathbb{R}} f \, d\mathbb{P}.$$

We can also evaluate

$$\int_{\mathbb{R}} f\varphi \, d\mu_0,$$

which is an integral with respec to Lebesgue measure over the real line. We want to show that

$$\int_{\mathbb{R}} f \, d\mathbb{P} = \int_{\mathbb{R}} f\varphi \, d\mu_0, \tag{4.5}$$

an equation which is suggested by the notation introduced in (4.4) (substitute $\frac{d\mathbb{P}}{d\mu_0}$ for $\varphi$ in (4.5) and "cancel" the $d\mu_0$). We include a proof of this because it allows us to illustrate the concept of the *standard machine* explained in Williams's book in Section 5.12, page 5.

The standard machine argument proceeds in four steps.

**Step 1.** Assume that $f$ is an *indicator function*, i.e., $f(x) = \mathbb{I}_A(x)$ for some Borel set $A \subseteq \mathbb{R}$. In that case, (4.5) becomes

$$\mathbb{P}(A) = \int_A \varphi \, d\mu_0.$$

This is true because it is the definition of $\mathbb{P}(A)$.

**Step 2.** Now that we know that (4.5) holds when $f$ is an indicator function, assume that $f$ is a *simple function*, i.e., a linear combination of indicator functions. In other words,

$$f(x) = \sum_{k=1}^{n} c_k h_k(x),$$

where each $c_k$ is a real number and each $h_k$ is an indicator function. Then

$$\begin{aligned}
\int_{\mathbb{R}} f \, d\mathbb{P} &= \int_{\mathbb{R}} \left[ \sum_{k=1}^{n} c_k h_k \right] d\mathbb{P} \\
&= \sum_{k=1}^{n} c_k \int_{\mathbb{R}} h_k \, d\mathbb{P} \\
&= \sum_{k=1}^{n} c_k \int_{\mathbb{R}} h_k \varphi \, d\mu_0 \\
&= \int_{\mathbb{R}} \left[ \sum_{k=1}^{n} c_k h_k \right] \varphi \, d\mu_0 \\
&= \int_{\mathbb{R}} f \varphi \, d\mu_0.
\end{aligned}$$

**Step 3.** Now that we know that (4.5) holds when $f$ is a simple function, we consider a general nonnegative function $f$. We can always construct a sequence of nonnegative simple functions $f_n, n = 1, 2, \ldots$ such that

$$0 \le f_1(x) \le f_2(x) \le f_3(x) \le \ldots \text{ for every } x \in \mathbb{R},$$

and $f(x) = \lim_{n\to\infty} f_n(x)$ for every $x \in \mathbb{R}$. We have already proved that

$$\int_{\mathbb{R}} f_n \, d\mathbb{P} = \int_{\mathbb{R}} f_n \varphi \, d\mu_0 \text{ for every } n.$$

We let $n \to \infty$ and use the Monotone Convergence Theorem on both sides of this equality to get

$$\int_{\mathbb{R}} f \, d\mathbb{P} = \int_{\mathbb{R}} f \varphi \, d\mu_0.$$

**Step 4.** In the last step, we consider an *integrable* function $f$, which can take both positive and negative values. By *integrable*, we mean that

$$\int_{\mathbb{R}} f^+ \, d\mathbb{P} < \infty, \quad \int_{\mathbb{R}} f^- \, d\mathbb{P} < \infty.$$

¿From Step 3, we have

$$\begin{aligned} \int_{\mathbb{R}} f^+ \, d\mathbb{P} &= \int_{\mathbb{R}} f^+ \varphi \, d\mu_0, \\ \int_{\mathbb{R}} f^- \, d\mathbb{P} &= \int_{\mathbb{R}} f^- \varphi \, d\mu_0. \end{aligned}$$

Subtracting these two equations, we obtain the desired result:

$$\begin{aligned} \int_{\mathbb{R}} f \, d\mathbb{P} &= \int_{\mathbb{R}} f^+ \, d\mathbb{P} - \int_{\mathbb{R}} f^- \, d\mathbb{P} \\ &= \int_{\mathbb{R}} f^+ \varphi \, d\mu_0 - \int_{\mathbb{R}} f^- \varphi \, d\mu_0 \\ &= \int_{R} f \varphi \, d\mu_0. \end{aligned}$$

## 1.5 Independence

In this section, we define and discuss the notion of independence in a general probability space $(\Omega, \mathcal{F}, \mathbb{P})$, although most of the examples we give will be for coin toss space.

### 1.5.1 Independence of sets

**Definition 1.15** We say that two sets $A \in \mathcal{F}$ and $B \in \mathcal{F}$ are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Suppose a random experiment is conducted, and $\omega$ is the outcome. The probability that $\omega \in A$ is $\mathbb{P}(A)$. Suppose you are not told $\omega$, but you are told that $\omega \in B$. Conditional on this information, the probability that $\omega \in A$ is

$$\mathbb{P}(A|B) \triangleq \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

The sets $A$ and $B$ are independent if and only if this conditional probability is the uncondidtional probability $\mathbb{P}(A)$, i.e., knowing that $\omega \in B$ does not change the probability you assign to $A$. This discussion is symmetric with respect to $A$ and $B$; if $A$ and $B$ are independent and you know that $\omega \in A$, the conditional probability you assign to $B$ is still the unconditional probability $\mathbb{P}(B)$.

Whether two sets are independent depends on the probability measure $\mathbb{P}$. For example, suppose we toss a coin twice, with probability $p$ for $H$ and probability $q = 1 - p$ for $T$ on each toss. To avoid trivialities, we assume that $0 < p < 1$. Then

$$\mathbb{P}\{HH\} = p^2, \ \mathbb{P}\{HT\} = \mathbb{P}\{TH\} = pq, \ \mathbb{P}\{TT\} = q^2. \tag{5.1}$$

Let $A = \{HH, HT\}$ and $B = \{HT, TH\}$. In words, $A$ is the set "$H$ on the first toss" and $B$ is the set "one $H$ and one $T$." Then $A \cap B = \{HT\}$. We compute

$$\begin{aligned}
I\!\!P(A) &= p^2 + pq = p,\\
I\!\!P(B) &= 2pq,\\
I\!\!P(A) I\!\!P(B) &= 2p^2 q,\\
I\!\!P(A \cap B) &= pq.
\end{aligned}$$

These sets are independent if and only if $2p^2 q = pq$, which is the case if and only if $p = \frac{1}{2}$.

If $p = \frac{1}{2}$, then $I\!\!P(B)$, the probability of one head and one tail, is $\frac{1}{2}$. If you are told that the coin tosses resulted in a head on the first toss, the probability of $B$, which is now the probability of a $T$ on the second toss, is still $\frac{1}{2}$.

Suppose however that $p = 0.01$. By far the most likely outcome of the two coin tosses is $TT$, and the probability of one head and one tail is quite small; in fact, $I\!\!P(B) = 0.0198$. However, if you are told that the first toss resulted in $H$, it becomes very likely that the two tosses result in one head and one tail. In fact, conditioned on getting a $H$ on the first toss, the probability of one $H$ and one $T$ is the probability of a $T$ on the second toss, which is $0.99$.

### 1.5.2 Independence of $\sigma$-algebras

**Definition 1.16** Let $\mathcal{G}$ and $\mathcal{H}$ be sub-$\sigma$-algebras of $\mathcal{F}$. We say that $\mathcal{G}$ and $\mathcal{H}$ are *independent* if every set in $\mathcal{G}$ is independent of every set in $\mathcal{H}$, i.e,

$$I\!\!P(A \cap B) = I\!\!P(A) I\!\!P(B) \text{ for every } A \in \mathcal{H},\ \ B \in \mathcal{G}.$$

**Example 1.14** Toss a coin twice, and let $I\!\!P$ be given by (5.1). Let $\mathcal{G} = \mathcal{F}_1$ be the $\sigma$-algebra determined by the first toss: $\mathcal{G}$ contains the sets

$$\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}.$$

Let $\mathcal{H}$ be the $\sigma$-albegra determined by the second toss: $\mathcal{H}$ contains the sets

$$\emptyset, \Omega, \{HH, TH\}, \{HT, TT\}.$$

These two $\sigma$-algebras are independent. For example, if we choose the set $\{HH, HT\}$ from $\mathcal{G}$ and the set $\{HH, TH\}$ from $\mathcal{H}$, then we have

$$\begin{aligned}
&I\!\!P\{HH, HT\} I\!\!P\{HH, TH\} = (p^2 + pq)(p^2 + pq) = p^2,\\
&I\!\!P\Big(\{HH, HT\} \cap \{HH, TH\}\Big) = I\!\!P\{HH\} = p^2.
\end{aligned}$$

No matter which set we choose in $\mathcal{G}$ and which set we choose in $\mathcal{H}$, we will find that the product of the probabilties is the probability of the intersection.

Example 1.14 illustrates the general principle that when the probability for a sequence of tosses is defined to be the product of the probabilities for the individual tosses of the sequence, then every set depending on a particular toss will be independent of every set depending on a different toss. We say that the different tosses are independent when we construct probabilities this way. It is also possible to construct probabilities such that the different tosses are not independent, as shown by the following example.

**Example 1.15** Define $I\!\!P$ for the individual elements of $\Omega = \{HH, HT, TH, TT\}$ to be

$$I\!\!P\{HH\} = \frac{1}{9},\ I\!\!P\{HT\} = \frac{2}{9},\ I\!\!P\{TH\} = \frac{1}{3},\ I\!\!P\{TT\} = \frac{1}{3},$$

and for every set $A \subseteq \Omega$, define $I\!\!P(A)$ to be the sum of the probabilities of the elements in $A$. Then $I\!\!P(\Omega) = 1$, so $I\!\!P$ is a probability measure. Note that the sets $\{H \text{ on first toss}\} = \{HH, HT\}$ and $\{H \text{ on second toss}\} = \{HH, TH\}$ have probabilities $I\!\!P\{HH, HT\} = \frac{1}{3}$ and $I\!\!P\{HH, TH\} = \frac{4}{9}$, so the product of the probabilities is $\frac{4}{27}$. On the other hand, the intersection of $\{HH, HT\}$ and $\{HH, TH\}$ contains the single element $\{HH\}$, which has probability $\frac{1}{9}$. These sets are not independent.

### 1.5.3 Independence of random variables

**Definition 1.17** We say that two random variables $X$ and $Y$ are *independent* if the $\sigma$-algebras they generate $\sigma(X)$ and $\sigma(Y)$ are independent.

In the probability space of three independent coin tosses, the price $S_2$ of the stock at time $2$ is independent of $\frac{S_3}{S_2}$. This is because $S_2$ depends on only the first two coin tosses, whereas $\frac{S_3}{S_2}$ is either $u$ or $d$, depending on whether the *third* coin toss is $H$ or $T$.

Definition 1.17 says that for independent random variables $X$ and $Y$, every set defined in terms of $X$ is independent of every set defined in terms of $Y$. In the case of $S_2$ and $\frac{S_3}{S_2}$ just considered, for example, the sets $\{S_2 = udS_0\} = \{HTH, HTT\}$ and $\left\{\frac{S_3}{S_2} = u\right\} = \{HHH, HTH, THH, TTH\}$ are indepedent sets.

Suppose $X$ and $Y$ are independent random variables. We defined earlier the measure induced by $X$ on $I\!\!R$ to be

$$\mathcal{L}_X(A) \triangleq I\!\!P\{X \in A\},\ A \subseteq I\!\!R.$$

Similarly, the measure induced by $Y$ is

$$\mathcal{L}_Y(B) \triangleq I\!\!P\{Y \in B\},\ B \subseteq I\!\!R.$$

Now the pair $(X, Y)$ takes values in the plane $I\!\!R^2$, and we can define the measure induced by the pair

$$\mathcal{L}_{X,Y}(C) = I\!\!P\{(X, Y) \in C\},\ C \subseteq I\!\!R^2.$$

The set $C$ in this last equation is a subset of the plane $I\!\!R^2$. In particular, $C$ could be a "rectangle", i.e, a set of the form $A \times B$, where $A \subseteq I\!\!R$ and $B \subseteq I\!\!R$. In this case,

$$\{(X, Y) \in A \times B\} = \{X \in A\} \cap \{Y \in B\},$$

and $X$ and $Y$ are independent if and only if

$$\begin{aligned} \mathcal{L}_{X,Y}(A \times B) &= I\!\!P\Big(\{X \in A\} \cap \{Y \in B\}\Big) \\ &= I\!\!P\{X \in A\} I\!\!P\{Y \in B\} \qquad (5.2) \\ &= \mathcal{L}_X(A)\mathcal{L}_Y(B). \end{aligned}$$

In other words, for independent random variables $X$ and $Y$, the *joint distribution* represented by the measure $\mathcal{L}_{X,Y}$ factors into the product of the *marginal distributions* represented by the measures $\mathcal{L}_X$ and $\mathcal{L}_Y$.

A *joint density* for $(X, Y)$ is a nonnegative function $f_{X,Y}(x, y)$ such that

$$\mathcal{L}_{X,Y}(A \times B) = \int_A \int_B f_{X,Y}(x, y)\, dx\, dy.$$

Not every pair of random variables $(X, Y)$ has a joint density, but if a pair does, then the random variables $X$ and $Y$ have *marginal densities* defined by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta)\, d\eta, \quad f_Y(y) \int_{-\infty}^{\infty} f_{X,Y}(\xi, y)\, d\xi.$$

These have the properties

$$\begin{aligned} \mathcal{L}_X(A) &= \int_A f_X(x)\, dx,\ A \subseteq I\!\!R, \\ \mathcal{L}_Y(B) &= \int_B f_Y(y)\, dy,\ B \subseteq I\!\!R. \end{aligned}$$

Suppose $X$ and $Y$ have a joint density. Then $X$ and $Y$ are independent variables if and only if the joint density is the product of the marginal densities. This follows from the fact that (5.2) is equivalent to independence of $X$ and $Y$. Take $A = (-\infty, x]$ and $B = (-\infty, y]$, write (5.1) in terms of densities, and differentiate with respect to both $x$ and $y$.

**Theorem 5.7** *Suppose $X$ and $Y$ are independent random variables. Let $g$ and $h$ be functions from $I\!\!R$ to $I\!\!R$. Then $g(X)$ and $h(Y)$ are also independent random variables.*

PROOF: Let us denote $W = g(X)$ and $Z = h(Y)$. We must consider sets in $\sigma(W)$ and $\sigma(Z)$. But a typical set in $\sigma(W)$ is of the form

$$\{\omega; W(\omega) \in A\} = \{\omega : g(X(\omega)) \in A\},$$

which is defined in terms of the random variable $X$. Therefore, this set is in $\sigma(X)$. (In general, we have that every set in $\sigma(W)$ is also in $\sigma(X)$, which means that $X$ contains at least as much information as $W$. In fact, $X$ can contain strictly more information than $W$, which means that $\sigma(X)$ will contain all the sets in $\sigma(W)$ and others besides; this is the case, for example, if $W = X^2$.)

In the same way that we just argued that every set in $\sigma(W)$ is also in $\sigma(X)$, we can show that every set in $\sigma(Z)$ is also in $\sigma(Y)$. Since every set in $\sigma(X)$ is independent of every set in $\sigma(Y)$, we conclude that every set in $\sigma(W)$ is independent of every set in $\sigma(Z)$. $\diamond$

**Definition 1.18** Let $X_1, X_2, \ldots$ be a sequence of random variables. We say that these random variables are *independent* if for every sequence of sets $A_1 \in \sigma(X_1), A_2 \in \sigma(X_2), \ldots$ and for every positive integer $n$,

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n).$$

### 1.5.4 Correlation and independence

**Theorem 5.8** *If two random variables $X$ and $Y$ are independent, and if $g$ and $h$ are functions from $\mathbb{R}$ to $\mathbb{R}$, then*

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}g(X) \cdot \mathbb{E}h(Y),$$

*provided all the expectations are defined.*

PROOF: Let $g(x) = \mathbb{I}_A(x)$ and $h(y) = \mathbb{I}_B(y)$ be indicator functions. Then the equation we are trying to prove becomes

$$\mathbb{P}\Big(\{X \in A\} \cap \{Y \in B\}\Big) = \mathbb{P}\{X \in A\}\mathbb{P}\{Y \in B\},$$

which is true because $X$ and $Y$ are independent. Now use the standard machine to get the result for general functions $g$ and $h$. $\diamond$

The *variance* of a random variable $X$ is defined to be

$$\text{Var}(X) \triangleq \mathbb{E}[X - \mathbb{E}X]^2.$$

The covariance of two random variables $X$ and $Y$ is defined to be

$$\begin{aligned}\text{Cov}(X, Y) &\triangleq \mathbb{E}\Big[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\Big] \\ &= \mathbb{E}[XY] - \mathbb{E}X \cdot \mathbb{E}Y.\end{aligned}$$

According to Theorem 5.8, for independent random variables, the covariance is zero. If $X$ and $Y$ both have positive variances, we define their *correlation coefficient*

$$\rho(X, Y) \triangleq \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

For independent random variables, the correlation coefficient is zero.

Unfortunately, two random variables can have zero correlation and still not be independent. Consider the following example.

**Example 1.16** Let $X$ be a standard normal random variable, let $Z$ be independent of $X$ and have the distribution $\mathbb{P}\{Z = 1\} = \mathbb{P}\{Z = -1\} = 0$. Define $Y = XZ$. We show that $Y$ is also a standard normal random variable, $X$ and $Y$ are uncorrelated, but $X$ and $Y$ are not independent.

The last claim is easy to see. If $X$ and $Y$ were independent, so would be $X^2$ and $Y^2$, but in fact, $X^2 = Y^2$ almost surely.

We next check that $Y$ is standard normal. For $y \in I\!\!R$, we have

$$\begin{aligned}
I\!\!P\{Y \le y\} &= I\!\!P\{Y \le y \text{ and } Z = 1\} + I\!\!P\{Y \le y \text{ and } Z = -1\} \\
&= I\!\!P\{X \le y \text{ and } Z = 1\} + I\!\!P\{-X \le y \text{ and } Z = -1\} \\
&= I\!\!P\{X \le y\} I\!\!P\{Z = 1\} + I\!\!P\{-X \le y\} I\!\!P\{Z = -1\} \\
&= \frac{1}{2} I\!\!P\{X \le y\} + \frac{1}{2} I\!\!P\{-X \le y\}.
\end{aligned}$$

Since $X$ is standard normal, $I\!\!P\{X \le y\} = I\!\!P\{X \le -y\}$, and we have $I\!\!P\{Y \le y\} = I\!\!P\{X \le y\}$, which shows that $Y$ is also standard normal.

Being standard normal, both $X$ and $Y$ have expected value zero. Therefore,

$$\text{Cov}(X, Y) = I\!\!E[XY] = I\!\!E[X^2 Z] = I\!\!E X^2 \cdot I\!\!E Z = 1 \cdot 0 = 0.$$

Where in $I\!\!R^2$ does the measure $\mathcal{L}_{X,Y}$ put its mass, i.e., what is the distribution of $(X, Y)$?

We conclude this section with the observation that for independent random variables, the variance of their sum is the sum of their variances. Indeed, if $X$ and $Y$ are independent and $Z = X + Y$, then

$$\begin{aligned}
\text{Var}(Z) &\triangleq I\!\!E\left[(Z - I\!\!E Z)^2\right] \\
&= I\!\!E\left(X + Y - I\!\!E X - I\!\!E Y)^2\right] \\
&= I\!\!E\left[(X - I\!\!E X)^2 + 2(X - I\!\!E X)(Y - I\!\!E Y) + (Y - I\!\!E Y)^2\right] \\
&= \text{Var}(X) + 2 I\!\!E[X - I\!\!E X] I\!\!E[Y - I\!\!E Y] + \text{Var}(Y) \\
&= \text{Var}(X) + \text{Var}(Y).
\end{aligned}$$

This argument extends to any finite number of random variables. If we are given independent random variables $X_1, X_2, \ldots, X_n$, then

$$\text{Var}(X_1 + X_2 + \cdots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n). \tag{5.3}$$

### 1.5.5 Independence and conditional expectation.

We now return to property (k) for conditional expectations, presented in the lecture dated October 19, 1995. The property as stated there is taken from Williams's book, page 88; we shall need only the second assertion of the property:

**(k)** If a random variable $X$ is independent of a $\sigma$-algebra $\mathcal{H}$, then

$$I\!\!E[X|\mathcal{H}] = I\!\!E X.$$

The point of this statement is that if $X$ is independent of $\mathcal{H}$, then the best estimate of $X$ based on the information in $\mathcal{H}$ is $I\!\!E X$, the same as the best estimate of $X$ based on no information.

To show this equality, we observe first that $\mathbb{E}X$ is $\mathcal{H}$-measurable, since it is not random. We must also check the partial averaging property

$$\int_A \mathbb{E}X \, d\mathbb{P} = \int_A X \, d\mathbb{P} \text{ for every } A \in \mathcal{H}.$$

If $X$ is an indicator of some set $B$, which by assumption must be independent of $\mathcal{H}$, then the partial averaging equation we must check is

$$\int_A \mathbb{P}(B) \, d\mathbb{P} = \int_A \mathbb{I}_B \, d\mathbb{P}.$$

The left-hand side of this equation is $\mathbb{P}(A)\mathbb{P}(B)$, and the right hand side is

$$\int_\Omega \mathbb{I}_A \mathbb{I}_B \, d\mathbb{P} = \int_\Omega \mathbb{I}_{A\cap B} \, d\mathbb{P} = \mathbb{P}(A \cap B).$$

The partial averaging equation holds because $A$ and $B$ are independent. The partial averaging equation for general $X$ independent of $\mathcal{H}$ follows by the standard machine.

### 1.5.6 Law of Large Numbers

There are two fundamental theorems about sequences of independent random variables. Here is the first one.

**Theorem 5.9 (Law of Large Numbers)** *Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables, each with expected value $\mu$ and variance $\sigma^2$. Define the sequence of averages*

$$Y_n \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n}, \;\; n = 1, 2, \ldots.$$

*Then $Y_n$ converges to $\mu$ almost surely as $n \to \infty$.*

We are not going to give the proof of this theorem, but here is an argument which makes it plausible. We will use this argument later when developing stochastic calculus. The argument proceeds in two steps. We first check that $\mathbb{E}Y_n = \mu$ for every $n$. We next check that Var$(Y_n) \to 0$ as $n \to 0$. In other words, the random variables $Y_n$ are increasingly tightly distributed around $\mu$ as $n \to \infty$.

For the first step, we simply compute

$$\mathbb{E}Y_n = \frac{1}{n}[\mathbb{E}X_1 + \mathbb{E}X_2 + \cdots + \mathbb{E}X_n] = \frac{1}{n} \underbrace{[\mu + \mu + \cdots + \mu]}_{n \text{ times}} = \mu.$$

For the second step, we first recall from (5.3) that the variance of the sum of independent random variables is the sum of their variances. Therefore,

$$\text{Var}(Y_n) = \sum_{k=1}^n \text{Var}\left(\frac{X_k}{n}\right) = \sum_{k=1}^n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

As $n \to \infty$, we have Var$(Y_n) \to 0$.

### 1.5.7 Central Limit Theorem

The Law of Large Numbers is a bit boring because the limit is nonrandom. This is because the denominator in the definition of $Y_n$ is so large that the variance of $Y_n$ converges to zero. If we want to prevent this, we should divide by $\sqrt{n}$ rather than $n$. In particular, if we again have a sequence of independent, identically distributed random variables, each with expected value $\mu$ and variance $\sigma^2$, but now we set

$$Z_n \triangleq \frac{(X_1 - \mu) + (X_2 - \mu) + \cdots + (X_n - \mu)}{\sqrt{n}},$$

then each $Z_n$ has expected value zero and

$$\text{Var}(Z_n) = \sum_{k=1}^{n} \text{Var}\left(\frac{X_k - \mu}{\sqrt{n}}\right) = \sum_{k=1}^{n} \frac{\sigma^2}{n} = \sigma^2.$$

As $n \to \infty$, the distributions of all the random variables $Z_n$ have the same degree of tightness, as measured by their variance, around their expected value $0$. The Central Limit Theorem asserts that as $n \to \infty$, the distribution of $Z_n$ approaches that of a normal random variable with mean (expected value) zero and variance $\sigma^2$. In other words, for every set $A \subset I\!\!R$,

$$\lim_{n \to \infty} I\!\!P\{Z_n \in A\} = \frac{1}{\sigma\sqrt{2\pi}} \int_A e^{-\frac{x^2}{2\sigma^2}}\, dx.$$

# Chapter 2

# Conditional Expectation

Please see Hull's book (Section 9.6.)

## 2.1 A Binomial Model for Stock Price Dynamics

Stock prices are assumed to follow this simple binomial model: The initial stock price during the period under study is denoted $S_0$. At each time step, the stock price either goes up by a factor of $u$ or down by a factor of $d$. It will be useful to visualize tossing a coin at each time step, and say that

- the stock price moves up by a factor of $u$ if the coin comes out heads ($H$), and
- down by a factor of $d$ if it comes out tails ($T$).

Note that we are not specifying the probability of heads here.

Consider a sequence of 3 tosses of the coin (See Fig. 2.1) The collection of all possible outcomes (i.e. sequences of tosses of length 3) is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THH, THT, TTH, TTT\}.$$

A typical sequence of $\Omega$ will be denoted $\omega$, and $\omega_k$ will denote the $k$th element in the sequence $\omega$. We write $S_k(\omega)$ to denote the stock price at "time" $k$ (i.e. after $k$ tosses) under the outcome $\omega$. Note that $S_k(\omega)$ depends only on $\omega_1, \omega_2, \ldots, \omega_k$. Thus in the 3-coin-toss example we write for instance,

$$S_1(\omega) \triangleq S_1(\omega_1, \omega_2, \omega_3) \triangleq S_1(\omega_1),$$

$$S_2(\omega) \triangleq S_2(\omega_1, \omega_2, \omega_3) \triangleq S_2(\omega_1, \omega_2).$$

Each $S_k$ is a *random variable* defined on the set $\Omega$. More precisely, let $\mathcal{F} = \mathcal{P}(\Omega)$. Then $\mathcal{F}$ is a $\sigma$-algebra and $(\Omega, \mathcal{F})$ is a measurable space. Each $S_k$ is an $\mathcal{F}$-measurable function $\Omega \to \mathbb{R}$, that is, $S_k^{-1}$ is a function $\mathcal{B} \to \mathcal{F}$ where $\mathcal{B}$ is the Borel $\sigma$-algebra on R. We will see later that $S_k$ is in fact

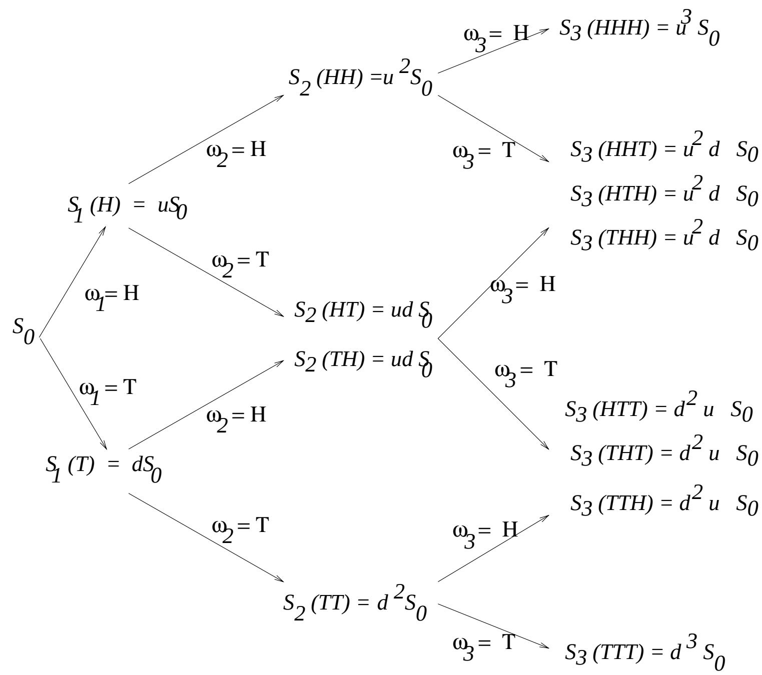



Figure 2.1: *A three coin period binomial model.*

measurable under a sub-$\sigma$-algebra of $\mathcal{F}$. Recall that the Borel $\sigma$-algebra $\mathcal{B}$ is the $\sigma$-algebra generated by the open intervals of R. In this course we will always deal with subsets of R that belong to $\mathcal{B}$.

For any random variable $X$ defined on a sample space $\Omega$ and any $y \in I\!R$, we will use the notation:

$$\{X \le y\} \stackrel{\triangle}{=} \{\omega \in \Omega; X(\omega) \le y\}.$$

The sets $\{X < y\}, \{X \ge y\}, \{X = y\}$, etc, are defined similarly. Similarly for any subset $B$ of $I\!R$, we define

$$\{X \in B\} \stackrel{\triangle}{=} \{\omega \in \Omega; X(\omega) \in B\}.$$

**Assumption 2.1** $u > d > 0$.

## 2.2 Information

**Definition 2.1 (Sets determined by the first $k$ tosses.)** We say that a set $A \subset \Omega$ is *determined by the first $k$ coin tosses* if, knowing only the outcome of the first $k$ tosses, we can decide whether the outcome of *all* tosses is in $A$. In general we denote the collection of sets determined by the first $k$ tosses by $\mathcal{F}_k$. It is easy to check that $\mathcal{F}_k$ is a $\sigma$-algebra.

Note that the random variable $S_k$ is $\mathcal{F}_k$-measurable, for each $k = 1, 2, \ldots, n$.

**Example 2.1** In the 3 coin-toss example, the collection $\mathcal{F}_1$ of sets determined by the first toss consists of:

1. $A_H \triangleq \{HHH, HHT, HTH, HTT\}$,
2. $A_T \triangleq \{THH, THT, TTH, TTT\}$,
3. $\phi$,
4. $\Omega$.

The collection $\mathcal{F}_2$ of sets determined by the first two tosses consists of:

1. $A_{HH} \triangleq \{HHH, HHT\}$,
2. $A_{HT} \triangleq \{HTH, HTT\}$,
3. $A_{TH} \triangleq \{THH, THT\}$,
4. $A_{TT} \triangleq \{TTH, TTT\}$,
5. The complements of the above sets,
6. Any union of the above sets (including the complements),
7. $\phi$ and $\Omega$.

■

**Definition 2.2 (Information carried by a random variable.)** Let $X$ be a random variable $\Omega \to I\!R$. We say that a set $A \subset \Omega$ is *determined by the random variable* $X$ if, knowing only the value $X(\omega)$ of the random variable, we can decide whether or not $\omega \in A$. Another way of saying this is that for every $y \in I\!R$, either $X^{-1}(y) \subset A$ or $X^{-1}(y) \cap A = \phi$. The collection of susbets of $\Omega$ determined by $X$ is a $\sigma$-algebra, which we call the $\sigma$-algebra generated by $X$, and denote by $\sigma(X)$.

If the random variable $X$ takes finitely many different values, then $\sigma(X)$ is generated by the collection of sets

$$\{X^{-1}(X(\omega)) | \omega \in \Omega\};$$

these sets are called the *atoms* of the $\sigma$-algebra $\sigma(X)$.

In general, if $X$ is a random variable $\Omega \to I\!R$, then $\sigma(X)$ is given by

$$\sigma(X) = \{X^{-1}(B); B \in \mathcal{B}\}.$$

**Example 2.2 (Sets determined by $S_2$)** The $\sigma$-algebra generated by $S_2$ consists of the following sets:

1. $A_{HH} = \{HHH, HHT\} = \{\omega \in \Omega; S_2(\omega) = u^2 S_0\}$,
2. $A_{TT} = \{TTH, TTT\} = \{S_2 = d^2 S_0\}$,
3. $A_{HT} \cup A_{TH} = \{S_2 = udS_0\}$,
4. Complements of the above sets,
5. Any union of the above sets,
6. $\phi = \{S_2(\omega) \in \phi\}$,
7. $\Omega = \{S_2(\omega) \in I\!R\}$.

■

## 2.3 Conditional Expectation

In order to talk about conditional expectation, we need to introduce a probability measure on our coin-toss sample space $\Omega$. Let us define

- $p \in (0,1)$ is the probability of $H$,
- $q \triangleq (1-p)$ is the probability of $T$,
- the coin tosses are *independent*, so that, e.g., $I\!\!P(HHT) = p^2 q$, etc.
- $I\!\!P(A) \triangleq \sum_{\omega \in A} I\!\!P(\omega)$, $\forall A \subset \Omega$.

**Definition 2.3 (Expectation.)**

$$I\!\!E X \triangleq \sum_{\omega \in \Omega} X(\omega) I\!\!P(\omega).$$

If $A \subset \Omega$ then

$$I_A(\omega) \triangleq \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

and

$$I\!\!E(I_A X) = \int_A X \, dI\!\!P = \sum_{\omega \in A} X(\omega) I\!\!P(\omega).$$

We can think of $I\!\!E(I_A X)$ as a *partial average* of $X$ over the set $A$.

### 2.3.1 An example

Let us estimate $S_1$, given $S_2$. Denote the estimate by $I\!\!E(S_1|S_2)$. From elementary probability, $I\!\!E(S_1|S_2)$ is a random variable $Y$ whose value at $\omega$ is defined by

$$Y(\omega) = I\!\!E(S_1|S_2 = y),$$

where $y = S_2(\omega)$. Properties of $I\!\!E(S_1|S_2)$:

- $I\!\!E(S_1|S_2)$ should depend on $\omega$, i.e., it is a *random variable*.
- If the value of $S_2$ is known, then the value of $I\!\!E(S_1|S_2)$ should also be known. In particular,
  - If $\omega = HHH$ or $\omega = HHT$, then $S_2(\omega) = u^2 S_0$. If we know that $S_2(\omega) = u^2 S_0$, then even without knowing $\omega$, we know that $S_1(\omega) = uS_0$. We define

    $$I\!\!E(S_1|S_2)(HHH) = I\!\!E(S_1|S_2)(HHT) = uS_0.$$

  - If $\omega = TTT$ or $\omega = TTH$, then $S_2(\omega) = d^2 S_0$. If we know that $S_2(\omega) = d^2 S_0$, then even without knowing $\omega$, we know that $S_1(\omega) = dS_0$. We define

    $$I\!\!E(S_1|S_2)(TTT) = I\!\!E(S_1|S_2)(TTH) = dS_0.$$

– If $\omega \in A = \{HTH, HTT, THH, THT\}$, then $S_2(\omega) = udS_0$. If we know $S_2(\omega) = udS_0$, then we do not know whether $S_1 = uS_0$ or $S_1 = dS_0$. We then take a weighted average:

$$I\!P(A) = p^2q + pq^2 + p^2q + pq^2 = 2pq.$$

Furthermore,

$$\begin{aligned}\int_A S_1\,dI\!P &= p^2quS_0 + pq^2uS_0 + p^2qdS_0 + pq^2dS_0 \\ &= pq(u+d)S_0\end{aligned}$$

For $\omega \in A$ we define

$$I\!E(S_1|S_2)(\omega) = \frac{\int_A S_1\,dI\!P}{I\!P(A)} = \tfrac{1}{2}(u+d)S_0.$$

Then

$$\int_A I\!E(S_1|S_2)\,dI\!P = \int_A S_1\,dI\!P.$$

In conclusion, we can write

$$I\!E(S_1|S_2)(\omega) = g(S_2(\omega)),$$

where

$$g(x) = \begin{cases} uS_0 & \text{if } x = u^2S_0 \\ \frac{1}{2}(u+d)S_0 & \text{if } x = udS_0 \\ dS_0 & \text{if } x = d^2S_0 \end{cases}$$

In other words, $I\!E(S_1|S_2)$ is random *only through dependence on* $S_2$. We also write

$$I\!E(S_1|S_2 = x) = g(x),$$

where $g$ is the function defined above.

The random variable $I\!E(S_1|S_2)$ has two fundamental properties:

- $I\!E(S_1|S_2)$ is $\sigma(S_2)$-measurable.
- For every set $A \in \sigma(S_2)$,

$$\int_A I\!E(S_1|S_2)\,dI\!P = \int_A S_1\,dI\!P.$$

### 2.3.2 Definition of Conditional Expectation

Please see Williams, p.83.

Let $(\Omega, \mathcal{F}, I\!P)$ be a probability space, and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. Let $X$ be a random variable on $(\Omega, \mathcal{F}, I\!P)$. Then $I\!E(X|\mathcal{G})$ is defined to be any random variable $Y$ that satisfies:

**(a)** $Y$ is $\mathcal{G}$-measurable,

**(b)** For every set $A \in \mathcal{G}$, we have the "partial averaging property"

$$\int_A Y\,d\mathbb{P} = \int_A X\,d\mathbb{P}.$$

**Existence.** There is always a random variable $Y$ satisfying the above properties (provided that $\mathbb{E}|X| < \infty$), i.e., conditional expectations always exist.

**Uniqueness.** There can be more than one random variable $Y$ satisfying the above properties, but if $Y'$ is another one, then $Y = Y'$ almost surely, i.e., $\mathbb{P}\{\omega \in \Omega; Y(\omega) = Y'(\omega)\} = 1.$

**Notation 2.1** For random variables $X, Y$, it is standard notation to write

$$\mathbb{E}(X|Y) \stackrel{\triangle}{=} \mathbb{E}(X|\sigma(Y)).$$

Here are some useful ways to think about $\mathbb{E}(X|\mathcal{G})$:

- A random experiment is performed, i.e., an element $\omega$ of $\Omega$ is selected. The value of $\omega$ is partially but not fully revealed to us, and thus we cannot compute the exact value of $X(\omega)$. Based on what we know about $\omega$, we compute an estimate of $X(\omega)$. Because this estimate depends on the partial information we have about $\omega$, it depends on $\omega$, i.e., $\mathbb{E}[X|Y](\omega)$ is a function of $\omega$, although the dependence on $\omega$ is often not shown explicitly.
- If the $\sigma$-algebra $\mathcal{G}$ contains finitely many sets, there will be a "smallest" set $A$ in $\mathcal{G}$ containing $\omega$, which is the intersection of all sets in $\mathcal{G}$ containing $\omega$. The way $\omega$ is partially revealed to us is that we are told it is in $A$, but not told which element of $A$ it is. We then define $\mathbb{E}[X|Y](\omega)$ to be the average (with respect to $\mathbb{P}$) value of $X$ over this set $A$. Thus, for all $\omega$ in this set $A$, $\mathbb{E}[X|Y](\omega)$ will be the same.

### 2.3.3 Further discussion of Partial Averaging

The partial averaging property is

$$\int_A \mathbb{E}(X|\mathcal{G})d\mathbb{P} = \int_A X\,d\mathbb{P}, \forall A \in \mathcal{G}. \tag{3.1}$$

We can rewrite this as

$$\mathbb{E}[I_A.\mathbb{E}(X|\mathcal{G})] = \mathbb{E}[I_A.X]. \tag{3.2}$$

Note that $I_A$ is a $\mathcal{G}$-measurable random variable. In fact the following holds:

**Lemma 3.10** *If $V$ is any $\mathcal{G}$-measurable random variable, then provided $\mathbb{E}|V.\mathbb{E}(X|\mathcal{G})| < \infty$,*

$$\mathbb{E}[V.\mathbb{E}(X|\mathcal{G})] = \mathbb{E}[V.X]. \tag{3.3}$$

**Proof:** To see this, first use (3.2) and linearity of expectations to prove (3.3) when $V$ is a *simple* $\mathcal{G}$-measurable random variable, i.e., $V$ is of the form $V = \sum_{k=1}^{n} c_k I_{A_K}$, where each $A_k$ is in $\mathcal{G}$ and each $c_k$ is constant. Next consider the case that $V$ is a nonnegative $\mathcal{G}$-measurable random variable, but is not necessarily simple. Such a $V$ can be written as the limit of an increasing sequence of simple random variables $V_n$; we write (3.3) for each $V_n$ and then pass to the limit, using the Monotone Convergence Theorem (See Williams), to obtain (3.3) for $V$. Finally, the general $\mathcal{G}$-measurable random variable $V$ can be written as the difference of two nonnegative random-variables $V = V^+ - V^-$, and since (3.3) holds for $V^+$ and $V^-$ it must hold for $V$ as well. Williams calls this argument the "standard machine" (p. 56). ∎

Based on this lemma, we can replace the second condition in the definition of a conditional expectation (Section 2.3.2) by:

**(b')** For every $\mathcal{G}$-measurable random-variable $V$, we have

$$\mathbb{E}[V.\mathbb{E}(X|\mathcal{G})] = \mathbb{E}[V.X]. \tag{3.4}$$

### 2.3.4 Properties of Conditional Expectation

Please see Willams p. 88. Proof sketches of some of the properties are provided below.

**(a)** $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$.

Proof: Just take $A$ in the partial averaging property to be $\Omega$.

The conditional expectation of $X$ is thus an unbiased estimator of the random variable $X$.

**(b)** If $X$ is $\mathcal{G}$-measurable, then

$$\mathbb{E}(X|\mathcal{G}) = X.$$

Proof: The partial averaging property holds trivially when $Y$ is replaced by $X$. And since $X$ is $\mathcal{G}$-measurable, $X$ satisfies the requirement (a) of a conditional expectation as well.

If the information content of $\mathcal{G}$ is sufficient to determine $X$, then the best estimate of $X$ based on $\mathcal{G}$ is $X$ itself.

**(c)** (Linearity)

$$\mathbb{E}(a_1X_1 + a_2X_2|\mathcal{G}) = a_1\mathbb{E}(X_1|\mathcal{G}) + a_2\mathbb{E}(X_2|\mathcal{G}).$$

**(d)** (Positivity) If $X \geq 0$ almost surely, then

$$\mathbb{E}(X|\mathcal{G}) \geq 0.$$

Proof: Take $A = \{\omega \in \Omega; \mathbb{E}(X|\mathcal{G})(\omega) < 0\}$. This set is in $\mathcal{G}$ since $\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable. Partial averaging implies $\int_A \mathbb{E}(X|\mathcal{G})\,d\mathbb{P} = \int_A X\,d\mathbb{P}$. The right-hand side is greater than or equal to zero, and the left-hand side is strictly negative, unless $\mathbb{P}(A) = 0$. Therefore, $\mathbb{P}(A) = 0$.

**(h)** (Jensen's Inequality) If $\phi : R \to R$ is convex and $\mathbb{E}|\phi(X)| < \infty$, then

$$\mathbb{E}(\phi(X)|\mathcal{G}) \geq \phi(\mathbb{E}(X|\mathcal{G})).$$

Recall the usual Jensen's Inequality: $\mathbb{E}\phi(X) \geq \phi(\mathbb{E}(X))$.

**(i)** (Tower Property) If $\mathcal{H}$ is a sub-$\sigma$-algebra of $\mathcal{G}$, then

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H}).$$

$\mathcal{H}$ is a sub-$\sigma$-algebra of $\mathcal{G}$ means that $\mathcal{G}$ contains more information than $\mathcal{H}$. If we estimate $X$ based on the information in $\mathcal{G}$, and then estimate the estimator based on the smaller amount of information in $\mathcal{H}$, then we get the same result as if we had estimated $X$ directly based on the information in $\mathcal{H}$.

**(j)** (Taking out what is known) If $Z$ is $\mathcal{G}$-measurable, then

$$\mathbb{E}(ZX|\mathcal{G}) = Z.\mathbb{E}(X|\mathcal{G}).$$

When conditioning on $\mathcal{G}$, the $\mathcal{G}$-measurable random variable $Z$ acts like a constant.

Proof: Let $Z$ be a $\mathcal{G}$-measurable random variable. A random variable $Y$ is $\mathbb{E}(ZX|\mathcal{G})$ if and only if

(a) $Y$ is $\mathcal{G}$-measurable;
(b) $\int_A Y\,d\mathbb{P} = \int_A ZX\,d\mathbb{P}, \forall A \in \mathcal{G}$.

Take $Y = Z.\mathbb{E}(X|\mathcal{G})$. Then $Y$ satisfies (a) (a product of $\mathcal{G}$-measurable random variables is $\mathcal{G}$-measurable). $Y$ also satisfies property (b), as we can check below:

$$\begin{aligned}
\int_A Y\,d\mathbb{P} &= \mathbb{E}(I_A.Y) \\
&= \mathbb{E}[I_A Z \mathbb{E}(X|\mathcal{G})] \\
&= \mathbb{E}[I_A Z.X] \text{ ((b') with } V = I_A Z \text{)} \\
&= \int_A ZX\,d\mathbb{P}.
\end{aligned}$$

**(k)** (Role of Independence) If $\mathcal{H}$ is independent of $\sigma(\sigma(X), \mathcal{G})$, then

$$\mathbb{E}(X|\sigma(\mathcal{G}, \mathcal{H})) = \mathbb{E}(X|\mathcal{G}).$$

In particular, if $X$ is independent of $\mathcal{H}$, then

$$\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(X).$$

If $\mathcal{H}$ is independent of $X$ and $\mathcal{G}$, then nothing is gained by including the information content of $\mathcal{H}$ in the estimation of $X$.

### 2.3.5 Examples from the Binomial Model

Recall that $\mathcal{F}_1 = \{\phi, A_H, A_T, \Omega\}$. Notice that $I\!E(S_2|\mathcal{F}_1)$ must be constant on $A_H$ and $A_T$.
Now since $I\!E(S_2|\mathcal{F}_1)$ must satisfy the partial averaging property,

$$\int_{A_H} I\!E(S_2|\mathcal{F}_1)\,dI\!P = \int_{A_H} S_2\,dI\!P,$$

$$\int_{A_T} I\!E(S_2|\mathcal{F}_1)\,dI\!P = \int_{A_T} S_2\,dI\!P.$$

We compute

$$\begin{aligned}\int_{A_H} I\!E(S_2|\mathcal{F}_1)\,dI\!P &= I\!P(A_H).I\!E(S_2|\mathcal{F}_1)(\omega)\\ &= pI\!E(S_2|\mathcal{F}_1)(\omega), \forall\omega \in A_H.\end{aligned}$$

On the other hand,

$$\int_{A_H} S_2\,dI\!P = p^2u^2S_0 + pqudS_0.$$

Therefore,

$$I\!E(S_2|\mathcal{F}_1)(\omega) = pu^2S_0 + qudS_0, \forall\omega \in A_H.$$

We can also write

$$\begin{aligned}I\!E(S_2|\mathcal{F}_1)(\omega) &= pu^2S_0 + qudS_0\\ &= (pu+qd)uS_0\\ &= (pu+qd)S_1(\omega), \forall\omega \in A_H\end{aligned}$$

Similarly,

$$I\!E(S_2|\mathcal{F}_1)(\omega) = (pu+qd)S_1(\omega), \forall\omega \in A_T.$$

Thus in both cases we have

$$I\!E(S_2|\mathcal{F}_1)(\omega) = (pu+qd)S_1(\omega), \forall\omega \in \Omega.$$

A similar argument one time step later shows that

$$I\!E(S_3|\mathcal{F}_2)(\omega) = (pu+qd)S_2(\omega).$$

We leave the verification of this equality as an exercise. We can verify the Tower Property, for instance, from the previous equations we have

$$\begin{aligned}I\!E[I\!E(S_3|\mathcal{F}_2)|\mathcal{F}_1] &= I\!E[(pu+qd)S_2|\mathcal{F}_2]\\ &= (pu+qd)I\!E(S_2|\mathcal{F}_1) \quad \text{(linearity)}\\ &= (pu+qd)^2S_1.\end{aligned}$$

This final expression is $I\!E(S_3|\mathcal{F}_1)$.

## 2.4 Martingales

The ingredients are:

- A probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
- A sequence of $\sigma$-algebras $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n$, with the property that $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \mathcal{F}$. Such a sequence of $\sigma$-algebras is called a *filtration*.
- A sequence of random variables $M_0, M_1, \ldots, M_n$. This is called a *stochastic process*.

Conditions for a martingale:

1. Each $M_k$ is $\mathcal{F}_k$-measurable. If you know the information in $\mathcal{F}_k$, then you know the value of $M_k$. We say that the process $\{M_k\}$ is *adapted* to the filtration $\{\mathcal{F}_k\}$.
2. For each $k$, $\mathbb{E}(M_{k+1}|\mathcal{F}_k) = M_k$. Martingales tend to go neither up nor down.

A *supermartingale* tends to go *down*, i.e. the second condition above is replaced by $\mathbb{E}(M_{k+1}|\mathcal{F}_k) \leq M_k$; a *submartingale* tends to go *up*, i.e. $\mathbb{E}(M_{k+1}|\mathcal{F}_k) \geq M_k$.

**Example 2.3 (Example from the binomial model.)** For $k = 1, 2$ we already showed that

$$\mathbb{E}(S_{k+1}|\mathcal{F}_k) = (pu + qd)S_k.$$

For $k = 0$, we set $\mathcal{F}_0 = \{\phi, \Omega\}$, the "trivial $\sigma$-algebra". This $\sigma$-algebra contains no information, and any $\mathcal{F}_0$-measurable random variable must be constant (nonrandom). Therefore, by definition, $\mathbb{E}(S_1|\mathcal{F}_0)$ is that constant which satisfies the averaging property

$$\int_\Omega \mathbb{E}(S_1|\mathcal{F}_0)\, d\mathbb{P} = \int_\Omega S_1\, d\mathbb{P}.$$

The right hand side is $\mathbb{E}S_1 = (pu + qd)S_0$, and so we have

$$\mathbb{E}(S_1|\mathcal{F}_0) = (pu + qd)S_0.$$

In conclusion,

- If $(pu + qd) = 1$ then $\{S_k, \mathcal{F}_k; k = 0, 1, 2, 3\}$ is a martingale.
- If $(pu + qd) \geq 1$ then $\{S_k, \mathcal{F}_k; k = 0, 1, 2, 3\}$ is a submartingale.
- If $(pu + qd) \leq 1$ then $\{S_k, \mathcal{F}_k; k = 0, 1, 2, 3\}$ is a supermartingale.

■

# Chapter 3

# Arbitrage Pricing

## 3.1 Binomial Pricing

Return to the binomial pricing model

Please see:

- Cox, Ross and Rubinstein, *J. Financial Economics*, **7**(1979), 229–263, and
- Cox and Rubinstein (1985), **Options Markets**, Prentice-Hall.

**Example 3.1 (Pricing a Call Option)** Suppose $u = 2, d = 0.5, r = 25\%$(interest rate), $S_0 = 50$. (In this and all examples, the interest rate quoted is per unit time, and the stock prices $S_0, S_1, \ldots$ are indexed by the same time periods). We know that

$$S_1(\omega) = \begin{cases} 100 & \text{if } \omega_1 = H \\ 25 & \text{if } \omega_1 = T \end{cases}$$

Find the value *at time zero* of a call option to buy one share of stock at time 1 for \$50 (i.e. the *strike price* is \$50).

The value of the call at time 1 is

$$V_1(\omega) = (S_1(\omega) - 50)^+ = \begin{cases} 50 & \text{if } \omega_1 = H \\ 0 & \text{if } \omega_1 = T \end{cases}$$

Suppose the option sells for \$20 at time 0. Let us construct a portfolio:

1. Sell 3 options for \$20 each. Cash outlay is $-\$60$.
2. Buy 2 shares of stock for \$50 each. Cash outlay is \$100.
3. Borrow \$40. Cash outlay is $-\$40$.

This portfolio thus requires no initial investment. For this portfolio, the cash outlay at time 1 is:

| | $\omega_1 = H$ | $\omega_1 = T$ |
|---|---|---|
| Pay off option | $\$150$ | $\$0$ |
| Sell stock | $-\$200$ | $-\$50$ |
| Pay off debt | $\$50$ | $\$50$ |
| | $-----$ | $-----$ |
| | $\$0$ | $\$0$ |

The *arbitrage pricing theory (APT)* value of the option at time 0 is $V_0 = 20$. ■

Assumptions underlying APT:

- Unlimited short selling of stock.
- Unlimited borrowing.
- No transaction costs.
- Agent is a "small investor", i.e., his/her trading does not move the market.

**Important Observation:** The APT value of the option does not depend on the probabilities of $H$ and $T$.

## 3.2 General one-step APT

Suppose a derivative security pays off the amount $V_1$ at time 1, where $V_1$ is an $\mathcal{F}_1$-measurable random variable. (This measurability condition is important; this is why it does not make sense to use some stock unrelated to the derivative security in valuing it, at least in the straightforward method described below).

- Sell the security for $V_0$ at time 0. ($V_0$ is to be determined later).
- Buy $\Delta_0$ shares of stock at time 0. ($\Delta_0$ is also to be determined later)
- Invest $V_0 - \Delta_0 S_0$ in the money market, at risk-free interest rate $r$. ($V_0 - \Delta_0 S_0$ might be negative).
- Then wealth at time 1 is

$$\begin{aligned} X_1 &\triangleq \Delta_0 S_1 + (1+r)(V_0 - \Delta_0 S_0) \\ &= (1+r)V_0 + \Delta_0(S_1 - (1+r)S_0). \end{aligned}$$

- We want to choose $V_0$ and $\Delta_0$ so that

$$X_1 = V_1$$

  *regardless of whether the stock goes up or down.*

The last condition above can be expressed by *two* equations (which is fortunate since there are *two* unknowns):

$$(1+r)V_0 + \Delta_0(S_1(H) - (1+r)S_0) = V_1(H) \tag{2.1}$$

$$(1+r)V_0 + \Delta_0(S_1(T) - (1+r)S_0) = V_1(T) \tag{2.2}$$

Note that this is where we use the fact that the derivative security value $V_k$ is a function of $S_k$, i.e., when $S_k$ is known for a given $\omega$, $V_k$ is known (and therefore non-random) at that $\omega$ as well. Subtracting the second equation above from the first gives

$$\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)}. \tag{2.3}$$

Plug the formula (2.3) for $\Delta_0$ into (2.1):

$$\begin{aligned}
(1+r)V_0 &= V_1(H) - \Delta_0(S_1(H) - (1+r)S_0) \\
&= V_1(H) - \frac{V_1(H) - V_1(T)}{(u-d)S_0}(u-1-r)S_0 \\
&= \frac{1}{u-d}[(u-d)V_1(H) - (V_1(H) - V_1(T))(u-1-r)] \\
&= \frac{1+r-d}{u-d}V_1(H) + \frac{u-1-r}{u-d}V_1(T).
\end{aligned}$$

We have already assumed $u > d > 0$. We now also assume $d \le 1+r \le u$ (otherwise there would be an arbitrage opportunity). Define

$$\tilde{p} \stackrel{\triangle}{=} \frac{1+r-d}{u-d}, \quad \tilde{q} \stackrel{\triangle}{=} \frac{u-1-r}{u-d}.$$

Then $\tilde{p} > 0$ and $\tilde{q} > 0$. Since $\tilde{p} + \tilde{q} = 1$, we have $0 < \tilde{p} < 1$ and $\tilde{q} = 1 - \tilde{p}$. Thus, $\tilde{p}, \tilde{q}$ are like probabilities. We will return to this later. Thus the price of the call at time 0 is given by

$$V_0 = \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)]. \tag{2.4}$$

## 3.3 Risk-Neutral Probability Measure

Let $\Omega$ be the set of possible outcomes from $n$ coin tosses. Construct a probability measure $\widetilde{I\!P}$ on $\Omega$ by the formula

$$\widetilde{I\!P}(\omega_1, \omega_2, \ldots, \omega_n) \stackrel{\triangle}{=} \tilde{p}^{\#\{j;\omega_j = H\}}\tilde{q}^{\#\{j;\omega_j = T\}}$$

$\widetilde{I\!P}$ is called the *risk-neutral probability measure*. We denote by $\widetilde{I\!E}$ the expectation under $\widetilde{I\!P}$. Equation 2.4 says

$$V_0 = \widetilde{I\!E}\left(\frac{1}{1+r}V_1\right).$$

**Theorem 3.11** *Under $\widetilde{I\!\!P}$, the discounted stock price process $\{(1+r)^{-k}S_k, \mathcal{F}_k\}_{k=0}^n$ is a martingale.*

**Proof:**

$$\begin{aligned}
&\widetilde{I\!\!E}[(1+r)^{-(k+1)}S_{k+1}|\mathcal{F}_k] \\
&= (1+r)^{-(k+1)}(\tilde{p}u+\tilde{q}d)S_k \\
&= (1+r)^{-(k+1)}\left(\frac{u(1+r-d)}{u-d}+\frac{d(u-1-r)}{u-d}\right)S_k \\
&= (1+r)^{-(k+1)}\frac{u+ur-ud+du-d-dr}{u-d}S_k \\
&= (1+r)^{-(k+1)}\frac{(u-d)(1+r)}{u-d}S_k \\
&= (1+r)^{-k}S_k.
\end{aligned}$$

### 3.3.1 Portfolio Process

The portfolio process is $\Delta = (\Delta_0, \Delta_1, \dots, \Delta_{n-1})$, where

- $\Delta_k$ is the number of shares of stock held between times $k$ and $k+1$.
- Each $\Delta_k$ is $\mathcal{F}_k$-measurable. (No insider trading).

### 3.3.2 Self-financing Value of a Portfolio Process $\Delta$

- Start with nonrandom initial wealth $X_0$, which need not be 0.
- Define recursively

$$\begin{aligned}
X_{k+1} &= \Delta_k S_{k+1} + (1+r)(X_k - \Delta_k S_k) && (3.1) \\
&= (1+r)X_k + \Delta_k(S_{k+1} - (1+r)S_k). && (3.2)
\end{aligned}$$

- Then each $X_k$ is $\mathcal{F}_k$-measurable.

**Theorem 3.12** *Under $\widetilde{I\!\!P}$, the discounted self-financing portfolio process value $\{(1+r)^{-k}X_k, \mathcal{F}_k\}_{k=0}^n$ is a martingale.*

**Proof:** We have

$$(1+r)^{-(k+1)}X_{k+1} = (1+r)^{-k}X_k + \Delta_k\left((1+r)^{-(k+1)}S_{k+1} - (1+r)^{-k}S_k\right).$$

Therefore,

$$\begin{aligned}
&\widetilde{I\!E}[(1+r)^{-(k+1)}X_{k+1}|\mathcal{F}_k] \\
&\quad = \widetilde{I\!E}[(1+r)^{-k}X_k|\mathcal{F}_k] \\
&\qquad +\widetilde{I\!E}[(1+r)^{-(k+1)}\Delta_k S_{k+1}|\mathcal{F}_k] \\
&\qquad -\widetilde{I\!E}[(1+r)^{-k}\Delta_k S_k|\mathcal{F}_k] \\
&\quad = (1+r)^{-k}X_k \quad \text{(requirement (b) of conditional exp.)} \\
&\qquad +\Delta_k\widetilde{I\!E}[(1+r)^{-(k+1)}S_{k+1}|\mathcal{F}_k] \quad \text{(taking out what is known)} \\
&\qquad -(1+r)^{-k}\Delta_k S_k \quad \text{(property (b))} \\
&\quad = (1+r)^{-k}X_k \quad \text{(Theorem 3.11)}
\end{aligned}$$

■

## 3.4 Simple European Derivative Securities

**Definition 3.1 ()** A *simple European derivative security* with expiration time $m$ is an $\mathcal{F}_m$-measurable random variable $V_m$. (Here, $m$ is less than or equal to $n$, the number of periods/coin-tosses in the model).

**Definition 3.2 ()** A simple European derivative security $V_m$ is said to be *hedgeable* if there exists a constant $X_0$ and a portfolio process $\Delta = (\Delta_0, \dots, \Delta_{m-1})$ such that the self-financing value process $X_0, X_1, \dots, X_m$ given by (3.2) satisfies

$$X_m(\omega) = V_m(\omega), \quad \forall\omega \in \Omega.$$

In this case, for $k = 0, 1, \dots, m$, we call $X_k$ the *APT value at time $k$ of $V_m$*.

**Theorem 4.13 (Corollary to Theorem 3.12)** *If a simple European security $V_m$ is hedgeable, then for each $k = 0, 1, \dots, m$, the APT value at time $k$ of $V_m$ is*

$$V_k \stackrel{\triangle}{=} (1+r)^k\widetilde{I\!E}[(1+r)^{-m}V_m|\mathcal{F}_k]. \tag{4.1}$$

**Proof:** We first observe that if $\{M_k, \quad \mathcal{F}_k; k = 0, 1, \dots, m\}$ is a martingale, i.e., satisfies the martingale property

$$\widetilde{I\!E}[M_{k+1}|\mathcal{F}_k] = M_k$$

for each $k = 0, 1, \dots, m-1$, then we also have

$$\widetilde{I\!E}[M_m|\mathcal{F}_k] = M_k, k = 0, 1, \dots, m-1. \tag{4.2}$$

When $k = m-1$, the equation (4.2) follows directly from the martingale property. For $k = m-2$, we use the tower property to write

$$\begin{aligned}
\widetilde{I\!E}[M_m|\mathcal{F}_{m-2}] &= \widetilde{I\!E}[\widetilde{I\!E}[M_m|\mathcal{F}_{m-1}]|\mathcal{F}_{m-2}] \\
&= \widetilde{I\!E}[M_{m-1}|\mathcal{F}_{m-2}] \\
&= M_{m-2}.
\end{aligned}$$

We can continue by induction to obtain (4.2).

If the simple European security $V_m$ is hedgeable, then there is a portfolio process whose self-financing value process $X_0, X_1, \ldots, X_m$ satisfies $X_m = V_m$. By definition, $X_k$ is the APT value at time $k$ of $V_m$. Theorem 3.12 says that

$$X_0, (1+r)^{-1}X_1, \ldots, (1+r)^{-m}X_m$$

is a martingale, and so for each $k$,

$$(1+r)^{-k}X_k = \widetilde{I\!E}[(1+r)^{-m}X_m|\mathcal{F}_k] = \widetilde{I\!E}[(1+r)^{-m}V_m|\mathcal{F}_k].$$

Therefore,

$$X_k = (1+r)^k\widetilde{I\!E}[(1+r)^{-m}V_m|\mathcal{F}_k].$$

■

## 3.5 The Binomial Model is Complete

Can a simple European derivative security always be hedged? It depends on the model. If the answer is "yes", the model is said to be *complete.* If the answer is "no", the model is called *incomplete.*

**Theorem 5.14** *The binomial model is complete. In particular, let $V_m$ be a simple European derivative security, and set*

$$V_k(\omega_1, \ldots, \omega_k) = (1+r)^k\widetilde{I\!E}[(1+r)^{-m}V_m|\mathcal{F}_k](\omega_1, \ldots, \omega_k), \tag{5.1}$$

$$\Delta_k(\omega_1, \ldots, \omega_k) = \frac{V_{k+1}(\omega_1, \ldots, \omega_k, H) - V_{k+1}(\omega_1, \ldots, \omega_k, T)}{S_{k+1}(\omega_1, \ldots, \omega_k, H) - S_{k+1}(\omega_1, \ldots, \omega_k, T)}. \tag{5.2}$$

*Starting with initial wealth $V_0 = \widetilde{I\!E}[(1+r)^{-m}V_m]$, the self-financing value of the portfolio process $\Delta_0, \Delta_1, \ldots, \Delta_{m-1}$ is the process $V_0, V_1, \ldots, V_m$.*

**Proof:** Let $V_0, \ldots, V_{m-1}$ and $\Delta_0, \ldots, \Delta_{m-1}$ be defined by (5.1) and (5.2). Set $X_0 = V_0$ and define the self-financing value of the portfolio process $\Delta_0, \ldots, \Delta_{m-1}$ by the recursive formula 3.2:

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)(X_k - \Delta_k S_k).$$

We need to show that

$$X_k = V_k, \quad \forall k \in \{0, 1, \ldots, m\}. \tag{5.3}$$

We proceed by induction. For $k = 0$, (5.3) holds by definition of $X_0$. Assume that (5.3) holds for some value of $k$, i.e., for each fixed $(\omega_1, \ldots, \omega_k)$, we have

$$X_k(\omega_1, \ldots, \omega_k) = V_k(\omega_1, \ldots, \omega_k).$$

We need to show that

$$X_{k+1}(\omega_1,\ldots,\omega_k,H) = V_{k+1}(\omega_1,\ldots,\omega_k,H),$$

$$X_{k+1}(\omega_1,\ldots,\omega_k,T) = V_{k+1}(\omega_1,\ldots,\omega_k,T).$$

We prove the first equality; the second can be shown similarly. Note first that

$$\begin{aligned}\widetilde{I\!\!E}[(1+r)^{-(k+1)}V_{k+1}|\mathcal{F}_k] &= \widetilde{I\!\!E}[\widetilde{I\!\!E}[(1+r)^{-m}V_m|\mathcal{F}_{k+1}]|\mathcal{F}_k]\\ &= \widetilde{I\!\!E}[(1+r)^{-m}V_m|\mathcal{F}_k]\\ &= (1+r)^{-k}V_k\end{aligned}$$

In other words, $\{(1+r)^{-k}V_k\}_{k=0}^n$ is a martingale under $\widetilde{I\!\!P}$. In particular,

$$\begin{aligned}V_k(\omega_1,\ldots,\omega_k) &= \widetilde{I\!\!E}[(1+r)^{-1}V_{k+1}|\mathcal{F}_k](\omega_1,\ldots,\omega_k)\\ &= \frac{1}{1+r}\left(\tilde{p}V_{k+1}(\omega_1,\ldots,\omega_k,H)+\tilde{q}V_{k+1}(\omega_1,\ldots,\omega_k,T)\right).\end{aligned}$$

Since $(\omega_1,\ldots,\omega_k)$ will be fixed for the rest of the proof, we simplify notation by suppressing these symbols. For example, we write the last equation as

$$V_k = \frac{1}{1+r}\left(\tilde{p}V_{k+1}(H)+\tilde{q}V_{k+1}(T)\right).$$

We compute

$$\begin{aligned}&X_{k+1}(H)\\ &= \Delta_k S_{k+1}(H) + (1+r)(X_k - \Delta_k S_k)\\ &= \Delta_k\left(S_{k+1}(H)-(1+r)S_k\right) + (1+r)V_k\\ &= \frac{V_{k+1}(H)-V_{k+1}(T)}{S_{k+1}(H)-S_{k+1}(T)}\left(S_{k+1}(H)-(1+r)S_k\right)\\ &\qquad +\tilde{p}V_{k+1}(H)+\tilde{q}V_{k+1}(T)\\ &= \frac{V_{k+1}(H)-V_{k+1}(T)}{uS_k - dS_k}\left(uS_k-(1+r)S_k\right)\\ &\qquad +\tilde{p}V_{k+1}(H)+\tilde{q}V_{k+1}(T)\\ &= \left(V_{k+1}(H)-V_{k+1}(T)\right)\left(\frac{u-1-r}{u-d}\right)+\tilde{p}V_{k+1}(H)+\tilde{q}V_{k+1}(T)\\ &= \left(V_{k+1}(H)-V_{k+1}(T)\right)\tilde{q}+\tilde{p}V_{k+1}(H)+\tilde{q}V_{k+1}(T)\\ &= V_{k+1}(H).\end{aligned}$$

■

# Chapter 4

# The Markov Property

## 4.1 Binomial Model Pricing and Hedging

Recall that $V_m$ is the given simple European derivative security, and the value and portfolio processes are given by:

$$V_k = (1+r)^k \widetilde{I\!E}[(1+r)^{-m} V_m | \mathcal{F}_k], \quad k = 0, 1, \ldots, m-1.$$

$$\Delta_k(\omega_1, \ldots, \omega_k) = \frac{V_{k+1}(\omega_1, \ldots, \omega_k, H) - V_{k+1}(\omega_1, \ldots, \omega_k, T)}{S_{k+1}(\omega_1, \ldots, \omega_k, H) - S_{k+1}(\omega_1, \ldots, \omega_k, T)}, \quad k = 0, 1, \ldots, m-1.$$

**Example 4.1 (Lookback Option)** $u = 2, d = 0.5, r = 0.25, S_0 = 4, \tilde{p} = \frac{1+r-d}{u-d} = 0.5, \tilde{q} = 1 - \tilde{p} = 0.5.$ Consider a simple European derivative security with expiration 2, with payoff given by (See Fig. 4.1):

$$V_2 = \max_{0 \le k \le 2} (S_k - 5)^+.$$

Notice that

$$V_2(HH) = 11, \;\; V_2(HT) = 3 \neq V_2(TH) = 0, \;\; V_2(TT) = 0.$$

The payoff is thus “path dependent”. Working backward in time, we have:

$$V_1(H) = \frac{1}{1+r}[\tilde{p} V_2(HH) + \tilde{q} V_2(HT)] = \frac{4}{5}[0.5 \times 11 + 0.5 \times 3] = 5.60,$$

$$V_1(T) = \frac{4}{5}[0.5 \times 0 + 0.5 \times 0] = 0,$$

$$V_0 = \frac{4}{5}[0.5 \times 5.60 + 0.5 \times 0] = 2.24.$$

Using these values, we can now compute:

$$\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)} = 0.93,$$

$$\Delta_1(H) = \frac{V_2(HH) - V_2(HT)}{S_2(HH) - S_2(HT)} = 0.67,$$

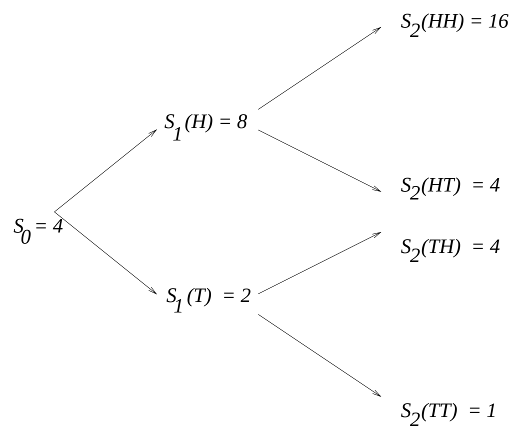



Figure 4.1: *Stock price underlying the lookback option.*

$$\Delta_1(T) = \frac{V_2(TH) - V_2(TT)}{S_2(TH) - S_2(TT)} = 0.$$

Working forward in time, we can check that

$$X_1(H) = \Delta_0 S_1(H) + (1+r)(X_0 - \Delta_0 S_0) = 5.59; \;\; V_1(H) = 5.60,$$

$$X_1(T) = \Delta_0 S_1(T) + (1+r)(X_0 - \Delta_0 S_0) = 0.01; \;\; V_1(T) = 0,$$

$$X_1(HH) = \Delta_1(H) S_1(HH) + (1+r)(X_1(H) - \Delta_1(H) S_1(H)) = 11.01; \;\; V_1(HH) = 11,$$

etc.

■

**Example 4.2 (European Call)** Let $u = 2, d = \frac{1}{2}, r = \frac{1}{4}, S_0 = 4, \tilde{p} = \tilde{q} = \frac{1}{2}$, and consider a European call with expiration time 2 and payoff function

$$V_2 = (S_2 - 5)^+.$$

Note that

$$V_2(HH) = 11, \; V_2(HT) = V_2(TH) = 0, \; V_2(TT) = 0,$$

$$V_1(H) = \frac{4}{5}[\tfrac{1}{2}.11 + \tfrac{1}{2}.0] = 4.40$$

$$V_1(T) = \frac{4}{5}[\tfrac{1}{2}.0 + \tfrac{1}{2}.0] = 0$$

$$V_0 = \frac{4}{5}[\tfrac{1}{2} \times 4.40 + \tfrac{1}{2} \times 0] = 1.76.$$

Define $v_k(x)$ to be the value of the call at time $k$ when $S_k = x$. Then

$$v_2(x) = (x-5)^+$$

$$v_1(x) = \frac{4}{5}[\tfrac{1}{2} v_2(2x) + \tfrac{1}{2} v_2(x/2)],$$

$$v_0(x) = \frac{4}{5}[\tfrac{1}{2} v_1(2x) + \tfrac{1}{2} v_1(x/2)].$$

In particular,

$$v_2(16) = 11, \; v_2(4) = 0, \; v_2(1) = 0,$$

$$v_1(8) = \frac{4}{5}[\tfrac{1}{2}.11 + \tfrac{1}{2}.0] = 4.40,$$

$$v_1(2) = \frac{4}{5}[\tfrac{1}{2}.0 + \tfrac{1}{2}.0] = 0,$$

$$v_0 = \frac{4}{5}[\tfrac{1}{2} \times 4.40 + \tfrac{1}{2} \times 0] = 1.76.$$

Let $\delta_k(x)$ be the number of shares in the hedging portfolio at time $k$ when $S_k = x$. Then

$$\delta_k(x) = \frac{v_{k+1}(2x) - v_{k+1}(x/2)}{2x - x/2}, \quad k = 0, 1.$$

■

## 4.2 Computational Issues

For a model with $n$ periods (coin tosses), $\Omega$ has $2^n$ elements. For period $k$, we must solve $2^k$ equations of the form

$$V_k(\omega_1, \ldots, \omega_k) = \frac{1}{1+r}[\tilde{p}V_{k+1}(\omega_1, \ldots, \omega_k, H) + \tilde{q}V_{k+1}(\omega_1, \ldots, \omega_k, T)].$$

For example, a three-month option has 66 trading days. If each day is taken to be one period, then $n = 66$ and $2^{66} \sim 7 \times 10^{19}$.

There are three possible ways to deal with this problem:

1. Simulation. We have, for example, that

   $$V_0 = (1+r)^{-n}\widetilde{I\!\!E}V_n,$$

   and so we could compute $V_0$ by simulation. More specifically, we could simulate $n$ coin tosses $\omega = (\omega_1, \ldots, \omega_n)$ under the risk-neutral probability measure. We could store the value of $V_n(\omega)$. We could repeat this several times and take the average value of $V_n$ as an approximation to $\widetilde{I\!\!E}V_n$.

2. Approximate a many-period model by a continuous-time model. Then we can use calculus and partial differential equations. We'll get to that.

3. Look for Markov structure. Example 4.2 has this. In period 2, the option in Example 4.2 has three possible values $v_2(16), v_2(4), v_2(1)$, rather than four possible values $V_2(HH), V_2(HT), V_2(TH), V_2(TT)$. If there were 66 periods, then in period 66 there would be 67 possible stock price values (since the final price depends only on the *number* of up-ticks of the stock price – i.e., heads – so far) and hence only 67 possible option values, rather than $2^{66} \sim 7 \times 10^{19}$.

## 4.3 Markov Processes

**Technical condition always present:** We consider only functions on R and subsets of R which are Borel-measurable, i.e., we only consider subsets $A$ of R that are in $\mathcal{B}$ and functions $g : \mathbb{R} \to \mathbb{R}$ such that $g^{-1}$ is a function $\mathcal{B} \to \mathcal{B}$.

**Definition 4.1 ()** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\{\mathcal{F}_k\}_{k=0}^n$ be a filtration under $\mathcal{F}$. Let $\{X_k\}_{k=0}^n$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. This process is said to be *Markov* if:

- The stochastic process $\{X_k\}$ is adapted to the filtration $\{\mathcal{F}_k\}$, and
- *(The Markov Property).* For each $k = 0, 1, \ldots, n-1$, the distribution of $X_{k+1}$ conditioned on $\mathcal{F}_k$ is the same as the distribution of $X_{k+1}$ conditioned on $X_k$.

### 4.3.1 Different ways to write the Markov property

**(a)** (Agreement of distributions). For every $A \in \mathcal{B} \stackrel{\triangle}{=} \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned} \mathbb{P}(X_{k+1} \in A | \mathcal{F}_k) &= \mathbb{E}[I_A(X_{k+1}) | \mathcal{F}_k] \\ &= \mathbb{E}[I_A(X_{k+1}) | X_k] \\ &= \mathbb{P}[X_{k+1} \in A | X_k]. \end{aligned}$$

**(b)** (Agreement of expectations of all functions). For every (Borel-measurable) function $h : \mathbb{R} \to \mathbb{R}$ for which $\mathbb{E}|h(X_{k+1})| < \infty$, we have

$$\mathbb{E}[h(X_{k+1}) | \mathcal{F}_k] = \mathbb{E}[h(X_{k+1}) | X_k].$$

**(c)** (Agreement of Laplace transforms.) For every $u \in \mathbb{R}$ for which $\mathbb{E}e^{uX_{k+1}} < \infty$, we have

$$\mathbb{E}\left[ e^{uX_{k+1}} \middle| \mathcal{F}_k \right] = \mathbb{E}\left[ e^{uX_{k+1}} \middle| X_k \right].$$

(If we fix $u$ and define $h(x) = e^{ux}$, then the equations in (b) and (c) are the same. However in (b) we have a condition which holds for *every* function $h$, and in (c) we assume this condition only for functions $h$ of the form $h(x) = e^{ux}$. A main result in the theory of Laplace transforms is that if the equation holds for every $h$ of this special form, then it holds for every $h$, i.e., (c) implies (b).)

**(d)** (Agreement of characteristic functions) For every $u \in \mathbb{R}$, we have

$$\mathbb{E}\left[ e^{iuX_{k+1}} | \mathcal{F}_k \right] = \mathbb{E}\left[ e^{iuX_{k+1}} | X_k \right],$$

where $i = \sqrt{-1}$. (Since $|e^{iux}| = |\cos x + \sin x| \leq 1$ we don't need to assume that $\mathbb{E}|e^{iux}| < \infty$.)

**Remark 4.1** In every case of the Markov properties where $I\!E[\dots|X_k]$ appears, we could just as well write $g(X_k)$ for some function $g$. For example, form (a) of the Markov property can be restated as:

> For every $A \in \mathcal{B}$, we have
>
> $$I\!P(X_{k+1} \in A|\mathcal{F}_k) = g(X_k),$$
>
> where $g$ is a function that depends on the set $A$.

Conditions (a)-(d) are equivalent. The Markov property as stated in (a)-(d) involves the process at a "current" time $k$ and one future time $k+1$. Conditions (a)-(d) are also equivalent to conditions involving the process at time $k$ and multiple future times. We write these apparently stronger but actually equivalent conditions below.

**Consequences of the Markov property.** Let $j$ be a positive integer.

**(A)** For every $A_{k+1} \subset I\!R, \dots, \mathcal{A}_{k+j} \subset I\!R$,

$$I\!P[X_{k+1} \in A_{k+1}, \dots, X_{k+j} \in A_{k+j}|\mathcal{F}_k] = I\!P[X_{k+1} \in A_{k+1}, \dots, X_{k+j} \in A_{k+j}|X_k].$$

**(A')** For every $A \in I\!R^j$,

$$I\!P[(X_{k+1}, \dots, X_{k+j}) \in A|\mathcal{F}_k] = I\!P[(X_{k+1}, \dots, X_{k+j}) \in A|X_k].$$

**(B)** For every function $h : I\!R^j \to I\!R$ for which $I\!E|h(X_{k+1}, \dots, X_{k+j})| < \infty$, we have

$$I\!E[h(X_{k+1}, \dots, X_{k+j})|\mathcal{F}_k] = I\!E[h(X_{k+1}, \dots, X_{k+j})|X_k].$$

**(C)** For every $u = (u_{k+1}, \dots, u_{k+j}) \in I\!R^j$ for which $I\!E|e^{u_{k+1}X_{k+1}+\dots+u_{k+j}X_{k+j}}| < \infty$, we have

$$I\!E[e^{u_{k+1}X_{k+1}+\dots+u_{k+j}X_{k+j}}|\mathcal{F}_k] = I\!E[e^{u_{k+1}X_{k+1}+\dots+u_{k+j}X_{k+j}}|X_k].$$

**(D)** For every $u = (u_{k+1}, \dots, u_{k+j}) \in I\!R^j$ we have

$$I\!E[e^{i(u_{k+1}X_{k+1}+\dots+u_{k+j}X_{k+j})}|\mathcal{F}_k] = I\!E[e^{i(u_{k+1}X_{k+1}+\dots+u_{k+j}X_{k+j})}|X_k].$$

Once again, every expression of the form $I\!E(\dots|X_k)$ can also be written as $g(X_k)$, where the function $g$ depends on the random variable represented by $\dots$ in this expression.

**Remark.** All these Markov properties have analogues for vector-valued processes.

**Proof that (b) $\Longrightarrow$ (A).** (with $j = 2$ in (A)) Assume (b). Then (a) also holds (take $h = I_A$). Consider

$$
\begin{aligned}
& I\!P[X_{k+1} \in A_{k+1}, X_{k+2} \in A_{k+2} | \mathcal{F}_k] \\
&= I\!E[I_{A_{k+1}}(X_{k+1}) I_{A_{k+2}}(X_{k+2}) | \mathcal{F}_k] \\
&\qquad \text{(Definition of conditional probability)} \\
&= I\!E[I\!E[I_{A_{k+1}}(X_{k+1}) I_{A_{k+2}}(X_{k+2}) | \mathcal{F}_{k+1}] | \mathcal{F}_k] \\
&\qquad \text{(Tower property)} \\
&= I\!E[I_{A_{k+1}}(X_{k+1}).I\!E[I_{A_{k+2}}(X_{k+2}) | \mathcal{F}_{k+1}] | \mathcal{F}_k] \\
&\qquad \text{(Taking out what is known)} \\
&= I\!E[I_{A_{k+1}}(X_{k+1}).I\!E[I_{A_{k+2}}(X_{k+2}) | X_{k+1}] | \mathcal{F}_k] \\
&\qquad \text{(Markov property, form (a).)} \\
&= I\!E[I_{A_{k+1}}(X_{k+1}).g(X_{k+1}) | \mathcal{F}_k] \\
&\qquad \text{(Remark 4.1)} \\
&= I\!E[I_{A_{k+1}}(X_{k+1}).g(X_{k+1}) | X_k] \\
&\qquad \text{(Markov property, form (b).)}
\end{aligned}
$$

Now take conditional expectation on both sides of the above equation, conditioned on $\sigma(X_k)$, and use the tower property on the left, to obtain

$$I\!P[X_{k+1} \in A_{k+1}, X_{k+2} \in A_{k+2} | X_k] = I\!E[I_{A_{k+1}}(X_{k+1}).g(X_{k+1}) | X_k]. \tag{3.1}$$

Since both

$$I\!P[X_{k+1} \in A_{k+1}, X_{k+2} \in A_{k+2} | \mathcal{F}_k]$$

and

$$I\!P[X_{k+1} \in A_{k+1}, X_{k+2} \in A_{k+2} | X_k]$$

are equal to the RHS of (3.1)), they are equal to each other, and this is property (A) with $j = 2$. ■

**Example 4.3** It is intuitively clear that the stock price process in the binomial model is a Markov process. We will formally prove this later. If we want to estimate the distribution of $S_{k+1}$ based on the information in $\mathcal{F}_k$, the only relevant piece of information is the value of $S_k$. For example,

$$\widetilde{I\!E}[S_{k+1} | \mathcal{F}_k] = (\tilde{p}u + \tilde{q}d)S_k = (1 + r)S_k \tag{3.2}$$

is a function of $S_k$. Note however that form (b) of the Markov property is stronger then (3.2); the Markov property requires that for *any* function $h$,

$$\widetilde{I\!E}[h(S_{k+1}) | \mathcal{F}_k]$$

is a function of $S_k$. Equation (3.2) is the case of $h(x) = x$.

Consider a model with 66 periods and a simple European derivative security whose payoff at time 66 is

$$V_{66} = \frac{1}{3}(S_{64} + S_{65} + S_{66}).$$

The value of this security at time 50 is

$$\begin{aligned} V_{50} &= (1+r)^{50}\widetilde{I\!\!E}[(1+r)^{-66}V_{66}|\mathcal{F}_{50}] \\ &= (1+r)^{-16}\widetilde{I\!\!E}[V_{66}|S_{50}], \end{aligned}$$

because the stock price process is Markov. (We are using form (B) of the Markov property here). In other words, the $F_{50}$-measurable random variable $V_{50}$ can be written as

$$V_{50}(\omega_1,\dots,\omega_{50}) = g(S_{50}(\omega_1,\dots,\omega_{50}))$$

for some function $g$, which we can determine with a bit of work. ■

## 4.4 Showing that a process is Markov

**Definition 4.2 (Independence)** Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, and let $\mathcal{G}$ and $\mathcal{H}$ be sub-$\sigma$-algebras of $\mathcal{F}$. We say that $\mathcal{G}$ and $\mathcal{H}$ are *independent* if for every $A\in\mathcal{G}$ and $B\in\mathcal{H}$, we have

$$I\!\!P(A\cap B) = I\!\!P(A)I\!\!P(B).$$

We say that a random variable $X$ is independent of a $\sigma$-algebra $\mathcal{G}$ if $\sigma(X)$, the $\sigma$-algebra generated by $X$, is independent of $\mathcal{G}$.

**Example 4.4** Consider the two-period binomial model. Recall that $\mathcal{F}_1$ is the $\sigma$-algebra of sets determined by the first toss, i.e., $\mathcal{F}_1$ contains the four sets

$$A_H \stackrel{\triangle}{=} \{HH,HT\},\ A_T \stackrel{\triangle}{=} \{TH,TT\},\ \phi,\ \Omega.$$

Let $\mathcal{H}$ be the $\sigma$-algebra of sets determined by the second toss, i.e., $\mathcal{H}$ contains the four sets

$$\{HH,TH\},\{HT,TT\},\phi,\Omega.$$

Then $\mathcal{F}_1$ and $\mathcal{H}$ are independent. For example, if we take $A=\{HH,HT\}$ from $\mathcal{F}_1$ and $B=\{HH,TH\}$ from $\mathcal{H}$, then $I\!\!P(A\cap B)=I\!\!P(HH)=p^2$ and

$$I\!\!P(A)I\!\!P(B) = (p^2+pq)(p^2+pq) = p^2(p+q)^2 = p^2.$$

Note that $\mathcal{F}_1$ and $S_2$ are not independent (unless $p=1$ or $p=0$). For example, one of the sets in $\sigma(S_2)$ is $\{\omega; S_2(\omega)=u^2S_0\}=\{HH\}$. If we take $A=\{HH,HT\}$ from $\mathcal{F}_1$ and $B=\{HH\}$ from $\sigma(S_2)$, then $I\!\!P(A\cap B)=I\!\!P(HH)=p^2$, but

$$I\!\!P(A)I\!\!P(B) = (p^2+pq)p^2 = p^3(p+q) = p^3.$$

■

The following lemma will be very useful in showing that a process is Markov:

**Lemma 4.15 (Independence Lemma)** *Let $X$ and $Y$ be random variables on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. Assume*

- *$X$ is independent of $\mathcal{G}$;*
- *$Y$ is $\mathcal{G}$-measurable.*

*Let $f(x,y)$ be a function of two variables, and define*

$$g(y) \triangleq \mathbb{E} f(X,y).$$

*Then*

$$\mathbb{E}[f(X,Y)|\mathcal{G}] = g(Y).$$

**Remark.** In this lemma and the following discussion, capital letters denote random variables and lower case letters denote nonrandom variables.

**Example 4.5 (Showing the stock price process is Markov)** Consider an $n$-period binomial model. Fix a time $k$ and define $X \triangleq \frac{S_{k+1}}{S_k}$ and $\mathcal{G} \triangleq \mathcal{F}_k$. Then $X = u$ if $\omega_{k+1} = H$ and $X = d$ if $\omega_{k+1} = T$. Since $X$ depends only on the $(k+1)$st toss, $X$ is independent of $\mathcal{G}$. Define $Y \triangleq S_k$, so that $Y$ is $\mathcal{G}$-measurable. Let $h$ be any function and set $f(x,y) \triangleq h(xy)$. Then

$$g(y) \triangleq \mathbb{E} f(X,y) = \mathbb{E} h(Xy) = ph(uy) + qh(dy).$$

The Independence Lemma asserts that

$$\begin{aligned}
\mathbb{E}[h(S_{k+1})|\mathcal{F}_k] &= \mathbb{E}[h\left(\frac{S_{k+1}}{S_k}.S_k\right)|\mathcal{F}_k] \\
&= \mathbb{E}[f(X,Y)|\mathcal{G}] \\
&= g(Y) \\
&= ph(uS_k) + qh(dS_k).
\end{aligned}$$

This shows the stock price is Markov. Indeed, if we condition both sides of the above equation on $\sigma(S_k)$ and use the tower property on the left and the fact that the right hand side is $\sigma(S_k)$-measurable, we obtain

$$\mathbb{E}[h(S_{k+1})|S_k] = ph(uS_k) + qh(dS_k).$$

Thus $\mathbb{E}[h(S_{k+1})|\mathcal{F}_k]$ and $\mathbb{E}[h(S_{k+1})|X_k]$ are equal and form (b) of the Markov property is proved.

Not only have we shown that the stock price process is Markov, but we have also obtained a formula for $\mathbb{E}[h(S_{k+1})|\mathcal{F}_k]$ as a function of $S_k$. This is a special case of Remark 4.1.

■

## 4.5 Application to Exotic Options

Consider an $n$-period binomial model. Define the *running maximum* of the stock price to be

$$M_k \triangleq \max_{1 \le j \le k} S_j.$$

Consider a simple European derivative security with payoff at time $n$ of $v_n(S_n, M_n)$.

**Examples:**

- $v_n(S_n, M_n) = (M_n - K)^+$ (Lookback option);
- $v_n(S_n, M_n) = I_{M_n \geq B}(S_n - K)^+$ (Knock-in Barrier option).

**Lemma 5.16** *The two-dimensional process $\{(S_k, M_k)\}_{k=0}^n$ is Markov. (Here we are working under the risk-neutral measure* **P**, *although that does not matter).*

**Proof:** Fix $k$. We have

$$M_{k+1} = M_k \vee S_{k+1},$$

where $\vee$ indicates the maximum of two quantities. Let $Z \triangleq \frac{S_{k+1}}{S_k}$, so

$$\widetilde{I\!P}(Z = u) = \tilde{p}, \quad \widetilde{I\!P}(Z = d) = \tilde{q},$$

and $Z$ is independent of $\mathcal{F}_k$. Let $h(x, y)$ be a function of two variables. We have

$$\begin{aligned} h(S_{k+1}, M_{k+1}) &= h(S_{k+1}, M_k \vee S_{k+1}) \\ &= h(ZS_k, M_k \vee (ZS_k)). \end{aligned}$$

Define

$$\begin{aligned} g(x, y) &\triangleq \widetilde{I\!E} h(Zx, y \vee (Zx)) \\ &= \tilde{p} h(ux, y \vee (ux)) + \tilde{q} h(dx, y \vee (dx)). \end{aligned}$$

The Independence Lemma implies

$$\widetilde{I\!E}[h(S_{k+1}, M_{k+1})|\mathcal{F}_k] = g(S_k, M_k) = \tilde{p} h(uS_k, M_k \vee (uS_k)) + \tilde{q} h(dS_k, M_k),$$

the second equality being a consequence of the fact that $M_k \wedge dS_k = M_k$. Since the RHS is a function of $(S_k, M_k)$, we have proved the Markov property (form (b)) for this two-dimensional process. ■

Continuing with the exotic option of the previous Lemma... Let $V_k$ denote the value of the derivative security at time $k$. Since $(1+r)^{-k} V_k$ is a martingale under $\widetilde{I\!P}$, we have

$$V_k = \frac{1}{1+r} \widetilde{I\!E}[V_{k+1}|\mathcal{F}_k], k = 0, 1, \ldots, n-1.$$

At the final time, we have

$$V_n = v_n(S_n, M_n).$$

Stepping back one step, we can compute

$$\begin{aligned} V_{n-1} &= \frac{1}{1+r} \widetilde{I\!E}[v_n(S_n, M_n)|\mathcal{F}_{n-1}] \\ &= \frac{1}{1+r} \left[\tilde{p} v_n(uS_{n-1}, uS_{n-1} \vee M_{n-1}) + \tilde{q} v_n(dS_{n-1}, M_{n-1})\right]. \end{aligned}$$

This leads us to define

$$v_{n-1}(x, y) \stackrel{\triangle}{=} \frac{1}{1+r}\left[\tilde{p} v_n(ux, ux \vee y) + \tilde{q} v_n(dx, y)\right]$$

so that

$$V_{n-1} = v_{n-1}(S_{n-1}, M_{n-1}).$$

The general algorithm is

$$v_k(x, y) = \frac{1}{1+r}\left[\tilde{p} v_{k+1}(ux, ux \vee y) + \tilde{q} v_{k+1}(dx, y)\right],$$

and the value of the option at time $k$ is $v_k(S_k, M_k)$. Since this is a simple European option, the hedging portfolio is given by the usual formula, which in this case is

$$\Delta_k = \frac{v_{k+1}(uS_k, (uS_k) \vee M_k) - v_{k+1}(dS_k, M_k)}{(u-d)S_k}$$

# Chapter 5

# Stopping Times and American Options

## 5.1 American Pricing

Let us first review the **European pricing formula in a Markov model**. Consider the Binomial model with $n$ periods. Let $V_n = g(S_n)$ be the payoff of a derivative security. Define by backward recursion:

$$\begin{aligned} v_n(x) &= g(x) \\ v_k(x) &= \frac{1}{1+r}[\tilde{p}v_{k+1}(ux) + \tilde{q}v_{k+1}(dx)]. \end{aligned}$$

Then $v_k(S_k)$ is the value of the option at time $k$, and the hedging portfolio is given by

$$\Delta_k = \frac{v_{k+1}(uS_k) - v_{k+1}(dS_k)}{(u-d)S_k}, \quad k = 0, 1, 2, \dots, n-1.$$

Now consider an American option. Again a function $g$ is specified. In any period $k$, the holder of the derivative security can "exercise" and receive payment $g(S_k)$. Thus, the hedging portfolio should create a wealth process which satisfies

$$X_k \geq g(S_k), \forall k, \quad \text{almost surely.}$$

This is because the value of the derivative security at time $k$ is at least $g(S_k)$, and the wealth process value at that time must equal the value of the derivative security.

**American algorithm.**

$$\begin{aligned} v_n(x) &= g(x) \\ v_k(x) &= \max\left\{\frac{1}{1+r}(\tilde{p}v_{k+1}(ux) + \tilde{q}v_{k+1}(dx)),\ g(x)\right\} \end{aligned}$$

Then $v_k(S_k)$ is the value of the option at time $k$.

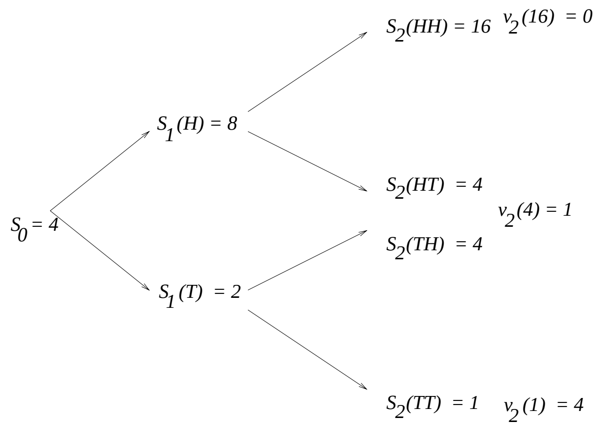



Figure 5.1: *Stock price and final value of an American put option with strike price 5.*

**Example 5.1** See Fig. 5.1. $S_0 = 4, u = 2, d = \frac{1}{2}, r = \frac{1}{4}, \tilde{p} = \tilde{q} = \frac{1}{2}, n = 2$. Set $v_2(x) = g(x) = (5-x)^+$. Then

$$\begin{aligned}
v_1(8) &= \max\left\{\frac{4}{5}\left[\tfrac{1}{2}.0 + \tfrac{1}{2}.1\right], (5-8)^+\right\} \\
&= \max\left\{\frac{2}{5}, 0\right\} \\
&= 0.40 \\
v_1(2) &= \max\left\{\frac{4}{5}\left[\tfrac{1}{2}.1 + \tfrac{1}{2}.4\right], (5-2)^+\right\} \\
&= \max\{2, 3\} \\
&= 3.00 \\
v_0(4) &= \max\left\{\frac{4}{5}\left[\tfrac{1}{2}.(0.4) + \tfrac{1}{2}.(3.0)\right], (5-4)^+\right\} \\
&= \max\{1.36, 1\} \\
&= 1.36
\end{aligned}$$

Let us now construct the hedging portfolio for this option. Begin with initial wealth $X_0 = 1.36$. Compute $\Delta_0$ as follows:

$$\begin{aligned}
0.40 &= v_1(S_1(H)) \\
&= S_1(H)\Delta_0 + (1+r)(X_0 - \Delta_0 S_0) \\
&= 8\Delta_0 + \frac{5}{4}(1.36 - 4\Delta_0) \\
&= 3\Delta_0 + 1.70 \Longrightarrow \Delta_0 = -0.43 \\
3.00 &= v_1(S_1(T)) \\
&= S_1(T)\Delta_0 + (1+r)(X_0 - \Delta_0 S_0) \\
&= 2\Delta_0 + \frac{5}{4}(1.36 - 4\Delta_0) \\
&= -3\Delta_0 + 1.70 \Longrightarrow \Delta_0 = -0.43
\end{aligned}$$

Using $\Delta_0 = -0.43$ results in

$$X_1(H) = v_1(S_1(H)) = 0.40, \quad X_1(T) = v_1(S_1(T)) = 3.00$$

Now let us compute $\Delta_1$ (Recall that $S_1(T) = 2$):

$$\begin{aligned}
1 &= v_2(4) \\
&= S_2(TH)\Delta_1(T) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)) \\
&= 4\Delta_1(T) + \frac{5}{4}(3 - 2\Delta_1(T)) \\
&= 1.5\Delta_1(T) + 3.75 \Longrightarrow \Delta_1(T) = -1.83 \\
4 &= v_2(1) \\
&= S_2(TT)\Delta_1(T) + (1+r)(X_1(T) - \Delta_1(T)S_1(T)) \\
&= \Delta_1(T) + \frac{5}{4}(3 - 2\Delta_1(T)) \\
&= -1.5\Delta_1(T) + 3.75 \Longrightarrow \Delta_1(T) = -0.16
\end{aligned}$$

We get different answers for $\Delta_1(T)$! If we had $X_1(T) = 2$, the value of the *European* put, we would have

$$1 = 1.5\Delta_1(T) + 2.5 \Longrightarrow \Delta_1(T) = -1,$$

$$4 = -1.5\Delta_1(T) + 2.5 \Longrightarrow \Delta_1(T) = -1,$$

■

## 5.2 Value of Portfolio Hedging an American Option

$$\begin{aligned}
X_{k+1} &= \Delta_k S_{k+1} + (1+r)(X_k - C_k - \Delta_k S_k) \\
&= (1+r)X_k + \Delta_k(S_{k+1} - (1+r)S_k) - (1+r)C_k
\end{aligned}$$

Here, $C_k$ is the amount "consumed" at time $k$.

- The discounted value of the portfolio is a *supermartingale*.
- The value satisfies $X_k \geq g(S_k),\, k = 0, 1, \ldots, n$.
- The value process is the smallest process with these properties.

When do you consume? If

$$\widetilde{I\!\!E}((1+r)^{-(k+1)} v_{k+1}(S_{k+1}) | \mathcal{F}_k] < (1+r)^{-k} v_k(S_k),$$

or, equivalently,

$$\widetilde{I\!\!E}(\frac{1}{1+r} v_{k+1}(S_{k+1}) | \mathcal{F}_k] < v_k(S_k)$$

and the holder of the American option does not exercise, then the seller of the option can consume to close the gap. By doing this, he can ensure that $X_k = v_k(S_k)$ for all $k$, where $v_k$ is the value defined by the American algorithm in Section 5.1.

In the previous example, $v_1(S_1(T)) = 3$, $v_2(S_2(TH)) = 1$ and $v_2(S_2(TT)) = 4$. Therefore,

$$\begin{aligned}
\widetilde{I\!E}[\frac{1}{1+r}v_2(S_2)|\mathcal{F}_1](T) &= \frac{4}{5}\left[\tfrac{1}{2}.1 + \tfrac{1}{2}.4\right] \\
&= \frac{4}{5}\left[\frac{5}{2}\right] \\
&= 2, \\
v_1(S_1(T)) &= 3,
\end{aligned}$$

so there is a gap of size 1. If the owner of the option does not exercise it at time one in the state $\omega_1 = T$, then the seller can consume 1 at time 1. Thereafter, he uses the usual hedging portfolio

$$\Delta_k = \frac{v_{k+1}(uS_k) - v_{k+1}(dS_k)}{(u-d)S_k}$$

In the example, we have $v_1(S_1(T)) = g(S_1(T))$. It is optimal for the owner of the American option to exercise whenever its value $v_k(S_k)$ agrees with its intrinsic value $g(S_k)$.

**Definition 5.1 (Stopping Time)** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{\mathcal{F}_k\}_{k=0}^n$ be a filtration. A *stopping time* is a random variable $\tau : \Omega \rightarrow \{0, 1, 2, \ldots, n\} \cup \{\infty\}$ with the property that:

$$\{\omega \in \Omega; \tau(\omega) = k\} \in \mathcal{F}_k, \ \forall k = 0, 1, \ldots, n, \infty.$$

**Example 5.2** Consider the binomial model with $n = 2, S_0 = 4, u = 2, d = \frac{1}{2}, r = \frac{1}{4}$, so $\tilde{p} = \tilde{q} = \frac{1}{2}$. Let $v_0, v_1, v_2$ be the value functions defined for the American put with strike price 5. Define

$$\tau(\omega) = \min\{k; v_k(S_k) = (5 - S_k)^+\}.$$

The stopping time $\tau$ corresponds to "stopping the first time the value of the option agrees with its intrinsic value". It is an optimal exercise time. We note that

$$\tau(\omega) = \begin{cases} 1 & \text{if } \omega \in A_T \\ 2 & \text{if } \omega \in A_H \end{cases}$$

We verify that $\tau$ is indeed a stopping time:

$$\begin{aligned}
\{\omega; \tau(\omega) = 0\} &= \phi \in \mathcal{F}_0 \\
\{\omega; \tau(\omega) = 1\} &= A_T \in \mathcal{F}_1 \\
\{\omega; \tau(\omega) = 2\} &= A_H \in \mathcal{F}_2
\end{aligned}$$

■

**Example 5.3 (A random time which is not a stopping time)** In the same binomial model as in the previous example, define

$$\rho(\omega) = \min\{k; S_k(\omega) = m_2(\omega)\},$$

where $m_2 \triangleq \min_{0 \le j \le 2} S_j$. In other words, $\rho$ stops when the stock price reaches its minimum value. This random variable is given by

$$\rho(\omega) = \begin{cases} 0 & \text{if } \omega \in A_H, \\ 1 & \text{if } \omega = TH, \\ 2 & \text{if } \omega = TT \end{cases}$$

We verify that $\rho$ is *not* a stopping time:

$$\begin{aligned} \{\omega; \rho(\omega) = 0\} &= A_H \notin \mathcal{F}_0 \\ \{\omega; \rho(\omega) = 1\} &= \{TH\} \notin \mathcal{F}_1 \\ \{\omega; \rho(\omega) = 2\} &= \{TT\} \in \mathcal{F}_2 \end{aligned}$$

■

## 5.3 Information up to a Stopping Time

**Definition 5.2** Let $\tau$ be a stopping time. We say that a set $A \subset \Omega$ is *determined by time* $\tau$ provided that

$$A \cap \{\omega; \tau(\omega) = k\} \in \mathcal{F}_k, \forall k.$$

The collection of sets determined by $\tau$ is a $\sigma$-algebra, which we denote by $\mathcal{F}_\tau$.

**Example 5.4** In the binomial model considered earlier, let

$$\tau = \min\{k; v_k(S_k) = (5 - S_k)^+\},$$

i.e.,

$$\tau(\omega) = \begin{cases} 1 & \text{if } \omega \in A_T \\ 2 & \text{if } \omega \in A_H \end{cases}$$

The set $\{HT\}$ is determined by time $\tau$, but the set $\{TH\}$ is not. Indeed,

$$\begin{aligned} \{HT\} \cap \{\omega; \tau(\omega) = 0\} &= \phi \in \mathcal{F}_0 \\ \{HT\} \cap \{\omega; \tau(\omega) = 1\} &= \phi \in \mathcal{F}_1 \\ \{HT\} \cap \{\omega; \tau(\omega) = 2\} &= \{HT\} \in \mathcal{F}_2 \end{aligned}$$

but

$$\{TH\} \cap \{\omega; \tau(\omega) = 1\} = \{TH\} \notin \mathcal{F}_1.$$

The atoms of $\mathcal{F}_\tau$ are

$$\{HT\},\ \{HH\},\ A_T = \{TH, TT\}.$$

■

**Notation 5.1 (Value of Stochastic Process at a Stopping Time)** If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $\{\mathcal{F}_k\}_{k=0}^n$ is a filtration under $\mathcal{F}$, $\{X_k\}_{k=0}^n$ is a stochastic process adapted to this filtration, and $\tau$ is a stopping time with respect to the same filtration, then $X_\tau$ is an $\mathcal{F}_\tau$-measurable random variable whose value at $\omega$ is given by

$$X_\tau(\omega) \triangleq X_{\tau(\omega)}(\omega).$$

**Theorem 3.17 (Optional Sampling)** *Suppose that* $\{Y_k, \mathcal{F}_k\}_{k=0}^{\infty}$ *(or* $\{Y_k, \mathcal{F}_k\}_{k=0}^{n}$*) is a submartingale. Let* $\tau$ *and* $\rho$ *be* bounded *stopping times, i.e., there is a nonrandom number* $n$ *such that*

$$\tau \le n, \ \rho \le n, \quad \textit{almost surely.}$$

*If* $\tau \le \rho$ *almost surely, then*

$$Y_\tau \le I\!\!E(Y_\rho | \mathcal{F}_\tau).$$

*Taking expectations, we obtain* $I\!\!EY_\tau \le I\!\!EY_\rho$*, and in particular,* $Y_0 = I\!\!EY_0 \le I\!\!EY_\rho$*. If* $\{Y_k, \mathcal{F}_k\}_{k=0}^{\infty}$ *is a supermartingale, then* $\tau \le \rho$ *implies* $Y_\tau \ge I\!\!E(Y_\rho|\mathcal{F}_\tau)$*.*
*If* $\{Y_k, \mathcal{F}_k\}_{k=0}^{\infty}$ *is a martingale, then* $\tau \le \rho$ *implies* $Y_\tau = I\!\!E(Y_\rho|\mathcal{F}_\tau)$*.*

**Example 5.5** In the example 5.4 considered earlier, we define $\rho(\omega) = 2$ for all $\omega \in \Omega$. Under the risk-neutral probability measure, the discounted stock price process $(\frac{5}{4})^{-k} S_k$ is a martingale. We compute

$$\widetilde{I\!\!E}\left[\left(\frac{4}{5}\right)^2 S_2 \middle| \mathcal{F}_\tau\right].$$

The atoms of $\mathcal{F}_\tau$ are $\{HH\}, \{HT\}$, and $A_T$. Therefore,

$$\begin{aligned}
\widetilde{I\!\!E}\left[\left(\frac{4}{5}\right)^2 S_2 \middle| \mathcal{F}_\tau\right](HH) &= \left(\frac{4}{5}\right)^2 S_2(HH), \\
\widetilde{I\!\!E}\left[\left(\frac{4}{5}\right)^2 S_2 \middle| \mathcal{F}_\tau\right](HT) &= \left(\frac{4}{5}\right)^2 S_2(HT),
\end{aligned}$$

and for $\omega \in A_T$,

$$\begin{aligned}
\widetilde{I\!\!E}\left[\left(\frac{4}{5}\right)^2 S_2 \middle| \mathcal{F}_\tau\right](\omega) &= \tfrac{1}{2}\left(\frac{4}{5}\right)^2 S_2(TH) + \tfrac{1}{2}\left(\frac{4}{5}\right)^2 S_2(TT) \\
&= \tfrac{1}{2} \times 2.56 + \tfrac{1}{2} \times 0.64 \\
&= 1.60
\end{aligned}$$

In every case we have gotten (see Fig. 5.2)

$$\widetilde{I\!\!E}\left[\left(\frac{4}{5}\right)^2 S_2 \middle| \mathcal{F}_\tau\right](\omega) = \left(\frac{4}{5}\right)^{\tau(\omega)} S_{\tau(\omega)}(\omega).$$

■

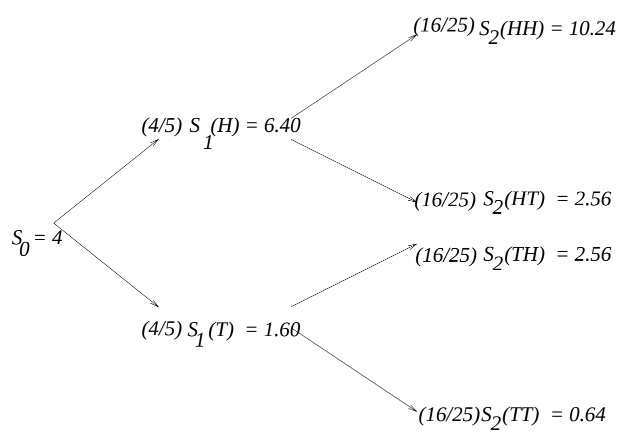



Figure 5.2: *Illustrating the optional sampling theorem.*

# Chapter 6

# Properties of American Derivative Securities

## 6.1 The properties

**Definition 6.1** An *American derivative security* is a sequence of non-negative random variables $\{G_k\}_{k=0}^n$ such that each $G_k$ is $\mathcal{F}_k$-measurable. The owner of an American derivative security can exercise at any time $k$, and if he does, he receives the payment $G_k$.

**(a)** The value $V_k$ of the security at time $k$ is

$$V_k = \max_{\tau} (1+r)^k \widetilde{I\!E}[(1+r)^{-\tau} G_\tau | \mathcal{F}_k],$$

where the maximum is over all stopping times $\tau$ satisfying $\tau \geq k$ almost surely.

**(b)** The discounted value process $\{(1+r)^{-k} V_k\}_{k=0}^n$ is the smallest supermartingale which satisfies

$$V_k \geq G_k, \forall k, \quad \text{almost surely.}$$

**(c)** Any stopping time $\tau$ which satisfies

$$V_0 = \widetilde{I\!E}[(1+r)^{-\tau} G_\tau]$$

is an optimal exercise time. In particular

$$\tau \stackrel{\triangle}{=} \min\{k; V_k = G_k\}$$

is an optimal exercise time.

**(d)** The hedging portfolio is given by

$$\Delta_k(\omega_1, \ldots, \omega_k) = \frac{V_{k+1}(\omega_1, \ldots, \omega_k, H) - V_{k+1}(\omega_1, \ldots, \omega_k, T)}{S_{k+1}(\omega_1, \ldots, \omega_k, H) - S_{k+1}(\omega_1, \ldots, \omega_k, T)}, k = 0, 1, \ldots, n-1.$$

**(e)** Suppose for some $k$ and $\omega$, we have $V_k(\omega) = G_k(\omega)$. Then the owner of the derivative security should exercise it. If he does not, then the seller of the security can immediately consume

$$V_k(\omega) - \frac{1}{1+r}\widetilde{I\!E}[V_{k+1}|\mathcal{F}_k](\omega)$$

and still maintain the hedge.

## 6.2 Proofs of the Properties

Let $\{G_k\}_{k=0}^n$ be a sequence of non-negative random variables such that each $G_k$ is $\mathcal{F}_k$-measurable. Define $T_k$ to be the set of all stopping times $\tau$ satisfying $k \le \tau \le n$ almost surely. Define also

$$V_k \stackrel{\triangle}{=} (1+r)^k \max_{\tau \in T_k} \widetilde{I\!E}\left[(1+r)^{-\tau} G_\tau | \mathcal{F}_k\right].$$

**Lemma 2.18** $V_k \ge G_k$ *for every* $k$.

**Proof:** Take $\tau \in T_k$ to be the constant $k$. ■

**Lemma 2.19** *The process* $\{(1+r)^{-k} V_k\}_{k=0}^n$ *is a supermartingale.*

**Proof:** Let $\tau^*$ attain the maximum in the definition of $V_{k+1}$, i.e.,

$$(1+r)^{-(k+1)} V_{k+1} = \widetilde{I\!E}\left[(1+r)^{-\tau^*} G_{\tau^*} | \mathcal{F}_{k+1}\right].$$

Because $\tau^*$ is also in $T_k$, we have

$$\begin{aligned}
\widetilde{I\!E}[(1+r)^{-(k+1)} V_{k+1} | \mathcal{F}_k] &= \widetilde{I\!E}\left[\widetilde{I\!E}[(1+r)^{-\tau^*} G_{\tau^*} | \mathcal{F}_{k+1}] | \mathcal{F}_k\right] \\
&= \widetilde{I\!E}[(1+r)^{-\tau^*} G_{\tau^*} | \mathcal{F}_k] \\
&\le \max_{\tau \in T_k} \widetilde{I\!E}\left[(1+r)^{-\tau} G_\tau | \mathcal{F}_k\right] \\
&= (1+r)^{-k} V_k.
\end{aligned}$$

■

**Lemma 2.20** *If* $\{Y_k\}_{k=0}^n$ *is another process satisfying*

$$Y_k \ge G_k,\, k = 0, 1, \ldots, n, \ \ \textit{a.s.},$$

*and* $\{(1+r)^{-k} Y_k\}_{k=0}^n$ *is a supermartingale, then*

$$Y_k \ge V_k,\, k = 0, 1, \ldots, n, \ \ \textit{a.s.}$$

**Proof:** The optional sampling theorem for the supermartingale $\{(1+r)^{-k}Y_k\}_{k=0}^n$ implies

$$\widetilde{I\!E}[(1+r)^{-\tau}Y_\tau|\mathcal{F}_k] \le (1+r)^{-k}Y_k, \forall \tau \in T_k.$$

Therefore,

$$\begin{aligned}
V_k &= (1+r)^k \max_{\tau\in T_k} \widetilde{I\!E}[(1+r)^{-\tau}G_\tau|\mathcal{F}_k] \\
&\le (1+r)^k \max_{\tau\in T_k} \widetilde{I\!E}[(1+r)^{-\tau}Y_\tau|\mathcal{F}_k] \\
&\le (1+r)^{-k}(1+r)^k Y_k \\
&= Y_k.
\end{aligned}$$

■

**Lemma 2.21** *Define*

$$\begin{aligned}
C_k &= V_k - \frac{1}{1+r}\widetilde{I\!E}[V_{k+1}|\mathcal{F}_k] \\
&= (1+r)^k \left\{(1+r)^{-k}V_k - \widetilde{I\!E}[(1+r)^{-(k+1)}V_{k+1}|\mathcal{F}_k]\right\}.
\end{aligned}$$

*Since* $\{(1+r)^{-k}V_k\}_{k=0}^n$ *is a supermartingale,* $C_k$ *must be non-negative almost surely. Define*

$$\Delta_k(\omega_1,\dots,\omega_k) = \frac{V_{k+1}(\omega_1,\dots,\omega_k,H) - V_{k+1}(\omega_1,\dots,\omega_k,T)}{S_{k+1}(\omega_1,\dots,\omega_k,H) - S_{k+1}(\omega_1,\dots,\omega_k,T)}.$$

*Set* $X_0 = V_0$ *and define recursively*

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)(X_k - C_k - \Delta_k S_k).$$

*Then*

$$X_k = V_k \ \ \forall k.$$

**Proof:** We proceed by induction on $k$. The induction hypothesis is that $X_k = V_k$ for some $k \in \{0,1,\dots,n-1\}$, i.e., for each fixed $(\omega_1,\dots,\omega_k)$ we have

$$X_k(\omega_1,\dots,\omega_k) = V_k(\omega_1,\dots,\omega_k).$$

We need to show that

$$X_{k+1}(\omega_1,\dots,\omega_k,H) = V_{k+1}(\omega_1,\dots,\omega_k,H),$$

$$X_{k+1}(\omega_1,\dots,\omega_k,T) = V_{k+1}(\omega_1,\dots,\omega_k,T).$$

We prove the first equality; the proof of the second is similar. Note first that

$$\begin{aligned}
&V_k(\omega_1,\dots,\omega_k) - C_k(\omega_1,\dots,\omega_k) \\
&= \frac{1}{1+r}\widetilde{I\!E}[V_{k+1}|\mathcal{F}_k](\omega_1,\dots,\omega_k) \\
&= \frac{1}{1+r}\left(\tilde{p}V_{k+1}(\omega_1,\dots,\omega_k,H) + \tilde{q}V_{k+1}(\omega_1,\dots,\omega_k,T)\right).
\end{aligned}$$

Since $(\omega_1, \ldots, \omega_k)$ will be fixed for the rest of the proof, we will suppress these symbols. For example, the last equation can be written simply as

$$V_k - C_k = \frac{1}{1+r}\left(\tilde{p}V_{k+1}(H) + \tilde{q}V_{k+1}(T)\right).$$

We compute

$$\begin{aligned}
X_{k+1}(H) &= \Delta_k S_{k+1}(H) + (1+r)(X_k - C_k - \Delta_k S_k) \\
&= \frac{V_{k+1}(H) - V_{k+1}(T)}{S_{k+1}(H) - S_{k+1}(T)}\left(S_{k+1}(H) - (1+r)S_k\right) \\
&\quad +(1+r)(V_k - C_k) \\
&= \frac{V_{k+1}(H) - V_{k+1}(T)}{(u-d)S_k}\left(uS_k - (1+r)S_k\right) \\
&\quad +\tilde{p}V_{k+1}(H) + \tilde{q}V_{k+1}(T) \\
&= (V_{k+1}(H) - V_{k+1}(T))\tilde{q} + \tilde{p}V_{k+1}(H) + \tilde{q}V_{k+1}(T) \\
&= V_{k+1}(H).
\end{aligned}$$

## 6.3 Compound European Derivative Securities

In order to derive the optimal stopping time for an American derivative security, it will be useful to study compound European derivative securities, which are also interesting in their own right.

A compound European derivative security consists of $n+1$ different simple European derivative securities (with the same underlying stock) expiring at times $0, 1, \ldots, n$; the security that expires at time $j$ has payoff $C_j$. Thus a compound European derivative security is specified by the process $\{C_j\}_{j=0}^n$, where each $C_j$ is $\mathcal{F}_j$-measurable, i.e., the process $\{C_j\}_{j=0}^n$ is adapted to the filtration $\{\mathcal{F}_k\}_{k=0}^n$.

**Hedging a short position (one payment)**. Here is how we can hedge a short position in the $j$'th European derivative security. The value of European derivative security $j$ at time $k$ is given by

$$V_k^{(j)} = (1+r)^k \widetilde{I\!E}[(1+r)^{-j}C_j|\mathcal{F}_k], \; k = 0, \ldots, j,$$

and the hedging portfolio for that security is given by

$$\Delta_k^{(j)}(\omega_1, \ldots, \omega_k) = \frac{V_{k+1}^{(j)}(\omega_1, \ldots, \omega_k, H) - V_{k+1}^{(j)}(\omega_1, \ldots, \omega_k, T)}{S_{k+1}^{(j)}(\omega_1, \ldots, \omega_k, H) - S_{k+1}^{(j)}(\omega_1, \ldots, \omega_k, T)}, k = 0, \ldots, j-1.$$

Thus, starting with wealth $V_0^{(j)}$, and using the portfolio $(\Delta_0^{(j)}, \ldots, \Delta_{j-1}^{(j)})$, we can ensure that at time $j$ we have wealth $C_j$.

**Hedging a short position (all payments).** Superpose the hedges for the individual payments. In other words, start with wealth $V_0 = \sum_{j=0}^n V_0^{(j)}$. At each time $k \in \{0, 1, \ldots, n-1\}$, first make the payment $C_k$ and then use the portfolio

$$\Delta_k = \Delta_k{}^{(k+1)} + \Delta_k{}^{(k+2)} + \ldots + \Delta_k{}^{(n)}$$

corresponding to all future payments. At the final time $n$, after making the final payment $C_n$, we will have exactly zero wealth.

Suppose you own a compound European derivative security$\{C_j\}_{j=0}^n$. Compute

$$V_0 = \sum_{j=0}^n V_0^{(j)} = \widetilde{IE}\left[\sum_{j=0}^n (1+r)^{-j} C_j\right]$$

and the hedging portfolio is $\{\Delta_k\}_{k=0}^{n-1}$. You can borrow $V_0$ and consume it immediately. This leaves you with wealth $X_0 = -V_0$. In each period $k$, *receive* the payment $C_k$ and then use the portfolio $-\Delta_k$. At the final time $n$, after receiving the last payment $C_n$, your wealth will reach zero, i.e., you will no longer have a debt.

## 6.4 Optimal Exercise of American Derivative Security

In this section we derive the optimal exercise time for the owner of an American derivative security. Let $\{G_k\}_{k=0}^n$ be an American derivative security. Let $\tau$ be the stopping time the owner plans to use. (We assume that each $G_k$ is non-negative, so we may assume without loss of generality that the owner stops at expiration – time $n$– if not before). Using the stopping time $\tau$, in period $j$ the owner will receive the payment

$$C_j = I_{\{\tau=j\}} G_j.$$

In other words, once he chooses a stopping time, the owner has effectively converted the American derivative security into a compound European derivative security, whose value is

$$\begin{aligned} V_0^{(\tau)} &= \widetilde{IE}\left[\sum_{j=0}^n (1+r)^{-j} C_j\right] \\ &= \widetilde{IE}\left[\sum_{j=0}^n (1+r)^{-j} I_{\{\tau=j\}} G_j\right] \\ &= \widetilde{IE}[(1+r)^{-\tau} G_\tau]. \end{aligned}$$

The owner of the American derivative security can borrow this amount of money immediately, if he chooses, and invest in the market so as to exaclty pay off his debt as the payments $\{C_j\}_{j=0}^n$ are received. Thus, his optimal behavior is to use a stopping time $\tau$ which maximizes $V_0^{(\tau)}$.

**Lemma 4.22** $V_0^{(\tau)}$ *is maximized by the stopping time*

$$\tau^* = \min\{k; V_k = G_k\}.$$

**Proof:** Recall the definition

$$V_0 \stackrel{\triangle}{=} \max_{\tau \in T_0} \widetilde{IE}\left[(1+r)^{-\tau} G_\tau\right] = \max_{\tau \in T_0} V_0^{(\tau)}$$

Let $\tau'$ be a stopping time which maximizes $V_0^{(\tau)}$, i.e., $V_0 = \widetilde{I\!E}\left[(1+r)^{-\tau'}G_{\tau'}\right]$. Because $\{(1+r)^{-k}V_k\}_{k=0}^n$ is a supermartingale, we have from the optional sampling theorem and the inequality $V_k \geq G_k$, the following:

$$\begin{aligned} V_0 &\geq \widetilde{I\!E}\left[(1+r)^{-\tau'}V_{\tau'}|\mathcal{F}_0\right] \\ &= \widetilde{I\!E}\left[(1+r)^{-\tau'}V_{\tau'}\right] \\ &\geq \widetilde{I\!E}\left[(1+r)^{-\tau'}G_{\tau'}\right] \\ &= V_0. \end{aligned}$$

Therefore,

$$V_0 = \widetilde{I\!E}\left[(1+r)^{-\tau'}V_{\tau'}\right] = \widetilde{I\!E}\left[(1+r)^{-\tau'}G_{\tau'}\right],$$

and

$$V_{\tau'} = G_{\tau'}, \quad \text{a.s.}$$

We have just shown that if $\tau'$ attains the maximum in the formula

$$V_0 = \max_{\tau \in T_0} \widetilde{I\!E}\left[(1+r)^{-\tau}G_\tau\right], \tag{4.1}$$

then

$$V_{\tau'} = G_{\tau'}, \quad \text{a.s.}$$

But we have defined

$$\tau^* = \min\{k; V_k = G_k\},$$

and so we must have $\tau^* \leq \tau' \leq n$ almost surely. The optional sampling theorem implies

$$\begin{aligned} (1+r)^{-\tau^*}G_{\tau^*} &= (1+r)^{-\tau^*}V_{\tau^*} \\ &\geq \widetilde{I\!E}\left[(1+r)^{-\tau'}V_{\tau'}|\mathcal{F}_{\tau^*}\right] \\ &= \widetilde{I\!E}\left[(1+r)^{-\tau'}G_{\tau'}|\mathcal{F}_{\tau^*}\right]. \end{aligned}$$

Taking expectations on both sides, we obtain

$$\widetilde{I\!E}\left[(1+r)^{-\tau^*}G_{\tau^*}\right] \geq \widetilde{I\!E}\left[(1+r)^{-\tau'}G_{\tau'}\right] = V_0.$$

It follows that $\tau^*$ also attains the maximum in (4.1), and is therefore an optimal exercise time for the American derivative security. ∎

# Chapter 7

# Jensen's Inequality

## 7.1 Jensen's Inequality for Conditional Expectations

**Lemma 1.23** *If $\varphi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}|\varphi(X)| < \infty$, then*

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}]).$$

*For instance, if $\mathcal{G} = \{\phi, \Omega\}$, $\varphi(x) = x^2$:*

$$\mathbb{E}X^2 \geq (\mathbb{E}X)^2.$$

**Proof:** Since $\varphi$ is convex we can express it as follows (See Fig. 7.1):

$$\varphi(x) = \max_{\substack{h \leq \varphi \\ h \text{ is linear}}} h(x).$$

Now let $h(x) = ax + b$ lie below $\varphi$. Then,

$$\begin{aligned} \mathbb{E}[\varphi(X)|\mathcal{G}] &\geq \mathbb{E}[aX + b|\mathcal{G}] \\ &= a\mathbb{E}[X|\mathcal{G}] + b \\ &= h(\mathbb{E}[X|\mathcal{G}]) \end{aligned}$$

This implies

$$\begin{aligned} \mathbb{E}[\varphi(X)|\mathcal{G}] &\geq \max_{\substack{h \leq \varphi \\ h \text{ is linear}}} h(\mathbb{E}[X|\mathcal{G}]) \\ &= \varphi(\mathbb{E}[X|\mathcal{G}]). \end{aligned}$$

■

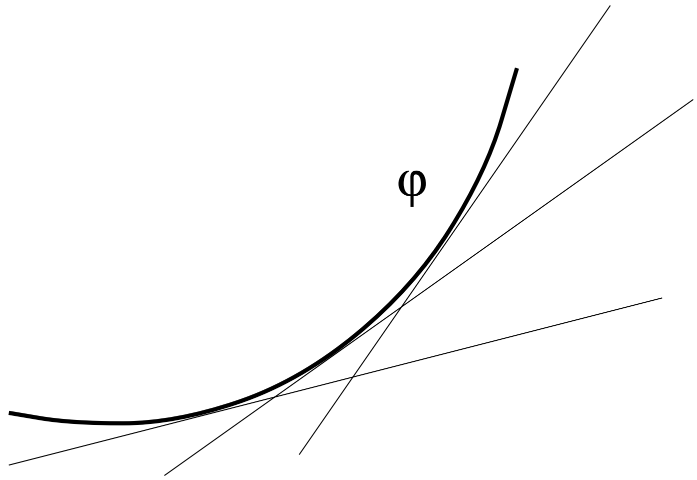



Figure 7.1: *Expressing a convex function as a max over linear functions.*

**Theorem 1.24** *If $\{Y_k\}_{k=0}^n$ is a martingale and $\phi$ is convex then $\{\varphi(Y_k)\}_{k=0}^n$ is a submartingale.*

**Proof:**

$$\begin{aligned} I\!E[\varphi(Y_{k+1})|\mathcal{F}_k] &\geq \varphi(I\!E[Y_{k+1}|\mathcal{F}_k]) \\ &= \varphi(Y_k). \end{aligned}$$

■

## 7.2 Optimal Exercise of an American Call

This follows from Jensen's inequality.

**Corollary 2.25** *Given a convex function $g : [0, \infty) \rightarrow I\!R$ where $g(0) = 0$. For instance, $g(x) = (x - K)^+$ is the payoff function for an American call. Assume that $r \geq 0$. Consider the American derivative security with payoff $g(S_k)$ in period $k$. The value of this security is the same as the value of the simple European derivative security with final payoff $g(S_n)$, i.e.,*

$$\widetilde{I\!E}\left[(1+r)^{-n} g(S_n)\right] = \max_{\tau} \widetilde{I\!E}\left[(1+r)^{-\tau} g(S_\tau)\right],$$

*where the LHS is the European value and the RHS is the American value. In particular $\tau = n$ is an optimal exercise time.*

**Proof:** Because $g$ is convex, for all $\lambda \in [0, 1]$ we have (see Fig. 7.2):

$$\begin{aligned} g(\lambda x) &= g(\lambda x + (1-\lambda).0) \\ &\leq \lambda g(x) + (1-\lambda).g(0) \\ &= \lambda g(x). \end{aligned}$$

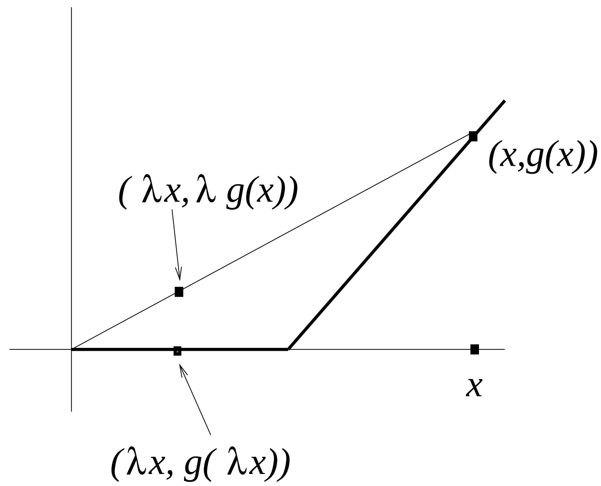



Figure 7.2: *Proof of Cor. 2.25*

Therefore,

$$g\left(\frac{1}{1+r}S_{k+1}\right) \le \frac{1}{1+r}g(S_{k+1})$$

and

$$\begin{aligned}\widetilde{I\!E}\left[(1+r)^{-(k+1)}g(S_{k+1})|\mathcal{F}_k\right] &= (1+r)^{-k}\widetilde{I\!E}\left[\frac{1}{1+r}g(S_{k+1})|\mathcal{F}_k\right]\\ &\ge (1+r)^{-k}\widetilde{I\!E}\left[g\left(\frac{1}{1+r}S_{k+1}\right)|\mathcal{F}_k\right]\\ &\ge (1+r)^{-k}g\left(\widetilde{I\!E}\left[\frac{1}{1+r}S_{k+1}|\mathcal{F}_k\right]\right)\\ &= (1+r)^{-k}g(S_k),\end{aligned}$$

So $\{(1+r)^{-k}g(S_k)\}_{k=0}^n$ is a submartingale. Let $\tau$ be a stopping time satisfying $0 \le \tau \le n$. The optional sampling theorem implies

$$(1+r)^{-\tau}g(S_\tau) \le \widetilde{I\!E}\left[(1+r)^{-n}g(S_n)|\mathcal{F}_\tau\right].$$

Taking expectations, we obtain

$$\begin{aligned}\widetilde{I\!E}\left[(1+r)^{-\tau}g(S_\tau)\right] &\le \widetilde{I\!E}\left(\widetilde{I\!E}\left[(1+r)^{-n}g(S_n)|\mathcal{F}_\tau\right]\right)\\ &= \widetilde{I\!E}\left[(1+r)^{-n}g(S_n)\right].\end{aligned}$$

Therefore, the value of the American derivative security is

$$\max_\tau \widetilde{I\!E}\left[(1+r)^{-\tau}g(S_\tau)\right] \le \widetilde{I\!E}\left[(1+r)^{-n}g(S_n)\right],$$

and this last expression is the value of the European derivative security. Of course, the LHS cannot be strictly less than the RHS above, since stopping at time $n$ is always allowed, and we conclude that

$$\max_\tau \widetilde{I\!E}\left[(1+r)^{-\tau}g(S_\tau)\right] = \widetilde{I\!E}\left[(1+r)^{-n}g(S_n)\right].$$

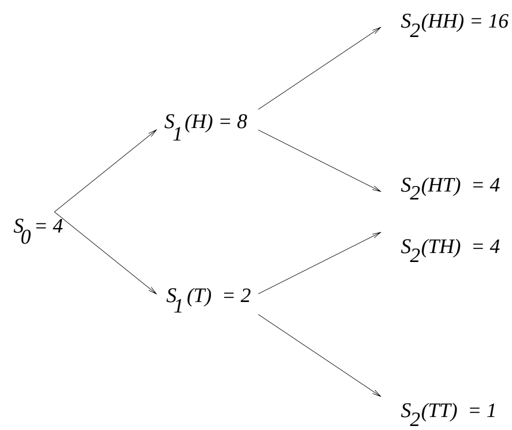



Figure 7.3: *A three period binomial model.*

■

## 7.3 Stopped Martingales

Let $\{Y_k\}_{k=0}^n$ be a stochastic process and let $\tau$ be a stopping time. We denote by $\{Y_{k\wedge\tau}\}_{k=0}^n$ the *stopped process*

$$Y_{k\wedge\tau(\omega)}(\omega),\ k=0,1,\dots,n.$$

**Example 7.1 (Stopped Process)** Figure 7.3 shows our familiar 3-period binomial example.

Define

$$\tau(\omega)=\begin{cases}1 & \text{if } \omega_1=T,\\ 2 & \text{if } \omega_1=H.\end{cases}$$

Then

$$S_{2\wedge\tau(\omega)}(\omega)=\begin{cases}S_2(HH)=16 & \text{if } \omega=HH,\\ S_2(HT)=4 & \text{if } \omega=HT,\\ S_1(T)=2 & \text{if } \omega=TH,\\ S_1(T)=2 & \text{if } \omega=TT.\end{cases}$$

■

**Theorem 3.26** *A stopped martingale (or submartingale, or supermartingale) is still a martingale (or submartingale, or supermartingale respectively).*

**Proof:** Let $\{Y_k\}_{k=0}^n$ be a martingale, and $\tau$ be a stopping time. Choose some $k\in\{0,1,\dots,n\}$. The set $\{\tau\le k\}$ is in $\mathcal{F}_k$, so the set $\{\tau\ge k+1\}=\{\tau\le k\}^c$ is also in $\mathcal{F}_k$. We compute

$$\begin{aligned}I\!\!E\left[Y_{(k+1)\wedge\tau}|\mathcal{F}_k\right] &= I\!\!E\left[I_{\{\tau\le k\}}Y_\tau+I_{\{\tau\ge k+1\}}Y_{k+1}|\mathcal{F}_k\right]\\ &= I_{\{\tau\le k\}}Y_\tau+I_{\{\tau\ge k+1\}}I\!\!E[Y_{k+1}|\mathcal{F}_k]\\ &= I_{\{\tau\le k\}}Y_\tau+I_{\{\tau\ge k+1\}}Y_k\\ &= Y_{k\wedge\tau}.\end{aligned}$$

■

# Chapter 8

# Random Walks

## 8.1 First Passage Time

Toss a coin infinitely many times. Then the sample space $\Omega$ is the set of all infinite sequences $\omega = (\omega_1, \omega_2, \ldots)$ of $H$ and $T$. Assume the tosses are independent, and on each toss, the probability of $H$ is $\frac{1}{2}$, as is the probability of $T$. Define

$$Y_j(\omega) = \begin{cases} 1 & \text{if } \omega_j = H, \\ -1 & \text{if } \omega_j = T, \end{cases}$$

$$\begin{aligned} M_0 &= 0, \\ M_k &= \sum_{j=1}^{k} Y_j, \; k = 1, 2, \ldots \end{aligned}$$

The process $\{M_k\}_{k=0}^{\infty}$ is a *symmetric random walk* (see Fig. 8.1) Its analogue in continuous time is *Brownian motion*.

Define

$$\tau = \min\{k \geq 0; M_k = 1\}.$$

If $M_k$ never gets to 1 (e.g., $\omega = (TTTT\ldots)$), then $\tau = \infty$. The random variable $\tau$ is called the *first passage time to 1*. It is the first time the number of heads exceeds by one the number of tails.

## 8.2 $\tau$ is almost surely finite

It is shown in a Homework Problem that $\{M_k\}_{k=0}^{\infty}$ and $\{N_k\}_{k=0}^{\infty}$ where

$$\begin{aligned} N_k &= \exp\left\{\theta M_k - k \log\left(\frac{e^{\theta} + e^{-\theta}}{2}\right)\right\} \\ &= e^{\theta M_k}\left(\frac{2}{e^{\theta} + e^{-\theta}}\right)^k \end{aligned}$$

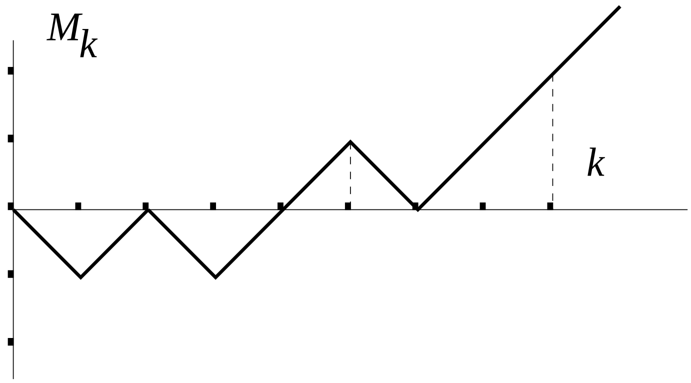



Figure 8.1: *The random walk process* $M_k$

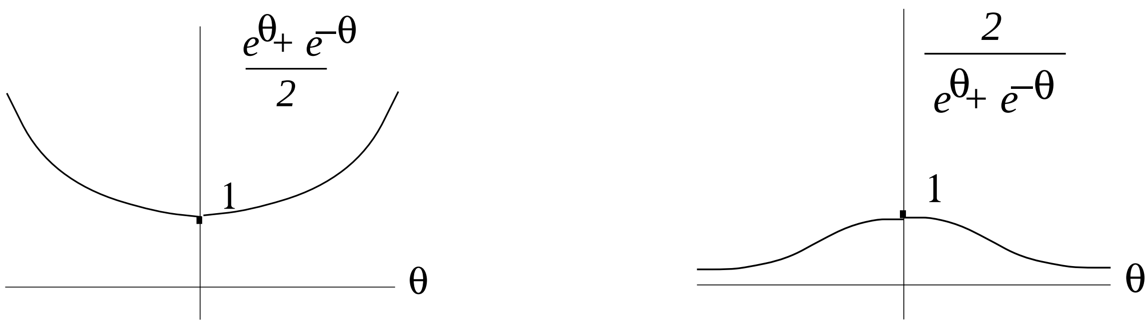



Figure 8.2: *Illustrating two functions of* $\theta$

are martingales. (Take $M_k = -S_k$ in part (i) of the Homework Problem and take $\theta = -\sigma$ in part (v).) Since $N_0 = 1$ and a stopped martingale is a martingale, we have

$$1 = I\!\!E N_{k\wedge\tau} = I\!\!E\left[e^{\theta M_{k\wedge\tau}}\left(\frac{2}{e^\theta + e^{-\theta}}\right)^{k\wedge\tau}\right] \tag{2.1}$$

for every fixed $\theta \in I\!\!R$ (See Fig. 8.2 for an illustration of the various functions involved). We want to let $k\to\infty$ in (2.1), but we have to worry a bit that for some sequences $\omega \in \Omega$, $\tau(\omega) = \infty$.

We consider fixed $\theta > 0$, so

$$\left(\frac{2}{e^\theta + e^{-\theta}}\right) < 1.$$

As $k\to\infty$,

$$\left(\frac{2}{e^\theta + e^{-\theta}}\right)^{k\wedge\tau} \to \begin{cases} \left(\frac{2}{e^\theta + e^{-\theta}}\right)^{\tau} & \text{if } \tau < \infty, \\ 0 & \text{if } \tau = \infty \end{cases}$$

Furthermore, $M_{k\wedge\tau} \le 1$, because we stop this martingale when it reaches 1, so

$$0 \le e^{\theta M_{k\wedge\tau}} \le e^\theta$$

and

$$0 \leq e^{\theta M_{k\wedge\tau}} \left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{k\wedge\tau} \leq e^{\theta}.$$

In addition,

$$\lim_{k\to\infty} e^{\theta M_{k\wedge\tau}} \left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{k\wedge\tau} = \begin{cases} e^{\theta}\left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{\tau} & \text{if } \tau<\infty, \\ 0 & \text{if } \tau=\infty. \end{cases}$$

Recall Equation (2.1):

$$I\!\!E\left[e^{\theta M_{k\wedge\tau}} \left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{k\wedge\tau}\right] = 1$$

Letting $k\to\infty$, and using the Bounded Convergence Theorem, we obtain

$$I\!\!E\left[e^{\theta} \left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{\tau} I_{\{\tau<\infty\}}\right] = 1. \tag{2.2}$$

For all $\theta \in (0,1]$, we have

$$0 \leq e^{\theta} \left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{\tau} I_{\{\tau<\infty\}} \leq e,$$

so we can let $\theta\downarrow 0$ in (2.2), using the Bounded Convergence Theorem again, to conclude

$$I\!\!E\left[I_{\{\tau<\infty\}}\right] = 1,$$

i.e.,

$$I\!\!P\{\tau<\infty\} = 1.$$

We know there are paths of the symmetric random walk $\{M_k\}_{k=0}^{\infty}$ which never reach level 1. We have just shown that these paths *collectively* have no probability. (In our infinite sample space $\Omega$, each path *individually* has zero probability). We therefore do not need the indicator $I_{\{\tau<\infty\}}$ in (2.2), and we rewrite that equation as

$$I\!\!E\left[\left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{\tau}\right] = e^{-\theta}. \tag{2.3}$$

## 8.3 The moment generating function for $\tau$

Let $\alpha \in (0,1)$ be given. We want to find $\theta > 0$ so that

$$\alpha = \left(\frac{2}{e^{\theta}+e^{-\theta}}\right).$$

Solution:

$$\alpha e^{\theta} + \alpha e^{-\theta} - 2 = 0$$

$$\alpha(e^{-\theta})^2 - 2e^{-\theta} + \alpha = 0$$

$$e^{-\theta} = \frac{1 \pm \sqrt{1-\alpha^2}}{\alpha}.$$

We want $\theta > 0$, so we must have $e^{-\theta} < 1$. Now $0 < \alpha < 1$, so

$$0 < (1-\alpha)^2 < (1-\alpha) < 1-\alpha^2,$$

$$1-\alpha < \sqrt{1-\alpha^2},$$

$$1-\sqrt{1-\alpha^2} < \alpha,$$

$$\frac{1-\sqrt{1-\alpha^2}}{\alpha} < 1$$

We take the negative square root:

$$e^{-\theta} = \frac{1-\sqrt{1-\alpha^2}}{\alpha}.$$

Recall Equation (2.3):

$$I\!\!E\left[\left(\frac{2}{e^{\theta}+e^{-\theta}}\right)^{\tau}\right] = e^{-\theta},\ \theta > 0.$$

With $\alpha \in (0,1)$ and $\theta > 0$ related by

$$\begin{aligned} e^{-\theta} &= \frac{1-\sqrt{1-\alpha^2}}{\alpha}, \\ \alpha &= \left(\frac{2}{e^{\theta}+e^{-\theta}}\right), \end{aligned}$$

this becomes

$$I\!\!E\alpha^{\tau} = \frac{1-\sqrt{1-\alpha^2}}{\alpha},\ 0 < \alpha < 1. \tag{3.1}$$

We have computed the *moment generating function* for the first passage time to 1.

## 8.4 Expectation of $\tau$

Recall that

$$I\!\!E\alpha^{\tau} = \frac{1-\sqrt{1-\alpha^2}}{\alpha}, \quad 0 < \alpha < 1,$$

so

$$\begin{aligned} \frac{d}{d\alpha} I\!\!E\alpha^{\tau} &= I\!\!E(\tau\alpha^{\tau-1}) \\ &= \frac{d}{d\alpha}\left(\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right) \\ &= \frac{1-\sqrt{1-\alpha^2}}{\alpha^2\sqrt{1-\alpha^2}}. \end{aligned}$$

Using the Monotone Convergence Theorem, we can let $\alpha \uparrow 1$ in the equation

$$\mathbb{E}(\tau\alpha^{\tau-1}) = \frac{1-\sqrt{1-\alpha^2}}{\alpha^2\sqrt{1-\alpha^2}},$$

to obtain

$$\mathbb{E}\tau = \infty.$$

Thus in summary:

$$\tau \stackrel{\triangle}{=} \min\{k; M_k = 1\},$$

$$\mathbb{P}\{\tau < \infty\} = 1,$$

$$\mathbb{E}\tau = \infty.$$

## 8.5 The Strong Markov Property

The random walk process $\{M_k\}_{k=0}^{\infty}$ is a Markov process, i.e.,

$$\begin{aligned} &\mathbb{E}\,[\text{ random variable depending only on } \quad M_{k+1}, M_{k+2}, \ldots \mid \ \mathcal{F}_k] \\ &\quad = \mathbb{E}\,[\text{ same random variable } \quad |M_k]\,. \end{aligned}$$

In discrete time, this Markov property implies the *Strong Markov property*:

$$\begin{aligned} &\mathbb{E}\,[\text{ random variable depending only on } \quad M_{\tau+1}, M_{\tau+2}, \ldots \mid \ \mathcal{F}_\tau] \\ &\quad = \mathbb{E}\,[\text{ same random variable } \quad | \ M_\tau]\,. \end{aligned}$$

for any almost surely finite stopping time $\tau$.

## 8.6 General First Passage Times

Define

$$\tau_m \stackrel{\triangle}{=} \min\{k \geq 0; M_k = m\}, \ m = 1, 2, \ldots$$

Then $\tau_2 - \tau_1$ is the number of periods between the first arrival at level 1 and the first arrival at level 2. The distribution of $\tau_2 - \tau_1$ is the same as the distribution of $\tau_1$ (see Fig. 8.3), i.e.,

$$\mathbb{E}\alpha^{\tau_2-\tau_1} = \frac{1-\sqrt{1-\alpha^2}}{\alpha}, \ \alpha \in (0,1).$$

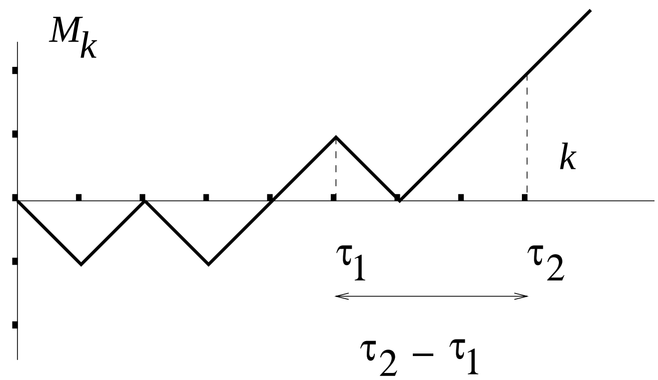



Figure 8.3: *General first passage times.*

For $\alpha \in (0,1)$,

$$\begin{aligned}
I\!\!E[\alpha^{\tau_2}|\mathcal{F}_{\tau_1}] &= I\!\!E[\alpha^{\tau_1}\alpha^{\tau_2-\tau_1}|\mathcal{F}_{\tau_1}] \\
&= \alpha^{\tau_1} I\!\!E[\alpha^{\tau_2-\tau_1}|\mathcal{F}_{\tau_1}] \\
&\qquad \text{(taking out what is known)} \\
&= \alpha^{\tau_1} I\!\!E[\alpha^{\tau_2-\tau_1}|M_{\tau_1}] \\
&\qquad \text{(strong Markov property)} \\
&= \alpha^{\tau_1} I\!\!E[\alpha^{\tau_2-\tau_1}] \\
&\qquad (M_{\tau_1} = 1, \ \text{not random}) \\
&= \alpha^{\tau_1}\left(\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right).
\end{aligned}$$

Take expectations of both sides to get

$$\begin{aligned}
I\!\!E\alpha^{\tau_2} &= I\!\!E\alpha^{\tau_1}.\left(\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right) \\
&= \left(\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right)^2
\end{aligned}$$

In general,

$$I\!\!E\alpha^{\tau_m} = \left(\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right)^m, \ \alpha \in (0,1).$$

## 8.7 Example: Perpetual American Put

Consider the binomial model, with $u = 2, d = \frac{1}{2}, r = \frac{1}{4}$, and payoff function $(5 - S_k)^+$. The risk neutral probabilities are $\tilde{p} = \frac{1}{2}$, $\tilde{q} = \frac{1}{2}$, and thus

$$S_k = S_0 u^{M_k},$$

where $M_k$ is a symmetric random walk under the risk-neutral measure, denoted by $\widetilde{I\!P}$. Suppose $S_0 = 4$. Here are some possible exercise rules:

**Rule 0:** Stop immediately. $\tau_0 = 0, V^{(\tau_0)} = 1$.

**Rule 1:** Stop as soon as stock price falls to 2, i.e., at time

$$\tau_{-1} \stackrel{\triangle}{=} \min\{k; M_k = -1\}.$$

**Rule 2:** Stop as soon as stock price falls to 1, i.e., at time

$$\tau_{-2} \stackrel{\triangle}{=} \min\{k; M_k = -2\}.$$

Because the random walk is symmetric under $\widetilde{I\!P}$, $\tau_{-m}$ has the same distribution under $\widetilde{I\!P}$ as the stopping time $\tau_m$ in the previous section. This observation leads to the following computations of value. **Value of Rule 1:**

$$\begin{aligned} V^{(\tau_{-1})} &= \widetilde{I\!E}\left[(1+r)^{-\tau_{-1}}(5 - S_{\tau_{-1}})^+\right] \\ &= (5-2)^+ I\!E\left[(\tfrac{4}{5})^{\tau_{-1}}\right] \\ &= 3.\frac{1 - \sqrt{1 - (\frac{4}{5})^2}}{\frac{4}{5}} \\ &= \frac{3}{2}. \end{aligned}$$

**Value of Rule 2:**

$$\begin{aligned} V^{(\tau_{-2})} &= (5-1)^+ \widetilde{I\!E}\left[(\tfrac{4}{5})^{\tau_{-2}}\right] \\ &= 4.(\tfrac{1}{2})^2 \\ &= 1. \end{aligned}$$

This suggests that the optimal rule is Rule 1, i.e., stop (exercise the put) as soon as the stock price falls to 2, and the value of the put is $\frac{3}{2}$ if $S_0 = 4$.

Suppose instead we start with $S_0 = 8$, and stop the first time the price falls to 2. This requires 2 down steps, so the value of this rule with this initial stock price is

$$(5-2)^+ \widetilde{I\!E}\left[(\tfrac{4}{5})^{\tau_{-2}}\right] = 3.(\tfrac{1}{2})^2 = \frac{3}{4}.$$

In general, if $S_0 = 2^j$ for some $j \geq 1$, and we stop when the stock price falls to 2, then $j-1$ down steps will be required and the value of the option is

$$(5-2)^+ \widetilde{I\!E}\left[(\tfrac{4}{5})^{\tau_{-(j-1)}}\right] = 3.(\tfrac{1}{2})^{j-1}.$$

We define

$$v(2^j) \stackrel{\triangle}{=} 3.(\tfrac{1}{2})^{j-1}, \; j = 1, 2, 3, \ldots$$

If $S_0 = 2^j$ for some $j \leq 1$, then the initial price is at or below 2. In this case, we exercise immediately, and the value of the put is

$$v(2^j) \stackrel{\triangle}{=} 5 - 2^j, \; j = 1, 0, -1, -2, \ldots$$

**Proposed exercise rule:** Exercise the put whenever the stock price is at or below 2. The value of this rule is given by $v(2^j)$ as we just defined it. Since the put is perpetual, the initial time is no different from any other time. This leads us to make the following:

**Conjecture 1** *The value of the perpetual put at time $k$ is $v(S_k)$.*

How do we recognize the value of an American derivative security when we see it?

There are three parts to the proof of the conjecture. We must show:

**(a)** $v(S_k) \geq (5 - S_k)^+ \; \forall k$,

**(b)** $\left\{(\frac{4}{5})^k v(S_k)\right\}_{k=0}^{\infty}$ is a supermartingale,

**(c)** $\{v(S_k)\}_{k=0}^{\infty}$ is the smallest process with properties (a) and (b).

**Note:** To simplify matters, we shall only consider initial stock prices of the form $S_0 = 2^j$, so $S_k$ is always of the form $2^j$, with a possibly different $j$.

**Proof: (a).** Just check that

$$v(2^j) \stackrel{\triangle}{=} 3.(\tfrac{1}{2})^{j-1} \geq (5 - 2^j)^+ \;\text{ for }\; j \geq 1,$$

$$v(2^j) \stackrel{\triangle}{=} 5 - 2^j \geq (5 - 2^j)^+ \;\text{ for }\; j \leq 1.$$

This is straightforward. ■

**Proof: (b).** We must show that

$$\begin{aligned} v(S_k) &\geq \widetilde{I\!E}\left[\tfrac{4}{5} v(S_{k+1}) | \mathcal{F}_k\right] \\ &= \tfrac{4}{5}.\tfrac{1}{2} v(2S_k) + \tfrac{4}{5}.\tfrac{1}{2} v(\tfrac{1}{2} S_k). \end{aligned}$$

By assumption, $S_k = 2^j$ for some $j$. We must show that

$$v(2^j) \geq \tfrac{2}{5} v(2^{j+1}) + \tfrac{2}{5} v(2^{j-1}).$$

If $j \geq 2$, then $v(2^j) = 3.(\frac{1}{2})^{j-1}$ and

$$\begin{aligned} &\tfrac{2}{5} v(2^{j+1}) + \tfrac{2}{5} v(2^{j-1}) \\ &\quad= \tfrac{2}{5}.3.(\tfrac{1}{2})^j + \tfrac{2}{5}.3.(\tfrac{1}{2})^{j-2} \\ &\quad= 3.\left[\tfrac{2}{5}.\frac{1}{4} + \tfrac{2}{5}\right](\tfrac{1}{2})^{j-2} \\ &\quad= 3.\tfrac{1}{2}.(\tfrac{1}{2})^{j-2} \\ &\quad= v(2^j). \end{aligned}$$

If $j = 1$, then $v(2^j) = v(2) = 3$ and

$$\begin{aligned}
&\tfrac{2}{5}v(2^{j+1}) + \tfrac{2}{5}v(2^{j-1}) \\
&\quad = \tfrac{2}{5}v(4) + \tfrac{2}{5}v(1) \\
&\quad = \tfrac{2}{5}.3.\tfrac{1}{2} + \tfrac{2}{5}.4 \\
&\quad = 3/5 + 8/5 \\
&\quad = 2\tfrac{1}{5} < v(2) = 3
\end{aligned}$$

There is a gap of size $\frac{4}{5}$.

If $j \le 0$, then $v(2^j) = 5 - 2^j$ and

$$\begin{aligned}
&\tfrac{2}{5}v(2^{j+1}) + \tfrac{2}{5}v(2^{j-1}) \\
&\quad = \tfrac{2}{5}(5 - 2^{j+1}) + \tfrac{2}{5}(5 - 2^{j-1}) \\
&\quad = 4 - \tfrac{2}{5}(4+1)2^{j-1} \\
&\quad = 4 - 2^j < v(2^j) = 5 - 2^j.
\end{aligned}$$

There is a gap of size 1. This concludes the proof of (b). ■

**Proof: (c).** Suppose $\{Y_k\}_{k=0}^n$ is some other process satisfying:

**(a')** $Y_k \ge (5 - S_k)^+ \ \forall k$,

**(b')** $\{(\frac{4}{5})^k Y_k\}_{k=0}^\infty$ is a supermartingale.

We must show that

$$Y_k \ge v(S_k) \ \forall k. \tag{7.1}$$

Actually, since the put is perpetual, every time $k$ is like every other time, so it will suffice to show

$$Y_0 \ge v(S_0), \tag{7.2}$$

provided we let $S_0$ in (7.2) be any number of the form $2^j$. With appropriate (but messy) conditioning on $\mathcal{F}_k$, the proof we give of (7.2) can be modified to prove (7.1).

For $j \le 1$,

$$v(2^j) = 5 - 2^j = (5 - 2^j)^+,$$

so if $S_0 = 2^j$ for some $j \le 1$, then (a') implies

$$Y_0 \ge (5 - 2^j)^+ = v(S_0).$$

Suppose now that $S_0 = 2^j$ for some $j \ge 2$, i.e., $S_0 \ge 4$. Let

$$\begin{aligned}
\tau &= \min\{k; S_k = 2\} \\
&= \min\{k; M_k = j - 1\}.
\end{aligned}$$

Then

$$\begin{aligned} v(S_0) &= v(2^j) = 3.(\tfrac{1}{2})^{j-1} \\ &= I\!E\left[(\tfrac{4}{5})^{\tau}(5 - S_\tau)^+\right]. \end{aligned}$$

Because $\{(\frac{4}{5})^k Y_k\}_{k=0}^{\infty}$ is a supermartingale

$$Y_0 \geq I\!E\left[(\tfrac{4}{5})^{\tau} Y_\tau\right] \geq I\!E\left[(\tfrac{4}{5})^{\tau}(5 - S_\tau)^+\right] = v(S_0).$$

■

**Comment on the proof of (c):** If the candidate value process is the actual value of a particular exercise rule, then (c) will be automatically satisfied. In this case, we constructed $v$ so that $v(S_k)$ is the value of the put at time $k$ if the stock price at time $k$ is $S_k$ and *if we exercise the put the first time ($k$, or later) that the stock price is 2 or less.* In such a situation, we need only verify properties (a) and (b).

## 8.8 Difference Equation

If we imagine stock prices which can fall at any point in $(0, \infty)$, not just at points of the form $2^j$ for integers $j$, then we can imagine the function $v(x)$, defined for all $x > 0$, which gives the value of the perpetual American put when the stock price is $x$. This function should satisfy the conditions:

**(a)** $v(x) \geq (K - x)^+, \ \forall x,$

**(b)** $v(x) \geq \frac{1}{1+r}\left[\tilde{p}v(ux) + \tilde{q}v(dx)\right], \ \forall x,$

**(c)** At each $x$, either (a) or (b) holds with equality.

In the example we worked out, we have

$$\text{For } j \geq 1 : v(2^j) = 3.(\tfrac{1}{2})^{j-1} = \frac{6}{2^j};$$

$$\text{For } j \leq 1 : v(2^j) = 5 - 2^j.$$

This suggests the formula

$$v(x) = \begin{cases} \frac{6}{x}, & x \geq 3, \\ 5 - x, & 0 < x \leq 3. \end{cases}$$

We then have (see Fig. 8.4):

**(a)** $v(x) \geq (5 - x)^+; \forall x,$

**(b)** $v(x) \geq \frac{4}{5}\left[\frac{1}{2}v(2x) + \frac{1}{2}v(\frac{x}{2})\right]$ for every $x$ except for $2 < x < 4$.

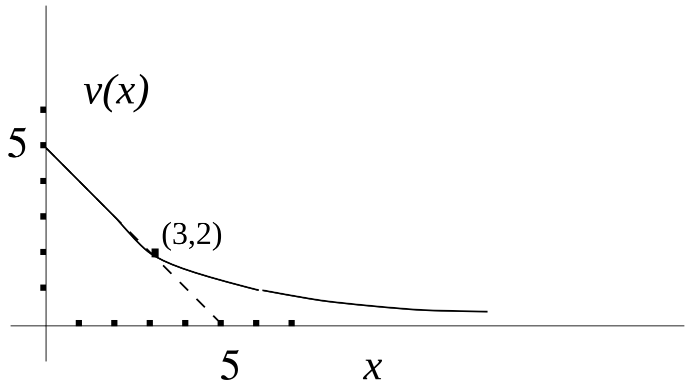



Figure 8.4: *Graph of* $v(x)$.

Check of condition (c):

- If $0 < x \leq 3$, then (a) holds with equality.
- If $x \geq 6$, then (b) holds with equality:

$$\tfrac{4}{5}\left[\tfrac{1}{2}v(2x) + \tfrac{1}{2}v(\frac{x}{2})\right] = \tfrac{4}{5}\left[\tfrac{1}{2}\frac{6}{2x} + \tfrac{1}{2}\frac{12}{x}\right] = \frac{6}{x}.$$

- If $3 < x < 4$ or $4 < x < 6$, then both (a) and (b) are strict. This is an artifact of the discreteness of the binomial model. This artifact will disappear in the continuous model, in which an analogue of (a) or (b) holds with equality at every point.

## 8.9 Distribution of First Passage Times

Let $\{M_k\}_{k=0}^{\infty}$ be a symetric random walk under a probability measure $I\!\!P$, with $M_0 = 0$. Defining

$$\tau = \min\{k \geq 0; M_k = 1\},$$

we recall that

$$I\!\!E\alpha^\tau = \frac{1 - \sqrt{1-\alpha^2}}{\alpha}, \quad 0 < \alpha < 1.$$

We will use this moment generating function to obtain the distribution of $\tau$. We first obtain the Taylor series expasion of $I\!\!E\alpha^\tau$ as follows:

$$
\begin{aligned}
f(x) &= 1-\sqrt{1-x}, \quad f(0)=0 \\
f'(x) &= \tfrac{1}{2}(1-x)^{-\frac{1}{2}}, \quad f'(0)=\tfrac{1}{2} \\
f''(x) &= \frac{1}{4}(1-x)^{-\frac{3}{2}}, \quad f''(0)=\frac{1}{4} \\
f'''(x) &= \frac{3}{8}(1-x)^{-\frac{5}{2}}, \quad f'''(0)=\frac{3}{8} \\
& \ldots \\
f^{(j)}(x) &= \frac{1\times 3\times\ldots\times(2j-3)}{2^j}(1-x)^{-\frac{(2j-1)}{2}}, \\
f^{(j)}(0) &= \frac{1\times 3\times\ldots\times(2j-3)}{2^j} \\
&= \frac{1\times 3\times\ldots\times(2j-3)}{2^j}.\frac{2\times 4\times\ldots\times(2j-2)}{2^{j-1}(j-1)!} \\
&= \left(\tfrac{1}{2}\right)^{2j-1}\frac{(2j-2)!}{(j-1)!}
\end{aligned}
$$

The Taylor series expansion of $f(x)$ is given by

$$
\begin{aligned}
f(x) &= 1-\sqrt{1-x} \\
&= \sum_{j=0}^{\infty}\frac{1}{j!}f^{(j)}(0)x^j \\
&= \sum_{j=1}^{\infty}\left(\tfrac{1}{2}\right)^{2j-1}\frac{(2j-2)!}{j!(j-1)!}x^j \\
&= \frac{x}{2}+\sum_{j=2}^{\infty}\left(\tfrac{1}{2}\right)^{2j-1}\frac{1}{(j-1)}\binom{2j-2}{j}x^j.
\end{aligned}
$$

So we have

$$
\begin{aligned}
\mathbb{E}\alpha^{\tau} &= \frac{1-\sqrt{1-\alpha^2}}{\alpha} \\
&= \frac{1}{\alpha}f(\alpha^2) \\
&= \frac{\alpha}{2}+\sum_{j=2}^{\infty}\left(\frac{\alpha}{2}\right)^{2j-1}\frac{1}{(j-1)}\binom{2j-2}{j}.
\end{aligned}
$$

But also,

$$
\mathbb{E}\alpha^{\tau}=\sum_{j=1}^{\infty}\alpha^{2j-1}\mathbb{P}\{\tau=2j-1\}.
$$

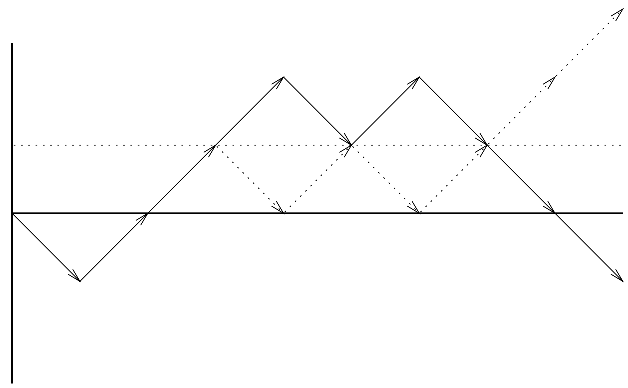


Figure 8.5: *Reflection principle.*

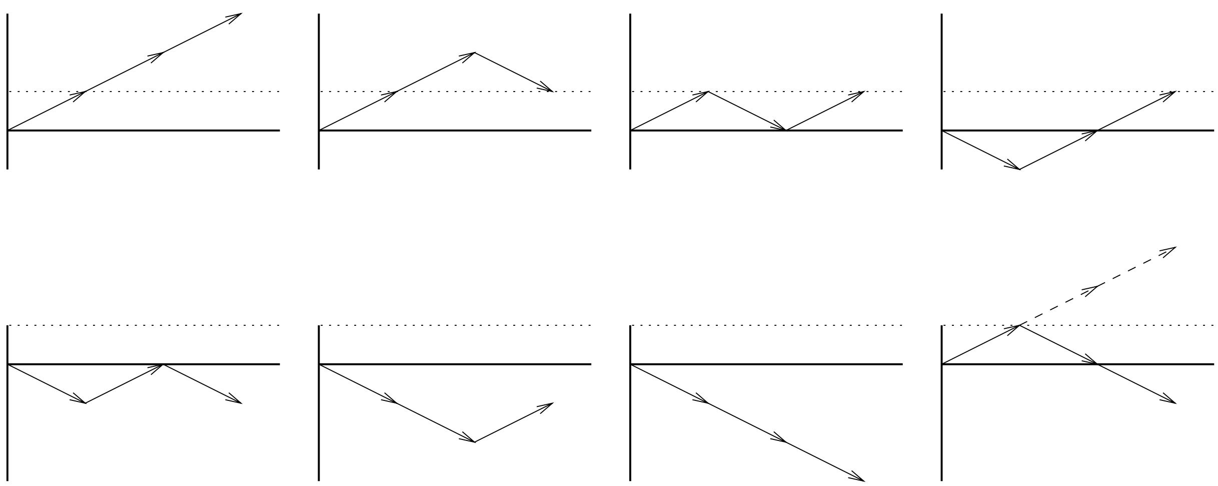


Figure 8.6: *Example with $j = 2$.*

Therefore,

$$
\begin{aligned}
I\!\!P\{\tau = 1\} &= \tfrac{1}{2}, \\
I\!\!P\{\tau = 2j-1\} &= \left(\frac{1}{2}\right)^{2j-1} \frac{1}{(j-1)} \binom{2j-2}{j}, \quad j = 2, 3, \ldots
\end{aligned}
$$

## 8.10 The Reflection Principle

To count how many paths reach level 1 by time $2j-1$, count all those for which $M_{2j-1} = 1$ and double count all those for which $M_{2j-1} \geq 3$. (See Figures 8.5, 8.6.)

In other words,

$$\begin{aligned}
I\!P\{\tau \le 2j-1\} &= I\!P\{M_{2j-1} = 1\} + 2I\!P\{M_{2j-1} \ge 3\} \\
&= I\!P\{M_{2j-1} = 1\} + I\!P\{M_{2j-1} \ge 3\} + I\!P\{M_{2j-1} \le -3\} \\
&= 1 - I\!P\{M_{2j-1} = -1\}.
\end{aligned}$$

For $j \ge 2$,

$$\begin{aligned}
I\!P\{\tau = 2j-1\} &= I\!P\{\tau \le 2j-1\} - I\!P\{\tau \le 2j-3\} \\
&= [1 - I\!P\{M_{2j-1} = -1\}] - [1 - I\!P\{M_{2j-3} = -1\}] \\
&= I\!P\{M_{2j-3} = -1\} - I\!P\{M_{2j-1} = -1\} \\
&= \left(\tfrac{1}{2}\right)^{2j-3} \frac{(2j-3)!}{(j-1)!(j-2)!} - \left(\tfrac{1}{2}\right)^{2j-1} \frac{(2j-1)!}{j!(j-1)!} \\
&= \left(\tfrac{1}{2}\right)^{2j-1} \frac{(2j-3)!}{j!(j-1)!} [4j(j-1) - (2j-1)(2j-2)] \\
&= \left(\tfrac{1}{2}\right)^{2j-1} \frac{(2j-3)!}{j!(j-1)!} [2j(2j-2) - (2j-1)(2j-2)] \\
&= \left(\tfrac{1}{2}\right)^{2j-1} \frac{(2j-2)!}{j!(j-1)!} \\
&= \left(\tfrac{1}{2}\right)^{2j-1} \frac{1}{(j-1)} \binom{2j-2}{j}.
\end{aligned}$$

# Chapter 9

# Pricing in terms of Market Probabilities: The Radon-Nikodym Theorem.

## 9.1 Radon-Nikodym Theorem

**Theorem 1.27 (Radon-Nikodym)** *Let P and $\widetilde{I\!\!P}$ be two probability measures on a space $(\Omega, \mathcal{F})$. Assume that for every $A \in \mathcal{F}$ satisfying $I\!\!P(A) = 0$, we also have $\widetilde{I\!\!P}(A) = 0$. Then we say that $\widetilde{I\!\!P}$ is* absolutely continuous *with respect to P. Under this assumption, there is a nonegative random variable $Z$ such that*

$$\widetilde{I\!\!P}(A) = \int_A Z \, dI\!\!P, \quad \forall A \in \mathcal{F}, \tag{1.1}$$

*and $Z$ is called the* Radon-Nikodym derivative *of $\widetilde{I\!\!P}$ with respect to P.*

**Remark 9.1** Equation (1.1) implies the apparently stronger condition

$$\widetilde{I\!\!E} X = I\!\!E[XZ]$$

for every random variable $X$ for which $I\!\!E|XZ| < \infty$.

**Remark 9.2** If $\widetilde{I\!\!P}$ is absolutely continuous with respect to P, and P is absolutely continuous with respect to $\widetilde{I\!\!P}$, we say that P and $\widetilde{I\!\!P}$ are *equivalent*. P and $\widetilde{I\!\!P}$ are equivalent if and only if

$$I\!\!P(A) = 0 \text{ exactly when } \widetilde{I\!\!P}(A) = 0, \quad \forall A \in \mathcal{F}.$$

If P and $\widetilde{I\!\!P}$ are equivalent and $Z$ is the Radon-Nikodym derivative of $\widetilde{I\!\!P}$ w.r.t. P, then $\frac{1}{Z}$ is the Radon-Nikodym derivative of P w.r.t. $\widetilde{I\!\!P}$, i.e.,

$$\widetilde{I\!\!E} X = I\!\!E[XZ] \;\; \forall X, \tag{1.2}$$

$$I\!\!E Y = \widetilde{I\!\!E}[Y.\frac{1}{Z}] \;\; \forall Y. \tag{1.3}$$

(Let $X$ and $Y$ be related by the equation $Y = XZ$ to see that (1.2) and (1.3) are the same.)

**Example 9.1 (Radon-Nikodym Theorem)** Let $\Omega = \{HH, HT, TH, TT\}$, the set of coin toss sequences of length 2. Let $P$ correspond to probability $\frac{1}{3}$ for $H$ and $\frac{2}{3}$ for $T$, and let $\widetilde{I\!P}$ correspond to probability $\frac{1}{2}$ for $H$ and $\frac{1}{2}$ for $T$. Then $Z(\omega) = \frac{\widetilde{I\!P}(\omega)}{I\!P(\omega)}$, so

$$Z(HH) = \frac{9}{4}, \;\; Z(HT) = \frac{9}{8}, \;\; Z(TH) = \frac{9}{8}, \;\; Z(TT) = \frac{9}{16}.$$

■

## 9.2 Radon-Nikodym Martingales

Let $\Omega$ be the set of all sequences of $n$ coin tosses. Let P be the market probability measure and let $\widetilde{I\!P}$ be the risk-neutral probability measure. Assume

$$I\!P(\omega) > 0, \; \widetilde{I\!P}(\omega) > 0, \; \forall \omega \in \Omega,$$

so that P and $\widetilde{I\!P}$ are equivalent. The Radon-Nikodym derivative of $\widetilde{I\!P}$ with respect to P is

$$Z(\omega) = \frac{\widetilde{I\!P}(\omega)}{I\!P(\omega)}.$$

Define the P-martingale

$$Z_k \stackrel{\triangle}{=} I\!E[Z|\mathcal{F}_k], \quad k = 0, 1, \ldots, n.$$

We can check that $Z_k$ is indeed a martingale:

$$\begin{aligned} I\!E[Z_{k+1}|\mathcal{F}_k] &= I\!E\left[I\!E[Z|\mathcal{F}_{k+1}]|\mathcal{F}_k\right] \\ &= I\!E[Z|\mathcal{F}_k] \\ &= Z_k. \end{aligned}$$

**Lemma 2.28** *If $X$ is $\mathcal{F}_k$-measurable, then* $\widetilde{I\!E}X = I\!E[XZ_k]$.

**Proof:**

$$\begin{aligned} \widetilde{I\!E}X &= I\!E[XZ] \\ &= I\!E\left[I\!E[XZ|\mathcal{F}_k]\right] \\ &= I\!E\left[X.I\!E[Z|\mathcal{F}_k]\right] \\ &= I\!E[XZ_k]. \end{aligned}$$

■

Note that Lemma 2.28 implies that if $X$ is $\mathcal{F}_k$-measurable, then for any $A \in \mathcal{F}_k$,

$$\widetilde{I\!E}[I_A X] = I\!E[Z_k I_A X],$$

or equivalently,

$$\int_A X \, d\widetilde{I\!P} = \int_A X Z_k \, dI\!P.$$

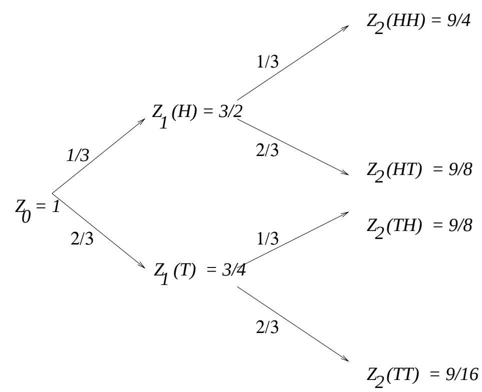



Figure 9.1: *Showing the $Z_k$ values in the 2-period binomial model example. The probabilities shown are for P, not $\widetilde{I\!P}$.*

**Lemma 2.29** *If $X$ is $\mathcal{F}_k$-measurable and $0 \le j \le k$, then*

$$\widetilde{I\!E}[X|\mathcal{F}_j] = \frac{1}{Z_j} I\!E[XZ_k|\mathcal{F}_j].$$

**Proof:** Note first that $\frac{1}{Z_j} I\!E[XZ_k|\mathcal{F}_j]$ is $\mathcal{F}_j$-measurable. So for any $A \in \mathcal{F}_j$, we have

$$\begin{aligned}
\int_A \frac{1}{Z_j} I\!E[XZ_k|\mathcal{F}_j]\, d\widetilde{I\!P} &= \int_A I\!E[XZ_k|\mathcal{F}_j]\, dI\!P \quad \text{(Lemma 2.28)} \\
&= \int_A XZ_k\, dI\!P \quad \text{(Partial averaging)} \\
&= \int_A X\, d\widetilde{I\!P} \quad \text{(Lemma 2.28)}
\end{aligned}$$

■

**Example 9.2 (Radon-Nikodym Theorem, continued)** We show in Fig. 9.1 the values of the martingale $Z_k$. We always have $Z_0 = 1$, since

$$Z_0 = I\!EZ = \int_\Omega Z\, dI\!P = \widetilde{I\!P}(\Omega) = 1.$$

■

## 9.3 The State Price Density Process

In order to express the value of a derivative security in terms of the market probabilities, it will be useful to introduce the following *state price density process*:

$$\zeta_k = (1+r)^{-k} Z_k, \; k = 0, \ldots, n.$$

We then have the following pricing formulas: For a **Simple European derivative security** with payoff $C_k$ at time $k$,

$$\begin{aligned} V_0 &= \widetilde{I\!E}\left[(1+r)^{-k}C_k\right] \\ &= I\!E\left[(1+r)^{-k}Z_kC_k\right] \quad \text{(Lemma 2.28)} \\ &= I\!E[\zeta_k C_k]. \end{aligned}$$

More generally for $0 \le j \le k$,

$$\begin{aligned} V_j &= (1+r)^j\widetilde{I\!E}\left[(1+r)^{-k}C_k|\mathcal{F}_j\right] \\ &= \frac{(1+r)^j}{Z_j} I\!E\left[(1+r)^{-k}Z_kC_k|\mathcal{F}_j\right] \quad \text{(Lemma 2.29)} \\ &= \frac{1}{\zeta_j} I\!E[\zeta_k C_k|\mathcal{F}_j] \end{aligned}$$

**Remark 9.3** $\{\zeta_j V_j\}_{j=0}^k$ is a martingale under P, as we can check below:

$$\begin{aligned} I\!E[\zeta_{j+1}V_{j+1}|\mathcal{F}_j] &= I\!E\left[I\!E[\zeta_k C_k|\mathcal{F}_{j+1}]|\mathcal{F}_j\right] \\ &= I\!E[\zeta_k C_k|\mathcal{F}_j] \\ &= \zeta_j V_j. \end{aligned}$$

Now for an **American derivative security** $\{G_k\}_{k=0}^n$:

$$\begin{aligned} V_0 &= \sup_{\tau\in T_0} \widetilde{I\!E}\left[(1+r)^{-\tau}G_\tau\right] \\ &= \sup_{\tau\in T_0} I\!E\left[(1+r)^{-\tau}Z_\tau G_\tau\right] \\ &= \sup_{\tau\in T_0} I\!E[\zeta_\tau G_\tau]. \end{aligned}$$

More generally for $0 \le j \le n$,

$$\begin{aligned} V_j &= (1+r)^j \sup_{\tau\in T_j} \widetilde{I\!E}\left[(1+r)^{-\tau}G_\tau|\mathcal{F}_j\right] \\ &= (1+r)^j \sup_{\tau\in T_j} \frac{1}{Z_j} I\!E\left[(1+r)^{-\tau}Z_\tau G_\tau|\mathcal{F}_j\right] \\ &= \frac{1}{\zeta_j} \sup_{\tau\in T_j} I\!E[\zeta_\tau G_\tau|\mathcal{F}_j]. \end{aligned}$$

**Remark 9.4** Note that

**(a)** $\{\zeta_j V_j\}_{j=0}^n$ is a supermartingale under P,

**(b)** $\zeta_j V_j \ge \zeta_j G_j \ \forall j,$

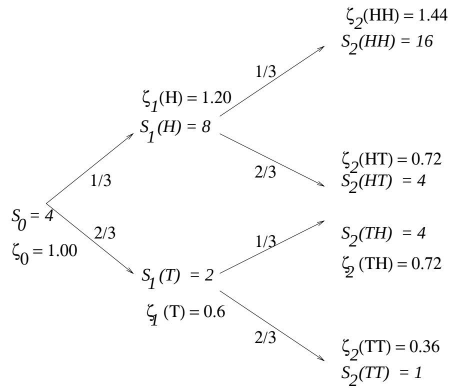



Figure 9.2: *Showing the state price values $\zeta_k$. The probabilities shown are for **P**, not $\widetilde{I\!P}$.*

**(c)** $\{\zeta_j V_j\}_{j=0}^n$ is the smallest process having properties (a) and (b).

We interpret $\zeta_k$ by observing that $\zeta_k(\omega) I\!P(\omega)$ is the value at time zero of a contract which pays \$1 at time $k$ if $\omega$ occurs.

**Example 9.3 (Radon-NikodymTheorem, continued)** We illustrate the use of the valuation formulas for European and American derivative securities in terms of market probabilities. Recall that $p = \frac{1}{3}$, $q = \frac{2}{3}$. The state price values $\zeta_k$ are shown in Fig. 9.2.

For a **European Call** with strike price 5, expiration time 2, we have

$$V_2(HH) = 11,\ \zeta_2(HH)V_2(HH) = 1.44 \times 11 = 15.84.$$

$$V_2(HT) = V_2(TH) = V_2(TT) = 0.$$

$$V_0 = \frac{1}{3} \times \frac{1}{3} \times 15.84 = 1.76.$$

$$\frac{\zeta_2(HH)}{\zeta_1(HH)} V_2(HH) = \frac{1.44}{1.20} \times 11 = 1.20 \times 11 = 13.20$$

$$V_1(H) = \frac{1}{3} \times 13.20 = 4.40$$

Compare with the risk-neutral pricing formulas:

$$V_1(H) = \tfrac{2}{5} V_1(HH) + \tfrac{2}{5} V_1(HT) = \tfrac{2}{5} \times 11 = 4.40,$$

$$V_1(T) = \tfrac{2}{5} V_1(TH) + \tfrac{2}{5} V_1(TT) = 0,$$

$$V_0 = \tfrac{2}{5} V_1(H) + \tfrac{2}{5} V_1(T) = \tfrac{2}{5} \times 4.40 = 1.76.$$

Now consider an **American put** with strike price 5 and expiration time 2. Fig. 9.3 shows the values of $\zeta_k(5 - S_k)^+$. We compute the value of the put under various stopping times $\tau$:

**(0)** Stop immediately: value is 1.

**(1)** If $\tau(HH) = \tau(HT) = 2,\ \tau(TH) = \tau(TT) = 1$, the value is

$$\frac{1}{3} \times \tfrac{2}{3} \times 0.72 + \tfrac{2}{3} \times 1.80 = 1.36.$$

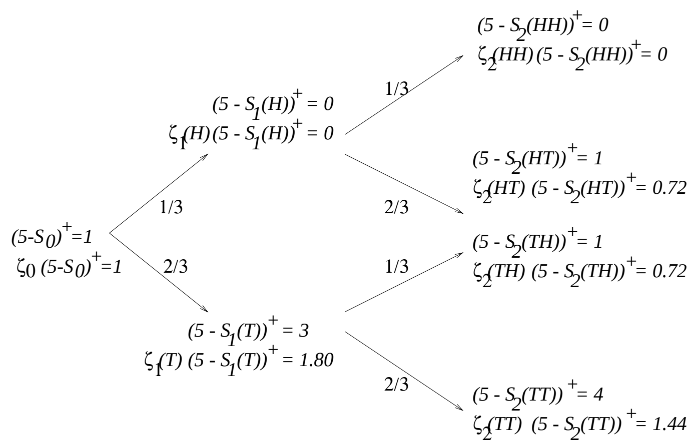



Figure 9.3: *Showing the values* $\zeta_k(5-S_k)^+$ *for an American put. The probabilities shown are for P, not* $\widetilde{I\!P}$.

**(2)** If we stop at time 2, the value is

$$\frac{1}{3} \times \tfrac{2}{3} \times 0.72 + \tfrac{2}{3} \times \frac{1}{3} \times 0.72 + \tfrac{2}{3} \times \tfrac{2}{3} \times 1.44 = 0.96$$

We see that (1) is optimal stopping rule. ■

## 9.4 Stochastic Volatility Binomial Model

Let $\Omega$ be the set of sequences of $n$ tosses, and let $0 < d_k < 1+r_k < u_k$, where for each $k$, $d_k, u_k, r_k$ are $\mathcal{F}_k$-measurable. Also let

$$\tilde{p}_k = \frac{1+r_k-d_k}{u_k-d_k}, \quad \tilde{q}_k = \frac{u_k-(1+r_k)}{u_k-d_k}.$$

Let $\widetilde{I\!P}$ be the risk-neutral probability measure:

$$\widetilde{I\!P}\{\omega_1 = H\} = \tilde{p}_0,$$

$$\widetilde{I\!P}\{\omega_1 = T\} = \tilde{q}_0,$$

and for $2 \le k \le n$,

$$\widetilde{I\!P}[\omega_{k+1} = H|\mathcal{F}_k] = \tilde{p}_k,$$

$$\widetilde{I\!P}[\omega_{k+1} = T|\mathcal{F}_k] = \tilde{q}_k.$$

Let P be the market probability measure, and assume $I\!P\{\omega\} > 0 \ \forall \omega \in \Omega$. Then P and $\widetilde{I\!P}$ are equivalent. Define

$$Z(\omega) = \frac{\widetilde{I\!P}(\omega)}{I\!P(\omega)} \quad \forall \omega \in \Omega,$$

$$Z_k = \mathbb{E}[Z|\mathcal{F}_k], \quad k = 0, 1, \ldots, n.$$

We define the *money market price process* as follows:

$$M_0 = 1,$$

$$M_k = (1 + r_{k-1})M_{k-1}, \; k = 1, \ldots, n.$$

Note that $M_k$ is $F_{k-1}$-measurable.

We then define the *state price process* to be

$$\zeta_k = \frac{1}{M_k} Z_k, \; k = 0, \ldots, n.$$

As before the portfolio process is $\{\Delta_k\}_{k=0}^{n-1}$. The self-financing value process (wealth process) consists of $X_0$, the non-random initial wealth, and

$$X_{k+1} = \Delta_k S_{k+1} + (1 + r_k)(X_k - \Delta_k S_k), \; k = 0, \ldots, n-1.$$

Then the following processes are martingales under $\widetilde{\mathbb{P}}$:

$$\left\{ \frac{1}{M_k} S_k \right\}_{k=0}^{n} \quad \text{and} \quad \left\{ \frac{1}{M_k} X_k \right\}_{k=0}^{n},$$

and the following processes are martingales under P:

$$\{\zeta_k S_k\}_{k=0}^{n} \quad \text{and} \quad \{\zeta_k X_k\}_{k=0}^{n}.$$

We thus have the following pricing formulas:

**Simple European derivative security** with payoff $C_k$ at time $k$:

$$\begin{aligned} V_j &= M_j \widetilde{\mathbb{E}} \left[ \frac{C_k}{M_k} \middle| \mathcal{F}_j \right] \\ &= \frac{1}{\zeta_j} \mathbb{E} \left[ \zeta_k C_k | \mathcal{F}_j \right] \end{aligned}$$

**American derivative security** $\{G_k\}_{k=0}^{n}$:

$$\begin{aligned} V_j &= M_j \sup_{\tau \in T_j} \widetilde{\mathbb{E}} \left[ \frac{G_\tau}{M_\tau} \middle| \mathcal{F}_j \right] \\ &= \frac{1}{\zeta_j} \sup_{\tau \in T_j} \mathbb{E} \left[ \zeta_\tau G_\tau | \mathcal{F}_j \right]. \end{aligned}$$

The usual hedging portfolio formulas still work.

## 9.5 Another Applicaton of the Radon-Nikodym Theorem

Let $(\Omega, \mathcal{F}, Q)$ be a probability space. Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$, and let $X$ be a non-negative random variable with $\int_\Omega X \; dQ = 1$. We construct the conditional expectation (under $Q$) of $X$ given $\mathcal{G}$. On $\mathcal{G}$, define two probability measures

$$I\!\!P(A) = Q(A) \quad \forall A \in \mathcal{G};$$

$$\widetilde{I\!\!P}(A) = \int_A X \, dQ \quad \forall A \in \mathcal{G}.$$

Whenever $Y$ is a $\mathcal{G}$-measurable random variable, we have

$$\int_\Omega Y \; dI\!\!P = \int_\Omega Y \; dQ;$$

if $Y = \mathbf{1}_A$ for some $A \in \mathcal{G}$, this is just the definition of $I\!\!P$, and the rest follows from the "standard machine". If $A \in \mathcal{G}$ and $I\!\!P(A) = 0$, then $Q(A) = 0$, so $\widetilde{I\!\!P}(A) = 0$. In other words, the measure $\widetilde{I\!\!P}$ is absolutely continuous with respect to the measure $\widetilde{I\!\!P}$. The Radon-Nikodym theorem implies that there exists a $\mathcal{G}$-measurable random variable $Z$ such that

$$\widetilde{I\!\!P}(A) \triangleq \int_A Z \; dI\!\!P \;\; \forall A \in \mathcal{G},$$

i.e.,

$$\int_A X \; dQ = \int_A Z \; dI\!\!P \;\; \forall A \in \mathcal{G}.$$

This shows that $Z$ has the "partial averaging" property, and since $Z$ is $\mathcal{G}$-measurable, it is the conditional expectation (under the probability measure $Q$) of $X$ given $\mathcal{G}$. The existence of conditional expectations is a consequence of the Radon-Nikodym theorem.

# Chapter 10

# Capital Asset Pricing

## 10.1 An Optimization Problem

Consider an agent who has initial wealth $X_0$ and wants to invest in the stock and money markets so as to maximize

$$\mathbb{E} \log X_n.$$

**Remark 10.1** Regardless of the portfolio used by the agent, $\{\zeta_k X_k\}_{k=0}^{\infty}$ is a martingale under P, so

$$\mathbb{E}\zeta_n X_n = X_0 \qquad (BC)$$

Here, (BC) stands for "Budget Constraint".

**Remark 10.2** If $\xi$ is any random variable satisfying (BC), i.e.,

$$\mathbb{E}\zeta_n \xi = X_0,$$

then there is a portfolio which starts with initial wealth $X_0$ and produces $X_n = \xi$ at time $n$. To see this, just regard $\xi$ as a simple European derivative security paying off at time $n$. Then $X_0$ is its value at time 0, and starting from this value, there is a hedging portfolio which produces $X_n = \xi$.

Remarks 10.1 and 10.2 show that the optimal $X_n$ for the capital asset pricing problem can be obtained by solving the following

**Constrained Optimization Problem:**
Find a random variable $\xi$ which solves:

$$\text{Maximize} \quad \mathbb{E} \log \xi$$

$$\text{Subject to} \quad \mathbb{E}\zeta_n \xi = X_0.$$

Equivalently, we wish to

$$\text{Maximize} \quad \sum_{\omega \in \Omega} (\log \xi(\omega))\, \mathbb{P}(\omega)$$

$$\text{Subject to} \quad \sum_{\omega\in\Omega} \zeta_n(\omega)\xi(\omega)\mathbb{P}(\omega) \ - \ X_0 = 0.$$

There are $2^n$ sequences $\omega$ in $\Omega$. Call them $\omega_1, \omega_2, \ldots, \omega_{2^n}$. Adopt the notation

$$x_1 = \xi(\omega_1), \ x_2 = \xi(\omega_2), \ \ldots, \ x_{2^n} = \xi(\omega_{2^n}).$$

We can thus restate the problem as:

$$\text{Maximize} \quad \sum_{k=1}^{2^n} (\log x_k)\mathbb{P}(\omega_k)$$

$$\text{Subject to} \quad \sum_{k=1}^{2^n} \zeta_n(\omega_k)x_k\mathbb{P}(\omega_k) \ - \ X_o = 0.$$

In order to solve this problem we use:

**Theorem 1.30 (Lagrange Multiplier)** *If $(x_1^*, \ldots, x_m^*)$ solve the problem*

$$\textit{Maxmize} \quad f(x_1, \ldots, x_m)$$

$$\textit{Subject to} \quad g(x_1, \ldots, x_m) = 0,$$

*then there is a number $\lambda$ such that*

$$\frac{\partial}{\partial x_k} f(x_1^*, \ldots, x_m^*) = \lambda \frac{\partial}{\partial x_k} g(x_1^*, \ldots, x_m^*), \quad k = 1, \ldots, m, \tag{1.1}$$

*and*

$$g(x_1^*, \ldots, x_m^*) = 0. \tag{1.2}$$

For our problem, (1.1) and (1.2) become

$$\frac{1}{x_k^*}\mathbb{P}(\omega_k) = \lambda\zeta_n(\omega_k)\mathbb{P}(\omega_k), \ k = 1, \ldots, 2^n, \tag{1.1'}$$

$$\sum_{k=1}^{2^n} \zeta_n(\omega_k)x_k^*\mathbb{P}(\omega_k) \ = \ X_0. \tag{1.2'}$$

Equation (1.1’) implies

$$x_k^* = \frac{1}{\lambda\zeta_n(\omega_k)}.$$

Plugging this into (1.2’) we get

$$\frac{1}{\lambda}\sum_{k=1}^{2^n} \mathbb{P}(\omega_k) = X_0 \implies \frac{1}{\lambda} = X_0.$$

Therefore,

$$x_k^* = \frac{X_0}{\zeta_n(\omega_k)},\ k = 1, \ldots, 2^n.$$

Thus we have shown that if $\xi^*$ solves the problem

$$\begin{array}{ll} \text{Maximize} & I\!\!E \log \xi \\ \text{Subject to} & I\!\!E(\zeta_n \xi) = X_0, \end{array} \tag{1.3}$$

then

$$\xi^* = \frac{X_0}{\zeta_n}. \tag{1.4}$$

**Theorem 1.31** *If $\xi^*$ is given by (1.4), then $\xi^*$ solves the problem (1.3).*

**Proof:** Fix $Z > 0$ and define

$$f(x) = \log x - xZ.$$

We maximize $f$ over $x > 0$:

$$f'(x) = \frac{1}{x} - Z = 0 \iff x = \frac{1}{Z},$$

$$f''(x) = -\frac{1}{x^2} < 0,\ \forall x \in I\!\!R.$$

The function $f$ is maximized at $x^* = \frac{1}{Z}$, i.e.,

$$\log x - xZ \le f(x^*) = \log \frac{1}{Z} - 1,\ \forall x > 0,\ \forall Z > 0. \tag{1.5}$$

Let $\xi$ be any random variable satisfying

$$I\!\!E(\zeta_n \xi) = X_0$$

and let

$$\xi^* = \frac{X_0}{\zeta_n}.$$

From (1.5) we have

$$\log \xi - \xi\left(\frac{\zeta_n}{X_0}\right) \le \log\left(\frac{X_0}{\zeta_n}\right) - 1.$$

Taking expectations, we have

$$I\!\!E \log \xi - \frac{1}{X_0} I\!\!E(\zeta_n \xi) \le I\!\!E \log \xi^* - 1,$$

and so

$$I\!\!E \log \xi \le I\!\!E \log \xi^*.$$

■

In summary, capital asset pricing works as follows: Consider an agent who has initial wealth $X_0$ and wants to invest in the stock and money market so as to maximize

$$\mathbb{E} \log X_n.$$

The optimal $X_n$ is $X_n = \frac{X_0}{\zeta_n}$, i.e.,

$$\zeta_n X_n = X_0.$$

Since $\{\zeta_k X_k\}_{k=0}^n$ is a martingale under P, we have

$$\zeta_k X_k = \mathbb{E}[\zeta_n X_n | \mathcal{F}_k] = X_0, \ k = 0, \ldots, n,$$

so

$$X_k = \frac{X_0}{\zeta_k},$$

and the optimal portfolio is given by

$$\Delta_k(\omega_1, \ldots, \omega_k) = \frac{\frac{X_0}{\zeta_{k+1}(\omega_1, \ldots, \omega_k, H)} - \frac{X_0}{\zeta_{k+1}(\omega_1, \ldots, \omega_k, T)}}{S_{k+1}(\omega_1, \ldots, \omega_k, H) - S_{k+1}(\omega_1, \ldots, \omega_k, T)}.$$

# Chapter 11

# General Random Variables

## 11.1 Law of a Random Variable

Thus far we have considered only random variables whose domain and range are discrete. We now consider a general random variable $X : \Omega \rightarrow I\!\!R$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that:

- $\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$.
- P is a probability measure on $\mathcal{F}$, i.e., $I\!\!P(A)$ is defined for every $A \in \mathcal{F}$.

A function $X : \Omega \rightarrow I\!\!R$ is a random variable if and only if for every $B \in \mathcal{B}(I\!\!R)$ (the $\sigma$-algebra of Borel subsets of R), the set

$$\{X \in B\} \stackrel{\triangle}{=} X^{-1}(B) \stackrel{\triangle}{=} \{\omega; X(\omega) \in B\} \in \mathcal{F},$$

i.e., $X : \Omega \rightarrow I\!\!R$ is a random variable if and only if $X^{-1}$ is a function from $\mathcal{B}(I\!\!R)$ to $\mathcal{F}$(See Fig. 11.1)

Thus any random variable $X$ induces a measure $\mu_X$ on the measurable space $(I\!\!R, \mathcal{B}(I\!\!R))$ defined by

$$\mu_X(B) = I\!\!P\left(X^{-1}(B)\right) \quad \forall B \in \mathcal{B}(I\!\!R),$$

where the probabiliy on the right is defined since $X^{-1}(B) \in \mathcal{F}$. $\mu_X$ is often called the *Law of* $X$ – in Williams' book this is denoted by $\mathcal{L}_X$.

## 11.2 Density of a Random Variable

The *density of* $X$ (if it exists) is a function $f_X : I\!\!R \rightarrow [0, \infty)$ such that

$$\mu_X(B) = \int_B f_X(x)\, dx \quad \forall B \in \mathcal{B}(I\!\!R).$$

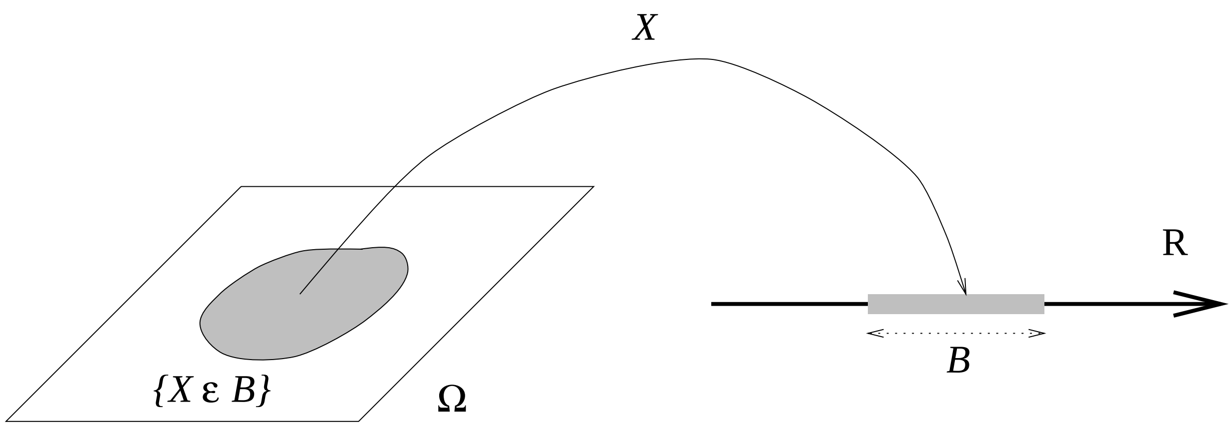



Figure 11.1: *Illustrating a real-valued random variable $X$.*

We then write

$$d\mu_X(x) = f_X(x)dx,$$

where the integral is with respect to the Lebesgue measure on R. $f_X$ is the Radon-Nikodym derivative of $\mu_X$ with respect to the Lebesgue measure. Thus $X$ has a density if and only if $\mu_X$ is absolutely continuous with respect to Lebesgue measure, which means that whenever $B \in \mathcal{B}(\mathbb{R})$ has Lebesgue measure zero, then

$$\mathbb{P}\{X \in B\} = 0.$$

## 11.3 Expectation

**Theorem 3.32 (Expectation of a function of $X$)** *Let $h : \mathbb{R} \to \mathbb{R}$ be given. Then*

$$\begin{aligned} \mathbb{E}h(X) &\stackrel{\triangle}{=} \int_\Omega h(X(\omega))\, d\mathbb{P}(\omega) \\ &= \int_{\mathbb{R}} h(x)\, d\mu_X(x) \\ &= \int_{\mathbb{R}} h(x) f_X(x)\, dx. \end{aligned}$$

**Proof:** (Sketch). If $h(x) = \mathbf{1}_B(x)$ for some $B \subset \mathbb{R}$, then these equations are

$$\begin{aligned} \mathbb{E}\mathbf{1}_B(X) &\stackrel{\triangle}{=} P\{X \in B\} \\ &= \mu_X(B) \\ &= \int_B f_X(x)\, dx, \end{aligned}$$

which are true by definition. Now use the “standard machine” to get the equations for general $h$.
■

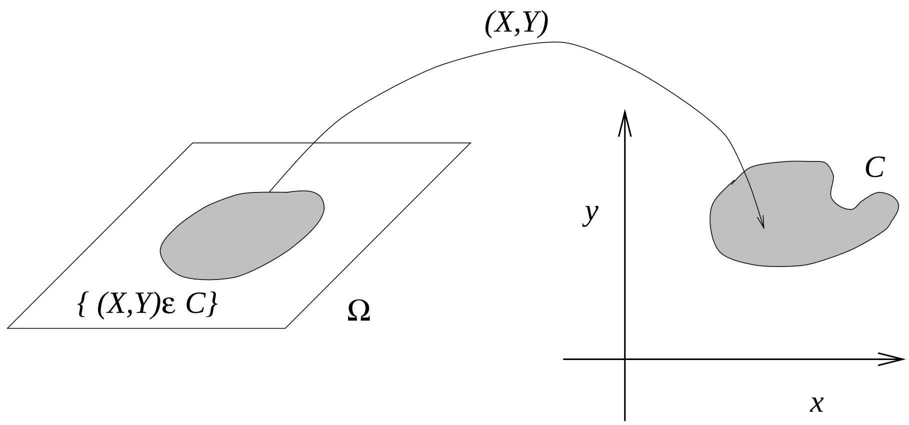



Figure 11.2: *Two real-valued random variables* $X, Y$.

## 11.4 Two random variables

Let $X, Y$ be two random variables $\Omega \rightarrow \mathbb{R}$ defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $X, Y$ induce a measure on $\mathcal{B}(\mathbb{R}^2)$ (see Fig. 11.2) called the *joint law of* $(X, Y)$, defined by

$$\mu_{X,Y}(C) \stackrel{\triangle}{=} \mathbb{P}\{(X,Y) \in C\} \quad \forall C \in \mathcal{B}(\mathbb{R}^2).$$

The *joint density of* $(X, Y)$ is a function

$$f_{X,Y} : \mathbb{R}^2 \rightarrow [0, \infty)$$

that satisfies

$$\mu_{X,Y}(C) = \iint_C f_{X,Y}(x,y)\, dxdy \quad \forall C \in \mathcal{B}(\mathbb{R}^2).$$

$f_{X,Y}$ is the Radon-Nikodym derivative of $\mu_{X,Y}$ with respect to the Lebesgue measure (area) on $\mathbb{R}^2$.

We compute the expectation of a function of $X, Y$ in a manner analogous to the univariate case:

$$\begin{aligned} \mathbb{E}k(X,Y) &\stackrel{\triangle}{=} \int_\Omega k(X(\omega), Y(\omega))\, d\mathbb{P}(\omega) \\ &= \iint_{\mathbb{R}^2} k(x,y)\, d\mu_{X,Y}(x,y) \\ &= \iint_{\mathbb{R}^2} k(x,y) f_{X,Y}(x,y)\, dxdy \end{aligned}$$

## 11.5 Marginal Density

Suppose $(X, Y)$ has joint density $f_{X,Y}$. Let $B \subset I\!\!R$ be given. Then

$$\begin{aligned}
\mu_Y(B) &= I\!\!P\{Y \in B\} \\
&= I\!\!P\{(X,Y) \in I\!\!R \times B\} \\
&= \mu_{X,Y}(I\!\!R \times B) \\
&= \int_B \int_{I\!\!R} f_{X,Y}(x,y)\, dxdy \\
&= \int_B f_Y(y)\, dy,
\end{aligned}$$

where

$$f_Y(y) \stackrel{\triangle}{=} \int_{I\!\!R} f_{X,Y}(x,y)\, dx.$$

Therefore, $f_Y(y)$ is the (marginal) density for $Y$.

## 11.6 Conditional Expectation

Suppose $(X, Y)$ has joint density $f_{X,Y}$. Let $h : I\!\!R \to I\!\!R$ be given. Recall that $I\!\!E[h(X)|Y] \stackrel{\triangle}{=} I\!\!E[h(X)|\sigma(Y)]$ depends on $\omega$ through $Y$, i.e., there is a function $g(y)$ ($g$ depending on $h$) such that

$$I\!\!E[h(X)|Y](\omega) = g(Y(\omega)).$$

How do we determine $g$?

We can characterize $g$ using *partial averaging:* Recall that $A \in \sigma(Y) \Longleftrightarrow A = \{Y \in B\}$ for some $B \in \mathcal{B}(I\!\!R)$. Then the following are equivalent characterizations of $g$:

$$\int_A g(Y)\, dI\!\!P = \int_A h(X)\, dI\!\!P \quad \forall A \in \sigma(Y), \tag{6.1}$$

$$\int_\Omega \mathbf{1}_B(Y) g(Y)\, dI\!\!P = \int_\Omega \mathbf{1}_B(Y) h(X)\, dI\!\!P \quad \forall B \in \mathcal{B}(I\!\!R), \tag{6.2}$$

$$\int_{I\!\!R} \mathbf{1}_B(y) g(y) \mu_Y(dy) = \iint\limits_{I\!\!R^2} \mathbf{1}_B(y) h(x)\, d\mu_{X,Y}(x,y) \quad \forall B \in \mathcal{B}(I\!\!R), \tag{6.3}$$

$$\int_B g(y) f_Y(y)\, dy = \int_B \int_{I\!\!R} h(x) f_{X,Y}(x,y)\, dxdy \quad \forall B \in \mathcal{B}(I\!\!R). \tag{6.4}$$

## 11.7 Conditional Density

A function $f_{X|Y}(x|y) : I\!R^2 \to [0, \infty)$ is called a *conditional density* for $X$ given $Y$ provided that for any function $h : I\!R \to I\!R$:

$$g(y) = \int_{I\!R} h(x) f_{X|Y}(x|y)\, dx. \tag{7.1}$$

(Here $g$ is the function satisfying

$$I\!E\,[h(X)|Y] = g(Y),$$

and $g$ depends on $h$, but $f_{X|Y}$ does not.)

**Theorem 7.33** *If $(X, Y)$ has a joint density $f_{X,Y}$, then*

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}. \tag{7.2}$$

**Proof:** Just verify that $g$ defined by (7.1) satisfies (6.4): For $B \in \mathcal{B}(I\!R)$,

$$\int_B \underbrace{\int_{I\!R} h(x) f_{X|Y}(x|y)\, dx}_{g(y)}\, f_Y(y)\, dy \;=\; \int_B \int_{I\!R} h(x) f_{X,Y}(x, y)\, dx dy.$$

**Notation 11.1** Let $g$ be the function satisfying

$$I\!E[h(X)|Y] = g(Y).$$

The function $g$ is often written as

$$g(y) = I\!E[h(X)|Y = y],$$

and (7.1) becomes

$$I\!E[h(X)|Y = y] = \int_{I\!R} h(x) f_{X|Y}(x|y)\, dx.$$

In conclusion, to determine $I\!E[h(X)|Y]$ (a function of $\omega$), first compute

$$g(y) = \int_{I\!R} h(x) f_{X|Y}(x|y)\, dx,$$

and then replace the dummy variable $y$ by the random variable $Y$:

$$I\!E[h(X)|Y](\omega) = g(Y(\omega)).$$

**Example 11.1 (Jointly normal random variables)** Given parameters: $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$. Let $(X, Y)$ have the joint density

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{x^2}{\sigma_1^2} - 2\rho\frac{x}{\sigma_1}\frac{y}{\sigma_2} + \frac{y^2}{\sigma_2^2} \right] \right\}.$$

The exponent is

$$\begin{aligned}
& -\frac{1}{2(1-\rho^2)}\left[\frac{x^2}{\sigma_1^2}-2\rho\frac{x}{\sigma_1}\frac{y}{\sigma_2}+\frac{y^2}{\sigma_2^2}\right] \\
&= -\frac{1}{2(1-\rho^2)}\left[\left(\frac{x}{\sigma_1}-\rho\frac{y}{\sigma_2}\right)^2+\frac{y^2}{\sigma_2^2}(1-\rho^2)\right] \\
&= -\frac{1}{2(1-\rho^2)}\frac{1}{\sigma_1^2}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right)^2-\frac{1}{2}\frac{y^2}{\sigma_2^2}.
\end{aligned}$$

We can compute the *Marginal density of* $Y$ as follows

$$\begin{aligned}
f_Y(y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right)^2}\,dx.e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}} \\
&= \frac{1}{2\pi\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}\,du.e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}} \\
&\qquad \text{using the substitution } u=\frac{1}{\sqrt{1-\rho^2}\sigma_1}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right),\ du=\frac{dx}{\sqrt{1-\rho^2}\sigma_1} \\
&= \frac{1}{\sqrt{2\pi}\,\sigma_2}e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}}.
\end{aligned}$$

Thus $Y$ is normal with mean 0 and variance $\sigma_2^2$.

**Conditional density.** From the expressions

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\frac{1}{\sigma_1^2}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right)^2}e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}},$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\,\sigma_2}e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}},$$

we have

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \frac{1}{\sqrt{2\pi}\,\sigma_1}\frac{1}{\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\frac{1}{\sigma_1^2}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right)^2}.
\end{aligned}$$

In the $x$-variable, $f_{X|Y}(x|y)$ is a normal density with mean $\frac{\rho\sigma_1}{\sigma_2}y$ and variance $(1-\rho^2)\sigma_1^2$. Therefore,

$$\mathbb{E}[X|Y=y]=\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx = \frac{\rho\sigma_1}{\sigma_2}y;$$

$$\begin{aligned}
&\mathbb{E}\left[\left(X-\frac{\rho\sigma_1}{\sigma_2}y\right)^2\middle|Y=y\right] \\
&= \int_{-\infty}^{\infty}\left(x-\frac{\rho\sigma_1}{\sigma_2}y\right)^2 f_{X|Y}(x|y)\,dx \\
&= (1-\rho^2)\sigma_1^2.
\end{aligned}$$

From the above two formulas we have the formulas

$$\mathbb{E}[X|Y] = \frac{\rho\sigma_1}{\sigma_2}Y, \tag{7.3}$$

$$\mathbb{E}\left[\left(X - \frac{\rho\sigma_1}{\sigma_2}Y\right)^2 \bigg| Y\right] = (1-\rho^2)\sigma_1^2. \tag{7.4}$$

Taking expectations in (7.3) and (7.4) yields

$$\mathbb{E}X = \frac{\rho\sigma_1}{\sigma_2}\mathbb{E}Y = 0, \tag{7.5}$$

$$\mathbb{E}\left[\left(X - \frac{\rho\sigma_1}{\sigma_2}Y\right)^2\right] = (1-\rho^2)\sigma_1^2. \tag{7.6}$$

Based on $Y$, the best estimator of $X$ is $\frac{\rho\sigma_1}{\sigma_2}Y$. This estimator is unbiased (has expected error zero) and the expected square error is $(1-\rho^2)\sigma_1^2$. No other estimator based on $Y$ can have a smaller expected square error (Homework problem 2.1). ■

## 11.8 Multivariate Normal Distribution

Please see Oksendal Appendix A.

Let $\mathbf{X}$ denote the column vector of random variables $(X_1, X_2, \ldots, X_n)^T$, and $\mathbf{x}$ the corresponding column vector of values $(x_1, x_2, \ldots, x_n)^T$. $\mathbf{X}$ has a multivariate normal distribution if and only if the random variables have the joint density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\sqrt{\det \mathbf{A}}}{(\mathbf{2}\pi)^{\mathbf{n/2}}} \exp\left\{-\tfrac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^{\mathbf{T}}.\mathbf{A}.(\mathbf{X}-\boldsymbol{\mu})\right\}.$$

Here,

$$\boldsymbol{\mu} \triangleq (\mu_1, \ldots, \mu_n)^T = \mathbb{E}\mathbf{X} \triangleq (\mathbb{E}X_1, \ldots, \mathbb{E}X_n)^T,$$

and $A$ is an $n \times n$ nonsingular matrix. $A^{-1}$ is the covariance matrix

$$A^{-1} = \mathbb{E}\left[(\mathbf{X}-\boldsymbol{\mu}).(\mathbf{X}-\boldsymbol{\mu})^T\right],$$

i.e. the $(i,j)$th element of $A^{-1}$ is $\mathbb{E}(X_i-\mu_i)(X_j-\mu_j)$. The random variables in $\mathbf{X}$ are independent if and only if $A^{-1}$ is diagonal, i.e.,

$$A^{-1} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2),$$

where $\sigma_j^2 = \mathbb{E}(X_j - \mu_j)^2$ is the variance of $X_j$.

## 11.9 Bivariate normal distribution

Take $n = 2$ in the above definitions, and let

$$\rho \triangleq \frac{\mathbb{E}(X_1 - \mu_1)(X_2 - \mu_2)}{\sigma_1 \sigma_2}.$$

Thus,

$$A^{-1} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

$$A = \begin{bmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & -\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ -\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{bmatrix},$$

$$\sqrt{\det A} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

and we have the formula from Example 11.1, adjusted to account for the possibly non-zero expectations:

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

## 11.10 MGF of jointly normal random variables

Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ denote a column vector with components in $\mathbb{R}$, and let $\mathbf{X}$ have a multivariate normal distribution with covariance matrix $A^{-1}$ and mean vector $\boldsymbol{\mu}$. Then the moment generating function is given by

$$\begin{aligned} \mathbb{E}e^{\mathbf{u}^T.\mathbf{X}} &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{u}^T.\mathbf{X}} f_{X_1,\,X_2,\,\ldots,\,X_n}(x_1, x_2, \ldots, x_n)\, dx_1 \ldots dx_n \\ &= \exp\left\{\tfrac{1}{2}\mathbf{u}^T A^{-1}\mathbf{u} + \mathbf{u}^T\boldsymbol{\mu}\right\}. \end{aligned}$$

If any $n$ random variables $X_1, X_2, \ldots, X_n$ have this moment generating function, then they are jointly normal, and we can read out the means and covariances. The random variables are jointly normal *and independent* if and only if for any real column vector $\mathbf{u} = (u_1, \ldots, u_n)^T$

$$\mathbb{E}e^{\mathbf{u}^T.\mathbf{X}} \triangleq \mathbb{E}\exp\left\{\sum_{j=1}^n u_j X_j\right\} = \exp\left\{\sum_{j=1}^n [\tfrac{1}{2}\sigma_j^2 u_j^2 + u_j\mu_j]\right\}.$$

# Chapter 12

# Semi-Continuous Models

## 12.1 Discrete-time Brownian Motion

Let $\{Y_j\}_{j=1}^n$ be a collection of independent, standard normal random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$, where P is the *market measure*. As before we denote the column vector $(Y_1, \ldots, Y_n)^T$ by $\mathbf{Y}$. We therefore have for any real colum vector $\mathbf{u} = (u_1, \ldots, u_n)^T$,

$$\mathbb{E} e^{\mathbf{u}^T \mathbf{Y}} = \mathbb{E} \exp\left\{ \sum_{j=1}^n u_j Y_j \right\} = \exp\left\{ \sum_{j=1}^n \tfrac{1}{2} u_j^2 \right\}.$$

Define the *discrete-time Brownian motion* (See Fig. 12.1):

$$\begin{aligned} B_0 &= 0, \\ B_k &= \sum_{j=1}^k Y_j, \; k = 1, \ldots, n. \end{aligned}$$

If we know $Y_1, Y_2, \ldots, Y_k$, then we know $B_1, B_2, \ldots, B_k$. Conversely, if we know $B_1, B_2, \ldots, B_k$, then we know $Y_1 = B_1, Y_2 = B_2 - B_1, \ldots, Y_k = B_k - B_{k-1}$. Define the filtration

$$\begin{aligned} \mathcal{F}_0 &= \{\phi, \Omega\}, \\ \mathcal{F}_k &= \sigma(Y_1, Y_2, \ldots, Y_k) = \sigma(B_1, B_2, \ldots, B_k), \; k = 1, \ldots, n. \end{aligned}$$

**Theorem 1.34** *$\{B_k\}_{k=0}^n$ is a martingale (under P).*

**Proof:**

$$\begin{aligned} \mathbb{E}\left[B_{k+1} | \mathcal{F}_k\right] &= \mathbb{E}\left[Y_{k+1} + B_k | \mathcal{F}_k\right] \\ &= \mathbb{E} Y_{k+1} + B_k \\ &= B_k. \end{aligned}$$

■

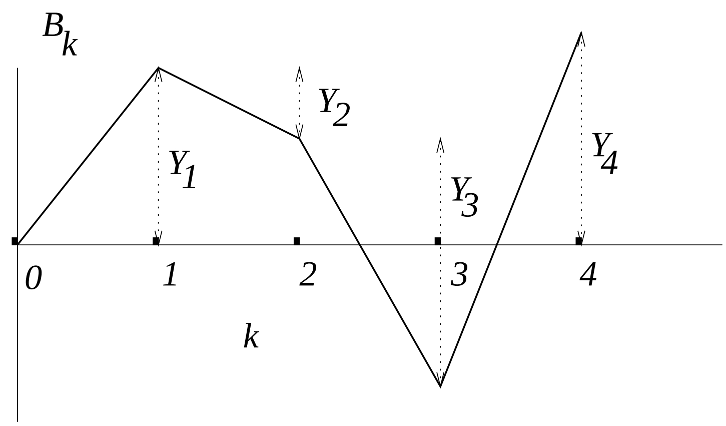



Figure 12.1: *Discrete-time Brownian motion.*

**Theorem 1.35** $\{B_k\}_{k=0}^n$ *is a Markov process.*

**Proof:** Note that

$$\mathbb{E}[h(B_{k+1})|\mathcal{F}_k] = \mathbb{E}[h(Y_{k+1} + B_k)|\mathcal{F}_k].$$

Use the Independence Lemma. Define

$$g(b) = \mathbb{E}h(Y_{k+1} + b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(y+b)e^{-\frac{1}{2}y^2}\ dy.$$

Then

$$\mathbb{E}[h(Y_{k+1} + B_k)|\mathcal{F}_k] = g(B_k),$$

which is a function of $B_k$ alone.

## 12.2 The Stock Price Process

Given parameters:

- $\mu \in \mathbb{R}$, the *mean rate of return.*
- $\sigma > 0$, the *volatility.*
- $S_0 > 0$, the initial stock price.

The *stock price process* is then given by

$$S_k = S_0 \exp\left\{\sigma B_k + (\mu - \tfrac{1}{2}\sigma^2)k\right\},\ k = 0, \ldots, n.$$

Note that

$$S_{k+1} = S_k \exp\left\{\sigma Y_{k+1} + (\mu - \tfrac{1}{2}\sigma^2)\right\},$$

$$\begin{aligned} I\!E[S_{k+1}|\mathcal{F}_k] &= S_k I\!E[e^{\sigma Y_{k+1}}|\mathcal{F}_k].e^{\mu-\frac{1}{2}\sigma^2} \\ &= S_k e^{\frac{1}{2}\sigma^2} e^{\mu-\frac{1}{2}\sigma^2} \\ &= e^{\mu} S_k. \end{aligned}$$

Thus

$$\mu = \log \frac{I\!E[S_{k+1}|\mathcal{F}_k]}{S_k} = \log I\!E\left[\frac{S_{k+1}}{S_k}\middle|\mathcal{F}_k\right],$$

and

$$\operatorname{var}\left(\log \frac{S_{k+1}}{S_k}\right) = \operatorname{var}\left(\sigma Y_{k+1} + (\mu - \tfrac{1}{2}\sigma^2)\right) = \sigma^2.$$

## 12.3 Remainder of the Market

The other processes in the market are defined as follows.

Money market process:

$$M_k = e^{rk},\ k = 0, 1, \ldots, n.$$

Portfolio process:

- $\Delta_0, \Delta_1, \ldots, \Delta_{n-1}$,
- Each $\Delta_k$ is $\mathcal{F}_k$-measurable.

Wealth process:

- $X_0$ given, nonrandom.
- $$\begin{aligned} X_{k+1} &= \Delta_k S_{k+1} + e^r(X_k - \Delta_k S_k) \\ &= \Delta_k(S_{k+1} - e^r S_k) + e^r X_k \end{aligned}$$
- Each $X_k$ is $\mathcal{F}_k$-measurable.

Discounted wealth process:

$$\frac{X_{k+1}}{M_{k+1}} = \Delta_k\left(\frac{S_{k+1}}{M_{k+1}} - \frac{S_k}{M_k}\right) + \frac{X_k}{M_k}.$$

## 12.4 Risk-Neutral Measure

**Definition 12.1** Let $\widetilde{I\!P}$ be a probability measure on $(\Omega, \mathcal{F})$, equivalent to the market measure P. If $\left\{\frac{S_k}{M_k}\right\}_{k=0}^{n}$ is a martingale under $\widetilde{I\!P}$, we say that $\widetilde{I\!P}$ is a *risk-neutral measure.*

**Theorem 4.36** *If $\widetilde{I\!P}$ is a risk-neutral measure, then every discounted wealth process $\left\{\frac{X_k}{M_k}\right\}_{k=0}^{n}$ is a martingale under $\widetilde{I\!P}$, regardless of the portfolio process used to generate it.*

**Proof:**

$$\begin{aligned}\widetilde{I\!E}\left[\frac{X_{k+1}}{M_{k+1}}\middle|\mathcal{F}_k\right] &= \widetilde{I\!E}\left[\Delta_k\left(\frac{S_{k+1}}{M_{k+1}}-\frac{S_k}{M_k}\right)+\frac{X_k}{M_k}\middle|\mathcal{F}_k\right]\\ &= \Delta_k\left(\widetilde{I\!E}\left[\frac{S_{k+1}}{M_{k+1}}\middle|\mathcal{F}_k\right]-\frac{S_k}{M_k}\right)+\frac{X_k}{M_k}\\ &= \frac{X_k}{M_k}.\end{aligned}$$

■

## 12.5 Risk-Neutral Pricing

Let $V_n$ be the payoff at time $n$, and say it is $\mathcal{F}_n$-measurable. Note that $V_n$ may be path-dependent. Hedging a short position:

- Sell the simple European derivative security $V_n$.
- Receive $X_0$ at time 0.
- Construct a portfolio process $\Delta_0, \ldots, \Delta_{n-1}$ which starts with $X_0$ and ends with $X_n = V_n$.
- If there is a risk-neutral measure $\widetilde{I\!P}$, then

$$X_0 = \widetilde{I\!E}\frac{X_n}{M_n} = \widetilde{I\!E}\frac{V_n}{M_n}.$$

**Remark 12.1** Hedging in this "semi-continuous" model is usually not possible because there are not enough trading dates. This difficulty will disappear when we go to the fully continuous model.

## 12.6 Arbitrage

**Definition 12.2** An *arbitrage* is a portfolio which starts with $X_0 = 0$ and ends with $X_n$ satisfying

$$I\!P(X_n \geq 0) = 1,\ I\!P(X_n > 0) > 0.$$

(P here is the market measure).

**Theorem 6.37 (Fundamental Theorem of Asset Pricing: Easy part)** *If there is a risk-neutral measure, then there is no arbitrage.*

**Proof:** Let $\widetilde{I\!\!P}$ be a risk-neutral measure, let $X_0 = 0$, and let $X_n$ be the final wealth corresponding to any portfolio process. Since $\left\{\frac{X_k}{M_k}\right\}_{k=0}^{n}$ is a martingale under $\widetilde{I\!\!P}$,

$$\widetilde{I\!\!E}\frac{X_n}{M_n} = \widetilde{I\!\!E}\frac{X_0}{M_0} = 0. \tag{6.1}$$

Suppose $I\!\!P(X_n \geq 0) = 1$. We have

$$I\!\!P(X_n \geq 0) = 1 \Longrightarrow I\!\!P(X_n < 0) = 0 \Longrightarrow \widetilde{I\!\!P}(X_n < 0) = 0 \Longrightarrow \widetilde{I\!\!P}(X_n \geq 0) = 1. \tag{6.2}$$

(6.1) and (6.2) imply $\widetilde{I\!\!P}(X_n = 0) = 1$. We have

$$\widetilde{I\!\!P}(X_n = 0) = 1 \Longrightarrow \widetilde{I\!\!P}(X_n > 0) = 0 \Longrightarrow I\!\!P(X_n > 0) = 0.$$

This is not an arbitrage. ■

## 12.7 Stalking the Risk-Neutral Measure

Recall that

- $Y_1, Y_2, \ldots, Y_n$ are independent, standard normal random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
- $S_k = S_0 \exp\left\{\sigma B_k + (\mu - \frac{1}{2}\sigma^2)k\right\}$.
- 

$$\begin{aligned} S_{k+1} &= S_0 \exp\left\{\sigma(B_k + Y_{k+1}) + (\mu - \tfrac{1}{2}\sigma^2)(k+1)\right\} \\ &= S_k \exp\left\{\sigma Y_{k+1} + (\mu - \tfrac{1}{2}\sigma^2)\right\}. \end{aligned}$$

Therefore,

$$\frac{S_{k+1}}{M_{k+1}} = \frac{S_k}{M_k}.\exp\left\{\sigma Y_{k+1} + (\mu - r - \tfrac{1}{2}\sigma^2)\right\},$$

$$\begin{aligned} I\!\!E\left[\frac{S_{k+1}}{M_{k+1}}\middle|\mathcal{F}_k\right] &= \frac{S_k}{M_k}.I\!\!E\left[\exp\left\{\sigma Y_{k+1}\right\}|\mathcal{F}_k\right].\exp\{\mu - r - \tfrac{1}{2}\sigma^2\} \\ &= \frac{S_k}{M_k}.\exp\{\tfrac{1}{2}\sigma^2\}.\exp\{\mu - r - \tfrac{1}{2}\sigma^2\} \\ &= e^{\mu - r}.\frac{S_k}{M_k}. \end{aligned}$$

If $\mu = r$, the market measure is risk neutral. If $\mu \neq r$, we must seek further.

$$
\begin{aligned}
\frac{S_{k+1}}{M_{k+1}} &= \frac{S_k}{M_k}.\exp\left\{\sigma Y_{k+1} + (\mu - r - \tfrac{1}{2}\sigma^2)\right\} \\
&= \frac{S_k}{M_k}.\exp\left\{\sigma (Y_{k+1} + \tfrac{\mu - r}{\sigma}) - \tfrac{1}{2}\sigma^2\right\} \\
&= \frac{S_k}{M_k}.\exp\left\{\sigma \tilde{Y}_{k+1} - \tfrac{1}{2}\sigma^2\right\},
\end{aligned}
$$

where

$$\tilde{Y}_{k+1} = Y_{k+1} + \tfrac{\mu - r}{\sigma}.$$

The quantity $\frac{\mu - r}{\sigma}$ is denoted $\theta$ and is called the *market price of risk*.

We want a probability measure $\widetilde{I\!P}$ under which $\tilde{Y}_1, \dots, \tilde{Y}_n$ are independent, standard normal random variables. Then we would have

$$
\begin{aligned}
\widetilde{I\!E}\left[\frac{S_{k+1}}{M_{k+1}}\middle|\mathcal{F}_k\right] &= \frac{S_k}{M_k}.\widetilde{I\!E}\left[\exp\{\sigma\tilde{Y}_{k+1}\}|\mathcal{F}_k\right].\exp\{-\tfrac{1}{2}\sigma^2\} \\
&= \frac{S_k}{M_k}.\exp\{\tfrac{1}{2}\sigma^2\}.\exp\{-\tfrac{1}{2}\sigma^2\} \\
&= \frac{S_k}{M_k}.
\end{aligned}
$$

**Cameron-Martin-Girsanov's Idea:** Define the random variable

$$Z = \exp\left[\sum_{j=1}^{n}(-\theta Y_j - \tfrac{1}{2}\theta^2)\right].$$

Properties of $Z$:

- $Z \geq 0$.
- $$
\begin{aligned}
I\!EZ &= I\!E\exp\left\{\sum_{j=1}^{n}(-\theta Y_j)\right\}.\exp\left\{-\frac{n}{2}\theta^2\right\} \\
&= \exp\left\{\frac{n}{2}\theta^2\right\}.\exp\left\{-\frac{n}{2}\theta^2\right\} = 1.
\end{aligned}
$$

Define

$$\widetilde{I\!P}(A) = \int_A Z\, dI\!P \quad \forall A \in \mathcal{F}.$$

Then $\widetilde{I\!P}(A) \geq 0$ for all $A \in \mathcal{F}$ and

$$\widetilde{I\!P}(\Omega) = I\!EZ = 1.$$

In other words, $\widetilde{I\!P}$ is a probability measure.

We show that $\widetilde{I\!P}$ is a risk-neutral measure. For this, it suffices to show that

$$\tilde{Y}_1 = Y_1 + \theta,\ \ldots,\ \tilde{Y}_n = Y_n + \theta$$

are independent, standard normal under $\widetilde{I\!P}$.

**Verification:**

- $Y_1, Y_2, \ldots, Y_n$: Independent, standard normal under IP, and

$$I\!E \exp\left[\sum_{j=1}^{n} u_j Y_j\right] = \exp\left[\sum_{j=1}^{n} \tfrac{1}{2} u_j^2\right].$$

- $\tilde{Y} = Y_1 + \theta,\ \ldots,\ \tilde{Y}_n = Y_n + \theta$.
- $Z > 0$ almost surely.
- $Z = \exp\left[\sum_{j=1}^{n} (-\theta Y_j - \frac{1}{2}\theta^2)\right]$,

$$\widetilde{I\!P}(A) = \int_A Z\, dI\!P \quad \forall A \in \mathcal{F},$$

$\widetilde{I\!E} X = I\!E(XZ)$ for every random variable $X$.

- Compute the moment generating function of $(\tilde{Y}_1, \ldots, \tilde{Y}_n)$ under $\widetilde{I\!P}$:

$$\begin{aligned}
\widetilde{I\!E} \exp\left[\sum_{j=1}^{n} u_j \tilde{Y}_j\right] &= I\!E \exp\left[\sum_{j=1}^{n} u_j (Y_j + \theta) + \sum_{j=1}^{n} (-\theta Y_j - \tfrac{1}{2}\theta^2)\right] \\
&= I\!E \exp\left[\sum_{j=1}^{n} (u_j - \theta) Y_j\right] . \exp\left[\sum_{j=1}^{n} (u_j \theta - \tfrac{1}{2}\theta^2)\right] \\
&= \exp\left[\sum_{j=1}^{n} \tfrac{1}{2} (u_j - \theta)^2\right] . \exp\left[\sum_{j=1}^{n} (u_j \theta - \tfrac{1}{2}\theta^2)\right] \\
&= \exp\left[\sum_{j=1}^{n} \left( (\tfrac{1}{2} u_j^2 - u_j \theta + \tfrac{1}{2}\theta^2) + (u_j \theta - \tfrac{1}{2}\theta^2) \right)\right] \\
&= \exp\left[\sum_{j=1}^{n} \tfrac{1}{2} u_j^2\right].
\end{aligned}$$

## 12.8 Pricing a European Call

Stock price at time $n$ is

$$\begin{aligned} S_n &= S_0 \exp\left\{\sigma B_n + (\mu - \tfrac{1}{2}\sigma^2)n\right\} \\ &= S_0 \exp\left\{\sigma \sum_{j=1}^{n} Y_j + (\mu - \tfrac{1}{2}\sigma^2)n\right\} \\ &= S_0 \exp\left\{\sigma \sum_{j=1}^{n} (Y_j + \tfrac{\mu - r}{\sigma}) - (\mu - r)n + (\mu - \tfrac{1}{2}\sigma^2)n\right\} \\ &= S_0 \exp\left\{\sigma \sum_{j=1}^{n} \tilde{Y}_j \ + \ (r - \tfrac{1}{2}\sigma^2)n\right\}. \end{aligned}$$

Payoff at time $n$ is $(S_n - K)^+$. Price at time zero is

$$\begin{aligned} \widetilde{I\!E}\frac{(S_n - K)^+}{M_n} &= \widetilde{I\!E}\left[e^{-rn}\left(S_0 \exp\left\{\sigma \sum_{j=1}^{n} \tilde{Y}_j \ + \ (r - \tfrac{1}{2}\sigma^2)n\right\} - K\right)^+\right] \\ &= \int_{-\infty}^{\infty} e^{-rn}\left(S_0 \exp\left\{\sigma b + (r - \tfrac{1}{2}\sigma^2)n\right\} - K\right)^+ . \frac{1}{\sqrt{2\pi n}} e^{-\frac{b^2}{2n^2}} \ db \end{aligned}$$

since $\sum_{j=1}^{n} \tilde{Y}_j$ is normal with mean 0, variance $n$, under $\widetilde{I\!P}$.

This is the *Black-Scholes* price. It does not depend on $\mu$.

# Chapter 13

# Brownian Motion

## 13.1 Symmetric Random Walk

Toss a fair coin infinitely many times. Define

$$X_j(\omega) = \begin{cases} 1 & \text{if} \quad \omega_j = H, \\ -1 & \text{if} \quad \omega_j = T. \end{cases}$$

Set

$$\begin{aligned} M_0 &= 0 \\ M_k &= \sum_{j=1}^{k} X_j, \qquad k \geq 1. \end{aligned}$$

## 13.2 The Law of Large Numbers

We will use the method of moment generating functions to derive the Law of Large Numbers:

**Theorem 2.38 (Law of Large Numbers:)**

$$\frac{1}{k} M_k \to 0 \quad \textit{almost surely, as} \quad k \to \infty.$$

**Proof:**

$$\begin{aligned}
\varphi_k(u) &= \mathbb{E}\exp\left\{\frac{u}{k}M_k\right\} \\
&= \mathbb{E}\exp\left\{\sum_{j=1}^{k}\frac{u}{k}X_j\right\} && \text{(Def. of } M_k.\text{)} \\
&= \prod_{j=1}^{k}\mathbb{E}\exp\left\{\frac{u}{k}X_j\right\} && \text{(Independence of the } X_j\text{'s)} \\
&= \left(\tfrac{1}{2}e^{\frac{u}{k}} + \tfrac{1}{2}e^{-\frac{u}{k}}\right)^k,
\end{aligned}$$

which implies,

$$\log\varphi_k(u) = k\log\left(\tfrac{1}{2}e^{\frac{u}{k}} + \tfrac{1}{2}e^{-\frac{u}{k}}\right)$$

Let $x = \frac{1}{k}$. Then

$$\begin{aligned}
\lim_{k\to\infty}\log\varphi_k(u) &= \lim_{x\to 0}\frac{\log\left(\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}\right)}{x} \\
&= \lim_{x\to 0}\frac{\frac{u}{2}e^{ux} - \frac{u}{2}e^{-ux}}{\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}} && \text{(L'Hôpital's Rule)} \\
&= 0.
\end{aligned}$$

Therefore,

$$\lim_{k\to\infty}\varphi_k(u) = e^0 = 1,$$

which is the m.g.f. for the constant 0. ■

## 13.3 Central Limit Theorem

We use the method of moment generating functions to prove the Central Limit Theorem.

**Theorem 3.39 (Central Limit Theorem)**

$$\frac{1}{\sqrt{k}}M_k \to \quad \textit{Standard normal, as} \quad k\to\infty.$$

**Proof:**

$$\begin{aligned}
\varphi_k(u) &= \mathbb{E}\exp\left\{\frac{u}{\sqrt{k}}M_k\right\} \\
&= \left(\tfrac{1}{2}e^{\frac{u}{\sqrt{k}}} + \tfrac{1}{2}e^{-\frac{u}{\sqrt{k}}}\right)^k,
\end{aligned}$$

so that,

$$\log \varphi_k(u) = k \log\left(\tfrac{1}{2}e^{\frac{u}{\sqrt{k}}} + \tfrac{1}{2}e^{-\frac{u}{\sqrt{k}}}\right).$$

Let $x = \frac{1}{\sqrt{k}}$. Then

$$\begin{aligned}
\lim_{k\to\infty} \log \varphi_k(u) &= \lim_{x\to 0} \frac{\log\left(\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}\right)}{x^2} \\
&= \lim_{x\to 0} \frac{\frac{u}{2}e^{ux} - \frac{u}{2}e^{-ux}}{2x\left(\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}\right)} && \text{(L'Hôpital's Rule)} \\
&= \lim_{x\to 0} \frac{1}{\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}} . \lim_{x\to 0} \frac{\frac{u}{2}e^{ux} - \frac{u}{2}e^{-ux}}{2x} \\
&= \lim_{x\to 0} \frac{\frac{u}{2}e^{ux} - \frac{u}{2}e^{-ux}}{2x} \\
&= \lim_{x\to 0} \frac{\frac{u^2}{2}e^{ux} - \frac{u^2}{2}e^{-ux}}{2} && \text{(L'Hôpital's Rule)} \\
&= \tfrac{1}{2}u^2.
\end{aligned}$$

Therefore,

$$\lim_{k\to\infty} \varphi_k(u) = e^{\frac{1}{2}u^2},$$

which is the m.g.f. for a standard normal random variable. ■

## 13.4 Brownian Motion as a Limit of Random Walks

Let $n$ be a positive integer. If $t \geq 0$ is of the form $\frac{k}{n}$, then set

$$B^{(n)}(t) = \frac{1}{\sqrt{n}} M_{tn} = \frac{1}{\sqrt{n}} M_k.$$

If $t \geq 0$ is not of the form $\frac{k}{n}$, then define $B^{(n)}(t)$ by linear interpolation (See Fig. 13.1).

Here are some properties of $B^{(100)}(t)$:

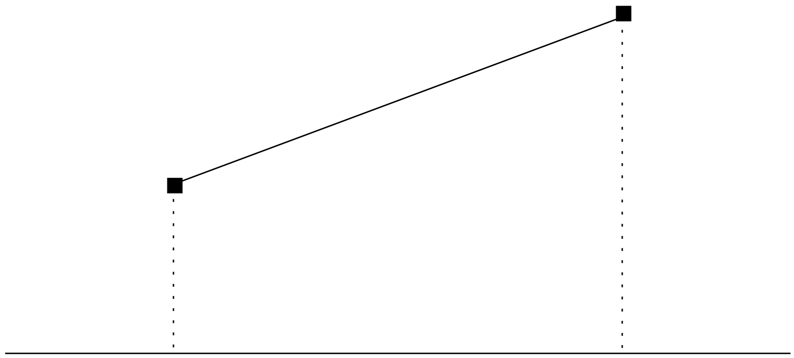



Figure 13.1: *Linear Interpolation to define* $B^{(n)}(t)$.

Properties of $\quad B^{(100)}(1)$ :

$$B^{(100)}(1) = \frac{1}{10}\sum_{j=1}^{100} X_j \qquad \text{(Approximately normal)}$$

$$\mathbb{E}B^{(100)}(1) = \frac{1}{10}\sum_{j=1}^{100} \mathbb{E}X_j = 0.$$

$$\text{var}(B^{(100)}(1)) = \frac{1}{100}\sum_{j=1}^{100} \text{var}(X_j) = 1$$

Properties of $\quad B^{(100)}(2)$ :

$$B^{(100)}(2) = \frac{1}{10}\sum_{j=1}^{200} X_j \qquad \text{(Approximately normal)}$$

$$\mathbb{E}B^{(100)}(2) = 0.$$

$$\text{var}(B^{(100)}(2)) = 2.$$

Also note that:

- $B^{(100)}(1)$ and $B^{(100)}(2) - B^{(100)}(1)$ are independent.
- $B^{(100)}(t)$ is a continuous function of $t$.

To get Brownian motion, let $n\to\infty$ in $B^{(n)}(t), \quad t \ge 0$.

## 13.5 Brownian Motion

(Please refer to Oksendal, Chapter 2.)

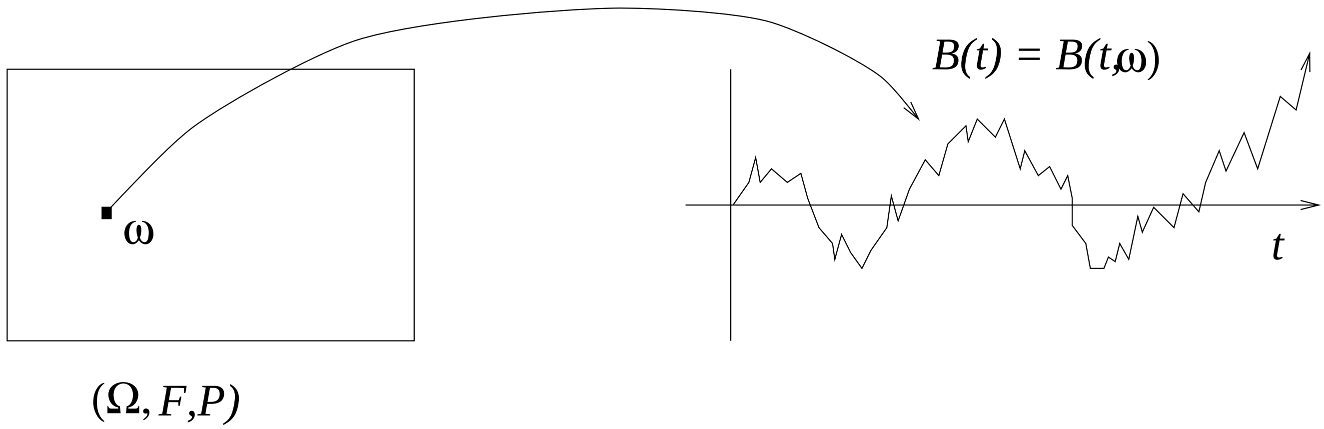



Figure 13.2: *Continuous-time Brownian Motion.*

A random variable $B(t)$ (see Fig. 13.2) is called a Brownian Motion if it satisfies the following properties:

1. $B(0) = 0$,
2. $B(t)$ is a continuous function of $t$;
3. $B$ has independent, normally distributed increments: If

$$0 = t_0 < t_1 < t_2 < \ldots < t_n$$

and

$$Y_1 = B(t_1) - B(t_0), \quad Y_2 = B(t_2) - B(t_1), \quad \ldots \quad Y_n = B(t_n) - B(t_{n-1}),$$

then

- $Y_1, Y_2, \ldots, Y_n$ are independent,
- $I\!\!E Y_j = 0 \qquad \forall j$,
- $\text{var}(Y_j) = t_j - t_{j-1} \qquad \forall j$.

## 13.6 Covariance of Brownian Motion

Let $0 \le s \le t$ be given. Then $B(s)$ and $B(t) - B(s)$ are independent, so $B(s)$ and $B(t) = (B(t) - B(s)) + B(s)$ are jointly normal. Moreover,

$$\begin{aligned}
I\!\!E B(s) &= 0, \qquad \text{var}(B(s)) = s, \\
I\!\!E B(t) &= 0, \qquad \text{var}(B(t)) = t, \\
I\!\!E B(s)B(t) &= I\!\!E B(s)[(B(t) - B(s)) + B(s)] \\
&= \underbrace{I\!\!E B(s)(B(t) - B(s))}_{0} + \underbrace{I\!\!E B^2(s)}_{s} \\
&= s.
\end{aligned}$$

Thus for any $s \geq 0$, $t \geq 0$ (not necessarily $s \leq t$), we have

$$\mathbb{E}B(s)B(t) = s \wedge t.$$

## 13.7 Finite-Dimensional Distributions of Brownian Motion

Let

$$0 < t_1 < t_2 < \ldots < t_n$$

be given. Then

$$(B(t_1), B(t_2), \ldots, B(t_n))$$

is jointly normal with covariance matrix

$$C = \begin{bmatrix} \mathbb{E}B^2(t_1) & \mathbb{E}B(t_1)B(t_2) & \ldots & \mathbb{E}B(t_1)B(t_n) \\ \mathbb{E}B(t_2)B(t_1) & \mathbb{E}B^2(t_2) & \ldots & \mathbb{E}B(t_2)B(t_n) \\ \cdots & \cdots & \cdots & \cdots \\ \mathbb{E}B(t_n)B(t_1) & \mathbb{E}B(t_n)B(t_2) & \ldots & \mathbb{E}B^2(t_n) \end{bmatrix}$$
$$= \begin{bmatrix} t_1 & t_1 & \ldots & t_1 \\ t_1 & t_2 & \ldots & t_2 \\ \cdots & \cdots & \cdots & \cdots \\ t_1 & t_2 & \ldots & t_n \end{bmatrix}$$

## 13.8 Filtration generated by a Brownian Motion

$$\{\mathcal{F}(t)\}_{t \geq 0}$$

Required properties:

- For each $t$, $B(t)$ is $\mathcal{F}(t)$-measurable,
- For each $t$ and for $t < t_1 < t_2 < \cdots < t_n$, the Brownian motion increments

  $$B(t_1) - B(t), \quad B(t_2) - B(t_1), \quad \ldots, \quad B(t_n) - B(t_{n-1})$$

  are *independent of* $\mathcal{F}(t)$.

Here is one way to construct $\mathcal{F}(t)$. First fix $t$. Let $s \in [0, t]$ and $C \in \mathcal{B}(\mathbb{R})$ be given. Put the set

$$\{B(s) \in C\} = \{\omega : B(s, \omega) \in C\}$$

in $\mathcal{F}(t)$. Do this for all possible numbers $s \in [0, t]$ and $C \in \mathcal{B}(\mathbb{R})$. Then put in every other set required by the $\sigma$-algebra properties.

This $\mathcal{F}(t)$ contains exactly the information learned by observing the Brownian motion upto time $t$. $\{\mathcal{F}(t)\}_{t \geq 0}$ is called the *filtration generated by the Brownian motion.*

## 13.9 Martingale Property

**Theorem 9.40** *Brownian motion is a martingale.*

**Proof:** Let $0 \le s \le t$ be given. Then

$$\begin{aligned} \mathbb{E}[B(t)|\mathcal{F}(s)] &= \mathbb{E}[(B(t)-B(s)) + B(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[B(t)-B(s)] + B(s) \\ &= B(s). \end{aligned}$$

■

**Theorem 9.41** *Let $\theta \in \mathbb{R}$ be given. Then*

$$Z(t) = \exp\left\{-\theta B(t) - \tfrac{1}{2}\theta^2 t\right\}$$

*is a martingale.*

**Proof:** Let $0 \le s \le t$ be given. Then

$$\begin{aligned} \mathbb{E}[Z(t)|\mathcal{F}(s)] &= \mathbb{E}\left[\exp\{-\theta(B(t)-B(s)+B(s)) - \tfrac{1}{2}\theta^2((t-s)+s)\}\middle|\mathcal{F}(s)\right] \\ &= \mathbb{E}\left[Z(s)\exp\{-\theta(B(t)-B(s)) - \tfrac{1}{2}\theta^2(t-s)\}\middle|\mathcal{F}(s)\right] \\ &= Z(s)\mathbb{E}\left[\exp\{-\theta(B(t)-B(s)) - \tfrac{1}{2}\theta^2(t-s)\}\right] \\ &= Z(s)\exp\left\{\tfrac{1}{2}(-\theta)^2\operatorname{var}(B(t)-B(s)) - \tfrac{1}{2}\theta^2(t-s)\right\} \\ &= Z(s). \end{aligned}$$

■

## 13.10 The Limit of a Binomial Model

Consider the $n$'th Binomial model with the following parameters:

- $u_n = 1 + \frac{\sigma}{\sqrt{n}}$. "Up" factor. ($\sigma > 0$).
- $d_n = 1 - \frac{\sigma}{\sqrt{n}}$. "Down" factor.
- $r = 0$.
- $\tilde{p}_n = \frac{1-d_n}{u_n-d_n} = \frac{\sigma/\sqrt{n}}{2\sigma/\sqrt{n}} = \frac{1}{2}$.
- $\tilde{q}_n = \frac{1}{2}$.

Let $\sharp_k(H)$ denote the number of $H$ in the first $k$ tosses, and let $\sharp_k(T)$ denote the number of $T$ in the first $k$ tosses. Then

$$\begin{aligned}\sharp_k(H) + \sharp_k(T) &= k,\\ \sharp_k(H) - \sharp_k(T) &= M_k,\end{aligned}$$

which implies,

$$\begin{aligned}\sharp_k(H) &= \tfrac{1}{2}(k + M_k)\\ \sharp_k(T) &= \tfrac{1}{2}(k - M_k).\end{aligned}$$

In the $n$'th model, take $n$ steps per unit time. Set $S_0^{(n)} = 1$. Let $t = \frac{k}{n}$ for some $k$, and let

$$S^{(n)}(t) = \left(1 + \frac{\sigma}{\sqrt{n}}\right)^{\frac{1}{2}(nt+M_{nt})} \left(1 - \frac{\sigma}{\sqrt{n}}\right)^{\frac{1}{2}(nt-M_{nt})}.$$

Under $\widetilde{I\!P}$, the price process $S^{(n)}$ is a martingale.

**Theorem 10.42** *As $n\to\infty$, the distribution of $S^{(n)}(t)$ converges to the distribution of*

$$\exp\{\sigma B(t) - \tfrac{1}{2}\sigma^2 t\},$$

*where $B$ is a Brownian motion. Note that the correction $-\frac{1}{2}\sigma^2 t$ is necessary in order to have a martingale.*

**Proof:** Recall that from the Taylor series we have

$$\log(1+x) = x - \tfrac{1}{2}x^2 + O(x^3),$$

so

$$\begin{aligned}\log S^{(n)}(t) &= \tfrac{1}{2}(nt + M_{nt})\log(1 + \frac{\sigma}{\sqrt{n}}) + \tfrac{1}{2}(nt - M_{nt})\log(1 - \frac{\sigma}{\sqrt{n}})\\
&= nt\left(\tfrac{1}{2}\log(1 + \frac{\sigma}{\sqrt{n}}) + \tfrac{1}{2}\log(1 - \frac{\sigma}{\sqrt{n}})\right)\\
&\quad + M_{nt}\left(\tfrac{1}{2}\log(1 + \frac{\sigma}{\sqrt{n}}) - \tfrac{1}{2}\log(1 - \frac{\sigma}{\sqrt{n}})\right)\\
&= nt\left(\tfrac{1}{2}\frac{\sigma}{\sqrt{n}} - \frac{1}{4}\frac{\sigma^2}{n} - \tfrac{1}{2}\frac{\sigma}{\sqrt{n}} - \frac{1}{4}\frac{\sigma^2}{n} + O(n^{-3/2})\right)\\
&\quad + M_{nt}\left(\tfrac{1}{2}\frac{\sigma}{\sqrt{n}} - \frac{1}{4}\frac{\sigma^2}{n} + \tfrac{1}{2}\frac{\sigma}{\sqrt{n}} + \frac{1}{4}\frac{\sigma^2}{n} + O(n^{-3/2})\right)\\
&= -\tfrac{1}{2}\sigma^2 t + O(n^{-\frac{1}{2}})\\
&\quad + \sigma\underbrace{\left(\frac{1}{\sqrt{n}}M_{nt}\right)}_{\to B_t} + \underbrace{\left(\frac{1}{n}M_{nt}\right)}_{\to 0} O(n^{-\frac{1}{2}})\end{aligned}$$

As $n\to\infty$, the distribution of $\log S^{(n)}(t)$ approaches the distribution of $\sigma B(t) - \frac{1}{2}\sigma^2 t$. ■

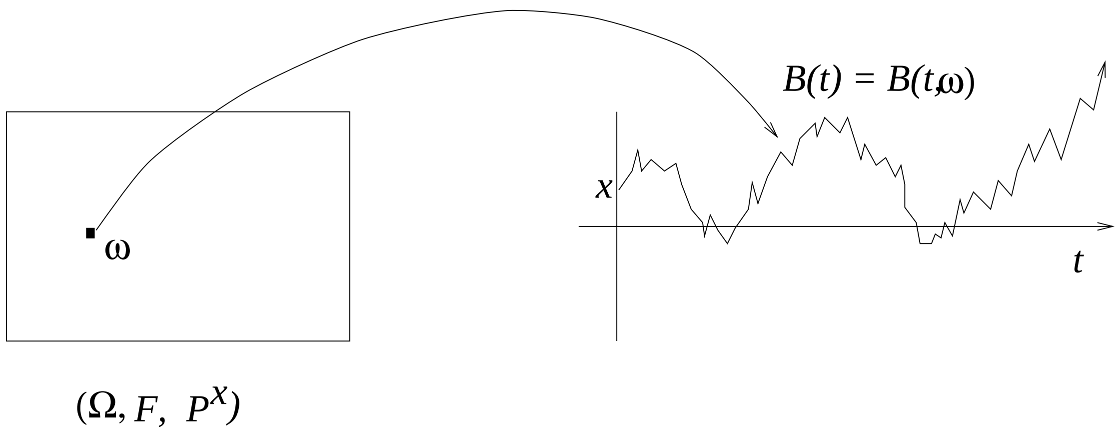



Figure 13.3: *Continuous-time Brownian Motion, starting at* $x \neq 0$.

## 13.11 Starting at Points Other Than 0

(The remaining sections in this chapter were taught Dec 7.)

For a Brownian motion $B(t)$ that starts at 0, we have:

$$I\!\!P(B(0) = 0) = 1.$$

For a Brownian motion $B(t)$ that starts at $x$, denote the corresponding probability measure by $I\!\!P^x$ (See Fig. 13.3), and for such a Brownian motion we have:

$$I\!\!P^x(B(0) = x) = 1.$$

Note that:

- If $x \neq 0$, then $I\!\!P^x$ puts all its probability on a completely different set from P.
- The distribution of $B(t)$ under $I\!\!P^x$ is the same as the distribution of $x + B(t)$ under P.

## 13.12 Markov Property for Brownian Motion

We prove that

**Theorem 12.43** *Brownian motion has the Markov property.*

**Proof:**

Let $s \geq 0, \quad t \geq 0$ be given (See Fig. 13.4).

$$I\!\!E\left[h(B(s+t))\Big|\mathcal{F}(s)\right] = I\!\!E\left[h(\underbrace{B(s+t) - B(s)}_{\text{Independent of } \mathcal{F}(s)} + \underbrace{B(s)}_{\mathcal{F}(s)\text{-measurable}})\Big|\mathcal{F}(s)\right]$$

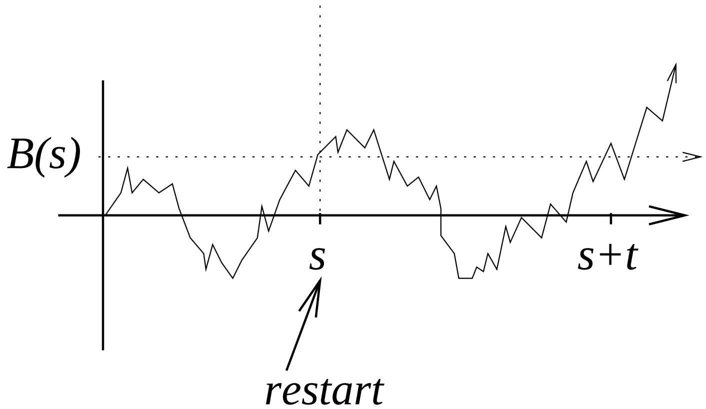



Figure 13.4: *Markov Property of Brownian Motion.*

Use the Independence Lemma. Define

$$\begin{aligned} g(x) &= \mathbb{E}\left[h(\,B(s+t) - B(s)\, + x\,)\right] \\ &= \mathbb{E}\left[h(\,x + \underbrace{B(t)}_{\text{same distribution as } B(s+t)-B(s)}\,)\right] \\ &= \mathbb{E}^x h(B(t)). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}\left[h\left(B(s+t)\,\right)\middle|\mathcal{F}(s)\right] &= g(B(s)) \\ &= E^{B(s)} h(B(t)). \end{aligned}$$

■

In fact Brownian motion has the *strong Markov property.*

**Example 13.1 (Strong Markov Property)** See Fig. 13.5. Fix $x > 0$ and define

$$\tau = \min\{t \geq 0; \quad B(t) = x\}.$$

Then we have:

$$\mathbb{E}\left[h(\,B(\tau+t)\,)\middle|\mathcal{F}(\tau)\right] = g(B(\tau)) = \mathbb{E}^x h(B(t)).$$

■

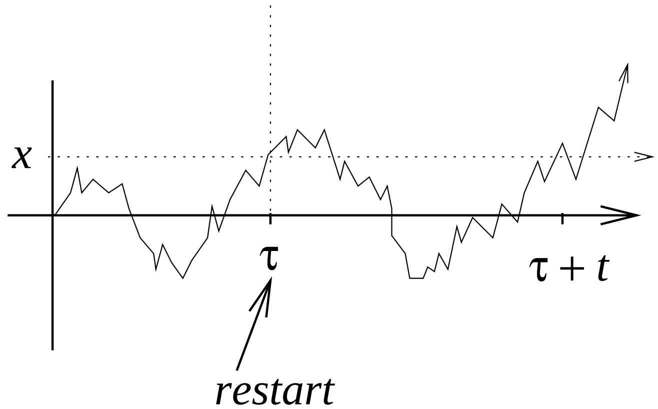



Figure 13.5: *Strong Markov Property of Brownian Motion.*

## 13.13 Transition Density

Let $p(t, x, y)$ be the probability that the Brownian motion changes value from $x$ to $y$ in time $t$, and let $\tau$ be defined as in the previous section.

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$$

$$g(x) = \mathbb{E}^x h(B(t)) = \int_{-\infty}^{\infty} h(y)p(t, x, y)\, dy.$$

$$\mathbb{E}\left[h(B(s+t))\middle|\mathcal{F}(s)\right] = g(B(s)) = \int_{-\infty}^{\infty} h(y)p(t, B(s), y)\, dy.$$

$$\mathbb{E}\left[h(B(\tau+t))\middle|\mathcal{F}(\tau)\right] = \int_{-\infty}^{\infty} h(y)p(t, x, y)\, dy.$$

## 13.14 First Passage Time

Fix $x > 0$. Define

$$\tau = \min\{t \geq 0; \quad B(t) = x\}.$$

Fix $\theta > 0$. Then

$$\exp\left\{\theta B(t \wedge \tau) - \tfrac{1}{2}\theta^2 (t \wedge \tau)\right\}$$

is a martingale, and

$$\mathbb{E} \exp\left\{\theta B(t \wedge \tau) - \tfrac{1}{2}\theta^2 (t \wedge \tau)\right\} = 1.$$

We have

$$\lim_{t\to\infty} \exp\left\{-\tfrac{1}{2}\theta^2(t\wedge\tau)\right\} = \begin{cases} e^{-\frac{1}{2}\theta^2\tau} & \text{if } \tau<\infty, \\ 0 & \text{if } \tau=\infty, \end{cases} \tag{14.1}$$

$$0 \le \exp\{\theta B(t\wedge\tau) - \tfrac{1}{2}\theta^2(t\wedge\tau)\} \le e^{\theta x}.$$

Let $t\to\infty$ in (14.1), using the Bounded Convergence Theorem, to get

$$I\!\!E\left[\exp\{\theta x - \tfrac{1}{2}\theta^2\tau\}\mathbf{1}_{\{\tau<\infty\}}\right] = 1.$$

Let $\theta\downarrow 0$ to get $I\!\!E\mathbf{1}_{\{\tau<\infty\}} = 1$, so

$$I\!\!P\{\tau<\infty\} = 1,$$

$$I\!\!E\exp\{-\tfrac{1}{2}\theta^2\tau\} = e^{-\theta x}. \tag{14.2}$$

Let $\alpha = \frac{1}{2}\theta^2$. We have the m.g.f.:

$$I\!\!E e^{-\alpha\tau} = e^{-x\sqrt{2\alpha}}, \qquad \alpha>0. \tag{14.3}$$

Differentiation of (14.3) w.r.t. $\alpha$ yields

$$-I\!\!E\left[\tau e^{-\alpha\tau}\right] = -\frac{x}{\sqrt{2\alpha}}e^{-x\sqrt{2\alpha}}.$$

Letting $\alpha\downarrow 0$, we obtain

$$I\!\!E\tau = \infty. \tag{14.4}$$

**Conclusion.** Brownian motion reaches level $x$ with probability 1. The expected time to reach level $x$ is infinite.

We use the Reflection Principle below (see Fig. 13.6).

$$\begin{aligned} I\!\!P\{\tau\le t, \quad B(t)<x\} &= I\!\!P\{B(t)>x\} \\ I\!\!P\{\tau\le t\} &= I\!\!P\{\tau\le t, B(t)<x\} + I\!\!P\{\tau\le t, B(t)>x\} \\ &= I\!\!P\{B(t)>x\} + I\!\!P\{B(t)>x\} \\ &= 2I\!\!P\{B(t)>x\} \\ &= \frac{2}{\sqrt{2\pi t}}\int_x^\infty e^{-\frac{y^2}{2t}}\,dy \end{aligned}$$

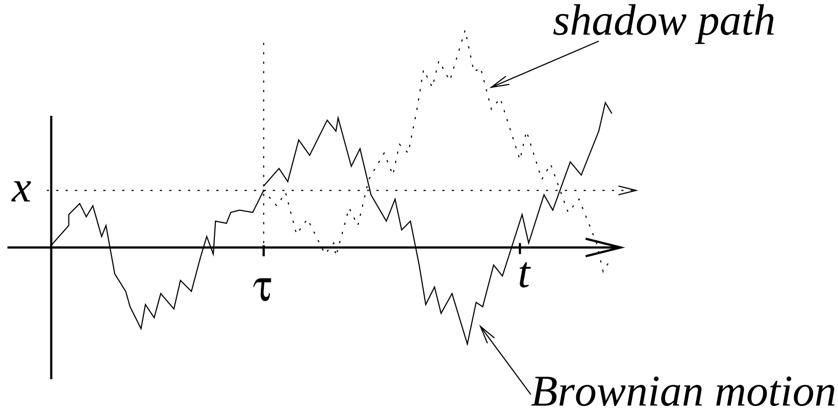



Figure 13.6: *Reflection Principle in Brownian Motion.*

Using the substitution $z = \frac{y}{\sqrt{t}}, \quad dz = \frac{dy}{\sqrt{t}}$ we get

$$\mathbb{P}\{\tau \le t\} = \frac{2}{\sqrt{2\pi}} \int_{\frac{x}{\sqrt{t}}}^{\infty} e^{-\frac{z^2}{2}} dz.$$

Density:

$$f_\tau(t) = \frac{\partial}{\partial t}\mathbb{P}\{\tau \le t\} = \frac{x}{\sqrt{2\pi t^3}} e^{-\frac{x^2}{2t}},$$

which follows from the fact that if

$$F(t) = \int_{a(t)}^{b} g(z)\, dz,$$

then

$$\frac{\partial F}{\partial t} = -\frac{\partial a}{\partial t} g(a(t)).$$

Laplace transform formula:

$$\mathbb{E}e^{-\alpha\tau} = \int_0^\infty e^{-\alpha t} f_\tau(t) dt = e^{-x\sqrt{2\alpha}}.$$

# Chapter 14

# The Itô Integral

The following chapters deal with *Stochastic Differential Equations in Finance*. References:

1. B. Oksendal, *Stochastic Differential Equations*, Springer-Verlag,1995
2. J. Hull, *Options, Futures and other Derivative Securities*, Prentice Hall, 1993.

## 14.1 Brownian Motion

(See Fig. 13.3.) $(\Omega, \mathcal{F}, \mathbb{P})$ is given, always in the background, even when not explicitly mentioned. **Brownian motion**, $B(t,\omega) : [0,\infty) \times \Omega \rightarrow I\!\!R$, has the following properties:

1. $B(0) = 0$; Technically, $I\!\!P\{\omega; B(0,\omega) = 0\} = 1$,
2. $B(t)$ is a continuous function of $t$,
3. If $0 = t_0 \leq t_1 \leq \ldots \leq t_n$, then the increments

$$B(t_1) - B(t_0),\ \ldots,\ B(t_n) - B(t_{n-1})$$

are *independent,normal*, and

$$I\!\!E[B(t_{k+1}) - B(t_k)] = 0,$$
$$I\!\!E[B(t_{k+1}) - B(t_k)]^2 = t_{k+1} - t_k.$$

## 14.2 First Variation

Quadratic variation is a measure of volatility. First we will consider *first variation*, $FV(f)$, of a function $f(t)$.

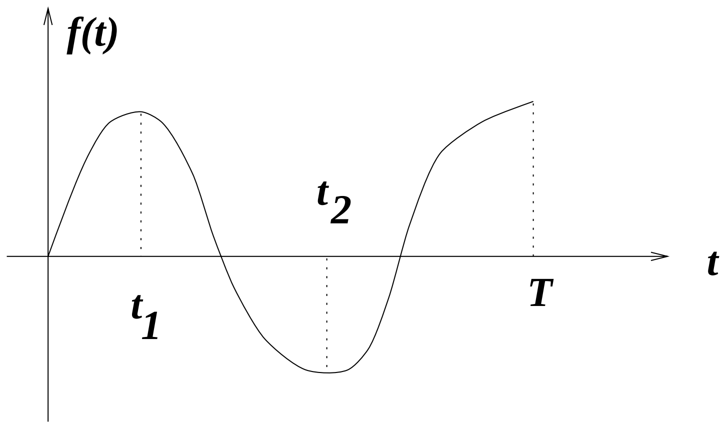



Figure 14.1: *Example function* $f(t)$.

For the function pictured in Fig. 14.1, the first variation over the interval $[0,T]$ is given by:

$$\begin{aligned} FV_{[0,T]}(f) &= [f(t_1)-f(0)] - [f(t_2)-f(t_1)] + [f(T)-f(t_2)] \\ &= \int_0^{t_1} f'(t)\,dt + \int_{t_1}^{t_2} (-f'(t))\,dt + \int_{t_2}^{T} f'(t)\,dt. \\ &= \int_0^T |f'(t)|\,dt. \end{aligned}$$

Thus, first variation measures the total amount of up and down motion of the path.

The general definition of first variation is as follows:

**Definition 14.1 (First Variation)** Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a *partition* of $[0,T]$, i.e.,

$$0 = t_0 \le t_1 \le \ldots \le t_n = T.$$

The *mesh* of the partition is defined to be

$$||\Pi|| = \max_{k=0,\ldots,n-1} (t_{k+1} - t_k).$$

We then define

$$FV_{[0,T]}(f) = \lim_{||\Pi|| \to 0} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|.$$

Suppose $f$ is differentiable. Then the Mean Value Theorem implies that in each subinterval $[t_k, t_{k+1}]$, there is a point $t_k^*$ such that

$$f(t_{k+1}) - f(t_k) = f'(t_k^*)(t_{k+1} - t_k).$$

Then

$$\sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| = \sum_{k=0}^{n-1} |f'(t_k^*)|(t_{k+1} - t_k),$$

and

$$\begin{aligned} FV_{[0,T]}(f) &= \lim_{||\Pi||\to 0} \sum_{k=0}^{n-1} |f'(t_k^*)|(t_{k+1} - t_k) \\ &= \int_0^T |f'(t)|\, dt. \end{aligned}$$

## 14.3 Quadratic Variation

**Definition 14.2 (Quadratic Variation)** The *quadratic variation* of a function $f$ on an interval $[0,T]$ is

$$\langle f\rangle(T) = \lim_{||\Pi||\to 0} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|^2.$$

**Remark 14.1 (Quadratic Variation of Differentiable Functions)** If $f$ is differentiable, then $\langle f\rangle(T) = 0$, because

$$\begin{aligned} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|^2 &= \sum_{k=0}^{n-1} |f'(t_k^*)|^2 (t_{k+1} - t_k)^2 \\ &\le ||\Pi||. \sum_{k=0}^{n-1} |f'(t_k^*)|^2 (t_{k+1} - t_k) \end{aligned}$$

and

$$\begin{aligned} \langle f\rangle(T) &\le \lim_{||\Pi||\to 0} ||\Pi||. \lim_{||\Pi||\to 0} \sum_{k=0}^{n-1} |f'(t_k^*)|^2 (t_{k+1} - t_k) \\ &= \lim_{||\Pi||\to 0} ||\Pi|| \int_0^T |f'(t)|^2\, dt \\ &= 0. \end{aligned}$$

**Theorem 3.44**

$$\langle B\rangle(T) = T,$$

*or more precisely,*

$$I\!P\{\omega \in \Omega; \langle B(.,\omega)\rangle(T) = T\} = 1.$$

*In particular, the paths of Brownian motion are not differentiable.*

**Proof:** (Outline) Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, T]$. To simplify notation, set $D_k = B(t_{k+1}) - B(t_k)$. Define the *sample quadratic variation*

$$Q_\Pi = \sum_{k=0}^{n-1} D_k^2.$$

Then

$$Q_\Pi - T = \sum_{k=0}^{n-1} [D_k^2 - (t_{k+1} - t_k)].$$

We want to show that

$$\lim_{||\Pi|| \to 0} (Q_\Pi - T) = 0.$$

Consider an individual summand

$$D_k^2 - (t_{k+1} - t_k) = [B(t_{k+1}) - B(t_k)]^2 - (t_{k+1} - t_k).$$

This has expectation 0, so

$$\mathbb{E}(Q_\Pi - T) = \mathbb{E} \sum_{k=0}^{n-1} [D_k^2 - (t_{k+1} - t_k)] = 0.$$

For $j \neq k$, the terms

$$D_j^2 - (t_{j+1} - t_j) \quad \text{and} \quad D_k^2 - (t_{k+1} - t_k)$$

are independent, so

$$\begin{aligned}
\operatorname{var}(Q_\Pi - T) &= \sum_{k=0}^{n-1} \operatorname{var}[D_k^2 - (t_{k+1} - t_k)] \\
&= \sum_{k=0}^{n-1} \mathbb{E}[D_k^4 - 2(t_{k+1} - t_k) D_k^2 + (t_{k+1} - t_k)^2] \\
&= \sum_{k=0}^{n-1} [3(t_{k+1} - t_k)^2 - 2(t_{k+1} - t_k)^2 + (t_{k+1} - t_k)^2] \\
&\qquad \text{(if } X \text{ is normal with mean 0 and variance } \sigma^2\text{, then } \mathbb{E}(X^4) = 3\sigma^4) \\
&= 2 \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 \\
&\leq 2||\Pi|| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \\
&= 2||\Pi||\, T.
\end{aligned}$$

Thus we have

$$\begin{aligned}
\mathbb{E}(Q_\Pi - T) &= 0, \\
\operatorname{var}(Q_\Pi - T) &\leq 2||\Pi||.T.
\end{aligned}$$

As $||\Pi||\to 0$, $\mathrm{var}(Q_\Pi - T)\to 0$, so

$$\lim_{||\Pi||\to 0}(Q_\Pi - T) = 0.$$

■

**Remark 14.2 (Differential Representation)** We know that

$$I\!\!E[(B(t_{k+1}) - B(t_k))^2 - (t_{k+1} - t_k)] = 0.$$

We showed above that

$$\mathrm{var}[(B(t_{k+1}) - B(t_k))^2 - (t_{k+1} - t_k)] = 2(t_{k+1} - t_k)^2.$$

When $(t_{k+1} - t_k)$ is small, $(t_{k+1} - t_k)^2$ is *very* small, and we have the approximate equation

$$(B(t_{k+1}) - B(t_k))^2 \simeq t_{k+1} - t_k,$$

which we can write informally as

$$dB(t)\ dB(t) = dt.$$

## 14.4 Quadratic Variation as Absolute Volatility

On any time interval $[T_1, T_2]$, we can sample the Brownian motion at times

$$T_1 = t_0 \le t_1 \le \ldots \le t_n = T_2$$

and compute the *squared sample absolute volatility*

$$\frac{1}{T_2 - T_1}\sum_{k=0}^{n-1}(B(t_{k+1}) - B(t_k))^2.$$

This is approximately equal to

$$\frac{1}{T_2 - T_1}[\langle B\rangle(T_2) - \langle B\rangle(T_1)] = \frac{T_2 - T_1}{T_2 - T_1} = 1.$$

As we increase the number of sample points, this approximation becomes exact. In other words, Brownian motion has *absolute volatility 1*.

Furthermore, consider the equation

$$\langle B\rangle(T) = T = \int_0^T 1\ dt, \qquad \forall T \ge 0.$$

This says that quadratic variation for Brownian motion accumulates at rate 1 *at all times along almost every path*.

## 14.5 Construction of the Itô Integral

The **integrator** is Brownian motion $B(t), t \geq 0$, with associated filtration $\mathcal{F}(t), t \geq 0$, and the following properties:

1. $s \leq t \Longrightarrow$ every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$,
2. $B(t)$ is $\mathcal{F}(t)$-measurable, $\forall t$,
3. For $t \leq t_1 \leq \ldots \leq t_n$, the increments $B(t_1) - B(t), B(t_2) - B(t_1), \ldots, B(t_n) - B(t_{n-1})$ are independent of $\mathcal{F}(t)$.

The **integrand** is $\delta(t), t \geq 0$, where

1. $\delta(t)$ is $\mathcal{F}(t)$-measurable $\forall t$ (i.e., $\delta$ is adapted)
2. $\delta$ is square-integrable:
$$\mathbb{E} \int_0^T \delta^2(t)\, dt < \infty, \qquad \forall T.$$

We want to define the **Itô Integral:**

$$I(t) = \int_0^t \delta(u)\, dB(u), \qquad t \geq 0.$$

**Remark 14.3 (Integral w.r.t. a differentiable function)** If $f(t)$ is a differentiable function, then we can define

$$\int_0^t \delta(u)\, df(u) = \int_0^t \delta(u) f'(u)\, du.$$

This won't work when the integrator is Brownian motion, because the paths of Brownian motion are not differentiable.

## 14.6 Itô integral of an elementary integrand

Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, T]$, i.e.,

$$0 = t_0 \leq t_1 \leq \ldots \leq t_n = T.$$

Assume that $\delta(t)$ is constant on each subinterval $[t_k, t_{k+1}]$ (see Fig. 14.2). We call such a $\delta$ an *elementary process*.

The functions $B(t)$ and $\delta(t_k)$ can be interpreted as follows:

- Think of $B(t)$ as the *price per unit share* of an asset at time $t$.

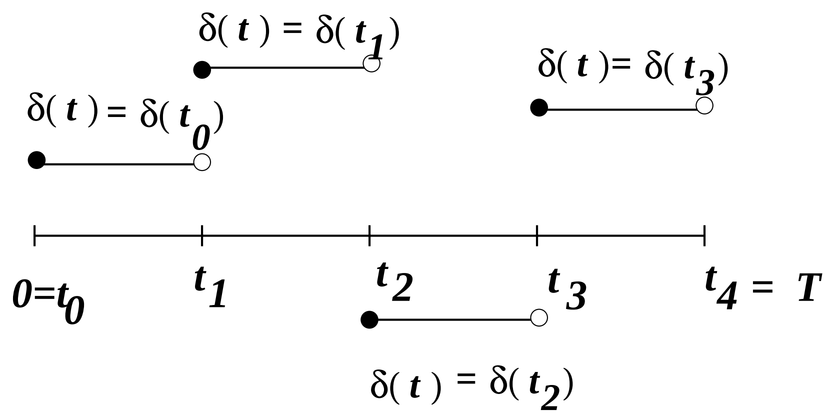



Figure 14.2: *An elementary function* $\delta$.

- Think of $t_0, t_1, \ldots, t_n$ as the *trading dates* for the asset.
- Think of $\delta(t_k)$ as the *number of shares of the asset acquired* at trading date $t_k$ and held until trading date $t_{k+1}$.

Then the Itô integral $I(t)$ can be interpreted as the *gain from trading* at time $t$; this gain is given by:

$$I(t) = \begin{cases} \delta(t_0)[B(t) - \underbrace{B(t_0)}_{=B(0)=0}], & 0 \leq t \leq t_1 \\ \delta(t_0)[B(t_1) - B(t_0)] + \delta(t_1)[B(t) - B(t_1)], & t_1 \leq t \leq t_2 \\ \delta(t_0)[B(t_1) - B(t_0)] + \delta(t_1)[B(t_2) - B(t_1)] + \delta(t_2)[B(t) - B(t_2)], & t_2 \leq t \leq t_3. \end{cases}$$

In general, if $t_k \leq t \leq t_{k+1}$,

$$I(t) = \sum_{j=0}^{k-1} \delta(t_j)[B(t_{j+1}) - B(t_j)] + \delta(t_k)[B(t) - B(t_k)].$$

## 14.7 Properties of the Itô integral of an elementary process

**Adaptedness** For each $t$, $I(t)$ is $\mathcal{F}(t)$-measurable.

**Linearity** If

$$I(t) = \int_0^t \delta(u)\, dB(u), \qquad J(t) = \int_0^t \gamma(u)\, dB(u)$$

then

$$I(t) \pm J(t) = \int_0^t (\delta(u) \pm \gamma(u))\, dB(u)$$

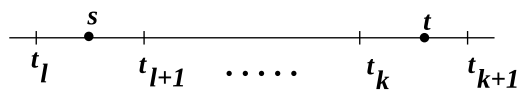



Figure 14.3: *Showing $s$ and $t$ in different partitions.*

and

$$cI(t) = \int_0^t c\delta(u)dB(u).$$

**Martingale** $I(t)$ is a martingale.

We prove the martingale property for the elementary process case.

**Theorem 7.45 (Martingale Property)**

$$I(t) = \sum_{j=0}^{k-1} \delta(t_j)[B(t_{j+1}) - B(t_j)] \; + \; \delta(t_k)[B(t) - B(t_k)], \qquad t_k \leq t \leq t_{k+1}$$

*is a martingale.*

**Proof:** Let $0 \leq s \leq t$ be given. We treat the more difficult case that $s$ and $t$ are in different subintervals, i.e., there are partition points $t_\ell$ and $t_k$ such that $s \in [t_\ell, t_{\ell+1}]$ and $t \in [t_k, t_{k+1}]$ (See Fig. 14.3).

Write

$$\begin{aligned} I(t) = \sum_{j=0}^{\ell-1} \delta(t_j)[B(t_{j+1}) - B(t_j)] \; + \; \delta(t_\ell)[B(t_{\ell+1}) - B(t_\ell)] \\ + \sum_{j=\ell+1}^{k-1} \delta(t_j)[B(t_{j+1}) - B(t_j)] \; + \; \delta(t_k)[B(t) - B(t_k)] \end{aligned}$$

We compute conditional expectations:

$$\begin{aligned} \mathbb{E}\left[\sum_{j=0}^{\ell-1} \delta(t_j)(B(t_{j+1}) - B(t_j)) \middle| \mathcal{F}(s)\right] &= \sum_{j=0}^{\ell-1} \delta(t_j)(B(t_{j+1}) - B(t_j)). \\ \mathbb{E}\left[\delta(t_\ell)(B(t_{\ell+1}) - B(t_\ell)) \middle| \mathcal{F}(s)\right] &= \delta(t_\ell)\left(\mathbb{E}[B(t_{\ell+1})|\mathcal{F}(s)] - B(t_\ell)\right) \\ &= \delta(t_\ell)[B(s) - B(t_\ell)] \end{aligned}$$

These first two terms add up to $I(s)$. We show that the third and fourth terms are zero.

$$
\begin{aligned}
\mathbb{E}\left[\sum_{j=\ell+1}^{k-1}\delta(t_j)(B(t_{j+1})-B(t_j))\Big|\mathcal{F}(s)\right] &= \sum_{j=\ell+1}^{k-1}\mathbb{E}\left[\mathbb{E}\left[\delta(t_j)(B(t_{j+1})-B(t_j))\Big|\mathcal{F}(t_j)\right]\Big|\mathcal{F}(s)\right]\\
&= \sum_{j=\ell+1}^{k-1}\mathbb{E}\left[\delta(t_j)\underbrace{(\mathbb{E}[B(t_{j+1})|\mathcal{F}(t_j)]-B(t_j))}_{=0}\Big|\mathcal{F}(s)\right]
\end{aligned}
$$

$$
\mathbb{E}\left[\delta(t_k)(B(t)-B(t_k))\Big|\mathcal{F}(s)\right] = \mathbb{E}\left[\delta(t_k)\underbrace{(\mathbb{E}[B(t)|\mathcal{F}(t_k)]-B(t_k))}_{=0}\Big|\mathcal{F}(s)\right]
$$

■

**Theorem 7.46 (Itô Isometry)**

$$
\mathbb{E}I^2(t) = \mathbb{E}\int_0^t \delta^2(u)\,du.
$$

**Proof:** To simplify notation, assume $t = t_k$, so

$$
I(t) = \sum_{j=0}^{k}\delta(t_j)[\underbrace{B(t_{j+1})-B(t_j)}_{D_j}]
$$

Each $D_j$ has expectation 0, and different $D_j$ are independent.

$$
\begin{aligned}
I^2(t) &= \left(\sum_{j=0}^{k}\delta(t_j)D_j\right)^2\\
&= \sum_{j=0}^{k}\delta^2(t_j)D_j^2 + 2\sum_{i<j}\delta(t_i)\delta(t_j)D_iD_j.
\end{aligned}
$$

Since the cross terms have expectation zero,

$$
\begin{aligned}
\mathbb{E}I^2(t) &= \sum_{j=0}^{k}\mathbb{E}[\delta^2(t_j)D_j^2]\\
&= \sum_{j=0}^{k}\mathbb{E}\left[\delta^2(t_j)\mathbb{E}\left[(B(t_{j+1})-B(t_j))^2\Big|\mathcal{F}(t_j)\right]\right]\\
&= \sum_{j=0}^{k}\mathbb{E}\delta^2(t_j)(t_{j+1}-t_j)\\
&= \mathbb{E}\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\delta^2(u)\,du\\
&= \mathbb{E}\int_0^t \delta^2(u)\,du
\end{aligned}
$$

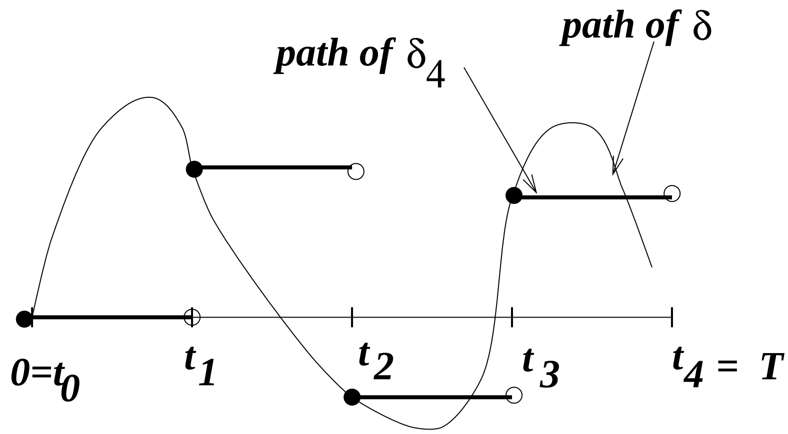



Figure 14.4: *Approximating a general process by an elementary process $\delta_4$, over* $[0,T]$.

■

## 14.8 Itô integral of a general integrand

Fix $T > 0$. Let $\delta$ be a process (not necessarily an elementary process) such that

- $\delta(t)$ is $\mathcal{F}(t)$-measurable, $\forall t \in [0,T]$,
- $I\!\!E \int_0^T \delta^2(t)\ dt < \infty$.

**Theorem 8.47** *There is a sequence of elementary processes* $\{\delta_n\}_{n=1}^{\infty}$ *such that*

$$\lim_{n \longrightarrow \infty} I\!\!E \int_0^T |\delta_n(t) - \delta(t)|^2\ dt = 0.$$

**Proof:** Fig. 14.4 shows the main idea.

■

In the last section we have defined

$$I_n(T) = \int_0^T \delta_n(t)\ dB(t)$$

for every $n$. We now define

$$\int_0^T \delta(t)\ dB(t) = \lim_{n \longrightarrow \infty} \int_0^T \delta_n(t)\ dB(t).$$

The only difficulty with this approach is that we need to make sure the above limit exists. Suppose $n$ and $m$ are large positive integers. Then

$$\begin{aligned}
\operatorname{var}(I_n(T) - I_m(T)) &= \mathbb{E}\left(\int_0^T [\delta_n(t) - \delta_m(t)]\, dB(t)\right)^2 \\
\text{(Itô Isometry:)} &= \mathbb{E}\int_0^T [\delta_n(t) - \delta_m(t)]^2\, dt \\
&= \mathbb{E}\int_0^T [\, |\delta_n(t) - \delta(t)| + |\delta(t) - \delta_m(t)|\, ]^2\, dt \\
((a+b)^2 \le 2a^2 + 2b^2 :) &\le 2\mathbb{E}\int_0^T |\delta_n(t) - \delta(t)|^2\, dt + 2\mathbb{E}\int_0^T |\delta_m(t) - \delta(t)|^2\, dt,
\end{aligned}$$

which is small. This guarantees that the sequence $\{I_n(T)\}_{n=1}^{\infty}$ has a limit.

## 14.9 Properties of the (general) Itô integral

$$I(t) = \int_0^t \delta(u)\, dB(u).$$

Here $\delta$ is any adapted, square-integrable process.

**Adaptedness.** For each $t$, $I(t)$ is $\mathcal{F}(t)$-measurable.

**Linearity.** If

$$I(t) = \int_0^t \delta(u)\, dB(u), \qquad J(t) = \int_0^t \gamma(u)\, dB(u)$$

then

$$I(t) \pm J(t) = \int_0^t (\delta(u) \pm \gamma(u))\, dB(u)$$

and

$$cI(t) = \int_0^t c\delta(u) dB(u).$$

**Martingale.** $I(t)$ is a martingale.

**Continuity.** $I(t)$ is a continuous function of the upper limit of integration $t$.

**Itô Isometry.** $\mathbb{E} I^2(t) = \mathbb{E}\int_0^t \delta^2(u)\, du$.

**Example 14.1 ()** Consider the Itô integral

$$\int_0^T B(u)\, dB(u).$$

We approximate the integrand as shown in Fig. 14.5

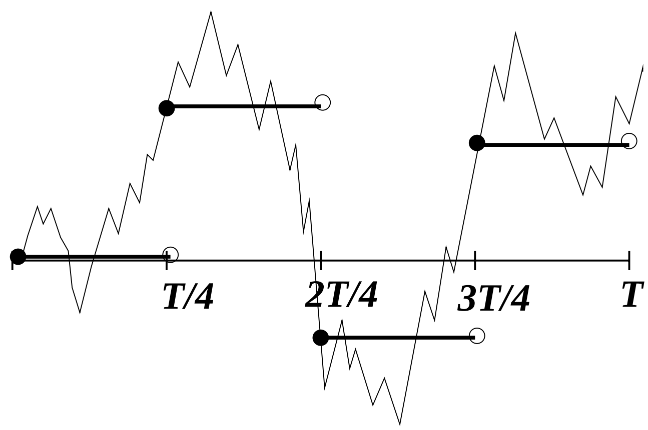



Figure 14.5: *Approximating the integrand* $B(u)$ *with* $\delta_4$, *over* $[0,T]$.

$$\delta_n(u) = \begin{cases} B(0)=0 & \text{if} \quad 0 \le u < T/n; \\ B(T/n) & \text{if} \quad T/n \le u < 2T/n; \\ \ldots & \\ B\left(\frac{(n-1)T}{T}\right) & \text{if} \quad \frac{(n-1)T}{n} \le u < T. \end{cases}$$

By definition,

$$\int_0^T B(u)\, dB(u) = \lim_{n\to\infty} \sum_{k=0}^{n-1} B\left(\frac{kT}{n}\right)\left[B\left(\frac{(k+1)T}{n}\right) - B\left(\frac{kT}{n}\right)\right].$$

To simplify notation, we denote

$$B_k \triangleq B\left(\frac{kT}{n}\right),$$

so

$$\int_0^T B(u)\, dB(u) = \lim_{n\to\infty} \sum_{k=0}^{n-1} B_k(B_{k+1} - B_k).$$

We compute

$$\begin{aligned} \tfrac{1}{2}\sum_{k=0}^{n-1}(B_{k+1}-B_k)^2 &= \tfrac{1}{2}\sum_{k=0}^{n-1} B_{k+1}^2 - \sum_{k=0}^{n-1} B_k B_{k+1} + \tfrac{1}{2}\sum_{k=0}^{n-1} B_k^2 \\ &= \tfrac{1}{2}B_n^2 + \tfrac{1}{2}\sum_{j=0}^{n-1} B_j^2 - \sum_{k=0}^{n-1} B_k B_{k+1} + \tfrac{1}{2}\sum_{k=0}^{n-1} B_k^2 \\ &= \tfrac{1}{2}B_n^2 + \sum_{k=0}^{n-1} B_k^2 - \sum_{k=0}^{n-1} B_k B_{k+1} \\ &= \tfrac{1}{2}B_n^2 - \sum_{k=0}^{n-1} B_k(B_{k+1} - B_k). \end{aligned}$$

Therefore,

$$\sum_{k=0}^{n-1} B_k(B_{k+1} - B_k) = \tfrac{1}{2}B_n^2 - \tfrac{1}{2}\sum_{k=0}^{n-1}(B_{k+1} - B_k)^2,$$

or equivalently

$$\sum_{k=0}^{n-1} B\left(\frac{kT}{n}\right)\left[B\left(\frac{(k+1)T}{n}\right) - B\left(\frac{kT}{n}\right)\right] = \tfrac{1}{2}B^2(T) - \tfrac{1}{2}\sum_{k=0}^{n-1}\left[B\left(\frac{(k+1)T}{n}\right)\left(\frac{k}{T}\right)\right]^2.$$

Let $n \to \infty$ and use the definition of quadratic variation to get

$$\int_0^T B(u)\ dB(u) = \tfrac{1}{2}B^2(T) - \tfrac{1}{2}T.$$

∎

**Remark 14.4 (Reason for the $\frac{1}{2}T$ term)** If $f$ is differentiable with $f(0) = 0$, then

$$\begin{aligned}\int_0^T f(u)\ df(u) &= \int_0^T f(u)f'(u)\ du \\ &= \tfrac{1}{2}f^2(u)\Big|_0^T \\ &= \tfrac{1}{2}f^2(T).\end{aligned}$$

In contrast, for Brownian motion, we have

$$\int_0^T B(u)dB(u) = \tfrac{1}{2}B^2(T) - \tfrac{1}{2}T.$$

The extra term $\frac{1}{2}T$ comes from the nonzero quadratic variation of Brownian motion. It has to be there, because

$$I\!\!E\int_0^T B(u)\ dB(u) = 0 \qquad \text{(Itô integral is a martingale)}$$

but

$$I\!\!E\tfrac{1}{2}B^2(T) = \tfrac{1}{2}T.$$

## 14.10 Quadratic variation of an Itô integral

**Theorem 10.48 (Quadratic variation of Itô integral)** *Let*

$$I(t) = \int_0^t \delta(u)\ dB(u).$$

*Then*

$$\langle I\rangle(t) = \int_0^t \delta^2(u)\ du.$$

This holds even if $\delta$ is not an elementary process. The quadratic variation formula says that at each time $u$, the *instantaneous absolute volatility* of $I$ is $\delta^2(u)$. This is the absolute volatility of the Brownian motion scaled by the size of the position (i.e. $\delta(t)$) in the Brownian motion. Informally, we can write the quadratic variation formula in differential form as follows:

$$dI(t)\ dI(t) = \delta^2(t)\ dt.$$

Compare this with

$$dB(t)\ dB(t) = dt.$$

**Proof:** (For an elementary process $\delta$). Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be the partition for $\delta$, i.e., $\delta(t) = \delta(t_k)$ for $t_k \le t \le t_{k+1}$. To simplify notation, assume $t = t_n$. We have

$$\langle I\rangle(t) = \sum_{k=0}^{n-1} \left[\langle I\rangle(t_{k+1}) - \langle I\rangle(t_k)\right].$$

Let us compute $\langle I\rangle(t_{k+1}) - \langle I\rangle(t_k)$. Let $\Xi = \{s_0, s_1, \ldots, s_m\}$ be a partition

$$t_k = s_0 \le s_1 \le \ldots \le s_m = t_{k+1}.$$

Then

$$\begin{aligned} I(s_{j+1}) - I(s_j) &= \int_{s_j}^{s_{j+1}} \delta(t_k)\ dB(u) \\ &= \delta(t_k)\left[B(s_{j+1}) - B(s_j)\right], \end{aligned}$$

so

$$\begin{aligned} \langle I\rangle(t_{k+1}) - \langle I\rangle(t_k) &= \sum_{j=0}^{m-1} \left[I(s_{j+1}) - I(s_j)\right]^2 \\ &= \delta^2(t_k) \sum_{j=0}^{m-1} \left[B(s_{j+1}) - B(s_j)\right]^2 \\ &\xrightarrow{||\Xi||\to 0} \delta^2(t_k)(t_{k+1} - t_k). \end{aligned}$$

It follows that

$$\begin{aligned} \langle I\rangle(t) &= \sum_{k=0}^{n-1} \delta^2(t_k)(t_{k+1} - t_k) \\ &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \delta^2(u)\ du \\ &\xrightarrow{||\Pi||\to 0} \int_0^t \delta^2(u)\ du. \end{aligned}$$

■

# Chapter 15

# Itô's Formula

## 15.1 Itô's formula for one Brownian motion

We want a rule to "differentiate" expressions of the form $f(B(t))$, where $f(x)$ is a differentiable function. If $B(t)$ were also differentiable, then the ordinary *chain rule* would give

$$\frac{d}{dt}f(B(t)) = f'(B(t))B'(t),$$

which could be written in differential notation as

$$\begin{aligned} df(B(t)) &= f'(B(t))B'(t)\ dt \\ &= f'(B(t))dB(t) \end{aligned}$$

However, $B(t)$ is not differentiable, and in particular has nonzero quadratic variation, so the correct formula has an extra term, namely,

$$df(B(t)) = f'(B(t))\ dB(t) + \tfrac{1}{2}f''(B(t))\ \underbrace{dt}_{dB(t)\ dB(t)}.$$

This is *Itô's formula in differential form.* Integrating this, we obtain *Itô's formula in integral form:*

$$f(B(t)) - \underbrace{f(B(0))}_{f(0)} = \int_0^t f'(B(u))\ dB(u) + \tfrac{1}{2}\int_0^t f''(B(u))\ du.$$

**Remark 15.1 (Differential vs. Integral Forms)** The mathematically meaningful form of Itô's formula is Itô's formula in integral form:

$$f(B(t)) - f(B(0)) = \int_0^t f'(B(u))\ dB(u) + \tfrac{1}{2}\int_0^t f''(B(u))\ du.$$

This is because we have solid definitions for both integrals appearing on the right-hand side. The first,

$$\int_0^t f'(B(u))\ dB(u)$$

is an *Itô integral*, defined in the previous chapter. The second,

$$\int_0^t f''(B(u))\ du,$$

is a *Riemann integral*, the type used in freshman calculus.

For paper and pencil computations, the more convenient form of Itô's rule is *Itô's formula in differential form:*

$$df(B(t)) = f'(B(t))\ dB(t) + \tfrac{1}{2}f''(B(t))\ dt.$$

There is an intuitive meaning but no solid definition for the terms $df(B(t))$, $dB(t)$ and $dt$ appearing in this formula. This formula becomes mathematically respectable only after we integrate it.

## 15.2 Derivation of Itô's formula

Consider $f(x) = \frac{1}{2}x^2$, so that

$$f'(x) = x, \quad f''(x) = 1.$$

Let $x_k, x_{k+1}$ be numbers. Taylor's formula implies

$$f(x_{k+1}) - f(x_k) = (x_{k+1} - x_k)f'(x_k) + \tfrac{1}{2}(x_{k+1} - x_k)^2 f''(x_k).$$

In this case, Taylor's formula to second order is *exact* because $f$ is a *quadratic function*.

In the general case, the above equation is only approximate, and the error is of the order of $(x_{k+1} - x_k)^3$. The total error will have limit zero in the last step of the following argument.

Fix $T > 0$ and let $\Pi = \{t_0, t_1, \dots, t_n\}$ be a partition of $[0, T]$. Using Taylor's formula, we write:

$$\begin{aligned}
&f(B(T)) - f(B(0)) \\
&= \tfrac{1}{2}B^2(T) - \tfrac{1}{2}B^2(0) \\
&= \sum_{k=0}^{n-1} \left[f(B(t_{k+1})) - f(B(t_k))\right] \\
&= \sum_{k=0}^{n-1} \left[B(t_{k+1}) - B(t_k)\right] f'(B(t_k)) + \tfrac{1}{2}\sum_{k=0}^{n-1} \left[B(t_{k+1}) - B(t_k)\right]^2 f''(B(t_k)) \\
&= \sum_{k=0}^{n-1} B(t_k) \left[B(t_{k+1}) - B(t_k)\right] + \tfrac{1}{2}\sum_{k=0}^{n-1} \left[B(t_{k+1}) - B(t_k)\right]^2 .
\end{aligned}$$

We let $||\Pi||\to 0$ to obtain

$$\begin{aligned} f(B(T)) - f(B(0)) &= \int_0^T B(u)\ dB(u) + \tfrac{1}{2}\underbrace{\langle B\rangle(T)}_{T} \\ &= \int_0^T f'(B(u))\ dB(u) + \tfrac{1}{2}\int_0^T \underbrace{f''(B(u))}_{1}\, du. \end{aligned}$$

This is Itô's formula in integral form for the special case

$$f(x) = \tfrac{1}{2}x^2.$$

## 15.3 Geometric Brownian motion

**Definition 15.1 (Geometric Brownian Motion)** Geometric Brownian motion is

$$S(t) = S(0)\exp\left\{\sigma B(t) + \left(\mu - \tfrac{1}{2}\sigma^2\right)t\right\},$$

where $\mu$ and $\sigma > 0$ are constant.

Define

$$f(t,x) = S(0)\exp\left\{\sigma x + \left(\mu - \tfrac{1}{2}\sigma^2\right)t\right\},$$

so

$$S(t) = f(t, B(t)).$$

Then

$$f_t = \left(\mu - \tfrac{1}{2}\sigma^2\right) f,\ \ f_x = \sigma f,\ \ f_{xx} = \sigma^2 f.$$

According to Itô's formula,

$$\begin{aligned} dS(t) &= df(t, B(t)) \\ &= f_t dt + f_x dB + \tfrac{1}{2} f_{xx} \underbrace{dBdB}_{dt} \\ &= (\mu - \tfrac{1}{2}\sigma^2) f\ dt + \sigma f\ dB + \tfrac{1}{2}\sigma^2 f\ dt \\ &= \mu S(t) dt + \sigma S(t)\ dB(t) \end{aligned}$$

Thus, *Geometric Brownian motion in differential form* is

$$dS(t) = \mu S(t)dt + \sigma S(t)\ dB(t),$$

and *Geometric Brownian motion in integral form* is

$$S(t) = S(0) + \int_0^t \mu S(u)\ du + \int_0^t \sigma S(u)\ dB(u).$$

## 15.4 Quadratic variation of geometric Brownian motion

In the integral form of Geometric Brownian motion,

$$S(t) = S(0) + \int_0^t \mu S(u)\ du + \int_0^t \sigma S(u)\ dB(u),$$

the Riemann integral

$$F(t) = \int_0^t \mu S(u)\ du$$

is differentiable with $F'(t) = \mu S(t)$. This term has zero quadratic variation. The Itô integral

$$G(t) = \int_0^t \sigma S(u)\ dB(u)$$

is not differentiable. It has quadratic variation

$$\langle G \rangle(t) = \int_0^t \sigma^2 S^2(u)\ du.$$

Thus the quadratic variation of $S$ is given by the quadratic variation of $G$. In differential notation, we write

$$dS(t)\ dS(t) = (\mu S(t)dt + \sigma S(t)dB(t))^2 = \sigma^2 S^2(t)\ dt$$

## 15.5 Volatility of Geometric Brownian motion

Fix $0 \le T_1 \le T_2$. Let $\Pi = \{t_0, \ldots, t_n\}$ be a partition of $[T_1, T_2]$. The *squared absolute sample volatility* of $S$ on $[T_1, T_2]$ is

$$\begin{aligned} \frac{1}{T_2 - T_1} \sum_{k=0}^{n-1} [S(t_{k+1}) - S(t_k)]^2 &\simeq \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \sigma^2 S^2(u)\ du \\ &\simeq \sigma^2 S^2(T_1) \end{aligned}$$

As $T_2 \downarrow T_1$, the above approximation becomes exact. In other words, the *instantaneous relative volatility* of $S$ is $\sigma^2$. This is usually called simply the *volatility* of $S$.

## 15.6 First derivation of the Black-Scholes formula

**Wealth of an investor.** An investor begins with nonrandom initial wealth $X_0$ and at each time $t$, holds $\Delta(t)$ shares of stock. Stock is modelled by a geometric Brownian motion:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t).$$

$\Delta(t)$ can be random, but must be adapted. The investor finances his investing by borrowing or lending at interest rate $r$.

Let $X(t)$ denote the wealth of the investor at time $t$. Then

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + r\left[X(t) - \Delta(t)S(t)\right]dt \\ &= \Delta(t)\left[\mu S(t)dt + \sigma S(t)dB(t)\right] + r\left[X(t) - \Delta(t)S(t)\right]dt \\ &= rX(t)dt \;+\; \Delta(t)S(t)\;\underbrace{(\mu - r)}_{\text{Risk premium}}\;dt \;+\; \Delta(t)S(t)\sigma dB(t). \end{aligned}$$

**Value of an option.** Consider an European option which pays $g(S(T))$ at time $T$. Let $v(t,x)$ denote the value of this option at time $t$ if the stock price is $S(t) = x$. In other words, the value of the option at each time $t \in [0,T]$ is

$$v(t, S(t)).$$

The differential of this value is

$$\begin{aligned} dv(t,S(t)) &= v_t dt + v_x dS + \tfrac{1}{2}v_{xx} dS\; dS \\ &= v_t dt + v_x\left[\mu S\; dt + \sigma S\; dB\right] + \tfrac{1}{2}v_{xx}\sigma^2 S^2\; dt \\ &= \left[v_t + \mu S v_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx}\right]\; dt + \sigma S v_x dB \end{aligned}$$

A hedging portfolio starts with some initial wealth $X_0$ and invests so that the wealth $X(t)$ at each time tracks $v(t,S(t))$. We saw above that

$$dX(t) = \left[rX + \Delta(\mu - r)S\right]\; dt + \sigma S\Delta dB.$$

To ensure that $X(t) = v(t,S(t))$ for all $t$, we equate coefficients in their differentials. Equating the $dB$ coefficients, we obtain the $\Delta$-*hedging rule*:

$$\Delta(t) = v_x(t, S(t)).$$

Equating the $dt$ coefficients, we obtain:

$$v_t + \mu S v_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx} = rX + \Delta(\mu - r)S.$$

But we have set $\Delta = v_x$, and we are seeking to cause $X$ to agree with $v$. Making these substitutions, we obtain

$$v_t + \mu S v_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx} = rv + v_x(\mu - r)S,$$

(where $v = v(t,S(t))$ and $S = S(t)$) which simplifies to

$$v_t + rSv_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx} = rv.$$

In conclusion, we should let $v$ be the solution to the *Black-Scholes partial differential equation*

$$v_t(t,x) + rxv_x(t,x) + \tfrac{1}{2}\sigma^2 x^2 v_{xx}(t,x) = rv(t,x)$$

satisfying the terminal condition

$$v(T,x) = g(x).$$

If an investor starts with $X_0 = v(0,S(0))$ and uses the hedge $\Delta(t) = v_x(t,S(t))$, then he will have $X(t) = v(t,S(t))$ for all $t$, and in particular, $X(T) = g(S(T))$.

## 15.7 Mean and variance of the Cox-Ingersoll-Ross process

The *Cox-Ingersoll-Ross* model for interest rates is

$$dr(t) = a(b - cr(t))dt + \sigma\sqrt{r(t)}\, dB(t),$$

where $a, b, c, \sigma$ and $r(0)$ are positive constants. In integral form, this equation is

$$r(t) = r(0) + a\int_0^t (b - cr(u))\, du + \sigma\int_0^t \sqrt{r(u)}\, dB(u).$$

We apply Itô's formula to compute $dr^2(t)$. This is $df(r(t))$, where $f(x) = x^2$. We obtain

$$\begin{aligned}
dr^2(t) &= df(r(t)) \\
&= f'(r(t))\, dr(t) + \tfrac{1}{2}f''(r(t))\, dr(t)\, dr(t) \\
&= 2r(t)\left[a(b - cr(t))\, dt + \sigma\sqrt{r(t)}\, dB(t)\right] + \left[a(b - cr(t))\, dt + \sigma\sqrt{r(t)}\, dB(t)\right]^2 \\
&= 2abr(t)\, dt - 2acr^2(t)\, dt + 2\sigma r^{\frac{3}{2}}(t)\, dB(t) + \sigma^2 r(t)\, dt \\
&= (2ab + \sigma^2)r(t)\, dt - 2acr^2(t)\, dt + 2\sigma r^{\frac{3}{2}}(t)\, dB(t)
\end{aligned}$$

**The mean of $r(t)$.** The integral form of the CIR equation is

$$r(t) = r(0) + a\int_0^t (b - cr(u))\, du + \sigma\int_0^t \sqrt{r(u)}\, dB(u).$$

Taking expectations and remembering that the expectation of an Itô integral is zero, we obtain

$$\mathbb{E}r(t) = r(0) + a\int_0^t (b - c\mathbb{E}r(u))\, du.$$

Differentiation yields

$$\frac{d}{dt}\mathbb{E}r(t) = a(b - c\mathbb{E}r(t)) = ab - ac\mathbb{E}r(t),$$

which implies that

$$\frac{d}{dt}\left[e^{act}\mathbb{E}r(t)\right] = e^{act}\left[ac\mathbb{E}r(t) + \frac{d}{dt}\mathbb{E}r(t)\right] = e^{act}ab.$$

Integration yields

$$e^{act}\mathbb{E}r(t) - r(0) = ab\int_0^t e^{acu}\, du = \frac{b}{c}(e^{act} - 1).$$

We solve for $\mathbb{E}r(t)$:

$$\mathbb{E}r(t) = \frac{b}{c} + e^{-act}\left(r(0) - \frac{b}{c}\right).$$

If $r(0) = \frac{b}{c}$, then $\mathbb{E}r(t) = \frac{b}{c}$ for every $t$. If $r(0) \neq \frac{b}{c}$, then $r(t)$ exhibits *mean reversion*:

$$\lim_{t\to\infty} \mathbb{E}r(t) = \frac{b}{c}.$$

**Variance of** $r(t)$. The integral form of the equation derived earlier for $dr^2(t)$ is

$$r^2(t) = r^2(0) + (2ab + \sigma^2)\int_0^t r(u)\,du - 2ac\int_0^t r^2(u)\,du + 2\sigma\int_0^t r^{\frac{3}{2}}(u)\,dB(u).$$

Taking expectations, we obtain

$$\mathbb{E}r^2(t) = r^2(0) + (2ab + \sigma^2)\int_0^t \mathbb{E}r(u)\,du - 2ac\int_0^t \mathbb{E}r^2(u)\,du.$$

Differentiation yields

$$\frac{d}{dt}\mathbb{E}r^2(t) = (2ab + \sigma^2)\mathbb{E}r(t) - 2ac\mathbb{E}r^2(t),$$

which implies that

$$\begin{aligned}\frac{d}{dt}e^{2act}\mathbb{E}r^2(t) &= e^{2act}\left[2ac\mathbb{E}r^2(t) + \frac{d}{dt}\mathbb{E}r^2(t)\right]\\ &= e^{2act}(2ab + \sigma^2)\mathbb{E}r(t).\end{aligned}$$

Using the formula already derived for $\mathbb{E}r(t)$ and integrating the last equation, after considerable algebra we obtain

$$\begin{aligned}\mathbb{E}r^2(t) &= \frac{b\sigma^2}{2ac^2} + \frac{b^2}{c^2} + \left(r(0) - \frac{b}{c}\right)\left(\frac{\sigma^2}{ac} + \frac{2b}{c}\right)e^{-act}\\ &\qquad + \left(r(0) - \frac{b}{c}\right)^2\frac{\sigma^2}{ac}e^{-2act} + \frac{\sigma^2}{ac}\left(\frac{b}{2c} - r(0)\right)e^{-2act}.\\ \operatorname{var} r(t) &= \mathbb{E}r^2(t) - (\mathbb{E}r(t))^2\\ &= \frac{b\sigma^2}{2ac^2} + \left(r(0) - \frac{b}{c}\right)\frac{\sigma^2}{ac}e^{-act} + \frac{\sigma^2}{ac}\left(\frac{b}{2c} - r(0)\right)e^{-2act}.\end{aligned}$$

## 15.8 Multidimensional Brownian Motion

**Definition 15.2 ($d$-dimensional Brownian Motion)** A *$d$-dimensional Brownian Motion* is a process

$$B(t) = (B_1(t), \ldots, B_d(t))$$

with the following properties:

- Each $B_k(t)$ is a one-dimensional Brownian motion;
- If $i \neq j$, then the processes $B_i(t)$ and $B_j(t)$ are independent.

Associated with a $d$-dimensional Brownian motion, we have a filtration $\{\mathcal{F}(t)\}$ such that

- For each $t$, the random vector $B(t)$ is $\mathcal{F}(t)$-measurable;
- For each $t \leq t_1 \leq \ldots \leq t_n$, the vector increments

  $$B(t_1) - B(t), \ldots, B(t_n) - B(t_{n-1})$$

  are independent of $\mathcal{F}(t)$.

## 15.9 Cross-variations of Brownian motions

Because each component $B_i$ is a one-dimensional Brownian motion, we have the informal equation

$$dB_i(t)\ dB_i(t) = dt.$$

However, we have:

**Theorem 9.49** *If $i \neq j$,*

$$dB_i(t)\ dB_j(t) = 0$$

**Proof:** Let $\Pi = \{t_0, \ldots, t_n\}$ be a partition of $[0, T]$. For $i \neq j$, define the *sample cross variation* of $B_i$ and $B_j$ on $[0, T]$ to be

$$C_\Pi = \sum_{k=0}^{n-1} \left[B_i(t_{k+1}) - B_i(t_k)\right]\left[B_j(t_{k+1}) - B_j(t_k)\right].$$

The increments appearing on the right-hand side of the above equation are all independent of one another and all have mean zero. Therefore,

$$I\!\!E C_\Pi = 0.$$

We compute $\mathrm{var}(C_\Pi)$. First note that

$$\begin{aligned} C_\Pi^2 = &\sum_{k=0}^{n-1} \Big[B_i(t_{k+1}) - B_i(t_k)\Big]^2 \Big[B_j(t_{k+1}) - B_j(t_k)\Big]^2 \\ &+ 2\sum_{\ell<k}^{n-1} \left[B_i(t_{\ell+1}) - B_i(t_\ell)\right]\left[B_j(t_{\ell+1}) - B_j(t_\ell)\right].\left[B_i(t_{k+1}) - B_i(t_k)\right]\left[B_j(t_{k+1}) - B_j(t_k)\right] \end{aligned}$$

All the increments appearing in the sum of cross terms are independent of one another and have mean zero. Therefore,

$$\begin{aligned} \mathrm{var}(C_\Pi) &= I\!\!E C_\Pi^2 \\ &= I\!\!E \sum_{k=0}^{n-1} \left[B_i(t_{k+1}) - B_i(t_k)\right]^2 \left[B_j(t_{k+1}) - B_j(t_k)\right]^2. \end{aligned}$$

But $[B_i(t_{k+1}) - B_i(t_k)]^2$ and $[B_j(t_{k+1}) - B_j(t_k)]^2$ are independent of one another, and each has expectation $(t_{k+1} - t_k)$. It follows that

$$\mathrm{var}(C_\Pi) = \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 \leq ||\Pi|| \sum_{k=0}^{n-1} (t_{k+1} - t_k) = ||\Pi||.T.$$

As $||\Pi|| \to 0$, we have $\mathrm{var}(C_\Pi) \to 0$, so $C_\Pi$ converges to the constant $I\!\!E C_\Pi = 0$. ∎

## 15.10 Multi-dimensional Itô formula

To keep the notation as simple as possible, we write the Itô formula for *two* processes driven by a *two*-dimensional Brownian motion. The formula generalizes to *any number* of processes driven by a Brownian motion of *any number* (not necessarily the same number) of dimensions.

Let $X$ and $Y$ be processes of the form

$$X(t) = X(0) + \int_0^t \alpha(u)\, du + \int_0^t \delta_{11}(u)\, dB_1(u) + \int_0^t \delta_{12}(u)\, dB_2(u),$$
$$Y(t) = Y(0) + \int_0^t \beta(u)\, du + \int_0^t \delta_{21}(u)\, dB_1(u) + \int_0^t \delta_{22}(u)\, dB_2(u).$$

Such processes, consisting of a nonrandom initial condition, plus a Riemann integral, plus one or more Itô integrals, are called *semimartingales*. The integrands $\alpha(u), \beta(u)$, and $\delta_{ij}(u)$ can be any adapted processes. The adaptedness of the integrands guarantees that $X$ and $Y$ are also adapted. In differential notation, we write

$$dX = \alpha\, dt + \delta_{11}\, dB_1 + \delta_{12}\, dB_2,$$
$$dY = \beta\, dt + \delta_{21}\, dB_1 + \delta_{22}\, dB_2.$$

Given these two semimartingales $X$ and $Y$, the quadratic and cross variations are:

$$\begin{aligned}
dX\, dX &= (\alpha\, dt + \delta_{11}\, dB_1 + \delta_{12}\, dB_2)^2, \\
&= \delta_{11}^2 \underbrace{dB_1\, dB_1}_{dt} + 2\delta_{11}\delta_{12} \underbrace{dB_1\, dB_2}_{0} + \delta_{12}^2 \underbrace{dB_2\, dB_2}_{dt} \\
&= (\delta_{11}^2 + \delta_{12}^2)^2\, dt, \\
dY\, dY &= (\beta\, dt + \delta_{21}\, dB_1 + \delta_{22}\, dB_2)^2 \\
&= (\delta_{21}^2 + \delta_{22}^2)^2\, dt, \\
dX\, dY &= (\alpha\, dt + \delta_{11}\, dB_1 + \delta_{12}\, dB_2)(\beta\, dt + \delta_{21}\, dB_1 + \delta_{22}\, dB_2) \\
&= (\delta_{11}\delta_{21} + \delta_{12}\delta_{22})\, dt
\end{aligned}$$

Let $f(t,x,y)$ be a function of three variables, and let $X(t)$ and $Y(t)$ be semimartingales. Then we have the corresponding Itô formula:

$$df(t,x,y) \;=\; f_t\, dt \;+\; f_x\, dX \;+\; f_y\, dY \;+\; \tfrac{1}{2}\left[f_{xx}\, dX\, dX + 2f_{xy}\, dX\, dY + f_{yy}\, dY\, dY\right].$$

In integral form, with $X$ and $Y$ as decribed earlier and with all the variables filled in, this equation is

$$\begin{aligned}
&f(t, X(t), Y(t)) - f(0, X(0), Y(0)) \\
&= \int_0^t [f_t + \alpha f_x + \beta f_y + \tfrac{1}{2}(\delta_{11}^2 + \delta_{12}^2) f_{xx} + (\delta_{11}\delta_{21} + \delta_{12}\delta_{22}) f_{xy} + \tfrac{1}{2}(\delta_{21}^2 + \delta_{22}^2) f_{yy}]\, du \\
&\quad + \int_0^t [\delta_{11} f_x + \delta_{21} f_y]\, dB_1 \;+\; \int_0^t [\delta_{12} f_x + \delta_{22} f_y]\, dB_2,
\end{aligned}$$

where $f = f(u, X(u), Y(u))$, for $i,j \in \{1,2\}$, $\delta_{ij} = \delta_{ij}(u)$, and $B_i = B_i(u)$.

# Chapter 16

# Markov processes and the Kolmogorov equations

## 16.1 Stochastic Differential Equations

Consider the *stochastic differential equation*:

$$dX(t) = a(t, X(t))\,dt + \sigma(t, X(t))\,dB(t). \tag{SDE}$$

Here $a(t,x)$ and $\sigma(t,x)$ are given functions, usually assumed to be continuous in $(t,x)$ and Lipschitz continuous in $x$,i.e., there is a constant $L$ such that

$$|a(t,x) - a(t,y)| \le L|x-y|, \qquad |\sigma(t,x) - \sigma(t,y)| \le L|x-y|$$

for all $t, x, y$.

Let $(t_0, x)$ be given. A *solution* to (SDE) with the *initial condition* $(t_0, x)$ is a process $\{X(t)\}_{t \ge t_0}$ satisfying

$$\begin{aligned} X(t_0) &= x, \\ X(t) &= X(t_0) + \int_{t_0}^{t} a(s, X(s))\,ds + \int_{t_0}^{t} \sigma(s, X(s))\,dB(s), \qquad t \ge t_0 \end{aligned}$$

The solution process $\{X(t)\}_{t \ge t_0}$ will be adapted to the filtration $\{\mathcal{F}(t)\}_{t \ge 0}$ generated by the Brownian motion. If you know the path of the Brownian motion up to time $t$, then you can evaluate $X(t)$.

**Example 16.1 (Drifted Brownian motion)** Let $a$ be a constant and $\sigma = 1$, so

$$dX(t) = a\,dt + dB(t).$$

If $(t_0, x)$ is given and we start with the initial condition

$$X(t_0) = x,$$

then

$$X(t) = x + a(t - t_0) + (B(t) - B(t_0)), \qquad t \geq t_0.$$

To compute the differential w.r.t. $t$, treat $t_0$ and $B(t_0)$ as constants:

$$dX(t) = a\ dt + dB(t).$$

■

**Example 16.2 (Geometric Brownian motion)** Let $r$ and $\sigma$ be constants. Consider

$$dX(t) = rX(t)\ dt + \sigma X(t)\ dB(t).$$

Given the initial condition

$$X(t_0) = x,$$

the solution is

$$X(t) = x \exp\left\{\sigma(B(t) - B(t_0)) + (r - \tfrac{1}{2}\sigma^2)(t - t_0)\right\}.$$

Again, to compute the differential w.r.t. $t$, treat $t_0$ and $B(t_0)$ as constants:

$$\begin{aligned} dX(t) &= (r - \tfrac{1}{2}\sigma^2)X(t)\ dt + \sigma X(t)\ dB(t) + \tfrac{1}{2}\sigma^2 X(t)\ dt \\ &= rX(t)\ dt + \sigma X(t)\ dB(t). \end{aligned}$$

■

## 16.2 Markov Property

Let $0 \leq t_0 < t_1$ be given and let $h(y)$ be a function. Denote by

$$\mathbb{E}^{t_0,x} h(X(t_1))$$

the expectation of $h(X(t_1))$, given that $X(t_0) = x$. Now let $\xi \in \mathbb{R}$ be given, and start with initial condition

$$X(0) = \xi.$$

We have the *Markov property*

$$\mathbb{E}^{0,\xi}\left[h(X(t_1))\middle|\mathcal{F}(t_0)\right] = \mathbb{E}^{t_0,X(t_0)} h(X(t_1)).$$

In other words, if you observe the path of the driving Brownian motion from time 0 to time $t_0$, and based on this information, you want to estimate $h(X(t_1))$, the only relevant information is the value of $X(t_0)$. You imagine starting the $(SDE)$ at time $t_0$ at value $X(t_0)$, and compute the expected value of $h(X(t_1))$.

## 16.3 Transition density

Denote by

$$p(t_0, t_1;\ x, y)$$

the density (in the $y$ variable) of $X(t_1)$, conditioned on $X(t_0) = x$. In other words,

$$\mathbb{E}^{t_0,x} h(X(t_1)) = \int_{\mathbb{R}} h(y) p(t_0, t_1;\ x, y)\ dy.$$

The Markov property says that for $0 \le t_0 \le t_1$ and for every $\xi$,

$$\mathbb{E}^{0,\xi} \left[ h(X(t_1)) \middle| \mathcal{F}(t_0) \right] = \int_{\mathbb{R}} h(y) p(t_0, t_1;\ X(t_0), y)\ dy.$$

**Example 16.3 (Drifted Brownian motion)** Consider the SDE

$$dX(t) = a\ dt + dB(t).$$

Conditioned on $X(t_0) = x$, the random variable $X(t_1)$ is normal with mean $x + a(t_1 - t_0)$ and variance $(t_1 - t_0)$, i.e.,

$$p(t_0, t_1;\ x, y) = \frac{1}{\sqrt{2\pi(t_1 - t_0)}} \exp\left\{ -\frac{(y - (x + a(t_1 - t_0)))^2}{2(t_1 - t_0)} \right\}.$$

Note that $p$ depends on $t_0$ and $t_1$ only through their difference $t_1 - t_0$. This is always the case when $a(t, x)$ and $\sigma(t, x)$ don't depend on $t$. ■

**Example 16.4 (Geometric Brownian motion)** Recall that the solution to the SDE

$$dX(t) = rX(t)\ dt + \sigma X(t)\ dB(t),$$

with initial condition $X(t_0) = x$, is Geometric Brownian motion:

$$X(t_1) = x \exp\left\{ \sigma(B(t_1) - B(t_0)) + (r - \tfrac{1}{2}\sigma^2)(t_1 - t_0) \right\}.$$

The random variable $B(t_1) - B(t_0)$ has density

$$\mathbb{P}\left\{ B(t_1) - B(t_0) \in db \right\} = \frac{1}{\sqrt{2\pi(t_1 - t_0)}} \exp\left\{ -\frac{b^2}{2(t_1 - t_0)} \right\}\ db,$$

and we are making the change of variable

$$y = x \exp\left\{ \sigma b + (r - \tfrac{1}{2}\sigma^2)(t_1 - t_0) \right\}$$

or equivalently,

$$b = \frac{1}{\sigma} \left[ \log \frac{y}{x} - (r - \tfrac{1}{2}\sigma^2)(t_1 - t_0) \right].$$

The derivative is

$$\frac{dy}{db} = \sigma y, \qquad \text{or equivalently,} \qquad db = \frac{dy}{\sigma y}.$$

Therefore,

$$
\begin{aligned}
p(t_0, t_1; x, y)\ dy &= I\!P\{X(t_1) \in dy\} \\
&= \frac{1}{\sigma y \sqrt{2\pi(t_1 - t_0)}} \exp\left\{ -\frac{1}{2(t_1 - t_0)\sigma^2} \left[ \log \frac{y}{x} - (r - \tfrac{1}{2}\sigma^2)(t_1 - t_0) \right]^2 \right\}\ dy.
\end{aligned}
$$

Using the transition density and a fair amount of calculus, one can compute the expected payoff from a European call:

$$
\begin{aligned}
I\!E^{t,x}(X(T) - K)^+ &= \int_0^\infty (y - K)^+ p(t, T; x, y)\ dy \\
&= e^{r(T-t)} x N\left( \frac{1}{\sigma\sqrt{T-t}} \left[ \log \frac{x}{K} + r(T-t) + \tfrac{1}{2}\sigma^2(T-t) \right] \right) \\
&\quad - K N\left( \frac{1}{\sigma\sqrt{T-t}} \left[ \log \frac{x}{K} + r(T-t) - \tfrac{1}{2}\sigma^2(T-t) \right] \right)
\end{aligned}
$$

where

$$
N(\eta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\eta} e^{-\frac{1}{2}x^2}\ dx = \frac{1}{\sqrt{2\pi}} \int_{-\eta}^{\infty} e^{-\frac{1}{2}x^2}\ dx.
$$

Therefore,

$$
\begin{aligned}
I\!E^{0,\xi}\left[ e^{-r(T-t)}(X(T) - K)^+ \middle| \mathcal{F}(t) \right] &= e^{-r(T-t)} I\!E^{t,X(t)}(X(T) - K)^+ \\
&= X(t) N\left( \frac{1}{\sigma\sqrt{T-t}} \left[ \log \frac{X(t)}{K} + r(T-t) + \tfrac{1}{2}\sigma^2(T-t) \right] \right) \\
&\quad - e^{-r(T-t)} K\ N\left( \frac{1}{\sigma\sqrt{T-t}} \left[ \log \frac{X(t)}{K} + r(T-t) - \tfrac{1}{2}\sigma^2(T-t) \right] \right)
\end{aligned}
$$

■

## 16.4 The Kolmogorov Backward Equation

Consider

$$
dX(t) = a(t, X(t))\ dt + \sigma(t, X(t))\ dB(t),
$$

and let $p(t_0, t_1; x, y)$ be the transition density. Then the Kolmogorov Backward Equation is:

$$
-\frac{\partial}{\partial t_0} p(t_0, t_1;\ x, y) = a(t_0, x) \frac{\partial}{\partial x} p(t_0, t_1;\ x, y) + \tfrac{1}{2}\sigma^2(t_0, x) \frac{\partial^2}{\partial x^2} p(t_0, t_1;\ x, y). \tag{KBE}
$$

The variables $t_0$ and $x$ in $(KBE)$ are called the *backward variables*.

In the case that $a$ and $\sigma$ are functions of $x$ alone, $p(t_0, t_1;\ x, y)$ depends on $t_0$ and $t_1$ only through their difference $\tau = t_1 - t_0$. We then write $p(\tau;\ x, y)$ rather than $p(t_0, t_1;\ x, y)$, and $(KBE)$ becomes

$$
\frac{\partial}{\partial \tau} p(\tau;\ x, y) = a(x) \frac{\partial}{\partial x} p(\tau;\ x, y) + \tfrac{1}{2}\sigma^2(x) \frac{\partial^2}{\partial x^2} p(\tau;\ x, y). \tag{KBE'}
$$

**Example 16.5 (Drifted Brownian motion)**

$$\begin{aligned}
dX(t) &= a\ dt + dB(t)\\
p(\tau;\ x,y) &= \frac{1}{\sqrt{2\pi\tau}}\exp\left\{-\frac{(y-(x+a\tau))^2}{2\tau}\right\}.\\
\frac{\partial}{\partial\tau}p = p_\tau &= \left(\frac{\partial}{\partial\tau}\frac{1}{\sqrt{2\pi\tau}}\right)\exp\left\{-\frac{(y-x-a\tau)^2}{2\tau}\right\}\\
&\quad -\left(\frac{\partial}{\partial\tau}\frac{(y-x-a\tau)^2}{2\tau}\right)\frac{1}{\sqrt{2\pi\tau}}\exp\left\{-\frac{(y-x-a\tau)^2}{2\tau}\right\}\\
&= \left[-\frac{1}{2\tau}+\frac{a(y-x-a\tau)}{\tau}+\frac{(y-x-a\tau)}{2\tau^2}\right]p.\\
\frac{\partial}{\partial x}p = p_x &= \frac{y-x-a\tau}{\tau}p.\\
\frac{\partial^2}{\partial x^2}p = p_{xx} &= \left(\frac{\partial}{\partial x}\frac{y-x-a\tau}{\tau}\right)p + \frac{y-x-a\tau}{\tau}p_x\\
&= -\frac{1}{\tau}p + \frac{(y-x-a\tau)^2}{\tau^2}p.
\end{aligned}$$

Therefore,

$$\begin{aligned}
ap_x + \tfrac{1}{2}p_{xx} &= \left[\frac{a(y-x-a\tau)}{\tau} - \frac{1}{2\tau} + \frac{(y-x-a\tau)^2}{2\tau^2}\right]p\\
&= p_\tau.
\end{aligned}$$

This is the Kolmogorov backward equation. ■

**Example 16.6 (Geometric Brownian motion)**

$$dX(t) = rX(t)\ dt + \sigma X(t)\ dB(t).$$

$$p(\tau;\ x,y) = \frac{1}{\sigma y\sqrt{2\pi\tau}}\exp\left\{-\frac{1}{2\tau\sigma^2}\left[\log\frac{y}{x} - (r-\tfrac{1}{2}\sigma^2)\tau\right]^2\right\}.$$

It is true but very tedious to verify that $p$ satisfies the KBE

$$p_\tau = rxp_x + \tfrac{1}{2}\sigma^2x^2p_{xx}.$$

■

## 16.5 Connection between stochastic calculus and KBE

Consider

$$dX(t) = a(X(t))\ dt + \sigma(X(t))\ dB(t). \tag{5.1}$$

Let $h(y)$ be a function, and define

$$v(t,x) = I\!\!E^{t,x}h(X(T)),$$

where $0 \leq t \leq T$. Then

$$\begin{aligned} v(t,x) &= \int h(y)\ p(T-t;\ x,y)\ dy, \\ v_t(t,x) &= -\int h(y)\ p_\tau(T-t;\ x,y)\ dy, \\ v_x(t,x) &= \int h(y)\ p_x(T-t;\ x,y)\ dy, \\ v_{xx}(t,x) &= \int h(y)\ p_{xx}(T-t;\ x,y)\ dy. \end{aligned}$$

Therefore, the Kolmogorov backward equation implies

$$\begin{aligned} &v_t(t,x) + a(x)v_x(t,x) + \tfrac{1}{2}\sigma^2(x)v_{xx}(t,x) = \\ &\qquad \int h(y)\left[-p_\tau(T-t;x,y) + a(x)p_x(T-t;x,y) + \tfrac{1}{2}\sigma^2(x)p_{xx}(T-t;x,y)\right]\ dy = 0 \end{aligned}$$

Let $(0,\xi)$ be an initial condition for the SDE (5.1). We simplify notation by writing $I\!\!E$ rather than $I\!\!E^{0,\xi}$.

**Theorem 5.50** *Starting at $X(0) = \xi$, the process $v(t, X(t))$ satisfies the martingale property:*

$$I\!\!E\left[v(t,X(t))\middle|\mathcal{F}(s)\right] = v(s,X(s)), \qquad 0 \leq s \leq t \leq T.$$

**Proof:** According to the Markov property,

$$I\!\!E\left[h(X(T))\middle|\mathcal{F}(t)\right] = I\!\!E^{t,X(t)}h(X(T)) = v(t,X(t)),$$

so

$$\begin{aligned} I\!\!E\left[v(t,X(t))|\mathcal{F}(s)\right] &= I\!\!E\left[\ I\!\!E\left[h(X(T))\middle|\mathcal{F}(t)\right]\ \middle|\mathcal{F}(s)\right] \\ &= I\!\!E\left[h(X(T))\middle|\mathcal{F}(s)\right] \\ &= I\!\!E^{s,X(s)}h(X(T)) \qquad \text{(Markov property)} \\ &= v(s,X(s)). \end{aligned}$$

■

Itô's formula implies

$$\begin{aligned} dv(t,X(t)) &= v_t dt + v_x dX + \tfrac{1}{2}v_{xx}dX\ dX \\ &= v_t dt + a v_x dt + \sigma v_x dB + \tfrac{1}{2}\sigma^2 v_{xx} dt. \end{aligned}$$

In integral form, we have

$$\begin{aligned} v(t,X(t)) = v(0,X(0)) \\ &+ \int_0^t \left[ v_t(u,X(u)) + a(X(u))v_x(u,X(u)) + \tfrac{1}{2}\sigma^2(X(u))v_{xx}(u,X(u)) \right] du \\ &+ \int_0^t \sigma(X(u))v_x(u,X(u))\, dB(u). \end{aligned}$$

We know that $v(t,X(t))$ is a martingale, so the integral $\int_0^t \left[ v_t + av_x + \frac{1}{2}\sigma^2 v_{xx} \right] du$ must be zero for all $t$. This implies that the integrand is zero; hence

$$v_t + av_x + \tfrac{1}{2}\sigma^2 v_{xx} = 0.$$

Thus by two different arguments, one based on the Kolmogorov backward equation, and the other based on Itô's formula, we have come to the same conclusion.

**Theorem 5.51 (Feynman-Kac)** *Define*

$$v(t,x) = \mathbb{E}^{t,x} h(X(T)), \qquad 0 \le t \le T,$$

*where*

$$dX(t) = a(X(t))\, dt + \sigma(X(t))\, dB(t).$$

*Then*

$$v_t(t,x) + a(x)v_x(t,x) + \tfrac{1}{2}\sigma^2(x)v_{xx}(t,x) = 0 \tag{FK}$$

*and*

$$v(T,x) = h(x).$$

The Black-Scholes equation is a special case of this theorem, as we show in the next section.

**Remark 16.1 (Derivation of KBE)** We plunked down the Kolmogorov backward equation without any justification. In fact, one can use Itô's formula to prove the Feynman-Kac Theorem, and use the Feynman-Kac Theorem to derive the Kolmogorov backward equation.

## 16.6 Black-Scholes

Consider the SDE

$$dS(t) = rS(t)\, dt + \sigma S(t)\, dB(t).$$

With initial condition

$$S(t) = x,$$

the solution is

$$S(u) = x \exp\left\{ \sigma(B(u) - B(t)) + (r - \tfrac{1}{2}\sigma^2)(u-t) \right\}, \qquad u \ge t.$$

Define

$$\begin{aligned} v(t,x) &= I\!\!E^{t,x} h(S(T)) \\ &= I\!\!E h\left(x \exp\left\{\sigma(B(T)-B(t)) + (r - \tfrac{1}{2}\sigma^2)(T-t)\right\}\right), \end{aligned}$$

where $h$ is a function to be specified later.

Recall the *Independence Lemma*: If $\mathcal{G}$ is a $\sigma$-field, $X$ is $\mathcal{G}$-measurable, and $Y$ is independent of $\mathcal{G}$, then

$$I\!\!E\left[h(X,Y)\middle|\mathcal{G}\right] = \gamma(X),$$

where

$$\gamma(x) = I\!\!E h(x,Y).$$

With geometric Brownian motion, for $0 \le t \le T$, we have

$$\begin{aligned} S(t) &= S(0)\exp\left\{\sigma B(t) + (r - \tfrac{1}{2}\sigma^2)t\right\}, \\ S(T) &= S(0)\exp\left\{\sigma B(T) + (r - \tfrac{1}{2}\sigma^2)T\right\} \\ &= \underbrace{S(t)}_{\mathcal{F}(t)\text{-measurable}} \quad \underbrace{\exp\left\{\sigma(B(T)-B(t)) + (r - \tfrac{1}{2}\sigma^2)(T-t)\right\}}_{\text{independent of } \mathcal{F}(t)} \end{aligned}$$

We thus have

$$S(T) = XY,$$

where

$$\begin{aligned} X &= S(t) \\ Y &= \exp\left\{\sigma(B(T)-B(t)) + (r - \tfrac{1}{2}\sigma^2)(T-t)\right\}. \end{aligned}$$

Now

$$I\!\!E h(xY) = v(t,x).$$

The independence lemma implies

$$\begin{aligned} I\!\!E\left[h(S(T))\middle|\mathcal{F}(t)\right] &= I\!\!E\left[h(XY)|\mathcal{F}(t)\right] \\ &= v(t,X) \\ &= v(t,S(t)). \end{aligned}$$

We have shown that

$$v(t, S(t)) = \mathbb{E}\left[h(S(T))\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T.$$

Note that the random variable $h(S(T))$ whose conditional expectation is being computed does not depend on $t$. Because of this, the tower property implies that $v(t, S(t))$, $0 \le t \le T$, is a martingale: For $0 \le s \le t \le T$,

$$\begin{aligned} \mathbb{E}\left[v(t, S(t))\middle|\mathcal{F}(s)\right] &= \mathbb{E}\left[\ \mathbb{E}\left[h(S(T))\middle|\mathcal{F}(t)\right]\ \middle|\mathcal{F}(s)\right] \\ &= \mathbb{E}\left[h(S(T))\middle|\mathcal{F}(s)\right] \\ &= v(s, S(s)). \end{aligned}$$

This is a special case of Theorem 5.51.

Because $v(t, S(t))$ is a martingale, the sum of the $dt$ terms in $dv(t, S(t))$ must be 0. By Itô's formula,

$$\begin{aligned} dv(t, S(t)) = &\left[v_t(t, S(t))\ dt + rS(t)v_x(t, S(t)) + \tfrac{1}{2}\sigma^2 S^2(t) v_{xx}(t, S(t))\right]\ dt \\ &+ \sigma S(t) v_x(t, S(t))\ dB(t). \end{aligned}$$

This leads us to the equation

$$v_t(t, x) + rxv_x(t, x) + \tfrac{1}{2}\sigma^2 x^2 v_{xx}(t, x) = 0, \qquad 0 \le t < T, \quad x \ge 0.$$

This is a special case of Theorem 5.51 (Feynman-Kac).

Along with the above partial differential equation, we have the *terminal condition*

$$v(T, x) = h(x), \qquad x \ge 0.$$

Furthermore, if $S(t) = 0$ for some $t \in [0, T]$, then also $S(T) = 0$. This gives us the *boundary condition*

$$v(t, 0) = h(0), \qquad 0 \le t \le T.$$

Finally, we shall eventually see that the value at time $t$ of a contingent claim paying $h(S(T))$ is

$$\begin{aligned} u(t, x) &= e^{-r(T-t)} \mathbb{E}^{t,x} h(S(T)) \\ &= e^{-r(T-t)} v(t, x) \end{aligned}$$

at time $t$ if $S(t) = x$. Therefore,

$$\begin{aligned} v(t, x) &= e^{r(T-t)} u(t, x), \\ v_t(t, x) &= -re^{r(T-t)} u(t, x) + e^{r(T-t)} u_t(t, x), \\ v_x(t, x) &= e^{r(T-t)} u_x(t, x), \\ v_{xx}(t, x) &= e^{r(T-t)} u_{xx}(t, x). \end{aligned}$$

Plugging these formulas into the partial differential equation for $v$ and cancelling the $e^{r(T-t)}$ appearing in every term, we obtain the *Black-Scholes partial differential equation*:

$$-ru(t,x) + u_t(t,x) + rxu_x(t,x) + \tfrac{1}{2}\sigma^2 x^2 u_{xx}(t,x) = 0, \qquad 0 \le t < T, \;\; x \ge 0. \tag{BS}$$

Compare this with the earlier derivation of the Black-Scholes PDE in Section 15.6.

In terms of the transition density

$$p(t,T;\; x,y) = \frac{1}{\sigma y\sqrt{2\pi(T-t)}} \exp\left\{ -\frac{1}{2(T-t)\sigma^2}\left[\log\frac{y}{x} - (r - \tfrac{1}{2}\sigma^2)(T-t)\right]^2 \right\}$$

for geometric Brownian motion (See Example 16.4), we have the "stochastic representation"

$$\begin{aligned} u(t,x) &= e^{-r(T-t)} \mathbb{E}^{t,x} h(S(T)) \qquad \text{(SR)} \\ &= e^{-r(T-t)} \int_0^\infty h(y)p(t,T;\; x,y)\, dy. \end{aligned}$$

In the case of a call,

$$h(y) = (y-K)^+$$

and

$$\begin{aligned} u(t,x) = x\; N&\left(\frac{1}{\sigma\sqrt{T-t}}\left[\log\frac{x}{K} + r(T-t) + \tfrac{1}{2}\sigma^2(T-t)\right]\right) \\ &- e^{-r(T-t)} K\; N\left(\frac{1}{\sigma\sqrt{T-t}}\left[\log\frac{x}{K} + r(T-t) - \tfrac{1}{2}\sigma^2(T-t)\right]\right) \end{aligned}$$

Even if $h(y)$ is some other function (e.g., $h(y) = (K-y)^+$, a put), $u(t,x)$ is still given by and satisfies the Black-Scholes PDE (BS) derived above.

## 16.7 Black-Scholes with price-dependent volatility

$$\begin{aligned} dS(t) &= rS(t)\, dt + \beta(S(t))\, dB(t), \\ v(t,x) &= e^{-r(T-t)} \mathbb{E}^{t,x} (S(T)-K)^+. \end{aligned}$$

The Feynman-Kac Theorem now implies that

$$-rv(t,x) + v_t(t,x) + rxv_x(t,x) + \tfrac{1}{2}\beta^2(x)v_{xx}(t,x) = 0, \qquad 0 \le t < T, \;\; x > 0.$$

$v$ also satisfies the *terminal condition*

$$v(T,x) = (x-K)^+, \qquad x \ge 0,$$

and the *boundary condition*

$$v(t,0) = 0, \qquad 0 \le t \le T.$$

An example of such a process is the following from J.C. Cox, *Notes on options pricing I: Constant elasticity of variance diffusions,* Working Paper, Stanford University, 1975:

$$dS(t) = rS(t)\ dt + \sigma S^{\delta}(t)\ dB(t),$$

where $0 \le \delta < 1$. The "volatility" $\sigma S^{\delta-1}(t)$ decreases with increasing stock price. The corresponding Black-Scholes equation is

$$\begin{aligned} -rv + v_t + rxv_x + \tfrac{1}{2}\sigma^2 x^{2\delta} v_{xx} &= 0, \qquad 0 \le t < T \quad x > 0; \\ v(t,0) &= 0, \qquad 0 \le t \le T \\ v(T,x) &= (x-K)^+, \qquad x \ge 0. \end{aligned}$$

# Chapter 17

# Girsanov's theorem and the risk-neutral measure

(Please see Oksendal, 4th ed., pp 145–151.)

**Theorem 0.52 (Girsanov, One-dimensional)** *Let $B(t), 0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}(t), 0 \leq t \leq T$, be the accompanying filtration, and let $\theta(t), 0 \leq t \leq T$, be a process adapted to this filtration. For $0 \leq t \leq T$, define*

$$
\begin{aligned}
\widetilde{B}(t) &= \int_0^t \theta(u)\, du + B(t), \\
Z(t) &= \exp\left\{ -\int_0^t \theta(u)\, dB(u) - \tfrac{1}{2}\int_0^t \theta^2(u)\, du \right\},
\end{aligned}
$$

*and define a new probability measure by*

$$
\widetilde{I\!\!P}(A) = \int_A Z(T)\, dI\!\!P, \qquad \forall A \in \mathcal{F}.
$$

*Under $\widetilde{I\!\!P}$, the process $\widetilde{B}(t), 0 \leq t \leq T$, is a Brownian motion.*

**Caveat:** This theorem requires a technical condition on the size of $\theta$. If

$$
I\!\!E \exp\left\{ \tfrac{1}{2}\int_0^T \theta^2(u)\, du \right\} < \infty,
$$

everything is OK.

We make the following remarks:

$Z(t)$ **is a matingale.** In fact,

$$
\begin{aligned}
dZ(t) &= -\theta(t)Z(t)\, dB(t) + \tfrac{1}{2}\theta^2(t)Z(t)\, dB(t)\, dB(t) - \tfrac{1}{2}\theta^2(t)Z(t)\, dt \\
&= -\theta(t)Z(t)\, dB(t).
\end{aligned}
$$

$\widetilde{I\!P}$ **is a probability measure.** Since $Z(0) = 1$, we have $I\!E Z(t) = 1$ for every $t \geq 0$. In particular

$$\widetilde{I\!P}(\Omega) = \int_\Omega Z(T)\, dI\!P = I\!E Z(T) = 1,$$

so $\widetilde{I\!P}$ is a probability measure.

$\widetilde{I\!E}$ **in terms of** $I\!E$**.** Let $\widetilde{I\!E}$ denote expectation under $\widetilde{I\!P}$. If $X$ is a random variable, then

$$\widetilde{I\!E} Z = I\!E\left[Z(T)X\right].$$

To see this, consider first the case $X = \mathbf{1}_A$, where $A \in \mathcal{F}$. We have

$$\widetilde{I\!E} X = \widetilde{I\!P}(A) = \int_A Z(T)\, dI\!P = \int_\Omega Z(T)\mathbf{1}_A\, dI\!P = I\!E\left[Z(T)X\right].$$

Now use Williams' "standard machine".

$\widetilde{I\!P}$ **and** $I\!P$**.** The intuition behind the formula

$$\widetilde{I\!P}(A) = \int_A Z(T)\, dI\!P \qquad \forall A \in \mathcal{F}$$

is that we want to have

$$\widetilde{I\!P}(\omega) = Z(T,\omega) I\!P(\omega),$$

but since $I\!P(\omega) = 0$ and $\widetilde{I\!P}(\omega) = 0$, this doesn't really tell us anything useful about $\widetilde{I\!P}$. Thus, we consider subsets of $\Omega$, rather than individual elements of $\Omega$.

**Distribution of** $\widetilde{B}(T)$**.** If $\theta$ is constant, then

$$\begin{aligned} Z(T) &= \exp\left\{-\theta B(T) - \tfrac{1}{2}\theta^2 T\right\} \\ \widetilde{B}(T) &= \theta T + B(T). \end{aligned}$$

Under $I\!P$, $B(T)$ is normal with mean 0 and variance $T$, so $\widetilde{B}(T)$ is normal with mean $\theta T$ and variance $T$:

$$I\!P(\widetilde{B}(T) \in d\tilde{b}) = \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(\tilde{b} - \theta T)^2}{2T}\right\}\, d\tilde{b}.$$

**Removal of Drift from** $\widetilde{B}(T)$**.** The change of measure from $I\!P$ to $\widetilde{I\!P}$ removes the drift from $\widetilde{B}(T)$. To see this, we compute

$$\begin{aligned} \widetilde{I\!E}\widetilde{B}(T) &= I\!E\left[Z(T)(\theta T + B(T))\right] \\ &= I\!E\left[\exp\left\{-\theta B(T) - \tfrac{1}{2}\theta^2 T\right\}\, (\theta T + B(T))\right] \\ &= \frac{1}{\sqrt{2\pi T}} \int_{-\infty}^{\infty} (\theta T + b) \exp\{-\theta b - \tfrac{1}{2}\theta^2 T\}\, \exp\left\{-\frac{b^2}{2T}\right\}\, db \\ &= \frac{1}{\sqrt{2\pi T}} \int_{-\infty}^{\infty} (\theta T + b) \exp\left\{-\frac{(b + \theta T)^2}{2T}\right\}\, db \\ (y = \theta T + b) &= \frac{1}{\sqrt{2\pi T}} \int_{-\infty}^{\infty} y \exp\left\{-\frac{y^2}{2}\right\}\, dy \quad \text{(Substitute } y = \theta T + b\text{)} \\ &= 0. \end{aligned}$$

We can also see that $\widetilde{I\!E}\widetilde{B}(T) = 0$ by arguing directly from the density formula

$$I\!P\left\{\widetilde{B}(t) \in d\tilde{b}\right\} = \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(\tilde{b}-\theta T)^2}{2T}\right\} \, d\tilde{b}.$$

Because

$$\begin{aligned} Z(T) &= \exp\{-\theta B(T) - \tfrac{1}{2}\theta^2 T\} \\ &= \exp\{-\theta(\widetilde{B}(T) - \theta T) - \tfrac{1}{2}\theta^2 T\} \\ &= \exp\{-\theta\widetilde{B}(T) + \tfrac{1}{2}\theta^2 T\}, \end{aligned}$$

we have

$$\begin{aligned} \widetilde{I\!P}\left\{\widetilde{B}(T) \in d\tilde{b}\right\} &= I\!P\left\{\widetilde{B}(T) \in d\tilde{b}\right\} \, \exp\left\{-\theta\tilde{b} + \tfrac{1}{2}\theta^2 T\right\} \\ &= \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(\tilde{b}-\theta T)^2}{2T} - \theta\tilde{b} + \tfrac{1}{2}\theta^2 T\right\} \, d\tilde{b}. \\ &= \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{\tilde{b}^2}{2T}\right\} \, d\tilde{b}. \end{aligned}$$

Under $\widetilde{I\!P}$, $\widetilde{B}(T)$ is normal with *mean zero* and variance $T$. Under $I\!P$, $\widetilde{B}(T)$ is normal with *mean* $\theta T$ and variance $T$.

**Means change, variances don't.** When we use the Girsanov Theorem to change the probability measure, means change but variances do not. Martingales may be destroyed or created. Volatilities, quadratic variations and cross variations are unaffected. Check:

$$d\widetilde{B}\; d\widetilde{B} = (\theta(t)\; dt + dB(t))^2 = dB.dB = dt.$$

## 17.1 Conditional expectations under $\widetilde{I\!P}$

**Lemma 1.53** *Let $0 \le t \le T$. If $X$ is $\mathcal{F}(t)$-measurable, then*

$$\widetilde{I\!E}X = I\!E[X.Z(t)].$$

**Proof:**

$$\begin{aligned} \widetilde{I\!E}X &= I\!E[X.Z(T)] = I\!E\left[\; I\!E[X.Z(T)|\mathcal{F}(t)]\;\right] \\ &= I\!E\left[X \; I\!E[Z(T)|\mathcal{F}(t)]\;\right] \\ &= I\!E[X.Z(t)] \end{aligned}$$

because $Z(t), 0 \le t \le T$, is a martingale under $I\!P$. ∎

**Lemma 1.54 (Baye's Rule)** *If $X$ is $\mathcal{F}(t)$-measurable and $0 \leq s \leq t \leq T$, then*

$$\widetilde{I\!\!E}[X|\mathcal{F}(s)] = \frac{1}{Z(s)} I\!\!E[XZ(t)|\mathcal{F}(s)]. \tag{1.1}$$

**Proof:** It is clear that $\frac{1}{Z(s)} I\!\!E[XZ(t)|\mathcal{F}(s)]$ is $\mathcal{F}(s)$-measurable. We check the partial averaging property. For $A \in \mathcal{F}(s)$, we have

$$\begin{aligned}
\int_A \frac{1}{Z(s)} I\!\!E[XZ(t)|\mathcal{F}(s)]\, d\widetilde{I\!\!P} &= \widetilde{I\!\!E}\left[\mathbf{1}_A \frac{1}{Z(s)} I\!\!E[XZ(t)|\mathcal{F}(s)]\right] \\
&= I\!\!E\left[\mathbf{1}_A I\!\!E[XZ(t)|\mathcal{F}(s)]\right] \qquad \text{(Lemma 1.53)} \\
&= I\!\!E\left[I\!\!E[\mathbf{1}_A XZ(t)|\mathcal{F}(s)]\right] \qquad \text{(Taking in what is known)} \\
&= I\!\!E[\mathbf{1}_A XZ(t)] \\
&= \widetilde{I\!\!E}[\mathbf{1}_A X] \qquad \text{(Lemma 1.53 again)} \\
&= \int_A X\, d\widetilde{I\!\!P}.
\end{aligned}$$

$\blacksquare$

Although we have proved Lemmas 1.53 and 1.54, we have not proved Girsanov's Theorem. We will not prove it completely, but here is the beginning of the proof.

**Lemma 1.55** *Using the notation of Girsanov's Theorem, we have the martingale property*

$$\widetilde{I\!\!E}[\widetilde{B}(t)|\mathcal{F}(s)] = \widetilde{B}(s), \qquad 0 \leq s \leq t \leq T.$$

**Proof:** We first check that $\widetilde{B}(t)Z(t)$ is a martingale under $I\!\!P$. Recall

$$\begin{aligned}
d\widetilde{B}(t) &= \theta(t)\, dt + dB(t), \\
dZ(t) &= -\theta(t)Z(t)\, dB(t).
\end{aligned}$$

Therefore,

$$\begin{aligned}
d(\widetilde{B}Z) &= \widetilde{B}\, dZ + Z\, d\widetilde{B} + d\widetilde{B}\, dZ \\
&= -\widetilde{B}\theta Z\, dB + Z\theta\, dt + Z\, dB - \theta Z\, dt \\
&= (-\widetilde{B}\theta Z + Z)\, dB.
\end{aligned}$$

Next we use Bayes' Rule. For $0 \leq s \leq t \leq T$,

$$\begin{aligned}
\widetilde{I\!\!E}[\widetilde{B}(t)|\mathcal{F}(s)] &= \frac{1}{Z(s)} I\!\!E[\widetilde{B}(t)Z(t)|\mathcal{F}(s)] \\
&= \frac{1}{Z(s)} \widetilde{B}(s)Z(s) \\
&= \widetilde{B}(s).
\end{aligned}$$

$\blacksquare$

**Definition 17.1 (Equivalent measures)** Two measures on the same probability space which have the same measure-zero sets are said to be *equivalent*.

The probability measures $I\!P$ and $\widetilde{I\!P}$ of the Girsanov Theorem are equivalent. Recall that $\widetilde{I\!P}$ is defined by

$$\widetilde{I\!P}(A) = \int Z(T)\ dI\!P, \qquad A \in \mathcal{F}.$$

If $I\!P(A) = 0$, then $\int_A Z(T)\ dI\!P = 0$. Because $Z(T) > 0$ for every $\omega$, we can invert the definition of $\widetilde{I\!P}$ to obtain

$$I\!P(A) = \int_A \frac{1}{Z(T)}\ d\widetilde{I\!P}, \qquad A \in \mathcal{F}.$$

If $\widetilde{I\!P}(A) = 0$, then $\int_A \frac{1}{Z(T)}\ dI\!P = 0$.

## 17.2 Risk-neutral measure

As usual we are given the **Brownian motion:** $B(t), 0 \le t \le T$, with filtration $\mathcal{F}(t), 0 \le t \le T$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We can then define the following.

**Stock price:**

$$dS(t) = \mu(t)S(t)\ dt + \sigma(t)S(t)\ dB(t).$$

The processes $\mu(t)$ and $\sigma(t)$ are adapted to the filtration. The stock price model is completely general, subject only to the condition that the paths of the process are continuous.

**Interest rate:** $r(t), 0 \le t \le T$. The process $r(t)$ is adapted.

**Wealth of an agent,** starting with $X(0) = x$. We can write the wealth process differential in several ways:

$$\begin{aligned}
dX(t) &= \underbrace{\Delta(t)\ dS(t)}_{\text{Capital gains from Stock}} + \underbrace{r(t)[X(t) - \Delta(t)S(t)]\ dt}_{\text{Interest earnings}} \\
&= r(t)X(t)\ dt + \Delta(t)[dS(t) - rS(t)\ dt] \\
&= r(t)X(t)\ dt + \Delta(t)\underbrace{(\mu(t) - r(t))}_{\text{Risk premium}}S(t)\ dt + \Delta(t)\sigma(t)S(t)\ dB(t) \\
&= r(t)X(t)\ dt + \Delta(t)\sigma(t)S(t)\left[\underbrace{\frac{\mu(t) - r(t)}{\sigma(t)}}_{\text{Market price of risk}=\theta(t)}\ dt + dB(t)\right]
\end{aligned}$$

**Discounted processes:**

$$\begin{aligned}
d\left(e^{-\int_0^t r(u)\,du} S(t)\right) &= e^{-\int_0^t r(u)\,du}[-r(t)S(t)\;dt + dS(t)] \\
d\left(e^{-\int_0^t r(u)\,du} X(t)\right) &= e^{-\int_0^t r(u)\,du}[-r(t)X(t)\;dt + dX(t)] \\
&= \Delta(t)d\left(e^{-\int_0^t r(u)\,du} S(t)\right).
\end{aligned}$$

**Notation:**

$$\begin{aligned}
\beta(t) = e^{\int_0^t r(u)\,du}, &\qquad \frac{1}{\beta(t)} = e^{-\int_0^t r(u)\,du}, \\
d\beta(t) = r(t)\beta(t)\;dt, &\qquad d\left(\frac{1}{\beta(t)}\right) = -\frac{r(t)}{\beta(t)}\;dt.
\end{aligned}$$

The discounted formulas are

$$\begin{aligned}
d\left(\frac{S(t)}{\beta(t)}\right) &= \frac{1}{\beta(t)}\left[-r(t)S(t)\;dt + dS(t)\right] \\
&= \frac{1}{\beta(t)}\left[(\mu(t) - r(t))S(t)\;dt + \sigma(t)S(t)\;dB(t)\right] \\
&= \frac{1}{\beta(t)}\sigma(t)S(t)\left[\theta(t)\;dt + dB(t)\right], \\
d\left(\frac{X(t)}{\beta(t)}\right) &= \Delta(t)\;d\left(\frac{S(t)}{\beta(t)}\right) \\
&= \frac{\Delta(t)}{\beta(t)}\;\sigma(t)S(t)\;[\theta(t)\;dt + dB(t)].
\end{aligned}$$

**Changing the measure.** Define

$$\widetilde{B}(t) = \int_0^t \theta(u)\;du + B(t).$$

Then

$$\begin{aligned}
d\left(\frac{S(t)}{\beta(t)}\right) &= \frac{1}{\beta(t)}\sigma(t)S(t)\;d\widetilde{B}(t), \\
d\left(\frac{X(t)}{\beta(t)}\right) &= \frac{\Delta(t)}{\beta(t)}\sigma(t)S(t)\;d\widetilde{B}(t).
\end{aligned}$$

Under $\widetilde{I\!\!P}$, $\frac{S(t)}{\beta(t)}$ and $\frac{X(t)}{\beta(t)}$ are martingales.

**Definition 17.2 (Risk-neutral measure)** A *risk-neutral measure* (sometimes called a *martingale measure*) is any probability measure, equivalent to the market measure $I\!\!P$, which makes all discounted asset prices martingales.

For the market model considered here,

$$\widetilde{I\!P}(A) = \int_A Z(T)\; dI\!P, \qquad A \in \mathcal{F},$$

where

$$Z(t) = \exp\left\{ -\int_0^t \theta(u)\; dB(u) - \tfrac{1}{2}\int_0^t \theta^2(u)\; du \right\},$$

is the unique risk-neutral measure. Note that because $\theta(t) = \frac{\mu(t)-r(t)}{\sigma(t)}$, we must assume that $\sigma(t) \neq 0$.

**Risk-neutral valuation.** Consider a contingent claim paying an $\mathcal{F}(T)$-measurable random variable $V$ at time $T$.

**Example 17.1**

$$V = (S(T) - K)^+, \qquad \text{European call}$$
$$V = (K - S(T))^+, \qquad \text{European put}$$
$$V = \left( \frac{1}{T} \int_0^T S(u)\; du - K \right)^+, \qquad \text{Asian call}$$
$$V = \max_{0 \leq t \leq T} S(t), \qquad \text{Look back}$$

■

If there is a hedging portfolio, i.e., a process $\Delta(t)$, $0 \leq t \leq T$, whose corresponding wealth process satisfies $X(T) = V$, then

$$X(0) = \widetilde{I\!E}\left[ \frac{V}{\beta(T)} \right].$$

This is because $\frac{X(t)}{\beta(t)}$ is a martingale under $\widetilde{I\!P}$, so

$$X(0) = \frac{X(0)}{\beta(0)} = \widetilde{I\!E}\left[ \frac{X(T)}{\beta(T)} \right] = \widetilde{I\!E}\left[ \frac{V}{\beta(T)} \right].$$

# Chapter 18

# Martingale Representation Theorem

## 18.1 Martingale Representation Theorem

See Oksendal, 4th ed., Theorem 4.11, p.50.

**Theorem 1.56** *Let $B(t), 0 \leq t \leq T$, be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}(t), 0 \leq t \leq T$, be the filtration* generated by this Brownian motion. *Let $X(t), 0 \leq t \leq T$, be a martingale (under $I\!\!P$) relative to this filtration. Then there is an adapted process $\delta(t), 0 \leq t \leq T$, such that*

$$X(t) = X(0) + \int_0^t \delta(u)\; dB(u), \qquad 0 \leq t \leq T.$$

*In particular, the paths of $X$ are continuous.*

**Remark 18.1** We already know that if $X(t)$ is a process satisfying

$$dX(t) = \delta(t)\; dB(t),$$

then $X(t)$ is a martingale. Now we see that if $X(t)$ is a martingale adapted to the filtration generated by the Brownian motion $B(t)$, i.e, the Brownian motion is the only source of randomness in $X$, then

$$dX(t) = \delta(t)\; dB(t)$$

for some $\delta(t)$.

## 18.2 A hedging application

**Homework Problem 4.5.** In the context of Girsanov's Theorem, suppse that $\mathcal{F}(t), 0 \leq t \leq T$, is the filtration generated by the Brownian motion $B$ (under $I\!\!P$). Suppose that $Y$ is a $\widetilde{I\!\!P}$-martingale. Then there is an adapted process $\gamma(t), 0 \leq t \leq T$, such that

$$Y(t) = Y(0) + \int_0^t \gamma(u)\; d\widetilde{B}(u), \qquad 0 \leq t \leq T.$$

$$\begin{aligned}
dS(t) &= \mu(t)S(t)\,dt + \sigma(t)S(t)\,dB(t),\\
\beta(t) &= \exp\left\{\int_0^t r(u)\,du\right\},\\
\theta(t) &= \frac{\mu(t)-r(t)}{\sigma(t)},\\
\widetilde{B}(t) &= \int_0^t \theta(u)\,du + B(t),\\
Z(t) &= \exp\left\{-\int_0^t \theta(u)\,dB(u) - \tfrac{1}{2}\int_0^t \theta^2(u)\,du\right\},\\
\widetilde{I\!P}(A) &= \int_A Z(T)\,dI\!P, \qquad \forall A \in \mathcal{F}.
\end{aligned}$$

Then

$$d\left(\frac{S(t)}{\beta(t)}\right) = \frac{S(t)}{\beta(t)}\sigma(t)\,d\widetilde{B}(t).$$

Let $\Delta(t), 0 \le t \le T,$ be a portfolio process. The corresponding wealth process $X(t)$ satisfies

$$d\left(\frac{X(t)}{\beta(t)}\right) = \Delta(t)\sigma(t)\frac{S(t)}{\beta(t)}\,d\widetilde{B}(t),$$

i.e.,

$$\frac{X(t)}{\beta(t)} = X(0) + \int_0^t \Delta(u)\sigma(u)\frac{S(u)}{\beta(u)}\,d\widetilde{B}(u), \qquad 0 \le t \le T.$$

Let $V$ be an $\mathcal{F}(T)$-measurable random variable, representing the payoff of a contingent claim at time $T$. We want to choose $X(0)$ and $\Delta(t), 0 \le t \le T$, so that

$$X(T) = V.$$

Define the $\widetilde{I\!P}$-martingale

$$Y(t) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T.$$

According to Homework Problem 4.5, there is an adapted process $\gamma(t), 0 \le t \le T$, such that

$$Y(t) = Y(0) + \int_0^t \gamma(u)\,d\widetilde{B}(u), \qquad 0 \le t \le T.$$

Set $X(0) = Y(0) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\right]$ and choose $\Delta(u)$ so that

$$\Delta(u)\sigma(u)\frac{S(u)}{\beta(u)} = \gamma(u).$$

With this choice of $\Delta(u), 0 \le u \le T$, we have

$$\frac{X(t)}{\beta(t)} = Y(t) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T.$$

In particular,

$$\frac{X(T)}{\beta(T)} = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(T)\right] = \frac{V}{\beta(T)},$$

so

$$X(T) = V.$$

The Martingale Representation Theorem guarantees the existence of a hedging portfolio, although it does not tell us how to compute it. It also justifies the risk-neutral pricing formula

$$\begin{aligned} X(t) &= \beta(t)\widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right] \\ &= \frac{\beta(t)}{Z(t)} I\!E\left[\frac{Z(T)}{\beta(T)}V\middle|\mathcal{F}(t)\right] \\ &= \frac{1}{\zeta(t)} I\!E\left[\zeta(T)V\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T, \end{aligned}$$

where

$$\begin{aligned} \zeta(t) &= \frac{Z(t)}{\beta(t)} \\ &= \exp\left\{-\int_0^t \theta(u)\, dB(u) - \int_0^t (r(u) + \tfrac{1}{2}\theta^2(u))\, du\right\} \end{aligned}$$

## 18.3 $d$-dimensional Girsanov Theorem

**Theorem 3.57 ($d$-dimensional Girsanov)** *• $B(t) = (B_1(t), \ldots, B_d(t)), 0 \le t \le T$, a $d$-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$;*

- *$\mathcal{F}(t), 0 \le t \le T$, the accompanying filtration, perhaps larger than the one generated by $B$;*
- *$\theta(t) = (\theta_1(t), \ldots, \theta_d(t)), 0 \le t \le T$, $d$-dimensional adapted process.*

*For $0 \le t \le T$, define*

$$\begin{aligned} \widetilde{B}_j(t) &= \int_0^t \theta_j(u)\, du + B_j(t), \qquad j = 1, \ldots, d, \\ Z(t) &= \exp\left\{-\int_0^t \theta(u).\, dB(u) - \tfrac{1}{2}\int_0^t ||\theta(u)||^2\, du\right\}, \\ \widetilde{I\!P}(A) &= \int_A Z(T)\, dI\!P. \end{aligned}$$

*Then, under $\widetilde{I\!\!P}$, the process*

$$\widetilde{B}(t) = (\widetilde{B}_1(t), \dots, \widetilde{B}_d(t)), \qquad 0 \le t \le T,$$

*is a $d$-dimensional Brownian motion.*

## 18.4 $d$-dimensional Martingale Representation Theorem

**Theorem 4.58**

- $B(t) = (B_1(t), \dots, B_d(t)), 0 \le t \le T$, *a $d$-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$;*
- $\mathcal{F}(t), 0 \le t \le T$, *the filtration* generated by the Brownian motion $B$.

*If $X(t), 0 \le t \le T$, is a martingale (under $I\!\!P$) relative to $\mathcal{F}(t), 0 \le t \le T$, then there is a $d$-dimensional adpated process $\delta(t) = (\delta_1(t), \dots, \delta_d(t))$, such that*

$$X(t) = X(0) + \int_0^t \delta(u).\, dB(u), \qquad 0 \le t \le T.$$

**Corollary 4.59** *If we have a $d$-dimensional adapted process $\theta(t) = (\theta_1(t), \dots, \theta_d(t))$, then we can define $\widetilde{B}, Z$ and $\widetilde{I\!\!P}$ as in Girsanov's Theorem. If $Y(t), 0 \le t \le T$, is a martingale under $\widetilde{I\!\!P}$ relative to $\mathcal{F}(t), 0 \le t \le T$, then there is a $d$-dimensional adpated process $\gamma(t) = (\gamma_1(t), \dots, \gamma_d(t))$ such that*

$$Y(t) = Y(0) + \int_0^t \gamma(u).\, d\widetilde{B}(u), \qquad 0 \le t \le T.$$

## 18.5 Multi-dimensional market model

Let $B(t) = (B_1(t), \dots, B_d(t)),\ 0 \le t \le T$, be a $d$-dimensional Brownian motion on some $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t),\ 0 \le t \le T$, be the *filtration generated by* $B$. Then we can define the following:

**Stocks**

$$dS_i(t) = \mu_i(t) S_i(t)\ dt + S_i(t) \sum_{j=1}^{d} \sigma_{ij}(t)\ dB_j(t), \qquad i = 1, \dots, m$$

**Accumulation factor**

$$\beta(t) = \exp\left\{ \int_0^t r(u)\ du \right\}.$$

Here, $\mu_i(t), \sigma_{ij}(t)$ and $r(t)$ are adpated processes.

**Discounted stock prices**

$$
\begin{aligned}
d\left(\frac{S_i(t)}{\beta(t)}\right) &= \underbrace{(\mu_i(t) - r(t))}_{\text{Risk Premium}} \frac{S_i(t)}{\beta(t)}\, dt + \frac{S_i(t)}{\beta(t)} \sum_{j=1}^{d} \sigma_{ij}(t)\, dB_j(t) \\
&\stackrel{?}{=} \frac{S_i(t)}{\beta(t)} \sum_{j=1}^{d} \sigma_{ij}(t) \underbrace{[\theta_j(t) + dB_j(t)]}_{d\widetilde{B}_j(t)}
\end{aligned} \tag{5.1}
$$

For 5.1 to be satisfied, we need to choose $\theta_1(t), \ldots, \theta_d(t)$, so that

$$
\sum_{j=1}^{d} \sigma_{ij}(t)\theta_j(t) = \mu_i(t) - r(t), \qquad i = 1, \ldots, m. \tag{MPR}
$$

**Market price of risk.** The market price of risk is an adapted process $\theta(t) = (\theta_1(t), \ldots, \theta_d(t))$ satisfying the system of equations (MPR) above. There are three cases to consider:

**Case I:** (Unique Solution). For Lebesgue-almost every $t$ and $I\!\!P$-almost every $\omega$, (MPR) has a *unique solution* $\theta(t)$. Using $\theta(t)$ in the $d$-dimensional Girsanov Theorem, we define a *unique risk-neutral probability measure* $\widetilde{I\!\!P}$. Under $\widetilde{I\!\!P}$, every discounted stock price is a martingale. Consequently, the discounted wealth process corresponding to any portfolio process is a $\widetilde{I\!\!P}$-martingale, and this implies that the market admits no arbitrage. Finally, the Martingale Representation Theorem can be used to show that every contingent claim can be hedged; the market is said to be *complete*.

**Case II:** (No solution.) If (MPR) has no solution, then there is *no risk-neutral probability measure* and the market admits *arbitrage*.

**Case III:** (Multiple solutions). If (MPR) has multiple solutions, then there are *multiple risk-neutral probability measures*. The market admits *no arbitrage*, but there are contingent claims which cannot be hedged; the market is said to be *incomplete.*

**Theorem 5.60 (Fundamental Theorem of Asset Pricing) Part I.** *(Harrison and Pliska, Martingales and Stochastic integrals in the theory of continuous trading,* Stochastic Proc. and Applications 11 *(1981), pp 215-260.):*
*If a market has a risk-neutral probability measure, then it admits no arbitrage.*

**Part II.** *(Harrison and Pliska, A stochastic calculus model of continuous trading: complete markets,* Stochastic Proc. and Applications *15 (1983), pp 313-316):*
*The risk-neutral measure is unique if and only if every contingent claim can be hedged.*

# Chapter 19

# A two-dimensional market model

Let $B(t) = (B_1(t), B_2(t)), 0 \le t \le T$, be a two-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}(t), 0 \le t \le T$, be the filtration generated by $B$.

In what follows, all processes can depend on $t$ and $\omega$, but are adapted to $\mathcal{F}(t), 0 \le t \le T$. To simplify notation, we omit the arguments whenever there is no ambiguity.

**Stocks:**

$$\begin{aligned} dS_1 &= S_1 \left[\mu_1 \; dt + \sigma_1 \; dB_1\right], \\ dS_2 &= S_2 \left[\mu_2 \; dt + \rho\sigma_2 \; dB_1 + \sqrt{1-\rho^2}\,\sigma_2 \; dB_2\right]. \end{aligned}$$

We assume $\sigma_1 > 0, \; \sigma_2 > 0, \; -1 \le \rho \le 1$. Note that

$$\begin{aligned} dS_1 \; dS_2 &= S_1^2\sigma_1^2 \; dB_1 \; dB_1 = \sigma_1^2 S_1^2 \; dt, \\ dS_2 \; dS_2 &= S_2^2\rho^2\sigma_2^2 \; dB_1 \; dB_1 + S_2^2(1-\rho^2)\sigma_2^2 \; dB_2 \; dB_2 \\ &= \sigma_2^2 S_2^2 \; dt, \\ dS_1 \; dS_2 &= S_1\sigma_1 S_2 \rho\sigma_2 \; dB_1 \; dB_1 = \rho\sigma_1\sigma_2 S_1 S_2 \; dt. \end{aligned}$$

In other words,

- $\frac{dS_1}{S_1}$ has instantaneous variance $\sigma_1^2$,
- $\frac{dS_2}{S_2}$ has instantaneous variance $\sigma_2^2$,
- $\frac{dS_1}{S_1}$ and $\frac{dS_2}{S_2}$ have instantaneous covariance $\rho\sigma_1\sigma_2$.

**Accumulation factor:**

$$\beta(t) = \exp\left\{\int_0^t r \; du\right\}.$$

The market price of risk equations are

$$\begin{aligned} \sigma_1\theta_1 &= \mu_1 - r \\ \rho\sigma_2\theta_1 + \sqrt{1-\rho^2}\sigma_2\theta_2 &= \mu_2 - r \end{aligned} \qquad \text{(MPR)}$$

The solution to these equations is

$$\theta_1 = \frac{\mu_1 - r}{\sigma_1},$$
$$\theta_2 = \frac{\sigma_1(\mu_2 - r) - \rho\sigma_2(\mu_1 - r)}{\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

provided $-1 < \rho < 1$.

Suppose $-1 < \rho < 1$. Then (MPR) has a unique solution $(\theta_1, \theta_2)$; we define

$$Z(t) = \exp\left\{-\int_0^t \theta_1 \, dB_1 - \int_0^t \theta_2 \, dB_2 - \tfrac{1}{2}\int_0^t (\theta_1^2 + \theta_2^2)\, du\right\},$$
$$\widetilde{I\!P}(A) = \int_A Z(T)\, dI\!P, \qquad \forall A \in \mathcal{F}.$$

$\widetilde{I\!P}$ is the *unique* risk-neutral measure. Define

$$\widetilde{B}_1(t) = \int_0^t \theta_1 \, du + B_1(t),$$
$$\widetilde{B}_2(t) = \int_0^t \theta_2 \, du + B_2(t).$$

Then

$$dS_1 = S_1\left[r\, dt + \sigma_1 \, d\widetilde{B}_1\right],$$
$$dS_2 = S_2\left[r\, dt + \rho\sigma_2 \, d\widetilde{B}_1 + \sqrt{1-\rho^2}\sigma_2 d\widetilde{B}_2\right].$$

We have changed the mean rates of return of the stock prices, but not the variances and covariances.

## 19.1 Hedging when $-1 < \rho < 1$

$$\begin{aligned}
dX &= \Delta_1 \, dS_1 + \Delta_2 \, dS_2 + r(X - \Delta_1 S_1 - \Delta_2 S_2)\, dt \\
d\left(\frac{X}{\beta}\right) &= \frac{1}{\beta}(dX - rX\, dt) \\
&= \frac{1}{\beta}\Delta_1(dS_1 - rS_1\, dt) + \frac{1}{\beta}\Delta_2(dS_2 - rS_2\, dt) \\
&= \frac{1}{\beta}\Delta_1 S_1\sigma_1 \, d\widetilde{B}_1 + \frac{1}{\beta}\Delta_2 S_2\left[\rho\sigma_2 \, d\widetilde{B}_1 + \sqrt{1-\rho^2}\sigma_2 \, d\widetilde{B}_2\right].
\end{aligned}$$

Let $V$ be $\mathcal{F}(T)$-measurable. Define the $\widetilde{I\!P}$-martingale

$$Y(t) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T.$$

The Martingale Representation Corollary implies

$$Y(t) = Y(0) + \int_0^t \gamma_1 \, d\widetilde{B}_1 + \int_0^t \gamma_2 \, d\widetilde{B}_2.$$

We have

$$\begin{aligned} d\left(\frac{X}{\beta}\right) &= \left(\frac{1}{\beta}\Delta_1 S_1 \sigma_1 + \frac{1}{\beta}\Delta_2 S_2 \rho \sigma_2\right) \, d\widetilde{B}_1 \\ &\quad + \frac{1}{\beta}\Delta_2 S_2 \sqrt{1-\rho^2}\sigma_2 \, d\widetilde{B}_2, \\ dY &= \gamma_1 \, d\widetilde{B}_1 + \gamma_2 \, d\widetilde{B}_2. \end{aligned}$$

We solve the equations

$$\begin{aligned} \frac{1}{\beta}\Delta_1 S_1 \sigma_1 + \frac{1}{\beta}\Delta_2 S_2 \rho \sigma_2 &= \gamma_1 \\ \frac{1}{\beta}\Delta_2 S_2 \sqrt{1-\rho^2}\sigma_2 &= \gamma_2 \end{aligned}$$

for the hedging portfolio $(\Delta_1, \Delta_2)$. With this choice of $(\Delta_1, \Delta_2)$ and setting

$$X(0) = Y(0) = \widetilde{I\!E} \frac{V}{\beta(T)},$$

we have $X(t) = Y(t), \; 0 \le t \le T$, and in particular,

$$X(T) = V.$$

Every $\mathcal{F}(T)$-measurable random variable can be hedged; the market is *complete*.

## 19.2 Hedging when $\rho = 1$

The case $\rho = -1$ is analogous. Assume that $\rho = 1$. Then

$$\begin{aligned} dS_1 &= S_1[\mu_1 \, dt + \sigma_1 \, dB_1] \\ dS_2 &= S_2[\mu_2 \, dt + \sigma_2 \, dB_1] \end{aligned}$$

The stocks are perfectly correlated.

The market price of risk equations are

$$\begin{aligned} \sigma_1 \theta_1 &= \mu_1 - r \\ \sigma_2 \theta_1 &= \mu_2 - r \end{aligned} \tag{MPR}$$

The process $\theta_2$ is free. There are two cases:

**Case I:** $\frac{\mu_1-r}{\sigma_1} \neq \frac{\mu_2-r}{\sigma_2}$. There is no solution to (MPR), and consequently, there is no risk-neutral measure. This market admits arbitrage. Indeed

$$\begin{aligned} d\left(\frac{X}{\beta}\right) &= \frac{1}{\beta}\Delta_1(dS_1 - rS_1\ dt) + \frac{1}{\beta}\Delta_2(dS_2 - rS_2\ dt) \\ &= \frac{1}{\beta}\Delta_1 S_1[(\mu_1 - r)\ dt + \sigma_1\ dB_1] + \frac{1}{\beta}\Delta_2 S_2[(\mu_2 - r)\ dt + \sigma_2\ dB_1] \end{aligned}$$

Suppose $\frac{\mu_1-r}{\sigma_1} > \frac{\mu_2-r}{\sigma_2}$. Set

$$\Delta_1 = \frac{1}{\sigma_1 S_1}, \quad \Delta_2 = -\frac{1}{\sigma_2 S_2}.$$

Then

$$\begin{aligned} d\left(\frac{X}{\beta}\right) &= \frac{1}{\beta}\left[\frac{\mu_1 - r}{\sigma_1}dt + dB_1\right] - \frac{1}{\beta}\left[\frac{\mu_2 - r}{\sigma_2}dt + dB_1\right] \\ &= \frac{1}{\beta}\underbrace{\left[\frac{\mu_1 - r}{\sigma_1} - \frac{\mu_2 - r}{\sigma_2}\right]}_{\text{Positive}} dt \end{aligned}$$

**Case II:** $\frac{\mu_1-r}{\sigma_1} = \frac{\mu_2-r}{\sigma_2}$. The market price of risk equations

$$\begin{aligned} \sigma_1\theta_1 &= \mu_1 - r \\ \sigma_2\theta_1 &= \mu_2 - r \end{aligned}$$

have the solution

$$\theta_1 = \frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2},$$

$\theta_2$ is free; there are infinitely many risk-neutral measures. Let $\widetilde{I\!P}$ be one of them.

**Hedging:**

$$\begin{aligned} d\left(\frac{X}{\beta}\right) &= \frac{1}{\beta}\Delta_1 S_1[(\mu_1 - r)\ dt + \sigma_1\ dB_1] + \frac{1}{\beta}\Delta_2 S_2[(\mu_2 - r)\ dt + \sigma_2\ dB_1] \\ &= \frac{1}{\beta}\Delta_1 S_1\sigma_1[\theta_1\ dt + dB_1] + \frac{1}{\beta}\Delta_2 S_2\sigma_2[\theta_1\ dt + dB_1] \\ &= \left(\frac{1}{\beta}\Delta_1 S_1\sigma_1 + \frac{1}{\beta}\Delta_2 S_2\sigma_2\right)\ d\widetilde{B}_1. \end{aligned}$$

Notice that $\widetilde{B}_2$ does not appear.

Let $V$ be an $\mathcal{F}(T)$-measurable random variable. If $V$ depends on $B_2$, then it can probably not be hedged. For example, if

$$V = h(S_1(T), S_2(T)),$$

and $\sigma_1$ or $\sigma_2$ depend on $B_2$, then there is trouble.

More precisely, we define the $\widetilde{I\!P}$-martingale

$$Y(t) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right], \qquad 0 \le t \le T.$$

We can write

$$Y(t) = Y(0) + \int_0^t \gamma_1 \, d\widetilde{B}_1 + \int_0^t \gamma_2 \, d\widetilde{B}_2,$$

so

$$dY = \gamma_1 \, d\widetilde{B}_1 + \gamma_2 \, d\widetilde{B}_2.$$

To get $d\left(\frac{X}{\beta}\right)$ to match $dY$, we must have

$$\gamma_2 = 0.$$

# Chapter 20

# Pricing Exotic Options

## 20.1 Reflection principle for Brownian motion

**Without drift.**

Define

$$M(T) = \max_{0 \le t \le T} B(t).$$

Then we have:

$$\begin{aligned}
I\!P\{M(T) > m, B(T) < b\} \\
&= I\!P\{B(T) > 2m - b\} \\
&= \frac{1}{\sqrt{2\pi T}} \int_{2m-b}^{\infty} \exp\left\{-\frac{x^2}{2T}\right\} dx, \quad m > 0,\ b < m
\end{aligned}$$

So the joint density is

$$\begin{aligned}
I\!P\{M(T) \in dm, B(T) \in db\} &= -\frac{\partial^2}{\partial m\, \partial b}\left(\frac{1}{\sqrt{2\pi T}} \int_{2m-b}^{\infty} \exp\left\{-\frac{x^2}{2T}\right\}\, dx\right) dm\, db \\
&= -\frac{\partial}{\partial m}\left(\frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(2m-b)^2}{2T}\right\}\right) dm\, db, \\
&= \frac{2(2m-b)}{T\sqrt{2\pi T}} \exp\left\{-\frac{(2m-b)^2}{2T}\right\} dm\, db, \quad m > 0, b < m.
\end{aligned}$$

**With drift.** Let

$$\widetilde{B}(t) = \theta t + B(t),$$

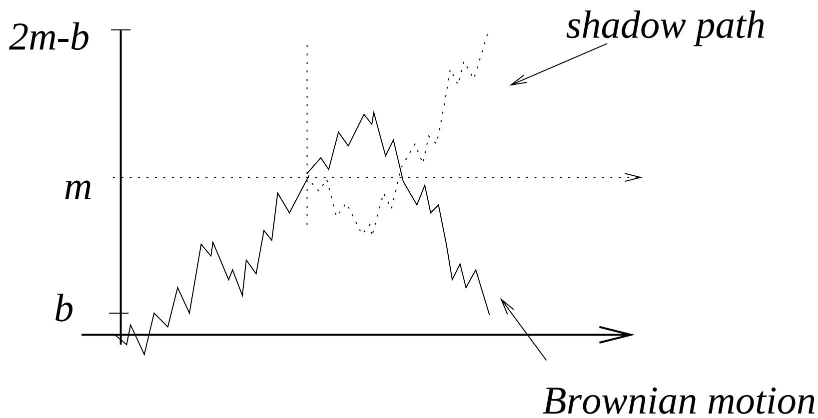



Figure 20.1: *Reflection Principle for Brownian motion without drift*

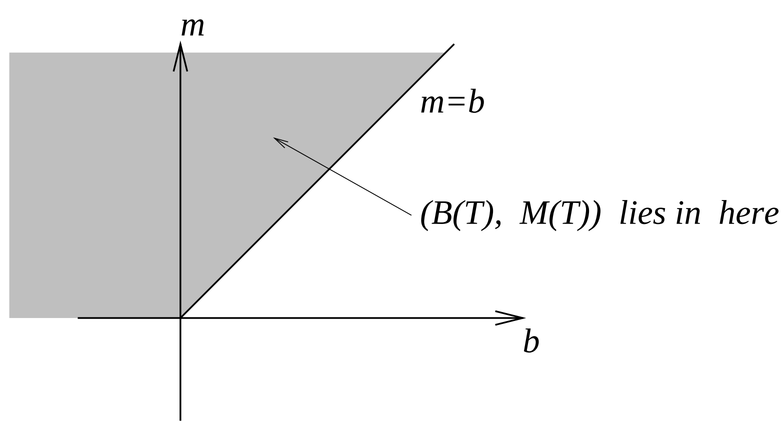



Figure 20.2: *Possible values of* $B(T), M(T)$.

where $B(t),\ 0 \le t \le T$, is a Brownian motion (without drift) on $(\Omega, \mathcal{F}, \mathbb{P})$. Define

$$\begin{aligned} Z(T) &= \exp\{-\theta B(T) - \tfrac{1}{2}\theta^2 T\} \\ &= \exp\{-\theta(B(T) + \theta T) + \tfrac{1}{2}\theta^2 T\} \\ &= \exp\{-\theta \widetilde{B}(t) + \tfrac{1}{2}\theta^2 T\}, \\ \widetilde{I\!P}(A) &= \int_A Z(T)\, dI\!P, \quad \forall A \in \mathcal{F}. \end{aligned}$$

Set$\widetilde{M}(T) = \max_{0 \le t \le T} \widetilde{B}(T)$.

Under $\widetilde{I\!P}$, $\widetilde{B}$ is a Brownian motion (without drift), so

$$\widetilde{I\!P}\{\widetilde{M}(T) \in d\tilde{m},\ \widetilde{B}(T) \in d\tilde{b}\} = \frac{2(2\tilde{m} - \tilde{b})}{T\sqrt{2\pi T}} \exp\left\{-\frac{(2\tilde{m} - \tilde{b})^2}{2T}\right\} d\tilde{m}\, d\tilde{b}, \quad \tilde{m} > 0,\ \tilde{b} < \tilde{m}.$$

Let $h(\tilde{m}, \tilde{b})$ be a function of two variables. Then

$$\begin{aligned} I\!E h(\widetilde{M}(T), \widetilde{B}(T)) &= \widetilde{I\!E} \frac{h(\widetilde{M}(T), \widetilde{B}(T))}{Z(T)} \\ &= \widetilde{I\!E}\left[h(\widetilde{M}(T), \widetilde{B}(T)) \exp\{\theta \widetilde{B}(T) - \tfrac{1}{2}\theta^2 T\}\right] \\ &= \int\limits_{\tilde{m}=0}^{\tilde{m}=\infty} \int\limits_{\tilde{b}=-\infty}^{\tilde{b}=\tilde{m}} h(\tilde{m}, \tilde{b})\ \exp\{\theta\tilde{b} - \tfrac{1}{2}\theta^2 T\}\ \widetilde{I\!P}\{\widetilde{M}(T) \in d\tilde{m},\ \widetilde{B}(T) \in d\tilde{b}\}. \end{aligned}$$

But also,

$$I\!E h(\widetilde{M}(T), \widetilde{B}(T)) = \int\limits_{\tilde{m}=0}^{\tilde{m}=\infty} \int\limits_{\tilde{b}=-\infty}^{\tilde{b}=\tilde{m}} h(\tilde{m}, \tilde{b})\ I\!P\{\widetilde{M}(T) \in d\tilde{m},\ \widetilde{B}(T) \in d\tilde{b}\}.$$

Since $h$ is arbitrary, we conclude that

(MPR)

$$\begin{aligned} &I\!P\{\widetilde{M}(T) \in d\tilde{m},\ \widetilde{B}(T) \in d\tilde{b}\} \\ &= \exp\{\theta\tilde{b} - \tfrac{1}{2}\theta^2 T\}\ \widetilde{I\!P}\{\widetilde{M}(T) \in d\tilde{m},\ \widetilde{B}(T) \in d\tilde{b}\} \\ &= \frac{2(2\tilde{m} - \tilde{b})}{T\sqrt{2\pi T}} \exp\left\{-\frac{(2\tilde{m} - \tilde{b})^2}{2T}\right\} . \exp\{\theta\tilde{b} - \tfrac{1}{2}\theta^2 T\} d\tilde{m}\, d\tilde{b}, \quad \tilde{m} > 0,\ \tilde{b} < \tilde{m}. \end{aligned}$$

## 20.2 Up and out European call.

Let $0 < K < L$ be given. The payoff at time $T$ is

$$(S(T) - K)^+ \mathbf{1}_{\{S^*(T) < L\}},$$

where

$$S^*(T) = \max_{0 \le t \le T} S(t).$$

To simplify notation, assume that $I\!P$ is already the risk-neutral measure, so the value at time zero of the option is

$$v(0, S(0)) = e^{-rT} I\!E\left[(S(T) - K)^+ \mathbf{1}_{\{S^*(T) < L\}}\right].$$

Because $I\!P$ is the risk-neutral measure,

$$\begin{aligned}
dS(t) &= rS(t)\,dt + \sigma S(t)\,dB(t) \\
S(t) &= S_0 \exp\{\sigma B(t) + (r - \tfrac{1}{2}\sigma^2)t\} \\
&= S_0 \exp\left\{\sigma\left[B(t) + \underbrace{\left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)}_{\theta} t\right]\right\} \\
&= S_0 \exp\{\sigma \widetilde{B}(t)\},
\end{aligned}$$

where

$$\begin{aligned}
\theta &= \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right), \\
\widetilde{B}(t) &= \theta t + B(t).
\end{aligned}$$

Consequently,

$$S^*(t) = S_0 \exp\{\sigma \widetilde{M}(t)\},$$

where,

$$\widetilde{M}(t) = \max_{0 \le u \le t} \widetilde{B}(u).$$

We compute,

$$\begin{aligned}
v(0, S(0)) &= e^{-rT} I\!E\left[(S(T) - K)^+ \mathbf{1}_{\{S^*(T) < L\}}\right] \\
&= e^{-rT} I\!E\left[\left(S(0)\exp\{\sigma \widetilde{B}(T)\} - K\right)^+ \mathbf{1}_{\{S(0)\exp\{\sigma \widetilde{M}(T)\} \,<\, L\}}\right] \\
&= e^{-rT} I\!E\left[\left(S(0)\exp\{\sigma \widetilde{B}(T)\} - K\right) \mathbf{1}_{\left\{\widetilde{B}(T) > \underbrace{\frac{1}{\sigma}\log\frac{K}{S(0)}}_{\tilde{b}},\ \widetilde{M}(T) < \underbrace{\frac{1}{\sigma}\log\frac{L}{S(0)}}_{\tilde{m}}\right\}}\right]
\end{aligned}$$

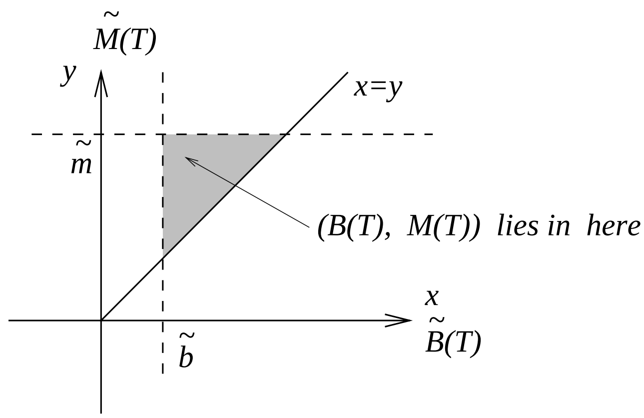



Figure 20.3: *Possible values of* $\widetilde{B}(T), \widetilde{M}(T)$.

We consider only the case

$$S(0) \le K < L, \quad \text{so} \quad 0 \le \tilde{b} < \tilde{m}.$$

The other case, $K < S(0) \le L$ leads to $\tilde{b} < 0 \le \tilde{m}$ and the analysis is similar.

We compute $\int_{\tilde{b}}^{\tilde{m}} \int_x^{\tilde{m}} \dots dy\ dx$:

$$\begin{aligned}
v(0, S(0)) &= e^{-rT} \int_{\tilde{b}}^{\tilde{m}} \int_x^{\tilde{m}} (S(0)\exp\{\sigma x\} - K) \frac{2(2y-x)}{T\sqrt{2\pi T}} \exp\left\{-\frac{(2y-x)^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\} dy\ dx \\
&= -e^{-rT} \int_{\tilde{b}}^{\tilde{m}} (S(0)\exp\{\sigma x\} - K) \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(2y-x)^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\} \Bigg|_{y=x}^{y=\tilde{m}} dx \\
&= e^{-rT} \int_{\tilde{b}}^{\tilde{m}} (S(0)\exp\{\sigma x\} - K) \frac{1}{\sqrt{2\pi T}} \Bigg[\exp\left\{-\frac{x^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\} \\
&\qquad - \exp\left\{-\frac{(2\tilde{m}-x)^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}\Bigg]\ dx \\
&= \frac{1}{\sqrt{2\pi T}} e^{-rT} S(0) \int_{\tilde{b}}^{\tilde{m}} \exp\left\{\sigma x - \frac{x^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}\ dx \\
&\quad - \frac{1}{\sqrt{2\pi T}} e^{-rT} K \int_{\tilde{b}}^{\tilde{m}} \exp\left\{-\frac{x^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}\ dx \\
&\quad - \frac{1}{\sqrt{2\pi T}} e^{-rT} S(0) \int_{\tilde{b}}^{\tilde{m}} \exp\left\{\sigma x - \frac{(2\tilde{m}-x)^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}\ dx \\
&\quad + \frac{1}{\sqrt{2\pi T}} e^{-rT} K \int_{\tilde{b}}^{\tilde{m}} \exp\left\{-\frac{(2\tilde{m}-x)^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}\ dx.
\end{aligned}$$

The standard method for all these integrals is to complete the square in the exponent and then recognize a cumulative normal distribution. We carry out the details for the first integral and just give the result for the other three. The exponent in the first integrand is

$$\begin{aligned}
&\sigma x - \frac{x^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T \\
&= -\frac{1}{2T}(x - \sigma T - \theta T)^2 + \tfrac{1}{2}\sigma^2 T + \sigma\theta T \\
&= -\frac{1}{2T}\left(x - \frac{rT}{\sigma} - \frac{\sigma T}{2}\right)^2 + rT.
\end{aligned}$$

In the first integral we make the change of variable

$$y = (x - rT/\sigma - \sigma T/2)/\sqrt{T}, \quad dy = dx/\sqrt{T},$$

to obtain

$$\begin{aligned}
&\frac{e^{-rT}S(0)}{\sqrt{2\pi T}}\int_{\tilde b}^{\tilde m}\exp\left\{\sigma x - \frac{x^2}{2T} + \theta x - \tfrac{1}{2}\theta^2 T\right\}dx \\
&= \frac{1}{\sqrt{2\pi T}}S(0)\int_{\tilde b}^{\tilde m}\exp\left\{-\frac{1}{2T}\left(x - \frac{rT}{\sigma} - \frac{\sigma T}{2}\right)^2\right\}\,dx \\
&= \frac{1}{\sqrt{2\pi T}}S(0). \int_{\frac{\tilde b}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}}^{\frac{\tilde m}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}} \exp\{-\frac{y^2}{2}\}\,dy \\
&= S(0)\left[N\left(\frac{\tilde m}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right) - N\left(\frac{\tilde b}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right)\right].
\end{aligned}$$

Putting all four integrals together, we have

$$\begin{aligned}
v(0, S(0)) = S(0)&\left[N\left(\frac{\tilde m}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right) - N\left(\frac{\tilde b}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right)\right] \\
&- e^{-rT}K\left[N\left(\frac{\tilde m}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right) - N\left(\frac{\tilde b}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right)\right] \\
&- S(0)\left[N\left(\frac{\tilde m}{\sqrt{T}} + \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right) - N\left(\frac{(2\tilde m - \tilde b)}{\sqrt{T}} + \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right)\right] \\
&+ \exp\left\{-rT + 2\tilde m\left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)\right\}\left[N\left(\frac{\tilde m}{\sqrt{T}} + \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right) - \right. \\
&\quad \left. N\left(\frac{(2\tilde m - \tilde b)}{\sqrt{T}} + \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right)\right],
\end{aligned}$$

where

$$\tilde b = \frac{1}{\sigma}\log\frac{K}{S(0)}, \quad \tilde m = \frac{1}{\sigma}\log\frac{L}{S(0)}.$$

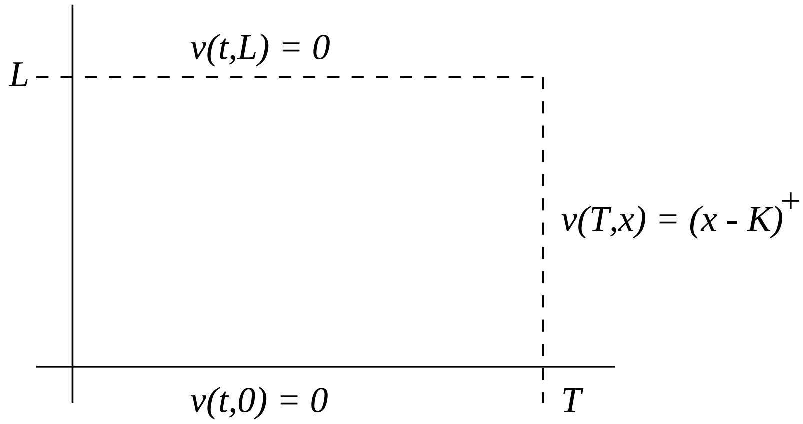



Figure 20.4: *Initial and boundary conditions.*

If we let $L \to \infty$ we obtain the classical Black-Scholes formula

$$\begin{aligned} v(0, S(0)) &= S(0)\left[1 - N\left(\frac{\tilde{b}}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right)\right] \\ &\quad - e^{-rT}K\left[1 - N\left(\frac{\tilde{b}}{\sqrt{T}} - \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right)\right] \\ &= S(0)N\left(\frac{1}{\sigma\sqrt{T}}\log\frac{S(0)}{K} + \frac{r\sqrt{T}}{\sigma} + \frac{\sigma\sqrt{T}}{2}\right) \\ &\quad - e^{-rT}KN\left(\frac{1}{\sigma\sqrt{T}}\log\frac{S(0)}{K} + \frac{r\sqrt{T}}{\sigma} - \frac{\sigma\sqrt{T}}{2}\right). \end{aligned}$$

If we replace $T$ by $T-t$ and replace $S(0)$ by $x$ in the formula for $v(0, S(0))$, we obtain a formula for $v(t,x)$, the value of the option at the time $t$ if $S(t) = x$. We have actually derived the formula under the assumption $x \le K \le L$, but a similar albeit longer formula can also be derived for $K < x \le L$. We consider the function

$$v(t,x) = I\!E^{t,x}\left[e^{-r(T-t)}(S(T)-K)^+ \mathbf{1}_{\{S^*(T)<L\}}\right], \quad 0 \le t \le T,\ 0 \le x \le L.$$

This function satisfies the *terminal condition*

$$v(T,x) = (x-K)^+, \quad 0 \le x < L$$

and the *boundary conditions*

$$\begin{aligned} v(t,0) &= 0, \quad 0 \le t \le T, \\ v(t,L) &= 0, \quad 0 \le t \le T. \end{aligned}$$

We show that $v$ satisfies the Black-Scholes equation

$$-rv + v_t + rxv_x + \tfrac{1}{2}\sigma^2 x^2 v_{xx}, \quad 0 \le t < T,\ 0 \le x \le L.$$

Let $S(0) > 0$ be given and define the *stopping time*

$$\tau = \min\{t \geq 0;\ S(t) = L\}.$$

**Theorem 2.61** *The process*

$$e^{-r(t\wedge\tau)}v(t\wedge\tau, S(t\wedge\tau)), \quad 0 \leq t \leq T,$$

*is a martingale.*

**Proof:** First note that

$$S^*(T) < L \Longleftrightarrow \tau > T.$$

Let $\omega \in \Omega$ be given, and choose $t \in [0,T]$. If $\tau(\omega) \leq t$, then

$$I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(t)\right](\omega) \;=\; 0.$$

But when $\tau(\omega) \leq t$, we have

$$v(t\wedge\tau(\omega), S(t\wedge\tau(\omega),\omega)) = v(t\wedge\tau(\omega), L) = 0,$$

so we may write

$$I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(t)\right](\omega) \;=\; e^{-r(t\wedge\tau(\omega))}v\left(t\wedge\tau(\omega),\ S(t\wedge\tau(\omega),\omega)\right).$$

On the other hand, if $\tau(\omega) > t$, then the Markov property implies

$$\begin{aligned}
&I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(t)\right](\omega)\\
&= I\!E^{t,S(t,\omega)}\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\right]\\
&= e^{-rt}v(t, S(t,\omega))\\
&= e^{-r(t\wedge\tau(\omega))}v\left(t\wedge\tau,\ S(t\wedge\tau(\omega),\omega)\right).
\end{aligned}$$

In both cases, we have

$$e^{-r(t\wedge\tau)}v(t\wedge\tau, S(t\wedge\tau)) = I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(t)\right].$$

Suppose $0 \leq u \leq t \leq T$. Then

$$\begin{aligned}
&I\!E\left[e^{-r(t\wedge\tau)}v(t\wedge\tau, S(t\wedge\tau))\middle|\mathcal{F}(u)\right]\\
&= I\!E\left[I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(t)\right]\middle|\mathcal{F}(u)\right]\\
&= I\!E\left[e^{-rT}(S(T)-K)^+\mathbf{1}_{\{S^*(T)<L\}}\middle|\mathcal{F}(u)\right]\\
&= e^{-r(u\wedge\tau)}v\left(u\wedge\tau,\ S(u\wedge\tau)\right).
\end{aligned}$$

■

For $0 \le t \le T$, we compute the differential

$$d\left(e^{-rt}v(t,S(t))\right) = e^{-rt}(-rv + v_t + rSv_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx})\ dt + e^{-rt}\sigma S v_x\ dB.$$

Integrate from 0 to $t \wedge \tau$:

$$\begin{aligned} e^{-r(t\wedge\tau)}v\left(t\wedge\tau, S(t\wedge\tau)\right) &= v(0,S(0)) \\ &\quad + \int_0^{t\wedge\tau} e^{-ru}(-rv + v_t + rSv_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx})\ du \\ &\quad + \underbrace{\int_0^{t\wedge\tau} e^{-ru}\sigma S v_x\ dB.}_{\text{A stopped martingale is still a martingale}} \end{aligned}$$

Because $e^{-r(t\wedge\tau)}v\left(t\wedge\tau, S(t\wedge\tau)\right)$ is also a martingale, the Riemann integral

$$\int_0^{t\wedge\tau} e^{-ru}(-rv + v_t + rSv_x + \tfrac{1}{2}\sigma^2 S^2 v_{xx})\ du$$

is a martingale. Therefore,

$$-rv(u,S(u)) + v_t(u,S(u)) + rS(u)v_x(u,S(u)) + \tfrac{1}{2}\sigma^2 S^2(u) v_{xx}(u,S(u)) = 0, \quad 0 \le u \le t\wedge\tau.$$

The PDE

$$-rv + v_t + rxv_x + \tfrac{1}{2}\sigma^2 x^2 v_{xx} = 0, \quad 0 \le t \le T,\ 0 \le x \le L,$$

then follows.

**The Hedge**

$$d\left(e^{-rt}v(t,S(t))\right) = e^{-rt}\sigma S(t) v_x(t,S(t))\ dB(t), \quad 0 \le t \le \tau.$$

Let $X(t)$ be the wealth process corresponding to some portfolio $\Delta(t)$. Then

$$d(e^{-rt}X(t)) = e^{-rt}\Delta(t)\sigma S(t)\ dB(t).$$

We should take

$$X(0) = v(0,S(0))$$

and

$$\Delta(t) = v_x(t,S(t)), \quad 0 \le t \le T\wedge\tau.$$

Then

$$\begin{aligned} X(T\wedge\tau) &= v(T\wedge\tau, S(T\wedge\tau)) \\ &= \begin{cases} v(T,S(T)) = (S(T)-K)^+ & \text{if } \tau > T \\ v(\tau,L) = 0 & \text{if } \tau \le T. \end{cases} \end{aligned}$$

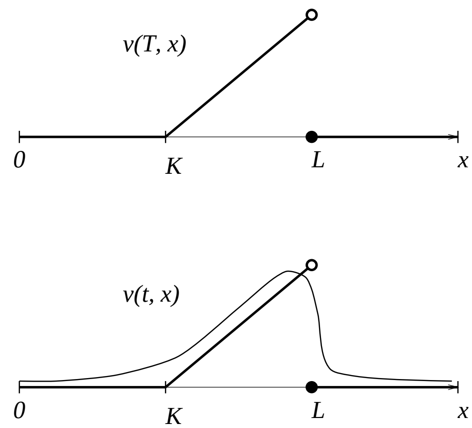



Figure 20.5: *Practial issue.*

## 20.3 A practical issue

For $t < T$ but $t$ near $T$, $v(t,x)$ has the form shown in the bottom part of Fig. 20.5.

In particular, the hedging portfolio

$$\Delta(t) = v_x(t, S(t))$$

can become very negative near the knockout boundary. The hedger is in an unstable situation. He should take a large short position in the stock. If the stock does not cross the barrier $L$, he covers this short position with funds from the money market, pays off the option, and is left with zero. If the stock moves across the barrier, he is now in a region of $\Delta(t) = v_x(t, S(t))$ near zero. He should cover his short position with the money market. This is more expensive than before, because the stock price has risen, and consequently he is left with no money. However, the option has "knocked out", so no money is needed to pay it off.

Because a large short position is being taken, a small error in hedging can create a significant effect. Here is a possible resolution.

Rather than using the boundary condition

$$v(t, L) = 0, \quad 0 \le t \le T,$$

solve the PDE with the boundary condition

$$v(t, L) + \alpha L v_x(t, L) = 0, \quad 0 \le t \le T,$$

where $\alpha$ is a "tolerance parameter", say $1\%$. At the boundary, $Lv_x(t, L)$ is the dollar size of the short position. The new boundary condition guarantees:

1. $Lv_x(t, L)$ remains bounded;
2. The value of the portfolio is always sufficient to cover a hedging error of $\alpha$ times the dollar size of the short position.

# Chapter 21

# Asian Options

Stock:

$$dS(t) = rS(t)\,dt + \sigma S(t)\,dB(t).$$

Payoff:

$$V = h\left(\int_0^T S(t)\,dt\right)$$

Value of the payoff at time zero:

$$X(0) = I\!\!E\left[e^{-rT} h\left(\int_0^T S(t)\,dt\right)\right].$$

Introduce an *auxiliary process* $Y(t)$ by specifying

$$dY(t) = S(t)\,dt.$$

With the initial conditions

$$S(t) = x, \quad Y(t) = y,$$

we have the solutions

$$S(T) = x\exp\left\{\sigma(B(T) - B(t)) + (r - \tfrac{1}{2}\sigma^2)(T - t)\right\},$$
$$Y(T) = y + \int_t^T S(u)\,du.$$

Define the undiscounted expected payoff

$$u(t,x,y) = I\!\!E^{t,x,y} h(Y(T)), \quad 0 \le t \le T,\ x \ge 0,\ y \in I\!\!R.$$

## 21.1 Feynman-Kac Theorem

The function $u$ satisfies the PDE

$$u_t + rxu_x + \tfrac{1}{2}\sigma^2 x^2 u_{xx} + xu_y = 0, \quad 0 \le t \le T,\ x \ge 0,\ y \in I\!R,$$

the terminal condition

$$u(T,x,y) = h(y), \quad x \ge 0,\ y \in I\!R,$$

and the boundary condition

$$u(t,0,y) = h(y), \quad 0 \le t \le T,\ y \in I\!R.$$

One can solve this equation. Then

$$v\left(t, S(t), \int_0^t S(u)\,du\right)$$

is the option value at time $t$, where

$$v(t,x,y) = e^{-r(T-t)} u(t,x,y).$$

The PDE for $v$ is

$$\begin{aligned} -rv + v_t + rxv_x + \tfrac{1}{2}\sigma^2 x^2 v_{xx} + xv_y &= 0, \\ v(T,x,y) &= h(y), \\ v(t,0,y) &= e^{-r(T-t)} h(y). \end{aligned} \tag{1.1}$$

One can solve this equation rather than the equation for $u$.

## 21.2 Constructing the hedge

Start with the stock price $S(0)$. The differential of the value $X(t)$ of a portfolio $\Delta(t)$ is

$$\begin{aligned} dX &= \Delta\,dS + r(X - \Delta S)\,dt \\ &= \Delta S(r\,dt + \sigma\,dB) + rX\,dt - r\Delta S\,dt \\ &= \Delta\sigma S\,dB + rX\,dt. \end{aligned}$$

We want to have

$$X(t) = v\left(t, S(t), \int_0^t S(u)\,du\right),$$

so that

$$\begin{aligned} X(T) &= v\left(T, S(0), \int_0^T S(u)\,du\right), \\ &= h\left(\int_0^T S(u)\,du\right). \end{aligned}$$

The differential of the value of the option is

$$
\begin{aligned}
dv\left(t, S(t), \int_0^t S(u)\, du\right) &= v_t dt + v_x dS + v_y S\, dt + \tfrac{1}{2} v_{xx}\, dS\, dS \\
&= (v_t + rSv_x + Sv_y + \tfrac{1}{2}\sigma^2 S^2 v_{xx})\, dt + \sigma S v_x\, dB \\
&= rv(t, S(t))\, dt + v_x(t, S(t))\, \sigma\, S(t)\, dB(t). \quad \text{(From Eq. 1.1)}
\end{aligned}
$$

Compare this with

$$dX(t) = rX(t)\, dt + \Delta(t)\, \sigma\, S(t)\, dB(t).$$

Take $\Delta(t) = v_x(t, S(t))$. If $X(0) = v(0, S(0), 0)$, then

$$X(t) = v\left(t, S(t), \int_0^t S(u)\, du\right), \quad 0 \le t \le T,$$

because both these processes satisfy the same stochastic differential equation, starting from the same initial condition.

## 21.3 Partial average payoff Asian option

Now suppose the payoff is

$$V = h\left(\int_\tau^T S(t)\, dt\right),$$

where $0 < \tau < T$. We compute

$$v(\tau, x, y) = I\!E^{\tau,x,y} e^{-r(T-\tau)} h(Y(T))$$

just as before. For $0 \le t \le \tau$, we compute next the value of a derivative security which pays off

$$v(\tau, S(\tau), 0)$$

at time $\tau$. This value is

$$w(t, x) = I\!E^{t,x} e^{-r(\tau-t)} v(\tau, S(\tau), 0).$$

The function $w$ satisfies the Black-Scholes PDE

$$-rw + w_t + rxw_x + \tfrac{1}{2}\sigma^2 x^2 w_{xx} = 0, \quad 0 \le t \le \tau,\ x \ge 0,$$

with terminal condition

$$w(\tau, x) = v(\tau, x, 0), \quad x \ge 0,$$

and boundary condition

$$w(t, 0) = e^{-r(T-t)} h(0), \quad 0 \le t \le T.$$

The hedge is given by

$$\Delta(t) = \begin{cases} w_x(t, S(t)), & 0 \le t \le \tau, \\ v_x\left(t, S(t), \int_\tau^t S(u)\, du\right), & \tau < t \le T. \end{cases}$$

**Remark 21.1** While no closed-form for the Asian option price is known, the Laplace transform (in the variable $\frac{\sigma^2}{4}(T - t)$) has been computed. See H. Geman and M. Yor, *Bessel processes, Asian options, and perpetuities,* Math. Finance 3 (1993), 349–375.

# Chapter 22

# Summary of Arbitrage Pricing Theory

A *simple European derivative security* makes a random payment at a time fixed in advance. The *value at time* $t$ of such a security is the amount of wealth needed at time $t$ in order to replicate the security by trading in the market. The *hedging portfolio* is a specification of how to do this trading.

## 22.1 Binomial model, Hedging Portfolio

Let $\Omega$ be the set of all possible sequences of $n$ coin-tosses. We have *no probabilities* at this point. Let $r \geq 0,\ u > r+1,\ d = 1/u$ be given. (See Fig. 2.1)

Evolution of the value of a portfolio:

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)(X_k - \Delta_k S_k).$$

Given a simple European derivative security $V(\omega_1, \omega_2)$, we want to start with a nonrandom $X_0$ and use a portfolio processes

$$\Delta_0,\ \Delta_1(H),\ \Delta_1(T)$$

so that

$$X_2(\omega_1, \omega_2) = V(\omega_1, \omega_2) \quad \forall \omega_1, \omega_2. \qquad \text{(four equations)}$$

There are four unknowns: $X_0$, $\Delta_0$, $\Delta_1(H)$, $\Delta_1(T)$. Solving the equations, we obtain:

$$
\begin{aligned}
X_1(\omega_1) &= \frac{1}{1+r}\left[\frac{1+r-d}{u-d}\underbrace{X_2(\omega_1,H)}_{V(\omega_1,H)} + \frac{u-(1+r)}{u-d}\underbrace{X_2(\omega_1,T)}_{V(\omega_1,T)}\right],\\
X_0 &= \frac{1}{1+r}\left[\frac{1+r-d}{u-d}X_1(H) + \frac{u-(1+r)}{u-d}X_1(T)\right],\\
\Delta_1(\omega_1) &= \frac{X_2(\omega_1,H)-X_2(\omega_1,T)}{S_2(\omega_1,H)-S_2(\omega_1,T)},\\
\Delta_0 &= \frac{X_1(H)-X_1(T)}{S_1(H)-S_1(T)}.
\end{aligned}
$$

The probabilities of the stock price paths are irrelevant, because we have a hedge which works on *every path.* From a practical point of view, what matters is that the paths in the model include all the possibilities. We want to find a description of the paths in the model. They all have the property

$$
\begin{aligned}
(\log S_{k+1} - \log S_k)^2 &= \left(\log \frac{S_{k+1}}{S_k}\right)^2\\
&= (\pm \log u)^2\\
&= (\log u)^2.
\end{aligned}
$$

Let $\sigma = \log u > 0$. Then

$$\sum_{k=0}^{n-1}(\log S_{k+1} - \log S_k)^2 = \sigma^2 n.$$

The paths of $\log S_k$ accumulate quadratic variation at rate $\sigma^2$ per unit time.

If we change $u$, then we change $\sigma$, and the pricing and hedging formulas on the previous page will give different results.

We reiterate that the probabilities are only introduced as an aid to understanding and computation. Recall:

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)(X_k - \Delta_k S_k).$$

Define

$$\beta_k = (1+r)^k.$$

Then

$$\frac{X_{k+1}}{\beta_{k+1}} = \Delta_k \frac{S_{k+1}}{\beta_{k+1}} + \frac{X_k}{\beta_k} - \Delta_k \frac{S_k}{\beta_k},$$

i.e.,

$$\frac{X_{k+1}}{\beta_{k+1}} - \frac{X_k}{\beta_k} = \Delta_k\left(\frac{S_{k+1}}{\beta_{k+1}} - \frac{S_k}{\beta_k}\right).$$

In continuous time, we will have the analogous equation

$$d\left(\frac{X(t)}{\beta(t)}\right) = \Delta(t)\; d\left(\frac{S(t)}{\beta(t)}\right).$$

If we introduce a probability measure $\widetilde{I\!P}$ under which $\frac{S_k}{\beta_k}$ is a martingale, then $\frac{X_k}{\beta_k}$ will also be a martingale, regardless of the portfolio used. Indeed,

$$\begin{aligned}\widetilde{I\!E}\left[\frac{X_{k+1}}{\beta_{k+1}}\middle|\mathcal{F}_k\right] &= \widetilde{I\!E}\left[\frac{X_k}{\beta_k} + \Delta_k\left(\frac{S_{k+1}}{\beta_{k+1}} - \frac{S_k}{\beta_k}\right)\middle|\mathcal{F}_k\right] \\ &= \frac{X_k}{\beta_k} + \Delta_k \underbrace{\left(\widetilde{I\!E}\left[\frac{S_{k+1}}{\beta_{k+1}}\middle|\mathcal{F}_k\right] - \frac{S_k}{\beta_k}\right)}_{=0}.\end{aligned}$$

Suppose we want to have $X_2 = V$, where $V$ is some $\mathcal{F}_2$-measurable random variable. Then we must have

$$\begin{aligned}\frac{1}{1+r}X_1 = \frac{X_1}{\beta_1} &= \widetilde{I\!E}\left[\frac{X_2}{\beta_2}\middle|\mathcal{F}_1\right] = \widetilde{I\!E}\left[\frac{V}{\beta_2}\middle|\mathcal{F}_1\right], \\ X_0 = \frac{X_0}{\beta_0} &= \widetilde{I\!E}\left[\frac{X_1}{\beta_1}\right] = \widetilde{I\!E}\left[\frac{V}{\beta_2}\right].\end{aligned}$$

To find the risk-neutral probability measure $\widetilde{I\!P}$ under which $\frac{S_k}{\beta_k}$ is a martingale, we denote $\tilde{p} = \widetilde{I\!P}\{\omega_k = H\}$, $\tilde{q} = \widetilde{I\!P}\{\omega_k = T\}$, and compute

$$\begin{aligned}\widetilde{I\!E}\left[\frac{S_{k+1}}{\beta_{k+1}}\middle|\mathcal{F}_k\right] &= \tilde{p}u\frac{S_k}{\beta_{k+1}} + \tilde{q}d\frac{S_k}{\beta_{k+1}} \\ &= \frac{1}{1+r}[\tilde{p}u + \tilde{q}d]\frac{S_k}{\beta_k}.\end{aligned}$$

We need to choose $\tilde{p}$ and $\tilde{q}$ so that

$$\begin{aligned}\tilde{p}u + \tilde{q}d &= 1+r, \\ \tilde{p} + \tilde{q} &= 1.\end{aligned}$$

The solution of these equations is

$$\tilde{p} = \frac{1+r-d}{u-d}, \quad \tilde{q} = \frac{u-(1+r)}{u-d}.$$

## 22.2 Setting up the continuous model

Now the stock price $S(t), 0 \le t \le T$, is a continuous function of $t$. We would like to hedge along every possible path of $S(t)$, but that is impossible. Using the binomial model as a guide, we choose $\sigma > 0$ and try to hedge along every path $S(t)$ for which the quadratic variation of $\log S(t)$ accumulates at rate $\sigma^2$ per unit time. These are the paths with volatility $\sigma^2$.

To generate these paths, we use Brownian motion, rather than coin-tossing. To introduce Brownian motion, we need a probability measure. However, the only thing about this probability measure which ultimately matters is the set of paths to which it assigns probability zero.

Let $B(t), 0 \le t \le T$, be a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For any $\rho \in I\!R$, the paths of

$$\rho t + \sigma B(t)$$

accumulate quadratic variation at rate $\sigma^2$ per unit time. We want to define

$$S(t) = S(0) \exp\{\rho t + \sigma B(t)\},$$

so that the paths of

$$\log S(t) = \log S(0) + \rho t + \sigma B(t)$$

accumulate quadratic variation at rate $\sigma^2$ per unit time. Surprisingly, the choice of $\rho$ in this definition is irrelevant. Roughly, the reason for this is the following: Choose $\omega_1 \in \Omega$. Then, for $\rho_1 \in I\!R$,

$$\rho_1 t + \sigma B(t, \omega_1), \quad 0 \le t \le T,$$

is a continuous function of $t$. If we replace $\rho_1$ by $\rho_2$, then $\rho_2 t + \sigma B(t, \omega_1)$ is a different function. However, there is an $\omega_2 \in \Omega$ such that

$$\rho_1 t + \sigma B(t, \omega_1) = \rho_2 t + \sigma B(t, \omega_2), \quad 0 \le t \le T.$$

In other words, regardless of whether we use $\rho_1$ or $\rho_2$ in the definition of $S(t)$, we will see the same paths. The mathematically precise statement is the following:

> If a set of stock price paths has a positive probability when $S(t)$ is defined by
>
> $$S(t) = S(0) \exp\{\rho_1 t + \sigma B(t)\},$$
>
> then this set of paths has positive probability when $S(t)$ is defined by
>
> $$S(t) = S(0) \exp\{\rho_2 t + \sigma B(t)\}.$$
>
> Since we are interested in hedging along every path, except possibly for a set of paths which has probability zero, the choice of $\rho$ is irrelevant.

The most *convenient* choice of $\rho$ is

$$\rho = r - \tfrac{1}{2}\sigma^2,$$

so

$$S(t) = S(0) \exp\{rt + \sigma B(t) - \tfrac{1}{2}\sigma^2 t\},$$

and

$$e^{-rt} S(t) = S(0) \exp\{\sigma B(t) - \tfrac{1}{2}\sigma^2 t\}$$

is a martingale under $I\!P$. With this choice of $\rho$,

$$dS(t) = rS(t)\, dt + \sigma S(t)\, dB(t)$$

and $I\!P$ is the risk-neutral measure. If a different choice of $\rho$ is made, we have

$$\begin{aligned} S(t) &= S(0)\exp\{\rho t + \sigma B(t)\}, \\ dS(t) &= \underbrace{(\rho + \tfrac{1}{2}\sigma^2)}_{\mu} S(t)\ dt + \sigma S(t)\ dB(t). \\ &= rS(t)\ dt + \sigma \underbrace{\left[\tfrac{\mu - r}{\sigma} dt + dB(t)\right]}_{d\widetilde{B}(t)}. \end{aligned}$$

$\widetilde{B}$ has the same paths as $B$. We can change to the risk-neutral measure $\widetilde{I\!P}$, under which $\widetilde{B}$ is a Brownian motion, and then proceed as if $\rho$ had been chosen to be equal to $r - \frac{1}{2}\sigma^2$.

## 22.3 Risk-neutral pricing and hedging

Let $\widetilde{I\!P}$ denote the risk-neutral measure. Then

$$dS(t) = rS(t)\ dt + \sigma S(t)\ d\widetilde{B}(t),$$

where $\widetilde{B}$ is a Brownian motion under $\widetilde{I\!P}$. Set

$$\beta(t) = e^{rt}.$$

Then

$$d\left(\frac{S(t)}{\beta(t)}\right) = \sigma \frac{S(t)}{\beta(t)} d\widetilde{B}(t),$$

so $\frac{S(t)}{\beta(t)}$ is a martingale under $\widetilde{I\!P}$.

Evolution of the value of a portfolio:

$$dX(t) = \Delta(t) dS(t) + r(X(t) - \Delta(t)S(t))\ dt, \tag{3.1}$$

which is equivalent to

$$\begin{aligned} d\left(\frac{X(t)}{\beta(t)}\right) &= \Delta(t) d\left(\frac{S(t)}{\beta(t)}\right) \\ &= \Delta(t)\sigma \frac{S(t)}{\beta(t)} d\widetilde{B}(t). \end{aligned} \tag{3.2}$$

Regardless of the portfolio used, $\frac{X(t)}{\beta(t)}$ is a martingale under $\widetilde{I\!P}$.

Now suppose $V$ is a given $\mathcal{F}(T)$-measurable random variable, the payoff of a simple European derivative security. We want to find the portfolio process $\Delta(T), 0 \le t \le T$, and initial portfolio value $X(0)$ so that $X(T) = V$. Because $\frac{X(t)}{\beta(t)}$ must be a martingale, we must have

$$\frac{X(t)}{\beta(t)} = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right], \quad 0 \le t \le T. \tag{3.3}$$

This is the *risk-neutral pricing formula.* We have the following sequence:

1. $V$ is given,
2. Define $X(t), 0 \le t \le T$, by (3.3) (not by (3.1) or (3.2), because we do not yet have $\Delta(t)$).
3. Construct $\Delta(t)$ so that (3.2) (or equivalently, (3.1)) is satisfied by the $X(t), 0 \le t \le T$, defined in step 2.

To carry out step 3, we first use the tower property to show that $\frac{X(t)}{\beta(t)}$ defined by (3.3) is a martingale under $\widetilde{I\!P}$. We next use the corollary to the Martingale Representation Theorem (Homework Problem 4.5) to show that

$$d\left(\frac{X(t)}{\beta(t)}\right) = \gamma(t)\, d\widetilde{B}(t) \tag{3.4}$$

for some proecss $\gamma$. Comparing (3.4), which we know, and (3.2), which we want, we decide to define

$$\Delta(t) = \frac{\beta(t)\gamma(t)}{\sigma S(t)}. \tag{3.5}$$

Then (3.4) implies (3.2), which implies (3.1), which implies that $X(t), 0 \le t \le T$, is the value of the portfolio process $\Delta(t), 0 \le t \le T$.

From (3.3), the definition of $X$, we see that the hedging portfolio must begin with value

$$X(0) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\right],$$

and it will end with value

$$X(T) = \beta(T)\widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(T)\right] = \beta(T)\frac{V}{\beta(T)} = V.$$

**Remark 22.1** Although we have taken $r$ and $\sigma$ to be constant, the risk-neutral pricing formula is still "valid" when $r$ and $\sigma$ are processes adapted to the filtration generated by $B$. If they depend on either $\widetilde{B}$ or on $S$, they are adapted to the filtration generated by $B$. The "validity" of the risk-neutral pricing formula means:

1. If you start with
$$X(0) = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\right],$$
then there is a hedging portfolio $\Delta(t), 0 \le t \le T$, such that $X(T) = V$;
2. At each time $t$, the value $X(t)$ of the hedging portfolio in 1 satisfies
$$\frac{X(t)}{\beta(t)} = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\middle|\mathcal{F}(t)\right].$$

**Remark 22.2** In general, when there are multiple assets and/or multiple Brownian motions, the risk-neutral pricing formula is valid provided there is a *unique risk-neutral measure.* A probability measure is said to be risk-neutral provided

- it has the same probability-zero sets as the original measure;
- it makes all the discounted asset prices be martingales.

To see if the risk-neutral measure is unique, compute the differential of all discounted asset prices and check if there is more than one way to define $\widetilde{B}$ so that all these differentials have only $d\widetilde{B}$ terms.

## 22.4 Implementation of risk-neutral pricing and hedging

To get a computable result from the general risk-neutral pricing formula

$$\frac{X(t)}{\beta(t)} = \widetilde{I\!E}\left[\frac{V}{\beta(T)}\bigg|\mathcal{F}(t)\right],$$

one uses the Markov property. We need to identify some *state variables*, the stock price and possibly other variables, so that

$$X(t) = \beta(t)\widetilde{I\!E}\left[\frac{V}{\beta(T)}\bigg|\mathcal{F}(t)\right]$$

is a function of these variables.

**Example 22.1** Assume $r$ and $\sigma$ are constant, and $V = h(S(T))$. We can take the stock price to be the state variable. Define

$$v(t,x) = \widetilde{I\!E}^{t,x}\left[e^{-r(T-t)}h(S(T))\right].$$

Then

$$\begin{aligned} X(t) &= e^{rt}\,\widetilde{I\!E}\left[e^{-rT}h(S(T))\Big|\mathcal{F}(t)\right] \\ &= v(t,S(t)), \end{aligned}$$

and $\frac{X(t)}{\beta(t)} = e^{-rt}v(t,S(t))$ is a martingale under $\widetilde{I\!P}$. ■

**Example 22.2** Assume $r$ and $\sigma$ are constant.

$$V = h\left(\int_0^T S(u)\,du\right).$$

Take $S(t)$ and $Y(t) = \int_0^t S(u)\,du$ to be the state variables. Define

$$v(t,x,y) = \widetilde{I\!E}^{t,x,y}\left[e^{-r(T-t)}h(Y(T))\right],$$

where

$$Y(T) = y + \int_t^T S(u)\,du.$$

Then

$$
\begin{aligned}
X(t) &= e^{rt}\ \widetilde{\mathbb{E}}\left[e^{-rT}h(S(T))\Big|\mathcal{F}(t)\right] \\
&= v(t, S(t), Y(t))
\end{aligned}
$$

and

$$
\frac{X(t)}{\beta(t)} = e^{-rt} v(t, S(t), Y(t))
$$

is a martingale under $\widetilde{\mathbb{P}}$. ■

**Example 22.3** (Homework problem 4.2)

$$
\begin{aligned}
dS(t) &= r(t, Y(t))\ S(t)dt + \sigma(t, Y(t))S(t)\ d\widetilde{B}(t), \\
dY(t) &= \alpha(t, Y(t))\ dt + \gamma(t, Y(t))\ d\widetilde{B}(t), \\
V &= h(S(T)).
\end{aligned}
$$

Take $S(t)$ and $Y(t)$ to be the state variables. Define

$$
v(t, x, y) = \widetilde{\mathbb{E}}^{t,x,y}\left[\underbrace{\exp\left\{-\int_t^T r(u, Y(u))\ du\right\}}_{\frac{\beta(t)}{\beta(T)}} h(S(T))\right].
$$

Then

$$
\begin{aligned}
X(t) &= \beta(t)\widetilde{\mathbb{E}}\left[\frac{h(S(T))}{\beta(T)}\Big|\mathcal{F}(t)\right] \\
&= \widetilde{\mathbb{E}}\left[\exp\left\{-\int_t^T r(u, Y(u))\ du\right\} h(S(T))\Big|\mathcal{F}(t)\right] \\
&= v(t, S(t), Y(t)),
\end{aligned}
$$

and

$$
\frac{X(t)}{\beta(t)} = \exp\left\{-\int_0^t r(u, Y(u))\ du\right\} v(t, S(t), Y(t))
$$

is a martingale under $\widetilde{\mathbb{P}}$. ■

In every case, we get an expression involving $v$ to be a martingale. We take the differential and set the $dt$ term to zero. This gives us a partial differential equation for $v$, and this equation must hold wherever the state processes can be. The $d\widetilde{B}$ term in the differential of the equation is the differential of a martingale, and since the martingale is

$$
\frac{X(t)}{\beta(t)} = X(0) + \int_0^t \Delta(u)\sigma\frac{S(u)}{\beta(u)}d\widetilde{B}(u)
$$

we can solve for $\Delta(t)$. This is the argument which uses (3.4) to obtain (3.5).

**Example 22.4 (Continuation of Example 22.3)**

$$\frac{X(t)}{\beta(t)} = \underbrace{\exp\left\{-\int_0^t r(u, Y(u))\, du\right\}}_{1/\beta(t)} v(t, S(t), Y(t))$$

is a martingale under $\widetilde{I\!P}$. We have

$$\begin{aligned}
d\left(\frac{X(t)}{\beta(t)}\right) &= \frac{1}{\beta(t)}\Big[-r(t, Y(t))v(t, S(t), Y(t))\, dt \\
&\qquad + v_t\, dt + v_x\, dS + v_y\, dY \\
&\qquad + \tfrac{1}{2}v_{xx}\, dS\, dS + v_{xy}\, dS\, dY + \tfrac{1}{2}v_{yy}\, dY\, dY\Big] \\
&= \frac{1}{\beta(t)}\Big[(-rv + v_t + rSv_x + \alpha v_y + \tfrac{1}{2}\sigma^2 S^2 v_{xx} + \sigma\gamma S v_{xy} + \tfrac{1}{2}\gamma^2 v_{yy})\, dt \\
&\qquad + (\sigma S v_x + \gamma v_y)\, d\widetilde{B}\Big]
\end{aligned}$$

The partial differential equation satisfied by $v$ is

$$-rv + v_t + rxv_x + \alpha v_y + \tfrac{1}{2}\sigma^2 x^2 v_{xx} + \sigma\gamma x v_{xy} + \tfrac{1}{2}\gamma^2 v_{yy} = 0$$

where it should be noted that $v = v(t, x, y)$, and all other variables are functions of $(t, y)$. We have

$$d\left(\frac{X(t)}{\beta(t)}\right) = \frac{1}{\beta(t)}[\sigma S v_x + \gamma v_y]\, d\widetilde{B}(t),$$

where $\sigma = \sigma(t, Y(t))$, $\gamma = \gamma(t, Y(t))$, $v = v(t, S(t), Y(t))$, and $S = S(t)$. We want to choose $\Delta(t)$ so that (see (3.2))

$$d\left(\frac{X(t)}{\beta(t)}\right) = \Delta(t)\sigma(t, Y(t))\frac{S(t)}{\beta(t)}\, d\widetilde{B}(t).$$

Therefore, we should take $\Delta(t)$ to be

$$\Delta(t) = v_x(t, S(t), Y(t)) + \frac{\gamma(t, Y(t))}{\sigma(t, Y(t))\, S(t)} v_y(t, S(t), Y(t)).$$

■

# Chapter 23

# Recognizing a Brownian Motion

**Theorem 0.62 (Levy)** *Let* $B(t), 0 \leq t \leq T$, *be a process on* $(\Omega, \mathcal{F}, \mathbb{P})$, *adapted to a filtration* $\mathcal{F}(t), 0 \leq t \leq T$, *such that:*

1. *the paths of* $B(t)$ *are continuous,*
2. $B$ *is a martingale,*
3. $\langle B \rangle(t) = t, 0 \leq t \leq T$, *(i.e., informally* $dB(t)\ dB(t) = dt$*).*

*Then* $B$ *is a Brownian motion.*

**Proof:** (Idea) Let $0 \leq s < t \leq T$ be given. We need to show that $B(t) - B(s)$ is normal, with mean zero and variance $t - s$, and $B(t) - B(s)$ is independent of $\mathcal{F}(s)$. We shall show that the *conditional moment generating function* of $B(t) - B(s)$ is

$$\mathbb{E}\left[ e^{u(B(t)-B(s))} \Big| \mathcal{F}(s) \right] = e^{\frac{1}{2}u^2(t-s)}.$$

Since the moment generating function characterizes the distribution, this shows that $B(t) - B(s)$ is normal with mean 0 and variance $t - s$, and conditioning on $\mathcal{F}(s)$ does not affect this, i.e., $B(t) - B(s)$ is independent of $\mathcal{F}(s)$.

We compute (this uses the continuity condition (1) of the theorem)

$$de^{uB(t)} = ue^{uB(t)} dB(t) + \tfrac{1}{2}u^2 e^{uB(t)} dB(t)\ dB(t),$$

so

$$e^{uB(t)} = e^{uB(s)} + \int_s^t ue^{uB(v)}\ dB(v) + \tfrac{1}{2}u^2 \int_s^t e^{uB(v)} \underbrace{dv.}_{\text{uses cond. 3}}$$

Now $\int_0^t ue^{uB(v)}dB(v)$ is a martingale (by condition 2), and so

$$\begin{aligned}
&\mathbb{E}\left[\int_s^t ue^{uB(v)}dB(v)\middle|\mathcal{F}(s)\right]\\
&= -\int_0^s ue^{uB(v)}dB(v) + \mathbb{E}\left[\int_0^t ue^{uB(v)}dB(v)\middle|\mathcal{F}(s)\right]\\
&= 0.
\end{aligned}$$

It follows that

$$\mathbb{E}\left[e^{uB(t)}\middle|\mathcal{F}(s)\right] = e^{uB(s)} + \tfrac{1}{2}u^2\int_s^t \mathbb{E}\left[e^{uB(v)}\middle|\mathcal{F}(s)\right]\,dv.$$

We define

$$\varphi(v) = \mathbb{E}\left[e^{uB(v)}\middle|\mathcal{F}(s)\right],$$

so that

$$\varphi(s) = e^{uB(s)}$$

and

$$\begin{aligned}
\varphi(t) &= e^{uB(s)} + \tfrac{1}{2}u^2\int_s^t \varphi(v)\,dv,\\
\varphi'(t) &= \tfrac{1}{2}u^2\varphi(t),\\
\varphi(t) &= ke^{\frac{1}{2}u^2 t}.
\end{aligned}$$

Plugging in $s$, we get

$$e^{uB(s)} = ke^{\frac{1}{2}u^2 s} \Longrightarrow k = e^{uB(s)-\frac{1}{2}u^2 s}.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[e^{uB(t)}\middle|\mathcal{F}(s)\right] &= \varphi(t) = e^{uB(s)+\frac{1}{2}u^2(t-s)},\\
\mathbb{E}\left[e^{u(B(t)-B(s))}\middle|\mathcal{F}(s)\right] &= e^{\frac{1}{2}u^2(t-s)}.
\end{aligned}$$

■

## 23.1 Identifying volatility and correlation

Let $B_1$ and $B_2$ be independent Brownian motions and

$$\frac{dS_1}{S_1} = r\ dt + \sigma_{11}\ dB_1 + \sigma_{12}\ dB_2,$$
$$\frac{dS_2}{S_2} = r\ dt + \sigma_{21}\ dB_1 + \sigma_{22}\ dB_2,$$

Define

$$\sigma_1 = \sqrt{\sigma_{11}^2 + \sigma_{12}^2},$$
$$\sigma_2 = \sqrt{\sigma_{21}^2 + \sigma_{22}^2},$$
$$\rho = \frac{\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}}{\sigma_1\sigma_2}.$$

Define processes $W_1$ and $W_2$ by

$$dW_1 = \frac{\sigma_{11}\ dB_1 + \sigma_{12}\ dB_2}{\sigma_1}$$
$$dW_2 = \frac{\sigma_{21}\ dB_1 + \sigma_{22}\ dB_2}{\sigma_2}.$$

Then $W_1$ and $W_2$ have continuous paths, are martingales, and

$$\begin{aligned} dW_1\ dW_1 &= \frac{1}{\sigma_1^2}(\sigma_{11}dB_1 + \sigma_{12}dB_2)^2 \\ &= \frac{1}{\sigma_1^2}(\sigma_{11}^2 dB_1\ dB_1 + \sigma_{12}^2 dB_2\ dB_2) \\ &= dt, \end{aligned}$$

and similarly

$$dW_2\ dW_2 = dt.$$

Therefore, $W_1$ and $W_2$ are Brownian motions. The stock prices have the representation

$$\frac{dS_1}{S_1} = r\ dt + \sigma_1\ dW_1,$$
$$\frac{dS_2}{S_2} = r\ dt + \sigma_2\ dW_2.$$

The Brownian motions $W_1$ and $W_2$ are correlated. Indeed,

$$\begin{aligned} dW_1\ dW_2 &= \frac{1}{\sigma_1\sigma_2}(\sigma_{11}dB_1 + \sigma_{12}dB_2)(\sigma_{21}dB_1 + \sigma_{22}dB_2) \\ &= \frac{1}{\sigma_1\sigma_2}(\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})\ dt \\ &= \rho\ dt. \end{aligned}$$

## 23.2 Reversing the process

Suppose we are given that

$$\frac{dS_1}{S_1} = r\ dt + \sigma_1 dW_1,$$
$$\frac{dS_2}{S_2} = r\ dt + \sigma_2 dW_2,$$

where $W_1$ and $W_2$ are Brownian motions with correlation coefficient $\rho$. We want to find

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

so that

$$\begin{aligned}
\Sigma\Sigma' &= \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \\
&= \begin{bmatrix} \sigma_{11}^2 + \sigma_{12}^2 & \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} \\ \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} & \sigma_{21}^2 + \sigma_{22}^2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}
\end{aligned}$$

A simple (but not unique) solution is (see Chapter 19)

$$\sigma_{11} = \sigma_1, \qquad \sigma_{12} = 0,$$
$$\sigma_{21} = \rho\sigma_2, \qquad \sigma_{22} = \sqrt{1-\rho^2}\ \sigma_2.$$

This corresponds to

$$\begin{aligned}
\sigma_1\ dW_1 &= \sigma_1 dB_1 \Longrightarrow dB_1 = dW_1, \\
\sigma_2\ dW_2 &= \rho\sigma_2\ dB_1 + \sqrt{1-\rho^2}\sigma_2\ dB_2 \\
&\Longrightarrow dB_2 = \frac{dW_2 - \rho\ dW_1}{\sqrt{1-\rho^2}}, \quad (\rho \neq \pm 1)
\end{aligned}$$

If $\rho = \pm 1$, then there is no $B_2$ and $dW_2 = \rho\ dB_1 = \rho\ dW_1$.

Continuing in the case $\rho \neq \pm 1$, we have

$$\begin{aligned}
dB_1\ dB_1 &= dW_1\ dW_1 = dt, \\
dB_2\ dB_2 &= \frac{1}{1-\rho^2}\Big(dW_2\ dW_2 - 2\rho\ dW_1\ dW_2 + \rho^2 dW_2\ dW_2\Big) \\
&= \frac{1}{1-\rho^2}\Big(dt - 2\rho^2\ dt + \rho^2\ dt\Big) \\
&= dt,
\end{aligned}$$

so both $B_1$ and $B_2$ are Brownian motions. Furthermore,

$$
\begin{aligned}
dB_1\ dB_2 &= \frac{1}{\sqrt{1-\rho^2}}\left(dW_1\ dW_2 - \rho dW_1\ dW_1\right) \\
&= \frac{1}{\sqrt{1-\rho^2}}\left(\rho\ dt - \rho\ dt\right) = 0.
\end{aligned}
$$

We can now apply an **Extension of Levy's Theorem** that says that Brownian motions with zero cross-variation are independent, to conclude that $B_1, B_2$ are independent Brownians.

# Chapter 24

# An outside barrier option

Barrier process:

$$\frac{dY(t)}{Y(t)} = \lambda\ dt + \sigma_1\ dB_1(t).$$

Stock process:

$$\frac{dS(t)}{S(t)} = \mu\ dt + \rho\sigma_2\ dB_1(t) + \sqrt{1-\rho^2}\,\sigma_2\ dB_2(t),$$

where $\sigma_1 > 0,\ \sigma_2 > 0,\ -1 < \rho < 1$, and $B_1$ and $B_2$ are independent Brownian motions on some $(\Omega, \mathcal{F}, \mathbb{P})$. The option pays off:

$$(S(T) - K)^+ \mathbf{1}_{\{Y^*(T) < L\}}$$

at time $T$, where

$$0 < S(0) < K, \quad 0 < Y(0) < L,$$

$$Y^*(T) = \max_{0 \le t \le T} Y(t).$$

**Remark 24.1** The option payoff depends on both the $Y$ and $S$ processes. In order to hedge it, we will need the money market and two other assets, which we take to be $Y$ and $S$. The risk-neutral measure must make the discounted value of every traded asset be a martingale, which in this case means the discounted $Y$ and $S$ processes.

We want to find $\theta_1$ and $\theta_2$ and define

$$d\widetilde{B}_1 = \theta_1\ dt + dB_1, \quad d\widetilde{B}_2 = \theta_2\ dt + dB_2,$$

so that

$$
\begin{aligned}
\frac{dY}{Y} &= r\ dt + \sigma_1 d\widetilde{B}_1 \\
&= r\ dt + \sigma_1\theta_1\ dt + \sigma_1\ dB_1, \\
\frac{dS}{S} &= r\ dt + \rho\sigma_2\ d\widetilde{B}_1 + \sqrt{1-\rho^2}\ \sigma_2 d\widetilde{B}_2 \\
&= r\ dt + \rho\sigma_2\ \theta_1\ dt + \sqrt{1-\rho^2}\ \sigma_2\theta_2\ dt \\
&\quad + \rho\sigma_2\ dB_1 + \sqrt{1-\rho^2}\ \sigma_2\ dB_2.
\end{aligned}
$$

We must have

$$
\lambda = r + \sigma_1\theta_1, \tag{0.1}
$$

$$
\mu = r + \rho\sigma_2\theta_1 + \sqrt{1-\rho^2}\ \sigma_2\theta_2. \tag{0.2}
$$

We solve to get

$$
\begin{aligned}
\theta_1 &= \frac{\lambda - r}{\sigma_1}, \\
\theta_2 &= \frac{\mu - r - \rho\sigma_2\theta_1}{\sqrt{1-\rho^2}\ \sigma_2}.
\end{aligned}
$$

We shall see that the formulas for $\theta_1$ and $\theta_2$ do not matter. What matters is that (0.1) and (0.2) uniquely determine $\theta_1$ and $\theta_2$. This implies the existence and uniqueness of the risk-neutral measure. We define

$$
\begin{aligned}
Z(T) &= \exp\left\{-\theta_1 B_1(T) - \theta_2 B_2(T) - \tfrac{1}{2}(\theta_1^2 + \theta_2^2)T\right\}, \\
\widetilde{I\!P}(A) &= \int_A Z(T)\ dI\!P, \quad \forall A \in \mathcal{F}.
\end{aligned}
$$

Under $\widetilde{I\!P}$, $\widetilde{B}_1$ and $\widetilde{B}_2$ are independent Brownian motions (Girsanov's Theorem). $\widetilde{I\!P}$ is the unique risk-neutral measure.

**Remark 24.2** Under both $I\!P$ and $\widetilde{I\!P}$, $Y$ has volatility $\sigma_1$, $S$ has volatility $\sigma_2$ and

$$
\frac{dY\ dS}{YS} = \rho\sigma_1\sigma_2\ dt,
$$

i.e., the correlation between $\frac{dY}{Y}$ and $\frac{dS}{S}$ is $\rho$.

The value of the option at time zero is

$$
v(0, S(0), Y(0)) = \widetilde{I\!E}\left[e^{-rT}(S(T)-K)^{+}\mathbf{1}_{\{Y^*(T)<L\}}\right].
$$

We need to work out a density which permits us to compute the right-hand side.

Recall that the *barrier process* is

$$\frac{dY}{Y} = r\ dt + \sigma_1\ d\widetilde{B}_1,$$

so

$$Y(t) = Y(0)\exp\left\{rt + \sigma_1\widetilde{B}_1(t) - \tfrac{1}{2}\sigma_1^2 t\right\}.$$

Set

$$\begin{aligned}\widehat{\theta} &= r/\sigma_1 - \sigma_1/2,\\ \widehat{B}(t) &= \widehat{\theta}t + \widetilde{B}_1(t),\\ \widehat{M}(T) &= \max_{0\le t\le T}\widehat{B}(t).\end{aligned}$$

Then

$$\begin{aligned}Y(t) &= Y(0)\exp\{\sigma_1\widehat{B}(t)\},\\ Y^*(T) &= Y(0)\exp\{\sigma_1\widehat{M}(T)\}.\end{aligned}$$

The joint density of $\widehat{B}(T)$ and $\widehat{M}(T)$, appearing in Chapter 20, is

$$\begin{aligned}&\widetilde{I\!P}\{\widehat{B}(T)\in d\hat{b}, \widehat{M}(T)\in d\hat{m}\}\\ &= \frac{2(2\hat{m}-\hat{b})}{T\sqrt{2\pi T}}\exp\left\{-\frac{(2\hat{m}-\hat{b})^2}{2T} + \widehat{\theta}\hat{b} - \tfrac{1}{2}\widehat{\theta}^2 T\right\}\ d\hat{b}\ d\hat{m},\\ &\quad \hat{m} > 0, \hat{b} < \hat{m}.\end{aligned}$$

*The stock process.*

$$\frac{dS}{S} = r\ dt + \rho\sigma_2 d\widetilde{B}_1 + \sqrt{1-\rho^2}\ \sigma_2 d\widetilde{B}_2,$$

so

$$\begin{aligned}S(T) &= S(0)\exp\{rT + \rho\sigma_2\widetilde{B}_1(T) - \tfrac{1}{2}\rho^2\sigma_2^2 T + \sqrt{1-\rho^2}\ \sigma_2\widetilde{B}_2(T) - \tfrac{1}{2}(1-\rho^2)\sigma_2^2 T\}\\ &= S(0)\exp\{rT - \tfrac{1}{2}\sigma_2^2 T + \rho\sigma_2\widetilde{B}_1(T) + \sqrt{1-\rho^2}\ \sigma_2\widetilde{B}_2(T)\}\end{aligned}$$

From the above paragraph we have

$$\widetilde{B}_1(T) = -\widehat{\theta}T + \widehat{B}(T),$$

so

$$S(T) = S(0)\exp\{rT + \rho\sigma_2\widehat{B}(T) - \tfrac{1}{2}\sigma_2^2 T - \rho\sigma_2\widehat{\theta}T + \sqrt{1-\rho^2}\ \sigma_2\widetilde{B}_2(T)\}$$

## 24.1 Computing the option value

$$
\begin{aligned}
v(0,S(0),Y(0)) &= \widetilde{I\!E}\left[e^{-rT}(S(T)-K)^{+}\mathbf{1}_{\{Y^*(T)<L\}}\right] \\
&= e^{-rT}\widetilde{I\!E}\Bigg[\left(S(0)\exp\left\{(r-\tfrac{1}{2}\sigma_2^2-\rho\sigma_2\widehat{\theta})T+\rho\sigma_2\widehat{B}(T)+\sqrt{1-\rho^2}\,\sigma_2\widetilde{B}_2(T)\right\}-K\right)^{+} \\
&\quad .\mathbf{1}_{\{Y(0)\exp[\sigma_1\widehat{M}(T)]<L\}}\Bigg]
\end{aligned}
$$

We know the joint density of $(\widehat{B}(T),\widehat{M}(T))$. The density of $\widetilde{B}_2(T)$ is

$$
\widetilde{I\!P}\{\widetilde{B}_2(T)\in d\tilde{b}\} = \frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{\tilde{b}^2}{2T}\right\}d\tilde{b}, \quad \tilde{b}\in I\!R.
$$

Furthermore, the pair of random variables $(\widehat{B}(T),\widehat{M}(T))$ is *independent* of $\widetilde{B}_2(T)$ because $\widetilde{B}_1$ and $\widetilde{B}_2$ are independent under $\widetilde{I\!P}$. Therefore, the joint density of the random vector $(\widetilde{B}_2(T),\widehat{B}(T),\widehat{M}(T))$ is

$$
\widetilde{I\!P}\{\widetilde{B}_2(T)\in d\tilde{b},\widehat{B}(T)\in d\hat{b},\widehat{M}(T)\in d\hat{m},\} = \widetilde{I\!P}\{\widetilde{B}_2(T)\in d\tilde{b}\}.\widetilde{I\!P}\{\widehat{B}(T)\in d\hat{b},\widehat{M}(T)\in d\hat{m}\}
$$

The option value at time zero is

$$
\begin{aligned}
&v(0,S(0),Y(0)) \\
&= e^{-rT}\int_0^{\frac{1}{\sigma_1}\log\frac{L}{Y(0)}}\int_{-\infty}^{\hat{m}}\int_{-\infty}^{\infty}\left(S(0)\exp\left\{(r-\tfrac{1}{2}\sigma_2^2-\rho\sigma_2\widehat{\theta})T+\rho\sigma_2\hat{b}+\sqrt{1-\rho^2}\sigma_2\tilde{b}\right\}-K\right)^{+} \\
&\quad .\frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{\tilde{b}^2}{2T}\right\} \\
&\quad .\frac{2(2\hat{m}-\hat{b})}{T\sqrt{2\pi T}}\exp\left\{-\frac{(2\hat{m}-\hat{b})^2}{2T}+\widehat{\theta}\hat{b}-\tfrac{1}{2}\widehat{\theta}^2T\right\} \\
&\quad .d\tilde{b}\;d\hat{b}\;d\hat{m}.
\end{aligned}
$$

The answer depends on $T, S(0)$ and $Y(0)$. It also depends on $\sigma_1,\sigma_2,\rho,r,K$ and $L$. It does not depend on $\lambda,\mu,\theta_1$, nor $\theta_2$. The parameter $\widehat{\theta}$ appearing in the answer is $\widehat{\theta}=\frac{r}{\sigma_1}-\frac{\sigma_1}{2}$.

**Remark 24.3** If we had not regarded $Y$ as a traded asset, then we would not have tried to set its mean return equal to $r$. We would have had only one equation (see Eqs (0.1),(0.2))

$$
\mu = r+\rho\sigma_2\theta_1+\sqrt{1-\rho^2}\;\sigma_2\theta_2 \tag{1.1}
$$

to determine $\theta_1$ and $\theta_2$. The nonuniqueness of the solution alerts us that some options cannot be hedged. Indeed, any option whose payoff depends on $Y$ cannot be hedged when we are allowed to trade only in the stock.

If we have an option whose payoff depends only on $S$, then $Y$ is superfluous. Returning to the original equation for $S$,

$$\frac{dS}{S} = \mu\ dt + \rho\sigma_2\ dB_1 + \sqrt{1-\rho^2}\ \sigma_2\ dB_2,$$

we should set

$$dW = \rho\ dB_1 + \sqrt{1-\rho^2} dB_2,$$

so $W$ is a Brownian motion under $I\!P$ (Levy's theorem), and

$$\frac{dS}{S} = \mu\ dt + \sigma_2 dW.$$

Now we have only Brownian motion, there will be only one $\theta$, namely,

$$\theta = \frac{\mu - r}{\sigma_2},$$

so with $d\widetilde{W} = \theta\ dt + dW$, we have

$$\frac{dS}{S} = r\ dt + \sigma_2\ d\widetilde{W},$$

and we are on our way.

## 24.2 The PDE for the outside barrier option

Returning to the case of the option with payoff

$$(S(T) - K)^+ \mathbf{1}_{\{Y^*(T) < L\}},$$

we obtain a formula for

$$v(t,x,y) = e^{-r(T-t)} \widetilde{I\!E}^{t,x,y} \left[ (S(T) - K)^+ \mathbf{1}_{\{\max_{t \le u \le T} Y(u)\ <\ L\}}, \right]$$

by replacing $T$, $S(0)$ and $Y(0)$ by $T-t$, $x$ and $y$ respectively in the formula for $v(0, S(0), Y(0))$. Now start at time 0 at $S(0)$ and $Y(0)$. Using the Markov property, we can show that the stochastic process

$$e^{-rt} v(t, S(t), Y(t))$$

is a martingale under $\widetilde{I\!P}$. We compute

$$\begin{aligned}
&d\left[e^{-rt} v(t, S(t), Y(t))\right] \\
&= e^{-rt}\Bigg[ \left( -rv + v_t + rSv_x + rYv_y + \tfrac{1}{2}\sigma_2^2 S^2 v_{xx} + \rho\sigma_1\sigma_2 SY v_{xy} + \tfrac{1}{2}\sigma_1^2 Y^2 v_{yy} \right)\ dt \\
&\quad + \rho\sigma_2 S v_x\ d\widetilde{B}_1 + \sqrt{1-\rho^2}\ \sigma_2 S v_x\ d\widetilde{B}_2 + \sigma_1 Y v_y d\widetilde{B}_1 \Bigg]
\end{aligned}$$

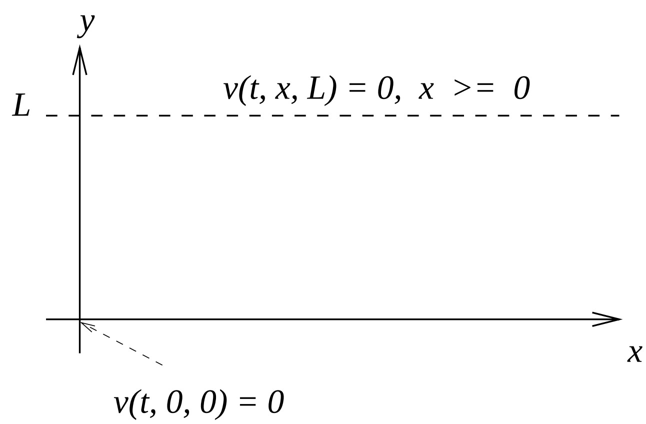



Figure 24.1: *Boundary conditions for barrier option. Note that $t \in [0, T]$ is fixed.*

Setting the $dt$ term equal to 0, we obtain the PDE

$$
\begin{aligned}
-rv + v_t + rxv_x + ryv_y + \tfrac{1}{2}\sigma_2^2 x^2 v_{xx} \\
+ \rho\sigma_1\sigma_2 xy v_{xy} + \tfrac{1}{2}\sigma_1^2 y^2 v_{yy} = 0, \\
0 \le t < T, \quad x \ge 0, \quad 0 \le y \le L.
\end{aligned}
$$

The terminal condition is

$$v(T, x, y) = (x - K)^+, \quad x \ge 0,\ 0 \le y < L,$$

and the boundary conditions are

$$
\begin{aligned}
v(t, 0, 0) = 0, \quad 0 \le t \le T, \\
v(t, x, L) = 0, \quad 0 \le t \le T, \quad x \ge 0.
\end{aligned}
$$

| | |
|---|---|
| $x = 0$<br>$-rv + v_t + ryv_y + \frac{1}{2}\sigma_1^2 y^2 v_{yy} = 0$ | $y = 0$<br>$-rv + v_t + rxv_x + \frac{1}{2}\sigma_2^2 x^2 v_{xx} = 0$ |
| This is the usual Black-Scholes formula in $y$. | This is the usual Black-Scholes formula in $x$. |
| The boundary conditions are<br>$v(t,0,L) = 0, \ v(t,0,0) = 0$;<br>the terminal condition is<br>$v(T,0,y) = (0-K)^+ = 0, \quad y \geq 0.$ | The boundary condition is<br>$v(t,0,0) = e^{-r(T-t)}(0-K)^+ = 0$;<br>the terminal condition is<br>$v(T,x,0) = (x-K)^+, \quad x \geq 0.$ |
| On the $x = 0$ boundary, the option value is $v(t,0,y) = 0, \quad 0 \leq y \leq L.$ | On the $y = 0$ boundary, the barrier is irrelevant, and the option value is given by the usual Black-Scholes formula for a European call. |

## 24.3 The hedge

After setting the $dt$ term to 0, we have the equation

$$\begin{aligned} & d\left[e^{-rt} v(t, S(t), Y(t))\right] \\ & = e^{-rt}\left[\rho\sigma_2 S v_x \, d\widetilde{B}_1 + \sqrt{1-\rho^2}\,\sigma_2 S v_x \, d\widetilde{B}_2 + \sigma_1 Y v_y d\widetilde{B}_1\right], \end{aligned}$$

where $v_x = v_x(t, S(t), Y(t))$, $v_y = v_y(t, S(t), Y(t))$, and $\widetilde{B}_1, \widetilde{B}_2, S, Y$ are functions of $t$. Note that

$$\begin{aligned} d\left[e^{-rt} S(t)\right] &= e^{-rt}\left[-rS(t)\,dt + dS(t)\right] \\ &= e^{-rt}\left[\rho\sigma_2 S(t)\,d\widetilde{B}_1(t) + \sqrt{1-\rho^2}\,\sigma_2 S(t)\,d\widetilde{B}_2(t)\right]. \\ d\left[e^{-rt} Y(t)\right] &= e^{-rt}\left[-rY(t)\,dt + dY(t)\right] \\ &= e^{-rt}\sigma_1 Y(t)\,d\widetilde{B}_1(t). \end{aligned}$$

Therefore,

$$d\left[e^{-rt} v(t, S(t), Y(t))\right] = v_x d[e^{-rt}S] + v_y d[e^{-rt}Y].$$

Let $\Delta_2(t)$ denote the number of shares of stock held at time $t$, and let $\Delta_1(t)$ denote the number of "shares" of the barrier process $Y$. The value $X(t)$ of the portfolio has the differential

$$dX = \Delta_2 dS + \Delta_1 dY + r[X - \Delta_2 S - \Delta_1 Y]\,dt.$$

This is equivalent to

$$d[e^{-rt}X(t)] = \Delta_2(t)d[e^{-rt}S(t)] + \Delta_1(t)d[e^{-rt}Y(t)].$$

To get $X(t) = v(t, S(t), Y(t))$ for all $t$, we must have

$$X(0) = v(0, S(0), Y(0))$$

and

$$\Delta_2(t) = v_x(t, S(t), Y(t)),$$
$$\Delta_1(t) = v_y(t, S(t), Y(t)).$$

# Chapter 25

# American Options

This and the following chapters form part of the course *Stochastic Differential Equations for Finance II.*

## 25.1 Preview of perpetual American put

$$dS = rS\,dt + \sigma S\,dB$$

Intrinsic value at time $t$ : $(K - S(t))^+$.

Let $L \in [0, K]$ be given. Suppose we exercise the first time the stock price is $L$ or lower. We define

$$\begin{aligned}
\tau_L &= \min\{t \geq 0; S(t) \leq L\}, \\
v_L(x) &= \mathbb{E}e^{-r\tau_L}(K - S(\tau_L))^+ \\
&= \begin{cases} K - x & \text{if } x \leq L, \\ (K - L)\mathbb{E}e^{-r\tau_L} & \text{if } x > L. \end{cases}
\end{aligned}$$

The plan is to comute $v_L(x)$ and then maximize over $L$ to find the optimal exercise price. We need to know the distribution of $\tau_L$.

## 25.2 First passage times for Brownian motion: first method

(Based on the reflection principle)

Let $B$ be a Brownian motion under $\mathbb{P}$, let $x > 0$ be given, and define

$$\tau = \min\{t \geq 0; B(t) = x\}.$$

$\tau$ is called the *first passage time to* $x$. We compute the distribution of $\tau$.

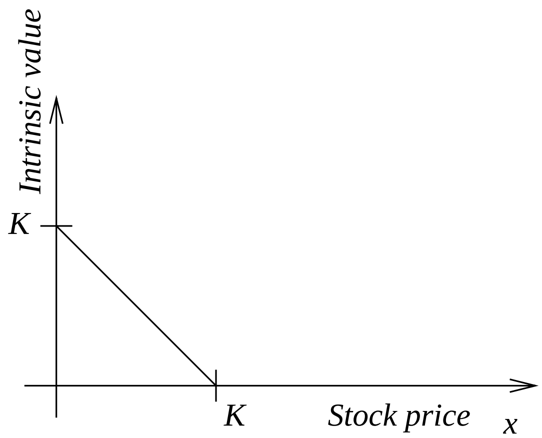



Figure 25.1: *Intrinsic value of perpetual American put*

Define

$$M(t) = \max_{0 \le u \le t} B(u).$$

From the first section of Chapter 20 we have

$$I\!\!P\{M(t) \in dm, B(t) \in db\} = \frac{2(2m-b)}{t\sqrt{2\pi t}} \exp\left\{-\frac{(2m-b)^2}{2t}\right\} dm\, db, \quad m > 0, b < m.$$

Therefore,

$$\begin{aligned} I\!\!P\{M(t) \ge x\} &= \int_x^\infty \int_{-\infty}^m \frac{2(2m-b)}{t\sqrt{2\pi t}} \exp\left\{-\frac{(2m-b)^2}{2t}\right\} db\, dm \\ &= \int_x^\infty \frac{2}{\sqrt{2\pi t}} \exp\left\{-\frac{(2m-b)^2}{2t}\right\} \Bigg|_{b=-\infty}^{b=m} dm \\ &= \int_x^\infty \frac{2}{\sqrt{2\pi t}} \exp\left\{-\frac{m^2}{2t}\right\} dm. \end{aligned}$$

We make the change of variable $z = \frac{m}{\sqrt{t}}$ in the integral to get

$$= \int_{x/\sqrt{t}}^\infty \frac{2}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz.$$

Now

$$\tau \le t \Longleftrightarrow M(t) \ge x,$$

so

$$\begin{aligned}
\mathbb{P}\{\tau \in dt\} &= \frac{\partial}{\partial t}\mathbb{P}\{\tau \le t\}\ dt \\
&= \frac{\partial}{\partial t}\mathbb{P}\left\{M(t) \ge x\right\}\ dt \\
&= \left[\frac{\partial}{\partial t}\int_{x/\sqrt{t}}^{\infty} \frac{2}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz\right]\ dt \\
&= -\frac{2}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2t}\right\} . \frac{\partial}{\partial t}\left(\frac{x}{\sqrt{t}}\right)\ dt \\
&= \frac{x}{t\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}\ dt.
\end{aligned}$$

We also have the Laplace transform formula

$$\begin{aligned}
\mathbb{E}e^{-\alpha\tau} &= \int_0^\infty e^{-\alpha t}\,\mathbb{P}\{\tau \in dt\} \\
&= e^{-x\sqrt{2\alpha}}, \quad \alpha > 0. \quad \text{(See Homework)}
\end{aligned}$$

Reference: Karatzas and Shreve, Brownian Motion and Stochastic Calculus, pp 95-96.

## 25.3 Drift adjustment

Reference: Karatzas/Shreve, *Brownian motion and Stochastic Calculus*, pp 196–197.

For $0 \le t < \infty$, define

$$\begin{aligned}
\widetilde{B}(t) &= \theta t + B(t), \\
Z(t) &= \exp\{-\theta B(t) - \tfrac{1}{2}\theta^2 t\}, \\
&= \exp\{-\theta \widetilde{B}(t) + \tfrac{1}{2}\theta^2 t\},
\end{aligned}$$

Define

$$\tilde{\tau} = \min\{t \ge 0; \widetilde{B}(t) = x\}.$$

We fix a finite time $T$ and change the probability measure "only up to $T$". More specifically, with $T$ fixed, define

$$\widetilde{\mathbb{P}}(A) = \int_A Z(T)\ dP, \quad A \in \mathcal{F}(T).$$

Under $\widetilde{\mathbb{P}}$, the process $\widetilde{B}(t), 0 \le t \le T$, is a (nondrifted) Brownian motion, so

$$\begin{aligned}
\widetilde{\mathbb{P}}\{\tilde{\tau} \in dt\} &= \mathbb{P}\{\tau \in dt\} \\
&= \frac{x}{t\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}\ dt, \quad 0 < t \le T.
\end{aligned}$$

For $0 < t \le T$ we have

$$\begin{aligned}
I\!\!P\{\tilde{\tau} \le t\} &= I\!\!E\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\right] \\
&= \widetilde{I\!\!E}\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\frac{1}{Z(T)}\right] \\
&= \widetilde{I\!\!E}\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\exp\{\theta\widetilde{B}(T) - \tfrac{1}{2}\theta^2 T\}\right] \\
&= \widetilde{I\!\!E}\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\widetilde{I\!\!E}\left[\exp\{\theta\widetilde{B}(T) - \tfrac{1}{2}\theta^2 T\}\middle|\mathcal{F}(\tilde{\tau}\wedge t)\right]\right] \\
&= \widetilde{I\!\!E}\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\exp\{\theta\widetilde{B}(\tilde{\tau}\wedge t) - \tfrac{1}{2}\theta^2(\tilde{\tau}\wedge t)\}\right] \\
&= \widetilde{I\!\!E}\left[\mathbf{1}_{\{\tilde{\tau}\le t\}}\exp\{\theta x - \tfrac{1}{2}\theta^2\tilde{\tau}\}\right] \\
&= \int_0^t \exp\{\theta x - \tfrac{1}{2}\theta^2 s\}\widetilde{I\!\!P}\{\tilde{\tau}\in ds\} \\
&= \int_0^t \frac{x}{s\sqrt{2\pi s}}\exp\left\{\theta x - \tfrac{1}{2}\theta^2 s - \frac{x^2}{2s}\right\}\,ds \\
&= \int_0^t \frac{x}{s\sqrt{2\pi s}}\exp\left\{-\frac{(x-\theta s)^2}{2s}\right\}\,ds.
\end{aligned}$$

Therefore,

$$I\!\!P\{\tilde{\tau}\in dt\} = \frac{x}{t\sqrt{2\pi t}}\exp\left\{-\frac{(x-\theta t)^2}{2t}\right\}\,dt, \quad 0 < t \le T.$$

Since $T$ is arbitrary, this must in fact be the correct formula for all $t > 0$.

## 25.4 Drift-adjusted Laplace transform

Recall the Laplace transform formula for

$$\tau = \min\{t \ge 0; B(t) = x\}$$

for nondrifted Brownian motion:

$$I\!\!E e^{-\alpha\tau} = \int_0^\infty \frac{x}{t\sqrt{2\pi t}}\exp\left\{-\alpha t - \frac{x^2}{2t}\right\}dt = e^{-x\sqrt{2\alpha}}, \quad \alpha > 0, x > 0.$$

For

$$\tilde{\tau} = \min\{t \ge 0; \theta t + B(t) = x\},$$

the Laplace transform is

$$\begin{aligned}
\mathbb{E}e^{-\alpha\tilde{\tau}} &= \int_0^\infty \frac{x}{t\sqrt{2\pi t}} \exp\left\{-\alpha t - \frac{(x-\theta t)^2}{2t}\right\}\, dt \\
&= \int_0^\infty \frac{x}{t\sqrt{2\pi t}} \exp\left\{-\alpha t - \frac{x^2}{2t} + x\theta - \tfrac{1}{2}\theta^2 t\right\}\, dt \\
&= e^{x\theta} \int_0^\infty \frac{x}{t\sqrt{2\pi t}} \exp\left\{-(\alpha + \tfrac{1}{2}\theta^2)t - \frac{x^2}{2t}\right\}\, dt \\
&= e^{x\theta - x\sqrt{2\alpha+\theta^2}}, \quad \alpha > 0, x > 0,
\end{aligned}$$

where in the last step we have used the formula for $\mathbb{E}e^{-\alpha\tau}$ with $\alpha$ replaced by $\alpha + \frac{1}{2}\theta^2$.
If $\tilde{\tau}(\omega) < \infty$, then

$$\lim_{\alpha\downarrow 0} e^{-\alpha\tilde{\tau}(\omega)} = 1;$$

if $\tilde{\tau}(\omega) = \infty$, then $e^{-\alpha\tilde{\tau}(\omega)} = 0$ for every $\alpha > 0$, so

$$\lim_{\alpha\downarrow 0} e^{-\alpha\tilde{\tau}(\omega)} = 0.$$

Therefore,

$$\lim_{\alpha\downarrow 0} e^{-\alpha\tilde{\tau}(\omega)} = \mathbf{1}_{\tilde{\tau}<\infty}.$$

Letting $\alpha\downarrow 0$ and using the Monotone Convergence Theorem in the Laplace transform formula

$$\mathbb{E}e^{-\alpha\tilde{\tau}} = e^{x\theta - x\sqrt{2\alpha+\theta^2}},$$

we obtain

$$\mathbb{P}\{\tilde{\tau} < \infty\} = e^{x\theta - x\sqrt{\theta^2}} = e^{x\theta - x|\theta|}.$$

If $\theta \geq 0$, then

$$\mathbb{P}\{\tilde{\tau} < \infty\} = 1.$$

If $\theta < 0$, then

$$\mathbb{P}\{\tilde{\tau} < \infty\} = e^{2x\theta} < 1.$$

(Recall that $x > 0$).

## 25.5 First passage times: Second method

(Based on martingales)

Let $\sigma > 0$ be given. Then

$$Y(t) = \exp\{\sigma B(t) - \tfrac{1}{2}\sigma^2 t\}$$

is a martingale, so $Y(t \wedge \tau)$ is also a martingale. We have

$$\begin{aligned} 1 &= Y(0 \wedge \tau) \\ &= I\!\!E Y(t \wedge \tau) \\ &= I\!\!E \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\}. \\ &= \lim_{t \to \infty} I\!\!E \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\}. \end{aligned}$$

We want to take the limit inside the expectation. Since

$$0 \;\le\; \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\} \;\le\; e^{x},$$

this is justified by the Bounded Convergence Theorem. Therefore,

$$1 = I\!\!E \lim_{t \to \infty} \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\}.$$

There are two possibilities. For those $\omega$ for which $\tau(\omega) < \infty$,

$$\lim_{t \to \infty} \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\} = e^{\sigma x - \frac{1}{2}\sigma^2 \tau}.$$

For those $\omega$ for which $\tau(\omega) = \infty$,

$$\lim_{t \to \infty} \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\} \le \lim_{t \to \infty} \exp\{\sigma x - \tfrac{1}{2}\sigma^2 t\} \;=\; 0.$$

Therefore,

$$\begin{aligned} 1 &= I\!\!E \lim_{t \to \infty} \exp\{\sigma B(t \wedge \tau) - \tfrac{1}{2}\sigma^2 (t \wedge \tau)\} \\ &= I\!\!E \left[ e^{\sigma x - \frac{1}{2}\sigma^2 \tau} \mathbf{1}_{\tau < \infty} \right] \\ &= I\!\!E e^{\sigma x - \frac{1}{2}\sigma^2 \tau}, \end{aligned}$$

where we understand $e^{\sigma x - \frac{1}{2}\sigma^2 \tau}$ to be zero if $\tau = \infty$.

Let $\alpha = \frac{1}{2}\sigma^2$, so $\sigma = \sqrt{2\alpha}$. We have again derived the Laplace transform formula

$$e^{-x\sqrt{2\alpha}} = I\!\!E e^{-\alpha\tau}, \quad \alpha > 0, x > 0,$$

for the first passage time for nondrifted Brownian motion.

## 25.6 Perpetual American put

$$\begin{aligned} dS &= rS\;dt + \sigma S\;dB \\ S(0) &= x \\ S(t) &= x \exp\{(r - \tfrac{1}{2}\sigma^2)t + \sigma B(t)\} \\ &= x \exp\left\{ \sigma \left[ \underbrace{\left( \frac{r}{\sigma} - \frac{\sigma}{2} \right)}_{\theta} t + B(t) \right] \right\}. \end{aligned}$$

Intrinsic value of the put at time $t$: $(K - S(t))^+$.

Let $L \in [0, K]$ be given. Define for $x \geq L$,

$$\begin{aligned}\tau_L &= \min\{t \geq 0;\ S(t) = L\} \\ &= \min\{t \geq 0;\ \theta t + B(t) = \frac{1}{\sigma}\log\frac{L}{x}\} \\ &= \min\{t \geq 0;\ -\theta t - B(t) = \frac{1}{\sigma}\log\frac{x}{L}\}\end{aligned}$$

Define

$$\begin{aligned}v_L &= (K - L)\mathbb{E}e^{-r\tau_L} \\ &= (K - L)\exp\left\{-\frac{\theta}{\sigma}\log\frac{x}{L} - \frac{1}{\sigma}\log\frac{x}{L}\sqrt{2r + \theta^2}\right\} \\ &= (K - L)\left(\frac{x}{L}\right)^{-\frac{\theta}{\sigma} - \frac{1}{\sigma}\sqrt{2r+\theta^2}}.\end{aligned}$$

We compute the exponent

$$\begin{aligned}-\frac{\theta}{\sigma} - \frac{1}{\sigma}\sqrt{2r + \theta^2} &= -\frac{r}{\sigma^2} + \tfrac{1}{2} - \frac{1}{\sigma}\sqrt{2r + \left(\frac{r}{\sigma} - \sigma/2\right)^2} \\ &= -\frac{r}{\sigma^2} + \tfrac{1}{2} - \frac{1}{\sigma}\sqrt{2r + \frac{r^2}{\sigma^2} - r + \sigma^2/4} \\ &= -\frac{r}{\sigma^2} + \tfrac{1}{2} - \frac{1}{\sigma}\sqrt{\frac{r^2}{\sigma^2} + r + \sigma^2/4} \\ &= -\frac{r}{\sigma^2} + \tfrac{1}{2} - \frac{1}{\sigma}\sqrt{\left(\frac{r}{\sigma} + \sigma/2\right)^2} \\ &= -\frac{r}{\sigma^2} + \tfrac{1}{2} - \frac{1}{\sigma}\left(\frac{r}{\sigma} + \sigma/2\right) \\ &= -\frac{2r}{\sigma^2}.\end{aligned}$$

Therefore,

$$v_L(x) = \begin{cases}(K - x), & 0 \leq x \leq L, \\ (K - L)\left(\frac{x}{L}\right)^{-2r/\sigma^2}, & x \geq L.\end{cases}$$

The curves $(K - L)\left(\frac{x}{L}\right)^{-2r/\sigma^2}$, are all of the form $Cx^{-2r/\sigma^2}$.

We want to choose the largest possible constant. The constant is

$$C = (K - L)L^{2r/\sigma^2},$$

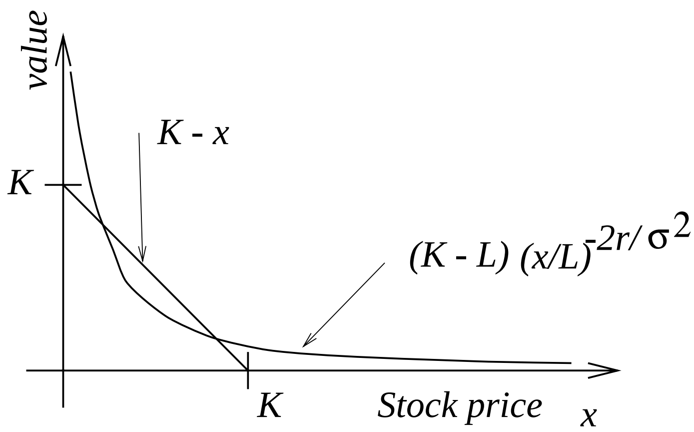



Figure 25.2: *Value of perpetual American put*

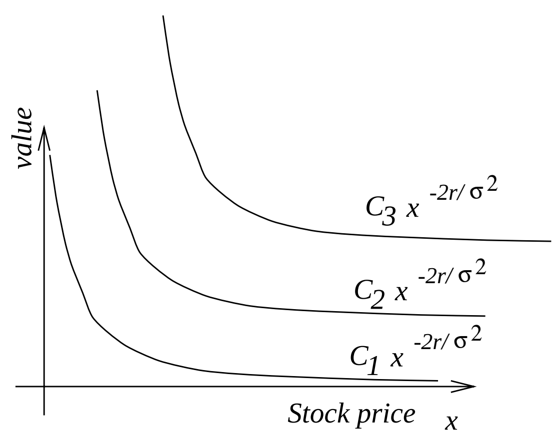



Figure 25.3: *Curves.*

and

$$\begin{aligned}\frac{\partial C}{\partial L} &= -L^{\frac{2r}{\sigma^2}} + \frac{2r}{\sigma^2}(K-L)L^{\frac{2r}{\sigma^2}-1} \\ &= L^{\frac{2r}{\sigma^2}}\left[-1 + \frac{2r}{\sigma^2}(K-L)\frac{1}{L}\right] \\ &= L^{\frac{2r}{\sigma^2}}\left[-\left(1+\frac{2r}{\sigma^2}\right) + \frac{2r}{\sigma^2}\frac{K}{L}\right].\end{aligned}$$

We solve

$$-\left(1+\frac{2r}{\sigma^2}\right) + \frac{2r}{\sigma^2}\frac{K}{L} = 0$$

to get

$$L = \frac{2rK}{\sigma^2 + 2r}.$$

Since $0 < 2r < \sigma^2 + 2r$, we have

$$0 < L < K.$$

Solution to the perpetual American put pricing problem (see Fig. 25.4):

$$v(x) = \begin{cases} (K-x), & 0 \le x \le L^*, \\ (K-L^*)\left(\frac{x}{L^*}\right)^{-2r/\sigma^2}, & x \ge L^*, \end{cases}$$

where

$$L^* = \frac{2rK}{\sigma^2 + 2r}.$$

Note that

$$v'(x) = \begin{cases} -1, & 0 \le x < L^*, \\ -\frac{2r}{\sigma^2}(K-L)^*(L^*)^{2r/\sigma^2} x^{-2r/\sigma^2 - 1}, & x > L^*. \end{cases}$$

We have

$$\begin{aligned}\lim_{x \downarrow L^*} v'(x) &= -2\frac{r}{\sigma^2}(K-L^*)\frac{1}{L^*} \\ &= -2\frac{r}{\sigma^2}\left(K - \frac{2rK}{\sigma^2+2r}\right)\frac{\sigma^2+2r}{2rK} \\ &= -2\frac{r}{\sigma^2}\left(\frac{\sigma^2+2r-2r}{\sigma^2+2r}\right)\frac{\sigma^2+2r}{2r} \\ &= -1 \\ &= \lim_{x \uparrow L^*} v'(x).\end{aligned}$$

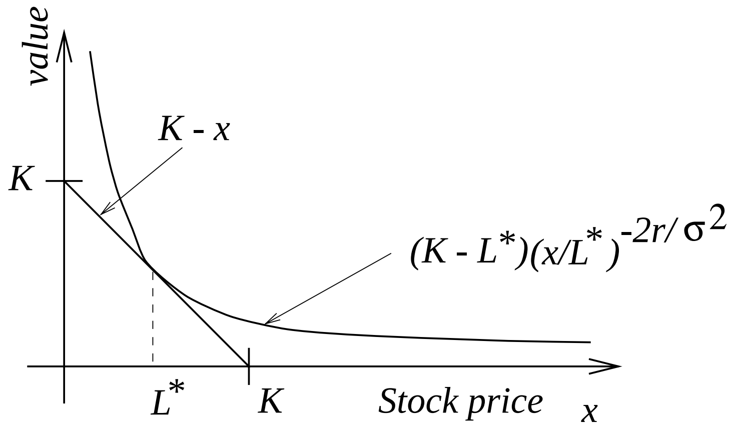



Figure 25.4: *Solution to perpetual American put.*

## 25.7 Value of the perpetual American put

Set

$$\gamma = \frac{2r}{\sigma^2}, \quad L^* = \frac{2rK}{\sigma^2 + 2r} = \frac{\gamma}{\gamma+1}K.$$

If $0 \le x < L^*$, then $v(x) = K - x$. If $L^* \le x < \infty$, then

$$v(x) = \underbrace{(K - L^*)(L^*)^\gamma}_{C} x^{-\gamma} \tag{7.1}$$

$$= I\!\!E^x \left[ e^{-r\tau}(K - L^*)^+ \mathbf{1}_{\{\tau<\infty\}} \right], \tag{7.2}$$

where

$$S(0) = x \tag{7.3}$$

$$\tau = \min\{t \ge 0;\ S(t) = L^*\}. \tag{7.4}$$

If $0 \le x < L^*$, then

$$-rv(x) + rxv'(x) + \tfrac{1}{2}\sigma^2 x^2 v''(x) \;=\; -r(K - x) + rx(-1) \;=\; -rK.$$

If $L^* \le x < \infty$, then

$$\begin{aligned}
&-rv(x) + rxv'(x) + \tfrac{1}{2}\sigma^2 x^2 v''(x) \\
&= C[-rx^{-\gamma} - rx\gamma x^{-\gamma-1} - \tfrac{1}{2}\sigma^2 x^2 \gamma(-\gamma-1)x^{-\gamma-2}] \\
&= Cx^{-\gamma}[-r - r\gamma - \tfrac{1}{2}\sigma^2\gamma(-\gamma-1)] \\
&= C(-\gamma-1)x^{-\gamma}\left[r - \tfrac{1}{2}\sigma^2\left(\frac{2r}{\sigma^2}\right)\right] \\
&= 0.
\end{aligned}$$

In other words, $v$ solves the *linear complementarity problem:* (See Fig. 25.5).

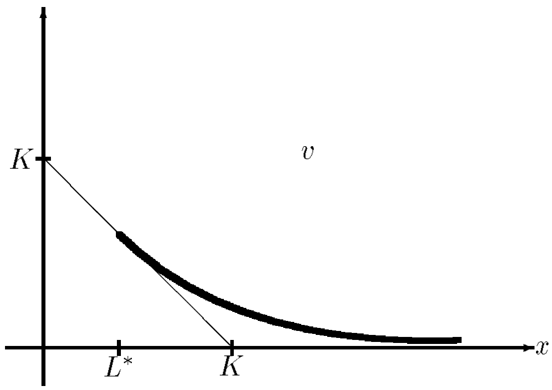



Figure 25.5: *Linear complementarity*

For all $x \in I\!R$, $x \neq L^*$,

$$rv - rxv' - \tfrac{1}{2}\sigma^2 x^2 v'' \geq 0, \qquad \text{(a)}$$

$$v \geq (K-x)^+, \qquad \text{(b)}$$

One of the inequalities (a) or (b) is an equality. (c)

The half-line $[0, \infty)$ is divided into two regions:

$$\mathcal{C} = \{x;\ v(x) > (K-x)^+\},$$
$$\mathcal{S} = \{x;\ rv - rxv' - \tfrac{1}{2}\sigma^2 x^2 v'' > 0\},$$

and $L^*$ is the boundary between them. If the stock price is in $\mathcal{C}$, the owner of the put should not exercise (should "continue"). If the stock price is in $\mathcal{S}$ or at $L^*$, the owner of the put should exercise (should "stop").

## 25.8 Hedging the put

Let $S(0)$ be given. Sell the put at time zero for $v(S(0))$. Invest the money, holding $\Delta(t)$ shares of stock and consuming at rate $C(t)$ at time $t$. The value $X(t)$ of this portfolio is governed by

$$dX(t) = \Delta(t)\ dS(t)\ +\ r(X(t) - \Delta(t)S(t))\ dt\ -\ C(t)\ dt,$$

or equivalently,

$$d(e^{-rt}X(t)) = -e^{-rt}C(t)\ dt + e^{-rt}\Delta(t)\sigma S(t)\ dB(t).$$

The discounted value of the put satisfies

$$\begin{aligned}
d\left(e^{-rt}v(S(t))\right) &= e^{-rt}\left[-rv(S(t)) + rS(t)v'(S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)v''(S(t))\right]\,dt \\
&\quad + e^{-rt}\sigma S(t)v'(S(t))\,dB(t) \\
&= -rKe^{-rt}\mathbf{1}_{\{S(t)<L^*\}}dt + e^{-rt}\sigma S(t)v'(S(t))\,dB(t).
\end{aligned}$$

We should set

$$\begin{aligned}
C(t) &= rK\mathbf{1}_{\{S(t)<L^*\}}, \\
\Delta(t) &= v'(S(t)).
\end{aligned}$$

**Remark 25.1** If $S(t) < L^*$, then

$$v(S(t)) = K - S(t), \quad \Delta(t) = v'(S(t)) = -1.$$

To hedge the put when $S(t) < L^*$, short one share of stock and hold $K$ in the money market. As long as the owner does not exercise, you can consume the interest from the money market position, i.e.,

$$C(t) = rK\mathbf{1}_{\{S(t)<L^*\}}.$$

Properties of $e^{-rt}v(S(t))$:

1. $e^{-rt}v(S(t))$ is a supermartingale (see its differential above).
2. $e^{-rt}v(S(t)) \geq e^{-rt}(K - S(t))^+, \quad 0 \leq t < \infty$;
3. $e^{-rt}v(S(t))$ is the smallest process with properties 1 and 2.

**Explanation of property 3.** Let $Y$ be a supermartingale satisfying

$$Y(t) \geq e^{-rt}(K - S(t))^+, \quad 0 \leq t < \infty. \tag{8.1}$$

Then property 3 says that

$$Y(t) \geq e^{-rt}v(S(t)), \quad 0 \leq t < \infty. \tag{8.2}$$

We use (8.1) to prove (8.2) for $t = 0$, i.e.,

$$Y(0) \geq v(S(0)). \tag{8.3}$$

If $t$ is not zero, we can take $t$ to be the initial time and $S(t)$ to be the initial stock price, and then adapt the argument below to prove property (8.2).

**Proof of (8.3), assuming $Y$ is a supermartingale satisfying (8.1)**:

**Case I:** $S(0) \leq L^*$. We have

$$Y(0) \underbrace{\geq}_{(8.1)} (K - S(0))^+ = v(S(0)).$$

**Case II:** $S(0) > L^*$: For $T > 0$, we have

$$
\begin{aligned}
Y(0) &\geq \mathbb{E}Y(\tau \wedge T) \quad \text{(Stopped supermartingale is a supermartingale)} \\
&\geq \mathbb{E}\left[Y(\tau \wedge T)\mathbf{1}_{\{\tau<\infty\}}\right]. \quad \text{(Since } Y \geq 0\text{)}
\end{aligned}
$$

Now let $T \to \infty$ to get

$$
\begin{aligned}
Y(0) &\geq \lim_{T\to\infty} \mathbb{E}\left[Y(\tau \wedge T)\mathbf{1}_{\{\tau<\infty\}}\right] \\
&\geq \mathbb{E}\left[Y(\tau)\mathbf{1}_{\{\tau<\infty\}}\right] \quad \text{(Fatou's Lemma)} \\
&\geq \mathbb{E}\left[e^{-r\tau}(K - \underbrace{S(\tau)}_{L^*})^+\mathbf{1}_{\{\tau<\infty\}}\right] \quad \text{(by 8.1)} \\
&= v(S(0)). \quad \text{(See eq. 7.2)}
\end{aligned}
$$

## 25.9 Perpetual American contingent claim

Intinsic value: $h(S(t))$.

Value of the American contingent claim:

$$
v(x) = \sup_{\tau} \mathbb{E}^x \left[e^{-r\tau} h(S(\tau))\right],
$$

where the supremum is over all stopping times.

Optimal exercise rule: Any stopping time $\tau$ which attains the supremum.

**Characterization of $v$:**

1. $e^{-rt}v(S(t))$ is a supermartingale;
2. $e^{-rt}v(S(t)) \geq e^{-rt}h(S(t)), \quad 0 < t < \infty$;
3. $e^{-rt}v(S(t))$ is the smallest process with properties 1 and 2.

## 25.10 Perpetual American call

$$
v(x) = \sup_{\tau} \mathbb{E}^x \left[e^{-r\tau}(S(\tau) - K)^+\right]
$$

**Theorem 10.63**

$$
v(x) = x \quad \forall x \geq 0.
$$

**Proof:** For every $t$,

$$\begin{aligned} v(x) &\geq \mathbb{E}^x\left[e^{-rt}(S(t)-K)^+\right] \\ &\geq \mathbb{E}^x\left[e^{-rt}(S(t)-K)\right] \\ &= \mathbb{E}^x\left[e^{-rt}S(t)\right] - e^{-rt}K \\ &= x - e^{-rt}K. \end{aligned}$$

Let $t \to \infty$ to get $v(x) \geq x$.

Now start with $S(0) = x$ and define

$$Y(t) = e^{-rt}S(t).$$

Then:

1. $Y$ is a supermartingale (in fact, $Y$ is a martingale);
2. $Y(t) \geq e^{-rt}(S(t)-K)^+, \quad 0 \leq t < \infty$.

Therefore, $Y(0) \geq v(S(0))$, i.e.,

$$x \geq v(x).$$

■

**Remark 25.2** No matter what $\tau$ we choose,

$$\mathbb{E}^x\left[e^{-r\tau}(S(\tau)-K)^+\right] \;<\; \mathbb{E}^x\left[e^{-r\tau}S(\tau)\right] \;\leq\; x \;=\; v(x).$$

There is no optimal exercise time.

## 25.11 Put with expiration

Expiration time: $T > 0$.

Intrinsic value: $(K - S(t))^+$.

Value of the put:

$$\begin{aligned} v(t,x) &= (\text{value of the put at time } t \text{ if } S(t) = x) \\ &= \sup_{\substack{\underbrace{t \leq \tau \leq T} \\ \tau:\text{stopping time}}} \mathbb{E}^x e^{-r(\tau-t)}(K - S(\tau))^+. \end{aligned}$$

See Fig. 25.6. It can be shown that $v, v_t, v_x$ are continuous across the boundary, while $v_{xx}$ has a jump.

Let $S(0)$ be given. Then

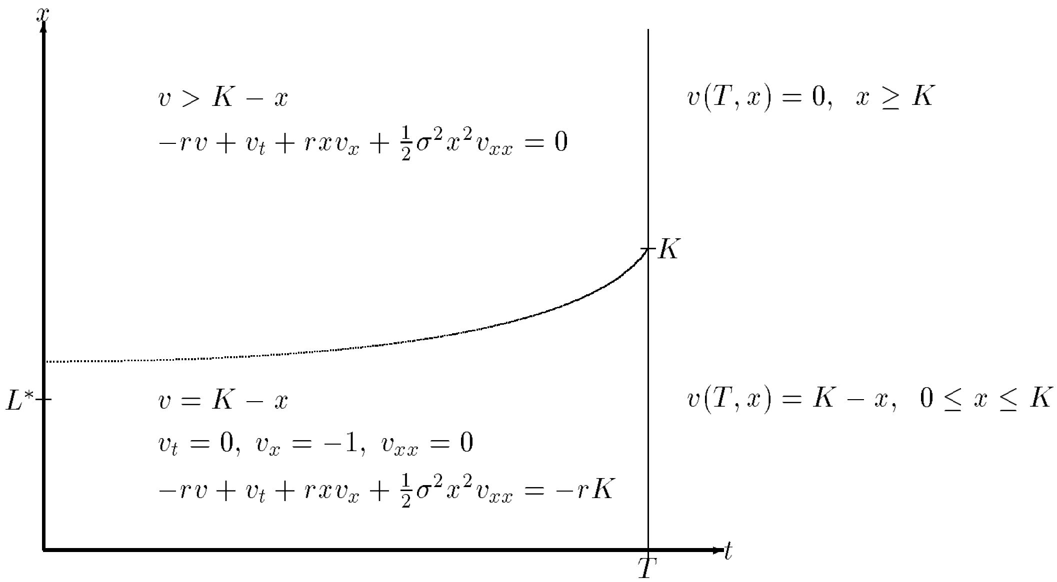



Figure 25.6: *Value of put with expiration*

1. $e^{-rt}v(t, S(t)), \quad 0 \leq t \leq T$, is a supermartingale;
2. $e^{-rt}v(t, S(t)) \geq e^{-rt}(K - S(t))^+, \quad 0 \leq t \leq T$;
3. $e^{-rt}v(t, S(t))$ is the smallest process with properties 1 and 2.

## 25.12 American contingent claim with expiration

Expiration time: $T > 0$.

Intrinsic value: $h(S(t))$.

Value of the contingent claim:

$$v(t, x) = \sup_{t \leq \tau \leq T} I\!E^x e^{-r(\tau - t)} h(S(\tau)).$$

Then

$$rv - v_t - rxv_x - \tfrac{1}{2}\sigma^2 x^2 v_{xx} \geq 0, \qquad \text{(a)}$$

$$v \geq h(x), \qquad \text{(b)}$$

At every point $(t, x) \in [0, T] \times [0, \infty)$, either (a) or (b) is an equality. (c)

**Characterization of $v$:** Let $S(0)$ be given. Then

1. $e^{-rt}v(t, S(t)), \quad 0 \le t \le T$, is a supermartingale;
2. $e^{-rt}v(t, S(t)) \ge e^{-rt}h(S(t))$;
3. $e^{-rt}v(t, S(t))$ is the smallest process with properties 1 and 2.

The optimal exercise time is

$$\tau = \min\{t \ge 0;\; v(t, S(t)) = h(S(t))\}$$

If $\tau(\omega) = \infty$, then there is no optimal exercise time along the particular path $\omega$.

# Chapter 26

# Options on dividend-paying stocks

## 26.1 American option with convex payoff function

**Theorem 1.64** *Consider the stock price process*

$$dS(t) = r(t)S(t)\ dt + \sigma(t)S(t)\ dB(t),$$

*where $r$ and $\sigma$ are processes and $r(t) \geq 0, \quad 0 \leq t \leq T$, a.s. This stock pays no dividends. Let $h(x)$ be a convex function of $x \geq 0$, and assume $h(0) = 0$. (E.g., $h(x) = (x - K)^+$). An American contingent claim paying $h(S(t))$ if exercised at time $t$ does not need to be exercised before expiration, i.e., waiting until expiration to decide whether to exercise entails no loss of value.*

**Proof:** For $0 \leq \alpha \leq 1$ and $x \geq 0$, we have

$$\begin{aligned} h(\alpha x) &= h((1-\alpha)0 + \alpha x) \\ &\leq (1-\alpha)h(0) + \alpha h(x) \\ &= \alpha h(x). \end{aligned}$$

Let $T$ be the time of expiration of the contingent claim. For $0 \leq t \leq T$,

$$0 \leq \frac{\beta(t)}{\beta(T)} = \exp\left\{-\int_t^T r(u)\ du\right\} \leq 1$$

and $S(T) \geq 0$, so

$$h\left(\frac{\beta(t)}{\beta(T)} S(T)\right) \leq \frac{\beta(t)}{\beta(T)} h(S(T)). \qquad (*)$$

Consider a European contingent claim paying $h(S(T))$ at time $T$. The value of this claim at time $t \in [0, T]$ is

$$X(t) = \beta(t)\ I\!E\left[\frac{1}{\beta(T)} h(S(T)) \middle| \mathcal{F}(t)\right].$$

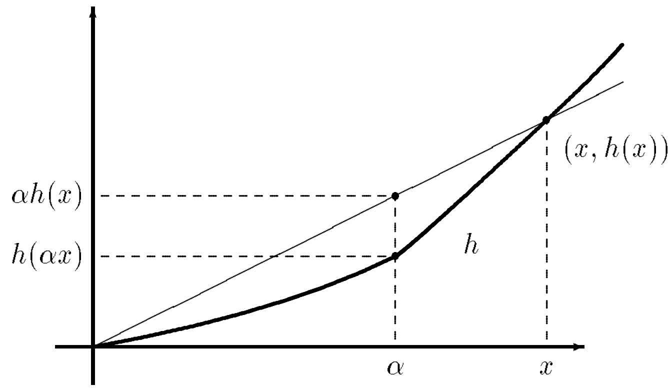



Figure 26.1: *Convex payoff function*

Therefore,

$$\begin{aligned}
\frac{X(t)}{\beta(t)} &= \frac{1}{\beta(t)} I\!E\left[\frac{\beta(t)}{\beta(T)} h(S(T)) \middle| \mathcal{F}(t)\right] \\
&\geq \frac{1}{\beta(t)} I\!E\left[h\left(\frac{\beta(t)}{\beta(T)} S(T)\right) \middle| \mathcal{F}(t)\right] \quad \text{(by (*))} \\
&\geq \frac{1}{\beta(t)} h\left(\beta(t)\ I\!E\left[\frac{S(T)}{\beta(T)} \middle| \mathcal{F}(t)\right]\right) \quad \text{(Jensen's inequality)} \\
&= \frac{1}{\beta(t)} h\left(\beta(t) \frac{S(t)}{\beta(t)}\right) \quad \left(\frac{S}{\beta} \text{ is a martingale}\right) \\
&= \frac{1}{\beta(t)} h(S(t)).
\end{aligned}$$

This shows that the value $X(t)$ of the European contingent claim dominates the intrinsic value $h(S(t))$ of the American claim. In fact, except in degenerate cases, the inequality

$$X(t) \geq h(S(t)), \quad 0 \leq t \leq T,$$

is strict, i.e., the American claim should not be exercised prior to expiration. ∎

## 26.2 Dividend paying stock

Let $r$ and $\sigma$ be constant, let $\delta$ be a "dividend coefficient" satisfying

$$0 < \delta < 1.$$

Let $T > 0$ be an expiration time, and let $t_1 \in (0,T)$ be the time of dividend payment. The stock price is given by

$$S(t) = \begin{cases} S(0)\exp\{(r - \frac{1}{2}\sigma^2)t + \sigma B(t)\}, & 0 \le t \le t_1, \\ (1-\delta)S(t_1)\exp\{(r - \frac{1}{2}\sigma^2)(t - t_1) + \sigma(B(t) - B(t_1))\}, & t_1 < t \le T. \end{cases}$$

Consider an American call on this stock. At times $t \in (t_1, T)$, it is not optimal to exercise, so the value of the call is given by the usual Black-Scholes formula

$$v(t,x) = xN(d_+(T-t,x)) \; - \; Ke^{-r(T-t)}N(d_-(T-t,x)), \quad t_1 < t \le T,$$

where

$$d_\pm(T-t,x) = \frac{1}{\sigma\sqrt{T-t}}\left[\log\frac{x}{K} + (T-t)(r \pm \sigma^2/2)\right].$$

At time $t_1$, immediately *after* payment of the dividend, the value of the call is

$$v(t_1, (1-\delta)S(t_1)).$$

At time $t_1$, immediately *before* payment of the dividend, the value of the call is

$$w(t_1, S(t_1)),$$

where

$$w(t_1,x) = \max\left\{(x-K)^+, \; v(t_1,(1-\delta)x\right\}.$$

**Theorem 2.65** *For $0 \le t \le t_1$, the value of the American call is $w(t, S(t))$, where*

$$w(t,x) = I\!\!E^{t,x}\left[e^{-r(t_1-t)}w(t_1, S(t_1))\right].$$

*This function satisfies the usual Black-Scholes equation*

$$-rw + w_t + rxw_x + \tfrac{1}{2}\sigma^2x^2w_{xx} = 0, \quad 0 \le t \le t_1, \; x \ge 0,$$

*(where $w = w(t,x)$) with terminal condition*

$$w(t_1,x) = \max\left\{(x-K)^+, \; v(t_1,(1-\delta)x)\right\}, \; x \ge 0,$$

*and boundary condition*

$$w(t,0) = 0, \quad 0 \le t \le T.$$

*The hedging portfolio is*

$$\Delta(t) = \begin{cases} w_x(t,S(t)), & 0 \le t \le t_1, \\ v_x(t,S(t)), & t_1 < t \le T. \end{cases}$$

**Proof:** We only need to show that an American contingent claim with payoff $w(t_1, S(t_1))$ at time $t_1$ need not be exercised before time $t_1$. According to Theorem 1.64, it suffices to prove

1. $w(t_1, 0) = 0$,

2. $w(t_1, x)$ is convex in $x$.

Since $v(t_1, 0) = 0$, we have immediately that

$$w(t_1, 0) = \max\left\{(0 - K)^+,\ v(t_1, (1-\delta)0)\right\} = 0.$$

To prove that $w(t_1, x)$ is convex in $x$, we need to show that $v(t_1, (1-\delta)x)$ is convex is $x$. Obviously, $(x - K)^+$ is convex in $x$, and the maximum of two convex functions is convex. The proof of the convexity of $v(t_1, (1-\delta)x)$ in $x$ is left as a homework problem. ■

## 26.3 Hedging at time $t_1$

Let $x = S(t_1)$.

**Case I:** $v(t_1, (1-\delta)x) \geq (x - K)^+$.
The option need not be exercised at time $t_1$ (should not be exercised if the inequality is strict). We have

$$\begin{aligned} w(t_1, x) &= v(t_1, (1-\delta)x), \\ \Delta(t_1) = w_x(t_1, x) &= (1-\delta)v_x(t_1, (1-\delta)x) = (1-\delta)\Delta(t_1+), \end{aligned}$$

where

$$\Delta(t_1+) = \lim_{t \downarrow t_1} \Delta(t)$$

is the number of shares of stock held by the hedge immediately after payment of the dividend. The post-dividend position can be achieved by reinvesting in stock the dividends received on the stock held in the hedge. Indeed,

$$\begin{aligned} \Delta(t_1+) &= \frac{1}{1-\delta}\Delta(t_1) = \Delta(t_1) + \frac{\delta}{1-\delta}\Delta(t_1) \\ &= \Delta(t_1) + \frac{\delta\Delta(t_1)S(t_1)}{(1-\delta)S(t_1)} \\ &= \text{\# of shares held when dividend is paid} + \frac{\text{dividends received}}{\text{price per share when dividend is reinvested}} \end{aligned}$$

**Case II:** $v(t_1, (1-\delta)x) < (x - K)^+$.
The owner of the option should exercise before the dividend payment at time $t_1$ and receive $(x - K)$. The hedge has been constructed so the seller of the option has $x - K$ before the dividend payment at time $t_1$. If the option is not exercised, its value drops from $x - K$ to $v(t_1, (1-\delta)x)$, and the seller of the option can pocket the difference and continue the hedge.

# Chapter 27

# Bonds, forward contracts and futures

Let $\{W(t), \mathcal{F}(t);\ 0 \le t \le T\}$ be a Brownian motion (Wiener process) on some $(\Omega, \mathcal{F}, \mathbb{P})$. Consider an asset, which we call a stock, whose price satisfies

$$dS(t) = r(t)S(t)\ dt + \sigma(t)S(t)\ dW(t).$$

Here, $r$ and $\sigma$ are adapted processes, and we have already switched to the risk-neutral measure, which we call $I\!\!P$. Assume that every martingale under $I\!\!P$ can be represented as an integral with respect to $W$.

Define the accumulation factor

$$\beta(t) = \exp\left\{\int_0^t r(u)\ du\right\}.$$

A zero-coupon bond, maturing at time $T$, pays 1 at time $T$ and nothing before time $T$. According to the risk-neutral pricing formula, its value at time $t \in [0, T]$ is

$$\begin{aligned} B(t,T) &= \beta(t)\ I\!\!E\left[\frac{1}{\beta(T)}\middle|\mathcal{F}(t)\right] \\ &= I\!\!E\left[\frac{\beta(t)}{\beta(T)}\middle|\mathcal{F}(t)\right] \\ &= I\!\!E\left[\exp\left\{-\int_t^T r(u)\ du\right\}\middle|\mathcal{F}(t)\right]. \end{aligned}$$

Given $B(t,T)$ dollars at time $t$, one can construct a portfolio of investment in the stock and money market so that the portfolio value at time $T$ is 1 almost surely. Indeed, for some process $\gamma$,

$$\begin{aligned}
B(t,T) &= \beta(t) \underbrace{I\!E\left[\frac{1}{\beta(T)}\middle|\mathcal{F}(t)\right]}_{\text{martingale}} \\
&= \beta(t)\left[I\!E\left(\frac{1}{\beta(T)}\right) + \int_0^t \gamma(u)\,dW(u)\right] \\
&= \beta(t)\left[B(0,T) + \int_0^t \gamma(u)\,dW(u)\right], \\
dB(t,T) &= r(t)\beta(t)\left[B(0,T) + \int_0^t \gamma(u)\,dW(u)\right]\,dt \;+\; \beta(t)\gamma(t)\,dW(t) \\
&= r(t)B(t,T)\,dt + \beta(t)\gamma(t)\,dW(t).
\end{aligned}$$

The value of a portfolio satisfies

$$\begin{aligned}
dX(t) &= \Delta(t)\,dS(t) + r(t)[X(t) - \Delta(t)S(t)]dt \\
&= r(t)X(t)\,dt + \Delta(t)\sigma(t)S(t)\,dW(t).
\end{aligned} \tag{*}$$

We set

$$\Delta(t) = \frac{\beta(t)\gamma(t)}{\sigma(t)S(t)}.$$

If, at any time $t$, $X(t) = B(t,T)$ and we use the portfolio $\Delta(u),\ t \le u \le T$, then we will have

$$X(T) = B(T,T) = 1.$$

If $r(t)$ is nonrandom for all $t$, then

$$\begin{aligned}
B(t,T) &= \exp\left\{-\int_t^T r(u)\,du\right\}, \\
dB(t,T) &= r(t)B(t,T)\,dt,
\end{aligned}$$

i.e., $\gamma = 0$. Then $\Delta$ given above is zero. If, at time $t$, you are given $B(t,T)$ dollars and you always invest only in the money market, then at time $T$ you will have

$$B(t,T)\exp\left\{\int_t^T r(u)\,du\right\} = 1.$$

If $r(t)$ is random for all $t$, then $\gamma$ is not zero. One generally has three different instruments: the stock, the money market, and the zero coupon bond. Any two of them are sufficient for hedging, and the two which are most convenient can depend on the instrument being hedged.

## 27.1 Forward contracts

We continue with the set-up for zero-coupon bonds. The $T$-*forward price* of the stock at time $t \in [0,T]$ is the $\mathcal{F}(t)$-measurable price, agreed upon at time $t$, for purchase of a share of stock at time $T$, chosen so the forward contract has value zero at time $t$. In other words,

$$\mathbb{E}\left[\frac{1}{\beta(T)}\left(S(T)-F(t)\right)\Big|\mathcal{F}(t)\right]=0, \quad 0 \le t \le T.$$

We solve for $F(t)$:

$$\begin{aligned}
0 &= \mathbb{E}\left[\frac{1}{\beta(T)}\left(S(T)-F(t)\right)\Big|\mathcal{F}(t)\right] \\
&= \mathbb{E}\left[\frac{S(T)}{\beta(T)}\Big|\mathcal{F}(t)\right] - \frac{F(t)}{\beta(t)}\mathbb{E}\left[\frac{\beta(t)}{\beta(T)}\Big|\mathcal{F}(t)\right] \\
&= \frac{S(t)}{\beta(t)} - \frac{F(t)}{\beta(t)}B(t,T).
\end{aligned}$$

This implies that

$$F(t)=\frac{S(t)}{B(t,T)}.$$

**Remark 27.1 (Value vs. Forward price)** The $T$-forward price $F(t)$ is *not* the value at time $t$ of the forward contract. The value of the contract at time $t$ is zero. $F(t)$ is the price agreed upon at time $t$ which will be paid for the stock at time $T$.

## 27.2 Hedging a forward contract

Enter a forward contract at time 0, i.e., agree to pay $F(0)=\frac{S(0)}{B(0,T)}$ for a share of stock at time $T$. At time zero, this contract has value 0. At later times, however, it does not. In fact, its value at time $t \in [0,T]$ is

$$\begin{aligned}
V(t) &= \beta(t)\ \mathbb{E}\left[\frac{1}{\beta(T)}(S(T)-F(0))\Big|\mathcal{F}(t)\right] \\
&= \beta(t)\ \mathbb{E}\left[\frac{S(T)}{\beta(T)}\Big|\mathcal{F}(t)\right] - F(0)\ \mathbb{E}\left[\frac{\beta(t)}{\beta(T)}\Big|\mathcal{F}(t)\right] \\
&= \beta(t)\frac{S(t)}{\beta(t)} - F(0)B(t,T) \\
&= S(t)-F(0)B(t,T).
\end{aligned}$$

This suggests the following hedge of a short position in the forward contract. At time 0, short $F(0)$ $T$-maturity zero-coupon bonds. This generates income

$$F(0)B(0,T)=\frac{S(0)}{B(0,T)}B(0,T)=S(0).$$

Buy one share of stock. This portfolio requires no initial investment. Maintain this position until time $T$, when the portfolio is worth

$$S(T) - F(0)B(T,T) = S(T) - F(0).$$

Deliver the share of stock and receive payment $F(0)$.

A short position in the forward could also be hedged using the stock and money market, but the implementation of this hedge would require a term-structure model.

## 27.3 Future contracts

Future contracts are designed to remove the risk of default inherent in forward contracts. Through the device of *marking to market*, the value of the future contract is maintained at zero at all times. Thus, either party can close out his/her position at any time.

Let us first consider the situation with discrete trading dates

$$0 = t_0 < t_1 < \ldots < t_n = T.$$

On each $[t_j, t_{j+1})$, $r$ is constant, so

$$\begin{aligned}\beta(t_{k+1}) &= \exp\left\{\int_0^{t_{k+1}} r(u)\,du\right\}\\ &= \exp\left\{\sum_{j=0}^{k} r(t_j)(t_{j+1} - t_j)\right\}\end{aligned}$$

is $\mathcal{F}(t_k)$-measurable.

Enter a future contract at time $t_k$, taking the long position, when the future price is $\Phi(t_k)$. At time $t_{k+1}$, when the future price is $\Phi(t_{k+1})$, you receive a payment $\Phi(t_{k+1}) - \Phi(t_k)$. (If the price has fallen, you make the payment $-(\Phi(t_{k+1}) - \Phi(t_k))$. ) The mechanism for receiving and making these payments is the *margin account* held by the broker.

By time $T = t_n$, you have received the sequence of payments

$$\Phi(t_{k+1}) - \Phi(t_k),\ \Phi(t_{k+2}) - \Phi(t_{k+1}),\ \ldots,\ \Phi(t_n) - \Phi(t_{n-1})$$

at times $t_{k+1}, t_{k+2}, \ldots, t_n$. The value at time $t = t_0$ of this sequence is

$$\beta(t)\ I\!\!E\left[\sum_{j=k}^{n-1} \frac{1}{\beta(t_{j+1})}\left(\Phi(t_{j+1}) - \Phi(t_j)\right)\Big|\mathcal{F}(t)\right].$$

Because it costs nothing to enter the future contract at time $t$, this expression must be zero almost surely.

The continuous-time version of this condition is

$$\beta(t)\ \mathbb{E}\left[\int_t^T \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right] = 0, \quad 0 \le t \le T.$$

Note that $\beta(t_{j+1})$ appearing in the discrete-time version is $\mathcal{F}(t_j)$-measurable, as it should be when approximating a stochastic integral.

**Definition 27.1** The $T$-*future price* of the stock is any $\mathcal{F}(t)$-adapted stochastic process

$$\{\Phi(t);\ 0 \le t \le T\},$$

satisfying

$$\Phi(T) = S(T) \text{ a.s., and} \tag{a}$$

$$\mathbb{E}\left[\int_t^T \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right] = 0, \quad 0 \le t \le T. \tag{b}$$

**Theorem 3.66** *The unique process satisfying (a) and (b) is*

$$\Phi(t) = \mathbb{E}\left[S(T)\middle|\mathcal{F}(t)\right], 0 \le t \le T.$$

**Proof:** We first show that (b) holds if and only if $\Phi$ is a martingale. If $\Phi$ is a martingale, then $\int_0^t \frac{1}{\beta(u)}\ d\Phi(u)$ is also a martingale, so

$$\begin{aligned}\mathbb{E}\left[\int_t^T \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right] &= \mathbb{E}\left[\int_0^t \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right] - \int_0^t \frac{1}{\beta(u)}\ d\Phi(u)\\ &= 0.\end{aligned}$$

On the other hand, if (b) holds, then the martingale

$$M(t) = \mathbb{E}\left[\int_0^T \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right]$$

satisfies

$$\begin{aligned}M(t) &= \int_0^t \frac{1}{\beta(u)}\ d\Phi(u) + \mathbb{E}\left[\int_t^T \frac{1}{\beta(u)}\ d\Phi(u)\middle|\mathcal{F}(t)\right]\\ &= \int_0^t \frac{1}{\beta(u)}\ d\Phi(u), \quad 0 \le t \le T.\end{aligned}$$

this implies

$$\begin{aligned}dM(t) &= \frac{1}{\beta(t)}\ d\Phi(t),\\ d\Phi(t) &= \beta(t)\ dM(t),\end{aligned}$$

and so $\Phi$ is a martingale (its differential has no $dt$ term).

Now define

$$\Phi(t) = \mathbb{E}\left[S(T)\middle|\mathcal{F}(t)\right], \quad 0 \le t \le T.$$

Clearly (a) is satisfied. By the tower property, $\Phi$ is a martingale, so (b) is also satisfied. Indeed, this $\Phi$ is the only martingale satisfying (a). ■

## 27.4 Cash flow from a future contract

With a forward contract, entered at time 0, the buyer agrees to pay $F(0)$ for an asset valued at $S(T)$. The only payment is at time $T$.

With a future contract, entered at time 0, the buyer receives a cash flow (which may at times be negative) between times 0 and $T$. If he still holds the contract at time $T$, then he pays $S(T)$ at time $T$ for an asset valued at $S(T)$. The cash flow received between times 0 and $T$ sums to

$$\int_0^T d\Phi(u) = \Phi(T) - \Phi(0) = S(T) - \Phi(0).$$

Thus, if the future contract holder takes delivery at time $T$, he has paid a total of

$$(\Phi(0) - S(T)) + S(T) = \Phi(0)$$

for an asset valued at $S(T)$.

## 27.5 Forward-future spread

Future price: $\Phi(t) = \mathbb{E}\left[S(T)\middle|\mathcal{F}(t)\right]$.

Forward price:

$$F(t) = \frac{S(t)}{B(t,T)} = \frac{S(t)}{\beta(t)\,\mathbb{E}\left[\frac{1}{\beta(T)}\middle|\mathcal{F}(t)\right]}.$$

Forward-future spread:

$$\begin{aligned}\Phi(0) - F(0) &= \mathbb{E}[S(T)] - \frac{S(0)}{\mathbb{E}\left[\frac{1}{\beta(T)}\right]}\\ &= \frac{1}{\mathbb{E}\left(\frac{1}{\beta(T)}\right)}\left[\mathbb{E}\left(\frac{1}{\beta(T)}\right)\mathbb{E}\left(S(T)\right) - \mathbb{E}\left(\frac{S(T)}{\beta(T)}\right)\right].\end{aligned}$$

If $\frac{1}{\beta(T)}$ and $S(T)$ are uncorrelated,

$$\Phi(0) = F(0).$$

If $\frac{1}{\beta(T)}$ and $S(T)$ are positively correlated, then

$$\Phi(0) \le F(0).$$

This is the case that a rise in stock price tends to occur with a fall in the interest rate. The owner of the future tends to receive income when the stock price rises, but invests it at a declining interest rate. If the stock price falls, the owner usually must make payments on the future contract. He withdraws from the money market to do this just as the interest rate rises. In short, the long position in the future is hurt by positive correlation between $\frac{1}{\beta(T)}$ and $S(T)$. The buyer of the future is compensated by a reduction of the future price below the forward price.

## 27.6 Backwardation and contango

Suppose

$$dS(t) = \mu S(t)\ dt + \sigma S(t)\ dW(t).$$

Define $\theta = \frac{\mu - r}{\sigma}$, $\widetilde{W}(t) = \theta t + W(t)$,

$$\begin{aligned} Z(T) &= \exp\{-\theta W(T) - \tfrac{1}{2}\theta^2 T\} \\ \widetilde{I\!P}(A) &= \int_A Z(T)\ dI\!P, \quad \forall A \in \mathcal{F}(T). \end{aligned}$$

Then $\widetilde{W}$ is a Brownian motion under $\widetilde{I\!P}$, and

$$dS(t) = rS(t)\ dt + \sigma S(t)\ d\widetilde{W}(t).$$

We have

$$\begin{aligned} \beta(t) &= e^{rt} \\ S(t) &= S(0)\exp\{(\mu - \tfrac{1}{2}\sigma^2)t + \sigma W(t)\} \\ &= S(0)\exp\{(r - \tfrac{1}{2}\sigma^2)t + \sigma \widetilde{W}(t)\} \end{aligned}$$

Because $\frac{1}{\beta(T)} = e^{-rT}$ is nonrandom, $S(T)$ and $\frac{1}{\beta(T)}$ are uncorrelated under $\widetilde{I\!P}$. Therefore,

$$\begin{aligned} \Phi(t) &= \widetilde{I\!E}[S(T)\Big|\mathcal{F}(t)] \\ &= F(t) \\ &= \frac{S(t)}{B(t,T)} = e^{r(T-t)}S(t). \end{aligned}$$

The expected future spot price of the stock under $I\!P$ is

$$\begin{aligned} I\!ES(T) &= S(0)e^{\mu T} I\!E\left[\exp\left\{-\tfrac{1}{2}\sigma^2 T + \sigma W(T)\right\}\right] \\ &= e^{\mu T}S(0). \end{aligned}$$

The future price at time $0$ is

$$\Phi(0) = e^{rT} S(0).$$

If $\mu > r$, then $\Phi(0) < \mathbb{E}S(T)$. This situation is called *normal backwardation* (see Hull). If $\mu < r$, then $\Phi(0) > \mathbb{E}S(T)$. This is called *contango.*

# Chapter 28

# Term-structure models

Throughout this discussion, $\{W(t);\ 0 \le t \le T^*\}$ is a Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\{F(t);\ 0 \le t \le T^*\}$ is the filtration generated by $W$.

Suppose we are given an adapted *interest rate process* $\{r(t);\ 0 \le t \le T^*\}$. We define the accumulation factor

$$\beta(t) = \exp\left\{\int_0^t r(u)\, du\right\}, \quad 0 \le t \le T^*.$$

In a term-structure model, we take the zero-coupon bonds ("zeroes") of various maturities to be the primitive assets. We assume these bonds are default-free and pay \$1 at maturity. For $0 \le t \le T \le T^*$, let

$$B(t,T) \ = \ \text{price at time } t \text{ of the zero-coupon bond paying \$1 at time } T.$$

**Theorem 0.67 (Fundamental Theorem of Asset Pricing)** *A term structure model is free of arbitrage if and only if there is a probability measure $\widetilde{I\!\!P}$ on $\Omega$ (a risk-neutral measure) with the same probability-zero sets as $I\!\!P$ (i.e.,* equivalent *to $I\!\!P$), such that for each $T \in (0, T^*]$, the process*

$$\frac{B(t,T)}{\beta(t)}, \quad 0 \le t \le T,$$

*is a martingale under $\widetilde{I\!\!P}$.*

**Remark 28.1** We shall always have

$$dB(t,T) = \mu(t,T)B(t,T)\, dt + \rho(t,T)B(t,T)\, dW(t), \quad 0 \le t \le T,$$

for some functions $\mu(t,T)$ and $\rho(t,T)$. Therefore

$$\begin{aligned} d\left(\frac{B(t,T)}{\beta(t)}\right) &= B(t,T)\, d\left(\frac{1}{\beta(t)}\right) + \frac{1}{\beta(t)}\, dB(t,T) \\ &= [\mu(t,T) - r(t)]\frac{B(t,T)}{\beta(t)}\, dt + \rho(t,T)\frac{B(t,T)}{\beta(t)}\, dW(t), \end{aligned}$$

so $I\!\!P$ is a risk-neutral measure if and only if $\mu(t,T)$, the mean rate of return of $B(t,T)$ under $I\!\!P$, is the interest rate $r(t)$. If the mean rate of return of $B(t,T)$ under $I\!\!P$ is not $r(t)$ at each time $t$ and for each maturity $T$, we should change to a measure $\widetilde{I\!\!P}$ under which the mean rate of return is $r(t)$. If such a measure does not exist, then the model admits an arbitrage by trading in zero-coupon bonds.

## 28.1 Computing arbitrage-free bond prices: first method

Begin with a stochastic differential equation (SDE)

$$dX(t) = a(t, X(t))\,dt + b(t, X(t))\,dW(t).$$

The solution $X(t)$ is the *factor*. If we want to have $n$-factors, we let $W$ be an $n$-dimensional Brownian motion and let $X$ be an $n$-dimensional process. We let the interest rate $r(t)$ be a function of $X(t)$. In the usual one-factor models, we take $r(t)$ to be $X(t)$ (e.g., Cox-Ingersoll-Ross, Hull-White).

Now that we have an interest rate process $\{r(t);\ 0 \le t \le T^*\}$, we define the zero-coupon bond prices to be

$$\begin{aligned} B(t,T) &= \beta(t)\ I\!\!E\left[\frac{1}{\beta(T)}\middle|\mathcal{F}(t)\right] \\ &= I\!\!E\left[\exp\left\{-\int_t^T r(u)\,du\right\}\middle|\mathcal{F}(t)\right], \quad 0 \le t \le T \le T^*. \end{aligned}$$

We showed in Chapter 27 that

$$dB(t,T) = r(t)B(t,T)\,dt + \beta(t)\gamma(t)\,dW(t)$$

for some process $\gamma$. Since $B(t,T)$ has mean rate of return $r(t)$ under $I\!\!P$, $I\!\!P$ is a risk-neutral measure and there is no arbitrage.

## 28.2 Some interest-rate dependent assets

**Coupon-paying bond:** Payments $P_1, P_2, \ldots, P_n$ at times $T_1, T_2, \ldots, T_n$. Price at time $t$ is

$$\sum_{\{k: t < T_k\}} P_k B(t, T_k).$$

**Call option on a zero-coupon bond:** Bond matures at time $T$. Option expires at time $T_1 < T$. Price at time $t$ is

$$\beta(t)\ I\!\!E\left[\frac{1}{\beta(T_1)}(B(T_1,T) - K)^+\middle|\mathcal{F}(t)\right], \quad 0 \le t \le T_1.$$

## 28.3 Terminology

**Definition 28.1 (Term-structure model)** Any mathematical model which determines, at least theoretically, the stochastic processes

$$B(t,T), \quad 0 \le t \le T,$$

for all $T \in (0, T^*]$.

**Definition 28.2 (Yield to maturity)** For $0 \le t \le T \le T^*$, the *yield to maturity* $Y(t,T)$ is the $\mathcal{F}(t)$-measurable random-variable satisfying

$$B(t,T) \exp\{(T-t)Y(t,T)\} = 1,$$

or equivalently,

$$Y(t,T) = -\frac{1}{T-t} \log B(t,T).$$

Determining

$$B(t,T), \quad 0 \le t \le T \le T^*,$$

is equivalent to determining

$$Y(t,T), \quad 0 \le t \le T \le T^*.$$

## 28.4 Forward rate agreement

Let $0 \le t \le T < T + \epsilon \le T^*$ be given. Suppose you want to borrow \$1 at time $T$ with repayment (plus interest) at time $T + \epsilon$, at an interest rate agreed upon at time $t$. To synthesize a *forward-rate agreement* to do this, at time $t$ buy a $T$-maturity zero and short $\frac{B(t,T)}{B(t,T+\epsilon)}$ $(T+\epsilon)$-maturity zeroes. The value of this portfolio at time $t$ is

$$B(t,T) - \frac{B(t,T)}{B(t,T+\epsilon)} B(t,T+\epsilon) = 0.$$

At time $T$, you receive \$1 from the $T$-maturity zero. At time $T+\epsilon$, you pay \$ $\frac{B(t,T)}{B(t,T+\epsilon)}$. The effective interest rate on the dollar you receive at time $T$ is $R(t,T,T+\epsilon)$ given by

$$\frac{B(t,T)}{B(t,T+\epsilon)} = \exp\{\epsilon\, R(t,T,T+\epsilon)\},$$

or equivalently,

$$R(t,T,T+\epsilon) = -\frac{\log B(t,T+\epsilon) - \log B(t,T)}{\epsilon}.$$

The *forward rate* is

$$f(t,T) = \lim_{\epsilon \downarrow 0} R(t,T,T+\epsilon) = -\frac{\partial}{\partial T} \log B(t,T). \tag{4.1}$$

This is the instantaneous interest rate, agreed upon at time $t$, for money borrowed at time $T$.

Integrating the above equation, we obtain

$$\begin{aligned}\int_t^T f(t,u)\,du &= -\int_t^T \frac{\partial}{\partial u}\log B(t,u)\,du\\ &= -\log B(t,u)\Big|_{u=t}^{u=T}\\ &= -\log B(t,T),\end{aligned}$$

so

$$B(t,T)=\exp\left\{-\int_t^T f(t,u)\,du\right\}.$$

You can agree at time $t$ to receive interest rate $f(t,u)$ at each time $u\in[t,T]$. If you invest \$ $B(t,T)$ at time $t$ and receive interest rate $f(t,u)$ at each time $u$ between $t$ and $T$, this will grow to

$$B(t,T)\exp\left\{\int_t^T f(t,u)\,du\right\}=1$$

at time $T$.

## 28.5 Recovering the interest $r(t)$ from the forward rate

$$\begin{aligned}B(t,T) &= I\!\!E\left[\exp\left\{-\int_t^T r(u)\,du\right\}\Big|\mathcal{F}(t)\right],\\ \frac{\partial}{\partial T}B(t,T) &= I\!\!E\left[-r(T)\exp\left\{-\int_t^T r(u)\,du\right\}\Big|\mathcal{F}(t)\right],\\ \frac{\partial}{\partial T}B(t,T)\Big|_{T=t} &= I\!\!E\left[-r(t)\Big|\mathcal{F}(t)\right]=-r(t).\end{aligned}$$

On the other hand,

$$\begin{aligned}B(t,T) &= \exp\left\{-\int_t^T f(t,u)\,du\right\},\\ \frac{\partial}{\partial T}B(t,T) &= -f(t,T)\exp\left\{-\int_t^T f(t,u)\,du\right\},\\ \frac{\partial}{\partial T}B(t,T)\Big|_{T=t} &= -f(t,t).\end{aligned}$$

Conclusion: $r(t)=f(t,t)$.

## 28.6 Computing arbitrage-free bond prices: Heath-Jarrow-Morton method

For each $T \in (0, T^*]$, let the forward rate be given by

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)\, du + \int_0^t \sigma(u,T)\, dW(u), \quad 0 \le t \le T.$$

Here $\{\alpha(u,T);\ 0 \le u \le T\}$ and $\{\sigma(u,T);\ 0 \le u \le T\}$ are adapted processes.
In other words,

$$df(t,T) = \alpha(t,T)\, dt + \sigma(t,T)\, dW(t).$$

Recall that

$$B(t,T) = \exp\left\{ -\int_t^T f(t,u)\, du \right\}.$$

Now

$$\begin{aligned} d\left\{ -\int_t^T f(t,u)\, du \right\} &= f(t,t)\, dt - \int_t^T df(t,u)\, du \\ &= r(t)\, dt - \int_t^T [\alpha(t,u)\, dt + \sigma(t,u)\, dW(t)]\, du \\ &= r(t)\, dt - \underbrace{\left[ \int_t^T \alpha(t,u)\, du \right]}_{\alpha^*(t,T)} dt - \underbrace{\left[ \int_t^T \sigma(t,u)\, du \right]}_{\sigma^*(t,T)} dW(t) \\ &= r(t)\, dt - \alpha^*(t,T)\, dt - \sigma^*(t,T)\, dW(t). \end{aligned}$$

Let

$$g(x) = e^x,\ g'(x) = e^x,\ g''(x) = e^x.$$

Then

$$B(t,T) = g\left( -\int_t^T f(t,u)\, du \right),$$

and

$$\begin{aligned} dB(t,T) &= dg\left( -\int_t^T f(t,u)\, du \right) \\ &= g'\left( -\int_t^T f(t,u)\, du \right) (r\, dt - \alpha^*\, dt - \sigma^*\, dW) \\ &\quad + \tfrac{1}{2} g''\left( -\int_t^T f(t,u)\, du \right) (\sigma^*)^2\, dt \\ &= B(t,T) \left[ r(t) - \alpha^*(t,T) + \tfrac{1}{2} (\sigma^*(t,T))^2 \right] dt \\ &\quad - \sigma^*(t,T) B(t,T)\, dW(t). \end{aligned}$$

## 28.7 Checking for absence of arbitrage

$I\!P$ is a risk-neutral measure if and only if

$$\alpha^*(t,T) = \tfrac{1}{2}\left(\sigma^*(t,T)\right)^2, \quad 0 \le t \le T \le T^*,$$

i.e.,

$$\int_t^T \alpha(t,u)\,du = \tfrac{1}{2}\left(\int_t^T \sigma(t,u)\,du\right)^2, \quad 0 \le t \le T \le T^*. \tag{7.1}$$

Differentiating this w.r.t. $T$, we obtain

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)\,du, \quad 0 \le t \le T \le T^*. \tag{7.2}$$

Not only does (7.1) imply (7.2), (7.2) also implies (7.1). This will be a homework problem.

Suppose (7.1) does not hold. Then $I\!P$ is not a risk-neutral measure, but there might still be a risk-neutral measure. Let $\{\theta(t);\ 0 \le t \le T^*\}$ be an adapted process, and define

$$\begin{aligned}
\widetilde{W}(t) &= \int_0^t \theta(u)\,du \;+\; W(t),\\
Z(t) &= \exp\left\{-\int_0^t \theta(u)\,dW(u) \;-\; \tfrac{1}{2}\int_0^t \theta^2(u)\,du\right\},\\
\widetilde{I\!P}(A) &= \int_A Z(T^*)\,dI\!P \quad \forall A \in \mathcal{F}(T^*).
\end{aligned}$$

Then

$$\begin{aligned}
dB(t,T) &= B(t,T)\left[r(t) - \alpha^*(t,T) + \tfrac{1}{2}(\sigma^*(t,T))^2\right]\,dt\\
&\quad - \sigma^*(t,T)B(t,T)\,dW(t)\\
&= B(t,T)\left[r(t) - \alpha^*(t,T) + \tfrac{1}{2}(\sigma^*(t,T))^2 + \sigma^*(t,T)\theta(t)\right]\,dt\\
&\quad - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t), \quad 0 \le t \le T.
\end{aligned}$$

In order for $B(t,T)$ to have mean rate of return $r(t)$ under $\widetilde{I\!P}$, we must have

$$\alpha^*(t,T) = \tfrac{1}{2}(\sigma^*(t,T))^2 + \sigma^*(t,T)\theta(t), \quad 0 \le t \le T \le T^*. \tag{7.3}$$

Differentiation w.r.t. $T$ yields the equivalent condition

$$\alpha(t,T) = \sigma(t,T)\sigma^*(t,T) + \sigma(t,T)\theta(t), \quad 0 \le t \le T \le T^*. \tag{7.4}$$

**Theorem 7.68 (Heath-Jarrow-Morton)** *For each $T \in (0,T^*]$, let $\alpha(u,T),\ 0 \le u \le T$, and $\sigma(u,T), 0 \le u \le T$, be adapted processes, and assume $\sigma(u,T) > 0$ for all $u$ and $T$. Let $f(0,T),\ 0 \le t \le T^*$, be a deterministic function, and define*

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)\,du + \int_0^t \sigma(u,T)\,dW(u).$$

*Then* $f(t,T),\ 0 \le t \le T \le T^*$ *is a family of forward rate processes for a term-structure model without arbitrage if and only if there is an adapted process* $\theta(t),\ 0 \le t \le T^*$*, satisfying (7.3), or equivalently, satisfying (7.4).*

**Remark 28.2** Under $I\!P$, the zero-coupon bond with maturity $T$ has mean rate of return

$$r(t) - \alpha^*(t,T) + \tfrac{1}{2}(\sigma^*(t,T))^2$$

and volatility $\sigma^*(t,T)$. The excess mean rate of return, above the interest rate, is

$$-\alpha^*(t,T) + \tfrac{1}{2}(\sigma^*(t,T))^2,$$

and when normalized by the volatility, this becomes the *market price of risk*

$$\frac{-\alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2}{\sigma^*(t,T)}.$$

The no-arbitrage condition is that this market price of risk at time $t$ does not depend on the maturity $T$ of the bond. We can then set

$$\theta(t) = -\left[\frac{-\alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2}{\sigma^*(t,T)}\right],$$

and (7.3) is satisfied.

(The remainder of this chapter was taught Mar 21)

Suppose the market price of risk does not depend on the maturity $T$, so we can solve (7.3) for $\theta$. Plugging this into the stochastic differential equation for $B(t,T)$, we obtain for every maturity $T$:

$$dB(t,T) = r(t)B(t,T)\,dt - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t).$$

Because (7.4) is equivalent to (7.3), we may plug (7.4) into the stochastic differential equation for $f(t,T)$ to obtain, for every maturity $T$:

$$\begin{aligned} df(t,T) &= [\sigma(t,T)\sigma^*(t,T) + \sigma(t,T)\theta(t)]\,dt + \sigma(t,T)\,dW(t) \\ &= \sigma(t,T)\sigma^*(t,T)\,dt + \sigma(t,T)\,d\widetilde{W}(t). \end{aligned}$$

## 28.8 Implementation of the Heath-Jarrow-Morton model

Choose

$$\begin{aligned} \sigma^*(t,T), &\quad 0 \le t \le T \le T^*, \\ \theta(t), &\quad 0 \le t \le T^*. \end{aligned}$$

These may be stochastic processes, but are usually taken to be deterministic functions. Define

$$\begin{aligned}
\alpha(t,T) &= \sigma(t,T)\sigma^*(t,T) + \sigma(t,T)\theta(t), \\
\widetilde{W}(t) &= \int_0^t \theta(u)\ du + W(t), \\
Z(t) &= \exp\left\{ -\int_0^t \theta(u)\ dW(u) - \tfrac{1}{2}\int_0^t \theta^2(u)\ du \right\}, \\
\widetilde{I\!P}(A) &= \int_A Z(T^*)\ dI\!P \quad \forall A \in \mathcal{F}(T^*).
\end{aligned}$$

Let $f(0,T),\ 0 \le T \le T^*$, be determined by the market; recall from equation (4.1):

$$f(0,T) = -\frac{\partial}{\partial T}\log B(0,T), \quad 0 \le T \le T^*.$$

Then $f(t,T)$ for $0 \le t \le T$ is determined by the equation

$$df(t,T) = \sigma(t,T)\sigma^*(t,T)\ dt + \sigma(t,T)\ d\widetilde{W}(t), \tag{8.1}$$

this determines the interest rate process

$$r(t) = f(t,t), \quad 0 \le t \le T^*, \tag{8.2}$$

and then the zero-coupon bond prices are determined by the initial conditions $B(0,T),\ 0 \le T \le T^*$, gotten from the market, combined with the stochastic differential equation

$$dB(t,T) = r(t)B(t,T)\ dt - \sigma^*(t,T)B(t,T)\ d\widetilde{W}(t). \tag{8.3}$$

Because all pricing of interest rate dependent assets will be done under the risk-neutral measure $\widetilde{I\!P}$, under which $\widetilde{W}$ is a Brownian motion, we have written (8.1) and (8.3) in terms of $\widetilde{W}$ rather than $W$. Written this way, it is apparent that neither $\theta(t)$ nor $\alpha(t,T)$ will enter subsequent computations. The only process which matters is $\sigma(t,T),\ 0 \le t \le T \le T^*$, and the process

$$\sigma^*(t,T) = \int_t^T \sigma(t,u)\ du, \quad 0 \le t \le T \le T^*, \tag{8.4}$$

obtained from $\sigma(t,T)$.

From (8.3) we see that $\sigma^*(t,T)$ is the volatility at time $t$ of the zero coupon bond maturing at time $T$. Equation (8.4) implies

$$\sigma^*(T,T) = 0, \quad 0 \le T \le T^*. \tag{8.5}$$

This is because $B(T,T) = 1$ and so as $t$ approaches $T$ (from below), the volatility in $B(t,T)$ must vanish.

In conclusion, to implement the HJM model, it suffices to have the initial market data $B(0,T),\ 0 \le T \le T^*$, and the volatilities

$$\sigma^*(t,T), \quad 0 \le t \le T \le T^*.$$

We require that $\sigma^*(t,T)$ be differentiable in $T$ and satisfy (8.5). We can then define

$$\sigma(t,T) = \frac{\partial}{\partial T}\sigma^*(t,T),$$

and (8.4) will be satisfied because

$$\sigma^*(t,T) = \sigma^*(t,T) - \sigma^*(t,t) = \int_t^T \frac{\partial}{\partial u}\sigma^*(t,u)\, du.$$

We then let $\widetilde{W}$ be a Brownian motion under a probability measure $\widetilde{I\!P}$, and we let $B(t,T),\ 0 \le t \le T \le T^*$, be given by (8.3), where $r(t)$ is given by (8.2) and $f(t,T)$ by (8.1). In (8.1) we use the initial conditions

$$f(0,T) = -\frac{\partial}{\partial T}\log B(0,T), \quad 0 \le T \le T^*.$$

**Remark 28.3** It is customary in the literature to write $W$ rather than $\widetilde{W}$ and $I\!P$ rather than $\widetilde{I\!P}$, so that $I\!P$ is the symbol used for the risk-neutral measure and no reference is ever made to the market measure. The only parameter which must be estimated from the market is the bond volatility $\sigma^*(t,T)$, and volatility is unaffected by the change of measure.

# Chapter 29

# Gaussian processes

**Definition 29.1 (Gaussian Process)** A *Gaussian process* $X(t)$, $t \geq 0$, is a stochastic process with the property that for every set of times $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, the set of random variables

$$X(t_1), X(t_2), \ldots, X(t_n)$$

is jointly normally distributed.

**Remark 29.1** If $X$ is a Gaussian process, then its distribution is determined by its *mean function*

$$m(t) = \mathbb{E}X(t)$$

and its *covariance function*

$$\rho(s,t) = \mathbb{E}[(X(s) - m(s)) \cdot (X(t) - m(t))].$$

Indeed, the joint density of $X(t_1), \ldots, X(t_n)$ is

$$\begin{aligned}&\mathbb{P}\{X(t_1) \in dx_1, \ldots, X(t_n) \in dx_n\} \\ &= \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left\{-\tfrac{1}{2}(\mathbf{x} - m(\mathbf{t})) \cdot \Sigma^{\mathbf{-1}} \cdot (\mathbf{x} - \mathbf{m}(\mathbf{t}))^{\mathbf{T}}\right\} \; dx_1 \ldots \; dx_n,\end{aligned}$$

where $\Sigma$ is the covariance matrix

$$\Sigma = \begin{bmatrix} \rho(t_1,t_1) & \rho(t_1,t_2) & \ldots & \rho(t_1,t_n) \\ \rho(t_2,t_1) & \rho(t_2,t_2) & \ldots & \rho(t_2,t_n) \\ \ldots & \ldots & \ldots & \ldots \\ \rho(t_n,t_1) & \rho(t_n,t_2) & \ldots & \rho(t_n,t_n) \end{bmatrix}$$

$\mathbf{x}$ is the row vector $[x_1, x_2, \ldots, x_n]$, $\mathbf{t}$ is the row vector $[t_1, t_2, \ldots, t_n]$, and $m(\mathbf{t}) = [m(t_1), m(t_2), \ldots, m(t_n)]$. The moment generating function is

$$\mathbb{E}\exp\left\{\sum_{k=1}^{n} u_k X(t_k)\right\} = \exp\left\{\mathbf{u} \cdot m(\mathbf{t})^{\mathbf{T}} + \tfrac{1}{2}\mathbf{u} \cdot \Sigma \cdot \mathbf{u}^{\mathbf{T}}\right\},$$

where $\mathbf{u} = [u_1, u_2, \ldots, u_n]$.

## 29.1 An example: Brownian Motion

Brownian motion $W$ is a Gaussian process with $m(t) = 0$ and $\rho(s,t) = s \wedge t$. Indeed, if $0 \le s \le t$, then

$$\begin{aligned}
\rho(s,t) &= I\!E\,[W(s)W(t)] = I\!E\left[W(s)\,(W(t) - W(s)) + W^2(s)\right] \\
&= I\!EW(s).I\!E\,(W(t) - W(s)) + I\!EW^2(s) \\
&= I\!EW^2(s) \\
&= s \wedge t.
\end{aligned}$$

To prove that a process is Gaussian, one must show that $X(t_1), \ldots, X(t_n)$ has either a density or a moment generating function of the appropriate form. We shall use the m.g.f., and shall cheat a bit by considering only two times, which we usually call $s$ and $t$. We will want to show that

$$I\!E\exp\{u_1X(s) + u_2X(t)\} = \exp\left\{u_1m_1 + u_2m_2 + \tfrac{1}{2}[u_1\ u_2]\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}\right\}.$$

**Theorem 1.69 (Integral w.r.t. a Brownian)** *Let $W(t)$ be a Brownian motion and $\delta(t)$ a nonrandom function. Then*

$$X(t) = \int_0^t \delta(u)\,dW(u)$$

*is a Gaussian process with $m(t) = 0$ and*

$$\rho(s,t) = \int_0^{s\wedge t} \delta^2(u)\,du.$$

**Proof:** (Sketch.) We have

$$dX = \delta\,dW.$$

Therefore,

$$\begin{aligned}
de^{uX(s)} &= ue^{uX(s)}\delta(s)\,dW(s) + \tfrac{1}{2}u^2e^{uX(s)}\delta^2(s)\,ds, \\
e^{uX(s)} &= e^{uX(0)} + \underbrace{u\int_0^s e^{uX(v)}\delta(v)\,dW(v)}_{\text{Martingale}} + \tfrac{1}{2}u^2\int_0^s e^{uX(v)}\delta^2(v)\,dv, \\
I\!Ee^{uX(s)} &= 1 + \tfrac{1}{2}u^2\int_0^s \delta^2(v)\,I\!Ee^{uX(v)}\,dv, \\
\frac{d}{ds}I\!Ee^{uX(s)} &= \tfrac{1}{2}u^2\delta^2(s)\,I\!Ee^{uX(s)}, \\
I\!Ee^{uX(s)} &= e^{uX(0)}\exp\left\{\tfrac{1}{2}u^2\int_0^s \delta^2(v)\,dv\right\} \qquad (1.1)\\
&= \exp\left\{\tfrac{1}{2}u^2\int_0^s \delta^2(v)\,dv\right\}.
\end{aligned}$$

This shows that $X(s)$ is normal with mean 0 and variance $\int_0^s \delta^2(v)\,dv$.

Now let $0 \le s < t$ be given. Just as before,

$$de^{uX(t)} = ue^{uX(t)}\delta(t)\ dW(t) + \tfrac{1}{2}u^2 e^{uX(t)}\delta^2(t)\ dt.$$

Integrate from $s$ to $t$ to get

$$e^{uX(t)} = e^{uX(s)} + u\int_s^t \delta(v)e^{uX(v)}\ dW(v) + \tfrac{1}{2}u^2\int_s^t \delta^2(v)e^{uX(v)}\ dv.$$

Take $I\!E[\ldots|\mathcal{F}(s)]$ conditional expectations and use the martingale property

$$\begin{aligned} I\!E\left[\int_s^t \delta(v)e^{uX(v)}\ dW(v)\middle|\mathcal{F}(s)\right] &= I\!E\left[\int_0^t \delta(v)e^{uX(v)}\ dW(v)\middle|\mathcal{F}(s)\right] - \int_0^s \delta(v)e^{uX(v)}\ dW(v) \\ &= 0 \end{aligned}$$

to get

$$\begin{aligned} I\!E\left[e^{uX(t)}\middle|\mathcal{F}(s)\right] &= e^{uX(s)} + \tfrac{1}{2}u^2\int_s^t \delta^2(v)I\!E\left[e^{uX(v)}\middle|\mathcal{F}(s)\right]\ dv \\ \frac{d}{dt}I\!E\left[e^{uX(t)}\middle|\mathcal{F}(s)\right] &= \tfrac{1}{2}u^2\delta^2(t)I\!E\left[e^{uX(t)}\middle|\mathcal{F}(s)\right], \quad t \ge s. \end{aligned}$$

The solution to this ordinary differential equation with initial time $s$ is

$$I\!E\left[e^{uX(t)}\middle|\mathcal{F}(s)\right] = e^{uX(s)}\exp\left\{\tfrac{1}{2}u^2\int_s^t \delta^2(v)\ dv\right\}, \quad t \ge s. \tag{1.2}$$

We now compute the m.g.f. for $(X(s), X(t))$, where $0 \le s \le t$:

$$\begin{aligned} I\!E\left[e^{u_1X(s)+u_2X(t)}\middle|\mathcal{F}(s)\right] &= e^{u_1X(s)}I\!E\left[e^{u_2X(t)}\middle|\mathcal{F}(s)\right] \\ &\overset{(1.2)}{=} e^{(u_1+u_2)X(s)}\exp\left\{\tfrac{1}{2}u_2^2\int_s^t \delta^2(v)\ dv\right\}, \\ I\!E\left[e^{u_1X(s)+u_2X(t)}\right] &= I\!E\left\{I\!E\left[e^{u_1X(s)+u_2X(t)}\middle|\mathcal{F}(s)\right]\right\} \\ &= I\!E\left\{e^{(u_1+u_2)X(s)}\right\}.\exp\left\{\tfrac{1}{2}u_2^2\int_s^t \delta^2(v)\ dv\right\} \\ &\overset{(1.1)}{=} \exp\left\{\tfrac{1}{2}(u_1+u_2)^2\int_0^s \delta^2(v)\ dv \ + \ \tfrac{1}{2}u_2^2\int_s^t \delta^2(v)\ dv\right\} \\ &= \exp\left\{\tfrac{1}{2}(u_1^2+2u_1u_2)\int_0^s \delta^2(v)\ dv \ + \ \tfrac{1}{2}u_2^2\int_0^t \delta^2(v)\ dv\right\} \\ &= \exp\left\{\tfrac{1}{2}[u_1\ u_2]\begin{bmatrix} \int_0^s \delta^2 & \int_0^s \delta^2 \\ \int_0^s \delta^2 & \int_0^t \delta^2 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right\}. \end{aligned}$$

This shows that $(X(s), X(t))$ is jointly normal with $I\!EX(s) = I\!EX(t) = 0$,

$$I\!EX^2(s) = \int_0^s \delta^2(v)\ dv, \qquad I\!EX^2(t) = \int_0^t \delta^2(v)\ dv,$$

$$I\!E[X(s)X(t)] = \int_0^s \delta^2(v)\ dv.$$

■

**Remark 29.2** The hard part of the above argument, and the reason we use moment generating functions, is to prove the normality. The computation of means and variances does not require the use of moment generating functions. Indeed,

$$X(t) = \int_0^t \delta(u)\, dW(u)$$

is a martingale and $X(0) = 0$, so

$$m(t) = \mathbb{E}X(t) = 0 \quad \forall t \geq 0.$$

For fixed $s \geq 0$,

$$\mathbb{E}X^2(s) = \int_0^s \delta^2(v)\, dv$$

by the Itô isometry. For $0 \leq s \leq t$,

$$\begin{aligned}
\mathbb{E}[X(s)(X(t) - X(s))] &= \mathbb{E}\left[\mathbb{E}\left\{X(s)(X(t) - X(s))\middle|\mathcal{F}(s)\right\}\right] \\
&= \mathbb{E}\left[X(s)\underbrace{\left(\mathbb{E}\left[X(t)\middle|\mathcal{F}(s)\right] - X(s)\right)}_{0}\right] \\
&= 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}[X(s)X(t)] &= \mathbb{E}[X(s)(X(t) - X(s)) + X^2(s)] \\
&= \mathbb{E}X^2(s) = \int_0^s \delta^2(v)\, dv.
\end{aligned}$$

If $\delta$ were a stochastic proess, the Itô isometry says

$$\mathbb{E}X^2(s) = \int_0^s \mathbb{E}\delta^2(v)\, dv$$

and the same argument used above shows that for $0 \leq s \leq t$,

$$\mathbb{E}[X(s)X(t)] = \mathbb{E}X^2(s) = \int_0^s \mathbb{E}\delta^2(v)\, dv.$$

However, when $\delta$ is stochastic, $X$ is not necessarily a Gaussian process, so its distribution is not determined from its mean and covariance functions.

**Remark 29.3** When $\delta$ is nonrandom,

$$X(t) = \int_0^t \delta(u)\, dW(u)$$

is also Markov. We proved this before, but note again that the Markov property follows immediately from (1.2). The equation (1.2) says that conditioned on $\mathcal{F}(s)$, the distribution of $X(t)$ depends only on $X(s)$; in fact, $X(t)$ is normal with mean $X(s)$ and variance $\int_s^t \delta^2(v)\, dv$.

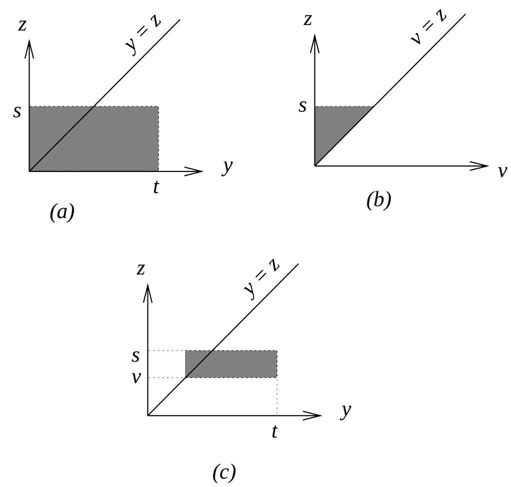



Figure 29.1: *Range of values of $y, z, v$ for the integrals in the proof of Theorem 1.70.*

**Theorem 1.70** *Let $W(t)$ be a Brownian motion, and let $\delta(t)$ and $h(t)$ be nonrandom functions. Define*

$$X(t) = \int_0^t \delta(u)\, dW(u), \quad Y(t) = \int_0^t h(u)X(u)\, du.$$

*Then $Y$ is a Gaussian process with mean function $m_Y(t) = 0$ and covariance function*

$$\rho_Y(s,t) = \int_0^{s\wedge t} \delta^2(v) \left(\int_v^s h(y)\, dy\right) \left(\int_v^t h(y)\, dy\right)\, dv. \tag{1.3}$$

**Proof:** (Partial) Computation of $\rho_Y(s,t)$: Let $0 \le s \le t$ be given. It is shown in a homework problem that $(Y(s), Y(t))$ is a jointly normal pair of random variables. Here we observe that

$$m_Y(t) = I\!\!E Y(t) = \int_0^t h(u)\, I\!\!E X(u)\, du = 0,$$

and we verify that (1.3) holds.

We have

$$
\begin{aligned}
\rho_Y(s,t) &= \mathbb{E}\left[Y(s)Y(t)\right] \\
&= \mathbb{E}\left[\int_0^s h(y)X(y)\ dy.\int_0^t h(z)X(z)\ dz\right] \\
&= \mathbb{E}\int_0^s\int_0^t h(y)h(z)X(y)X(z)\ dy\ dz \\
&= \int_0^s\int_0^t h(y)h(z)\mathbb{E}\left[X(y)X(z)\right]\ dy\ dz \\
&= \int_0^s\int_0^t h(y)h(z)\int_0^{y\wedge z}\delta^2(v)\ dv\ dy\ dz \\
&= \int_0^s\int_z^t h(y)h(z)\left(\int_0^z\delta^2(v)\ dv\right)\ dy\ dz \\
&\quad + \int_0^s\int_y^s h(y)h(z)\left(\int_0^y\delta^2(v)\ dv\right)\ dz\ dy \quad \text{(See Fig. 29.1(a))} \\
&= \int_0^s h(z)\left(\int_z^t h(y)\ dy\right)\left(\int_0^z\delta^2(v)\ dv\right)\ dz \\
&\quad + \int_0^s h(y)\left(\int_y^s h(z)\ dz\right)\left(\int_0^y\delta^2(v)\ dv\right)\ dy \\
&= \int_0^s\int_0^z h(z)\delta^2(v)\left(\int_z^t h(y)\ dy\right)\ dv\ dz \\
&\quad + \int_0^s\int_0^y h(y)\delta^2(v)\left(\int_y^s h(z)\ dz\right)\ dv\ dy \\
&= \int_0^s\int_v^s h(z)\delta^2(v)\left(\int_z^t h(y)\ dy\right)\ dz\ dv \\
&\quad + \int_0^s\int_v^s h(y)\delta^2(v)\left(\int_y^s h(z)\ dz\right)\ dy\ dv \quad \text{(See Fig. 29.1(b))} \\
&= \int_0^s\delta^2(v)\left(\int_v^s\int_z^t h(y)h(z)\ dy\ dz\right)\ dv \\
&\quad + \int_0^s\delta^2(v)\left(\int_v^s\int_y^s h(y)h(z)\ dz\ dy\right)\ dv \\
&= \int_0^s\delta^2(v)\left(\int_v^s\int_v^t h(y)h(z)\ dy\ dz\right)\ dv \quad \text{(See Fig. 29.1(c))} \\
&= \int_0^s\delta^2(v)\left(\int_v^s h(y)\ dy\right)\left(\int_v^t h(z)\ dz\right)\ dv \\
&= \int_0^s\delta^2(v)\left(\int_v^s h(y)\ dy\right)\left(\int_v^t h(y)\ dy\right)\ dv
\end{aligned}
$$

■

**Remark 29.4** Unlike the process $X(t) = \int_0^t \delta(u)\ dW(u)$, the process $Y(t) = \int_0^t X(u)\ du$ is neither Markov nor a martingale. For $0 \le s < t$,

$$\begin{aligned}
I\!E[Y(t)|\mathcal{F}(s)] &= \int_0^s h(u)X(u)\,du + I\!E\left[\int_s^t h(u)X(u)\,du\middle|\mathcal{F}(s)\right] \\
&= Y(s) + \int_s^t h(u) I\!E[X(u)\Big|\mathcal{F}(s)]\,du \\
&= Y(s) + \int_s^t h(u)X(s)\,du \\
&= Y(s) + X(s)\int_s^t h(u)\,du,
\end{aligned}$$

where we have used the fact that $X$ is a martingale. The conditional expectation $I\!E[Y(t)|\mathcal{F}(s)]$ is not equal to $Y(s)$, nor is it a function of $Y(s)$ alone.

# Chapter 30

# Hull and White model

Consider

$$dr(t) = (\alpha(t) - \beta(t)r(t))\ dt + \sigma(t)\ dW(t),$$

where $\alpha(t)$, $\beta(t)$ and $\sigma(t)$ are nonrandom functions of $t$.

We can solve the stochastic differential equation. Set

$$K(t) = \int_0^t \beta(u)\ du.$$

Then

$$\begin{aligned} d\left(e^{K(t)}r(t)\right) &= e^{K(t)}\Big(\beta(t)r(t)\ dt + dr(t)\Big) \\ &= e^{K(t)}\left(\alpha(t)\ dt + \sigma(t)\ dW(t)\right). \end{aligned}$$

Integrating, we get

$$e^{K(t)}r(t) = r(0) + \int_0^t e^{K(u)}\alpha(u)\ du + \int_0^t e^{K(u)}\sigma(u)\ dW(u),$$

so

$$r(t) = e^{-K(t)}\left[r(0) + \int_0^t e^{K(u)}\alpha(u)\ du + \int_0^t e^{K(u)}\sigma(u)\ dW(u)\right].$$

From Theorem 1.69 in Chapter 29, we see that $r(t)$ is a Gaussian process with mean function

$$m_r(t) = e^{-K(t)}\left[r(0) + \int_0^t e^{K(u)}\alpha(u)\ du\right] \tag{0.1}$$

and covariance function

$$\rho_r(s,t) = e^{-K(s)-K(t)}\int_0^{s\wedge t} e^{2K(u)}\sigma^2(u)\ du. \tag{0.2}$$

The process $r(t)$ is also Markov.

We want to study $\int_0^T r(t)\,dt$. To do this, we define

$$X(t) = \int_0^t e^{K(u)}\sigma(u)\,dW(u), \quad Y(T) = \int_0^T e^{-K(t)}X(t)\,dt.$$

Then

$$\begin{aligned}
r(t) &= e^{-K(t)}\left[r(0) + \int_0^t e^{K(u)}\alpha(u)\,du\right] + e^{-K(t)}X(t),\\
\int_0^T r(t)\,dt &= \int_0^T e^{-K(t)}\left[r(0) + \int_0^t e^{K(u)}\alpha(u)\,du\right]\,dt + Y(T).
\end{aligned}$$

According to Theorem 1.70 in Chapter 29, $\int_0^T r(t)\,dt$ is normal. Its mean is

$$I\!E\int_0^T r(t)\,dt = \int_0^T e^{-K(t)}\left[r(0) + \int_0^t e^{K(u)}\alpha(u)\,du\right]\,dt, \tag{0.3}$$

and its variance is

$$\begin{aligned}
\operatorname{var}\left(\int_0^T r(t)\,dt\right) &= I\!E Y^2(T)\\
&= \int_0^T e^{2K(v)}\sigma^2(v)\left(\int_v^T e^{-K(y)}\,dy\right)^2\,dv.
\end{aligned}$$

The price at time 0 of a zero-coupon bond paying \$1 at time $T$ is

$$\begin{aligned}
B(0,T) &= I\!E\exp\left\{-\int_0^T r(t)\,dt\right\}\\
&= \exp\left\{(-1)I\!E\int_0^T r(t)\,dt + \tfrac{1}{2}(-1)^2\operatorname{var}\left(\int_0^T r(t)\,dt\right)\right\}\\
&= \exp\Bigg\{-r(0)\int_0^T e^{-K(t)}\,dt - \int_0^T\int_0^t e^{-K(t)+K(u)}\alpha(u)\,du\,dt\\
&\qquad + \tfrac{1}{2}\int_0^T e^{2K(v)}\sigma^2(v)\left(\int_v^T e^{-K(y)}\,dy\right)^2\,dv\Bigg\}\\
&= \exp\{-r(0)C(0,T) - A(0,T)\},
\end{aligned}$$

where

$$\begin{aligned}
C(0,T) &= \int_0^T e^{-K(t)}\,dt,\\
A(0,T) &= \int_0^T\int_0^t e^{-K(t)+K(u)}\alpha(u)\,du\,dt - \tfrac{1}{2}\int_0^T e^{2K(v)}\sigma^2(v)\left(\int_v^T e^{-K(y)}\,dy\right)^2\,dv.
\end{aligned}$$

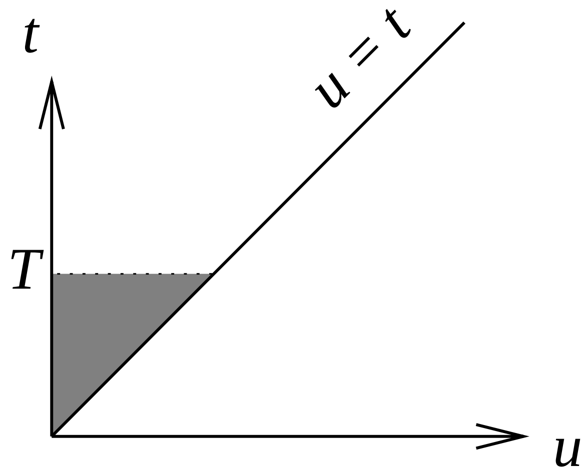



Figure 30.1: *Range of values of* $u, t$ *for the integral.*

## 30.1 Fiddling with the formulas

Note that (see Fig 30.1)

$$
\begin{aligned}
&\int_0^T \int_0^t e^{-K(t)+K(u)} \alpha(u)\, du\, dt \\
&= \int_0^T \int_u^T e^{-K(t)+K(u)} \alpha(u)\, dt\, du \\
(y=t;\ v=u) &= \int_0^T e^{K(v)} \alpha(v) \left( \int_v^T e^{-K(y)}\, dy \right) dv.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
A(0,T) &= \int_0^T \left[ e^{K(v)} \alpha(v) \left( \int_v^T e^{-K(y)}\, dy \right) - \tfrac{1}{2} e^{2K(v)} \sigma^2(v) \left( \int_v^T e^{-K(y)}\, dy \right)^2 \right] dv, \\
C(0,T) &= \int_0^T e^{-K(y)}\, dy, \\
B(0,T) &= \exp\left\{ -r(0)C(0,T) - A(0,T) \right\}.
\end{aligned}
$$

Consider the price at time $t \in [0,T]$ of the zero-coupon bond:

$$
B(t,T) = I\!E \left[ \exp\left\{ -\int_t^T r(u)\, du \right\} \middle| \mathcal{F}(t) \right].
$$

Because $r$ is a Markov process, this should be random only through a dependence on $r(t)$. In fact,

$$
B(t,T) = \exp\left\{ -r(t)C(t,T) - A(t,T) \right\},
$$

where

$$A(t,T) = \int_t^T \left[ e^{K(v)}\alpha(v) \left( \int_v^T e^{-K(y)}\, dy \right) - \tfrac{1}{2} e^{2K(v)}\sigma^2(v) \left( \int_v^T e^{-K(y)}\, dy \right)^2 \right] dv,$$

$$C(t,T) = e^{K(t)} \int_t^T e^{-K(y)}\, dy.$$

The reason for these changes is the following. We are now taking the initial time to be $t$ rather than zero, so it is plausible that $\int_0^T \ldots\, dv$ should be replaced by $\int_t^T \ldots\, dv$. Recall that

$$K(v) = \int_0^v \beta(u)\, du,$$

and this should be replaced by

$$K(v) - K(t) = \int_t^v \beta(u)\, du.$$

Similarly, $K(y)$ should be replaced by $K(y) - K(t)$. Making these replacements in $A(0,T)$, we see that the $K(t)$ terms cancel. In $C(0,T)$, however, the $K(t)$ term does not cancel.

## 30.2 Dynamics of the bond price

Let $C_t(t,T)$ and $A_t(t,T)$ denote the partial derivatives with respect to $t$. From the formula

$$B(t,T) = \exp\left\{ -r(t)C(t,T) - A(t,T) \right\},$$

we have

$$\begin{aligned} dB(t,T) &= B(t,T)\left[ -C(t,T)\, dr(t) - \tfrac{1}{2}C^2(t,T)\, dr(t)\, dr(t) - r(t)C_t(t,T)\, dt - A_t(t,T)\, dt \right] \\ &= B(t,T)\Big[ -C(t,T)\,(\alpha(t) - \beta(t)r(t))\, dt \\ &\qquad - C(t,T)\sigma(t)\, dW(t) - \tfrac{1}{2}C^2(t,T)\sigma^2(t)\, dt \\ &\qquad - r(t)C_t(t,T)\, dt - A_t(t,T)\, dt \Big]. \end{aligned}$$

Because we have used the risk-neutral pricing formula

$$B(t,T) = I\!E\left[ \exp\left\{ -\int_t^T r(u)\, du \right\} \middle| \mathcal{F}(t) \right]$$

to obtain the bond price, its differential must be of the form

$$dB(t,T) = r(t)B(t,T)\, dt + (\ldots)\, dW(t).$$

Therefore, we must have

$$-C(t,T)\left(\alpha(t)-\beta(t)r(t)\right)-\tfrac{1}{2}C^2(t,T)\sigma^2(t)-r(t)C_t(t,T)-A_t(t,T) \;=\; r(t).$$

We leave the verification of this equation to the homework. After this verification, we have the formula

$$dB(t,T)=r(t)B(t,T)\;dt-\sigma(t)C(t,T)B(t,T)\;dW(t).$$

In particular, the volatility of the bond price is $\sigma(t)C(t,T)$.

## 30.3 Calibration of the Hull & White model

Recall:

$$\begin{aligned}
dr(t) &= (\alpha(t)-\beta(t)r(t))\;dt+\sigma(t)\;dB(t),\\
K(t) &= \int_0^t \beta(u)\;du,\\
A(t,T) &= \int_t^T \left[e^{K(v)}\alpha(v)\left(\int_v^T e^{-K(y)}\;dy\right)-\tfrac{1}{2}e^{2K(v)}\sigma^2(v)\left(\int_v^T e^{-K(y)}\;dy\right)^2\right]\;dv,\\
C(t,T) &= e^{K(t)}\int_t^T e^{-K(y)}\;dy,\\
B(t,T) &= \exp\left\{-r(t)C(t,T)-A(t,T)\right\}.
\end{aligned}$$

Suppose we obtain $B(0,T)$ for all $T\in[0,T^*]$ from market data (with some interpolation). Can we determine the functions $\alpha(t)$, $\beta(t)$, and $\sigma(t)$ for all $t\in[0,T^*]$? Not quite. Here is what we can do.

We take the following input data for the calibration:

1. $B(0,T),\ 0\le T\le T^*$;
2. $r(0)$;
3. $\alpha(0)$;
4. $\sigma(t),\ 0\le t\le T^*$ (usually assumed to be constant);
5. $\sigma(0)C(0,T),\ 0\le T\le T^*$, i.e., the volatility at time zero of bonds of all maturities.

**Step 1.** From 4 and 5 we solve for

$$C(0,T)=\int_0^T e^{-K(y)}\;dy.$$

We can then compute

$$\begin{aligned} \frac{\partial}{\partial T} C(0,T) &= e^{-K(T)} \\ \Longrightarrow\ K(T) &= -\log \frac{\partial}{\partial T} C(0,T), \\ \frac{\partial}{\partial T} K(T) &= \frac{\partial}{\partial T} \int_0^T \beta(u)\, du = \beta(T). \end{aligned}$$

We now have $\beta(T)$ for all $T \in [0, T^*]$.

**Step 2.** From the formula

$$B(0,T) = \exp\{-r(0)C(0,T) - A(0,T)\},$$

we can solve for $A(0,T)$ for all $T \in [0, T^*]$. Recall that

$$A(0,T) = \int_0^T \left[ e^{K(v)} \alpha(v) \left( \int_v^T e^{-K(y)}\, dy \right) - \tfrac{1}{2} e^{2K(v)} \sigma^2(v) \left( \int_v^T e^{-K(y)}\, dy \right)^2 \right] dv.$$

We can use this formula to determine $\alpha(T),\ 0 \le T \le T^*$ as follows:

$$\begin{aligned} \frac{\partial}{\partial T} A(0,T) &= \int_0^T \left[ e^{K(v)} \alpha(v) e^{-K(T)} - e^{2K(v)} \sigma^2(v) e^{-K(T)} \left( \int_v^T e^{-K(y)}\, dy \right) \right] dv, \\ e^{K(T)} \frac{\partial}{\partial T} A(0,T) &= \int_0^T \left[ e^{K(v)} \alpha(v) - e^{2K(v)} \sigma^2(v) \left( \int_v^T e^{-K(y)}\, dy \right) \right] dv, \\ \frac{\partial}{\partial T} \left[ e^{K(T)} \frac{\partial}{\partial T} A(0,T) \right] &= e^{K(T)} \alpha(T) - \int_0^T e^{2K(v)} \sigma^2(v)\ e^{-K(T)}\, dv, \\ e^{K(T)} \frac{\partial}{\partial T} \left[ e^{K(T)} \frac{\partial}{\partial T} A(0,T) \right] &= e^{2K(T)} \alpha(T) - \int_0^T e^{2K(v)} \sigma^2(v)\, dv, \\ \frac{\partial}{\partial T} \left[ e^{K(T)} \frac{\partial}{\partial T} \left[ e^{K(T)} \frac{\partial}{\partial T} A(0,T) \right] \right] &= \alpha'(T) e^{2K(T)} + 2\alpha(T)\beta(T) e^{2K(T)} - e^{2K(T)} \sigma^2(T), \quad 0 \le T \le T^*. \end{aligned}$$

This gives us an ordinary differential equation for $\alpha$, i.e.,

$$\alpha'(t) e^{2K(t)} + 2\alpha(t)\beta(t) e^{2K(t)} - e^{2K(t)} \sigma^2(t) = \text{known function of } t.$$

From assumption 4 and step 1, we know all the coefficients in this equation. From assumption 3, we have the initial condition $\alpha(0)$. We can solve the equation numerically to determine the function $\alpha(t),\ 0 \le t \le T^*$.

**Remark 30.1** The derivation of the ordinary differential equation for $\alpha(t)$ requires three differentiations. Differentiation is an unstable procedure, i.e., functions which are close can have very different derivatives. Consider, for example,

$$\begin{aligned} f(x) &= 0 \quad \forall x \in I\!\!R, \\ g(x) &= \frac{\sin(1000x)}{100} \quad \forall x \in I\!\!R. \end{aligned}$$

Then

$$|f(x) - g(x)| \leq \frac{1}{100} \quad \forall x \in I\!R,$$

but because

$$g'(x) = 10\cos(1000x),$$

we have

$$|f'(x) - g'(x)| = 10$$

for many values of $x$.

Assumption 5 for the calibration was that we know the volatility at time zero of bonds of all maturities. These volatilities can be implied by the prices of options on bonds. We consider now how the model prices options.

## 30.4 Option on a bond

Consider a European call option on a zero-coupon bond with strike price $K$ and expiration time $T_1$. The bond matures at time $T_2 > T_1$. The price of the option at time 0 is

$$\begin{aligned}
&I\!E\left[e^{-\int_0^{T_1} r(u)\,du}\,(B(T_1,T_2)-K)^+\right] \\
&= I\!E e^{-\int_0^{T_1} r(u)\,du}(\exp\{-r(T_1)C(T_1,T_2)-A(T_1,T_2)\}-K)^+. \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x}\Big(\exp\{-yC(T_1,T_2)-A(T_1,T_2)\}-K\Big)^+ f(x,y)\,dx\,dy,
\end{aligned}$$

where $f(x,y)$ is the joint density of $\left(\int_0^{T_1} r(u)\,du,\ r(T_1)\right)$.

We observed at the beginning of this Chapter (equation (0.3)) that $\int_0^{T_1} r(u)\,du$ is normal with

$$\begin{aligned}
\mu_1 \stackrel{\triangle}{=} I\!E\left[\int_0^{T_1} r(u)\,du\right] &= \int_0^{T_1} I\!E r(u)\,du \\
&= \int_0^{T_1}\left[r(0)e^{-K(v)} + e^{-K(v)}\int_0^v e^{K(u)}\alpha(u)\,du\right]dv, \\
\sigma_1^2 \stackrel{\triangle}{=} \mathrm{var}\left[\int_0^{T_1} r(u)\,du\right] &= \int_0^{T_1} e^{2K(v)}\sigma^2(v)\left(\int_v^{T_1} e^{-K(y)}\,dy\right)^2 dv.
\end{aligned}$$

We also observed (equation (0.1)) that $r(T_1)$ is normal with

$$\begin{aligned}
\mu_2 &\stackrel{\triangle}{=} I\!E r(T_1) = r(0)e^{-K(T_1)} + e^{-K(T_1)}\int_0^{T_1} e^{K(u)}\alpha(u)\,du, \\
\sigma_2^2 &\stackrel{\triangle}{=} \mathrm{var}\,(r(T_1)) = e^{-2K(T_1)}\int_0^{T_1} e^{2K(u)}\sigma^2(u)\,du.
\end{aligned}$$

In fact, $\left(\int_0^{T_1} r(u)\,du,\; r(T_1)\right)$ is jointly normal, and the covariance is

$$\begin{aligned}
\rho\sigma_1\sigma_2 &= I\!E\left[\int_0^{T_1} (r(u) - I\!Er(u))\; du.\, (r(T_1) - I\!Er(T_1))\right] \\
&= \int_0^{T_1} I\!E[(r(u) - I\!Er(u))\; (r(T_1) - I\!Er(T_1))]\, du \\
&= \int_0^{T_1} \rho_r(u, T_1)\; du,
\end{aligned}$$

where $\rho_r(u, T_1)$ is defined in Equation 0.2.
The option on the bond has price at time zero of

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x}\Big(\exp\{-yC(T_1, T_2) - A(T_1, T_2)\} - K\Big)^{+} \\
&\qquad \cdot \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{x^2}{\sigma_1^2} + \frac{2\rho x y}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right]\right\}\; dx\; dy. \qquad (4.1)
\end{aligned}$$

The price of the option at time $t \in [0, T_1]$ is

$$\begin{aligned}
&I\!E\left[e^{-\int_t^{T_1} r(u)\,du}\,(B(T_1, T_2) - K)^{+}\Big|\mathcal{F}(t)\right] \\
&\qquad = I\!E\left[e^{-\int_t^{T_1} r(u)\,du}(\exp\{-r(T_1)C(T_1, T_2) - A(T_1, T_2)\} - K)^{+}\Big|\mathcal{F}(t)\right] \qquad (4.2)
\end{aligned}$$

Because of the Markov property, this is random only through a dependence on $r(t)$. To compute this option price, we need the joint distribution of $\left(\int_t^{T_1} r(u)\,du,\; r(T_1)\right)$ conditioned on $r(t)$. This pair of random variables has a jointly normal conditional distribution, and

$$\begin{aligned}
\mu_1(t) &= \mathbb{E}\left[\int_t^{T_1} r(u)\,du \middle| \mathcal{F}(t)\right] \\
&= \int_t^{T_1}\left[r(t)e^{-K(v)+K(t)} + e^{-K(v)}\int_t^v e^{K(u)}\alpha(u)\,du\right]dv, \\
\sigma_1^2(t) &= \mathbb{E}\left[\left(\int_t^{T_1} r(u)\,du - \mu_1(t)\right)^2 \middle| \mathcal{F}(t)\right] \\
&= \int_t^{T_1} e^{2K(v)}\sigma^2(v)\left(\int_v^{T_1} e^{-K(y)}\,dy\right)^2 dv, \\
\mu_2(t) &= \mathbb{E}\left[r(T_1)\middle| r(t)\right] \\
&= r(t)e^{-K(T_1)+K(t)} + e^{-K(T_1)}\int_t^{T_1} e^{K(u)}\alpha(u)\,du, \\
\sigma_2^2(t) &= \mathbb{E}\left[(r(T_1)-\mu_2(t))^2\middle|\mathcal{F}(t)\right] \\
&= e^{-2K(T_1)}\int_t^{T_1} e^{2K(u)}\sigma^2(u)\,du, \\
\rho(t)\sigma_1(t)\sigma_2(t) &= \mathbb{E}\left[\left(\int_t^{T_1} r(u)\,du - \mu_1(t)\right)(r(T_1)-\mu_2(t))\middle|\mathcal{F}(t)\right] \\
&= \int_t^{T_1} e^{-K(u)-K(T_1)}\int_t^u e^{2K(v)}\sigma^2(v)\,dv\,du.
\end{aligned}$$

The variances and covariances are not random. The means are random through a dependence on $r(t)$.

Advantages of the Hull & White model:

1. Leads to closed-form pricing formulas.
2. Allows calibration to fit initial yield curve exactly.

Short-comings of the Hull & White model:

1. One-factor, so only allows parallel shifts of the yield curve, i.e.,

   $$B(t,T) = \exp\{-r(t)C(t,T) - A(t,T)\},$$

   so bond prices of all maturities are perfectly correlated.

2. Interest rate is normally distributed, and hence can take negative values. Consequently, the bond price

   $$B(t,T) = \mathbb{E}\left[\exp\left\{-\int_t^T r(u)\,du\right\}\middle|\mathcal{F}(t)\right]$$

   can exceed 1.

# Chapter 31

# Cox-Ingersoll-Ross model

In the Hull & White model, $r(t)$ is a Gaussian process. Since, for each $t$, $r(t)$ is normally distributed, there is a positive probability that $r(t) < 0$. The Cox-Ingersoll-Ross model is the simplest one which avoids negative interest rates.

We begin with a $d$-dimensional Brownian motion $(W_1, W_2, \ldots, W_d)$. Let $\beta > 0$ and $\sigma > 0$ be constants. For $j = 1, \ldots, d$, let $X_j(0) \in I\!\!R$ be given so that

$$X_1^2(0) + X_2^2(0) + \ldots + X_d^2(0) \geq 0,$$

and let $X_j$ be the solution to the stochastic differential equation

$$dX_j(t) = -\tfrac{1}{2}\beta X_j(t)\ dt + \tfrac{1}{2}\sigma\ dW_j(t).$$

$X_j$ is called the *Orstein-Uhlenbeck* process. It always has a drift toward the origin. The solution to this stochastic differential equation is

$$X_j(t) = e^{-\frac{1}{2}\beta t}\left[X_j(0) + \tfrac{1}{2}\sigma \int_0^t e^{\frac{1}{2}\beta u}\ dW_j(u)\right].$$

This solution is a Gaussian process with mean function

$$m_j(t) = e^{-\frac{1}{2}\beta t} X_j(0)$$

and covariance function

$$\rho(s,t) = \frac{1}{4}\sigma^2 e^{-\frac{1}{2}\beta(s+t)} \int_0^{s\wedge t} e^{\beta u}\ du.$$

Define

$$r(t) \stackrel{\triangle}{=} X_1^2(t) + X_2^2(t) + \ldots + X_d^2(t).$$

If $d = 1$, we have $r(t) = X_1^2(t)$ and for each $t$, $I\!\!P\{r(t) > 0\} = 1$, but (see Fig. 31.1)

$$I\!\!P\left\{\text{There are infinitely many values of } t > 0 \text{ for which } r(t) = 0\right\} = 1$$

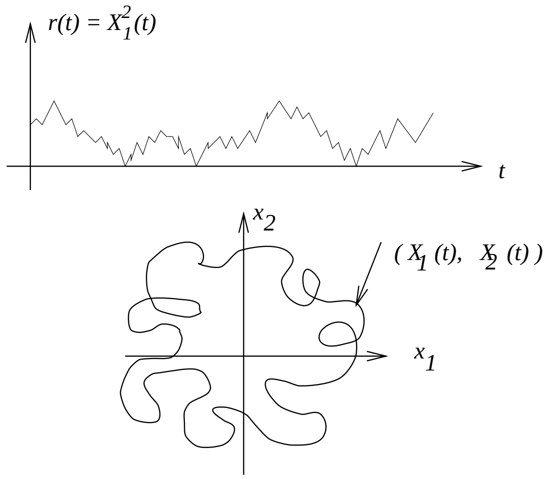



Figure 31.1: $r(t)$ *can be zero.*

If $d \geq 2$, (see Fig. 31.1)

$$\mathbb{P}\{\text{There is at least one value of } t > 0 \text{ for which } r(t) = 0\} \;=\; 0.$$

Let $f(x_1, x_2, \ldots, x_d) = x_1^2 + x_2^2 + \ldots + x_d^2$. Then

$$f_{x_i} = 2x_i, \quad f_{x_i x_j} = \begin{cases} 2 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Itô's formula implies

$$\begin{aligned} dr(t) &= \sum_{i=1}^{d} f_{x_i}\, dX_i + \tfrac{1}{2} \sum_{i=1}^{d} f_{x_i x_i}\, dX_i\, dX_i \\ &= \sum_{i=1}^{d} 2X_i \left(-\tfrac{1}{2}\beta X_i\, dt + \tfrac{1}{2}\sigma\, dW_i(t)\right) + \sum_{i=1}^{d} \frac{1}{4}\sigma^2\, dW_i\, dW_i \\ &= -\beta r(t)\, dt + \sigma \sum_{i=1}^{d} X_i\, dW_i + \frac{d\sigma^2}{4}\, dt \\ &= \left(\frac{d\sigma^2}{4} - \beta r(t)\right) dt + \sigma \sqrt{r(t)} \sum_{i=1}^{d} \frac{X_i(t)}{\sqrt{r(t)}}\, dW_i(t). \end{aligned}$$

Define

$$W(t) = \sum_{i=1}^{d} \int_0^t \frac{X_i(u)}{\sqrt{r(u)}}\, dW_i(u).$$

Then $W$ is a martingale,

$$dW = \sum_{i=1}^{d} \frac{X_i}{\sqrt{r}}\, dW_i,$$
$$dW\; dW = \sum_{i=1}^{d} \frac{X_i^2}{r}\, dt = dt,$$

so $W$ is a Brownian motion. We have

$$dr(t) = \left(\frac{d\sigma^2}{4} - \beta r(t)\right)\, dt + \sigma\sqrt{r(t)}\; dW(t).$$

The *Cox-Ingersoll-Ross (CIR) process* is given by

$$dr(t) = (\alpha - \beta r(t))\; dt + \sigma\sqrt{r(t)}\; dW(t),$$

We define

$$d = \frac{4\alpha}{\sigma^2} > 0.$$

If $d$ happens to be an integer, then we have the representation

$$r(t) = \sum_{i=1}^{d} X_i^2(t),$$

but we do not require $d$ to be an integer. If $d < 2$ (i.e., $\alpha < \frac{1}{2}\sigma^2$), then

$$I\!P\{\text{There are infinitely many values of } t > 0 \text{ for which } r(t) = 0\} \;=\; 1.$$

This is not a good parameter choice.

If $d \geq 2$ (i.e., $\alpha \geq \frac{1}{2}\sigma^2$), then

$$I\!P\{\text{There is at least one value of } t > 0 \text{ for which } r(t) = 0\} \;=\; 0.$$

With the CIR process, one can derive formulas under the assumption that $d = \frac{4\alpha}{\sigma^2}$ is a positive integer, and they are still correct even when $d$ is not an integer.

For example, here is the distribution of $r(t)$ for fixed $t > 0$. Let $r(0) \geq 0$ be given. Take

$$X_1(0) = 0,\; X_2(0) = 0,\; \ldots,\; X_{d-1}(0) = 0,\; X_d(0) = \sqrt{r(0)}.$$

For $i = 1, 2, \ldots, d-1$, $X_i(t)$ is normal with mean zero and variance

$$\rho(t,t) = \frac{\sigma^2}{4\beta}(1 - e^{-\beta t}).$$

$X_d(t)$ is normal with mean

$$m_d(t) = e^{-\frac{1}{2}\beta t}\sqrt{r(0)}$$

and variance $\rho(t,t)$. Then

$$r(t) = \underbrace{\rho(t,t)\sum_{i=1}^{d-1}\left(\frac{X_i(t)}{\sqrt{\rho(t,t)}}\right)^2}_{\text{Chi-square with } d-1=\frac{4\alpha-\sigma^2}{\sigma^2} \text{ degrees of freedom}} + \underbrace{X_d^2(t)}_{\text{Normal squared and independent of the other term}} \tag{0.1}$$

Thus $r(t)$ has a *non-central chi-square distribution.*

## 31.1 Equilibrium distribution of $r(t)$

As $t \to \infty$, $m_d(t) \to 0$. We have

$$r(t) = \rho(t,t)\sum_{i=1}^{d}\left(\frac{X_i(t)}{\sqrt{\rho(t,t)}}\right)^2.$$

As $t \to \infty$, we have $\rho(t,t) = \frac{\sigma^2}{4\beta}$, and so the limiting distribution of $r(t)$ is $\frac{\sigma^2}{4\beta}$ times a chi-square with $d = \frac{4\alpha}{\sigma^2}$ degrees of freedom. The chi-square density with $\frac{4\alpha}{\sigma^2}$ degrees of freedom is

$$f(y) = \frac{1}{2^{2\alpha/\sigma^2}\Gamma\left(\frac{2\alpha}{\sigma^2}\right)} y^{\frac{2\alpha-\sigma^2}{\sigma^2}} e^{-y/2}.$$

We make the change of variable $r = \frac{\sigma^2}{4\beta}y$. The limiting density for $r(t)$ is

$$\begin{aligned} p(r) &= \frac{4\beta}{\sigma^2}.\frac{1}{2^{2\alpha/\sigma^2}\Gamma\left(\frac{2\alpha}{\sigma^2}\right)}\left(\frac{4\beta}{\sigma^2}r\right)^{\frac{2\alpha-\sigma^2}{\sigma^2}} e^{-\frac{2\beta}{\sigma^2}r} \\ &= \left(\frac{2\beta}{\sigma^2}\right)^{\frac{2\alpha}{\sigma^2}} \frac{1}{\Gamma\left(\frac{2\alpha}{\sigma^2}\right)} r^{\frac{2\alpha-\sigma^2}{\sigma^2}} e^{-\frac{2\beta}{\sigma^2}r}. \end{aligned}$$

We computed the mean and variance of $r(t)$ in Section 15.7.

## 31.2 Kolmogorov forward equation

Consider a Markov process governed by the stochastic differential equation

$$dX(t) = b(X(t))\,dt + \sigma(X(t))\,dW(t).$$

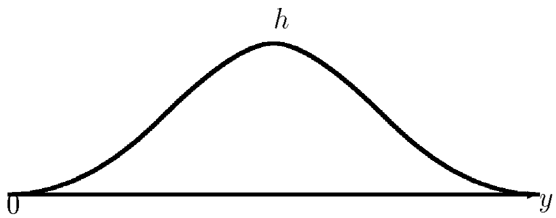



Figure 31.2: *The function* $h(y)$

Because we are going to apply the following analysis to the case $X(t) = r(t)$, we assume that $X(t) \geq 0$ for all $t$.

We start at $X(0) = x \geq 0$ at time 0. Then $X(t)$ is random with density $p(0,t,x,y)$ (in the $y$ variable). Since 0 and $x$ will not change during the following, we omit them and write $p(t,y)$ rather than $p(0,t,x,y)$. We have

$$I\!\!E h(X(t)) = \int_0^\infty h(y)p(t,y)\,dy$$

for any function $h$.

The Kolmogorov forward equation (KFE) is a partial differential equation in the "forward" variables $t$ and $y$. We derive it below.

Let $h(y)$ be a smooth function of $y \geq 0$ which vanishes near $y = 0$ and for all large values of $y$ (see Fig. 31.2). Itô's formula implies

$$dh(X(t)) = \left[h'(X(t))b(X(t)) + \tfrac{1}{2}h''(X(t))\sigma^2(X(t))\right]\,dt + h'(X(t))\sigma(X(t))\,dW(t),$$

so

$$\begin{aligned}h(X(t)) &= h(X(0)) + \int_0^t \left[h'(X(s))b(X(s)) + \tfrac{1}{2}h''(X(s))\sigma^2(X(s))\right]\,ds + \\ &\quad \int_0^t h'(X(s))\sigma(X(s))\,dW(s), \\ I\!\!E h(X(t)) &= h(X(0)) + I\!\!E\int_0^t \left[h'(X(s))b(X(s))\,dt + \tfrac{1}{2}h''(X(s))\sigma^2(X(s))\right]\,ds,\end{aligned}$$

or equivalently,

$$\int_0^\infty h(y)p(t,y)\,dy = h(X(0)) + \int_0^t\int_0^\infty h'(y)b(y)p(s,y)\,dy\,ds + \tfrac{1}{2}\int_0^t\int_0^\infty h''(y)\sigma^2(y)p(s,y)\,dy\,ds.$$

Differentiate with respect to $t$ to get

$$\int_0^\infty h(y)p_t(t,y)\,dy = \int_0^\infty h'(y)b(y)p(t,y)\,dy + \tfrac{1}{2}\int_0^\infty h''(y)\sigma^2(y)p(t,y)\,dy.$$

Integration by parts yields

$$\int_0^\infty h'(y)b(y)p(t,y)\,dy = \underbrace{h(y)b(y)p(t,y)\Big|_{y=0}^{y=\infty}}_{=0} - \int_0^\infty h(y)\frac{\partial}{\partial y}\left(b(y)p(t,y)\right)\,dy,$$

$$\begin{aligned}\int_0^\infty h''(y)\sigma^2(y)p(t,y)\,dy &= \underbrace{h'(y)\sigma^2(y)p(t,y)\Big|_{y=0}^{y=\infty}}_{=0} - \int_0^\infty h'(y)\frac{\partial}{\partial y}\left(\sigma^2(y)p(t,y)\right)\,dy \\ &= \underbrace{-h(y)\frac{\partial}{\partial y}\left(\sigma^2(y)p(t,y)\right)\Big|_{y=0}^{y=\infty}}_{=0} + \int_0^\infty h(y)\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(t,y)\right)\,dy.\end{aligned}$$

Therefore,

$$\int_0^\infty h(y)p_t(t,y)\,dy = -\int_0^\infty h(y)\frac{\partial}{\partial y}\left(b(y)p(t,y)\right)\,dy + \tfrac{1}{2}\int_0^\infty h(y)\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(t,y)\right)\,dy,$$

or equivalently,

$$\int_0^\infty h(y)\left[p_t(t,y) + \frac{\partial}{\partial y}\left(b(y)p(t,y)\right) - \tfrac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(t,y)\right)\right]\,dy = 0.$$

This last equation holds for every function $h$ of the form in Figure 31.2. It implies that

$$p_t(t,y) + \frac{\partial}{\partial y}\left((b(y)p(t,y)\right) - \tfrac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(t,y)\right) = 0. \tag{KFE}$$

If there were a place where (KFE) did not hold, then we could take $h(y) > 0$ at that and nearby points, but take $h$ to be zero elsewhere, and we would obtain

$$\int_0^\infty h\left[p_t + \frac{\partial}{\partial y}(bp) - \tfrac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2 p)\right]\,dy \neq 0.$$

If the process $X(t)$ has an equilibrium density, it will be

$$p(y) = \lim_{t\to\infty} p(t,y).$$

In order for this limit to exist, we must have

$$0 = \lim_{t\to\infty} p_t(t,y).$$

Letting $t\to\infty$ in (KFE), we obtain the equilibrium Kolmogorov forward equation

$$\frac{\partial}{\partial y}\left(b(y)p(y)\right) - \tfrac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(y)\right) = 0.$$

When an equilibrium density exists, it is the unique solution to this equation satisfying

$$p(y) \geq 0 \quad \forall y \geq 0,$$
$$\int_0^\infty p(y)\ dy = 1.$$

## 31.3 Cox-Ingersoll-Ross equilibrium density

We computed this to be

$$p(r) = Cr^{\frac{2\alpha-\sigma^2}{\sigma^2}} e^{-\frac{2\beta}{\sigma^2}r},$$

where

$$C = \left(\frac{2\beta}{\sigma^2}\right)^{\frac{2\alpha}{\sigma^2}} \frac{1}{\Gamma\left(\frac{2\alpha}{\sigma^2}\right)}.$$

We compute

$$\begin{aligned}
p'(r) &= \frac{2\alpha-\sigma^2}{\sigma^2}.\frac{p(r)}{r} - \frac{2\beta}{\sigma^2}p(r) \\
&= \frac{2}{\sigma^2 r}\left(\alpha - \tfrac{1}{2}\sigma^2 - \beta r\right)p(r), \\
p''(r) &= -\frac{2}{\sigma^2 r^2}\left(\alpha - \tfrac{1}{2}\sigma^2 - \beta r\right)p(r) + \frac{2}{\sigma^2 r}(-\beta)p(r) + \frac{2}{\sigma^2 r}\left(\alpha - \tfrac{1}{2}\sigma^2 - \beta r\right)p'(r) \\
&= \frac{2}{\sigma^2 r}\left(-\frac{1}{r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r) - \beta + \frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)^2\right)\ p(r)
\end{aligned}$$

We want to verify the equilibrium Kolmogorov forward equation for the CIR process:

$$\frac{\partial}{\partial r}\left((\alpha-\beta r)p(r)\right) - \tfrac{1}{2}\frac{\partial^2}{\partial r^2}(\sigma^2 r p(r)) = 0. \tag{EKFE}$$

Now

$$\begin{aligned}
\frac{\partial}{\partial r}\left((\alpha-\beta r)p(r)\right) &= -\beta p(r) + (\alpha-\beta r)p'(r),\\
\frac{\partial^2}{\partial r^2}(\sigma^2 r p(r)) &= \frac{\partial}{\partial r}(\sigma^2 p(r) + \sigma^2 r p'(r))\\
&= 2\sigma^2 p'(r) + \sigma^2 r p''(r).
\end{aligned}$$

The LHS of (EKFE) becomes

$$\begin{aligned}
&-\beta p(r) + (\alpha-\beta r)p'(r) - \sigma^2 p'(r) - \tfrac{1}{2}\sigma^2 r p''(r)\\
&= p(r)\Bigg[-\beta + (\alpha - \beta r - \sigma^2)\frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)\\
&\qquad + \frac{1}{r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r) + \beta - \frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)^2\Bigg]\\
&= p(r)\Bigg[(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)\frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)\\
&\qquad - \tfrac{1}{2}\sigma^2\frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)\\
&\qquad + \frac{1}{r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r) - \frac{2}{\sigma^2 r}(\alpha - \tfrac{1}{2}\sigma^2 - \beta r)^2\Bigg]\\
&= 0,
\end{aligned}$$

as expected.

## 31.4 Bond prices in the CIR model

The interest rate process $r(t)$ is given by

$$dr(t) = (\alpha - \beta r(t))\ dt + \sigma\sqrt{r(t)}\ dW(t),$$

where $r(0)$ is given. The bond price process is

$$B(t,T) = \mathbb{E}\left[\exp\left\{-\int_t^T r(u)\ du\right\}\Big|\mathcal{F}(t)\right].$$

Because

$$\exp\left\{-\int_0^t r(u)\ du\right\} B(t,T) = \mathbb{E}\left[\exp\left\{-\int_0^T r(u)\ du\right\}\Big|\mathcal{F}(t)\right],$$

the tower property implies that this is a martingale. The Markov property implies that $B(t,T)$ is random only through a dependence on $r(t)$. Thus, there is a function $B(r,t,T)$ of the three dummy variables $r, t, T$ such that the *process* $B(t,T)$ is the *function* $B(r,t,T)$ evaluated at $r(t), t, T$, i.e.,

$$B(t,T) = B(r(t), t, T).$$

Because $\exp\left\{-\int_0^t r(u)\,du\right\} B(r(t),t,T)$ is a martingale, its differential has no $dt$ term. We compute

$$
\begin{aligned}
&d\left(\exp\left\{-\int_0^t r(u)\,du\right\} B(r(t),t,T)\right)\\
&= \exp\left\{-\int_0^t r(u)\,du\right\}\Big[-r(t)B(r(t),t,T)\,dt + B_r(r(t),t,T)\,dr(t) +\\
&\qquad \tfrac{1}{2}B_{rr}(r(t),t,T)\,dr(t)\,dr(t) + B_t(r(t),t,T)\,dt\Big].
\end{aligned}
$$

The expression in $[\ldots]$ equals

$$
\begin{aligned}
&= -rB\,dt + B_r(\alpha-\beta r)\,dt + B_r\sigma\sqrt{r}\,dW\\
&\qquad + \tfrac{1}{2}B_{rr}\sigma^2 r\,dt + B_t\,dt.
\end{aligned}
$$

Setting the $dt$ term to zero, we obtain the partial differential equation

$$
\begin{aligned}
-\,rB(r,t,T) + B_t(r,t,T) + (\alpha-\beta r)B_r(r,t,T) + \tfrac{1}{2}\sigma^2 r B_{rr}(r,t,T) = 0,&\\
0 \le t < T,\quad r \ge 0.&\qquad (4.1)
\end{aligned}
$$

The terminal condition is

$$B(r,T,T) = 1,\quad r \ge 0.$$

Surprisingly, this equation has a closed form solution. Using the Hull & White model as a guide, we look for a solution of the form

$$B(r,t,T) = e^{-rC(t,T)-A(t,T)},$$

where $C(T,T)=0,\ A(T,T)=0$. Then we have

$$
\begin{aligned}
B_t &= (-rC_t - A_t)B,\\
B_r &= -CB,\quad B_{rr} = C^2 B,
\end{aligned}
$$

and the partial differential equation becomes

$$
\begin{aligned}
0 &= -rB + (-rC_t - A_t)B - (\alpha-\beta r)CB + \tfrac{1}{2}\sigma^2 rC^2 B\\
&= rB(-1 - C_t + \beta C + \tfrac{1}{2}\sigma^2 C^2) - B(A_t + \alpha C)
\end{aligned}
$$

We first solve the ordinary differential equation

$$-1 - C_t(t,T) + \beta C(t,T) + \tfrac{1}{2}\sigma^2 C^2(t,T) = 0;\quad C(T,T) = 0,$$

and then set

$$A(t,T) = \alpha\int_t^T C(u,T)\,du,$$

so $A(T,T)=0$ and

$$A_t(t,T) = -\alpha C(t,T).$$

It is tedious but straightforward to check that the solutions are given by

$$C(t,T) = \frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t)) + \frac{1}{2}\beta\sinh(\gamma(T-t))},$$

$$A(t,T) = -\frac{2\alpha}{\sigma^2}\log\left[\frac{\gamma e^{\frac{1}{2}\beta(T-t)}}{\gamma\cosh(\gamma(T-t)) + \frac{1}{2}\beta\sinh(\gamma(T-t))}\right],$$

where

$$\gamma = \tfrac{1}{2}\sqrt{\beta^2+2\sigma^2}, \quad \sinh u = \frac{e^u - e^{-u}}{2}, \quad \cosh u = \frac{e^u + e^{-u}}{2}.$$

Thus in the CIR model, we have

$$I\!\!E\left[\exp\left\{-\int_t^T r(u)\,du\right\}\middle|\mathcal{F}(t)\right] = B(r(t),t,T),$$

where

$$B(r,t,T) = \exp\left\{-rC(t,T) - A(t,T)\right\}, \quad 0 \le t < T, \ \ r \ge 0,$$

and $C(t,T)$ and $A(t,T)$ are given by the formulas above. Because the coefficients in

$$dr(t) = (\alpha - \beta r(t))\,dt + \sigma\sqrt{r(t)}\,dW(t)$$

do not depend on $t$, the function $B(r,t,T)$ depends on $t$ and $T$ only through their difference $\tau = T-t$. Similarly, $C(t,T)$ and $A(t,T)$ are functions of $\tau = T-t$. We write $B(r,\tau)$ instead of $B(r,t,T)$, and we have

$$B(r,\tau) = \exp\left\{-rC(\tau) - A(\tau)\right\}, \quad \tau \ge 0, \ \ r \ge 0,$$

where

$$C(\tau) = \frac{\sinh(\gamma\tau)}{\gamma\cosh(\gamma\tau) + \frac{1}{2}\beta\sinh(\gamma\tau)},$$

$$A(\tau) = -\frac{2\alpha}{\sigma^2}\log\left[\frac{\gamma e^{\frac{1}{2}\beta\tau}}{\gamma\cosh(\gamma\tau) + \frac{1}{2}\beta\sinh(\gamma\tau)}\right],$$

$$\gamma = \tfrac{1}{2}\sqrt{\beta^2+2\sigma^2}.$$

We have

$$B(r(0),T) = I\!\!E\exp\left\{-\int_0^T r(u)\,du\right\}.$$

Now $r(u) > 0$ for each $u$, almost surely, so $B(r(0),T)$ is strictly decreasing in $T$. Moreover,

$$B(r(0),0) = 1,$$

$$\lim_{T\to\infty} B(r(0),T) = I\!\!E \exp\left\{-\int_0^\infty r(u)\,du\right\} = 0.$$

But also,

$$B(r(0),T) = \exp\left\{-r(0)C(T) - A(T)\right\},$$

so

$$r(0)C(0) + A(0) = 0,$$
$$\lim_{T\to\infty}\left[r(0)C(T) + A(T)\right] = \infty,$$

and

$$r(0)C(T) + A(T)$$

is strictly inreasing in $T$.

## 31.5 Option on a bond

The value at time $t$ of an option on a bond in the CIR model is

$$v(t, r(t)) = I\!\!E\left[\exp\left\{-\int_t^{T_1} r(u)\,du\right\}(B(T_1,T_2) - K)^+ \bigg| \mathcal{F}(t)\right],$$

where $T_1$ is the expiration time of the option, $T_2$ is the maturity time of the bond, and $0 \le t \le T_1 \le T_2$. As usual, $\exp\left\{-\int_0^t r(u)\,du\right\} v(t, r(t))$ is a martingale, and this leads to the partial differential equation

$$-rv + v_t + (\alpha - \beta r)v_r + \tfrac{1}{2}\sigma^2 r v_{rr} = 0, \quad 0 \le t < T_1, \;\; r \ge 0.$$

(where $v = v(t,r)$.) The terminal condition is

$$v(T_1, r) = \left(B(r, T_1, T_2) - K\right)^+, \quad r \ge 0.$$

Other European derivative securities on the bond are priced using the same partial differential equation with the terminal condition appropriate for the particular security.

## 31.6 Deterministic time change of CIR model

*Process time scale:* In this time scale, the interest rate $r(t)$ is given by the constant coefficient CIR equation

$$dr(t) = (\alpha - \beta r(t))\,dt + \sigma\sqrt{r(t)}\,dW(t).$$

*Real time scale:* In this time scale, the interest rate $\hat{r}(\hat{t})$ is given by a time-dependent CIR equation

$$d\hat{r}(\hat{t}) = (\hat{\alpha}(\hat{t}) - \hat{\beta}(\hat{t})\hat{r}(\hat{t}))\,d\hat{t} + \hat{\sigma}(\hat{t})\sqrt{\hat{r}(\hat{t})}\,d\hat{W}(\hat{t}).$$

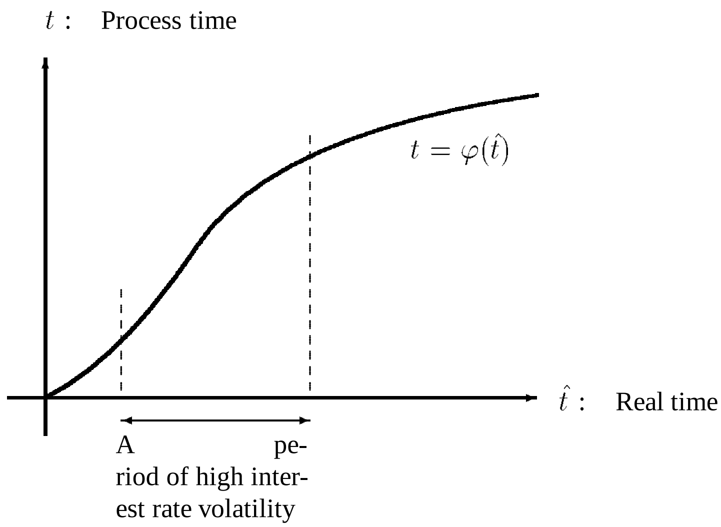



Figure 31.3: *Time change function.*

There is a strictly increasing time change function $t = \varphi(\hat{t})$ which relates the two time scales (See Fig. 31.3).

Let $\hat{B}(\hat{r}, \hat{t}, \hat{T})$ denote the price at real time $\hat{t}$ of a bond with maturity $\hat{T}$ when the interest rate at time $\hat{t}$ is $\hat{r}$. We want to set things up so

$$\hat{B}(\hat{r}, \hat{t}, \hat{T}) = B(r, t, T) = e^{-rC(t,T)-A(t,T)},$$

where $t = \varphi(\hat{t})$, $T = \varphi(\hat{T})$, and $C(t,T)$ and $A(t,T)$ are as defined previously.

We need to determine the relationship between $\hat{r}$ and $r$. We have

$$B(r(0), 0, T) = I\!\!E \exp\left\{-\int_0^T r(t)\; dt\right\},$$
$$B(\hat{r}(0), 0, \hat{T}) = I\!\!E \exp\left\{-\int_0^{\hat{T}} \hat{r}(\hat{t})\; d\hat{t}\right\}.$$

With $T = \varphi(\hat{T})$, make the change of variable $t = \varphi(\hat{t})$, $dt = \varphi'(\hat{t})\; d\hat{t}$ in the first integral to get

$$B(r(0), 0, T) = I\!\!E \exp\left\{-\int_0^{\hat{T}} r(\varphi(\hat{t}))\varphi'(\hat{t})\; d\hat{t}\right\},$$

and this will be $B(\hat{r}(0), 0, \hat{T})$ if we set

$$\boxed{\hat{r}(\hat{t}) = r(\varphi(\hat{t}))\; \varphi'(\hat{t}).}$$

## 31.7 Calibration

$$\begin{aligned}
\hat{B}(\hat{r}(\hat{t}), \hat{t}, \hat{T}) &= B\left(\frac{\hat{r}(\hat{t})}{\varphi'(\hat{t})}, \varphi(\hat{t}), \varphi(\hat{T})\right) \\
&= \exp\left\{-\hat{r}(\hat{t})\frac{C(\varphi(\hat{t}), \varphi(\hat{T}))}{\varphi'(\hat{t})} - A(\varphi(\hat{t}), \varphi(\hat{T}))\right\} \\
&= \exp\left\{-\hat{r}(\hat{t})\hat{C}(\hat{t}, \hat{T}) - \hat{A}(\hat{t}, \hat{T})\right\},
\end{aligned}$$

where

$$\begin{aligned}
\hat{C}(\hat{t}, \hat{T}) &= \frac{C(\varphi(\hat{t}), \varphi(\hat{T}))}{\varphi'(\hat{t})} \\
\hat{A}(\hat{t}, \hat{T}) &= A(\varphi(\hat{t}), \varphi(\hat{T}))
\end{aligned}$$

do *not* depend on $\hat{t}$ and $\hat{T}$ only through $\hat{T} - \hat{t}$, since, in the real time scale, the model coefficients are time dependent.

Suppose we know $\hat{r}(0)$ and $\hat{B}(\hat{r}(0), 0, \hat{T})$ for all $\hat{T} \in [0, \hat{T}^*]$. We calibrate by writing the equation

$$\hat{B}(\hat{r}(0), 0, \hat{T}) = \exp\left\{-\hat{r}(0)\hat{C}(0, \hat{T}) - \hat{A}(0, \hat{T})\right\},$$

or equivalently,

$$-\log \hat{B}(\hat{r}(0), 0, \hat{T}) = \frac{\hat{r}(0)}{\varphi'(0)} C(\varphi(0), \varphi(\hat{T})) + A(\varphi(0), \varphi(\hat{T})).$$

Take $\alpha, \beta$ and $\sigma$ so the equilibrium distribution of $r(t)$ seems reasonable. These values determine the functions $C, A$. Take $\varphi'(0) = 1$ (we justify this in the next section). For each $\hat{T}$, solve the equation for $\varphi(\hat{T})$:

$$-\log \hat{B}(\hat{r}(0), 0, \hat{T}) = \hat{r}(0)C(0, \varphi(\hat{T})) + A(0, \varphi(\hat{T})). \tag{*}$$

The right-hand side of this equation is increasing in the $\varphi(\hat{T})$ variable, starting at 0 at time $0$ and having limit $\infty$ at $\infty$, i.e.,

$$\begin{aligned}
\hat{r}(0)C(0, 0) + A(0, 0) = 0, \\
\lim_{T \to \infty} [\hat{r}(0)C(0, T) + A(0, T)] = \infty.
\end{aligned}$$

Since $0 \le -\log \hat{B}(\hat{r}(0), 0, \hat{T}) < \infty$, (*) has a unique solution for each $\hat{T}$. For $\hat{T} = 0$, this solution is $\varphi(0) = 0$. If $\hat{T}_1 < \hat{T}_2$, then

$$-\log \hat{B}(r(0), 0, \hat{T}_1) < -\log \hat{B}(r(0), 0, \hat{T}_2),$$

so $\varphi(\hat{T}_1) < \varphi(\hat{T}_2)$. Thus $\varphi$ is a strictly increasing time-change-function with the right properties.

## 31.8 Tracking down $\varphi'(0)$ in the time change of the CIR model

Result for general term structure models:

$$-\frac{\partial}{\partial T}\log B(0,T)\Big|_{T=0} = r(0).$$

Justification:

$$\begin{aligned}
B(0,T) &= I\!\!E \exp\left\{-\int_0^T r(u)\,du\right\}.\\
-\log B(0,T) &= -\log I\!\!E \exp\left\{-\int_0^T r(u)\,du\right\}\\
-\frac{\partial}{\partial T}\log B(0,T) &= \frac{I\!\!E\left[r(T)e^{-\int_0^T r(u)\,du}\right]}{I\!\!E e^{-\int_0^T r(u)\,du}}\\
-\frac{\partial}{\partial T}\log B(0,T)\Big|_{T=0} &= r(0).
\end{aligned}$$

In the real time scale associated with the calibration of CIR by time change, we write the bond price as

$$\hat{B}(\hat{r}(0),0,\hat{T}),$$

thereby indicating explicitly the initial interest rate. The above says that

$$-\frac{\partial}{\partial \hat{T}}\log \hat{B}(\hat{r}(0),0,\hat{T})\Big|_{\hat{T}=0} = \hat{r}(0).$$

The calibration of CIR by time change requires that we find a strictly increasing function $\varphi$ with $\varphi(0)=0$ such that

$$-\log \hat{B}(\hat{r}(0),0,\hat{T}) = \frac{1}{\varphi'(0)}\hat{r}(0)C(\varphi(\hat{T})) + A(\varphi(\hat{T})), \quad \hat{T}\geq 0, \tag{cal}$$

where $\hat{B}(\hat{r}(0),0,\hat{T})$, determined by market data, is strictly increasing in $\hat{T}$, starts at 1 when $\hat{T}=0$, and goes to zero as $\hat{T}\to\infty$. Therefore, $-\log \hat{B}(\hat{r}(0),0,\hat{T})$ is as shown in Fig. 31.4.

Consider the function

$$\hat{r}(0)C(T) + A(T),$$

Here $C(T)$ and $A(T)$ are given by

$$\begin{aligned}
C(T) &= \frac{\sinh(\gamma T)}{\gamma\cosh(\gamma T) + \frac{1}{2}\beta\sinh(\gamma T)},\\
A(T) &= -\frac{2\alpha}{\sigma^2}\log\left[\frac{\gamma e^{\frac{1}{2}\beta T}}{\gamma\cosh(\gamma T)+\frac{1}{2}\beta\sinh(\gamma T)}\right],\\
\gamma &= \tfrac{1}{2}\sqrt{\beta^2+2\sigma^2}.
\end{aligned}$$

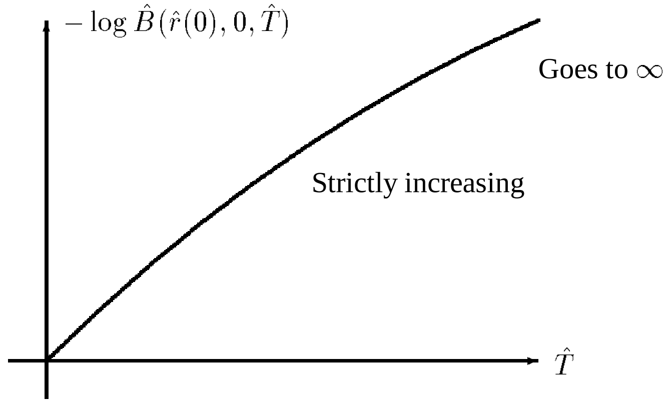



Figure 31.4: *Bond price in CIR model*

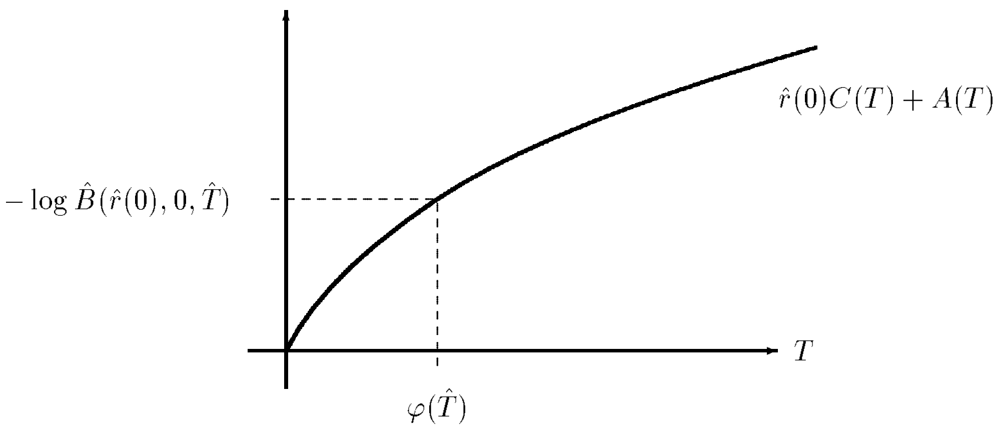



Figure 31.5: *Calibration*

The function $\hat{r}(0)C(T) + A(T)$ is zero at $T = 0$, is strictly increasing in $T$, and goes to $\infty$ as $T\to\infty$. This is because the interest rate is positive in the CIR model (see last paragraph of Section 31.4).

To solve (cal), let us first consider the related equation

$$-\log \hat{B}(\hat{r}(0), 0, \hat{T}) = \hat{r}(0)C(\varphi(\hat{T})) + A(\varphi(\hat{T})). \qquad \text{(cal')}$$

Fix $\hat{T}$ and define $\varphi(\hat{T})$ to be the unique $T$ for which (see Fig. 31.5)

$$-\log \hat{B}(\hat{r}(0), 0, \hat{T}) = \hat{r}(0)C(T) + A(T)$$

If $\hat{T} = 0$, then $\varphi(\hat{T}) = 0$. If $\hat{T}_1 < \hat{T}_2$, then $\varphi(\hat{T}_1) < \varphi(\hat{T}_2)$. As $\hat{T}\to\infty$, $\varphi(\hat{T})\to\infty$. We have thus defined a time-change function $\varphi$ which has all the right properties, except it satisfies (cal') rather than (cal).

We conclude by showing that $\varphi'(0) = 1$ so $\varphi$ also satisfies (cal). From (cal') we compute

$$\begin{aligned}\hat{r}(0) &= -\frac{\partial}{\partial \hat{T}} \log \hat{B}(\hat{r}(0), 0, \hat{T})\Big|_{\hat{T}=0} \\ &= \hat{r}(0) C'(\varphi(0))\varphi'(0) + A'(\varphi(0))\varphi'(0) \\ &= \hat{r}(0) C'(0)\varphi'(0) + A'(0)\varphi'(0).\end{aligned}$$

We show in a moment that $C'(0) = 1$, $A'(0) = 0$, so we have

$$\hat{r}(0) = \hat{r}(0)\varphi'(0).$$

Note that $\hat{r}(0)$ is the initial interest rate, observed in the market, and is striclty positive. Dividing by $\hat{r}(0)$, we obtain

$$\varphi'(0) = 1.$$

Computation of $C'(0)$:

$$\begin{aligned}C'(\tau) &= \frac{1}{\left(\gamma\cosh(\gamma\tau) + \frac{1}{2}\beta\sinh(\gamma\tau)\right)^2}\Big[\gamma\cosh(\gamma\tau)\left(\gamma\cosh(\gamma\tau) + \tfrac{1}{2}\beta\sinh(\gamma\tau)\right) \\ &\qquad - \sinh(\gamma\tau)\left(\gamma^2\sinh(\gamma\tau) + \tfrac{1}{2}\beta\gamma\cosh(\gamma\tau)\right)\Big] \\ C'(0) &= \frac{1}{\gamma^2}\left[\gamma(\gamma+0) - 0(0 + \tfrac{1}{2}\beta\gamma)\right] = 1.\end{aligned}$$

Computation of $A'(0)$:

$$\begin{aligned}A'(\tau) &= -\frac{2\alpha}{\sigma^2}\left[\frac{\gamma\cosh(\gamma\tau) + \frac{1}{2}\beta\sinh(\gamma\tau)}{\gamma e^{\beta\tau/2}}\right] \\ &\qquad \times \frac{1}{\left(\gamma\cosh(\gamma\tau) + \frac{1}{2}\beta\sinh(\gamma\tau)\right)^2}\Big[\frac{\beta\gamma}{2}e^{\beta\tau/2}\left(\gamma\cosh(\gamma\tau) + \tfrac{1}{2}\beta\sinh(\gamma\tau)\right) \\ &\qquad - \gamma e^{\beta\tau/2}\left(\gamma^2\sinh(\gamma\tau) + \tfrac{1}{2}\beta\gamma\cosh(\gamma\tau)\right)\Big], \\ A'(0) &= -\frac{2\alpha}{\sigma^2}\left[\frac{\gamma+0}{\gamma}\right]\frac{1}{(\gamma+0)^2}\left[\frac{\beta\gamma}{2}(\gamma+0) - \gamma(0 + \tfrac{1}{2}\beta\gamma)\right] \\ &= -\frac{2\alpha}{\sigma^2}\cdot\frac{1}{\gamma^2}\left[\frac{\beta\gamma^2}{2} - \tfrac{1}{2}\beta\gamma^2\right] \\ &= 0.\end{aligned}$$

# Chapter 32

# A two-factor model (Duffie & Kan)

Let us define:

$$X_1(t) = \text{Interest rate at time } t$$
$$X_2(t) = \text{Yield at time } t \text{ on a bond maturing at time } t + \tau_0$$

Let $X_1(0) > 0$, $X_2(0) > 0$ be given, and let $X_1(t)$ and $X_2(t)$ be given by the coupled stochastic differential equations

$$dX_1(t) = (a_{11}X_1(t) + a_{12}X_2(t) + b_1)\ dt + \sigma_1\sqrt{\beta_1X_1(t) + \beta_2X_2(t) + \alpha}\ \ dW_1(t), \tag{SDE1}$$

$$dX_2(t) = (a_{21}X_1(t) + a_{22}X_2(t) + b_2)\ dt + \sigma_2\sqrt{\beta_1X_1(t) + \beta_2X_2(t) + \alpha}\ (\rho\ dW_1(t) + \sqrt{1-\rho^2}\ dW_2(t)), \tag{SDE2}$$

where $W_1$ and $W_2$ are independent Brownian motions. To simplify notation, we define

$$Y(t) \stackrel{\triangle}{=} \beta_1X_1(t) + \beta_2X_2(t) + \alpha,$$
$$W_3(t) \stackrel{\triangle}{=} \rho W_1(t) + \sqrt{1-\rho^2}W_2(t).$$

Then $W_3$ is a Brownian motion with

$$dW_1(t)\ dW_3(t) = \rho\ dt,$$

and

$$dX_1\ dX_1 = \sigma_1^2 Y\ dt, \quad dX_2\ dX_2 = \sigma_2^2 Y\ dt, \quad dX_1\ dX_2 = \rho\sigma_1\sigma_2 Y\ dt.$$

## 32.1 Non-negativity of $Y$

$$
\begin{aligned}
dY &= \beta_1\, dX_1 + \beta_2\, dX_2 \\
&= (\beta_1 a_{11} X_1 + \beta_1 a_{12} X_2 + \beta_1 b_1)\, dt + (\beta_2 a_{21} X_1 + \beta_2 a_{22} X_2 + \beta_2 b_2)\, dt \\
&\qquad + \sqrt{Y}(\beta_1 \sigma_1\, dW_1 + \beta_2 \rho \sigma_2\, dW_1 + \beta_2 \sqrt{1-\rho^2}\sigma_2\, dW_2) \\
&= [(\beta_1 a_{11} + \beta_2 a_{21}) X_1 + (\beta_1 a_{12} + \beta_2 a_{22}) X_2]\, dt + (\beta_1 b_1 + \beta_2 b_2)\, dt \\
&\qquad + (\beta_1^2 \sigma_1^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2 + \beta_2^2\sigma_2^2)^{\frac{1}{2}} \sqrt{Y(t)}\, dW_4(t)
\end{aligned}
$$

where

$$
W_4(t) = \frac{(\beta_1\sigma_1 + \beta_2\rho\sigma_2) W_1(t) + \beta_2\sqrt{1-\rho^2}\sigma_2 W_2(t)}{\sqrt{\beta_1^2\sigma_1^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2 + \beta_2^2\sigma_2^2}}
$$

is a Brownian motion. We shall choose the parameters so that:

**Assumption 1:** For some $\gamma$, $\quad \beta_1 a_{11} + \beta_2 a_{21} = \gamma\beta_1, \quad \beta_1 a_{12} + \beta_2 a_{22} = \gamma\beta_2.$

Then

$$
\begin{aligned}
dY &= [\gamma\beta_1 X_1 + \gamma\beta_2 X_2 + \alpha\gamma]\, dt + (\beta_1 b_1 + \beta_2 b_2 - \alpha\gamma)\, dt \\
&\qquad + (\beta_1^2\sigma_1^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2 + \beta_2^2\sigma_2^2)^{\frac{1}{2}} \sqrt{Y}\, dW_4 \\
&= \gamma Y\, dt + (\beta_1 b_1 + \beta_2 b_2 - \alpha\gamma)\, dt + (\beta_1^2\sigma_1^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2 + \beta_2^2\sigma_2^2)^{\frac{1}{2}} \sqrt{Y}\, dW_4.
\end{aligned}
$$

From our discussion of the CIR process, we recall that $Y$ will stay strictly positive provided that:

**Assumption 2:** $\quad Y(0) = \beta_1 X_1(0) + \beta_2 X_2(0) + \alpha > 0,$

and

**Assumption 3:** $\quad \beta_1 b_1 + \beta_2 b_2 - \gamma\alpha \geq \frac{1}{2}(\beta_1^2\sigma_1^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2 + \beta_2^2\sigma_2^2).$

Under Assumptions 1,2, and 3,

$$
Y(t) > 0, \quad 0 \leq t < \infty, \text{ almost surely,}
$$

and (SDE1) and (SDE2) make sense. These can be rewritten as

$$
dX_1(t) = (a_{11} X_1(t) + a_{12} X_2(t) + b_1)\, dt + \sigma_1 \sqrt{Y(t)}\, dW_1(t), \tag{SDE1'}
$$

$$
dX_2(t) = (a_{21} X_1(t) + a_{22} X_2(t) + b_2)\, dt + \sigma_2 \sqrt{Y(t)}\, dW_3(t). \tag{SDE2'}
$$

## 32.2 Zero-coupon bond prices

The value at time $t \le T$ of a zero-coupon bond paying \$1 at time $T$ is

$$B(t,T) = I\!E\left[\exp\left\{-\int_t^T X_1(u)\,du\right\}\Big|\mathcal{F}(t)\right].$$

Since the pair $(X_1, X_2)$ of processes is Markov, this is random only through a dependence on $X_1(t), X_2(t)$. Since the coefficients in (SDE1) and (SDE2) do not depend on time, the bond price depends on $t$ and $T$ only through their difference $\tau = T - t$. Thus, there is a function $B(x_1, x_2, \tau)$ of the dummy variables $x_1, x_2$ and $\tau$, so that

$$B(X_1(t), X_2(t), T-t) = I\!E\left[\exp\left\{-\int_t^T X_1(u)\,du\right\}\Big|\mathcal{F}(t)\right].$$

The usual tower property argument shows that

$$\exp\left\{-\int_0^t X_1(u)\,du\right\} B(X_1(t), X_2(t), T-t)$$

is a martingale. We compute its stochastic differential and set the $dt$ term equal to zero.

$$\begin{aligned}
&d\left(\exp\left\{-\int_0^t X_1(u)\,du\right\} B(X_1(t), X_2(t), T-t)\right)\\
&= \exp\left\{-\int_0^t X_1(u)\,du\right\}\Big[-X_1 B\,dt + B_{x_1}\,dX_1 + B_{x_2}\,dX_2 - B_\tau\,dt\\
&\qquad\qquad + \tfrac{1}{2}B_{x_1x_1}\,dX_1\,dX_1 + B_{x_1x_2}\,dX_1\,dX_2 + \tfrac{1}{2}B_{x_2x_2}\,dX_2\,dX_2\Big]\\
&= \exp\left\{-\int_0^t X_1(u)\,du\right\}\Big[\Big(-X_1 B + (a_{11}X_1 + a_{12}X_2 + b_1)B_{x_1} + (a_{21}X_1 + a_{22}X_2 + b_2)B_{x_2} - B_\tau\\
&\qquad\qquad + \tfrac{1}{2}\sigma_1^2 Y B_{x_1x_1} + \rho\sigma_1\sigma_2 Y B_{x_1x_2} + \tfrac{1}{2}\sigma_2^2 Y B_{x_2x_2}\Big)\,dt\\
&\qquad\qquad + \sigma_1\sqrt{Y}B_{x_1}\,dW_1 + \sigma_2\sqrt{Y}B_{x_2}\,dW_3\Big]
\end{aligned}$$

The partial differential equation for $B(x_1, x_2, \tau)$ is

$$\begin{aligned}
-x_1B - B_\tau + (a_{11}x_1 + a_{12}x_2 + b_1)B_{x_1} + (a_{21}x_1 + a_{22}x_2 + b_2)B_{x_2} + \tfrac{1}{2}\sigma_1^2(\beta_1x_1 + \beta_2x_2 + \alpha)B_{x_1x_1}&\\
+ \rho\sigma_1\sigma_2(\beta_1x_1 + \beta_2x_2 + \alpha)B_{x_1x_2} + \tfrac{1}{2}\sigma_2^2(\beta_1x_1 + \beta_2x_2 + \alpha)B_{x_2x_2} = 0.& \quad \text{(PDE)}
\end{aligned}$$

We seek a solution of the form

$$B(x_1, x_2, \tau) = \exp\{-x_1C_1(\tau) - x_2C_2(\tau) - A(\tau)\},$$

valid for all $\tau \ge 0$ and all $x_1, x_2$ satisfying

$$\beta_1x_1 + \beta_2x_2 + \alpha > 0. \qquad (*)$$

We must have
$$B(x_1, x_2, 0) = 1, \quad \forall x_1, x_2 \quad \text{satisfying (*)},$$
because $\tau = 0$ corresponds to $t = T$. This implies the initial conditions
$$C_1(0) = C_2(0) = A(0) = 0. \tag{IC}$$
We want to find $C_1(\tau), C_2(\tau), A(\tau)$ for $\tau > 0$. We have
$$\begin{aligned}
B_\tau(x_1, x_2, \tau) &= \left[-x_1 C_1'(\tau) - x_2 C_2'(\tau) - A'(\tau)\right] B(x_1, x_2, \tau),\\
B_{x_1}(x_1, x_2, \tau) &= -C_1(\tau) B(x_1, x_2, \tau),\\
B_{x_2}(x_1, x_2, \tau) &= -C_2(\tau) B(x_1, x_2, \tau),\\
B_{x_1 x_1}(x_1, x_2, \tau) &= C_1^2(\tau) B(x_1, x_2, \tau),\\
B_{x_1 x_2}(x_1, x_2, \tau) &= C_1(\tau) C_2(\tau) B(x_1, x_2, \tau),\\
B_{x_2 x_2}(x_1, x_2, \tau) &= C_2^2(\tau) B(x_1, x_2, \tau).
\end{aligned}$$
(PDE) becomes
$$\begin{aligned}
0 = B(x_1, x_2, \tau)\Big[&-x_1 + x_1 C_1'(\tau) + x_2 C_2'(\tau) + A'(\tau) - (a_{11} x_1 + a_{12} x_2 + b_1) C_1(\tau)\\
&- (a_{21} x_1 + a_{22} x_2 + b_2) C_2(\tau)\\
&+ \tfrac{1}{2}\sigma_1^2(\beta_1 x_1 + \beta_2 x_2 + \alpha) C_1^2(\tau) + \rho\sigma_1\sigma_2(\beta_1 x_1 + \beta_2 x_2 + \alpha) C_1(\tau) C_2(\tau)\\
&+ \tfrac{1}{2}\sigma_2^2(\beta_1 x_1 + \beta_2 x_2 + \alpha) C_2^2(\tau)\Big]\\
= x_1 B(x_1, x_2, \tau)\Big[&-1 + C_1'(\tau) - a_{11} C_1(\tau) - a_{21} C_2(\tau)\\
&+ \tfrac{1}{2}\sigma_1^2 \beta_1 C_1^2(\tau) + \rho\sigma_1\sigma_2\beta_1 C_1(\tau) C_2(\tau) + \tfrac{1}{2}\sigma_2^2\beta_1 C_2^2(\tau)\Big]\\
+ x_2 B(x_1, x_2, \tau)\Big[&C_2'(\tau) - a_{12} C_1(\tau) - a_{22} C_2(\tau)\\
&+ \tfrac{1}{2}\sigma_1^2 \beta_2 C_1^2(\tau) + \rho\sigma_1\sigma_2\beta_2 C_1(\tau) C_2(\tau) + \tfrac{1}{2}\sigma_2^2\beta_2 C_2^2(\tau)\Big]\\
+ B(x_1, x_2, \tau)\Big[&A'(\tau) - b_1 C_1(\tau) - b_2 C_2(\tau)\\
&+ \tfrac{1}{2}\sigma_1^2 \alpha C_1^2(\tau) + \rho\sigma_1\sigma_2\alpha C_1(\tau) C_2(\tau) + \tfrac{1}{2}\sigma_2^2\alpha C_2^2(\tau)\Big]
\end{aligned}$$
We get three equations:
$$C_1'(\tau) = 1 + a_{11} C_1(\tau) + a_{21} C_2(\tau) - \tfrac{1}{2}\sigma_1^2\beta_1 C_1^2(\tau) - \rho\sigma_1\sigma_2\beta_1 C_1(\tau) C_2(\tau) - \tfrac{1}{2}\sigma_2^2\beta_1 C_2^2(\tau), \tag{1}$$
$$C_1(0) = 0;$$
$$C_2'(\tau) = a_{12} C_1(\tau) + a_{22} C_2(\tau) - \tfrac{1}{2}\sigma_1^2\beta_2 C_1^2(\tau) - \rho\sigma_1\sigma_2\beta_2 C_1(\tau) C_2(\tau) - \tfrac{1}{2}\sigma_2^2\beta_2 C_2^2(\tau), \tag{2}$$
$$C_2(0) = 0;$$
$$A'(\tau) = b_1 C_1(\tau) + b_2 C_2(\tau) - \tfrac{1}{2}\sigma_1^2\alpha C_1^2(\tau) - \rho\sigma_1\sigma_2\alpha C_1(\tau) C_2(\tau) - \tfrac{1}{2}\sigma_2^2\alpha C_2^2(\tau), \tag{3}$$
$$A(0) = 0;$$

We first solve (1) and (2) simultaneously numerically, and then integrate (3) to obtain the function $A(\tau)$.

## 32.3 Calibration

Let $\tau_0 > 0$ be given. The value at time $t$ of a bond maturing at time $t + \tau_0$ is

$$B(X_1(t), X_2(t), \tau_0) = \exp\{-X_1(t)C_1(\tau_0) - X_2(t)C_2(\tau_0) - A(\tau_0)\}$$

and the yield is

$$-\frac{1}{\tau_0} \log B(X_1(t), X_2(t), \tau_0) = \frac{1}{\tau_0} \left[X_1(t)C_1(\tau_0) + X_2(t)C_2(\tau_0) + A(\tau_0)\right].$$

But we have set up the model so that $X_2(t)$ is the yield at time $t$ of a bond maturing at time $t + \tau_0$. Thus

$$X_2(t) = \frac{1}{\tau_0} \left[X_1(t)C_1(\tau_0) + X_2(t)C_2(\tau_0) + A(\tau_0)\right].$$

This equation must hold for every value of $X_1(t)$ and $X_2(t)$, which implies that

$$C_1(\tau_0) = 0, \;\; C_2(\tau_0) = \tau_0, \;\; A(\tau) = 0.$$

We must choose the parameters

$$a_{11}, a_{12}, b_1; \;\; a_{21}, a_{22}, b_2; \;\; \beta_1, \beta_2, \alpha; \;\; \sigma_1, \rho, \sigma_2;$$

so that these three equations are satisfied.

# Chapter 33

# Change of numéraire

Consider a Brownian motion driven market model with time horizon $T^*$. For now, we will have one asset, which we call a "stock" even though in applications it will usually be an interest rate dependent claim. The price of the stock is modeled by

$$dS(t) = r(t)\ S(t)\ dt + \sigma(t)S(t)\ dW(t), \tag{0.1}$$

where the interest rate process $r(t)$ and the volatility process $\sigma(t)$ are adapted to some filtration $\{\mathcal{F}(t);\ 0 \le t \le T^*\}$. $W$ is a Brownian motion relative to this filtration, but $\{\mathcal{F}(t);\ 0 \le t \le T^*\}$ may be larger than the filtration generated by $W$.

This is *not* a geometric Brownian motion model. We are particularly interested in the case that the interest rate is stochastic, given by a term structure model we have not yet specified.

We shall work only under the risk-neutral measure, which is reflected by the fact that the mean rate of return for the stock is $r(t)$.

We define the *accumulation factor*

$$\beta(t) = \exp\left\{\int_0^t r(u)\ du\right\},$$

so that the discounted stock price $\frac{S(t)}{\beta(t)}$ is a martingale. Indeed,

$$d\left(\frac{S(t)}{\beta(t)}\right) = \frac{S(t)}{\beta(t)}\sigma(t)\ dW(t).$$

The zero-coupon bond prices are given by

$$\begin{aligned} B(t,T) &= I\!E\left[\exp\left\{-\int_t^T r(u)\ du\right\}\middle|\mathcal{F}(t)\right] \\ &= I\!E\left[\frac{\beta(t)}{\beta(T)}\middle|\mathcal{F}(t)\right], \end{aligned}$$

so

$$\frac{B(t,T)}{\beta(t)} = \mathbb{E}\left[\frac{1}{\beta(T)}\middle|\mathcal{F}(t)\right]$$

is also a martingale (tower property).

The $T$-*forward price* $F(t,T)$ of the stock is the price set at time $t$ for delivery of one share of stock at time $T$ with payment at time $T$. The value of the forward contract at time $t$ is zero, so

$$\begin{aligned}
0 &= \mathbb{E}\left[\frac{\beta(t)}{\beta(T)}\left(S(T) - F(t,T)\right)\middle|\mathcal{F}(t)\right] \\
&= \beta(t)\mathbb{E}\left[\frac{S(T)}{\beta(T)}\middle|Ft\right] - F(t,T)\mathbb{E}\left[\frac{\beta(t)}{\beta(T)}\middle|\mathcal{F}(t)\right] \\
&= \beta(t)\frac{S(t)}{\beta(t)} - F(t,T)B(t,T) \\
&= S(t) - F(t,T)B(t,T)
\end{aligned}$$

Therefore,

$$F(t,T) = \frac{S(t)}{B(t,T)}.$$

**Definition 33.1 (Numéraire)** Any asset in the model whose price is always strictly positive can be taken as the numéraire. We then denominate all other assets in units of this numéraire.

**Example 33.1 (Money market as numéraire)** The money market could be the numéraire. At time $t$, the stock is worth $\frac{S(t)}{\beta(t)}$ units of money market and the $T$-maturity bond is worth $\frac{B(t,T)}{\beta(t)}$ units of money market. ■

**Example 33.2 (Bond as numéraire)** The $T$-maturity bond could be the numéraire. At time $t \le T$, the stock is worth $F(t,T)$ units of $T$-maturity bond and the $T$-maturity bond is worth 1 unit. ■

We will say that a probability measure $\mathbb{P}_N$ is *risk-neutral for the numéraire* $N$ if every asset price, divided by $N$, is a martingale under $\mathbb{P}_N$. The original probability measure $\mathbb{P}$ is risk-neutral for the numéraire $\beta$ (Example 33.1).

**Theorem 0.71** *Let $N$ be a numéraire, i.e., the price process for some asset whose price is always strictly positive. Then $\mathbb{P}_N$ defined by*

$$\mathbb{P}_N(A) = \frac{1}{N(0)}\int_A \frac{N(T^*)}{\beta(T^*)}\, d\mathbb{P}, \quad \forall A \in \mathcal{F}(T^*),$$

*is risk-neutral for $N$.*

*Note:* $I\!\!P$ and $I\!\!P_N$ are equivalent, i.e., have the same probability zero sets, and

$$I\!\!P(A) = N(0) \int_A \frac{\beta(T^*)}{N(T^*)}\, dI\!\!P_N, \quad \forall A \in \mathcal{F}(T^*).$$

**Proof:** Because $N$ is the price process for some asset, $N/\beta$ is a martingale under $I\!\!P$. Therefore,

$$\begin{aligned} I\!\!P_N(\Omega) &= \frac{1}{N(0)} \int_\Omega \frac{N(T^*)}{\beta(T^*)}\, dI\!\!P \\ &= \frac{1}{N(0)}.I\!\!E\left[\frac{N(T^*)}{\beta(T^*)}\right] \\ &= \frac{1}{N(0)} \frac{N(0)}{\beta(0)} \\ &= 1, \end{aligned}$$

and we see that $I\!\!P_N$ is a probability measure.

Let $Y$ be an asset price. Under $I\!\!P$, $Y/\beta$ is a martingale. We must show that under $I\!\!P_N$, $Y/N$ is a martingale. For this, we need to recall how to combine conditional expectations with change of measure (Lemma 1.54). If $0 \le t \le T \le T^*$ and $X$ is $\mathcal{F}(T)$-measurable, then

$$\begin{aligned} I\!\!E_N\left[X\middle|\mathcal{F}(t)\right] &= \frac{N(0)\beta(t)}{N(t)} I\!\!E\left[\frac{N(T)}{N(0)\beta(T)} X\middle|\mathcal{F}(t)\right] \\ &= \frac{\beta(t)}{N(t)} I\!\!E\left[\frac{N(T)}{\beta(T)} X\middle|\mathcal{F}(t)\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} I\!\!E_N\left[\frac{Y(T)}{N(T)}\middle|\mathcal{F}(t)\right] &= \frac{\beta(t)}{N(t)} I\!\!E\left[\frac{N(T)}{\beta(T)} \frac{Y(T)}{N(T)}\middle|\mathcal{F}(t)\right] \\ &= \frac{\beta(t)}{N(t)} \frac{Y(t)}{\beta(t)} \\ &= \frac{Y(t)}{N(t)}, \end{aligned}$$

which is the martingale property for $Y/N$ under $I\!\!P_N$. ■

## 33.1 Bond price as numéraire

Fix $T \in (0, T^*]$ and let $B(t,T)$ be the numéraire. The risk-neutral measure for this numéraire is

$$\begin{aligned} I\!\!P_T(A) &= \frac{1}{B(0,T)} \int_A \frac{B(T,T)}{\beta(T)}\, dI\!\!P \\ &= \frac{1}{B(0,T)} \int_A \frac{1}{\beta(T)}\, dI\!\!P \quad \forall A \in \mathcal{F}(T). \end{aligned}$$

Because this bond is not defined after time $T$, we change the measure only "up to time $T$", i.e., using $\frac{1}{B(0,T)}\frac{B(T,T)}{\beta(T)}$ and only for $A \in \mathcal{F}(T)$.

$I\!\!P_T$ is called the $T$-*forward measure.* Denominated in units of $T$-maturity bond, the value of the stock is

$$F(t,T) = \frac{S(t)}{B(t,T)}, \quad 0 \le t \le T.$$

This is a martingale under $I\!\!P_T$, and so has a differential of the form

$$dF(t,T) = \sigma_F(t,T)F(t,T)\,dW_T(t), \quad 0 \le t \le T, \tag{1.1}$$

i.e., a differential without a $dt$ term. The process $\{W_T;\ 0 \le t \le T\}$ is a Brownian motion under $I\!\!P_T$. We may assume without loss of generality that $\sigma_F(t,T) \ge 0$.

We write $F(t)$ rather than $F(t,T)$ from now on.

## 33.2 Stock price as numéraire

Let $S(t)$ be the numéraire. In terms of this numéraire, the stock price is identically 1. The risk-neutral measure under this numéraire is

$$I\!\!P_S(A) = \frac{1}{S(0)} \int_A \frac{S(T^*)}{\beta(T^*)}\,dI\!\!P, \quad \forall A \in \mathcal{F}(T^*).$$

Denominated in shares of stock, the value of the $T$-maturity bond is

$$\frac{B(t,T)}{S(t)} = \frac{1}{F(t)}.$$

This is a martingale under $I\!\!P_S$, and so has a differential of the form

$$d\left(\frac{1}{F(t)}\right) = \gamma(t,T)\left(\frac{1}{F(t)}\right)\,dW_S(t), \tag{2.1}$$

where $\{W_S(t);\ 0 \le t \le T^*\}$ is a Brownian motion under $I\!\!P_S$. We may assume without loss of generality that $\gamma(t,T) \ge 0$.

**Theorem 2.72** *The volatility $\gamma(t,T)$ in (2.1) is equal to the volatility $\sigma_F(t,T)$ in (1.1). In other words, (2.1) can be rewritten as*

$$d\left(\frac{1}{F(t)}\right) = \sigma_F(t,T)\left(\frac{1}{F(t)}\right)\,dW_S(t), \tag{2.1'}$$

**Proof:** Let $g(x) = 1/x$, so $g'(x) = -1/x^2,\ g''(x) = 2/x^3$. Then

$$\begin{aligned}
d\left(\frac{1}{F(t)}\right) &= dg(F(t)) \\
&= g'(F(t))\ dF(t) + \tfrac{1}{2}g''(F(t))\ dF(t)\ dF(t) \\
&= -\frac{1}{F^2(t)}\sigma_F(t,T)F(t,T)\ dW_T(t) + \frac{1}{F^3(t)}\sigma_F^2(t,T)F^2(t,T)\ dt \\
&= \frac{1}{F(t)}\left[-\sigma_F(t,T)\ dW_T(t) + \sigma_F^2(t,T)\ dt\right] \\
&= \sigma_F(t,T)\left(\frac{1}{F(t)}\right)[-dW_T(t) + \sigma_F(t,T)\ dt].
\end{aligned}$$

Under $I\!\!P_T$, $-W_T$ is a Brownian motion. Under this measure, $\frac{1}{F(t)}$ has volatility $\sigma_F(t,T)$ and mean rate of return $\sigma_F^2(t,T)$. The change of measure from $I\!\!P_T$ to $I\!\!P_S$ makes $\frac{1}{F(t)}$ a martingale, i.e., it changes the mean return to zero, but the change of measure does not affect the volatility. Therefore, $\gamma(t,T)$ in (2.1) must be $\sigma_F(t,T)$ and $W_S$ must be

$$W_S(t) = -W_T(t) + \int_0^t \sigma_F(u,T)\ du.$$

∎

## 33.3 Merton option pricing formula

The price at time zero of a European call is

$$\begin{aligned}
V(0) &= I\!\!E\left[\frac{1}{\beta(T)}(S(T)-K)^+\right] \\
&= I\!\!E\left[\frac{S(T)}{\beta(T)}\mathbf{1}_{\{S(T)>K\}}\right] - KI\!\!E\left[\frac{1}{\beta(T)}\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S(0)\int_{\{S(T)>K\}}\frac{S(T)}{S(0)\beta(T)}\ dI\!\!P - KB(0,T)\int_{\{S(T)>K\}}\frac{1}{B(0,T)\beta(T)}\ dI\!\!P \\
&= S(0)I\!\!P_S\{S(T)>K\} - KB(0,T)I\!\!P_T\{S(T)>K\} \\
&= S(0)I\!\!P_S\{F(T)>K\} - KB(0,T)I\!\!P_T\{F(T)>K\} \\
&= S(0)I\!\!P_S\left\{\frac{1}{F(T)} < \frac{1}{K}\right\} - KB(0,T)I\!\!P_T\{F(T)>K\}.
\end{aligned}$$

This is a completely general formula which permits computation as soon as we specify $\sigma_F(t,T)$. If we assume that $\sigma_F(t,T)$ is a constant $\sigma_F$, we have the following:

$$\begin{aligned}
\frac{1}{F(T)} &= \frac{B(0,T)}{S(0)} \exp\left\{\sigma_F W_S(T) - \tfrac{1}{2}\sigma_F^2 T\right\}, \\
I\!\!P_S\left(\frac{1}{F(T)} < \frac{1}{K}\right) &= I\!\!P_S\left\{\sigma_F W_S(T) - \tfrac{1}{2}\sigma_F^2 T < \log \frac{S(0)}{KB(0,T)}\right\} \\
&= I\!\!P_S\left\{\frac{W_S(T)}{\sqrt{T}} < \frac{1}{\sigma_F\sqrt{T}} \log \frac{S(0)}{KB(0,T)} + \tfrac{1}{2}\sigma_F\sqrt{T}\right\} \\
&= N(\rho_1),
\end{aligned}$$

where

$$\rho_1 = \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{S(0)}{KB(0,T)} + \tfrac{1}{2}\sigma_F^2 T\right].$$

Similarly,

$$\begin{aligned}
F(T) &= \frac{S(0)}{B(0,T)} \exp\left\{\sigma_F W_T(T) - \tfrac{1}{2}\sigma_F^2 T\right\}, \\
I\!\!P_T\{F(T) > K\} &= I\!\!P_T\left\{\sigma_F W_T(T) - \tfrac{1}{2}\sigma_F^2 T > \log \frac{KB(0,T)}{S(0)}\right\} \\
&= I\!\!P_T\left\{\frac{W_T(T)}{\sqrt{T}} > \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{KB(0,T)}{S(0)} + \tfrac{1}{2}\sigma_F^2 T\right]\right\} \\
&= I\!\!P_T\left\{\frac{-W_T(T)}{\sqrt{T}} < \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{S(0)}{KB(0,T)} - \tfrac{1}{2}\sigma_F^2 T\right]\right\} \\
&= N(\rho_2),
\end{aligned}$$

where

$$\rho_2 = \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{S(0)}{KB(0,T)} - \tfrac{1}{2}\sigma_F^2 T\right].$$

If $r$ is constant, then $B(0,T) = e^{-rT}$,

$$\begin{aligned}
\rho_1 &= \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{S(0)}{K} + (r + \tfrac{1}{2}\sigma_F^2)T\right], \\
\rho_2 &= \frac{1}{\sigma_F\sqrt{T}}\left[\log \frac{S(0)}{K} + (r - \tfrac{1}{2}\sigma_F^2)T\right],
\end{aligned}$$

and we have the usual Black-Scholes formula. When $r$ is not constant, we still have the explicit formula

$$V(0) = S(0)N(\rho_1) - KB(0,T)N(\rho_2).$$

As this formula suggests, if $\sigma_F$ is constant, then for $0 \le t \le T$, the value of a European call expiring at time $T$ is

$$V(t) = S(t)N(\rho_1(t)) - KB(t,T)N(\rho_2(t)),$$

where

$$\rho_1(t) = \frac{1}{\sigma_F\sqrt{T-t}}\left[\log\frac{F(t)}{K} + \tfrac{1}{2}\sigma_F^2(T-t)\right],$$
$$\rho_2(t) = \frac{1}{\sigma_F\sqrt{T-t}}\left[\log\frac{F(t)}{K} - \tfrac{1}{2}\sigma_F^2(T-t)\right].$$

This formula also suggests a hedge: at each time $t$, hold $N(\rho_1(t))$ shares of stock and short $KN(\rho_2(t))$ bonds.

We want to verify that this hedge is *self-financing.* Suppose we begin with \$ $V(0)$ and at each time $t$ hold $N(\rho_1(t))$ shares of stock. We short bonds as necessary to finance this. Will the position in the bond always be $-KN(\rho_2(t))$? If so, the value of the portfolio will always be

$$S(t)N(\rho_1(t)) - KB(t,T)N(\rho_2(t)) = V(t),$$

and we will have a hedge.

Mathematically, this question takes the following form. Let

$$\Delta(t) = N(\rho_1(t)).$$

At time $t$, hold $\Delta(t)$ shares of stock. If $X(t)$ is the value of the portfolio at time $t$, then $X(t) - \Delta(t)S(t)$ will be invested in the bond, so the number of bonds owned is $\frac{X(t)-\Delta(t)}{B(t,T)}S(t)$ and the portfolio value evolves according to

$$dX(t) = \Delta(t)\ dS(t) + \frac{X(t)-\Delta(t)}{B(t,T)}S(t)\ dB(t,T). \tag{3.1}$$

The value of the option evolves according to

$$\begin{aligned} dV(t) &= N(\rho_1(t))\ dS(t) + S(t)\ dN(\rho_1(t)) + dS(t)\ dN(\rho_1(t)) \\ &\quad - KN(\rho_2(t))\ dB(t,T) - K\ dB(t,T)\ dN(\rho_2(t)) - KB(t,T)\ dN(\rho_2(t)). \end{aligned} \tag{3.2}$$

If $X(0) = V(0)$, will $X(t) = V(t)$ for $0 \le t \le T$?

Formulas (3.1) and (3.2) are difficult to compare, so we simplify them by a change of numéraire. This change is justified by the following theorem.

**Theorem 3.73** *Changes of numéraire affect portfolio values in the way you would expect.*

**Proof:** Suppose we have a model with $k$ assets with prices $S_1, S_2, \ldots, S_k$. At each time $t$, hold $\Delta_i(t)$ shares of asset $i$, $i = 1, 2, \ldots, k-1$, and invest the remaining wealth in asset $k$. Begin with a nonrandom initial wealth $X(0)$, and let $X(t)$ be the value of the portfolio at time $t$. The number of shares of asset $k$ held at time $t$ is

$$\Delta_k(t) = \frac{\left(X(t) - \sum_{i=1}^{k-1}\Delta_i(t)S_i(t)\right)}{S_k(t)},$$

and $X$ evolves according to the equation

$$
\begin{aligned}
dX &= \sum_{i=1}^{k-1} \Delta_i \; dS_i + \left( X - \sum_{i=1}^{k-1} \Delta_i S_i \right) \frac{dS_k}{S_k} \\
&= \sum_{i=1}^{k} \Delta_i \; dS_i.
\end{aligned}
$$

Note that

$$
X_k(t) = \sum_{i=1}^{k} \Delta_i(t) S_i(t),
$$

and we only get to specify $\Delta_1, \ldots, \Delta_{k-1}$, not $\Delta_k$, in advance.

Let $N$ be a numéraire, and define

$$
\widehat{X}(t) = \frac{X(t)}{N(t)}, \quad \widehat{S_i}(t) = \frac{S_i(t)}{N(t)}, \quad i = 1, 2, \ldots, k.
$$

Then

$$
\begin{aligned}
d\widehat{X} &= \frac{1}{N} \; dX + X \; d\left(\frac{1}{N}\right) + dX \; d\left(\frac{1}{N}\right) \\
&= \frac{1}{N} \sum_{i=1}^{k} \Delta_i \; dS_i + \left( \sum_{i=1}^{k} \Delta_i S_i \right) d\left(\frac{1}{N}\right) + \sum_{i=1}^{k} \Delta_i \; dS_i \; d\left(\frac{1}{N}\right) \\
&= \sum_{i=1}^{k} \Delta_i \left( \frac{1}{N} \; dS_i + S_i d\left(\frac{1}{N}\right) + dS_i \; d\left(\frac{1}{N}\right) \right) \\
&= \sum_{i=1}^{k} \Delta_i \; d\widehat{S_i}.
\end{aligned}
$$

Now

$$
\begin{aligned}
\Delta_k &= \frac{\left( X - \sum_{i=1}^{k-1} \Delta_i S_i \right)}{S_k} \\
&= \frac{\left( X/N - \sum_{i=1}^{k-1} \Delta_i S_i/N \right)}{S_k/N} \\
&= \frac{\widehat{X} - \sum_{i=1}^{k-1} \Delta_i \widehat{S_i}}{\widehat{S_k}}.
\end{aligned}
$$

Therefore,

$$
d\widehat{X} = \sum_{i=1}^{k} \Delta_i \; d\widehat{S_i} + \left( \widehat{X} - \sum_{i=1}^{k-1} \Delta_i \widehat{S_i} \right) \frac{d\widehat{S_k}}{\widehat{S_k}}
$$

This is the formula for the evolution of a portfolio which holds $\Delta_i$ shares of asset $i$, $i = 1, 2, \ldots, k-1$, and all assets and the portfolio are denominated in units of $N$. ■

We return to the European call hedging problem (comparison of (3.1) and (3.2)), but we now use the zero-coupon bond as numéraire. We still hold $\Delta(t) = N(\rho_1(t))$ shares of stock at each time $t$. In terms of the new numéraire, the asset values are

$$\begin{aligned} \text{Stock:} \quad & \frac{S(t)}{B(t,T)} = F(t), \\ \text{Bond:} \quad & \frac{B(t,T)}{B(t,T)} = 1. \end{aligned}$$

The portfolio value evolves according to

$$d\widehat{X}(t) = \Delta(t)\; dF(t) + (\widehat{X}(t) - \Delta(t))\frac{d(1)}{1} = \Delta(t)\; dF(t). \tag{3.1'}$$

In the new numéraire, the option value formula

$$V(t) = N(\rho_1(t))S(t) - KB(t,T)N(\rho_2(t))$$

becomes

$$\widehat{V}(t) = \frac{V(t)}{B(t,T)} = N(\rho_1(t))F(t) - KN(\rho_2(t)),$$

and

$$d\widehat{V} = N(\rho_1(t))\; dF(t) + F(t)\; dN(\rho_1(t)) + dN(\rho_1(t))\; dF(t) - K\; dN(\rho_2(t)). \tag{3.2'}$$

To show that the hedge works, we must show that

$$F(t)\; dN(\rho_1(t)) + dN(\rho_1(t))\; dF(t) - K\; dN(\rho_2(t)) = 0.$$

This is a homework problem.

# Chapter 34

# Brace-Gatarek-Musiela model

## 34.1 Review of HJM under risk-neutral $I\!P$

$$\begin{aligned} f(t,T) &= \quad \text{Forward rate at time } t \text{ for borrowing at time } T. \\ df(t,T) &= \quad \sigma(t,T)\sigma^*(t,T)\ dt + \sigma(t,T)\ dW(t), \end{aligned}$$

where

$$\sigma^*(t,T) = \int_t^T \sigma(t,u)\ du$$

The interest rate is $r(t) = f(t,t)$. The bond prices

$$\begin{aligned} B(t,T) &= I\!E\left[\exp\left\{-\int_t^T r(u)\ du\right\}\middle|\mathcal{F}(t)\right] \\ &= \exp\left\{-\int_t^T f(t,u)\ du\right\} \end{aligned}$$

satisfy

$$dB(t,T) = r(t)\ B(t,T)\ dt - \underbrace{\sigma^*(t,T)}_{\text{volatility of } T\text{-maturity bond.}} B(t,T)\ dW(t).$$

To implement HJM, you specify a function

$$\sigma(t,T), \quad 0 \le t \le T.$$

A simple choice we would like to use is

$$\sigma(t,T) = \sigma f(t,T)$$

where $\sigma > 0$ is the constant "volatility of the forward rate". This is not possible because it leads to

$$\begin{aligned} \sigma^*(t,T) &= \sigma \int_t^T f(t,u)\ du, \\ df(t,T) &= \sigma^2 f(t,T)\left(\int_t^T f(t,u)\ du\right)\ dt + \sigma f(t,T)\ dW(t), \end{aligned}$$

and Heath, Jarrow and Morton show that solutions to this equation explode before $T$.

The problem with the above equation is that the $dt$ term grows like the square of the forward rate. To see what problem this causes, consider the similar deterministic ordinary differential equation

$$f'(t) = f^2(t),$$

where $f(0) = c > 0$. We have

$$\begin{aligned}
\frac{f'(t)}{f^2(t)} &= 1, \\
-\frac{d}{dt}\frac{1}{f(t)} &= 1, \\
-\frac{1}{f(t)} + \frac{1}{f(0)} &= \int_0^t 1\, du = t \\
-\frac{1}{f(t)} &= t - \frac{1}{f(0)} = t - 1/c = \frac{ct-1}{c}, \\
f(t) &= \frac{c}{1-ct}.
\end{aligned}$$

This solution explodes at $t = 1/c$.

## 34.2 Brace-Gatarek-Musiela model

New variables:

Current time $t$

Time to maturity $\tau = T - t$.

Forward rates:

$$r(t,\tau) = f(t, t+\tau), \quad r(t,0) = f(t,t) = r(t), \tag{2.1}$$

$$\frac{\partial}{\partial \tau} r(t,\tau) = \frac{\partial}{\partial T} f(t, t+\tau) \tag{2.2}$$

Bond prices:

$$\begin{aligned}
D(t,\tau) &= B(t, t+\tau) && (2.3)\\
&= \exp\left\{-\int_t^{t+\tau} f(t,v)\, dv\right\} \\
(u = v - t;\ du = dv): \quad &= \exp\left\{-\int_0^{\tau} f(t,t+u)\, du\right\} \\
&= \exp\left\{-\int_0^{\tau} r(t,u)\, du\right\} \\
\frac{\partial}{\partial \tau} D(t,\tau) &= \frac{\partial}{\partial T} B(t, t+\tau) = -r(t,\tau) D(t,\tau). && (2.4)
\end{aligned}$$

We will now write $\sigma(t,\tau) = \sigma(t, T-t)$ rather than $\sigma(t,T)$. In this notation, the HJM model is

$$df(t,T) = \sigma(t,\tau)\sigma^*(t,\tau)\; dt + \sigma(t,\tau)\; dW(t), \tag{2.5}$$
$$dB(t,T) = r(t)B(t,T)\; dt - \sigma^*(t,\tau)B(t,T)\; dW(t), \tag{2.6}$$

where

$$\sigma^*(t,\tau) = \int_0^\tau \sigma(t,u)\; du, \tag{2.7}$$
$$\frac{\partial}{\partial\tau}\sigma^*(t,\tau) = \sigma(t,\tau). \tag{2.8}$$

We now derive the differentials of $r(t,\tau)$ and $D(t,\tau)$, analogous to (2.5) and (2.6) We have

$$\begin{aligned} dr(t,\tau) &= \underbrace{df(t,t+\tau)}_{\text{differential applies only to first argument}} + \frac{\partial}{\partial T}f(t,t+\tau)\; dt \\ &\overset{(2.5),(2.2)}{=} \sigma(t,\tau)\sigma^*(t,\tau)\; dt + \sigma(t,\tau)\; dW(t) + \frac{\partial}{\partial\tau}r(t,\tau)\; dt \\ &\overset{(2.8)}{=} \frac{\partial}{\partial\tau}\left[r(t,\tau) + \tfrac{1}{2}(\sigma^*(t,\tau))^2\right]\; dt + \sigma(t,\tau)\; dW(t). \end{aligned} \tag{2.9}$$

Also,

$$\begin{aligned} dD(t,\tau) &= \underbrace{dB(t,t+\tau)}_{\text{differential applies only to first argument}} + \frac{\partial}{\partial T}B(t,t+\tau)\; dt \\ &\overset{(2.6),(2.4)}{=} r(t)\; B(t,t+\tau)\; dt - \sigma^*(t,\tau)B(t,t+\tau)\; dW(t) - r(t,\tau)D(t,\tau)\; dt \\ &\overset{(2.1)}{=} [r(t,0) - r(t,\tau)]\, D(t,\tau)\; dt - \sigma^*(t,\tau)D(t,\tau)\; dW(t). \end{aligned} \tag{2.10}$$

## 34.3 LIBOR

Fix $\delta > 0$ (say, $\delta = \frac{1}{4}$ year). \$ $D(t,\delta)$ invested at time $t$ in a $(t+\delta)$-maturity bond grows to \$ 1 at time $t+\delta$. $L(t,0)$ is defined to be the corresponding rate of simple interest:

$$\begin{aligned} D(t,\delta)(1+\delta L(t,0)) &= 1, \\ 1 + \delta L(t,0) &= \frac{1}{D(t,\delta)} = \exp\left\{\int_0^\partial r(t,u)\; du\right\}, \\ L(t,0) &= \frac{\exp\left\{\int_0^\partial r(t,u)\; du\right\} - 1}{\delta}. \end{aligned}$$

## 34.4 Forward LIBOR

$\delta > 0$ is still fixed. At time $t$, agree to invest \$ $\frac{D(t,\tau+\delta)}{D(t,\tau)}$ at time $t+\tau$, with payback of \$1 at time $t+\tau+\delta$. Can do this at time $t$ by shorting $\frac{D(t,\tau+\delta)}{D(t,\tau)}$ bonds maturing at time $t+\tau$ and going long one bond maturing at time $t+\tau+\delta$. The value of this portfolio at time $t$ is

$$-\frac{D(t,\tau+\delta)}{D(t,\tau)}D(t,\tau)+D(t,\tau+\delta)=0.$$

The *forward LIBOR* $L(t,\tau)$ is defined to be the simple (forward) interest rate for this investment:

$$\begin{aligned}
\frac{D(t,\tau+\delta)}{D(t,\tau)}\left(1+\delta L(t,\tau)\right)&=1,\\
1+\delta L(t,\tau)=\frac{D(t,\tau)}{D(t,\tau+\delta)}&=\frac{\exp\left\{-\int_0^\tau r(t,u)\,du\right\}}{\exp\left\{-\int_0^{\tau+\delta} r(t,u)\,du\right\}}\\
&=\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\},\\
L(t,\tau)&=\frac{\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\}-1}{\delta}.
\end{aligned} \tag{4.1}$$

Connection with forward rates:

$$\begin{aligned}
\frac{\partial}{\partial\delta}\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\}\Bigg|_{\delta=0}&=r(t,\tau+\delta)\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\}\Bigg|_{\delta=0}\\
&=r(t,\tau),
\end{aligned}$$

so

$$\begin{aligned}
f(t,t+\tau)=r(t,\tau)&=\lim_{\delta\downarrow 0}\frac{\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\}-1}{\delta}\\
L(t,\tau)&=\frac{\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\,du\right\}-1}{\delta},\quad \delta>0 \text{ fixed.}
\end{aligned} \tag{4.2}$$

$r(t,\tau)$ is the continuously compounded rate. $L(t,\tau)$ is the simple rate over a period of duration $\delta$.

We cannot have a log-normal model for $r(t,\tau)$ because solutions explode as we saw in Section 34.1. For fixed positive $\delta$, we *can* have a log-normal model for $L(t,\tau)$.

## 34.5 The dynamics of $L(t,\tau)$

We want to choose $\sigma(t,\tau),\ t\geq 0,\ \tau\geq 0$, appearing in (2.5) so that

$$dL(t,\tau)=(\ldots)\,dt+L(t,\tau)\,\gamma(t,\tau)\,dW(t)$$

for some $\gamma(t,\tau),\ t \geq 0, \tau \geq 0$. This is the BGM model, and is a subclass of HJM models, corresponding to particular choices of $\sigma(t,\tau)$.

Recall (2.9):

$$dr(t,\tau) = \frac{\partial}{\partial u}\left[r(t,u) + \tfrac{1}{2}(\sigma^*(t,u))^2\right]\ dt + \sigma(t,u)\ dW(t).$$

Therefore,

$$\begin{aligned} d\left(\int_\tau^{\tau+\delta} r(t,u)\ du\right) &= \int_\tau^{\tau+\delta} dr(t,u)\ du \\ &= \int_\tau^{\tau+\delta} \frac{\partial}{\partial u}\left[r(t,u) + \tfrac{1}{2}(\sigma^*(t,u))^2\right]\ du\ dt + \int_\tau^{\tau+\delta} \sigma(t,u)\ du\ dW(t) \\ &= \left[r(t,\tau+\delta) - r(t,\tau) + \tfrac{1}{2}(\sigma^*(t,\tau+\delta))^2 - \tfrac{1}{2}(\sigma^*(t,\tau))^2\right]\ dt \\ &\qquad + [\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]\ dW(t) \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} dL(t,\tau) &\overset{(4.1)}{=} d\left[\frac{\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\ du\right\} - 1}{\delta}\right] \\ &= \frac{1}{\delta}\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\ du\right\}\ d\int_\tau^{\tau+\delta} r(t,u)\ du \\ &\qquad + \frac{1}{2\delta}\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\ du\right\}\ \left(d\int_\tau^{\tau+\delta} r(t,u)\ du\right)^2 \\ &\overset{(4.1),(5.1)}{=} \frac{1}{\delta}[1+\delta L(t,\tau)] \times \\ &\qquad \times \Big\{[r(t,\tau+\delta) - r(t,\delta) + \tfrac{1}{2}(\sigma^*(t,\tau+\delta))^2 - \tfrac{1}{2}(\sigma^*(t,\tau))^2]\ dt \\ &\qquad\qquad + [\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]\ dW(t) \\ &\qquad\qquad + \tfrac{1}{2}[\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]^2\ dt\Big\} \\ &= \frac{1}{\delta}[1+\delta L(t,\tau)]\Big\{[r(t,\tau+\delta) - r(t,\delta)]\ dt \\ &\qquad\qquad + \sigma^*(t,\tau+\delta)[\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]\ dt \\ &= \qquad\qquad +[\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]\ dW(t)\Big\}. \end{aligned} \tag{5.2}$$

But

$$\frac{\partial}{\partial \tau}L(t,\tau) = \frac{\partial}{\partial \tau}\left[\frac{\exp\left\{\int_\tau^{\tau+\delta} r(t,u)\ du\right\} - 1}{\delta}\right]$$
$$= \exp\left\{\int_\tau^{\tau+\delta} r(t,u)\ du\right\}.[r(t,\tau+\delta) - r(t,\delta)]$$
$$= \frac{1}{\delta}[1+\delta L(t,\tau)][r(t,\tau+\delta) - r(t,\delta)].$$

Therefore,

$$dL(t,\tau) = \frac{\partial}{\partial \tau}L(t,\tau)\ dt + \frac{1}{\delta}[1+\delta L(t,\tau)][\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)].[\sigma^*(t,\tau+\delta)\ dt + dW(t)].$$

Take $\gamma(t,\tau)$ to be given by

$$\gamma(t,\tau)L(t,\tau) = \frac{1}{\delta}[1+\delta L(t,\tau)][\sigma^*(t,\tau+\delta) - \sigma^*(t,\tau)]. \tag{5.3}$$

Then

$$dL(t,\tau) = [\frac{\partial}{\partial \tau}L(t,\tau) + \gamma(t,\tau)L(t,\tau)\sigma^*(t,\tau+\delta)]\ dt + \gamma(t,\tau)L(t,\tau)\ dW(t). \tag{5.4}$$

Note that (5.3) is equivalent to

$$\sigma^*(t,\tau+\delta) = \sigma^*(t,\tau) + \frac{\delta L(t,\tau)\gamma(t,\tau)}{1+\delta L(t,\tau)}. \tag{5.3'}$$

Plugging this into (5.4) yields

$$dL(t,\tau) = \left[\frac{\partial}{\partial \tau}L(t,\tau) + \gamma(t,\tau)L(t,\tau)\sigma^*(t,\tau) + \frac{\delta L^2(t,\tau)\gamma^2(t,\tau)}{1+\delta L(t,\tau)}\right]\ dt$$
$$+\gamma(t,\tau)L(t,\tau)\ dW(t). \tag{5.4'}$$

## 34.6 Implementation of BGM

Obtain the initial *forward LIBOR curve*

$$L(0,\tau), \quad \tau \geq 0,$$

from market data. Choose a *forward LIBOR volatility function* (usually nonrandom)

$$\gamma(t,\tau), \quad t \geq 0,\ \tau \geq 0.$$

Because LIBOR gives no rate information on time periods smaller than $\delta$, we must also choose a *partial bond volatility function*

$$\sigma^*(t,\tau), \quad t \geq 0,\ 0 \leq \tau < \delta$$

for maturities less than $\delta$ from the current time variable $t$.

With these functions, we can for each $\tau \in [0,\delta)$ solve (5.4') to obtain

$$L(t,\tau), \quad t \geq 0,\ 0 \leq \tau < \delta.$$

Plugging the solution into (5.3'), we obtain $\sigma^*(t,\tau)$ for $\delta \leq \tau < 2\delta$. We then solve (5.4') to obtain

$$L(t,\tau), \quad t \geq 0,\ \delta \leq \tau < 2\delta,$$

and we continue recursively.

**Remark 34.1** BGM is a special case of HJM with HJM's $\sigma^*(t,\tau)$ generated recursively by (5.3'). In BGM, $\gamma(t,\tau)$ is usually taken to be nonrandom; the resulting $\sigma^*(t,\tau)$ is random.

**Remark 34.2** (5.4) (equivalently, (5.4')) is a stochastic *partial* differential equation because of the $\frac{\partial}{\partial\tau}L(t,\tau)$ term. This is not as terrible as it first appears. Returning to the HJM variables $t$ and $T$, set

$$K(t,T) = L(t,T-t).$$

Then

$$dK(t,T) = dL(t,T-t) - \frac{\partial}{\partial\tau}L(t,T-t)\ dt$$

and (5.4) and (5.4') become

$$\begin{aligned} dK(t,T) &= \gamma(t,T-t)K(t,T)\left[\sigma^*(t,T-t+\delta)\ dt + dW(t)\right] \\ &= \gamma(t,T-t)K(t,T)\left[\sigma^*(t,T-t)\ dt + \frac{\delta K(t,T)\gamma(t,T-t)}{1+\delta K(t,T)}\ dt + dW(t)\right]. \end{aligned} \tag{6.1}$$

**Remark 34.3** From (5.3) we have

$$\gamma(t,\tau)L(t,\tau) = [1+\delta L(t,\tau)]\frac{\sigma^*(t,\tau+\delta)-\sigma^*(t,\tau)}{\delta}.$$

If we let $\delta{\downarrow}0$, then

$$\gamma(t,\tau)L(t,\tau) \to \frac{\partial}{\partial\delta}\sigma^*(t,\tau+\delta)\bigg|_{\delta=0} = \sigma(t,\tau),$$

and so

$$\gamma(t,T-t)K(t,T) \to \sigma(t,T-t).$$

We saw before (eq. 4.2) that as $\delta{\downarrow}0$,

$$L(t,\tau) \to r(t,\tau) = f(t,t+\tau),$$

so

$$K(t,T) \to f(t,T).$$

Therefore, the limit as $\delta \downarrow 0$ of (6.1) is given by equation (2.5):

$$df(t,T) = \sigma(t,T-t)\left[\sigma^*(t,T-t)\ dt + dW(t)\right].$$

**Remark 34.4** Although the $dt$ term in (6.1) has the term $\frac{\delta\gamma^2(t,T-t)K^2(t,T)}{1+K(t,T)}$ involving $K^2$, solutions to this equation do not explode because

$$\begin{aligned} \frac{\delta\gamma^2(t,T-t)K^2(t,T)}{1+\delta K(t,T)} &\le \frac{\delta\gamma^2(t,T-t)K^2(t,T)}{\delta K(t,T)} \\ &\le \gamma^2(t,T-t)K(t,T). \end{aligned}$$

## 34.7 Bond prices

Let $\beta(t) = \exp\left\{\int_0^t r(u)\ du\right\}$. From (2.6) we have

$$\begin{aligned} d\left(\frac{B(t,T)}{\beta(t)}\right) &= \frac{1}{\beta(t)}[-r(t)B(t,T)\ dt + dB(t,T)] \\ &= -\frac{B(t,T)}{\beta(t)}\sigma^*(t,T-t)\ dW(t). \end{aligned}$$

The solution $\frac{B(t,T)}{\beta(t)}$ to this stochastic differential equation is given by

$$\frac{B(t,T)}{\beta(t)B(0,T)} = \exp\left\{-\int_0^t \sigma^*(u,T-u)\ dW(u) - \tfrac{1}{2}\int_0^t (\sigma^*(u,T-u))^2\ du\right\}.$$

This is a martingale, and we can use it to switch to the *forward measure*

$$\begin{aligned} I\!\!P_T(A) &= \frac{1}{B(0,T)}\int_A \frac{1}{\beta(T)}\ dI\!\!P \\ &= \int_A \frac{B(T,T)}{\beta(T)B(0,T)}\ dI\!\!P \quad \forall A \in \mathcal{F}(T). \end{aligned}$$

Girsanov's Theorem implies that

$$W_T(t) = W(t) + \int_0^t \sigma^*(u,T-u)\ du, \quad 0 \le t \le T,$$

is a Brownian motion under $I\!\!P_T$.

## 34.8 Forward LIBOR under more forward measure

From (6.1) we have

$$\begin{aligned} dK(t,T) &= \gamma(t,T-t)K(t,T)\left[\sigma^*(t,T-t+\delta)\;dt + dW(t)\right] \\ &= \gamma(t,T-t)K(t,T)\;dW_{T+\delta}(t), \end{aligned}$$

so

$$K(t,T) = K(0,T)\exp\left\{\int_0^t \gamma(u,T-u)\;dW_{T+\delta}(u) - \tfrac{1}{2}\int_0^t \gamma^2(u,T-u)\;du\right\}$$

and

$$\begin{aligned} K(T,T) &= K(0,T)\exp\left\{\int_0^T \gamma(u,T-u)\;dW_{T+\delta}(u) - \tfrac{1}{2}\int_0^T \gamma^2(u,T-u)\;du\right\} \\ &= K(t,T)\exp\left\{\int_t^T \gamma(u,T-u)\;dW_{T+\delta}(u) - \tfrac{1}{2}\int_t^T \gamma^2(u,T-u)\;du\right\}. \end{aligned} \tag{8.1}$$

We assume that $\gamma$ is nonrandom. Then

$$X(t) = \int_t^T \gamma(u,T-u)\;dW_{T+\delta}(u) - \tfrac{1}{2}\int_t^T \gamma^2(u,T-u)\;du \tag{8.2}$$

is normal with variance

$$\rho^2(t) = \int_t^T \gamma^2(u,T-u)\;du$$

and mean $-\frac{1}{2}\rho^2(t)$.

## 34.9 Pricing an interest rate caplet

Consider a floating rate interest payment settled in arrears. At time $T+\delta$, the floating rate interest payment due is $\delta L(T,0) = \delta K(T,T)$, the LIBOR at time $T$. A caplet protects its owner by requiring him to pay only the cap $\delta c$ if $\delta K(T,T) > \delta c$. Thus, the value of the caplet at time $T+\delta$ is $\delta(K(T,T)-c)^+$. We determine its value at times $0 \le t \le T+\delta$.

**Case I:** $T \le t \le T+\delta$.

$$\begin{aligned} C_{T+\delta}(t) &= I\!\!E\left[\frac{\beta(t)}{\beta(T+\delta)}\delta(K(T,T)-c)^+ \middle| \mathcal{F}(t)\right] \\ &= \delta(K(T,T)-c)^+ I\!\!E\left[\frac{\beta(t)}{\beta(T+\delta)}\middle|\mathcal{F}(t)\right] \\ &= \delta(K(T,T)-c)^+ B(t,T+\delta). \end{aligned} \tag{9.1}$$

**Case II:** $0 \le t \le T$.
Recall that

$$I\!P_{T+\delta}(A) = \int_A Z(T+\delta)\, dI\!P, \quad \forall A \in \mathcal{F}(T+\delta),$$

where

$$Z(t) = \frac{B(t, T+\delta)}{\beta(t) B(0, T+\delta)}.$$

We have

$$\begin{aligned} C_{T+\delta}(t) &= I\!E\left[\frac{\beta(t)}{\beta(T+\delta)}\delta(K(T,T)-c)^+ \Big| \mathcal{F}(t)\right] \\ &= \delta B(t, T+\delta)\underbrace{\frac{\beta(t)B(0,T+\delta)}{B(t,T+\delta)}}_{\frac{1}{Z(t)}} I\!E\left[\underbrace{\frac{B(T+\delta,T+\delta)}{\beta(T+\delta)B(0,T+\delta)}}_{Z(T+\delta)}(K(T,T)-c)^+\Big|\mathcal{F}(t)\right] \\ &= \delta B(t,T+\delta) I\!E_{T+\delta}\left[(K(T,T)-c)^+ \Big| \mathcal{F}(t)\right] \end{aligned}$$

From (8.1) and (8.2) we have

$$K(T,T) = K(t,T)\exp\{X(t)\},$$

where $X(t)$ is normal under $I\!P_{T+\delta}$ with variance $\rho^2(t) = \int_t^T \gamma^2(u, T-u)\, du$ and mean $-\frac{1}{2}\rho^2(t)$. Furthermore, $X(t)$ is independent of $\mathcal{F}(t)$.

$$C_{T+\delta}(t) = \delta B(t,T+\delta) I\!E_{T+\delta}\left[(K(t,T)\exp\{X(t)\}-c)^+\Big|\mathcal{F}(t)\right].$$

Set

$$\begin{aligned} g(y) &= I\!E_{T+\delta}\left[(y\exp\{X(t)\}-c)^+\right] \\ &= y\, N\left(\frac{1}{\rho(t)}\log\frac{y}{c} + \tfrac{1}{2}\rho(t)\right) - c\, N\left(\frac{1}{\rho(t)}\log\frac{y}{c} - \tfrac{1}{2}\rho(t)\right). \end{aligned}$$

Then

$$C_{T+\delta}(t) = \delta\ B(t,T+\delta)\ g(K(t,T)), \quad 0 \le t \le T-\delta. \tag{9.2}$$

In the case of constant $\gamma$, we have

$$\rho(t) = \gamma\sqrt{T-t},$$

and (9.2) is called the *Black caplet formula.*

## 34.10 Pricing an interest rate cap

Let

$$T_0 = 0,\ T_1 = \delta,\ T_2 = 2\delta,\ \dots,\ T_n = n\delta.$$

A cap is a series of payments

$$\delta(K(T_k, T_k) - c)^+ \quad \text{at time } T_{k+1},\ k = 0, 1, \dots, n-1.$$

The value at time $t$ of the cap is the value of all remaining caplets, i.e.,

$$C(t) = \sum_{k: t \le T_k} C_{T_k}(t).$$

## 34.11 Calibration of BGM

The interest rate caplet $c$ on $L(0,T)$ at time $T+\delta$ has time-zero value

$$C_{T+\delta}(0) = \delta B(0, T+\delta)\ g(K(0,T)),$$

where $g$ (defined in the last section) depends on

$$\int_0^T \gamma^2(u, T-u)\ du.$$

Let us suppose $\gamma$ is a deterministic function of its second argument, i.e.,

$$\gamma(t,\tau) = \gamma(\tau).$$

Then $g$ depends on

$$\int_0^T \gamma^2(T-u)\ du = \int_0^T \gamma^2(v)\ dv.$$

If we know the caplet price $C_{T+\delta}(0)$, we can "back out" the squared volatility $\int_0^T \gamma^2(v)\ dv$. If we know caplet prices

$$C_{T_0+\delta}(0), C_{T_1+\delta}(0), \dots, C_{T_n+\delta}(0),$$

where $T_0 < T_1 < \dots < T_n$, we can "back out"

$$\int_0^{T_0} \gamma^2(v)\ dv,\ \ \int_{T_0}^{T_1} \gamma^2(v)\ dv = \int_0^{T_1} \gamma^2(v)\ dv - \int_0^{T_0} \gamma^2(v)\ dv,$$
$$\dots,\ \int_{T_{n-1}}^{T_n} \gamma^2(v)\ dv. \quad (11.1)$$

In this case, we may assume that $\gamma$ is constant on each of the intervals

$$(0, T_0),\ (T_0, T_1),\ \dots,\ (T_{n-1}, T_n),$$

and choose these constants to make the above integrals have the values implied by the caplet prices. If we know caplet prices $C_{T+\delta}(0)$ for all $T \geq 0$, we can "back out" $\int_0^T \gamma^2(v)\,dv$ and then differentiate to discover $\gamma^2(\tau)$ and $\gamma(\tau) = \sqrt{\gamma^2(\tau)}$ for all $\tau \geq 0$.

To implement BGM, we need both $\gamma(\tau),\ \tau \geq 0$, and

$$\sigma^*(t,\tau), \quad t \geq 0,\ 0 \leq \tau < \delta.$$

Now $\sigma^*(t,\tau)$ is the volatility at time $t$ of a zero coupon bond maturing at time $t+\tau$ (see (2.6)). Since $\delta$ is small (say $\frac{1}{4}$ year), and $0 \leq \tau < \delta$, it is reasonable to set

$$\sigma^*(t,\tau) = 0, \quad t \geq 0, \ \ 0 \leq \tau < \delta.$$

We can now solve (or simulate) to get

$$L(t,\tau), \quad t \geq 0, \tau \geq 0,$$

or equivalently,

$$K(t,T), \quad t \geq 0, T \geq 0,$$

using the recursive procedure outlined at the start of Section 34.6.

## 34.12 Long rates

The long rate is determined by long maturity bond prices. Let $n$ be a large fixed positive integer, so that $n\delta$ is 20 or 30 years. Then

$$\begin{aligned}
\frac{1}{D(t,n\delta)} &= \exp\left\{\int_0^{n\delta} r(t,u)\,du\right\} \\
&= \prod_{k=1}^{n} \exp\left\{\int_{(k-1)\delta}^{k\delta} r(t,u)\,du\right\} \\
&= \prod_{k=1}^{n} [1 + \delta L(t,(k-1)\delta)],
\end{aligned}$$

where the last equality follows from (4.1). The long rate is

$$\frac{1}{n\delta}\log\frac{1}{D(t,n\delta)} = \frac{1}{n\delta}\sum_{k=1}^{n}\log[1+\delta L(t,(k-1)\delta)].$$

## 34.13 Pricing a swap

Let $T_0 \geq 0$ be given, and set

$$T_1 = T_0 + \delta, \ \ T_2 = T_0 + 2\delta, \ \ \ldots, \ \ T_n = T_0 + n\delta.$$

The swap is the series of payments

$$\delta(L(T_k,0)-c) \quad \text{at time } T_{k+1}, k=0,1,\ldots,n-1.$$

For $0 \le t \le T_0$, the value of the swap is

$$\sum_{k=0}^{n-1} I\!\!E\left[\frac{\beta(t)}{\beta(T_{k+1})}\delta(L(T_k,0)-c)\Big|\mathcal{F}(t)\right].$$

Now

$$1+\delta L(T_k,0)=\frac{1}{B(T_k,T_{k+1})},$$

so

$$L(T_k,0)=\frac{1}{\delta}\left[\frac{1}{B(T_k,T_{k+1})}-1\right].$$

We compute

$$\begin{aligned}
&I\!\!E\left[\frac{\beta(t)}{\beta(T_{k+1})}\delta(L(T_k,0)-c)\Big|\mathcal{F}(t)\right]\\
&=I\!\!E\left[\frac{\beta(t)}{\beta(T_{k+1})}\left(\frac{1}{B(T_k,T_{k+1})}-1-\delta c\right)\Big|\mathcal{F}(t)\right]\\
&=I\!\!E\left[\frac{\beta(t)}{\beta(T_k)B(T_k,T_{k+1})}\underbrace{I\!\!E\left[\frac{\beta(T_k)}{\beta(T_{k+1})}\Big|\mathcal{F}(T_k)\right]}_{B(T_k,T_{k+1})}\Big|\mathcal{F}(t)\right]-(1+\delta c)B(t,T_{k+1})\\
&=I\!\!E\left[\frac{\beta(t)}{\beta(T_{k})}\Big|\mathcal{F}(t)\right]-(1+\delta c)B(t,T_{k+1})\\
&=B(t,T_k)-(1+\delta c)B(t,T_{k+1}).
\end{aligned}$$

The value of the swap at time $t$ is

$$\begin{aligned}
&\sum_{k=0}^{n-1} I\!\!E\left[\frac{\beta(t)}{\beta(T_{k+1})}\delta(L(T_k,0)-c)\Big|\mathcal{F}(t)\right]\\
&=\sum_{k=0}^{n-1}[B(t,T_k)-(1+\delta c)B(t,T_{k+1})]\\
&=B(t,T_0)-(1+\delta c)B(t,T_1)+B(t,T_1)-(1+\delta c)B(t,T_2)+\ldots+B(t,T_{n-1})-(1+\delta c)B(t,T_n)\\
&=B(t,T_0)-\delta cB(t,T_1)-\delta cB(t,T_2)-\ldots-\delta cB(t,T_n)-B(t,T_n).
\end{aligned}$$

The forward swap rate $w_{T_0}(t)$ at time $t$ for maturity $T_0$ is the value of $c$ which makes the time-$t$ value of the swap equal to zero:

$$w_{T_0}(t)=\frac{B(t,T_0)-B(t,T_n)}{\delta\left[B(t,T_1)+\ldots+B(t,T_n)\right]}.$$

In contrast to the cap formula, which depends on the term structure model and requires estimation of $\gamma$, the swap formula is generic.